# Evaluation & Prediction on the second-hand car market in Romania

No. 3, February 2024

# EvaCar

The primary objective of this study is to highlight trends in the Romanian second-hand car market. The first specific goal is to conduct an analytical, quantitative, and qualitative study on existing data, providing an overview of the big picture for a comprehensive understanding of the current market situation, without additional analyses. The second specific goal is to develop a tool that predicts future trends in the market and car prices based on specific characteristics. To accomplish these specific goals, this study will address a series of key business questions.

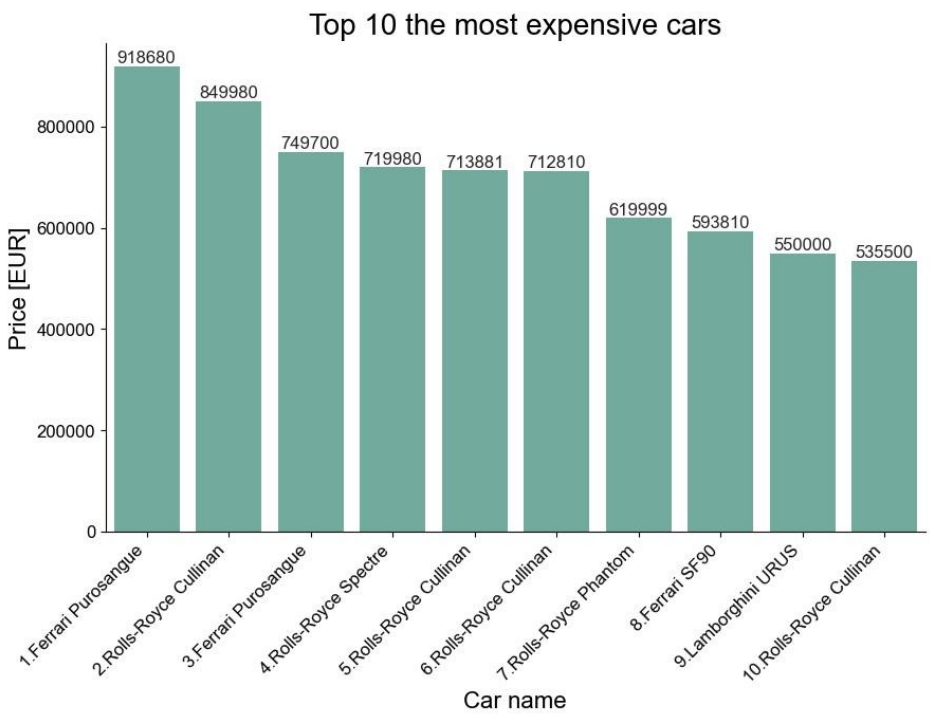
The following business questions are addressed in this report:
- How many records were utilized in the analysis?
- What is the most expensive/cheapest car?
- Which manufacturer/model has the highest/lowest number of cars?
- Which manufacturer/model has the highest/lowest average price?
- In which county/city are the most cars being sold?
- Which type of fuel has the highest/lowest sales?
- What are the largest/smallest engine capacity, mileage, and power figures?
- How many cars were manufactured in the current year?
- How are features correlated with each other?
- What is the predicted price of the car based on specific characteristics?

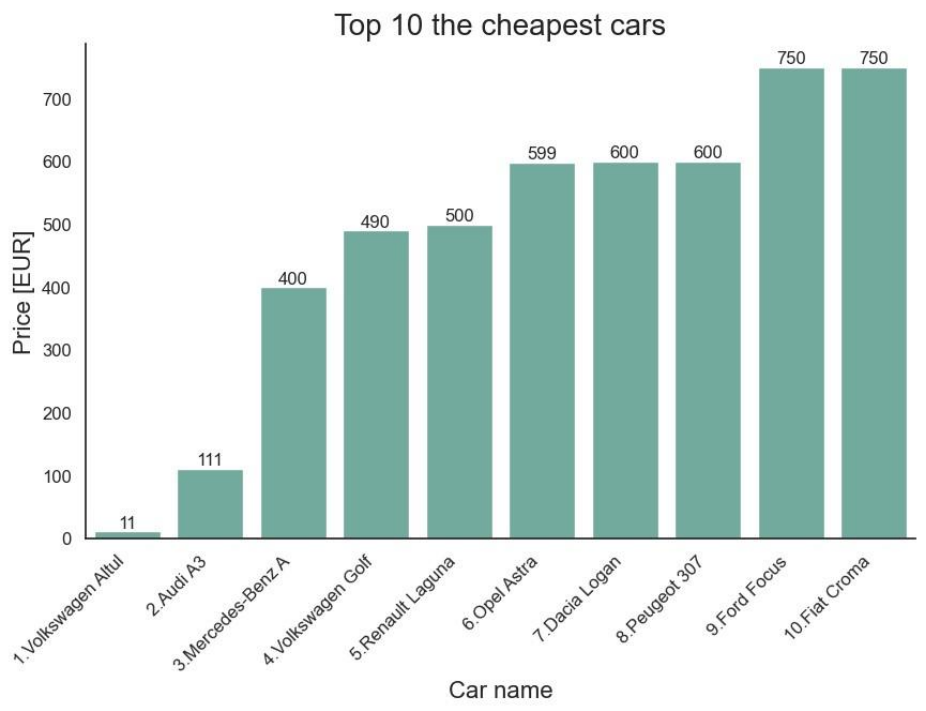
The dataset used in the study was obtained:
- from one single source, car seller platforms - Autovit.
- using web scraping process.
- data available on **10th February 2024** from the car seller web page.

```
RangeIndex: 38136 entries, 0 to 38135
Data columns (total 10 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   manufacturer  38136 non-null  object
 1   model         38136 non-null  object
 2   mileage       38136 non-null  int32
 3   capacity      38136 non-null  int32
 4   power         38136 non-null  int32
 5   year          38136 non-null  int32
 6   fuel          38136 non-null  object
 7   county        38136 non-null  object
 8   city          38136 non-null  object
 9   price         38136 non-null  int32
```

