**-**

**Advance Data Analysis**

**On**

**Cars Dataset**

**Name :- Adnan Baig**

**Batch :- 251**

* **Analysis 1:--**

**1) Perform general Data analysis:-**

\* Importing Libraries

\* Importing Dataset

\* Data Cleaning and Manipulation

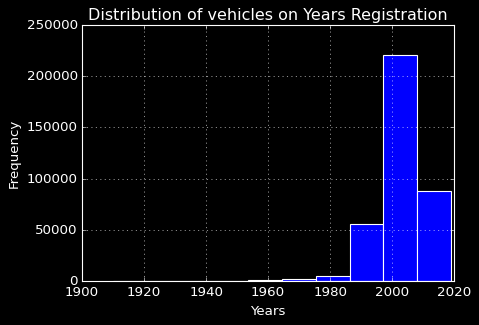
\* Filling missing values

\* Handling with structural error

\* Handling with Typecasting

\* Columns proper sequence

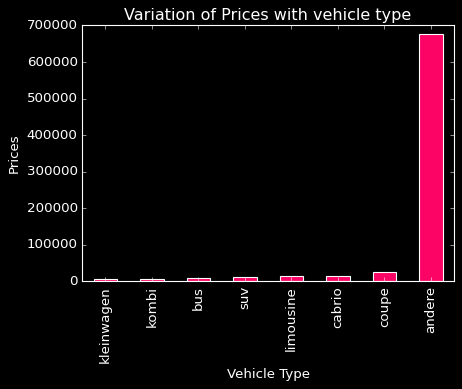
**2) Can you tell me the Distribution of Vehicles based on Year of Registration with the help of a plot :-**

****

\* Above bar plot is showing the distribution of Vehicle based on Registration Year , the maximum vehicles are having registration year 2000.

\* There less vehicles with old registration years.

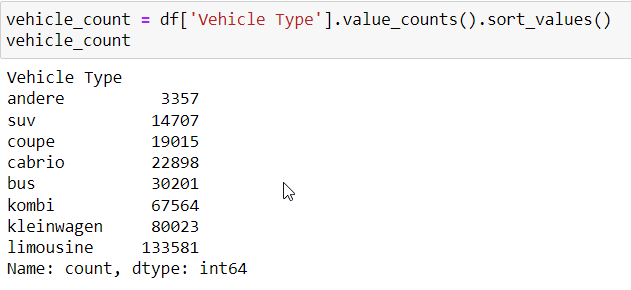
**3) Create a plot based on the Variation of the price range by the vehicle type :-**

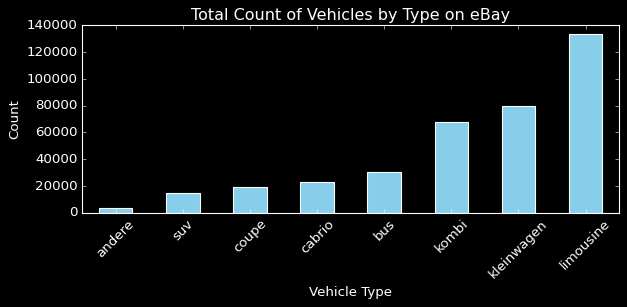
****

\* Above plot is showing there is high average price in andere type of vehicle.

\* Rest of vehicle types are less prices.

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

****

****

\* Above plot is showing there is maximum and near about 130000 vehicles are limousine type vehicle

\* Less vehicle are from andere type of vehicle

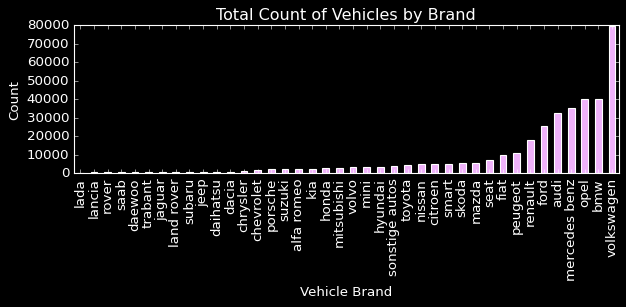
**5) Is there any relationship between dollar price and kilometer? (Explain with appropriate analysis**

**88**

\* From above scatter plot and heatmap we can see there is -0.2 negative week relationship between registration years and prices, prices is not influencing by years.

