Assignment 4 - Dashboard and Story Development using Cognos Analytics

Tanushree Kumar | tanushree.kumar@outlook.com

DATA 610 - Fall 2023

Dr. George Cross

Introduction

The dataset that I will be using for dashboard and story development using Cognos Analytics is the Used Car Prices (Reddy, 2018). The dataset contains 1436 observations across 12 columns on various attributes of the vehicles. They are: ‘Price’ (price of the car); ‘Age’ (age of the car); ‘KM’ (car mileage in kilometers); ‘FuelType’ (petrol, diesel, or CNG); ‘HP (Horsepower); ‘MetColor’ (is the car a metallic color (1 = yes, 0 = no); equivalent to ‘MetColorType’); ‘Automatic’ (is the car automatic (1 = yes, 0 = no); equivalent to ‘AutoType’); ‘CC’ (volume of the cylinder in cubic centimeters), ‘Doors’ (number of doors); ‘Weight’ (weight of the car). While the dataset has many attributes, some of them are unnecessary for the purpose of this assignment. Similar to Assignment 3, I removed the following columns as they had no significant impact: ‘HP’, ‘MetColor’, ‘Automatic’, ‘Doors’, ‘Autotype’, and ‘MetColorType’. A snapshot of the original dataset can be seen in Figure 1 in Appendix A.

The objective of this assignment is to develop an insightful dashboard and to discuss the insights gained from the visualizations. Then use that dashboard to make a Story that presents the information in a visually appealing way.

Dashboard Presentation and Discussion

Two key findings were gained from Assignment 3. The first is that ‘Age' is a crucial predictor for 'Price,' and the combination of 'Age,' 'KM,' 'CC,' and 'HP' provides the strongest predictive strength. The second is that 'Price' is the most important predictor for 'Age,' and the combination of 'Price,' 'CC,' 'KM,' and 'FuelType' has high predictive strength.

I generated 5 additional insights to further understand the dataset:

- Effect of ‘FuelType’ on ‘Age’: Diesel and CNG cars may predict a lower age than petrol cars, impacting the inventory strategy.
- Influence of ‘Weight’ on ‘Price’: Although removed from the model, understanding the impact of 'Weight' on 'Price' could provide insights into customer preferences.
- Market price trends: Explore how market prices for new and used cars, as mentioned in Domonoske (2023), align with the dataset.
- Correlation between ‘HP’ and ‘Price’: Analyze if there's a correlation between 'HP' and 'Price' that could influence the purchasing decision.
- ‘Age’ and Depreciation: Investigate the relationship between 'Age' and the rate of depreciation, considering the changing trends in the market.

The objective of the Effect of ‘FuelType’ on ‘Age’ is to understand how the type of fuel (petrol, diesel, or CNG) affects the predicted age of used cars. This was done by examining the decision tree rules and visualizations related to 'FuelType' in the context of predicting 'Age.' I then checked to see if diesel and CNG cars tend to have a lower predicted age compared to Petrol cars. The potential implications of this are if diesel and CNG cars predict a lower age, a used car dealership might consider adjusting inventory to include more of these types. This strategy could be based on the assumption that younger cars are generally more attractive to customers.

The Influence of ‘Weight’ on ‘Price’ could offer insights into customer preferences. By examining the historical data and patterns related to 'Weight' and 'Price', we can consider whether heavier cars tended to have higher or lower prices. This could provide insights into whether customers associate heavier cars with higher value, safety, or other factors. Understanding customer preferences related to 'Weight' could guide marketing strategies. For instance, if customers prefer lighter cars, which can contribute to better fuel efficiency, emphasizing this feature in marketing campaigns might be beneficial.

Investigating how the dataset aligns with the market trends mentioned in Domonoske (2023) regarding new and used car prices would also be helpful. By comparing the average prices of the used cars in the dataset with the market trends mentioned in the article, we can look for any notable deviations or similarities. If the dataset aligns with market trends, it validates the dataset's representativeness. If there are discrepancies, further investigation is needed to understand the factors contributing to these differences. Other factors need to be accounted for first, such as the year of the dataset, the origin, which car model and company was used for data collection, and is the data from a reputable source.

