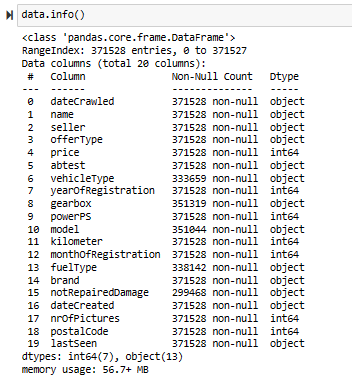
# Report c:-

## C-1 The memory usage of the data is around 6.1 mb.How can we reduce the memory usage of the data set?



In [13]:

data.columns

Out[13]:

Index(['dateCrawled', 'name', 'seller', 'offerType', 'price', 'abtest',

'vehicleType', 'yearOfRegistration', 'gearbox', 'powerPS', 'model',

'kilometer', 'monthOfRegistration', 'fuelType', 'brand',

'notRepairedDamage', 'dateCreated', 'nrOfPictures', 'postalCode',

'lastSeen'],

dtype='object')

In [15]:

data.drop(["nrOfPictures"],axis**=**1,inplace**=True**)

In [17]:

data.drop(["dateCrawled"],axis**=**1,inplace**=True**)

In [19]:

data.drop(["dateCreated"],axis**=**1,inplace**=True**)

In [21]:

data.drop(["postalCode"],axis**=**1,inplace**=True**)

In [23]:

data.drop(["lastSeen"],axis**=**1,inplace**=True**)

In [26]:

data["price"]**=**data["price"].astype(np.int32)

In [27]:

data["kilometer"]**=**data["kilometer"].astype(np.int32)

In [30]:

data["monthOfRegistration"]**=**data["monthOfRegistration"].astype(np.int8)

A screenshot of a computer

Description automatically generated

* From this the memory usage had reduced from 56.7 to 40.0 mb by making appropriate corrections.

## C-2. What is the Average price of vehicle by fuel type and gearbox type.Give a plot?

A screenshot of a computer

Description automatically generated

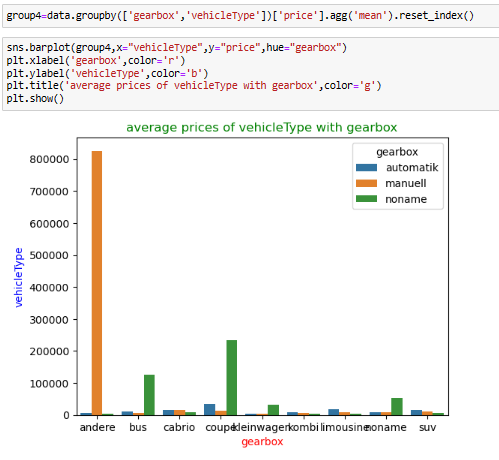
A screenshot of a computer

Description automatically generated

A graph with green and orange lines

Description automatically generated

## C-3 What is the Average power of a vehicle by vehicle type and gearbox type.Give a plot?



* In automatic coupe has highest power ps and least for kleinwagen.
* In manuell andere has highest power ps and least for kombi.

## C-4. What is the Average price of a vehicle by brand as well as vehicle type.Use heatmap to explain this ?



