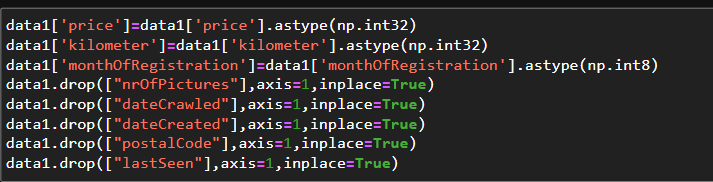
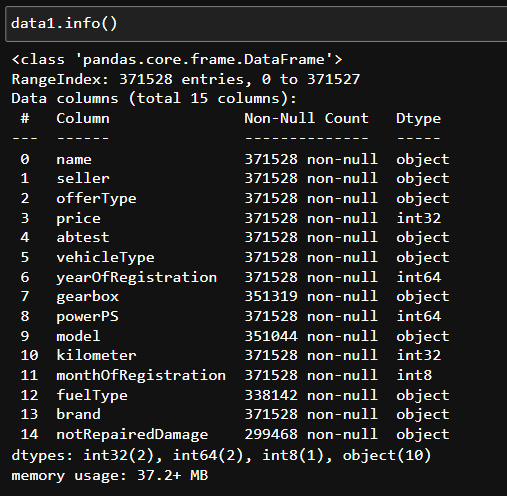
ANALYSIS – 3

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 data size from the data we have to use some techiniques in data using numpy and pandas first of all we have to know the memory size of our data to know that we use the code as “datac.info()”.
* The size of the data is 56.690799713134766 i.e., approximately equal to 56.7 MB.
* To reduce the size of our data we can change the datatypes of the columns based on the precedence of the type of data and requirements as shown below.



* After running the above code some of the unnecessary columns are dropped which are not used to our analysis. So that our analysis will be easier and memory is also decreased.
* Now, we can see the memory is also decreased from 57 MB to 32 MB.
* The info is shown in the below figure after doing manipulation.



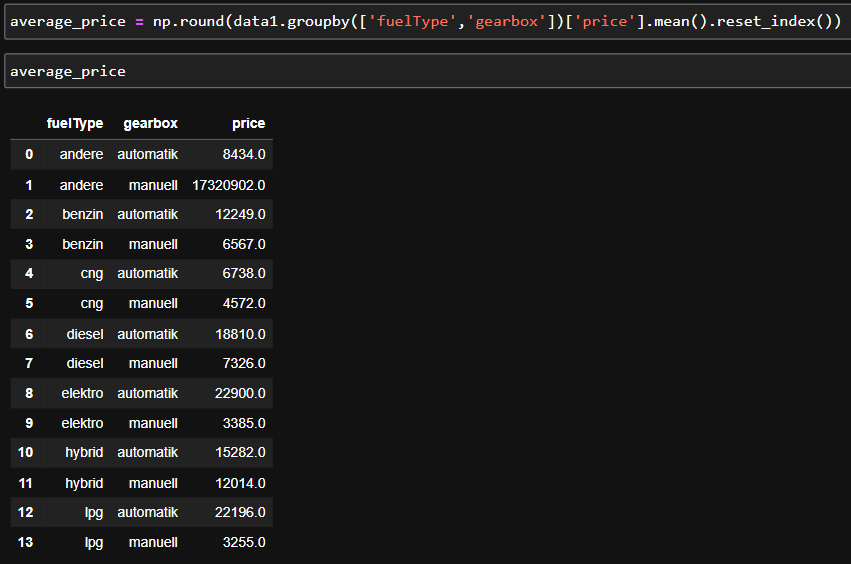
* We can see the memory usage in the below that is 37.2 MB.

