

Advanced-Data Analysis on Germany Used Cars

Prepared By: Sathish Vanga

Date: 23-12-2023

Introduction:

The "Exploratory Analysis of Used Car Listings in Germany" is a comprehensive investigation into a dataset capturing the intricacies of the German used car market as reflected on eBay.

Data Set Overview:

dateCrawled: Date and time when the car listing was first crawled.

name: Name of the car.

seller: Indicates whether the seller is private or commercial.

offerType: Indicates whether the listing is an offer or a request.

price: Price of the car.

abtest: A/B test category.

vehicleType: Type of vehicle (e.g., coupe, SUV, etc.).

yearOfRegistration: Year the car was first registered.

gearbox: Type of gearbox (manual, automatic, etc.).

powerPS: Power of the car in PS.

model: Model of the car.

kilometer: The distance the car has been driven in kilometers.

monthOfRegistration: Month the car was registered.

fuelType: Type of fuel the car uses.

brand: Brand of the car.

notRepairedDamage: Indicates if the car has damage that is not yet repaired.

dateCreated: Date when the car listing was created.

nrOfPictures: Number of pictures in the ad.

postalCode: Postal code location of the vehicle.

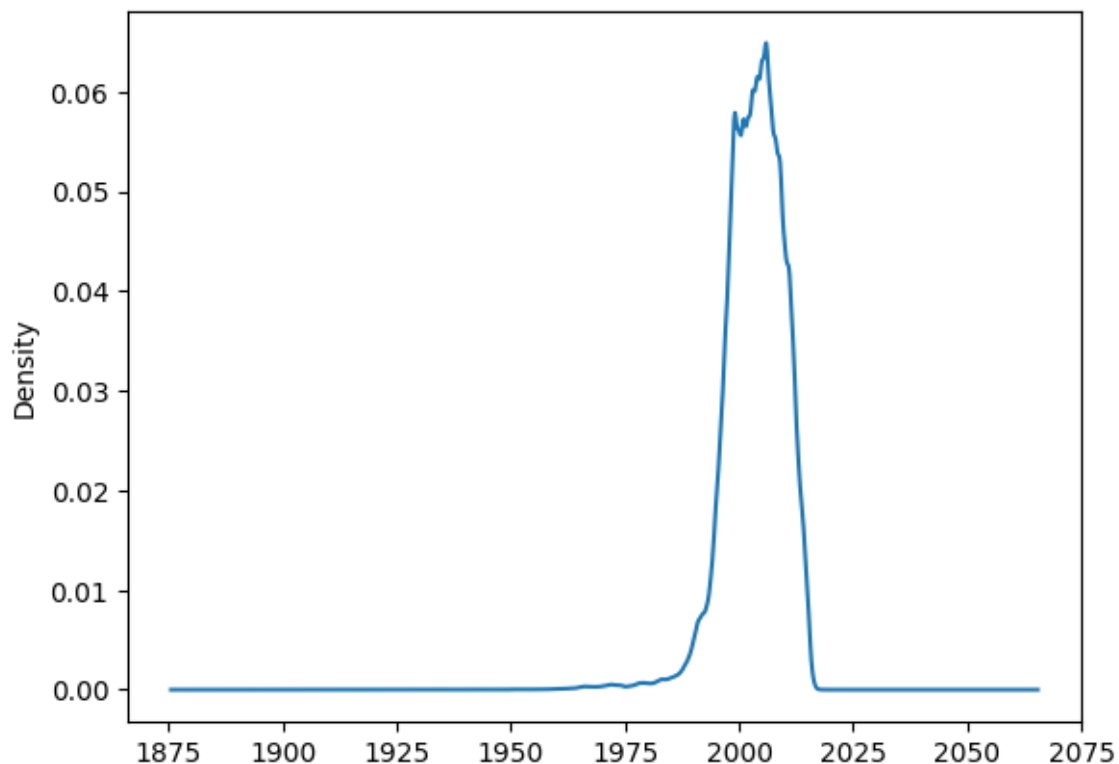
lastSeen: Date and time the crawler saw this ad last online.

