UNIVERSIDADE DE COIMBRA

FACULDADE DE CIÊNCIAS E TECNOLOGIAS

ADVANCED DATA VISUALIZATION

Project Proposal

Mariana Paulino - 2020190448

Rui Alexandre Tapadinhas - 2018283200

Coimbra - Portugal

2023/2024

Index

1   Introduction

Nowadays, cars are used for every minor or major journey that a person has. Whether it is to go to and from work, to the supermarket, basically living your daily life, or, to go on holiday or business trips. Our goal for this project is to build a dashboard where the user can get a more in-depth look into cars and their manufacturers. To achieve our final goal, we have to develop a few visualizations that can provide information and make it available for everyone to understand it.

2   Problem and Solutions

The problem we came across is when searching for information on a car or manufacturer, the information is not centralized and can be a little confusing for a normal person who does not know a lot about cars. Our goal to mitigate this issue is to centralize all the information we can, gathering all the details and then making it easy to comprehend through some visualizations and graphics using Plotly and Dash. Recurring to the visualizations created, we aim to create a Car Comparison Tool, allowing the user to pick two specific cars and compare them in specific fields that can be selected by them in our Dashboard.

3   Related Work

As an inspiration for our project, we based our goals on some existing work as Ze Perfs [5] that allows people to compare cars. And, Ultimate Specs [4] that has lots of data and information but lacks on the visual information and for users that don't understand that much it can be hard to digest so much text.

4   Data

4.1   Description

The dataset "Car Specification Dataset" [1] contains nearly 30,000 model variations from 124 car brands. Each row has information about the engine, power, dimensions, segment, fuel, fuel efficiency and other details more specific to each model. There are some attributes specific to only some types of engine, so there are some null values, making the dataset need pre-processing.

4.2   Attributes

- Model
- Serie
- Manufacturer
- Body style
- Segment
- Production years
- Cylinders

- Displacement
- Power
- Torque
- Electrical motor power
- Electrical motor torque
- Fuel Info

- Top Speed
- Acceleration 0-62 Mph (0-100 kph)
- Drive Type
- Gearbox
- Brakes
- Tire Size

- Dimension
- Front/rear Track
- Wheelbase
- Ground Clearance
- Aerodynamics
- Turning circle

- Cargo Volume
- Weight
- Consumption

- Emissions
- Power pack
- Top speed (elec-

trical)
- EV Range
- Battery Maxi-

mum Capacity
- Specification summary

5   Design Requirements

5.1   Plots

For this project the initial idea is to have little plots that complement each other and alone can give details but together gather more conclusions about the car or manufacturer chosen. The idea is to have a multi page Dashboard for our graphics to be more organized and for the user to get a better look into whether it is the brand or the specific car or characteristics. Some of the plots we are considering to include in our dashboard are enumerated in the list below:

- **Average Maximum Speed by Brand by Year**

  Line plot with the yearly evolution of the average maximum speed by brands. This visualization would help the user to understand and compare the evolution of each brand.

- **Average Power by Brand by Year**

  Line plot with the yearly evolution of the average power output by brands. As in the previous visualization, this would help the user to understand and compare the evolution of each brand.

- **Comparison of Specific Cars**

  In this option the user is going to receive all the attributes a car presents in our dataset and is going to be able to compare two or more vehicles. Here, the user is going to select the car he wants to compare first and then the second one. This choice is going to be in order of attributes such as Manufacturer, Model, Engine Specifications, Year, and Traction Type if needed so that the user only selects one car at a time and compare the desired ones.

- **Compare Average Combined Consumption by Brand**

  This visualization has the objective to help the user pick a brand based on its consumption, by displaying the average combined consumption by brand, after choosing the fuel type.

- **Manufacturer Count per Country**

  For this plot, we aim to get the country where the manufacturer is based and then plot a world map where the user can see the density of brands per country. This is the less informative graph of our Dashboard but a good one to know where the cars are from. The user can select two or more brands and be able to see where they are from.

- **Power vs Displacement per Fuel Type**

  By plotting these two variables graphically after selecting the fuel type, we aim to compare the power per litre (displacement), which is an interesting measure for the potential and technology of the engine.

- **Top Speed vs Power**

  We would implement a scatter plot where each point represents a model specification, facilitating comparison between the Top Speed and Power of each model.

- **Combined Consumption x Segment or Model**

    In this plot the aim is to be able to visualize the combined miles per gallon consumption translating it into litres per 100 kilometres, with the segment that a car belongs to or the model since there are different models and segments inside one manufacturer.

    Segment Comparison: This is the default view of this graphic. Here it is only going to be presented the average combined consumption per segment.

    Model Comparison: The user chooses a segment and in the plot various models appear (from the segment selected) from different brands.

5.2   Color Palette

To develop the dashboard, we have chosen a colour palette that has distinct colours and some colours to complement various details on the graphs. These colours are vivid enough to be easily separated and not be confused with others.
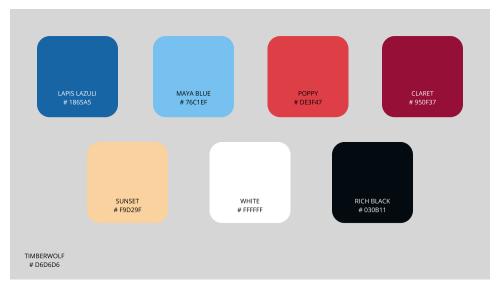

Figure 1 - Chosen Palette for this Project

In order to pick these colours we went to the Coolors [2] website, that is a generator of palettes, and then we personalized it to be this vivid.

5.3   Fonts

The chosen fonts for this project were Palatino and Aspira, since we saw a good combination of the two fonts, making texts and headers easy to see for every user. Figure 2 is a representation of both fonts in light colour background. Figure 3 is the same representation but in dark colour backgrounds. With these two figures, we could easily understand if they went well together and their readability, which for this initial stage is good for our project. This set of fonts was generated with the help of Monotype [3] website.

Palatino

This is a sample title. The quick brown fox jumps over the lazy dog.

Aspira

This is a sample paragraph and simple text. The quick brown fox jumps over the lazy dog.

Palatino

This is a sample title. The quick brown fox jumps over the lazy dog.

Aspira

This is a sample paragraph and simple text. The quick brown fox jumps over the lazy dog.

Figures 2 and 3 - Chosen Fonts for this Project

6   Conclusion

For this project to be successful, we hope that we can prepare the dataset correctly and then create the best visualizations to help the user understand more about cars.

The final goal is for people who don't understand a lot of the characteristics of cars and their correlations to be able to comprehend them and help them make decisions. And, for the ones who already know a lot about cars, to help them explore and find trends, by displaying the information in an easier way to digest and analyse it.

References

[1] *Car Specification Dataset.* https://www.kaggle.com/datasets/usefashrfi/car-specification-dataset.

[2] *Coolors.* https://coolors.co//1865a5-76c1ef-f9d29f-de3f47-950f37-ffffff-d6d6d6-030b11.

[3] *Monotype.* https://https://www.monotype.com/font-pairing#/playground?fontPair1=Palatino&fontPair2=Aspira.

[4] *Ultimate Specs.* https://www.ultimatespecs.com/.

[5] *Ze Perfs.* https://zeperfs.com/pt/.