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

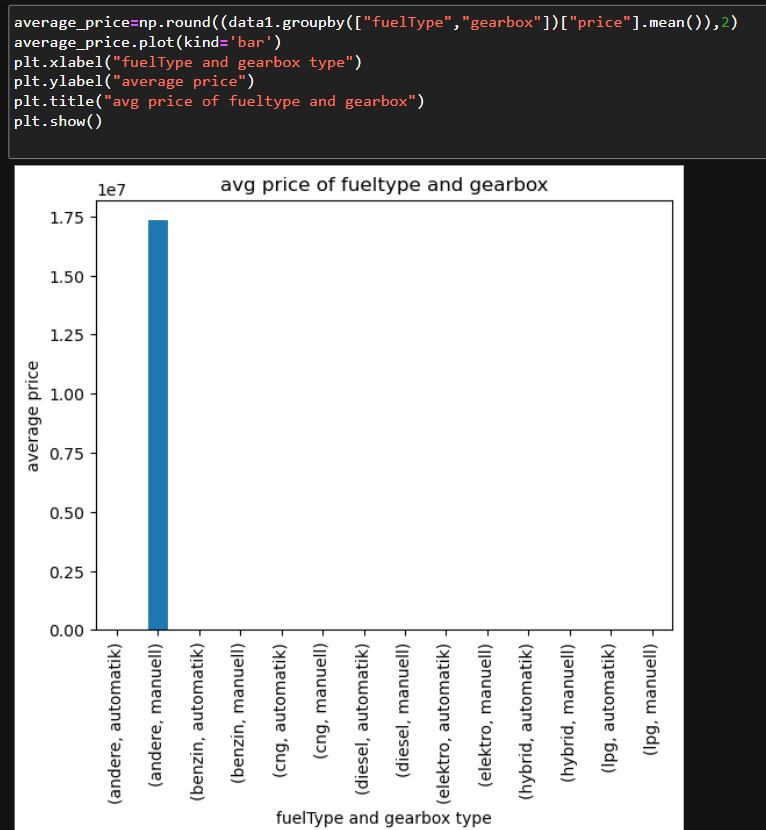
* To show the average price of vehicles based on the fuel type and gear type we should use bar plot for this
* To collect the average prices of both the fuel type and gear box at a time
* We should use the groupby method and in the groupby method we should use aggregate function as mean for the average for the price column
* This is the code for the data to be collected ….

“average\_price=data1.groupby(["fuel Type","gearbox"])["price"].mean()”

* This will gives us the data as shown below.



* From the above figure we can see the average price of the fueltype and gearbox columns.
* From the figure below we can say that, The fuel name containing “andere” and “gearbox” as manuell has the highest average prize among all the vehicle combinations of vehicle type and gear box aswell.
* Due to the outlier in the fuel type the average price is effected so that the only one bar is at the highest and all other are lower.

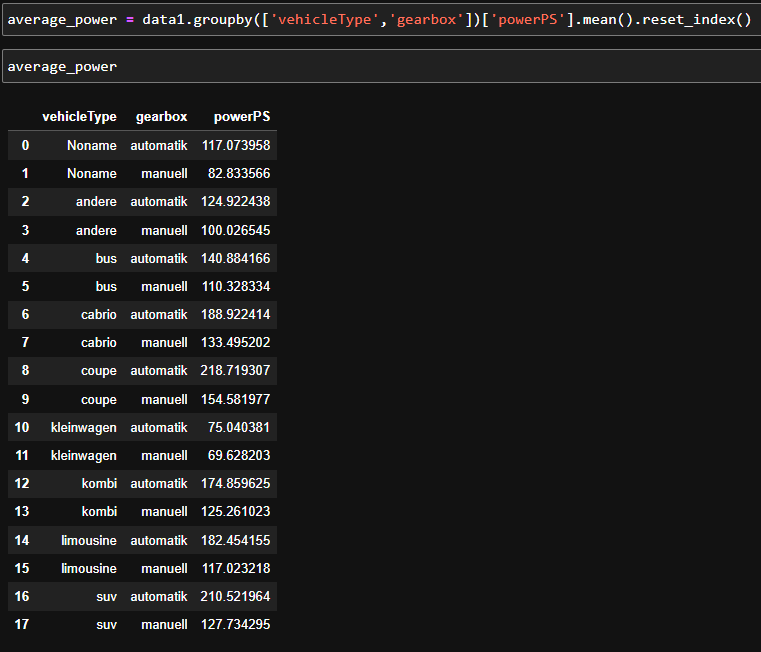


3.What is the Average power of a vehicle by vehicle type and gearbox type.Give a plot

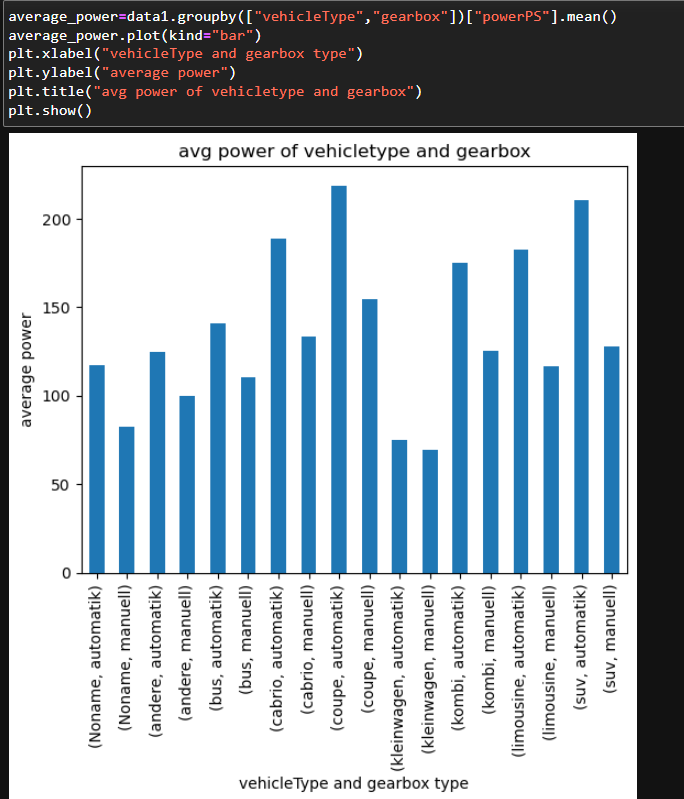
* To show the average power of vehicles based on the vehicle type and gear type we should use bar plot for this
* To collect the average power of both the vehicle type and gear box at a time.
* We should use the groupby method and in the groupby method we should use aggregate function as mean for the average for the power column.
* This is the code for the data to be collected ….

