Assignment 2 — Tableau Application

Sourabh D. Khot (Student ID 002754952)

College of Professional Studies, Northeastern University

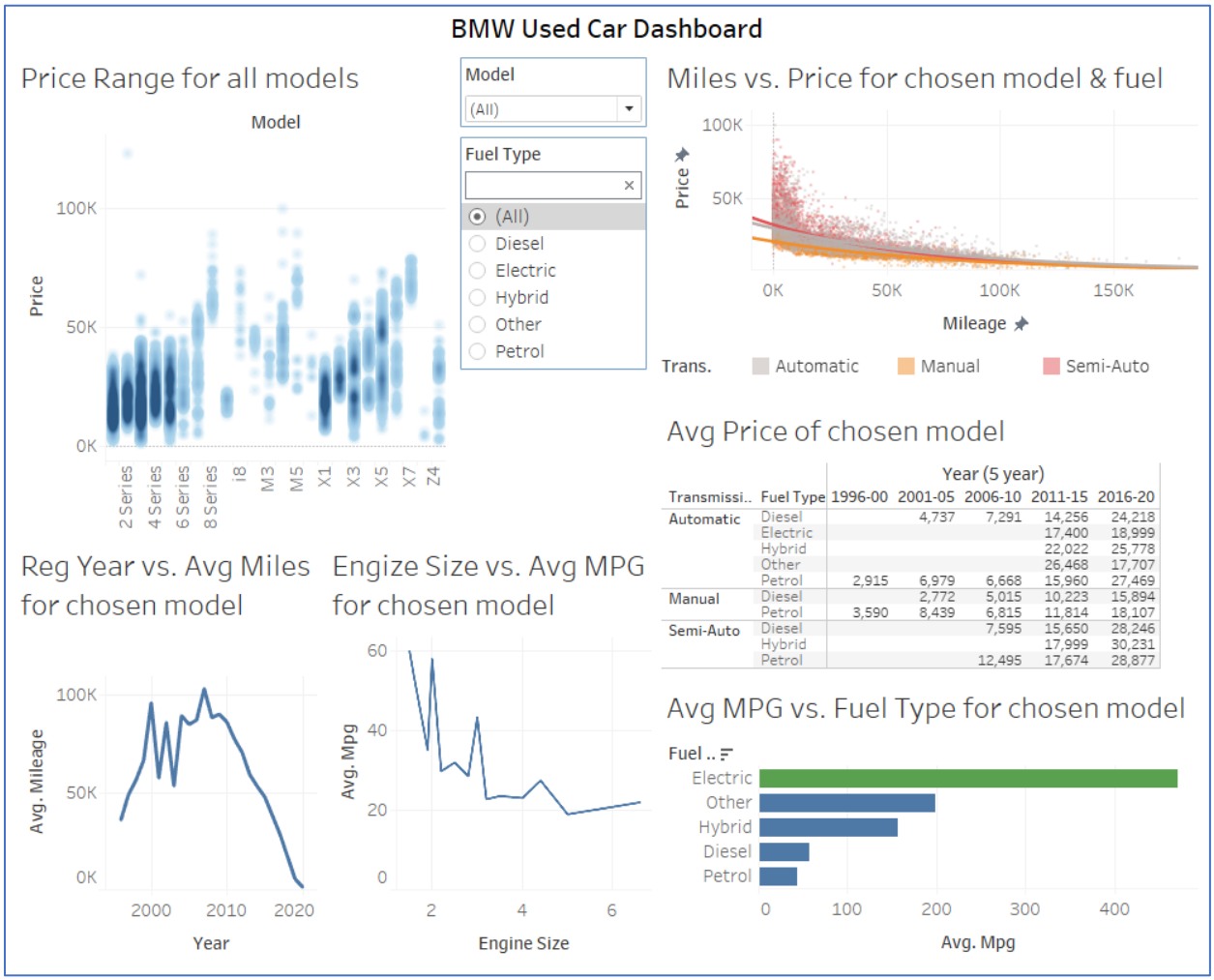
ALY 6070: Communication and Visualization for Data Analytics CRN 81182

Professor Venkata Duvvuri

Jun 16, 2022

Rationale

I created the following dashboard using the 'BMW used car listing' dataset from Kaggle.com. The research questions are what price to list (seller), what price to expect (buyer), and how vital characteristics are interlinked with each other and the price.