The analysis was conducted on a dataset comprising 38.136 records, encompassing 10 columns—5 numerical and 5 categorical.

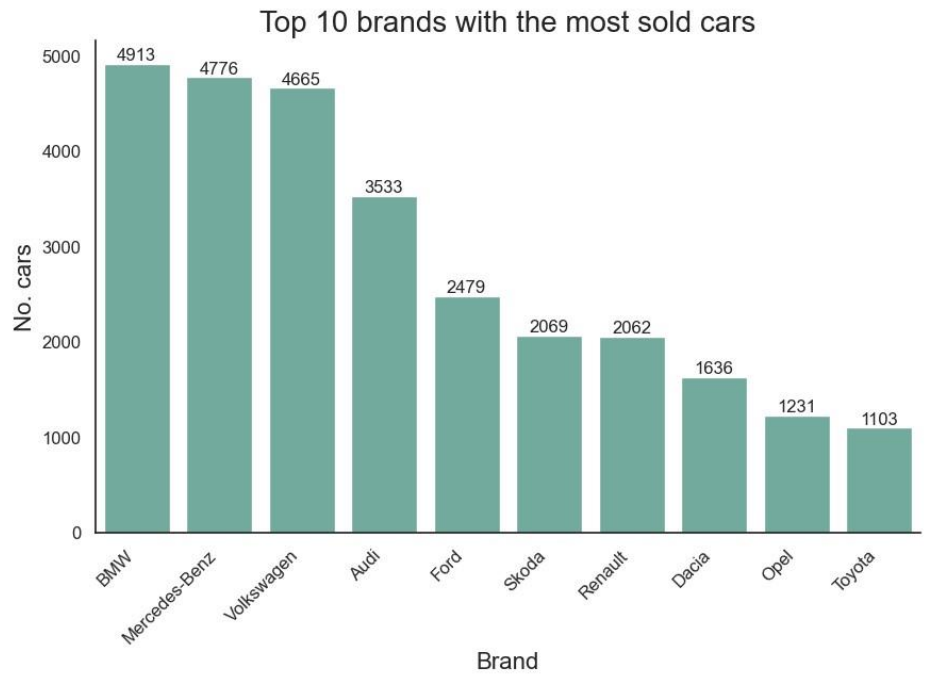
In January 2024, the most expensive car in the analysis was the Ferrari Purosangue, priced at a staggering 918.680 EUR. Notably, the top 10 most expensive cars were all within the vicinity of 600.000 Euros.
On the other end of the spectrum, the least expensive car in the analysis was the Audi A3, available at a modest cost of 111 EUR. The first 10 least expensive cars were all priced around 650 Euros, reflecting the affordable range in this segment.