Analysis 1

1.1 Perform general Data analysis

- Identify Missing Values
- Handle Missing Values
- Identify and Handle Outliers
- Check for Duplicates
- DataType Conversion
- Identify and Handle Structural Errors

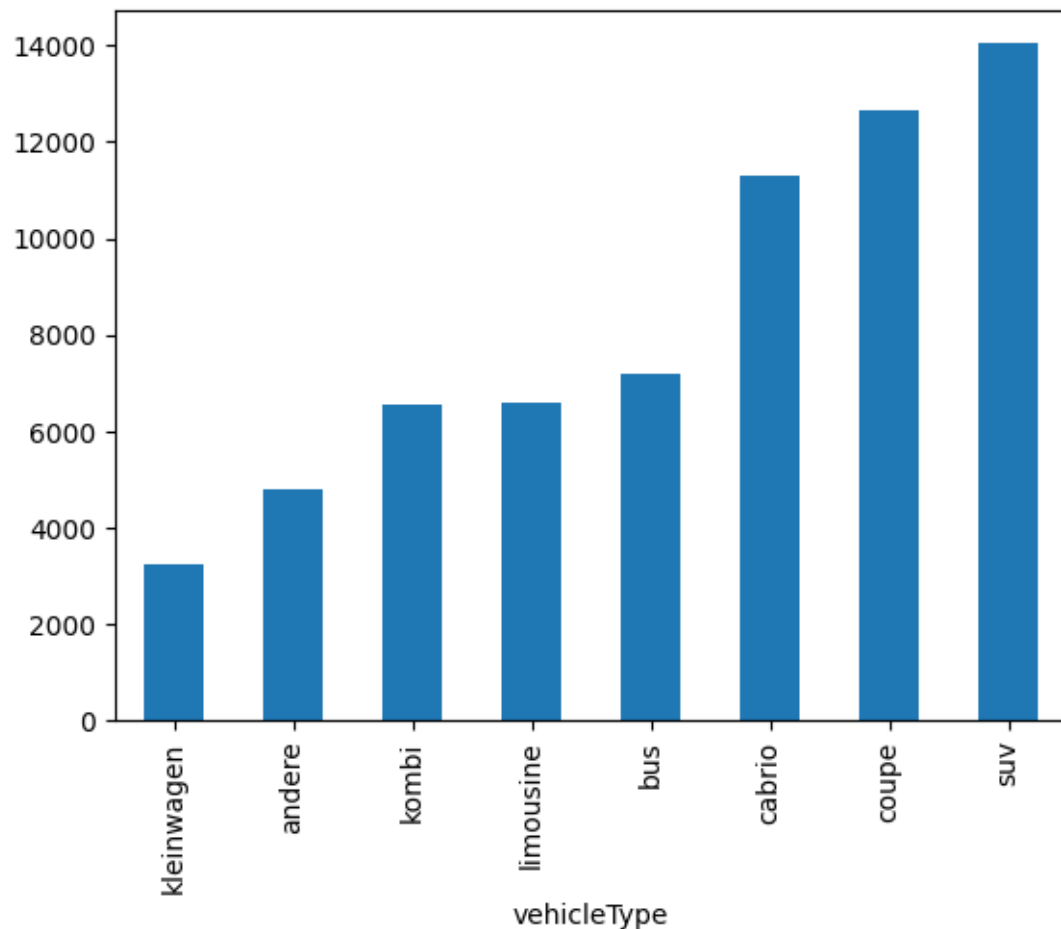
1.2 Can you tell me the Distribution of Vehicles based on Year of Registration with the help of a plot



Conclusion:

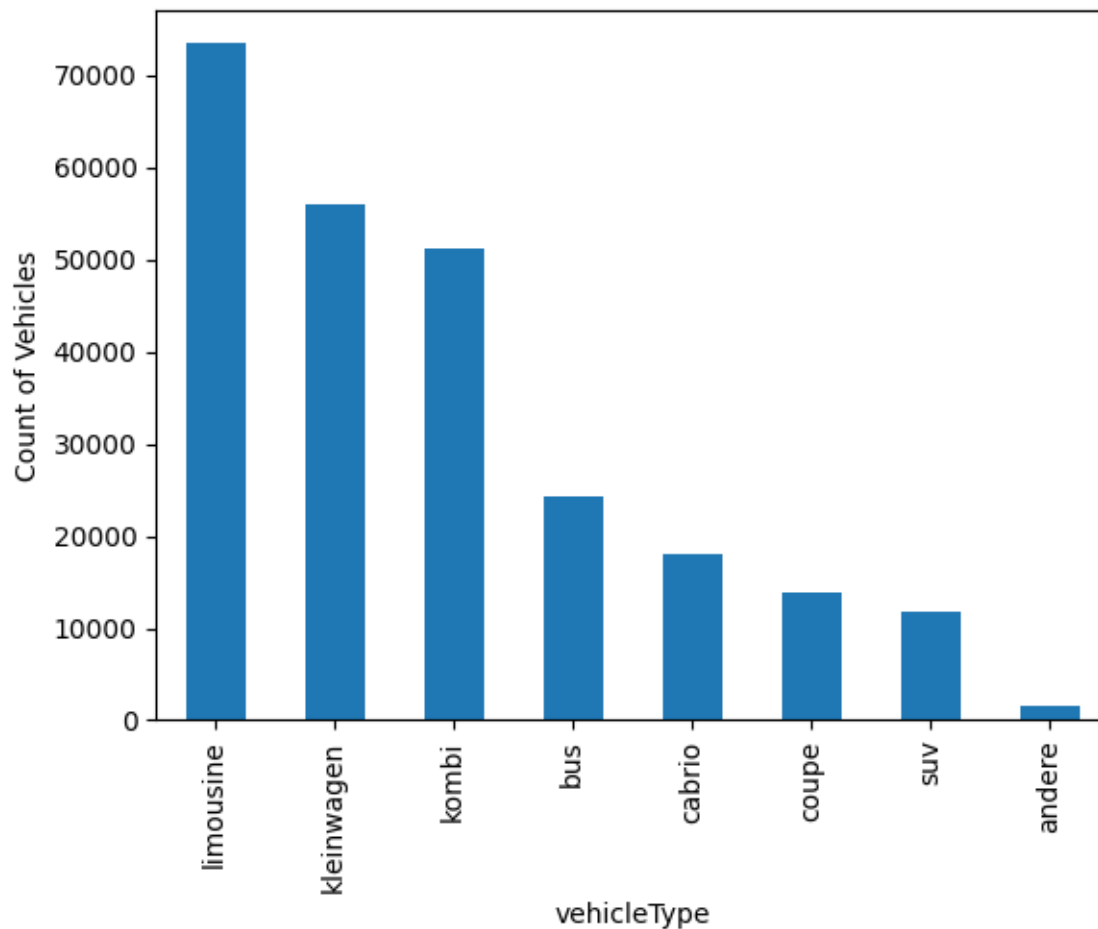
- yearofRegistration is **Left Skewed Distribution** because more data points on the right side

1.3 Create a plot based on the Variation of the price range by the vehicle type



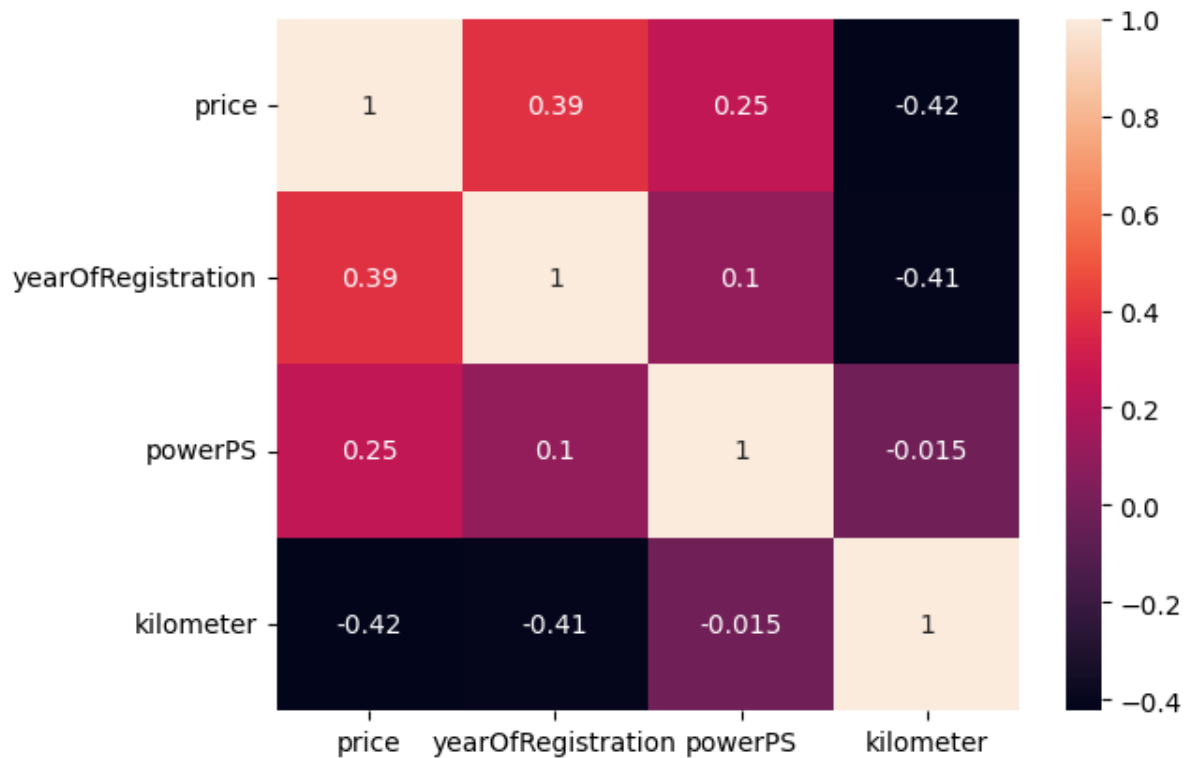
The price range varies based on the vehicle type, with SUVs having the highest price range, coupes coming in at the second highest, and Kleinwagens having the lowest price range.

