UNIVERSITY
INSTITUTE
OF LISBON

Classic car spare parts for restoration:
International supply chain and environmental impact

Maria Inês Monteiro Celorico

Master in Data Science

Supervisor:
Doctor Fernando Brito e Abreu, Associate Professor,
Iscte-IUL

Co-supervisor:
Doctor Vasco Amaral, Associate Professor,
FCT-NOVA

September, 2024

[This page has been intentionally left blank]

TECHNOLOGY
AND ARCHITECTURE

BUSINESS
SCHOOL

Department of Quantitative Methods for Management and Economics
Department of Information Science and Technology

Classic car spare parts for restoration:
International supply chain and environmental impact

Maria Inês Monteiro Celorico

Master in Data Science

Supervisor:
Doctor Fernando Brito e Abreu, Associate Professor,
Iscte-IUL

Co-supervisor:
Doctor Vasco Amaral, Associate Professor,
FCT-NOVA

September, 2024

[This page has been intentionally left blank]

Acknowledgements

I am truly grateful to all those who have supported me during the completion of my master's thesis and have contributed to my ongoing education.

First and foremost, I extend my heartfelt gratitude to my advisor, Professor Fernando Brito e Abreu, for his invaluable guidance and mentorship. His instruction in various software tools significantly enhanced the quality of this thesis, and his unwavering availability and direction under challenging issues were indispensable in every aspect of this work. I also express my sincere appreciation to my co-advisor, Professor Vasco Amaral, for his constructive feedback and guidance throughout this journey. His expertise and insights were equally crucial to the development of this thesis.

I am also thankful to the Raimundo Branco Lda restoration shop for cooperating in providing essential guidance and information, especially to Olga and Sérgio. Additionally, thanks to the PhD student António, who generously shared his environmental expertise and technical tools used for long hours. To my colleagues, José and Frederico, thank you for your support and assistance throughout this journey.

Lastly, I am deeply thankful to my family and friends, whose unwavering support and constant presence helped me stay grounded and motivated during this academic endeavor.

[This page has been intentionally left blank]

Abstract

Preserving classic cars' historical, cultural, and economic value requires careful selection of original or accurately reproduced parts to ensure authenticity. A rapid systematic literature review found no studies classifying the classic car market by international supplier networks, consumer preferences, or environmental impact, with research focusing primarily on the broader automotive or used car markets.

This dissertation examines the international distribution of classic car manufacturers and parts suppliers, analyzes market segmentation based on brand popularity, transportation impact, and supplier availability factors, and explores sustainable practices. The case of a Loures-based restoration shop, Raimundo Branco Lda, is used for illustration.

Web scraping techniques were used on the Superclassics and Restoparts websites to collect data on classic cars, while Google search hits captured brand popularity. Traffic impact data was collected from the World-Cities dataset, Geopy API, IATA, and Atmosfair websites.

Descriptive analysis assessed the distribution of manufacturers and suppliers and the materials used in classic car body parts. Based on these factors, K-Means clustering was applied to create brand groups.

Results indicate that suppliers of classic car parts are diverse in Europe and North America, with Germany leading the world. Consumer preferences align with supplier availability, and Porsche and Volkswagen brands best fit these factors. In addition, modern technologies and recycled materials can be used to produce parts, maintaining authenticity while promoting sustainability.

Keywords: classic cars; parts suppliers; environmental impact; consumer preferences; web scraping; K-Means clustering;

Resumo

A preservação do valor histórico, cultural e económico dos automóveis clássicos exige uma seleção cuidadosa de peças originais ou reproduzidas com precisão para garantir a autenticidade. Uma revisão sistemática rápida da literatura não encontrou estudos a classificar o mercado de carros clássicos em termos de redes de fornecedores internacionais, preferências dos consumidores, ou impacto ambiental, com a investigação a centrar-se principalmente nos mercados mais amplos de automóveis novos ou usados.

Esta dissertação examina a distribuição internacional dos fabricantes e dos fornecedores de peças sobresselentes para automóveis clássicos, analisa a segmentação do mercado com base nem fatores como a popularidade da marca, impacto do transporte e disponibilidade dos fornecedores, e explora as práticas sustentáveis. O caso de uma oficina de restauro sediada em Loures, a Raimundo Branco Lda, é utilizado para ilustração.

Foram utilizadas técnicas de raspagem da Web nos sítios Web da Superclassics e da Restoparts para recolher dados sobre automóveis clássicos, enquanto que os resultados da pesquisa no Google captaram a popularidade da marca. Os dados sobre o impacto do transporte foram obtidos a partir do conjunto de dados World-Cities, da API Geopy, dos sítios Web da IATA e da Atmosfair.

A análise descritiva avaliou a distribuição dos fabricantes e fornecedores, bem como os materiais utilizados nas peças da carroçaria dos automóveis clássicos. Com base nestes factores, foi aplicado o algoritmo de agrupamento K-Means para criar grupos de marcas.

Os resultados indicam que os fornecedores de peças para automóveis clássicos podem ser encontrados tanto na Europa como na América do Norte, com a Alemanha a liderar a nível mundial. As preferências dos consumidores alinham-se com a disponibilidade dos fornecedores, e as marcas Porsche e Volkswagen são as que melhor se enquadram nos factores estudados. Além disso, podem ser utilizadas tecnologias modernas e materiais reciclados para produzir peças, mantendo a autenticidade e promovendo a sustentabilidade.

Palavras-chave: carros clássicos; fornecedores de peças; impacto ambiental; preferências do consumidor; web scraping; K-Means clustering;

[This page has been intentionally left blank]

Contents

List of Figures

List of Tables

[This page has been intentionally left blank]

CHAPTER

1. Introduction

Contents

In this chapter, an overview of classic car restoration and its challenges is presented. The motivation section outlines the problem-solution in this study. Additionally, describes the objectives along with a brief methodology of the research questions. Finally, the alignment of the dissertation chapters with the CRISP-DM phases is presented.

[This page has been intentionally left blank]

Chapter 1
Introduction

1.1 Context

1.1.1 Restoration of classic cars

According to the International Council of Museums - Committee for Conservation (ICOM-CC), restoration is "all actions directly applied to a single and stable item to facilitate its appreciation, understanding, and use. These actions are only carried out when the item has lost part of its significance or function through past alteration or deterioration. They are based on respect for the original material. Most often, such actions modify the appearance of the item"[49]. The Federation Internationale Vehicules Anciens (FIVA) considers classic car restoration as an invasive operation in which pieces from other vehicles can be used for reconstruction, ensuring the maximum respect for the original manufacturing [40].

While many well-known museums and conservation institutes do not primarily focus on classic cars, the Canadian Conservation Institute (CCI) provides guidelines for preserving vehicles and FIVA [24, 40]. Their technical specificity puts the long-term preservation of vehicle collections at risk. Technical objects are made of highly varied materials, requiring multiple engineering processes from different specialization fields and processes challenging to trace back in history, which may imply the reconstruction of missing parts [40]. These materials are prone to deterioration, such as corrosion, wear, and physical damage, but on a large scale [24].

According to Delagneau et al., we can not see a car as a painting. Cars have mechanical parts and move on the road; if we applied the same rules to them as a classic object d'art, it would be hazardous for the car and probably put at risk the driver's safety [29]. Period spare parts and historic parts taken from similar vehicles can be employed to maintain functionality. Additionally, replicated parts made of modern industrial materials produced with present-day techniques are highly used these days [40].

The restoration of classic cars leaves us with another relevant debate, encompassing the question of whether the restoration of a functional car should strive to maintain the original while maintaining visible signs of restoration to preserve its authenticity. Parts from similar vehicles or replicated may not be the same visually as the original model, especially when made from artificial materials such as leather, fabric, or nickel plating. Even when authentic materials are used, restoration could lead to an imprecise interpretation [40].

Also, using authentic materials poses a significant issue for classic car restoration shops in preserving the cultural heritage. There is a substantial lack of information about the original materials used in classic cars. According to Delagneau et al., the manufacturer does not make most parts of the classic car, the original specification is not known or is ignored, and the quality of the reproduction is not monitored [29]. For this purpose, restoration must always be documented for the benefit of future curators and owners [40].

1.1.2 Challenges in classic car industry

Classic cars were defined in the market as such in the 1970s [14]. According to the International Forum on the Authenticity of Historical Vehicles by FIVA, a classic car can be of historical and cultural value, predominately determined by the authenticity of the car [29]. Also, according to the Historic & Classic Vehicles Alliance (HCVA) and Automóvel Club de Portugal (ACP), classic cars have no specific age requirement. They are considered as such if they hold historical, aesthetic, or cultural value [2, 44]. For example, the Herbie car, famous for its appearance in a 1960s Disney movie, is a classic due to its cultural value [33]. Or, the BMW 3.0 CSL Batmobile, adapted from a luxury sports car to a race car [12]. A brief history of the top 7 best-performing brands in this dissertation is presented in Appendix A.

Some challenges were encountered in the classic car industry. For some years, vehicles with great historical interest, such as classic cars, were being destroyed by their owners or sold for breaking up purposes [40]. Some cases, such as the classic cars that won the Le Mans race, underwent significant changes, risking their authenticity. Because of that, using the chassis plate with 'historic' numbers was often questioned [14].

The issue of classic car destruction is exacerbated by the incentives for disposing of cars older than ten years. The proposed addition to the 2024 Portugal state budget, as outlined in Article 120ºA, offers a sum of 4500 euros for the purchase of another vehicle and disposal of the older car or a reduction in car taxes, particularly for electric or hybrid cars [82]. This requires a higher priority on integrating eco-friendly solutions into classic car restoration to prevent their disposal.

Various eco-friendly measures can be implemented within the classic car supply chain market, from raw material extraction to parts transportation. Studies have indicated that remanufacturing, recycling, reuse, greenhouse gas emissions, and transportation are the primary sources of pollutant emissions [23, 80]. Furthermore, alternative materials, such as fiber composites, have been suggested to minimize environmental impact in automotive parts production [4].

It is also important to consider consumer preferences for specific classic car brands. Several factors contribute to changes in the market. For instance, the manufacturer and brand type can significantly influence car pricing, which in turn impacts the overall market [6, 13, 30, 31, 75]. From this perspective, it is important to assess whether the availability of parts suppliers aligns with the interest of collectors and to determine if improvements are needed for specific classic car brands.

1.2 Motivation

Several challenges emerge from the experience of the Raimundo Branco Lda (RB) classic car restoration shop in Frielas industrial zone. The authenticity of classic cars is essential for preserving their historical, cultural, and economic value. The careful selection of genuine or accurately reproduced parts contributes to the restored classic car's overall quality, and accuracy. While internationally sourced parts are often more economical (even considering the shipping costs) than locally handcrafted ones, obtaining authentic parts for each model is often challenging, and the craftsmanship required to maintain the car's authenticity is rare.

In addition, the carbon footprint associated with the transportation of the parts plays a significant role in managing the market, as suppliers are often far from restoration shops [80]. Another related problem is the risk of classic car destruction, particularly due to disposal regulations in Portugal. These concerns highlight the need for a systematic study of parts availability and sustainable solutions to preserve classic cars.

While studies exist on the broader automotive industry, little research focuses on classic cars. This study aims to fill this gap by collecting data on classic car brands and models, suppliers, and the transportation impact of parts. It also analyzes the manufacturers, brand popularity, and bodywork materials to explore the market demand and consumer preferences and identify sustainable practices.

The findings from this study can benefit various stakeholders: suppliers may improve their sustainable practices, collectors can gain insights into the best deals, and the availability of parts, and restoration shops can enhance car authenticity by identifying original parts suppliers and adopting new, more sustainable restoration technologies.

1.3 Objectives and research questions

The main aim of this study is to explore the international distribution of manufacturers and parts suppliers for classic cars and to evaluate how factors such as brand popularity, transportation impact, and supplier availability influence the different brands of classic cars.

Our secondary objectives involve proposing sustainable practices, including modern technologies, to produce spare parts for classic cars while addressing concerns about authenticity. To assess the suitability of these technologies, it is essential first to understand the original materials used to produce these parts.

To achieve these goals, the following research questions (RQ) are posed:

- **RQ1:** To what extent are classic car manufacturers and parts suppliers distributed internationally, and how do their geographic locations and classic car brands they endorse compare?
- **RQ2:** Is there a relationship between the popularity of classic car brands and the parts supplier availability on the international market?
- **RQ3:** Which of the most popular classic car brands best align with the transportation impact and the suppliers' availability for parts supply at a specific restoration shop location?
- **RQ4:** How can sustainable methods be applied to the materials used in classic car bodywork parts, balancing the preservation of authenticity with modern technologies?

By web scrapping and cross-referencing various sources, it was possible to gather the information needed for the study.

Regarding RQ1, an exploratory analysis was employed to visualize the international distribution of the manufacturers and parts suppliers for classic car brands, thereby identifying countries with a higher concentration.

Regarding RQ2, parts availability was assessed by counting the number of suppliers for each brand, and the popularity of classic car brands was determined using Google search hits. These two factors were compared to identify any discrepancies between demand and supply.

Regarding RQ3, a multi-criteria analysis of classic cars was conducted using a K-means clustering algorithm. This analysis focused on transportation impact, the popularity of brands, and the availability of parts suppliers. Concerning the impact of transportation, carbon emissions associated with air travel for the transportation of parts between international suppliers and a restoration shop in Loures were considered. The previous analysis used supplier count and Google search hits for the brand popularity and supplier availability factors.

Regarding RQ4, an exploratory analysis was used to identify the most commonly used materials in bodywork parts for classic cars. This involved analyzing the composition of the part and the type of finish applied to it.

1.4 Dissertation organization

This dissertation is structured according to the CRISP-DM (Cross-Industry Standard Process for Data Mining) model, which guided data acquisition and analysis steps. Each phase of the CRISP-DM process is distributed in the different thesis chapters, as detailed in Table 1.1.

This paper is organized as follows: the next section 2 describes the rapid systematic literature review, section 3 presents the methodology used for the data acquisition and analysis, 4 presents the outcome of the analysis, section 5 discusses the outcomes, and section 6 covers the conclusions, limitations and the outline of future work.

Table 1.1: Integration of CRISP-DM phases into thesis chapters

CRISP-DM Phase	**Description of CRISP-DM Phase**	**Thesis Chapter**	**Chapter Section**
Business Understanding	Addressing the business problem and defining the research goals	Chapter 1 (Introduction) Chapter 2 (Rapid Systematic Literature Review) Chapter 3 (Methodology)	1.1 Context 1.2 Motivation 1.3 Objectives and Research Questions 2.3 Review of Selected Articles 3.1 Illustration Case
Data Understanding	Familiarity with collected data, verification of data quality, and descriptive analysis	Chapter 2 (Rapid Systematic Literature Review) Chapter 3 (Methodology) Chapter 4 (Results)	2.1 Methodology 2.2 Selected Articles and Reference Sources 3.2 Data Collection 3.3.3 Data Description 4.1 Descriptive Analysis
Data Preparation	Data cleaning and transformation	Chapter 3 (Methodology)	3.3.1 Data Cleaning 3.3.2 Data Transformation
Modeling	Application of modeling techniques	Chapter 3 (Methodology) Chapter 4 (Results)	3.4 Data Modeling 4.2 Modeling
Evaluation	Evaluation of results and possible improvements	Chapter 5 (Discussion)	5.1 International Distribution 5.2 Segmentation of Classic Car Brands 5.3 Top Classic Car Brands 5.4 Material Composition and Surface Treatments
Deployment	Model implementation and final report	Chapter 6 (Conclusion and Future Work)	6.1 Conclusion 6.2 Future Work

CHAPTER 2.

Rapid systematic literature review

Contents

This chapter contains a rapid systematic literature review conducted using the Parsifal tool. It presents the different sections of the tool and the criteria adopted. Accepted articles are briefly analyzed, and their review is subsequently presented, followed by an alignment of the current research with the research questions.

[This page has been intentionally left blank]

Chapter 2
Rapid systematic literature review

2.1 Methodology

To select articles for this study, Parsifal was used, a tool designed to assist researchers in conducting systematic literature reviews. It enables us to construct string searches, choose articles based on abstracts, and evaluate them based on predefined criteria. The criteria defined with this tool are presented next.

The sources section included academic databases such as Dimensions, IEEE Digital Library, ISI Web of Science, and Scopus.

In the "keyword and Synonyms"section, the search string was formulated based on specific keywords and their synonyms (Table 2.1), with each keyword designed to represent a particular aspect of the study (Population: subject of the study; Intervention: actions to produce results; Outcome: desired effect).

Table 2.1: Keywords and synonyms used in Parsifal for the search strings.

Keyword	Synonyms	Related to
Auto parts	automotive spare parts car spare part second-hand car used car etc.	Population
Data science	classification clustering machine learning text mining etc.	Intervention
Sustainability	supply chain environment impact pattern recognition price prediction etc.	Outcome

The final search string obtained was: *("auto parts"OR "car spare part"OR "automotive spare parts"OR "second-hand car"OR "used car"OR "used cars") AND ("data science"OR "classification"OR "cluster"OR "clustering"OR "data mining"OR "decision tree"OR "forecast*"OR "machine learning"OR "neuronal network"OR "regression"OR "text mining"OR "time series"OR "web crawling"OR "web scraping") AND ("sustainab*"OR "brand"OR "country"OR "supply chain"OR "LCA"OR "carbon footprint"OR "environment impact"OR "pattern recognition"OR "price prediction"OR "supplier")* .

This search string was filtered by the publication type, language, and published year. The publication type allows us to acquire more precise and higher-quality papers. The readership base is also expanded by choosing a global language and facilitating international collaboration. Finally, the publication year ensures information on the most recent modeling and evaluation techniques. These filters selected articles, conference papers, and book chapters written in English from 2016 to 2024.

Based on this search string, applied to the previously mentioned databases, a set of papers was imported to Parsifal: 56 Dimensions, 47 IEEE, 97 Scopus, and 29 Web of Science.

There were 229 papers total, 74 duplicates, and 155 articles. Then, these articles were selected based on the title and abstract, with some exclusion criteria: the paper is out of the automotive context, it applies to electric vehicles, it applies to passenger cars, it is not a primary study, and it is not written in English (as some papers had passed the previous filter). A total of 33 papers were rejected, with 122 remaining articles.

To evaluate the quality of accepted papers, a set of questions is made for the classification under a weighted scale (0 - strongly disagree; 1 - disagree; 2 - neither agree nor disagree; 3 - agree; 4 - strongly agree):

- Does the article's goals align with our research questions?
- Was the methodology easily described?
- Was the data collected/used clearly described?
- Were the modeling techniques and evaluation metrics directly described?
- Does the results answer the goals of the study?

Each paper can receive a score between 0 and 20 with this weighted scale. Sixty articles with a score equal to or higher than 15 passed to the extraction step. In addition to the Parsifal score, journal rank (JCR or SJR) and citations per year were used. Regarding journal rank, SJR is based on Scopus citations and evaluates journal quality using the SJR metric. JCR, based on Web of Science citations, utilizes the impact factor metric (JIF). Between JCR and JIF, the one with the highest ranking was selected. Citations per year were calculated based on Google Scholar citations divided by the years the article has been published. After the evaluation step, articles were exported to a CSV file with some details (author, year, category, features, methods, data split, best result, etc.), with each line representing an article.

2.2 Selected articles and reference sources

For this systematic literature review, 60 scientific articles were selected, of which 19 were published in journals, 40 were presented as conference papers, and 1 was a book chapter.

Based on figures 2.1 and 2.2, it is evident that the majority of the articles were retrieved from the Scopus database (42.4%), while the lowest proportion originated from the ISI Web of Science database (12.7%). Moreover, a comparable number of articles were accepted from each database. However, a smaller percentage of articles were accepted by the IEEE database.

However, the current literature focuses mainly on the used car market, making it difficult to conduct a reliable comparative analysis for our study. Nevertheless, some similar assumptions were found.

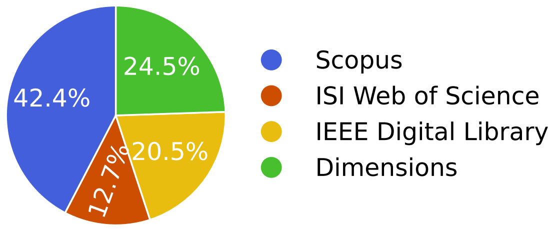

Figure 2.1: Parsifal selected articles by database.

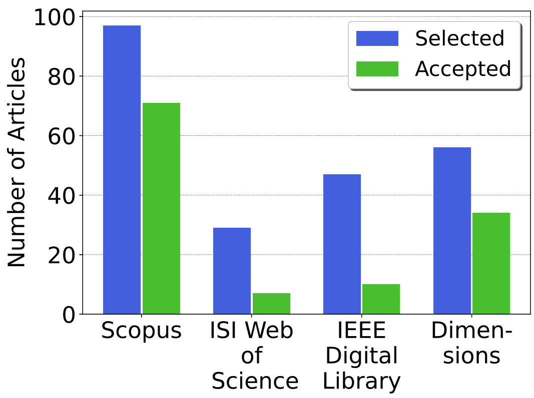

Figure 2.2: Parsifal selected vs. accepted articles by database.

The following section presents a summary of the selected articles, firstly dedicated to the information collected online, secondly to the exploratory analysis made by authors from their datasets, and finally to the modelling techniques applied. This information can be mainly found in the appendices section as well (Tables B.1, B.2, and B.3).

2.3 Review of selected articles

2.3.1 Collecting vehicle information online

Web scraping allows for the automation of gathering secondary data from websites, reducing the time required to collect data and enabling the saving of the data into local files or databases [22, 39, 68]. The first step in web scraping involves identifying the URLs of the pages to be scraped. Next, an HTTP request is sent to the website's server at that URL. The server's response contains an HTML web page with the requested content, which is then parsed to extract the relevant information, and the data is saved to a file or database [85].

Some authors use web crawling technology to extract information from websites. Web crawlers or spiders systematically browse web pages and follow links. Starting from a specific URL, they follow the hyperlinks on that web page to visit other pages until it reaches a certain number of web pages or criteria. In this process, multiple pages are crawled. This technology can collect car listings from various websites [102].

In fourteen articles, web crawling/scraping techniques were used to predict prices in the used car market in Romania, Morocco, and China. Most of the data from the websites was organized in tables, except for the study by Brahimi, where online advertisements were analyzed, and data was retrieved in textual documents [20].

The authors utilized Beautiful Soup and Selenium libraries for web crawling/scraping techniques. The Beautiful Soup library parses and extracts data from HTML web pages. At the same time, the Selenium library also allows for interaction with web pages, such as clicking buttons, inputting data, and more [43, 103]. The Selenium library also includes the WebDriver tool for interacting with web pages and facilitating web scraping and browser automation tasks [92]. The Beautiful Soup and requests libraries can be used together to make HTTP requests and parse HTML content [13].

Instead of using code to scrap information from websites, Putro and Indrawati used ”Parsehub”, a tool that allows for automatic extraction of data from web pages without writing code, offering an easier way to perform web crawling/scraping [77].

The collected dataset features from web scraping techniques typically included information from automotive sellers about car features (e.g., brand, model, year, origin), with only one study incorporating car parts data (e.g., doors, roof, bumpers) [103]. Some difficulties were encountered in extracting data from the web pages. Some content did not perfectly align with the feature names on the website. For example, the brand feature had titles the seller wrote, which needed to be matched with a list of all car brands [31]. Additionally, another study found that brands are often underrepresented [30], so the preprocessing analysis should be done carefully.

2.3.2 Exploring trends and patterns in vehicle characteristics

Before applying modeling algorithms, an exploratory analysis of the dataset was conducted across thirty-nine articles. The authors examined dataset features using descriptive statistics and correlation analysis to understand their relationships and strengths. They studied different aspects of their datasets, ranging from market trends to the key factors influencing the prediction of their model.

