Open Educational Resources   Queensborough Community College

2018

Mathematics in Contemporary Society - Chapter 3 (Spring 2018)

Patrick J. Wallach
Queensborough Community College

[How does access to this work benefit you? Let us know!](#)
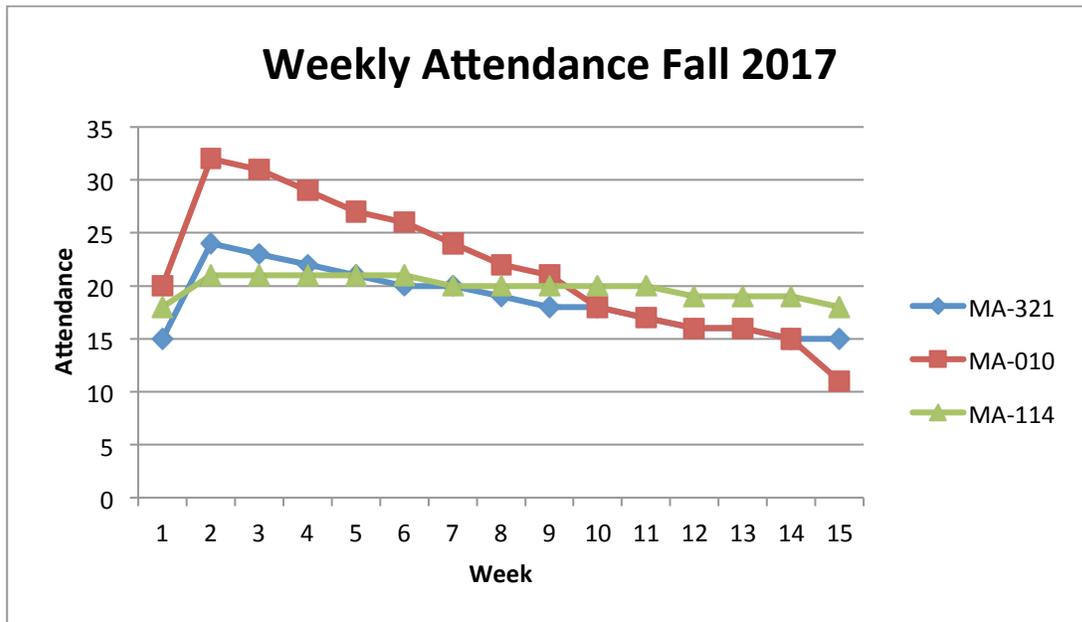
Chapter 3

Time-Series Diagrams

A **time-series diagram** is a histogram, bar graph or line chart in which the horizontal axis is categories of time.

Example: Suppose Professor Hernandez tracks the weekly attendance of his courses. The following time-series diagram is created:

Each line represents the attendance for a different course. Therefore, we can determine the attendance for any given week for the entire semester. During Week 1, the attendance for MA-321 is 15 students; the attendance for MA-010 is 20 students and the attendance for MA-114 is 18 students. During Week 2, all values go up. MA-321 has 24 students; MA-010 has 32 students; MA-114 has 21 students. We can determine the attendance for any other week as well.

Question 1: What is the attendance for each class during week 9?

What is also important to the reader are the overall trends for the entire length of the semester. We can observe that:

1)   The attendance rises in the second week and drops afterward for all courses. This makes sense, because students register for courses often until sometime in the second week. After this, it is no longer possible to register for a course. Attendance can only decrease, as students drop or leave the course.

2)   MA-010 has the greatest rise and fall in attendance (from 20, up to 32, then all the way down to 11 in Week 15). This class shows the greatest **variation** (a topic to be discussed in more detail later) between attendance values for the 15 weeks. The professor teaching this course may want to address this problem in the future.

3)   MA-114 shows the least change in attendance throughout the semester, with weekly attendance between 18 and 21 students. MA-114 has the least variation. Perhaps this course is much easier for students than MA-010.

4)   MA-321 has greater variation than MA-114 but less than MA-010. Attendance rises from 15 to 24 in the second week, and slowly decreases to 15 students by the end of the semester.

Question 2:   Susan owns stock in three companies: Apex Inc., Acme Inc., and Global Inc. At the end of the year, she examines the following time-series diagram:

a) Describe the activity of each stock separately.
b) Is there an overall trend for all of the stocks?
c) Which stock shows the most variation? Which shows the least?
d) Which stock was the best investment in 2017?
e) Overall, how did Susan do with her stock portfolio for the year? Is there anything she could have done to increase her profit?
f) If you had $1000 to invest for 2018, which stock would you choose? Why?

Note: In describing a time-series diagram, we shouldn't use terms such as "left-skewed", "right-skewed", "normal", etc. These terms generally apply to histograms and line charts created from frequency tables.

