Forecasting Demand for Electric Vehicles in the United States

Aavasna Rupakheti

July 2023

Contents

Introduction

Although the first successful electric car was introduced earlier than a car equipped with an internal combustion engine, their popularity has grown only over the last decade. As we can see in Figure 1, the market for electric vehicles (EVs) in the world has experienced rapid growth in recent years.

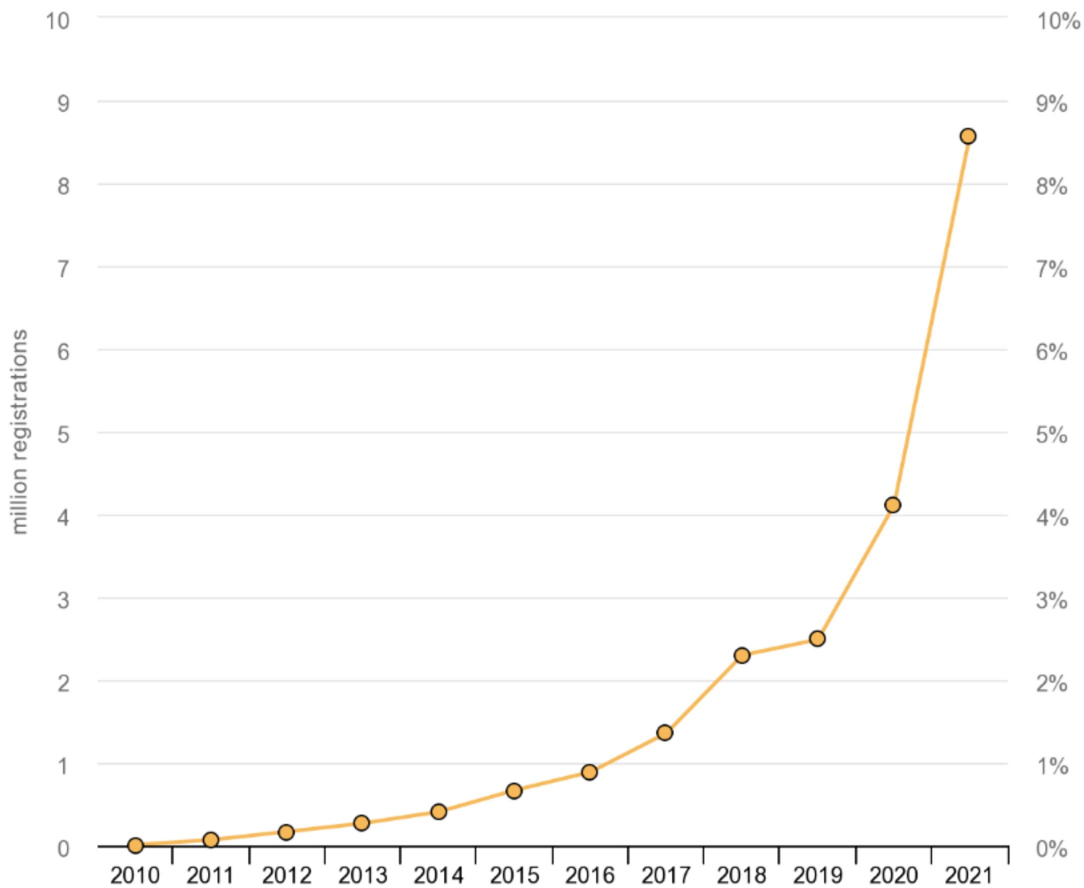


Figure 1: Figure 1: Global sales and sales market share of electric cars, 2010-2021, IEA, Paris.

Similar trend is seen in the United States where electric car sales have increased significantly, rising from 0.2 percent of total car sales in 2011 to 4.6 percent in 2021 (Colato & Ice, 2023). This growth can be attributed to heightened environmental concerns, evolving government policies, increased availability of EV models, improved cost competitiveness compared to conventional gas vehicles, advancements in vehicle range, technological advancements, and shifting consumer preferences. This rapid growth has also sparked considerable interest and debate regarding its future demand and implications for sustainable transportation. This project aims to address the key question of forecasting demand for electric vehicles, which holds great significance in shaping the transition to a greener and more sustainable future.

The importance of this research question is emphasized by the increasing global concerns over carbon emissions and the need to reduce reliance on fossil fuels. The transport sector, particularly gasoline-based vehicles, has emerged as a major contributor to carbon emissions. As such, understanding the factors driving EV demand and accurately forecasting its future trajectory is crucial for policymakers, industry stakeholders, and researchers alike.

To investigate this question, a combination of advanced forecasting techniques will be utilized. This includes time series analysis methods such as auto-regressive integrated moving average (ARIMA) models, and multivariate regression analysis. By employing these techniques, we can capture the historical patterns, trends, and underlying factors that influence electric vehicle demand.

The results of this research have the potential to provide valuable insights into the future demand for electric vehicles. By analyzing historical data and factors influencing EV adoption, we can identify key drivers of demand and project future sales figures. This information can guide industry stakeholders in their strategic planning, investment decisions, and product development, ensuring they align with market demand and consumer preferences.

Furthermore, the policy implications of our results are significant. Forecasting electric vehicle sales in the United States is of utmost importance due to the transformative potential of the EV market. Accurate demand forecasting can assist policymakers in formulating effective policies and incentives to promote EV adoption. It can inform decisions related to infrastructure development, charging station networks, and financial incentives to encourage consumers to switch to electric vehicles. Additionally, understanding the potential market size and growth patterns can aid in optimizing resource allocation, preventing overproduction or stockouts, setting ambitious yet attainable targets for reducing carbon emissions and achieving sustainable transportation goals.

Literature Review

Studies conducted previously on forecasting electric vehicles demand provide valuable insights into the drivers and barriers of electric vehicle adoption, demand forecasting, and market development strategies. This literature review synthesizes the key findings from these papers and highlights their contributions to the understanding of electric vehicle sales in other countries and how this might help in forecasting electric vehicle sales and demand in the United States.

Rezvani, Jansson, and Bodin (2015) analyze the drivers and barriers of consumer adoption of plug-in electric vehicles. Their study emphasizes the high cost of electric vehicles as a major barrier, while moral satisfaction and environmental awareness positively influence purchase decisions. The authors also identify pro-environmental attitudes, symbolic meanings, identity considerations, innovativeness, and emotions as key drivers of electric vehicle adoption. The paper highlights the importance of understanding these factors for policymakers and industry stakeholders to accelerate consumer electric vehicle adoption.

