X11 DECOMPOSITION

DS 809 - Time Series Analysis

Final Project

Team Members: Ben Stafford, Sai Venkat, Vikas Sabbani

Introduction

Time series analysis plays a crucial role in understanding and forecasting various phenomena characterized by sequential data points collected over time. One of the widely used methods for time series decomposition is the X11 decomposition technique. In this paper, we delve into the application of the X11 decomposition method for analyzing and forecasting retail car sales data.

Our dataset comes from the US Census Bureau and spans from 1992 to 2023, offering a comprehensive view of retail sales across various industries within the US. To narrow our focus and tailor our insights, we selected a particular industry, new cars sales, to help extract meaningful patterns and trends within a specific sector. Additionally, we considered a recent timeframe, concentrating on the years from 2000 - 2023 rather than the entire dataset dating back to 1992. The idea behind this was to provide clearer insights for forecasting within the 21st century.

The X11 decomposition method is an approach developed by the U.S. Census Bureau for decomposing a time series into its underlying trend, seasonal, and irregular components. Unlike simpler decomposition methods, such as the additive or multiplicative decomposition, X11 offers greater flexibility by incorporating both additive and multiplicative adjustments. It also employs sophisticated techniques to handle irregular fluctuations and outliers in the data.

The dataset under consideration in this study comprises monthly retail car sales data from the year 2000 to 2023. Retail car sales represent a crucial economic indicator, reflecting consumer spending patterns, economic conditions, and market trends. The time series nature of this data, characterized by monthly observations over a 23-year period, makes it suitable for analysis using X11 decomposition.

The relevance of the X11 decomposition method for this dataset lies in its ability to effectively capture the underlying patterns and seasonal variations present in retail car sales data. By decomposing the time series into its components, we can better understand the long-term trends, seasonal fluctuations, and irregular variations in car sales over time. Furthermore, the flexibility of the X11 method allows for robust adjustment and forecasting, making it a suitable choice for analyzing retail car sales over time.

Method

A time series data comprises three components: a trend, seasonal and residual component. The decomposition method is used to extract these three components to help improve the understanding of the time series, but it can also be used to improve forecast accuracy.

(Yaffee, 2000, iii) In the classical decomposition method, there are two types of additive and multiplicative decomposition. Additive decomposition is generally used when the seasonal variation is independent of the trend, whereas the multiplicative component is used when the seasonal variation is proportional to the trend.

An alternative to using a multiplicative decomposition is to first transform the data until the variation in the series appears to be stable over time, then use an additive decomposition.

When a log transformation has been used, this is equivalent to using a multiplicative decomposition because:

$$\mathbf{yt = St \times Tt \times Rt} \text{ is equivalent to } \mathbf{logyt = logSt + logTt + log\ Rt}$$

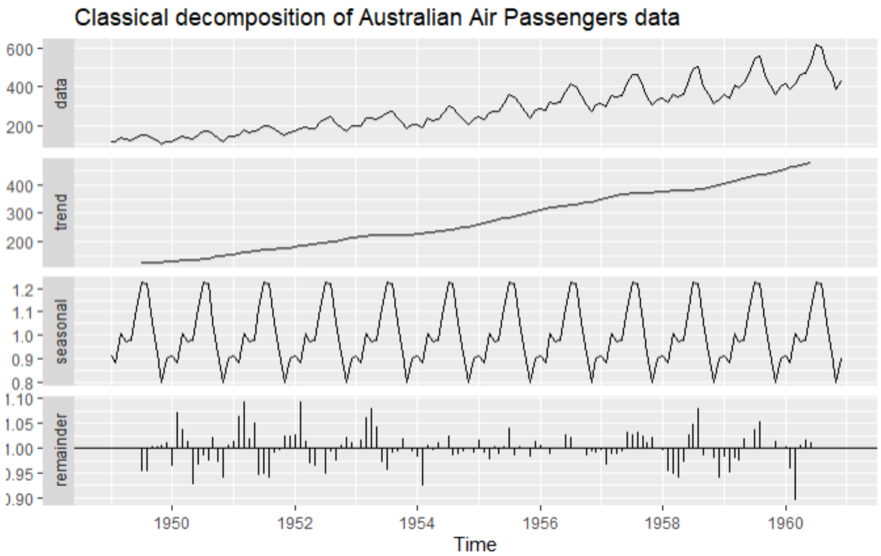
Classical decomposition originated in the 1920s and was widely used until the 1950s. It still forms the basis of many time series decomposition methods, so it is important to understand how it works.