Why did you choose the types of visualizations that you did?

Each chart was tailored to the critical questions that a buyer or seller may have. A buyer, especially one who is fixated on one car model, will always ask for what range of prices their model is being sold. The density chart best describes the range as shown in the first chart for price range of each BMW model.

The boxes with blue borders represent inputs of car model and fuel type, that a buyer or seller has to enter. Price has the highest correlation with mileage, and the scatterplot shows this exponential trend, which can help buyers and sellers arrive at a desired price. A buyer or seller can select their required fuel type and obtain three trend lines for each transmission type.

The line chart shows the relationship between continuous aggregate numbers and categories well. The bottom left line chart on the year of car registration vs. average miles driven. This can help buyers get an idea of this relationship to decide their preferences. The bottom center and bottom right charts are related to fuel economy. A buyer can understand the fuel economy vs. engine size (performance) tradeoff with the line chart. A sorted vertical bar plot was chosen for average MPG vs. Fuel Type so a buyer can easily compare fuel economies of various fuel types. The highest fuel economy (MPG equivalent) is for electric vehicles, which is phenomenal and hence highlighted in green.

Finally, a buyer or seller will ask for the average price for various model variants. These prices will make more sense after knowing the range of registration years. A table is a simple and great way to demonstrate cross-tabulate data for a combination of features. Hence, the average price is shown in table format for cars registered groups of half decades.

How are the visualizations effective and address the gestalt and design principles discussed in the course?

These visualizations are compelling in answering each question a buyer or seller may have. The gestalt principles are addressed in the following ways.

- Proximity – Each chart element is close to the respective chart. E.g., the legend of the scatterplot is placed just below it.
- Similarity – is used in the scatterplot to show that each transmission is similar color.

- Enclosure – Input sections are enclosed in blue to highlight that they are editable.
- Closure – Chart borders are removed to reduce cognitive load as they are understood.
- Continuity – Axis from vertical bar plots are removed as it is implied due to continuity from where the bars start.
- Connection – Line charts are used instead of bar charts for two relationships because the connection between variables is better understood through this type.

How do the visualizations answer the research/business question?

The visualization directly questions users may have, as answered in the first question above. In particular, the density plot tells a buyer what will be the range of the model they are interested. The price to set for a seller can be determined by getting an idea from the scatterplot for the chosen model and fuel type. A buyer can understand tradeoffs of miles vs. registration year and between performance, fuel economy, and fuel type.

What story do the visualizations tell?

The visualizations highlight several aspects as part of its story from a research point of view that can benefit sellers, buyers, or others related to this industry. These aspects are given below, sorted by level of importance.

- Price of a car decreases exponentially for the miles driven. In general, older cars have more miles on them, so it can also be said that price drops exponentially to car age.
- Powerful cars are fuel guzzlers, which is obvious. However, what is astonishing is that electric ones are more than double as efficient as hybrids, and eight times as petrol ones. Thus, electric vehicles can save a lot over long distances.
- As per this sample, diesel cars were introduced after 2000, and hybrid, other fuel, and electric cars came after 2010. Similarly, semi-auto transmission cars came after 2005.

References

APA Style Table: APA.org. (n.d.). Retrieved from https://apastyle.apa.org/style-grammar-guidelines/tables-figures/tables

BMW used car listing: Kaggle.com. (n.d.). Retrieved from https://www.kaggle.com/datasets/mysarahmadbhat/bmw-used-car-listing

Canvas ALY6070: Assignment 2 Tableau Application. (n.d.). Retrieved from https://northeastern.instructure.com/courses/110061/assignments/1351346

Canvas ALY6070: Lesson 4-1 to 4-8 Storytelling with Tableau: Importance of Context and Application. (n.d.). Retrieved from https://northeastern.instructure.com/courses/110061/pages/module-4-introduction

Knaflic, C. N. (2015). *Storytelling with Data - A Data Visualization Guide for Business Professionals.* New Jersey: Wiley.

<End of Report>