Chapter 1: Absolute and Relative Quantities

Objectives:
1. Distinguish between absolute and relative quantities.
2. Identify different types of relative quantities.
3. Analyze real data sets in Excel, using absolute and relative quantities.
4. Identify appropriate uses and misuses of both absolute and relative quantities.

There are two ways to measure numerical data, especially if the goal is to measure the least and greatest occurrence of some quantifiable variable.

Absolute and Relative Quantities Basics

The absolute quantity is a measure of the absolute occurrence of the variable. It is a “sheer” number. It tells how many or how much.

A relative quantity is generally an absolute quantity divided by some other quantity. It is the measure of an absolute quantity in relation to some other quantity. A relative quantity is often found by computing part / whole (part divided by whole). Often we convert relative numbers into percentages by multiplying by 100 (or moving the decimal two places to the right). Remember: \( \frac{1}{4} = 0.25 = 25\% \)

Types of Relative Quantities: RATIOS, FRACTIONS, DECIMALS, and PERCENTAGES are used when comparing same units of data, (examples: number of cell phones per person, percent of adults that are HIV positive) and RATES are used to compare quantities with different units (examples: miles per hour, number of crimes per 1000 people).

Example 1
Which of the following statements refer to relative quantities and which refer to absolute quantities?

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Relative</th>
<th>Absolute</th>
</tr>
</thead>
<tbody>
<tr>
<td>According to FIFA, the total attendance at the 2015 Women’s World Cup was 1,353,506.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than half of the students at ECC are women.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The ratio of full-time students to part-time students at ECC is about 3 to 7.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>According to a March 2015 article in Rock and Ice: The Climber’s Magazine, a climber’s chance of dying while trying to summit Annapurna in Nepal is around 35%.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One out of every 10 people are left-handed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>According to the Chicago Tribune*, officials estimate that about 250,000 people attended the Women’s March on Chicago on January 21, 2017.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The United States incarcerates 716 for every 100,000 residents, more than any other country, according to <a href="http://www.prisonpolicy.org/global">www.prisonpolicy.org/global</a>.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

More on Absolute Quantities

Example 2 The Excel file [https://faculty.elgin.edu/nscherger/courses/MTH104/World HIV Adults by Country 2013.xlsx](https://faculty.elgin.edu/nscherger/courses/MTH104/World HIV Adults by Country 2013.xlsx) contains data on the number of adults with HIV in countries around the world. Sort the dataset to determine the countries with the largest numbers of people living with HIV.

Even though we do not hear as much about the HIV epidemic in countries outside of Africa, several non-African countries (China, India, and the US) made the top ten list. Why should it not be surprising that they are in the top ten?

If we wanted to compare the severity of the HIV epidemic of different countries, do you think that the absolute number of people living with HIV is the best way to do that? Explain why or why not.

Example 3 The Excel file ([http://faculty.elgin.edu/nscherger/Courses/MTH104/Cellular Phones by Country.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Cellular Phones by Country.xlsx)) contains data on the number of cellular phones in countries around the world. One could reorganize the data by the number of cellular phones in descending order.

To sort in Excel: First, click on a cell in the column you wish to sort by. Then on the home ribbon at the far right, click on Sort & Filter, choosing the sort shown. (For more advanced sort options, go to the Data tab.)

One would find that ___________________________________ have the most cellular phones.

However, ________________________________________________________________.

Thus, another way to look at the data would be with relative quantities.
Group Activity 1: Sports Injuries I

1. For each pair of sports below, identify the one that your group considers more hazardous to its participants. Don't belabor this.
   a. bicycle riding or football?
   b. soccer or ice hockey?
   c. swimming or skateboarding?

2. Open the file on sports injuries (http://faculty.elgin.edu/nscherger/Courses/MTH104/Sports Injuries 2012.xlsx), which contains recent data on the number of participants in some of the most common sports in the US and the number of sports related injuries treated in US hospital emergency rooms.

   Sort the data on the number of injuries from greatest to least.

   Copy the sorted table in your Word document.

3. Using the number of injuries as a measure of the risk of getting hurt, which sports are the most hazardous in each of the three comparisons?
   a. bicycle riding or football?
   b. soccer or ice hockey?
   c. swimming or skateboarding?

4. Of all the sports, which is the
   a. least hazardous according to this measure?
   b. most hazardous according to this measure?

5. Are the quantities used in 3 and 4 absolute or relative measures?

6. Do you think the number of injuries is the best way to reflect the danger of a sport? Explain.
Chapter 1: Absolute and Relative Quantities (Continued)

More on Relative Quantities

Example 4 In the previous Excel file (http://faculty.elgin.edu/nscherger/Courses/MTH104/Cellular Phones by Country.xlsx), one could calculate the number of cellular phones per capita by creating a column D and dividing the number of cellular phones (column B) by the population of the country (column C).

In cell D6, enter: \( \frac{B6}{C6} \).

Then, hover the cursor over the small square in the lower right corner of cell D6, and hold down with your left mouse button and drag to fill the column. (Alternatively, you can double-click there to fill the column.)

If that calculation is done and column D is sorted, ____________________________ have the most cellular phones per capita.

Comparing Absolute and Relative Quantities

The above example shows a very different picture looking at absolute and relative quantities. Which is more informative? It depends on the purpose for which you are using the data.

An example of when the absolute quantities are more useful:

An example of when the relative quantities are more useful:
Unit 1: Number Sense
Chapter 1: Absolute and Relative Quantities

How and why did certain countries “move” when sorting by absolute quantity (number of cell phones) vs. relative quantity (cell phones per capita)?

The United States which had the 3rd highest number of cellular phones also has the 3rd highest adult population in the world, so it is not very surprising that it has a high number of cell phones. Once the population is taken account by calculating the relative quantity, the United States drops down in the list (to 115th), with a little less than 1 cell phone per person (0.972). The high absolute quantity yielded a low relative quantity because the number you divided by (the population) was so large. Conversely, there are smaller populated countries that top the relative quantity list, because while they do not have that many actual cell phones, their populations are so much smaller, that the number of cell phones actual outnumbers their smaller populations.

Often when dealing with information from entities of varying sizes, a relative quantity is more informative. By calculating the relative quantity, you are taking the differing populations into account.

In the previous example, we left the relative quantities in their decimal form; however, there will be times when you will be asked to express these relative quantities as percentages, which will be shown in the next example.

Example 5 Open the file on state poverty (https://faculty.elgin.edu/nscherger/Courses/Mth104/Poverty by State 2013.xlsx), which contains data on the number of persons below the poverty level in 2013.

a. Sort the data set to determine the states with the most people below poverty.

b. Should we conclude that these states are necessarily the “poorest?”

c. Determine the percentage of each state that is below the poverty line.
We begin the same way by dividing. So, in cell D4, enter =B4/C4.

Now, there are several ways to convert values into percentages:

- Multiply by 100, which will always result in ...
- Moving the decimal place to the right two places, or
- In Excel, either:
  - Multiply by 100 (*100)
  - Use the “%” button, which is in home ribbon and looks like
  - Right click and select Format Cells and then within the Number tab, select Percentage

Next, a quick way to add decimal places is to use the "Increase Decimal" button (or if needed, to express fewer decimals click on the "Decrease Decimal" button).

For this course, when asked for percentages, show two decimal places. (EX: 18.90%)

Finally, hover the cursor over the small square in the lower right corner of cell D4, and hold down with your left mouse button and drag to fill the column. (Alternatively, you can double-click there to fill the column.)

Now sort the data set by this new column to determine the states with the highest percentage of its population below poverty.

d. Which measure do you think most accurately describes the “poorest” states: absolute or relative (percentage)?

e. Determine the percentage of all persons below poverty.

To do so, we need to first sum columns B and C and then divide the sum from column B by the sum from column C.

In Excel: One way to sum a column is to highlight the data you would like to sum and then click the sum button from the home tab in the ribbon on the top: $\sum$ AutoSum. (Note: Many times you may not even need to highlight the data, if you are summing all the data in a column, Excel typically assumes this.)

So, the percentage of all persons below poverty is __________.
Example 6 In the Excel file https://faculty.elgin.edu/nscherger/courses/MTH104/World HIV Adults by Country 2013.xlsx, create a column D that gives the “Percent of the Population with HIV.” Express the values in this column as percentages, with two decimal places. Sort the dataset to determine the countries with the largest percentage of their population with HIV. Notice that we get a different list of countries here than we did when we previously determined the countries that had the largest absolute number of adults living with HIV.

For example, India ranked ______ according to the absolute number of HIV cases, with ______________ people having HIV; however, its ranking dropped to ______ according to the percentage of its population with HIV, with_________ percent of its population having HIV.

Why did India’s ranking drop so significantly when looking at the percent of its population with HIV?

As another example, Kiribati ranked ______ according to the absolute number of HIV cases, with just___________ people having HIV; however, its ranking jumped up to ______ according to the percentage of its population with HIV, with_________ percent of its population having HIV.

Why did Kiribati’s ranking increase so significantly when looking at the percent of its population in poverty?

Using the data we have here, what is the percentage of the global population that have HIV?
Group Activity 2: Sports Injuries II

1. Open the file on sports injuries (http://faculty.elgin.edu/nscherger/Courses/MTH104/Sports Injuries 2012.xlsx), which contains recent data on the number of participants in some of the most common sports in the US and the number of sports related injuries treated in US hospital emergency rooms.

Fill column D with the ratio of the number of injuries to the number of participants. This ratio is called the rate of injuries per participant.

(Note that after you divide, but before you convert to a percent, you may see a value of 0, which might seem wrong; however, this 0 is due to Excel rounding down. For example, 0.0123 might be rounded down to 0 initially. However, once you covert to a percentage and add on two decimals, you will see non-zero results that make more sense. For example, the 0.0123, which may have initially been rounded down to 0, will now be displayed as 1.23%.)

Express these ratios as percentages, with two decimals.

Sort the data according to the injury rate (column D) from greatest to least. Copy this sorted table into your Word document.

2. Using skateboarding as an example, state and explain what this measure means as a ratio. To help you, begin by filling in the blanks below with numbers and completing the sentence.

“There is approximately ____ injury(s) for every _____ skateboarders.”

3. Using this measure, which sport is the most hazardous in each of the three comparisons below?
   a. bicycle riding or football?
   b. soccer or ice hockey?
   c. swimming or skateboarding?

4. Of all the sports, which is the
   a. least hazardous according to this measure?
   b. most hazardous according to this measure?

5. Are the quantities used in 3 and 4 absolute or relative quantities?

6. Which measure do you think is better to reflect the risk of injury in a sport: the number of injuries (as in activity 1) or the injury rate (as in activity 2)?

7. Reread the statement about the data set:

“The file on sports injuries contains recent data on the number of participants in some of the most common sports in the US and the number of sports related injuries treated in US hospital emergency rooms.”
Now, what other information might you like to know to provide a more complete picture of sports injuries?

8. Calculate the injury rate, as a percentage, of all listed sports.

(You will need to add all the data in columns B and C to do this. Refer to the process illustrated at the end of the notes before this activity.)

State the injury rate (as a percentage, of all listed sports) in your Word document.

The first two group activities exemplify:

Key Idea

Quantitative measures often shed light on an issue but almost always are incomplete and limited.

Using quantitative methods correctly and recognizing their limitations are equally important.
Assignment 1: Absolute and Relative Quantities

Complete Assignment #1 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. This assignment will utilize the sort feature in Excel extensively. Start early, so if you have questions, you will have time to ask. All values and percentages should show 2 decimal places, unless otherwise specified.

1. Open the file on unemployment, ([http://faculty.elgin.edu/nscherger/Courses/MTH104/US Unemployment Data March 2019.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/US Unemployment Data March 2019.xlsx)), which contains data on the number of people unemployed by state. Sort the states by the number of people unemployed, from greatest to least. Copy and paste the first 5 rows of this sorted table into your Word document.

2. Are the values (number of people unemployed) in (#1) absolute or relative quantities?

3. Without any knowledge of states’ unemployment, why were the states that topped the list in (#1) not very surprising?

4. Create a column D (don’t forget to give it an appropriate title) with percent of labor force that is unemployed. Put it in percent form and show 2 decimal places. Now sort the data in column D, from greatest to least, and copy and paste the first 5 rows of this sorted table into your Word document.

5. Are the values you created in column D absolute or relative quantities?

6. If you were writing an article on states with an unemployment problem, would it make more sense to focus on states with the highest number of unemployed individuals (as in the states listed in 1) or would it make more sense to discuss the states with the highest percent of unemployed (as in the states listed in 4)? Explain your answer in complete sentences.
Chapter 2: Absolute and Relative (Percent) Change

Objectives:
1. Distinguish between and compute absolute and relative (percent) change.
2. Analyze real data sets in Excel, computing both absolute and relative quantities.
3. Identify appropriate uses and misuses of both absolute and relative change.

There are a couple ways to look at changes in numerical data: looking simply at the raw change in the values or by looking at the change in the values, relative to its original value.

Absolute Change
We use absolute change to describe the actual increase or decrease from a reference (or older/earlier) value to a new (or later) value:

\[ \text{Absolute Change} = \text{new value} - \text{reference value} = \]

Relative (Percent) Change
We use relative change to compare the absolute change to the reference (old) value:

\[ \text{Relative (Percent) Change} = \frac{\text{Absolute Change}}{\text{Reference Value}} = \frac{\text{New Value} - \text{Reference Value}}{\text{Reference Value}} = \]

Relative change is most often expressed by a percent, where the fraction is converted to a decimal number (by doing the division indicated) and then to a percentage using any of the methods discussed in the last section.

Example 1: During a 6-month period a company’s stock doubled in price from $10 to $20.

a. Calculate the absolute change in the stock price.

b. Calculate the relative (as a percent) change in the stock price.

c. Will you get the same answer if the stock started at $5 and doubled to $10?

Key Idea: When a quantity doubles, or is multiplied by a factor of _____, that is equivalent to a percent change of __________.
Example 2] With the rise of streaming services, many people are now cancelling their traditional cable subscriptions. According to a report from the Leichtman Research Group, 95,000 customers “cut the cable cord” in 2013 and in 2014, another 125,000 customers followed suit and also terminated their cable subscription.

a. How many more customers “cut the cable cord” in 2014 than in 2013?

b. By how many percent did the number of cable cancellations rise from 2013 to 2014?

Example 3] In 1991, in the US, there were 20,159 alcohol-related fatalities. Twenty years later, in 2011, there were 9,878 alcohol-related fatalities. What percent decrease is this?

Example 4] The website [www.hockeydb.com](http://www.hockeydb.com) reports that the per game average attendance of the Chicago Blackhawks during their 2000-2001 season was 14,996 and during their 2014-2015 season, it was 21,769.

a. What absolute increase does this represent? When might this be useful?

b. What percent increase does this represent? When might this be useful?
Example 5 The Excel file https://faculty.elgin.edu/nscherger/courses/MTH104/US immigration by region 1980 1990.xlsx contains data on immigration to the US from around the world. Sort the dataset to determine the regions with the largest absolute increase in their immigration population over this 10 year timespan.

Then, create a column E and compute the percentage change in the immigration population from 1980 to 1990. Express the values in this column as percentages, with two decimal places. Sort the dataset to determine the regions with the largest percentage increase in their immigration population over this 10 year timespan.

Notice that we get a different lists we get in looking at absolute increase versus percentage increase.

As an example, from 1980 to 1990, the number of immigrants from Eastern Africa rose by __________, which would have placed Eastern Africa in tenth place in terms of the absolute increase of immigrants; however, when looking at the percent increase, Eastern Africa came in ________ place, with a growth of _______ percent from 1980 to 1990.

Explain why Eastern Africa’s percent increase is so much higher than other regions that had a much larger absolute increase of immigrants
Group Activity 3: State Populations

In this activity, you will study the change in population in the states from 2000 to 2010.

1. Before proceeding to look at the data, briefly discuss among your group which states would you guess will have greatest growth? Don’t belabor this.

2. Open the file on state populations (http://faculty.elgin.edu/nscherger/Courses/MTH104/State Populations 1970-2010.xlsx), which contains official decennial census population figures for all states from 1970 to 2010.

   Fill column G with the difference between the 2010 and 2000 populations for each state.

   Sort the states according to their absolute change from greatest to least.

   Copy and paste the first ten rows of this sorted table into your Word document.

3. List the three states that had the largest increase (positive change) in population from 2000 to 2010. List the names and population increases of these states. (Be careful! Note that the population figures in the table are in thousands.)

4. Are the increases listed in (3) reflected of absolute change or relative change?

5. Fill column H with the percentage change in population from 2000 to 2010 for each state.

   Sort the states according to their percentage change, from greatest to least.

   Copy and paste the first ten rows of this sorted table into your Word document.

6. List the three states that had the largest percentage increase in population from 2000 to 2010. List the names and population percentage increases of these states.

7. Are the increases listed in (6) reflected of absolute change or relative change?

8. Why does the first list (in number 3) differ so much from the second list (in number 6)?
Chapter 3: Percentage Problems

Objective: 1. Solve various types of percentage problems.

There are a variety of types of problems that arise involving percentages that can range from basic to a bit trickier, if we’re not careful.

Basic “Percentage Of” Problems

“Part/Whole” Problems
Note that we’ve already been working with these types of problems. For example, in activity 2, when you computed the injury rate, you computed # injured (part) / total # of participants (whole).

Let’s look at another quick example.

Example 1 You recently got job at a restaurant as a server, where last night, you and trainer (an experienced server) worked as a team on the same tables all night, as part of your training. At the end of the night, the trainer told you that servers share their tips with those who work bussing the tables and then proceeds to give each of the two bussers $28.50 each, takes $76, and gives you $57. What percent of the total tips were you given?

“Translation” Problems

Recall: OF typically indicates _______________

IS typically indicates _______________

Example 2 Florida was a key swing state in the 2016 election. According to the New York Times*, Donald Trump received 4,617,886 votes, Hillary Clinton received 4,504,975 votes, and there were 378,756 votes for other candidates. What percent of the popular vote did Donald Trump receive? Express your answer as a percent, rounded to the nearest tenth (one decimal).

*nytimes.com/2017/01/29/business/5-million-for-a-super-bowl-ad-another-million-or-more-to-market-the-ad.html?r=0
Example 3

a. Use the first bullet in the clip to find the total number of homes in Lawrence.

b. Use the second bullet in the clip to find the total number of homes in Lawrence.

c. What do you notice?

Example 4
Open poverty in Illinois file (http://faculty.elgin.edu/nscherger/Courses/MTH104/Poverty by County in Illinois 2013.xlsx), and compute the number of individuals below poverty for each county.

Percent Change / Percent Growth or Loss / Percent More Than or Less than / Percent Increase or Decrease (also contrasted with Times More Than or Less Than)

Recall from the last section that Percent Change = \( \frac{\text{Absolute Change}}{\text{Reference Value}} \)

Again, note that we’ve already been working with these types of problems. For example, in activity 3, you computed the percentage change in state populations from 2000 to 2010.

Let’s look at some more varied examples.
Example 5  Compute the correct percent increase for the comic below.

Example 6  Andre makes $10/hr and Brianna makes $20/hr.

a.  By what percent is Brianna’s hourly wage higher than Andre’s hourly wage?

b.  By what percent is Andre’s hourly wage less than Brianna’s hourly wage?

c.  How many times greater (or by what factor) is Brianna’s hourly wage higher than Andre’s hourly wage?
Example 7  According to the World Health Organization, the life expectancy in the US is 79 years and in Canada, it is 82.

a. By what percent is the life expectancy of the US lower than the life expectancy of Canada?

b. By what percent is the life expectancy of Canada higher than the life expectancy of the US?

c. How many times greater is the life expectancy of Canada than that of the US? (OR- “By what factor is the life expectancy of Canada greater than that of the US?”)

Example 8  According to the Chicago Tribune*, about 250,000 people attended the Women’s March on Chicago on January 21, 2017. Other large non-sport related gatherings in Chicago include the Apollo 11 astronauts parade/rally (in 1969) with about 2 million and the Chicago Pride parade (in 2015) with about 1 million.


a. How many more people attended the Chicago Pride parade than the Women’s March?

b. By what percent was the attendance for the Chicago Pride parade smaller than the Apollo 11 parade?
Example 9] According to the New York Times*, the average price of a 30-second commercial during the 2017 Super Bowl is $5 million. According to Sports Illustrated*, the average price of a 30-second commercial during Game 7 of the historic 2016 World Series was about $500,000.

*nytimes.com/2017/01/29/business/5-million-for-a-super-bowl-ad-another-million-or-more-to-market-the-ad.html?_r=0

a. How many times greater is the cost of a Superbowl commercial than a World Series commercial?

b. By what percent is the cost of a Superbowl commercial larger than a World Series commercial?

Successive Percentage Change

One technique for successive percentage change problems is to:

1. Choose any original quantity (100 is nice to work with) and compute the first percent change (using \( \text{original} \pm \% \cdot \text{Original} = \text{New} \), noting to use addition for percent increase and subtraction for percent decrease).

2. Compute the second percent change, but this time starting the end result from step 1.

3. Finally, compute the percent change (using \( \frac{\text{New} - \text{Original}}{\text{Original}} \)), where the “New” is the end result from step 2 and the “Original” is the original quantity you chose in step 1.

Example 10] Your boss has told you that times are tough in the small business in which you work and that for the slower winter months, your salary will be decreased by 10%; however, he says not to worry, because in the spring, he will increase your salary 10%. What is the overall percent change to your salary?

Example 11] According to the Chicago Tribune*, Metra increased its fares in 2016 by about 2%, followed by an approximate 5.8% increase in fares in 2017 (that just went into effect on February 1st). What is the overall percent change in fares over these past two years?

Reverse Percentage Change

One technique is to use the relationship \( \text{original} \pm \% \cdot \text{original} = \text{new} \), noting that here, the “New” value will be given and it is the “Original” value that is the unknown.

**Example 11** If it costs a consumer $749 to buy the new iPhone 6 Plus (without a carrier subsidy), and Apple marks up the iPhone 6 by about 247.4%, what does it cost Apple to make an iPhone 6 Plus?

**Example 12** According to the Chicago Tribune*, the approximate 5 million people who attended the 2016 Cubs World Series parade/rally, was a 150% increase over the attendance at the 2015 Blackhawks Stanley Cup parade/rally. How many people attended the 2015 Blackhawks parade/rally?

Group Activity 4: Percentages

1. Open the file on poverty in Illinois (http://faculty.elgin.edu/nscherger/Courses/MTH104/Poverty by County in Illinois 2013.xlsx) (from class today). Use the sort feature to determine the three counties with the largest percent of people in poverty. Copy and paste the first three rows into your Word document.

2. Use the sort feature to determine the three counties with the greatest number of people in poverty. Copy and paste the first three rows into your Word document.

3. Suppose that the Illinois State Legislature has $2 million earmarked to help the poor in Illinois. Who do you think should get the most money - the counties with the highest percentage of their population in poverty (so those in (1)) or the counties with the highest number of individuals in poverty (so those in (2))? Explain your choice.

   (NOTE: You will build on these answers in your individual assignment #2.)

The remaining exercises do not require the use of Excel; however, for full credit, you must show all work and correctly label your answer. (Sometimes this label may just be a percent symbol.)

4. On the 2014-15 Chicago Bulls roster, the top 5 players’ (Rose, Boozer, Noah, Gibson, and Gasol) salaries totaled $60,240,876. The remaining 10 players’ salaries totaled $20,773,806. The top 5 players make up what percentage of the total team salaries?

5. Chicago now has the highest sales tax of any major US city, at 10.25%. Currently, the sales tax is 8.25% in Elgin. You plan to purchase a new 55” curved flat screen television for $2350 after the new year, during the post-holiday sales.
   a. How much would you pay in Elgin?
   b. How much would you pay in Chicago?
Unit 1: Number Sense
Chapter 3: Percentage Problems

6. Average household credit card debt was $19,000 in January 2009 and has dropped to $15,706 in January 2015. By what percent did the average household credit card debt decrease over those 6 years?

7. The average life expectancy of Nigeria in 2013 was 52.5. The average life expectancy of Mexico was 77.5.
   a. How many times larger is the life expectancy of Mexico than the life expectancy of Nigeria?
   b. By what percent is the life expectancy of Mexico larger than the life expectancy of Nigeria?
   c. By what percent is the life expectancy of Nigeria smaller than the life expectancy of Mexico?
8. Chicago’s population decreased by 6.92% from 2000 to 2010 to 2,692,598. What was Chicago’s population in 2000?

9. The value of stocks can vary dramatically. Suppose a stock increases in value by 50% one year. The next year it decreases by 50%. Compute by what percent the value of the stock has changed over the two year period?
Assignment 2: Absolute and Relative Quantities and Change & Other Percentage Problems

Complete Assignment #2 in MyOpenMath and the problems below.

1. We have already looked at the file on poverty in Illinois
   (http://faculty.elgin.edu/nscherger/Courses/MTH104/Poverty by County in Illinois 2013.xlsx).

   In group activity 4, you were asked to if the Illinois State Legislature had $2 million earmarked to help
   the poor in Illinois, who you thought should get the most money - the counties with the highest
   percentage of their population in poverty or the counties with the highest number of individuals in
   poverty. Hopefully you answered that the counties that have the most number of individuals in poverty
   have the most need, so they deserved the most money.

   One way to fairly distribute the money, so that the counties with the most number of people in poverty
   get the most money, is to determine the percent of the total poverty population that resides in each
   county (and then give that county that percent of the money).

   a. To do this, create a column D with the number of individuals below poverty for each county
      (which we did together in class). Be sure to title the column appropriately.

      As you did in group activity 4, #2, sort from greatest to least according to the number of
      individuals below poverty in each county.
      Copy and paste the first three rows of this sorted table into your Word document.

   b. Next, find the total number of people in poverty by summing column D (use AutoSum). State
      this total poverty population.

   c. Now, determine the percent of the total poverty population that resides in each county. To do
      this, create a column E (with the appropriate title) by computing:

      \[
      \text{PART \ (the number in poverty in each state, which are the values in the column you created in part a)}
      \times \text{WHOLE \ (the total poverty population, which is the value found in part b)}
      \]

      *While you should be using a cell reference for the numerator, FOR THE DENOMINATOR, TYPE
      THE ACTUAL NUMBER and do not use a cell reference (because we do not want the
      denominator to change to other cells as we fill the column).

      Copy and paste the first 5 rows of this table into your Word document.

   d. Finally, create a column F (with an appropriate title) that will determine how much of the $2
      million would go to each county by multiplying the percent of the total poverty population of
      each county (in column E) by $2 million. Copy and paste the first 5 rows of this sorted table into
      your Word document.
2. Open the file on frequent moviegoers (http://faculty.elgin.edu/nscherger/Courses/MTH104/Frequent Movie Goers by Age 2010-14.xlsx). This file contains data from 2010 to 2014 on the number of frequent moviegoers by age group.
   a. Create a column G (don’t forget to give it an appropriate title) with the absolute change in the number of moviegoers from 2011 (note this is 2011 and not 2010) to 2014. Then sort the age groups by this column G (from greatest to least) and copy this sorted table into your Word document.
   b. Create a column H (don’t forget to give it an appropriate title) with the percent change in the number of moviegoers from 2011 to 2014. Now sort by column H (from greatest to least) and copy this sorted table into your Word document.
   c. Complete the following statement:
      
The 50-59 age group increased its number of frequent moviegoers by ___________ from 2011 to 2014, which was a percent growth of ________.
      
      Over this same time period, the 60+ age group added even more frequent moviegoers, increasing by ___________ from 2011 to 2014, but that resulted in the smaller percent growth of ________.
   d. Explain the statement in (c). In other words, why did the 60+ age group gain more moviegoers than the 50-59 age group, but the 60+ group had a smaller percent growth?

© Scott Adams, Inc./Dist. by UFS, Inc.
Unit 1 Review

This first small test covers material presented in Chapters 1-3.

The following represents a list of the major topics you should be able to do.

1. Sort data sets in Excel to answer various questions.
2. Distinguish between absolute and relative quantities and discuss situations where one type of measure may be more appropriate than the other.
3. Compute a relative quantity (i.e., a ratio left in decimal form or expressed as a percentage when appropriate), including in Excel for an entire spreadsheet of data and interpret relative quantities.
4. Sum columns of data in Excel and use those sums where appropriate to compute particular relative quantities (as in number 6 in group activity 1).
5. Distinguish between absolute and relative (percent) change and discuss situations where one type of measure may be more appropriate than the other.
6. Compute absolute and relative (percent) change, including in Excel for an entire spreadsheet of data.
7. Compute other percentage-related problems in for an entire spreadsheet of data.
8. Solve all the various types of percentage problems discussed in class, including but not limited to “successive percentage change” problems and “reverse percentage change” problems.
1. Indicate which of the following statements refer to relative quantities and which refer to absolute quantities.

   a. Millennials now match Baby Boomers with the portion of the voting electorate that they comprise, with each generation now making up about 31% of the voting-eligible population

   b. August 2016 was Chicago’s worst in decades, with 90 murders

   c. The U.S. death rate rose to 729.5 deaths per 100,000 people in 2015, one of the few increases in the death rate in the last number of decades.

   d. As of August 31, 2016, there were a total of 2722 Zika cases in the US.

2. When Illinois passed a “stop-gap budget” in June of 2017, the state’s colleges received money in the following three ways: $655 million went to the nine state universities, $114 million went to the state’s community colleges, and $151 million went to tuition grants for low-income students that were promised last year but not paid. What percent of the total amount of money that went to colleges did the community colleges receive?

3. The number of Democrats who voted in the presidential primaries in 2016 was about 21% less than in 2008, when 38,111,341 Democrats who voted during the presidential primaries. How many Democrats voted in this year’s presidential primaries?
Unit 1: Number Sense

Unit Summary

4. In the 2016 summer Olympics, the US won 121 medals, China won 70 medals, Great Britain won 67 medals, Russia won 56 medals, Germany and France each won 42 medals, and Japan won 41 medals.

   a. How many more medals did China win than Germany?

   b. By what percent is the number of medals China won larger than the number of medals Germany won?

   c. How many times more medals did the US win than Japan?

5. In 2016, the number of Republicans who voted in the presidential primaries was 31,108,968, which is 62% more than the number of Republicans who voted in the presidential primaries in 2012. How many Republicans voted in the presidential primaries in 2012?

6. While shopping, you find a coat you like that was originally marked down 30% off and now it is on clearance with an additional 15% off. By what overall percent is the coat marked down from its original price?
Use Excel to Answer the Remaining Questions

Open the file entitled “Murder Rates 2000-2014 by State” and determine the following.

7. First, notice that the homicide data is given as a “Rate per 100,000 Population.” Why is it important to use a rate here and NOT the absolute number of homicides when comparing different states?

8. Determine the states with the highest homicide rate in 2010 by sorting the table from largest to smallest, according to the homicide rate in 2010. Copy and paste the first three rows of the Excel spreadsheet into your Word document.

9. Determine the states with the highest homicide rate in 2014 by sorting the table from largest to smallest, according to the homicide rate in 2014. Copy and paste the first three rows of the Excel spreadsheet into your Word document.

10. Create column D that gives the “Absolute Change in Homicide Rate.”

Determine the states with the largest absolute decrease in their homicide rate from 2010 to 2014 by sorting the table from smallest to largest, according to this new column D. Copy and paste the first three rows of the Excel spreadsheet into your Word document.

11. Create column E that gives the “Percent Change in Homicide Rate.” Express these percentages to two decimal places.

Determine the states with the largest percent decrease in their homicide rate from 2010 to 2014 by sorting the table from smallest to largest, according to this new column E. Copy and paste the first three rows of the Excel spreadsheet into your Word document.

12. Notice that the states that had the largest absolute decrease in their homicide rates were not the same cities with the largest percent decrease in their homicide rates.

For example, let’s compare the District of Columbia and Massachusetts. Washington DC had the largest absolute decrease in homicide rate and Massachusetts was seventh; however, when looking at the percent decrease, Massachusetts the greatest percent decrease and Washington DC moved fifth.

Explain why this occurs.
Chapter 4: Types of Data and Graphical Displays of Data

Objective:
1. Determine if data is categorical or numerical / quantitative.
2. Interpret pie charts, bar and multiple bar charts, and scatter plots.
3. In Excel, create pie charts, bar and multiple bar charts, and scatterplots, for appropriate data sets.
4. Describe scatterplots/line graphs.
5. Identify features of misleading and generally poorly constructed graphs, including actual graphs used in the media.

Types of Data

Categorical Data
A set of data is said to be categorical if the values can be sorted according to non-overlapping categories based on some qualitative trait. Some examples:

Numerical (Quantitative) Data
Numerical data (or quantitative data) is data measured or identified on a numerical/quantitative scale. Some examples:

Types of Graphs

Pie Charts
- Used to summarize categorical data.
- Consists of a circle, divided into segments, where each segment represents a particular category.
- The area of each segment is proportional to the number of cases in that category.
- Only used where both the categories and the quantities each add up to a whole.
- CATEGORIES MUST BE DISJOINT!

The most common errors with pie charts:
- Categories that do not make a whole
- Categories that overlap

Key Idea
Creating pie charts in Excel

**Example 1** Consider the data set on the ECC’s student population by ethnicity (http://faculty.elgin.edu/nscherger/Courses/MTH104/ECC Students by Ethnicity.xlsx).

**Creating a pie chart for the ethnicity of ECC students**
- Highlight the data in the appropriate columns (so, cells A5-A11 and cells B5-B11).
- From the insert ribbon, select 2-D Pie.

**Other options**

**Legend:** To delete the legend (the labels, which here list the ethnicities), simply tap on it and hit delete.

**Resizing:** To resize the pie chart, tap on the “invisible” white square around the pie chart and then use your cursor to make that box (the plot area) larger or smaller.

**Title:** If a title box does not appear, to add a title, click on the chart elements “plus sign” in the upper right of the chart and select Chart Title. After you see the title box above your chart, simply double-click to edit the title. Here, let’s enter “ECC Students by Ethnicity 2013-14.”

**Labels:** To add labels and percentages directly on your pie chart, again go to chart elements and now select Data Labels then click on the arrow to right of Data Labels and select More Options.

A menu will appear at the right.

Under label options, check category name, percentage, and show leader lines. (The category name is especially important if you are printing in black-and-white.)

Under number, because we have some values that are less than 1%, let’s choose percentage, so we can specify one decimal.

**Note:** To “clean up” the graph, you can often just click and drag to make it more readable. You will want to experiment to get efficient at making nice charts.

To copy and paste your graph into a Word document, simply right click on the area and select copy and then in your Word document, right click and select paste. (One good idea is to select
the last paste option, as a picture, so if we change things in Excel, it won’t change our graph in Word.)

Once in Word, when you tap on the chart you will see “Picture Tools” and the top and can make further adjustments.

Here is the final graph:

**EXCEL Check:** Verify the percents shown are correct by computing them for yourself on the spreadsheet!

**Bar Charts**

- Can be displayed horizontally or vertically and they are usually drawn with a gap between the bars (whereas, the bars of a histogram, which is a special type of bar chart, are drawn with no gap).
- Used to summarize quantitative or categorical data.

For categorical data, there must be an associated quantitative variable (typically, the number of cases in that category).

**Key Idea**

It displays the data using a number of rectangles of equal width, each of which represents a particular category. The length (and hence, the area) of each rectangle is proportional to the number of cases in the category it represents (for example, age group, religion, etc...).
Creating bar (column) charts in Excel

**Example 2** Consider the data set on the Planned Parenthood services
(http://faculty.elgin.edu/nscherger/Courses/MTH104/Planned Parenthood Services.xlsx).

---

**Creating a bar chart for Planned Parenthood services**

- Highlight the data in the appropriate columns (so, cells A5-A10 and cells B5-B10).
- From the insert ribbon, select the first 2-D Column.

**Other options**

**Elements:** As with pie charts, clicking on the chart elements “plus sign” in the upper right of the chart will reveal many options. Typically, we want to minimally select axes, axis titles, chart title, and gridlines. (In particular, getting into the habit of always including a chart title and both axes titles is important.)

**Axis titles:** Now, double-click on the title box and axes title boxes and enter appropriate titles. For this example, enter “Planned Parenthood Services 2011-2012” for the title, “Type of Service” for the horizontal axis title, and “Number” for the vertical axis title.

**Other editing:**

- If you click on any box with text, you can then alter its appearance (font, color, etc...).
- If you right click on the vertical axis numbers and then select Format Axis, you will get different options, such as specifying your own minimum, maximum, number of increments, or units.

Let’s say we wanted to display in units of millions, then we could select the option shown and then change our horizontal axis title to “Number in Millions.”

---

Here is the graph:
Creating multiple bar (column) charts in Excel

Example 3 Consider the data set on the movie attendance by age (http://faculty.elgin.edu/nscherger/Courses/MTH104/Frequent Movie Goers by Age 2010-14.xlsx).

Creating a multiple bar chart for movie attendance in 2010, 2012, and 2014:

- Highlight the data in the appropriate columns (so, cells A6-A12, B6-B12, D6-D12, and F6-F12), where to skip columns C and E and still select columns D and F, hold down the control button and highlight.

- From the insert ribbon, select the first 2-D Column.

Important note!
With a double-bar chart, you must LEAVE THE LEGEND and edit it to reflect the appropriate category names. To edit the legend, click on it and initially the entire legend (series 1, series 2, and series 3) is selected. Then click again on just series 1, so that only series 1 is selected.

Now, right click and choose Select Data. In the box that appears, with series 1 highlighted, select and then type in an appropriate label (here, 2010). Then do the same for series 2 and series 3.

You should get a graph that looks similar to the following:
There are advantages and disadvantages to multiple-bar charts.

- Their main advantage is their succinctness and the ability they afford to make comparisons within categories and across categories; they are best used in printed works so that a viewer can study them carefully.
- Their disadvantage is that they often present far too much information to viewers of presentations; it is hard to make a single, clear point with them, and presenters tend not to leave them up long enough to absorb the information fully.

**Misleading / “Bad” Graphs**

Consider the following misleading and bad graphs. Let's discuss what makes them bad and/or misleading.

**Example 4** The next two graphs were release by CNN and were in regards to the Terri Schiavo case and a court’s decision to allow her husband to remove her feeding tube. Which graph is misleading? Why? How are the differing impressions created?
Example 5 What is misleading about this graph, from an advertisement for Zantac?

Example 6 This ad not only appeared in the Chicago Tribune but also appeared on billboards along the Kennedy Expressway. Again, what is misleading about this graph and why?
Example 7 Observe the following pie chart. Do you have any mathematical critiques?

Enter the data from this pie chart as follows into Excel and create a pie chart showing the percents. What do you observe?

<table>
<thead>
<tr>
<th>Source</th>
<th>Amount (in millions):</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEG Corp</td>
<td>$50</td>
</tr>
<tr>
<td>City Government</td>
<td>$265</td>
</tr>
<tr>
<td>Maloof Family</td>
<td>$85</td>
</tr>
<tr>
<td>Unknown</td>
<td>$10</td>
</tr>
</tbody>
</table>
Group Activity 5: Pie Charts and Bar Graphs

1. Open the file, http://faculty.elgin.edu/nscherger/Courses/MTH104/IL Population by Age.xlsx, which shows the population of Illinois in 2010 classified by age. Make a pie chart of this data. Remember to include an appropriate title and to display the category names and the percentages (not the values) on the pie chart. Copy this pie chart into your Word document.

2. To verify that Excel’s percentages are correct in our pie chart from (1), create a column C (with an appropriate title) which contains the percentage of total population for each age category. (Note: You first need to find the total population, which is the sum of column B.) Express the percentages in column C with two decimal places, and paste the resulting table into your Word document.

3. Consider the data below along with the pie chart made from that data.

<table>
<thead>
<tr>
<th>ECC Students</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>7786</td>
</tr>
<tr>
<td>Female</td>
<td>9251</td>
</tr>
<tr>
<td>White, non-Hispanic</td>
<td>8127</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6219</td>
</tr>
<tr>
<td>Asian / Pacific Islander</td>
<td>1090</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>954</td>
</tr>
<tr>
<td>American Indian / Alaskan, Native Hawaiian, or Pacific Island</td>
<td>102</td>
</tr>
<tr>
<td>Non-resident Alien</td>
<td>102</td>
</tr>
<tr>
<td>Unknown</td>
<td>443</td>
</tr>
</tbody>
</table>

Explain what is wrong with the pie chart created.