The Jupyter Notebook web-based platform is used for data analysis. It enables the construction of notes along the code, allows the visualization of outputs in different parts of the code, and supports the visualization of graphs and tables, among others [103]. With some libraries in this platform, such as Scikit-learn and Pandas, data analysis and models are constructed [51].

Descriptive statistics were made to understand the main characteristics of the dataset features. Used car datasets mainly included cars with manual transmission and running on diesel or petrol fuel [71]. The most popular brands found in the datasets were Volkswagen, BMW, Audi, Renault, and Mercedes-Benz (Figure 2.3) [22, 30, 31]. However, some other authors also mentioned Skoda, Opel, Fiat, Ford, Maruti, Toyota, and Hyundai as popular brands in their datasets [41, 64, 92]. Regarding cars with higher mileage, the leading brands are Mercedes-Benz, BMW, Volkswagen, Opel, Skoda, and Toyota [13, 22]. This indicates that these cars are among the most frequently used on the road, and may show more wear and require additional maintenance.

From an international perspective, the United States and Germany have well-developed used car markets, reaching 68.75% and 65.55% of the total trade, respectively (Figure 2.4). On the other hand, China's used car market only constitutes 22.48% of the total trading but is experiencing rapid development [64].

Manufacturers can significantly influence car pricing and impact the car market [6]. The type of brand can also impact car prices, particularly for brands such as BMW, Mercedes-Benz, Land Rover, Porsche, Audi, and Volkswagen. Although other brands like Rolls-Royce, Lamborghini, Toyota, and Jaguar, were also identified, they were found in smaller portions [13, 30, 31, 75].

To predict the price of a used car, some researchers have analyzed the relationship between different features to understand better how each feature impacts the selling price. The analysis revealed that various factors such as general information (brand, model, and production year), performance and usage (age, kilometers driven/mileage, transmission, engine displacement,

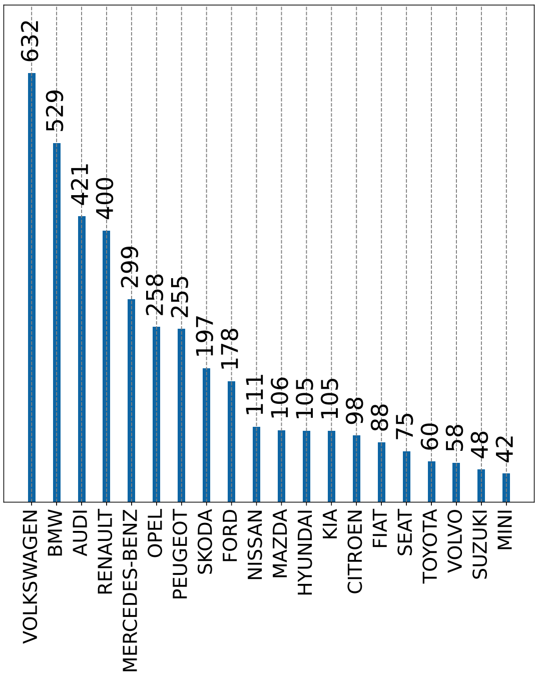

Figure 2.3: Brand's popularity. Adapted from Bukvić et al. [22].

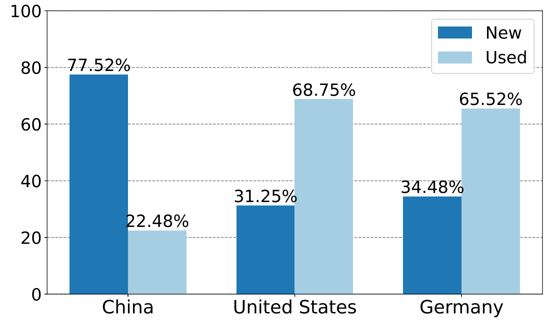

Figure 2.4: Used car trading proportion of China, the United States, and Germany. Adapted from Ma [64].

cylinders, fuel tank, and odometer), and additional information (color, number of passengers, assembly, seller type, and city) significantly influence the price of the car. Based on the studies, the most notable correlations with the price feature were found to be: engine (0.66), peak power (0.65), curb weight (0.59), wheelbase (0.58), seller type individual (-0.55), year (0.53), and transmission (0.53) [3, 13, 43, 61, 70, 71, 83, 93, 95]. These features highly influence the car price. According to Alhakamy et al., Bhatt et al., Rane et al., the following features are positively related to the price of cars: individual sellers, automatic transmission, diesel type, drive wheels, and zero owners [6, 17, 78]. This means that with those features increase, the car price also increases.

2.3.3 Modeling approaches for vehicle market analysis

After completing data exploration, it is time to input the data into a model. In the used car market, the focus has been on developing price-prediction models. A few studies also consider consumer preferences and environmental impact analysis to enhance manufacturer and supplier logistics. This section was divided into supervised and unsupervised techniques.

2.3.3.1 Supervised learning techniques

The supervised learning technique uses labeled data. The model learns over time through training algorithms. These algorithms can be used for classification or regression problems, such as generating supplier classifications and price predictions.

Several studies have used this technique to enhance decision-making in the used car market and enable on-demand sales. In 2019, Skribans and Hegerty analyzed thousands of used cars in Latvia and Germany. They applied an econometric model to visualize the relationship between the price and the main characteristics of the cars. The model explains 89% of the variance in market prices. The authors agree that exported used vehicles are profitable for Germany and Eastern Europe [88]. Early in 2023, Dutulescu et al. predicted the used car prices in two different countries (Germany and Romania) by analyzing relationships between car features and images through a convolution neural network model, which achieved a 95% accuracy. They also measured the most important features of the model: transmission type, age, and engine capacity. Authors advise standardizing data and extracting the car manufacturer and model directly from the images [30]. The authors H. Zhang and Shaprapawad et al. standardize data before applying a model. They use the min-max algorithm to eliminate the influence of different magnitudes between features [86, 106]. Amellal et al. improved the price prediction method with sentiment analysis addiction. They quantified customer sentiments from the textual data on ”customer comment”. It is made a fusion of the sentiment analysis with the price prediction to achieve a more accurate data-driven decision-making [8]. Other authors made the price prediction of used cars using different techniques, where the best performance models were varied, from BI-GRU ($R^2 = 0.99$) [60], linear regression ($R^2 = 0.99$) [10], Random forest ($R^2 = 0.99$) [68], extra tree regression ($R^2 = 0.98$) [99], XGBoost ($R^2 = 0.97$) [103], to neural network ($R^2 = 0.96$) [77], among others. Some different approaches were made for the price prediction, such as the making of a categorical feature for the price [39], choosing his price levels such as ’500-2000’ from the calculation of the used car factors influencing their cost [95]. Also, a study on the analysis of techniques for treating missing values in a used car dataset was made. The MICE technique had the best result for the random forest regression model, achieving 94.07% of accuracy [84]. To predict the future price of used cars, [78] forecasted their selling cost. The decision tree obtained the best-performing model ($R^2 = 0.85$). The authors also recommend training on data clusters rather than the entire dataset to improve performance. Also, a similar analysis was made by [45] in 2023, using random forest model (R^2 squared $= 0.95$) [45], and by Nandan and Ghosh, with XGBoost model ($R^2 = 0.87$) [70].

From some studies, conclusions were made for the used car price market, such as in the last years, there was an increase in the used cars selling price [57], damaged cars after being repaired, increase in the selling price [75], and that the most expensive seller type is distinguishable the dealer, instead of the individual [95].

Relating to manufacturer and parts supplier logistics, there were few studies. In 2019, Rojniruttikul explores four strategic supply chain management for auto parts market (information, transportation, facilities, and inventory). He applied a regression model, which showed that all four strategies positively affected the competitive advantage in the auto parts market, especially transportation [80]. A few years later, in 2020, Ishak and Wijaya classify the auto parts suppliers based on desired criteria. They pretend to answer to some difficulties, such as the selection of suppliers, delays in the receipt, and the lack of raw materials. In selecting suppliers, there are essential attributes such as quality, quantity of materials, reliability, and price. The model was made with a decision tree (accuracy = 90.85%), allowing they to observe the relations between each criterion (Figure 2.5) [50].

Also, a lead time forecast and anomaly detection study was made for a Morocco automotive supplier responsible for importing and distributing automotive parts. They provide a more accurate and reliable lead time forecast to reduce stock-out risk, minimize costs, and improve

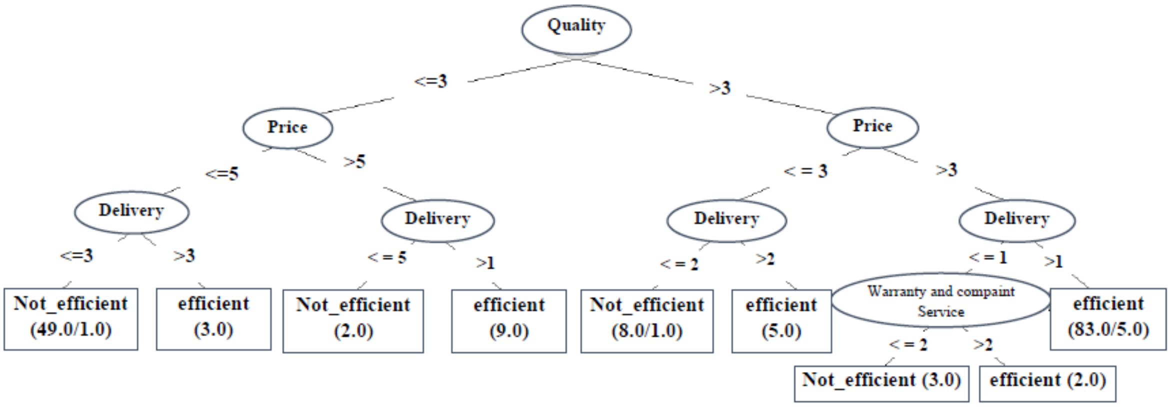

Figure 2.5: Classification of the auto parts suppliers based on quality, price, and delivery criteria [50].

transportation planning [7]. A predictive system model using smart, innovative technologies was presented to evaluate automotive parts. It was found that the significance of defects depends on the country of manufacture: Airbags are the most likely to fail in Japanese-made cars; South Koreans have parking brake and external lightning problems; German cars have mainly airbags problems; and American cars have speed sensor defect [62].

Regarding consumer preferences, Lyu and Li identified the influencing factors of consumers' purchase of re-manufactured parts from questionnaires. They found that internal factors such as product cognition, trust, and environmental friendliness could regulate the relationship between external factors such as product attributes, public publicity, and the purchase behavior factor [63]. In 2023, Ma analyzed the consumer preference for developing China's used car market. Analysis revealed an increase in car sales until 2019 and then suffered a decrease, probably due to the pandemic and the international market in 2022 [64]. In another study, a customer segmentation analysis was made. Regarding the car sales dealers for used cars, a clustering technique was applied [35].

Regarding the environmental impact, Akhshik et al. predicted the greenhouse emissions of replacing fiber composites in automotive parts for a life cycle assessment with a linear regression model (accuracy = 95%) [4].

2.3.3.2 Unsupervised learning techniques

The unsupervised learning technique uses unlabeled data to discover new patterns. Algorithms such as clustering, association, and dimensionality reduction can be employed for tasks like profile identification, car popularity analysis, or life cycle assessment.

For environmental impact analysis, in 2022, Caliskan et al. evaluates one of the 12th goals for sustainable manufacture - RC&P. They compare automotive manufacturers according to the inclusion of this goal in their sustainability reports. According to the experts' choices, they have produced 13 responsible consumption and production dimensions. Text mining was used to measure weights for each report, and sensitivity analysis was implemented to investigate the effect of weights on ranks. The most weighted dimensions were re-manufacturing, recycling, reuse, and greenhouse gas emissions. According to ranked dimensions, there were five top companies:

Peugeot, General Motors, Toyota, Ford, and Nissan (Figure 2.6). Also, the countries with the most greenhouse gas emissions are China, the USA, Japan, Germany, and South Korea [23].

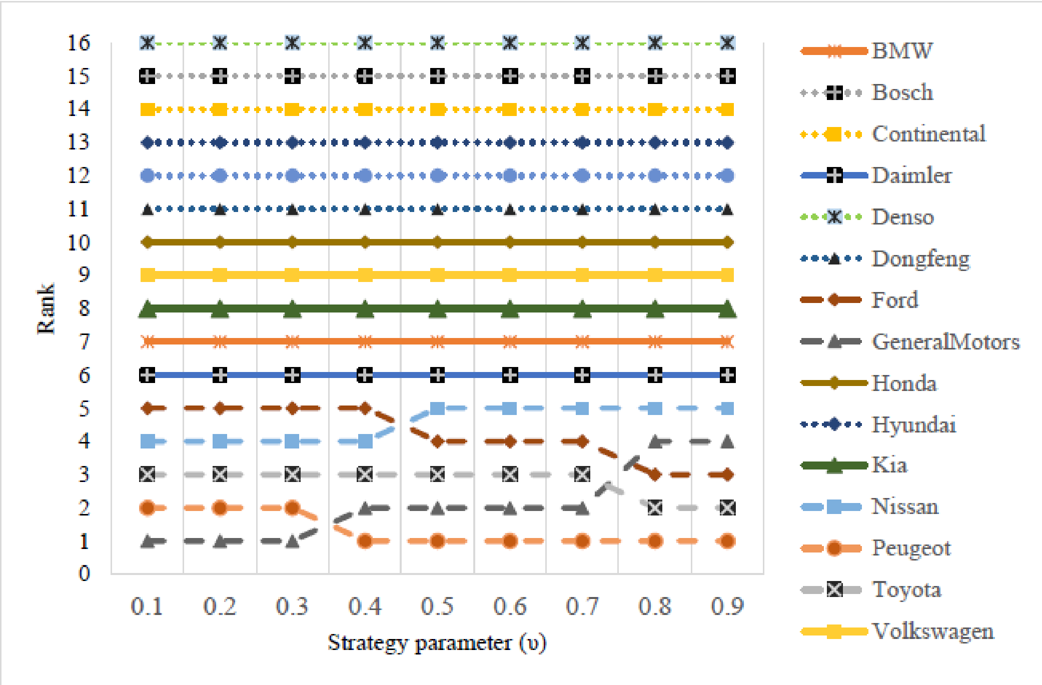

Figure 2.6: Top suppliers based on weight dimensions [23].

Some studies used clustering models to identify group characteristics. Farhan and Heikal applied this model to identify customer profiles and needs to develop targeted marketing strategies for each customer group. They analyzed the characteristics and preferences of customers to identify distinct profiles. The K-Means algorithm was used for clustering, and the elbow curve was employed to determine the optimal number of clusters. Analysis of variance (ANOVA), and p-value were used to enhance the quality and significance of cluster groups. Some conclusions drawn include that the Toyota 2018-2019 model was the most popular among buyers and that placing dealers in strategic locations near the work areas of private employees is an important factor [35].

In 2023, to increase the profit of the PQR showroom in Indonesia, Bhaskara and Wasesa offers a solution using descriptive and prescriptive analysis with K-means clustering. From clustering results, the authors accomplish that the fastest cars sold were LCGC, the preferable car was MPY Toyota Avanza, and the best profit car was SUV Toyota Fortuner [16].

Another study by Vesovic et al. in 2022 investigates the car motor life cycle in Serbia and Montenegro for better decision-making. He uses three clustering models based on McQueen and Hartigan algorithms. They argue that data should be organized with some properties: high level of intra-cluster similarity, low level of inter-cluster similarity, and natural grouping among data points. They state that the findings should be tested using regression tests, such as the sum of squared errors (SSE), adjusted Rand index (ARI), variation of information (VI), or normalized mutual information (NMI). Analysis of variance (ANOVA), the F-test, and the P-value can be used to determine statistical significance and compare group clusters. ANOVA assesses whether there are significant differences between the means of different groups by comparing the variance between groups to the variance within groups. The F-test evaluates the ratio of variance between groups to variance within groups. The P-value indicates the probability of observing the expected values, assuming the null hypothesis is true. The clustering results showed that the maintenance cost for the end-of-life phase's operating cost is medium-high (Figure 2.7). They affirm that maintenance costs are no longer sustainable as soon as the car reaches 200.000km driven [94].

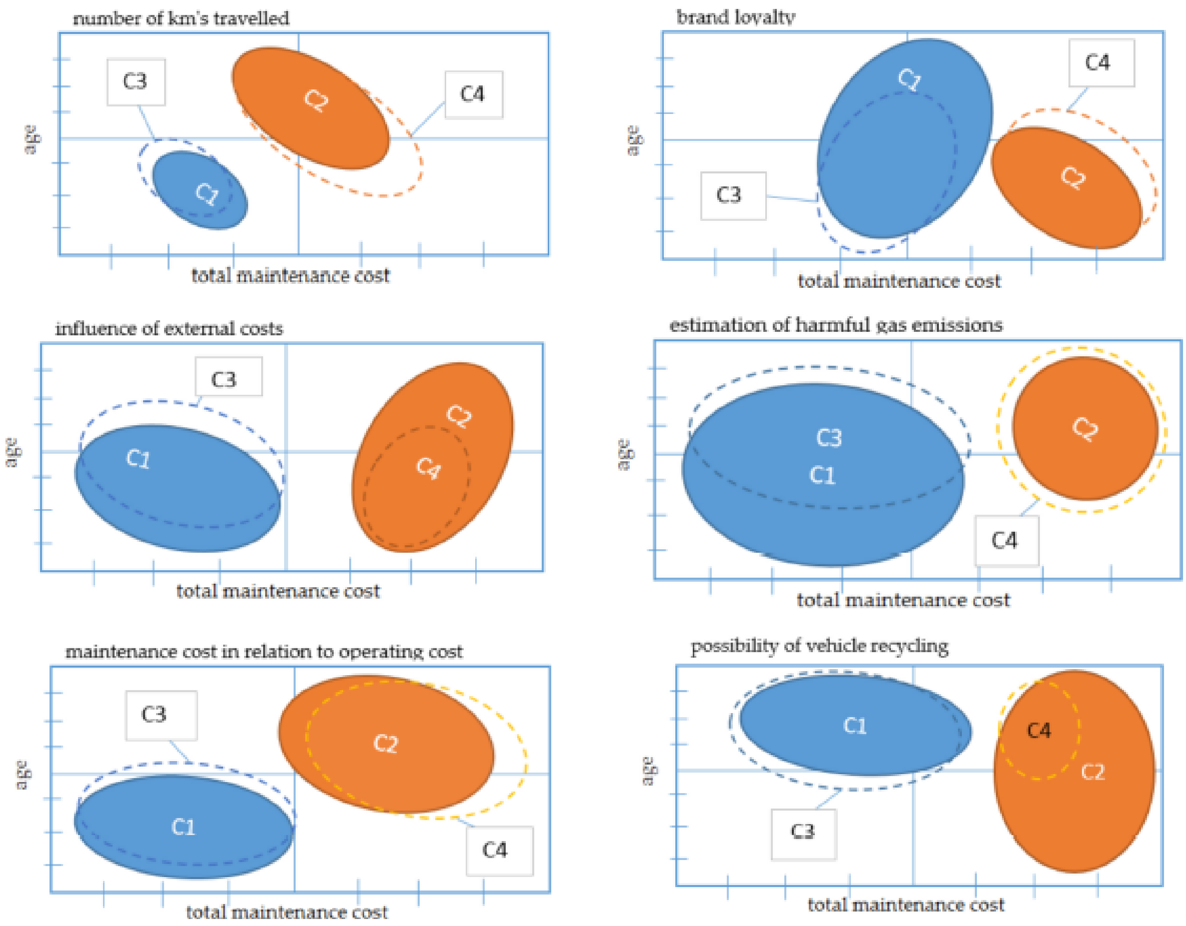

Figure 2.7: Clustering results between the two observed life-cycle phases [94].

2.3.4 Alignment with research questions

Regarding the first research goal, there is little information on the international manufacturers and parts suppliers regarding their geographic distribution and related brands. Only a study of the used car market could relate to market distribution. However, it focuses on analyzing consumer preferences for the trading market [64]. According to the article, the USA and Germany have well-developed markets, indicating a higher interest in used cars among consumers in those countries. This information could be relevant to analyzing suppliers' availability in our investigation.

Regarding the second research goal, there is limited information on the brand popularity and the availability of suppliers for specific brands. This data is only accessible in the used car market [22, 30, 31]. These articles concluded that Volkswagen and Mercedes-Benz were the most frequent brands in the dataset. On the other hand, Volkswagen, BMW, and Toyota stand out as the most popular brands for owners. This information could be relevant for exploring the brands' manufacturer dataset and compare to Google search hits analysis.

Regarding the third research goal, many studies evaluated the environmental impacts of the used car market. Some studies assess the polluting emissions of supplier companies. One of them concludes that China, the USA, Japan, Germany, and South Korea companies are the ones that produce the most greenhouse emissions [23], which means that Europe is not the most pollutant continent, but Asia. The authors also observed that the companies most correlated with the RC&P sustainable goal were Peugeot, General Motors, Toyota, Ford, and Nissan [23]. Regarding supply chain management, transportation was the factor that mostly positively affected the auto parts market [80]. Regarding car exportation from Latvia, the most profitable countries were identified as Germany and Eastern Europe [88]. These findings could help highlight the meaning of transportation in terms of environmental impact and also help understand which brands, based on the supplier brand companies and locations, could achieve better sustainable results. Regarding the clustering technique, a study provides a detailed description of a motor life cycle. Even though it is not the main focus of this thesis, the analysis is comprehensive and includes a detailed description of tests to validate the model Vesovic et al. This study was the best guide for evaluating the clustering technique in this thesis. The article possesses a 17 score in Parsifal and was published in a Q2 journal.

Regarding the fourth research goal, no relevant information was found using data science techniques with the specific search string provided earlier. However, many studies were found regarding the use of new technologies for used cars and also classic cars, which are mentioned in the future work section [9, 28, 47, 73, 89, 104]. On the other hand, there is a lack of research on the material composition or finishing used during classic car restoration processes.

CHAPTER 3.

Methodology

Contents

This chapter presents the illustration case used for this study. The methodology for collecting the different data, as well as information on the scraped websites. The chapter also covers data processing, including the cleaning techniques, and final features format. Finally, it identifies the model techniques used.

[This page has been intentionally left blank]

Chapter 3
Methodology

3.1 Illustration case

For illustration, the RB classic cars restoration shop is considered, located in Frielas, in the Loures region of Portugal, where a digital transformation initiative has been underway for some time in collaboration with our two universities [36, 42, 69]. RB has been operating for over forty years and specializes in classic car bodywork, being committed to employing the best preservation and restoration practices to ensure authenticity and quality, as defined in the Charter of Turin [40], including the local handicraft of spare parts, as mentioned in our research questions.

RB's location was used as a benchmark for measuring air travel's carbon footprint for classic car spare suppliers worldwide.

The approach employed for the study is presented in the following subsections, covering data collection, preprocessing, and modeling. The data pipeline is represented in Figure 3.1, and the project code is available on the GitHub platform.

3.2 Data collection

3.2.1 Initial web search and website criteria

An extensive Google search was conducted to gather information on classic car manufacturers, suppliers, and bodywork spare parts through classic car datasets. However, these datasets were scarce and provided insufficient information on brands and models.