1.4 Find out the Total count of vehicles by type available on eBay for sale. As well as create a visualization for the client



The Limousine vehicle type has the highest count on eBay for sale, followed by Kleinwagen, and the lowest count is for Andere vehicle types

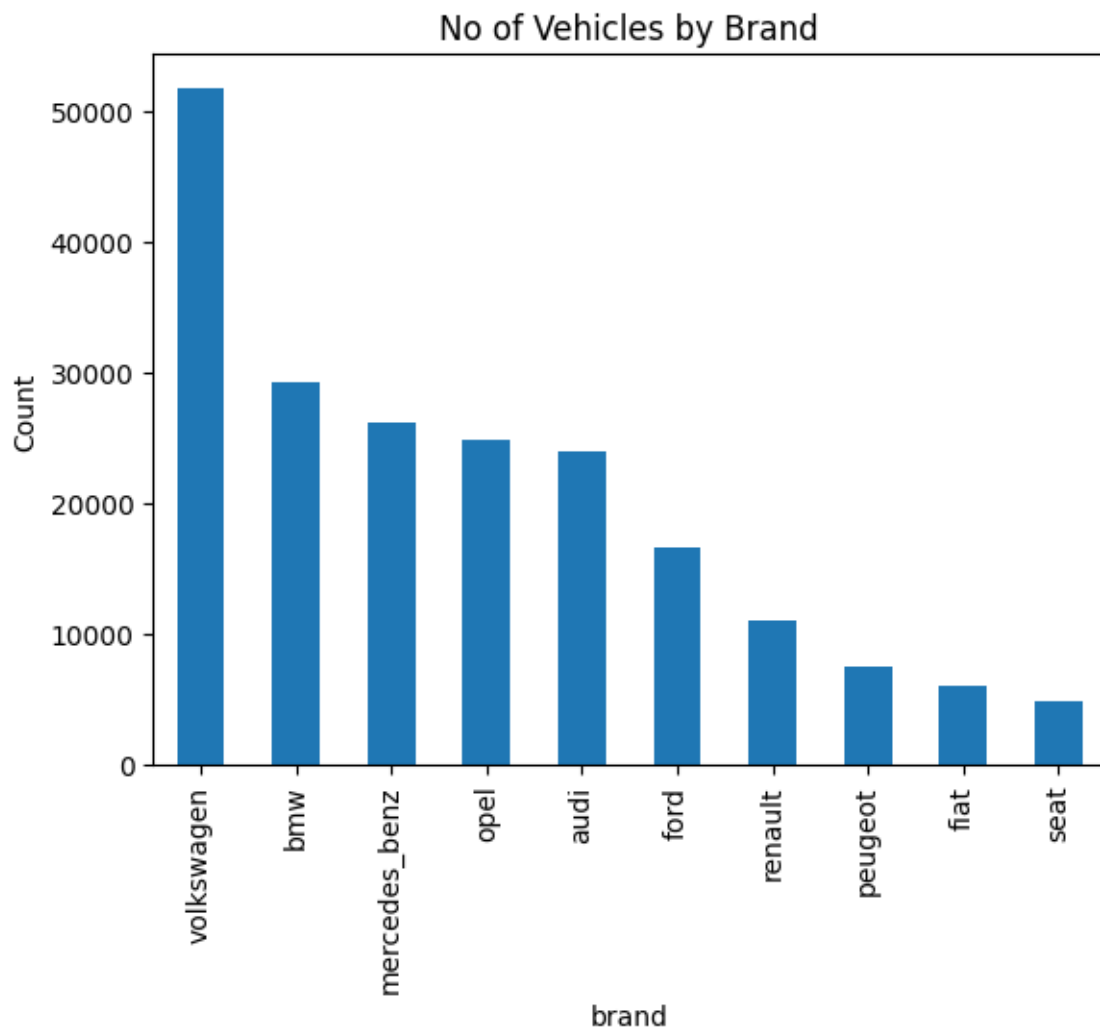
1.5 Is there any relationship between dollar price and kilo meter?
(Explain with appropriate analysis)