The study on forecasting the demand for electric cars in Morocco by *Chachdi et al. (2019)* focuses on understanding the determinants of future electric vehicles purchases in the country. The findings reveal that age, daily mileage, and ownership of non-electric cars significantly influence the demand for electric cars. The research provides valuable insights for decision-makers to stimulate the growth of the EV sector in Morocco and underscores the need for effective strategies and infrastructure development.

Kovárník and Staňková (2021) examined the significant factors influencing the sales of BEVs (Battery Electric Vehicles) in different European countries. They employ multiple regression models and vector auto-regressive models to capture the complete history of the electric car market. The findings reveal that while there is a common factor in the BEV market across countries, the identified factors vary significantly among individual countries. The study highlights the importance of factors such as sales of internal combustion cars, average wage, GDP per capita, battery price, oil prices, brand perception, and psychological factors. The paper concludes that considering both country-specific and common factors is crucial for understanding the electric vehicle market.

Sang and Bekhet (2015) delve into the factors shaping the development of the electric vehicle market in Malaysia, shedding light on a country where EV adoption is still in its early stages. Through a survey of 1000 private vehicle drivers, the study identifies significant factors such as social influences, performance attributes, financial benefits, environmental concerns, demographics, infrastructure readiness, and government interventions that impact electric vehicle adoption in Malaysia. The study provides insights for forming marketing strategies in the Malaysian electric car market based on these identified factors. The research contributes to the understanding of electric vehicle technology diffusion and highlights the importance of public awareness, government support, and tailored marketing approaches for electric vehicle adoption.

Lastly, *Gnann, et. all (2015)* presents a comprehensive analysis of the market evolution of plug-in electric vehicles (PEVs) in Germany until 2020. They highlight the importance of energy prices, state compensation for electric car prices, and flexible credit policies as crucial factors influencing the market's deployment. The authors emphasize the high uncertainty of market evolution and recommend that state policies be dynamically adapted to respond to changing market conditions. Monetary policy options, such as special depreciation allowances and purchase subsidies, are identified as effective measures to stimulate PEV adoption, particularly in commercial fleets. The research also emphasizes the importance of considering non-monetary factors, such as users' willingness to pay more for innovative technology. Overall, this study provides valuable insights into the complexities of PEV market diffusion and emphasizes the significance of policy measures in promoting sustainable transportation in Germany.

Based on the analysis of publications, it is evident that the world's electric car market deployment is influenced by several key determinants, encompassing technical aspects such as the improvement of vehicle range and efficiency, economic factors including the cost of the vehicle and availability of charging stations, as well as environmental considerations such as consumers' environmental awareness and opportunities to reduce pollution.

In conclusion, these papers collectively contribute to the understanding of the factors influencing EV adoption and sales and demand forecasting. The studies emphasize the role of cost, environmental awareness, psychological factors, policy measures, and demographic characteristics in shaping consumer behavior towards electric vehicles affecting demand. The findings provide valuable insights for policymakers, industry stakeholders, and researchers seeking to promote sustainable transportation systems and facilitate the successful integration of electric vehicles into national markets.

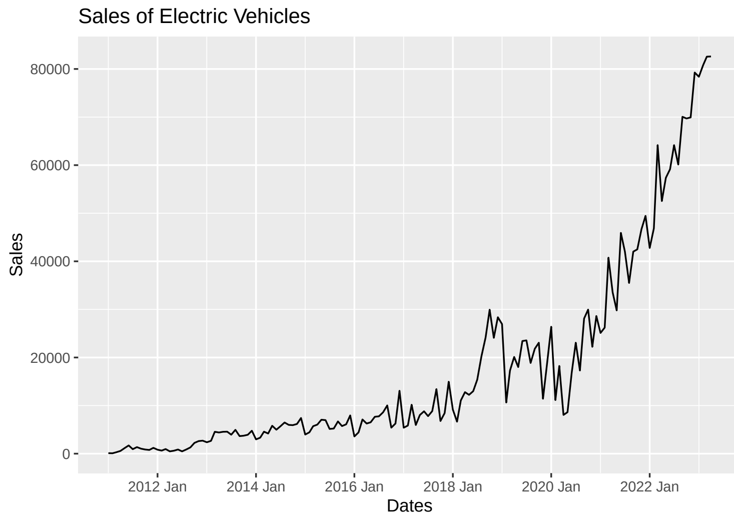
Data and Time Series Characteristics

The **main series of interest** is the sale of electric vehicles per month from 2011 to April 2023. The data is the total number of sales in the United States and is in monthly frequency. The data source is: https://www.anl.gov/esia/reference/light-duty-electric-drive-vehicles-monthly-sales-updates-historical-data

We are only interested in all-electric vehicles which are known as Battery Electric Vehicles (denoted by BEV in the data set) and not hybrid electric vehicles.

Trend in Sales of Electric Vehicles

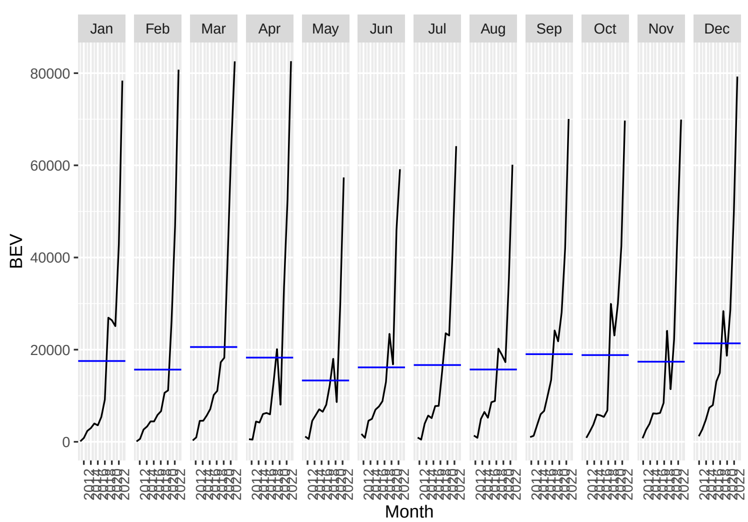
```
ev_ts |> autoplot(BEV) +
  labs(title = "Sales of Electric Vehicles",
       x = "Dates",
       y = "Sales")
```



As we can see in the figure above, the trend for sales for electric vehicles is positive. The series up to the year 2017 is not that interesting because the trend is not that strong and it does not have much seasonality. After 2017, we can see a big overall rise in sales for electric vehicles. It does show that sales went down a little in Feb-April in 2020 which makes sense because of the pandemic.

