**Project 1**

1. For this project, you are going to focus on analyzing time series. Here are the steps you will need to take to complete this project. *I suggest reading through these steps fully before beginning your project to make sure you understand everything that is required*:
   1. You need to find data sets with a time variable and at least one variable that you can analyze over time. For example, think of a data set with home prices. You could analyze the price of the homes over time if you are given the day, month, year, etc. However, it would not make as much sense to plot something such as number of bedrooms or bathrooms over time, as these don’t really make sense as a variable that might change over time. Think carefully when looking for data sets about something that you might be interested in analyzing over time. Each data set needs to include at least fifteen rows of data.
   2. You will need five separate data sets that match the criteria in part a: they need a time component and a variable that makes sense to analyze over time.
   3. You must provide me with a source for where you found each data set by providing the links in the project document. You can use websites such as statcrunch (click data sets on the home screen; if you find a data set that works, you can click ctrl+a to highlight it all, click ctrl+c to copy it and then paste it into an excel sheet using ctrl+v-you do NOT need to purchase anything to use the data sets on this website), Kaggle, or any other reliable source.
   4. Please put all of your data sets in Excel with each one on its own individual sheet.
   5. You need to plot each of the time series in a line plot.
   6. What type of time series is each of your plots? Explain why you feel they follow that type of time series. Is there a trend? Seasonality? Is it random? What may be influencing those factors?
   7. For each of the data sets, do not include the last three rows of data in the next steps. We will use these to compare to your predictions. Cut them out and paste them lower in the Excel sheet.
   8. For three of your time series, please use a three-period moving average and forecast the next three values in the time series. Compare this is the three values you did not include. Does the three-period moving average forecast these time series accurately?
   9. For the fourth and fifth time series, use simple linear regression with the time component as an input variable and forecast the next three values in the time series. Is the time component significant? Compare this to the three values you did not include. Does your prediction with regression forecast these time series accurately?
   10. Please create a word document with one section for each time series. Each section should include the link to the data set for that section, screenshots of the plots, screenshots of the forecasting, and all analysis done. Please also submit your Excel file.
   11. Any indication of sharing data sets or cheating will result in both students receiving a zero on this project. Please find your own data and do your own interpretations.