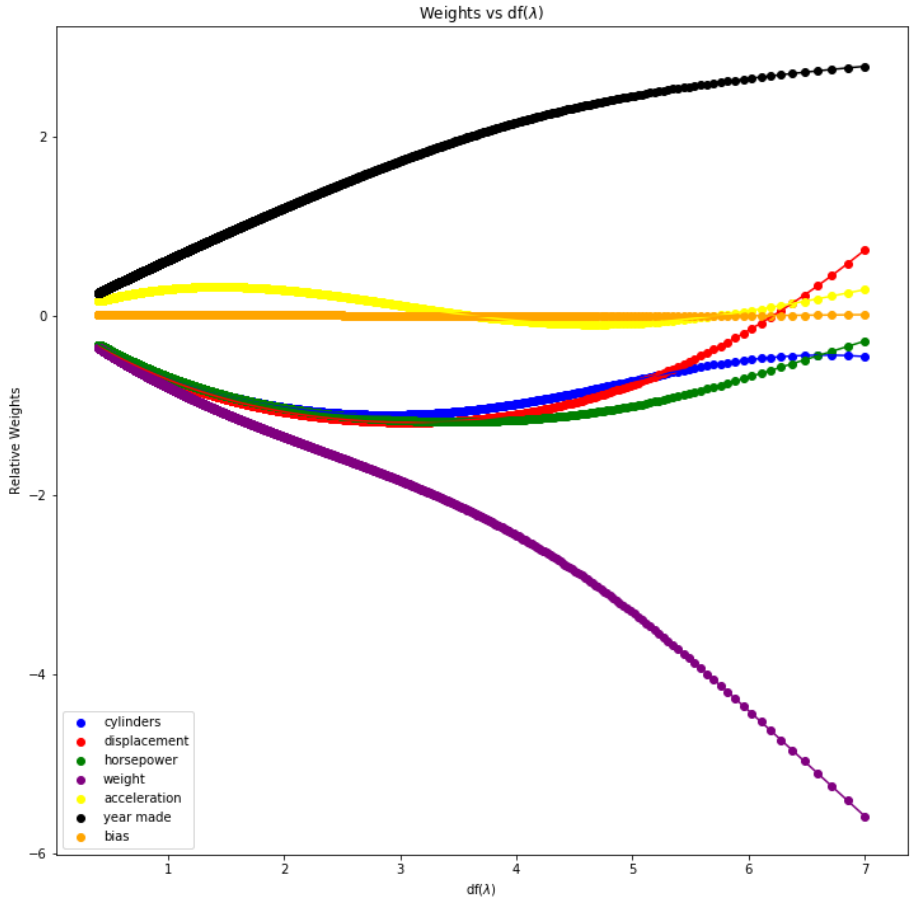
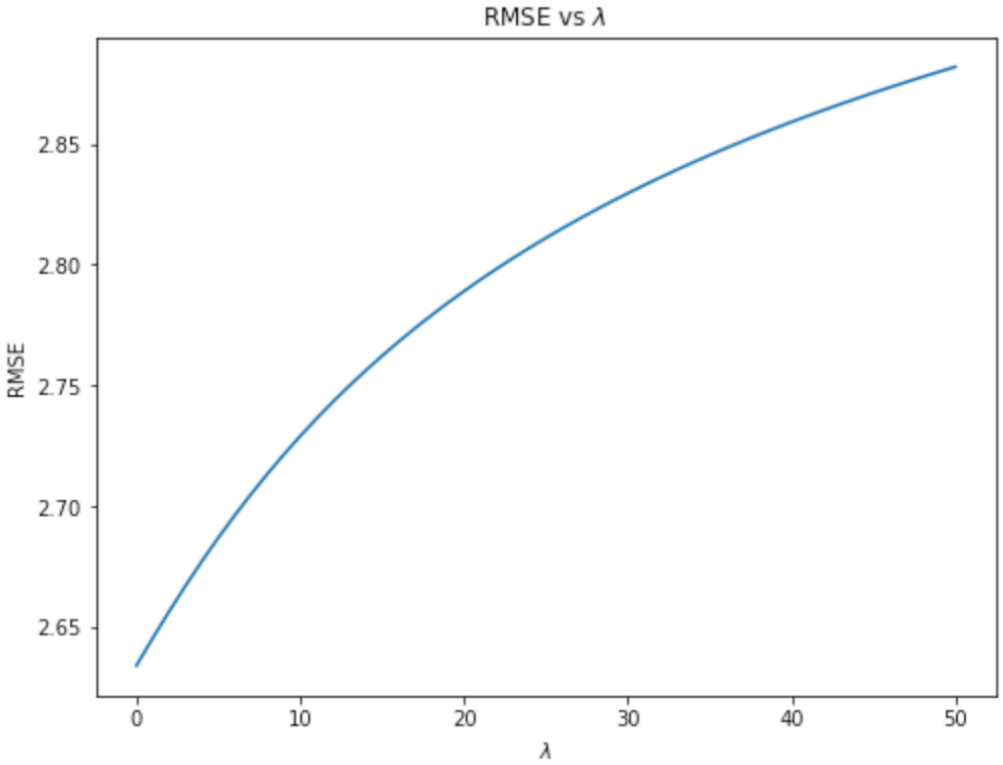
Question 3
Part 1.
A.



