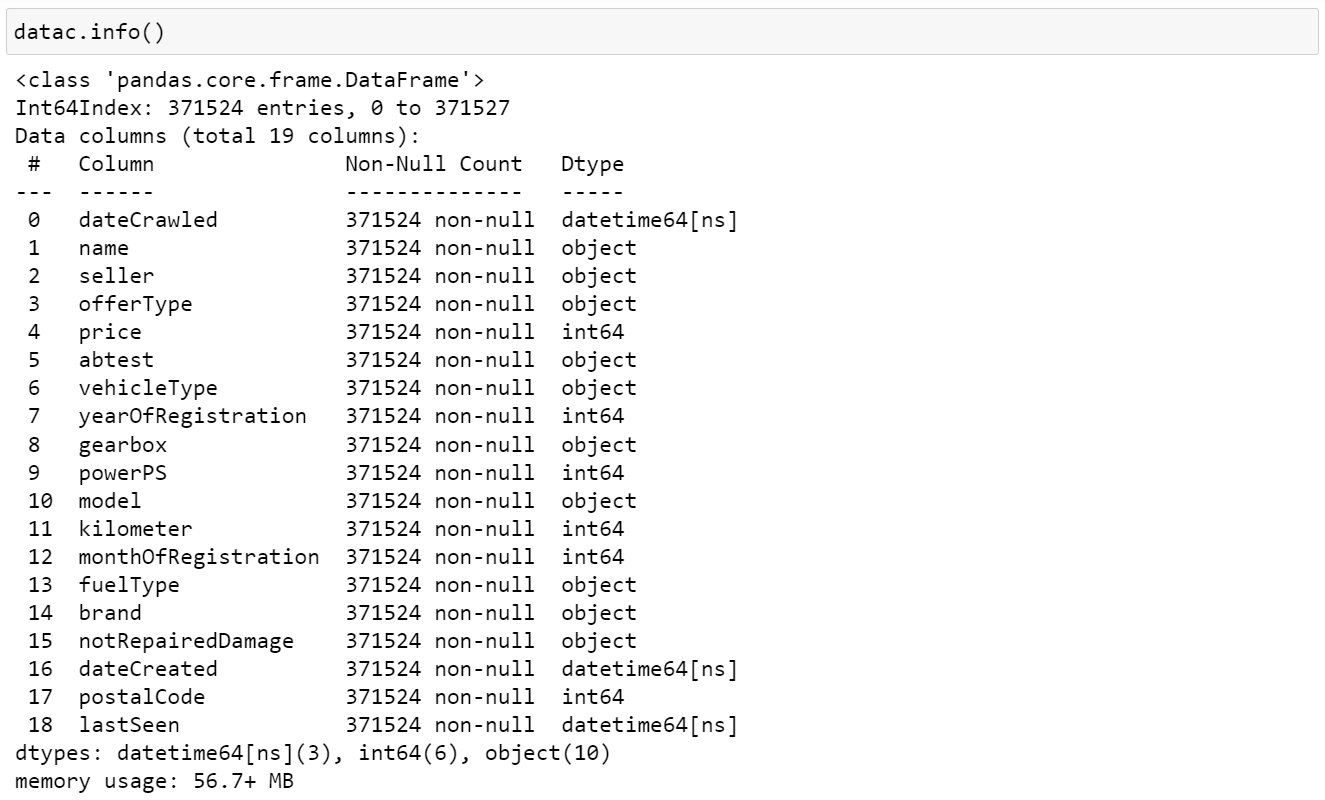
**Report**

**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:**

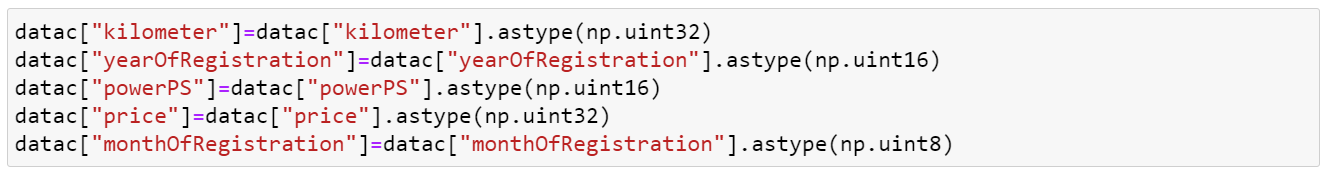
* Firstly, we have to check the memory of data set by using data.info ().