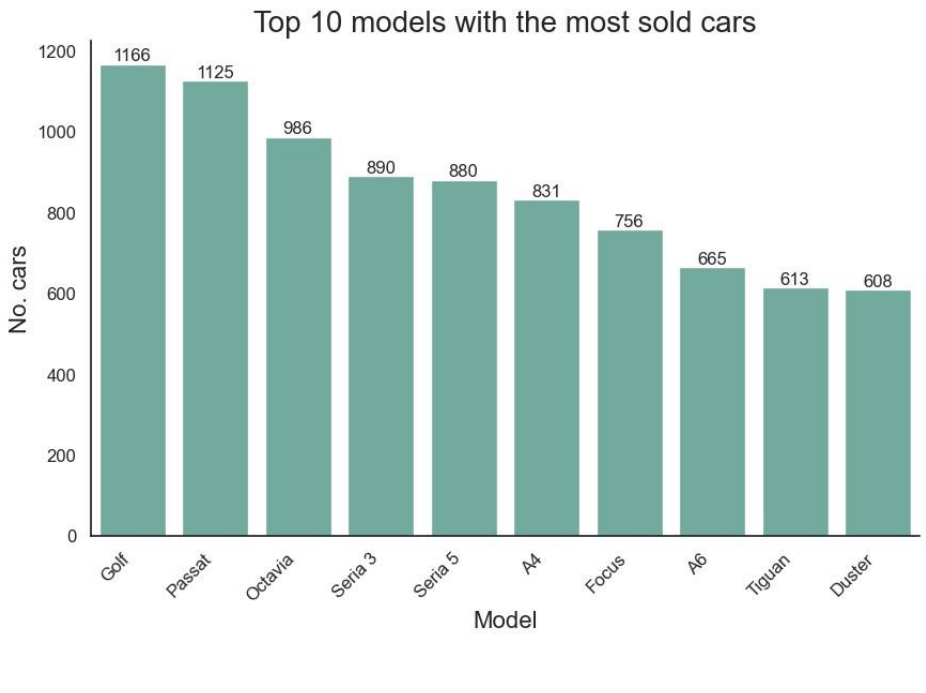


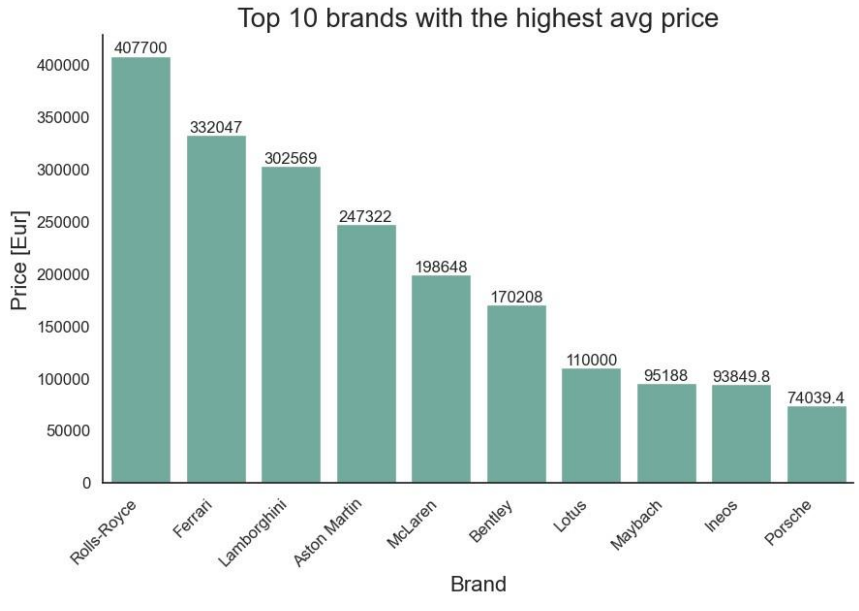

In terms of quantity, BMW emerges as the leading manufacturer in the dataset, boasting an impressive 4.913 cars. Among the various models, the one that dominates in sheer numbers is the Golf, accounting for a substantial 1.166 cars.

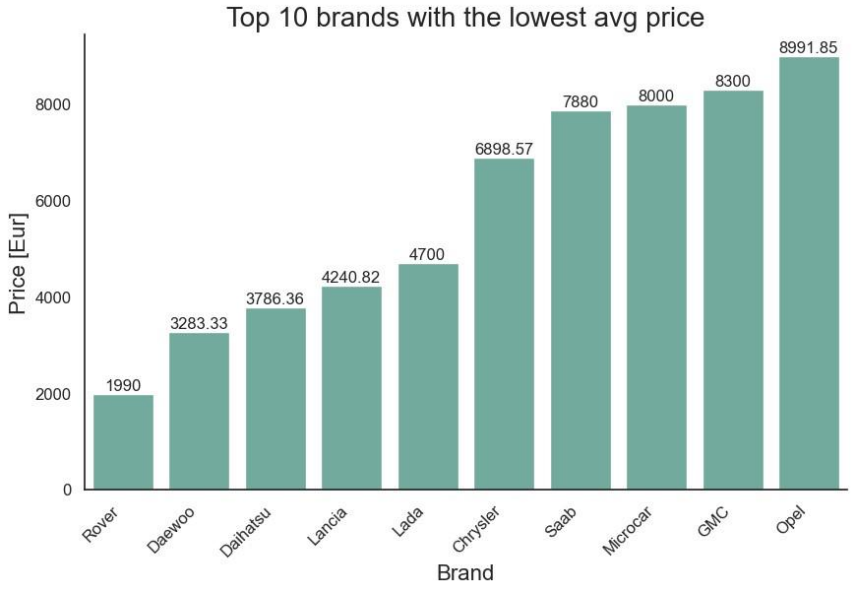
Conversely, Lada holds the distinction of having the fewest cars among manufacturers, with a mere single car. On the model level, the title of the least represented goes to the CSX, also with just one car, reflecting a unique and less common presence in the dataset.