4. Suppose that this data on population by age is given to you as:

<table>
<thead>
<tr>
<th>Age</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-14 Years</td>
<td>123456</td>
</tr>
<tr>
<td>15-24 Years</td>
<td>2345678</td>
</tr>
<tr>
<td>25-49 Years</td>
<td>98765432</td>
</tr>
<tr>
<td>50-65 Years</td>
<td>654321</td>
</tr>
</tbody>
</table>

What criteria for making pie charts should prevent you from making a pie chart with this data?
5. Open the file, http://faculty.elgin.edu/nscherger/Courses/MTH104/Fundraising by Political Party 1987-2012.xlsx, which shows amount of money raised by political party from 1987-2012. Create a bar chart (what Excel calls a “column chart”) showing just the Democrats. (Be sure to include a title and x and y axes labels.) Copy and paste the chart into your Word document.

6. Create a bar chart (what Excel calls a “column chart”) showing just the Republicans. (Highlight column A data and then hold down the control button to skip over column B and be able to also highlight column C; Be sure to include a title and x and y axes labels.) Copy and paste the chart into your Word document.

7. Create a multiple bar chart (what Excel calls a “column chart”) showing Democrats and Republicans. (Be sure to include a title and x and y axes labels AND an appropriately labeled legend.) Copy and paste the chart into your Word document.

8. Which do you think is better here – two separate bar charts (as in 5 and 6) or one multiple bar chart (as in 7)? Explain.

9. Examine the multiple bar chart from 7. What do you notice about how the Democrats and Republicans compared to one another up through the early 2000s and what has happened over the last number of years?

10. Can your group come up with any reasons for this change you noticed over the last number of years (in number 9)?  (*Background:* The Republicans historically did better with fundraising because of an advanced direct mail system they had established which consistently raised millions in small contributions. So, can you think about what the Democrats did differently than the Republicans in fundraising since 2007?  *Hint:* President Obama was elected in 2008 and 2012 and the change reflected in the chart is something that the Democrats and Obama’s campaign were widely publicized as being better at than the Republicans.)

11. What is misleading about the bar chart at the right? Specifically, what is done to give the appearance of the great difference between the bars?
Chapter 4: Types of Data and Graphical Displays of Data (continued)

Scatterplots and Line Graphs

- Scatterplots and line graphs are a way of summarizing data, often used in exploratory data analysis to illustrate the major features of the distribution of the data in a convenient form.
- Scatterplots and line graphs should “tell a story” (often over time, as the independent variable) as you look at the graph from left to right.
- Vocabulary used to “tell the story” includes:
  - Increasing – graph is going *uphill* as you look from left to right.
  - Decreasing – graph is going *downhill* as you look from left to right.
  - Relative Minimums and Maximums – low and high points of the graph in a local neighborhood. 
    ie, the *valleys* and *peaks* of the graph.
  - Absolute Minimums and Maximums – smallest and largest values on the entire graph.
  - Increasing and Decreasing at Faster or Slower Rates – judged by the steepness of the increase or decrease.
  - Periodic – graphs that have a pattern that repeats.

Creating a scatterplot or line graph in Excel

**Example 8** Consider the data set on the Chicago census from 1840-2010 ([http://faculty.elgin.edu/nscherger/Courses/MTH104/Chicago Population 1840-2010.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Chicago Population 1840-2010.xlsx)).

To create a scatterplot or line graph for Chicago’s historical census data:

- Highlight the data in the appropriate columns (so, cells A2-A19 and B2-B19).
- From the insert ribbon, select the first scatterplot option. You should now have a basic scatterplot.
- As before, click on the chart elements “plus sign” in the upper right of the chart and select axes, axis titles, chart title, and gridlines.
  Now add appropriate titles and labels to the axes.
- You will want to experiment to get efficient at making nice charts. For example, with this graph, let’s right click on the horizontal axis numbers and then select Format Axis, where you will get different options. We could now specify our minimum as 1820, our maximum as 2020 maximum, and the major units as 20.
Unit 2: Statistics
Chapter 4: Types of Data and Graphical Displays of Data

- You should now have a graph that looks like the one below.

![Chicago Population Graph](image)

- Some scatterplots are easier to read and interpret as line graphs.
To do so, tap on your chart in Excel and observe the Chart Tools that appear on the ribbon at the top. From the Design tab, select Change Chart Type.

Then, from the box that appears, select the second scatterplot option, which is Scatter with Smooth Lines and Markers.

![Change Chart Type Dialog](image)

Now, your graph should look like the one that follows:
Describing scatterplots and line graphs

To describe scatterplots and line graphs, “tell the story from left to right,” using the proper terminology from above (increasing, decreasing, etc...). For example, a good description for the above graph might read:

“Chicago’s population started at its absolute minimum of just 4,470 in 1840 and grew slowly from 1850 to 1880. From 1880 to 1930, it continued to increase, but did so at a faster rate. The population dipped a bit in 1940, but then increased to its absolute maximum of 3,620,962 in 1950. The population began decreasing after that, reaching a relative minimum of 2,783,726 in 1990 and then increasing to a relative maximum of 2,896,016 in 2000 and ending a bit lower at 2,695,598 in 2010.”

Example 9 Identify the relative and absolute extrema on the graph.
Example 10 Use the Growth of Incarceration graph below to answer the following questions, using phrases like relatively stable, decreasing slowly, decreasing quickly, increasing slowly, or increasing quickly.

![The Growth Of Incarceration Graph](image)

a. Describe what was generally happening to the US imprisonment rate from from 1880-early 1920s and from 1940-1960.

b. Describe what was happening to the US imprisonment rate from from early 1920s-1940 and from 1970-1980.

c. Describe what was happening to the US imprisonment rate from from 1980-late 1990s.
Example 11 Consider the graph below of the 12-month percent change for the costs of tuition and fees as compared to the change in cost for all items.

Identify the absolute and relative maximums and minimums for the 12-month percent change in the cost of college tuition and fees.

Absolute Maximum ________________________________
Absolute Minimum ________________________________
Relative Maximum ________________________________
Relative Maximum ________________________________
Relative Minimum ________________________________
Relative Minimum ________________________________
Relative Minimum ________________________________
Relative Minimum ________________________________
Periodic data in scatterplots and line graphs

**Example 12** Consider the data set on the average high temperatures each month in Elgin from 2000 through 2015, which is available at: [http://faculty.elgin.edu/nscherger/Courses/MTH104/Elgin Temperature.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Elgin Temperature.xlsx).

Create a line graph (so, a scatterplot, selecting the second option) of the data from June 2011 through June 2015. As usual, create an appropriate title and axes’ labels.

You should get a graph that looks like the one below.

![Elgin Average High Temperatures](image)

Observe the cyclical nature of this graph. That is, how its behavior appears to “repeat” every 12 months. We would describe temperature as *periodic* phenomenon, and the value of 12 is what we would say is the *period* of this graph.

There are a plethora of real-world data that is periodic in nature, including sun spots, sound waves, and electromagnetic radiation.

**Misleading / “Bad” Graphs**

Consider the following misleading and bad graphs. Let’s discuss what makes them bad and/or misleading.

**Example 13** Compare the two graphs below. Which graph is bad/misleading? Why? How are the differing impressions created?
Example 14: Observe this line graph. Do you have any mathematical critiques?

Enter the data from this line graph as follows into Excel and create a line graph. What do you observe?

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abortions</td>
<td>289,750</td>
<td>327,000</td>
</tr>
<tr>
<td>Cancer Screenings &amp; Other Prevention Services</td>
<td>2,007,371</td>
<td>935,573</td>
</tr>
</tbody>
</table>

For more information, read this detailed breakdown from PolitiFact.org: [http://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading/]
Group Activity 6: Scatterplots and Line Graphs

1. Open the file, [http://faculty.elgin.edu/nscherger/Courses/MTH104/US motor vehicle theft rates 1980-2012.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/US motor vehicle theft rates 1980-2012.xlsx), which contains motor vehicle theft rate for the US from 1980 to the near present. Make a scatterplot (select an option under scatterplot that still shows the data points, so the first or second option) of the data (be sure to include a title and appropriate labels for the x and y axes) and paste it into your Word document.

2. Refer to your graph in (1) and identify the absolute extrema. (Remember that absolute extrema are the smallest and largest values on the entire graph and they may or may not be at the location of a relative extrema.)
   a. Give both the year and rate of the absolute minimum.
   b. Give both the year and rate of the absolute maximum.

3. Open the file, [http://faculty.elgin.edu/nscherger/Courses/MTH104/Number of Motorcycles 1960-2012.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Number of Motorcycles 1960-2012.xlsx), which contains the number of motorcycles in the US from 1960 to 2012. Make a scatterplot (select an option under scatterplot that still shows the data points, so the first or second option) of the data (be sure to include a title and appropriate labels for the x and y axes) and paste it into your Word document.

4. Carefully describe your graph in (3), “telling the story over time from left to right,” by filling in each of the following blanks using specific years, values from the Excel table, and vocabulary (such increasing, decreasing, and absolute/relative minimum and maximum).

   The number of motorcycles started at an absolute ______________ of __________ motorcycles in the year _______. The number of motorcycles then __________ over the years from _______ to _______, reaching a ______________ ______________ of __________ motorcycles in the year _______. After that, the number of motorcycles then __________ over the years from _______ to _______, reaching a ______________ ______________ of __________ motorcycles in the year _______. The number of motorcycles then increased, reaching an absolute ______________ of __________ motorcycles in the year _______.

5. Open the file, [http://faculty.elgin.edu/nscherger/Courses/MTH104/Births to Unmarried Mothers ages 25-29 1940-2013.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Births to Unmarried Mothers ages 25-29 1940-2013.xlsx), which contains the number of births to unmarried mothers, ages 25-29 in the US, from 1940-2013. Make a scatterplot (select an option under scatterplot that still shows the data points, so the first or second option) of the data (be sure to include a title and appropriate labels for the x and y axes) and paste it into your Word document.

6. Refer to your graph in (5) and describe each of the intervals below as increasing slowly, increasing quickly, relatively stable, decreasing slowly, or decreasing quickly.
   a. 1940-1975
   b. 1975-1990
   c. 1990-1997
   d. 1997-2009
Assignment 3: Graphical Displays of Data

Complete Assignment #3 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. Consider the data in the table below on property crimes in Illinois in 2013, which is available at http://faculty.elgin.edu/nscherger/Courses/MTH104/Property Crime in Illinois 2013.xlsx.

<table>
<thead>
<tr>
<th>Property Crimes in Illinois in 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Arson</td>
</tr>
<tr>
<td>Burglary</td>
</tr>
<tr>
<td>Motor Vehicle Theft</td>
</tr>
<tr>
<td>Other Theft</td>
</tr>
</tbody>
</table>

Assuming that these are the only four types of property crimes, a pie chart is appropriate. Thus, make a pie chart to display the data. Include an appropriate title and display category names and percentages. Paste this pie chart into your Word document.


   a. Make two single-bar charts that would inform the reader about violent crime rates in Illinois by region in 1983 and 2013. Include appropriate titles and axes labels. Paste these charts into your Word document.

   b. Make a double-bar chart that would inform the reader about violent crime rates in Illinois by region in 1983 and 2013. Include appropriate titles and axes labels AND an appropriately labeled legend. Paste this chart into your Word document.

   c. How do the violent crime rates in 1983 compare with the rates in 2013 in each of the regions?

   d. Do you think the double-bar graph or two single-bar graphs are a better way to graphically illustrate the violent crime rate trends? Explain.

3. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Heroin Deaths 1999-2013.xlsx, which contains data on national overdose deaths from heroin from 1999 to 2013.

   a. Make a scatterplot (select an option under scatterplot that still shows the data points, so the first or second option) of the data (be sure to include a title and appropriate labels for the x and y axes) and paste it into your Word document.

   b. Refer to your graph in (a) and describe each of the intervals below as increasing slowly, increasing quickly, relatively stable, decreasing slowly, or decreasing quickly.

      i. 1999-2006
      ii. 2006-2009
      iii. 2010-2013

Graphs are like jokes: If you have to explain them, they have failed. -Sally Bigwood
Chapter 5: Measures of Central Tendency (or Measures of Average)

**Objective:**
1. Define the various measures of central tendency / average (mean, median, and mode).
2. Compute mean, median, and mode of data sets by hand and in Excel.
3. Analyze the varying results of different measures of average, including determining which measure is more robust against outliers and how different measures of average can be used and misused to convey certain impressions.

The term *average* is one you hear often. For example “grade point average” and “average salary” both use the term *average*, but likely, they would use different types of computations to describe them.

There are three common used *averages* or *measures of central tendency* to describe a data set:

**Mean**
What most people are referring to when they say average.

To find the mean by hand:

\[
\text{mean} = \bar{x} = \frac{\text{sum of all data values}}{\text{number of items in data set}}
\]

**Example 1**
Grade point average is a mean, defined as: 

\[
\text{GPA} = \frac{\text{sum of grade pts}}{\text{number of credits}}
\]

Compute your GPA if you 1 A, 2 B’s, 1 C, and 1 D, where all the classes where 3 credits, except for the C, which was 4 credits. (Note: You earn 4 grade points for an A, 3 for a B, 2 for a C, and 1 for a D.)

**Median**
What is commonly referred to as the middle value.

To find the median by hand:

- Put the data in order from least to greatest.
- If there are an odd number of items in the data set, then the median is the middle value.
- If there are an even number of items in the data set, then it is the mean (average) of the middle two values.

**Example 2**
Often terms like “average salary” or “average price of a home” are using the median.
Unit 2: Statistics
Chapter 5: Measures of Central Tendency (or Measures of Average)

Compute the median price of a home on a block where homes are valued at: $185000, $208000, $197000, $585000, $221000, and $215000.

Mode
The most frequently occurring number(s) in a data set.

Example 3 Working at a shoe store and ordering sizes, you wouldn’t care about mean or median size, but you would want to know about the mode.

Compute the mode of the following recorded shoe sizes that were sold during a particular week:
5, 5.5, 6, 6.5, 6.5, 7, 7, 7, 7.5, 7.5, 8, 8, 8, 8, 8.5, 8.5, 8.5, 9, 9, 9.5, 9.5, 10, 10.5

As with many topics in this course, we want to know how to compute these values for real-life data sets, and real-life data sets can often be quite large. We can do this in Excel by using functions.

To find measures of central tendency in Excel:
- Go to an empty cell and type =
- Tap on the function tab, which looks like The function box should appear.
- In the category box, either leave the category as “All” or you can select “Statistical.” Either way, scroll down and select appropriate measure of central tendency shown below:
  - For Mean: Select AVERAGE in Excel
  - For Median: Select MEDIAN in Excel
  - For Mode: Select MODE.SNGL in Excel

Now, the dialogue box for appropriate function should appear.

- For all of these functions, in the “Number” box, highlight the cells.
  (For example, if you have values in A1 through A20, you should have A1:A20 in the “Number” box.)

Example 4 Consider the data set on the life expectancy of various countries (http://faculty.elgin.edu/nscherger/Courses/MTH104/Life Expectancy by Country.xlsx). Let’s compute the mean and median life expectancy of all of the countries listed in 2013.
Unit 2: Statistics
Chapter 5: Measures of Central Tendency (or Measures of Average)

Following the above directions, you should get to the following for mean and median:

Mean = _______________  Median = _______________

**Example 5** Consider the data set on the 2015 salaries of the Orlando City Soccer Club (http://faculty.elgin.edu/nscherger/Courses/MTH104/Orlando City Soccer Salaries 2015.xlsx).

a. First observe the data set. Looking at all of their salaries, if you were to describe the “average” salary of all the Orlando City Soccer Club players, what seems reasonable?

b. Now, as we did in the first example, let’s compute the mean and median.

   Mean =

   Median =

c. Which do you think is a better reflection of the “average” salary of this group? Explain.

   Here, the_____________ is a better reflection, because the majority of the players have salaries...

   In fact, when looking at the mean, there is only

d. Suppose MF Kaká has a phenomenal season and he gets a raise to $10 million (FYI – according to http://www.businessinsider.com/lionel-messi-highest-paid-soccer-player-footballer-2015-3, Barcelona’s Messi earned $39 million in salary from Barcelona and another $31 million in endorsements!).
Unit 2: Statistics
Chapter 5: Measures of Central Tendency (or Measures of Average)

Observe that the ___________ is unchanged, but the ___________ is now even less representative of the Orlando City’s ‘average’ salary.

What you are observing here was the effects of outliers.

**Outliers**
- Value(s) in a data set that is much higher or lower than most the other values.
- Outliers (especially in smaller data sets) can greatly impact statistics, especially measures of average.

<table>
<thead>
<tr>
<th>Key Idea</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ___________ is much more affected by outliers than the ___________.</td>
</tr>
<tr>
<td>Another way to say this is that the ___________ is more robust against the effects of outliers.</td>
</tr>
</tbody>
</table>

The impact of outliers can sometimes be used in the media. For example, suppose that a similar data set (although beit slightly less extreme) represented a particular group of government workers or teachers in a particular district. The media could report the *average* salary using the mean (which is not a lie); however, in reality, there may only be a small number of people who even make that amount.

The impact of small outliers can work in the opposite way. For example, suppose you are taking a class where your entire grade is based on an average of 5 exams. On the first 4 exams you earn: 84, 76, 81, and 79, which averages to 320/4 = 80, but then you completely miss the last exam and get a 0. Now, your average is 320/5, which is a 64!

Note: A method for computing outliers will be discussed in the next section.
Group Activity 7 Part 1: Measures of Central Tendency

   a. Calculate the mean salary.
   b. Calculate the median salary.

2. Suppose Derrick Rose had been paid $28,862,876 (instead of $18,862,876).
   a. What would the mean have been in that situation?
   b. What would the median have been in that situation?

3. Suppose Derrick Rose had been paid $28,862,876 and now Carlos Boozer had been paid $23,550,000 (instead of $13,550,000).
   a. What would the mean have been in that situation?
   b. What would the median have been in that situation?

4. Observing the changes in (2) and (3), which measure (mean or median) is more robust against (so, not as impacted by) the effect of outliers?

5. Can it happen in a dataset that almost every data point is above the mean? If it can, give an example. If not, explain why not.

6. Can it happen in a dataset that almost every data point is above the median? If it can, give an example. If not, explain why not.

   a. State the mean and the median (you already have from 1) for the Chicago Bulls.
   b. Find the mean and median for the LA Lakers.
   c. Based on the results from (a) and (b), compare the average (remember, mean and median are both measures of average) salaries of the Chicago Bulls and the LA Lakers.
Chapter 6: Measures of Spread (or Measures of Variation)

Objectives:
1. Compute various measures of spread / variation, including range, quartiles and the five-number summary mean, and interquartile range by hand and using Excel.
2. Answer questions from a given five-number summary.
3. Compute the standard deviation using Excel.
4. Predict data sets' comparative values of standard deviation.
5. Determine outliers of a data set.

Another way to describe a data set is to talk about how close the numbers are or how spread out they are. This concept is known as spread or variation. For example, two classes could both have a median test score of 80, but one class's scores could all be in the 70's and 80's and another class's scores could range from the 60's through the 90's.

Variation (or spread) describes how widely data values are spread out about the center and there are a number of ways to describe this spread:

Range
The simplest way to describe the spread or variation of a data set is by computing the difference between the largest and smallest value in a data set:

\[ \text{Range} = \text{maximum} - \text{minimum} \]

Example 1 If the values of a home on a block are $185000, $208000, $197000, $585000, $221000, and $215000, compute the range.

The range alone does not reveal much about a data set, especially if there were outliers.

A more complete reflection of the spread of a data set is using...
Quartiles and the Five-Number Summary

The five-number summary divides the data set into four quarters by finding the following:

1. Minimum – smallest number in the data set
2. First Quartile ($Q_1$) – median of the lower half of data
3. Median (or Second Quartile) – middle value in the data set
4. Third Quartile ($Q_3$) – median of the upper half of data
5. Maximum – largest number in the data set

<table>
<thead>
<tr>
<th>25% of data</th>
<th>25% of data</th>
<th>25% of data</th>
<th>25% of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum value</td>
<td>$Q_1$</td>
<td>$Q_2$</td>
<td>$Q_3$</td>
</tr>
</tbody>
</table>

In a five-number summary, 25% of the data falls within in each quartile. Specifically, 25% of the data falls between the min and $Q_1$; 25% falls between $Q_1$ and the median; 25% falls between the median and $Q_3$; 25% falls between $Q_3$ and the max.

**Finding the 5-Number Summary in Excel**

- Go to an empty cell and type =
- Tap on the function tab, which looks like $\text{fx}$
  - The function box should appear.
- In the category box, either leave the category as “All” or you can select “Statistical.”
  - Either way, scroll down and select appropriate measure shown below:
    - For Minimum: Select MIN in Excel
    - For First Quartile ($Q_1$): Select QUARTILE.EXC in Excel
    - For Median: Select MEDIAN
    - For Third Quartile ($Q_3$): Select QUARTILE.EXC in Excel
    - For Maximum: Select MAX in Excel
- Now, the dialogue box for appropriate function should appear.
- For the Min, Max, or Median functions, in the “Number” box, highlight the cells.
  - For the Quartile command, in the “Array” box, highlight the cells and in the “Quart” box, enter the appropriate number (1 for 1st quartile and 3 for 3rd quartile).
Unit 2: Statistics
Chapter 7: The Normal Distribution

Example 2  Compute the five-number summary and the range for the data set on the 2015 salaries of the Orlando City Soccer Club (http://faculty.elgin.edu/nscherger/Courses/MTH104/Orlando City Soccer Salaries 2015.xlsx).

Minimum: _____

$1^{st}$ Quartile ($Q_1$): ______

Median: _____

$3^{rd}$ Quartile ($Q_3$): ______

Maximum: ______

Range: ______

Example 3  Suppose that the following Five-Number Summary represents the salaries of teachers in a particular school district.

Minimum: $34,500

$1^{st}$ Quartile ($Q_1$): $43,300

Median: $55,400

$3^{rd}$ Quartile ($Q_3$): $79,520

Maximum: $105,900

a. What percentage of salaries are less than $43,300?

b. What percentage of salaries are between $43,300 and $79,520?

c. If there are 154 teachers in this school district, how many make more than $55,400?

d. What percent of teachers make more than $65,000?

e. What percent of teachers make less than $34,000?
Interquartile Range (IQR)

The IQR gives the range of the middle 50% of the data and is found by computing the difference between the third and first quartiles:

\[ IQR = Q_3 - Q_1 \]

Example 4 Compute the interquartile range (IQR) of teacher salaries from example 3.

Outliers

Now that we have discussed quartiles and IQR, we can discuss a method for computing outliers.

There is not hard and fast rule for determining outliers; however, a common practice is if a data point is more than the \( Q_3 \) plus 1.5 times the IQR or if a data point is less than the \( Q_1 \), minus 1.5 times the IQR it can be called an outlier.

So, we will define an outlier to be:

- any data point less than \((Q_1 - 1.5\times IQR)\) \(\text{[lower cutoff for outliers]}\)
- or
- any data point more than \((Q_3 + 1.5\times IQR)\) \(\text{[upper cutoff for outliers]}\)

Example 5 Using this above definition, determine the lower and upper cutoffs for outliers and then identify any outliers in the data set on Orlando City’s salaries.

Boxplots

One handy way to summarize a set of data is a boxplot. It uses the five-number summary to create a “box” in the middle, with what are often called “whiskers” at the edges. Boxplots are also commonly modified to indicated possible outliers.

Note: Boxplots are difficult to render in Excel, so we’ll only be asking you to read and interpret them.
Standard Deviation

This is the most common single-number that is used to describe the spread/variation of a data set. For this course, you will not have to compute the standard deviation by hand. Instead, we will only use Excel to compute it. It is more important that you understand the concept of standard deviation.

The standard deviation is a measure of how far the data is spread around the mean. Roughly, for our purposes, it can be thought of as the average of all of the differences from the mean, or the average distance a data point is from the mean.

Finding the Standard Deviation in Excel

- Go to an empty cell and type =
- Tap on the function tab, which looks like $\text{f}$.
The function box should appear.
- In the category box, either leave the category as “All” or you can select “Statistical.” Either way, scroll down and select appropriate measure of standard deviation shown below:
  
  STDEV.P – if your data set represents the entire population of data
  
  (for example, if we are computing the standard deviation of the Orlando City’s salaries to describe the spread of the Orlando City’s salaries themselves)

  OR

  STDEV.S – if your data set represents a sample from the larger population of data

  (for example, if we are computing the standard deviation of the Orlando City’s salaries as a way of estimating the spread of all US Soccer Club salaries)

Example 6

Compute the standard deviation of the of the Orlando City’s salaries.

Thus far, we have been analyzing the Orlando City’s salaries (by computing various measures of central tendency and spread) as a way of describing and discussing this particular soccer club. There has been no discussion of using this data as a sample to make a prediction for some larger population. Thus, here we will use ___________ to compute the standard deviation.

Thus, the standard deviation is
Example 7 Again, let’s suppose that MF Kaká has a phenomenal season and he gets a raise to $10 million. Compute the standard deviation for this data set.

Note: It should make sense that the standard deviation in the second example is larger than the standard deviation in the first set, because the second data set was more spread out.

Key Idea
The larger the standard deviation, the more spread out the data set is; the smaller the standard deviation, the more tightly clustered the data.

Example 8 The box plots below represent the five-number summaries for the test scores of two different classes. Which class would you predict would have the larger standard deviation? Why?

And finally, an appropriate comic to end the section:

Image source: /xkcd.com/539/
Group Activity 7 Part 2: Measures of Spread

   a. Find the range and standard deviation for the Chicago Bulls.
   b. Find the range and standard deviation for the LA Lakers.
   c. Based on the results from (a) and (b), compare the spread (remember, range and standard deviation are both measures of spread) of salaries of the Chicago Bulls and the LA Lakers.

9. Now, let’s look at their 5-Number Summaries.
   a. Compute the 5-Number Summary for the Chicago Bulls.
   b. Compute the 5-Number Summary for the LA Lakers.
   c. Based on the results from (a) and (b), does the team that you identified as having the greater spread in #8c also appear to have a greater spread when you look at their 5-Number Summaries?

10. Before mathematically computing any possible outliers, discuss with your group and state any players who think are possibly outliers on each team.

11. Using the formula for the cutoffs for outliers in the previous section of notes, determine the following.
   a. Outliers among the Chicago Bulls would be any salaries less than what value (in other words, what is the lower cutoff for outliers)?
   b. Outliers among the Chicago Bulls would be any salaries greater than what value (in other words, what is the upper cutoff for outliers)?
   c. Based on the results of (a) and (b), are any players outliers on the Chicago Bulls? If so, who?
   d. Outliers among the LA Lakers would be any salaries less than what value (in other words, what is the lower cutoff for outliers)?
   e. Outliers among the LA Lakers would be any salaries greater than what value (in other words, what is the upper cutoff for outliers)?
   f. Based on the results of (d) and (e), are any players outliers on the LA Lakers? If so, who?
12. Old Faithful geyser in Yellowstone National Park is pictured at the right. This geyser erupts about every hour with some consistency, hence its name. The file contains data on the length of the eruption and the interval between eruptions.

Use the given 5-Number Summary (below) on the time between eruptions (in minutes) below to answer the following questions, if possible.

Minimum = 42
1st Quartile (Q_1) = 60
Median = 75
3rd Quartile (Q_3) = 81
Maximum = 95

a. What percent of the intervals between eruptions take less than 1 hour and 15 minutes?
b. What percent of the intervals between eruptions take more than 1 hour?
c. What percent of the intervals between eruptions take more than 1½ hours?

13. Consider the boxplot shown below for the percent of students who are college ready in each Illinois high school.

a. Use the graph to estimate the 5-number summary.
b. Use your estimates to calculate the lower and upper cutoffs.
c. Complete this sentence: The middle 50% of all Illinois schools have between _____ and _____ of their seniors qualifying as college ready.
Assignment 4: Measures of Average and Spread

Complete Assignment #4 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Age At Inauguration.xlsx.
   a. Find the mean age at inauguration.
   b. Find the median age at inauguration.
   c. Find the average spread of the ages about the mean. (You should be using the statistic here that is interpreted as “the average spread about the mean.”)
   d. Donald Trump was 70 at inauguration. About how many standard deviations is he above the mean?

2. Read the comic.

To help explain the joke, suppose there are 5 salaries in Mr. Stingy’s company: 4 workers and himself. Give an example of what all 5 salaries could be for the median salary to be less than $40,000 and mean salary to be over $90,000.

3. Suppose that the five-number summary for salaries of full-time teachers in school districts A and B are given below. Determine the following, if possible.

   School District A:
   Minimum: $37,500
   1st Quartile: $45,300
   Median: $53,400
   3rd Quartile: $67,120
   Maximum: $92,900

   School District B:
   Minimum: $38,500
   1st Quartile: $42,200
   Median: $61,400
   3rd Quartile: $80,120
   Maximum: $102,200

   a. What percentage of teachers in district A make less than $53,400?
   b. What percentage of teachers in district B make between $42,200 and $80,120?
   c. What percentage of teachers in district B make less than $50,000?
   d. If there are 108 teachers in district A, how many teachers make more than $67,120?
   e. Determine the ranges of salaries for each school district.
   f. Determine the interquartile ranges of salaries in each school district.
   g. While you cannot compute the standard deviations exactly without the entire data set, which district do you think would have the larger standard deviation? Explain.
Chapter 7: The Normal Distribution

Objectives:
1. Identify characteristics and properties (including the 68/95/99.7 rule) of the normal curve.
2. Give examples of normally distributed data.
3. Solve normal distribution problems by hand, finding both probabilities/percentages and data values, utilizing the 68/95/99.7 rule.
4. Use Excel to solve normal distribution problems, finding both probabilities/percentages and data values, that cannot be determined from 68/95/99.7 rule.

The Normal / Bell-Shaped Curve: Basic Idea
Suppose we went out and obtained the IQ scores of 40 randomly selected adults and then made a histogram of the data and obtained the following:

Suppose we went out and obtained the IQ scores of 80 randomly selected adults and then made a histogram of the data and obtained the following:
Finally, suppose we went out one last time and obtained IQ scores of 220 people and then made a histogram of the data and obtained the following:

Observe that as we increase the sample size, the histograms more closely resembles a bell-shaped curve, which is what we call the normal curve.

(Note: The data in the above graphs, although close to realistic, was not actual data.)

**Example 1**

Here’s an example of a histogram, which displays the weights of seventh graders. The histogram can be observed to be approximately bell-shaped and is overlayed with a bell-shaped / normal curve.

Although we have not yet discussed characteristics of the normal curve, what would you predict is the average weight of this group of seventh graders?

**Example 2**

The next graph at the right illustrates that the normal curves can take on different shapes and still be bell-shaped / normal.

Again, although we have not yet discussed characteristics of the normal curve, how would you describe the differences in these graphs? (Use terms such as mean and standard deviation.)
Normal Distribution: Definitions and Properties

- The **normal distribution** is the symmetrical bell-shaped distribution with a single peak.
- This peak is the mean, median and the mode of the distribution.
- Data in a normal distribution is more clustered around the mean and data that is further away is less common, which gives the **tails**.
- The reason we are talking about normal curves is because so many phenomena are normally distributed. Here are just a few examples:
  - heights, weights, and many biological measures
  - batting averages, quarterback ratings, and many sports statistics
  - IQ scores, SAT scores, and most standardized tests

Beyond the basic definition and conditions of a normal distribution, the primary properties of a normal distribution that you will need to know are encompassed in the 68/95/99.7 Rule.

**68/95/99.7 Rule for a Normal Distribution**

68% of all data falls within one standard deviation of the mean:

95% of all data falls within two standard deviations of the mean:

99.7% of all data falls within three standard deviations of the mean:

Now, also recall that 50% of the data lies above and below the mean. Combine this symmetry with the 68/95/99.7 rule and we can fill in the percentage of data in each group below:
Now, add the above percentages to obtain the cumulative percentiles:

**Key Idea**
The above chart’s values will remain the same in EVERY normal distribution.

**Example 3** Suppose that the heights of American women are normally distributed with a mean of 65 inches (5 feet 5 inches) and a standard deviation of 2.5 inches. Fill in the boxes on the normal curve below and answer the questions that follow.
Note: For questions involving the normal distribution, it is very beneficial to make a sketch of the desired region. Some of these sketches will initially be included, but in cases where they are not, you need to make the sketch. (This practice will be especially helpful as examples get more difficult.)

a. If a woman is 5 feet 10 inches, what percentile is she?

b. What is the height of a woman in the 16\textsuperscript{th} percentile?

c. If a woman is randomly selected, what is the probability that she will be taller than 5 feet?
Example 4 IQ scores are normally distributed with a mean of 100 and a standard deviation of 15. Fill in the boxes on the normal curve below and answer the questions that follow.

a. What percent of people would you expect to have an IQ between 70 and 130?

b. What is the IQ of someone who is in the 84th percentile?

c. What is the probability that a randomly selected person will have an IQ more than 115?
Unit 2: Statistics  
Chapter 7: The Normal Distribution

d. In a group of 280 people, how many would you expect to have an IQ less than 70?

![Graph](image)

e. What percentile is someone who has an IQ of 110?

![Graph](image)

**Solving normal distribution problems that can’t be answered exactly from the 68/95/99.7 rule**

For questions like those in (e), where the answer cannot be determined exactly from 68/95/99.7, one can either:

1. Use the *standardized normal distribution* (which has a mean of 0 and standard deviation of 1) by converting the data values to *standardized scores*, called *z-scores* (using the formula using $z = \frac{x-\mu}{\sigma}$), and then one can use statistical tables (found in the appendices of most statistics books) for the *standardized normal distribution*.

For example, in the normal distribution for IQ, then mean of 100 is converted to 0 and the standard deviation of 15 is converted to 1. Then, IQ scores are converted to *z-scores*. For example, an IQ of 115 would be a *z-score* of $(115-100)/15 = 1$ and an IQ of 70 would be *z-score* $(70-100)/15 = -2$.

To move beyond this basic introduction the *standardized normal distribution* and *z-scores*, it would take much more discussion and formulas to be able go further. You should simply be aware of this concept.

2. Use Excel functions, which is what we will be using:

**Find the probability less than a given score in a normal distribution**

- Go to an empty cell and type `=`.  
- Tap on the function tab, which looks like ![Function](image).  
  The function box should appear.  
- In the category box, either leave the category as “All” or you can select “Statistical.”  
  Either way, scroll down and select *NORM.DIST*.  
- In the dialogue box that appears, enter the required information (always entering “TRUE” in the last entry).

**Key Idea**

Remember that the **percentile** of a certain value is equivalent to the probability that a randomly chosen individual is *less than* that value.
**Example 4** continued... (IQs, which are normally distributed where the mean is 100 and the standard deviation is 15)

f. What percentile is someone in who has an IQ of 110?

Following the directions above, the result is:

![Function Arguments](image)

So, an IQ of 110 is the ___________ percentile.

g. If a person is randomly selected, what is the probability that their IQ is between 83 and 103?

h. What percentile is Albert Einstein, if his IQ was 160?

To find the Score that corresponds to a given probability (percentile) in a normal distribution in Excel:

- Go to an empty cell and type =.
- Tap on the function tab, which looks like $\text{fx}$.
  The function box should appear.
- In the category box, either leave the category as “All” or you can select “Statistical.” Either way, scroll down and select `NORM.INV`.
- In the dialogue box that appears, enter the required information (always entering the percentile as a cumulative probability; for example, to find the 35th percentile, enter 0.35).
Example 4 continued... (IQs, which are normally distributed where the mean is 100 and the standard deviation is 15)

i. What is the IQ score of someone in the 75th percentile?

Following the directions above, the result is:

![Graph showing the normal distribution and the 75th percentile]

So, the IQ of someone in the 75th percentile is __________.

j. What is the interquartile range (IQR) of IQ scores?

Recall that the IQR = Q₃ – Q₁ and we already know Q₃ = ________

So, we need to find Q₁ = ________

Thus, the interquartile range = Q₃ – Q₁ = ________

k. According to [http://www.us.mensa.org/learn/about/](http://www.us.mensa.org/learn/about/), to qualify for Mensa, an individual must be in the top 2 percent of the general population on an accepted standardized intelligence test. Thus, what would be the cutoff IQ score to join Mensa?

![Graph showing the normal distribution and the 2nd percentile]
Group Activity 8: Normal Distribution

1. For a particular model of car, the fuel economy obtained in miles per gallon (mpg) in highway driving conditions follows a normal distribution with mean 28 mpg and standard deviation 1.5 mpg. Use this information to answer the following questions. Complete all boxes below for this example.

Complete all boxes below for this example.

For 2-4, use the completed table in (1) above (NOT Excel), answer the following questions, if possible. (Note: You may not be able to answer all these questions.)

2. What percent of all of these models produced will get
   a. less than 26.5 mpg on the highway?
   b. at least (so, more than) 29.5 mpg on the highway

3. You just purchased this model of car.
   a. The probability is 0.84 that you will get less than what fuel economy on the highway?
   b. What is the probability that your car will get between 26.5 mpg and 31 mpg on the highway?

4. A car dealer just purchase 25 of these cars. How many of these cars can be expected to get
   a. less than 26.5 mpg on the highway?
   b. between 26.5 mpg and 29.5 mpg on the highway?
   c. at least 26 mpg on the highway?

Use Excel to answer the questions 5-8.

5. Suppose that a group of science majors’ GPAs have an approximate the normal distribution with mean 2.97 and standard deviation 0.34. Use Excel to answer the following questions.
   a. What percent of these science majors have GPAs less than 3.25?
   b. What percent of these science majors have GPAs less than 2.75?
Unit 2: Statistics
Chapter 7: The Normal Distribution

c. Use the results of (a) and (b) to determine what percent of these science majors have GPAs between 2.75 and 3.25.
d. Suppose this school is highly competitive and they automatically tell that bottom 5% they cannot return after the first year. All students with GPAs less than what value can be expected to not return?

6. The length of elephant pregnancies from conception to birth varies according to a distribution that is approximately normal with mean 525 days and standard deviation 32 days.
   a. What percent of pregnancies last between 510 and 540 days?
   b. How many days do the shortest 10% of all pregnancies last?
   c. What percentage (expressed as a percentage with two decimals) of pregnancies last more than 600 days?

7. Wingspans of adult herons have approximate normal distribution with mean 125 cm and standard deviation 12 cm.
   a. What is the median wingspan?
   b. What are the first and third quartiles of the wingspans? (Recall: the first quartile is the 25\text{th} percentile and the third quartile is the 75\text{th} percentile.)
   c. Use the results from (b) to obtain the IQR (Inter-Quartile Range) for heron wingspans.

8. When robins’ eggs are weighed, it turns out that they vary according to approximately the normal distribution with mean 17.5 mg and standard deviation 3.5 mg.
   a. What percent of robins’ eggs weigh between 10 mg and 25 mg?
   b. How large are the largest 2% of all robins’ eggs?
Chapter 8: Correlation & Causation

Objectives:
1. Estimate the correlation coefficient of scatterplots that exhibit positive, negative, or nearly no correlation.
2. Find the correlation coefficient in Excel.
3. Observe the effect of outliers on correlation.
4. Discuss the relationship between correlation and causation.

Correlation

Correlation measures the strength and direction of a relationship between two variables. Scatterplots are a good way to visually make an educated guess as to the possible correlation between two variables. Visit http://digitalfirst.bfwpub.com/stats_applet/stats_applet_5_correg.html and create some different scatterplots and see if you can make any conclusions about the value of the correlation coefficient (what we will later call r) for your various scatterplots.

Here are some samples:
Here are some more scatterplots (from http://catalog.flatworldknowledge.com/bookhub/reader/3318?e=fwk-shafer-ch10_s02), along with their correlation coefficient ($r$).

Conclusions/Basic Facts about Correlation:
The correlation coefficient $r$ tells whether or not there is a linear relationship between two variables.

- $r$ is always between $-1$ and $1$.
- A positive $r$ indicates a positive linear relationship between the variables and a negative $r$ indicates a negative linear relationship.
- An $r$ value of $-1$ or $1$ is only in the case of a perfect linear relationship.
- The closer the $r$ value is to $-1$ or $1$, the stronger the linear association.
- An $r$ value of $0$ indicates no linear association.

Key Idea

The correlation coefficient $r$ tells whether or not there is a linear relationship.

Caution: “Curved” associations are not reflected in $r$, no matter how strong.
How close does $R$ need to be to 1 or -1?

