**Port Authority of New York and New Jersey.**

**Forecasting Buses and Passengers.**

**Executive Summary**

1. **Determine when the weekday departures will be past 125,000 passenger departures and 3,900 bus departures.**

To forecast "The Port Authority of New York and New Jersey" the number of buses and passengers between December 2020 and September 2023. We gathered data on population, gas prices, temperature, and climate. We consolidated the data into an Excel document and then imported it into Power BI for analysis. As part of the project, data was consolidated and forecasting was done to obtain insights. To forecast the data, we used a trend line. Based on rigorous analysis utilizing predictive modeling, our projections indicate that passenger numbers are anticipated to surpass 125,000 on 06/23/2025 and bus departures to reach 3,900 on 04/30/2029. Our forecast is grounded in a meticulous analysis facilitated through Power BI’s line chart functionality.

1. **Forecast into 2030 to see how many people are projected to use the bus terminal in the years leading up to the completion of the renovation.**

Specific questions mention future projections like “By 2030, we can see that 3900 buses depart, and by 2026, we can see that the passenger departure count will cross 125,000.” It seems these are projections derived from the FORECAST.ETS function applied to our historical data. The exact details of how these specific forecasts are calculated would depend on the data and the parameters used in our FORECAST.ETS function call. We used exponential smoothing with the help of Excel. Based on the datasets, we used the FORECAST.ETS function in Microsoft Excel to perform time-series forecasting. The FORECAST.ETS function in Excel utilizes historical time series data (x-values and corresponding y-values) to predict future values based on the specified algorithm. The output of this function at a given time point will provide an estimated y-value (forecasted value) corresponding to that x-value.

1. **Develop three forecasting models, train them, test them to see which one works best.**

The time series forecasting technique in Power BI serves as the foundation for the forecasting model that we used in our analysis. In particular, we made use of Power BI’s built-in forecasting features for line charts, which use exponential smoothing and ARIMA (Autoregressive Integrated Moving Average) algorithms to estimate future values based on past trends. This model was selected because of its efficacy, readability, and simplicity when applied to time series data. An appropriate basis for capturing trends and patterns throughout time is provided by the time series technique, given the nature of the forecasting task and the availability of past data. Furthermore, Power BI’s integrated design makes it easier to explore and visualize data in a way that meets stakeholders’ practical demands.

**Data integration work**

In the process of data integration, we meticulously combined and processed diverse datasets to ensure a comprehensive and accurate representation of the information needed for forecasting. This involved cleaning, transforming, and merging data from the source. Our approach prioritized data quality, consistency, and relevance, ensuring that the integrated data set was robust and ready for analysis.

1. The forecasts were produced by utilizing patterns and trends seen in previous data. Based on the observed historical patterns, we projected future trends using Power BI’s line charting feature. In order to inform our forecasts, this required locating trends, seasonality, and any other recurrent patterns in the previous data. The predictive models that Power BI employed made it easier to anticipate future values and offered insights into possible outcomes.
2. We added a few more elements to our data model in order to better understand the behavior of this data. These variables could affect passenger counts and bus departures, and they could include things like the climate, gas prices, and holidays. This new data came from a variety of sources, including industry publications, company records, and other databases. The integration of these individual elements facilitated a more complex and comprehensive analysis.

**Models and Tools Used:**

In "The Port Authority of New York and New Jersey" forecasting of bus and passengers’ departures project we have used the three methods like ARIMA (Autoregressive Integrated Moving Average) model, Microsoft Power BI and Exponential Smoothing. For ARIMA model we used R programming and for exponential smoothing we used Excel.