Then the creation of new datasets using web scraping techniques was considered. The information contents that drove our choice in selecting target websites were as follows:

- mainly European classic cars (as the illustration case is in Europe);
- manufacturer features such as their brand, model, year, and country;
- supplier features such as their 'location';
- bodywork spare parts features such as their dimensions, weight, and type of material;
- a significant number of classic car brands, models, and bodywork parts (N=20 per subgroup in clustering analysis [27];

Typically, information about classic car manufacturers and their restoration shops is found in separate sources, complicating the search process. Fortunately, information on classic car manufacturers and restoration shops from a single unofficial website was gathered, making it easier to compare classic car models. On the other hand, a different website for information about spare parts suppliers were used, particularly those specializing in bodywork components. This information about bodywork components is essential for our illustration case, which involves

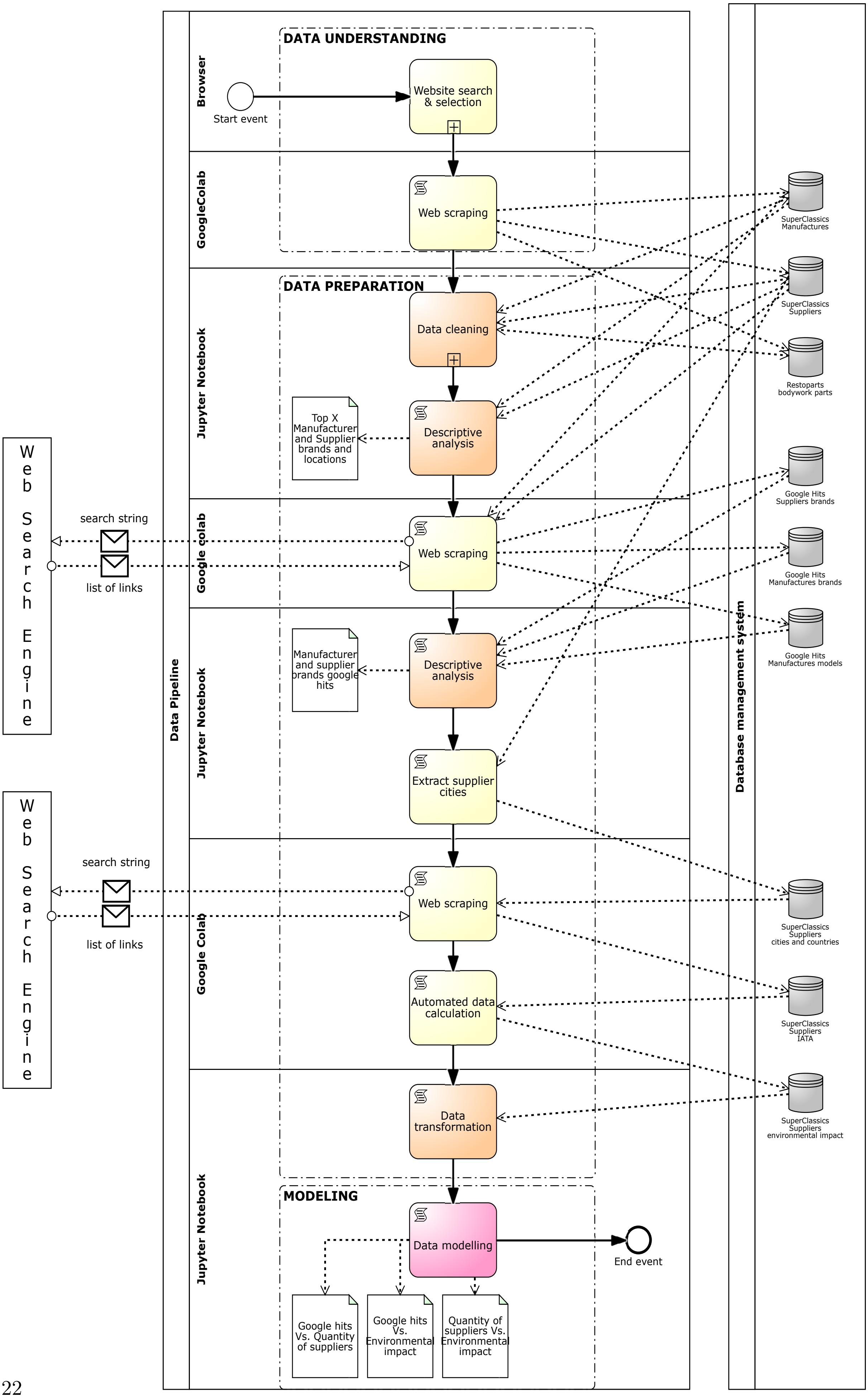

Figure 3.1: Methodology phases of data understanding, preparation, and modeling.

restoring classic cars, including repairing and replacing bodywork parts such as front and rear panels, roofs, hoods, doors, fenders, bumpers, etc. Although this website provided valuable information on composition materials and finishing processes (such as painting and protective coatings), comparing this information with data on classic car models and restoration shops may not be feasible due to the separation of sources.

3.2.2 Context and extracted pages from Superclassics and Restoparts websites

According to the abovementioned criteria, the SuperClassics website was selected to gather information on manufacturers and restoration shops. In contrast, the Restoparts website was chosen for bodywork parts information.

The public SuperClassics website was established in 2001. It provides information on classic cars, mainly from Europe. The website offers comprehensive information such as brand, model, submodel, origin of production, and years of production. Regarding restoration shops, it also provides information on their addresses. The website's structure has changed, and the manufacturers' page retrieved from web scraping in December 2023 for the manufacturer dataset is not currently the same, at least since March 2024. However, the information on classic car manufacturers likely does not need to be updated, as the origin of production for classic cars remains unchanged and does not affect the current project. For the supplier dataset, the suppliers' page retrieved in December 2023 is maintained. This website is designed for classic car owners interested in restoration activities. It offers access to suppliers' locations and contacts based on specific brands or spare part categories. Additionally, the website could interest collectors seeking to acquire a new classic car model, join a particular classic car brand or country's club, or explore museums in different locations.

The Restoparts website regards a producer of classic car spare parts operating since 1982. Restoparts claims to use the most advanced manufacturing techniques and technologies, from casting and moulding to machining and finishing, to ensure the highest industry standards of authenticity, value, fit, finishing, and availability. This website provides comprehensive information on each spare part, including dimensions, weight, type of material, type of finishing, and more. For the spare parts dataset, a specific filter was applied on the website to retrieve information only on bodywork parts, i.e., those used at RB, our illustration case. The sheet metal & body panels page was chosen. This website is designed for classic car enthusiasts and collectors to restore and maintain their classic cars. It provides a wide range of parts and accessories, presenting information on the materials, techniques implemented, the classic car models that can fit, and the price of each part.

The following subsection explains how the datasets were compiled from these two websites using web scraping techniques.

3.2.3 Web scraping techniques

The information from the manufacturers' page, suppliers' page, and sheet metal & body panels page was obtained using web scraping techniques. Due to the substantial amount of data to be processed, the cloud environment Google Colab was used for web scraping, employing libraries such as Beautiful Soup, Selenium WebDriver, and Scrapy.

The scraped datasets were exported in CSV format and imported into the Jupyter Notebook platform for data preprocessing and modelling. During preprocessing, an initial data cleaning was performed, the dataset features were described, and data normalization was carried out, as explained in the following subsection.

3.2.4 Data integration

Before data preprocessing, an initial data description of features is presented.

The classic car manufacturers' dataset is composed of 2272 classic car data points and five categorical features of type object: "Brand", "Model", "Submodel", "Origin", and "Year". The classic car suppliers' dataset comprises 3097 data points, each with three categorical features of type object: "Brand", "Title,"and "Address". Finally, the bodywork parts' dataset presents 804 data points and ten categorical features: "Name", "Product_type", "Composition", "Finish", "Weight", "Length", "Width", "Height", "Models", and "Price".

Additional data was still needed for the classic car suppliers' dataset to achieve our goals. This data includes the brands' popularity and transportation impact between suppliers, which is essential to our clustering analysis.

Firstly, Google search hits were performed. This algorithm is based on the number of user searches for each brand and model presented in the datasets. Subsequently, this data was added to both the classic car manufacturers and supplier datasets according to each brand or model presented in each dataset.

Secondly, the transportation impact between supplier locations and our illustration case in Loures was measured. The cities and countries were added to the classic car suppliers dataset through the supplier address feature, acquired from the world-cities dataset via datapackage API. An initial web search was made to identify the IATA codes from supplier cities. For each city, the IATA code was obtained (unique three-letter code assigned to each airport) from the IATA website. For towns without airports, this code was retrieved based on the nearest city airport using the Geopy API. A website search was made for the carbon footprint assessment between city airports. Subsequently, the Atmosfair website, a carbon emissions calculator for air travel, was used. On this website, the IATA codes of the supplier cities were used as the departure airports, and the IATA code of Lisbon district (for the Loures municipality) was used as the destination airport for a one-way journey with a single passenger (to ensure consistent data across all cities).

The classic car supplier dataset received additional features based on the gathered information. The final features of the datasets are described in the next section.

3.3 Data Preprocessing

The data cleaning and transformation processes are presented in the data preprocessing phase. As mentioned earlier, an initial data cleaning was performed on the datasets. Subsequently, data integration was carried out to incorporate new features into the suppliers' dataset, followed by a second data cleaning and transformation. Next, the data description indicates the final features used in the project for each dataset. Finally, the chosen normalization algorithm applied to both the manufacturer's and suppliers' datasets is described, as it is an essential step in preparing the data for modeling.

3.3.1 Data cleaning

The following steps describe the cleaning process for manufacturers, suppliers, and bodywork spare parts datasets.

3.3.1.1 Duplicated data

After analyzing the three datasets, no duplicated data was identified.

3.3.1.2 Missing values treatment

In the manufacturers' dataset, the "Brand"feature had one missing value, the "Model"feature had 10 missing values, the "Submodel"feature had 272 missing values, and the "Origin"feature had 14 missing values. These classic car records were not removed from the dataset because they contained other features with crucial information. The missing value in the "Brand"feature was identified as the Datsun 280 ZX Turbo, based on data from other features within this record. The record was then updated with the correct information. In the suppliers' dataset, there were no missing values. In the bodywork spare parts dataset, the "Product_type"feature had 37 missing values, the "Composition"feature had 58 missing values, and the "Finish"feature had 139 missing values. These bodywork spare parts records were not removed from the dataset since they could contain other crucial information for the project.

3.3.1.3 Data type conversion

After extracting features and their values into a dataset and exporting it to CSV format, all features from the three datasets were found to be categorical and of type "object". Therefore, each feature needed to be converted to the appropriate data type. In the manufacturers' dataset, the "Brand", "Model", "Submodel", and "Origin"were converted to categorical "string"data type as represented text, and "Start_year"and "End_year"features to numerical "int64"data type, as they represented integer numbers. In the suppliers' dataset, the "Brand", "Supplier", "Supplier_type", "Address", "Country", "City", "IATA", and "City_airport"features were converted to categorical "string"data type, as they contained text. On the other hand, "Version", "climate_impact_kg_co2", "flight_distance_km", "co2_emissions_kg", "climate_impact_of_contrails_ozone_formation_et-
_ kg", and "fuel_consumption_l"features were converted to numerical "int64"data type as they contained integer numbers. In the bodywork spare parts' dataset, the "Name", "Product_type", "Composition", "Finish", and "Models"features were converted to categorical "string"data types as they contained text, and "Weight", "Length", "Width", "Height", and "Price"features were converted to numerical "float64"data type as they contained numbers up to two decimal places.

3.3.1.4 Inconsistencies treatment

In the manufacturers' dataset, the ”Brand”feature contained syntax problems such as numeric values (”16”, ”18”, ”206”, ”208”, ”246”and ”308”), data inconsistencies (Citroem / Citroen, Daf / DAF, etc.), and unconventional names (Simca 9). Initially, accents were removed, and only the first letter was capitalized using the Unicode library. Additionally, hyphens were removed for ambiguous entries by adding a space and transforming the first letter of the second word into a capital letter. Acronyms were transformed only into capital letters (e.g., BMW, NSU, etc.). Most of the numeric values in the ”Brand”feature could be corrected thanks to other feature information such as the model, submodel, origin, and year: ”206”records referred to both Ferrari Dino 206 GT and Ferrari 206 SP; ”208”indicated Ferrari 208 GT4; ”246”had four records, primarily Ferrari Dino 246 GT and one Ferrari Dino 246 GTS; ”308”was identified as the Ferrari 308 GT4; ”Simca 9”was recognized as the Simca 9 classic car without a submodel. Numeric values such as ”16”and ”18”lacked information to identify a classic car model, representing only the type of automatic transmission. Consequently, five records were removed from the dataset. The ”Origin”feature in the same dataset had inconsistent country name formats (e.g., italy/Italy, Zweden/Sweden), which were altered for data consistency. One record, labeled ”1977”, had an incorrect name due to a missing country. After assessing the brand, model, and year, it was determined that the classic car was produced in Germany.

In the bodywork spare parts dataset, many records required case-by-case investigation. Some ”Composition”names included combinations of materials (e.g. ”Plastic/Aluminum”, ”Metal/Rubber”, etc.) as well as different types of steel (e.g. ”braided steel”, ”stamped steel”, etc.). Additionally, some parts used more than one composition material, such as the cowl induction system made of polyamide and aluminum, highlighting the importance of a thorough case-by-case investigation. The ”Finish”names contained a lot of repeated information (e.g. ”black”, ”black painting”, black coating”, etc.), which had to be manually altered to a single name.

3.3.2 Data transformation

3.3.2.1 Features creation

The ”Year”feature was adjusted from the manufacturers' dataset. Classic cars' start and end years were initially stored in a single column in ”XXX-XXX”format. To resolve this, non-numeric values, except for hyphens, were removed. The years were then split based on the hyphen into two new features: ”Start Year”and ”End Year.”

Some features in the suppliers' dataset had to be renamed. The ”Title”feature contained the supplier name, so it was renamed ”Supplier”. In the feature ”Category”, the Abarth brand often included a ”+”symbol followed by a number to indicate different versions of modified cars. To ensure data consistency, a ”Version”feature was added to the dataset to capture these versions. Additionally, there were records where the ”spare parts”name was associated with a brand in the ”Category”feature, resulting in multiple records for the same brand in the dataset. To handle this, a ”Supplier Type”feature was introduced. In this new feature, suppliers offering spare parts were labeled ”spare parts”, while others were categorized as ”service”, as they had a more general service. Finally, records with the same brand, supplier type, and version were grouped by supplier for enhanced data uniformity. The ”number of suppliers”feature for suppliers’ availability was also created for the data modeling. This feature represents the total count of unique suppliers associated with each classic car model. Other features were used in modeling, such as ”Google hits”for brand popularity, which represents 49 unique classic car brands, and ”CO2 impact”for the transportation impact, which means an average of the ”co2_emissions_kg”feature calculated for each classic car brand, considering that there are multiple suppliers for each brand.

3.3.2.2 Data units conversion

In the bodywork spare parts dataset, a data transformation was performed to convert pounds (lb) to kilograms (kg) using a factor of 1/0.453592 for the ”Weight”feature, inches (in) to millimeters (mm) using a factor of 1/25.4 for the ”Length”, ”Width”, and ”Height”features, and United States dollars (USD) to euros (EUR) using a factor of 1/0.85 for the ”Price”feature. Finally, all values were rounded to two decimal units. The features were renamed by deleting the initial scale indicators such as ”lb”, ”in”, and ”$”at the end of the names, which were then renamed to ”kg”, ”mm”, and ”€”respectively.

3.3.2.3 Data normalization

Features such as ”CO2 impact”, ”Google hits”, and ”number of suppliers”were normalized for the suppliers’ dataset to provide consistency across different scales. Data was normalized with the Min-Max algorithm, which performs a linear transformation to the data on a (0,1) scale based on the minimum and maximum values of each feature:

$$x_{\text{norm}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \qquad (3.1)$$

where:
x - original value
x_{min} - minimum feature value
x_{max} - maximum feature value
x_{norm} - normalized value

3.3.3 Data description

3.3.3.1 Features specifications

After the data preprocessing, datasets were exported into CSV format. The final features, their description, data type, and scale can be visualized in appendixes in tables C.1, C.2, and C.3.

The classic car manufacturers’ dataset contains 2267 data points and includes six features, four categorical and two numerical. The classic car suppliers’ dataset consists of 2915 data points, each with fourteen features, eight categorical and six numerical. Finally, the bodywork parts dataset consists of 2740 data points and ten features, five of which are categorical and five of which are numerical.

3.3.3.2 Descriptive statistics

Descriptive statistics have been used to outline key characteristics of the datasets. The manufacturer’s dataset includes 85 classic car brands, 479 models, and 1338 submodels. Citroen is the most frequent brand, producing the most classic cars. Italy leads the 12 different countries with 781 classic cars produced. The production years of classic cars range from 1928 to 2018, which means that even some recent models are now considered classic cars (Appendixes - Table D.1). Concerning the "Start_Year"feature, one should note that there is a lot of subjectivity when defining a classic car. Therefore, all vehicles have been maintained independently of their manufacturing year for this classic car website, as they may possess an intrinsic historical and cultural value to our society - according to the Historic & Classic Vehicles Alliance [44].

The suppliers' dataset consists of 2840 suppliers. Joop Stolze Classic Cars is the most frequent supplier, present on 9 classic car models. Among the 53 supplier countries, Germany has the most suppliers, totaling 533, with 525 located near Zwickau city. When comparing the climate impact between city suppliers and the Loures illustration case, the distance ranges from 520 km to 19647 km, with a fuel consumption ranging from 20.000.000 L to 1.078.000.000 L. The carbon emissions during these flights range from 73.000.000 kg to 2.725.000.000 kg (Appendixes - Table D.2 and D.3). The dataset for bodywork parts comprises 804 different parts, predominantly encompassing quarter panels. The classic car parts are primarily in steel, representing 630 out of the total bodywork parts available on the website. Notably, 494 of these parts feature an EDP coating, an electro-deposition painting that provides a protective layer guarding against corrosion, thus fortifying the material against environmental factors such as humidity and dust. These parts are compatible with 445 unique classic car models, with the Chevrolet Monte Carlo from 1970 to 1972 emerging as the most prevalent, being featured 111 times amidst the group of parts offered on the website (Appendixes - Table D.4).

3.4 Data modeling

To address our research questions, the K-Means clustering algorithm was employed. This algorithm is based on minimizing the sum of squares of distances between data points, particularly within our analysis of the classic car brands. The distances were computed based on the characteristics under examination, comprising the carbon footprint mean, Google hits, and the number of suppliers. Each brand was allocated to a cluster group based on its proximity to the cluster centroid. Each model's number of clusters was accomplished using the elbow method and silhouette score. The elbow method serves to identify the optimal number of clusters by locating the inflection point on the curve. At the same time, the silhouette score evaluates the distribution of data points within each cluster group. An approximation value near 1 indicates a strong fit, whereas proximity to -1 denotes a weaker fit.

CHAPTER 4.

Results

Contents

This chapter presents the results of the analysis after data pre-processing. The distribution of manufacturers' and suppliers' by country and classic car brands, as well as the most popular brands and models, are presented with descriptive analysis. The K-Means clustering algorithm is implemented with the three factors, and evaluated with ANOVA test.

[This page has been intentionally left blank]

Chapter 4
Results

4.1 Descriptive analysis

After the descriptive statistics to summarize our datasets' characteristics, a descriptive analysis was conducted to identify trends. Also, in this analysis, some of the research questions will be addressed.

The locations of manufacturers and suppliers are explored to compare countries with a higher concentration of classic car brands, which answers the RQ1. Data retrieved by Google search hits identifies trends, such as the top 10 most popular classic car brands and models. These trends will be compared to the brands with the most suppliers for RQ2. The bodywork parts dataset was also explored to identify trends in the material composition and finishing types used to implement new technologies, which addresses RQ4.

4.1.1 The international classic car manufacturers

International geographic locations, specifically continents, and countries of classic car origin of production, were identified for each model. Europe is leading in international classic car manufacture (90.1%), Asia comes in second as the largest manufacturer (5.3%), and finally, America (4.6%) (Figure 4.2). Regarding countries, Italy has the most manufacturers, accounting for 738 models, followed by Germany and France (Figure 4.1). These figures highlight the significant contribution of Europe to the classic car supplier industry.

Given our current understanding of the manufacturers across different locations, it would be beneficial to gain insight into the classic car brands currently produced on a large scale. The dataset includes 94 unique brands, among which Citroen stands out as the most common, with 335 classic car records, followed by Lancia and then Alfa Romeo with 265 and 198 records, respectively (Figure 4.3). Upon analyzing these results, valuable insights were obtained for comparing the geographic locations of manufacturers and suppliers and brands' availability.

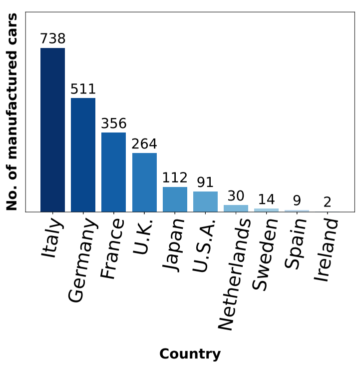

Figure 4.1: Top 10 classic car manufacturers by country.

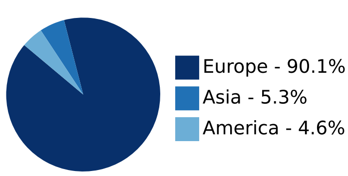

Figure 4.2: Classic car manufacturers by continent.

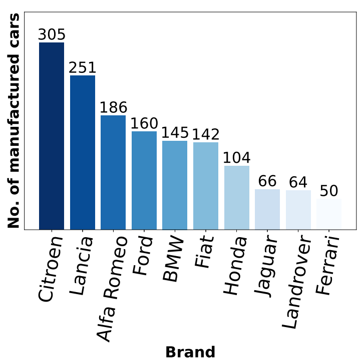

Figure 4.3: Top 10 classic car manufacturers by brand.

4.1.2 The international classic parts suppliers

As in the manufacturers' analysis, international geographic locations, including continents and countries, were identified as the classic car brand suppliers. European suppliers dominate as the primary continent in global sourcing of classic parts (51.8%), followed by America (47%), and then Asia (1.2%) contributing with a smaller number of suppliers (Figure 4.5).

From the 53 different countries and 563 different cities, it becomes evident that Germany is the primary country source of classic parts, with 533 suppliers. The United States follows closely as the second-largest contributor with 441 suppliers, and the Netherlands is the third with 404 suppliers. These countries emerged as the most predominant sources of classic car parts (Figure 4.4).

The classic car supplier dataset includes 77 unique brands. The analysis of the top 10 brands with the most suppliers unveils that Porsche stands out, followed by Volkswagen and then Mercedes-Benz (Figure 4.6). This means there is a more prosperous market of suppliers for these brands, making them a more affordable option for future replacement of classic auto parts.

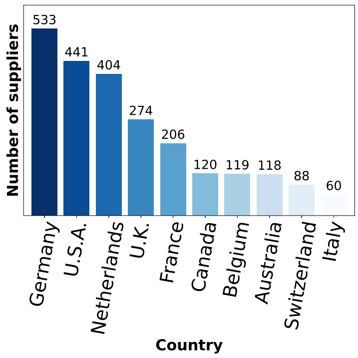

Figure 4.4: Top 10 classic car suppliers by country.

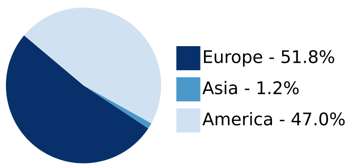

Figure 4.5: Classic car suppliers by continent.

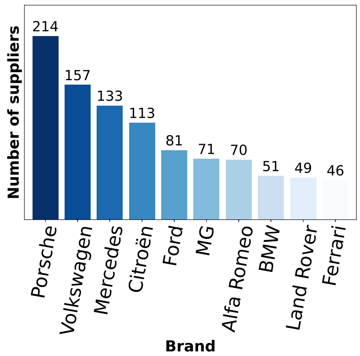

Figure 4.6: Top 10 classic car suppliers by brand.