Writing Assignment:

Find a news article (in the newspaper or on the internet) that contains a statistical display of some type (frequency table, bar graph, pie chart, histogram, line chart, or time-series diagram). Write a 1-1½ page paper about the article, focusing on the following questions:

1) What is the article about?
2) What information is displayed by the graph?
3) Describe whether or not the information in the article is supported by the information in the graph. If it is, give details. If it is not, explain how the graph gives a different message.
4) Is there anything that could have been done to make the graph better and more useful for whoever reads the news article?

You can submit your paper by e-mail or place it in Blackboard. You should provide a copy of the article itself as well. If it is a newspaper article, you can scan it and send the file. If it is an internet article, give me the link where it can be found.

More Complicated Graphs

We've already seen some of the basic graphical displays of data sets (bar graph, histogram, pie chart, line chart). We can now move on to fancier, more detailed displays that we can find in most newspapers and on the internet.

A **multiple bar graph** allows for comparison of two or more data sets under the same categories.

Consider the following multiple bar graph:

QCC Students with Cell Phones 2010 - 2015
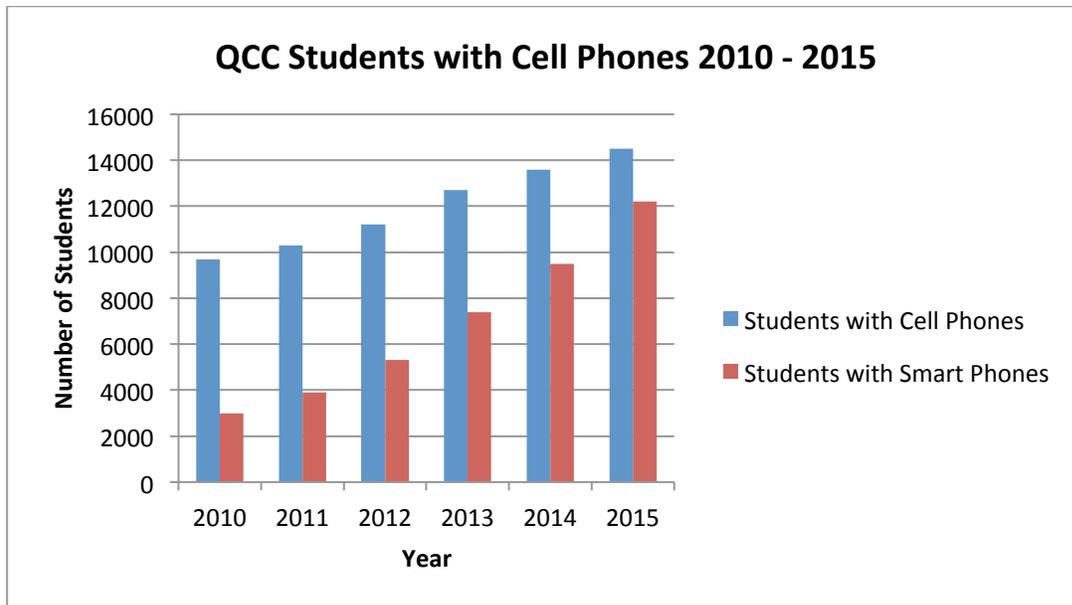
Looking at the graph, we see that, for each year from 2010-2015, the left bar represents **QCC Students with Cell Phones** and the right bar represents **QCC Students with Smart Phones**. It is clear that each data set is increasing over time. In 2010, 9700 students have cell phones and 3000 of them have cell phones that are smart phones. (Note that a smart phone is always a cell phone but not vice versa.) Five years later, in 2015, 14500 students have cell phones and 12200 of them have cell phones are smart phones. There is a 50% increase in cell phones and a 307% increase in smart phones.

But there's another story being told—the percentage of students with cell phones that are also smart phones is increasing at a rapid pace. In 2010, 3000 of 9700 students have cell phones that are also smart phones. That's 3000/9700=31%. In 2011, it's 3900/10300=38%. In 2012, it's 5300/11200=47%. In 2013, it's 7400/12700=58%. In 2014, it's 9500/13600=70%. Finally, in 2015, it's 12200/14500=84%. This tells us what we should already know—it's becoming increasingly important for all students (really everyone) to have cell phones that are smart phones. We expect that, in the future, both bars will increase (to get closer and closer to the entire QCC student population of about 16,000 students) and the height difference between the bars will continue to decrease (to get closer to 100% of all student's cell phones are also smart phones).

Question 3:   Give an example of two (or more) data sets that could be compared by using a multiple bar graph. What sort of trends would you expect to see from this multiple bar graph?

