**Homework #10 (Due Nov 5 11:59 PM)**

IST 3420 - Fall 2017, Chen

**Name**: \_\_\_\_Adam Forestier\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Hypothesis Testing, Simple Regression (20 points)**

***Task 1: Hypothesis Testing***

1. Use the mtcars dataset. We focus on the following two variables:

* mpg: miles per gallon
* am: transmission (0 = automatic, 1 = manual)

1. We want to visualize the difference of mpg between two types of transmission. Draw a boxplot of mpg grouped by transmission type (two boxplots on the same figure, each representing one transmission type). Paste your plot in the following box. (2 points)

|  |
| --- |
|  |

1. What conclusion could you get from the above boxplot? Explain in the following box. (2 points)

|  |
| --- |
| Vehicles with manual transmission type typically has higher miles per gallon than vehicles with automatic transmission type. This can be determined because every part of the 5 number summary is higher for manual transmission type than automatic transmission type; the minimum, first quartile, median, third quartile, maximum is higher for manual transmission type. |

1. Now we want to use hypothesis testing method to test if manual transmission cars have better miles per gallon than automatic transmission cars. Specify your null and alternative hypothesis in the following box. (1 point)

|  |
| --- |
| Ho: mean(automatic transmission cars miles per gallon) <= mean(miles per gallon)  Ha: mean(manual transmission cars miles per gallon) > mean(miles per gallon) |

1. What is the appropriate hypothesis testing method? Why? Explain in the following box. (2 points)

|  |
| --- |
| Two sample t-testing because the two sample t-test is used when we want to compare two groups. We are comparing manual to automatic transmission. It will also be unpaired hypothesis testing because there are different amounts of observations and the data is not paired. |

1. Use the appropriate method to test the hypothesis. Paste your hypothesis result and explain the result in the following box. (3 points)

|  |
| --- |
| Hypothesis Result:    Interpretation:  Because the p value is 0.001 < P < 0.01, there is Strong evidence against the null hypothesis in favor of the alternative. We can conclude that the average mpg of is higher for manual transmission than for automatic transmission cars. |

***Task 2: Simple Regression***

1. Use the mtcars dataset. We focus on the following two variables:

* mpg: miles per gallon
* wt: weight (lb/1000)

1. Regress mpg on wt, i.e., use mpg as dependent variable, wt as independent variable. Use linear regression model to fit the data. (2 points)
2. Use stargazer() function (in stargazer package) to beautify regression output, paste your regression result in the following box. (1 point)

|  |
| --- |
|  |

1. What is the relationship between wt and mpg? Explain below. (2 points)

|  |
| --- |
| mpg and wt are negatively correlated. The higher the weight, the less mpg the car will likely get. The effect of height on shoe size is not significant (p-value > 0.05). If weight increases by one thousand pounds, mpg would decrease by -5.3445 per unit on average. |

1. What is the regression model? If a car weighs as 5500 pounds, what is the predicted mpg of this car? Explain below. (2 points).

|  |
| --- |
| Regression formula/model: MPG = 37.2851 – 5.3445(wt)  37.2851 – 5.3445(5500/1000) = 7.89035 miles per gallon. |

1. What is the coefficient of determination? Interpret its meaning below. (2 points)

|  |
| --- |
| The coefficient of determination is 0.7528. The coefficient of determination, denoted by R2, is the proportion of variance that can be depicted in the dependent variable from the independent variable. This number can range from 0-1, with 0 making it impossible to predict and 1 making it predictable without error. With a score of .7528, it means that approximately 75 percent of the variance in Y is predictable. |

***Homework Submission***

1. Upload this document with your answers to “Homework 10” on Canvas.
2. Upload your R Markdown file (for both task 1 and task 2) to “Homework 10” on Canvas. (1 point)