

ASSIGNMENT -2

Find the missing values in each column.

Unnamed: 0 0

Name 0

Location 0

Year 0

Kilometers_Driven 0

Fuel_Type 0

Transmission 0

Owner_Type 0

Mileage 2

Engine 36

Power 36

Seats 38

New_Price 5032

Price 0

dtype: int64

- We use the median to fill in the missing values in the Mileage, Engine, and Power columns. This approach was made because of the average exceptions, ensuring that it accurately represents the data's core tendency. Using the median, we reduce the impact of extreme values and provide more representative estimates for missing data.
- We must use Mean to replace missing values in the Seat column, as Mode depends on the most often occurring number of seats.
- We removed the entire column as the new price column has more than 50% missing values, which could introduce insufficient data and cause wrong analysis.