Seasonality in Sales of Electric Vehicles

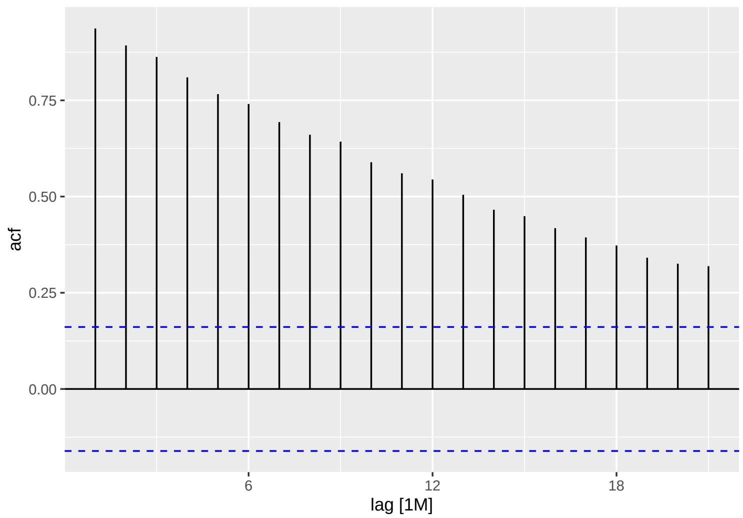
```
# a discussion of seasonality
ev_ts |> gg_subseries(BEV)
```



We can look at the figure above to examine the seasonality in the sales of electric vehicles. The seasonality in sales of EV is similar to the sales of other automobiles. The colder winter months often have slower sales, as consumers are less motivated to go out when the temperatures drop. We can see that sales start to see a spike in march when the weather warms up, perhaps coupled with the arrival of tax returns making people more apt to spend money on a car.

Autocorrelation Plot for Sales of Electric Vehicles

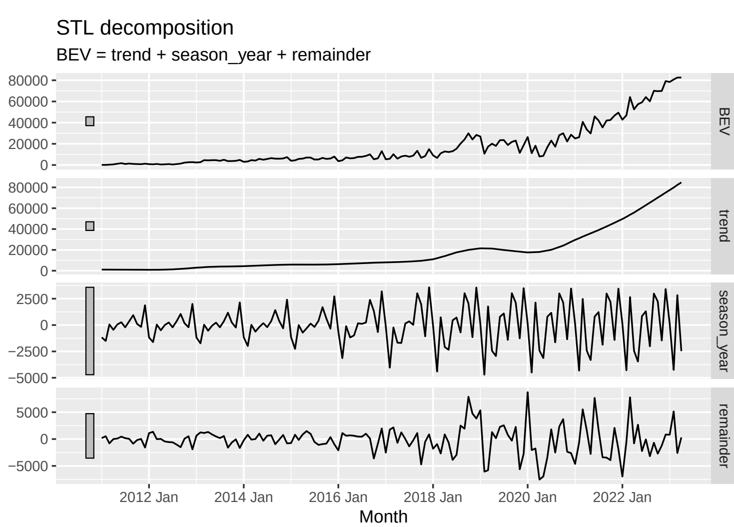
```
ev_ts |>
  ACF(BEV) |>
  autoplot()
```



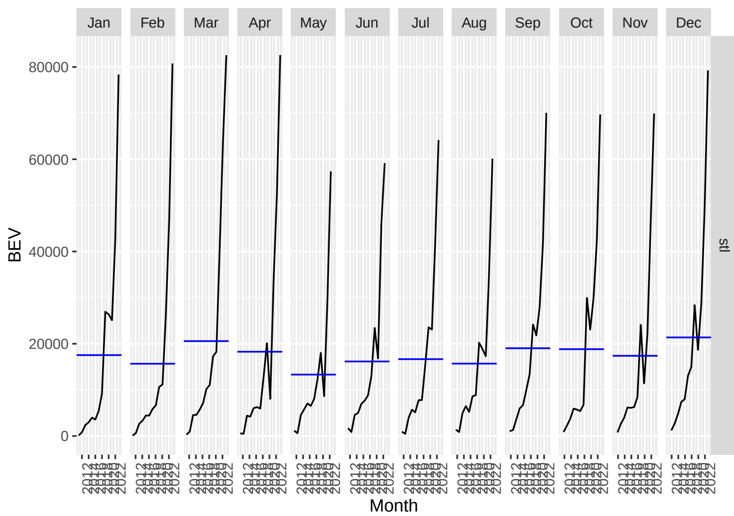
The above correlogram substantiates the assertions I made earlier, that the sales of electric vehicles has a strong trend. When data has a trend, the auto-correlations for small lags tend to be large and positive. When data is seasonal, the auto correlations will show some peaks, which is also true in the above figure.

Time Series Decomposition for Electric Vehicle Sales

```
dcmp <- ev_ts %>%
  model(stl = STL(BEV))
components(dcmp) %>% autoplot()
```



```
components(dcmp) %>% gg_subseries(BEV)
```

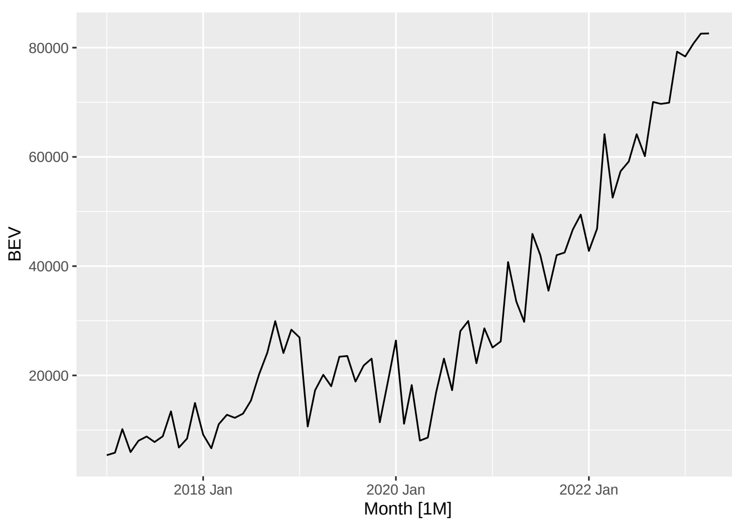


This decomposition shows that the variation in seasonality increases with time. But the seasonality might be affected by the lower sales at the lower end of the data set so we are going to focus on sales from 2017.

