[1]

Used Car Sales vs. New Car Sales

Diamond Burton, Member, IEEE
Department of Computer Systems Technology
North Carolina A&T State University
Greensboro, NC, USA

***Abstract*— The following term paper for Statistical methods, CST 702, will dissect the findings from the Maryland car sales dataset that has been recorded for over 20 years, specifically from 2002 to August 2022. Often, car buyers consider which car would be a better investment for them. The following dataset will explore the sales for both used and new cars in Maryland, and also for forecasting for 2023.**

***Index Terms*— Analysis, Forecasting, Seasonality, Maryland**

I. INTRODUCTION

A. CASE

Recognizing risk factors that may come from purchasing a car is imperative. For example, a customer may be someone who prefers a car with no history and has brand bells and whistles, while another customer may prefer a car that has a low depreciation value with at least one prior owner. Data analysis allows stakeholders like customers to make profound decisions. Within the past few years since the height of the pandemic, used cars have taken a turn and began surging over new cars. Some people believe this is due to the new interest in EVs (electric vehicles), while others may believe it is due to the Ukraine and Russia wars. [5] However, it has been proven through analysis that used cars will surpass new cars in 2023 based on forecasting statistics.

B. DATA

Big data is defined as large, raw data detailed by its volume, variety, velocity, and veracity. The data we will observe is coming from www.data.gov. The publisher is opendata.maryland.gov This data was recorded from 2002 through August 2022. The rows included every year parallel to every calendar month. The columns include (Month, Year, New, Used, Total Sales New, and Total Sales Used.) There were 248 rows of data and this data is still being updated in real-time. The used/new car sales were recorded by Maryland sales but it also has the total sales for the country being recorded as well. We see an increase in used car sales over time with this data. Many factors caused the data to surge based on season, months, and the significance of events throughout the years and we will see specifically how the sales were impacted by the pandemic.

II. RELATED WORK

It was extremely helpful to research for a more in-depth understanding of how car sales work and the demand for used cars and why used cars are having a surge over new cars. Except for vintage vehicles such as Porsches, cars & trucks depreciate as soon as it is driven out of the lot. New cars lose around 20% of their value within the first year of ownership and 60% within five years on the road. [1] According to Forbes, in recent years, there has been a shortage of new cars, fueled by supply chain issues. They are in high demand and the prices are reflecting as such. The average used car cost around $28,000 at the end of 2021. [1] This is 28% more than the same point a year before 2021. Buying used cars primarily as a cost-saving measure but in today's times, this isn't the case anymore. Used cars are in high demand and pre-orders are already piling up for 2023. [4] Many customers have the huge question of whether not to buy a used or new car. This article explains the pros and cons of both. When buying a new car the pros are: No research history, and new car deals but the cons included the expense of the car and the price of sales taxes. On the other hand, used cars are less expensive and won't take a huge depreciation hit. However, there's often no warranty and financing is more expensive. [4] As of November 2022, new car prices are rising due to an ongoing shortage and increase in raw market Al costs. According to JP Chase Morgan, new vehicle prices are up 6.3% within the last year. This fielded the demand for more used cars in 2020 which went up 42.5% in 2022 v in 2020. [5]JP Morgan explains that car prices are rising due to global supply chain issues. Which was exacerbated by the Russia-Ukraine crisis.

This further supports the dataset that I used for the analysis. I observed 248 rows of data and six columns. The dataset columns were (year, month, new car sales in Maryland, used car sales in Maryland, total sales new, and total sales used.) This data was recorded from 2002 up until August 2022. The types of transportation recorded were cars, trucks, preowned, etc. in Maryland. [3]

A. PROBLEM STATEMENT

Data analysis is used in pretty much every industry today for stakeholders to make the most logical decisions. Cars are a human resource as we use them for transportation and day-to-day errands. Stakeholders in the industrial industry will use big data to give suggestions to manufacturers on production to reduce inventory waste. The following data records show how used and new cars have been selling over time. I've noticed that used car sales have increased over time while new car sales surprisingly have not. This is shocking considering new cars tend to be more expensive and the assumption is that

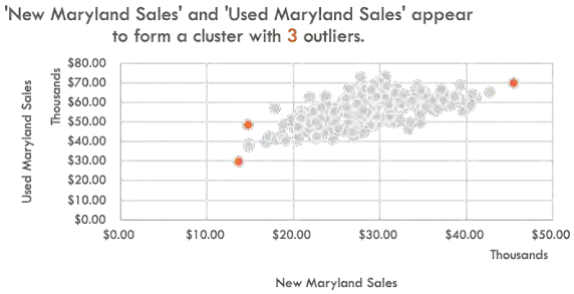
[1].

customers want to drive off the lot with a new car, but this data proves differently. I want to predict how sales will look for both used and new car sales for manufacturers to predict how much inventory they need and other stakeholders to make necessary marketing plans for sales. To help analyze the data, I researched the following questions:

1. Which year had the highest sales? Used & New
2. Which year had the lowest sales? Used & New
3. Compared to the total car sales in the United States, what percentage did Maryland sell in the summer of 2021? June-August or the Winter months: October-December
4. What were the average sales for used cars in Maryland through the 20 years?
5. What were the average sales for new cars in Maryland throughout the 20 years?