Consider the next graph:

Sales By Apex Inc. Salespeople for 2017
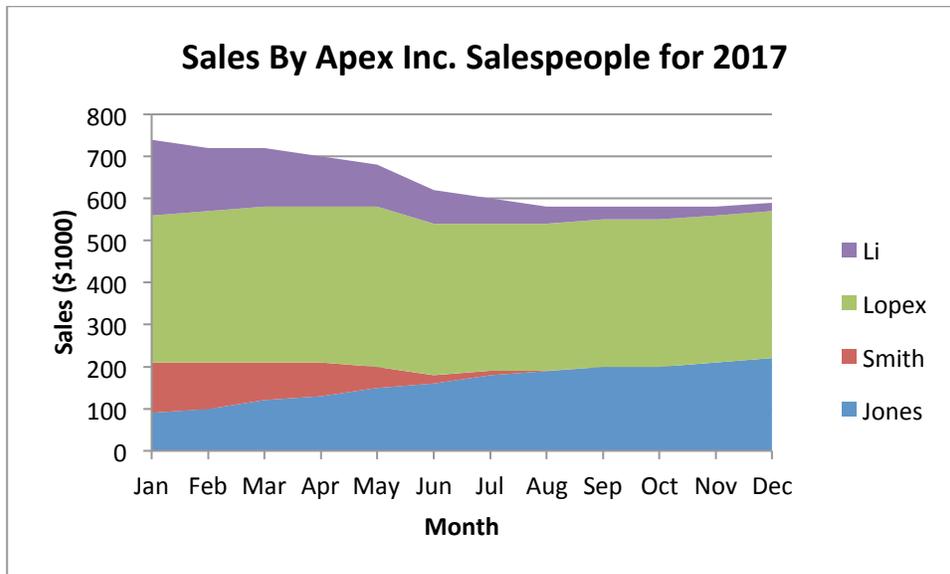
A **stack plot** displays the information of two or more data sets by stacking frequencies, one on top of the other. The graph represents the sales totals for Apex's Inc. sales team for the year 2015. It takes a few looks to figure it out; so let's look at the month of January:

In January, the sales for Jones (in $1000 of dollars) is 90 (that's the height, or thickness, of the lowest blue bar). The sales for Smith was 120. Why? The height, or thickness, of the two bars together is about 210. The lower bar represents Jones' sales of 90; the difference is 210-90=120 for Smith's sales. Lopez's January sales are about 350 (560 total minus 210 for the lower two bars). Finally, Li's sales are about 180 (740-560 for the lower three bars). We can check and see that 90+120+350+180=740.

The stack plot gives us the ability to see individual results (for each salesperson) as well as the total (sales for entire company) in situation where both pieces of information are relevant.

We can examine totals in other months:

In February, the sales totals are as follows:

Jones	100
Smith	110 (210-100)
Lopez	360 (570-210)
Li	150 (720-570)
Total	720

Question 4: What are the sales for the four salespeople in March?

In December, the sales totals are as follows:

Jones	220
Smith	0 (bar has disappeared)
Lopez	350 (570-220)
Li	20 (590-570)
Total	590

We are also interested in general trends for the stack plot as a whole:

We see that Jones' sales have increased throughout the year. This may suggest that she has been a great success for the company and continues to find new ways to find customers for the company. It seems the ongoing increase in sales suggests a bright future for this employee.

Smith's sales steadily decreased and disappeared by the middle of the year. We can only hope this employee is no longer with the company. (We can speculate that Jones took over his customers but we cannot know this for certain.)

Lopez's sales increased for part of the year, but declined to approximately the January amount by the end of the year. Is this cause for concern, or should we be happy with her reliability?

Li's sales decreased, becoming low but steady for the second half of the year. Perhaps his role in the company has changed (to a managerial position perhaps) and sales is not as important. Still, he remains a part of the sales team.

A cause for alarm is that monthly sales as a whole have declined for the company over the course of the year. Is this just a part of the industry that Apex is in?

Question 5:   Give an example of three (or more) data sets that could be compared by using a stack plot. What sort of trends would you expect to see from the stack plot?

Geographical data

Sometimes, we are interested in examining information for a geographical region. The region we are interested in becomes part of the graph.

Consider the following web site:

http://www.gasbuddy.com/gb_gastemperaturemap.aspx

This graph measures gasoline prices by county by state. If you right click on any county, you get the average price of a gallon of gasoline in that county. For example (in February 2018), the price of gas is around $2.92 in New York City and $3.59 in Los Angeles, CA but only $2.54 in Minnesota. Gasoline prices are also represented by colors.

We can also spot general trends. We see that the western states tend to have higher gasoline prices while some of the cheapest gas can be found in the center of the country. Gas prices seem to correspond with state boundaries as well (why do you think this is true?).

Question 6:   Right click around the country. What is the highest price of gas that you find? What is the lowest price of gas?

The only limit to this graph is that it suggests that gas prices are uniform, the same throughout the county. This may be somewhat true in some area, but there are often vast differences within a county (just drive around). We would probably agree that, in Queens, the price of gas is not $2.92 throughout the borough.

A **contour map** shows variation throughout a geographical region.