To determine whether there is a significant correlation between the 'HP' of a car and its predicted 'Price', we can calculate the correlation coefficient between 'HP' and 'Price'. Using scatter plots and even regression analysis can help visualize the relationship. A strong positive correlation might suggest that customers are willing to pay more for higher horsepower, influencing pricing and marketing strategies, and the stock a user car dealership might keep.

Exploring the relationship between the 'Age' of a used car and its rate of depreciation would be interesting considering the evolving trends in the market. This would be done by calculating the depreciation rate for different age groups. These rates would then be compared with industry benchmarks or historical data to identify any significant changes or patterns. Insights into the age-depreciation relationship can change pricing strategies. For example, if cars within a certain age range depreciate more slowly, pricing adjustments could be made to reflect this trend. Since newer used cars depreciate more quickly than older used cars, we can get concrete facts for validation.

To assess the impact of 'FuelType' on 'Age', I used a sunburst diagram to showcase the influence of 'FuelType' on predicted 'Age.' I wanted to see if diesel and CNG cars predict a lower age, guiding inventory strategies towards younger, potentially more attractive cars. This proved to be right as the average age of diesel and CNG cars was 64.78 years while the average age of petrol cars was 72.06. The 8-year gap makes quite a difference in the price, quality, and life of the car. The diagram can be seen in Figure 2 in Appendix A.

To see how 'Weight' influences 'Price', I chose to make a bubble plot to understand if heavier cars correlate with higher prices. Depending on the result, a used car dealership could tailor its marketing based on customer preferences related to vehicle weight. Through the chart, I was able to observe that 'Price' moderately drives 'Weight' (45%). The value of 'Weight' is unusually high when the values of 'Price' are 30811.00 to < 31374.00 and 31937.00 and above. The overall trend shows that the price increases while the weight of the car stays in a certain zone. This can be seen as 9417.00 to < 9980.00, 7728.00 to < 8291.00, and 8854.00 to < 9417.00 are the most frequently occurring categories of Price with a count of 532 items (37 % of the total). The diagram can be seen in Figure 3 in Appendix A.

To understand the 'HP'-'Price' correlation, I used a scatter plot to display the relationship to assess if higher horsepower correlates with higher prices, which can influence pricing and marketing strategies. I observed a strong correlation between the two as 'Price' increases as 'HP' increases. This positive association indicates a strong linear relationship: Price = 12259 + 92.77 * HP. The diagram can be seen in Figure 4 in Appendix A.

For the last relationship of 'Age'-Depreciation dynamics, I chose to visualize this by using a bar chart to illustrate the rate of depreciation for different age groups to see how 'Age' influences depreciation. The results of this can advise pricing strategies based on age-related depreciation patterns accordingly. The bar chart shows a downward trend in the price of a car as its age increases. The average values of Price range from a maximum of 31,591.667 (when Age

is 4) to a minimum of 7,592.241 (when Age is 74). Something to note however is that the value of ‘Price’ is unusually high when ‘Age’ is 4, when one would expect this for a car that is only a year old. We can also see that ‘Age’ strongly drives ‘Price’ (86%). The diagram can be seen in Figure 5 in Appendix A.

The dashboard enhances understanding of the dataset and its implications in many ways. First, it provides us with a holistic market view. The combination of visualizations provides a comprehensive view of the used car market, allowing for a holistic understanding of pricing, age prediction, market trends, fuel type impact on age, weight-price dynamics, HP-price correlation, and age-depreciation relationships. The dashboard gives users the ability to dynamically explore the impact of different variables on prices and age prediction, fostering an interactive and engaging learning experience through the sunburst, bubble, and scatter plots. Decision-makers can make more informed decisions as they’ll be able to derive actionable insights for pricing strategies, inventory management, and marketing efforts by synthesizing information from multiple visualizations. The dashboard facilitates real-time adaptation to market changes, enabling the dealership to make timely decisions aligned with current market dynamics and customer preferences. This gives them the flexibility needed to manage an efficient inventory and adjust it as needed. By aligning dataset prices with market trends, the dashboard ensures the validation of dataset representativeness, enhancing confidence in decision-making. Combined with a user-friendly interface, the dashboard ensures accessibility for stakeholders with varying levels of data literacy. This will help promote the widespread adoption of data-driven decision-making within the organization. The full dashboard can be seen in Figure 6.1 and Figure 6.2 in Appendix A.