4.1.3 The classic car bodywork parts

This section explores the production of bodywork spare parts for classic cars. The material compositions and finishing types of the bodywork parts' dataset are represented in Figure 4.7 and 4.8 respectively. It can be observed that the primary material used in the bodywork parts is steel, possibly produced using the hot rolled technique (85.6%), followed by glass fiber-reinforced plastic (7.6%), and rubber (2.6%). This indicates that steel is the most commonly used metal material, with zinc (0.9%) and aluminum (0.7%) used in smaller quantities in bodywork parts. Additionally, different forms of steel production are utilized, such as low-alloyed or converted steel, depending on the mechanical properties and cost requirements for each part. Plastic is not one of the primary materials, such as rubber, or polymers such as polypropylene (0.9%) or acrylonitrile-butadiene-styrene (ABS) (0.5%). Jute is a natural fiber plant used to produce certain classic car parts (0.4%). Regarding the paintings and coatings in the finishing stage, the electro-deposit primer (EDP) coating is the most prevalent (85.4%). Electrostatic paint is the next most common finishing type, a painting method used for classic parts production (7.6%). In addition to EDP coating, other corrosion protectors are employed for classic car parts, including zinc coating (2.6%), also known as ”galvanic coating”, and finally, electroplating chrome coating (0.9%). Other finishing types that are primarily aesthetic include black chrome (0.7%) and enameling (0.4%), which are used to achieve stylish finishes. Additionally, some coatings, such as seal natural rubber (0.9%), and plastic film (0.5%) may serve either aesthetic or surface protection purposes.

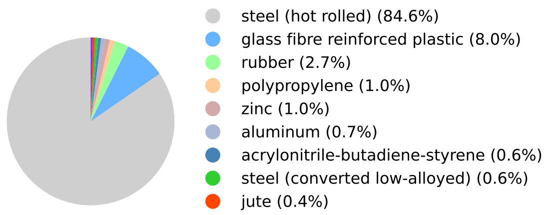

Figure 4.7: Bodywork classic parts material composition.

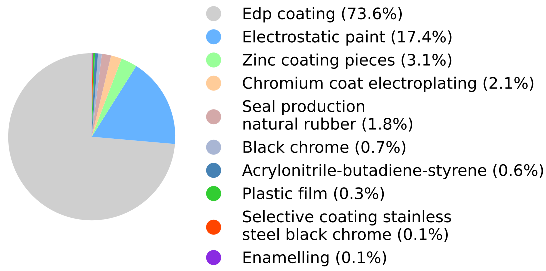

Figure 4.8: Bodywork classic parts finishing types.

4.1.4 The top 10 most popular classic car brands and models

The following analysis involves Google search hits to find trends in classic car brands and models retrieved from our manufacturers’ and suppliers’ datasets. These trends will reveal our top 10 most popular classic car brands and models. Google Hits is a powerful tool that allows one to track and analyze Google searches. By leveraging this tool, valuable insights were gained into the classic car brands and models most interesting to web surfers. This information is crucial to understand consumer preferences. According to our analysis of classic car brands, Ford, BMW, Toyota, Mercedes-Benz, and Volkswagen are the most sought-after brands by web surfers (Figure 4.9). However, Toyota does not belong to our manufacturer dataset; it is only in the supplier dataset. Although this classic car brand may interest our collectors, this manufacturer’s website does not list any Toyota models. Additionally, the supplier dataset did not include three of the most sought-after brands: Honda, Audi, and Jeep.

On the other hand, when it comes to classic car models, the Mercedes Benz 200, Mercedes Benz 500, and Mercedes Benz 250 are the most frequently searched by internet users. These models have generated the most online interest and might capture the attention of car enthusiasts and collectors alike (Table 4.1). For a broader analysis, the top 3 classic car models with the most Google hits of the top 10 most searched Google hits classic car brands was conducted. There is indistinguishable higher interest in the Mercedes-Benz 200, 500, and 250 models and BMW 3.0. For instance, the Ford 17M, Audi Front, Porsche 911, Volkswagen Golf, Chevrolet Corvette, Honda Civic, and Jeep Grand generated significant interest (Figure E.1).

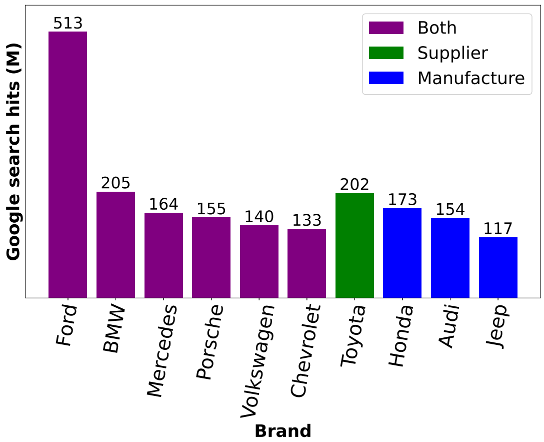

Figure 4.9: Top 10 most popular classic car brands.

Table 4.1: Top 10 most popular classic car models.

Model	Google search hits (M)
Mercedes Benz 200	326
Mercedes Benz 500	284
Mercedes Benz 250	240
BMW 3.0	196
Mercedes Benz 300	188
Mercedes Benz 600	135
Ford 17 M	93.8
Audi Front	93.4
Ford Mustang	84.4
Mercedes Benz 170	74.7

4.2 Modeling

In this section, the K-Means clustering algorithm is implemented to address our RQ3. The classic car brands are grouped based on three key factors: 1 - transportation impact; 2 - suppliers' availability; and 3 - brand popularity. The brand is treated as the dependent variable, with the three factors as independent features influencing cluster formation. Three models were developed to thoroughly assess all the factors: factor 1 vs factor 2, factor 2 vs factor 3, and factor 1 vs factor 3. These models are each presented in the following subsections.

4.2.1 Transportation Impact Vs. Suppliers' availability

Regarding the transportation impact and the number of suppliers available for each classic car brand, the elbow method shows that data should be split into three clusters, according to the inflection point equal to three (Figure F.1) with a threshold of 1.0. From the data distribution visualization (Figure F.4) for each cluster group (2, 3, 4, and 5) and the silhouette score, it is possible to verify that the best data distribution is possibly made with three clusters (silhouette score = 0.70).

On the K-Means clustering analysis with three clusters, three groups were defined, according to the carbon footprint mean (CF mean) and the number of suppliers (No. of suppliers) for each brand. To analyze the overall statistical significance of the clusters, analysis of variance (ANOVA test) determines the differences within and between features (Table 4.2). Regarding the CF mean feature, there is a significant statistical difference in the transportation impact between classic car brands when considering the number of suppliers ($F = 60.44$; p-value < 0.01). In the number of suppliers, it is also observed a significant difference, concerning the transportation impact ($F = 77.92$; p-value < 0.01).

To analyze the intrinsic characteristics within each cluster group, cluster centroids, and the distribution of data points represented by a bounding circle are presented in Figure 4.10. Cluster 1 is distinguished by a high number of suppliers and a low transportation impact. Cluster 2 represents the largest portion of the sample and is characterized by a lower transportation impact and number of suppliers. A high transportation impact and a lower number of suppliers characterize cluster 3.

An ANOVA test was also used to assess each feature's total and explained variance for each cluster group (Table F.1). Cluster 1 has a small variance in the CF mean feature (SST = 0.003) and a medium variance in the No. of suppliers feature (SST = 0.120), suggesting that this cluster is predominantly defined by a lower transportation impact, which varies across the number of suppliers for each brand. Cluster 2 is defined by high variance in CF mean and medium variance in the No. of suppliers features (SST = 0.539; SST = 0.282, respectively), indicating that this cluster is characterized by a lower number of suppliers, varying across the transportation impact. Cluster 3 shows lower variance for both the CF mean and No. of suppliers features (SST= 0.036; SST $<$ 0.001, respectively), compared to the other clusters. This cluster is indicative of a very low or non-existent number of suppliers, with a significantly high transportation impact. The findings can be summarized as follows (Table G.1):

- Cluster 1: Well-supplied eco-friendly brands representing classic cars with a significant number of suppliers and a lower transportation impact. These cars include Volkswagen, Porsche, and Citroen brands.

- Cluster 2: Poorly-supplied eco-friendly brands, with the highest sample of classic cars (42 classic car brands), featuring a lower transportation impact, but not representing the most supplied classic car brands.

- Cluster 3: Poorly-supplied non-eco-friendly brands with the negative aspect of a higher transportation impact and a lower number of suppliers. This cluster includes Subaru, Holden, Daimler, and Reliant classic car brands.

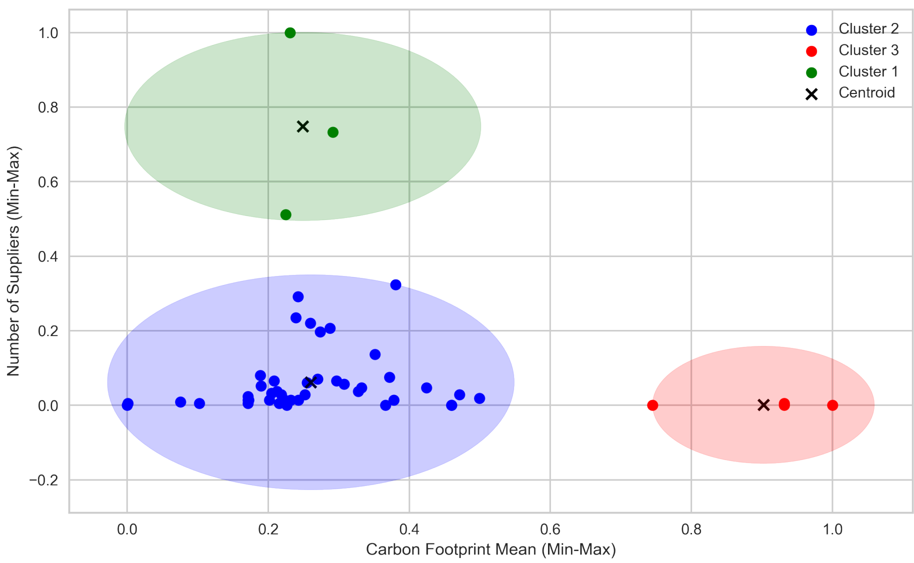

Figure 4.10: K-Means clustering (number of suppliers Vs. carbon footprint mean (CO_2/kg))

4.2.2 Transportation Impact Vs. Brand popularity

Regarding the transportation impact and brand popularity, the elbow method shows that data should be split into three clusters, according to the inflection point equal to three (Figure F.2) and the threshold of 1.0. From the data distribution results (Figure F.5) for each cluster group (2, 3, 4, and 5) and the silhouette score, it is possible to verify that the best data distribution is made with three clusters (silhouette score = 0.59).

On the K-Means clustering analysis with three clusters, three groups were formed for the CF mean measured to each brand's transportation impact and the brand's popularity (Google hits). ANOVA test determines the differences within and between cluster groups (Table 4.2). Regarding the CF mean feature, there is a significant statistical difference in transportation impact between clusters ($F = 60.46$; p-value < 0.01). The Google hits feature also shows a substantial difference in brand popularity between clusters ($F = 54.87$; p-value < 0.01).

Intrinsic characteristics in each cluster group were analyzed (Figure 4.11): Cluster 1 is the top choice, as it has a low transportation impact and a high brand popularity. Cluster 2 is characterized by the largest portion of the sample, representing a lower transportation impact and brand popularity. Cluster 3 is the least preferable, with a high transportation impact and lower brand popularity.

ANOVA test for each feature's total variance and explained variance for each cluster group (Table F.2) shows that: Cluster 1 is characterized by a higher variance in the Google hits feature ($SST = 0.173$) and a lower variance in the CF mean feature ($SST = 0.011$). This suggests that this cluster is predominantly defined by a lower transportation impact, which varies across popularity. Cluster 2 is defined by similar variances for the CF mean and Google hits features ($SST = 0.530$; $SST = 0.440$, respectively). This indicates that a lower transportation impact characterizes this cluster, but it is also due to lower brand popularity. Cluster 3 shows lower variance for both the CF mean and Google hits features ($SST= 0.036$; $SST= 0.012$, respectively), compared to the other clusters. This suggests that this cluster is more homogeneous than others and is characterized by higher transportation impact and lower brand popularity. The findings indicate the following (Table G.2):

- Cluster 1: Popular eco-friendly brands that represent a consistently lower transportation impact, with variation across the popularity of the brands. This cluster includes Ford, BMW, Toyota, Volkswagen, Porsche, and Chevrolet classic car brands.
- Cluster 2: Unpopular eco-friendly brands, constituting the largest portion of the sample (39 classic car brands), with a positive lower transportation impact, but not representing the most popular classic car brands.
- Cluster 3: Unpopular non-eco-friendly brands with the negative aspect of a higher transportation impact, as well as lower brand popularity. This cluster includes Subaru, Holden, Daimler, and Reliant classic car brands.

4.2.3 Suppliers' availability Vs. brand popularity

Regarding the suppliers' availability and the brand popularity, the elbow method shows that data should be split into three clusters, according to the inflection point equal to three (Figure F.3)

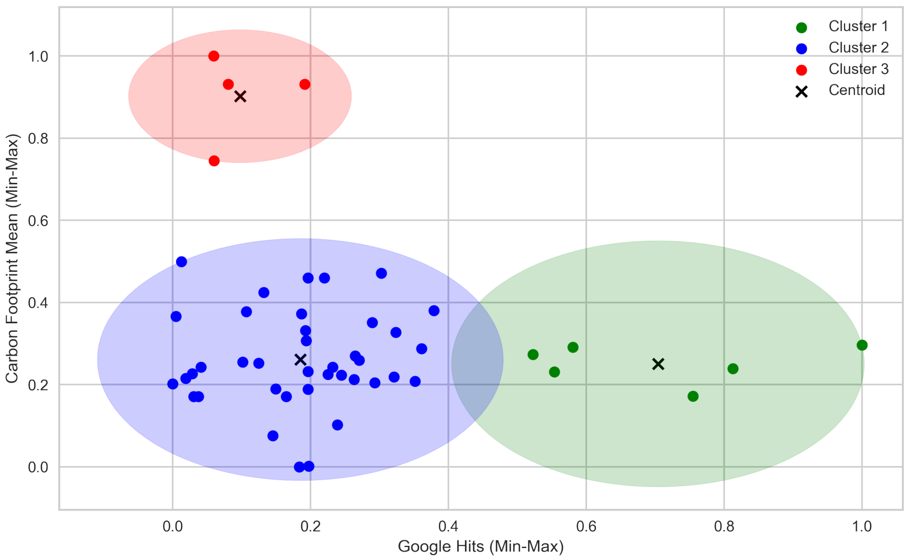

Figure 4.11: K-Means clustering (carbon footprint mean (CO_2/kg) Vs. Google search hits)

with a threshold of 1.0. From the data distribution results (Figure F.6) for each cluster group (2, 3, 4, and 5) and the silhouette score, it is possible to verify that the best data distribution is made with three clusters (silhouette score = 0.68).

On the K-Means clustering analysis with three clusters, three groups were defined for the brands' popularity and the number of suppliers for each brand.

ANOVA test determined the differences within and between cluster groups (Table 4.2). Regarding the number of suppliers, there is a significant statistical difference in the suppliers' availability for classic car brands when considering the brands' popularity ($F = 80.24$; p-value < 0.01). The Google hits feature also shows a significant difference for classic car brands, concerning the number of suppliers ($F = 49.39$; p-value < 0.01). Regarding intrinsic characteristics in each cluster group (Figure 4.12): Cluster 1 shows the highest number of suppliers and medium brand popularity. Cluster 2 indicates a lower number of suppliers for the most popular brands. Cluster 3 is characterized by the largest portion of the sample, representing a lower number of suppliers and the lowest brand popularity.

In ANOVA test (Table F.3): Cluster 1 is characterized by a medium variance in the number of suppliers (SST = 0.120) and a lower variance in the Google hits feature (SST = 0.078), suggesting that this cluster is defined by medium brand popularity, which varies across the number of suppliers. Cluster 2 is defined by a medium variance for the Google hits feature (SST = 0.116) and a small variance for the number of suppliers feature (SST = 0.033), indicating that this cluster is characterized by a lower number of suppliers, varying across brand popularity. Cluster 3 shows higher variance for both Google hits and the number of suppliers features (SST= 0.478; SST= 0.240, respectively), suggesting the most heterogeneous cluster, characterized by a low number of suppliers and brand popularity. The findings indicate the following (Table G.3):

- Cluster 1: Moderately-popular well-supplied brands that represent medium brand popularity and a high number of suppliers. This cluster includes Volkswagen, Porsche, and Citroen classic car brands.

- Cluster 2: Popular poorly-supplied brands, characterized by high brand popularity and a lower number of suppliers. This cluster includes Ford, BMW, Toyota, and Chevrolet classic car brands.
- Cluster 3: Unpopular poorly-supplied brands with the negative aspect of both the lowest number of suppliers and brand popularity, representing the largest portion of the sample (42 classic car brands).

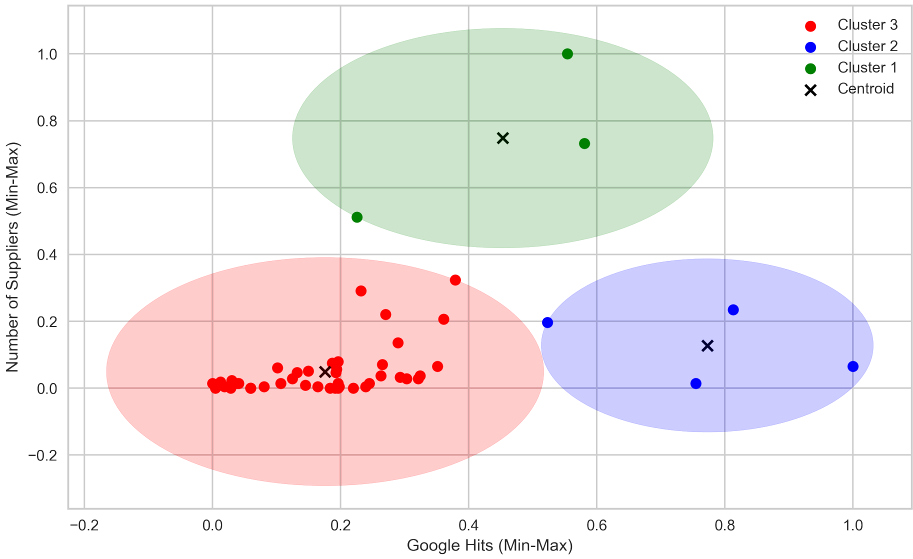

Figure 4.12: K-Means clustering (number of suppliers Vs. Google search hits)

Table 4.2: ANOVA variance tests for clustering results

Mod.*	**Feature**	**Sum of squares**			**Mean of squares**		**F test**
		Total variance (SST)	Between groups (SSB)	Within groups (SSW)	Between groups (MSB)	Within groups (MSW)	***p*-value**
1	CF mean	2.094	1.517	0.577	0.758	0.013	F 60.44 p-value < 0.01
1	No. of suppliers	1.762	1.361	0.402	0.680	0.009	F 77.92 p-value < 0.01
2	CF mean	2.094	1.517	0.577	0.759	0.013	F 60.46 p-value < 0.01
2	Google hits	2.116	1.491	0.625	0.745	0.014	F 54.87 p-value < 0.01
3	No. of suppliers	1.762	1.370	0.393	0.685	0.009	F 80.24 p-value < 0.01
3	Google hits	2.116	1.443	0.672	0.722	0.015	F 49.39 p-value < 0.01

Note: * Mod. = Model: 1 - transportation impact Vs. suppliers' availability; 2 - transportation impact Vs. brand popularity; 3 - suppliers' availability Vs. brand popularity.

CHAPTER

5. Discussion

Contents

This chapter presents the most important findings that address the research questions, along with suggesting improvements in data collection and analysis techniques.

[This page has been intentionally left blank]

Chapter 5
Discussion

5.1 International distribution of classic car manufacturers and suppliers

Descriptive analyses were conducted to assess the international distribution of manufacturers and suppliers. Although comprehensive data was collected, supplementing information from other websites could enhance the supplier network. The descriptive analysis reveals that most classic cars are manufactured in Europe, accounting for 90.1% of the production, suggesting that Europe has historical significance in classic car production. Italy emerges as the leading country in classic car manufacturing, producing 780 different classic car models. Germany and France follow Italy.

Regarding international classic car suppliers for parts and restoration services, Europe and North America are the primary sources, with 51.8% and 47% of the supply, respectively. Germany leads the way in Europe with over 500 suppliers for various classic car brands. According to the rapid systematic literature review, Germany is rapidly developing in the used car market [64], and apparently, so in the classic car market. The Netherlands is Europe's second-largest classic car supplier, followed by France. The United Kingdom also stands out, reaching numerous supplier cities.

In our results, the Citroen brand is known for its classic car manufacturing, offering a wide range of models. However, there is a greater demand for spare parts and restoration services for classic cars from brands such as Porsche, Volkswagen, and Mercedes-Benz. Citroen models may not be as widely sought after or valuable to stakeholders. Porsche, Volkswagen, and Mercedes-Benz are more popular brands, appearing in the top 10 Google search results, and generate a larger market for suppliers.

5.2 Segmentation of classic car brands: consumer preferences, transportation impact, and supplier availability

The analysis of consumer preferences was based on data obtained through Google search hits, revealing that the most interesting classic car brands for stakeholders were Ford, BMW, Toyota, Mercedes-Benz, and Porsche, which aligned with the top 10 brands from our supplier's dataset. Classic car brands' interest is possibly related to suppliers' availability. Also, most used car datasets reviewed presented these brands. Regarding classic car models, three Mercedes-Benz models are in the top three (200, 250, and 500) for stakeholders' interest. The BMW 3.0, Ford 17 M, Audi Front, and Ford Mustang are among the top 10 most interesting models. These models are also related to the most popular brands. Additionally, sentiment analysis of online reviews could complement consumer preferences regarding brand popularity in this case.

For the segmentation of classic car brands, three factors were evaluated: consumer preferences, transportation impact, and supplier availability. As reviewed in the literature, transportation significantly influences supply chain management in the auto parts market [80]. This project aimed to assess the impact of transportation on international suppliers using a Loures-based classic car restoration shop as an illustration case.

The factors evaluated required data collection through web scraping techniques. Web scraping techniques presented challenges, particularly with dynamic pages constantly altering the content, making it difficult to assess the data. To address this issue, the Selenium library was used to capture dynamic content. Extra caution was taken, as some web pages might not allow these techniques or require permission. Fortunately, the websites that have been assessed were public and did not have restrictive policies. A cloud environment was employed to scrap information from websites, which took hours to complete. At times, the data had to be split into parts to run separately to avoid constraints during scraping. The information collected enabled the construction of a comprehensive dataset encompassing classic car brands and models, the international location of parts' suppliers, brand popularity, and the calculation of the transportation impact for delivery at the location of the illustration case.

The K-Means clustering algorithm is effective at identifying patterns and segmenting the market. By grouping data points of classic car brands with similar characteristics, clusters of classic car brands were created. Three K-Means clustering models were successfully applied to the datasets, resulting in a moderate cluster distribution that rounded the 0.5 silhouette score threshold, as defined by Dalmaijer et al. [27]. While this score indicates a good degree of cluster separation, it also suggests that the cluster groups may not be entirely distinct. K-Means algorithm has the limitation of assuming spherical clusters, which might not accurately represent the data structure. To improve the results, future work could consider using more advanced methods such as density-based spatial clustering of applications with noise (DBSCAN) or Gaussian Mixture Models to better capture the underlying distribution of the data. Additionally, increasing the sample size is recommended, as 49 data points may not be sufficient to accurately represent three cluster groups. ANOVA tests were used to assess the variance between clusters. All three models exhibited significant variance (p-value < 0.01) among clusters.