* **Analysis 2:--**

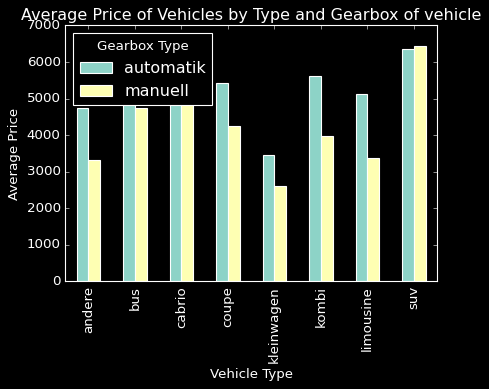
1. **Can you tell me No of Vehicles by Brand Available on ebay for sale with the help of visualization**

****

**\*** Above plot is showing no of vehicle based on vehicles brand , there is maximum vehicle from Volkswagen near about 80000

1. **What is the Average price for vehicles based on the type of vehicle as well as on the type of gearbox. Explain me with both numerical and visualization analysis**

****

****

**\*** Above plot is showing manual gearbox from andere vehicle type having maximum average price

**3) What is the marginal probability of private seller**

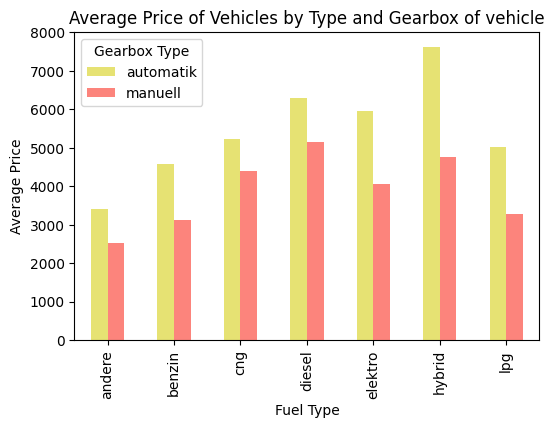
\* Marginal probability of 'private' seller in the 'seller' column is ---> 0.9999919212809617

* **Analysis 3:--**

**1) The memory usage of the data is around 6.1 mb.How can we reduce the memory usage of the data set?**

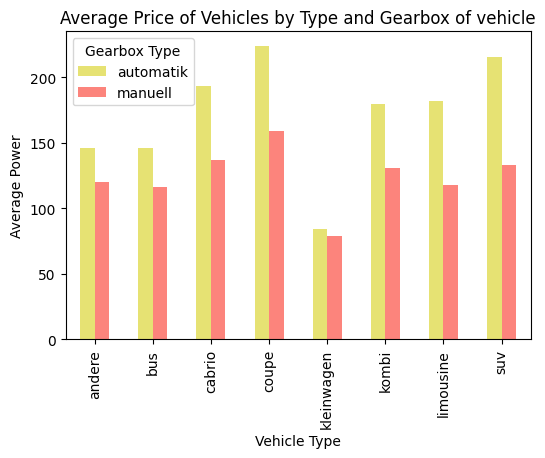
\* We can reduce the usage of the data by using appropriate and proper datatypes because its contains different bytes for different datatypes.

**2) What is the Average price of vehicle by fuel type and gearbox type. Give a plot**

****

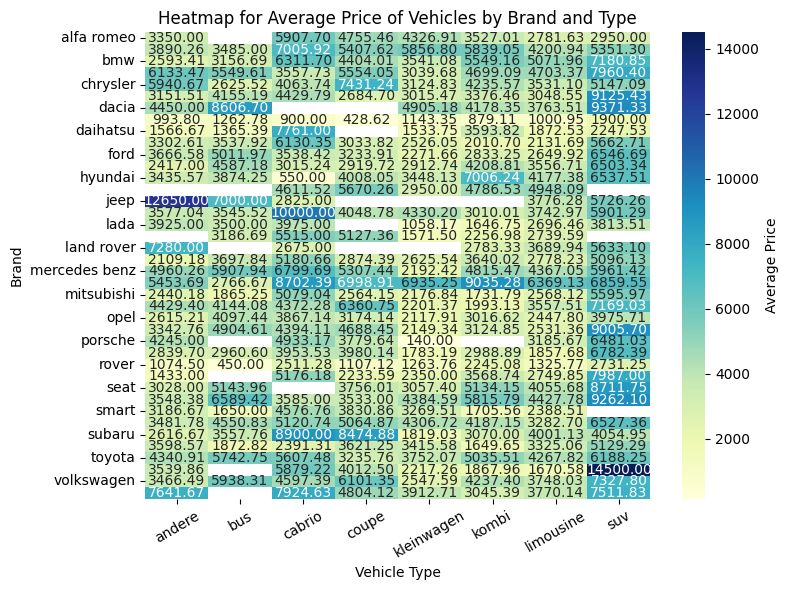
\* Above plot is showing manual gearbox and andere fuel type having maximum avg price

1. **What is the Average power of a vehicle by vehicle type and gearbox type.Give a plot**

****

\* Above plot is showing automatic gearbox from coupe vehicle type having maximum avg power more than 200 power ps

**4) What is the Average price of a vehicle by brand as well as vehicle type. Use heatmap to explain this.**

****