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?

* To reduce the memory size we can use many different cases like

🡪Removing the duplicated rows by using the code as   
 datac.drop(datac[duplicated],inplace=True)

🡪removing unwanted useless columns based on the requirement

Datac.drop[datac[“column name”],inplace= True)

Which will permanently removes the column based on the parameter called inplace = True

🡪we can also reduce the data memory size by using the content called type conversion

By accessing the columns and changing the datatype to the columns by using the code as

Datac[“column name”].astype(int8) or Datac[“column name”].astype(int16)

Datac[“column name”].astype(int32) or Datac[“column name”].astype(float16)

Datac[“column name”].astype(float32) etc.,

As the requirement of our dataset we can use these concepts to reduce the data memory set

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

🡪To know the average price for fuel type and gearbox at a time

🡪We can collect the data using groupby method in dataframe as

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

🡪Finally it will gives the all average prices of fuel type and gear box type

To represent this in plotting I use the bar plot as using the code

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

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

plt.xlabel(["fuel Type and gearbox type"])

plt.ylabel(["average price"])

plt.title("avg price of fuel type and gearbox")

plt.show()

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

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

To know the average power for vehicle type and gearbox at a time

We can collect the data using groupby method in dataframe as

avg\_ power =datac.groupby(["vehicleType","gearbox"])[" power "].mean()

Finally it will gives the all average power of vehicle type and gear box type

To represent this in plotting I use the bar plot as using the code

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

avg\_ power.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 power on another axis

1. What is the Average price of a vehicle by brand as well as vehicle type.Use heatmap to explain this?

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

avg\_ price =datac.groupby(["vehicleType","brand"])["price"].mean()

* Finally it will gives the all average power of vehicle type and brand

To represent this in plotting I use the heatmap as using the code as

plt.figure(figsize=(10,10))

sns.heatmap(data=avg\_price,annot=True,fmt=".4g")

plt.title("Average price of a vehicle by brand as well as vehicle type")

plt.show()

the above code gives us a heatmap plot which gives the average prices for the vehicle type and brand at a time