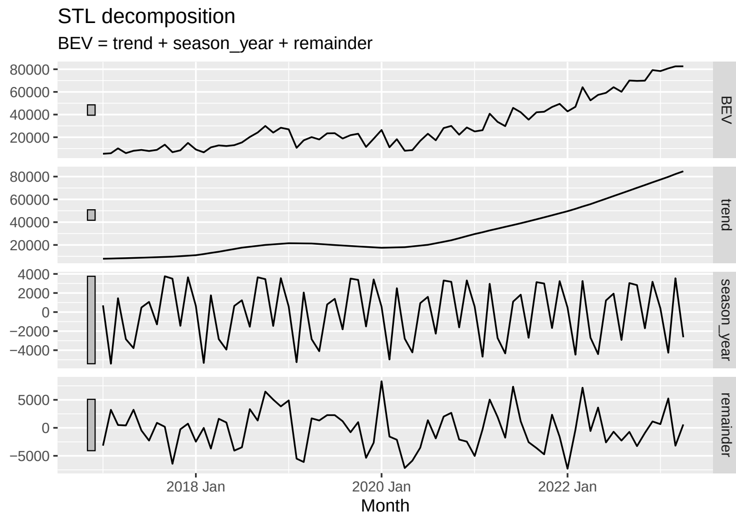
Time Series Decomposition for sales after 2017

In order to understand the trend for the sales of electric vehicles more closely, we are going to look at data after 2017.

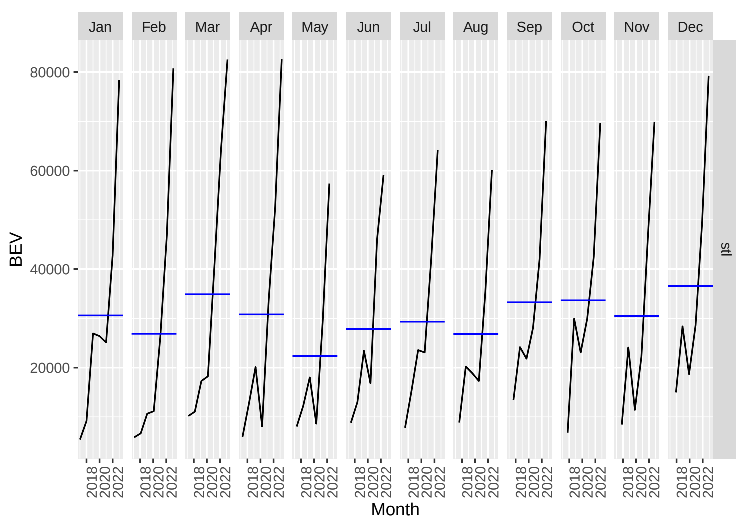
```
ev_ts_2017 |> autoplot(BEV)
```




```
dcmp2 <- ev_ts_2017 %>%
  model(stl = STL(BEV))
components(dcmp2) %>% autoplot()
```




```
components(dcmp2) %>% gg_subseries(BEV)
```

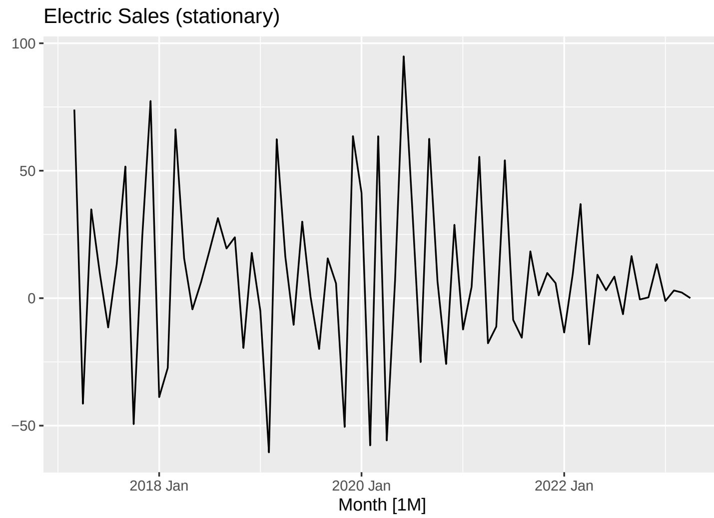


By selecting this period, we can see that there is pretty strong seasonality in the data.

Dealing with seasonality in the series

We can see if getting the percentage change in sales can make the data set stationary.

```
ev_ts_2017 %>%
  autoplot(sales) +
  labs(y="", title="Electric Sales (stationary)")
```



The data definitely does not have a strong trend anymore. We can do a formal test of seasonality just to verify what we see in the above graph.

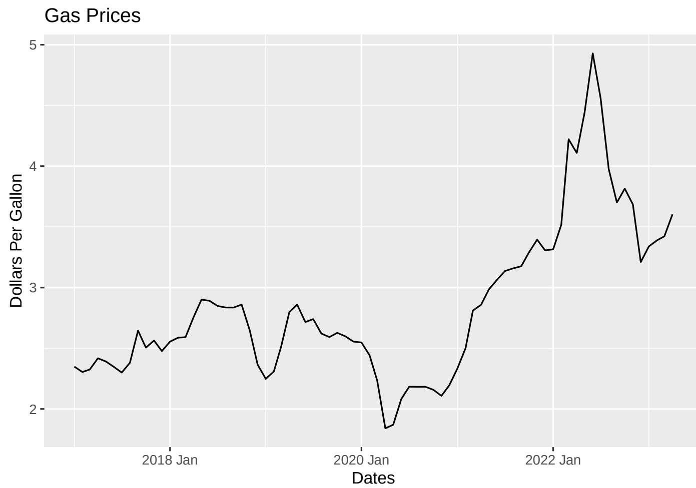
```
ev_ts_2017 %>%
  features(sales, unitroot_kpss)
## # A tibble: 1 x 2
##   kpss_stat kpss_pvalue
##       <dbl>       <dbl>
## 1     0.151         0.1
```

Since the p-value > 0.05: we can say that the transformed sales data is stationary.

