Project Report -1

ANALYSIS-1

1.Perform general Data analysis?

SUMMARY

- Firstly, I have performed general data analysis like data collection, data cleaning, Exploratory data analysis (EDA) and Reporting.
- Check the Null values with the help of isnull().sum() and Filling Null values with mean ,median or mode depend on other column
- Check the data type and change the datatypes depends on column data.
- Remove the duplicate values with the use drop function.
- Secondly, after complete data cleaning your analysing is good because of cleaning.

2. Can you tell me the Distribution of Vehicles based on Year of Registration with the help of a plot?

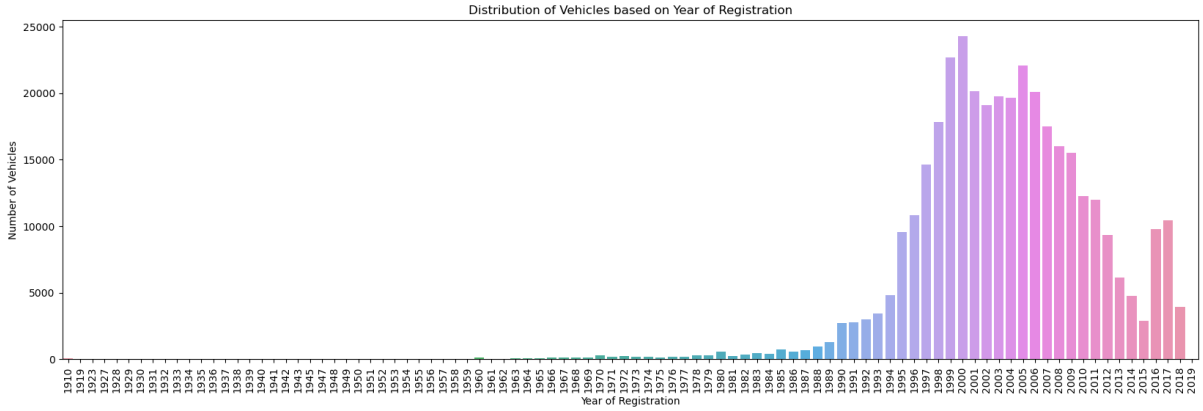
SUMMARY

- I have observed that the distribution of vehicles based on the year of registration

```
plt.figure(figsize=(20, 6))
sns.countplot(data=data, x='yearOfRegistration')
plt.title('Distribution of Vehicles based on Year of Registration')
plt.xlabel('Year of Registration')
plt.ylabel('Number of Vehicles')
plt.xticks(rotation=90)  # Rotate x-labels for better readability
plt.show()
```

- I am analysing the data show the visualize.
- I am took the help of matplot.lib and seaborn with import the two and using count plot.
- I have observed that the distribution of vehicles based on the year of registration provides valuable insights into the age composition of the autos. Additionally, the distribution can reveal trends in the market, such as the prevalence of vintage cars or influx of newer models. For instance, a higher concentration of vehicles might be

observed in certain years, indicating popular models. Overall, understanding this distribution is crucial for both buyers and sellers, helping them make informed decisions about pricing demand, and market trends.