A natural question is how close the value needs to be to 1 or -1 to be significant. Unfortunately, there is not a clear answer. It depends on many factors. For example, if there are only a few data points, then one needs a value closer to 1 or -1 to conclude that there is a correlation. Often, natural sciences can control variables more easily and therefore have higher standards than in the social sciences. Here are some “crude” guidelines:

-1.0 to -0.9 Very strong, negative linear correlation
-0.9 to -0.7 Strong, negative linear correlation
-0.7 to -0.4 Moderate, negative linear correlation
-0.4 to -0.2 Weak, negative linear correlation
-0.2 to 0.0 Negligible, negative linear correlation
0.0 to 0.2 Negligible, positive linear correlation
0.2 to 0.4 Weak, positive linear correlation
0.4 to 0.7 Moderate, positive linear correlation
0.7 to 0.9 Strong, positive linear correlation
0.9 to 1.0 Very strong, positive linear correlation

Example 1 Consider the data on murder and nonnegligent manslaughter and percent of households with guns in the US (http://faculty.elgin.edu/nscherger/Courses/MTH104/US murder and percent of houses with guns.xlsx). We can use Excel to make a scatterplot and to get the correlation coefficient $r$.

First, making the scatterplot, we would see:

From the graph, the prediction would be that while not overwhelming, there appears to be somewhat of a positive correlation between the two variables.

To find the Correlation Coefficient ($r$) between two data sets in Excel:

- Go to an empty cell and type =.
- Tap on the function tab, which looks like . The function box should appear.
- In the category box, either leave the category as “All” or you can select “Statistical.” Either way, scroll down and select CORREL.
- In the dialogue box that appears, highlight the first column of data for Array 1 and highlight the second column of data for Array 2.
Example 1 continued...

Now, computing the correlation coefficient:

![Function Arguments](image)

So, the correlation is $r = \text{________}

*(Discussion to follow under correlation and causation)*

Example 2 Consider the following data on the number of cigarettes smoked and the birthweight of a child *(http://faculty.elgin.edu/nscherger/Courses/MTH104/Cigarettes Birthweight.xlsx)*.

Again, first, complete a scatterplot, as shown:

From the graph, the prediction would be that there is a negative correlation between the two variables.

Next, as before, find the correlation coefficient, where it is discovered that $r = \text{________}$.

*(Discussion to follow under correlation and causation)*
Correlation and Causation

Example 1 continued...

Consider our first example. Do you think an increase in the number of households who have guns cause more crime? If yes, explain your reasoning. If no, what could be some other variables at play here?

Example 2 continued...

How about our second example. Does an increase in cigarettes/day lead to smaller babies? In particular, does more cigarettes/day lead to smaller babies on a per gestation day basis?

Further study of this correlation has shown that what is really happening is that smoking mothers tend to have higher birthweight babies on per gestation day basis, but smoking mothers' children tend to be born earlier.

Example 3 Let’s consider one more example. The following data is on mean annual temperature of a region and mortality rate due to breast cancer (http://faculty.elgin.edu/nscherger/Courses/MTH104/Breast Cancer Temperature.xls).

By plotting the data and computing the correlation coefficient, which is $r = 0.88$, we can see that there is a positive correlation.

Do you think a warmer climate is a likely cause for more breast cancer fatalities? If yes, explain your reasoning. If no, what could be some other variables at play here?

Example 4 While at your regular dental checkup, your dentist tells you that daily flossing is positively correlated with happiness. Do you think there is a reasonable case for causation here? That is, does the act of daily flossing cause people to be happy? If yes, explain your reasoning. If not, identify a lurking or confounding variable(s).
Correlation and Causation Possibilities

Consider the following possibilities that have correlation (shown by a dashed line) between variables A and B.

- **Causality:** One variable, say A, actually causes the change in B.
  
  This can be depicted in the following diagram where the yellow arrow represents causality and the dotted line represents an observed correlation.

  EX: Smoking has been shown to *cause* more premature births.

- **Sheer Coincidence:** A and B really don't have anything to do with each other but happen to go up or down simultaneously.

  EX: US spending on science, space, and technology correlates strongly with suicides by hanging. (You can see the graph at [http://www.tylervigen.com/spurious-correlations](http://www.tylervigen.com/spurious-correlations) and it shows $r=0.9979$).

- **Common underlying cause or causes:** A is correlated with B, but there is a third factor C (the common underlying cause) that causes the changes in both A and B. (C is sometimes called a “lurking” variable.)

  EX: It has been shown that there is a positive correlation between ice cream sales and crime rates. As ice cream sales go up, crime rates go up. Is there a yet to be discovered compound in ice cream that cause people to commit crime? Do people suffering from high crime rates seek succor in the comfort food ice cream? Clearly not. So, what is the lurking variable here?

- There is one other similar scenario to mention, which is "confounding" variables.

  What most authors mean by "confounding" is something like:

  A is correlated with B and you suspect A causes B, but C is also correlated with B, and C is correlated with A too. In this case, the effects of A and C are confounded: you can’t separate them out.

  EX: Parents’ education level (A) is correlated with students’ performance on standardized exams (B). Student performance (B) is correlated with family income (C) as well. But parents’ education level (A) is confounded with family income (C), because parents with more education tend to earn more and vice versa. The two variables are confounded and can’t be separated out without further study.

**Key Idea**

**CORRELATION DOES NOT IMPLY CAUSATION!**

There might be (and often is) underlying lurking/confounding variables.
Effects of Outliers

The effect of an outlier can be surprisingly significant on correlation, especially if the data set is small. For example, consider the following two scatterplots, noting the difference just one point can make!

Points: 6
Correlation coefficient: 0.9710

Points: 7
Correlation coefficient: 0.2939
Unit 2: Statistics
Chapter 8: Correlation & Causation

Group Activity 9: Correlation

1. Determine if the following data sets would have a positive correlation, a negative correlation, or no correlation (so, correlation close to 0). (Note: This exercise is referring to the correlation coefficient, r.)

   a. 
   
   ![Graph A](image)

   b. 
   
   ![Graph B](image)

   c. 
   
   ![Graph C](image)

2. The following statements indicate a real correlation between two variables. Do you think that these show a reasonable case for causation as well? If you answer yes, explain your reasoning. If you answer no, give an example of a reasonable “lurking” or “confounding” variable. So, start all your responses with ONE of the following:

   Yes, I think there is a reasonable case for causation, because...
   ~OR~
   No, I do not think there is a reasonable case for causation. The most likely lurking/confounding variable is....

   a. As the fall semester of classes is coming to an end in December, two ECC students get jobs at a local store. After the ECC students are hired, the store’s sales increase dramatically! Did the hiring of the ECC students cause the increase in sales?

   b. An insurance company collected data on the average speed of its drivers and the number of accidents and found a positive correlation. Does increased speed cause more accidents?
Unit 2: Statistics
Chapter 8: Correlation & Causation

3. Open the file, http://faculty.elgin.edu/nscherger/Courses/MTH104/MTH 096 Hmwk and Course Grades.xlsx, which contains actual grade data from my MTH 096 class from the spring 2015. It gives each student’s end-of-semester homework average and their final course average. (Please note that homework is only 10% of their course grade.
   a. Obtain a scatterplot of the MTH 096 students’ course grades versus their homework averages and paste it into your Word document.
   b. Does it look like there will be a positive correlation, a negative correlation, or no correlation (so, correlation close to 0)?
   c. Determine the actual correlation coefficient, r, between MTH 096 students’ course grades and their homework averages.
   d. According to the standards in the last section of notes, how would we classify this correlation?
   e. Do you think there is a likely case for causation here? In other words, does working on and doing well on algebra homework cause greater success? As in 2, start your response with ONE of the following:

   Yes, I think there is a reasonable case for causation, because...
   ~OR~
   No, I do not think there is a reasonable case for causation. The most likely lurking/confounding variable is....

Assignment 5: Normal Distribution and Correlation & Causation

Complete Assignment # 5 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Life Expectancy and Percent Urban by Country.xlsx. The file contains data on countries’ life expectancy and the percent of their population that is urban.
   a. Obtain a scatterplot of the countries’ life expectancy and the percent of their population that is urban. Include an appropriate title and axes labels. Copy and paste this plot into your Word document.
   b. Does it look like there will be a positive correlation, a negative correlation, or no correlation (so, correlation close to 0)?
   c. Determine the actual correlation coefficient, \( r \), between countries’ life expectancy and the percent of their population that is urban.
   d. According to the standards in our notes, how would we classify this correlation?

2. The following statements indicate a real correlation exists between two variables. Do you think that these show a reasonable case for causation as well? If you say yes, explain your answer. If you say no, then give at least one possible “lurking” or “confounding” variable. So, start all your responses with ONE of the following:

   Yes, I think there is a reasonable case for causation, because…
   ~ OR ~
   No, I do not think there is a reasonable case for causation. The most likely lurking/confounding variable is….

   a. A health club went under new management after the start of the year. During their first month under new management, the sales of new memberships soared! Did the new management cause the increase in sales?
   b. There is a negative correlation between amount children’s screen time (television, video games, phones, computers, tablets, etc..) and length of their attention span. (That is, the greater their screen time, the shorter their attention span.) Does screen time cause reduced attention span?

Carl Friedrich Gauss, who is most commonly credited with the bell-shaped curve, appeared on the German ten-mark banknote from 1989 to the end of 2001 in his honor.
This second test covers material presented in Chapters 5-8.

The following represents a list of the major topics you should be able to do.

1. Distinguish between categorical and numerical data.
2. Create pie charts, when appropriate.
3. Critique pie charts, including identifying situations when it is not appropriate to create a pie chart.
4. Create bar charts (including multiple bar charts) and scatterplots / line graphs.
5. Critique bar charts and scatterplots / line graphs, including identifying specific features of how a graph can be altered to alter its appearance to convey a particular impression.
6. Find measures of average (mean, median, and mode) of a data set, and identify scenarios when one measure of average may be more appropriate than another.
7. State which measure of average is more robust against outliers and which is not and explain why.
8. Find measures of spread (range, five-number summary, interquartile range, and standard deviation).
9. Answer questions from a given five-number summary.
10. Explain that the standard deviation represents the “average spread about the mean” or the “average distance a data point is from the mean” and be able to make predictions about how two data sets standard deviation would compare, based on how spread out the data is.
11. State and apply the 68/95/99.7 rule to answer questions about the normal distribution.
12. Given the percentiles that result from the 68/95/99.7 rule, answer questions about a given normal distribution.
13. Use Excel to answer questions about a normal distribution.
14. Identify positive correlation, negative correlation or no correlation from a graph.
15. Critically distinguish between correlation and causation in data.
1. Critique the given bar graph from the Conservative Party of Britain (Source: https://twitter.com/Conservatives/status/407945493700812800/photo/1).

2. The five-number summaries for two different classes’ final exams are given below. Answer the questions that follow as precisely as possible given the information below.

Class 1 Final Exam 5-Number Summary:
- Minimum: 45
- 1st Quartile: 60
- Median: 70
- 3rd Quartile: 79
- Maximum: 87

Class 2 Final Exam 5-Number Summary:
- Minimum: 65
- 1st Quartile: 72
- Median: 79
- 3rd Quartile: 85
- Maximum: 92

   a. What percent of students in Class 1 scored below 70?

   b. What percent of students in Class 2 scored below 70?

   c. What percent of students in Class 2 scored above 85?

   d. 75% of all students in Class 1 scored over what value?
Unit 2: Statistics
Unit Review

e. What are the range and interquartile range for Class 1?

f. What are the range and interquartile range for Class 2?

g. Based on the given information, which class do you think would the larger standard deviation? Explain your reasoning.

3. A recent study by the Center for Research on Media, Technology, and Health at the University at Pittsburgh found that people who report spending the most time on social media also have an increased perception of social isolation.
   a. Do you believe there is a likely case for causality? *(That is, does using social media more cause social isolation?)*
      
      YES or NO
   
   b. If you answered yes, explain your opinion and reasoning for possible causality. If you answered no, state one possible confounding or lurking variable.
4. Suppose that SAT scores are normally distributed with a mean of 500 and a standard deviation of 100.
   a. 68% of all students will score between what two values?

   b. 95% of all students will score between what two values?

   c. 99.7% of all students will score between what two values?

   d. Fill in the data in the table below.

   For parts (e) through (h), answer using the table above.

   e. An SAT score of 400 is in what percentile?

   f. What percent of SAT scores is above 300?

   g. What percent of SAT scores are between 300 and 500?

   h. 16% of all SAT scores are above what score?
Use Excel to Answer the Remaining Questions

5. The bar graph at the right it is not an accurate depiction of the data.
   a. Enter the following data into Excel and create an appropriate bar graph. Be sure to include a title and axis labels. Copy and paste this bar graph into your Word document. (2 pts)

<table>
<thead>
<tr>
<th>Years</th>
<th>Number of Barrels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970-1980</td>
<td>3,600,000</td>
</tr>
<tr>
<td>1990-2000</td>
<td>440,000</td>
</tr>
<tr>
<td>2000-2004</td>
<td>1,000</td>
</tr>
</tbody>
</table>

b. What do you now notice was wrong with the appearance of given bar graph? (So, what about your correct graph looks different?)

c. Can you identify one other issue with the original graph that makes it a bad graph? (Hint: Look at the years.)

6. Use the file titled “US alcohol related traffic fatalities over time” for the following.
   a. Make a scatterplot of the data, an appropriate chart title and labels on your axes. Copy and paste this graph into your Word document.

   b. What is the absolute maximum and absolute minimum (both the “when” and “what” for these points)?

   c. Observe that the graph is generally decreasing over time.
      i. During what years was the graph decreasing quickly? (Your answer should be roughly a 10 year time span)
      ii. During what years was the graph decreasing slowly? (Your answer should be roughly a 10 year time span)
7. Open the file entitled “Old Faithful” and determine the following.
   
   a. The mean number of minutes (interval) between eruptions.
   
   b. The median number of minutes (interval) between eruptions.
   
   c. The standard deviation of the number of minutes (interval) between eruptions.

8. Suppose that SAT scores are normally distributed with a mean of 500 and a standard deviation of 100. Use Excel to determine the following.
   
   a. What percentile is a score of 550?
   
   b. What percent of scores would be over 475?
   
   c. If somebody scored in the 80th percentile, what is their SAT score?
   
   d. The top 5% of scorers have a SAT score that is over what value?
Chapter 9: Hypothesis Testing

Objectives:
1. Define and identify null hypotheses and alternative hypotheses.
2. Determine the four outcomes of hypothesis testing.
3. Given an alpha level of significance and a p-value of a study, determine if the null hypothesis is rejected or failed to be rejected.
4. Identify type I and type II error.

Hypotheses

A hypothesis is a proposition that's being investigated that is yet to be proved.

It is an assumption about 1 or more of the population parameters that will be accepted or rejected on the basis of information from a sample from that population.

This process is called hypothesis testing.

Two Hypotheses

- The null hypothesis (H₀) is the hypothesis being tested and it assumes no relationship between the quantities. It’s the “is no effect” hypothesis.
- The alternative hypothesis or research hypothesis (Hₐ) is the hypothesis accepted if the null hypothesis is rejected. It’s the “is effect” hypothesis.

Example 1

Suppose we are interested in finding out if there is a difference in effectiveness in 2 brands of cold medicine, X and Z. Identify the null and alternative hypotheses.

Null hypothesis (H₀):

There ______________________________ in the effectiveness between brand X and brand Z.

Alternative hypothesis (Hₐ):

There ______________________________ in the effectiveness between brand X and brand Z.

Four Outcomes of Hypothesis Testing

There are 4 possible outcomes in hypothesis testing:

1. The null hypothesis is actually true and is not rejected. 😊
2. The null hypothesis is actually true and is rejected. 😞 This is a type I error.*
3. The null hypothesis is actually false and is rejected. 😞
4. The null hypothesis is actually false and is not rejected. 😊 This is a type II error.

*The type I error is typically the error we are most concerned with, because when a null hypothesis is rejected, often some sort of action is implemented.
Table representation of the 4 outcomes of hypothesis testing:

<table>
<thead>
<tr>
<th>REALITY</th>
<th>( H_0 ) is actually TRUE</th>
<th>( H_0 ) is actually FALSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 ) REJECTED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0 ) NOT REJECTED</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example 2: Consider the 4 outcomes of hypothesis testing, with the example of a court case, where

\( H_0: \)

\( H_A: \)

Fill the box in with the comparison between what is actually true and what a judge or jury decided.

<table>
<thead>
<tr>
<th>REALITY</th>
<th>( H_0 ) is actually TRUE: Defendant is actually________</th>
<th>( H_0 ) is actually FALSE: Defendant is actually________</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 ) REJECTED: Defendant found __________</td>
<td>TYPE I ERROR</td>
<td>CORRECT</td>
</tr>
<tr>
<td>( H_0 ) NOT REJECTED: Defendant found __________</td>
<td>CORRECT</td>
<td>TYPE II ERROR</td>
</tr>
</tbody>
</table>
The Level of Significance (alpha level) and the p-value

Both the level of significance (alpha level) and the p-value give the probability of rejecting a true null hypothesis; equivalently, they give the probability of a type I error. However, there IS a difference between these two quantities.

The level of significance (alpha level) is the value the researchers conducting the study set BEFORE the study. It’s the probability they are comfortable with that they reject a true null hypothesis. $\alpha$ is typically 0.01 or 0.05.

The p-value is the value determined AFTER the researchers conduct the study and run the appropriate statistical test, and it finds exactly the probability of a type I error (a true null hypothesis was rejected).

**Key Idea**

If the p-value $< \alpha$ level of significant, then the null hypothesis is rejected.

**Example 3** A study is done to see if there’s a difference in the mean test score from students from two different countries.

a. Complete the null and alternative hypotheses.

Null hypothesis ($H_0$): The mean test scores from the two countries

Alternative hypothesis ($H_A$): The mean test scores from the two countries

b. The level of significance is set by the researchers as $\alpha = 0.05$. The study is completed and the appropriate statistical test is applied, and the p-value is 0.0243. Will the researchers reject or fail to reject the null hypothesis?

Since the p-value _____ alpha level set before the study, the researchers will ___________ the null hypothesis, concluding that ____________________________ statistically significant difference in the mean test scores from the two countries.

c. If $\alpha$ had been set to equal 0.01, would the researchers reject or fail to reject the null hypothesis?

Since the p-value _____ alpha level set before the study, the researchers would _____________ the null hypothesis at the 0.01 level of significance.
OPTIONAL Further Analysis of Hypothesis Testing

Example 4 Recall that we previously worked on an example about IQ scores, where we learned they are normally distributed with a mean of 100 and a standard deviation of 15. An elite school claims that its students have a higher than average intelligence, as measure through IQ testing.

a. Suppose that a student at this school is randomly selected and the student’s IQ is 105. What is the probability that a randomly selected student has an IQ that is at least 105?

As we did before, we will use NORM.DIST:

So, the probability that 1 randomly selected student would have an IQ of at least 105 is $1 - 0.63 = 0.37$.

So, there is a 37% chance that an individual IQ of at least 105 occurs just by chance.

b. One student doesn’t tell us much though. So, suppose we take a random sample of 10 of their students and determine that their average IQ score was also 105. Logically, do you think the chances of an average of 10 students having an IQ of at least 105 would be greater than or less than the chances of just 1 student having an IQ of at least 105?

Once we start to takes samples, we can’t use the normal distribution exactly as we did before, because we are talking about the average of a sample and not just one value. This is where we would use a statistical test called a Z-Test, which slightly adjusts the normal distribution to be able to use it with varying sample sizes and not just singleton values.

Let’s return to our example of a sample of 10 students and use the TI-84 to run this Z-test and determine the probability of obtaining an average IQ of at least 105.

On the TI-84, go to STAT→TESTS→1:Z-Test

Then, enter the appropriate information, as follows:

Select “Calculate” at the bottom and you should get a p-value = 0.1459.

So, there is a 14.6% chance that a sample of 10 individuals with an average IQ of at least 105 occurs just by chance.

If you select “Draw”, you will see a sketch that looks like the one at the right.

Notice how much smaller this probability region of 0.146 is compared to the probability region of 0.369 was in part (a).
c. Now suppose we take a random sample of 20 of their students and determine that their average IQ score was also 105. Again, logically, do you think the chances of an average of 20 students having an IQ of at least 105 would be greater than or less than the chances of 10 students having an average IQ of at least 105?

As in part (b), let’s use the TI-84 to run this Z-test and determine the probability of obtaining an average IQ of at least 105 in a sample of 20 students.

On the TI-84, go to STAT→TESTS→1:Z-Test. Then, enter the appropriate information, just as in (b), but now enter 20 for n.

Select “Calculate” at the bottom and you should get a p-value = 0.068.

So, there is a 6.8% chance that a sample of 20 individuals with an average IQ of at least 105 occurs just by chance.

If you select “Draw”, you will see a sketch that looks like the one at the right.

Again, notice how much smaller this probability region of 0.068 is compared to the probability region of 0.146 was in part (b).

d. Now that we have gotten the feel for how (with the same mean IQ score), the increase in sample size leads to a smaller p-value, let’s go through a hypothesis testing example from beginning to end.

In order to test the school’s claim that their students have an above average IQ, we decide to go in and test 25 randomly selected students and we set our alpha level to be 0.05 level of significance.

\[ H_0: \text{Mean} = 100 \]

\[ H_A: \text{Mean} > 100 \] (according to the school’s claim)

Using the TI-84, run the Z-test. Selecting “Draw” and the sketch should look like:

Selecting “Calculate,” the p-value = 0.04783

So, there is a 4.78% chance that a sample of 25 individuals with an average IQ of at least 105 occurs just by chance.

Returning to our hypothesis test, because the p-value = 0.04783 < 0.05 = alpha (level of significance), we will reject our null hypothesis and accept the alternative hypothesis.
e. Class discussion questions:

Do you think the school’s educational methods are causing higher IQs?

Even though the difference the students’ average IQs (105) and the general population (100) is statistically significant, do you feel it is practically significant?

Note: This example is meant to serve as an overview of how a specific statistical test might work with hypothesis testing. Thorough development and treatment of hypothesis testing (and statistical tests used in hypothesis testing) are beyond the scope of this course and would be covered in a statistics course.
Group Activity 10: Hypothesis Testing

1. An insurance company is interested in knowing if automobile accident rates are affected by the value of the automobile.
   a. State the null hypothesis.
   b. State the alternative hypothesis.

2. A study was done by Roger A Foltz, entitled “Academic Achievement of Student-Athletes.” He had several null hypotheses.

   Each null hypothesis was tested at the 0.05 level of significance.

   a. One null hypothesis was: “The differences among the mean GPAs of athletes according to year in high school will not be statistically significant.”

      He found that freshmen athletes’ mean GPA was 2.7, sophomore athletes’ mean GPA was 2.8, junior athletes’ mean GPA was 2.8, and senior athletes’ mean GPA was 2.7.

      After conducting his statistical test, he determined the p-value associated with these differences to be 0.8728.

      Did he reject or fail to reject his null hypothesis? How do you know?

   b. In your own words and in practical terms of the problem, what is the conclusion from this statistical test in (a)?

   c. Another null hypothesis was: “The differences among the mean GPAs of athletes according to gender will not be statistically significant.”

      He found that female athletes’ mean GPA was 3.0 and male athletes’ mean GPA was 2.6.

      After conducting his statistical test, he determined the p-value associated with this difference to be 0.001.

      Did he reject or fail to reject his null hypothesis? How do you know?

   d. In your own words and in practical terms of the problem, what is the conclusion from this statistical test in (c)?

3. A study was done by Hannah Marie Fernsler, entitled “A Comparison between the Test Scores of Third Grade Children Who Receive Drama in Place of Traditional Social Studies Instruction and Third Grade Children Who Receive Traditional Social Studies Instruction.”

   In the abstract to her article, Fernsler explains:
   “Social studies can be hard for students because it asks students to learn abstract subjects. It was felt that students would gain more understanding of the discipline of social studies if they interacted with
the content. When students take an active role in learning, they usually internalize the material. A study sought to see whether test scores might improve when they used dramatic supplements in addition to traditional instruction. Participants, 30 third graders in an eastern Tennessee public school, completed the study designed to determine whether students who received drama in place of traditional social studies instruction would achieve higher test scores than students who only received traditional social studies instruction. The school’s principal divided the third grade students into classes before the beginning of the school year. They were divided into classes equally by ability level. Two classes participated: Group A, the experimental group, received 6 weeks of social studies instruction, using drama as the primary medium; while Group B, the control group, received 6 weeks of traditional social studies instruction. After 6 weeks, both classes took the same social studies test, and a t-test compared the data of each group.”

a. Her null hypothesis was: “The test scores between students who receive drama in place of traditional instruction in social studies and students who receive only traditional instruction in social studies will not be different at the 0.05 level of significance.”

After 6 weeks, both classes took the same social studies test, and a statistical test (called a t-test) compared the data of each group.

She found that the mean test score of the control group (those that received only the traditional social studies instruction) was 88.33 and the mean test score of the experimental group (those that received the drama with their social studies curriculum) was 93.33.

Fernsler reported that the null hypothesis was rejected.

What do you know about how large or how small the p-value associated with this statistical test is?

b. In your own words and in practical terms of the problem, what is the conclusion from this statistical test in (a)?

4. In general (so not for any particular example), the alpha level of significance and the p-value both give the probability of what?

5. If both the alpha level of significance and the p-value describe the same probability, then how do they differ?

6. Why do we tend to be more worried about type I error than type II error?
Chapter 10: The Tipping Point (Optional)

Objectives:
1. Identify examples of human’s tendency to apply linear thinking, where it is not appropriate.
2. Explain what Gladwell means by “the tipping point” and identify the concept in everyday thinking.

The Tipping Point
By Malcolm Gladwell

The New Yorker
June 3, 1996

DEPT. OF DISPUTATION

Why is the city suddenly so much safer—could it be that crime really is an epidemic?

1.

As you drive east on Atlantic Avenue, through the part of New York City that the Police Department refers to as Brooklyn North, the neighborhoods slowly start to empty out: the genteel brownstones of the western part of Brooklyn give way to sprawling housing projects and vacant lots. Bedford-Stuyvesant is followed by Bushwick, then by Brownsville, and, finally, by East New York, home of the Seventy-fifth Precinct, a 5.6-square-mile tract where some of the poorest people in the city live. East New York is not a place of office buildings or parks and banks, just graffiti-covered bodegas and hair salons and auto shops. It is an economically desperate community destined, by most accounts, to get more desperate in the years ahead—which makes what has happened there over the past two and a half years all the more miraculous. In 1993, there were a hundred and twenty-six homicides in the Seven-Five, as the police call it. Last year, there were forty-four. There is probably no other place in the country where violent crime has declined so far, so fast.

Once the symbol of urban violence, New York City is in the midst of a strange and unprecedented transformation. According to the preliminary crime statistics released by the F.B.I. earlier this month, New York has a citywide violent-crime rate that now ranks it a hundred and thirty-sixth among major American cities, on a par with Boise, Idaho. Car thefts have fallen to seventy-one thousand, down from a hundred and fifty thousand as recently as six years ago. Burglaries have fallen from more than two hundred thousand in the early nineteen-eighties to just under seventy-five thousand in 1995. Homicides are now at the level of the early seventies, nearly half of what they were in 1990. Over the past two and a half years, every precinct in the city has recorded double-digit decreases in violent crime. Nowhere, however, have the decreases been sharper than Brooklyn North, in neighborhoods that not long ago were all but written off to drugs and violence. On the streets of the Seven-Five today, it is possible to see signs of everyday life that would have been unthinkable in the early nineties. There are now ordinary people on the streets at dusk—small children riding their bicycles, old people on benches and stoops, people coming out of the subways alone. “There was a time when it wasn’t uncommon to hear rapid fire, like you would hear somewhere in the jungle in Vietnam,” Inspector Edward A. Mezzadri, who commands the Seventy-fifth Precinct, told me. “You would hear that in Bed-Stuy and Brownsville and,
particularly, East New York all the time. I don't hear the gunfire anymore. I've been at this job one year and twelve days. The other night when I was going to the garage to get my car, I heard my first volley. That was my first time."

But what accounts for the drop in crime rates? William J. Bratton—who as the New York City Police Commissioner presided over much of the decline from the fall of 1994 until his resignation, this spring—argues that his new policing strategies made the difference: he cites more coordination between divisions of the N.Y.P.D., more accountability from precinct commanders, more arrests for gun possession, more sophisticated computer-aided analysis of crime patterns, more aggressive crime prevention. In the Seven-Five, Mezzadri has a team of officers who go around and break up the groups of young men who congregate on street corners, drinking, getting high, and playing dice—and so remove what was once a frequent source of violent confrontations. He says that he has stepped up random "safety checks" on the streets, looking for drunk drivers or stolen cars. And he says that streamlined internal procedures mean that he can now move against drug-selling sites in a matter of days, where it used to take weeks. "It's aggressive policing," he says. "It's a no-nonsense attitude. Persistence is not just a word, it's a way of life."

All these changes make good sense. But how does breaking up dice games and streamlining bureaucracy cut murder rates by two-thirds? Many criminologists have taken a broader view, arguing that changes in crime reflect fundamental demographic and social trends—for example, the decline and stabilization of the crack trade, the aging of the population, and longer prison sentences, which have kept hard-core offenders off the streets. Yet these trends are neither particularly new nor unique to New York City; they don't account for why the crime rate has dropped so suddenly here and now. Furthermore, whatever good they have done is surely offset, at least in part, by the economic devastation visited on places like Brownsville and East New York in recent years by successive rounds of federal, state, and city social-spending cuts.

It's not that there is any shortage of explanations, then, for what has happened in New York City. It's that there is a puzzling gap between the scale of the demographic and policing changes that are supposed to have affected places like the Seven-Five and, on the other hand, the scale of the decrease in crime there. The size of that gap suggests that violent crime doesn't behave the way we expect it to behave. It suggests that we need a new way of thinking about crime, which is why it may be time to turn to an idea that has begun to attract serious attention in the social sciences: the idea that social problems behave like infectious agents. It may sound odd to talk about the things people do as analogous to the diseases they catch. And yet the idea has all kinds of fascinating implications. What if homicide, which we often casually refer to as an epidemic, actually is an epidemic, and moves through populations the way the flu bug does? Would that explain the rise and sudden decline of homicide in Brooklyn North?

2.

When social scientists talk about epidemics, they mean something very specific. Epidemics have their own set of rules. Suppose, for example, that one summer a thousand tourists come to Manhattan from Canada carrying an untreatable strain of twenty-four-hour flu. The virus has a two-per-cent infection rate, which is to say that one out of every fifty people who come into close contact with someone carrying it catches the bug himself. Let's say that fifty is also exactly the number of people the average Manhattanite—in the course of riding the subways and mingling with colleagues at work—comes into contact with every day. What we have, then, given the recovery rate, is a disease in equilibrium. Every day, each carrier passes on the virus to a new person. And the next day
those thousand newly infected people pass on the virus to another thousand people, so that throughout the rest of the summer and the fall the flu chugs along at a steady but unspectacular clip.

But then comes the Christmas season. The subways and buses get more crowded with tourists and shoppers, and instead of running into an even fifty people a day, the average Manhattanite now has close contact with, say, fifty-five people a day. That may not sound like much of a difference, but for our flu bug it is critical. All of a sudden, one out of every ten people with the virus will pass it on not just to one new person but to two. The thousand carriers run into fifty-five thousand people now, and at a two-per-cent infection rate that translates into eleven hundred new cases the following day. Some of those eleven hundred will also pass on the virus to more than one person, so that by Day Three there are twelve hundred and ten Manhattanites with the flu and by Day Four thirteen hundred and thirty-one, and by the end of the week there are nearly two thousand, and so on up, the figure getting higher every day, until Manhattan has a full-blown flu epidemic on its hands by Christmas Day.

In the language of epidemiologists, fifty is the "tipping point" in this epidemic, the point at which an ordinary and stable phenomenon—a low-level flu outbreak—can turn into a public-health crisis. Every epidemic has its tipping point, and to fight an epidemic you need to understand what that point is. Take AIDS, for example. Since the late eighties, the number of people in the United States who die of AIDS every year has been steady at forty thousand, which is exactly the same as the number of people who are estimated to become infected with H.I.V. every year. In other words, AIDS is in the same self-perpetuating phase that our Canadian flu was in, early on; on the average, each person who dies of AIDS infects, in the course of his or her lifetime, one new person.

That puts us at a critical juncture. If the number of new infections increases just a bit—if the average H.I.V. carrier passes on the virus to slightly more than one person—then the epidemic can tip upward just as dramatically as our flu did when the number of exposed people went from fifty to fifty-five. On the other hand, even a small decrease in new infections can cause the epidemic to nosedive. It would be as if the number of people exposed to our flu were cut from fifty to forty-five a day—a change that within a week would push the number of flu victims down to four hundred and seventy-eight.

Nobody really knows what the tipping point for reducing AIDS may be. Donald Des Jarlais, an epidemiologist at Beth Israel Hospital, in Manhattan, estimates that halving new infections to twenty thousand a year would be ideal. Even cutting it to thirty thousand, he says, would probably be enough. The point is that it's not some completely unattainable number. "I think people think that to beat AIDS everybody has to either be sexually abstinent or use a clean needle or a condom all the time," Des Jarlais said. "But you don't really need to completely eliminate risk. If over time you can just cut the number of people capable of transmitting the virus, then our present behavior-change programs could potentially eradicate the disease in this country."

That's the surprising thing about epidemics. They don't behave the way we think they will behave. Suppose, for example, that the number of new H.I.V. infections each year was a hundred thousand, and by some heroic AIDS-education effort you managed to cut that in half. You would expect the size of the epidemic to also be cut in half, right? This is what scientists call a linear assumption—the expectation that every extra increment of effort will produce a corresponding improvement in result. But epidemics aren't linear. Improvement does not correspond directly to effort. All that matters is the tipping point, and because fifty thousand is still above that point, all these heroics will come to naught. The epidemic would still rise. This is the fundamental lesson of nonlinearity. When it comes to fighting epidemics, small changes—like bringing new infections down to thirty
thousand from forty thousand-can have huge effects. And large changes-like reducing new infections to fifty thousand from a hundred thousand-can have small effects. It all depends on when and how the changes are made.

The reason this seems surprising is that human beings prefer to think in linear terms. Many expectant mothers, for example, stop drinking entirely, because they've heard that heavy alcohol use carries a high risk of damaging the fetus. They make the perfectly understandable linear assumption that if high doses of alcohol carry a high risk, then low doses must carry a low- but still unacceptable-risk. The problem is that fetal-alcohol syndrome isn't linear. According to one study, none of the sixteen problems associated with fetal-alcohol syndrome show up until a pregnant woman starts regularly consuming more than three drinks a day. But try telling that to a neurotic nineties couple.

I can remember struggling with these same theoretical questions as a child, when I tried to pour ketchup on my dinner. Like all children encountering this problem for the first time, I assumed that the solution was linear: that steadily increasing hits on the base of the bottle would yield steadily increasing amounts of ketchup out the other end. Not so, my father said, and he recited a ditty that, for me, remains the most concise statement of the fundamental nonlinearity of everyday life: Tomato ketchup in a bottle—None will come and then the lot'll be.

What does this have to do with the murder rate in Brooklyn? Quite a bit, as it turns out, because in recent years social scientists have started to apply the theory of epidemics to human behavior. The foundational work in this field was done in the early seventies by the economist Thomas Schelling, then at Harvard University, who argued that "white flight" was a tipping-point phenomenon. Since that time, sociologists have actually gone to specific neighborhoods and figured out what the local tipping point is. A racist white neighborhood, for example, might empty out when blacks reach five per cent of the population. A liberal white neighborhood, on the other hand, might not tip until blacks make up forty or fifty per cent. George Galster, of the Urban Institute, in Washington, argues that the same patterns hold for attempts by governments or developers to turn a bad neighborhood around. "You get nothing until you reach the threshold," he says, "then you get boom."

Another researcher, David Rowe, a psychologist at the University of Arizona, uses epidemic theory to explain things like rates of sexual intercourse among teen-agers. If you take a group of thirteen-year-old virgins and follow them throughout their teen-age years, Rowe says, the pattern in which they first have sex will look like an epidemic curve. Non-virginity starts out at a low level, and then, at a certain point, it spreads from the precocious to the others as if it were a virus.

Some of the most fascinating work, however, comes from Jonathan Crane, a sociologist at the University of Illinois. In a 1991 study in the American Journal of Sociology, Crane looked at the effect the number of role models in a community—the professionals, managers, teachers whom the Census Bureau has defined as "high status"—has on the lives of teen-agers in the same neighborhood. His answer was surprising. He found little difference in teen-pregnancy rates or school-dropout rates in neighborhoods with between forty and five per cent of high-status workers. But when the number of professionals dropped below five per cent, the problems exploded. For black school kids, for example, as the percentage of high-status workers falls just 2.2 percentage points—from 5.6 per cent to 3.4 per cent—dropout rates more than double. At the same tipping point, the rates of childbearing for teen-age girls—which barely move at all up to that point—nearly double as well.
The point made by both Crane and Rowe is not simply that social problems are contagious— that non-virgins spread sex to virgins and that when neighborhoods decline good kids become infected by the attitudes of dropouts and teen-age mothers. Their point is that teen-age sex and dropping out of school are contagious in the same way that an infectious disease is contagious. Crane's study essentially means that at the five-per-cent tipping point neighborhoods go from relatively functional to wildly dysfunctional virtually overnight. There is no steady decline: a little change has a huge effect. The neighborhoods below the tipping point look like they've been hit by the Ebola virus.

It is possible to read in these case studies a lesson about the fate of modern liberalism. Liberals have been powerless in recent years to counter the argument that their policy prescriptions don't work. A program that spends, say, an extra thousand dollars to educate inner-city kids gets cut by Congress because it doesn't raise reading scores. But if reading problems are nonlinear the failure of the program doesn't mean—as conservatives might argue—that spending extra money on inner-city kids is wasted. It may mean that we need to spend even more money on these kids so that we can hit their tipping point. Hence liberalism's crisis. Can you imagine explaining the link between tipping points and big government to Newt Gingrich? Epidemic theory, George Galster says, "greatly complicates the execution of public policy. . . . You work, and you work, and you work, and if you haven't quite reached the threshold you don't seem to get any payoff. That's a very tough situation to sustain politically."

At the same time, tipping points give the lie to conservative policies of benign neglect. In New York City, for example, one round of cuts in, say, subway maintenance is justified with the observation that the previous round of cuts didn't seem to have any adverse consequences. But that's small comfort. With epidemic problems, as with ketchup, nothing comes and then the lot'll.

4.

Epidemic theory, in other words, should change the way we think about whether and why social programs work. Now for the critical question: Should it change the way we think about violent crime as well? This is what a few epidemiologists at the Centers for Disease Control, in Atlanta, suggested thirteen years ago, and at the time no one took them particularly seriously. "There was just a small group of us in an old converted bathroom in the sub-subbasement of Building Three at C.D.C.," Mark L. Rosenberg, who heads the Centers' violence group today, says. "Even within C.D.C., we were viewed as a fringe group. We had seven people and our budget was two hundred thousand dollars. People were very skeptical." But that was before Rosenberg's group began looking at things like suicide and gunshot wounds in ways that had never quite occurred to anyone else. Today, bringing epidemiological techniques to bear on violence is one of the hottest ideas in criminal research. "We've got a hundred and ten people and a budget of twenty-two million dollars," Rosenberg says. "There is interest in this all around the world now."

The public-health approach to crime doesn't hold that all crime acts like infectious disease. Clearly, there are neighborhoods where crime is simply endemic—where the appropriate medical analogy for homicide is not something as volatile as aids but cancer, a disease that singles out its victims steadily and implacably. There are, however, times and places where the epidemic model seems to make perfect sense. In the United States between the early sixties and the early seventies, the homicide rate doubled. In Stockholm between 1950 and 1970, rape went up three hundred per cent, murder and attempted murder went up six hundred per cent, and robberies a thousand per cent. That's not cancer; that's aids.
An even better example is the way that gangs spread guns and violence. "Once crime reaches a certain level, a lot of the gang violence we see is reciprocal," Robert Sampson, a sociologist at the University of Chicago, says. "Acts of violence lead to further acts of violence. You get defensive gun ownership. You get retaliation. There is a nonlinear phenomenon. With a gang shooting, you have a particular act, then a counter-response. It's sort of like an arms race. It can blow up very quickly."

