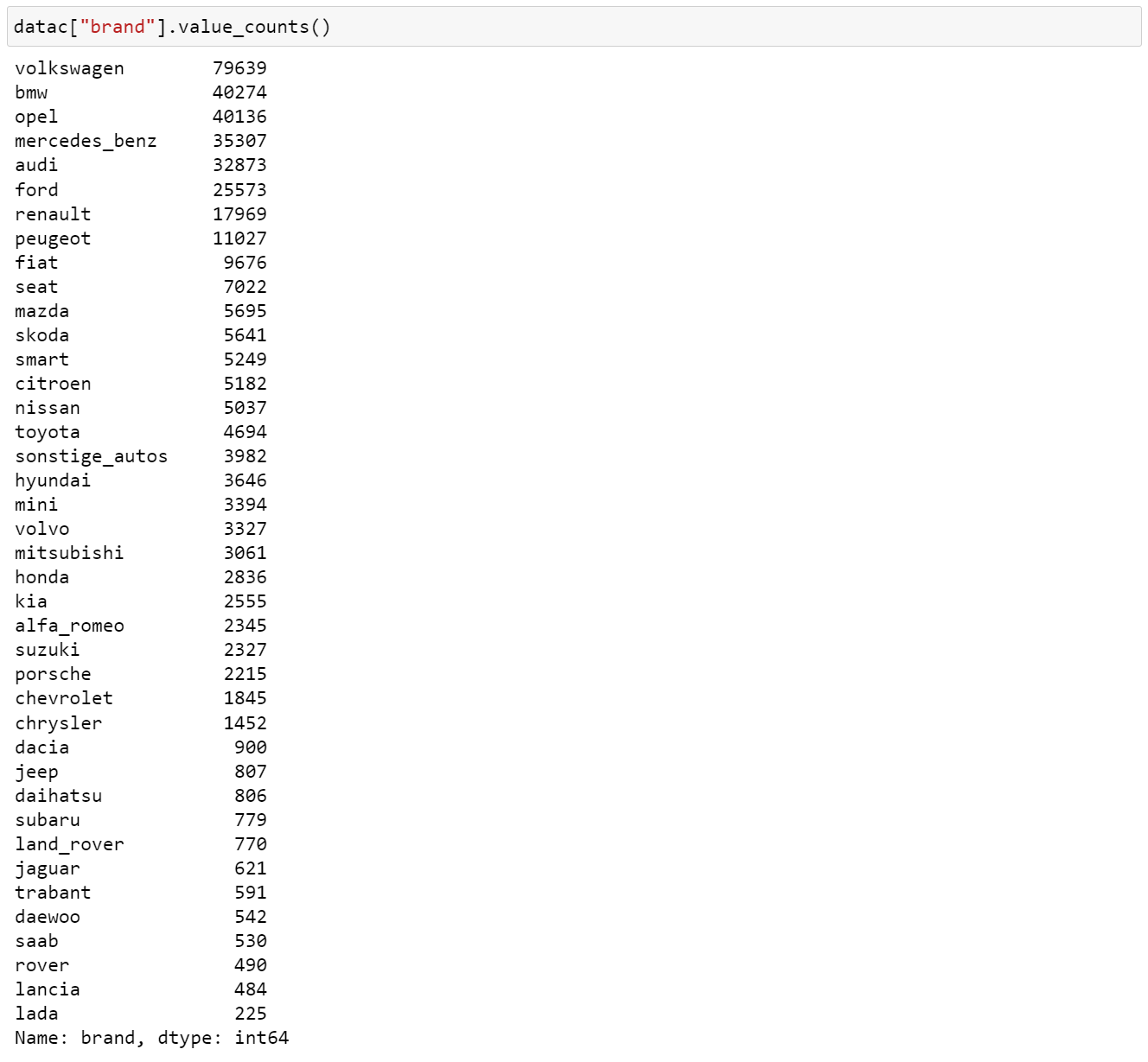
**Report**

**Analysis-2:**

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

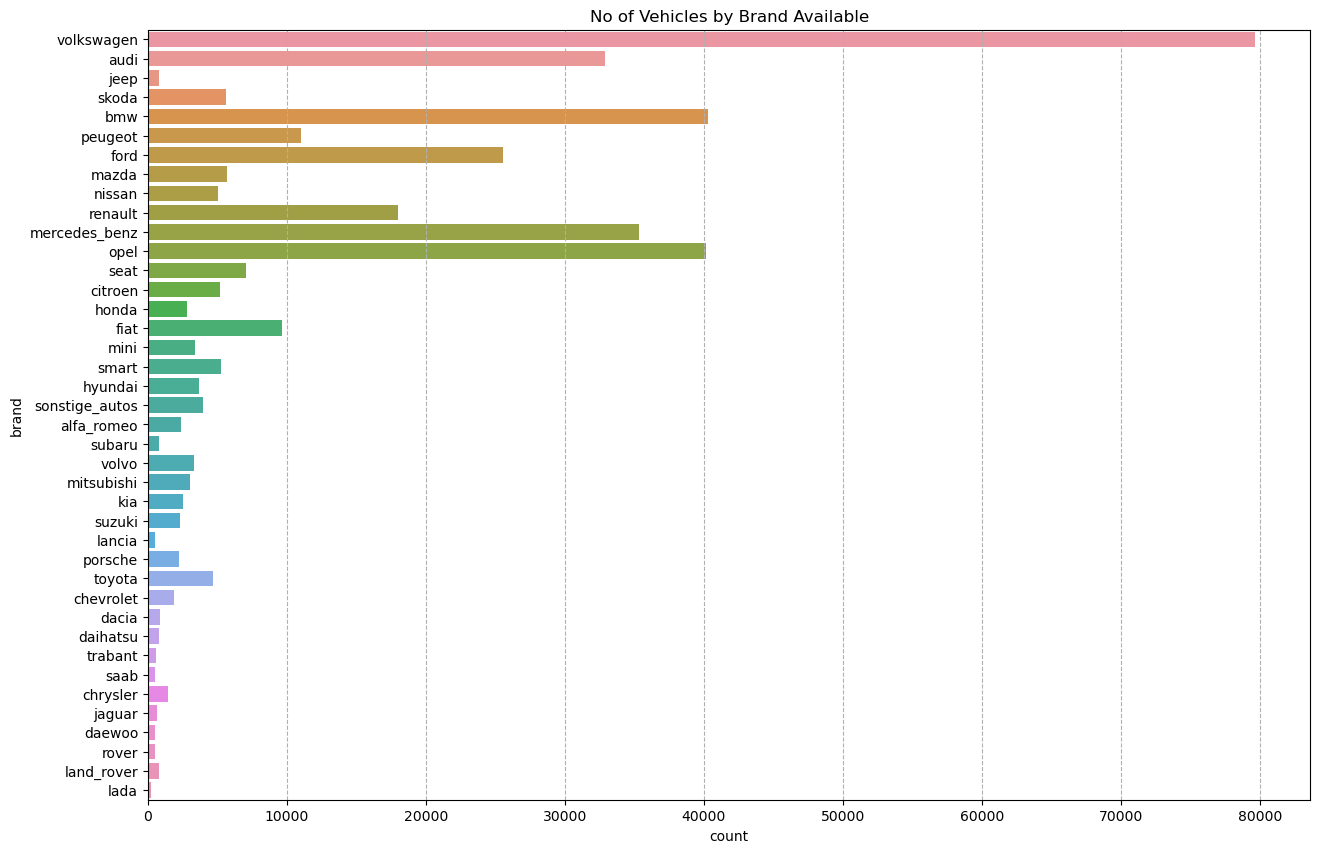
* The no of vehicles by brand available on eBay gives the information of no of vehicles available in different brands on eBay for sale.
* For that first we have to access brand column and apply value\_counts (), then the output is:

****

* For visualization, we used count plot to visualize the number of vehicles by brand available on eBay. Below is the code:

****

* The visualization for the number of vehicles by brand available on eBay is



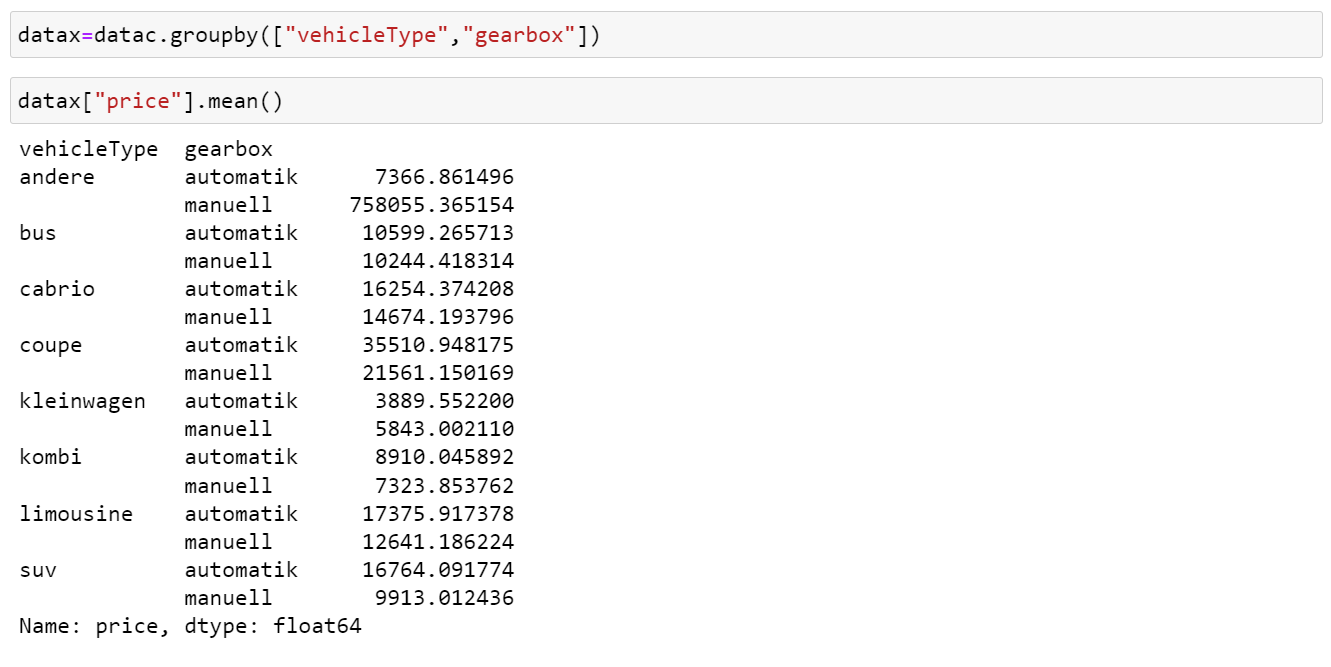
* The y-axis represents the different brands of vehicles, and x-axis represents the count of vehicles that are registered from different brand for sale.
* Out of all brands available, Volkswagen brand has highest no of vehicles 79639, by

brand available on eBay for sale.

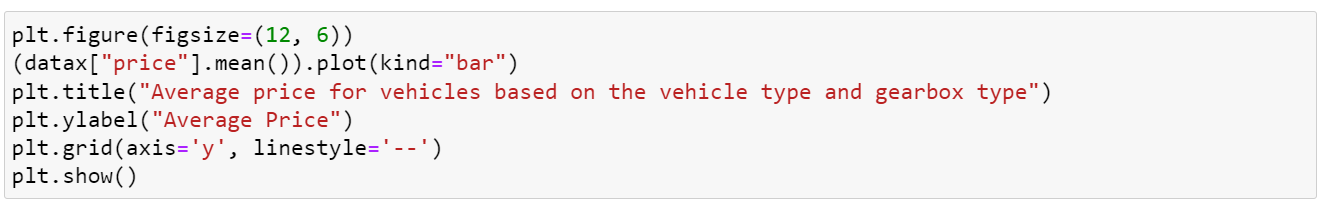
* Lada brand has lowest no of vehicles 225, by brand available on eBay for sale.