Based on the results of these cluster groups, the transportation impact, suppliers' availability, and brand popularity (based on Google hits) were discretized into three categorical levels: low, medium, and high (Table H.1). Table 5.1 shows the classic car brands that achieved the best results. The original numeric values for each discretized feature can be found in the appendices section in Table H.2. These top 7 brands have been discussed in detail in the accompanying literature, available in the appendix **??** for those who are interested.

In market behavior studies, these insights are critical for market segmentation. Porsche and Volkswagen have emerged as the most appealing classic car brands for Portugal buyers. They have a substantial transportation impact on Lisbon and a vast network of parts suppliers and restoration services. Moreover, these brands have a strong presence in Google search results, making them more attractive to stakeholders seeking well-supported, sustainable, and easily accessible options in the classic car market. Porsche stands out with the highest number of suppliers and is the only brand with a high level of popularity, followed by Volkswagen, which has a medium level of popularity. Notably, Porsche and Volkswagen primarily source their parts from Germany, the leading supplier country in our analysis.

Another interesting group is comprised of the BMW and Toyota brands, which exhibit lower transportation impact for parts and boast high popularity among stakeholders. Although these brands do not have a higher level of suppliers' availability in Europe, their lower transportation impact and popularity suggest the potential for improvement in the European supply sector for BMW and Toyota classic brands.

Furthermore, the top 10 classic brands with a lower transportation impact to Lisbon were identified: FacelVega, Riley, Allard, Alvis, Panhard, Simca, NSU, DKW, Detomaso, and Bugatti. These brands likely have significant geographical proximity between their European suppliers, facilitating the transportation and availability of classic car parts to Portugal.

Table 5.1: Classic car sustainability by supplier brand (discretized values)

Brand	Most Supplied Country	Transportation impact	Suppliers' availability	Popularity
BMW	Germany	low	low	high
Citroen	The Netherlands	low	medium	low
Chevrolet	United States	medium	low	medium
Ford	Germany	medium	low	high
Porsche	Germany	low	high	high
Toyota	Germany	low	low	high
Volkswagen	Germany	medium	medium	high

5.3 Material composition and surface treatments of classic car bodywork parts

The authenticity of a classic car is a significant concern in classic car restoration shops and gallery museums [40]. Understanding material composition and surface treatments is crucial for maintaining their original cultural, historical, and aesthetic value.

Budd's introduction of all-steel body technology in the 1920s revolutionized production efficiency, leading to cost reduction and quality improvement of today's classic cars. Steel replaced wood for automobile bodies [74]. This explains why steel (85.6%) is the primary material used in the bodywork parts of classic cars produced after the 1920s in our dataset. It indicates that steel is the most commonly used metal in producing bodywork parts for classic cars in this study.

The chassis and frames of historic vehicles were initially made of forged steel or steel sheet metal. However, manufacturers have altered these material compositions over time, causing changes that can be detected analytically. This includes using ”electric steel”in the 1950s, aluminum added to reduce gases in the steel, and recycled steel scrap, etc. [91]

In 1923, Fisher pioneered lacquer instead of paint and varnish for bodies. This innovation made the close car more accessible to the general public, reducing the work time from four weeks to six hours and consequently reducing the car price [74]. In our study of the bodywork parts painting process, it was found that electrostatic paint (7.6%) is the most widely used technique for classic cars. Black chrome (0.7%) and enameling (0.4%) are sometimes used for stylish finishes.

The finish stage of many bodywork parts contains an Electrophoretically Deposited Paint (EDP) coating (85.4%), to protect against corrosion. Other corrosion protectors are also used, although in smaller quantities for bodywork parts, such as zinc coating, also known as ”galvanic coating”(2.6%), and electroplating chrome coating (0,9%).

Moreover, the subjective nature of restoration presents a critical consideration, wherein decisions often revolve around upholding the originality versus integration of non-original modern materials. According to Tutt and Hoffmann, some additives and pigments have been banned from paints due to health concerns. They must be replaced by modern substitutes, which can be analytically detectable but not damaging to the material [91]. However, it is important to use original materials in restoration whenever possible. They can be identified in non-destructive ways, such as pigments detected with a probe emitting UV radiation and fluorescence measurement of the coating (Figure 5.1).

According to [91], paint repairs should remain distinguishable from the original surfaces in the long term, but it is not clear to what extent. This topic is also important and should be further discussed, for instance in forums such as the FIVA.

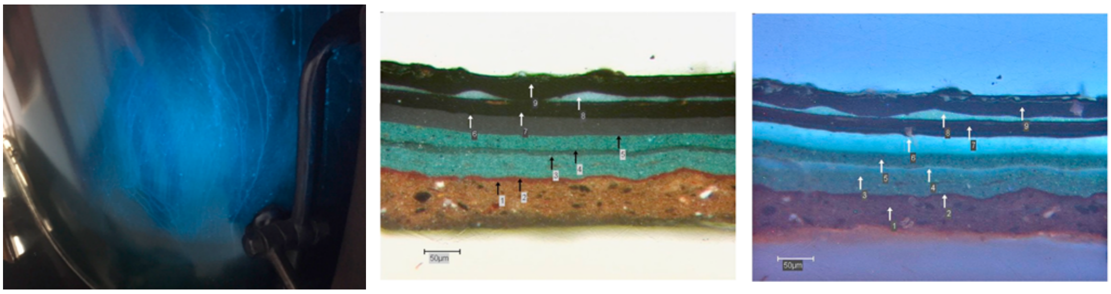

Figure 5.1: UV radiation shows sags and brush marks in the ”patina”of a Rolls-Royce Silver Ghost (left), Microscopic cross-section of a 1934 Mercedes 500 K where first paint was overpainted twice in black (layers 7-9): coating layers in the visible light spectrum (center) and UV radiation of 365-380 nm (right). Adapted from Tutt & Hoffman, 2022 [91]

.

Chapter

6.

Conclusion and future work

Contents

This chapter presents the final remarks based on the research questions, business implementation, and limitations found and suggests future analysis according to the retrieved data.

[This page has been intentionally left blank]

Chapter 6
Conclusion and future work

6.1 Conclusion

In conclusion, this study highlights several significant findings related to the international distribution of classic car manufacturers and spare parts suppliers, brand popularity, the transportation impact of parts on the Loures illustrative restore case, and concerns about authenticity in classic car restoration.

Europe is the primary classic car manufacturer, accounting for 90.1% of the market production, revealing its historical significance in the automotive industry. Major producers include Italy, Germany, and France, with Europe and North America as primary classic car parts and restoration services suppliers.

Answering RQ1, there has been a shift in the classic car market from the original European manufacturers to the suppliers extending to the USA. The fact that the classic car brands with a greater variety of manufactured models correspond to those with more spare parts suppliers (Citroën, Ford, BMW, Alfa Romeo, Land Rover, and Ferrari) indicates that demand matches supply.

Germany, the Netherlands, and France are the leading spare parts suppliers in the sector, with Germany having the most presence, as supported by the literature and the results obtained.

Regarding RQ2, some conclusions were drawn. Contrary to the Citroen brand, Porsche and Volkswagen are trendy, making them significant in the spare parts supplier market. On the other hand, the Citroen brand has a high presence in the spare parts supply market but is also less popular. It can be concluded that the availability of parts suppliers does not entirely relate to the popularity of the classic car brand, prompting a need for a change in the market. Additionally, brand popularity and the value of classic car models are significantly influenced by previous ownership, iconic appearances in movies, and race participation.

Regarding RQ3, market segmentation analysis reveals that Porsche and Volkswagen are the most sustainable classic car brands for Portuguese buyers due to their low transportation impact on Lisbon. These classic car brands also have extensive supplier networks and are highly popular. Despite having fewer suppliers on this website, BMW and Toyota also stand out for their popularity, with the research indicating a substantial number of spare parts suppliers for BMW.

Furthermore, the research provides valuable insights into the material composition and surface treatments used in classic cars, emphasizing the importance of maintaining authenticity during restoration. Answering RQ4, these materials can be replaced, and new technologies have been implemented to produce classic car parts.

From a business perspective, these findings offer valuable insights for stakeholders in the classic car market. Brands with well-established spare parts supplier networks and lower transportation impact of parts to Lisbon, such as Porsche, Volkswagen, and BMW, exhibit significant market potential. Additionally, a clear opportunity exists to expand supplier networks for the Toyota brand in Europe. While the availability of spare parts is important, we should recognize that many other factors influence the classic car market. Other key factors include the models' rarity, design, and ownership history. In contrast to the used car market, where spare parts availability is crucial for daily use, the classic car market caters to higher-end buyers interested in an art or passion investment. Leveraging these insights allows businesses to optimize some of their strategies in the classic car market, ensuring sustainable growth and meeting consumer demand for classic car restoration services.

6.2 Limitations

Although our insights were based on Google search results for the most popular brands and models for collectors, some limitations were encountered with the classic car manufacturers and spare parts suppliers datasets as they were derived from a single website. It is important to divide car models into two categories: classic cars and used cars. Classic cars are typically featured in events such as concours d'Élegance or auctions. On the other hand, used cars are sold on online websites. Also, our attempts to compare our results with other studies were hindered by their focus on used cars rather than classic cars. However, a relatively good comparison for our research was found in the specific literature analysis for individual classic car brands in the discussion section, regarding the international locations of spare parts suppliers, and brands' popularity [22, 23, 30, 31, 64].

6.3 Future work

This study requires future work, which was unfeasible due to time constraints. A comparison of the environmental impact of producing specific classic car bodywork parts in a restoration shop such as Raimundo Branco Lda for different car models would be of great interest. It would be valuable to compare the life cycle of these parts, considering the carbon footprint associated with parts degradation versus the carbon footprint of production and maintenance processes. It may be more sustainable to maintain a classic car that has low mileage rather than leave it to deteriorate, which could cause environmental damage. This could be achieved by evaluating the manufacturing processes of specific parts in a restoration shop and considering processes in a database to assess the environmental impact (e.g., Ecoinvent) using life cycle assessment software. These part assessments could be combined with classic car models using an online parts dataset that links parts with models.

Another interesting subject is that classic cars have benefited from 3D metal printing with polymer materials used for parts like the Porsche 964 crank arm and Jaguar E-type heater ducts, and metals for components of the Mercedes 300 SL coupe and Shelby Cobra [28]. This new technology could benefit classic cars with more sustainable practices using recyclable materials and enabling local production of hard-to-find parts [9, 73]. As equipment costs for 3D metal printing decrease, it will be crucial to compare their environmental impact against alternatives like international sourcing and local handcrafting, considering cost, sustainability, and authenticity.

Last but not least, there is a need for further research on the parts supply chain for the classic car market, as existing studies focus exclusively on the used car market.

Bibliography

[1] E. Abele, T. Meyer, U. Näher, G. Strube, and R. Sykes. *Global production: A handbook for strategy and implementation*. springer, Jan. 2008, pp. 1–401. DOI: 10.1007/978-3-540-71653-2.

[2] ACP. *O que é um Clássico*. Automóvel Club de Portugal (ACP). URL: https://www.acp.pt/inicio (visited on 09/19/2024).

[3] M. Ahtesham and J. Zulfiqar. “Used Car Price Prediction with Pyspark.” In: *Lecture Notes in Networks and Systems* 454 (2022), pp. 169–179. DOI: 10.1007/978-3-031-01942-5_17.

[4] M. Akhshik, A. Bilton, J. Tjong, C. V. Singh, O. Faruk, and M. Sain. “Prediction of greenhouse gas emissions reductions via machine learning algorithms: Toward an artificial intelligence-based life cycle assessment for automotive lightweighting.” In: *Sustainable Materials and Technologies* 31 (2022). DOI: 10.1016/j.susmat.2021.e00370.

[5] F. A. Alghifari, R. Andreswari, and E. Sutovo. “Used cars price prediction in DKI Jakarta using extreme gradient boosting and Bayesian optimization algorithm.” In: *Proceedings of the International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)*. 2022, pp. 01–05. DOI: 10.1109/icadeis56544.2022.10037301.

[6] A. Alhakamy, A. Alhowaity, A. A. Alatawi, and H. Alsaadi. “Are Used Cars More Sustainable? Price Prediction Based on Linear Regression.” In: *Sustainability* 15.2 (2023), p. 911. DOI: 10.3390/su15020911.

[7] A. Amellal, I. Amellal, H. Seghiouer, and M. R. Ech-Charrat. “Improving Lead Time Forecasting and Anomaly Detection for Automotive Spare Parts with A Combined CNN-LSTM Approach.” In: *Operations and Supply Chain Management* 16.2 (2023), pp. 265–278. DOI: 10.31387/oscm0530388.

[8] A. Amellal, M. Amellal, H. Seghiouer, and M. R. Ech-Charrat. “BERT-Enhanced Deep Gaussian Regression: A Unified Approach for Sentiment Analysis and Price Prediction.” In: *Proceedings of the Colloquium in Information Science and Technology, CIST*. IEEE, 2023, pp. 541–546. DOI: 10.1109/CiSt56084.2023.10409882.

[9] K. L. Ameta, V. S. Solanki, V. Singh, A. P. Devi, R. Chundawat, and S. Haque. “Critical appraisal and systematic review of 3D & 4D printing in sustainable and environment-friendly smart manufacturing technologies.” In: *Sustainable Materials and Technologies* 34 (Dec. 2022), e00481. ISSN: 22149937. DOI: 10.1016/j.susmat.2022.e00481.

[10] D. Ather, N. Chaudhary, G. Singh, Z. Tanveer Beig, and R. Kler. "Enhancing Used Automobile Valuations: A Data-Cleaning and Linear Regression Approach for Predicting Prices In Competitive Market." In: *Proceedings of the 4th International Conference on Computation, Automation and Knowledge Management, ICCAKM 2023*. IEEE, 2023. DOI: 10.1109/ICCAKM58659.2023.10449606.

[11] F. E. Ayo, J. B. Awotunde, S. Misra, S. A. Ajagbe, and N. Mishra. "A Rule-Based Deep Learning Method for Predicting Price of Used Cars." In: *Lecture Notes in Electrical Engineering* 998 LNEE (2023), pp. 845–857. DOI: 10.1007/978-981-99-0047-3_71.

[12] J. Barlow. *BMW 3.0 CSL*. TopGear, n.d. URL: https://www.topgear.com/car-news/motorsport/top-gears-coolest-racing-cars-bmw-30-csl (visited on 09/12/2024).

[13] A. Barlybayev, A. Sankibayev, Y. Kadyr, N. Amangeldy, and T. Sabyrov. "Predicting Used-Vehicle Resale Value in Developing Markets: Application of Machine Learning Models to the Kazakhstan Car Market." In: *Ingenierie des Systemes d'Information* 28.5 (2023). DOI: 10.18280/isi.280512.

[14] G. Belcher. "12. Values in World Heritage Sites." In: Open Book Publishers, Oct. 2014, pp. 181–195. DOI: 10.11647/OBP.0047.12.

[15] T. Bhaskar, S. A. Shiney, S. B. Rani, K. Maheswari, S. Ray, and V. Mohanavel. "Usage of Ensemble Regression Technique for Product Price Prediction." In: *Proceedings of the 4th International Conference on Inventive Research in Computing Applications (ICIRCA 2022)*. IEEE, 2022, pp. 1439–1445. DOI: 10.1109/ICIRCA54612.2022.9985521.

[16] M. R. S. Bhaskara and M. Wasesa. "Enhancing Car Showroom Profitability through K-Means Clustering for Customer Segmentation." In: *International Journal of Current Science Research and Review* 6.7 (2023). DOI: 10.47191/ijcsrr/v6-i7-106.

[17] N. S. Bhatt, T. Nath Pandey, S. R. Reddy, B. Jayasurya, B. B. Dash, and S. Shekhar Patra. "An Empirical Analysis of Machine Learning Algorithms for Used Car Price Prediction System." In: *Proceedings of the Global Conference on Information Technologies and Communications (GCITC 2023)*. IEEE, 2023. DOI: 10.1109/GCITC60406.2023.10426270.

[18] J. Birkinshaw. "Ecosystem businesses are changing the rules of strategy." In: *Harvard business review* 8 (2019), pp. 1–7. URL: https://hbr.org/2019/08/ecosystem-businesses-are-changing-the-rules-of-strategy.

[19] *BMW*. Classic.com, n.d. URL: https://www.classic.com/m/bmw/ (visited on 09/12/2024).

[20] B. Brahimi. "Arabic Text Mining for Price Prediction of Used Cars and Equipment." In: *Computacion y Sistemas* 26.2 (2022), pp. 1015–1025. DOI: 10.13053/CyS-26-2-3185.

[21] Z. Brock. "Design and Cars." In: (2007). URL: https://phoenixfiles.olin.edu/do/332/iiif/9eb79e53-7bb0-4c2f-9e27-43c6fb23e512/full/full/0/Zachary_Brock_final_version_2007.pdf.

[22] L. Bukvić, J. P. Škrinjar, T. Fratrović, and B. Abramović. "Price Prediction and Classification of Used-Vehicles Using Supervised Machine Learning." In: *Sustainability* 14.24 (Dec. 2022), p. 17034. ISSN: 2071-1050. DOI: 10.3390/su142417034.

[23] A. Caliskan, O. Ozturkoglu, and Y. Ozturkoglu. "Ranking of Responsible Automotive Manufacturers According to Sustainability Reports Using PROMETHEE and VIKOR Methods." In: *Advanced Sustainable Systems* 6.6 (2022). DOI: 10.1002/adsu.202100301.

[24] CCI. *Caring for vehicles - Preventive conservation guidelines for collections*. Access date: 17 sep. 2023. URL: https://www.canada.ca/en/conservation-institute/services/preventive-conservation/guidelines-collections/vehicles.html.

[25] Citroen. *CITROEN : 100 ANOS DE AUDÁCIA E INOVAÇÕES*. Consulted today. 2024. URL: https://www.citroen.pt/universo-citroen/historia.html (visited on 09/13/2024).

[26] M. V. Conlin and L. Jolliffe. *Automobile Heritage and Tourism*. Routledge, 2016. ISBN: 9781315436210.

[27] E. S. Dalmaijer, C. L. Nord, and D. E. Astle. "Statistical power for cluster analysis." In: *BMC bioinformatics* 23.1 (2022), p. 205. DOI: https://doi.org/10.1186/s12859-022-04675-1.

[28] E. Dalpadulo, A. Petruccioli, F. Gherardini, and F. Leali. "A Review of Automotive Spare-Part Reconstruction Based on Additive Manufacturing." In: *Journal of Manufacturing and Materials Processing* 6 (6 Oct. 2022), p. 133. ISSN: 2504-4494. DOI: 10.3390/jmmp6060133.

[29] C. Delagneau, R. Loi, F. Beltrami, G. Formini, R. G. Rossi, V. Molari, and M. Speziali. *Forum International sur l'Authenticité du Véhicule Historique*. Torino, Palazzo dell'Arsenale: Fédération Internationale des Véhicules Anciens (FIVA), ASI, 2004. URL: https://www.fiva.org/storage/Documents/Technical%20Commission/Forum-on-authenticity.pdf?v20240611083646.

[30] A. Dutulescu, A. Catruna, S. Ruseti, D. Iorga, V. Ghita, L.-M. Neagu, and M. Dascalu. "Car Price Quotes Driven by Data-Comprehensive Predictions Grounded in Deep Learning Techniques." In: *Electronics (Switzerland)* 12.14 (2023). DOI: 10.3390/electronics12143083.

[31] A. Dutulescu, M. Iamandei, L.-M. Neagu, S. Ruseti, V. Ghita, and M. Dascalu. "What is the Price of Your Used Car? Automated Predictions using XGBoost and Neural Networks." In: *Proceedings of the 24th International Conference on Control Systems and Computer Science (CSCS)*. May 2023, pp. 418–425. DOI: 10.1109/CSCS59211.2023.00072.

[32] k. Dziczek. *Contribution of General Motors' Manufacturing Plants to the Economies of Ten States and the United States in 2013 and 2014*. Tech. rep. Ann Arbor, Michigan: Center for Automotive Research, 2014. URL: https://www.cargroup.org/publication/contribution-of-general-motors-manufacturing-plants-to-the-economies-of-ten-states-and-the-united-states-in-2013-and-2014/.

[33] D. Evans. *History Of Herbie.* Access date: 21 oct. 2023. 2000. URL: https://www.herbiemania.com/vwt.htm.

[34] I. C. M. Eye and A. Ankokas. "Smoke signals." In: (2018). URL: https://www.ankokas.com/pubs/177.pdf.

[35] M. Farhan and J. Heikal. "Used Car Customer Segmentation Using K-Means Clustering Model With SPSS Program: Case Study Caroline.Id." In: *Jurnal Indonesia Sosial Sains* 5.3 (2024). DOI: 10.59141/jiss.v5i03.1042.

[36] F. Ferreira, V. Amaral, and F. Brito e Abreu. "Digital twinning for smart restoration of classic cars." In: Procedia Computer Science 232 (2024), pp. 2521–2530. ISSN: 1877-0509. DOI: 10.1016/j.procs.2024.02.070.

[37] D. J. Fogarty. "The Nostalgic Effect and Compounded Annual Growth Rate Used to Determine Classic Car Survival Rates." In: *Unpublished manuscript* (2018). URL: https://ijessr.com/uploads/ijessr_01_14.pdf.

[38] K. Garrity. "The American Engine." In: *Legacy* 18.1 (2018). URL: https://opensiuc.lib.siu.edu/legacy/vol18/iss1/6.

[39] E. Gegic, B. Isakovic, D. Keco, Z. Masetic, and J. Kevric. "Car price prediction using machine learning techniques." In: *TEM Journal* 8.1 (2019), pp. 113–118. DOI: 10.18421/TEM81-16.

[40] K. Gibbins. *Charter of Turin Handbook.* Tech. rep. Version 2. Fédération Internationale des Véhicules Anciens (FIVA), Nov. 2018. URL: https://fiva.org/download/turin-charter-handbook-updates-2019-english-version/.

[41] M. Gollapalli, T. A. Alqahtani, D. H. Alhamed, M. R. Alnassar, A. M. Alajmi, Y. H. Alali, M. M. Abdulqader, and A. Saadeldeen. "Intelligent Modelling Techniques for Predicting Used Cars Prices in Saudi Arabia." In: *Mathematical Modelling of Engineering Problems* 10.1 (2023), pp. 139–148. DOI: 10.18280/MMEP.100115.

[42] R. Gomes, V. Amaral, and F. Brito e Abreu. "Combining Different Data Sources for IIoT-Based Process Monitoring." In: *Proceedings of the International Conference on Information Technology and Applications (ICITA'2023).* Vol. 614. Lecture Notes in Networks and Systems. Singapore: Springer Nature Singapore, 2023, pp. 111–121. ISBN: 978-981-19-9331-2. DOI: https://doi.org/10.1007/978-981-19-9331-2_10.

[43] M. Hankar, M. Birjali, and A. Beni-Hssane. "Used Car Price Prediction using Machine Learning: A Case Study." In: *Proceedings of the 11th International Symposium on Signal, Image, Video and Communications, ISIVC 2022 - Conference Proceedings.* IEEE, 2022. DOI: 10.1109/ISIVC54825.2022.9800719.

[44] HCVA. *FAQ.* Access date: 17 sep. 2023. The Historic & Classic Vehicles Alliance. URL: https://www.hcva.co.uk/10/hcva-frequently-asked-questions.

[45] B. Hemendiran and P. N. Renjith. "Predicting the Prices of the Used Cars using Machine Learning for Resale." In: *Proceedings of the IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS).* IEEE, Feb. 2023, pp. 1–5. DOI: 10.1109/SCEECS57921.2023.10063133.

[46] S. Hoffmann. “The integration of classic cars as an alternative investment in wealth management environments and the possible influence of behavioral finance.” In: *European Journal of Multidisciplinary Studies* 3.3 (2018), pp. 6–14. DOI: https://doi.org/10.26417/ejms.v3i3.p6-14.