How quickly? Between 1982 and 1992, the number of gang-related homicides in Los Angeles County handled by the L.A.P.D. and the County Sheriff’s Department went from a hundred and fifty-eight to six hundred and eighteen. A more interesting number, however, is the proportion of those murders which resulted from drive-by shootings. Between 1979 and 1986, that number fluctuated, according to no particular pattern, between twenty-two and fifty-one: the phenomenon, an epidemiologist would say, was in equilibrium. Then, in 1987, the death toll from drive-bys climbed to fifty-seven, the next year to seventy-one, and the year after that to a hundred and ten; by 1992, it had reached two hundred and eleven. At somewhere between fifty and seventy homicides, the idea of drive-by shootings in L.A. had become epidemic. It tipped. When these results were published last fall in the Journal of the American Medical Association, the paper was entitled "The Epidemic of Gang-Related Homicides in Los Angeles County from 1979 Through 1994." The choice of the word "epidemic" was not metaphorical. "If this were a disease," H. Range Hutson, the physician who was the leading author on the study, says, "you would see the government rushing down here to assess what infectious organism is causing all these injuries and deaths."

Some of the best new ideas in preventing violence borrow heavily from the principles of epidemic theory. Take, for example, the so-called "broken window" hypothesis that has been used around the country as the justification for cracking down on "quality of life" crimes like public urination and drinking. In a famous experiment conducted twenty-seven years ago by the Stanford University psychologist Philip Zimbardo, a car was parked on a street in Palo Alto, where it sat untouched for a week. At the same time, Zimbardo had an identical car parked in a roughly comparable neighborhood in the Bronx, only in this case the license plates were removed and the hood was propped open. Within a day, it was stripped. Then, in a final twist, Zimbardo smashed one of the Palo Alto car's windows with a sledgehammer. Within a few hours, that car, too, was destroyed. Zimbardo's point was that disorder invites even more disorder—that a small deviation from the norm can set into motion a cascade of vandalism and criminality. The broken window was the tipping point.

The broken-window hypothesis was the inspiration for the cleanup of the subway system conducted by the New York City Transit Authority in the late eighties and early nineties. Why was the Transit Authority so intent on removing graffiti from every car and cracking down on the people who leaped over turnstiles without paying? Because those two "trivial" problems were thought to be tipping points-broken windows-that invited far more serious crimes. It is worth noting that not only did this strategy seem to work—since 1990, felonies have fallen more than fifty per cent—but one of its architects was the then chief of the Transit Police, William Bratton, who was later to take his ideas about preventing crime to the city as a whole when he became head of the New York Police Department.

Which brings us to North Brooklyn and the Seventy- fifth Precinct. In the Seven-Five, there are now slightly more officers than before. They stop more cars. They confiscate more guns. They chase away more street-corner loiterers. They shut down more drug markets. They have made a series of what seem, when measured against the extraordinary decline in murders, to be small changes. But it is the nature of nonlinear phenomena that
sometimes the most modest of changes can bring about enormous effects. What happened to the murder rate may not be such a mystery in the end. Perhaps what William Bratton and Inspector Mezzadri have done is the equivalent of repairing the broken window or preventing that critical ten or fifteen thousand new H.I.V. infections. Perhaps Brooklyn-and with it New York City-has tipped.

Note: This article was later developed into the best-selling book, *The Tipping Point*, by Malcolm Gladwell.
(Optional) Group Activity 11: The Tipping Point

In section 2, Gladwell gives a hypothetical scenario of the spread of a flu virus. By studying his example we can get a better handle on what he means by the tipping point and nonlinear phenomena.

Read this summary of the initial scenario. It is imperative that you understand this before moving on!

One thousand infected people visit New York City. During the day each person comes into close contact with 50 people. But, the infection rate of this particular virus is only 2% (meaning only 2% of the people who encounter the virus will actually become ill).

Thus, each person infects: \(0.02 \times 50 = 1\) new person (check this calculation for yourself).

So, the initial 1000 people infect 1000 new people.

On day two, since the virus is only a 24 hour flu, the original group of people is no longer sick, and, therefore, the total number of infected people remains constant at 1000.

1. During the Christmas season, Gladwell suggests, the average person will come into close contact with more people (since there are more people shopping). He suggests the number will rise from 50 to 55. Assuming the infection rate stays at 2% and starting with 1000 infected, as described above, set up an Excel spreadsheet to model the number infected.

To help get you started, set up the following:

<table>
<thead>
<tr>
<th>Day</th>
<th>Number Infected</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

The calculation for the number infection on day 1 is: \(1000 \times 55 \times 0.02 = 1100\) new people.

Now, you need to develop a formula (using cell reference(s) where appropriate) to perform this given computation and to then fill the second column. (Hint: If you did the formula correctly there should be 304,482 infected people on day 60.)

Paste the table for the first month into your Word document.

2. Based on your table, how long does it take for the number of infected people to double?

3. Make a scatterplot of the data. Paste the graph into your Word document.

4. Gladwell made the assumption that the average person would come into close contact with 50 people on an average day and 55 people during the Christmas season. Let’s look at what would happen if these numbers were raised or lowered.
Unit 2: Statistics  
Chapter 10: The Tipping Point (Optional)

a. Play with the numbers on your spreadsheet (in the formula for column 2) by changing the number 55 to something lower than 50. Paste a new scatterplot into your Word document that results from coming into contact with less than 50 people.

b. Now, play with the numbers on your spreadsheet (in the formula for column 2) by changing the number to something higher than 55. Paste a new scatterplot into your Word document that results from coming into contact with more than 55 people.

5. Gladwell discusses many concepts, such as linear thinking, nonlinearity, and the tipping point. Answer the following questions in your own words.
   a. What does Gladwell mean in the 7th paragraph of section 2 when he writes, “human beings prefer to think in linear terms”? In particular, how does Gladwell’s example involving pregnant women and alcohol consumption illustrate linear thinking?
   b. Explain the term tipping point in your own words.
   c. How does the expression – the straw that broke the camel's back – illustrate the concept of the tipping point?
Chapter 11: Linear Relationships and Regression

Objectives:
1. Determine if a relationship is linear, given a table of values.
2. If a relationship is linear, find its slope and equation.
3. Use Excel to obtain the least-squares regression line.
4. Interpret the $R^2$-value associated with the regression lines.
5. Use the regression line to interpolate and extrapolate values.
6. Identify limitations of mathematical models.

Mathematical Modeling: An Overview

We all need to plan for the future, whether it is getting ready for tomorrow's class, getting ready for winter, or getting ready for retirement. Businesses, governments, schools, and organizations of all kinds especially need to plan. In science, we seek to predict certain quantities when we know others and to discover relationships between variables. Mathematical modeling arises out of the need to plan, predict, and explore relationships.

- A **mathematical model** is an equation, graph, or algorithm that fits some real data reasonably well and that can be used to make predictions.
- Predictions come in two kinds. In both cases, one can predict $x$ given $y$, as well as $y$ given $x$.
  - **Extrapolations** are predictions outside the range of existing data.
  - **Interpolations** are predictions made in between existing data points.

**Key Idea**

When making extrapolations, the further one goes from the actual data, the less confident one becomes about one's predictions. A prediction very far out from the data may even end up being correct, but even so, we have hold back our confidence because we don't know if the model will apply any longer.

Linear Relationships

Facts you need to recall about linear functions:

- A linear function is one whose graph is a straight line.
- Slope gives a measure for how steep the line is.
  - $\text{Slope} = \frac{\text{rise}}{\text{run}} = \frac{\text{vertical change}}{\text{horizontal change}} = \frac{\text{change in } y}{\text{change in } x}$.
  - $\text{Slope} = \frac{y_2 - y_1}{x_2 - x_1}$, where $(x_1, y_1)$ and $(x_2, y_2)$ are two points on the line.

- Algebraically, linear functions can be written in the slope-intercept form: $y = mx + b$, where
  - $b$ is the $y$-intercept (where the graph crosses the $y$-axis)
  - $m$ is the slope
- To find the equation of the line, you can use the point-slope formula: $y - y_1 = m(x - x_1)$, where
  - $m$ is the slope and
  - $(x_1, y_1)$ is any point on the line.

While these mechanics are important, you should now be able to understand the essence of linearity. That is:
A linear function is one with a constant rate of change, which is its slope.

This means that

**Key Idea**

A linear relationship is one whose output values change by a constant (additive or subtractive) rate of change (for a fixed change in input values).

- If the slope is not written as a fraction (as will be the case with most of our applications), think “the graph rises or falls by $m$ units for every 1-unit move to the right.”

**Example 1**

Determine if the following relationship is linear. If so, what is the rate of change and what is the equation?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-4</td>
</tr>
<tr>
<td>10</td>
<td>-1</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>

In Excel, you determine if a relationship is linear similarly, by filling the neighboring column with the following:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>Rate of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>$(B3-B2)/(A3-A2)$</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

After typing in this formula, “click and drag” to fill the column and look to see if the column is constant.
Example 2 Determine if the following relationship is linear. If so, what is the rate of change and what is the equation?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
</tr>
</tbody>
</table>

Rate of change:

\[
\frac{4 - 1}{2 - 0} = \frac{3}{2}
\]

\[
\frac{16 - 4}{4 - 2} = 6
\]

\[
\frac{36 - 16}{6 - 4} = 10
\]

Because the rate of change is not constant, the relationship is not linear. (In fact, you would not need to find all the slopes, if you noticed that the change in $x$ was the same, but the change in $y$ was not constant.)

Example 3 Determine if the following relationship is linear. If so, what is the rate of change and what is the equation?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
</tr>
</tbody>
</table>

Rate of change:

\[
\frac{9 - 5}{4 - 3} = 4
\]

\[
\frac{11 - 9}{6 - 4} = 1
\]

\[
\frac{17 - 11}{9 - 6} = 2
\]

Because the rate of change is not constant, the relationship is not linear. (Notice that here, the change in $x$ was not the same, so you cannot just conclude that the relationship is not linear but simply looking at the change in $y$. You did need to actually find the slopes here.)

Example 4 Determine if the following relationships are linear. If so, what is the rate of change and what is the equation?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>4.5</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>9.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>4</td>
<td>31.25</td>
</tr>
</tbody>
</table>

Because the rate of change is not constant, the relationship is not linear.
Example 5 Use the graph given at the right to answer the following.

a. State and interpret the y-intercept, in practical terms of the problem.

b. State and interpret the slope, in practical terms of the problem.

c. State the equation of the line.

d. Use the equation to determine the resale value of the phone in 3 years.
Example 6  Studies of the metabolism of alcohol (e.g., Wilkinson et al., Journal of Pharmacokinetics and Biopharmaceutics 5(3):207-224, 1977) consistently show that blood alcohol content (BAC), after rising rapidly after ingesting alcohol, declines linearly. For example, in Wilkinson's study, BAC in a fasting person rose to about 0.018 % BAC after a single drink in half an hour. After an hour the level had dropped to 0.01 % BAC. Assuming that BAC continues to decline linearly, approximately when will BAC drop to 0.002%?

First observe that you are given two points on your line and could even be written in table form:

<table>
<thead>
<tr>
<th>$x$ (hrs)</th>
<th>$y$ (BAC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.018</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>?</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Linear Regression

Having now reviewed the generalities about linear functions as well as the concept of modeling, also recall what was discussed last week about correlation. As we saw, real data is seldom perfectly linear, but suppose we have a data set that we plot and see is reasonably linear. The natural question should be: How do we put all this together? That is, how does one get a linear model (or equation of the data)? The standard approach is to use a "best-fit" line. The one we use is called the least squares regression line. (As you will see below, the overall concept is relatively intuitive.)

Go to the website [http://digitalfirst.bfwpub.com/stats_applet/stats_applet_5_correg.html](http://digitalfirst.bfwpub.com/stats_applet/stats_applet_5_correg.html) and create a reasonably linear set of data points. For example, consider the data set:
Next, select “Draw Line” and try to draw in a line (click and drag) that you think (intuitively) best fits your data. Consider the line created at the below:

Intuitively, you probably find yourself trying to draw in a line that “bisects” the points.

Finally, click on the box to “show least-squares line” to see how close you were. For this example, the line drawn by intuition (in green) is similar to the actual least-squares line (in blue).

What exactly is going on here? Are we just trying to “bisect” all the points? And, what is the “least-squares” name all about? To explain the concept of the least-squares regression line, consider the data set, such as the one below:

(The data can be found at http://faculty.elgin.edu/nscherger/Courses/MTH104/Spruce Height DBH.xls.)

There appears to be a linear trend to the data and if one were to find the correlation coefficient in Excel (as we did before, here, entering =CORREL(B7:B42,C7:C42)), the result is approximately \( r = 0.81 \), a strong, positive correlation.

Thus, it makes sense to find a regression line to the data.
The least squares regression line is the line \( f(x) = mx + b \), such that the sum of the squares of vertical errors is as small as possible. The idea is to find the \( m \) and the \( b \) that minimize the sum for the set of fixed data points.

Consider the two graphs below, where lines were drawn in that “looked” pretty good.

This final graph below is the actual least-squares line which minimizes the sum of the squares of the errors.

As you can see, Excel can give both the equation of the line as well as the coefficient of determination, \( R^2 \), which will be discussed later in more detail.
Example 6 Consider the data on the number of cases of Lyme Disease in the US from 1990 through 2013 ([http://faculty.elgin.edu/nscherger/Courses/MTH104/Lyme Disease Totals 1990-2013.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/Lyme Disease Totals 1990-2013.xlsx)). Find the equation and graph the least-squares regression line, including the coefficient of determination, $R^2$.

Finding the least-squares regression line in Excel
- First, make a scatterplot of the data.

![Lyme Disease in the United States](image)

- Next, click on Chart Elements and check Trendline at the bottom and then go down and select More Options, as shown.

- Once you do, you should get options for your trendline that will appear at the right. Select Trendline Options, as shown.

- Linear should be selected, by default. Then, look to the bottom and select the options to display the equation and the $R^2$, as shown.

- The result should be:
Before we look at some examples using this linear model, let’s discuss the $R^2$ value.

**The coefficient of determination, $R^2$**

- In broadest terms, $R^2$ is a standard measure ranging from 0 to 1 that reveals how well the regression equation fits the data.
- More precisely, $R^2$ is the percentage of variance common to both variables. Stated a bit more simply, $R^2$ is the percentage of the variance of $y$ that is accounted for by the line.
  - If $R^2 = 1$, then 100% of the variance in $y$ is explained by the line, so we have a perfect fit.
  - If $R^2$ is close to 0, then virtually none of the variance in $y$ is accounted by the line, so we have a terrible fit.

Example 6 continued...

The $R^2$ value for the least-squares regression line for Lyme Disease was 0.8287, so because this $R^2$ value is close to 1, we can initially observe that the line fits the data quite well.

More specifically, we would say that about 83% of the variance in $y$ (number of Lyme Disease cases) is explained by line (in other words, accounted for in the model).

Example 6 continued...

Now that we have observed that the linear model fits the Lyme Disease data well (as seen graphically and numerically through the $R^2$ value of 0.8287), let’s make a prediction for the number of cases in 2017. (This is extrapolation.)

Because $x = 0$ corresponded to 1990, the year 2017 will correspond to $x = 27$. Subbing $x = 27$ into the regression equation, we get:

$$y = 855.41(27) + 8827.4 \approx 31,923$$

cases predicted by the model for 2017.

It can also be a good idea to look at the graph and extend the line out to the year 2017 to visualize this prediction.

Go to Trendline Options (as shown on the last page to display the equation and the $R^2$) and forecast forward by 5 units. The result is:
Note that our prediction of 31,923 cases in 2017 (x=27) looks reasonable.

Let’s look at one final example to illustrate a cautionary note about extrapolating.

**Example 7** Consider the data on public university tuition in Illinois
([http://faculty.elgin.edu/nscherger/Courses/MTH104/IL Public Univ Tuition 1992-2012.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/IL Public Univ Tuition 1992-2012.xlsx)). Create a column C that is the “Years after 1990.” Then, copy the tuition data into column D.

a. Using columns C and D, find the equation and graph the least-squares regression line, forecasting forward by 4 periods. Include the coefficient of determination, \( R^2 \). Overall, does the linear equation appear to model the data well?

b. Use the linear regression equation to predict the average annual tuition and fees at Illinois public universities in 2015.

c. The actual average annual tuition and fees at Illinois public universities in 2015 was $13,983. How far off was the linear model’s prediction?
Example 8 Consider the data about alcohol-related traffic fatalities in the US (http://faculty.elgin.edu/nscherger/Courses/MTH104/US alcohol related traffic fatalities over time.xlsx). When one fits a linear regression to the data, the following graph is obtained:

[Graph showing a linear regression line with R² = 0.9074]

Observe the high R² value of 0.9074, showing that the linear model represents the data well. So, let’s use the equation to predict (extrapolate) the alcohol-related traffic fatality rate (per 100 million miles) for 2015.

\[ y = -0.0478(2015) + 96.289 \]

\[ y = -0.028 \]

So, there are -0.028 traffic fatalities (per 100 million miles)? How can there be a negative number of fatalities? Or, even if it was 0….is it realistic to think that there would be no alcohol-related traffic fatalities in 2015? Looking at the data itself, if you had to predict what happened with the number of alcohol-related traffic fatalities over the last ten years (from 2005-2015), what would you guess? Do you think it would follow the regression line? Explain.

The lesson here is that you can easily make poor or even ridiculous predictions (even with models with high R² values) if you don’t pay attention to the data itself and also test your answers for reasonableness. Further, while one can make predictions far from the original data, the confidence in those predictions made must decrease the further out one goes.
Group Activity 12: Linear Relationships and Regression

1. Determine if each of the following tables represents a linear relationship. If it is linear, express the relationship as an equation (in the final form of $y = mx + b$).

   a. 
   
   Snow Crickets' Chirping
   
<table>
<thead>
<tr>
<th>Temperature (F)</th>
<th>Rate (chirps/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>70</td>
<td>120</td>
</tr>
<tr>
<td>75</td>
<td>140</td>
</tr>
<tr>
<td>80</td>
<td>160</td>
</tr>
</tbody>
</table>

   b. 
   
   Annual Value of a $1000 Investment
   
<table>
<thead>
<tr>
<th>Years</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$1000.00</td>
</tr>
<tr>
<td>1</td>
<td>$1061.68</td>
</tr>
<tr>
<td>2</td>
<td>$1127.16</td>
</tr>
<tr>
<td>3</td>
<td>$1196.68</td>
</tr>
<tr>
<td>4</td>
<td>$1270.49</td>
</tr>
<tr>
<td>5</td>
<td>$1348.85</td>
</tr>
</tbody>
</table>
Before starting your first semester at ECC, you wanted to purchase a new computer, but did not have the money to do so. Your aunt and uncle agreed to lend you the money, provided that you make regular monthly payments of a set amount until the amount is paid off. The linear function graph at the right represents the amount of money that you owe your aunt and uncle over time.

a. What is the y-intercept?

b. Interpret the y-intercept in practical terms of the problem.

c. What is the rate of change (ie, the slope) for this function?

d. Interpret the slope in practical terms of the problem.

e. What is the linear equation for this model?

f. Use the linear equation to how much money you still owe your aunt and uncle after 9 months. Show work.

g. Use the linear equation to determine how many months it will take until you owe less than 500. Show work.
   a. Make a scatterplot of the data, including a title and axes labels. Observe the linear trend of the data and thus, add a linear trendline. Include the equation for the trendline and the R-squared value. Paste this chart in your Word document.
   b. Do you think the linear model you obtained in (a) is a good fit to the data? Mathematically justify your answer (so discuss the R²).
   c. Use the linear **equation** to predict (or extrapolate) the mortality rate from colon and related cancers in 2015. Show work and label your answer.
   d. Does this prediction seem reasonable?

   a. Make a scatterplot of the data, including a title and axes labels. Add a linear trendline. Include the equation for the trendline and the R-squared value. Also, forecast the line forward by 2 years. Paste this chart in your Word document.
   b. Use the linear **equation** to predict (or extrapolate) the number of female deaths from opioids in 2015. Show work and label your answer.
   c. Does this prediction seem reasonable? (Hint: Look at the trend of the data itself over the last few years.)
Assignment 6: Linear Relationships and Modeling

Complete Assignment #6 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

   a. Create a new column D for years, where 0 corresponds to the beginning of 1985 (so, your first values should be 2, 4, ...). Then, create a new column E that is the sum of the spending by Democrats and Republicans. Copy and paste this table into your Word document.
   b. Make a scatterplot of the data, including a title and axes labels. Observe that while there are minor fluctuations to the data, there is a general upward trend that could be modeled by a linear function. Thus, add a linear trendline to your plot, including the equation and the $R^2$ value. Also, forecast forward by 5 units. Copy and paste this plot into in your Word document.
   c. Use this linear model (equation) to predict the amount of spending in 2015-16. Show work and label your answer.
   d. Does this prediction seem reasonable?

   a. Create a new column C for years, where 0 corresponds to 1850 (so, your first value should be 10). Then, it will be easiest to just recopy the population data from column B into column D. Copy and paste this table into your Word document.
   b. Make a scatterplot of the data, including a title and axes labels. Then, add a linear trendline to your plot, including the equation and the $R^2$ value. Also, forecast forward by 10 units. Copy and paste this plot into in your Word document.
   c. Compare the overall fit of the model in both #2 (political fundraising) and with this model in #3 (Elgin population), incorporating the $R^2$ values into your answer.
   d. Use this linear model (equation) to predict the Elgin Population in 2020. Show work and label your answer.
   e. Does this prediction seem reasonable? Explain.

Important Observation: Even though the $R^2$ was higher in the linear model in #3 than in the linear model in #2, the prediction in #2c made more practical sense with the data than did this prediction in #3d.

Thus, once again, we are seeing that...

We cannot rely on $R^2$ alone to determine the predictive value of a model.

Key Idea
MY HOBBY: EXTRAPOLATING

As you can see, by late next month you'll have over four dozen husbands. Better get a bulk rate on wedding cake.
Chapter 12: Non-Linear (Polynomial) Functions and Regression

Objectives:
1. Recall function basics, including the definition, notation, identifying independent and dependent variables, and determining domains and ranges.
2. From a scatterplot, decide the appropriate polynomial regression to fit.
3. In Excel, perform polynomial regression.
4. Interpret $R^2$ values.
5. Determine limitations of regression models, such as illogical predications from models with high $R^2$ values.

Function Basics
Facts you need to recall about functions:

- A function is a rule that assigns to each element of one set (domain) another set (range).
- A function can also be thought of as a set of ordered pairs (which could also be written as a table of values) in which no first coordinate is repeated.
- The domain is the set of independent variable.
  - In the $xy$-coordinate plane, the independent variable is ______.
- The range is the set of dependent variable.
  - In the $xy$-coordinate plane, the dependent variable is ______.
- Standard function notation: $y = f(x)$

While these facts are important, you should now be able to understand these and other generalities of functions from the perspective of real-world application problems.

Key Idea
Just as understanding linearity (linear functions) was the building blocks for linear modeling, understanding functions (in general, extending to include nonlinear functions) will be the building blocks of mathematical modeling (which will now be extended to include nonlinear functions).
Thus, now thinking about functions from a more application-based approach, you should also know:

- A function expresses how a ____________ variable changes with respect to an ____________ variable.

In other words, the dependent variable is a function of the independent variable, where ordered pairs are written as \((\text{independent variable}, \text{dependent variable})\).

- The domain is the set of all input values that “make sense” for the ____________ variable.
- The range is the set of all output values that result from the determined domain.
- Function notation: \(f\) (independent variable) = dependent variable.
- Finally, you should also recall that there are 3 common ways that we express functions:
  1. ____________ Form
  2. ____________ Form
  3. ____________ Form

Consider the following example, which will utilize the language and notation of functions.

**Example 1** Consider the data on US cancer mortality rates ([http://faculty.elgin.edu/nscherger/Courses/MTH104/US cancer mortality by type 1981-2013.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/US cancer mortality by type 1981-2013.xlsx)). (See worksheet 2 for just lung cancer data.)

<table>
<thead>
<tr>
<th>Year</th>
<th>Lung, Trachea, and Bronchus Mortality Rates (per 100,000 population) in the US</th>
<th>Year</th>
<th>Lung, Trachea, and Bronchus Mortality Rates (per 100,000 population) in the US</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>50.2</td>
<td>1998</td>
<td>57.1</td>
</tr>
<tr>
<td>1982</td>
<td>51.8</td>
<td>1999</td>
<td>55.5</td>
</tr>
<tr>
<td>1983</td>
<td>52.6</td>
<td>2000</td>
<td>56.1</td>
</tr>
<tr>
<td>1984</td>
<td>53.6</td>
<td>2001</td>
<td>55.4</td>
</tr>
<tr>
<td>1985</td>
<td>54.6</td>
<td>2002</td>
<td>55.0</td>
</tr>
<tr>
<td>1986</td>
<td>55.2</td>
<td>2003</td>
<td>54.1</td>
</tr>
<tr>
<td>1987</td>
<td>56.4</td>
<td>2004</td>
<td>53.3</td>
</tr>
<tr>
<td>1988</td>
<td>57.2</td>
<td>2005</td>
<td>52.7</td>
</tr>
<tr>
<td>1989</td>
<td>58.1</td>
<td>2006</td>
<td>51.5</td>
</tr>
<tr>
<td>1990</td>
<td>59.0</td>
<td>2007</td>
<td>50.6</td>
</tr>
<tr>
<td>1991</td>
<td>59.2</td>
<td>2008</td>
<td>49.5</td>
</tr>
<tr>
<td>1992</td>
<td>59.1</td>
<td>2009</td>
<td>48.4</td>
</tr>
<tr>
<td>1993</td>
<td>59.3</td>
<td>2010</td>
<td>47.6</td>
</tr>
<tr>
<td>1994</td>
<td>58.6</td>
<td>2011</td>
<td>46.0</td>
</tr>
<tr>
<td>1995</td>
<td>58.4</td>
<td>2012</td>
<td>44.9</td>
</tr>
<tr>
<td>1996</td>
<td>57.9</td>
<td>2013</td>
<td>43.4</td>
</tr>
<tr>
<td>1997</td>
<td>57.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• In this relation, each input (year) has exactly one output (fatality rate); therefore, the relation is a function.
• If thought of as ordered pairs (year, fatality rate), observe that no first coordinate is repeated; therefore, again we conclude the relation is a function.
• Fatality rate is the________________________variable, as it depends on ____________________.
• Year is the________________________variable, because each year passes independently of the fatality rate. (Note: Units of time will often be an independent variable.)
• This function is given to us in table form. Expressing the function in graphical form would mean obtaining the scatterplot.

![Graph: Lung, Trachea, and Bronchus Mortality Rates in the US]

• Determine the domain and range.

Note: To express this function in equation form, we will need to utilize mathematical modeling; however, it should be clear that this data would not be efficiently modeled by a linear equation; therefore, we need to discuss non-linear functions and non-linear regression.

[Example 2] Use the graph to answer the following.

a. What is the independent variable?

b. What is the dependent variable?

c. What is the domain?

d. What is the range?

![Graph: Number of Tadpoles in Different Levels of Ph for Water]
Choosing an initial model

The previous example (on lung cancer mortality) had one “turn” in the data, so a degree 2 (quadratic) polynomial was an appropriate choice. However, if there are two “turns” in the scatterplot, then a degree 3 (cubic) polynomial may be a better choice. The images below give guidance as to when to choose particular models:

<table>
<thead>
<tr>
<th>Model Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
</tr>
<tr>
<td>Quadratic</td>
</tr>
<tr>
<td>Cubic</td>
</tr>
</tbody>
</table>

**Example 3** Cassie is a researcher at the Department of Agriculture and has noticed that farmers have very different crop yields depending on the region of the farm. She thinks that this has to do with the different climates in each region. In order to test her idea, she collected data on crop yield and the average summer temperatures and she obtains the data below.

![Graph showing crop yield vs. average summer temperature](image)

a. What would be the most appropriate type of regression equation (linear, quadratic, or cubic) to fit to this data?

b. Cassie finds the regression equation: \( y = -0.06x^2 + 2.2x - 14 \). Using the regression equation, predict the crop yield if the average summer temperature is 13.5°C.

**Quadratic Regression**

Lines, by their defining property have a constant rate of change. However, other functions do not have a constant rate of change and are therefore broadly called non-linear functions. We will begin our first example with a quadratic regression, which is a degree 2 polynomial.
Example 4  Recall the graph shown at the right from Example 1. It is clear that this data should not be modeled by a line.

This is why it is extremely important to look at the data and decide what is reasonable when fitting regression equations.

After all, Excel will fit a linear regression line, which is:

$$y = -0.265x + 58.439$$

It is obvious from both the graph and the $R^2$ value that this line has absolutely no predicative value; however, unless we graph the data and/or observe the $R^2$, we may miss just how meaningless the equation $y = -0.265x + 58.439$ really is.

Thus, a non-linear regression equation would be better for this data.
Finding a quadratic regression equation in Excel:
As before, click on Chart Elements and go down to the Trendline option. (Another way, which would have worked with linear regression as well, is to right-click on any data point and choose the option “Add Trendline.”)

The same menu appears at the right side, as shown.

This time, choose “Polynomial” and select 2 as the order of the polynomial (which should be the default).

Before hitting ok, check the boxes to display the equation and the $R^2$ value.

The result is:

Observe both visually and by the very high $R^2$ value that this quadratic model is a much more appropriate model!

Example 4 continued... We can now use the model to make predictions.

For example, to predict the lung cancer mortality rate for the year 2015, we substitute $x = 35$ (because $x = 0$ corresponds to 1980).

We get: $y = \boxed{\text{expression}} \approx \boxed{\text{value}}$

Note that this prediction makes sense mathematically (because the model had a very high $R^2$ value of 0.9755) and it also makes sense reasonably, when one looks at the data.
Cubic and Higher Degree Polynomial Regression

Example 5 Consider the data on the unemployment rate in the US (http://faculty.elgin.edu/nscherger/Courses/MTH104/US unemployment data 2003-2014.xlsx), which yields the following scatterplot.

Visually, observe that a linear or a quadratic model would not make sense; however, a cubic model seems more appropriate.

To fit a cubic model, follow the same directions as with the quadratic regression, with only difference being that the order of the polynomial should be 3. The result is:

\[
y = -0.0374x^3 + 0.9139x^2 - 6.346x + 18.144
\]

\[
R^2 = 0.8079
\]

a. Use the model to extrapolate the unemployment rate for 2015 (which is \(x = 15\), because 0 is the year 2000):

\[
y = -0.0374(15)^3 + 0.9139(15)^2 - 6.346(15) + 18.144 = 2.36\%
\]
Unit 3: Math Modeling
Chapter 12: Non-Linear (Polynomial) Functions and Regression

This seems a bit low. Looking at the cubic model, it appears as if it is decreasing more quickly than the data shows.

b. Use the model to interpolate the unemployment rate midyear of 2009 (which is \( x = 9.5 \), because 0 is the year 2000):

\[
y = -0.0374(9.5)^3 + 0.9139(9.5)^2 - 6.346(9.5) + 18.144 = 8.27\%
\]

Looking at the graph, this 8.27% seems too low, as the data points for 2009 and 2010 are just under 10%.

So, while the degree-3 polynomial was the appropriate initial choice, these last two examples show that the model isn’t doing a very good job of interpolating and extrapolating. As we go back and look at the curve and the data, it does seem as if they don’t line up close enough in most areas to do a reliable job of making predictions.

**Key Idea**

Thus, an important CAUTION regarding regression:

When performing regression, we must always look at the model mathematically (observing the R^2) and reasonably. Look at the predictions being made from the equation logically (observing the graph’s original data) to decide if the results are reasonable.

Looking at the R^2 alone is never enough and cannot replace our own critical eye.

To illustrate this point further, consider if we raise the degree to the highest possible in Excel, which is degree 6.

\[
y = -0.000192x^6 + 0.012148x^5 - 0.298159x^4 + 3.591225x^3 - 22.074027x^2 + 65.050822x - 66.227622
\]

Now, we get an \( R^2 = 0.924 \), which certainly seems great and it is; however, suppose we were primarily interested in making a prediction for the unemployment rate in 2016 (so, \( x = 16 \)):

\[
y = -0.00019(16)^6 + 0.012148(16)^5 - 0.298159(16)^4 + 3.591225(16)^3 + 22.074027(16)^2 + 65.050822(16) - 66.227622 = 10.02\%
\]

Does 10.02% seem reasonable for 2016?
Looking at the downward trend of the data since 2000, we have no reason to think that data will turn back upward. So, what’s happening here? With the high $R^2$ of 0.924, shouldn’t this model do a better job?

This issue often occurs with higher-order polynomial regressions – especially if you don’t extend the graph to observe what the equation does in the area that you are trying to predict. Here, if we extend the regression equation, we see:

We now see that while the equation looked pretty good within the data points themselves, the degree 6 regression equation actually changes to increasing after the last data point in 2014.

Thus, even though this model has a high $R^2$ of 0.924, it has absolutely no predictive value!

**Key Idea**

Predictions from models need to make sense both mathematically AND logically!
Group Activity 13: Non-Linear Functions and Regression

Consider the following data on speed and fuel mileage recorded for a particular car.

<table>
<thead>
<tr>
<th>Speed mph (x)</th>
<th>Fuel Economy mpg (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td>30</td>
<td>46</td>
</tr>
<tr>
<td>40</td>
<td>48</td>
</tr>
<tr>
<td>50</td>
<td>45</td>
</tr>
<tr>
<td>60</td>
<td>41</td>
</tr>
<tr>
<td>70</td>
<td>39</td>
</tr>
<tr>
<td>80</td>
<td>38</td>
</tr>
</tbody>
</table>

1. What is the independent variable?

2. What is the dependent variable?

3. Enter the data in Excel and then make a scatterplot of the data, including a title and axes labels. Paste this graph into your Word document.

4. What is the domain?

5. What is the range?

6. Perform a quadratic regression (a degree 2 polynomial) to the data. Include the equation and the $R^2$ value and forecast forward by 10 units. Paste this graph into your Word document.

7. Use the quadratic regression equation to interpolate the fuel economy at a speed of 55 mph. Show work and label your answer.

8. Does this prediction in (7) seem reasonable? Explain. (Hint: Look at the actual data near 55 mph.)

9. In an effort to get a model that is a better fit to the data, perform a cubic regression (a degree 3 polynomial) to the data. Include the equation and the $R^2$ value and forecast forward by 10 units. ALSO, GET 6 DECIMALS ON THIS REGRESSION EQUATION BY TAPPING ON THE EQUATION AND IN THE MENU AT THE RIGHT, UNDER LABEL OPTIONS - NUMBER, CHANGE FROM “GENERAL” TO “NUMBER” AND THEN YOU CAN ENTER 6 AS THE NUMBER OF DECIMALS DESIRED. Paste this graph into your Word document.
10. Use the cubic regression equation to interpolate the fuel economy at a speed of 55 mph. Show work and label your answer. *(Hint: It’s easiest to copy and paste the equation from the graph into a cell in Excel, then modify it as needed.)*

11. Does this prediction in (10) seem reasonable? Explain.

12. Compare the overall fit of the quadratic model with the cubic model, incorporating the $R^2$ values into your answer.

13. Use the quadratic regression equation to extrapolate the fuel economy at a speed of 90 mph. Show work and label your answer.

14. Use the cubic regression equation to extrapolate the fuel economy at a speed of 90 mph. Show work and label your answer.

15. Which model (quadratic or cubic) seems more reasonable for 90 mph? *(Hint: Think about should be happening to gas mileage for very fast speeds.)*

16. Is your answer for (15) the same model you said had a better overall fit (in 12)?

What occurred here is why...

**Key Idea**

It is important to always think about these regression equations and their predictions mathematically and logically! Sometimes the model with the better overall fit and higher $R^2$ is NOT the best choice for predictive purposes.
Assignment 7: Non-Linear Relationships and Modeling

Complete Assignment #7 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. Consider the data on public university tuition in Illinois (http://faculty.elgin.edu/nscherger/Courses/MTH104/IL Public Univ Tuition 1992-2012.xlsx). Create a column C that is the “Years after 1990.” Then, copy the tuition data into column D.
   a. Using columns C and D, make a scatterplot of the data, including an appropriate title and axes labels. Then, find the quadratic regression equation, forecasting forward by 4 periods. Include the coefficient of determination, $R^2$. Copy and paste this graph into your Word document.
   b. Overall, does the quadratic equation appear to model the data well? Mathematically justify your answer.
   c. Use the quadratic regression equation to predict the average annual tuition and fees at Illinois public universities in 2015.
   d. The actual average annual tuition and fees at Illinois public universities in 2015 was $13,983. How far off was the quadratic model’s prediction?

2. Observe the scatterplot below and its corresponding cubic regression equation.

   \[ y = 0.00043182x^3 - 0.0717x^2 + 3.5233x - 5.5714 \]

   a. Use the regression equation to estimate the fuel economy when traveling at 58 mph. Show work and label your answer.
   b. Do you think this is a good estimate? Explain.
   c. Was the prediction you made in part a an example of interpolation or extrapolation?
   d. Use the regression equation to estimate the fuel economy when traveling at 85 mph. Show work and label your answer.
   e. Do you think this is a good estimate? Explain.

Key Idea
It is important to always think about these regression equations and their predictions mathematically and logically! Sometimes the model with the better overall fit and higher $R^2$ is NOT the best choice for predictive purposes. We must always look at HOW we are using a model to decide what model is best.
Mathematical modeling is about rules - the rules of reality. What distinguishes a mathematical model from, say, a poem, a song, a portrait or any other kind of 'model', is that the mathematical model is an image or picture of reality painted with logical symbols instead of with words, sounds or watercolors.

-John Casti
Chapter 13: Exponential Growth and Decay and Regression

Objectives:
1. Determine if a relationship is linear, exponential, or neither, given a table of values.
2. If a relationship is exponential, find its multiplicative factor and equation.
3. Solve problems involving half-life and doubling time.
4. Use Excel to perform exponential regression.
5. Interpret the $R^2$-value associated with the regression.
6. Use the regression line to interpolate and extrapolate values.

Exponential Relationships: An Introduction

Recall that a linear relationship is one where for a fixed change in $x$, $y$ increases or decreases by a fixed (additive/subtractive) quantity.

Whereas,

Key Idea

An exponential relationship is one whose output values change by a constant multiplicative factor (for a fixed change in input values).

This multiplicative factor is called a growth factor if it is greater than 1 and a decay factor if it is between 0 and 1.

Example 1

Determine if the following relationship is exponential. If so, what is the growth or decay factor?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>15</td>
<td>4.5</td>
</tr>
<tr>
<td>20</td>
<td>13.5</td>
</tr>
</tbody>
</table>

First, observe that the changes in the $x$-values is fixed (=5). Thus, we want to see if the $y$ gets multiplied by a fixed amount each time.

From the 1$^{\text{st}}$ point to the 2$^{\text{nd}}$ point, we compute $1.5/0.5 = 3$.
(So, $0.5 \times 3 = 1.5$.)

From the 2$^{\text{nd}}$ point to the 3$^{\text{rd}}$ point, we compute $4.5/1.5 = 3$.
(So, $1.5 \times 3 = 4.5$.)

From the 3$^{\text{rd}}$ point to the 4$^{\text{th}}$ point, we compute $13.5/4.5 = 3$.
(So, $4.5 \times 3 = 13.5$.)

Therefore, this is an exponential relationship with a growth factor of 3.
Example 2: Determine if the following relationship is exponential. If so, what is the growth or decay factor?

\[
\begin{array}{c|c}
 x & y \\
 0 & 0 \\
 1 & 1 \\
 2 & 4 \\
 3 & 9 \\
\end{array}
\]

**Factor:**

- \[ \frac{1}{0} = \text{DNE} \]
- \[ \frac{4}{1} = 4 \]
- \[ \frac{9}{4} = 2.25 \]

First, observe that the changes in the \(x\)-values is fixed (=1). Thus, we want to see if the \(y\) gets multiplied by a fixed amount each time.

From the 1\(^{st}\) point to the 2\(^{nd}\) point, we compute \(1/0 = \text{undefined}\).

From the 2\(^{nd}\) point to the 3\(^{rd}\) point, we compute \(4/1 = 4\).

From the 3\(^{rd}\) point to the 4\(^{th}\) point, we compute \(9/4 = 2.25\).