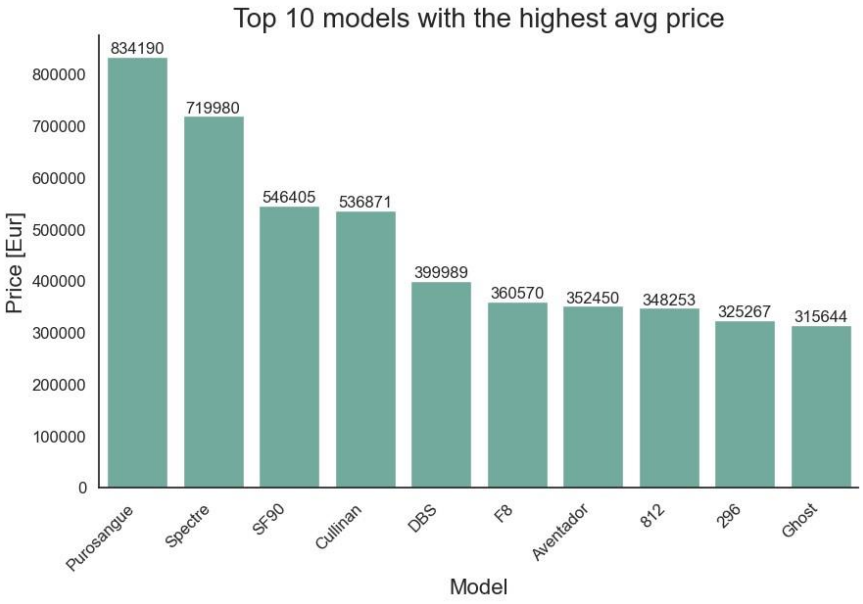



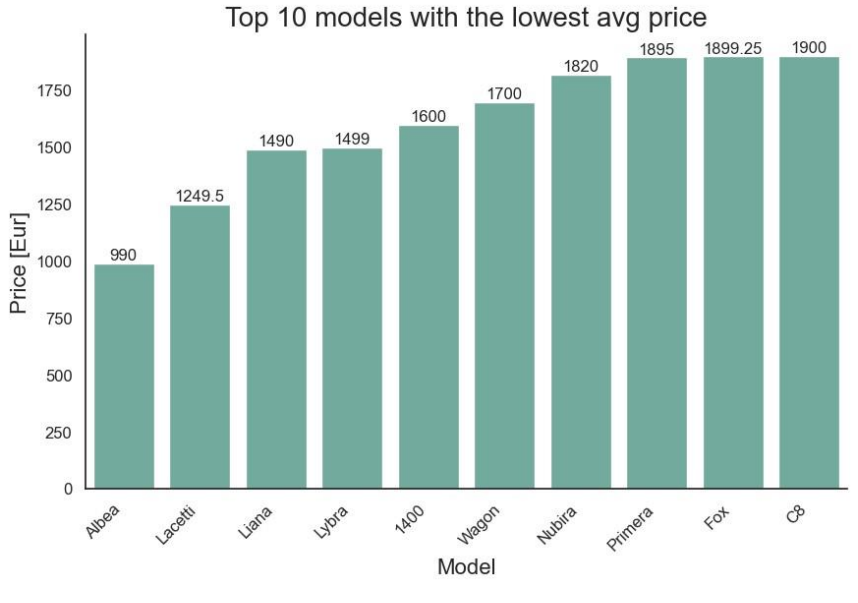
When considering average prices, Rolls-Royce stands out as the most luxurious manufacturer, commanding an impressive average price of 407.700 EUR per car. On the other end of the spectrum, Rover emerges as the most budget-friendly manufacturer, with an average price of 1.990 EUR per car, making it an accessible choice for cost-conscious consumers.