Story Presentation and Analysis

For the Story, I divided it into the following sections to create a flow. The Introduction includes a brief overview of the dataset and its relevance. The Predictive Model Recap includes the visualizations from Assignment 3, such as the spiral and sunburst diagram. The recap summarizes the key findings gained from the models and diagrams. The Additional Insights presents the new insights that were the focus of this assignment. The Dashboard Visualizations is a walkthrough of the dashboard. The Decision Tree Impact discusses the impact of decision tree rules on strategic decision-making, as discussed briefly in Assignment 3. The last section of Future Exploration has the goal to explore 'KM' as a target variable and its resulting implications.

The story delves into the intricate dynamics of the used car market, where every data point unveils a narrative of pricing, age, and strategic insights. The story of the dataset is that of a journey, which begins with deciphering the puzzle of used car pricing. The spiral visualization highlights that 'Age' stands out as the most influential predictor, trumping others by more than five times. As we explore the decision tree, a nuanced understanding emerges — 'Age,' 'KM,' 'CC,' and 'HP' intricately weave the narrative of a car's price.

Shifting gears, our attention turns to predicting a car's age. Surprisingly, 'Price' emerges as the paramount predictor, wielding influence over five times more than any other field. The decision tree reveals intriguing combinations, 'FuelType' is added to the collection, predicting age dynamics. Diesel and CNG cars display a younger age compared to their petrol counterparts as ‘FuelType’ plays a pivotal role in predicting a car's age, impacting inventory strategies. Staying in sync with market trends ensures competitive pricing and market relevance.

We next investigate the interconnection between 'HP' and 'Price.' The scatter plot showcases the potential correlation, revealing patterns that can shape the marketing strategies to emphasize high horsepower. Cruising through the end, we explore the relationship between 'Age'

and depreciation rates. In a market undergoing shifts, understanding how age impacts depreciation is paramount. The dynamic bar chart brings this to life, aiding strategic decisions on pricing and inventory management.

As our journey concludes, the insights gained from Cognos Analytics transform data points into a comprehensive story. This interactive narrative empowers the used car dealership with actionable strategies, from pricing decisions influenced by age and fuel type to marketing tactics guided by horsepower correlations.

Dashboard and Story Implementation in a Used Car Dealership

In Assignment 3, I discussed how the rules derived from the decision tree model could be used in an organization, in this context, a used car dealership and its inventory. I will be continuing with the same business niche for the utilization of the dashboard and story. There are many advantages of presenting findings as an interactive story. The first one is that the dealership could use a dynamic inventory strategy: The interactive story provides real-time insights into factors influencing the pricing and age of used cars, aiding in the dynamic adjustment of inventory strategies. For example, understanding the impact of 'FuelType' on age prediction allows the dealership to reevaluate its inventory mix to potentially favor diesel and CNG cars due to their lower predicted age. The dealership could explore the market price trends to align its pricing strategies with the broader market. This ensures that the prices set for used cars are competitive and reflective of the current market conditions, as highlighted in Domonoske (2023). The interactive exploration of 'Weight''s impact on 'Price' could offer insights into customer preferences. Understanding that customers correlate vehicle weight with value can enhance the dealership's procurement and marketing strategies. By performing correlation analysis and analyzing the correlation between 'HP' and 'Price' interactivity allows the dealership to uncover patterns that influence purchasing decisions. For instance, if there's a strong correlation, the dealership may emphasize high horsepower as a selling point in marketing campaigns. The dealership could also consider the age of the cars more strategically. They could investigate the relationship between 'Age' and the rate of depreciation in a dynamic market from earlier. If there's a shift in the trend indicating that certain age ranges depreciate slower, the dealership can adjust pricing strategies accordingly.

Although there are many advantages, there are also some disadvantages to this approach. As with any new software and methodologies being implemented, there will be a learning curve that could potentially impact the adoption and utilization of the presented insights. Developing an interactive story and maintaining its dynamic features can be resource-intensive. The dealership needs to be open to the idea that they will have to allocate time and resources for staff training and continuous updates. There's also the possibility of data overload as too much interactivity or complex visualizations might overwhelm users with information and can hinder effective decision-making. It is crucial that there's a balance between detail and simplicity. Although this data might not be deemed as sensitive information, privacy and security should also be taken into consideration. The dealership must implement robust security measures to protect customer and business data, which would also require extra resource allocation. Overall, presenting insights through an interactive story can revolutionize inventory management, pricing strategies, and marketing efforts. The advantages, such as dynamic decision support and market trend alignment, outweigh the disadvantages, but careful consideration must be given to user training, simplicity, and data security. The interactive story serves as a powerful tool for strategic decision-making in the ever-evolving used car market.