****

* Here, the memory usage of the data set is 56.7+ MB.
* To reduce the memory, we can drop the columns which are not useful for our analysis. Below is the code:

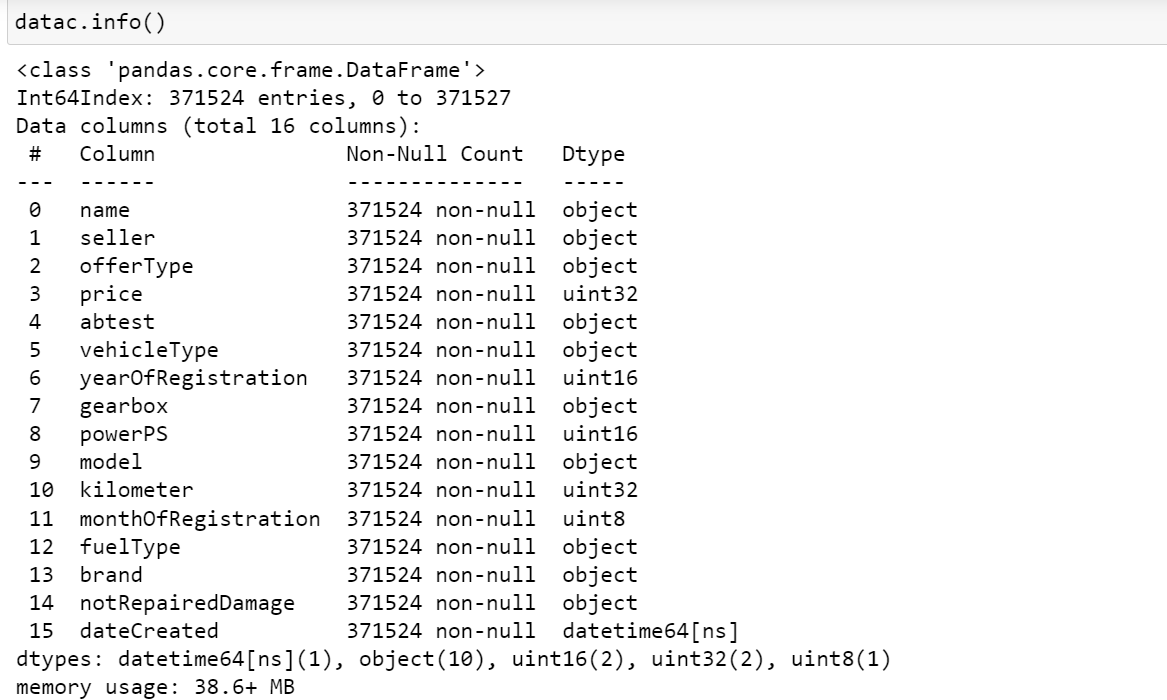
****

* Convert the data types to reduce the memory of the dataset.
* Numerical columns have the data type int64 which is large for certain columns. So, reduce them to int8 (or) int16 (or) int32.
* As the columns price, kilometer and other columns won’t be negative we can reduce the column data type to uint8 (or) uint16 (or) uint32.

****

* Now we can look at the size of memory weather it is reduced or not by using the

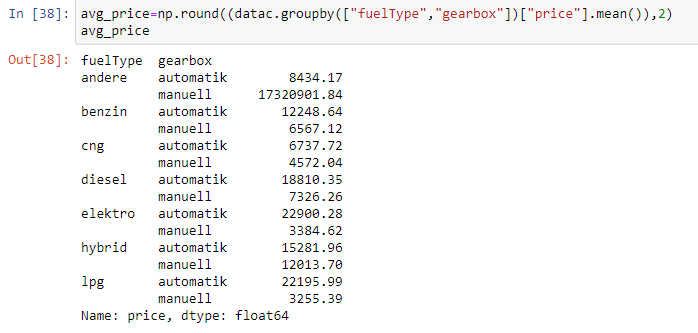
code as:

****

* Here, the memory usage of the data set is reduced from 56.7+ MB to 38.6+ MB.
* We have saved the memory usage of 18+ MB.