(Hyndman and Athanasopoulos, 2020, ii) The X11 decomposition process refers to the method used to extract seasonal and trend components from a time series. It is a commonly used technique in time series analysis and forecasting. The specific algorithms used in the X11 decomposition process include:

- Seasonal Adjustment: This step involves removing the seasonal component from the time series data. The X11 method typically uses a combination of moving averages and ratios to estimate the seasonal component.
- Trend Estimation: The trend component represents the long-term pattern or direction of the time series. Various algorithms can be used to estimate the trend, such as centered moving averages or regression analysis.
- Irregular Component: Also known as the residual component, this represents random fluctuations or noise in the time series data. It is calculated by subtracting the estimated seasonal and trend components from the original time series.
- Multiplicative Decomposition: In some cases, the X11 method employs a multiplicative decomposition approach, where the seasonal component is expressed as a proportion or ratio of the original time series. This can be useful when the seasonal patterns vary in magnitude over time.
- Moving Averages: Moving averages are frequently used in the X11 decomposition process to smooth out fluctuations and identify underlying patterns in the time series data.
- Ratio-to-Moving-Average (RMA) Method: This method is used to estimate the seasonal component by calculating ratios of observed data to moving averages. These ratios are then adjusted to account for irregularities and outliers in the data.

Toy example of Classical decomposition versus X11 decomposition on the Australian Passengers dataset within R, using the seas() function:

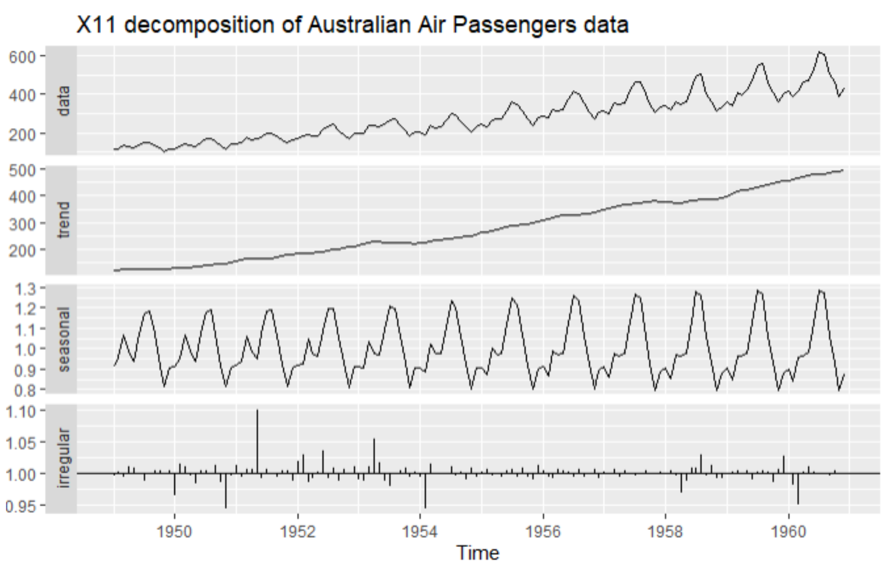
(Figure 1)




```
AirPassengers %>% seas(x11="") -> toy

autoplot(toy) +
  ggtitle("X11 decomposition of Australian Air Passengers data")
```

(Figure 2)



Application

The first step in a classical decomposition is to use a moving average method to estimate the trend-cycle, so we begin by discussing moving averages.