Check out:
http://weather.unisys.com/surface/sfc_con.php?image=te&inv=0&t=cur&expanddiv=hide_bar

This measures the temperature throughout the United States at a given time. Temperatures within 5 degree intervals (between 50° and 55°, between 55° and 60° etc.) are marked by different colors. The color boundaries represent the exact temperature. The line marked 40º is the boundary between temperatures in the low 40s and temperatures in the high 30s.

Looking at the graph, we see interesting variations in temperature throughout the country:

Obviously, there are warmer temperatures in the southern parts of the country. But we also notice that warmer temperatures seem to go up the coastline on both sides of the country. This is because coastal regions are often warmer than inland regions, because of the warmth received from the nearby water. For the same reason, we see colder temperatures dipping further south in the center of the country.

A contour map can be more useful than the previous map in some circumstances. We can appreciate the temperature variation within California (and other states as well), which usually has several temperature bars running through the state!

Question 7:   Identify two other states with at least three different temperature bars present in the state.

Graphic Distortion

The use of computers to make graphical displays can sometimes lead to distortions of the information being presented. These distortions may often be unnoticed, but their impact shouldn't be ignored.

Example 1: Consider the following display:
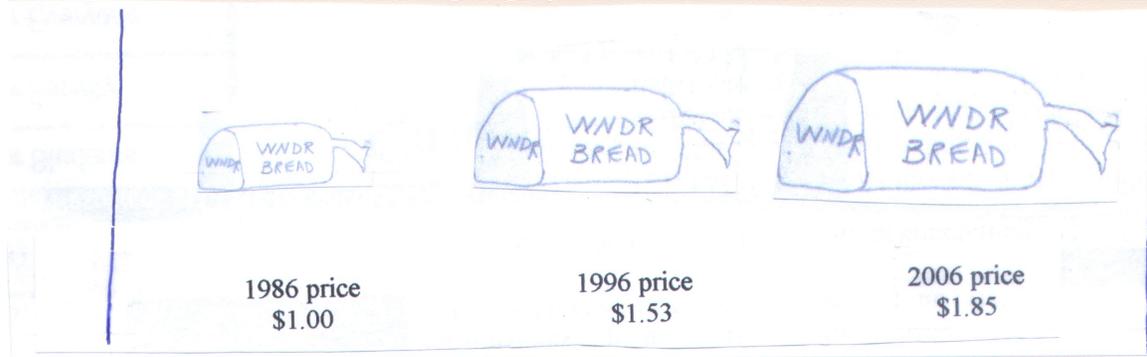
This graphic display shows the impact of inflation in the last 20 years. This is shown by increasing the area of each loaf of bread over time.

The 1996 bread is marked $1.53. What this means is that the same amount of bread cost that much more in 1996. In 2006, it now costs $1.85 to buy the same loaf.

While the numerical amounts are fairly presented, the sizes of the bills are not. The 1996 bread is larger than 50% of the area of the 1986 bread. The 2006 bread is more than double the area of the original loaf. The display is overemphasizing the impact of inflation.

Why is this done? Certainly, with computers, images can be increased or decreased to any percentage of the original size. The reason here is to further emphasize the point— inflation is on the rise! Just look at how big the bread is getting! While making the point, the information being presented is misleading.

Question 8:   What could be another example where a graphic display could be used to give a distorted message from what is actually true? (You don't have to draw anything, of course.)
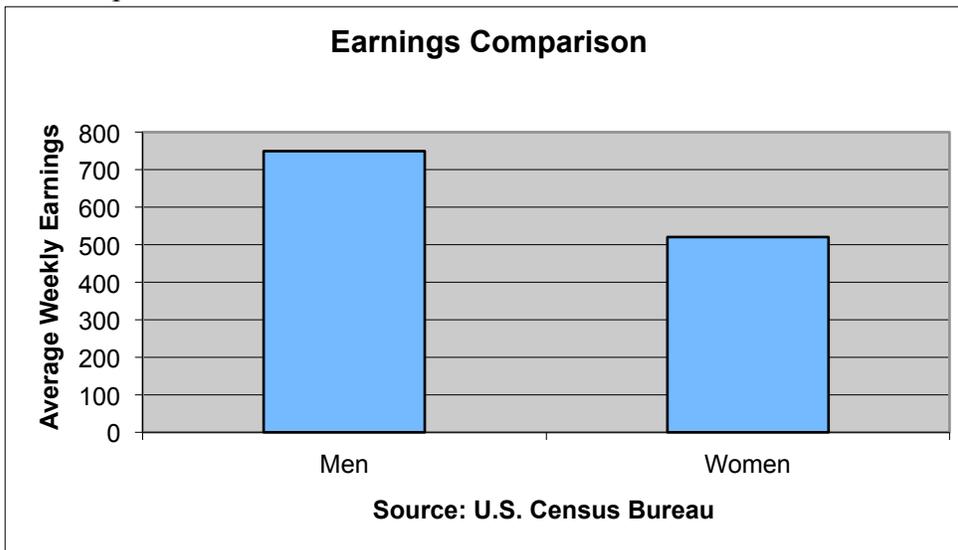
Example 2:   Consider the following graph:

Earnings Comparison

Source: U.S. Census Bureau
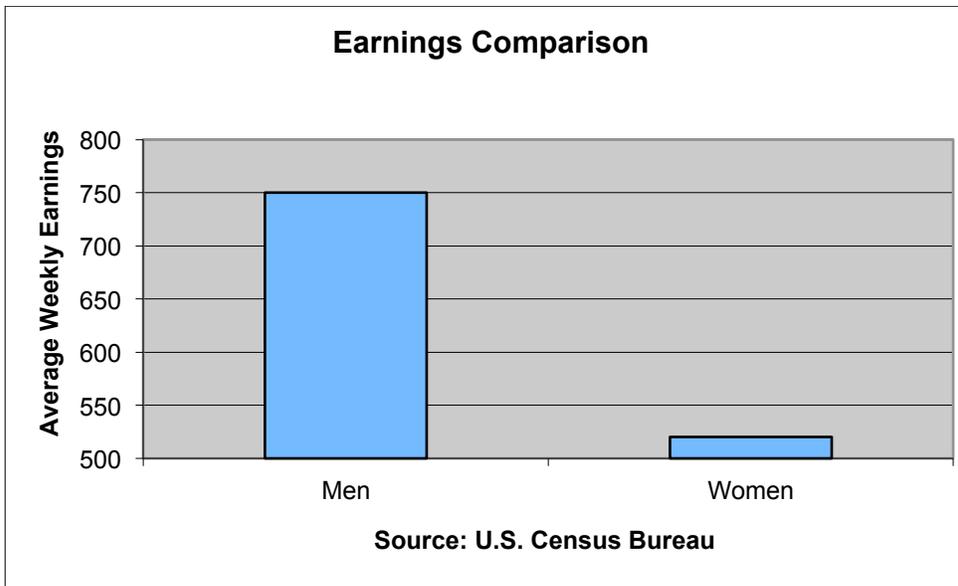
The graph displays the large difference between average weekly earnings for men and women. The height of the bar for Men is at $750, while the bar for Women is around $520.

But is the presentation fair? The height of the Men bar appears to be about 10 times the height of the Women bar. We can see that this is because the earnings axis starts at $500, not $0. The difference between earnings of men and women appears to be enormous. This is not really fair, although the numbers given are correct.

A fairer presentation would have looked like this:

Earnings Comparison

Source: U.S. Census Bureau

The difference in earnings is still readily apparent. While the difference is obviously an issue to be addressed, it should be presented correctly.

Example 3: Consider the following graph:

World Population Growth, Past and Future
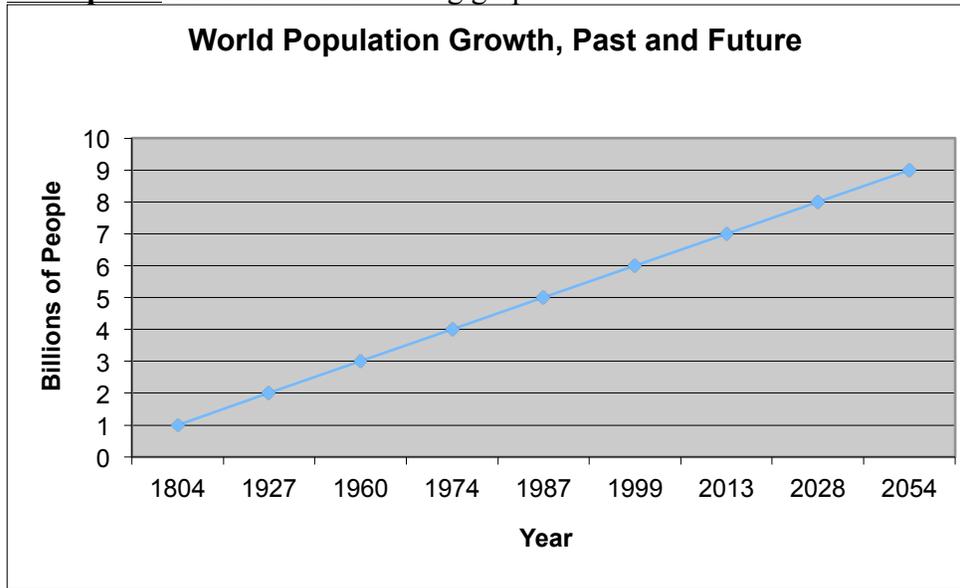
While it is interesting to view the changes in world population for the last 200 years, along with a prediction for the next 50, the information is again not presented fairly. We are given the impression that population has been increasing at a steady, linear (like a straight line) rate. This is not really the case.