[47] Y. Hou, D. Myung, J. K. Park, J. Min, H.-R. Lee, A. A. El-Aty, and M.-G. Lee. “A Review of Characterization and Modelling Approaches for Sheet Metal Forming of Lightweight Metallic Materials.” In: *Materials* 16 (2 Jan. 2023), p. 836. ISSN: 1996-1944. DOI: 10.3390/ma16020836.

[48] J. Huang, S. N. Saw, W. Feng, Y. Jiang, R. Yang, Y. Qin, and L. S. Seng. “A Latent Factor-Based Bayesian Neural Networks Model in Cloud Platform for Used Car Price Prediction.” In: *IEEE Transactions on Engineering Management* (2023), pp. 1–11. DOI: 10.1109/TEM.2023.3270301.

[49] ICOM-CC. *Terminology for conservation.* Access date: 17 sep. 2023. URL: https://www.icom-cc.org/en/terminology-for-conservation.

[50] A. Ishak and T. Wijaya. “Spare part supplier selection model using decision tree classification techniques: J48 Algorithm.” In: *IOP Conference Series: Materials Science and Engineering* 801.1 (May 2020), p. 012118. DOI: 10.1088/1757-899X/801/1/012118.

[51] J. S. Jhala and D. Anand. “Comparative Analysis of Supervised Learning Algorithms for Valuating Used Car Prices.” In: *Proceedings of the International Conference on Advancement in Computation & Computer Technologies (InCACCT).* 2023, pp. 265–270. DOI: 10.1109/incacct57535.2023.10141827.

[52] C. Jin. “Price Prediction of Used Cars Using Machine Learning.” In: *Proceedings of the International Conference on Emergency Science and Information Technology (ICESIT).* IEEE, Nov. 2021, pp. 223–230. DOI: 10.1109/ICESIT53460.2021.9696839.

[53] S. Joest. “The Citroen History after Shutdown of the Conservatoire: Quo Vadis?” In: *Amicale Citroen International* (July 2024). URL: https://www.amicale-citroen-internationale.org/2024/citroen-history-quo-vadis/.

[54] J. R. Krishnan and V. Selvaraj. “Predicting Resale Car Prices Using Machine Learning Regression Models With Ensemble Techniques.” In: *Proceedings of the AIP Conference.* Vol. 2516. American Institute of Physics Inc., 2022. DOI: 10.1063/5.0108560.

[55] A. Kumar, R. B. Chinnam, and A. Murat. “Hazard rate models for core return modeling in auto parts remanufacturing.” In: *International Journal of Production Economics* 183.B, SI (Jan. 2017). Cited by: 26, pp. 354–361. ISSN: 0925-5273. DOI: 10.1016/j.ijpe.2016.07.002.

[56] D. Laurs and L. Renneboog. “My kingdom for a horse (or a classic car).” In: *Journal of International Financial Markets, Institutions and Money* 58 (Jan. 2019), pp. 184–207. ISSN: 10424431. DOI: 10.1016/j.intfin.2018.10.002.

[57] Y. Li, Y. Li, and Y. Liu. "Research on used car price prediction based on random forest and LightGBM." In: *Proceedings of the IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA)*. Oct. 2022, pp. 539–543. DOI: 10.1109/ICDSCA56264.2022.9988116.

[58] E. Liu, J. Li, A. Zheng, H. Liu, and T. Jiang. "Research on the Prediction Model of the Used Car Price in View of the PSO-GRA-BP Neural Network." In: *Sustainability (Switzerland)* 14.15 (2022). DOI: 10.3390/su14158993.

[59] C. Longani, S. Prasad Potharaju, and S. Deore. "Price Prediction for Pre-Owned Cars Using Ensemble Machine Learning Techniques." In: *Recent Trends in Intensive Computing*. 2021. DOI: 10.3233/apc210194.

[60] C. Lu, G. Shang, L. Xu, H. Shao, and B. Zhang. "Sales Volume Forecast of Typical Auto Parts Based on BiGRU: A Case Study." In: *Proceedings of the E3S Web of Conferences*. Vol. 409. ICMSEM, 2023. DOI: 10.1051/e3sconf/202340904008.

[61] S. L. Y. Luo, A. M. J. Kanaan, R. K. Ayyasamy, G. C. Meng, B.-Y. Ooi, and C. M. Tyng. "Revolutionizing the Malaysian Used Car Market: A Machine Learning Approach to Transparent Pricing." In: *Proceedings of the 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS)*. Sept. 2023, pp. 154–159. DOI: 10.1109/AiDAS60501.2023.10284620.

[62] E. V. Lyapuntsova and T. A. Boyko. "Development of a Predictive System Model Using Big Data in the Transport Sector." In: *Proceedings of the E3S Web of Conferences*. Vol. 403. EDP Sciences, 2023. DOI: 10.1051/e3sconf/202340307025.

[63] Z. Lyu and Y. Li. "Construction and empirical analysis of remanufactured parts acceptance degree system based on slice inverse regression method." In: *Proceedings of the ACM International Conference Proceeding Series*. ACM, 2021, pp. 2132–2139. DOI: 10.1145/3482632.3484114.

[64] X. Ma. "Analysis of the Development Prospect of China's Used Automobile Market." In: *BCP Business & Management* 38 (2023), pp. 2863–2869. DOI: 10.54691/bcpbm.v38i.4201.

[65] A. Madar and N. A. Neacșu. "Quality Management-a Factor for Improving Sustainability in the Automotive Industry." In: *Ovidius" University Annals* 20.1 (2020), pp. 688–697. ISSN: 2393-3127. URL: https://stec.univ-ovidius.ro/html/anale/RO/2020/Section%204/20.pdf.

[66] D. R. Manna, G. Marco, B. L. Khalil, S. Meier, et al. "Sustainable markets: Case study of Toyota motor sales, USA, inc." In: *Journal of Business Case Studies (JBCS)* 7.3 (2011), pp. 63–72. DOI: https://doi.org/10.19030/jbcs.v7i3.4265.

[67] N. Monburinon, P. Chertchom, T. Kaewkiriya, S. Rungpheung, S. Buya, and P. Boonpou. "Prediction of prices for used car by using regression models." In: *Proceedings of the 5th International Conference on Business and Industrial Research (ICBIR)*. May 2018, pp. 115–119. DOI: 10.1109/ICBIR.2018.8391177.

[68] Z. A. Msiza and P. A. Owolawi. "Advancing Used Car Price Prediction in South Africa: An Empirical Examination of Machine Learning Techniques." In: *International Conference on Artificial Intelligence and its Applications* 2023 (2023), pp. 189–196. DOI: 10.59200/icarti.2023.027.

[69] J. Murta, V. Amaral, and F. Brito e Abreu. "Trusted Evidences on the Digital Transformation of Classic Cars Restoration." In: *Proceedings of the 5th International Conference on Industry 4.0 and Smart Manufacturing (ISM 2023)*. Vol. 232. Procedia Computer Science. 2024, pp. 254–263. DOI: 10.1016/j.procs.2024.01.025.

[70] M. Nandan and D. Ghosh. "Pre-owned car price prediction by employing machine learning techniques." In: *Journal of Decision Analytics and Intelligent Computing* 3.1 (2023), pp. 167–184. DOI: 10.31181/jdaic10008102023n.

[71] C. V. Narayana, C. L. Likhitha, S. Bademiya, and K. Kusumanjali. "Machine Learning Techniques to Predict the Price of Used Cars: Predictive Analytics in Retail Business." In: *Proceedings of the 2nd International Conference on Electronics and Sustainable Communication Systems, ICESC 2021*. IEEE, 2021, pp. 1680–1687. DOI: 10.1109/ICESC51422.2021.9532845.

[72] C. V. Narayana, N. O. G. Madhuri, A. Nagasindhu, M. Aksha, and C. Naveen. "Second Sale Car Price Prediction using Machine Learning Algorithm." In: *Proceedings of the 7th International Conference on Communication and Electronics Systems (ICCES 2022)*. IEEE, 2022, pp. 1171–1177. DOI: 10.1109/ICCES54183.2022.9835872.

[73] M. R. Nichols. "How does the automotive industry benefit from 3D metal printing?" In: *Metal Powder Report* 74 (5 Oct. 2019), pp. 257–258. ISSN: 0026-0657. DOI: 10.1016/j.mprp.2019.07.002.

[74] P. Nieuwenhuis and P. Wells. "The all-steel body as a cornerstone to the foundations of the mass production car industry." In: *Industrial and Corporate Change* 16.2 (2007), pp. 183–211. DOI: 10.1093/icc/dtm001. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-34447554437&doi=10.1093%2ficc%2fdtm001&partnerID=40&md5=a033bcbd1a8799341e8269ad43623ac8.

[75] N. Pal, P. Arora, P. Kohli, D. Sundararaman, and S. S. Palakurthy. "How Much Is My Car Worth? A Methodology for Predicting Used Cars' Prices Using Random Forest." In: *Advances in Information and Communication Networks*. 2018, pp. 413–422. DOI: 10.1007/978-3-030-03402-3_28.

[76] Porsche Museum. *Exhibition*. 2024. URL: https://www.porsche.com/international/aboutporsche/porschemuseum/aboutthemuseum/exhibition/.

[77] B. E. Putro and D. Indrawati. "Data Mining Analytics Application for Estimating Used Car Price during the Covid-19 Pandemic in Indonesia." In: *Jurnal Ilmiah Teknik Industri* 21.2 (2022), pp. 149–161. DOI: 10.23917/jiti.v21i2.18975.

[78] M. Rane, M. Patil, N. Rane, and A. Amune. “Random Forest Regressor Approach for Predicting Resale Value of used Vehicles (RFRVP).” In: *Proceedings of the Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*. 2023, pp. 1–6. DOI: 10.1109/ICAECT57570.2023.10118285.

[79] B. Rieger. “From People’s Car to New Beetle: The Transatlantic Journeys of the Volkswagen Beetle.” In: *The Journal of American History* (2010).

[80] N. Rojniruttikul. “Development of strategic supply chain management for autoparts and components manufacturing company.” In: *Proceedings of the ACM International Conference Proceeding Series*. ACM, 2019, pp. 52–55. DOI: 10.1145/3328886.3328902.

[81] N. Sanguansaringkhan and H. Liu. “Forecasting and inventory model for genuine automotive spare parts wholesale distributor.” In: *Proceedings of the 2020 IISE Annual Conference*. 2020, pp. 1282–1287. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85105610203&partnerID=40&md5=a33ebfa24d2291cb94d976e53635e4ce.

[82] J. Sarmento, H. Oliveira, H. arneir, B. Coimbra, D. Pacheco, H. M. de Carvalho, and A. Simões. *Proposta de Lei n.º 109/XV/2.ª: Orçamento do Estado para 2024*. Proposta de Aditamento. 2023.

[83] S. K. Satapathy, R. Vala, and S. Virpariya. “An Automated Car Price Prediction System Using Effective Machine Learning Techniques.” In: *2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*. May 2022, pp. 402–408. DOI: 10.1109/CISES54857.2022.9844350.

[84] S. Shah, M. Telrandhe, P. Waghmode, and S. Ghane. “Imputing missing values for Dataset of Used Cars.” In: *Proceedings of the 2nd Asian Conference on Innovation in Technology (ASIANCON)*. Aug. 2022, pp. 1–5. DOI: 10.1109/ASIANCON55314.2022.9908600.

[85] N. Shanti, A. Assi, H. Shakhshir, and A. Salman. “Machine Learning-Powered Mobile App for Predicting Used Car Prices.” In: *Proceedings of the ACM International Conference Proceeding Series*. ACM, 2021, pp. 52–60. DOI: 10.1145/3502300.3502307.

[86] S. Shaprapawad, P. Borugadda, and N. Koshika. “Car Price Prediction: An Application of Machine Learning.” In: *Proceedings of the 6th International Conference on Inventive Computation Technologies (ICICT 2023)*. IEEE, 2023, pp. 242–248. DOI: 10.1109/ICICT57646.2023.10134142.

[87] J. Sharma and S. Kumar Mitra. “Developing a used car pricing model applying Multivariate Adaptive regression Splines approach.” In: *Expert Systems with Applications* 236 (2024). DOI: 10.1016/j.eswa.2023.121277.

[88] V. Skribans and S. W. Hegerty. “Data mining application for used-car market analysis.” In: *Proceedings of the 10th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2019)*. Vol. 1. International Institute of Informatics and Systemics, 2019, pp. 35–40. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066024286&partnerID=40&md5=bbe8046dc159e3b10c70d85b9ee69a43.

[89] Y. Song, Y. Yan, R. Zhang, D. Xu, and F. Wang. "Manufacture of the die of an automobile deck part based on rapid prototyping and rapid tooling technology." In: *Journal of Materials Processing Technology* 120.1-3 (2002), pp. 237–242. DOI: 10.1016/S0924-0136(01)01165-7.

[90] P. Tom. "Composite Materials and their Recycling." In: *i-manager's Journal on Material Science* 4.2 (2016), p. 37. DOI: 10.26634/jms.4.2.8129.

[91] G. Tutt and S. Hoffmann. "Forensics in motion-Historic vehicles, genuine or fake?" In: *Forensic Science International: Synergy* 4 (2022), p. 100218. DOI: https://doi.org/10.1016/j.fsisyn.2022.100218.

[92] F. Uysal. "Comparative Analysis of Various Machine Learning and Deep Learning Approaches for Car Resale Price Prediction in the Turkish Market." In: *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi* (2023). DOI: 10.28948/ngumuh.1353526.

[93] J. Varshitha, K. Jahnavi, and C. Lakshmi. "Prediction Of Used Car Prices Using Artificial Neural Networks And Machine Learning." In: *Proceedings of the International Conference on Computer Communication and Informatics, ICCCI 2022*. IEEE, 2022. DOI: 10.1109/ICCCI54379.2022.9740817.

[94] V. Vesovic, D. Jovanovic, M. Arsic, Z. Avramovic, S. Sofijanic, B. Djorovic, N. Gospic, N. Tomovic, D. Milosevic, M. Dobric, S. Arsic, D. Kostadinovic, S. Kalac, V. Peulic, T. Fazekas, D. Rosulj, D. M. Mitrovic, J. Hodzic, M. Prelevic, and M. Andric. "Identification of the Right Moment for Motor Vehicle Replacement—Life-Cycle Analysis in Serbia and Montenegro." In: *Sustainability (Switzerland)* 14.5 (2022). DOI: 10.3390/su14052744.

[95] V. Viswanatha, A. Ramachandra, H. Vachan, S. S. Shetty, et al. "Predicting the Price of used Cars using Machine Learning." In: *Proceedings of the International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)*. Oct. 2023, pp. 1–6. DOI: 10.1109/EASCT59475.2023.10393486.

[96] Volkswagen. *50 Years of Golf: Volkswagen Shows a Rare Duo at the Bremen Classic Motorshow*. Wolfsburg, 2024. URL: https://www.volkswagen-newsroom.com/en/press-releases/50-years-of-the-golf-volkswagen-shows-a-rare-duo-at-the-bremen-classic-motorshow-18095.

[97] Volkswagen. *A World Bestseller Celebrates Its 50th Birthday – Volkswagen Started Production of the First Golf on 29 March 1974*. Wolfsburg, 2024. URL: https://www.volkswagen-newsroom.com/en/press-releases/a-world-bestseller-celebrates-its-50th-birthday-volkswagen-started-production-of-the-first-golf-on-29-march-1974-18313.

[98] Volkswagen. *History*. URL: https://www.volkswagen-newsroom.com/en/history-3693.

[99] F. Wang, X. Zhang, and Q. Wang. "Prediction of Used Car Price Based on Supervised Learning Algorithm." In: *Proceedings of the International Conference on Networking, Communications and Information Technology (NetCIT)*, pp. 143–147. DOI: 10.1109/NetCIT54147.2021.00036.

[100] J. P. Womack, D. T. Jones, and D. Roos. *The machine that changed the world: The story of lean production-Toyota's secret weapon in the global car wars that is now revolutionizing world industry*. Simon and Schuster, 2007.

[101] D. R. Woodward and P. Guimarães. "BMW in South Carolina: An Update." In: *Business & Economic Review* 55.2 (2009), pp. 13–21.

[102] W. Xuping, M. Haohao, F. Yanwei, M. Lisha, H. Weiliang, and C. Miaomiao. "Research on Used Car Valuation Model Based on SVM Fusion Learning." In: *Proceedings of the 2023 Asia Conference on Advanced Robotics, Automation, and Control Engineering (ARACE 2023)*. IEEE, 2023, pp. 142–147. DOI: 10.1109/ARACE60380.2023.00029.

[103] S. Yilmaz and I. H. Selvi. "Price Prediction Using Web Scraping and Machine Learning Algorithms in the Used Car Market." In: *Sakarya University Journal of Computer and Information Sciences* 6.2 (2023), pp. 140–148. DOI: 10.35377/saucis...1309103.

[104] S. S. M. Yusoh, D. A. Wahab, and A. H. Azman. "Analysis of automotive component design for reparation using additive manufacturing technology." In: *International Journal of Integrated Engineering* 12 (5 2020), pp. 20–26. ISSN: 2229838X. DOI: 10.30880/IJIE.2020.12.05.003.

[105] H. Zhang. "Prediction of Used Car Price Based on LightGBM." In: *Proceedings of the 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering, AEMCSE 2022*. IEEE, Apr. 2022, pp. 327–332. DOI: 10.1109/AEMCSE55572.2022.00073.

[106] N. Zhang, K. Kang, F. Zhang, and Q. Feng. "Risk Warning and Prevention of Supply Chain in Automobile Industry Based on Data Mining Technology." In: *Proceedings of the International Conference on Industrial IoT, Big Data and Supply Chain (IIoTBDSC)*. Sept. 2022, pp. 278–282. DOI: 10.1109/IIoTBDSC57192.2022.00058.

Appendix A.

Top 7 classic car brands: brief research

During the dissertation analysis, the K-means clustering algorithm yielded the best results for the top 7 classic car brands, and brief research was conducted on these brands to gain insights into market analysis and brand popularity.

BMW

Bayerische Motoren Werke (BMW), established in 1916, is the fifth most manufactured classic car brand with the highest number of unique submodels.

According to a study by Abele et al., BMW is one of the most successful automotive companies, with diverse manufacturing strategies that involve producing cars and large components in high-cost locations with highly skilled labor forces. In figure A.1, it is observed that most of BMW's production is concentrated in Europe (Germany, Austria, and the UK) [1].

Specific models from this brand, such as BMW 3.0, mainly the CSL submodel (Figure A.2, had only 165 models produced in the 1960s and created with Alpina's help, was produced in body styles including sedan, coupe, wagon, and convertible. Initially a luxury sports car, it was adapted for car races in 1972 with batmobile submodel.

BMW 3.0 is currently being sold with 181 cars at an average price of $25,765. However, the BMW 3.0 CSL is not currently available for sale, probably due to its rarity. Nonetheless, one of these models was sold in 2021 for $40,500, highlighting the intrinsic cultural value of this particular BMW model [19].

Chevrolet

After the Second World War, Chevrolet became a symbol of American industry and was very popular[37], which could be why it is prevalent in our dataset as well.

The 1953 Chevrolet Corvette, the most popular model of Chevrolet in our analysis, was the first mass-produced car made with glass fiber reinforced polymer (GFRP), incorporating a steel chassis [90]. This material helped reduce the vehicle's weight, enhancing its performance as a

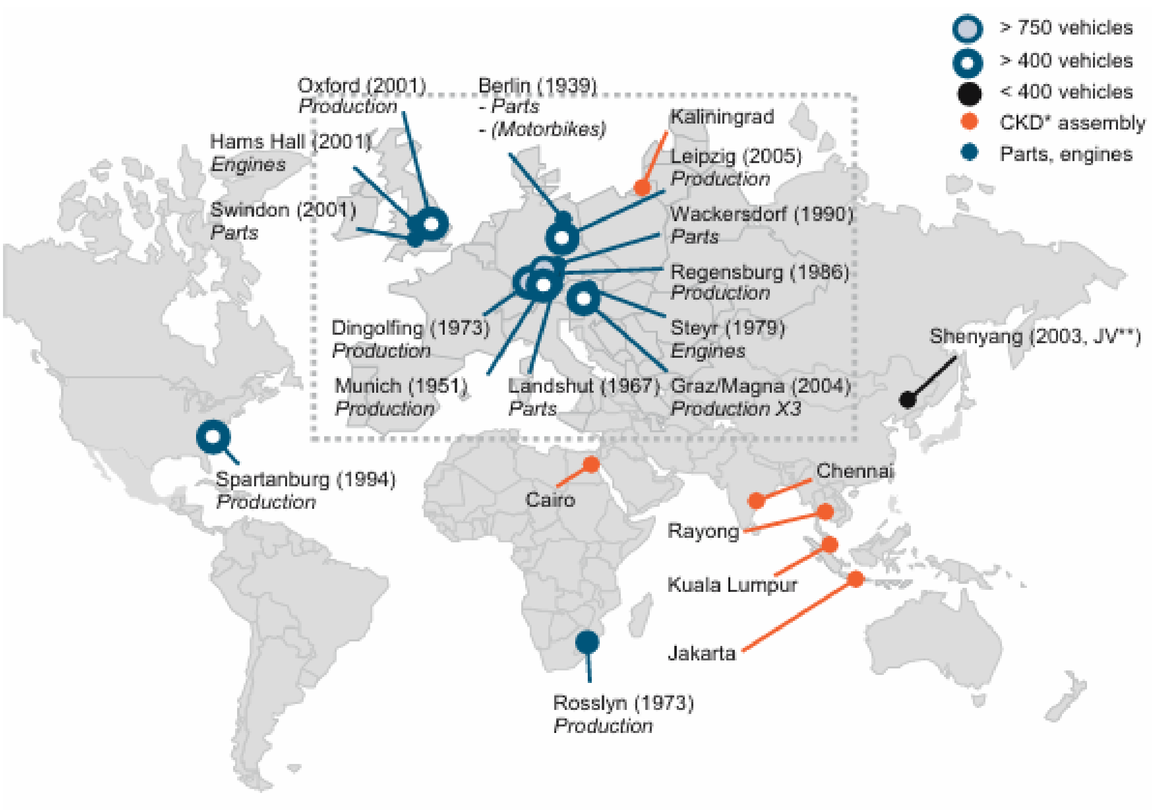

Figure A.1: Production network of the BMW Group. Adapted from *global production*: A handbook for strategy and implementation (p.123) by Abele et al., 2008, Springer.

Figure A.2: BMW 3.0 CSL (left) and BMW 3.0 CSL Batmobile (right). Adapted from *TopGear*, by Jason Barlow [12].

sports car in America, competing with some of the finest sports cars worldwide [38].

Citroen

Citroen is a French automotive manufacturer founded by André Citroen in 1919. Car enthusiast and clubs are organized worldwide to share their passion. In the 1970s, a global network called Amicale Citroen Internationale (ACI) was established, connecting 70000 club members in 47 countries to each other (Figure A.3). This makes it the world's largest non-profit vehicle enthusiast network of a brand [53]. This indicates its significant human interest.

It is known for its hydropneumatic suspension of the Citroen DS19, the gangster limousine, and the 2 CV models (Figure A.4). In the words of Joest: "Let's preserve it together in a way that makes it enjoyable to experience and makes it part of our identity"[53].