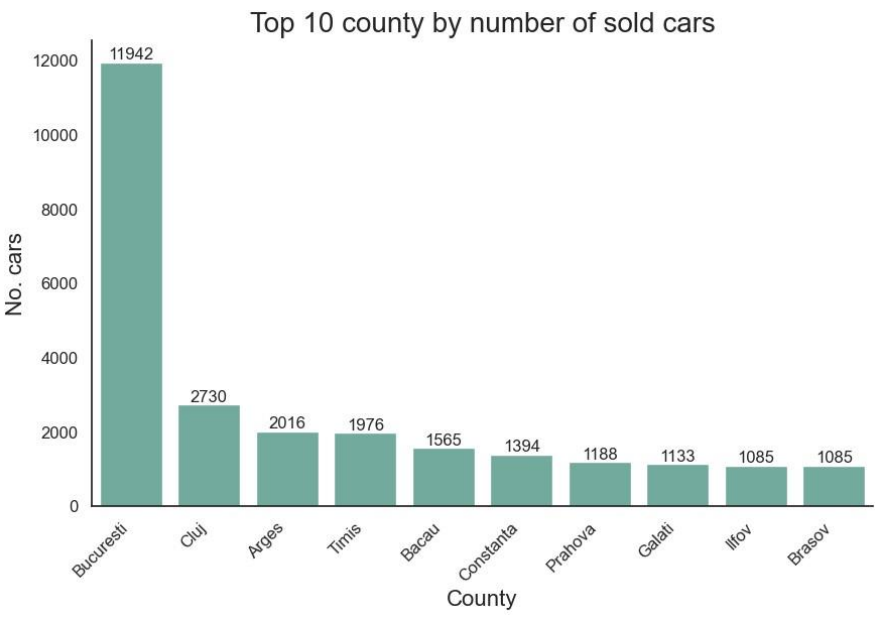



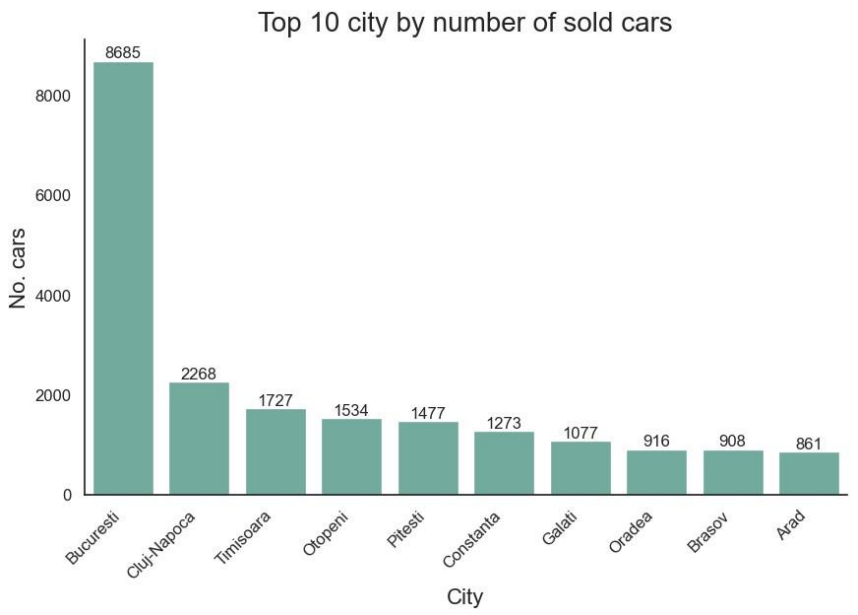

When it comes to average model prices, the Purosangue claims the title of the most opulent, boasting an average price of 834.190 EUR per car. On the flip side, the Albea takes the crown as the most economical model, with an average price of a modest 990 EUR per car, presenting an affordable option for budget-conscious buyers.



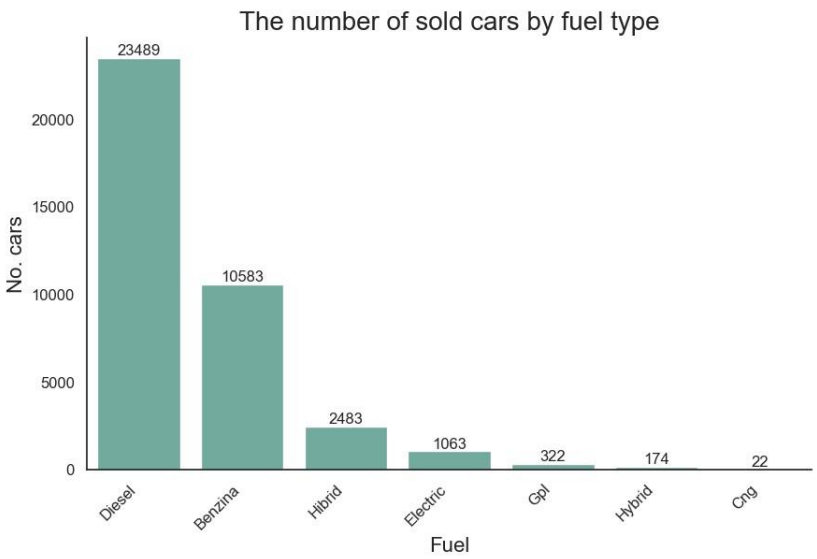

Bucuresti - Ilfov county takes the lead in terms of the highest number of cars, boasting a total of 11.942 vehicles. Drilling down further, within Bucuresti city alone, there are 8.685 cars, making it the city with the most significant concentration of vehicles in the dataset.