The relationship between price and kilometer is -0.4. i.e **Negative Weak Relationship**

Analysis 2

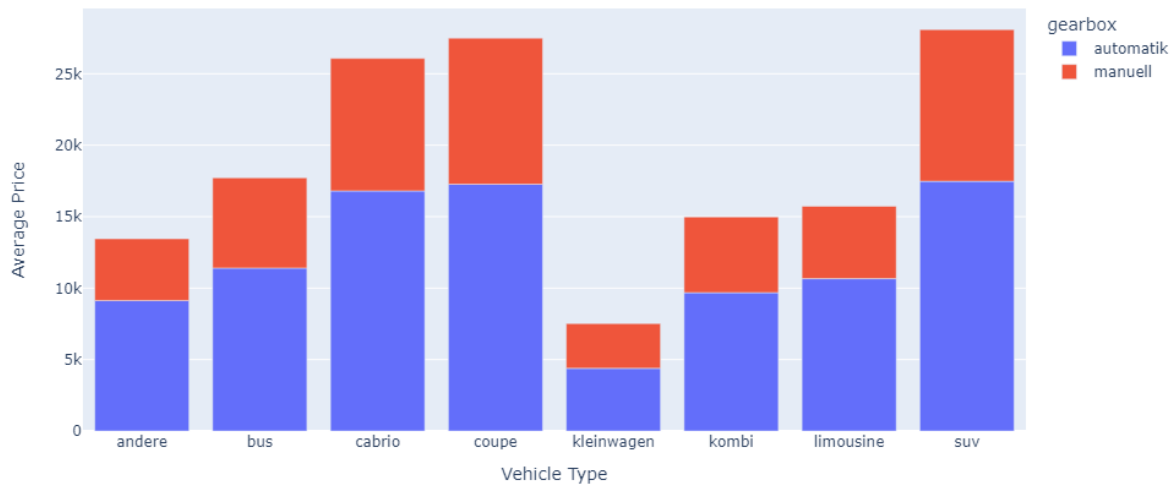
2.1 Can you tell me No of Vehicles by Brand Available on ebay for sale with the help of visualization



Volkswagen has the highest sales among the available vehicles on eBay.

2.2 What is the average price for vehicles based on the type of vehicle and the type of gearbox? Explain to me with both numerical and visualization analysis.

Average Price Based on Vehicle Type and Gearbox



The Average price for vehicles varies based on the type of vehicle and the type of gearbox, with SUVs having the highest average price