Independent variables that might help obtain better forecasts include:

- *Gas prices:* Higher gas prices can positively influence the demand for electric vehicles. As gas prices increase, consumers may seek alternatives to traditional gasoline-powered vehicles, and electric vehicles become more attractive due to their potential for lower operating costs. Electric vehicles offer the advantage of utilizing electricity, which generally costs less per mile than gasoline, making them an appealing option for cost-conscious consumers. The data is originally weekly but converted to a monthly frequency using average. The unit is Dollars per Gallon and the source is FRED: https://fred.stlouisfed.org/series/GASREGW#

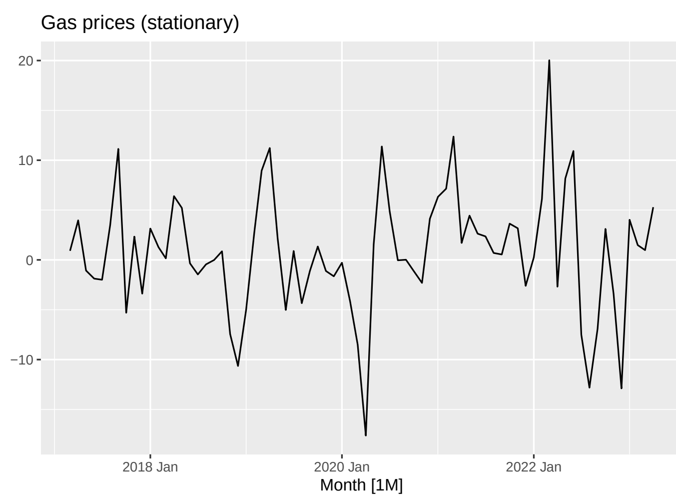
  ```
  gas_prices |> autoplot(price) +
    labs(title = "Gas Prices",
         x = "Dates",
         y = "Dollars Per Gallon")
  ```



We obtain stationary gas prices by taking the percentage change in price. This was checked using the KPSS test as well.

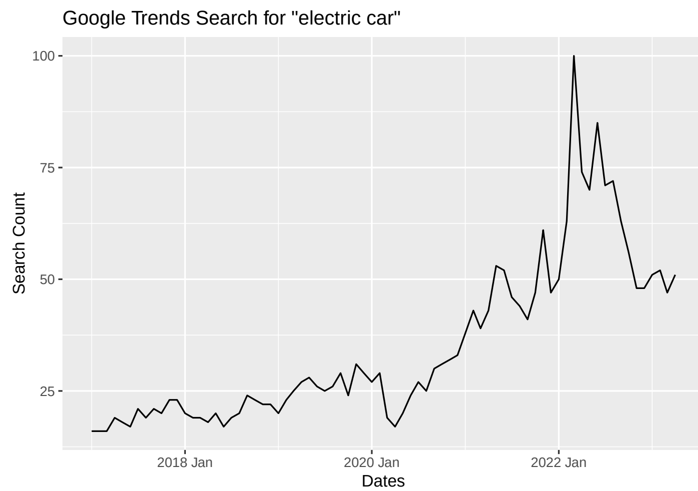
```
gas_prices <- gas_prices %>%
  mutate(price_change = 100 * (price / lag(price) - 1))
gas_prices <-gas_prices %>%
  slice(-c(1,2))

gas_prices%>%
  autoplot(price_change) +
  labs(y="", title="Gas prices (stationary)")
```



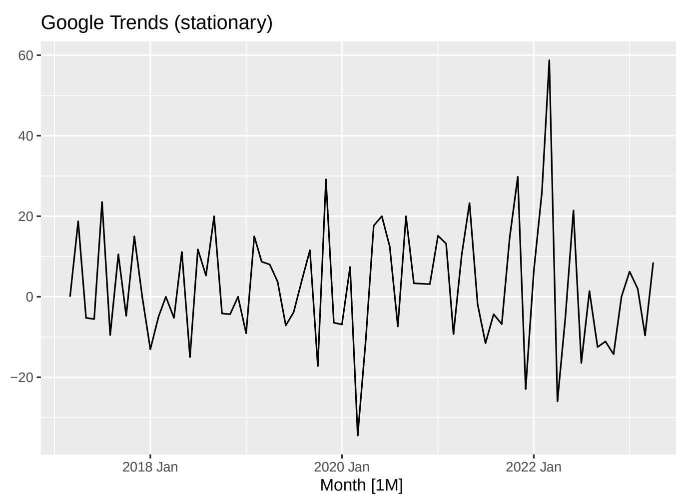
- *Google trend data:* Google trend data provides insights into the popularity and interest in electric vehicles. Increased search volume for terms related to electric vehicles on Google can be indicative of growing consumer curiosity and awareness. Higher Google trend data suggests that more people are actively seeking information about electric vehicles, which can translate into increased sales and demand. It reflects evolving consumer preferences and serves as an early indicator of potential shifts in the market. The data is for the term "electric car" for the same time period and frequency as the variable of interest. Similar results were seen for the search term (EV). The source is: https://trends.google.com/trends/explore?date=2011-01-01%202023-04-30&geo=US&q=electric%20car&hl=en

```r
google_trend |> autoplot(search) +
  labs(title = 'Google Trends Search for "electric car"',
       x = "Dates",
       y = "Search Count")
```



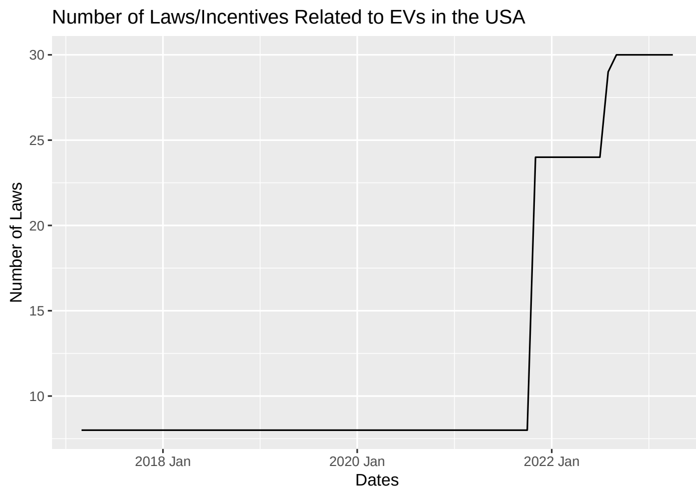
Here, we obtain the stationary series by transforming it into a percentage change as well. This was also confirmed using the KPSS test.

```r
google_trend%>%
  autoplot(trend) +
  labs(y="", title="Google Trends (stationary)")
```