References

Bytenskaya, Y. (2021a, March 6). Cognos Analytics - How to create new dashboard. Www.youtube.com. https://www.youtube.com/watch?v=0-fOLArlyfw&list=PL2r2WGYKOnJWvwjzgazpro6MzdmvLZqBj&index=20

Bytenskaya, Y. (2021b, March 14). Cognos Analytics - Export dashboard as a story. Www.youtube.com. https://www.youtube.com/watch?v=an93WQsmyH8&list=PL2r2WGYKOnJWvwjzgazpro6MzdmvLZqBj&index=23

Domonoske, C. (2023, March 18). *Why car prices are still so high — and why they are unlikely to fall anytime soon*. NPR. https://www.npr.org/2023/03/18/1163278082/car-prices-used-cars-electric-vehicles-pandemic#:~:text=July%209%2C%202021.-

IBM Cognos Analytics Version 11.0.0 Dashboards and Stories User Guide IBM. (n.d.). https://www.ibm.com/docs/en/SSEP7J_11.0.0/com.ibm.swg.ba.cognos.ug_ca_dshb.doc/ug_ca_dshb.pdf

Reddy, N. (2018). usedcarprices. Www.kaggle.com. https://www.kaggle.com/datasets/nitinreddy98/usedcarprices/data

Appendix A

All mentioned tables and figures throughout the paper can be found here.

Figure 1

A snippet of the Used Car Prices dataset

Row Id	Price	Age	KM	FuelType	HP	MetColor	Automatic	CC	Doors	Weight	AutoType	MetColorType
1	13500	23	46986	Diesel	90	1	0	2000	3	1165	Manual	Metcolor
2	13750	23	72937	Diesel	90	1	0	2000	3	1165	Manual	Metcolor
3	13950	24	41711	Diesel	90	1	0	2000	3	1165	Manual	Metcolor
4	14950	26	48000	Diesel	90	0	0	2000	3	1165	Manual	NonMetcolor
5	13750	30	38500	Diesel	90	0	0	2000	3	1170	Manual	NonMetcolor
6	12950	32	61000	Diesel	90	0	0	2000	3	1170	Manual	NonMetcolor
7	16900	27	94612	Diesel	90	1	0	2000	3	1245	Manual	Metcolor
8	18600	30	75889	Diesel	90	1	0	2000	3	1245	Manual	Metcolor
9	21500	27	19700	Petrol	192	0	0	1800	3	1185	Manual	NonMetcolor
10	12950	23	71138	Diesel	69	0	0	1900	3	1105	Manual	NonMetcolor

