```
[84] # Importing libraries
      import pandas as pd
      import matplotlib.pyplot as plt
      import numpy as np
      import seaborn as sns
      #Importing dataset,
      carData=pd.read_csv('/content/drive/My Drive/data.csv')
      carData.head()
[86]
      #Data types
      carData.dtypes
      #Statistical summary
      carData.describe()
      #Shape
 [88]
      carData.shape
 [89] #NULL values
       d=carData.isnull().sum()
```

From the above output we can clearly see that there are maximum number of null value

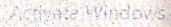
+ Code + Text



From the above output we can clearly see that there are maximum number of null values in the 'Market Category' column. Thus, for the ease of things we opt to remove it from the labelled columns. Also, we see that various more parameters do not have much significance in determining the prices of the cars. Thus dropping those columns won't make much of a difference in the processing of the model.

- [90] #Eliminating the insignificant columns carData = carData.drop(['Engine Fuel Type', 'Number of Doors', 'Market Category'], axis = 1) carData.head()
- [91] carData.shape
- [92] carData.rename(columns = { "Engine HP": "HP", "Engine Cylinders": "Cylinders", "Transmission Type": "Transmission", "Driven_Wheels": "Drive Mode", "highway MPG": "MPG-H", "city mpg": "MPG-C", "MSRP": "Price"}, inplace = True)
- [93] carData.drop duplicates()
- [94] #NULL values print(carData.isnull().sum())
- [95] carData=carData.dropna() carData.count()
- carData.isnull().sum()