2.3 What is the Marginal Probability of a Private Seller?

The marginal Probability of the Private Seller is 0.999992002303366

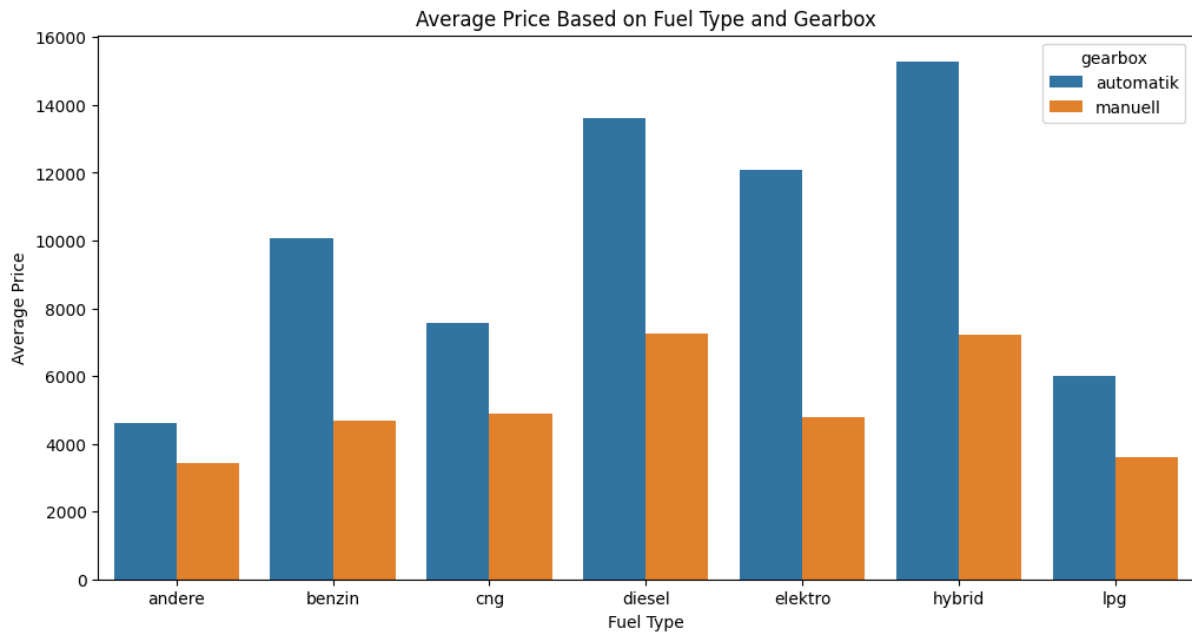
Analysis 3

3.1 The memory usage of the data is around 6.1 mb. How can we reduce the memory usage of the dataset?

Here we can reduce memory usage by dropping the data:

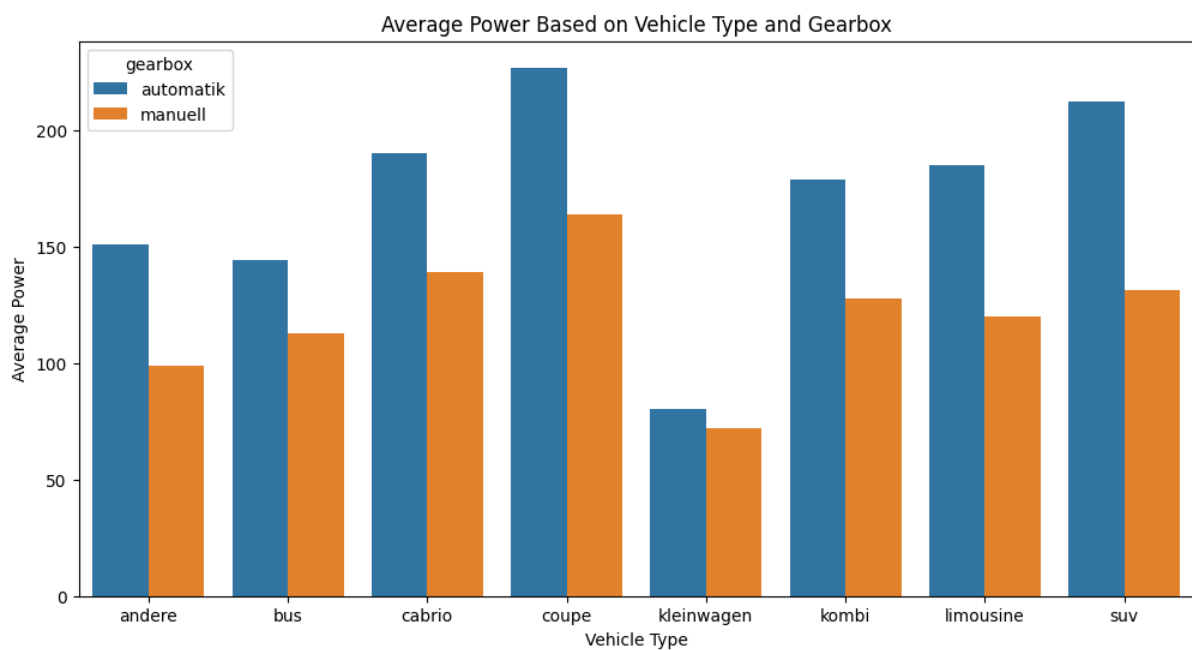
- Changing the Datatype of the data as int to float and float to int.
- Optimize textdata.
- Eliminate columns not crucial for analysis to reduce data size.

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



The average price of vehicles varies based on fuel type and gearbox type, with hybrid fuel type having the highest average price, followed by diesel, and the lowest being for vehicles with andere(other) fuel types.

3.3 What is the average power of a vehicle by vehicle type and gearbox type? Give a plot.



The average power of vehicles varies based on their type and gearbox type, with coupe vehicles having the highest average power, followed by SUVs, and Kleinwagen having the lowest average power.

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