With these questions, I was able to form my hypothesis on which car will surpass sales, the estimates, and a sales trajectory for 2023. I hypothesize that used car sales will surpass new car sales.

III. DATA ANALYSIS AND RESULTS



. ***Figure 1: 'New Maryland Sales' and 'Used Maryland Sales'***

Figure 1: 'New Maryland Sales' and 'Used Maryland Sales'

Based on the entire analysis from 248 rows of data, and scholarly research, I was able to focus on the relationship between used car sales and new car sales. Figure 1 shows that used car sales and new car sales have a linear relationship with each other except for the three outliers colored in orange. The cluster is formed between $15,000 and $40,000. We see that the highest sales for new cars were in 2005, with $45,547. The highest for used cars was in 2016 with 73,163.00. The lowest sale for new cars was in 2020 with $13,700. While the lowest sale for used cars was in 2020 with $29,420. Figure 2 shows the sum of average sales between used and new cars. Used cars sell more, on average.

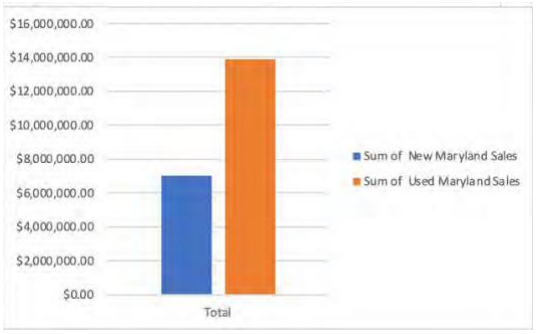


Figure 2 : Average sales for used and new

Based on regression statistics, the multiple R is at 61% showing that the y variable (used car sales) is highly accounted for how new car sales perform. 38%% of variance in used car sales, specifically account for it. Figure 3 shows an example of a line fit plot, showing the predicted used car sales. I wanted to predict used car sales based on new car sales since the two variables have a linear relationship. Based on this model, used car sales will continue to increase and the difference between one of the points in the plots is very significant. One of the used car sales was at 62,811 and was predicted to rise to 65,915. The higher the new car sale, the more used cars sell.

Figure 3: Predicted Used Car Sales

Figures 4 and 5, show the sales of used and new cars over time. There is a lot of seasonality in both models, which shows the ups and downs on the trend lines. This is due to cars selling more or less in certain seasons. Based on forecasting analytics, used car sales will continue to increase through 2023 and new car sales will decrease.

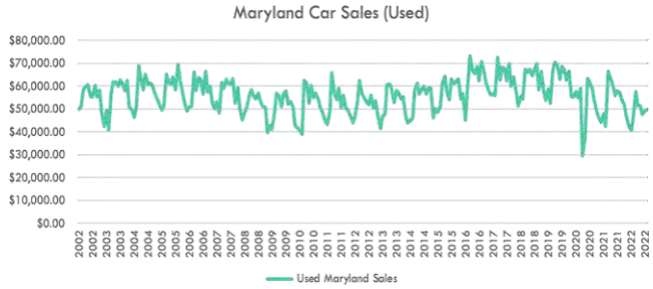


Figure 5: Sales over time for used cars

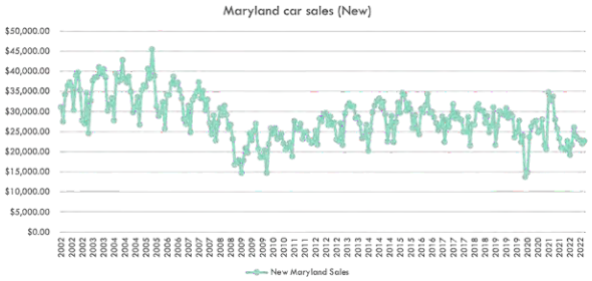
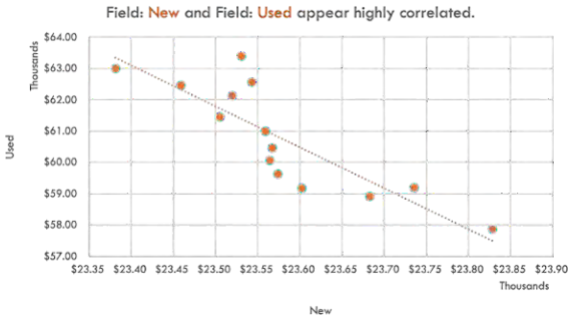


