ISYE 7406 - Homework 1 Sample

Student Name

Date

Introduction

In this area of the report, consider what kind of questions you're trying to answer with the analysis, or what kind of questions you can answer. What is it that you're doing in this analysis, its importance, and what does it contribute to?

Exploratory Data Analysis (EDA)

Use this section to help the reader better understand the data you're working with here. What do some of the columns mean, and what does some of the distribution look like? Use this area to communicate your understanding of the data and important qualitative and quantitative characteristics that might be important to share to help the reader better understand. Use visuals here to share characteristics that you observe, like multicollinearity, missing data, and outliers

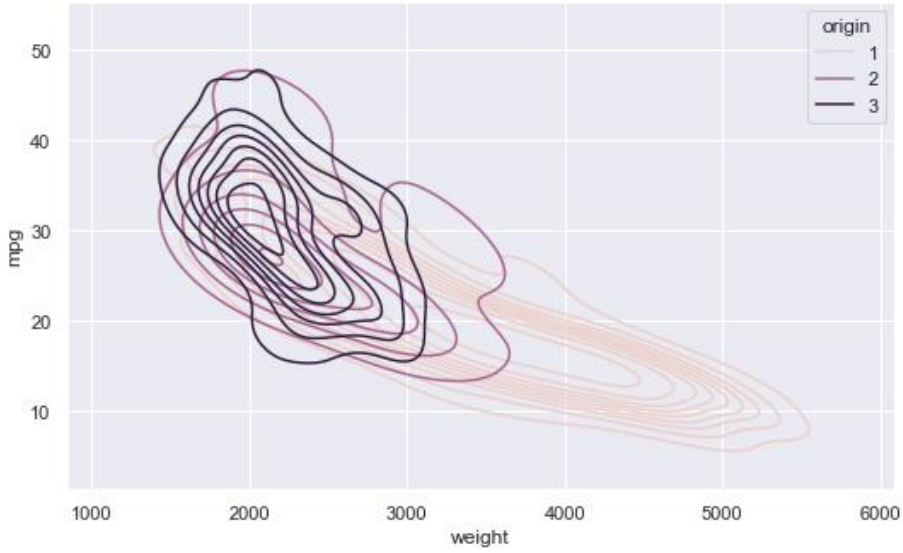

Figure 1: Label your plots to easily reference them and describe briefly what we see here. I.e. Kernel Density Plot of a car's MPG per weight and country of origin)

Methodology

A key aspect of the methodology is that someone should be able to replicate what you did in your analysis. Describe your analysis process here. What models were chosen, why were they chosen, and did you tune any parameters for the models? Why or why not? Did you cross-validate, and how did you evaluate to determine significance between results or models? Do not put your results in your methods section.

Do not make a common mistake of copying and pasting code and console outputs here. Share specific sections and very concise information as needed. It would not be appropriate to share a summary output only to discuss the r-squared.

Results

Put your results here! Use visuals and/or tables to professionally demonstrate your results. If you did not scale your visuals well when the results are really close, consider using a table for others to easily identify results.

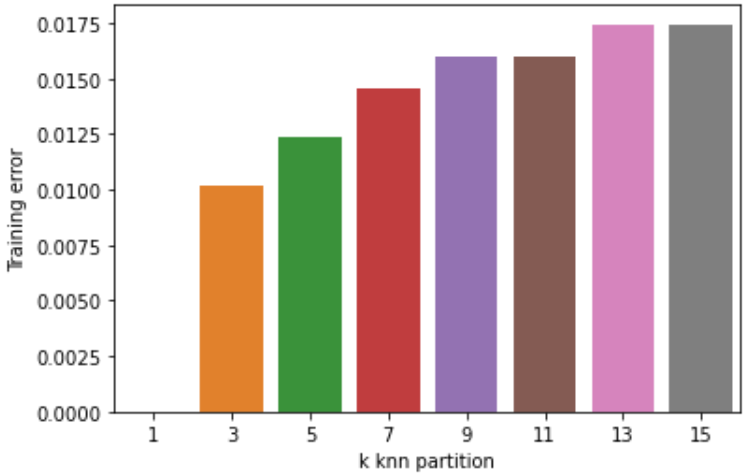

Figure 2: Always use a clear and concise figure label, and try labeling your figures: “Testing Error per K on KNN Partition (Evaluated with Training Data)”

Common mistakes include copying and pasting console print outs of code in your body report. DO NOT DO THIS.

Findings/Conclusion

This is where you put everything together. Looping back to your EDA to discuss surprises or patterns that you derived from start to beginning. This is the area to justify and discuss your results in detail. Don't tell us "X and Y model were the best", tell us how they were the best – was this reasoning derived from a qualitative area from your EDA, or perhaps your evaluation and modeling technique and results? What are some real world implications?