Assignment 2

The type of the dataset is a table, in which the items are list of different cars and the attributes are their name, price, mileage (mileage per gallon - MPG), weight and repair condition.

The attributes' type:
- Name: categorical
- Price: quantitative
- Mileage (MPG): quantitative
- Weight: quantitative
- Repair Condition: ordinal

The dataset is formatted as a .csv file. Data processing and all visualizations are done with python and plotly.

Task 1 - Locate the car with the highest MPG
Task 1 Abstraction: Browse Extremes
Task 1 Visualization: a bar graph that shows cars' names at x-axis and MPG at y-axis. In this graph, we use lines as the mark and the length of the lines is the channel of the mileage variable. We can easily see that Honda Civic Coupe has the highest MPG of 36.

Car Mileage
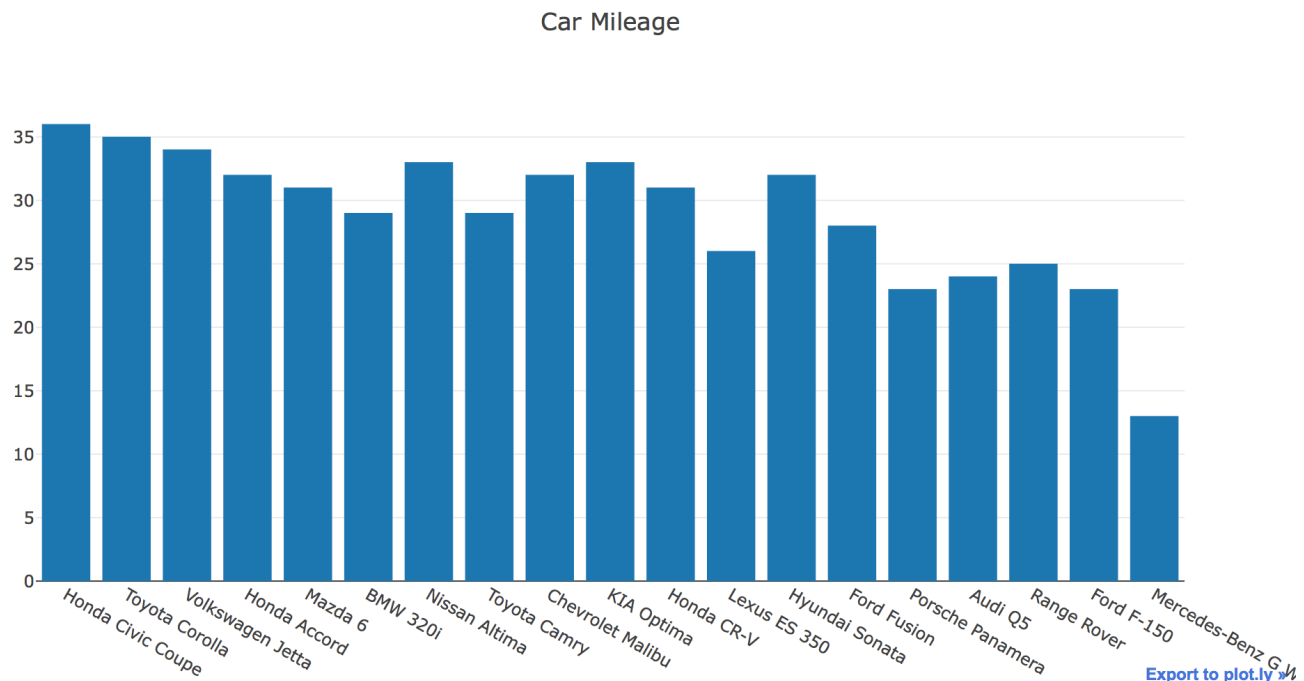
## Task 2 - Find out if there's a trend/correlation between the car's weight and its MPG
Task 2 Abstraction: Explore Trend / Explore Correlation
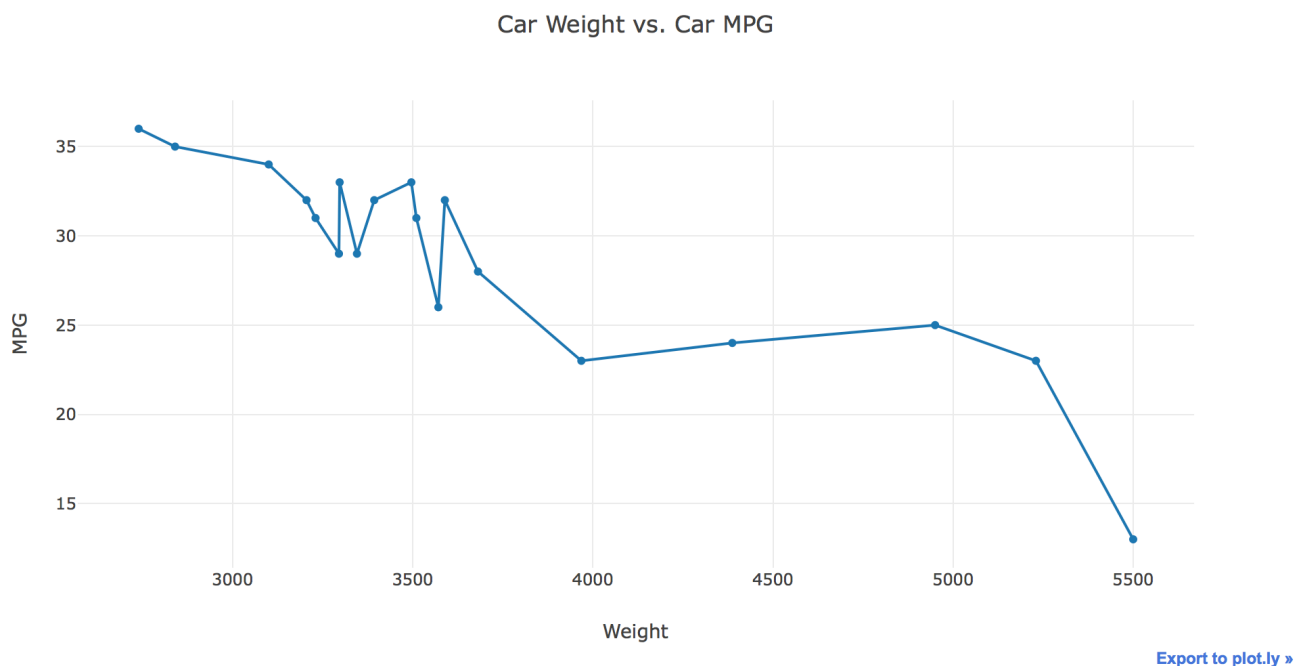Task 2 Visualization:

There are 2 graphs (both line graphs) to achieve this task:
- In the first graph, we can see a trend that the MPG tends to decrease when the car weight increases. For both the weight and mileage variable, we use points as the mark, while the

channels are horizontal position. Points are used as it will help us clearly visualize the values on 2 different y-axis.
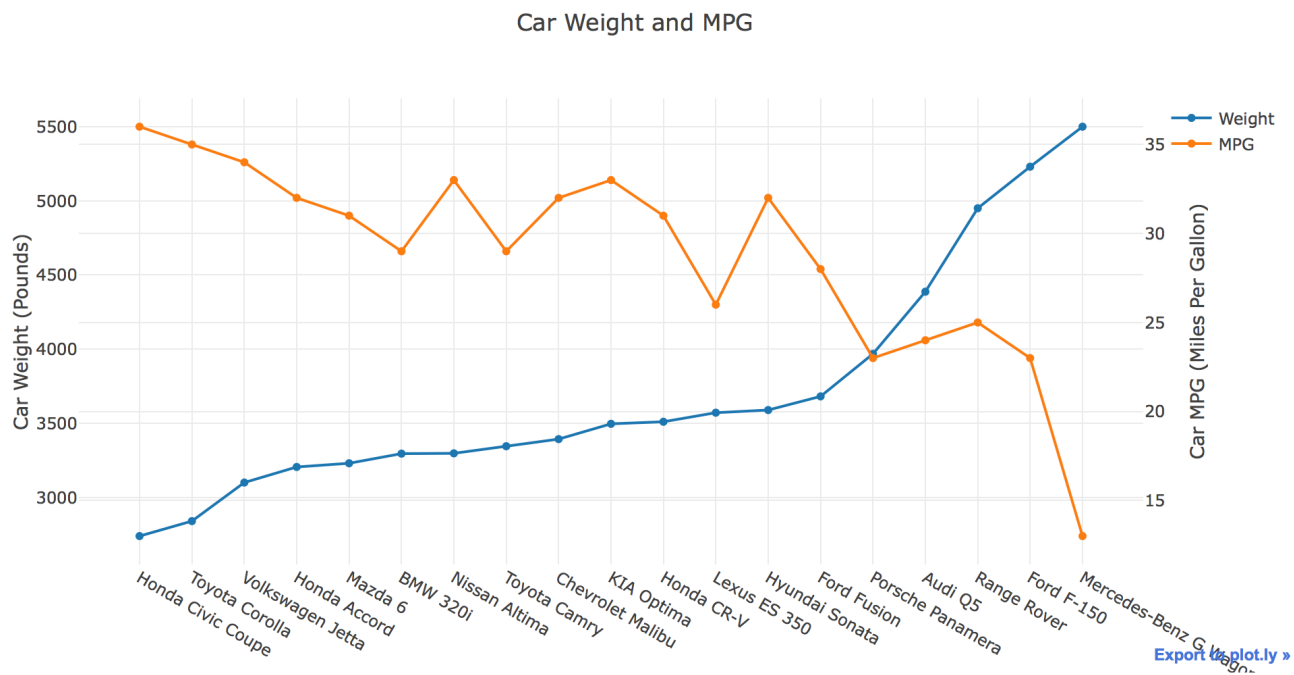
Car Weight and MPG