It takes 123 years for the population to increase from 1 billion to 2 billion, but only 33 years to increase from 2 billion to 3 billion. It takes only 14 years for the population to increase from 3 billion to 4 billion; 13 years to increase from 4 billion to 5 billion and 12 years to increase from 5 billion to 6 billion. After 2013, future one billion increments occur in 15 and 26 years (but this is only a prediction—it may take less or more time for these increases).

A fairer presentation of the data could look like this:

World Population Growth, Past and Future
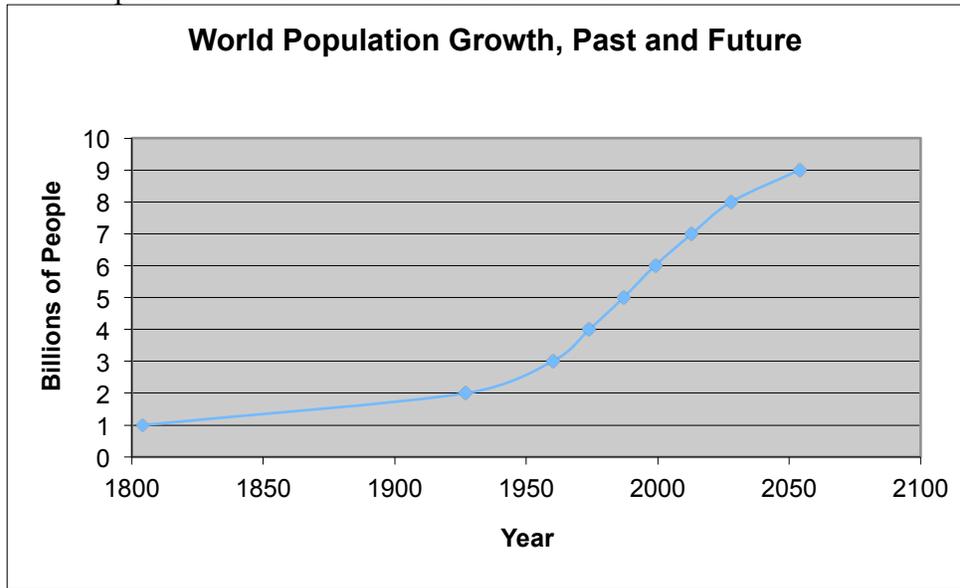
This shows the rapid increase in population growth, particularly in the last 50 years. Note how different the graph is now.

Question 9:   Describe the growth in the graph above. What does it look like? (Hint: Ever go skiing?) Can you think of anything else that might grow in this fashion?

We will see later in the class that this type of rapid growth is called exponential growth.

Graphic Confusion

The appeal to pack large amounts of information into a colorful, detailed graph is hard to resist.
Example 4: Consider the following:

Sales By Apex Inc. Salespeople for 2017
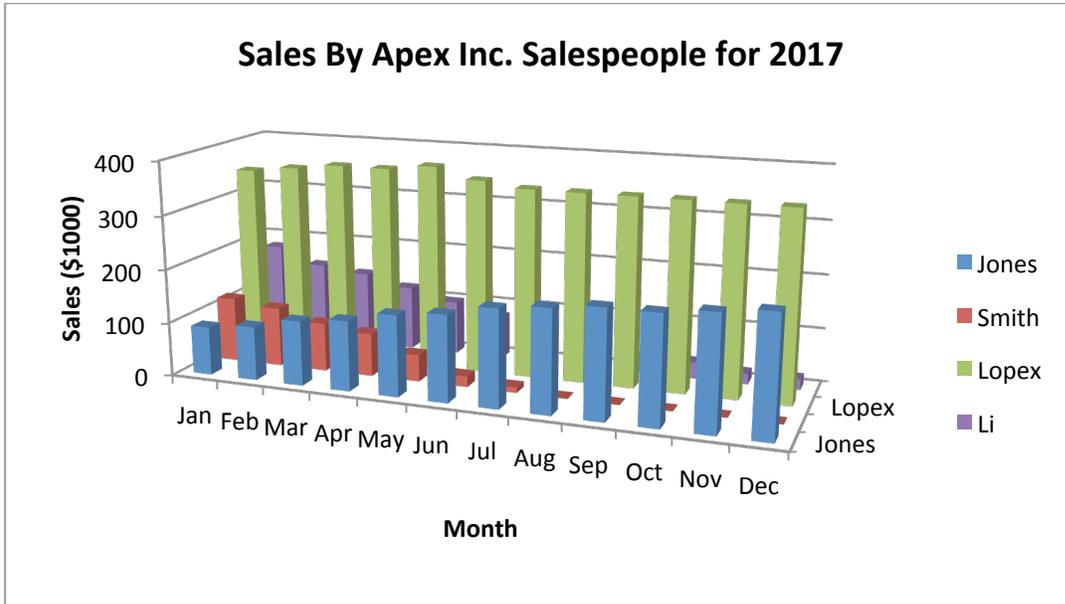
This information should look familiar. It the same information from the previous stack plot. Instead of stacking the results, a three dimensional image has been created with each salesperson being represented separately.