- *Laws and policies:* Favorable laws and policies can significantly impact the sales and demand for electric vehicles. Government initiatives such as tax credits, rebates, or subsidies for purchasing electric vehicles incentivize consumers to choose electric over conventional vehicles. Additionally, regulations promoting clean transportation or stricter emission standards can create a conducive environment for electric vehicle adoption, driving up sales as consumers are encouraged to make environmentally friendly choices. The data for laws implemented on the federal level needs to be filtered from the following source: https://afdc.energy.gov/data/10360

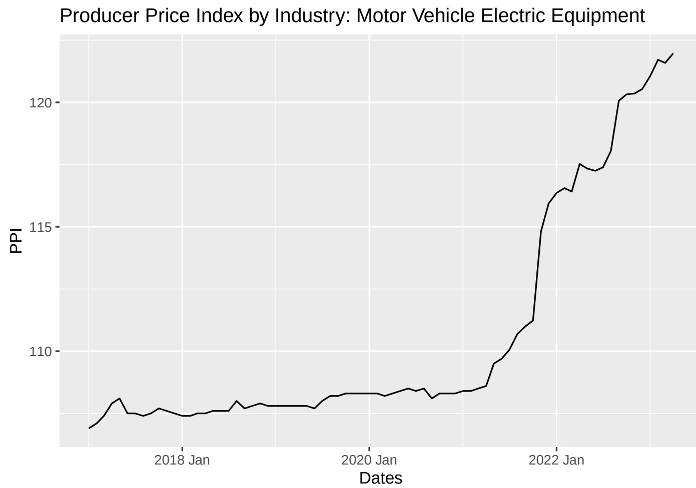
```r
laws |> autoplot(law_count) +
  labs(title = 'Number of Laws/Incentives Related to EVs in the USA ',
       x = "Dates",
       y = "Number of Laws")
```



As we can see, there was a sharp spike in the number of laws and incentives related to electric vehicles in the US around November, 2021 which makes sense because the Build Back Better Act was passed that month. This act included tax incentives and investment to spur consumer demand in electric vehicles. In order to reflect this change better, we are going to convert this variable into a dummy variable before and after the Build Back Better Act to reflect effect of policy changes.

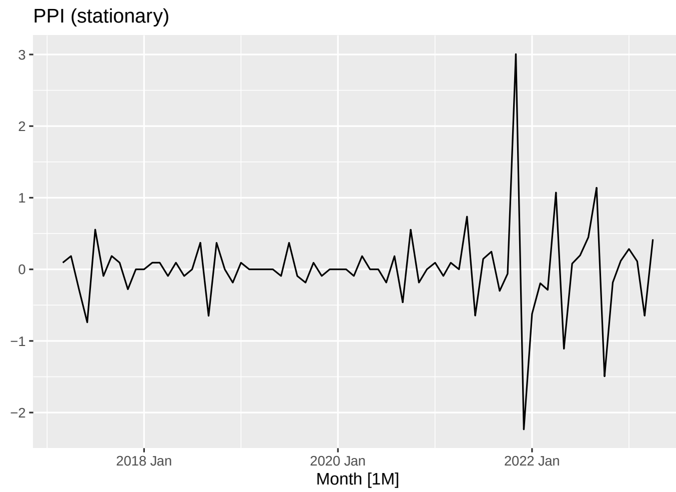
- *Producer Price Index (PPI)*: The PPI measures changes in the selling prices received by producers. For electric vehicles, the PPI for Motor Vehicle Electric Equipment focuses on pricing trends in the industry's supply chain. It provides insights into production costs and pricing dynamics. Monitoring the PPI helps forecasters understand potential impacts on sales. Rising PPI suggests higher costs, which may affect pricing and sales volumes. Declining PPI indicates cost reductions or efficiency improvements, potentially leading to competitive pricing and increased sales. The base year is Index Dec 2003=100 and the source is FRED: https://fred.stlouisfed.org/series/PCU3363233632

```r
PPI |> autoplot(price) +
  labs(title = 'Producer Price Index by Industry: Motor Vehicle Electric Equipment ',
       x = "Dates",
       y = "PPI")
```



Here, just the percentage transformation did not make the series stationary, so, we had to take the first difference as well.

```
PPI <- PPI %>%
  mutate(ppi_change = 100 * (price / lag(price) - 1))
PPI <- PPI |>
  mutate( ppi_change = difference(ppi_change))
PPI <- PPI %>%
  slice(-c(1,2))
PPI%>%
  autoplot(ppi_change) +
  labs(y="", title="PPI (stationary)")
```



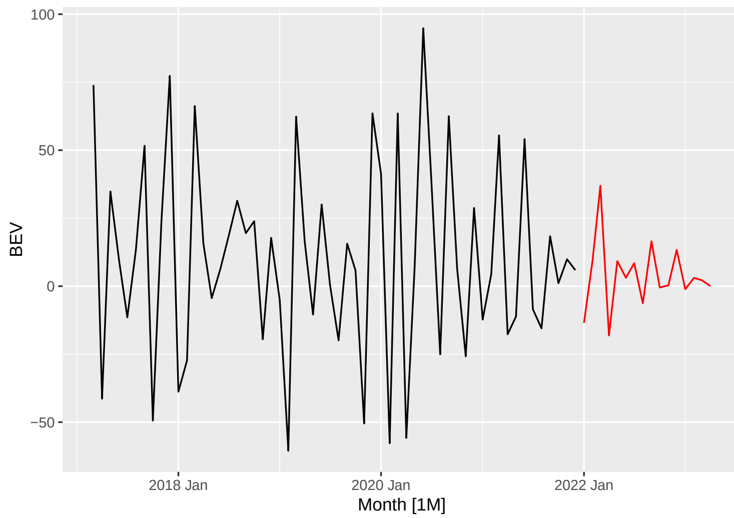
Here, we can still see that the variation is obviously higher after 2021. But none of the transformations really helped stabilize that and make the entire data stationary, so this is the best that was obtained.