Moving Average Smoothing:

Moving average smoothing is a calculation used to analyze data points by creating a series of averages from different subsets of the full data set. It assists in identifying trends, patterns, and underlying structures within the data by reducing noise and fluctuations. The primary purpose of moving average smoothing is to identify trends by smoothing out short-term fluctuations, and it can also be utilized to make predictions about future trends.

There are 3 types of moving averages:

- Simple Moving Average (SMA):
  - Which calculates the average of a specified number of data points over a defined period.
- Exponential Moving Average (EMA):
  - Gives more weight to recent data points, making it more responsive to changes.
- Weighted Moving Average (WMA):
  - Assigns different weights to different data points.

(Deppa, 2019, i) There are many alternative methods for decomposing a time series without explicitly employing moving averages, few of them are Local Regression Methods, Exponential Smoothing and Statistical Models such as ARIMA (Auto Regressive Integrated Moving Average).

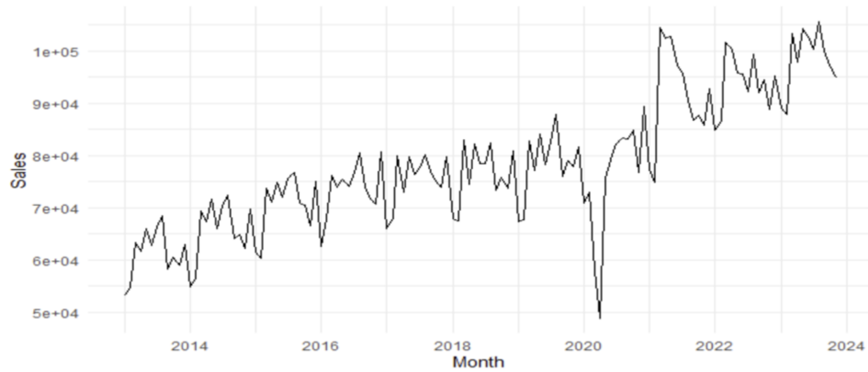
While classical decomposition is still widely used, it is not recommended, as there are now several much better methods. Some of the problems with classical decomposition are summarized below.

- The estimate of the trend-cycle is unavailable for the first few and last few observations.
- The trend-cycle estimate tends to over-smooth rapid rises and falls in the data.
- Classical decomposition methods assume that the seasonal component repeats from year to year. For many series, this is a reasonable assumption, but for some longer series it is not.
- There are certain time periods within the dataset where the values or observations deviate significantly from the expected pattern due to outliers, the classical method is not robust to these kinds of unusual values.
- The X-11 decomposition method is designed to address most of the issues mentioned above.
- X-11 is designed for seasonal adjustment, making it particularly suitable for datasets that exhibit regular seasonal patterns.
- If the data has non-repeating or irregular seasonal patterns, X-11 is designed to handle such situations.

Let’s explore our data to check if X11 decomposition is suitable for our data.

Car Retail Sales Over Time:

(Figure 3)



Stationary Test:

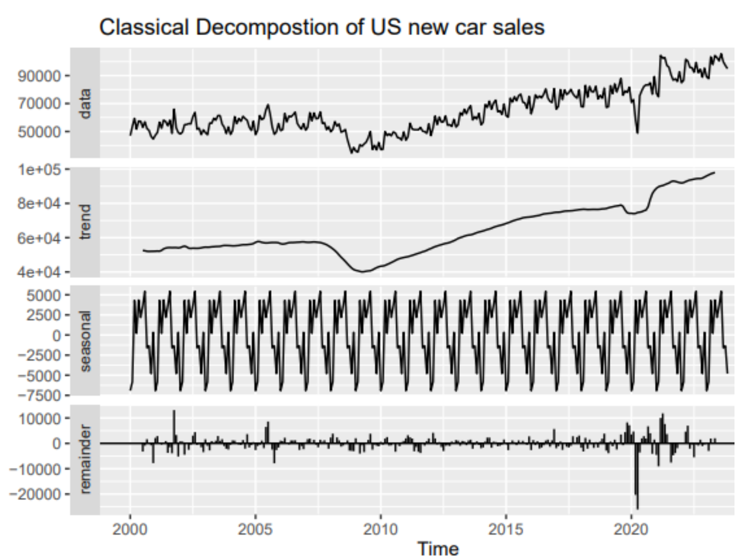
```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  48811   69522   76711   78182   86205  105627 
Warning: p-value smaller than printed p-value
	Augmented Dickey-Fuller Test

data:  ts_data
Dickey-Fuller = -4.9463, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

The negative Dickey-Fuller test statistic and the low p-value provide evidence that your time series is likely stationary. The Augmented Dickey-Fuller test supports the hypothesis that there is no unit root in your data, suggesting stationarity.

Classical Decomposition:

(Figure 4)



An increase in the sales over time signifies an upward trend. A temporary deviation or anomaly is suggested by the lower bump in the middle of the trend, possibly resulting from specific events, factors, or external influences impacting the ongoing upward movement. When we review the dates that this dip occurred, we see this was the same period as the Great Recession 2008-2010.

Residuals represent the remaining data after removing the trend and seasonal components. The presence of a deep groove suggests unexplained variation during that specific time, potentially attributed to outliers or irregularities.