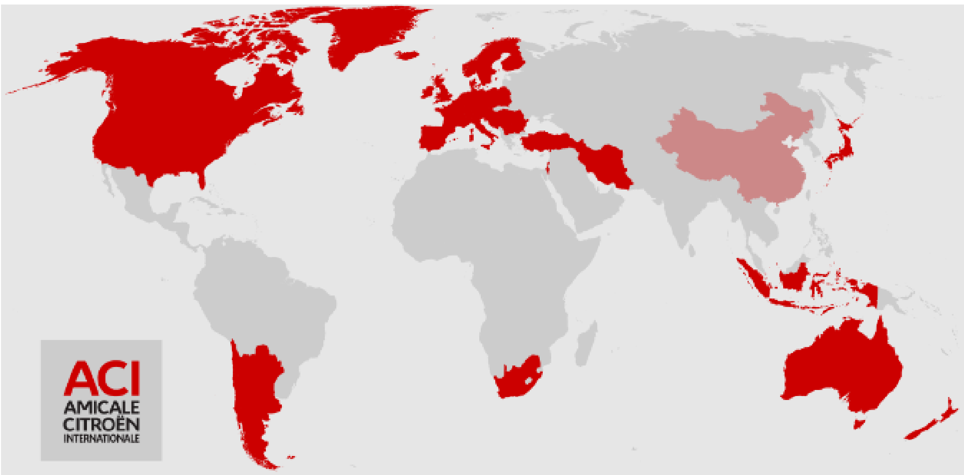

Figure A.3: ACI country network map. Adapted from *ACI*.

Figure A.4: Citroen DS19 (left), Citroen 2CV (center), and Citroen limousine (right). Adapted from *classic*.

Ford

Henry Ford pioneered high-volume mass production of classic brand Ford in the 1920s [100]. The Ford brand has a strong market presence, with 48651 units for sale on classics.com, generating a dollar volume of 1.7 billion [19].

Much literature is available on the Ford Model T and Ford Mustang (Figure A.5), due to its collector's interest. In the eyes of Americans, the classic Ford Model T was viewed as a small car, contrary to European perceptions. By 1915, Ford was recognized as the leading auto manufacturer in England. The Ford Mustang has been the best-selling car since the Model T [100].

By 1964, one million Ford Mustangs had been sold within the first eighteen months of its production. This iconic model was featured in "Bullitt,"with Steve McQueen behind the wheel [34]. The Ford Mustang is considered the quintessential American muscle car, and its popularity has endured for over forty years. It remains a highly sought-after model in the United States, frequently appearing at classic car shows, restoration clubs, and on open roads [21].

Figure A.5: Ford Model T (left), and Ford Mustang (right). Adapted from *classic*.

Porsche

Ferdinand Porsche was involved with the Porsche brand right from the start. In 1898, he worked on the Egger-Lohner C.2, an electric car now displayed at the Porsche Museum (Figure A.6). The original manufacturers of the Porsche brand still operate classic car shops that provide spare parts and offer factory restorations [46].

The brand's performance in adding value can be influenced by its previous owners. For example, the Porsche 911 used by Steve McQueen was sold for six times the model's value [46]. According to a Porsche Index by Laurs and Renneboog, this classic car brand can provide a return of up to 9.44%. Porsche is popular among collectors. The Porsche 911 Carrera RS (Figure A.6) is considered one of the ten best-looking classic cars of all time, as ranked by Total Car Score [56].

Figure A.6: Porsche Egger lohner C.2 (left), Porsche factory (center), and Porsche 911 Carrera RS (right). Adapted from *porsche*.

Toyota

The Toyota Motor Company was founded in Nagakute, Japan, in 1933, possibly by Sakichi Toyoda, who was a developer of loom machines [26]. In 1957, Toyota was also established in the United States and became the third best-selling import brand in the country, with the model Corona designed explicitly for the American market [66].

Toyota pursues an international manufacturing footprint strategy and is on track to become the most prominent global original equipment manufacturer (OEM) [1]. The company has also implemented environmental sustainability initiatives to protect the environment and its consumers [66].

Volkswagen

In 1937, the Volkswagen headquarters in Wolfsburg, Germany, commenced operations for the Kafer project, commonly known as the Beetle model. Initially conceived as the ”people's car”at the behest of Adolf Hitler, the vehicle aimed to offer a simple and affordable transportation solution for families. However, production was disrupted by the outbreak of war. After the war, the British administration ordered the production of a Volkswagen saloon, with the type 1 (Kafer) and type 2 models added in 1950, leading to a significant German economic revival [98].

The success of the Volkswagen Beetle in the United States symbolized strong transatlantic ties between Germany and America. Notably, the model exerted considerable cultural influence, attesting to the judicious American endorsement of European products and highlighting Germany's postwar recovery [79].

Volkswagen ranks among the foremost global automotive entities in terms of units sold and has adeptly cultivated extensive networks of suppliers and distributors for over five decades. The company has negotiated with its suppliers to safeguard uniform quality standards across the German automotive industry [18, 65].

The Volkswagen Golf, introduced in 1974, swiftly attained iconic status. The first million units were sold within two years, rendering it the highest-selling European car ever. The Golf's ascendancy has surpassed that of the Beetle and permeated society, emerging as a popular choice for driving schools and consumers, regardless of fuel type. The EA276 model, originating in 1969 and drawing inspiration from the Beetle, served as a predecessor to the Golf, subsequently evolving into another bestseller (Figure A.7) [96, 97]. This elucidates the preeminent status of the Golf model as the most sought-after classic Volkswagen brand, as ascertained through our comprehensive analysis utilizing Google hits.


Figure A.7: Picnic on the Klieversberg 1956 with Volkswagen Beetle (left), adapted from *volkswagen*, and Volkswagen Golf, and Volkswagen EA276 for Bremen Classic Motorshow (right), adapted from *volkswagen*.

[This page has been intentionally left blank]

Appendix B.

Rapid Systematic Literature Review

This appendix presents a summary of the 60 accepted articles by Parsifal for the rapid systematic literature review. These articles are arranged into three tables in alphabetic order. Each article includes the following information:

- **Topic**: topic of analysis (e.g., price prediction)
- **Car type**: type of car evaluated in the analysis (e.g., used car)
- **Features**: the main dataset features (e.g., brand, model, year)
- **Methods**: the methods used for data collection and data analysis (e.g., web scraping, modeling)
- **Data split**: the split of the data when applied (e.g., 70% train, 30%test)
- **Best result**: the model with performance metrics when applied (e.g., linear regression accuracy (acc.) = 90%)
- **Quality**: article quality evaluated with Parsifal
- **Cits. year**: article Google citations per year, where the number of citations is divided by the number of years since article publication
- **Journal rank**: SJR or JCR article best rank

Table B.1: Summary of articles in systematic literature review (authors A - G)

Reference*	Topic	Car type	Features	Methods	Data Split	Best Result	Quality	Cits./ year	Journal Rank
Akhshik et al, 2022 (J) [4]	env., parts	vehicle		descriptive analysis		linear regression (acc. = 95%)	15.0	0.0	
Alghifari et al., 2022 (J) [5]	brands, price	used car		descriptive analysis, scraping		linear regression (acc. = 95%)	18.0	4.0	Q2
Alhakamy et al, 2023 (J) [6]	price	used car	model, year, engine, fuel type, cylinders, transmission, wheels, doors, market category, vehicle size, vehicle style, mileage	descriptive analysis	80% train, 20% test	linear regression	16.0	2.0	Q2
Amellal et al, 2023a (C) [8]	price	vehicle		modelling		sentiment analysis + DGP (MSE = 0.0015)	16.0	0.0	
Amellal et al, 2023b (J) [7]	logistic parts	vehicle		correlation, feature sel., modelling		LSTM autoencoder with OCSVM (acc. = 0.94)	15.0	3.0	Q2
Anthesham & Zulfikar, 2022 (C) [3]	price	used car	adno, name, price, year, location, mileage, city, transmission, engine capacity, engine type, color, assembly, body type, URL, last updated	correlation, descriptive analysis, feature sel., modelling		gradient boost tree regression (R^2 = 89.67%)	15.0	4.0	
Ather et al, 2023 (C) [10]	price	used car	make, model, year, price, kilometers driven, fuel type	descriptive analysis, feature sel., modelling		linear regression ($R^2 = 0.9986$)	15.0	0.0	
Ayo et al, 2023 (C) [11]	price	used car		modelling		ANN (RMSE = 1.564)	18.0	0.0	
Barlybayev et al, 2023 (J) [13]	brands, price	used car	price, manufact., year, transm., local, generation, body, fuel, mileage, color, etc.	correlation, descriptive analysis, modelling, scraping		bagged trees (R = 94)	16.0	0.0	
Bhaskar et al., 2022 (C) [15]	price	used car	car type, selling price, present price, km driven, fuel type, seller type, transmission, owner	descriptive analysis, modelling		ensemble regression (acc. = 94.47%)	15.0	12.0	
Bhaskara & Wasesa, 2023 (J) [16]	brands, price	used car	transmission, fuel type, cc, days, age, car type, profit			clustering	15.0	0.0	Q4
Bhatt et al, 2023 (C) [17]	price	used car	local, year, manuf., model, cond., cyl., fuel, transm., size, color, lat, long, price, etc.	correlation, descriptive analysis, modelling		XGBoost ($R^2 = 0.9297$)	17.0		
Brahimi, 2022 (J) [20]	price	used car	year, mileage, seats, price	modelling, scraping		linear regression	17.0	0.0	Q3
Bukvic et al., 2022 (J) [22]	brands, price	used car		descriptive analysis, scraping		linear regression (acc. = 95%)	18.0	4.0	Q2
Caliskan et al, 2022(J) [23]	brands, logistic, parts	vehicle		descriptive analysis, modelling			16.0	4.0	Q1
Dutulescu et al, 2023a (J) [30]	brands, price	used car	brand, model, year, mileage, engine, gearbox, fuel, transmission, car, shape, color, add., images, price	descriptive analysis, modelling, scraping	80% train, 20% test	neural network (R^2 = 0.95%)	17.0	2.0	Q2
Dutulescu et al, 2023b (C) [31]	brands, price	used car	manufacturer, model, year, mileage, fuel, engine, capacity, transmission, add., price	descriptive analysis, modelling, scraping	80% train, 20% test	neural network (R^2 = 0.96%)	15.0	2.0	
Farhan & Heikal, 2023 (C) [35]	logistic	used car	generations, gender, brand, transmission, customer satisfaction, company, payment, branches	modelling	80% train, 20% test	clustering	15.0	0.0	Q4
Gegic et al, 2019 (J) [39]	price	used car	brand, model, fuel, year, color	descriptive analysis, feature sel., modelling, scraping		SVM + neural network (acc.= 87.30%)	18.0	21.0	Q3
Gollapalli et al, 2023 (J) [41]	price	used car	make, type, year, origin, color, options, engine size, fuel, gear mileage, local, price, negotiable	descriptive analysis, modelling	70% train, 30% test	random forest ($R^2 = 0.82$)	18.0	2.0	Q4

Note: * (J) - journal paper; (C) - conference paper; (Bc) - book chapter.

Table B.2: Summary of articles in systematic literature review (authors H - P)

Reference*	Topic	Car type	Features	Methods	Data Split	Best Result	Quality	Cits./ year	Journal Rank
Hankar et al, 2022 (C) [43]	price	used car	mileage, fuel type, production year, brand, model, fiscal power	correlation, modelling, scraping		XGBoost ($R^2 = 0.8$)	17.0	6.0	
Hemendiran& Renjith, 2023 (C) [45]	price	used car		modelling		random forest ($R^2 = 0.946$)	15.0	1.0	
Huang et al, 2023 (J) [48]	price	used car		modelling		bayesian neural network	15.0	0.0	Q1
Ishak & Wijaya, 2020 (C) [50]	logistic, parts	vehicle		modelling		decision tree (acc. = 90.85%)	16.0	1.0	
Jhala & Anand, 2023 (C) [51]	price	used car	new car, price, year, km driven, transmission, fuel	correlation, descriptive analysis, feature sel., modelling		random forest (acc. = 94.10%)	16.0	0.0	
Jin, 2021 (C) [52]	price	used car	mileage, year of manufact., fuel consumption, transmission, road tax, fuel, engine size	modelling		random forest (acc. = 90.41%)	15.0	4.0	
Krishnan & Selvaraj, 2022 (C) [54]		used car		descriptive analysis, modelling		XGBoost ($R^2 = 87.0\%$)	15.0	1.0	
Kumar et al., 2023 (C) [55]	price	used car	name, local, year, fuel, transm., owner, km driven, mileage, engine, power, seats, price	descriptive analysis, modelling		random forest (acc. = 86%)	16.0	1.0	
Li et al, 2022 (C) [57]	price	used car	name, sales time, transaction price, new car quotation, transaction address, mileage, licensing time, displacement, gearbox, emission standard	descriptive analysis, modelling, scraping		random forest ($R^2 = 0.936$)	15.0	4.0	
Liu et al, 2022 (J) [58]	price	used car		correlation		PSO-GRA-BPNN ($R^2 = 0.984$)	15.0	4.0	
Longani et al, 2021 (Bc) [59]	price	used car	year, seller type, km driven, owner, fuel, gearbox, engine cap., seats, price, etc.	modelling, scraping		XGBoost	15.0	4.0	
Lu et al, 2023 (C) [60]	price, parts	vehicle		modelling		Bi-GRU ($R^2 = 0.99$)		1.0	
Luo et al, 2023 (C) [61]	price	used car	name, local, cond., make, model, year, mileage, fuel, seats, transmission, engine, front & rear brakes, front & rear susp., front & rear tire price, etc.	correlation, scraping		XGBoost ($R^2 = 96\%$)	15.0	0.0	
Lyapuntsova and Boyko et al, 2023 (C) [62]	brands, parts	vehicle		descriptive analysis			15.0	1.0	
Lyu & Li, 2021 (C) [63]	env.	vehicle				SIR regression	15.0	0.0	
Ma, 2023 (C) [64]	brands, price, geo. dist.	used car		descriptive analysis			15.0	0.0	
Monburion et al, 2018(C) [67]	price	used car	gearbox, model, brand, fuel, vehicle type	correlation, descriptive analysis		gradient boosted regression tree (MAE = 0.28)	17.0	12.0	
Msiza & Owolawi, 2023 (C) [68]	price	used car		correlation, feature sel., modelling, scraping		random forest ($R^2 = 0.988$)	18.0	0.0	
Nandan & Ghosh, 2023 (J) [70]	price	used car	local, year, manufac., model, cond., cylinders, fuel, transmission, drive, size, color, lat, long, price, etc.	correlation, descriptive analysis, feature sel., modelling	90% train, 10% test	XGBoost (acc. = 86.87%)	17.0	0.0	
Narayana et al, 2021 (C) [71]	brands, price	used car	model, age, fuel, type of seller, transm., km driven	correlation, descriptive analysis		random forest (acc. = 85%)	17.0	7.0	
Narayana et al, 2022 (C) [72]	price	used car	name, year, km driven, fuel, seller type, transmission, owner, sell price, seats, steering type	correlation, descriptive analysis, modelling	80% train, 20% test	random forest (acc. = 90%)	16.0	2.0	

Note: * (J) - journal paper; (C) - conference paper; (Bc) - book chapter.

Table B.3: Summary of articles in systematic literature review (authors P - Z)

Reference*	Topic	Car type	Features	Methods	Data Split	Best Result	Quality	Cits./ year	Journal Rank
Pal et al, 2018 (C) [75]	brands, price	used car		descriptive analysis, modelling		random forest (acc. = 83.62%)	18.0	15.0	
Putro & Indrawati, 2022 (J) [77]	price	used car		modelling, scraping		KNN (acc. = 95.46%)	18.0	0.0	
Rane et al, 2023 (C) [78]	price	used car	present price, selling price, year, mileage, owner, fuel, seller type, transmission	descriptive analysis, modelling		random forest	15.0	0.0	
Rojniruttikul, 2019 (C) [80]	logistic	vehicle	gender, age, experience, management level	correlation, modelling		linear regression	15.0	0.0	
Sanguansaringjhan & Liu, 2020 (C) [81]	price, parts	vehicle		modelling		exponential smoothing	16.0	0.0	
Satapathy et al, 2022 (C) [83]	price	used car		correlation, descriptive analysis, modelling		XGBoost	17.0	3.0	
Shah et al, 2022 (C) [84]	price	used car		correlation, descriptive analysis, modelling		random forest + MICE (acc. = 94.07%)	17.0	0.0	
Shanti et al., 2021 (C) [85]	price	used car		modelling, scraping		gradient boosting (acc. = 89.72%)	18.0	2.0	
Shaprapawad et al, 2023 (C) [86]	price	used car	name, location, year, kilometers driven, fuel type, transmission, owner, mileage, engine, power, seats, new price, price	modelling	90% train, 10% val.	SVM (R^2 = 95.27%)	17.0	3.0	
Sharma & Mitra, 2024 (J) [87]	price	used car		descriptive analysis, modelling		MARS (acc. = 78.33%)	17.0	3.0	Q1
Skribans & Hegerty, 2019 (C) [88]	price	used car		modelling		econometric model (89%)	15.0	0.0	
Uysal, 2023 (C) [92]	brands, price	used car		descriptive analysis, modelling, scraping		extra tree ($R^2 = 0.956$)	18.0	0.0	
Varshitha et al., 2022 (C) [93]	price	used car	name, year, selling price, present price, mileage, fuel, seller type, transmission, owner	correlation, modelling		random forest ($R^2 = 0.7726$)	17.0	10.0	
Vesovic et al, 2022 (J) [94]	env., parts	vehicle		modelling		clustering (acc. = 95%)	17.0	1.0	Q2
Viswanatha et al, 2023 (C) [95]	price	used car	name, year, purchase price, current price, mileage, fuel, seller type, transmission, owner, etc.	correlation, descriptive analysis, modelling	-	linear regression ($R^2 = 0.829$)	15.0	0.0	
Wang et al, 2021 (C) [99]	price	used car	name, location, year, kilometers driven, fuel, transmission, owner, mileage, power, seats, new price, price	modelling	80% train, 20% test	extra tree (acc. = 98.07%)	16.0	2.0	
Xuping et al, 2023 (C) [102]	price	used car		modelling		SVM	16.0	0.0	
Yilmaz & Selvi, 2023 (J) [103]	price, parts	used car	price, brand, model, fuel, transm., engine power, engine capacity, car damage status, right rear fender, rear hood, left rear fender, right rear door, right front door, roof, left rear door, left front door, right front fender, engine hood, left front fender, front bumper, rear bumper	feature sel., modelling, scraping	80% train, 20% test	XGBoost ($R^2 = 0.973$)	17.0	1.0	
Zhang, 2022 (C) [105]	price	used car	sale ID, name, reg date, model, brand, body type, fuel type, transmission, power, kilometer driven, not repaired damage, region code, offer type, create date, price, seller	descriptive analysis, feature sel.,		lightGBM	15.0	1.0	

Note: (J) - journal paper; (C) - conference paper; (Bc) - book chapter.

APPENDIX C.

Features description

This appendix presents the features description for the manufacturers, suppliers, and bodywork parts datasets. Each feature includes:

- **Feature**: Name of the feature in the dataset.
- **Description**: Description of the feature.
- **Data type**: Type of data used in analysis. Categorical features are ”string”for words or phrases, while numeric features are ”int64”for integers or ”float64”for floating-point.
- **Scale**: The feature’s scale. Categorical features are ”nominal”(no specific order) or ”ordinal”(specific order), and numeric features are ”ratio”(with absolute zero) or ”interval”(without absolute zero).

Table C.1: Description of features in the manufacturers’ dataset.

Feature	Description	Data type	Scale
Brand	brand of the classic car	string	nominal
Model	model of the classic car	string	nominal
Submodel	submodel of the classic car	string	nominal
Start Year	start year of the manufacture of the classic car	int64	interval
End Year	end year of the manufacture of the classic car	int64	interval
Origin	origin of the manufacture of the classic car	string	nominal

Table C.2: Description of features in the suppliers’ dataset.

Feature	Description	Data type	Scale
Category	supplier focus: brands, categories, or market segments	string	nominal
Version	version of the modified classic car	int64	ordinal
Supplier	name of the supplier	string	nominal
Supplier_type	supplier type (spare parts / general services)	string	nominal
Country	supplier country location	string	nominal
City	supplier city location	string	nominal
Address	address of the supplier	string	nominal
IATA	three-letter code of the nearest airport to the supplier	string	nominal
City_airport	city of the nearest airport	string	nominal
Climate_impact_kg_co2	total CO_2 emissions and climate effects (CO_2/kg)	int64	ratio
Flight_distance_km	flight distance (km)	int64	ratio
Co2_emissions_kg	CO_2 emissions related to fuel consumption (CO_2/kg)	int64	ratio
Climate_impact_of_contrails_ozone_formation _etc_kg	total non-CO_2 emissions and climate effects (kg)	int64	ratio
Fuel_consumption_l	fuel consumed (L)	int64	ratio

Table C.3: Description of features in the bodywork spare parts' dataset.

Feature	**Description**	**Data type**	**Scale**
Name	name of the spare part	string	nominal
Product type	type of product	string	nominal
Composition	type of composition of spare part material	string	nominal
Finish	type of finish applied to the spare part	string	nominal
Weight	weight of the spare part (kg)	float64	ratio
Length	length of the spare part (mm)	float64	ratio
Width	width of the spare part (mm)	float64	ratio
Height	height of the spare part (mm)	float64	ratio
Price	price of the spare part (€)	float64	ratio
Models	list of compatible classic car models	string	nominal

APPENDIX D.

Descriptive statistics

In this appendix, you will find descriptive statistics for each feature from manufacturers, suppliers, and bodywork parts datasets. This includes the count of observations (count), the count of unique values (unique), the most frequent value (top), and the count of the most frequent value (freq) for categorical features, as well as the minimum (min) and maximum (max) values for numeric features.

Table D.1: Descriptive statistics for features in the manufacturers' dataset.

	Brand	**Model**	**Submodel**	**Origin**	**Start_year**	**End_year**
count	2267	2266	2000	2258	2267	2137
unique	85	479	1338	12	N/A	N/A
top	Citroen	CX	GT	Italy	N/A	N/A
freq	336	97	23	781	N/A	N/A
min	N/A	N/A	N/A	N/A	1928	1931
max	N/A	N/A	N/A	N/A	2016	2018

Table D.2: Descriptive statistics for features in the suppliers' dataset.

	Cat*	**Ver***	**Sup***	**Sup typ***	**Country**	**City**	**Address**	**IATA**	**Cty apt***
count	2915	88	2915	2915	2652	2652	2915	2915	2915
unique	89	N/A	2840	2	53	563	2803	338	356
top	classic car dealer	N/A	joop stolze classic cars	service	germany	zwickau	lierweg (...)	HOF	Hof
freq	656	N/A	9	2284	533	525	9	525	527
min	N/A	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A
max	N/A	9	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: * Cat - Category (brands, categories, or market segments); Ver - Version; Sup - Supplier; Sup typ - Supplier_type; Cty apt - City_airport.

Table D.3: Descriptive statistics for features in the suppliers' dataset (Cont.)

	climate_impact _kg_co2	**flight_distance _km**	**co2_emissions _kg**	**climate_impact _ozone_etc_kg**	**Fuel_consumption _l**
count	356	7	2267	442	2828
min	149.000	520.000	73.000	73.000	20.000
max	919	929	999	987	985

Table D.4: Descriptive statistics for features in the bodywork parts' dataset.

	Name	**Weight (kg)**	**Length (mm)**	**Width (mm)**	**Height (mm)**	**Price (€)**	**Mods***	**Prod. typ***	**Comp***	**Finish**
count	804	804	804	804	804	804	804	804	804	804
unique	804	N/A	N/A	N/A	N/A	N/A	222	77	16	11
top	drain plug (...)	N/A	N/A	N/A	N/A	N/A	Chev Monte Carlo '70-'72	quarter panels	steel (hot rolled)	EDP coat
freq	1	N/A	N/A	N/A	N/A	N/A	42	108	630	494
min	N/A	0.01	0.01	0.01	0.01	0.84	N/A	N/A	N/A	N/A
max	N/A	63.50	2540	1778	800	935	N/A	N/A	N/A	N/A

Note: * Mod. - Models; Prod. typ. - Product_type; Comp. - Composition.

APPENDIX E.

