Stat 021 Homework 6

Suzanne Thornton

Due: Friday, Nov. 8, 12:00pm

Instructions: A **pdf** version of your homework must be submitted to Gradescope by **noon** on the due date. The course passcode is **MPKJ4Z**. If you are having trouble getting your *.Rmd* file to compile, you need to get help with this **before** the due date.

You are allowed to hand in **only one** late homework assignment throughout the semester. If you need to hand in this particular assignment late, you must let me know via email by noon on the due date.

You are encouraged to study with your peers to help complete the homework assignments but no copying is allowed. If I see that two or more homework assignments are copied, all students involved will receive a grade of 0 on that assignment and will forfeit (perhaps retroactively) the opportunity to hand in a late homework.

Q0) To help you with some programming tools you will need on your final project, please complete the R *swirl* tutorial on writing functions. You can access this tutorial by typing the following commands in the R console window:

```
install.packages("swirl")
library("swirl")
swirl()
```

Then, the tutuorial will ask what to call you so enter your name and next type

```
main()
```

Make the following sequence of selections:

- 1: R Programming: The basics of programming in R
- 1: R Programming
- 9: Functions

Please complete this tutorial up until you get to the part about binary operators (this occurs at about 94% of the way through). Although you don't get points for this problem, it will dramatically help you with your final project and the material in this tutorial is fair game for future homework assignments.

Q1) Read the data uploaded to Moodle called "mileage.csv". This data describles the gasoline mileage performance for 32 automobiles. Use this data to answer the following questions. Before fitting any models make sure the data is being correctly read into R.

a) Build a linear regression model relating gasoline mileage, y to engine displacement x_1 and the type of transmission, x_2. (Note that transmission type is a binary categorical variable.) Does the type of transmission significantly affect the mileage performance? Justify your answer. (4 points)

b) Modify the model developed in part a to include an interaction between engine displacement and the type of transmission. What is the average effect on gasoline mileage when the engine is automatic? What is the average effect on gasoline mileage when the engine is manual? (4 points)

c) Build a linear regression model relating gasoline mileage, y, to vehicle weight x_3 and the type of transmission x_2 . Does the type of transmission significantly affect the mileage performance? Justify your answer. (4 points)

d) Modify the model developed in part c to include an interaction between vehicle weight and the type of transmission. What is the average effect on gasoline mileage when the transmission is automatic? What is the average effect on gasoline mileage when the transmission is manual? (4 points)

e) Based off of the results for parts (a)-(d), what terms do you think should be included in the final regression model and why? (4 points)