Lab 1 Complete

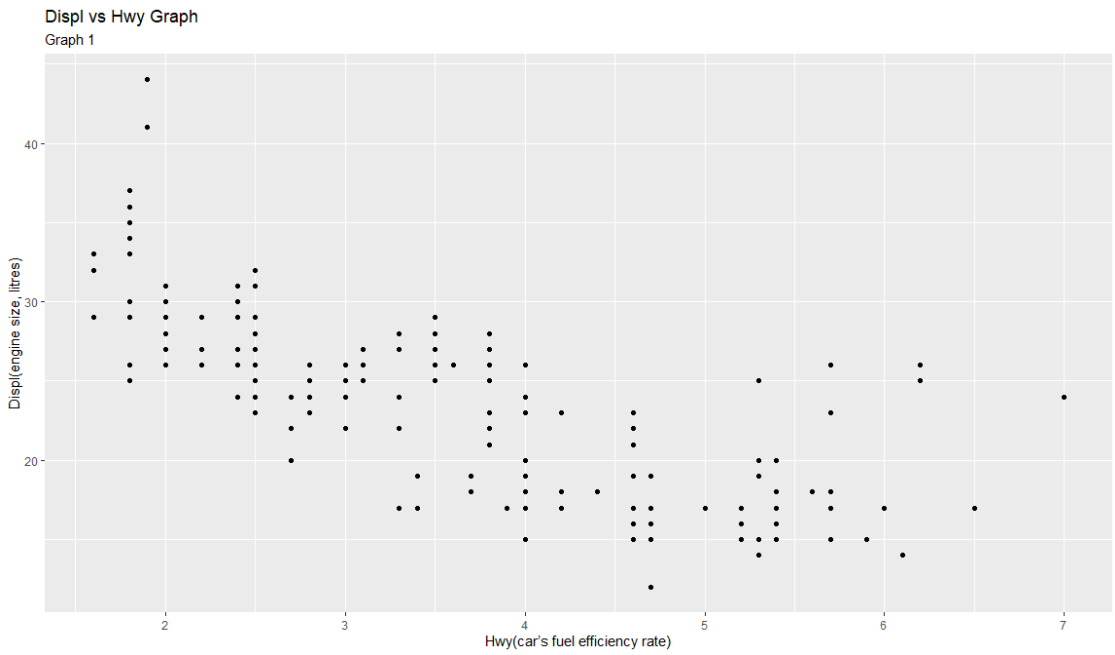
Team Awesome - Sayem Lincoln, Joshua Schwimmer, John Townshend.

1/14/2020

Contains Exercise 1a, Exercise 1b, and Exercise 2.