- 

3. Create a plot based on the Variation of the price range by the vehicle type?

SUMMARY

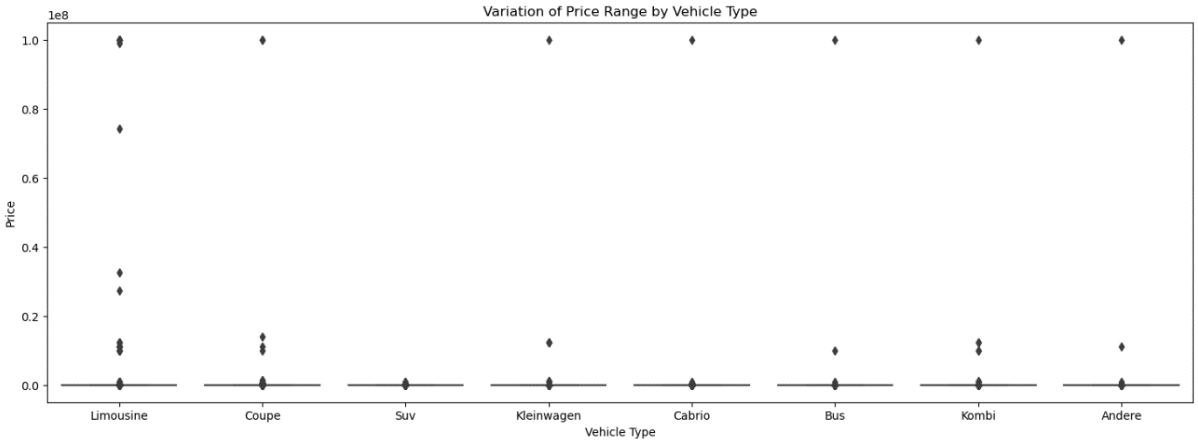
- I have observed that the variation of of the price range by the vehicle type.

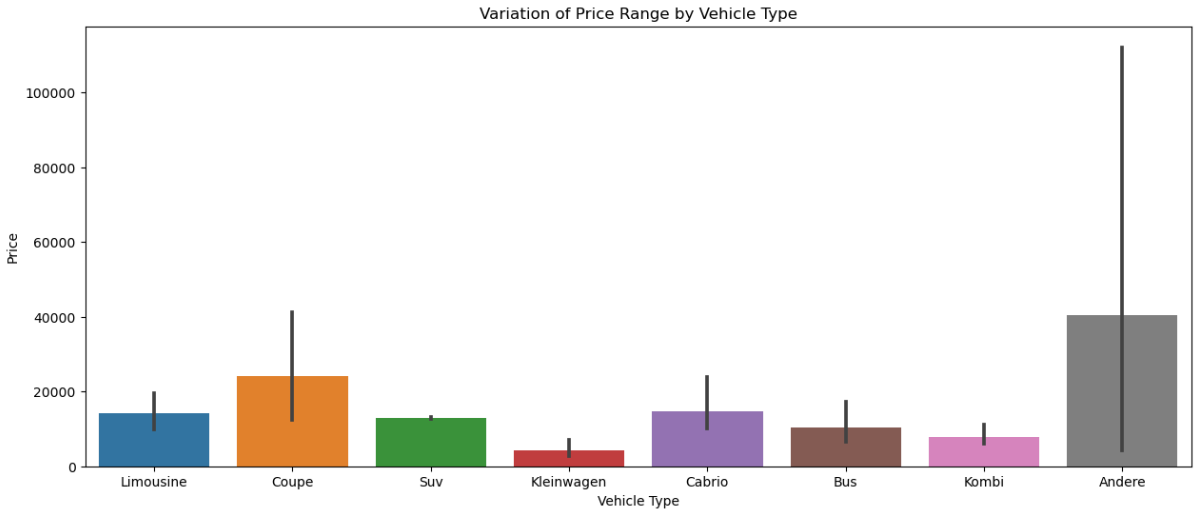
```
plt.figure(figsize=(18, 6))
sns.boxplot(data=data, x='vehicleType', y='price')
plt.title('Variation of Price Range by Vehicle Type')
plt.xlabel('Vehicle Type')
plt.ylabel('Price')
plt.show()
```

```
plt.figure(figsize=(15, 6))
sns.barplot(data=data, x='vehicleType', y='price')
plt.title('Variation of Price Range by Vehicle Type')
plt.xlabel('Vehicle Type')
plt.ylabel('Price')
plt.show()
```

- I am analysing the data show the visualize.
- I am took the help of matplot.lib and seaborn with import the two and using box and bar plot .

- The variation of the price range by vehicle type provides valuable insights into the pricing dynamics across different auto categories. For instance, luxurious types like “Coupe” and “SUV” tend to have higher average prices, reflecting their premium status. In contrast, more practical types like “Kleinwagen” (Compact Cars) have considerably lower average prices, suggesting is a key factor for buyers in these categories. This analysis aids both sellers and buyers in understanding the market trends. and making well-informed purchasing decisions.





