CSC 8016 Machine Learning Data Science

Assignment #2 Problem set#2

As continue working with dataset related with used cars in one of the car dealerships in the city, you have the 5 cars which recorded the high demand in the shop, you will still be working with the dataset of used cars in this assignment too, but the difference now, the dataset of 160 used cars is represent the cars which are available in the car dealership's inventory.

The excel sheet DealerShipUsedCars.csv, which attached with the assignment#2 problem set#2 will be used.

To start working with this assignment:

Step1: Working with dataset:

- Read the PDF file: Managing Data with R carefully, **please skip any visualization section**.
  (You may be read some sections and not using now, but we will use in future assignments, so it's good to keep this file)
- Download the excel file with name DealerShipUsedCars.csv, which attached with the assignment#2 problem set#2, and save it in your R workplace.
- Explore the data set.
- Load the dataset into R studio.
- Create a frame for the dataset.
- Save the file as R file under your name_assignment2.
- Submit your R file only.

Step2: Answer the following Question:

1. Show all numerical variable in the dataset.
2. Show the price range of the cars and the difference.
3. Using **one command** to show the number of cars with the same make.
4. Using the prop.table() function to show the proportion of the each car's make of the whole, and give your explanation.
5. Show which car's color is the highest present in an inventory.
6. Using CrossTable function to show how many cars from each Make have the color of (Red, Yellow, Black, Gray, White, Silver).
7. Do you think there is a relation between the price and the year of the cars? Show that with the results with scatterplot and give your analysis.
8. Save all the above step in the R file and upload to Black board.

CSC 8016 Machine Learning Data Science

Assignment #2 Problem set#2

Note: To keep your grade, submit your assignment on time.