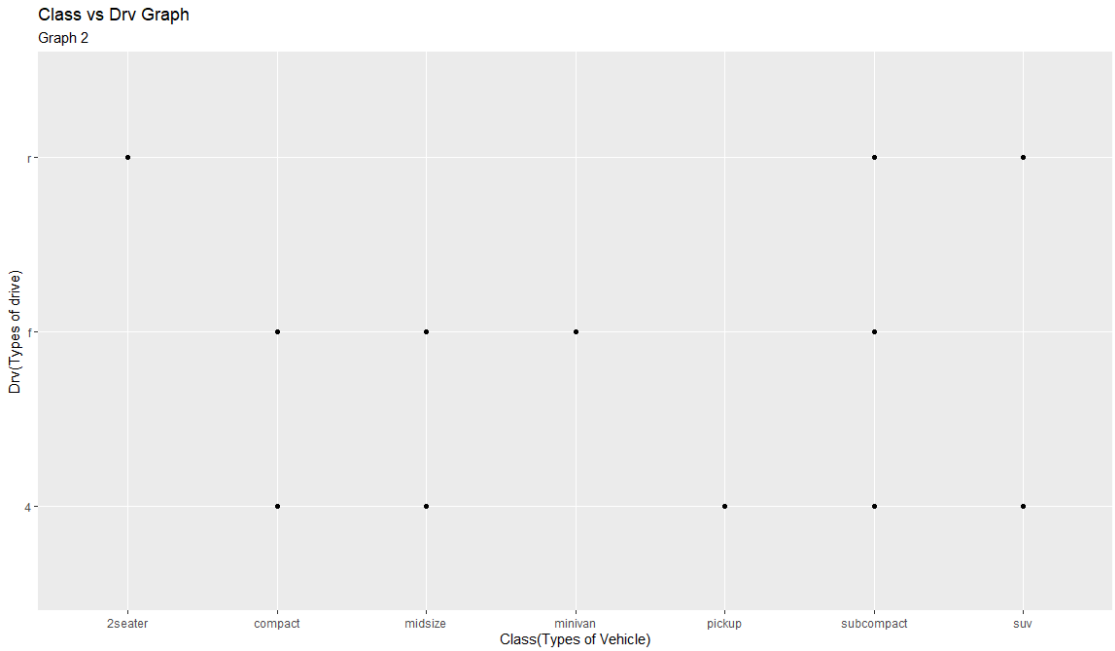
Exercise 1a

```
plot1
```



Does it capture the intuitive relationship you expected? Ans - Yes, it does capture the ituitive relationship, as displ is a car's engine size, in litres and hwy is a car's fuel efficiency on the highway, in miles per gallon (mpg). A car with a low fuel efficiency consumes more fuel than a car with a high fuel efficiency when the cars travel the same distance. So, a car with a smaller displ should have a smaller hwy value - which the Displ vs Hwy Graph (Graph 1) shows.

```
plot2
```

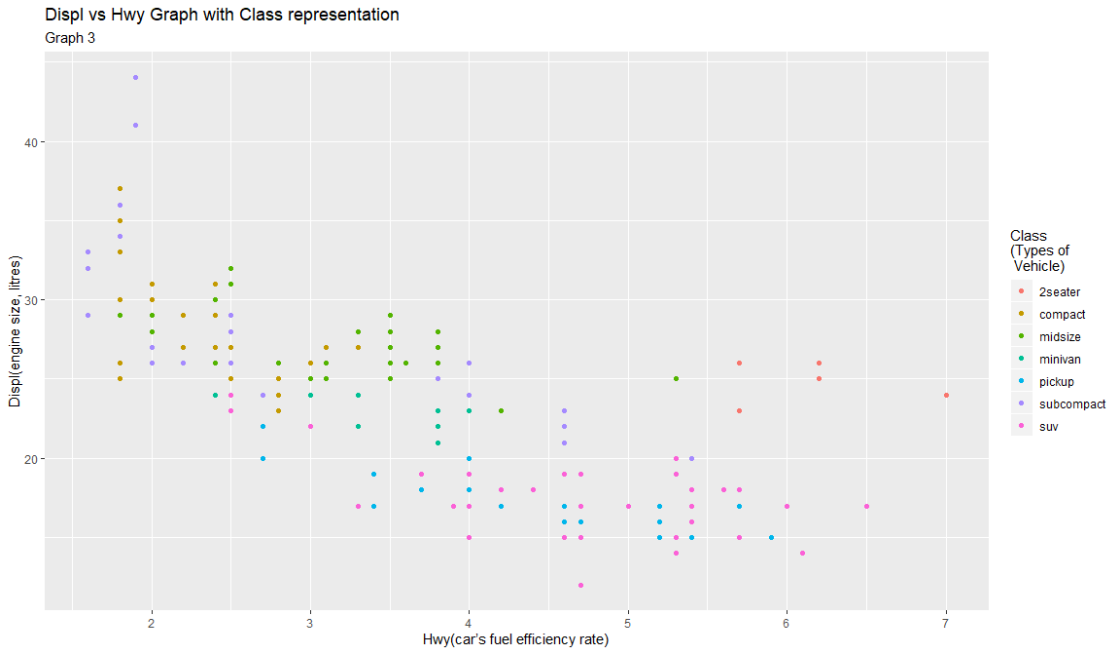


What happens if you make a scatterplot of class vs drv? Ans- For the drv data column- f = front-wheel drive, r = rear wheel drive, 4 = 4wd. And the class data column gives us the type of vehicles that are in the dataset. So when Drv vs Class Graph (Graph 2) is plotted we can see the type of vehicles plotted against the theree different type of drives (front wheel drive, rear wheel drive, 4 wheel drive).

