UNIVERSITY OF AMSTERDAM

EXERCISE 1: CARS DATASET

Visualizing the Evolution of Car Design

November 8, 2024

Lecturer:
dr. Rob Belleman

Student:
Kattelijn Bouma
12862444

Course:
Scientific Visualization and Virtual Reality

Course code:
5284SVVR6Y

Abstract

This report presents an interactive 3D visualization of the evolution of car design from 1970 to 1982. The visualization enables viewers the opportunity to explore shifts in design priorities during a transformative period in the global automotive industry.

1 Introduction

The automotive industry has seen significant changes over the years, influenced by technological advances, fuel efficiency requirements, and consumer preferences [Lall, 1980]. The 1970s brought a particularly pivotal change. Following the oil crisis of 1973, concerns over fuel efficiency and emissions grew, causing the United States to establish stricter environmental standards. Japanese and European cars, which were known for their fuel efficiency, grew in popularity. This report presents a visualization of a dataset containing specifications of 392 cars from the 1970s to the early 1980s. Using principles from Bertin [Bertin, 1983] and Mackinlay [Mackinlay, 1986], we created a visualization designed to explore the evolution of car design.

2 Methods

The dataset `cars.csv` consists of 392 rows and 7 columns, with each row representing a car and each column providing specific vehicle attributes. Table 1 summarizes the dataset columns.

Column Name	Data Type	Unique Values	Example Value
model	nominal	301	chevrolet chevelle malibu
MPG	quantitative (ratio)	127	18
cylinders	nominal	5	4
horsepower	quantitative (ratio)	93	130
weight	quantitative (ratio)	346	3504
year	quantitative (interval)	13	70
origin	nominal	3	US

Table 1: Summary of dataset columns, classifying each column as nominal, ordinal or quantitative data, along with the number of unique values in each column and an example value.

Our aim is to create a visualization that demonstrates the evolution of car design during the post-oil-crisis period, with a particular focus on examining the impact of regulations on fuel efficiency. To achieve this, fuel efficiency (MPG), horsepower and weight are given the highest priority. As these three metrics are of the quantitative type, they are best encoded in a visualization using position [Mackinlay, 1986]. We position MPG along the x-axis, horsepower along the y-axis, and weight along the z-axis.

Next in priority is the year of manufacture, which we encode by adding a temporal dimension [Mackinlay, 1988] to our visualization, using an animation frame for each year.

According to Mackinlay's importance ordering, color most accurately represents nominal data, second only to position. The origin of each car is therefore displayed through color. The number of cylinders is encoded through the size of the points.

Lastly, the model name appears as hover text, making it accessible without overwhelming the visualization.

2.1 Implementation

The implementation of the visualization is done using the Python Plotly package, which allows for the creation of interactive 3D plots. We used the `plotly.express.scatter_3d` function to create a 3D scatter plot, mapping MPG, horsepower, and weight onto the x, y, and z axes, respectively. The evolution is captured by using the `animation_frame` parameter, which animates the data points across different years. The `color` parameter encodes the origin of each car and `size` is used to represent the number of cylinders. To enhance the visualization, a static background layer is added, consisting of all data points with reduced opacity, allowing the animated points to stand out while still preserving the overall distribution of the dataset. The final interactive plot is then saved as an HTML file.

3 Results

Figure 1 shows a snapshot of our visualization. The visualization reveals several clear trends.

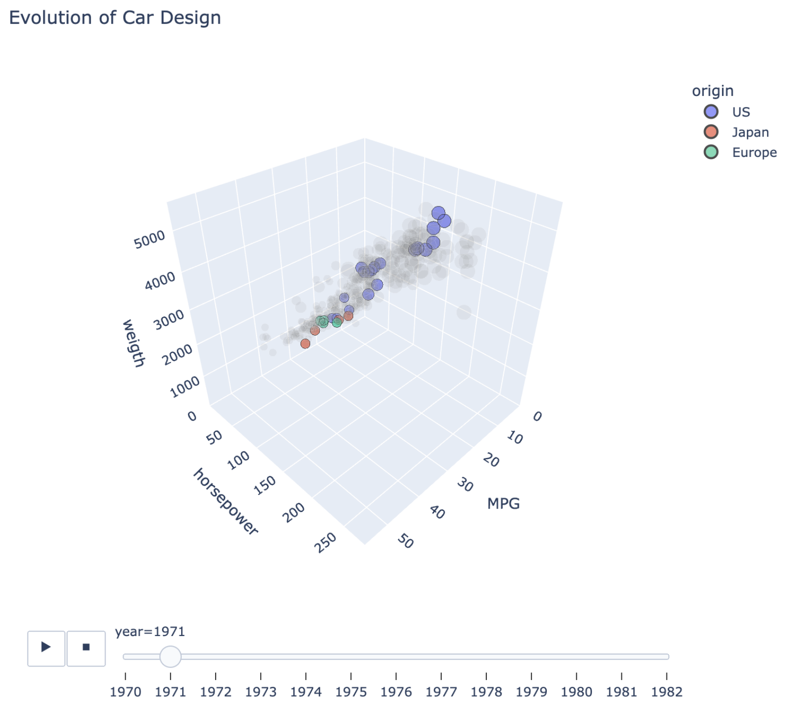

Figure 1: Snapshot of the visualization.

Over time, there seems to be a shift from heavier American models with higher horsepower and more cylinders toward lighter, more fuel-efficient cars with better miles per gallon.

4 Discussion

Our visualization successfully encodes the dataset's seven dimensions, achieving the goal of analysis within a single plot. By utilizing a 3D scatter plot, the visualization communicates the relationships between fuel efficiency (MPG), horsepower, weight, and other relevant car attributes. The use of animation allows for the temporal evolution of these design characteristics to be observed. However, some limitations and opportunities for improvement remain.

Firstly, the readability of the 3D scatter plot can be challenging, as the overlapping data points and perspective distortions may make it difficult to identify patterns. However, we attempted to address this challenge by making the visualization interactive, allowing users to rotate, zoom, and explore different perspectives.

Additionally, one could argue that the visual encoding of two of the dataset's columns might be improved. The use of point size to represent the number of cylinders may not be the most appropriate choice. Since we have categorized the cylinder count as nominal data, a better approach might involve using shape to represent the number of cylinders. Furthermore, while the model names are available through hover text, they are not directly visible in the plot.

Another potential limitation involves browser compatibility. So far, we have successfully tested the visualization in Chrome, but with less success in Safari. Future work could include testing the visualization across different browsers and optimizing performance to ensure a better user experience.

References

[Bertin, 1983] Bertin, J. (1983). Semiology of graphics.

[Lall, 1980] Lall, S. (1980). The international automotive industry and the developing world. *World Development*, 8(10):789–812.

[Mackinlay, 1986] Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *Acm Transactions On Graphics (Tog)*, 5(2):110–141.

[Mackinlay, 1988] Mackinlay, J. (1988). Applying a theory of graphical presentation to the graphic design of user interfaces. In *Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software*, pages 179–189.