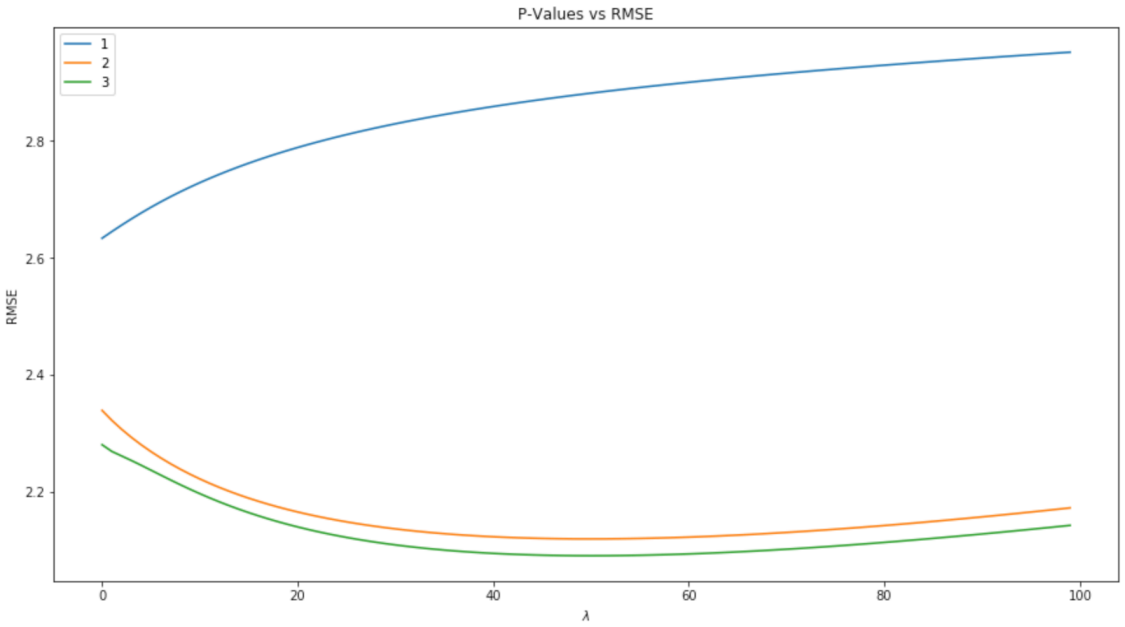
B. The two features that stand out are the weight and the year made features. The weight of the year made feature is highly positive, which means that increasing this feature would increase the output. The weight of the "weight" feature is strongly negative, which means that by increasing the weight, we will be reducing the output. In terms of relative weights, these features are much higher than the other features and have more significance on the output than the other features. The weights of these features also remain positive (for year made) and negative (for weight) even as degrees of freedom of lambda is reduced.

C.



The figure shows that the RMSE of the predictions on the test set is monotonically increasing with increasing lambda. Thus, for this particular problem, the choice of lambda=0 works best. This means that linear regression for this problem outperforms ridge regression.
\

Part 2



Even though RMSE of p=3 is lower than RMSE of p=2, the computation involved with p=3 is higher. However, I will still choose p=3, if computation power is not an issue.

A lambda value of around 50 in this case gives least RMSE on the test data if p=3. Thus, if polynomial features are used, I will choose ridge regression with p=3