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

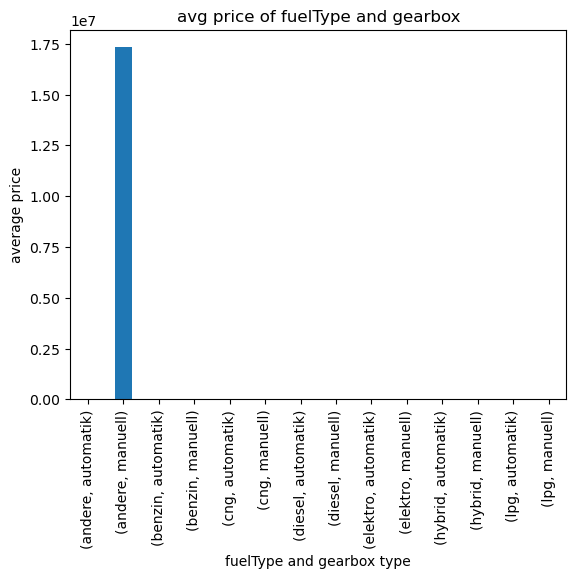
* Average price for vehicles based on the fuel type and gearbox type provides insights about average price for fuel type and gearbox both at a time.
* We can get the average price for both at a time by using groupby method.
* For that apply mean to get the average price for vehicles based on fuel type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



* Fuel type andere and gearbox type manuell has the highest average price 1.199950e+07.
* Fuel type kleinwagen and gearbox type manuell has the lowest average power 3.236602e+03.
* For visualization, we used bar plot to visualize the average price for fuel type and gearbox at a time. The code is



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



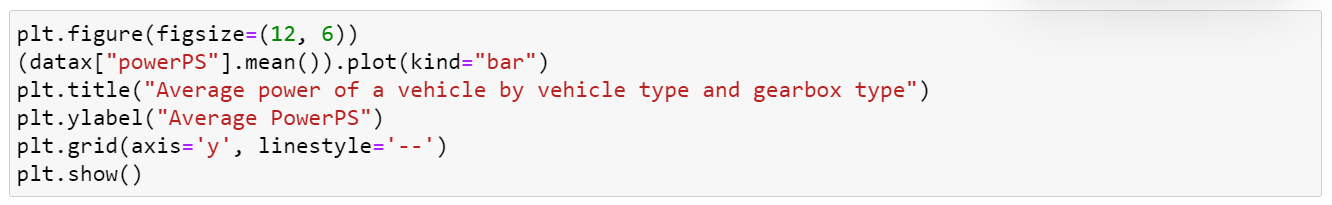
* The x-axis represents the fuel type and gearbox type, and the y-axis will represent the average price.
* Due to the outlier in the vehicle type the average price is affected. So that one bar is highest and all other are lower compared to that.

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

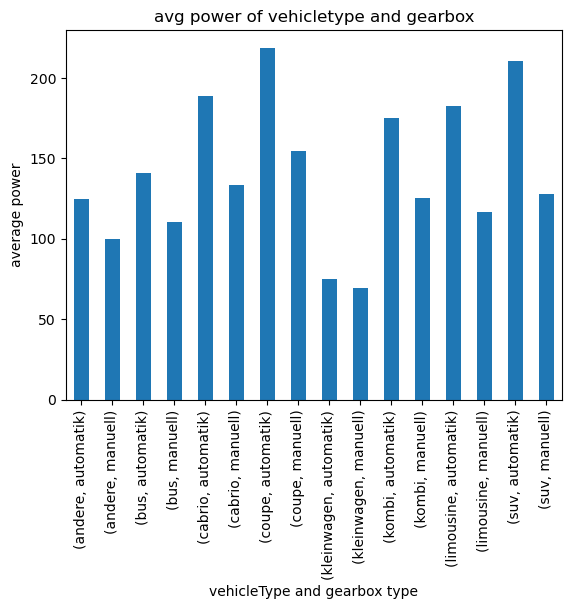
* Average power for vehicles based on the vehicle type and gearbox type provides insights about average price for vehicle type and gearbox both at a time.
* We can get the average power for both at a time by using groupby method.
* For that apply mean to get the average power for vehicles based on vehicle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

### 

* Vehicle type coupe and gearbox type automatik has the highest average power 218.723713.
* Vehicle type kleinwagen gearbox type manuell has the lowest average power 68.242905.
* For visualization, we used bar plot to visualize the average power 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 power.

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

### Average price for vehicles based on the brand and vehicle type provides insights about average price for brand and vehicle type 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 brand. Below is the code:

### 

### Now we have to use crosstab with index as brand and column as vehicle type and values as price and aggregate function as mean.