* ARIMA models are widely used for time series forecasting due to their ability to capture

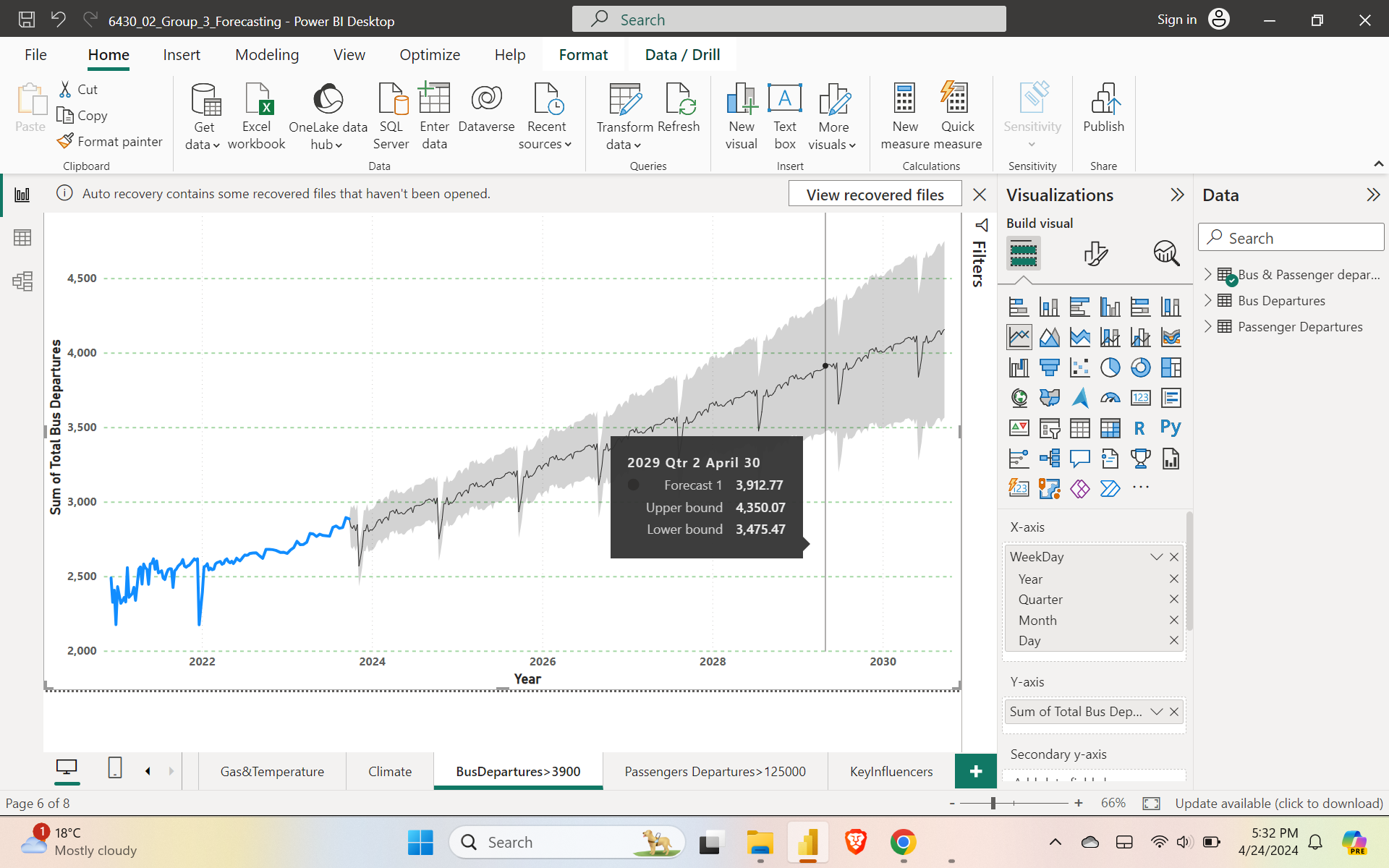
both autoregressive and moving average components of the data. Autoregressive (AR) component, captures the relationship between an observation and a number of lagged observations. For example, if the number of passengers or bus departures today is related to the number of passengers or departures on previous days. Integrated (I) component, represents differencing, which is the process of removing trend and seasonality from the data to make it stationary. Stationarity is important because many time series models, including ARIMA, assume that the data is stationary. Moving Average (MA) component, captures the dependency between an observation and a residual error from a moving average model applied to lagged observations.

* Exponential smoothing method is simple and intuitive approaches to time series forecasting. They involve updating a forecast as new observations become available, with more recent observations receiving higher weights. Holt-Winters' method is an extension of exponential smoothing that includes components for level, trend, and seasonality. By assigning higher weights to recent data, exponential smoothing puts more emphasis on capturing the latest trends and patterns.