Because there is not a constant multiplicative factor, this is not exponential relationship.

Example 3: Determine if the following relationship is exponential. If so, what is the growth or decay factor?

\[
\begin{array}{c|c}
 x & y \\
 0 & 192 \\
 1 & 96 \\
 2 & 48 \\
 3 & 24 \\
\end{array}
\]

**Factor:**

- \[ \frac{96}{192} = 0.5 \]
- \[ \frac{48}{96} = 0.5 \]
- \[ \frac{24}{48} = 0.5 \]

First, observe that the changes in the \(x\)-values is fixed (=1). Thus, we want to see if the \(y\) gets multiplied by a fixed amount each time.

From the 1\(^{st}\) point to the 2\(^{nd}\) point, we compute \(96/192 = 0.5\).

(So, \(192 \times 0.5 = 96\).)

From the 2\(^{nd}\) point to the 3\(^{rd}\) point, we compute \(48/96 = 0.5\).

(So, \(96 \times 0.5 = 48\).)

From the 3\(^{rd}\) point to the 4\(^{th}\) point, we compute \(24/48 = 0.5\).

(So, \(48 \times 0.5 = 24\).)

Therefore, this is an exponential relationship with a decay factor of 0.5.

In Excel, you determine if a relationship is exponential similarly, by filling the neighboring column with the following:

<table>
<thead>
<tr>
<th>(x)</th>
<th>(y)</th>
<th>Growth/Decay Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>192</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>96</td>
<td>=B3/B2</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

After typing in this formula, “click and drag” to fill the column and look to see if the column is constant.
Unit 3: Math Modeling
Chapter 13: Exponential Growth and Decay and Regression

Example 4
Determine if the following relationship is exponential. If so, what is the growth or decay factor?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4.5</td>
</tr>
<tr>
<td>3</td>
<td>6.75</td>
</tr>
<tr>
<td>4</td>
<td>10.125</td>
</tr>
</tbody>
</table>

Example 5
Determine if the following relationship is exponential. If so, what is the growth or decay factor?

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

Technically, the growth or decay factor is the multiplicative factor by which $y$ increases per unit change in $x$. Therefore, if $x$ is not increasing by the same amount (as it was in the prior examples), the growth factor is $\sqrt[\frac{y_2}{y_1}]{\frac{y_2}{y_1}}^{1/(x_2-x_1)}$. This is analogous to the slope in a linear relationship.
Graphs of exponential relationships

Observe that exponential functions increase (or decrease) very rapidly on a portion of their graph, but exponential growth often starts very slow and exponential decay often ends very slow.

Where do these relationships arise? All over the place. Here are just a few:

- Populations tend to grow exponentially not linearly
- When an object cools (e.g., a pot of soup on the dinner table), the temperature decreases exponentially toward the ambient temperature
- Radioactive substances decay exponentially
- Money accumulating in a bank at a fixed rate of interest increases exponentially
- Viruses and even rumors tend to spread exponentially through a population (at first)

A final, fundamental observation:

Key Idea

If a quantity grows by a fixed percentage change, it grows exponentially.

Why?

Suppose a quantity \( A \) is growing by \( r\% \) each year. After one year, \( A \) will become \( A + Ar \). Now,

\[
A + Ar = A(1 + r)
\]

So, \( A \) has been multiplied by the quantity \( 1 + r \). If \( A(1 + r) \) is in turn increased by \( r \) percent, it will be multiplied by \( 1 + r \) again. So after two years, we have \( A(1 + r)^2 \).
The general pattern that results:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$A$</td>
</tr>
<tr>
<td>1</td>
<td>$A(1 + r)$</td>
</tr>
<tr>
<td>2</td>
<td>$A(1 + r)^2$</td>
</tr>
<tr>
<td>3</td>
<td>$A(1 + r)^3$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$n$</td>
<td>$A(1 + r)^n$</td>
</tr>
</tbody>
</table>

For example, if a population is growing by 5% per year, then the population is being multiplied by 1.05 each year.

Note: If a quantity $A$ is decreasing by $r\%$, then it $A$ becomes $A - Ar = A(1 - r)$, and all of the $(1 + r)$ factors would then be replaced by $(1 - r)$ factors.

Example 4 The US population has been growing by about 0.8% each year, and in 2000 the population of the US was 282 million. In Excel, set up the table below, which illustrates that we get the same answer either way.

<table>
<thead>
<tr>
<th>Year</th>
<th>Population by adding percent</th>
<th>Population by multiplying by growth factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>282</td>
<td>282</td>
</tr>
<tr>
<td>2001</td>
<td>=B2+B2*0.008</td>
<td>=C2*(1.008)</td>
</tr>
<tr>
<td>2002</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

Example 5 Suppose you have a shot of expresso at 2pm in the afternoon for an afternoon boost that contains 64 mg of caffeine, and it is eliminated from the body at a rate of 11.5% per hour. Set up an Excel spreadsheet to model this situation and extend your spreadsheet to determine how many mg of caffeine are in your body at 10pm when you are trying to go to sleep.
Group Activity 14: The Power of Exponential Growth

1. Determine if each of the following tables represents a relationship that is linear, exponential or neither. (These tables can be found in Excel file [http://faculty.elgin.edu/nscherger/MTH104/Linear Exponential Or Neither.xlsx](http://faculty.elgin.edu/nscherger/MTH104/Linear Exponential Or Neither.xlsx). Each table is on a separate sheet within this file; click on the worksheet tabs on the bottom of the window to display each one.)

   a. Snow Crickets’ Chirping
      | Temperature (F) | Rate (chirps/min) |
      |----------------|-------------------|
      | 50             | 40                |
      | 55             | 60                |
      | 60             | 80                |
      | 65             | 100               |
      | 70             | 120               |
      | 75             | 140               |
      | 80             | 160               |

   b. Annual Value of $1000 Invested At 6% Compounded Monthly
      | Years | Value       |
      |-------|-------------|
      | 0     | $1000.00    |
      | 1     | $1061.68    |
      | 2     | $1127.16    |
      | 3     | $1196.68    |
      | 4     | $1270.49    |
      | 5     | $1348.85    |

c. Decay of 100 mg of Strontium 90
<table>
<thead>
<tr>
<th>Years</th>
<th>Amount (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.0000</td>
</tr>
<tr>
<td>10</td>
<td>78.07092</td>
</tr>
<tr>
<td>20</td>
<td>60.95068</td>
</tr>
<tr>
<td>30</td>
<td>47.58476</td>
</tr>
<tr>
<td>40</td>
<td>37.14986</td>
</tr>
<tr>
<td>50</td>
<td>29.00323</td>
</tr>
</tbody>
</table>
Unit 3: Math Modeling
Chapter 13: Exponential Growth and Decay and Regression

2. The population of the Russia in 1998 was 147 million; the population of Nigeria was 106 million. Russia’s population was decreasing by approximately 0.1% annually; the population of Nigeria was growing at 3% annually.

Using Excel, determine the year when Nigeria’s population would exceed Russia’s, assuming that the current trend continues.

(Hint: Create a 3-column table with years in the first column, Russia’s population in the second column, and Nigeria’s population in the third column. See the last example in the notes before this activity for two different ways these columns can be computed.)

Copy and paste the table into your Word document and state your answer.

3. Suppose a magic genie gives you a choice. He will give you $1000 each day in the month of March or he will give you 1¢ on the March 1st, 2¢ on March 2nd, 4¢ on March 3rd, doubling the amount each day.

a. Without thinking about it too hard, which would you pick?

b. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Genie.xlsx. Determine how to use Excel to fill the table for all 31 days. Paste the completed table into your Word document.

<table>
<thead>
<tr>
<th>Day</th>
<th>Amount given on each day under option 1</th>
<th>Total for option 1 on each day</th>
<th>Amount given on each day under option 2</th>
<th>Total for option 2 on each day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1000</td>
<td>$1000</td>
<td>$0.01</td>
<td>$0.01</td>
</tr>
<tr>
<td>2</td>
<td>$1000</td>
<td>$2000</td>
<td>$0.02</td>
<td>$0.03</td>
</tr>
<tr>
<td>3</td>
<td>$1000</td>
<td>$3000</td>
<td>$0.04</td>
<td>$0.07</td>
</tr>
<tr>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

c. Now using the table, which deal would you pick?

d. At what point in your table is the amount in the fifth column in your table larger that the amount in the third column?

e. Plot the two total columns under each option on the same chart as separate series. Paste this graph into your word document. (To do so you need to select the data in columns A, C, and E only and make a scatterplot.) In order to get a nicer view, adjust the scale of the y-axis, by right-clicking on the y-axis and under scale, enter a maximum of about 100,000.
Chapter 13: Exponential Growth and Decay and Regression (continued)

The Exponential Equation:  \( y = Ab^x \)

Recall that an exponential function is one where for a fixed change in \( x \), \( y \) gets multiplied by a fixed amount.

Based on the above discussion, where we saw that if a quantity \( A \) is increasing or decreasing by \( r \% \), we get the following (where the + is for increasing and the – is for decreasing):

<table>
<thead>
<tr>
<th>( x )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( A )</td>
</tr>
<tr>
<td>1</td>
<td>( A(1 + r) )</td>
</tr>
<tr>
<td>2</td>
<td>( A(1 + r)^2 )</td>
</tr>
<tr>
<td>3</td>
<td>( A(1 + r)^3 )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( n )</td>
<td>( A(1 + r)^n )</td>
</tr>
</tbody>
</table>

Thus, the exponential equation (or function) for a quantity \( y \) that is growing or decaying exponentially is

\[ y = Ab^x = A(1 \pm r)^x, \]

where

- \( A \) = the initial value (of \( y \) when \( x = 0 \)),
- \( b = \frac{1 \pm r}{1} \), where \( r \) is the fractional growth rate, typically given as a \%
  (must be converted to decimal form)

- For Exponential Growth: \( 1 + r \rightarrow b > 1 \)
- For Exponential Decay: \( 1 - r \rightarrow 0 < b < 1 \)

\( x \) = time (must be the same unit of time as used for the rate)

Example 6 Let’s look at Example 4 again, considering the US population. (Notice that one could solve the problem by either writing the model, as below, or by extending the table in Excel. You should be able to do both.)

The US population has been growing by about 0.8% each year, and in 2000 the population of the US was 282 million. Write the exponential model and use it to predict the population in the US in 2050.

Filling in the model, \( y = 282(1 + 0.008)^x \), so \( y = 282(1.008)^x \) will give the population (in million) \( x \) years after 2000.

Then, 2050 will be when \( x = 50 \): \( y = 282(1.008)^{50} \approx 420.02555 \) million or about 420,025,550.
Example 7 Suppose a drug breaks down in the body at a rate of 17% per hour. If the initial amount in the bloodstream is 10 milligrams, how much is left in the body after 5 hours?

Filling in the model, $y = 10(1 - 0.17)^x$, so $y = 10(0.83)^x$ will give the amount (in milligrams) in $x$ hours.

Then, when $x = 5$: $y = 10(0.83)^5 \approx \underline{\hspace{2cm}}$ milligrams.

If using Excel to solve, your table would have started as (using either of the last 2 columns) and extended the table to hour 5.

<table>
<thead>
<tr>
<th>Hour</th>
<th>Amount</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>=B2-B2*0.17</td>
<td>=B2*(0.83)</td>
</tr>
<tr>
<td>2</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

Example 8 Let’s look at the Example 5 from earlier in the section about taking a shot of espresso at 2pm that contains 64 mg of caffeine, where it is eliminated from the body at a rate of 11.5% per hour. Write out the the exponential model and use it to determine how many mg of caffeine are in your body at 10pm when you are trying to go to sleep.

Doubling Time and Half-Life
Both doubling time and half-life are just specific examples of exponential growth and decay that occur frequently in application problems, so we tend to talk about them specifically.

- Doubling Time is the amount of time it takes a population to double. Or, to say it differently, it is the amount of time for a population to grow by 100%.
- Half Life is the amount of time it takes a population to be cut in half. Or, to say it differently, it is the amount of time for a population to decline by 50%.

Note: We will primarily be using Excel to solve problems involving doubling time and half-life.
Example 9  The number of cells in a tumor is doubling every 3 months. If the tumor begins with just one cell, how many cells will there be after 3 years?

The Excel table should begin as:

<table>
<thead>
<tr>
<th>Month</th>
<th>Number of Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>=B2*2</td>
</tr>
</tbody>
</table>

Extend the table until the 36th month to see there are then ___________ cells in the tumor.

Example 10  Check out this headline from back in 2014, in the height of the Ebola crisis:

The article reads that the number of Ebola cases in West Africa could double every twenty days if there was no intervention. If we start with just one case of Ebola, set up an Excel spreadsheet to determine how many cases would there be after 360 days (which is about a year).

Example 11  The current population of a threatened animal species is estimated to be 950,000, but is thought to be declining with a half-life of 20 years. (a) How many animals will there be in 80 years? (b) How long until the animal species would be extinct?

The Excel table should begin as:

<table>
<thead>
<tr>
<th>Years</th>
<th>Population of Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>950000</td>
</tr>
<tr>
<td>20</td>
<td>=B2/2</td>
</tr>
<tr>
<td>40</td>
<td>:</td>
</tr>
</tbody>
</table>
a. Extend the table to 80 years to see that there are _______________ animals.

b. To see when the animal species is extinct, you might first think to extend the table until the population is 0, but observing the table in Excel and thinking critically about successive division by 2, and you should realize that you will never get to 0.

However, when you look at the extended table in Excel, you should see that after ___________ years, there is less than 1 animal left (0.90599), so we can safely say that the population will be extinct after years. (In reality, it would probably happen much sooner.)

Example 12: Once again, let’s return to the example about taking a shot of espresso at 2pm that contains 64 mg of caffeine, where it is eliminated from the body at a rate of 11.5% per hour. Set up the Excel spreadsheet to model this problem and determine the half-life of the caffeine in this scenario?

Example 13: The half-life of Carbon 14 is about 5730 years and translates to a decay rate of about 1.202% every 100 years. The Dead Sea Scrolls contain about 78% of the normally occurring amount of Carbon 14 in them. Using the decay rate given, approximate the date of the Dead Sea Scrolls.

The Excel table should begin as:

<table>
<thead>
<tr>
<th>Years After Death</th>
<th>% of Carbon Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>100</td>
<td>=B2 – 0.01202*B2</td>
</tr>
<tr>
<td>200</td>
<td>:</td>
</tr>
</tbody>
</table>

Extend the table to see that the Dead Sea Scrolls are somewhere between ________ and ________ years old.

(More exact current estimates for their date are between 150BC and 5BC.)

As you will see in other examples in your group activities, Carbon 14 is not the only substance used to date artifacts.
Regression with an Exponential Model

As has been previously covered, when fitting an equation to a data set, a scatterplot is obtained and then the most appropriate model (such as linear and quadratic) is obtained.

Now, an additional choice (beyond linear and polynomial) is available: exponential regression; however, be very careful with extrapolating with exponential models, because they do grow (or decrease) so quickly. You want to be sure that such growth (or decay) is reasonable for the data.

When using Excel to fit an exponential model, the equation will be given to you in the form of $y = Ae^{kx}$, instead of how we have been writing the equations as $y = Ab^x$. Do not let the form $y = Ae^{kx}$ confuse you. It is simply another way of expressing an exponential model that utilizes the number $e$, which is just approximately 2.718.

Example 13  The growth of fruit flies in a laboratory environment are recorded in the table below.

<table>
<thead>
<tr>
<th>Day Number</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Flies</td>
<td>111</td>
<td>122</td>
<td>134</td>
<td>147</td>
<td>161</td>
<td>177</td>
<td>195</td>
<td>214</td>
<td>235</td>
<td>258</td>
<td>283</td>
</tr>
</tbody>
</table>

a. Is the relationship exactly exponential, nearly exponential, or not very exponential at all?

Enter the data into 2 columns in Excel and in the 3rd column, divide that day’s population by the previous day’s population (as with the opening examples in this section). Observe that while there is not an exact multiplicative factor; all of the factors are very close to ____________.

This indicates the data is ________________________.

b. Predict the number of flies on the 20th day.

Because the data quite nearly exponential, it makes good sense to fit an exponential equation.

Perform all of the same steps as before; however, when adding the trendline, select exponential (remember to select the equation and the $R^2$).

The result should be:

![Fruit Fly Population](image)

Finally, use the exponential model $y = 110.99884e^{0.09367x}$, we can let $x = 20$ and predict the number of flies: $y = 110.99884e^{0.09367(20)} = \text{about 723 fruit flies}$
### Group Activity 15: Exponential Models

1. A cup of coffee in a covered mug is initially 180°F and cools down exponentially by about 1% per minute. Determine the exponential model/equation \( y = A(1 \pm r)^x = Ab^x \).

2. Using the model/equation from (1), determine the prediction for the temperature of the coffee after 6 minutes. Show work and label your answer.

3. Now set up an Excel spreadsheet to determine the temperature of the coffee after 6 minutes. Note that you should get the same answer as you did in (2).
   (Hint: Look back at example 5 in the previous section of notes to see an example of how we first solved with the equation approach, as you did in 1 and 2, and then we set up an Excel spreadsheet, as you are being asked to do here.)

   Copy and paste this Excel table into your Word document and state your answer.

4. Extend your Excel spreadsheet to predict when the temperature of the coffee would be 100°F. State your answer. (You do not need to copy the extended table.)

5. Using the exponential model/equation from (1), predict for the temperature of the coffee after 2.5 hours. (Hint: Remember your model is in terms of minutes.) Show work and label your answer.

6. If the temperature of the room is 72°F, how accurate do you think your prediction in (5) is? Explain.

7. Some of the most famous Cro-Magnon cave paintings are located in Lascaux, France. Charcoal found in the cave has approximately 14% of the carbon 14 found in living wood.

   Using the fact that Carbon 14 decays approximately 1.202% every 100 years, estimate the age of the paintings. (Hint: See example 8 in the previous section of notes.)

   Copy the first 10 rows of your Excel table and then state your final answer (from the full table) as to the age of the paintings.

8. Rock samples brought back from the moon by Apollo astronauts have approximately 58% of the original amount of Uranium 238 present.

   Using the fact that Uranium decays approximately 7.41% every 500 million years, estimate the age of the rocks.

   Copy the first 10 rows of your Excel table and then state your final answer (from the full table) as to the age of the rocks.
**Assignment 8: Exponential Relationships**

**Complete Assignment # 8 in MyOpenMath and the problems below.**

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. In order to understand the prehistory of the Hawaiian island of Lana'i better, anthropologists Maria Sweeney, Melinda Allen, and Boyd Dixon used radiocarbon dating on charcoal found in an ancient dwelling site, the Kaunolu Village National Historic Landmark, the largest archeological complex on the island.

   In one of their samples, they found that approximately 94% of the original carbon 14 remained. Using the fact that Carbon 14 decays by 1.202% every 100 years, determine the approximate age of this sample.

   In your Word document, paste the Excel table you made to come up with your answer.

2. A student strained her knee in an intramural volleyball game. Her doctor prescribed an anti-inflammatory drug to reduce the swelling. She is to take a 440 mg tablet of the drug every 8 hours for 7 days. Her kidneys filter, or remove 60% of the drug from her body every 8 hours.

   a. Let’s walk through what is happening by hand, by performing the necessary computations to fill in the following blanks.

   She takes her first pill and will now have 440 mg of the drug in her system. After 8 hours, her kidneys filter out ________ mg of the drug, leaving ________ mg of the drug in her body. Then, she takes her second pill and thus, she will now have ________ mg of the drug in her system.

   b. One way to set up and extend this process in Excel would be by doing the following.

<table>
<thead>
<tr>
<th>Time (hours)</th>
<th>Amount filtered by kidneys (mg)</th>
<th>Amount left after filtering by kidneys (mg)</th>
<th>New dose (mg)</th>
<th>Total amount in body (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>440</td>
<td>440</td>
</tr>
<tr>
<td>8</td>
<td>=E2*.60</td>
<td>=E2-B3</td>
<td>440</td>
<td>=C3+D3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Extend this table for the full course of the 7 days she is taking the drug, and copy and paste this table into your Word document.
c. What is the maximum amount of drug in the student’s body over the course of the 7 days?

d. Make a scatterplot of the total amount of drug in the student’s body (so, you will need to highlight the first and last columns only) over just the first 3 days. As always, include a title and axes labels. Copy and paste this graph into your Word document.

e. Describe the graph.

f. Create a new table OR modify and extend the current table to show the amount of drug remaining in the student's body for three days after she stops taking the drug (that is, for days 8, 9 and 10), and her kidneys continue to flush out 60% of the drug every 8 hours. Paste this Excel table into your Word document.

g. Is the data in the table in (f) linear or exponential? Determine this as you did in (1). Copy and paste the Excel table that shows that the data is linear or exponential and state your final answer. If it is linear, state the slope. If it is exponential, state the growth or decay rate.

---

"The mathematics of uncontrolled growth are frightening. A single cell of the bacterium E. coli would, under ideal circumstances, divide every twenty minutes. That is not particularly disturbing until you think about it, but the fact is that bacteria multiply geometrically: one becomes two, two become four, four become eight, and so on. In this way it can be shown that in a single day, one cell of E. coli could produce a super-colony equal in size and weight to the entire planet Earth.

Chapter 14: The Consumer Price Index

Objectives:
1. Compare prices in different years by converting to constant dollars, using the consumer price index (CPI), and interpret those results.
2. Use Excel to convert a table of values to convert prices over time to constant dollars and interpret those results.
3. Use the CPI to determine inflation.
4. Analyze graphs in nominal and constant dollars.

The Consumer Price Index: A Way to Compare Prices in Different Years

Inflation is a decline in the value of money in relation to the goods that it can buy and is a pervasive economic phenomenon. It is so pervasive that it is very difficult to compare this year’s prices to last year’s, much less compare prices over decades.

How can we compare prices of item in different years?

Why are so many advocating for a higher minimum wage?

The answer is the use of price indices such as the consumer price index or CPI. Below is a simplified explanation of how the indices are created and how we can use them.

Economists choose a base year and determine the prices of a “bundle” of goods: food, clothing, housing costs, transportation costs, services, entertainment in varying proportions. The proportions for the index we are using, CPI(U), obtained from http://www.bls.gov/cpi/cpiri_2013.pdf, are:

<table>
<thead>
<tr>
<th>Components of the CPI(U)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing</td>
<td>41.4%</td>
</tr>
<tr>
<td>Transportation</td>
<td>16.4%</td>
</tr>
<tr>
<td>Food</td>
<td>14.9%</td>
</tr>
<tr>
<td>Medical Care</td>
<td>7.6%</td>
</tr>
<tr>
<td>Education &amp; Communication</td>
<td>7.1%</td>
</tr>
<tr>
<td>Recreation</td>
<td>5.8%</td>
</tr>
<tr>
<td>Apparel</td>
<td>3.4%</td>
</tr>
<tr>
<td>Other</td>
<td>3.4%</td>
</tr>
</tbody>
</table>

The cost of this bundle in the base year is assigned an index number of 100. The cost of the same bundle is then determined in a different year.

For example, say that 1982 was the base year and the cost of the bundle of goods in 1982 was $1103.46. In 1983 that same bundle cost $1138.91. This means in terms of the bundle chosen,

\[ 1103.46 \times (1982) = 1138.91 \times (1983) \]
Proportionally, this means
\[ 1 \frac{1982\$}{1103.46} = \frac{1138.91}{1103.46} \]
\[ = 1.032(1983\$) \]

Since we assigned the index 100 to 1982, we assign the index 103.2 to 1983.

Following a very similar process, the official CPI results in the table below.

**Using the official CPI table**

The Table below shows the official CPI since 1982, and the full table is available at [http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx](http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx).

<table>
<thead>
<tr>
<th>Year</th>
<th>CPI (1982-1984=100)</th>
<th>Year</th>
<th>CPI (1982-1984=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>96.5</td>
<td>1999</td>
<td>166.6</td>
</tr>
<tr>
<td>1983</td>
<td>99.6</td>
<td>2000</td>
<td>172.2</td>
</tr>
<tr>
<td>1984</td>
<td>103.9</td>
<td>2001</td>
<td>177.1</td>
</tr>
<tr>
<td>1985</td>
<td>107.6</td>
<td>2002</td>
<td>179.9</td>
</tr>
<tr>
<td>1986</td>
<td>109.6</td>
<td>2003</td>
<td>184.0</td>
</tr>
<tr>
<td>1987</td>
<td>113.6</td>
<td>2004</td>
<td>188.9</td>
</tr>
<tr>
<td>1988</td>
<td>118.3</td>
<td>2005</td>
<td>195.3</td>
</tr>
<tr>
<td>1989</td>
<td>124.0</td>
<td>2006</td>
<td>201.6</td>
</tr>
<tr>
<td>1990</td>
<td>130.7</td>
<td>2007</td>
<td>207.3</td>
</tr>
<tr>
<td>1991</td>
<td>136.2</td>
<td>2008</td>
<td>215.3</td>
</tr>
<tr>
<td>1992</td>
<td>140.3</td>
<td>2009</td>
<td>214.5</td>
</tr>
<tr>
<td>1993</td>
<td>144.5</td>
<td>2010</td>
<td>218.1</td>
</tr>
<tr>
<td>1994</td>
<td>148.2</td>
<td>2011</td>
<td>224.9</td>
</tr>
<tr>
<td>1995</td>
<td>152.4</td>
<td>2012</td>
<td>229.6</td>
</tr>
<tr>
<td>1996</td>
<td>156.9</td>
<td>2013</td>
<td>233.0</td>
</tr>
<tr>
<td>1997</td>
<td>160.5</td>
<td>2014</td>
<td>236.7</td>
</tr>
<tr>
<td>1998</td>
<td>163.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The beauty of this table is that we can easily compare any two years prices.

**Example 1**

We can see that in 1990, it would cost $130.70 for goods and services costing $99.60 in 1983.

To state differently, in 1990, $130.70 would buy the same goods and services (on average) as $99.60 would in 1983.

Finally, we can say $99.60 in 1983 is equivalent to $130.70 in 1990, or 99.60 (1983 $) is equivalent to 130.70 (1990 $).

Further interpretation of the CPI allows the comparison of prices in different years.
Example 2 Why should we convert to constant dollars?

According to http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPM0_PTE_NUS_DPG&f=A, the average price of gasoline in 2008 was $3.30 per gallon on average. In 2014, it averaged $3.44. Was gasoline more expensive or less expensive in 2014, relative to the time?

On the face of it, it seems that gas is more expensive in 2014.

Prices from a given year are called nominal or current prices.

Nominally, gas was more expensive in 2014.

To compare the prices taking into account the changing value of money, we convert one of the prices to the same year as the other. Usually, we convert forward to the more recent year, but one could also convert back to the earlier year.

In this example, the goal is to convert $3.30 2008 dollars to its equivalent amount in 2014 dollars.

The CPI table reveals: 215.30 2008 dollars is equivalent to 236.70 2014 dollars.

The goal here is: 3.30 2008 dollars is equivalent to ??? 2014 dollars

One way to solve is to set up the proportion:

\[
\frac{236.70 \text{ (2014 dollars)}}{215.30 \text{ (2008 dollars)}} = \frac{x \text{ (2014 dollars)}}{3.30 \text{ (2008 dollars)}}
\]

Solving this proportion (by multiplying both sides by 3.30), we get:

\[
3.30 \text{ (2008 dollars)} \cdot \frac{x \text{ (2014 dollars)}}{3.30 \text{ (2008 dollars)}} = \frac{236.70 \text{ (2014 dollars)}}{215.30 \text{ (2008 dollars)}} \cdot 3.30 \text{ (2008 dollars)}
\]

Notice that on the right, the labels (2008 dollars) reduce and we are left with the desired unit (or label) of 2014 dollars.

\[
3.30 \text{ (2008 dollars)} \cdot \frac{x \text{ (2014 dollars)}}{3.30 \text{ (2008 dollars)}} = \frac{236.70 \text{ (2014 dollars)}}{215.30 \text{ (2008 dollars)}} \cdot 3.30 \text{ (2008 dollars)}
\]

Now, we have:

\[
x \text{ (2014 dollars)} = \frac{236.70-3.30}{215.30} \text{ (2014 dollars)}
\]

\[
x \text{ (2014 dollars)} = 3.628 \ldots \text{ (2014 dollars)}
\]

Thus, $3.30 in 2008 was equivalent to $3.63 in 2014.

In other words, when Americans paid $3.30 per gallon for gasoline in 2008, it was equivalent to someone paying $3.63 in 2014, which is more than what they were actually paying in 2014 ($3.44).

So gasoline was actually cheaper in 2014 than it was in 2008, relative to the time.

The equivalent price in a different year is in real or constant dollars.
In the last example, economists would say that the price of gasoline in 2008 was $3.63 in constant 2014 dollars. The process of converting is often called converting to constant dollars.

Another way to conceptualize the conversion to constant dollars (which will be especially useful when converting a whole series of values in Excel) is to focus on the meaning of the ratio of the CPI's. In the case of the example,

\[ \frac{3.30 \text{ (2008 dollars)}}{215.30 \text{ (2008 dollars)}} \cdot \frac{236.70 \text{ (2014 dollars)}}{215.30 \text{ (2008 dollars)}} = x \text{ (2014 dollars)} \]

This ratio in bold represents how many times more one 2008 dollar was worth in 2014. Multiplying the 2008 price by this ratio yields how much that 2008 price was worth in 2014.

In general,

\[ \text{any price (in its original (yr E) $)} \cdot \frac{\text{newer (yr N) CPI value}}{\text{earlier (yr E) CPI value}} = x \text{ (in constant N $)} \]

This above formula for conversion is what we will be primarily using, especially in Excel when we are converting an entire spreadsheet of values to constant dollars.

**Example 3** In 2014 (and still currently), US Senators make an annual salary of $174,000. Twenty years ago in 1994, US Senators made an annual salary of $133,600. Which salary was better, relative to the time?

**Example 4** Consider the data on the median price of new single-family homes ([http://faculty.elgin.edu/nscherger/Courses/MTH104/Median Price of New Homes 1979-2009.xls](http://faculty.elgin.edu/nscherger/Courses/MTH104/Median Price of New Homes 1979-2009.xls)).

The graph in nominal dollars is:

Looking at the graph above, one sees the general increases in prices from 1979 through 2009, after which we see the prices start to decrease (marking the beginning of the housing market decline). However,

**Key Idea** In order to compare costs on varying years, we need to convert all the prices to constant dollars. Data and graphs in constant dollars will give a more accurate understanding of prices.
First, paste the CPI values for the years 1979-2009 in column C, next to the nominal prices, as shown:

<table>
<thead>
<tr>
<th>Year</th>
<th>Median Sales Price</th>
<th>CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>$62,900.00</td>
<td>72.6</td>
</tr>
<tr>
<td>1980</td>
<td>$64,600.00</td>
<td>82.4</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>2008</td>
<td>$232,100.00</td>
<td>215.3</td>
</tr>
<tr>
<td>2009</td>
<td>$216,700.00</td>
<td>214.5</td>
</tr>
</tbody>
</table>

Next, compute the costs in constant 2009 dollars in column D by multiplying by the ratio of the CPI’s involved, as discussed in the yellow highlighted formula on the last page.

Note that the value for the CPI in 2009 will be the same in all cells, so it is typed as a constant, but the other values are cell-referenced, so we can conveniently fill that column.

Then, fill column D, and the result should be:

<table>
<thead>
<tr>
<th>Year</th>
<th>Median Sales Price</th>
<th>CPI</th>
<th>Median Sales Price in Constant 2009$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>$62,900.00</td>
<td>72.6</td>
<td>=B6*214.5/C6</td>
</tr>
<tr>
<td>1980</td>
<td>$64,600.00</td>
<td>82.4</td>
<td>$185,840.91</td>
</tr>
<tr>
<td>1981</td>
<td>$68,900.00</td>
<td>90.9</td>
<td>$168,163.83</td>
</tr>
<tr>
<td>1982</td>
<td>$69,300.00</td>
<td>96.5</td>
<td>$154,039.90</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>2008</td>
<td>$232,100.00</td>
<td>215.3</td>
<td>$231,237.58</td>
</tr>
<tr>
<td>2009</td>
<td>$216,700.00</td>
<td>214.5</td>
<td>$216,700.00</td>
</tr>
</tbody>
</table>

Now obtain a scatterplot in constant 2009 dollars:

![New Single Family Homes Sold in the US (In Constant 2009 Dollars)](image)

The graph in constant dollars tells a different and more realistic story of median home prices.
Interpreting a constant dollars graph:

- Regions where a constant dollars graph is increasing:
  From 1985-1989 and from 2000-2005, the nominal home prices were increasing, and the constant dollars were also increasing, which means these home prices were increasing more than the increases of other goods and services over this time period.

- Regions where a constant dollars graph is constant:
  From 1992-1997, the nominal home prices were increasing, but the constant dollars were relatively constant, which means these home prices were increasing at a rate that was just keeping pace with the increases of other goods and services during this time period.

- Regions where a constant dollars graph is decreasing:
  From 1979-1982 and from 2005-2007, the nominal home prices were increasing, but the constant dollars were decreasing, which means these home prices were increasing less than the increases of other goods and services over this time period.

Less commonly,
  From 1989-1992, the nominal home prices were relatively constant, but the constant dollars was decreasing, which means that the actual value of these unchanged home prices was decreasing when compared to the other rising costs of the time period.
  From 2007-2009, the nominal home prices were actually decreasing and the constant dollars were decreasing more quickly than the other regions of decrease described above, which means that the actual value of these decreasing home prices was dropping even faster than what those nominal decreases show, because other costs were continuing to rise during this time.
Example 5 The graph below shows the salaries of United States Senators in constant 2012 dollars.

Noting that in nominal dollars, the salaries for United States Senators have either remained the same or increased from 1940 to 2012 (see https://www.senate.gov/artandhistory/history/common/briefing/senate_salaries.htm), answer the following:

a. Were the salary increases for US Senators rising less than, more than, or about the same as what other goods and services were increasing from 1997-2004?

b. Were the salary increases for US Senators rising less than, more than, or about the same as what other goods and services were increasing from 1969-1975?

c. Were the salary increases for US Senators rising less than, more than, or about the same as what other goods and services were increasing from 1989-1991?

Inflation Rate is the percentage change in the annual CPI.

Example 5 The inflation rate in 1996 was \[
\frac{1996\ CPI - 1995\ CPI}{1995\ CPI} = \frac{162.6 - 157.9}{157.9} = 3.0\% \]
Group Activity 16: CPI

1. These questions require the CPI table at http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx. A portion of this table is copied at the right for your convenience.
   a. The official definition of “poverty level” is supposed to be adjusted for the effects of inflation. For a family of four, the poverty level was an income of $3,743 in 1970 and $5,050 in 1975. Convert the 1970 poverty level into 1975 dollars.
   
   
<table>
<thead>
<tr>
<th>YEAR</th>
<th>CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>38.8</td>
</tr>
<tr>
<td>1971</td>
<td>40.5</td>
</tr>
<tr>
<td>1972</td>
<td>41.8</td>
</tr>
<tr>
<td>1973</td>
<td>44.4</td>
</tr>
<tr>
<td>1974</td>
<td>49.3</td>
</tr>
<tr>
<td>1975</td>
<td>53.8</td>
</tr>
</tbody>
</table>

   b. Which poverty level income ($3,743 in 1970 and $5,050 in 1975) has greater buying power in their respective years?

2. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Minimum Wage 1938-2013.xlsx, which contains the minimum wage rate for the US starting in 1938 when it was instituted. (Note: An excellent source of information on the minimum wage in the United States is Department of Labor’s webpage on the topic, http://www.dol.gov/dol/topic/wages/minimumwage.htm. Please note that many states also have minimum wage laws. In cases where an employee is subject to both the state and federal minimum wage laws, the employee is entitled to the higher of the two minimum wages.)
   a. Make a scatterplot (choose the 2nd option, with connecting lines) of this data, including a title and x and y axes labels. Copy and paste this plot into your Word document.
   b. Identify two periods of time in which the minimum wage rate increased significantly.
   c. Identify two (long) periods of time in which the minimum wage did not increase at all.

3. The graph at the right was widely distributed in September of 1997 when the minimum wage hike was instituted. While it is very attractive, the graph contains a serious error. Compare this graph with the one that you just (correctly) made in (2a).

   What is the error in this graph? What impression does this graph give?

4. Is the graph you made in (2a) an accurate portrayal of the buying power of the minimum wage? Explain.
5. We want to make a graph of the minimum wage rate in constant 2014 dollars. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx, a file that contains the consumer price index (CPI-U base year 1982-84) from 1913 to the present.

Paste the CPI values from 1938 to 2014 in the column next to the minimum wage column in the minimum wage file you opened in (2).

Using the ideas discussed in the notes in the previous section, add a column which computes the minimum wage in constant 2014 dollars.

Copy the first ten rows of this table into your Word document.

6. Make a scatterplot (choose the 2nd option, with connecting lines) of the minimum wage in constant 2014 dollars, including a title and x and y axes labels. Copy and paste this plot into your Word document.

7. If you were earning the minimum wage, what year do you wish you were living in? Why?

8. Look at years in your first graph in (2a) when the actual minimum wage rate did not change (for example, the periods of time you answered in 2c.) Now look at the graph you just made in (6) and describe what happened to the minimum wage rate in constant dollars during those years when it remained the same in nominal dollars. Explain why this pattern occurs.

9. This data seems to be difficult for the media to handle. Time Magazine on Nov. 29, 1999 published a graph of the very same data, which is also incorrect.

Compare their constant dollars graph (their dollars graph is in 1999 dollars) with the constant graph that you just (correctly) made in (6).

Describe what is wrong with their constant dollars graph.
Assignment 9: Consumer Price Index

Complete Assignment # 9 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. The movie industry has been accused of dramatically over-inflating the cost of tickets. Open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/Movie Ticket Prices 1983-2014.xlsx.
   a. Create a scatterplot of the data, including appropriate axes labels and a title (remember...this data is in what we call nominal dollars). Copy and paste this graph into your Word document.
   b. What is the general impression of the graph?

2. Now suppose that you work for the movie industry and have just seen the graph that you created in (#1a) in a newspaper article in which the industry was criticized for over-inflating the cost of tickets. Remembering what you learned about CPI from your liberal arts mathematics course, you decide to submit to the newspaper a more accurate graph of the cost of movie ticket prices in CONSTANT 2014 DOLLARS.
   a. In column C of the Movie Ticket Prices spreadsheet from (#1), paste the CPI values (at http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx) for the corresponding years. Then, in column D, perform the necessary operations to fill this column with the ticket prices in constant 2014 dollars. Copy and paste this table into your Word document.
   b. Now, create a scatterplot of movie ticket prices in constant 2014 dollars. Include axes labels and a title. Paste this graph into your Word document. Observe how its general impression is very different than that of the first graph.

3. Recall that the inflation rate is defined to be the percentage change in the annual CPI. For example, the CPI in 1998 was 163.0 while the CPI in 1999 was 166.6. The inflation rate for 1999 was therefore (166.6-163.0)/163.0 or 2.2%. The inflation rate is important for many reasons. High inflation means that the money we earn and have saved up in the past is worth less. Interest rates for houses, cars, and credit cards rise if the inflation rate is high. For a variety of reasons, stock prices tend to fall even on hints of higher inflation. Increases for many government entitlement programs are tied to the inflation rate.
   a. Again, open the file http://faculty.elgin.edu/nscherger/Courses/MTH104/CPI.xlsx. Add a new column to the table that contains the annual inflation for each year. Paste the resulting table only for the last 40 years 1974-2014 into your Word document.
   b. Make a scatterplot (choose the 2nd option with the connecting lines) of the inflation rate from 1974 to 2014. Paste it in your Word document.
   c. Although there are always fluctuations, how would you generally compare the inflation rates of the 1970s and 1980s with the inflation rates of the 1990s and 2000s?

If I had to populate an asylum with people certified insane, I’d just pick ‘em from all those who claim to understand inflation.

-Will Rogers
Appendix A
Practice Test Answers

Unit 3 Topics

This test covers material presented in Chapters 11-14.