Application of Forecasting Methods

First, we are going to split our series into training and testing sets. Here, we have selected the period from January 2017 - December 2021 as the training data and January 2022 - April 2023 as the test data.

```r
autoplot(train, BEV) +
  autolayer(test,BEV,color= "red")
```

Plotting train and test data



Comparing Models

After comparing several seasonal decomposition models, auto-arima models, arima regression models and specified arima models, we were left with the following 9 models.

```r
fit_sales %>% pivot_longer(everything(), names_to = "Model name",
                     values_to = "Specification")
## # A mable: 9 x 2
## # Key:     Model name [9]
##   `Model name`                          Specification
##   <chr>                                       <model>
## 1 model1                      <STL decomposition model>
## 2 model2                      <STL decomposition model>
## 3 model3                      <STL decomposition model>
## 4 model4       <LM w/ ARIMA(0,0,1)(1,0,0)[12] errors>
## 5 model5                         <ARIMA(2,0,1) w/ mean>
## 6 model6                         <ARIMA(2,0,2) w/ mean>
## 7 model7                    <LM w/ ARIMA(2,0,2) errors>
## 8 model8       <LM w/ ARIMA(2,0,2)(1,0,0)[12] errors>
## 9 model9                    <LM w/ ARIMA(2,0,2) errors>
```

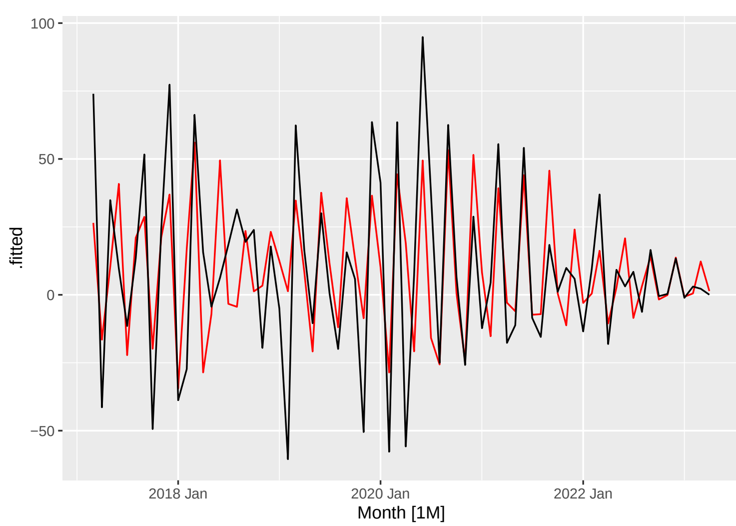
```r
fc_sales <- fit_sales %>% forecast(test)
accuracy(fc_sales, test)
## # A tibble: 9 x 10
##   .model .type      ME  RMSE   MAE     MPE  MAPE  MASE RMSSE    ACF1
##   <chr>  <chr>   <dbl> <dbl> <dbl>   <dbl> <dbl> <dbl> <dbl>   <dbl>
## 1 model1 Test    -4.49  30.2  23.5   1778.  2354.   NaN   NaN -0.423
## 2 model2 Test     1.57  31.2  22.9   1191.  1826.   NaN   NaN -0.383
```

```
## 3 model3 Test   3.28   30.6  22.7   1669.  2303.   NaN   NaN -0.450
## 4 model4 Test  -4.97   16.7  14.6  -1362.  1986.   NaN   NaN -0.0726
## 5 model5 Test  -6.20   14.0  11.2  -1379.  1861.   NaN   NaN -0.442
## 6 model6 Test  -5.86   11.6   9.54 -1065.  1434.   NaN   NaN -0.501
## 7 model7 Test   0.0852  9.41  7.13   -61.3  191.   NaN   NaN -0.514
## 8 model8 Test   1.44   17.9  14.8   -998.  1991.   NaN   NaN -0.248
## 9 model9 Test  -5.39   12.9  10.4  -1621.  2058.   NaN   NaN -0.356
```

As we can see in the table, model7 performs the best in terms of RMSE scores. RMSE is one of the metrics used to evaluate the accuracy of a model where the smaller the RMSE score, the better. Model 7 was a Multivariable Linear Regression Model with ARIMA (2,0,2) errors. The only predictor used here was 'laws'. The next best model here was model6 which was ARIMA(2,0,2). The third best model was a Multivariate Regression Model with 'Google trend' as the predictor and errors modeled using ARIMA(2,0,2).

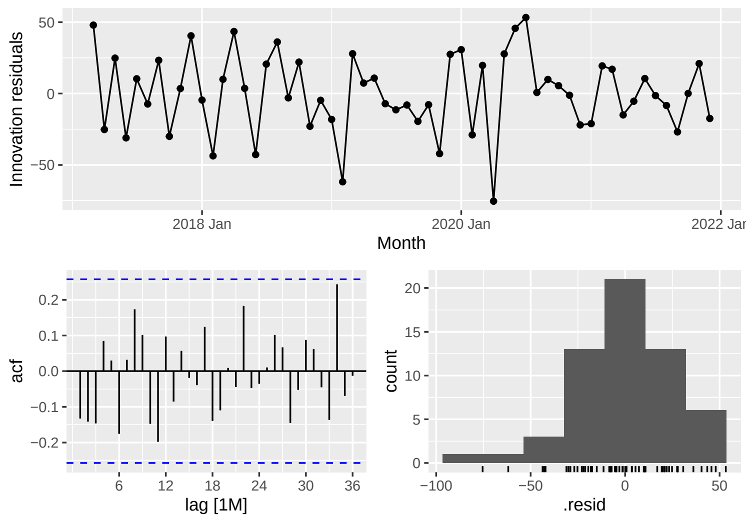
Fitted values for Model 7

```
fit <- dataset |>
  model(model7 = ARIMA(BEV ~ pdq(2,0,2) + laws))
fitted(fit) %>%
  autoplot(.fitted, color = "red") +
  autolayer(dataset, BEV)
```



Residual Analysis

```
fit_sales %>% select(model7) %>% gg_tsresiduals(lag=36)
```



- assumption 1: the correlogram shows that all autocorrelation coefficients are within the limits, where one is almost reaching the limit
- assumption 2: residuals are mostly fluctuating around zero so probably et terms have mean zero
- assumption 3: from the residual plot, it seems like the variation is pretty constant except for March 2020, which is expected because of the pandemic and this probably needs special attention
- assumption 4: The residuals are not normally distributed.