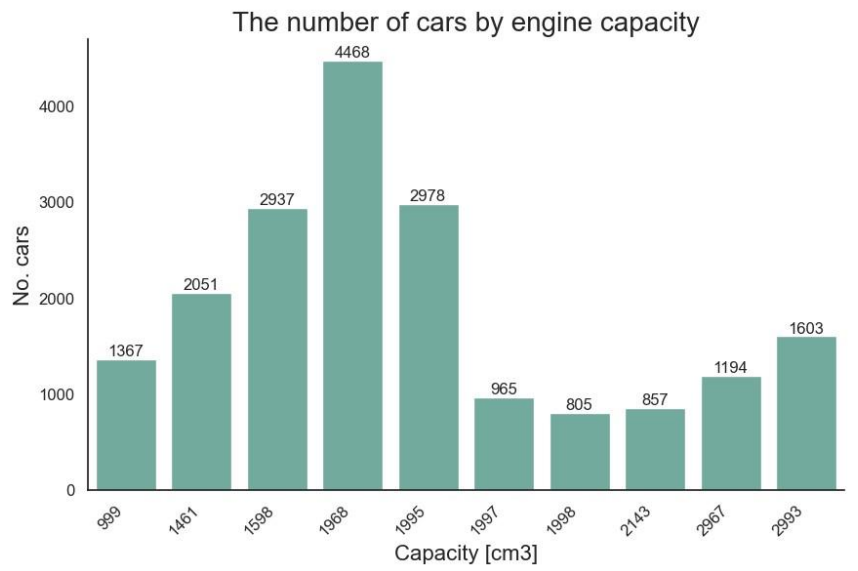


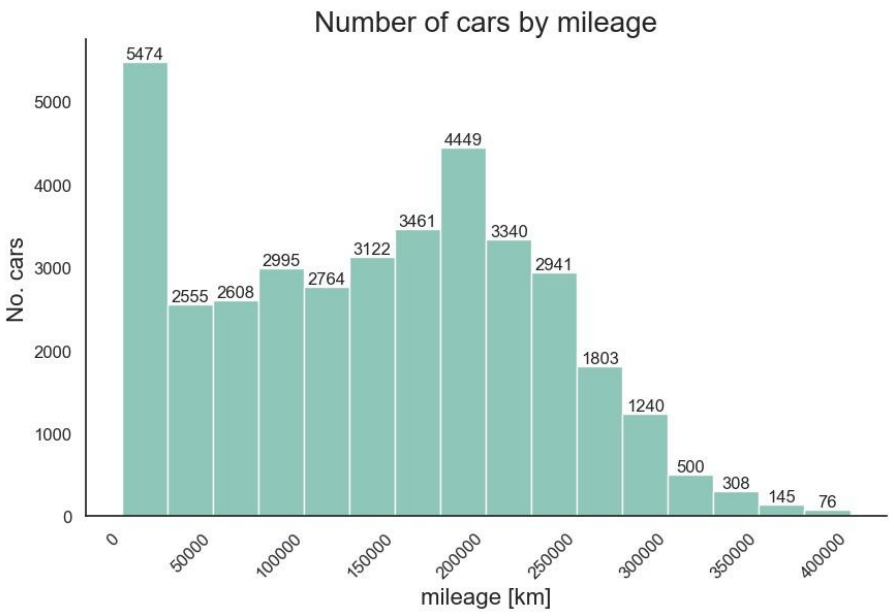

Diesel fuel takes the lead as the most preferred choice, powering a total of 23.485 cars. On the opposite end of the spectrum, the least common choice is Benzina + CNG fuel, with a mere 22 cars in the dataset.

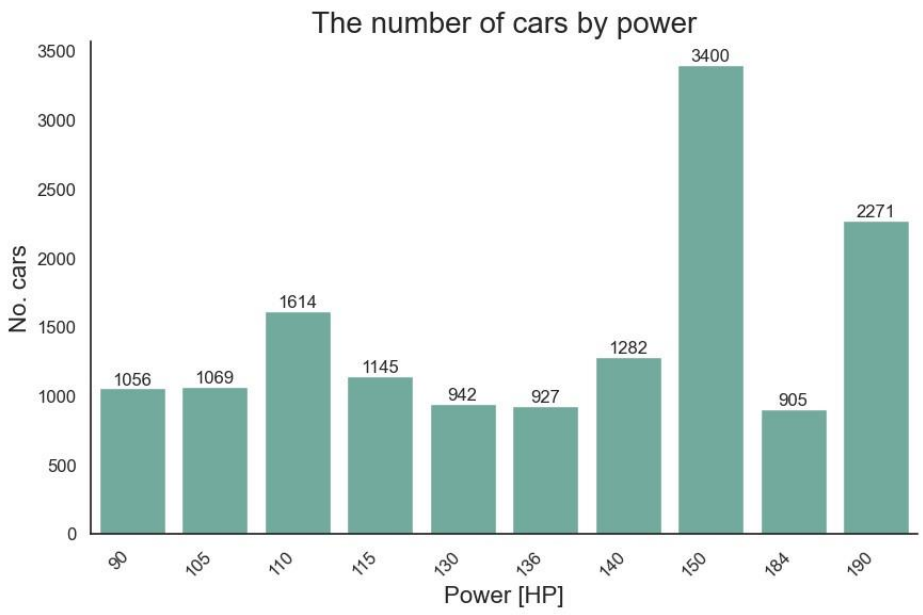
When it comes to engine capacity, the Lincoln Continental boasts an impressive maximum of 7.500 $cm^3$, showcasing its powerful performance. In contrast, the Opel Vectra holds the record for the smallest engine capacity, measuring at a nimble 400 $cm^3$, highlighting its efficiency and compact design.