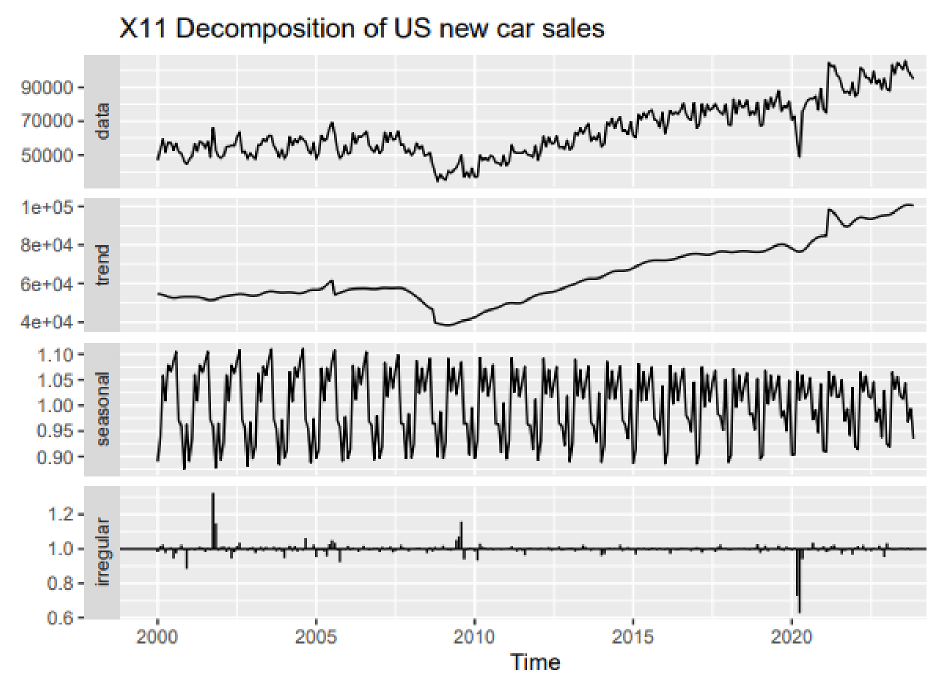
Throughout the entire time series, the magnitude of the seasonal pattern remains relatively constant.

The classical decomposition method might not completely capture the variation where the trend graph exhibits a bump, as indicated by the deep groove in the residuals.

So, now let's use the X11 decomposition to capture irregular patterns, outliers, and specific events that might not be adequately represented by the classical decomposition method.

X11 Decomposition:

(Figure 5)



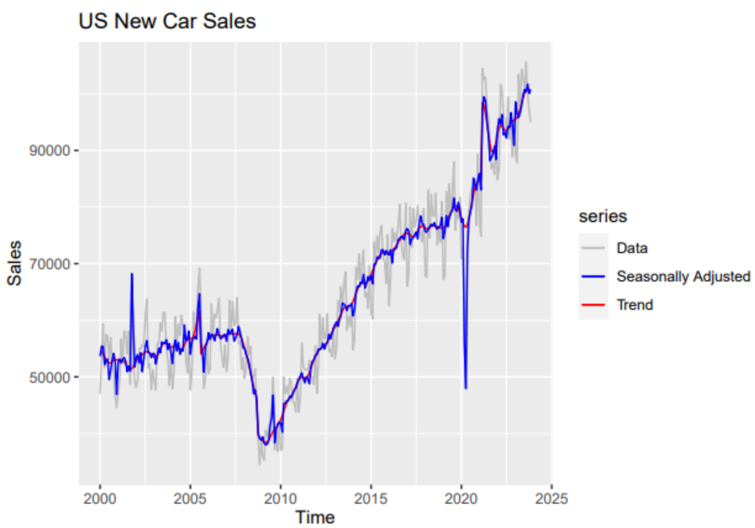
A major benefit to the X11 method is that, unlike classical decomposition, it does not drop the endpoints due to the moving average calculation. This leads to higher accuracy and time saved when plotting the adjusted model.

Through the X11 method, we are given the following coefficients representing the seasonal effects each month has on the data. From this we can deduce that, on average, October has the higher number of new car sales and April has the lowest.

