

Business analytics

Analysis of car advertisement data

Alina Ivanova, Moritz Hangen, Simon Gosch

2024/03/24

Contents

List of Figures	3
List of Tables	4
1 Introduction	5
2 Theoretical background	6
3 Methodology	7
4 Findings and discussion	8
4.1 Model usage	8
4.1.1 Assessing expectations	8
4.1.2 Usage of example data	8
4.2 Findings	9
4.2.1 In line with presumptions	9
4.2.2 Outliers	9
5 Conclusion	10
5.1 Further research topics	10
5.2 Résumé	11
Bibliography	12

List of Figures

List of Tables

4.1	Assessing expectations	8
4.2	Assessing model for business use cases	8

1 Introduction

[1]

2 Theoretical background

3 Methodology

4 Findings and discussion

4.1 Model usage

4.1.1 Assessing expectations

Registration year	Mileage	Horsepower	Width (mm)	Length (mm)	Average mpg	Top speed (mph)	Predicted price
2019	60000	135	2027	4284	49	116	17596
2019	130000	135	2027	4284	49	116	15267

Table 4.1: Assessing expectations

4.1.2 Usage of example data

Registration year	Mileage	Horsepower	Width (mm)	Length (mm)	Average mpg	Top speed (mph)	Predicted price (£)
2014	180000	110	1799	4204	45	110	5601
2016	150000	120	2027	4255	48	112	12130
2018	80000	130	2027	4255	50	115	16136
2015	190000	115	1799	4204	44	108	5819

Table 4.2: Assessing model for business use cases

4.2 Findings

4.2.1 In line with presumptions

Mileage \leftrightarrow Price

Engine size \leftrightarrow Price

4.2.2 Outliers

Engine size \leftrightarrow Price

Width \leftrightarrow Price

Year \leftrightarrow Price

TODO include example of higher influence of mileage the older the car is

5 Conclusion

5.1 Further research topics

Given the scale of the available data, more potential research questions may be examined.

Lack of sales data regarding advertisement

Firstly, while the given dataset includes conclusive data regarding the advertisement price of cars in the used vehicle market, there is no indicator given, whether the car was actually sold at the price that it has been advertised for.

Nevertheless, while one may anticipate a few dealerships to over- / undershoot their prices, considering the scale of the dataset, that effect evens out for the overall market. However, evaluating whether there is a trend that pre-owned cars are systematically under- / overpriced in commercials is only possible if you compare the given data set to real sales information.

Inter-model comparison of findings

Additionally, the analysis is currently only valid for the small subset of the data including Volkswagen (VW) Golf. To levitate generalizability to the whole used vehicle market and to assess potential disparities as well as similarities in between car models, expanding the scope to the entirety of available data is obligatory.

Assess value depreciation

Having inter-model differences evaluated, a possible further research topic is the comparison of each car's new price to its future advertisement prices. For each model in the advertisement dataset, there is a corresponding data point in the basic information table that contains, among others, the manufacturer's suggested retail price (MSRP).

Given that information, the following questions could be analyzed:

- Which model retains the most value compared to its MSRP?
- Given five years of use, which car's price decreased the most?
- Is there a correlation between value depreciation and the manufacturer of the car?

This information can be useful for customers considering the purchase of a new car in order to assess its potential resale value in the future.

5.2 Résumé

The conducted analysis provides a data driven solution to the business problem of a competitive advertisement price of a given VW Golf. After preparing the raw data, key influences have been separated using Pearson-correlation and used to train a linear regression model to predict a suitable advertisement price.

Considering the parameters

- Registration year
- Mileage (mi)
- Horsepower
- Width (mm)
- Length (mm)
- Average mpg
- Top speed (mph)

the model accurately predicts a suitable future advertisement price with $R^2 = 0.927$, solving the business problem therewith. The end results align with logical assumptions, for instance that newer cars sell for higher prices. Nonetheless, they provide further insight by revealing each factor's contribution to the total amount.

Provided that outcome, further potential research areas are examined and left open for future analysis.

Bibliography

- [1] *Industrie-Roboter - Hirata Engineering Europe GmbH*. URL: <https://www.hirata.de/de/produkte/scara-roboter> (visited on 02/02/2024).