“avg\_ power =data1.groupby(["vehicleType","gearbox"])["price"].mean()”

this will gives us the data as shown below,



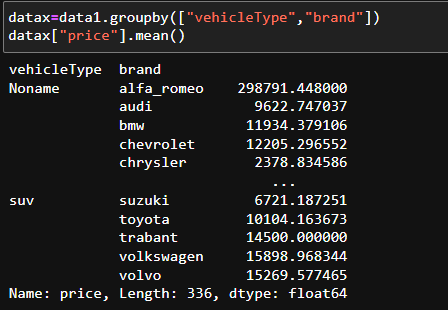
* Here shows that the all the average power about the combination of vehicletype and gearbox.
* This can also be shown by using the bar plot as the code.



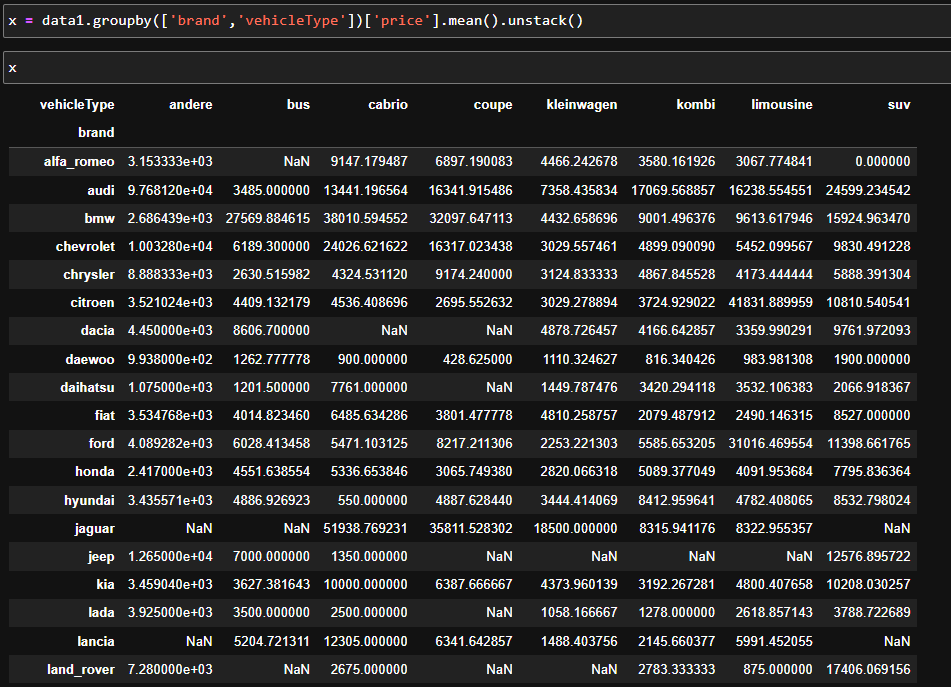
* From the figure above we can say that
* The vehicle name containing coupe and gearbox as “automatic” has the highest average power of 230 among all the vehicle combinations of vehicle type and gear box.

