**A.1. perform general data analysis**.

1.First point data set has 371528 data points this data set is about the different vehicles available for sale on eBay.

2.Data set has total 20 columns in that 7 are in 64 data type and remaining 13 has object data type.

3. it takes total 56.7 + MB memory space.

4. After cleaning the entire data set by using fillna() and changing some columns to date time data type Now it has 3 columns which have datetime data type and 7 columns which have int 64 data type and 10 columns which have object data type.

5. By using descriptive statistics I can tell that maximum of prize is 2.147e+09 ,minimum price is 0 and standard deviation is 3.5 8E + 06.

6. In the column power Ps the maximum power is 20,000 and the minimum power is 0 and the standard deviation is 192.139.

7. In the column kilometre maximum kilometres is 15 K and minimum kilometre is zero understand deviation is 3.712.

**A.2. Can you tell me the distribution of vehicles based on year of registration with the help of a plot.**

A graph of distribution of vehicles

Description automatically generated

1. For this question I used the cleaned dataset and I imported that using the PD dot read CSV function. to plots the distribution of vehicles based on year of registration I used histogram.

2. From that plot I can tell that in the year 2000 more vehicles are produced nearly 25 kg vehicles are produced in the year 2000.

3. N in the year 1961 to 64 very less amount of vehicles are produced.

**A.3. Create a plot day stood on the variation of the price range by the vehicle type.**

1. to plot the variation of the price range by the vehicle type I used bar plot , on axis I used vehicle type and on y axis I used price.

2. I can summarise but with the help of the plot andere brand has more price and SUV has very less price.

**A.4. Find out total count of vehicles by type available on eBay for sale as well as create a visualization for the client.**

1. To find the total count of vehicles by type I used group by function with the help of the group by function I got a new data and I named the data as da.

2.to plot that I used bar plot, from the data which I got after applying the group by function I take vehicle type on X axis and count on y-axis.

3. From the plot I can tell that Limousine waiter tip has more number of vehicles with the count of 95894 and andere vehicle type has least number of vehicles with the count of 3357.

**A.5. Is there any relationship between dollar price and kilometre explain with appropriate analysis.**

1. To find the relationship between price and the kilometre I used Pearson correlation coefficient.

2. With the help of Pearson correlation coefficient I can tell that the relationship between price and the kilometre is as the price increases the kilometre decreases because the Pearson correlation coefficient value is -0.00063538378, and P value is 0.698545355.

**B.1. Can you tell me number of vehicles by brand available on eBay for sale with the help of visualization.**

1. To find the total count of vehicles by type I used group by function from that group by data I have created a new data frame named data and then I plot a bar plot and that. And on the bar plot on axis I take brand and y axis I take count.

2. From the plot I can tell that Volkswagen brand has more number of vehicle available on eBay for sale and account is 79640.

3. The brand lada has less number of vehicles available on eBay for sale with the count of 225.

**B.2. What is the average price for vehicles based on the type of vehicle as well as on the type of gearbox.**

1. For average price for vehicle based on vehicle type I used group by function. From that group by data I have created a new data frame named data vehicles and then I plot a bar plot on that. In the plot on X axis I take vehicle type and on y axis I take mean of prices.

2. From the plot I can tell that average price for vehicles based on vehicle type is high for Andere vehicle with the 677329.

3. For the vehicle Klein wagon has less price with 5695.

4. For average price of vehicles based on gearbox type I used group by function from that group by data I have created a new data frame named data gearbox by using bar plot have created a visualization for that.

5. In the plot on X axis I used gearbox type and on Y axis I used mean of prices.

6. From the plot I can tell that the vehicles which dont have many type has more average price when compared to others with 44019, and d automatic has less average price weight 15146.

**B.3. What is the marginal probability of private seller.**

1. to find the marginal probability of private seller the formula will be the count of private seller/total count of the seller.

2. By using above formula I got marginal probability value as 0.999999.

**C.1. The memory usage of the data is around 6.1 MB how can we reduce the memory usage of the data set.**

1. Reduce the memory usage of the dataset I have changed the data type of some columns which are in 64 I changed them into int 32 and I have deleted some unnecessary columns like date crawled ,not repaired damage, last seen and number of pictures. Because of those manipulation the memory usage of the data set has decreased.

**C.2. What is the average price of vehicle by fuel type and gearbox type and give a plot.**

1. To Create a visualization for above question I have used bar plot, in the bar plot on X axis I have used fuel type and on y axis I have used average price, and I have group by them with the gearbox type.

2. With the help of the plot I can tell that and there a fuel type which is manual gearbox type has more average price when compared to all others.

**C.3. What is the average power of a vehicle by vehicle type and gearbox type give a plot.**

1. To create a visualization for above question I have used bar plot. in the bar plot on X axis I have used vehicle type and on y axis I have used average price, and I grouped them with the gearbox type.

2. With the help of the plot I can say that coupe vehicle type has more average price when compared to other vehicle types and in that coupe vehicle type with automatic gearbox type has more average price than other 2 gearbox types.

**C.4. What is the average price of a vehicle by brand as well as vehicle type use heat map to explain this.**

1. to plot a heat map visualization first I have created a cross table for mean of prices based on brand and vehicle type.

2. With the help of that cross table I have created a heat map where on X axis we take vehicle type and on y axis we take brand.