Figure 5: Sales over time for new cars



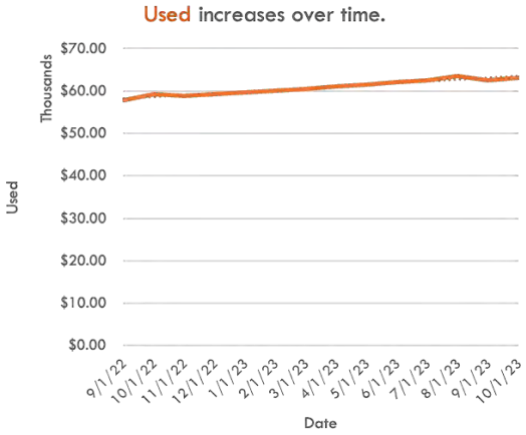


Figure 6: Forecast used and new car sales

II. Conclusion

Transportation is a huge stimulus in the economy as though it is an imperative resource that keeps businesses going. We have analyzed the relationship between used and new cars as well as forecasted the performance of how used cars will sell in 2023. This will allow merchants, salesmen, and other stakeholders to make logical decisions on what to market to buyers. Used cars in Maryland are expected to increase and surpass new cars in 2023 based on forecasting statistics by $5,612.83. New cars will decrease by $1,214 while used cars will increase by $6,827.42.

A. Suggestions to Stakeholders and Manufacturers

Since used car sales are projected to increase in sales, based on the data, I would suggest the following:

1. SLOW DOWN ON MAKING INCREASING AMOUNTS OF NEW CARS
2. FOCUS ON MARKETING USED CARS NO OLDER THAN 3 YEARS OLD
3. ELECTRIC CARS ARE BEGINNING TO PICK UP SALES, IT IS A GOOD TIME TO FOCUS ON HOW TO INCORPORATE THE BENEFITS FROM USED CARS INTO EV'S (ELECTRIC VEHICLES) FOR THE SURGE THAT IS COMING IN THE NEXT TEN YEARS

B. Future Work

The information presented here will suggest how the market will look for future sales in 2023 and further. The market for used cars will be more in demand and will be greatly impacted by the surge in EV's and other economic factors. Stakeholders can also utilize this information to impact when they can stop making gas vehicles altogether as new technology arises, or even how to upgrade technology in gas cars to meet the needs of buyers.

REFERENCES

[1] J. Gorzelany, "Mutual appreciation: Used cars that sell for more than new ones," *Forbes*, 09-Feb-2022. [Online]. Available: https://www.forbes.com/sites/jimgorzelany/2022/02/07/mutual-appreciation-used-cars-that-sell-for-more-than-new-ones/?sh=6e6cff9d4841. [Accessed: 27-Nov-2022].

[2] Evans, J.R. *Statistics, Data Analysis, and Decision Modeling*, 2013.

[3] "MVA Vehicle Sales Counts by Month for Calendar Year 2002 through August 2022 | Open Data | opendata.maryland.gov," *opendata.maryland.gov*. https://opendata.maryland.gov/Transportation/MVA-Vehicle-Sales-Counts-by-Month-for-Calendar-Yea/un65-7ipd (accessed Nov. 28, 2022).

[4] "New cars vs. used cars | U.S. news," *New Cars Vs. Used Cars*, 29-Jan-2019. [Online]. Available: https://cars.usnews.com/cars-trucks/advice/new-cars-vs-used-cars. [Accessed: 28-Nov-2022].

[5] J.P. Morgan Chase, "When will car prices drop?: J.P. Morgan Research," *J.P. Morgan*, 22-Aug-2022. [Online]. Available: https://www.jpmorgan.com/insights/research/when-will-car-prices-drop. [Accessed: 27-Nov-2022].

[6] Qi Zhang, Hongfei Zhan, Junhe Yu, "Car Sales Analysis Based on the Application of Big Data, Procedia Computer Science" https://doi.org/10.1016/j.procs.2017.03.137. [Accessed: 29-Nov-2022]

[7] Berkovec, J. *"New Car Sales and Used Car Stocks: A Model of the Automobile Market"* https://doi.org/10.2307/2555410. [Accessed: 29-Nov-2022]

Diamond K. Burton Burton was raised in Raleigh, NC. From being a knight at Knightdale High School to becoming a first-year graduate, at UNCG, and now an illustrious Aggie at A&T State University. Her journey started in 2014 when she graduated at 17 years old and went to Louisburg College to complete General Ed classes. She had the pressure of being the first grad in her family, but this pressure is what also kept her motivated. On May 6, 2019, I graduated with my B.A. in Communication Studies and English and soon was offered a career in education. Since 2019, she has been an educator, teaching ELA and Math. As teaching wasn't the top profession of choice, she soon grew a love for kids and helping them reach their highest potential. As an educator, she saw the political issues in the education system and how teachers and students are treated and decided to step away. However, she also had a love of the different aspects of teaching such as analyzing data, data entry, etc. With this passion, she noticed that there are people who analyze and truly study data for a living. This knowledge led her to want to enter the field of technology to become a Data Scientist. Because she did not have a professional background, education, or experience in technology, she knew that she was going to have to go back to school to get her master's. North Carolina A&T was the first school that came to mind. She heard so many good things about the program and she wanted the best support with the best professors to help me reach my goals.