-   In the second graph, we explore that the car weight and MPG have a moderate negative correlation. For both the weight and mileage variable, we also use points as the mark and the channels are horizontal and vertical position. Points are used in this position because we're trying to explore a correlation between 2 variables, and the channel used is because we have a quantitative attribute at both the x and y axis.
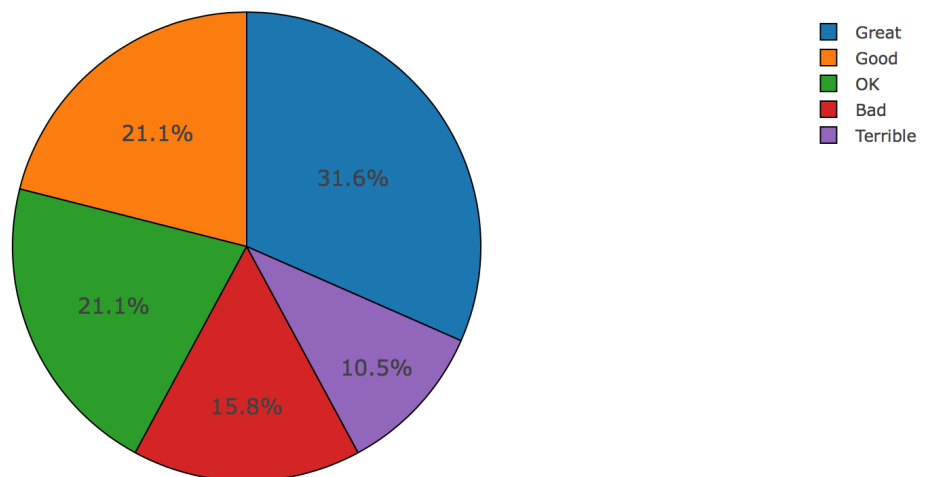
Car Weight vs. Car MPG

Other Visualizations:

In the graph below, we visualize the car's price and mileage, grouped by its condition. For the weight variable, we use line mark, while we use point mark for the MPG. 2 different mark types are used for 2 different variables to help us easily differentiate between the 2 variables. Color hue channel is used to help visualize the different categories of the car's condition.

Price and Mileage Grouped by Condition
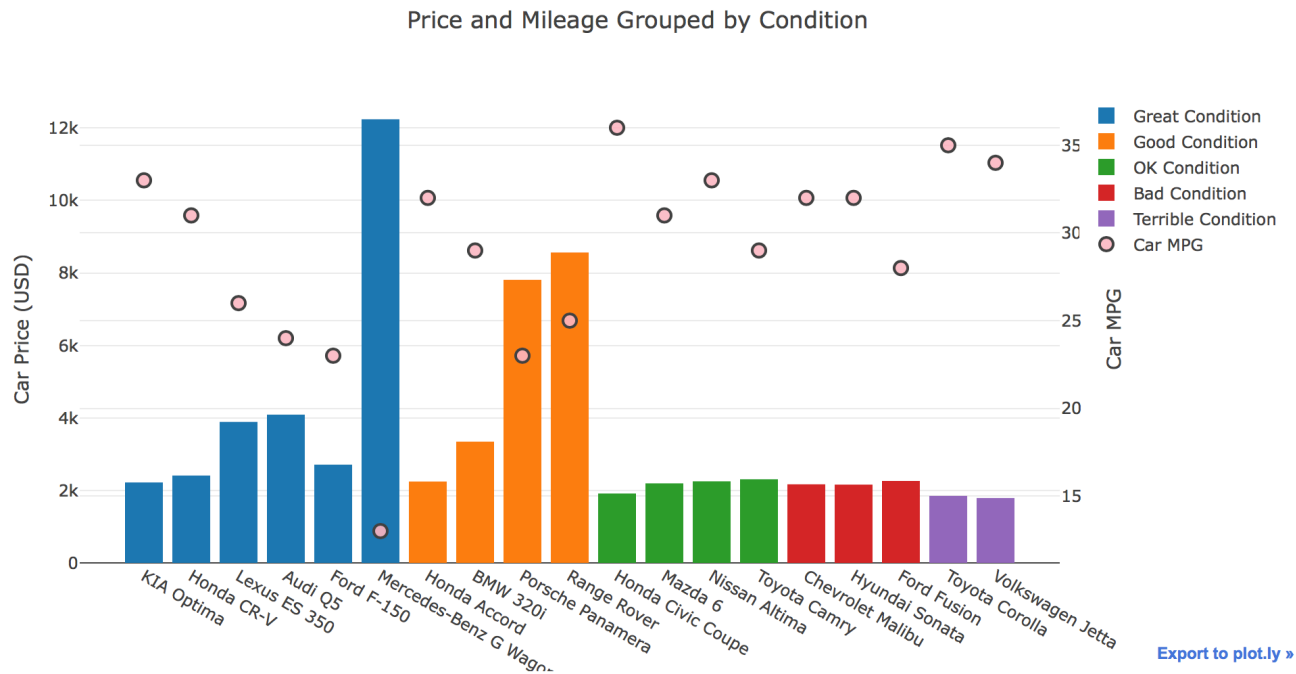
In the last visualization, we simply show the percentage (distribution) of each "repair condition" with a pie chart. We use area mark and are, color channel to clearly visualize different condition categories and how much each category makes up.

Car Condition Distribution.