4. Find out Total count of vehicles by type available on ebay for sale .As well as create a visualization for the client?

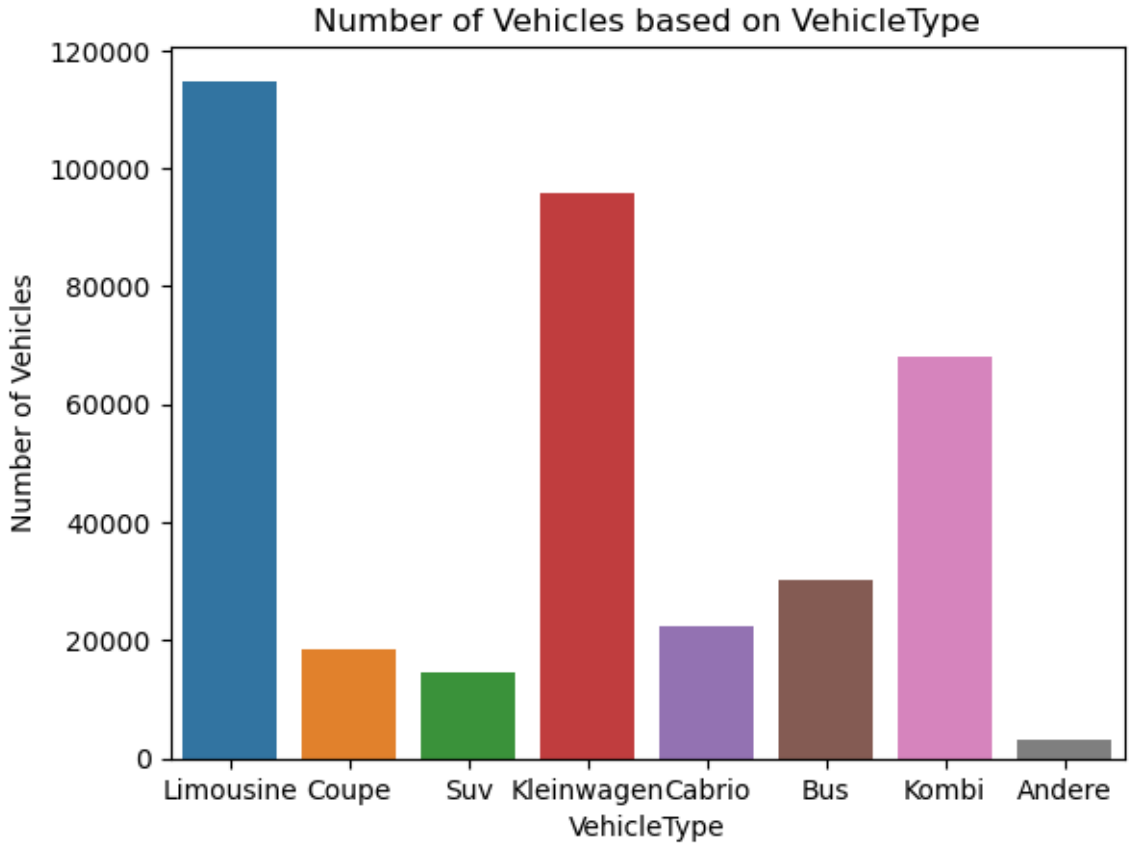
SUMMARY

- I have observed that the Total count of vehicles by type available on ebay for sale .

- I am took the help of value_counts() function because of total count we are asking I am doing value count function.

```
sns.countplot(data=data,x="vehicleType")
plt.title('Number of Vehicles based on VehicleType')
plt.xlabel('VehicleType')
plt.ylabel('Number of Vehicles')
plt.show()
```

- I am analysing the data show the visualize.
- I am took the help of matplot.lib and seaborn with import the two ,count plot is used.
- The total count of vehicles by type available on e-bay for sale provides a comprehensive overview of the platform's inventory. "Kleinwagen" and sedans might dominate indicating their high market demand and availability. "Carbio" and "SUV" might have lesser listings reflecting their limited market presence. Understanding these are crucial for market analysis.