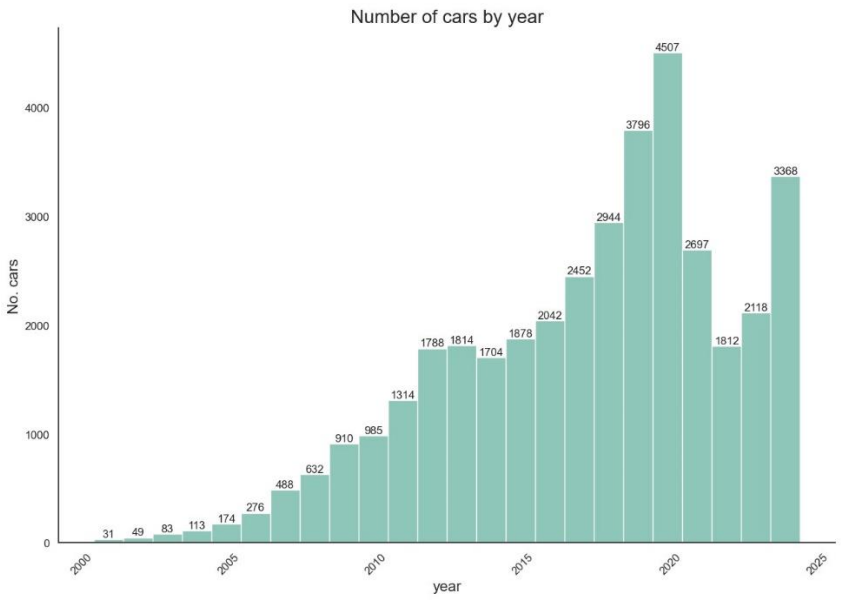



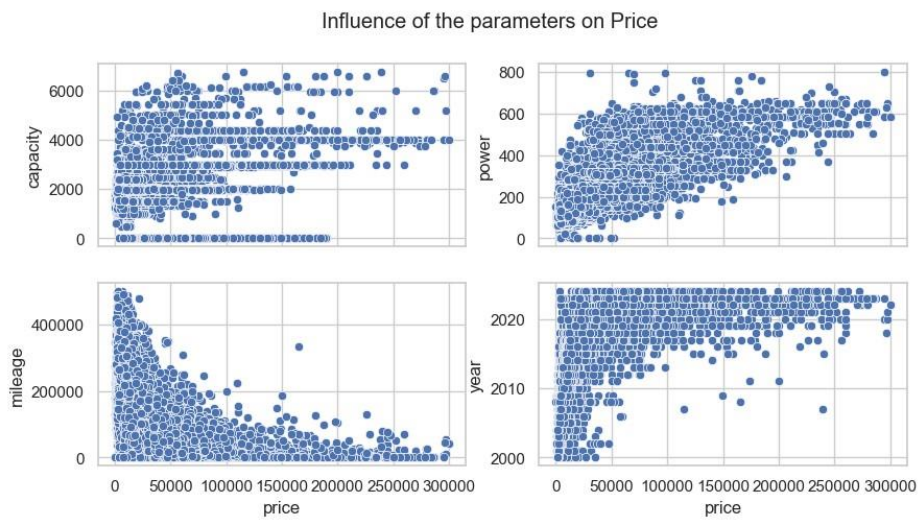
In the realm of mileage, the BMW Seria 5 boasts an impressive maximum of 2.714.900 kilometers, symbolizing its enduring and robust performance. On the flip side, the Alfa Romeo Tonale claims the record for the minimum mileage, standing at a fresh and pristine 0 kilometers.
Shifting gears to power, the Tesla Model X takes the lead with an awe-inspiring maximum power of 1.020 horsepower, showcasing its exceptional strength and performance. In contrast, the Unknown holds the distinction of the minimum power, with a modest yet functional 3 horsepower, underlining its efficiency and practical design.