Descriptive analysis

This appendix presents the descriptive analysis of the top 3 most popular models for each most popular classic car brands (Ford, BMW, Mercedes-Benz, Porsche, Volkswagen, Chevrolet, Honda, Audi, and Jeep). The popularity of brands and models was retrieved by Google search hits analysis.

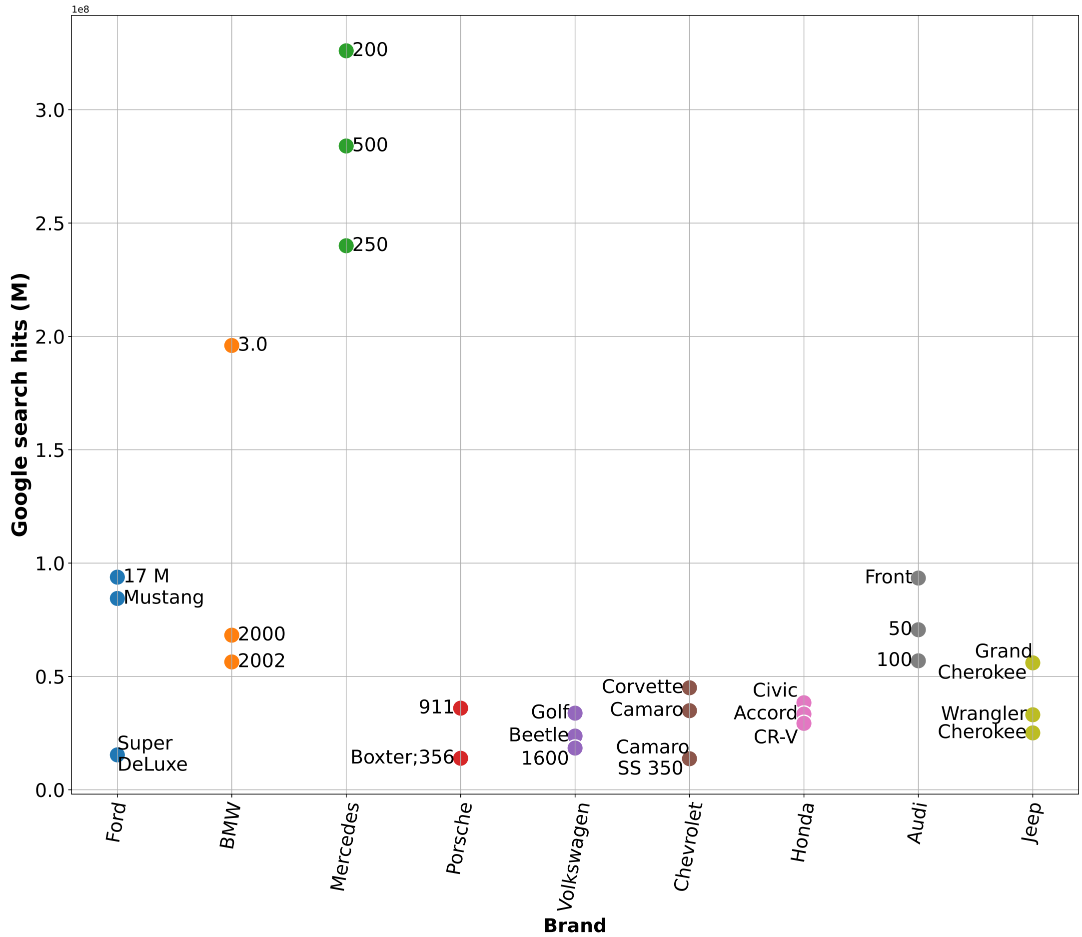

Figure E.1: Top 3 most popular classic car models for most popular classic car brands.

[This page has been intentionally left blank]

APPENDIX

F.

K-Means clustering metrics and evaluation

This appendix presents the results of the K-Means clustering metrics obtained for the three models. The elbow curve, silhouette score, and data point distribution were used to select the number of clusters. After choosing the number of clusters, the ANOVA test evaluates the final distribution for each feature in each clustering model.

F.1 Elbow curve

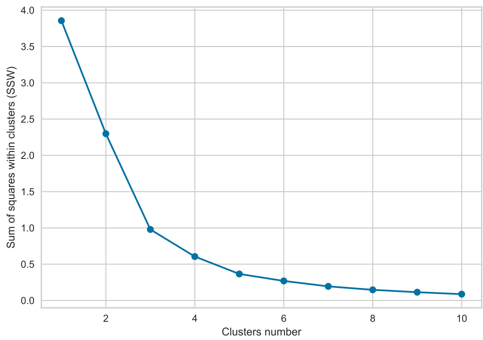

Figure F.1: Number of suppliers Vs. carbon footprint mean (CO_2/kg)

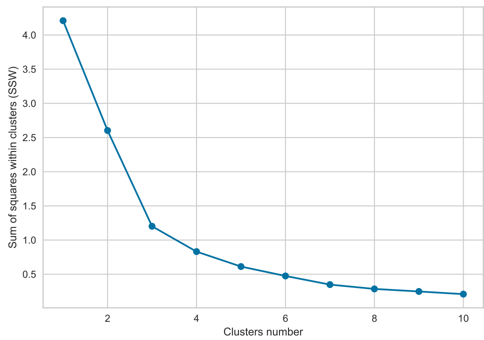

Figure F.2: Carbon footprint mean (CO_2/kg) Vs. Google Hits for classic car brands

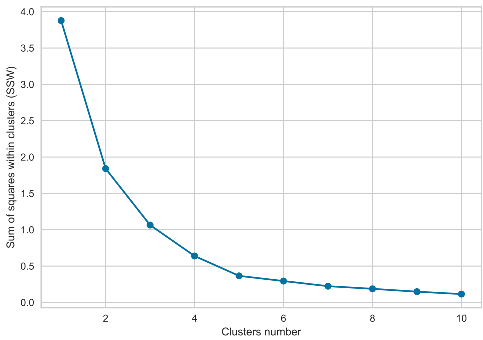

Figure F.3: Number of suppliers Vs. Google Hits for classic car brands

F.2 Silhouette score and data points distribution

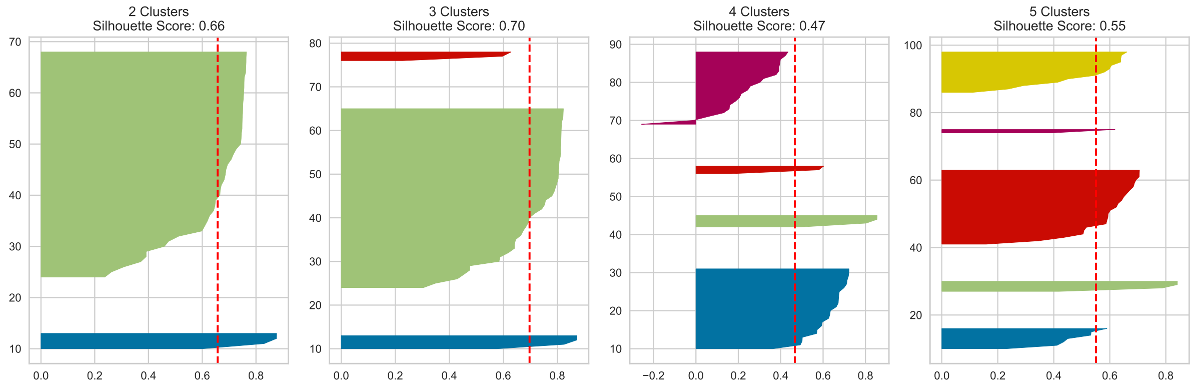

Figure F.4: Data distribution (number of suppliers Vs. carbon footprint mean (CO_2/kg) for each cluster group (2, 3, 4 and 5 respectively)

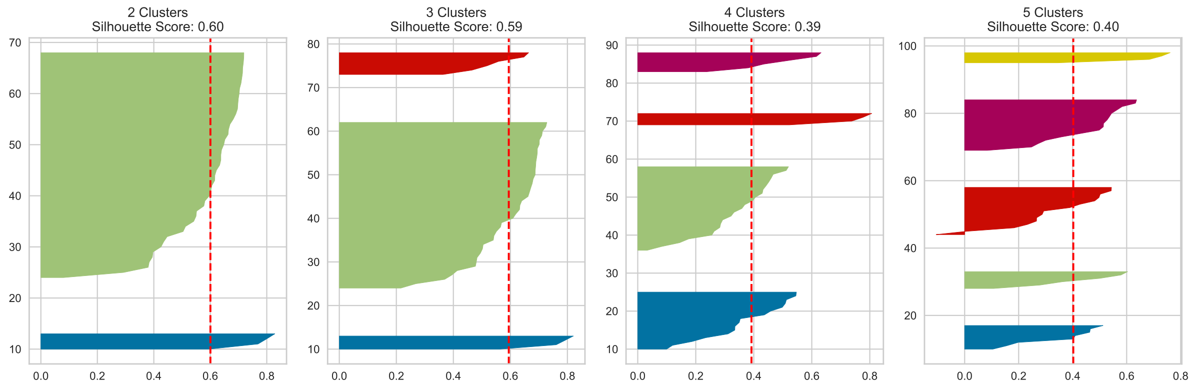

Figure F.5: Data distribution (carbon footprint mean (CO_2/kg) Vs. Google Hits for classic car brands) for each cluster group (2, 3, 4 and 5 respectively)

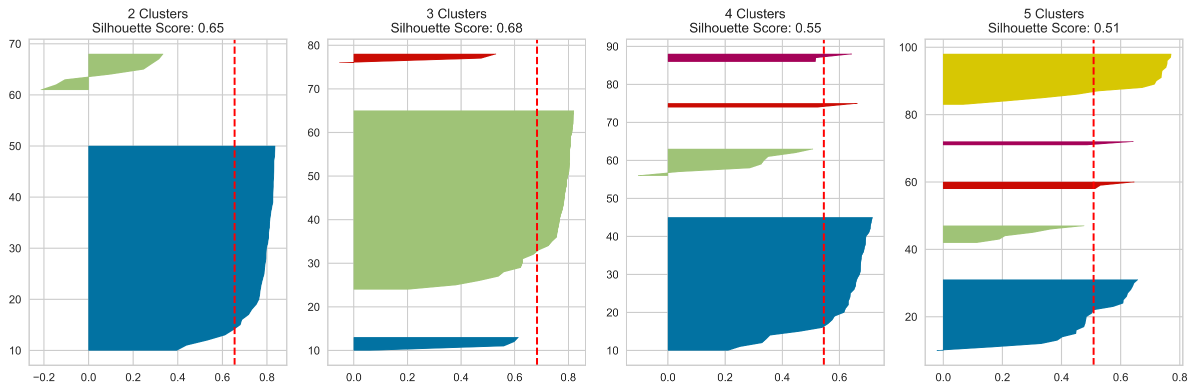

Figure F.6: Data distribution (number of suppliers Vs. Google Hits for classic car brands) for each cluster group (2, 3, 4 and 5 respectively)

F.3 ANOVA test for clustering results

Table F.1: Transportation impact Vs. Suppliers' availability

Cls.*	Feature	Sum of Squares			Mean of Squares	
		Total variance (SST)	Between groups (SSB)	Within groups (SSW)	Between groups (MSB)	Within groups (MSW)
C1	CF mean	0.003	0.022	< 0.001	0.011	< 0.001
C1	No. of suppliers	0.120	0.027	0.111	0.013	0.002
C2	CF mean	0.539	0.304	0.437	0.152	< 0.001
C2	No. of suppliers	0.282	0.371	0.158	0.185	0.003
C3	CF mean	0.036	0.029	0.026	0.015	< 0.001
C3	No. of suppliers	< 0.001	0.035	< 0.001	0.018	< 0.001

Note: * Cls. = Number of cluster group.

Table F.2: Transportation impact Vs. Brands' popularity

Cls.*	Feature	Sum of Squares			Mean of Squares	
		Total variance (SST)	Between groups (SSB)	Within groups (SSW)	Between groups (MSB)	Within groups (MSW)
C1	CF mean	0.011	0.043	< 0.001	0.022	< 0.001
C1	Google hits	0.173	0.029	0.163	0.015	0.004
C2	CF mean	0.530	0.282	0.436	0.141	0.010
C2	Google hits	0.440	0.189	0.377	0.095	0.008
C3	CF mean	0.036	0.029	0.026	0.015	< 0.001
C3	Google hits	0.012	0.019	0.006	0.010	< 0.001

Note: * Cls. = Number of cluster group.

Table F.3: Suppliers' availability Vs. Brands' popularity

Cls.*	Feature	Sum of Squares			Mean of Squares	
		Total variance (SST)	Between groups (SSB)	Within groups (SSW)	Between groups (MSB)	Within groups (MSW)
C1	No. of suppliers	0.120	0.026	0.111	0.013	0.002
C1	Google hits	0.078	0.020	0.072	0.010	0.002
C2	No. of suppliers	0.033	0.035	0.022	0.017	< 0.001
C2	Google hits	0.116	0.027	0.107	0.014	0.002
C3	No. of suppliers	0.240	0.364	0.119	0.182	0.003
C3	Google hits	0.478	0.284	0.377	0.095	0.008

Note: * Cls. = Number of cluster group.

Appendix

G.

Results of K-Means clustering for classic car brands with normalized features

In this appendix, you will find the normalized values using the min-max (mm) algorithm of each feature (number of suppliers, carbon footprint mean, Google search hits), along with the corresponding cluster group (1, 2, or 3) for each classic car brand that was evaluated. Three tables are provided, one for each clustering model.

Table G.1: Number of suppliers Vs. carbon footprint mean by cluster

Brand	**mm number of suppliers**	**mm carbon footprint mean**	**Cluster**
Volkswagen	0.732	0.291	1
Porsche	1.000	0.231	1
Citroen	0.511	0.225	1
MG	0.324	0.380	2
Ferrari	0.207	0.288	2
Morgan	0.066	0.208	2
Mazda	0.038	0.328	2
Renault	0.028	0.219	2
Rover	0.028	0.471	2
Peugeot	0.033	0.205	2
Triumph	0.136	0.352	2
Land Rover	0.221	0.260	2
Fiat	0.070	0.270	2
Austin	0.038	0.212	2
Cadillac	0.014	0.223	2
Aston Martin	0.005	0.102	2
Alfa Romeo	0.291	0.242	2
Chrysler	0.000	0.460	2
Riley	0.005	0.001	2
Buick	0.000	0.460	2
Pontiac	0.014	0.232	2
Opel	0.080	0.189	2
Maserati	0.056	0.307	2
Bentley	0.047	0.332	2
Lotus	0.075	0.372	2
Range Rover	0.000	0.000	2
Škoda	0.005	0.172	2
Saab	0.052	0.190	2
Bugatti	0.009	0.075	2
Morris	0.047	0.425	2
Datsun	0.028	0.252	2
Abarth	0.014	0.378	2
Lancia	0.061	0.255	2
De Tomaso	0.014	0.243	2
DKW	0.014	0.172	2
NSU	0.023	0.172	2
Simca	0.000	0.226	2
Panhard	0.005	0.215	2
Alvis	0.019	0.500	2
Allard	0.000	0.366	2
Facel Vega	0.014	0.202	2
Ford	0.066	0.297	2
BMW	0.235	0.239	2
Toyota	0.014	0.172	2
Chevrolet	0.197	0.273	2
Subaru	0.000	0.932	3
Holden	0.005	0.932	3
Daimler	0.000	0.745	3
Reliant	0.000	1.000	3

Table G.2: Google Hits Vs. carbon footprint mean by cluster

Brand	mm Google search hits	mm carbon footprint mean	Cluster
Ford	1.000	0.297	1
BMW	0.813	0.239	1
Toyota	0.755	0.172	1
Volkswagen	0.581	0.291	1
Porsche	0.554	0.231	1
Chevrolet	0.523	0.273	1
MG	0.379	0.380	2
Ferrari	0.361	0.288	2
Morgan	0.351	0.208	2
Mazda	0.324	0.328	2
Renault	0.321	0.219	2
Rover	0.303	0.471	2
Peugeot	0.293	0.205	2
Triumph	0.289	0.352	2
Land Rover	0.270	0.260	2
Fiat	0.265	0.270	2
Austin	0.263	0.212	2
Cadillac	0.245	0.223	2
Aston Martin	0.239	0.102	2
Alfa Romeo	0.232	0.242	2
Citroën	0.225	0.225	2
Chrysler	0.220	0.460	2
Riley	0.198	0.001	2
Buick	0.196	0.460	2
Pontiac	0.196	0.232	2
Opel	0.196	0.189	2
Maserati	0.194	0.307	2
Bentley	0.193	0.332	2
Lotus	0.187	0.372	2
Range Rover	0.184	0.000	2
Škoda	0.165	0.172	2
Saab	0.149	0.190	2
Bugatti	0.145	0.075	2
Morris	0.132	0.425	2
Datsun	0.125	0.252	2
Abarth	0.107	0.378	2
Lancia	0.102	0.255	2
De Tomaso	0.041	0.243	2
DKW	0.037	0.172	2
NSU	0.030	0.172	2
Simca	0.029	0.226	2
Panhard	0.019	0.215	2
Alvis	0.013	0.500	2
Allard	0.005	0.366	2
Facel Vega	0.000	0.202	2
Subaru	0.191	0.932	3
Holden	0.080	0.932	3
Daimler	0.060	0.745	3
Reliant	0.059	1.000	3

Table G.3: Number of suppliers Vs. Google search hits by cluster

Brand	**mm number of suppliers**	**mm Google search hits**	**Cluster**
Citroën	0.512	0.225	1
Volkswagen	0.732	0.581	1
Porsche	1.000	0.554	1
MG	0.324	0.379	2
Ferrari	0.207	0.361	2
Morgan	0.066	0.351	2
Mazda	0.038	0.324	2
Renault	0.028	0.321	2
Rover	0.028	0.303	2
Peugeot	0.033	0.293	2
Triumph	0.136	0.289	2
Land Rover	0.221	0.270	2
Fiat	0.070	0.265	2
Austin	0.038	0.263	2
Cadillac	0.014	0.245	2
Aston Martin	0.005	0.239	2
Alfa Romeo	0.291	0.232	2
Subaru	0.000	0.191	2
Lotus	0.075	0.187	2
Range Rover	0.000	0.184	2
Škoda	0.005	0.165	2
Saab	0.052	0.149	2
Bugatti	0.009	0.145	2
Morris	0.047	0.132	2
Datsun	0.028	0.125	2
Abarth	0.014	0.107	2
Lancia	0.061	0.102	2
Holden	0.005	0.080	2
Daimler	0.000	0.060	2
Reliant	0.000	0.059	2
De Tomaso	0.014	0.041	2
DKW	0.014	0.037	2
NSU	0.023	0.030	2
Simca	0.000	0.029	2
Panhard	0.005	0.019	2
Alvis	0.019	0.013	2
Allard	0.000	0.005	2
Facel Vega	0.014	0.000	2
Ford	0.066	1.000	3
BMW	0.235	0.813	3
Toyota	0.014	0.755	3
Chevrolet	0.197	0.523	3

APPENDIX H.

Categorization of transportation impact, suppliers' availability, and popularity

The features are discretized in this appendix into three categorical levels (low, medium, and high). Each discretized feature corresponds to a specific factor evaluated for the RQ3 (Table H.1):

- **Carbon footprint mean (CO_2/kg)**: transportation impact.
- **Number of suppliers**: suppliers' availability.
- **Google search hits (M)** : brand popularity.

For the best-performing classic car brands in these factors (BMW, Citroen, Chevrolet, Ford, Porsche, Toyota, and Volkswagen), original values and the leading supplier country are presented in Table H.2. Finally, these brands have a detailed discussion complemented by the literature to understand their supplier market availability and brand popularity.

Table H.1: Categorical levels of factors

Categorical level	Carbon footprint mean	Number of suppliers	Google hits (M)
Low	0 - 750	0 - 100	0 - 100
Medium	751 - 900	101 - 200	101 - 120
High	> 901	> 201	> 121

Table H.2: Factor values and the leading supplier country for the best-performing brands

Brand	Leading supplier country	Carbon footprint mean (CO_2/kg)	Number of suppliers	Google hits (M)
BMW	Germany	713.31	51	183
Chevrolet	United States	790.74	43	118
Citroen	The Netherlands	680.15	110	51.3
Ford	Germany	843.93	15	225
Porsche	Germany	693.77	214	125
Toyota	Germany	560.50	4	170
Volkswagen	Germany	831.71	157	131

BMW

BMW ranks in the top 10 classic car brands regarding service or parts supplier availability and is the second most popular classic car brand in our datasets.

BMW comprises a significant diversity of models and a high manufacturing level, with most of the production concentrated in Europe (Germany, Austria, and the UK) [1]. Our findings confirm this large-scale production in Europe and indicate that BMW is an eco-friendly brand in terms of transportation between international suppliers and Lisbon. This suggests the existence of a broad network of available suppliers for BMW classic cars near Europe.

In addition to the widespread production network, BMW is also one of the most preferable brands in the used car market [22, 30, 31], which has an impact on prices [13, 30, 31, 75] and has higher mileage compared to other selling brands [13, 22]. This popularity can be explained by two fundamental reasons, according to Woodward and Guimarães: its extensive supplier network and a relatively large direct payroll, primarily spent at local businesses [101].

Our analysis also observed that one BMW model, the BMW 3.0, is the fourth most popular classic car model, probably due to being considered a luxury car competitor and rarity [12].

Chevrolet

The Chevrolet brand is produced by General Motors (GM) in the United States of America. GM has a 107-year history in manufacturing and selling motors and parts [32].

Chevrolet is the classic car brand with the most parts sold on the restoparts website. It is observed that the Chevrolet classic car brand has a high production mass, with 59082 cars sold on the classic.com website, compared to only 24231 BMW classic cars [19]. This could explain why Chevrolet is the brand with most parts on the restoparts website and potentially the reason for its position in medium suppliers due to the significant global market demand.

Citroen

Citroen as a significant human interest and that it is a well-established European brand.

Our analysis found that Citroen is the most frequent brand manufacturer, meaning it has a significant diversity of models. Citroen has a century-long heritage, with the Citroen conservatory Aunlnay-sous-bois, a museum that contains more than 400 historic Citroen models. This may also explain why it is the fourth most supplied brand due to its long history of car variety and parts [25].

Ford

The mass-production of classic brand Ford in the 1920s [100] could explain the enduring popularity of the classic car brand, leading it to the top position in our dataset.

In the 1930s, Ford established manufacturing systems in England, Germany, and France [100], which may explain the significant production in Germany in our dataset.

Porsche

The brand was established in Zuffenhausen, Germany [76], which could explain its environmentally friendly image in our analysis, as it makes transportation easier in Europe.

As the brand's popularity can be influenced by its previous owners, such as those Porsche used by Steve McQueen [46], this can explain its ranking sixth in our Google hits analysis. This particular model is also the most popular for the Porsche brand.

Toyota

Toyota still manufactures over 50 percent of its cars in its home base, Japan [1], which may explain its substantial presence in our results, where it ranked third in Google hits.

Although Toyota is recognized as one of the top suppliers in the industry [23], our analysis suggests that it may not be the best supplier in Europe, where its presence is not as strong as in America and Asia.

Volkswagen

Factors such as the transatlantic tie between Germany and America, and the Beetle cultural influence [79], can explain why Volkswagen's classic car is our study's second most supplied brand. It is also known for being well-supplier and eco-friendly, contributing to more accessible transportation in Europe.

Volkswagen is one of the most automotive entities in terms of units sold, suppliers, and distributors [18, 65]. This sustained appeal of the classic car brand is discernible from our research findings, positioning it as the eighth most popular classic brand in our dataset.

[This page has been intentionally left blank]

iscte
UNIVERSITY
INSTITUTE
OF LISBON

Classic car spare parts for restoration: International supply chain and environmental impact

Inês Celorico