The same information is there, but is it better? We may wonder if a three-dimensional portrayal of this information is really necessary. The stack plot is easier to read and gives us total sales without additional effort.

Example 5: Consider the following graph:
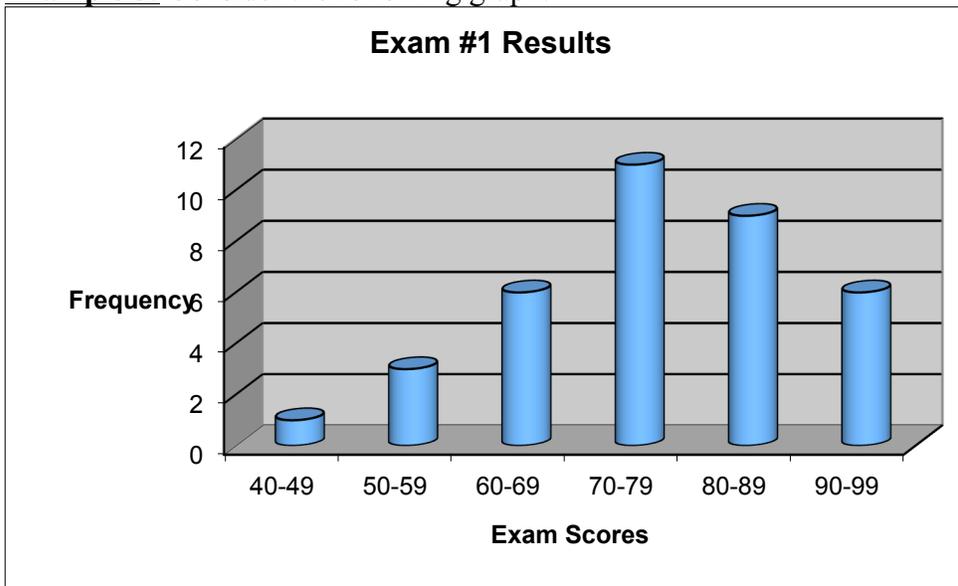
Take a quick look at the graph. Again, while the three dimensional look may give the bar graph a visual appeal, the data being presented has no need of it. Often, technology is used to give basic information a more exciting appearance. Is it really necessary though?

Homework:

For 1)-4), identify data sets that could be described graphically by a:
  1)   Multiple bar graph
  2)   Stack plot
  3)   Color coded county map of New York
  4)   Contour map

For 5)-8), do the following (for graphs found on the internet):
  a) Identify the kind of graph being used.
  b) Make two specific observations from the graph.
  c) Make two general observations of overall trends from the graph.

  5) http://www.weather.gov/forecasts/graphical/sectors/
  6) http://www.cnn.com/election/2016/results
  7) https://finance.yahoo.com/quote/AMZN?p=AMZN
  8) http://abcnews.go.com/US/shark-attacks-told-graphs/story?id=31779076

# Lab Assignment #3— Other Graphs
Due _____

If we create more complicated frequency tables, our graphs will become more complicated. Consider the following study:

IQ Score	Children	Teenagers	Adults
90-99	4	2	1
100-109	9	12	13
110-119	15	19	22
120-129	11	9	8
130-139	8	6	5
140-149	3	2	1
	50	50	50

(Source: IQ Exams taken on the same people at age 10, 15 and 25)

If there are multiple frequencies, we can create a multiple bar graph:
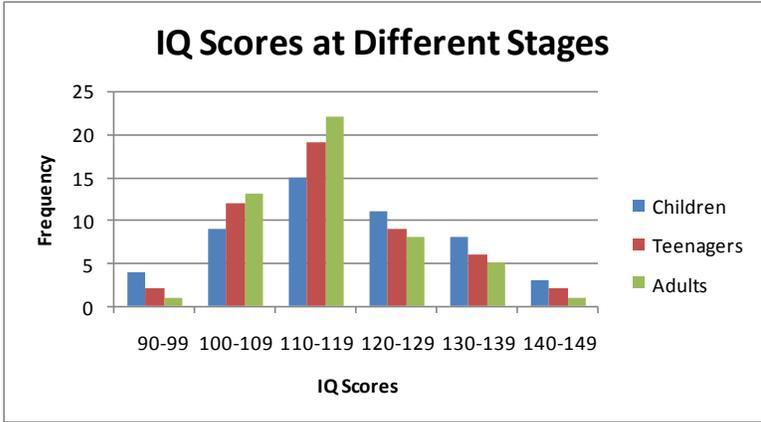
All you need to do is highlight the entire frequency table (not the totals) and *(by following the instructions of Lab #2)* Excel will do the rest. The color coding enables the reader to distinguish between the different sets.

