**Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Business Analytics: Take Home Final Exam**

Huggins Industries {HI}, an import/export company and purveyor of fine latex products (Seinfeld reference) based in Durango, CO, has a number of critical big data questions. Your job, as a high-priced consultant, is to answer them and explain your results. Good luck.

1. Demand for a product appears to be declining over the last 2 years (24 months). This data is online. Draw a scatter diagram to show this decline. Now, use Excel to fit a curve to the line with r2 ≥ 95% and list the equation and r2 value. HI plans to discontinue the product when demand falls below 100 units per month. Use the curve you developed to predict when this will happen. Approximately in which month will HI discontinue this product? {*Predictive and Prescriptive Analytics*} **Print the scatter diagram with the curve and equation and r2 on it and identify the month for discontinuation.**
2. a) While HI is headquartered here in Durango {81301}, the company has distributors all over the US. The CEO visits these locations regularly in his Gulfstream G650 and wants to know the distance between each pair of cities. Set up a 6x6 table and determine the pair-wise distances between Durango, CO; Missoula, MT; Olympia, WA; Claremont, CA; Ann Arbor, MI; and Honolulu, HI. {*Descriptive Analytics*} **Print out this table.** b) Occasionally, inventories become unbalanced at the six locations above. This is currently the case where Durango has 10000 units, Missoula has 2000, Olympia has 5000, Claremont has 12000, Ann Arbor has 7000 and Honolulu has none. Assume that you may transship products from one location to another, and the cost per product is $0.001 per mile. What is the cheapest way to transship products amongst the six locations to balance the inventories at all of them? {*Prescriptive Analytics*} **What is the minimum cost and describe how much should be shipped from where to where.**
3. HI has noticed that the longer potential customers browse on their website, the more likely they are to actually spend money. This data is online, with how many seconds the potential customer browsed and whether they made a purchase (1) or not (0). Use solver to develop a logistic curve that fits this data. You may or may not want to group the data first. {*Descriptive and Predictive Analytics*} **Graph and print a scatter diagram with the data points and the logistic curve on it. What are the best-fitting coefficients β0 and β1?**
4. Finally, HI would like to be able to predict how well new employees will perform over time. The data online has employees’ performance ratings one year after hire as well as four factors they were rated on when they were hired. First, use the standard Tukey’s rule and eliminate any outliers in the performance column. {*Descriptive Analytics*} **What is the range of typical values and are there any outliers?** After eliminating any outlier(s), run multiple regression on the data with performance as the y-variable. {*Descriptive Analytics*} **What is the multiple r value? Interpret the r2 value. According to the p-values, which factor should clearly be eliminated?** Eliminate the factor that is just noise and re-run the regression. {*Predictive Analytics*} **Rounding to two decimal places, what is the multiple regression equation that predicts performance based on the remaining relevant factors?** Use this equation to estimate performance ratings for the data then **calculate the MAD and MAPE.**