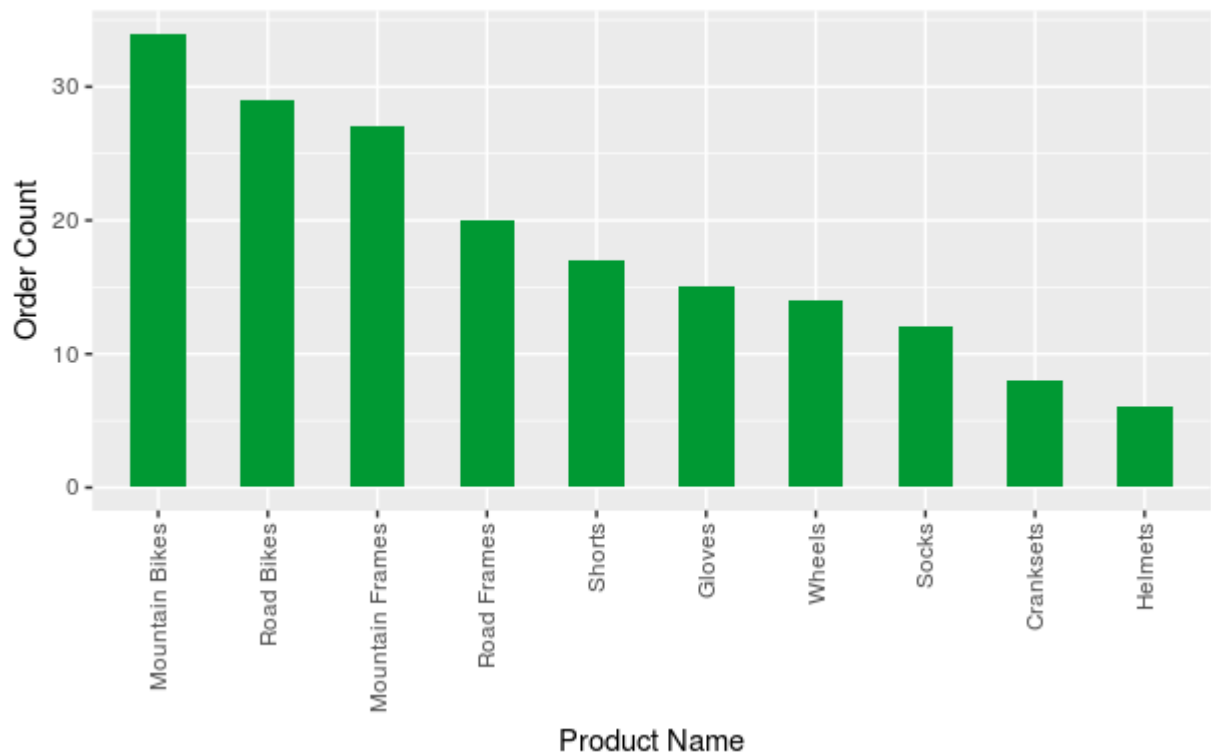


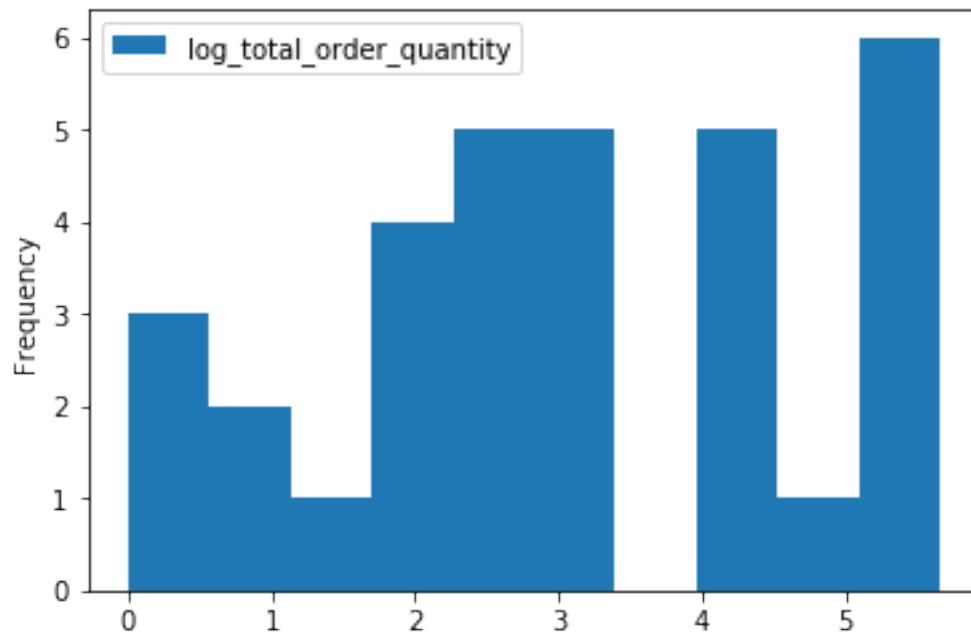
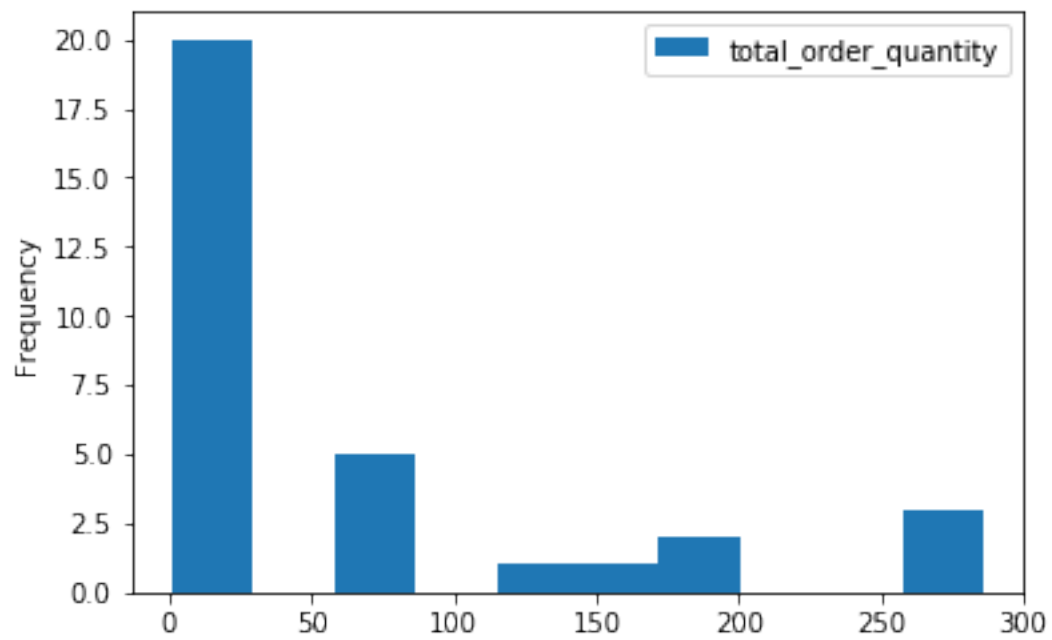
Part 1 Problem 2

state	non_bike_group_sales	bike_group_sales
California	\$257,985.69	\$88,531.92
New Mexico		\$15,275.20
Nevada		\$7,330.90
Utah		\$665.43
Colorado	\$14,017.91	