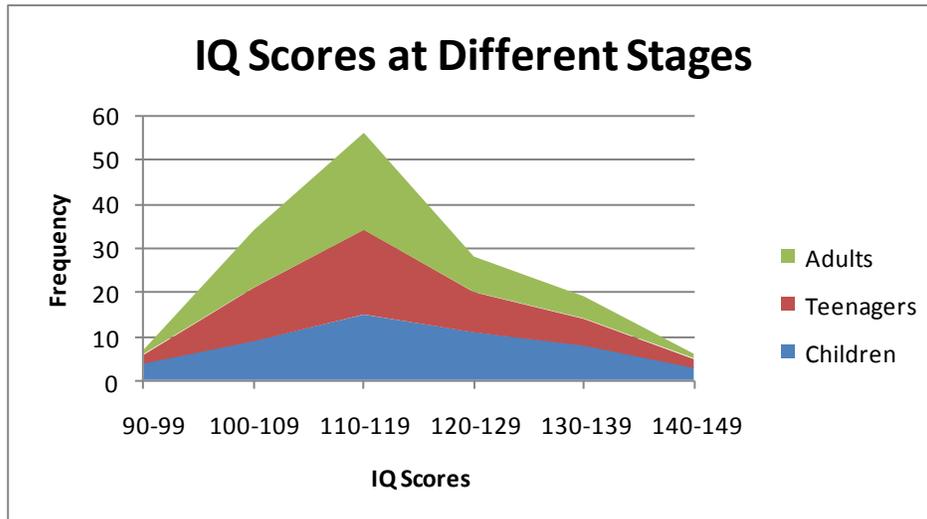A multiple line chart is also possible:
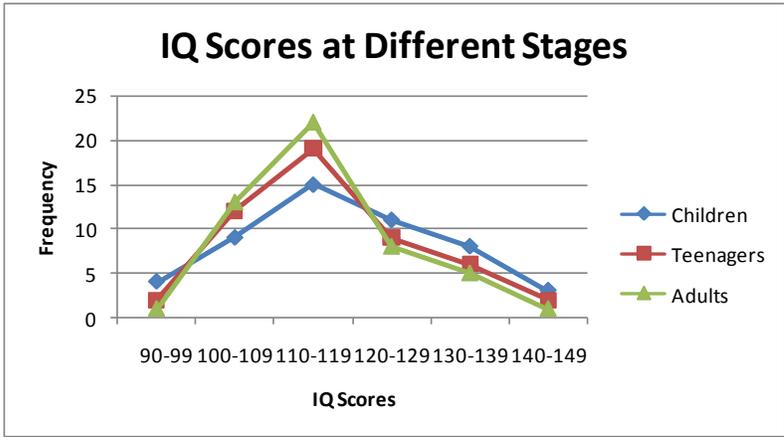
We can also stack the frequencies to make other graphs. For example:
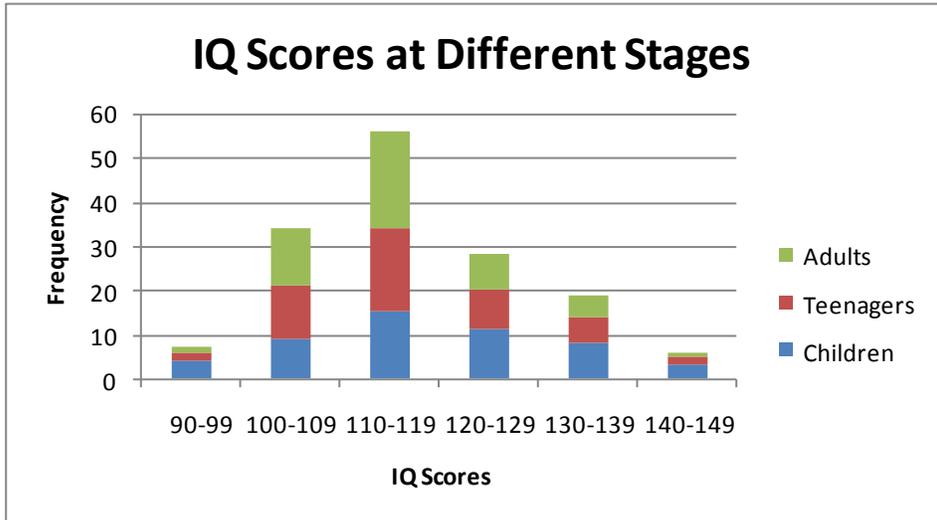
(This kind of graph is sometimes called a stacked bar graph.)
or:

(This kind of graph is called a stackplot.)

Create the four graphs and print them. Answer the following questions:

1) Describe an example of a data set in which using either the multiple bar graph or the multiple line chart to summarize the data would be more useful.
2) Describe an example of a data set in which using either the stacked bar graph or the stackplot to summarize the data would be more useful.