Olympic Data Analysis
Project 1

1. Relationship between age and sport and/or medals (Andrea Colin)
   a. In women's gymnastics for the years 1896 through 2016, a larger number of younger athletes have competed. In addition, a larger number of younger athletes had won a medal. However, both graphs showing the number of athletes and number of medalists for women's gymnastics for those years show that the data is skewed to the right. The standard deviation for athletes vs age is 8.6 and the standard deviation for medalists is 7.6 which isn't vastly different. Also, both data sets show the mean for ages to be 25 for athletes competing and 24 for medalists. This data suggests that age does not have a significant prediction for who will win a medal in women's gymnastics. The overlaying graph suggests that younger athletes tend to win more medals because there are just more younger athletes competing. If it did, the graph for medalists would have a skinnier distribution than for athletes and it would be much more skewed. This does not conclude that age does not have an affect on gymnastics performance. It only concludes, once qualified for the olympics, age isn't quite the determining factor most people believe it to be.

2. Gender representation at games from 1980s – 2016 (Elizabeth Christian)
   a. From 1980 on, both male and female athletes were increasing in representation at the games. Despite the increases in both groups, there's a notable gender disparity, with male athletes consistently outnumbering their female counterparts in each Olympic Games session of the years included in the data analysis. When examining the initial line graphs showing the time series trends of male and female representation, a more nuanced picture was shown once the games split to seasons in 1994. The Summer olympic games continued to have both groups with positive slopes and increasing over time at a similar rate. However, the Winter olympic games show a different story with the number of male athletes decreasing with a negative slope while the female athletes are increasing with a positive slope. Looking closer at the trend of male/female representation and the direction of the future of the games.

3. Major countries correlation of participation in sports (Nickolaus Johanson)
   a. To effectively draw a conclusion from this question we first had to filter out the data by winter and summer Olympic games. Once the data was broken out into the two categories I was able to determine the top 5 sports for each of the seasonal games, which created a baseline for what would be the most popular sports. I then took the data set and filtered it by the three most populated countries in each region of the world (Europe, Asia and North America) and broke down what the most popular sports were for each region. When viewing the data, we were able to see that there was strong correlation between the sports that were played in each region. While the top 5 sports in each region varied in their 1-5 ranking we were able to see a large level of popularity for the

same sports in each region. However, there were a few sports that showed some differences in each region such as rowing, fencing, ski jumping and short track speed skating. To conclude we can see the strong correlations between the same sports but looking deeper at each region we can see the level of popularity seems to change and shows us that each region is different in their own way.

4. How average height affects country's ability to medal for men's and women's basketball (Matthew Sturt)

   a. In answering this question, first the data had to be filtered down from the large data set of all athletes to only basketball, and then split mens and womens results from each other. Then, the data was separated between medalists and non medalists. The graphs seen are average heights of teams that did medal (average of all three teams) and those that did not. The key conclusions were as follows: In Women's Basketball, there never seemed to be any correlation between height and medal chances, as some years the average height of those who did not medal was even larger than those who medaled. When looking at the Men's side, it was a bit different, as in the most early recordings, the average height of the teams that did medal were significantly larger than those who did not. This disparity did somewhat diminish as time went on, as since the year 2000, the average heights have been similar to one another. This sort of 'leveling off' is most likely related to the overall popularity of basketball world wide.