### 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:

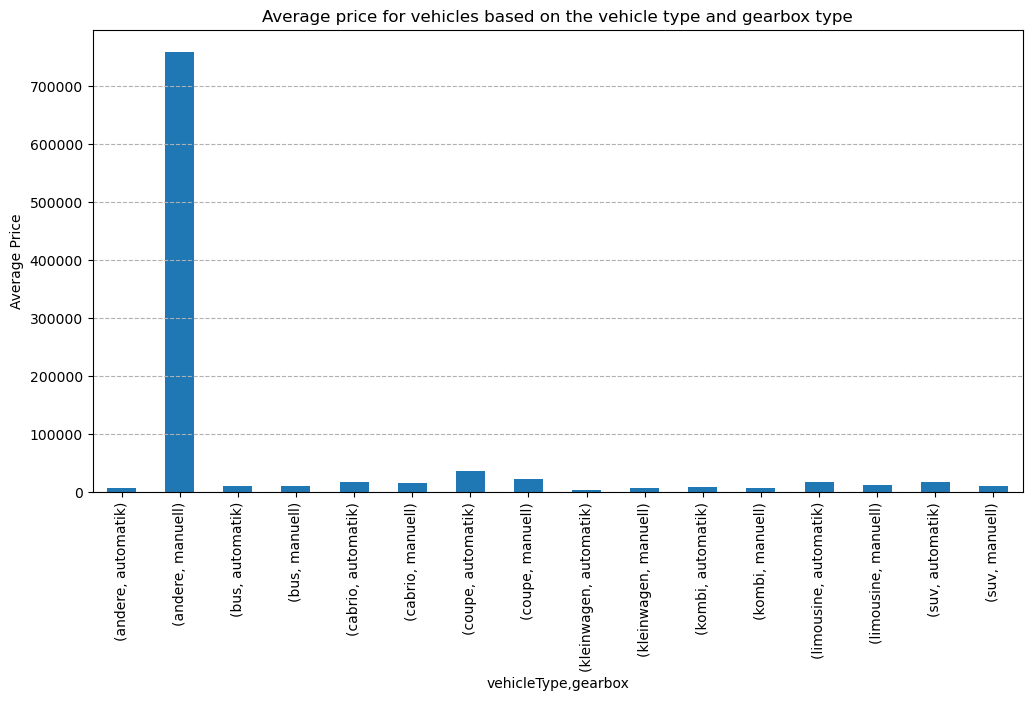
* Average price for vehicles based on the vehicle type and gearbox type provides insights about average price for vehicle type and gearbox both at a time.
* We can get the average price for both at a time by using groupby method.
* For that apply mean to get the average price for vehicles based on vehicle type and gear box. Below is the code:



* Vehicle type andere and gearbox type manuell has the highest average price 758055.365.
* Vehicle type kleinwagen and gearbox type automatic has the lowest average price 3889.552.
* For visualization, we used bar plot to visualize the average price for vehicle type and gearbox at a time. The code is:



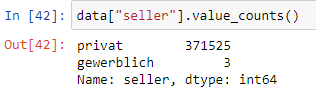
* The visualization for the average price for vehicles based on vehicle type and gear box is



* the x-axis represents the vehicle type and gearbox type, and the y-axis will represent the average price.
* Due to the outlier in the vehicle type the average price is affected. So that one bar is highest and all other are lower compared to that.

### What is the marginal probability of private seller:

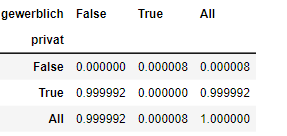
* To get marginal probability for privat seller, we have to access the seller column.
* First, we should know about the all types of sales in the seller column for that we need to collect the seller column data and apply value\_counts (). This will give you the information as



* Here we have 371525 counts of private sellers and only 3 gewerblich sellers in the data.
* Now we have to use crosstab by giving seller column for both index and column, where rownames as privat and colnames as gewerblich.
* But to know the marginal probability of the privat column we have to use the Parameter (margin =True).
* To get the marginal probability for the data, we use the code as:

**pd.crosstab (index=datac["seller"]=="privat",columns=data["seller"]=="gewerblich",rownames=["privat"],colnames=["gewerblich"],normalize="all",margins=True)**

* Given normalize= “all” to get the marginal probability.
* When we use the parameter in the code margin =True, the margin option will give the marginal probability.
* This will give the information as follows:



* This will represent that the gewerblich as a column and private as a row.
* **So, marginal probability for privat seller is 0.999992.**
* The marginal probability for not a private seller is 0.000008
* The marginal probability for gewerblich seller is 0.000008
* The marginal probability for not a gewerblich seller is 0.999992