Figure 2

Sunburst diagram for the effect of 'FuelType' on 'Age

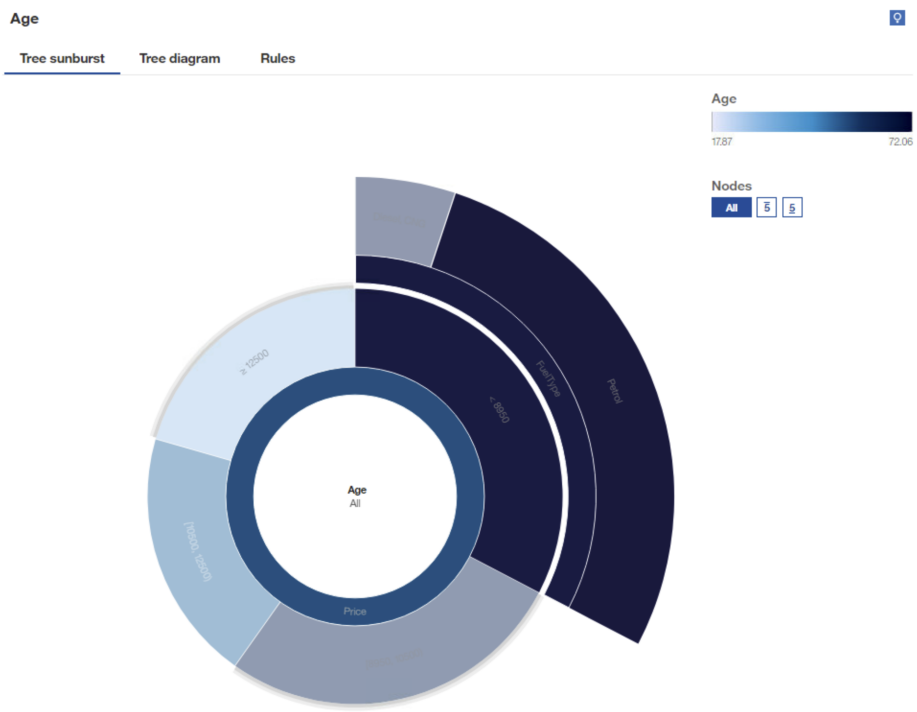

Figure 3

Bubble chart of the influence of 'Weight' on 'Price'

