**Asssignment 2**

**Question 1:** **Look for the missing values in all the columns and either impute them (replace with mean, median, or mode) or drop them. Justify your action for this task.**

A screenshot of a computer

Description automatically generated

Updating the dataset with replacing the missing values.

A screenshot of a computer

Description automatically generated

**Question 2: Remove the units from some of the attributes and only keep the numerical values (for example remove kmpl from “Mileage”, CC from “Engine”, bhp from “Power”, and lakh from “New\_price”).**

A screenshot of a computer

Description automatically generated

**Question 3: Change the categorical variables (“Fuel\_Type” and “Transmission”) into numerical one hot encoded values.**

A screenshot of a computer

Description automatically generated

**Question 4: Create one more feature and add this column to the dataset (you can use mutate function in R for this). For example, you can calculate the current age of the car by subtracting “Year” value from the current year.**

A screenshot of a car

Description automatically generated

**Question 5: Perform select, filter, rename, mutate, arrange and summarize with group by operations (or their equivalent operations in python) on this dataset.**

1. Selected data:

A screen shot of a computer

Description automatically generated

1. Renaming:

A black screen with white text

Description automatically generated

1. Arrange:

A screen shot of a computer

Description automatically generated

1. Summarizing using heatmap:

A screenshot of a computer generated diagram

Description automatically generated