```
Call:
seas(x = car_sales_ts)

Coefficients:
              Mon                Tue                Wed                Thu                Fri                Sat
         0.003420           0.001475           0.003591           0.003978           0.006316           0.008585
       AO2000.Dec         AO2001.Oct         AO2001.Nov         LS2005.Aug         LS2008.Oct         AO2009.Aug
        -0.124306           0.285719           0.135157          -0.144204          -0.147927           0.145188
       AO2020.Mar         AO2020.Apr         LS2021.Mar MA-Nonseasonal-01    MA-Seasonal-12
        -0.314736          -0.454703           0.153351           0.232627           0.683838
```

(Figure 6)



Conclusion

The X11 decomposition method offers several strengths for analyzing time series data, such as monthly car retail sales. However, it also has limitations that researchers should consider when interpreting the results.

Strengths of the X11 Decomposition Method:

1. Comprehensive Decomposition: X11 decomposition provides a comprehensive breakdown of time series data into its trend, seasonal, and irregular components. This allows researchers to better understand the underlying patterns and fluctuations in the data.
2. Flexibility: The method offers flexibility in adjusting for both additive and multiplicative seasonal effects, making it suitable for a wide range of time series datasets with different characteristics.
3. Established Methodology: X11 decomposition has been widely used and studied in various fields, providing a well-established framework for time series analysis.

Weaknesses of the X11 Decomposition Method:

1. Sensitivity to Outliers: X11 decomposition may be sensitive to outliers or extreme values in the data, leading to potentially biased decomposition results.
2. Complexity: The X11 decomposition method involves several steps and parameters that require careful consideration and expertise to implement correctly. This complexity may pose challenges for inexperienced users.
3. Assumption of Stationarity: X11 decomposition assumes that the underlying components of the time series (trend, seasonal, and irregular) are stationary over time. In practice, this assumption may not always hold true, particularly for dynamic or rapidly changing datasets.

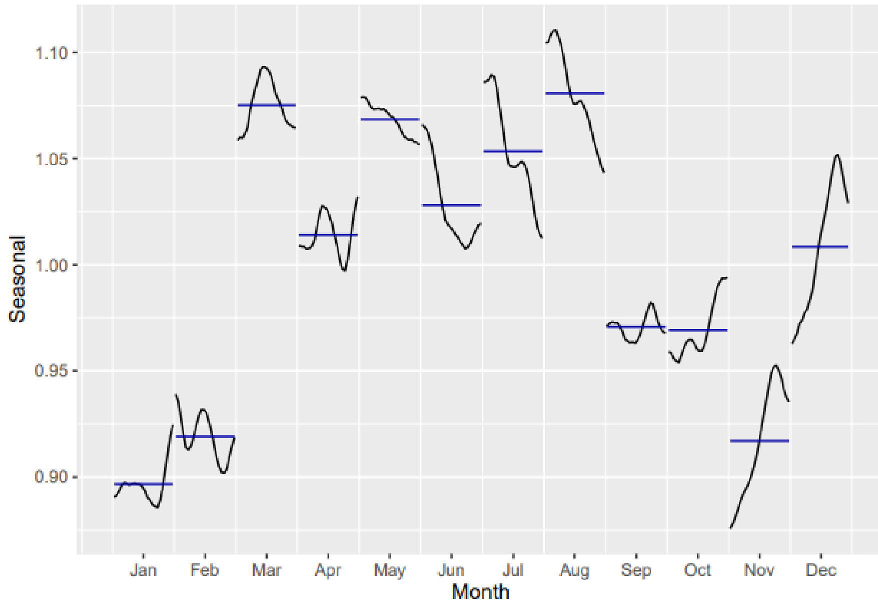
Insights Gained from Analysis of Monthly Car Retail Sales:

1. Seasonal Patterns (Figure 7): The decomposition reveals seasonal patterns in car sales, highlighting months or periods of the year when sales tend to peak or decline. This information can be valuable for sales forecasting and inventory management.
2. Long-Term Trends (Figure 8): The trend component of the decomposition provides insights into the overall growth or decline in car sales over time, allowing stakeholders to identify long-term market trends and make strategic business decisions. It also provides a good barometer for US economic health as a whole.
3. Irregular Fluctuations: By analyzing the irregular component, researchers can identify unexpected or irregular fluctuations in car sales that may be attributed to external factors such as economic conditions, marketing campaigns, or industry events.
4. Future Forecast (Figure 9): Forecasting the next 12 months of sales, we identified May as being the month that is projected to have the lowest car sales. As consumers we could target this month as the best time to buy a new car. Car retailers will understand the trends in their market and will likely also project May as being one of the slowest sales months of the year and likely enact a stimulus plan to mitigate this drop in sales. This could come in the form of promotional offers or deep discounts and could be a great time for consumers to buy to get the best rate of the year.

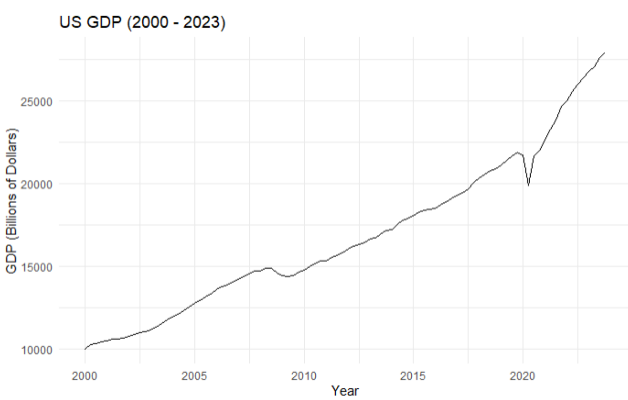
In conclusion, while the X11 decomposition method offers valuable insights into the underlying patterns and fluctuations in monthly car retail sales data, researchers should be mindful of its limitations and interpret the results with caution. By leveraging the strengths of the method and addressing its weaknesses, analysts can gain valuable insights to inform decision-making and strategic planning in the automotive industry.

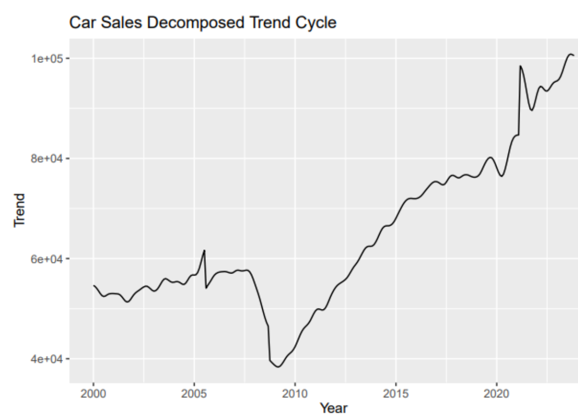
Appendix

(Figure 7)

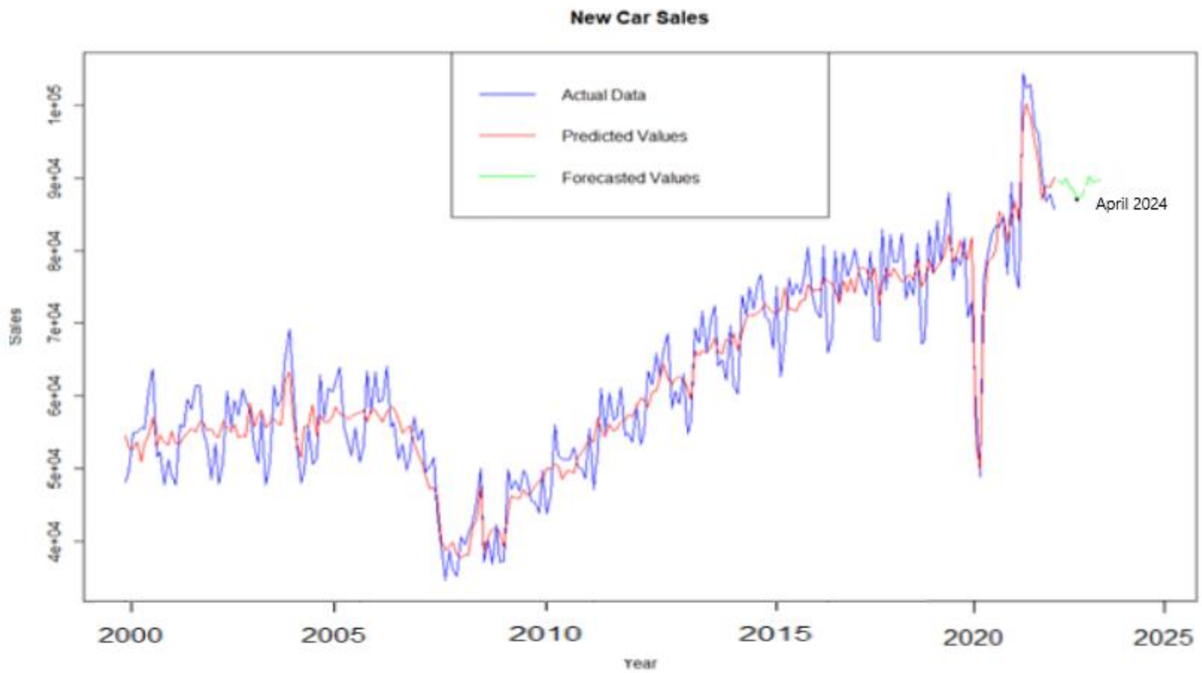


(Figure 8)





(Figure 9)



References

i. Deppa, B. (n.d.). Time Series Decomposition. Retrieved from https://course1.winona.edu/bdeppa/FIN%20335/Handouts/Time_Series_Decomposition.html, 2019

ii. Hyndman, Rob J., and George Athanasopoulos. "X11: Decomposition of Time Series Data." Forecasting: Principles and Practice, 2nd ed., 2020, https://otexts.com/fpp2/x11.html.

iii. Yaffee, Robert A. Time Series Analysis and Forecasting: With Applications of SAS and SPSS. Academic Press, 2000.

iv. Planas, Christophe. "The analysis of seasonality in economic statistics: a survey of recent developments". Qüestiió. 1998, vol.22, núm.1