5. Is there any relationship between dollar price and kilometer? (Explain with appropriate analysis).

SUMMARY

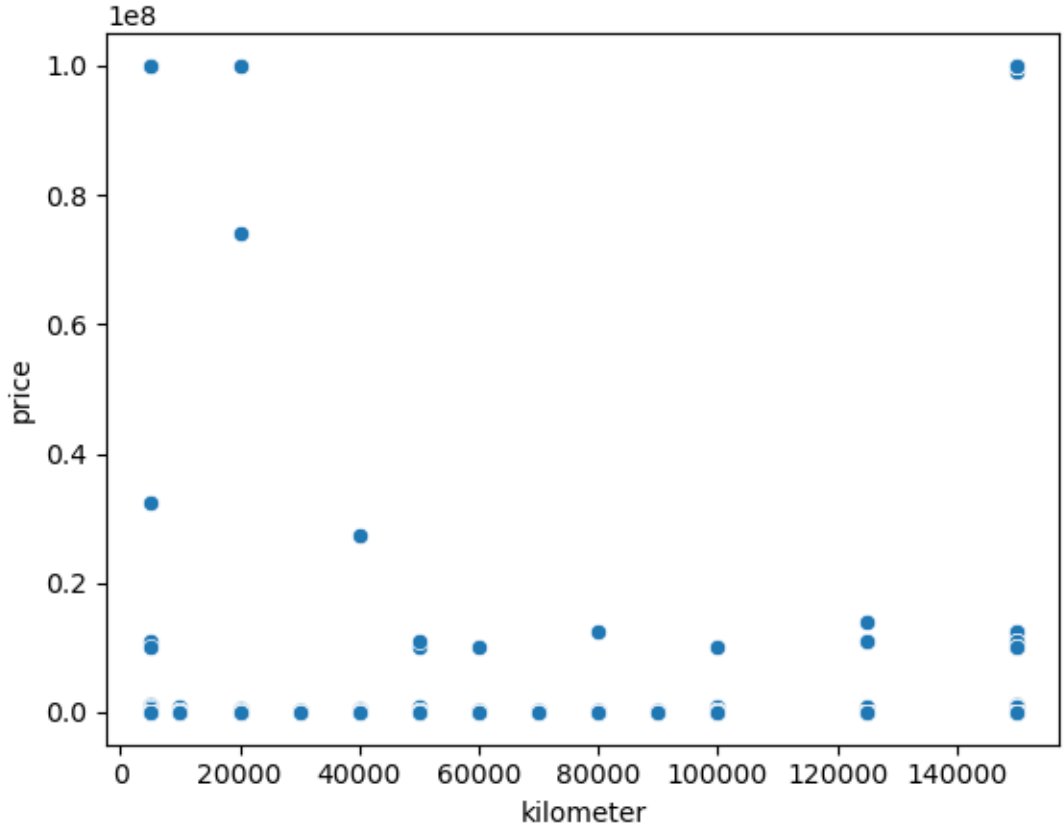
- I have observed that the relationship between dollar and kilometer .
- I am using correlation function only identify the relation ship between the two column must and should use corr().

```
correlation=data["price"].corr(data["kilometer"])
correlation
```

```
-0.007683223435760035
```

```
sns.scatterplot(x='kilometer', y='price', data=data)
plt.show()
```

- I am analysing the data show the visualize.
- I am took the help of matplot.lib and seaborn with import both ,scatter plot is used.



- The correlation coefficient ranges from -1 to 1.
    1. If the correlation coefficient is close to 1, with
    2. indicates a strong positive correlation.
    3. If the correlation coefficient is close to -1, it indicates a strong negative correlation.
    4. If the correlation coefficient is close to 0, it indicates a weak or no correlation.

- Here I have observed correlation is close to 0, between price and kilometer. If the points on the scatter plot are widely dispersed and do not follow a clear pattern, it suggests a weak or no correlation between the two variables.