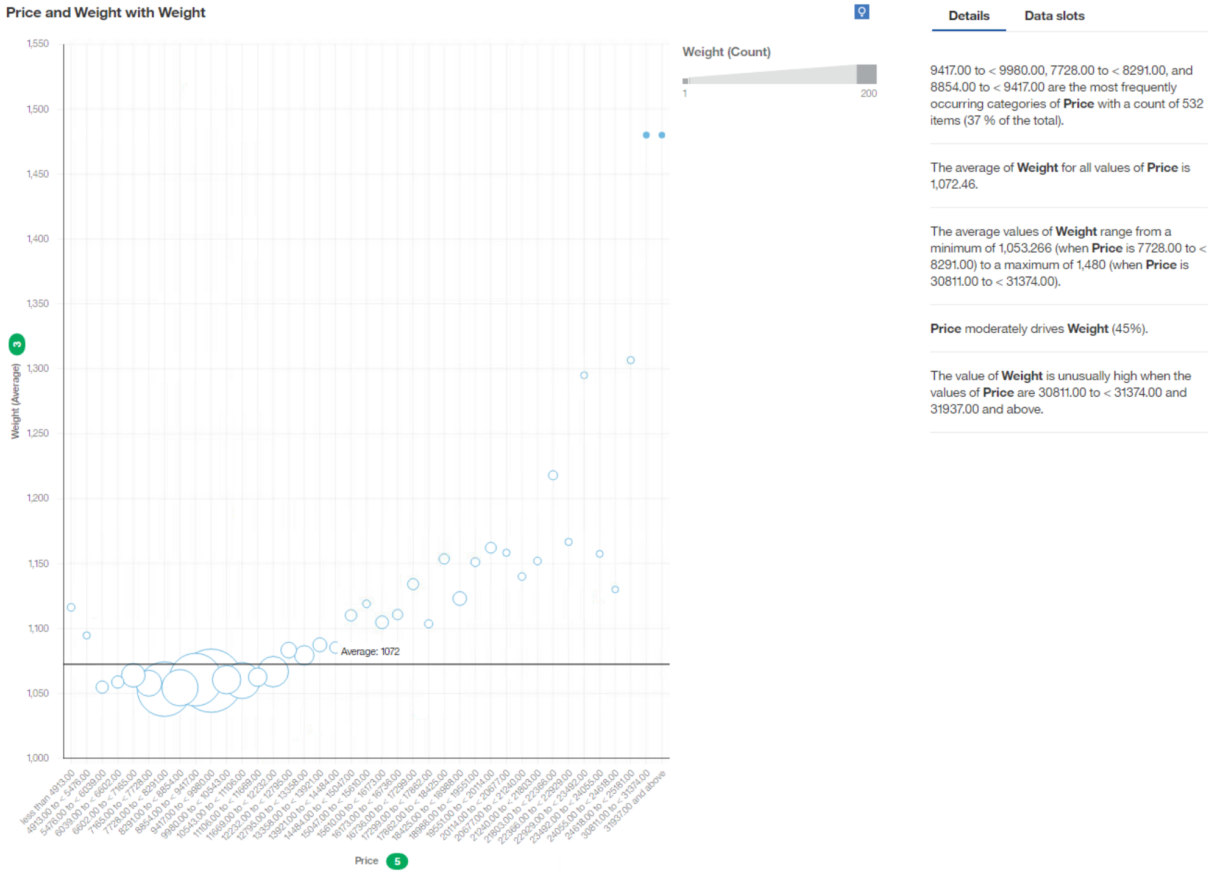

Figure 4

Scatter plot showing the correlation between 'HP' and 'Price

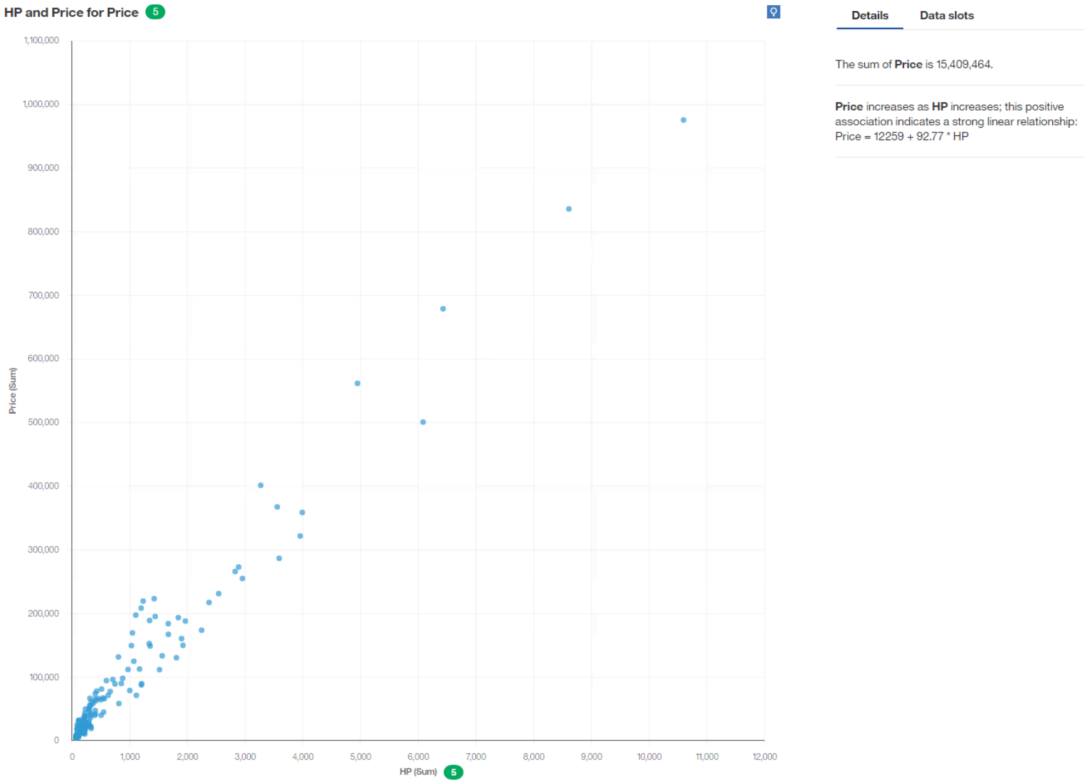

Figure 5

Bar graph showing 'Age' and depreciation of value