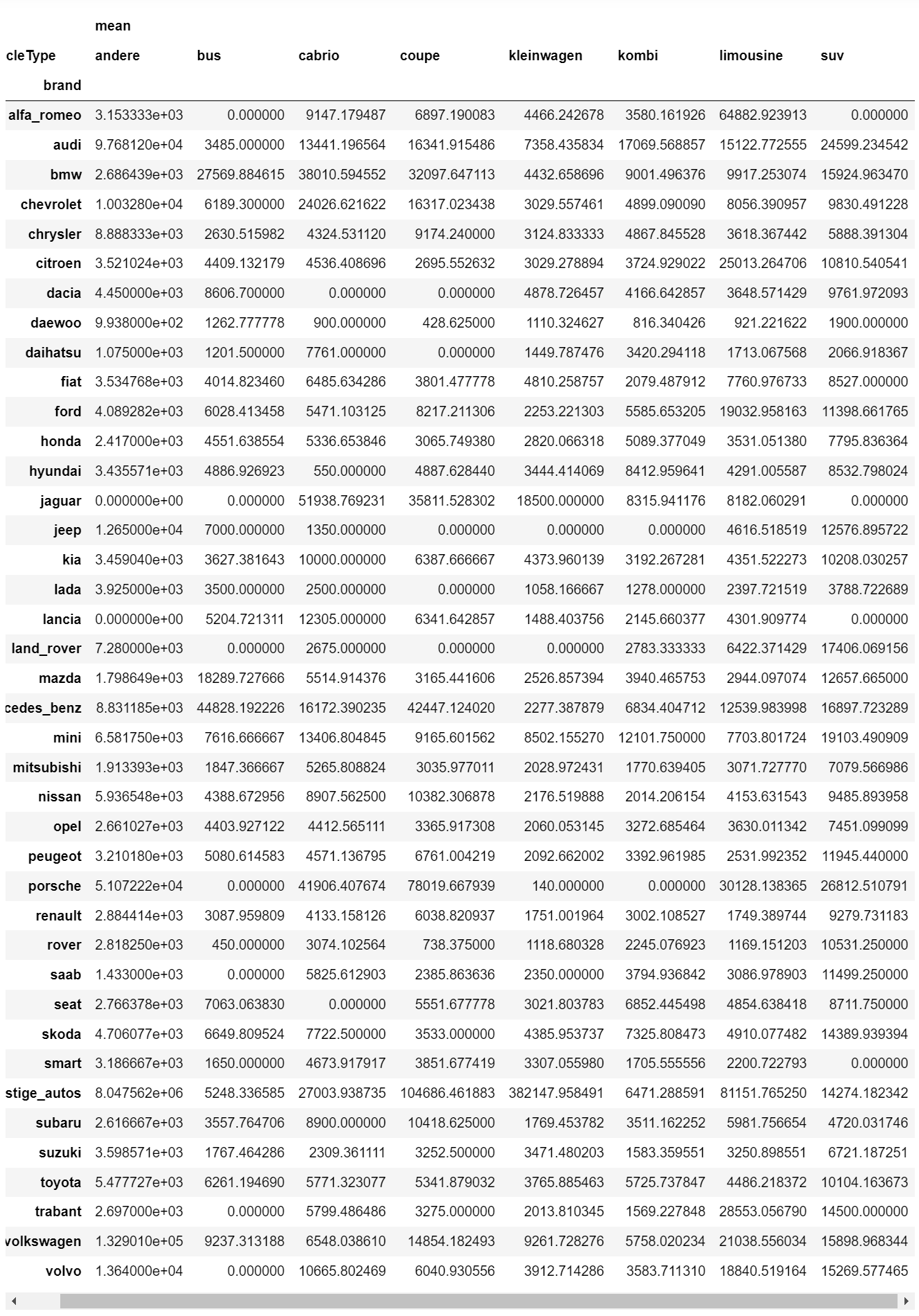
### There are few null values, fill them with 0. The code is:

### average\_price=pd.DataFrame (pd.crosstab(index=datac["brand"],columns=datac["vehicleType"],values=datac["price"],aggfunc=["mean"]))

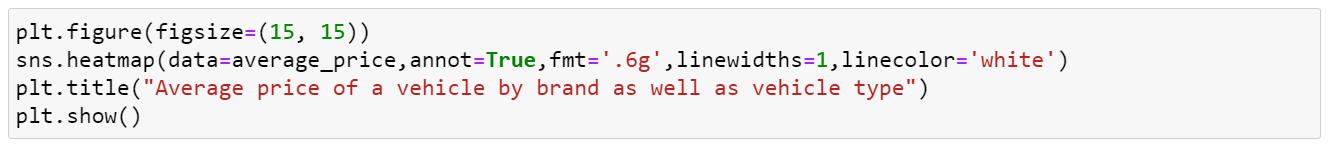
### average\_price.fillna(0,inplace=True)

### average\_price

* It gives the information as below:



* For visualization, we used bar heatmap to visualize the average price for brand and vehicle type at a time. The code is below:



* The heat map visualization for the average price for vehicles based on vehicle type and brand is