1. Examine data (presented in a table format) and decide if the data represent a linear relationship.
2. Write a linear model/equation (find the equation of a line); Identify and interpret the slope and y-intercept.
3. Create a scatterplot of data and use Excel to find the linear trendline and $R^2$ value.
4. Use the linear model/trendline to interpolation and extrapolation.
5. Use function vocabulary of independent variable, dependent variable, domain, and range, including identifying each of those, given a graph.
6. Create a scatterplot of data and use Excel to find the quadratic trendline (degree 2 polynomial) or a cubic (degree 3 polynomial) and $R^2$ value.
7. Determine the most reasonable choice for a model to fit to a scatterplot.
8. Critically evaluate the appropriateness of predictions made from regression equations (linear and polynomial); this should include an analysis of both the $R^2$ value and what your “common sense” tells by looking at the data itself.
9. Determine if data in a table is linear, exponential, or neither, including identifying the slope (if linear) and the growth/decay factor (if exponential).
10. Write the basic equation for exponential models: $y=Ab^x$, where $b=1+r$; Answer questions using this model.
11. Use Excel to create tables showing models of exponential growth or decay over a period of time.
12. Use Excel to create tables for the specific applications of exponential growth and decay of including doubling time and half life, including Carbon 14 dating.
13. Use the CPI to compare prices in two different years, by converting prices forward from one year in the past to a more recent year.
14. Use the CPI to convert an entire series of prices to constant dollars in Excel.
15. Analyze and compare graphs in nominal dollars and constant dollars, as discussed in class.
16. Compute the inflation rate, which is the percentage change in the annual CPI.
Practice Test 3

1. Use the graph at the right to answer the following questions.
   a. What is the \( y \)-intercept of the line?
   b. Interpret the \( y \)-intercept, in practical terms of the problem.
   c. What is the slope of the line?
   d. Interpret this slope, in practical terms of the problem.
   e. State the equation of this line.
   f. Use this equation to determine the total daily pay of someone who had $430 in sales. Show your work and round your answer to the nearest penny.

2. Use the graph at the right to answer the questions below.
   a. What is the independent variable?
   b. What is the dependent variable?
   c. What is the domain? (You will need to estimate values.)
   d. What is the range? (You will need to estimate values.)
   e. What would be the most appropriate type of regression equation (linear, quadratic, or cubic) to fit to this data?
3. The value of an investment has shown a consistent rate of growth of 3.5% each year. The initial investment was $4500.
   a. Write the exponential model/equation.
   b. Use your equation from part (a) to determine how much the investment will be worth in 25 years. Round your answer to the nearest penny.

4. The presidential salary in 1974 was $250,000. Forty years later in 2014, it was $400,000.
   a. Convert the 1974 salary to constant 2014 dollars. Show work.
   b. Which salary was better, 1974 or 2014 (relative to the time)? Explain your reasoning.

5. Use the graph to answer the questions below.
   a. In the late 1980s, the (nominal) price for a Superbowl commercial did increase (the nominal prices were 1986=$550,000; 1987=$600,000; 1988=$645,000; 1989=$675,000; 1990=$700,000); however, observe that in the constant 2014 dollars graph shown, the price of a commercial remained relatively constant. What does that say about the increases seen in the late 1980s?
   b. In the late 1990s, the (nominal) price for a Superbowl commercial was also increasing (the nominal prices were 1996=$1,085,000; 1997=$1,200,000; 1998=$1,300,000; 1999=$1,600,000; 1990=$1,900,000). Here, observe that the constant 2014 dollars graph, the price of a commercial was also increasing. What does that say about the increases seen in the late 1990s?
Use Excel to Answer the Remaining Questions

6. Open the file https://faculty.elgin.edu/nscherger/courses/MTH104/Lyme Disease Totals 1990-2013.xlsx, which has the number of cases of Lyme Disease in the US from 1990 to 2013.
   a. Make a scatterplot of the data, using columns B and C. Include a title and axes labels.

      Add a linear trendline to your plot. Be sure to display the equation and the $R^2$-value. Copy and paste this plot into your Word document.

      b. Use your linear regression equation to predict how many cases of Lyme Disease there would be in 2015 (noting that $x = 0$ corresponds to 1990). Show work, round your answer to the nearest integer, and label your answer.

7. The following data represents the cost of a senior citizen admission to a movie and the annual profit in thousands that the theater obtains from senior citizens at that price point.

<table>
<thead>
<tr>
<th>Cost of a Senior Citizen Movie Admission</th>
<th>Theater's Annual Profit from Senior Citizens in Thousands</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3.00</td>
<td>$308</td>
</tr>
<tr>
<td>$3.50</td>
<td>$352</td>
</tr>
<tr>
<td>$4.00</td>
<td>$376</td>
</tr>
<tr>
<td>$4.50</td>
<td>$385</td>
</tr>
<tr>
<td>$5.00</td>
<td>$370</td>
</tr>
<tr>
<td>$5.50</td>
<td>$338</td>
</tr>
</tbody>
</table>

   c. Make a scatterplot of the data. Include a title and axes labels.

      Add a quadratic regression equation to your plot. Be sure to display the equation and the $R^2$-value. Copy and paste this plot into your Word document.

   d. Use your quadratic regression equation to predict what the theater’s projected annual profit from senior citizens would be if the cost of admission was $4.25. Show work, round your answer to the nearest thousandth, and label your answer.

8. Which equation fit its corresponding data set better, the linear equation in number 1 or the quadratic equation in number 2? Mathematically justify your answer.

9. According to globalhealthfacts.org, in 2007, the population of South Africa was 43,997,828 and was decreasing by 0.46% per year; The population of Kenya was 36,913,721 and was increasing by 2.8% per year. Set up 3 columns in Excel (year, South Africa’s population, and Kenya’s population). If these rates continued to be true, when would Kenya’s population exceed South Africa’s population?

   State the year, and copy and paste the portion of the table up to the row where you determined your answer into your Word document.
10. Using the same data set that you just created in exercise 4, extend your table as far as necessary to determine the following. (You only need to state your answer here and you do not need to copy and paste the table from Excel.)

e. What is the doubling time for Kenya’s population?

f. What is the half-life for South Africa’s population?
Chapter 15: Doubling Time & Rule of 70

Objectives:
1. Use Excel spreadsheets to identify doubling times.
2. Recognize the approximate pattern between the growth rate and the doubling time, known as the rule of 70.
3. Use the Rule of 70 to approximately doubling time.

Doubling Time
Recall that doubling time is the amount of time it takes a population to double. Or, to say it differently, it is the amount of time for a population to grow by 100%.

Key Idea
If a quantity is growing exponentially it has constant doubling time.

Mathematically it is reasonably straightforward to show that if a quantity is growing exponentially it has constant doubling time:

Start with the general exponential model, \( y = Ab^x \) and let \( b > 1 \), so we have exponential growth.

Then, remembering that \( b \) is a constant, there must be some number \( k \) such that \( b^k = 2 \).

\( \text{(For example, if } b = 4, \text{ then } k = \frac{1}{2}, \text{ because } 4^{\frac{1}{2}} = \sqrt{4} = 2. \)\)

Working with the original model using some basic algebra, observe that:

\[
\begin{align*}
y &= Ab^x \\
    &= Ab^{\frac{x}{k}} \\
    &= A(b^{k})^{\frac{x}{k}}
\end{align*}
\]

Substituting 2 for \( b^k \):

\[
\begin{align*}
y &= A(2)^{\frac{x}{k}}
\end{align*}
\]

The last expression means that if \( x \) is increased by \( k \) then \( y \) will be multiplied by 2, no matter where we start.

So, there is a constant doubling time = \( k \).

Explicitly, say you are at \((x_0, y_0)\):

\[
\begin{align*}
y_0 &= A(2)^{\frac{x_0}{k}}
\end{align*}
\]

If \( x_0 \) is increased by \( k \), then \( y \) becomes \( y = A(2)^{\frac{x_0+k}{k}} = A(2)^{\frac{x_0}{k} + 1} = A(2)^{\frac{x_0}{k}} \cdot 2 = y_0 \cdot 2 \)
Now, let’s look at some examples with data to illustrate the phenomenon.

**Example 1**

A town’s population is growing by 7% per year and it currently has 1000 residents.

Create a table in Excel to discover the doubling time and observe that it is a constant doubling time.

By creating a table in Excel with the year and the population, one should get the table at the right.

It can be seen from the rows in bold that the population doubles to 2000 somewhere between 10 and 11 years.

It can then be observed that this doubling time is constant, because the population doubles again to 4000 somewhere between 20 and 21 years and again to 8000 somewhere between 30 and 31 years.

**Example 2**

If oil consumption in the US is increasing at a rate of about 2% per year. What is its doubling time? By what percent will oil consumption increase in a decade?

Again, by creating a table in Excel (which is started at the right) and extending the table, one should see that the price doubles from 1 (where any number could have been used as the initial value) to 2 between 35 and 36 years.

For the second question, after 10 years, the price has increased from 1 to about 1.22, which is an increase of 22%.
Doubling Time and the Rule of Seventy

Observe the last two examples:

When the rate was 7%, the doubling time was about 10.
When the rate was 2%, the doubling time was about 35.

Can you make a hypothesis about the relationship between rate and doubling time?

Let’s add one more example to make the hypothesis become clearer:

Example 3

Suppose that a city’s population is increasing by 10% per year.
What is the doubling time?

The Excel spreadsheet (shown at the right) indicates that the population would double somewhere between 7 and 8 years.

Now observe the following three examples:

When the rate was 7%, the doubling time was about 10.
When the rate was 2%, the doubling time was about 35.
When the rate was 10%, the doubling time was about 7.

Can you make a hypothesis about the relationship between rate and doubling time?

It seems that ______ divided by ______ approximately gives the doubling time.

This relationship between percentage growth and doubling time known as the rule of seventy...

Rule of 70:

If a quantity is growing exponentially (at a rate of \( x\% \) per time period), the doubling time for that quantity is approximately \( 70/x \).

Caution: Unlike most mathematical formulas, the Rule of 70 is based on this \( x\% \) NOT being converted to its decimal form.

Example 4 If an investment increases the amount of money invested at 5% per year, approximately how long will it take it to double your money?
Group Activity 17: Introduction to Savings, Doubling Time, and Rule of 70

One of the places all of us encounter exponential growth is in banking and finance. Almost everyone has a savings account in a bank, and almost everyone needs a loan to make a large purchase (such as a car) at some time. Our next unit will study finance in detail.

When you put money in a bank, the bank does not stash it away in a safe (not for very long in any case). The bank usually lends it out. In exchange for letting the bank use your money, the bank pays you interest, a percentage of what you deposited paid at regular intervals. If you do not withdraw the money, you will earn interest on your interest, a process called compounding.

Compounding is exponential. To see why, suppose you have $A$ dollars in your account, and suppose you will receive $r\%$ interest at the next compounding period. Your balance after the interest payment is $A + r^*A$ or $A*(1+r)$. Thus, receiving $r\%$ interest is equivalent to increasing by a factor of $(1+r)$ [i.e., multiplying by $(1+r)$].

This activity explores the phenomenon of exponential growth in this context.

1. Suppose you invest $10,000. Make a table like the one below showing the value of this investment at 1%, 2%, 3%, ..., interest assuming you do not withdraw anything for 25 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>1%</th>
<th>2%</th>
<th>3%</th>
<th>4%</th>
<th>5%</th>
<th>6%</th>
<th>7%</th>
<th>8%</th>
<th>9%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$10,000</td>
</tr>
<tr>
<td>1</td>
<td>$10,100</td>
<td>$10,200</td>
<td>$10,300</td>
<td>$10,400</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>$10,201</td>
<td>$10,402</td>
<td>$10,603</td>
<td>$10,804</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>$10,402</td>
<td>$10,810</td>
<td>$11,222</td>
<td>$11,636</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Copy this table into your Word document.

2. After 25 years, how many more times is the 10% investment worth than at the start?

3. How many more times is the 1% investment worth after 25 years?

4. Make a bar graph with B1 to K1 as the x-axis labels and the value of the investment after 25 years for the column heights (i.e., B27 to K27).

Hint: It is probably easiest to select only the range B28 to K28 and then the x-axis should appear with the appropriate values 1 through 10.

It is very important here to include appropriate axes labels that tell the reader that the x-axis is interest rate as a percent and the y-axis is the value of the investment in dollars. Also include an appropriate title indicating that this graph shows the value of the investment after 25 years.

Copy and paste this plot into your Word document.
5. When making investments or taking out loans, should you be concerned about differences in interest rates? Why or why not?

6. Study the last column (10%) carefully. Approximately how long does it take for the investment to double (reach $20,000)? After it reaches, $20,000 how long does it take for the investment to double again (reach $40,000)? How long does it take to double again?

7. The previous question might cause you to surmise that an investment always takes the same amount of time to double no matter how large the investment is. In fact, as discussed in class: Anything that grows exponentially has a constant doubling time; however, the doubling time varies according to the rate of increase.

Complete the following table. You may need to extend some columns in your Excel table in (1) to be able to fill in some of the doubling times below.

<table>
<thead>
<tr>
<th>Annual Percentage Rate of Increase</th>
<th>Approximate Doubling Time (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>9%</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
</tr>
</tbody>
</table>

Include the above table in your Word document.

8. The table you made is well known to bankers and investors alike. It is what was discussed in class and is known as the Rule of 70. To see it, you can take 70/APR (the rate in the first column) and you should approximately get the answer you gave in the second column. The Rule of 70 applies to any form of exponential growth and is useful for making estimations.

For example, in its early years, the US grew nearly exponentially at approximately 2.8% per year. The population of the US according to the first census in 1790 was 3.9 million.

a. Using the rule of 70, what was the approximate doubling time for the population of the US during this time?

b. Repeatedly use this doubling time to estimate the population of the US in 1865.
Chapter 16: Savings Accounts and Compounding

Objectives:
1. State the difference between simple and compound interest.
2. Model various compound interest growth problems involving lump-sum investments, in Excel.
3. Identify the difference between APR and APY.
4. Determine APY.
5. Answer questions comparing accounts (involving APR and APY).

Simple Interest Versus Compound Interest

Most of these finance-related concepts will be most easily understand through looking at examples. Therefore, let’s start with a simple case.

**Example 1** Suppose you have $1000 that you are going to invest at a rate of 4% interest per year.

- How much will you have in interest at the end of the year?

- If you continue to earn interest in this same way ($40/year), how much money would you have total by the end of 5 years?

This first example is what is known as:

**Simple interest**, where interest is only calculated based on the value of the original amount of the investment, which called the principal. *(Think of simple interest as the bank writing you a check for your interest and sending it home to you.)*

Now suppose that instead of the bank sending you a check at the end of each year, they just deposit the interest that you earned into your savings account. Let’s consider what would happen with the same example:

**Example 2** Suppose you have $1000 that you are going to invest at a rate of 4% interest per year. The first year, you still earn .04*1000 = $40 in interest, just like before. But now suppose the bank deposits the $40 into your account. How much interest would you earn and what would your balance be at the end of the 2nd year? 3rd year? 4th year? 5th year?

A nice way to solve this problem would be to set up an Excel spreadsheet like:

<table>
<thead>
<tr>
<th>Year</th>
<th>Balance at Beginning of Year</th>
<th>Interest Earned</th>
<th>Balance at End of Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,000.00</td>
<td>=B2*.04</td>
<td>=B2+C2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If you then fill in the columns, you should get:

<table>
<thead>
<tr>
<th>Year</th>
<th>Balance at Beginning of Year</th>
<th>Interest Earned</th>
<th>Balance at End of Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,000.00</td>
<td>$40.00</td>
<td>$1,040.00</td>
</tr>
<tr>
<td>2</td>
<td>$1,040.00</td>
<td>$41.60</td>
<td>$1,081.60</td>
</tr>
<tr>
<td>3</td>
<td>$1,081.60</td>
<td>$43.26</td>
<td>$1,124.86</td>
</tr>
<tr>
<td>4</td>
<td>$1,124.86</td>
<td>$44.99</td>
<td>$1,169.86</td>
</tr>
<tr>
<td>5</td>
<td>$1,169.86</td>
<td>$46.79</td>
<td>$1,216.65</td>
</tr>
</tbody>
</table>

This second example is what is known as:

**Compound interest**, where interest is calculated based on both the original principal and the interest that was added in. *(Here, you are earning interest on your interest.)*

- Which do you prefer? Simple or Compound Interest?
  
  It should be clear that compound interest will earn you more money.

- You may be wondering, which type do banks typically use?
  
  Typically, banks will use compound interest; however, savings bonds (typically issued by US treasury with a guaranteed rate and also certain tax benefits) typically pay simple interest.

**Compound Interest in Depth**

Now suppose that instead of the bank calculating your interest just once a year and depositing it into your account, they tell you that interest will be **compounded quarterly**, or 4 times a year. Let’s consider what would happen with the same example:

**Example 3** Suppose you have $1000 that you are going to invest at a rate of 4% interest per year, compounded quarterly. This does not mean that the bank will give you 4% interest 4 times a year (a common misunderstanding). Instead the quarterly interest rate = 4%/4 = 1%.

*Generally, you will divide the interest rate by the number of compoundings per year to find the interest rate that is used per “term.”*

Again, without knowing any formulas, you can use your basic knowledge of Excel to model this problem. You would start by setting up an Excel spreadsheet like:

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,000.00</td>
<td>=.04/4*B2</td>
<td>=B2+C2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>=D2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Realize that as we fill in the spreadsheet, year 1 will be at the end of the 4th quarter, year 2 will be at the end of the 8th quarter, etc... . Although this problem is a more involved (and would be much more complicated by hand), with Excel, it just means extending the table to the 20th quarter, as shown:
### Unit 4: Finance
#### Chapter 16: Savings Accounts and Compounding

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,000.00</td>
<td>$10.00</td>
<td>$1,010.00</td>
</tr>
<tr>
<td>2</td>
<td>$1,010.00</td>
<td>$10.10</td>
<td>$1,020.10</td>
</tr>
<tr>
<td>3</td>
<td>$1,020.10</td>
<td>$10.20</td>
<td>$1,030.30</td>
</tr>
<tr>
<td>4</td>
<td>$1,030.30</td>
<td>$10.30</td>
<td>$1,040.60</td>
</tr>
<tr>
<td>5</td>
<td>$1,040.60</td>
<td>$10.41</td>
<td>$1,051.01</td>
</tr>
<tr>
<td>6</td>
<td>$1,051.01</td>
<td>$10.51</td>
<td>$1,061.52</td>
</tr>
<tr>
<td>7</td>
<td>$1,061.52</td>
<td>$10.62</td>
<td>$1,072.14</td>
</tr>
<tr>
<td>8</td>
<td>$1,072.14</td>
<td>$10.72</td>
<td>$1,082.86</td>
</tr>
<tr>
<td>9</td>
<td>$1,082.86</td>
<td>$10.83</td>
<td>$1,093.69</td>
</tr>
<tr>
<td>10</td>
<td>$1,093.69</td>
<td>$10.94</td>
<td>$1,104.62</td>
</tr>
<tr>
<td>11</td>
<td>$1,104.62</td>
<td>$11.05</td>
<td>$1,115.67</td>
</tr>
<tr>
<td>12</td>
<td>$1,115.67</td>
<td>$11.16</td>
<td>$1,126.83</td>
</tr>
<tr>
<td>13</td>
<td>$1,126.83</td>
<td>$11.27</td>
<td>$1,138.09</td>
</tr>
<tr>
<td>14</td>
<td>$1,138.09</td>
<td>$11.38</td>
<td>$1,149.47</td>
</tr>
<tr>
<td>15</td>
<td>$1,149.47</td>
<td>$11.49</td>
<td>$1,160.97</td>
</tr>
<tr>
<td>16</td>
<td>$1,160.97</td>
<td>$11.61</td>
<td>$1,172.58</td>
</tr>
<tr>
<td>17</td>
<td>$1,172.58</td>
<td>$11.73</td>
<td>$1,184.30</td>
</tr>
<tr>
<td>18</td>
<td>$1,184.30</td>
<td>$11.84</td>
<td>$1,196.15</td>
</tr>
<tr>
<td>19</td>
<td>$1,196.15</td>
<td>$11.96</td>
<td>$1,208.11</td>
</tr>
<tr>
<td>20</td>
<td>$1,208.11</td>
<td>$12.08</td>
<td>$1,220.19</td>
</tr>
</tbody>
</table>

- **Which do you prefer?** Interest compounded annually or interest compounded quarterly?
  
  It should be clear that the more often interest is compounded, the more interest you will earn. The differences become more dramatic over longer lengths of time as well as when you continue to deposit more money in your account, which we will revisit later.

- **You may be wondering, which type do banks typically use?**
  
  Compounding varies and could be annually, semi-annually, quarterly, monthly, etc… . It is important to ask this question, because simply looking at the interest rate can be deceiving.
Unit 4: Finance  
Chapter 16: Savings Accounts and Compounding

Annual Percentage Rate (APR) and Annual Percentage Yield (APY)

Some people mistakenly think that these two terms are the same. Along with knowing how often interest is compounded, knowing the difference between APR and APY is important when comparing accounts.

- **The annual percentage rate (APR)** (sometimes called the nominal rate) is what was referred to in the previous examples as the *annual interest rate* (which was the 4%).

However, as we observed, depending on how often interest is compounded, varying amounts of interest is made (under the same APR).

In our examples, under the same APR of 4%, when interest was compounded annually, $40 was made in interest after 1 year; when interest was compounded quarterly, $40.60 was made in interest after 1 year. In order to compare how much money was actually accrued in interest, let’s determine the percent growth of each after 1 year. (Remember: percent change = change in value / original value.)

When interest was compounded annually, the annual percent growth was 4%.

When interest was compounded quarterly, the annual percent growth = ________

Thus we actually yielded 4.06% on our principal when interest was compounded quarterly; whereas, we yielded 4% on our principal when interest was compounded annually.

- Computing this rate that is actually yielded in 1 year from an investment is what is known as the **annual percentage yield (APY).**

- **APR vs APY**
  - The **annual percentage rate (APR)** (or the **nominal rate**) is the annual interest rate without taking into account compound interest.
  - The **annual percentage yield (APY)** (or the **effective rate**) is the actual percent which the balance increases in one year (taking into account the compound interest).
    \[
    \text{APY} = \frac{\text{change in balance}}{\text{original balance}} \quad \text{~OR~} \quad \frac{\text{amount earned in interest}}{\text{principal}}
    \]
  - Comparing APR and APY
    - If interest is compounded annually: \( \text{APY} = \text{APR} \)
    - If interest is compounded more than once a year: \( \text{APY} > \text{APR} \)
  - Neither APR nor APY depend on the starting principal.
  - You may be wondering, which type do banks typically use?

Often, they publish both; however, sometimes, if they want to make their rate look higher, they will use the APY. If you are comparing two different accounts, you should ask to see the APY of both, because this is what your money actually earns.
Example 4 You have a choice between two different savings plans. The first advertises an APR of 5% and interest is compounded quarterly. The second advertises they compound semi-annually and have an APY of 5%. Without computing anything, which would you choose and why?

Example 5 If in the previous example the second bank advertises an APY of 5% with interest compounded monthly, would that change your answer? Why?

Example 6 Now suppose you have the following two choices. The first advertises an APY of 6.05%, where interest is compounded monthly. The second advertises an APR of 6% and interest is compounded monthly. Which do you choose and why?

Again, we want to compare APY’s, and the first is given. Looking at the second choice, we know it will be higher than 6%, but now we need some exact computation to be able to compare with the first APY of 6.05%. Again, let’s set up an Excel spreadsheet. Because APR and APY do not depend on the principal, we can use any amount; let’s use $100. Remember, we need to use the monthly interest rate = 6%/12 = 0.5%. and set up the following table:

<table>
<thead>
<tr>
<th>Month</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$100.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>=D2</td>
<td>=.06/12*B2</td>
<td>=B2+C2</td>
</tr>
</tbody>
</table>

Filling in the table:

<table>
<thead>
<tr>
<th>Month</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$100.00</td>
<td>$0.50</td>
<td>$100.50</td>
</tr>
<tr>
<td>2</td>
<td>$100.50</td>
<td>$0.50</td>
<td>$101.00</td>
</tr>
<tr>
<td>3</td>
<td>$101.00</td>
<td>$0.51</td>
<td>$101.51</td>
</tr>
<tr>
<td>4</td>
<td>$101.51</td>
<td>$0.51</td>
<td>$102.02</td>
</tr>
<tr>
<td>5</td>
<td>$102.02</td>
<td>$0.51</td>
<td>$102.53</td>
</tr>
<tr>
<td>6</td>
<td>$102.53</td>
<td>$0.51</td>
<td>$103.04</td>
</tr>
<tr>
<td>7</td>
<td>$103.04</td>
<td>$0.52</td>
<td>$103.55</td>
</tr>
<tr>
<td>8</td>
<td>$103.55</td>
<td>$0.52</td>
<td>$104.07</td>
</tr>
<tr>
<td>9</td>
<td>$104.07</td>
<td>$0.52</td>
<td>$104.59</td>
</tr>
<tr>
<td>10</td>
<td>$104.59</td>
<td>$0.52</td>
<td>$105.11</td>
</tr>
<tr>
<td>11</td>
<td>$105.11</td>
<td>$0.53</td>
<td>$105.64</td>
</tr>
<tr>
<td>12</td>
<td>$105.64</td>
<td>$0.53</td>
<td>$106.17</td>
</tr>
</tbody>
</table>

Thus, when interest was compounded monthly, the APY is _______.

The first APY was 6.05%, and the second APY was ______ %; therefore, choose ____________.
Formulas
As shown in the above examples, “simple problems” can be handled as easily as “complex problems” in Excel by just extending Excel spreadsheets (which you will be doing in your assignments).

In this class, you will be required to use Excel spreadsheet and Excel functions (covered in the next section) to complete problems.

However, there are also formulas that you may have seen before and may choose to use to verify your results. They are as follows:

Simple Interest Formula:
\[ I = PRT \]
- \( I \) = amount made in interest
- \( P \) = principal
- \( R \) = annual interest rate (APR) (or APY, because this is simple interest)
- \( T \) = time (in years)

Compound Interest Formula:
\[ A = P \left(1 + \frac{R}{N}\right)^{NT} \]
- \( A \) = account balance after \( T \) years
- \( P \) = principal
- \( R \) = annual interest rate (APR)
- \( N \) = number of compoundings per year
- \( T \) = time (in years)

Annual Percentage Yield (APY):
\[ APY = \left(1 + \frac{R}{N}\right)^{N} - 1 \]
- \( APY \) = annual percentage yield
- \( R \) = annual interest rate (APR)
- \( N \) = number of compoundings per year
Group Activity 18: Savings Accounts and APR vs APY

1. Suppose that you deposit $500 at an annual percentage rate of 6.0%, compounded annually.
   a. What is your account balance after one year?
   b. What is your account balance after 10 years? Complete this problem in Excel, copying the table you created into your Word document.

2. How long would it take an investment to triple at an annual percentage rate of 6.75% compounded annually? Complete this problem in Excel, copying the table you created into your Word document.

3. Suppose that you invest $2000 at an annual percentage rate of 5.25%, compounded monthly. Complete the following in Excel.
   a. What is your account balance after one year? Copy the Excel table into your Word document.
   b. What is your account balance after 10 years? Copy only the row from your Excel table where you determined your answer.
   c. What is the annual percentage yield for this account? Utilize a couple of the cell entries from your Excel table to compute the APY, showing your work.

4. Compare the accumulated balance in three accounts that all start with an initial deposit of $1000. All three accounts have an annual percentage rate of 5.5%, but the first account compounds interest annually, the second account compounds interest quarterly while the third account compounds interest monthly. Make tables that show the accumulated balance in all three accounts for one year. You'll have three tables, each of different length.
   a. Paste your tables into your Word document.
   b. Calculate the annual percentage yield for each account (as you did in 3c). Show your work.
   c. Do you want an account that compounds interest annually, quarterly or monthly (given that the annual interest rate is the same)?

5. Suppose you had a choice between two investments. The first compounds interest annually at an annual percentage rate of 4.50%. The second compounds interest monthly at an annual percentage rate 4.45%. Which is the higher yielding investment? Mathematically justify your answer.
Assignment 10: Savings Accounts and APR vs APY

Complete Assignment # 10 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. Suppose that you invest $1500 at an annual percentage rate of 6.0%, compounded quarterly.
   a. What is your account balance after one year? Create and copy an Excel table into your Word document.
   b. What is the annual percentage yield (APY) for this account? Utilize a couple of the cell entries from your Excel table to compute the APY, showing your work.
   c. What is your account balance after 10 years? Copy the last two rows from your Excel table and state your answer.

2. You now wish to deposit the $2300 and you found an investment that is offering an annual percentage rate of 6.2%, compounded semi-annually.
   a. What is your account balance after 10 years? Create and copy an Excel table into your Word document.
   b. What is the annual percentage yield (APY) for this account? Utilize a couple of the cell entries from your Excel table to compute the APY, showing your work.

3. Assuming you had the same amount of money to invest, which account is better, 6.0% compounded quarterly in (#1) or 6.2% compounded semi-annually in (#2)? Explain your answer.
   (Hint: You should be able to answer utilizing some of the results in 1 and 2.)

"Why should I save for a rainy day? I don't want one."
Chapter 17: Savings Plans and Investments

Objectives:

1. Use Excel functions to determine the future value of a principal investment, the future value of investments with regular payments, and the APY.
2. Use Excel spreadsheets and functions to determine a principal investment or a regular payment needed to plan for a particular savings goal.

Basic Savings Accounts: Using Excel Functions

Basic concepts of savings accounts, including compounding interest were discussed in the last section. Many examples were looking at what happens to the value of a principal investment under different plans. In nearly all of the examples, Excel spreadsheets were used to answer the questions, which you still need to know how to do.

Now, these same questions will be answered using Excel functions.

Example 1: Suppose you have $1000 that you are going to invest at an APR of 4%, where interest is compounded quarterly. How much will you have after 5 years?

In last section, this example was solved with spreadsheets. Now, it will be shown with an Excel function. Because the goal is to find the future value of this investment, that is the function we will use, which is abbreviated FV. Here's how

Finding the future value of an investment in Excel

- Go to an empty cell and type =
- Tap on the function tab, which looks like .
  The function box should appear.
- In the category box, either leave the category as “All” or you can select “Financial.” Either way, scroll down and select the function FV.
  Now, the dialogue box for appropriate function should appear.
- Observe that there is a description at the bottom of the function. Also notice when the cursor is in each box, directions for how to fill in that box are shown.
  - For “Rate,” remember to divide it by the number of compoundings, as before. Note that you can type .04/4 or 4%/4.
  - For “Nper,” the number of payments, enter the the number of compoundings per year * the number of years. Here, type 4*5.
  - For “Pmt,” it is the payment deposited each period. Here, there is no money deposited each period, so enter 0.
  - For “Pv,” the present value, this is the principal. Here, enter 1000.
  - For “Type,” this has to do with the time of the payment, which we will leave as at the end, so leave this box blank.
Notice that the answer in the lower left-hand corner as “formula result = -1220.19.” Once “OK” is selected, the result will display as 1220.19 (red means negative). Ignoring the negative value (a quirk in Excel) and compare this answer with what was computed in the spreadsheet in last section’s notes and observe the same value of $1220.19.

Example 2 This example is from your last group activity. Suppose that you deposit $500 in a bank that offers an annual percentage rate of 6.0%, compounded monthly. How much will you have in 10 years?

Following the same directions as last time, the result should be:

Thus, the amount in 10 years would be $909.70.
Example 3  This example is in the notes from last section. You have a choice between two different savings plans. The first advertises an APY of 6.05%, where interest is compounded monthly. The second advertises an APR of 6% and interest is compounded monthly. Which do you choose and why?

Recall that we need to compare APY’s, and the first is given.

Looking at the second choice, its APY will be higher than 6%, but now we need its exact computation to be able to compare with the first APY of 6.05%.

Remember that the APY is also sometimes called the effective rate, so this time the function we are going to use is “EFFECT.”

Finding the Annual Percentage Yield (Effective Rate) in Excel

- Go to an empty cell and type =
- Tap on the function tab, which looks like $\text{fx}$
- The function box should appear.
- In the category box, either leave the category as “All” or you can select “Financial.” Either way, scroll down and select the function EFFECT.
- Now, the dialogue box for appropriate function should appear.
- Observe that there is a description at the bottom of the function. Also notice when the cursor is in each box, directions for how to fill in that box are shown.
  - For “Nominal_rate,” recall that this is the same as APR, so here enter 0.06 or 6%.
  - For “Npery,” the number of compoundings per year, here, enter 12.

Notice that the answer in the lower left-hand corner as “formula result = 0.0616...”

Once “OK” is selected, the result will display as 0.0616..., which we need to convert to a percent showing two decimals, which is 6.17% (same result as last section).

Thus, the APY of the second option is about 6.17% (same result as in last section) and 6.17% > 6.05%; therefore, we would choose the second option.
Still using these same concepts as last time, let’s consider a new type of example, where we are now trying to reach a desired goal.

**Example 4** Suppose you are planning to buy a car in five years and want to be able to put $2000 down. In order to encourage you to save your money, your parents have agreed to pay you an APR of 12% (which is MUCH HIGHER than anything you would find at a bank these days), where interest is compounded quarterly. How much would you have to give your parents today to have $2000 in five years?

a. One way to solve this problem in Excel would be to set up a spreadsheet (as covered in the last section) and then use trial-and-error with the principal amount to get to $2000 after 5 years.

The spreadsheet would be set up as:

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TRIAL AND ERROR AMOUNT</td>
<td>=B2*.03</td>
<td>=B2+C2</td>
</tr>
<tr>
<td>2</td>
<td>=D2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>Goal: 2000</td>
<td></td>
</tr>
</tbody>
</table>

Now, start type in different amounts for the principal into cell B2 and fill all the columns, until the result is 2000 in cell D20. After some trial-and-error, we can get to:

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Balance at Beginning</th>
<th>Interest Earned</th>
<th>Balance at End</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,107.50</td>
<td>$33.23</td>
<td>$1,140.73</td>
</tr>
<tr>
<td>2</td>
<td>$1,140.73</td>
<td>$34.22</td>
<td>$1,174.95</td>
</tr>
<tr>
<td>3</td>
<td>$1,174.95</td>
<td>$35.25</td>
<td>$1,210.20</td>
</tr>
<tr>
<td>:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>$1,885.44</td>
<td>$56.56</td>
<td>$1,942.01</td>
</tr>
<tr>
<td>20</td>
<td>$1,942.01</td>
<td>$58.26</td>
<td>$2,000.27</td>
</tr>
</tbody>
</table>

So, it looks like a principal investment of about $1107.50 would grow to about $2000 in 5 years when the APR is 12% and interest is compounded quarterly.

b. There is also another way to solve this problem in Excel, using formulas. Observe that the type of problem is similar to examples 1 and 2, where the FV (future value) function was used. The difference is now we would like tell Excel the future value and have it report the principal. Here’s how:
Using the Goal Seek Function to find the required principal amount in Excel

- Go to an empty cell, say A1, and type =.
- Tap on the function tab, which looks like .  The function box should appear.
- In the category box, either leave the category as “All” or you can select “Financial.” Either way, scroll down and select the function FV.

Now, the dialogue box for appropriate function should appear.

- Observe that there is a description at the bottom of the function. Also notice when the cursor is in each box, directions for how to fill in that box are shown.
  - For “Rate,” remember to divide it by the number of compoundings, as before. Note that you can type .12/4 or 12%/4.
  - For “Nper,” the number of payments, enter the the number of compoundings per year * the number of years. Here, type 4*5.
  - For “Pmt,” it is the payment deposited each period. Here, there is no money deposited each period, so enter 0.
  - For “Pv,” the present value, this is the principal, which is the desired goal for this problem. Thus, tap on an empty cell or type in the the cell reference of an empty cell, say A2. (Think of this cell reference like a variable in algebra that we want to solve for.)
  - For “Type,” again leave this box blank.

After selecting OK, cell A1 will display $0.00 (which makes sense, because the future value of nothing, which is what is in A2, is still nothing).

- Next, select the “Data” tab from the top ribbon, and then select “What If Analysis” and finally choose “Goal Seek,” as shown below.
After selecting “Goal Seek,” The Goal Seek dialogue box should appear.

- For “Set Cell,” enter the cell where you did the FV function, which here was A1.
- For “To value,” enter the desired goal, which here is 2000.
- For “By changing cell,” enter the cell that you chose as your cell reference when doing the Pv portion of the FV function, which here was A2.

After hitting OK, the result -1107.35 should appear in the cell that you cell-referenced, which here was A2. Again, ignoring the negative (a quirk of Excel), the conclusion is that you would need to give your parents $1107.35 now to get to $2000 in 5 years at an APR of 12%, compounded quarterly.

Investments and Planning for Retirement: An Overview

All the examples discussed thus far are based on investing a principal amount and letting it grow, without adding anything to it. This is not very realistic. After all, if you are saving for anything, a car, a house, retirement, etc…, you typically put away a smaller amount every month or every week.

In the last example you saw that if you were to give your parents $1107.35 at 12% APR, compounded quarterly, it would grow to $2000 in 5 years. That is the power of compounding. Although you will not find 12% at banks in today’s economy, the point is still the same (although less dramatic) with smaller APRs. [You also typically earn higher rates over the long term by investing in mutual funds (a diverse group of stocks and bonds).]

Where the power of compounding becomes really impressive is when looking at how much can be earned over the years if a small amount of money is put away each month. This concept is often referred to as paying yourself first. This lesson may be one of the most important that you learn all semester!
Consider the following cartoon.

Cartoon Comments:
The older worker (Boomer) did not plan well for retirement and has to work to supplement his retirement check. The Social Security system spends money taken from current workers to pay the benefits of today’s retirees. That explains the younger worker's (Generation X) response to the older worker. Money is taken from his paycheck to fund the older worker’s social security check.

In the near future, there will not be enough people working to pay all retirees at today’s rate. The Social Security system will go bankrupt, unless there are drastic changes made. Social Security’s options are increase the money taken from the workers’ paychecks or decrease the amount of money paid to retirees. Neither option is popular.

In addition, retirement plans (401k’s or other defined contribution plans) are not as common in companies as they once were. Further, even in those companies that do offer retirement plans, often, many eligible employees do not participate.

Key Idea

In this lesson, you will learn why it is important to participate in retirement plans at work and/or establishing your own retirement fund.
Investments and Planning for Retirement: Examples

Example 5  At the age of 30, Malcolm starts an IRA (Individual Retirement Account) to save for retirement. He decides he can comfortably deposit $100 at the end of each month.

a. If he earns an APR of 6%, where interest is paid monthly, how much will he have when he retires 35 years later, at age 65?

Let’s first solve this problem using spreadsheets. It will be very similar to before, but with an additional column for adding in the deposit at the end of each month. Remember that interest will be 6%/12 = .06/12.

The spreadsheet would be set up as:

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Interest Earned</th>
<th>New Deposit</th>
<th>New Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>=.06/12*B2</td>
<td>100</td>
<td>=B2+C2+D2</td>
</tr>
<tr>
<td>2</td>
<td>E2</td>
<td>=.06/12*B3</td>
<td>100</td>
<td>=B3+C3+D3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Then, because we want to know the balance after 35 years, that will be 35*12 = 420 months later. The result is:

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Interest Earned</th>
<th>New Deposit</th>
<th>New Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0.5</td>
<td>100</td>
<td>200.5</td>
</tr>
<tr>
<td>3</td>
<td>200.5</td>
<td>1.0025</td>
<td>100</td>
<td>301.5025</td>
</tr>
<tr>
<td>4</td>
<td>301.5025</td>
<td>1.5075125</td>
<td>100</td>
<td>403.0100125</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>418</td>
<td>140058.1335</td>
<td>700.2906675</td>
<td>100</td>
<td>140858.4242</td>
</tr>
<tr>
<td>419</td>
<td>140858.4242</td>
<td>704.291209</td>
<td>100</td>
<td>141662.7163</td>
</tr>
<tr>
<td>420</td>
<td>141662.7163</td>
<td>708.3135815</td>
<td>100</td>
<td>142471.0299</td>
</tr>
</tbody>
</table>

Therefore, by putting away just $100 a month, after 35 years, Malcolm will have about $142,471.03.

b. How much money did Malcolm actually deposit and how much was earned from interest?

Malcolm deposited $100 for 420 months, so he deposited ________________________.

Thus, the remaining amount of ________________________ was from interest!

That is the power of compounding!

c. Solve the same exercise, using the function future value (FV).