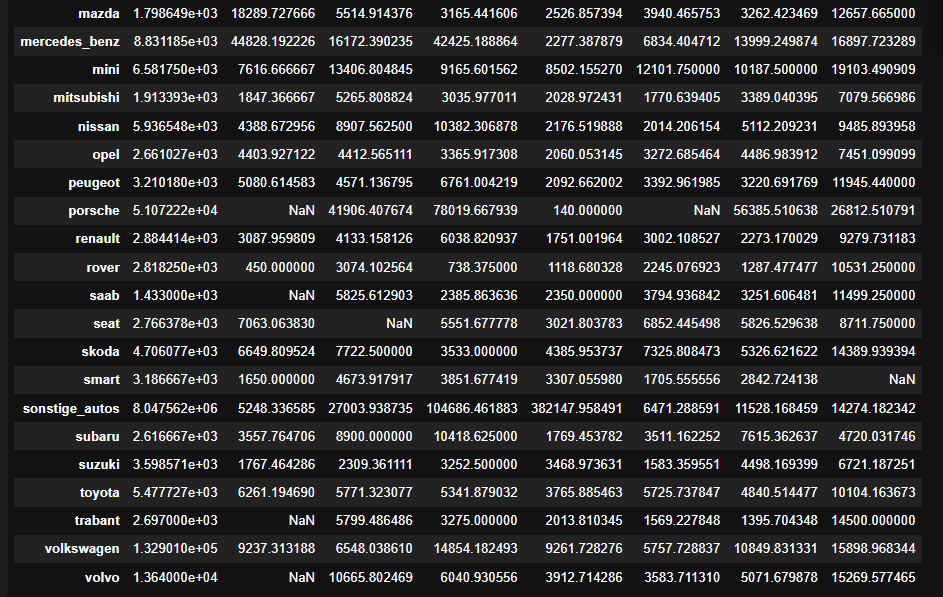
4. What is the Average price of a vehicle by brand as well as vehicle type? Use heatmap to explain this.

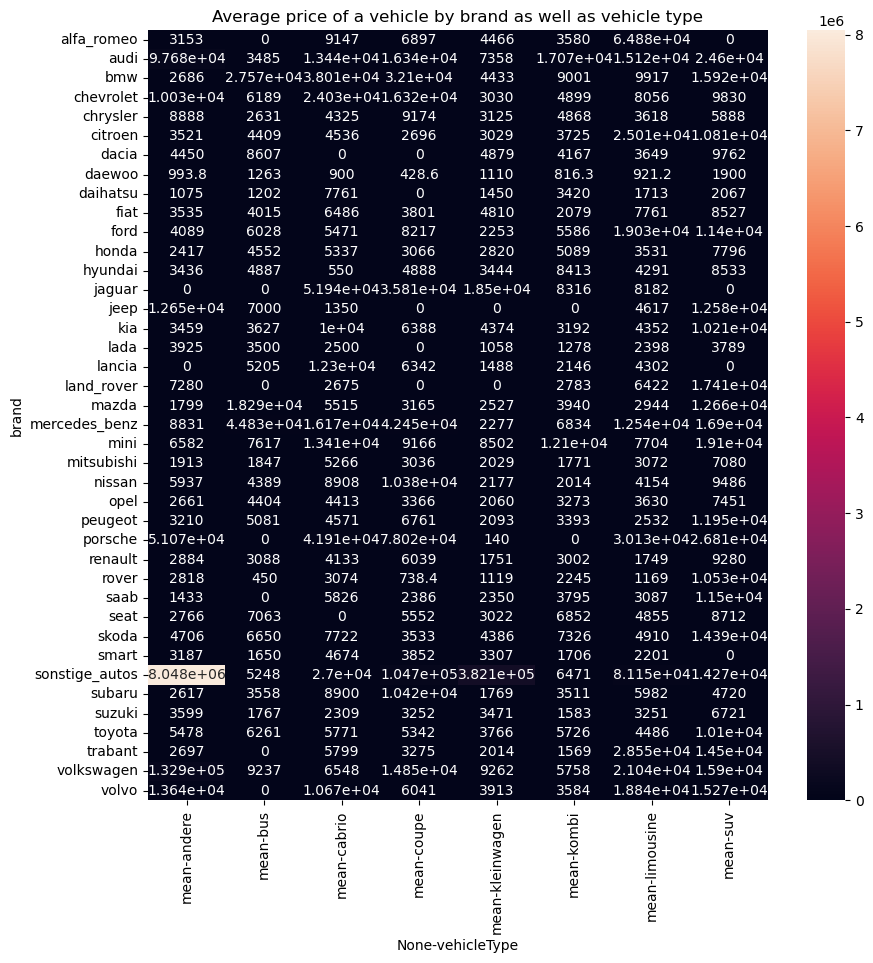
* To show the average price of vehicles based on the vehicle type and brand by using heatmap plot for this.
* To collect the average prices of both the vehicle type and brand at a time.
* We should use the groupby method and in the groupby method we should use aggregate function as mean for the average for the price column.
* The below is code for the data we have to collect.
* “avg\_price=data1.groupby(["vehicleType","brand"])["price"].mean()”
* This will gives us the data as shown below.



* This above picture shows you the average prices of all the brand and vehicle types as well at a time.
* This can also be shown by using the plot heatmap as the code.
* But before plotting the data we have some null values inside the data so we have to fill the null values with zero as follows.







* The heat map shows the brand and vehicle type combined price mean at a time.