We can also perform a formal test of autocorrelation using the **Ljung-Box test:**

```
augment(fit_sales) %>% features(.innov, ljung_box, lag=24, dof=5)
## # A tibble: 9 x 3
##   .model lb_stat lb_pvalue
##   <chr>    <dbl>     <dbl>
## 1 model1    28.8   0.0695
## 2 model2    33.1   0.0234
## 3 model3    36.4   0.00952
## 4 model4    39.3   0.00407
## 5 model5    18.4   0.499
## 6 model6    22.9   0.242
## 7 model7    23.2   0.229
## 8 model8    25.5   0.145
## 9 model9    19.6   0.419
```

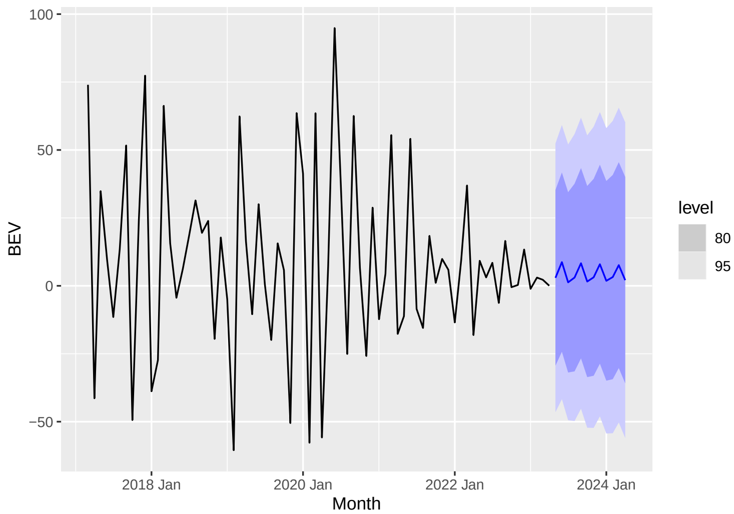
As a decision rule, if lb_pvalue is above 0.05 we can say that residuals are white noise, which is the case here for model 7 (**lb__pvalue** = 0.229168).

Obtaining Forecasts

Now, we can obtain forecasts for the next 12 months using Model 7.

```
fit <- dataset |>
  model(ARIMA(BEV ~ pdq(2, 0, 2) + laws))

forecast(fit, future_values) |>
  autoplot(dataset)
```



Conclusion

In conclusion, the forecasting model utilized in this project has demonstrated its ability to accurately capture the seasonality of electric vehicle sales data. The results obtained provide valuable insights into the trends and patterns of the market, enabling informed decision-making and planning for various stakeholders.

While the model performs moderately well, there is a lot of room for improvement. One potential avenue for enhancing the model's predictive power lies in identifying additional important predictors. The inclusion of such variables could potentially address the limitations observed in other models that utilized predictors, which yielded less accurate results. Notably, factors such as battery prices and real electricity prices, which directly impact operational costs for customers, have the potential to significantly influence electric vehicle demand. Furthermore, the availability and growth of EV charging stations, for which monthly data was not accessible in this study, could also play a crucial role in shaping consumer behavior and adoption rates.

The contribution of this forecasting project is significant, as it provides valuable insights into the seasonality of electric vehicle sales. By understanding and anticipating these patterns, policymakers, manufacturers, and investors can make informed decisions regarding production, marketing, and infrastructure development. Moreover, this research contributes to the broader discussion on sustainable transportation and its potential to mitigate climate change.

However, it is important to acknowledge the limitations of this work. While the model captures seasonality well, there are likely other factors beyond the scope of this study that affect electric vehicle demand. Future research should aim to explore these factors and incorporate them into the forecasting model to further enhance its accuracy and applicability. Additionally, the availability of more data, such as real-time battery prices, growth of charging infrastructure and changing customer sentiments would undoubtedly contribute to a more comprehensive understanding of the market dynamics.

In conclusion, this project provides valuable insights into electric vehicle sales forecasting, highlighting the importance of understanding seasonality and the potential for further improvement. The findings and recommendations presented here serve as a foundation for future research, contributing to the ongoing efforts in advancing sustainable transportation and facilitating the transition to electric mobility.

Citations

Chachdi, Aziza, Bouchra Rahmouni, Jamal Zahi, and Badr Ikken. "Forecasting the Demand for Electric Cars in Morocco." International Journal of Smart Grid and Clean Energy (2019): 191-200.https://doi.org/10.12720/sgce.8.2.191-200.

Gnann, Till, Patrick Plötz, André Kühn, and Martin Wietschel. "Modelling Market Diffusion of Electric Vehicles with Real World Driving Data – German Market and Policy Options." Transportation Research Part A: Policy and Practice 77 (July 2015): 95–112. https://doi.org/10.1016/j.tra.2015.04.001.

Javier Colato and Lindsey Ice. "Charging into the future: the transition to electric vehicles." Beyond the Numbers: Employment & Unemployment, vol. 12, no. 4 , U.S. Bureau of Labor Statistics (February 2023). https://www.bls.gov/opub/btn/volume-12/charging-into-the-future-the-transition-to-electric-vehicles.htm

Kovárník, Richard, and Michaela Staňková. "Determinants of Electric Car Sales in Europe." LOGI – Scientific Journal on Transport and Logistics 12, no. 1 (2021): 214–25. https://doi.org/10.2478/logi-2021-0020.

Rezvani, Zeinab, Johan Jansson, and Jan Bodin. "Advances in Consumer Electric Vehicle Adoption Research: A Review and Research Agenda." Transportation Research Part D: Transport and Environment 34 (January 2015): 122–36. https://doi.org/10.1016/j.trd.2014.10.010.

Sang, Yew-Ngin, and Bekhet, Hussain. "Exploring Factors Influencing Electric Vehicle Usage Intention: An Empirical Study in Malaysia." International Journal of Business and Society 16 (2015): 57-74.https://doi.org/10.33736/ijbs.554.2015.