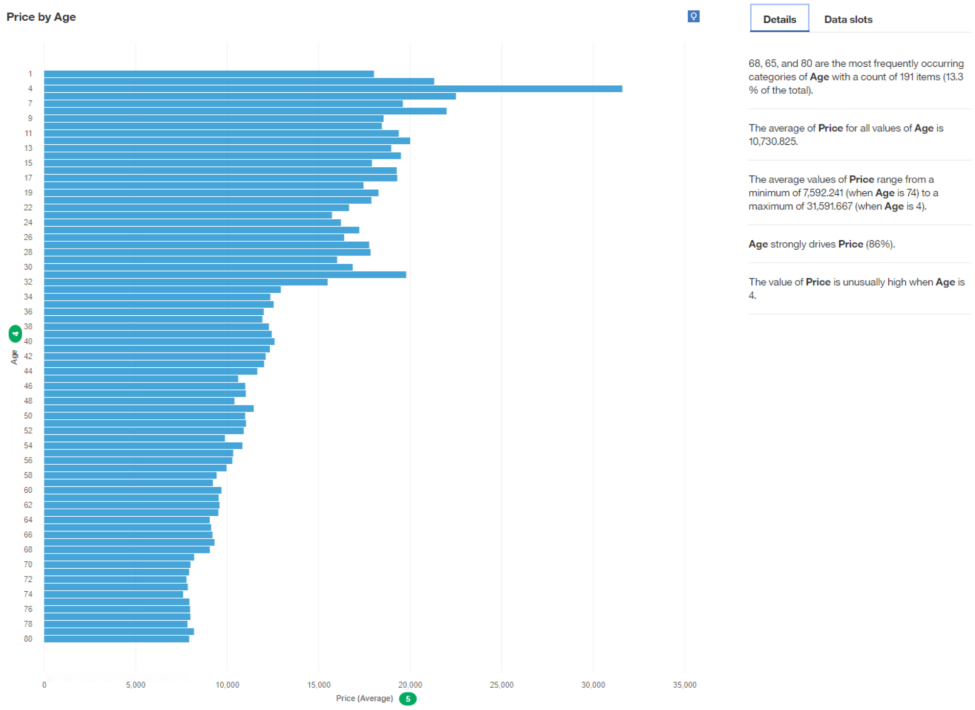

Figure 6.1

Dashboard created for the Used Car dataset

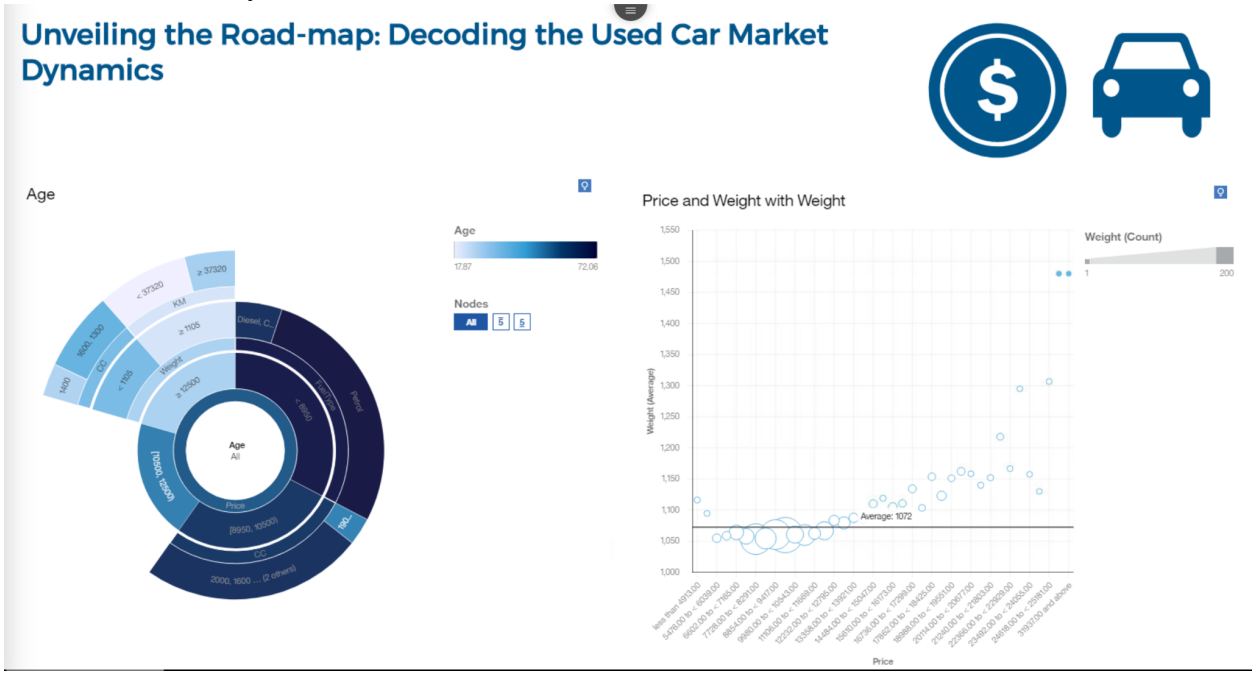

Figure 6.2

Dashboard created for the Used Car dataset