This time for “Nper,” enter 35*12, and for “Pmt,” because $100 is deposited each month, enter 100. Also, for “Pv,” the present value, because no lump sum amount is being invested, enter 0:
Unit 4: Finance  
Chapter 17: Savings Plans and Investments

Example 6  Suppose Malcolm, despite the good advice of his liberal arts mathematics teacher, did not start saving until he was 45. Recognizing he waited a little too long, he decides to deposit $200 at the end of each month.

a. How much will he have when he retires in 20 years? (Again assume the IRA has an APR of 6% with interest paid monthly.)

Using the function future value (FV).

This time for “Nper,” enter 20*12 and for “Pmt,” enter 200.

So, even though Malcolm doubled the amount he was depositing each month, he only has $92,408.18.
b. How much would Malcolm have to deposit each month if he wanted to have the $142,471 (the amount he ended up with in the previous exercise when he started at age 30) in his IRA?

The most efficient way to solve this type of problem using the “Goal Seek” feature (initially shown in example 4 part b).

We will follow the same process shown in example 4 part b; however, here the unknown quantity is the monthly payment, so it is in this “Pmt” box where an empty cell (such as A2) needs to be entered, as shown below.

![Image of Goal Seek process]

Then, access Goal Seek and enter:

![Image of Goal Seek dialog box]

After entering OK, the result shows that Malcolm needs to invest about $308.35 each month!

Think about these last two examples:

- If you start at age 30, you only need to put away $100 a month, for a total investment of $100*420 months = $42,000 of your money, to have $142,471 by age 65.
- If you wait until age 45, you have to more than triple your monthly deposit to about $308.35, for a total investment of $308.35*240 months = $74,004 of your money, to obtain that same amount by age 65!
Example 7: Now suppose that Malcolm, taking the good advice of his liberal arts mathematics teacher, decides to starting saving at the age of 20. How much will he need to put away each month to get to that same $142,471?

Again, let’s use the “Goal Seek” feature.

Starting with the FV function:

Then, access Goal Seek:

After you hit OK, you should see that Malcolm only needs to invest about $51.69 each month to have $142,471 when he retires. That’s only about $13 a week!

Think about this last example with the previous two examples:

- If you start at age 20, you only need to put away $51.69 a month, for a total investment of $51.69*540 months = $27,912.60 of your money, to have $142,471 by age 65.
- When you started at age 30, you needed to put away $100 a month, for a total investment of $100*420 months = $42,000 of your money, to have $142,471 by age 65.
- If you wait until age 45, you have to more than triple your monthly deposit to about $308.35, for a total investment of $308.35*240 months = $74,004 of your money, to obtain $142,471 by age 65.

Key Idea

There is no substitute for time when saving money!
Group Activity 19: Savings Plans, Investments, and Planning for Retirement

Exercises 1-4 are similar group activity 18’s exercises; however, instead of creating spreadsheets to solve, use the Excel functions we discussed (FV and Effect) to solve.

1. Suppose that you deposit $1500 into a bank account that offers an APR of 5.75%, compounded quarterly. Use the FV function to determine how much you will have in your account after 10 years.

2. Suppose that you deposit $1500 into a bank account that offers an APR 5.65%, compounded monthly. Use the FV function to determine how much you will have in your account after 10 years.

3. Which account pays better: 1 or 2?

4. Use the Effect function to determine the APY’s of each account. Does this verify your result from (3)?

5. In order to save money for when you plan to move out on your own after college, you plan to get a full time job this summer. In order to encourage you to save your money, your grandparents told you that they will pay you an APR of 9% (which you could not find in any bank!). Suppose you anticipating graduating college in 5 years and by that time, you hope the money you make from this summer will grow to $5,000. How much will you need to make this summer to reach $5000 in 5 years in each account described? 
   (Hint: You need to use the FV function and the “Goal Seek” tool to solve, where you are looking for the one-time “lump-sum” amount.)
   a. An account with annual compounding and an annual percentage rate of 9%.
   b. An account with quarterly compounding and an annual percentage rate of 9%.

Exercises 6-8 are about investments where you are making regular deposits. Read them carefully, because some exercises may ask you to create a spreadsheet to solve and others make ask you to use the Excel functions to solve.

6. As a new parent, you have decided to start investing for your child’s college education. With a new baby, finances are tight, so you decide to cut out that gourmet coffee everyday and save about $5 a day. This amounts to about $150 a month that you feel you can easily save. Assuming you find a college fund that pays an APR of 6.5%, where interest is paid monthly, how much will you have in 18 years when your child is ready for college? Create an Excel spreadsheet to solve this problem. Copy the first three rows and the last three rows of your spreadsheet into your Word document. (You may want to check yourself by using the FV function.)

7. Looking at your result in 6, you decide this will not be enough money. You want to have $100,000 saved for your child’s college education. At the same rate, how much would you need to put away each month to get to $100,000 in 18 years? (Hint: Use the FV function and the “Goal Seek” tool to solve.)

8. Suppose you waited until your child was 5 years old before you started saving. How much would you need to save each month to get to $100,000 by the time s/he was 18?
Assignment 11: Savings Accounts and Planning for Retirement

Complete Assignment #11 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

Exercises 1-2 are about investments where you are making regular monthly deposits. Read them carefully, because some exercises may ask you to create a spreadsheet to solve and others make ask you to use the Excel functions to solve.

1. You are going to start planning for retirement early.
   a. You begin to deposit $200 dollars at the end of the month, every month for 36 years, into an account that pays 6.2% interest compounded monthly. How much money will be in the account when you retire? Make an Excel spreadsheet to solve this problem. Copy only the first three rows and the last three rows into your Word document.
   b. How much of your own money did you actually deposit in (a)?

2. Suppose that you waited until you were middle-aged to start saving for retirement.
   a. Realizing that you will be saving for half as long, you decide to double your monthly deposits, hoping that this will “even things out.” So, you deposit $400 a month, every month for 18 years, into an account that pays 6.2% interest compounded monthly. How much money will be in the account when you retire? Did you end up with about the same amount (as in 3)? (You may use an Excel function exclusively here and not set up a spreadsheet.)
   b. How much of your own money did you actually deposit in (a)?

Take a moment to look back over these different scenarios in #1-2, and once again realize...

Key Idea

There is no substitute for time when saving money!
Chapter 18: Loans including Credit Cards

Objectives:

1. Use Excel spreadsheets to model credit card balances when making minimum payments.
2. Use Excel spreadsheets and functions to determine the payment needed to pay off a loan in a specified amount of time.
3. Determine amount paid in interest under various credit cards.
4. Identify dangers of only making minimum monthly payments and other dangers associated with credit cards.

There are many types of loans that you will have in your life, such as car loans, college loans, credit card loans, and mortgages. We will be focusing on loans with a constant interest.

Credit Cards

You have all heard about the dangers of credit card debt. Here, you will see first-hand just how long it can take to pay off a high interest credit card loan.

From [http://www.creditcards.com](http://www.creditcards.com), observe the following information on APRs of credit cards by different categories:

<table>
<thead>
<tr>
<th>Category</th>
<th>APR</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Average</td>
<td>15.00%</td>
</tr>
<tr>
<td>Low Interest</td>
<td>11.62%</td>
</tr>
<tr>
<td>Business</td>
<td>12.85%</td>
</tr>
<tr>
<td>Student</td>
<td>13.14%</td>
</tr>
<tr>
<td>Balance Transfer</td>
<td>14.12%</td>
</tr>
<tr>
<td>Airline</td>
<td>15.10%</td>
</tr>
<tr>
<td>Reward</td>
<td>15.14%</td>
</tr>
<tr>
<td>Cash Back</td>
<td>15.27%</td>
</tr>
<tr>
<td>Instant Approval</td>
<td>18.00%</td>
</tr>
<tr>
<td>Bad Credit</td>
<td>22.73%</td>
</tr>
</tbody>
</table>

In terms of minimum payments, credit cards typically either use the interest charged plus 1% of the remaining balance or a simple percent (2%-5%, typically on the lower end of this range) of the total balance. In either scenario, they also set a fixed amount that your minimum payment cannot go lower than.

One final comment regarding the calculation of interest on credit cards: There are slight differences in when (what time of the month / payment cycle) credit cards charge interest; therefore, the answers we model here could vary slightly from card to card – especially when it comes to the final payment / payoff balance.
Example 1: Suppose that when you started school this semester, you had no way to pay for your tuition, books, and a computer, which total $2800. You decide to put this amount on your credit card with an annual interest rate of 15%. Assume that you make no other purchases, and you pay the minimum payment of 2.5% per month and not less than $20.

a. What is your balance after 3 years? Set up a spreadsheet in Excel to solve.

Notes:
- Because the interest is computed monthly, you need to divide the interest rate by 12.
- To enter the “MAX” function, you can either type it exactly as you see it both, or you can insert the function like we were doing last week. Here, it is probably just as easy to simply type it in.

The spreadsheet would be set up as follows and then filled in.

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning CC Balance</th>
<th>Monthly Payment</th>
<th>Interest Charge</th>
<th>Ending CC Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2800</td>
<td>=MAX(0.025*B2,20)</td>
<td>=0.15/12*B2</td>
<td>=B2-C2+D2</td>
</tr>
<tr>
<td>2</td>
<td>=E2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After 3 years (or 36 months), the result should be:

| 36    | $1,802.84            | 45.07094775     | 22.53547387     | $1,780.30        |

So, the balance after 3 years is still $1780.30.

b. How long would it take you to pay off the balance?

Extending the table, we should see that:

| 154   | $46.77               | 20              | 0.584574628     | $27.35           |
| 155   | $27.35               | 20              | 0.341881811     | $7.69            |
| 156   | $7.69                | 7.69            | 0.096155333     | ($12.21) 0       |

By the 155th month, the balance was $7.69, so in reality the last payment would only need to be $7.69 and the credit card would be paid off by the 156th month.

Think about that. It took 156 months or 13 years to pay off a balance of $2800 if you only make the minimum payment!

c. How much interest did you pay?

To find how much interest was paid, sum column C. One way to sum is to highlight the cells and hit the sum button in the top ribbon. Additionally, the SUM function can be used, as shown below.

The result should be $2315.86.

That’s right....You paid $2315.86 in interest for a loan that was for $2800!
Example 2 Suppose you have another lower interest credit card that has an annual interest rate of 11.62%. How long would it have taken to pay off the balance of $2800? Assume that you make no other purchases, and you pay the minimum payment of 2.5% per month (not less than $20).

By changing the interest rate to 0.1162/12 in the Excel spreadsheet, we should see:

<table>
<thead>
<tr>
<th>131</th>
<th>$39.45</th>
<th>20</th>
<th>$0.38</th>
<th>$19.83</th>
</tr>
</thead>
<tbody>
<tr>
<td>132</td>
<td>$19.83</td>
<td>20</td>
<td>$0.19</td>
<td>$0.03</td>
</tr>
</tbody>
</table>

After 132 months or 11 years the loan is paid. While taking almost 11 years to pay off $2800 is still terrible, notice the difference from example 1. When the interest rate was 15%, it took 13 years!

In terms of the amount of interest paid, in exercise 1, you paid $2315.86. Here, if you sum column D, you should get $1481.02, which means you paid over $834.84 more in interest under a rate of 15%.

You must pay attention to the interest rates on your credit cards!

Example 3 Now suppose you are using this lower interest credit card with its annual interest rate of 11.62%. Having learned your lesson about minimum payments, you decide that you want to pay the loan off in 1 year. What should your monthly payment be?

a. One way to solve is to use trial-and-error with your spreadsheet.

Set up the spreadsheet:

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning CC Balance</th>
<th>Monthly Payment</th>
<th>Interest Charge</th>
<th>Ending CC Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2800.00</td>
<td>GUESSES</td>
<td>=0.1162/12*B2</td>
<td>=B2-C2+D2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>=E2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By filling in guesses in cell C2 and extending your table, your goal should be to get to an ending balance of 0 in cell E12. After some trial-and-error, you should get the following when there is a monthly payment of $249:

| 11    | $481.91             | 249             | $4.67           | $237.57          |
| 12    | $237.57             | 249             | $2.30           | ($9.13) 0        |

So with monthly payments of $249 (where the last month you would only need to pay $237.57), the balance would be paid off after 1 year.

b. Another more efficient way to determine the monthly payment needed to pay off the balance in 1 year is by using the function “PMT.”

- Go to an empty cell and type =
- Tap on the function tab, which looks like $f$  

The function box should appear.
In the category box, either leave the category as “All” or you can select “Financial.” Either way, scroll down and select the function PMT.

Now, the dialogue box for appropriate function should appear.

Observe that there is a description at the bottom of the function. Also notice when the cursor is in each box, directions for how to fill in that box are shown.

- For “Rate,” the monthly rate, enter 11.62%/12 or 0.1162/12.
- For “Nper,” the number of payments, enter 12 (in general, it’s the number of years*12).
- For “Pv,” the present value or the amount of the loan, enter 2800.
- “FV” is the future value and is left blank or enter 0.
- “Type” is when the payment is made, which is almost always at the end, so leave this blank.

![Function Arguments](image)

Calculates the payment for a loan based on constant payments and a constant interest rate.

The monthly payment can be seen in the lower left corner as well as when clicking OK.

So, with monthly payments of $248.28, the balance would be paid off after 1 year.

c. Compute the total amount of money paid and the amount paid in interest.

The total amount paid is about $248.28*12 = $2979.35

The interest paid = total paid – principal = $2979.35 – $2800 = $179.35

**Tips to avoid credit card trouble**

- Use only one card.
- Pay balances in full each month, or at least make as large of payments as possible.
- Don’t use your credit card for a cash advance unless it is an emergency.
- **Know your credit card rate.** Watch for teaser rates. This caution is especially true with store credit cards.
Example 4. Consider a scenario where you are in need of a major appliance and you are at the Lowe’s checkout and are told that if you apply for their Lowes credit card:

**5% off**

or Special Financing**

Choose 5% off every day,* or select 6 months special financing with a $299 minimum purchase** when you use your Lowe’s Consumer Credit Card.

Knowing you will need time to pay off this appliance, you read that the special financing will give you 6 months free financing!

However, if you look further into the fine print, you read:

**No interest if paid in full within 6 months.** Offer applies to purchase or order of $299 or more made on your Lowe’s® Consumer Credit Card. Interest will be charged to your account from the purchase date if the promotional purchase is not paid in full within 6 months.

That’s right – if you have not paid off the appliance purchase within the 6 months, you will be charged interest from first day!

Now, investigating even further, you read that: Standard APR is 24.99%

So, here’s the catch. It is free financing, ONLY if you can pay it off in 6 months. If are not able to do so for some reason, then you are going to pay 24.99% APR from day 1 and you may have been better off putting the purchase on your standard credit card.

Convince yourself by modeling in Excel the difference in interest paid on a $1500 purchase, making $50 payments, on a credit card with an APR of 25% versus and APR of 15%. What’s the difference in interest?

Let’s end with one final example on a car loan.
Example 5: You are financing $13000 for a used car.

According to [http://www.bankrate.com]:

You are able to secure a loan for 2.95% for 48 months (4 years).

a. What will your monthly payment be?

Again, using the PMT function:

Thus, the monthly payment will be $287.46.

b. You tell the car dealer that while you really like the car, you were hoping to keep your minimum payment under $200. He tells you “no problem,” and that you can do a 72 month loan (6 years) and then your minimum payment is only $197.23, as shown below:

Again, using the PMT function:

You love the car and think that the dealer was nice to give you a longer loan so you can afford the monthly payment, but you decide to think about it.
c. How much will you pay in interest with each loan?

d. What other factors should you consider in decision about the 72 month car loan for a used car?

e. Set up an amortization table for the loan in part (B), which shows how much of each payment is going towards your principal and how much is going towards interest.

The table will start off the same, but we will add a couple columns to observe how our monthly payment is being used:

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Monthly Payment</th>
<th>Interest Charge</th>
<th>Ending Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$13,000.00</td>
<td>197.23</td>
<td>0.0295/12*B2</td>
<td>=B2-C2+D2</td>
</tr>
<tr>
<td>2</td>
<td>=E2</td>
<td></td>
<td></td>
<td>=D2</td>
</tr>
</tbody>
</table>

Fill the spreadsheet and what do you notice about how much of each payment is going towards your principal and how much is going towards interest?

This trend will be even more significant when looking at mortgages, which is next.
### Group Activity 20: Loans, including Credit Cards

1. Suppose you are buying a car and want to borrow $20,000 over a 5-year period at an annual percentage rate of 7.5%.
   a. Use the PMT function in Excel to determine your monthly payment.
   b. Determine the total amount paid over the term of the loan. *Hint: You should not be creating an Excel spreadsheet here.*
   c. Determine how much you will pay in interest over the term of the loan. *Hint: You should not be creating an Excel spreadsheet here.*

2. You make a purchase for $2,800 on your credit card which has an annual interest rate of 19%. You pay the minimum payment of 2.5% per month and not less than $20. You make no additional purchases.
   a. Set up an Excel spreadsheet, as shown below.

<table>
<thead>
<tr>
<th>Month</th>
<th>Credit card balance at beginning of month</th>
<th>Payment at end of month</th>
<th>Interest for month</th>
<th>Credit card balance at end of month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2,800.00</td>
<td>=MAX(B2*0.025,20)</td>
<td>=B2*0.19/12</td>
<td>=B2-C2+D2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>=E2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   What is your credit card balance after 1 year? Copy the first 12 months of your table into your Word document.
   b. If you only paid the minimum each month (and didn’t make any additional purchases), how long would it take to pay off the balance? Copy the last two rows from your table into your Word document.
   c. How much interest would you pay altogether?

3. You only use your credit card in the case of an emergency. You charge $2,000 in auto repairs to credit card which has an annual interest rate of 9.75%.
   a. Use the PMT function to determine what your monthly payment should be if you would like to pay off your balance in full after one year.
   b. What would be your total payments on the credit card loan? *Hint: You should not be creating an Excel spreadsheet here.*

4. You take a vacation once a year and charge it to your credit card and pay it off in 6 months. You have the option to apply two credit cards. Credit card A has an annual rate of 9.99% with no annual fee and credit card B has an annual rate of 8.99% with a $20 annual fee (note that you have to pay the full annual fee – even if you pay the balance off in 6 months). The total cost of your vacation was $3,350.
   a. Use the PMT function to determine your monthly payment with credit card A.
   b. How much will you pay in total with credit card A?
   c. Use the PMT function to determine your monthly payment with credit card B.
   d. How much will you pay in total with credit card B (don’t forget to add in the annual fee here)?
   e. Which credit card do you choose for this situation?
Chapter 19: Loans – including Mortgages

Objectives:
1. Use Excel functions to determine the payment needed to pay off a mortgage in a specified amount of time.
2. Determine amount paid in interest over the course of a mortgage.
3. Use Excel spreadsheets to set amortization tables.
4. Compare different mortgage loan options, including differing closing costs and points.

Mortgages
If you have a fixed-rate mortgage (the interest rate is guaranteed for the entire length of the loan), then mortgages are very similar to any other loan (car loans, student loans, credit card loans, etc...).

However, there are a few more concepts you should be aware of.

- In order to get a home mortgage loan, lenders typically want a down payment, which is usually between 10% and 20% of the total cost of the home.
- There are typically additional closing costs associated with mortgages. For example, some lenders will charge points (sometimes called origination fees) which are one-time fees paid on the amount of the loan. [For example, on a $200,000 loan, if a lender charges 1 point, then you must pay 1% of 200,000, which is $2000.]

Mortgages can get very complicated, as there are many different types. This discussion is meant as an overview.

Example 1
You want to buy a condo that is $150,000. You are going to put 20% down, so you will only need a loan for $150,000 – 0.20*150,000 = 150,000 – 30,000 = $120,000. If a bank offers a 30-year fixed-rate mortgage at an APR of 6.4%, determine the (a) monthly payment, (b) the amount paid in full over the course of the loan, and (c) the amount paid in interest. Then (d) create an amortization table.

a. As with other loans, use the PMT function to determine the monthly payment.

- Go to an empty cell and type =
- Tap on the function tab, which looks like $f$x
  The function box should appear.
- In the category box, either leave the category as “All” or you can select “Financial.”
  Either way, scroll down and select the function PMT.
  Now, the dialogue box for appropriate function should appear.
- Observe that there is a description at the bottom of the function. Also notice when the cursor is in each box, directions for how to fill in that box are shown.
  o For “Rate,” the monthly rate, enter 6.4%/12 or 0.064/12.
  o For “Nper,” the number of payments, enter 30*12 (in general, it’s the number of years*12).
  o For “Pv,” the present value or the amount of the loan, enter 120,000.
Unit 4: Finance
Chapter 19: Loans – including Mortgages

- “FV” is the future value and is left blank or enter 0.
- “Type” is when the payment is made, which is almost always at the end, so leave this blank.

The monthly payment can be seen in the lower left corner as well as when clicking OK.

So, the monthly payments are about $750.61.

b. The total amount paid over the course of the loan will be 760.61*360 = $270,219.60.

c. The amount paid in interest will be $270,219.60 – $120,000 = $150,219.60.
   Notice that you end up paying more in interest than the actual house cost.

d. Using the monthly payment you computed ($750.61), construct a table in Excel to verify the loan is paid off in 30 years (360 months.) Make this table a true amortization table, showing the amount of each payment going to interest and to principal.

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Monthly Payment</th>
<th>Interest Charge</th>
<th>Ending Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$120,000.00</td>
<td>750.61</td>
<td>=0.064/12*B2</td>
<td>=B2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$120,000.00</td>
<td>=E2</td>
<td>=E2</td>
</tr>
</tbody>
</table>

The first few rows should fill as:

1. $120,000.00 750.61 640 $119,889.39 640 110.61
2. $119,889.39 750.61 639.41008 $119,778.19 639.41008 111.19992
3. $119,778.19 750.61 638.8170138 $119,666.40 638.8170138 111.7929862

The last few rows should fill as:

358 $2,224.92 750.61 11.86622047 $1,486.17 11.86622047 738.7437795
359 $1,486.17 750.61 7.926253648 $743.49 7.926253648 742.6837464
360 $743.49 750.61 3.965273668 ($3.16) 3.965273668 746.6447263

This document is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. Elgin Community College Mathematics Department, 2019.
Observe that yes, the loan would be paid off in 30 years.

Also observe the trend in the amount of money going to interest and to the principal each month. Early on, very little money is going towards paying down the actual principal loan amount.

For example, after 5 years of $750.61 monthly payments for a total amount paid of ______________, but if you sum the last column (of amount of payment to principal), you’ve only paid ______________ in principal!

This can be a problem if you need to sell your home – especially if home values haven’t appreciated, which is what so many Americans ran into. In fact, for many folks, the values of their homes depreciated, and they ended up owing more on their loan than their house was worth! This is called being “upside down” on your mortgage.

e. One month you notice that you seem to have a little more money than usual at the end of the month and realize that you got an “extra check” that month.

That is, you get paid every two weeks, which usually means you get 2 checks a month (which is how you have structured your monthly budget needs). BUT, TECHNICALLY, you get $26 checks each year. THUS, THERE ARE TWO MONTHS WHEN YOU BASICALLY GET AN “EXTRA” CHECK.

Because of this, you have decided that you can afford to make one extra payment per year. However, for ease here with our spreadsheet, let’s say we spread that extra $750.61 over the whole year, so we will pay an extra $62.55 each month.

Thus, making payments of $813.16, by what month is your mortgage paid off?

How much time and money did you save?

The conventional wisdom with mortgages is that 1 extra payment a year can take about_______ years off your mortgage!
Example 2 You recently sold your condo and want to buy a house that is $220,000. From the sale of your condo and the money you have saved, you are able to put 20% down, so you will only need a loan for $220,000 – 0.20 * $220,000 = $176,000. You have two different offers to consider:

Bank A: 30-year 7% fixed-rate loan with closing costs of $500 plus 1 point.

Bank B: 30-year 6.9% fixed-rate loan with closing costs of $800 plus 2 points.

Which will cost you less at the end of the loan?

Let’s organize the costs in a table form: *(Recall to find the monthly payment, use the PMT function.)*

**BANK A:**

<table>
<thead>
<tr>
<th>Total from Payments</th>
<th>Closing Costs</th>
<th>Points</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(monthly pmt*360)</td>
<td>500</td>
<td>.01*176,000</td>
<td>$423,794.80</td>
</tr>
<tr>
<td>1170.93*360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>421,534.80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**BANK B:**

<table>
<thead>
<tr>
<th>Total from Payments</th>
<th>Closing Costs</th>
<th>Points</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(monthly pmt*360)</td>
<td>800</td>
<td>.02*176,000</td>
<td>$421,610.40</td>
</tr>
<tr>
<td>1159.14*360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>417,290.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Although these loans resulted in similar amounts at the end, Bank B is really the better deal.

As previously mentioned, mortgages can get very complicated, as there are many different types. All of these examples were using fixed-rate mortgages.

Another type of loan is called an adjustable-rate mortgage (or ARM). These rates are often lower than fixed-rate mortgages, but are not guaranteed for the life of the loan. So, you let’s say you get a 5 year 4% ARM. This means your interest rate is only 4% for the first 5 years, but after that you have to re-finance your loan. This can be risky, because interest rates may be much higher in 5 years.

Other risky loans that some people ran into problems with were people who took out interest only loans, where for the first number of years (say 5 for example), the payments are only on the interest. This means that people had low monthly payments for the first number of years; however, they were not paying down their principal at all.

These are just some of the different types of programs that are out there now. It used to be that to buy a house, you had to do what was described above: Save 20% to put down and then get a fixed-rate mortgage. While now it can be nice that there are more options to buy homes, you have to be very careful that you understand the risks of the program you are signing up for.

**The moral of the story:** Do NOT just look at your monthly payment when making a decision about a home mortgage, and make sure you understand the “worse-case scenario” with non-traditional mortgage programs.
Group Activity 21: Loans, including Mortgages

1. Suppose that you have a student loan of $25,000 with an annual percentage rate of 7.25% for 12 years.
   a. Use the PMT function in Excel to determine your monthly payment.
   b. Determine the total amount paid over the term of the loan. *(Hint: You should not be creating an Excel spreadsheet here.)*
   c. Determine how much you will pay in interest over the term of the loan. *(Hint: You should not be creating an Excel spreadsheet here.)*
   d. Set up an Excel spreadsheet, as shown:

<table>
<thead>
<tr>
<th>Month</th>
<th>Loan balance at beginning of month</th>
<th>Payment at end of month</th>
<th>Interest for month</th>
<th>Loan balance at end of month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$25,000</td>
<td>Monthly payment (#1)</td>
<td>=B2*0.0725/12</td>
<td>=B2-C2+D2</td>
</tr>
<tr>
<td>2</td>
<td>=E2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Extend the spreadsheet and verify that the loan will be paid off in 12 years, copying and pasting the first three and last three rows of the table into your Word document.
   e. Suppose that you would like pay the loan off in 7 years instead of 12. What monthly payments will you need to make assuming the same interest rate?
   f. Determine the total amount paid over the term of the 7 year loan in (e). *(Hint: You should not be creating an Excel spreadsheet here.)*
   g. How much do you save paying off the loan in 7 years versus the 12 year loan?

2. You want to buy a condominium that is $160,000.
   You are going to put 20% down, which means you are going to pay 0.20*$160,000 = $32,000 up front and then you will need a loan for the remaining $160,000 – $32,000 = $128,000.
   a. You have a 30-year fixed-rate mortgage at an APR of 6.5%. Determine the monthly payment.
   b. Determine the total amount paid over the term of the loan. *(Hint: You should not be creating an Excel spreadsheet here.)*
   c. Determine how much you will pay in interest over the term of the loan. *(Hint: You should not be creating an Excel spreadsheet here.)*

3. You are purchasing a house for $200,000. You have two options:
   Bank A: 30-year 6.8% fixed-rate loan with closing costs of $500 plus 1 point.
   Bank B: 30-year 6.5% fixed-rate loan with closing costs of $1500 with no points.
   a. You are going to put 20% down, so how much money are you paying up front?
   b. Thus, you need a mortgage loan for how much money?
   c. How much will you pay in total with bank A? (Fill in the table below, as we did in the notes, to find this total.)
Unit 4: Finance
Chapter 19: Loans – including Mortgages

<table>
<thead>
<tr>
<th>Total from Payments (monthly pmt*360)</th>
<th>+</th>
<th>Closing Costs</th>
<th>+</th>
<th>Points (that % point * principal loan amt)</th>
<th>=</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d. How much will you pay in total with bank B? (Fill in the table below, as we did in the notes, to find this total.)

<table>
<thead>
<tr>
<th>Total from Payments (monthly pmt*360)</th>
<th>+</th>
<th>Closing Costs</th>
<th>+</th>
<th>Points (that % point * principal loan amt)</th>
<th>=</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

e. Which bank will you choose?
Assignment 12: Loans

Complete Assignment # 12 in MyOpenMath and the problems below.

Directions: This assignment should be typed and answers to questions should be in complete sentences, free of grammatical and spelling errors. Start early, so if you have questions, you will have time to ask.

1. You find a car you really want that is priced at $14,000. The car dealership says they will drop the price to $12,000, if you take their financing. They offer you 6.75% for 4 years. You tell them you will think about it.
   a. Construct a table in Excel with columns for the month, the balance at the beginning of the month, the monthly payment, the interest for the month, and the balance at the end of the month, and verify that the loan is paid off in 4 years. Copy the first two and last two rows of the table into your Word document.

2. Assume that you have a balance of $1500 on a credit card that carries an annual percentage rate of 18%.
   a. You pay the minimum payment of 2.5% per month (not less than $20), and make no additional purchases on your card. Construct a table in Excel with columns for the month, the balance at the beginning of the month, the monthly payment, the interest for the month, and the balance at the end of the month. How long will it take you to pay off the $1500 if you only make the minimum payment? Copy the first two and last two rows of your table into your Word document.
   b. How much did you pay in total?
   c. Realizing that you should make more than the minimum monthly payment, you decide to pay $100 a month. How long does it take you to pay off the balance now?
   d. How much did you pay in total?
   e. Now suppose that while you are making $100 monthly payments, you start charging an additional $50 per month. Assume that this charge is made at the end of the month and interest is calculated on the balance before the charge is added. Make a table that shows month, beginning balance, payment, interest, charge amount, and end balance. How long will it take you to pay off your credit card now? Copy the first two and last two rows of your table into your Word document.
   f. Repeat exercise (e), but assume you only make monthly payments of $80, how long will it take to pay off the credit card debt?

3. You are buying a house that is $300,000 and will be financing (so getting a loan for) $240,000. You have a 30-year 6% fixed-rate loan.
   a. How much will you pay per month?
   b. How much will you pay over the course of your entire loan?
   c. Now suppose that you decide that you can afford an extra $50 a week to put towards your house, and you start making payments that are $200 more than the original monthly payment (from part a). Make a table that shows month, beginning balance, payment, interest, and end balance. How long will it take you to pay off your house if you pay an additional $200 each month? Copy the first two and last two rows of your table into your Word document.
d. How much will you pay over the course of your entire loan (paying the additional $200 / month)?

e. How much will you save over the course of your entire loan (paying the additional $200 / month)?
Unit 4: Finance
Unit Review

Unit 4 Topics

This test covers material presented in Chapters 15-19.

The following represents a list of the major “new” topics you should be able to do.

Finance: Saving and Investing

- Relate compound interest to exponential growth; therefore, there is a constant doubling time.
  - Use the Rule of Seventy to estimate the time to double (but you have to use APY).
- APY versus APR:
  1. APR = nominal rate
  2. APY = effective rate
  3. APY > APR
  4. APY = APR only when interest is compounded annually
  5. Find APY:
     a. From a spreadsheet (Divide the interest earned in the first year by the original principal = (new-old)/old) AND
     b. Use EFFECT function
- Model savings (both lump-sum deposits and regular deposits) using:
  o a spreadsheet AND
  o the Future Value Function (FV)
- Determine how much to invest to get a certain value (“goal”) in the future, using the “FV” function with the “GOAL SEEK” process to determine
  o How much to invest as a lump-sum deposit to get to a certain goal AND
  o How much to invest in regular payments to get to a certain goal.

Finance: Loans (Including Credit Cards, Mortgages, and Car Loans)

- Model loans (especially credit cards with their minimum payments) in an Excel spreadsheet.
  o Sum payment column to find total amount paid on a loan.
  o Sum interest column (or subtract the initial loan amount from total paid) to find amount paid in interest.
- Determine a monthly payment, using PMT function.
  o Multiply monthly payment by the number of payments to find total amount paid on a loan.
  o Subtract initial loan amount from total paid to find amount paid in interest.
- For mortgages,
  o Compare two different plans, which may include closing costs and points (which is a percent of the loan amount).
  o Create amortization tables, illustrating how much of each payment goes toward interest and how much goes towards the principal.
EXCEL IS NOT NECESSARY FOR EXERCISES 1-4.

1. You are trying to decide between investing some money in two different accounts described below.

   Savings Account A: APR = 2.5%, where interest is compounded monthly
   Savings Account B: APR = 2.5%, where interest is compounded quarterly

   Which account do you choose and why?

2. You are trying to decide between investing some money in two different accounts described below.

   Savings Account A: APY = 3.4%, where interest is compounded monthly
   Savings Account B: APR = 3.4%, where interest is compounded quarterly

   Which account do you choose and why?

3. You are buying a condo that is $160,000 and you have enough money to make a 20% down payment. How much will you need a mortgage loan for? Show work.
4. You are buying a house and will be taking out a traditional 30-year home mortgage loan for $175,000.

   a. Determine the total cost of the plan A, with its APR = 6%, below. Show work.

      Mortgage Plan A: Monthly Payment = $1049.21  Closing Costs = $500 with no points

   b. Determine the total cost of plan B, with its APR of 5.75%, below. Show work.

      Mortgage Plan B: Monthly Payment = $1021.25  Closing Costs = $1500 plus 2 points

   c. Which plan would you choose?
Use Excel to Answer the Following Questions

5. You are investing $6500 in a savings account with an APR of 3.5%, where interest is compounded quarterly. Set up an Excel spreadsheet, with columns as shown below.

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Beginning Balance</th>
<th>Interest</th>
<th>End Balance</th>
</tr>
</thead>
</table>

Determine the balance in your account after 5 years.

Copy the first three and last three rows into your Word document and state your final answer.

6. Determine the APY of the account in #5 account, utilizing a couple of values from your table in (a). Show your work and round your percent to the nearest hundredth (for example 4.23%).

7. You have decided to deposit $150 a month, every month, into a savings account that pays 5.75% with interest compounded monthly. Set up an Excel spreadsheet, with columns as shown below.

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Interest</th>
<th>New Deposit</th>
<th>End Balance</th>
</tr>
</thead>
</table>

Determine the balance in your account after 20 years.

a. Copy the last three rows into your Word document and state your final answer.

b. How much money did you actually deposit?

c. How much money did you make in interest?

8. You have decided to deposit $120 a month, every month, into a savings account that pays 5.75% with interest compounded monthly.

a. Use an Excel function (so you do NOT need to set up a spreadsheet) to determine the balance in your account after 25 years.

b. How much money did you actually deposit?

c. How much money did you make in interest?

9. Based on your results from 7 and 8, would it be better to deposit $150 monthly for 20 years or $120 monthly for 25 years? Explain.

10. Suppose in 30 years you wanted to have $500,000 in an account which pays 6.5% with interest compounded monthly. What would your monthly payment need to be? (Notice that you are planning for a goal here...hint..hint..)
11. You are planning to buy a used car and will need to finance a loan for $7000. The car dealership offers you a 4.25% loan for 3 years.

   a. Determine the monthly payment, using the appropriate Excel function.
   b. Determine the total amount you will pay on the loan. (You should not be setting up a spreadsheet here.)
   c. Determine the amount paid in interest. (You should not be setting up a spreadsheet here.)

12. A couple of years after purchasing this car, it is need of some necessary repairs, totally $1000. Unfortunately, your only option is to use your credit card, which has an APR of 17%. Set up an Excel spreadsheet, with columns as shown below.

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Payment</th>
<th>Interest</th>
<th>End Balance</th>
</tr>
</thead>
</table>

   a. Assuming you that you make no other purchases and you pay only the minimum payment (the maximum 2.5% of the balance or $20), determine how long it would take you to pay off your credit card. Copy the first two and the last two rows of your Excel table into your Word document.

   b. How much did you pay in total?

   c. Realizing that paying only the minimum payment will cost you a lot of money in interest, you decide that you can make $45 monthly payments. Determine how long it would take you to pay off your credit card making monthly payments of $45. Copy the last two rows of your Excel table into your Word document.

   d. How much did you pay in total?

   e. How much do you save in interest by making $45 monthly payments over just making the minimum monthly payment?
Chapter 20: Interpreting Polls (Sampling, Margin of Error, and Confidence Intervals)

Objectives:
1. Discuss random sampling, as it relates to polling.
2. Determine the relationship between sample size and margin of error in polls.
3. Interpret margin of error and confidence intervals of polls.

Random Sampling
It is inconvenient, expensive, and often impossible to poll an entire population; therefore, polls and surveys use samples and generalize those samples to the entire population of interest.

Example 1
Suppose we wanted to know ECC students’ opinions regarding the classes offered at ECC. Let’s say we decided to use our class as a sample with the hopes of generalizing to the entire ECC student body.

a. Today in class, there are:
   
   _______ total students
   _______ female students  \rightarrow  _______ % are female students
   _______ male students  \rightarrow  _______ % are male students

   _______ full time students  \rightarrow  _______ % are full time students
   _______ part time students  \rightarrow  _______ % are part time students

   _______ students are under 23  \rightarrow  _______ % of students are under 23
   _______ students are 23 or older  \rightarrow  _______ % of students are 23 or older

b. Let’s compare our class sample to the general ECC student body, which

<table>
<thead>
<tr>
<th></th>
<th>Our Class</th>
<th>ECC</th>
</tr>
</thead>
<tbody>
<tr>
<td>% are female students</td>
<td></td>
<td>54.3%</td>
</tr>
<tr>
<td>% are male students</td>
<td></td>
<td>45.7%</td>
</tr>
<tr>
<td>% are full time students</td>
<td></td>
<td>31.4%</td>
</tr>
<tr>
<td>% are part time students</td>
<td></td>
<td>68.6%</td>
</tr>
<tr>
<td>% students under 23</td>
<td></td>
<td>50.9%</td>
</tr>
<tr>
<td>% students 23 and older</td>
<td></td>
<td>49.1%</td>
</tr>
</tbody>
</table>

c. How representative was our class in terms of gender, FT/PT status, and age? (*Note, we could also look at other demographics, such as ethnicity, socio-economic status, etc...*)
d. Why might our class not be a representative sample for all ECC student body?

This example is meant to be a brief introduction to what may be the most important aspect of any poll or survey: obtaining a representative sample. A statistics course would spend a lot of time discussing ways of obtaining a truly random and representative sample of a given population.

For example, national polling used to be done with phone calls to people’s landlines; however, with the rise of mobile phones and the increase of the number of homes that no longer have landlines, you will read in the surveys below that they now use a combination of both landlines and cell phones in their polling.

Sample Size, Margin of Error, and Confidence Intervals
Consider the three polls below, reading the fine print under the survey methods. Pay particular attention to the sample size, margin of error, and confidence intervals.

Example 2

<table>
<thead>
<tr>
<th>Those With Postgraduate Degrees Most Likely to Be Lifelong Learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>On a 5-point scale, where 5 means &quot;strongly agree&quot; and 1 means &quot;strongly disagree,&quot; please rate your level of agreement with the following items: You learn or do something interesting every day.</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Some postgraduate study/Postgraduate degree</td>
</tr>
<tr>
<td>Four-year bachelor's degree</td>
</tr>
<tr>
<td>Some college/Technical school/Two-year associate degree from a college</td>
</tr>
<tr>
<td>High school or less</td>
</tr>
</tbody>
</table>

Gallup-Healthways Well-Being Index, Jan. 2, 2014-June 1, 2015

GALLUP

Survey Methods

Results are based on telephone interviews conducted Jan. 2, 2014-June 1, 2015, as part of the Gallup-Healthways Well-Being Index survey, with a random sample of 251,163 adults, aged 18 and older, living in all 50 U.S. states and the District of Columbia. For results based on the total sample of Americans, the margin of sampling error is ±0.2 percentage point at the 95% confidence level, including computed design effects for weighting.

Each sample of national adults includes a minimum quota of 50% cellphone respondents and 50% landline respondents, with additional minimum quotas by time zone within region. Landline and cellular telephone numbers are selected using random-digit-dial methods.