벍































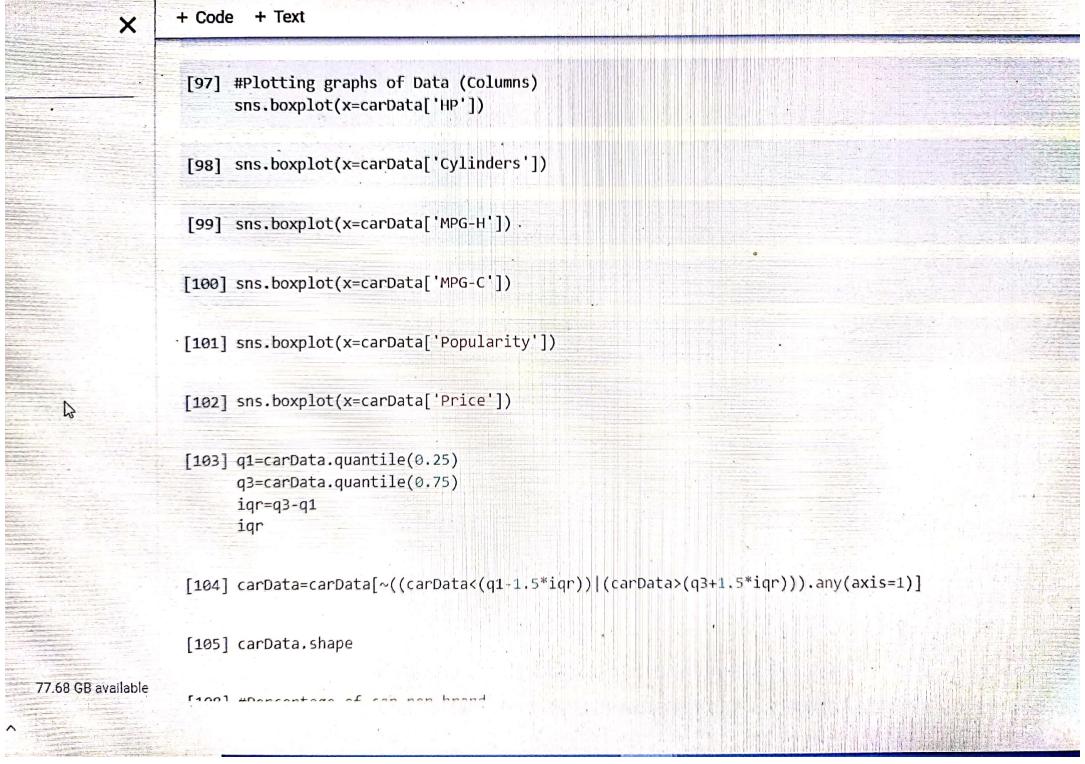


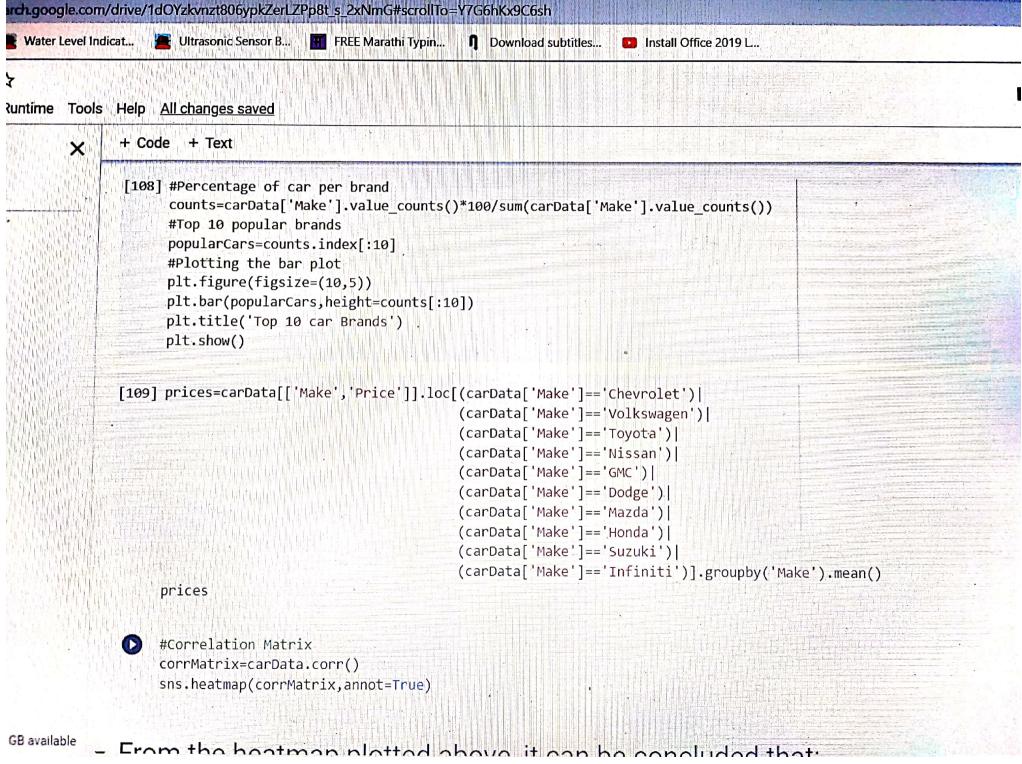














- + Code + Text
- From the heatmap plotted above, it can be concluded that:
 - >>> Price is positively dependent on features and Horse Power(HP) and Year
 - >>> The features HP and Cylinders are positively dependent on each other
 - i.e. if number of Cylinders are increased, the HP also increases.
 - >>>MPG-H and MPG-C have strong negative correlation with Cylinders.
 - i.e. if number of cylinders are increased, MPG-H and MPG-C decreases.
 - [115] sns.barplot(carData['Year'],carData['Price'])
 - sns.barplot(carData['HP'],carDita['Price'])
 - sns.barplot(carData['Cylinders'],carData['Price'])
 - sns.barplot(carData['MPG H'], carData['Price'])

Tools Help All changes saved + Code + Text × sns.barplot(carbata[Cylinders],carbata[Price]) sns.barplot(carData['MPG-H'],carData['Price']) [118] sns.barplot(carData['MPG-C'],carData['Price']) sns.barplot(carData['Popularity'],carData['Price'])