Why is the plot not useful? Ans- The plot Drv vs Class does not correlte to the plot of Displ vs Hwy, as both of the graphs look at different types of data so a correlation between these different forms of data can exist but certainly doesn't have to, so a causcation correlation hypothesis between the two graphs cannot be stated with certainity, making the plot not useful.

Exercise 1b

```
plot3
```

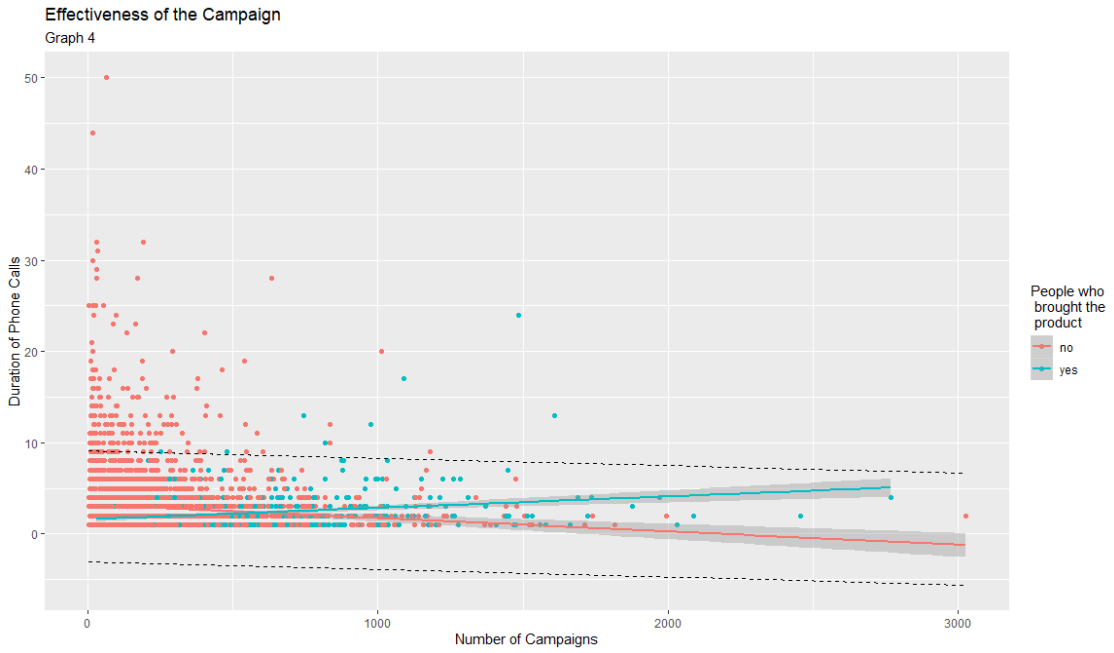


We can conclude that most of the pickup, subcompact, and suv type cars have bigger engine size in liters and they are not that efficient in saving fuel, but there are some subcompoact and suv cars that have samller engine size and is efficient at saving fuel which implies that these cars could be hybrid whereas the other cars on the bottom righthand side of Graph 3 are fuel burning cars with bigger engine size and with lower efficiency rate.

And the rest of the type of cars are scattered in between and to the upper-left side of the graph (Graph 3), which shows that these cars have good to moderate efficieny in contrast to small and moderate engine size. So, it can be inferred that cars with smaller engines are more fuel efficicient, whereas cars with bigger engines consumes more oil compared to smaller ones, and that Displ and Hwy are proportional to eah other, where a corrleation and causation analysis can be drawn up between these two types of data.