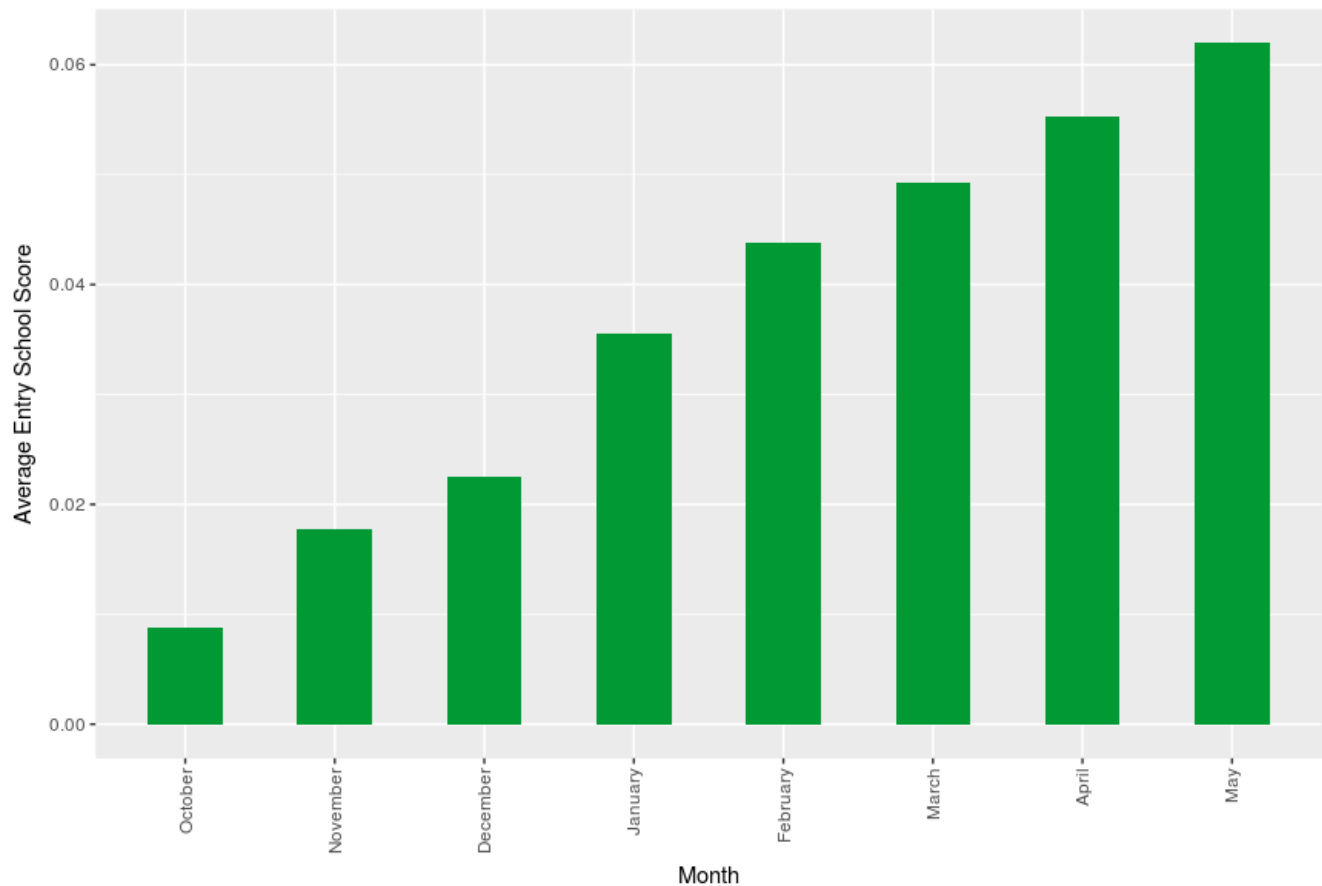
Part 1 Problem 3



Part 1 Problem 4a & 4b



## Part 2 Problem 2



## Part 3

I chose the *predict\_class\_assignment* dataset; I made three models for this instead of two (mostly because I had the code already written). I chose to train a decision tree, random forest, and a support vector machine. Of the three, my decision tree was the most accurate with ~75%. I split my data into test and train, giving 75% of the data to the training set. I then made predictions with each of my three models with the test set and compared my predictions to the actuals. I also outputted confusion matrices that detail where my predictions are failing. These models were really simple, as per the instructions but given the data I'm not sure that I could have done much more with it apart from building neural network. The main hardship with improving any one of these machine learning models is that I have no way to gain any domain knowledge about the data set (all the columns are the same) so feature engineering would be a random shot in the dark.