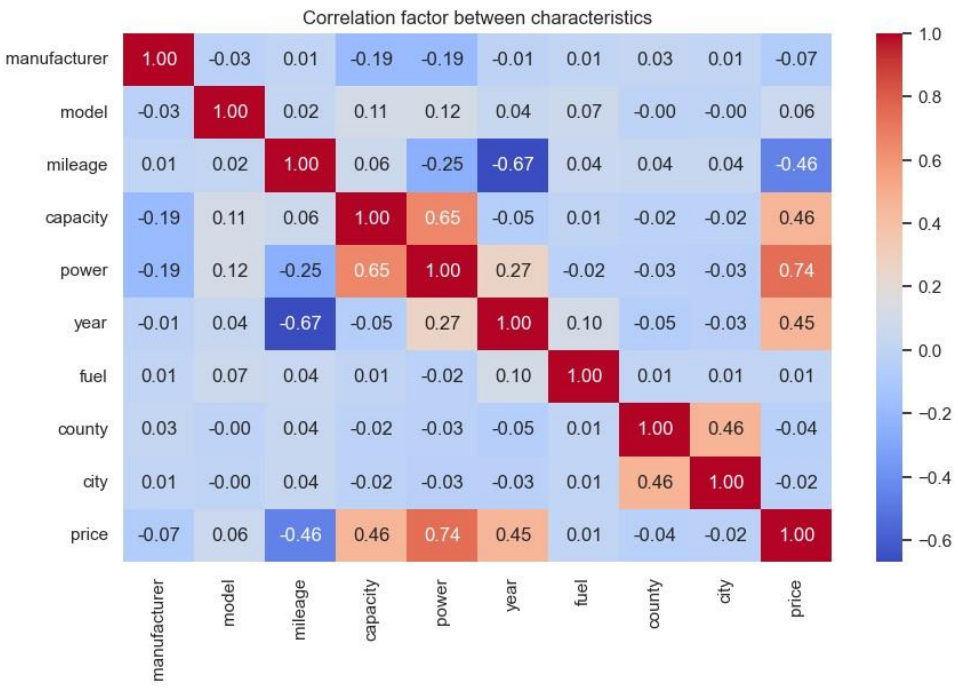



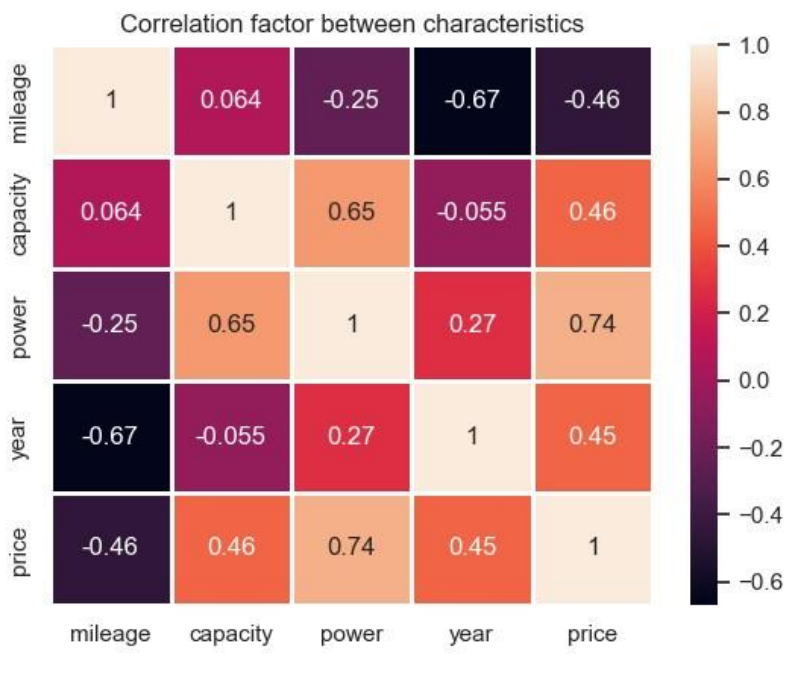

In the year 2024, a total of 330 cars rolled off the manufacturing lines, reflecting a snapshot of recent production.
Examining the factors impacting pricing, the engine's power emerges as the primary influencer, exerting the most significant impact on the car's price. Meanwhile, both mileage and the year of manufacturing play pivotal roles in determining the price, with their influences balancing each other in a reverse direction. Notably, engine capacity, while a factor, appears to have the smallest effect on the overall pricing dynamics.




The paramount factor influencing the price is the engine's power, showcasing a robust correlation coefficient of 0.74. This strong correlation underscores the significant impact that power has on determining the overall pricing dynamics in the dataset.




For the car with follow speicifications:
manufacturer: Dacia.
model: Logan.
mileage: 100.000 km.
capacity: 1.461 cm3.
power: 115 HP.
year: 2020.
fuel: Diesel.
the estimated price is: 9.362 EUR.

EvaCar

Data available: 10.02.2024
Report produced: 11.02.2024
GitHub repository: https://github.com/romanmurzac/EvaCar