Exercise 2

```
plot4
```



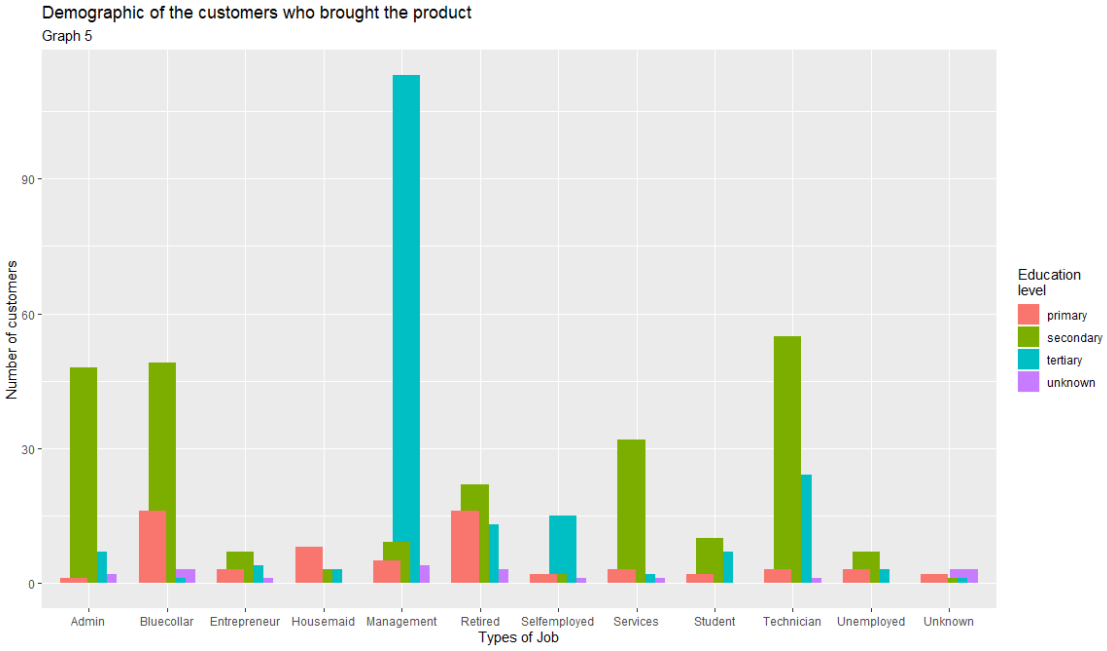
The above graph (Graph 4) shows the effectiveness of the campaign. The graph illustrates duration of the phone calls against the number of campaigns and shows us whether that campaign was successsful or not - this is shown by the number of people who bought the product and the people who did not (color coding represented on the left side of Graph 4).

And then we constructed a linear model from the given data, the linear model is showcased by the slightly curved lines which is our linear model's regression line. The grey area known as the confidence bands that covers both lines are the confidence intervals for each regrression line.

Then the dashed black lines known as the prediction bands represents the 95% prediction intervals, again for both types of data, so the prediction interval here gives us a prediction with a 95% probability that a future observation will be contained within the prediction interval. Conversely, there is also a 5% probability that the next observation will not be contained within the interval.

So, for a future capaign where the bank might offer a similar product, the bank should target and focus on conducting the campaign within the 95% prediction bands, as being within this range offers the bank the best probable outcome. By doing this, the bank can maximixe its utilities and the campaign can be more productive.

```
plot5
```



The above graph (Graph 5) illustrates the demographic of the customers who brought the product from this campaign, we displayed the customer's information and background by their types of job and by the level of education they had. So, for a future campaign the bank should target certain types of people who are buyers of this type of product so that they can again sell it to these type of customers and get the highest amount of sells and maximize their productivity.