ANALYSIS-2

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

* To know the data for no of vehicles by brand on sale

We can collect the data of brand from the dataframe as

datac["brand"].value\_counts()

* It will gives the count of each brand available on sale

To visualize the data I used the count plot as

sns.countplot(y=datac["brand"])

* Will gives the count of each and every individual brand as a bar representation in the plot

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

* To know the average price foe vehicle type and gearbox at a time
* We can collect the data using groupby method in dataframe as

avg\_price=datac.groupby(["vehicleType","gearbox"])["price"].mean()

* Finally it will gives the all average prices of vehicle type and gear box type
* To represent this in plotting I use the bar plot as using the code

avg\_price=datac.groupby(["vehicleType","gearbox"])["price"].mean()

avg\_price.plot(kind="bar")

plt.xlabel(["vehicleType and gearbox type"])

plt.ylabel(["average price"])

plt.title("avg price of vehicletype and gearbox")

plt.show()

* It will gives the combination of each individual vehicle type and gearbox as a bar and mean price on another axis

1. What is the marginal probability of private seller

* To find the marginal probability of seller==”private”

We have to collect all the data of seller column

* To get the marginal probability we have to crosstab the seller column into two different sets as index to “private” and column as “gewerblitch”
* The code is as follows

pd.crosstab(index=datac["seller"]=="privat",columns=data["seller"]=="gewerblich",normalize="all",margins=True)# marginal probability

🡪where in this code we given the data as one set as private and another as gewerblitch

* Finally it gives the all cross tab values based on the parameter normalize==”all”
* But to know the marginal probability of the private column we have to use margin =True to get the intersected values in the data
* At finally it gave the marginal probability value as 0.999992