Sample Size =
Margin of Error =
Confidence Interval =
Unit 5: Supplementary Materials
Chapter 20: Interpreting Polls (Sampling, Margin of Error, and Confidence Intervals)

Example 3

Frequent Exercise Rates Among Various Demographic Groups in June 2015
% Who exercised for 30+ minutes 3 or more days in the last week

<table>
<thead>
<tr>
<th>GENDER</th>
<th>% Frequently exercised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>58.4</td>
</tr>
<tr>
<td>Female</td>
<td>52.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AGE</th>
<th>% Frequently exercised</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-29</td>
<td>64.5</td>
</tr>
<tr>
<td>30-44</td>
<td>55.4</td>
</tr>
<tr>
<td>45-64</td>
<td>52.9</td>
</tr>
<tr>
<td>65+</td>
<td>49.7</td>
</tr>
</tbody>
</table>

June 1-30, 2015
Gallup-Healthways Well-Being Index

GALLUP
Survey Methods

Results are based on telephone interviews conducted June 1-30, 2015, as part of the Gallup-Healthways Well-Being Index survey, with a random sample of 14,883 adults, aged 18 and older, living in all 50 U.S. states and the District of Columbia. For results based on the total sample of national adults, the margin of sampling error is ±1 percentage point at the 95% confidence level. All reported margins of sampling error include computed design effects for weighting.

Each sample of national adults includes a minimum quota of 50% cell phone respondents and 50% landline respondents, with additional minimum quotas by time zone within region. Landline and cellular telephone numbers are selected using random-digit-dial methods.

Example 4

Donald Trump Favorable Ratings -- Recent Seven-Day Averages
Based on U.S. adults

<table>
<thead>
<tr>
<th></th>
<th>Favorable</th>
<th>Unfavorable</th>
<th>No opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>Jul 22-28, 2015</td>
<td>31</td>
<td>57</td>
<td>12</td>
</tr>
<tr>
<td>Jul 15-21, 2015</td>
<td>33</td>
<td>54</td>
<td>13</td>
</tr>
<tr>
<td>Jul 8-14, 2015</td>
<td>31</td>
<td>57</td>
<td>12</td>
</tr>
</tbody>
</table>

Gallup Daily tracking
Survey Methods

The results reported in this article are based on seven-day averages of telephone interviews conducted on the Gallup U.S. Daily survey in July 2015, including July 8-14, July 15-21 and July 22-28. Each week’s average contains a random sample of approximately 1,170 adults, aged 18 and older, living in all 50 U.S. states and the District of Columbia. For results based on the total sample of national adults, the margin of sampling error is ±4 percentage points at the 95% confidence level.
What do you notice about the relationship between the sample size and the margin of error?

Also observe that with the margin of error (which is typically shown with the poll or survey), there is an associated confidence level (which is often not shown). For example, here is how polls and surveys are often shown in the media:

**TRUMP’S UP**

Who’s your first choice for the Republican nomination?

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donald Trump</td>
<td>17%</td>
</tr>
<tr>
<td>Jeb Bush</td>
<td>14%</td>
</tr>
<tr>
<td>Scott Walker</td>
<td>8%</td>
</tr>
<tr>
<td>Ted Cruz</td>
<td>6%</td>
</tr>
<tr>
<td>Marco Rubio</td>
<td>5%</td>
</tr>
<tr>
<td>Ben Carson</td>
<td>4%</td>
</tr>
<tr>
<td>Rand Paul</td>
<td>4%</td>
</tr>
<tr>
<td>Mike Huckabee</td>
<td>4%</td>
</tr>
<tr>
<td>Chris Christie</td>
<td>3%</td>
</tr>
</tbody>
</table>

(source: USA TODAY / Suffolk University Poll of 349 likely Republican primary/caucus voters, taken July 9-12. Margin of error ± 5.25 percentage points. Other candidates backed by 1% or less. Frank Pompey, USA TODAY)

**USA TODAY Snapshots®**

Types of transportation Americans want the federal government to pay more attention to:

- Trains/light rail: 56%
- Roads: 27%
- Buses: 21%
- Bike paths: 15%
- Sidewalks: 14%

(source: Hart Research Associates survey of 1,005 adults in January for National Association of Realtors and Transportation for America. Margin of error: ± 3.1 percentage points)

**Bay News 9**

Would you support or oppose a state constitutional amendment to legalize medical marijuana?

- Support: 68%
- Oppose: 24%
- Not Sure: 8%

(source: AM 820 News)

Note: Multiple responses allowed.

By Anne R. Carey and Sam Ward, USA TODAY
Notice in all the above examples that the margin of error is shown; however, the confidence level is not. There are two concepts at work here:

- **Precision** = Margin of Error
- **Confidence** = how sure they are that the actual population parameter will fall in the given margin of error.

Many people have a correct intuition when they read polls and the reported margin of error; however, what they often don’t realize (1) how small the sample size often is and (2) that there is an associated confidence with that margin of error. *(For example, I can be 100% confident that the mean score on our next test will be between 0-100; however, if I want to state that I think the mean will be in a smaller margin, say 65-85, then I can no longer be 100% confident, even if I have given a similar test in a similar class many times before.)*

**Example 5** Consider this example from a television report, stating 53% of the people most recently polled believed the economy was good, with the margin of error reported at the bottom right of the graphic as +3%. The confidence level (though not reported here) is likely 95%.

Thus, the interpretation of this poll would be that they are 95% confident that the true population parameter is 53% + 3%, so somewhere from 50%-56%.

So, when polls (such as presidential polls) are discussed and you hear reporters say that it is too close to call, because it is “within the margin of error,” also keep in mind that there is still probably a 5% chance that the margin of error itself could be off – something that is often not discussed.
(Optional) Group Activity 22: Sampling, Margin of Error and Confidence Intervals

1. Each group in the class has a bag of candy, which we will treat as samples. Record the following information from your sample below and on the board. We will use all of the candy as our population and you can see how close your sample comes to the proportions in the entire population.
   a. Record how many of each color as well as the total quantity.
   b. Based on your sample, what proportion of each color would you estimate there is within the population?
   c. Now record how many there were of each color in the entire class as well as the total quantity in the entire class.
   d. Based on this population data, what proportion of each color were there within the population?
   e. How close was your sample to modeling the proportions in the full population?

2. The graph at the right shows Americans’ satisfaction with government’s work in healthcare from 2001-2015.

   Using the last poll shown in 2015 (56% dissatisfied and 43% satisfied) observing the margin of error and the confidence interval described in the first paragraph under “survey methods,” answer the following questions:
   a. State the smallest percent of “Dissatisfied’s” and the greatest percent of “Satisfied’s” that is within the margin of error.
   b. State the greatest percent of “Dissatisfied’s” and the smallest percent of “Satisfied’s” that is within the margin of error.
   c. What is the percent chance that the true opinion breakdown of all Americans does NOT fall somewhere between what you listed in part (a) to what you listed in part (b)?
3. Recently, Scotland had a referendum where its citizens voted on if they wanted to be an independent country. What is wrong with the CNN poll at the right?

4. You can see the polling numbers on Scottish independence that were recorded over a roughly one month period.

   The last poll, from early September, shows 51% “Yes” and 49% “No”.

   Reading below, see that the margin of error is ±3%.

   a. State the greatest percent of “Yes’s” and the smallest percent of “No’s” that is within the margin of error.

   b. Is “No” winning within the margin of error? If so, state the greatest percent of “No’s” and the smallest percent of “Yes’s” that is within the margin of error.

5. The actual vote took place September 18th (so very shortly after the last poll above).

   Are the final results shown within the margin of the poll in (6)?
   If not, explain how this is possible.
Chapter 21: Voting Theory: Determining Winners with Preferential Ballots

Objectives:
1. Determine the winner, using preferential ballots, with plurality, plurality-with-elimination, Borda-Count, and pairwise comparison techniques.
2. Recognize limitations of various techniques.
3. Identify real-life examples of these methods.

Voting is not as simple as it may seem.

In some situations, such as our electoral college system for electing a president, there are very apparent disparities, which we will cover in the section; however, even in seemingly small, simple situations (where everybody gets 1 vote that is counted equally), the method we use can often lead to different results.

The goal is to find a technique that is always fair; however, there is no such technique. Different situations may lend themselves more naturally to the different techniques we are going to discuss.

Preference Schedules
- A preference ballot asks voters to rank their choices by preference.
- A preference schedule is the most efficient way to summarize those results from preference ballots.

Example 1 Suppose a scholarship committee of 19 individuals are reviewing applicants for a lucrative scholarship. The committee has narrowed it down to 3 students (Alicia, Brenden, and Candice) and they have now all voted, using preference ballots, which are summarized in the preference table below.

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>A</th>
<th>A</th>
<th>B</th>
<th>B</th>
<th>C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Total Votes (for that ballot):</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Who do you think should get the scholarship? Why?
Method 1: Plurality Method

The most first place votes wins (choice with most first place votes is said to have *plurality*).

**Example 2** Who should get the scholarship in example 1 under the plurality method?

Number of first place votes for A:

Number of first place votes for B:

Number of first place votes for C:

Therefore, has plurality and thus would win under the plurality method.

Did the winner have a majority?

If you are Brenden, make a mathematical argument why you think you should win and not Alicia.

---

Method 2: Plurality with Elimination Method

- Count the number of first place votes.
- If first place has majority: Done – they win.
- If no majority: Eliminate candidate with fewest first place votes.

The goal with plurality with elimination is to obtain a choice that has the majority.
Example 3  Who should get the scholarship in example 1 under the plurality with elimination method?

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>A</th>
<th>A</th>
<th>B</th>
<th>B</th>
<th>C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

Total Votes (for that ballot): 5 3 3 4 3 1

Round 1:

Round 2:

Therefore, wins under the plurality with elimination method.

Method 3: Borda-Count Method

Points assigned to order of choices. The goal here is the best compromise candidate.

Example 4  Who should get the scholarship in example 1 under Borda-Count method?

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>A</th>
<th>A</th>
<th>B</th>
<th>B</th>
<th>C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

Total Votes (for that ballot): 5 3 3 4 3 1

Borda Points for A:

Borda Points for B:

Borda Points for C:

Therefore, wins under the Borda-Count method.
Method 4: Pairwise Comparisons

Form all head-to-head competitions and tally who wins the most of these pairwise competitions.

**Example 5** Who should get the scholarship in example 1 under pairwise comparisons?

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>A</th>
<th>A</th>
<th>B</th>
<th>B</th>
<th>C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Total Votes (for that ballot):</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

A vs B: 
\[
\begin{align*}
A &= \\
B &= \\
C &= 
\end{align*}
\]

Therefore, wins under the pairwise comparisons.

**Real World Examples**

- Academy Awards “Best Picture” Oscar – Plurality with Elimination:


  “The academy uses preferential voting, as it does to determine nominees in most other races, to determine the winner of Best Picture.

  This method of voting was reintroduced in 2009, when the academy went to 10 Best Picture nominees for the first time since 1943 and was kept in place in 2010 when the number of nominees was to be somewhere between five and 10. The academy believes this "best allows the collective judgment of all voting members to be most accurately represented."

  The preferential method was first used in 1934 when there were 12 Best Picture nominees (there had been between three and 10 in the first six years of the Academy Awards) and was used the following year when there were again 12 nominees, from 1936 to 1943 when there were 10 nominees, and in both 1944 and 1945 when there were just five contenders.

  Voters rank the Best Picture nominees. If one nominee garners more than 50% of the first place votes, it will win Best Picture.

  If, as is more likely, no nominee reaches this threshold, the film with the fewest first place votes is eliminated, with its ballots being reapportioned to the second place choice.
Should no film cross the required 50% + one ballot threshold, the film with the fewest first place votes is again eliminated, with its ballots being apportioned to the next choice still in play (i.e., if the second place choice is no longer in the running, then the ballot would be reapportioned to the third place choice and so on.)

This process of elimination and reapportion continues until one film reaches at least 50% + one ballots. That is the Best Picture winner.”

- Heisman Trophy (MVP for college football) - Borda Count.
  From http://www.wou.edu/~bblair10/blairweb/heisman.html:
  “To determine who will be the recipient, a vote is conducted with 870 select members of the media each receiving one vote, all previous Heisman Trophy winners getting a vote and one vote being awarded to the general public's vote. Voters select a first, second and third choice from all active college football players for that season.

  Then, each players first place votes is multiplied by three, their second place votes are multiplied by two and first place votes are multiplied by one. Then, the three numbers are added together for the player's grand total.”

- Olympic Venues – variation of plurality with elimination.
  “Each of the 100-plus IOC members gets one vote, and after every round the city with the fewest votes is eliminated. In the competition for the 2012 Games, Moscow expired in the first round, New York in the second. Madrid was the top vote-getter in Round 2, but it got the ax in Round 3. London edged out Paris, 54–50, in the final vote.

  Let’s suppose Chicago really is the favorite for 2016: It’s the highest-ranked city on average, and it beats all the others in head-to-head votes. Voters can rank the four candidates in 24 different ways, but say they sort themselves into five equal-size factions with the preferences in SET A.

  As one can easily confirm, Chicago is highest-rated city on average, and it beats all the others in paired votes. Everybody likes Chicago a lot—but nobody likes it most of all, and under the IOC’s rule Chicago is the first to go. Tokyo is the eventual winner.

  But the IOC’s rule can also work to Chicago’s advantage, even if it’s not deserving. Imagine that the five blocs instead rank the cities as in SET B.

  Now Chicago wins the 2016 Games–Round 1 eliminates Tokyo, Round 2 Rio and Round 3 Madrid—even though, given that set of preferences, Chicago is a poor choice. Three blocs rank it near the bottom; two think it’s worst of all. It’s not the highest
rated city on average—Madrid is. And the only city Chicago beats in a paired vote is Madrid.

But maybe all this really shows is that the IOC should scrap its current rule and come up with another. Wouldn’t it be better to choose the highest-ranked city on average? Well, not so fast. Let’s imagine the IOC’s voting blocs have the orderings in SET C. Tokyo’s average ranking is the highest, so it prevails under the new rule.

But suppose the IOC had made Doha a finalist instead of Rio, and imagine the voting blocs rank the four cities this way, as in SET D.

Nothing has changed in the rankings except Rio’s replacement by Doha: the relative ratings of Chicago, Madrid and Tokyo are exactly the same. But everything has changed in the outcome: Chicago now has a higher average than Tokyo, and it wins the 2016 Games. Thank you, Doha!

Okay, so maybe the IOC picked the wrong rule again. What if it awards the Games to the city that beats all the others in head-to-head votes? Sorry. Paired votes can “cycle,” yielding, for example, Chicago over Rio, Rio over Madrid, Madrid over Tokyo and Tokyo over Chicago. Nobody wins!

For centuries philosophers, mathematicians, political scientists and economists have searched for the best method of voting. Fifty-eight years ago the economist Kenneth Arrow (later a Nobel laureate) decided to see whether any voting rule could avoid the problems we’ve illustrated. Fix them all at once, he found, and you get—a dictatorship. One voter calls the shots every time. Arrow’s “impossibility theorem” demonstrates that no system of voting always gives the “right” result.

In the Olympics of voting, there are no gold medalists.”
1. Suppose you are on a hiring committee at work that consists of a total of 11 individuals. The committee brought the top three candidates in for interviews and the committee members have ranked their preferences for the new hire as follows:

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>Candidate A</th>
<th>Candidate C</th>
<th>Candidate B</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>Candidate C</td>
<td>Candidate A</td>
<td>Candidate A</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>Candidate B</td>
<td>Candidate B</td>
<td>Candidate C</td>
</tr>
<tr>
<td>Total Votes:</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

a. Who will get the job if votes are tabulated with the plurality method?

b. Who will get the job if votes are tabulated with the plurality-with-elimination method? (Show work, rewriting any additional preference schedule needed, until you get to a majority.)

c. Who will get the job if votes are tabulated with the Borda-Count method? (Show the Borda points that each candidate earns.)

d. Who will get the job if votes are tabulated with the pairwise competitions method? (Show the results of each pairwise competition.)
2. You live in a condominium and the homeowner’s association is electing a new president, and there are three people running. There are a total of 45 units in your building, but 4 of them are currently occupied by renters, who do not get a vote. Thus, there a total of 41 voting members and their votes are as follows.

<table>
<thead>
<tr>
<th>1st Choice:</th>
<th>Candidate B</th>
<th>Candidate C</th>
<th>Candidate C</th>
<th>Candidate A</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Choice:</td>
<td>Candidate A</td>
<td>Candidate B</td>
<td>Candidate A</td>
<td>Candidate C</td>
</tr>
<tr>
<td>3rd Choice:</td>
<td>Candidate C</td>
<td>Candidate A</td>
<td>Candidate B</td>
<td>Candidate B</td>
</tr>
<tr>
<td>Total Votes:</td>
<td>9</td>
<td>16</td>
<td>4</td>
<td>12</td>
</tr>
</tbody>
</table>

a. Who will win if votes are tabulated with the plurality method?

b. Who will get the job if votes are tabulated with the plurality-with-elimination method? (Show work, rewriting any additional preference schedule needed, until you get to a majority.)

c. Who will get the job if votes are tabulated with the Borda-Count method? (Show the Borda points that each candidate earns.)

d. Who will get the job if votes are tabulated with the pairwise competitions method? (Show the results of each pairwise competition.)
Chapter 22: Voting Theory: United States Voting Systems (Pros and Cons)

Objectives:

1. Explain how senators and representatives are elected to the Senate and House of Representatives.
2. Explain gerrymandering and its potential impact on congressional elections.
3. Discuss pros and cons of the presidential primary process.
4. Explain how the electoral college system is used to elect the US President.
5. Identify pros and cons of the electoral college system.

Congress

Senate – 2 senators per state (so 100 total); elected by state popular vote for 6 year terms

Pros? Cons?

Who are Illinois’ current senators?

House of Representatives – number determined by the state population (currently 435 total); elected by the congressional district* popular vote for 2 year terms

Pros? Cons?

How many representatives does Illinois currently have?

Who is Elgin’s’ current representative?

*Congressional districts

- supposed to have roughly the same population
- state governor and legislature determine the boundaries (using US Census)
- if new census shows significant changes, the boundaries can be changed – known as s the redistricting

Why might redistricting be controversial?
**Gerrymandering** - redistricting in a way that benefits a specific party's interests


**Examples**
Do the districts seem reasonable?

**Here are Illinois’ Congressional districts.**

**Here are Chicagoland’s districts.**

**How about this district in North Carolina?**

**How about this district in Maryland?**
Presidential Primary Process

- States hold primaries (state run elections) and caucuses (party run events) over the first half of the year (typically from January to June) to select the Republican and Democratic nominees for President.
  - Actually voting for delegates for the party’s national convention that will vote accordingly.
- 2016 Primary and Caucus Schedule (as of 7/13/15 from http://www.uspresidentialelectionnews.com/2016-presidential-primary-schedule-calendar/)

<table>
<thead>
<tr>
<th>Date</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday, February 1</td>
<td>Iowa caucus</td>
</tr>
<tr>
<td>Tuesday, February 9</td>
<td>New Hampshire</td>
</tr>
<tr>
<td>Saturday, February 20</td>
<td>South Carolina</td>
</tr>
<tr>
<td>Tuesday, February 23</td>
<td>Nevada caucus</td>
</tr>
<tr>
<td>Tuesday, March 1 (Super Tuesday)</td>
<td>Alabama, Arkansas, Colorado caucuses, Georgia, Massachusetts, Minnesota caucuses, North Carolina, Oklahoma, Tennessee, Texas, Vermont, Virginia</td>
</tr>
<tr>
<td>Saturday, March 5</td>
<td>Louisiana, Nebraska (Dem caucus)</td>
</tr>
<tr>
<td>Tuesday, March 8</td>
<td>Hawaii caucus (GOP), Mississippi, Michigan</td>
</tr>
<tr>
<td>Sunday, March 13</td>
<td>Puerto Rico (GOP)</td>
</tr>
<tr>
<td>Tuesday, March 15</td>
<td>Ohio, Florida, Illinois, Missouri</td>
</tr>
<tr>
<td>Tuesday, March 22</td>
<td>Arizona, Utah</td>
</tr>
<tr>
<td>Saturday, March 26</td>
<td>Hawaii caucus (Dems)</td>
</tr>
<tr>
<td>Tuesday, April 5</td>
<td>Wisconsin</td>
</tr>
<tr>
<td>Tuesday, April 26</td>
<td>Connecticut, Delaware, Maryland, Pennsylvania, Rhode Island</td>
</tr>
<tr>
<td>Tuesday, May 3</td>
<td>Indiana</td>
</tr>
<tr>
<td>Tuesday, May 10</td>
<td>Nebraska (GOP primary), West Virginia</td>
</tr>
<tr>
<td>Tuesday, May 17</td>
<td>Kentucky, Oregon</td>
</tr>
<tr>
<td>Sunday, June 5</td>
<td>Puerto Rico (Dem)</td>
</tr>
<tr>
<td>Tuesday, June 7</td>
<td>California, Montana, New Jersey, New Mexico, South Dakota</td>
</tr>
</tbody>
</table>
Unit 5: Supplementary Materials
Chapter 22: Voting Theory: United States Voting Systems (Pros and Cons)

<table>
<thead>
<tr>
<th>Date</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday, June 14</td>
<td>Washington, DC</td>
</tr>
<tr>
<td>States with no firm dates</td>
<td>New York, North Dakota, Utah, Colorado, Idaho, Kansas, Maine, Washington, Wyoming</td>
</tr>
</tbody>
</table>

Note: This order stays roughly the same (especially Iowa and New Hampshire) every election.

Pros? Cons?

The following editorial is from four elections ago, but it is as relevant today as it was then. It was an editorial that appeared in the Chicago Tribune:

**Illinois primary is an exercise in futility**
January 26, 2004 | By David Orr, Cook County Clerk.

Chicago — Every vote should count. Too bad they won’t for Illinois voters who will likely have no say in determining the presidential nominees this year.

An increasingly front-loaded nominating schedule means that 34 states will have conducted primaries or caucuses before Illinois gets a turn March 16. By that time, the frontrunners from both parties will probably have secured the required delegates needed to lock up the nominations.

Illinois not only loses relevance and influence under the current setup, but the Land of Lincoln also disenfranchises voters from the nominating process and threatens to depress turnout in local races.

Illinois isn’t alone. Even if the nominations are still up for grabs for Illinois voters in mid-March, voters from many other states will take part in primaries in which the presidential nominees are a foregone conclusion. This has prompted several states to take action.

To save money, seven states have cancelled their primaries this year, replacing them with caucuses that are funded by the parties. But a caucus system puts the decision-making into the hands of the party faithful and hard-core activists. This results in fewer voters who are expressing significantly narrowed opinions.

Other states have moved their primaries to earlier dates so their voters will have more impact. But accelerating the primary season places hardships on underfunded, underdog candidates with fewer resources and less time to get their message across to voters, and it favors more established, well-heeled candidates.

Despite obvious shortcomings, we shouldn’t scrap the primary system altogether. Instead we should develop a national or regional primary system that would eventually give all states a voice in anointing party nominees.

**Under a regional primary system, the nation would be divided into four regions of comparable populations that would hold primary elections a month apart beginning in February. Every four years the order would rotate, giving each region a shot at going first, thereby distributing regional influence and significance.**
This sensible alternative would provide greater fairness and ample time for candidates to debate broader issues that would address concerns of a more diverse electorate in the beginning of the process.

Until then, the presidential primary in Illinois will remain an exercise in futility, as we are forced to let voters from other states decide who will represent us.

**Electoral College**

- States have a number of electoral votes that equals the number of senators (2) plus the number of representatives (see earlier discussion).
  - So, that is 100 + 435 + The District of Columbia gets 3 electoral votes = 538 electoral votes.
  - A majority of electoral votes are needed (so 270) to win.
  - The actual electors “pledge” to vote for the candidate elected in their state.
  - That’s right – ALL electoral votes for a state go to the winning candidate, regardless of whether or not that candidate won by a landslide or a narrow margin.

- In the 2012 presidential, Illinois’ results were:

<table>
<thead>
<tr>
<th>Presidential Candidate</th>
<th>Vice Presidential Candidate</th>
<th>Political Party</th>
<th>Popular Vote</th>
<th>Electoral Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barack H. Obama</td>
<td>Joseph R. Biden, Jr.</td>
<td>Democratic</td>
<td>3,019,512</td>
<td>20</td>
</tr>
<tr>
<td>Willard Mitt Romney</td>
<td>Paul Ryan</td>
<td>Republican</td>
<td>2,135,216</td>
<td>0</td>
</tr>
</tbody>
</table>

Pros? Cons?
According to http://freedomslighthouse.net/election/2014-2016/2016-presidential-election-electoral-vote-map-polls-projections/, downloaded 5/19/15, the 2016 projections are

Pros? Cons?

Different perspectives on the Electoral College:


“The Electoral College is as relevant today as it was in 1789. The Founders realized the danger of subjecting fundamental rights to the whim of the majority, so they devised a system that protects smaller rural states (like Oklahoma). Each state is apportioned a number of electors based on population, as expressed in the number of its members in Congress, and the political parties nominate electors who are expected, but not required, to cast their votes for that party’s presidential candidate.

In the modern era, this system ensures political candidates can’t ignore smaller states in their quest for the magic 270 electoral votes. The alternative now promoted would encourage presidential candidates to focus exclusively on the big cities in the big states: California, Texas, New York, Massachusetts, Florida and Illinois. No candidate would ever waste his time in New Hampshire, or in Kansas or Montana (or Oklahoma). Residents of the flyover states could read about it in the newspapers.
The popular-vote movement was revived in 2000 after Al Gore won 500,000 more votes for president than George W. Bush, but Mr. Bush put together 271 electoral votes, just enough to win. A group of sore losers formed the National Popular Vote, funded by several wealthy and liberal businessmen, to shutter the Electoral College. They were joined by left-wing groups such as Common Cause, which seeks to eliminate procedural safeguards to citizens’ rights, such as the Senate filibuster.”


“If you live in New York, you can be forgiven for failing to pay attention to the presidential campaign because your vote, in all likelihood, will be meaningless. Have you seen Mitt Romney or Barack Obama around your neighborhood recently? You won’t, unless you live in Cincinnati or Miami and are just visiting in New York.

The reason we and many other states’ residents don’t matter in a presidential election year is a weird, vestigial voting system called the Electoral College. This system, which was unfortunately set up in the Constitution, has made much mischief over the past 200-plus years and promises to make much more, perhaps as soon as this November.

Nate Silver, the statistician who works for the New York Times analyzing poll data in his blog, FiveThirtyEight, wrote a fascinating post early last week about the chances of the 2012 presidential election ending up in a 269-269 tie in the Electoral College. The chance of a tie is small — about 0.6 percent, according to Mr. Silver — but significant. An Electoral College tie would throw the election into the House of Representatives, which would most likely end up in a Romney victory, and create enough gall among Democrats to poison U.S. politics for the next decade.

In a country where millions of people can vote by phone for their favorite singer, we should not be using a system most people don’t understand that could result in a tie for the country’s most important job.

And a tie is not the worst-case scenario. **We experienced the worst-case scenario in 2000, when most people voted for one candidate — Al Gore — but the other one — George W. Bush — won in the Electoral College.**

In 1824, 1876, 1888 and 2000, the candidate with the most votes lost in the Electoral College. Four times, the American people have had to live with a president most of us voted against.

If you live in a state like New York, which is reliably Democratic, you can expect to be ignored during a presidential election. Since the winner takes all of a state’s electoral votes (except in Maine and Nevada), individual votes mean nothing in states where the winning party can be predicted.

**Voters in swing states like Florida and Ohio, where the parties’ popularity is neck and neck, get catered to like prodigal children every four years.** Presidents have many favors to give. As long as we have the Electoral College, those favors will get steered to a handful of states, especially in the years, like this one, when a president is trying to get re-elected.
Also, because electors are not awarded in proportion to a state’s population, but are based on the size of its congressional delegation, states without many people wield more proportional power in the Electoral College than populous states like New York.

You can argue that, if we switch to a popular vote for presidential elections, candidates will ignore states without many people, like Wyoming. To the extent that happens, it will be fair. Candidates will naturally gravitate to population centers like the big cities of Texas, California, Florida and the Northeast Corridor, and that is as it should be.

In a popular vote, the majority will rule every time, which we believe is better than a system in which, sometimes, the minority will rule.

The Electoral College is out of date. It is unfair. It is prone to fiascos. It will be hard to get rid of, since it was established in the Constitution, but we should start work now on an amendment.”
Appendix A
Practice Test Answers

Practice Test Answers

Practice Test 1 Answers
1. a. relative  b. absolute  c. relative  d. absolute
2. \[
\frac{114}{655 + 114 + 151} \approx 12.4\%
\]
3. \[38,111,341 - 0.21 \times 34,111,341 = 0.79 \times 38,111,341 = 30,107,959\] primary voters
4. a. 28
   b. \[
   \frac{70 - 42}{42} \approx 67%
   \]
   c. \[
   \frac{121}{41} \approx 2.95 \text{ times as many medals}
   \]
5. 31,108,968 is 162% of the number in 2012, so \[31,108,968 = 1.62(2012\ \text{total})\]. The 2012 total is thus \[
\frac{31,108,968}{1.62} \approx 19,203,067\] primary voters in 2012
6. Final price = (original price) – (30% of original price) – (15% of sale price).
   Let’s let \(x = \) original price. Then the final price is:
   \[
   \text{Final price} = x - 0.3x - 0.15(x - 0.3x) = 0.7x - 0.15(0.7x) = 0.7x - 0.105x = 0.595x.
   \]
   This means the final price is 59.5% of the original price.
7. Different states have different populations, so looking at total murders is not appropriate to compare states.
8. The District of Columbia had the highest rate in 2010 of 21.8 murders per 100,000 residents.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2010</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>District of Columbia</td>
<td>41.8</td>
<td>21.8</td>
<td>15.9</td>
</tr>
<tr>
<td>Louisiana</td>
<td>12.5</td>
<td>11.0</td>
<td>10.3</td>
</tr>
<tr>
<td>Maryland</td>
<td>8.1</td>
<td>7.4</td>
<td>6.1</td>
</tr>
</tbody>
</table>

9. The District of Columbia also had the highest rate in 2014 of 15.9 murders per 100,000 residents.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2010</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>District of Columbia</td>
<td>41.8</td>
<td>21.8</td>
<td>15.9</td>
</tr>
<tr>
<td>Louisiana</td>
<td>12.5</td>
<td>11.0</td>
<td>10.3</td>
</tr>
<tr>
<td>Mississippi</td>
<td>9</td>
<td>6.9</td>
<td>8.6</td>
</tr>
</tbody>
</table>

10. The District of Columbia also had the largest absolute decrease in its homicide rates – declining from 21.8 per 100,000 residents in 2010 to 15.9 per 100,000 residents in 2014.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2010</th>
<th>2014</th>
<th>Absolute Change in Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>District of Columbia</td>
<td>41.8</td>
<td>21.8</td>
<td>15.9</td>
<td>-5.9</td>
</tr>
<tr>
<td>New Mexico</td>
<td>7.4</td>
<td>6.8</td>
<td>4.8</td>
<td>-2</td>
</tr>
<tr>
<td>Arizona</td>
<td>7</td>
<td>6.4</td>
<td>4.7</td>
<td>-1.7</td>
</tr>
</tbody>
</table>
11. Massachusetts had the largest percent decrease in the rate, decreasing 39.39% from 2010 to 2014.

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2010</th>
<th>2014</th>
<th>Absolute Change in Rate</th>
<th>Percent Change in Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massachusetts</td>
<td>2</td>
<td>3.3</td>
<td>2</td>
<td>-1.3</td>
<td>-39.39%</td>
</tr>
<tr>
<td>Connecticut</td>
<td>2.9</td>
<td>3.7</td>
<td>2.4</td>
<td>-1.3</td>
<td>-35.14%</td>
</tr>
<tr>
<td>New York</td>
<td>5</td>
<td>4.5</td>
<td>3.1</td>
<td>-1.4</td>
<td>-31.11%</td>
</tr>
</tbody>
</table>

12. This apparent contradiction occurs because Washington, D.C. had a much larger absolute value to begin with (21.8 murders per 100,000 residents in 2010), but the percent change was greater in Massachusetts, because the rate was so much smaller to begin with (3.3 murders per 100,000 residents in 2010).
Appendix A
Practice Test Answers

Practice Test 2 Answers

1. The scale in the left-hand portion is inaccurate. The blue bar is triple the size of the purple, but only 1.1 million more (29.9 million versus 28.8 million).

2.
   a. 50%
   b. Between 25% and 50%
   c. 25%
   d. 79
   e. Range = 87 – 45 = 42; IQR = 79 – 60 = 19
   f. Range = 92 – 65 = 27; IQR = 85 – 72 = 13
   g. Class 1 has the larger standard deviation, as its range and inter-quartile range are both larger than those for Class 2.

3.
   a. No, this was just an observational study (“...people who report spending...”).
   b. Answers may vary. A lurking variable is one which would cause someone to spend more time on social media and also increase the perception of social isolation.

4.
   a. 400 and 600
   b. 300 and 700
   c. 200 and 800
   d. 200, 300, 400, 500, 600, 700, 800
   e. 16th percentile
   f. 97.5%
   g. 47.5%
   h. 600

5.
   a.

   ![Graph](image)

   b. The vertical scale is not proportional.
   c. The time intervals are also not equivalent – the first two represent decades, while the latter only represents a 5-year interval.
Appendix A
Practice Test Answers

6. 
   a.

b. The absolute maximum was 1.64 fatalities per 100 million vehicle miles driven in 1982.

c. 
   i. The graph decreases quickly from roughly 1982 to 1992.
   ii. The graph decreases slowly from roughly 1994 to 2004.

7. 
   a. 71.0 minutes
   b. 75 minutes
   c. 12.8 minutes

8. 
   a. 69th
   b. 59.9%
   c. 584
   d. 664
Appendix A  
Practice Test Answers

Practice Test 3 Answers

1. 
   a. (0,75) 
   b. The base pay for a salesperson with $0 in sales is $75/day. 
   c. \( m = 0.25 \) 
   d. For each additional $1 in sales, the total pay increases $0.25. Equivalently, each $100 increase in sales results in an increase of $25 in total pay. 
   e. \( y = 0.25x + 75 \) 
   f. \( y = 0.25(430) + 75 = \$182.50 \)

2. 
   a. Peak temperature 
   b. Average units of electricity used 
   c. [0,32] 
   d. [6,48] 
   e. Quadratic

3. 
   a. \( A = 4500 \cdot 1.035^t \) 
   b. \( A = 4500 \cdot 1.035^{25} \approx \$10,634.60 \)

4. 
   a. \( \frac{\$250,000}{236.7} = \frac{49.3}{1} \Rightarrow \text{the value in constant 2014 dollars is } \$1,200,304.26. \) 
   b. Clearly the 1974 salary was better relative to the time.

5. 
   a. This means that the prices of the commercials in the late 1980s were increasing at about the same as the prices for other goods and services. 
   b. This means the increases in the prices of commercials in the late 1990s were more than the increases in the prices of other goods and services.

6. 
   a. 
   
   
   
   

   b. About 855.41(25) + 8827.4 = 30,213 cases.
Appendix A
Practice Test Answers

7. 
   a.

   ![Graph of annual profit from senior citizens]

   b. About $382,749, so about $382,749.

8. The quadratic in #2 fits much closer, $R^2$ is nearly equal to 1.

9. 

<table>
<thead>
<tr>
<th>Year</th>
<th>South Africa</th>
<th>Kenya</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>43,997,828</td>
<td>36,913,721</td>
</tr>
<tr>
<td>2008</td>
<td>43,795,438</td>
<td>37,947,305</td>
</tr>
<tr>
<td>2009</td>
<td>43,593,979</td>
<td>39,009,830</td>
</tr>
<tr>
<td>2010</td>
<td>43,393,447</td>
<td>40,102,105</td>
</tr>
<tr>
<td>2011</td>
<td>43,193,837</td>
<td>41,224,964</td>
</tr>
<tr>
<td>2012</td>
<td>42,995,145</td>
<td>42,379,263</td>
</tr>
<tr>
<td>2013</td>
<td>42,797,368</td>
<td>43,565,882</td>
</tr>
</tbody>
</table>

   Kenya passed South Africa some time during 2012.

10. 
    a. In about 2032, so 25 years.

    b. In about 2157, so 150 years.
Practice Test 4 Answers

1. We would want the first account, because of the more frequent compounding.

2. We would want the second account – the first account is the annual percentage yield, where the compounding has already been factored in, but the second is only the annual percentage rate.

3. 20% down means you need 80%. 80% of $160,000 = 0.8*160,000 = $128,000.

4. 
   a. Total cost = cost from payments + closing costs + points
      = $1,049.21*12*30 + $500 = $378,215.60
   b. Total cost = cost from payments + closing costs + points
      = $1,021.25*12*30 + $1,500 + 0.02*$175,000 = $372,650
   c. We would choose the second. The additional closing costs were worth the lower rate in the long run.

5.

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Beginning</th>
<th>Interest</th>
<th>Ending</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$6,500</td>
<td>$56.88</td>
<td>$6,556.88</td>
</tr>
<tr>
<td>2</td>
<td>$6,556.88</td>
<td>$57.37</td>
<td>$6,614.25</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>18</td>
<td>$7,537.61</td>
<td>$65.95</td>
<td>$7,603.56</td>
</tr>
<tr>
<td>19</td>
<td>$7,603.56</td>
<td>$66.53</td>
<td>$7,670.10</td>
</tr>
<tr>
<td>20</td>
<td>$7,670.10</td>
<td>$67.11</td>
<td>$7,737.21</td>
</tr>
</tbody>
</table>

6. APY = ($6,730.50 - $6,500)/$6,500 = 3.55%

7. 
   a.

<table>
<thead>
<tr>
<th>Month</th>
<th>Beginning Balance</th>
<th>Interest</th>
<th>New Deposit</th>
<th>End Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>238</td>
<td>$65,885.85</td>
<td>$315.70</td>
<td>$150</td>
<td>$66,351.56</td>
</tr>
<tr>
<td>239</td>
<td>$66,351.56</td>
<td>$317.93</td>
<td>$150</td>
<td>$66,819.49</td>
</tr>
<tr>
<td>240</td>
<td>$66,819.49</td>
<td>$320.18</td>
<td>$150</td>
<td>$67,289.67</td>
</tr>
</tbody>
</table>

   b. You deposited $150*12*20 = $36,000
   c. Total interest is $67,289.67 - $36,000 = $31,289.67.

8. 
   a. =FV(5.75%/12,12*25,120) = $80,032.15.
   b. You deposited $120*12*25 = $36,000
   c. Total interest is $80,032.15 - $36,000 = $44,032.15.

9. Clearly option 2 is better – the compounding makes all the difference!

10. =PMT(0.065/12,12*30,0,500000) = $452.01
Appendix A
Practice Test Answers

11. 
   a. \(\text{PMT}(0.0425/12,3*12,7000,0) = $207.45\)
   b. \(207.45*3*12 = $7,468.10\)
   c. Interest = $468.10

12. 
   a. 
      \[
      \begin{array}{c|c|c|c|c}
        \text{Month} & \text{Beginning Balance} & \text{Payment} & \text{Interest} & \text{End Balance} \\
        \hline
        1 & $1,000 & $25.00 & $14.17 & $989.17 \\
        2 & $989.17 & $24.73 & $14.01 & $978.45 \\
        \vdots & \vdots & \vdots & \vdots & \vdots \\
        79 & $37.80 & $20.00 & $0.54 & $18.34 \\
        80 & $18.34 & $18.60 & $0.26 & $0.00 \\
      \end{array}
      \]
   
   b. Total spent = $1,650.44
   
   c. 
      \[
      \begin{array}{c|c|c|c|c}
        \text{Month} & \text{Beginning Balance} & \text{Payment} & \text{Interest} & \text{End Balance} \\
        \hline
        1 & $1,000 & $45.00 & $14.17 & $969.17 \\
        2 & $969.17 & $45.00 & $13.73 & $937.90 \\
        \vdots & \vdots & \vdots & \vdots & \vdots \\
        26 & $82.71 & $45.00 & $1.17 & $38.89 \\
        27 & $38.89 & $39.44 & $0.55 & $0.00 \\
      \end{array}
      \]
   
   d. Total spent = $1,209.44
   
   e. Total saved = $1,650.44 - $1,209.44 = $441