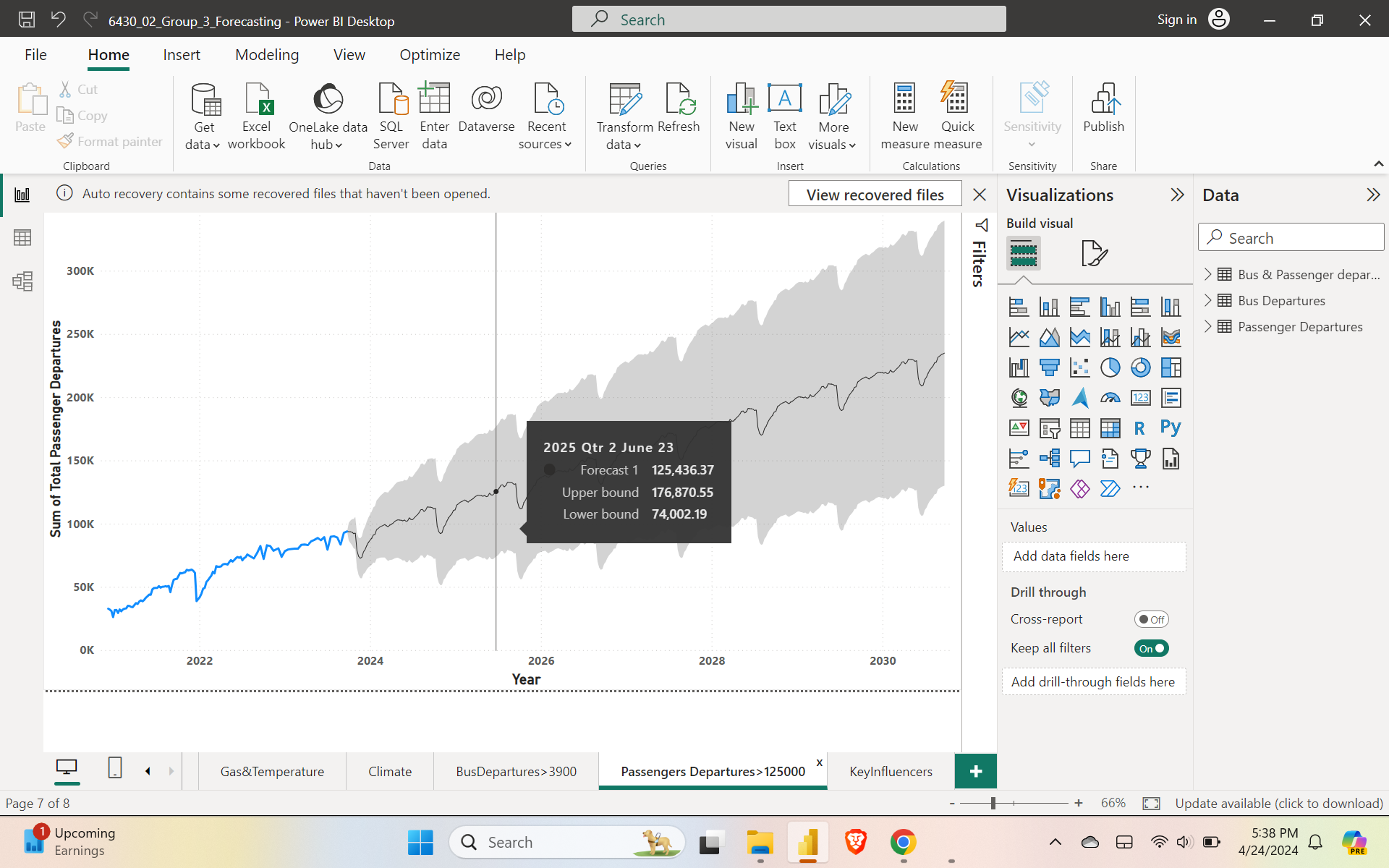
**Output Visualizations:**

**Microsoft Power BI Outputs**

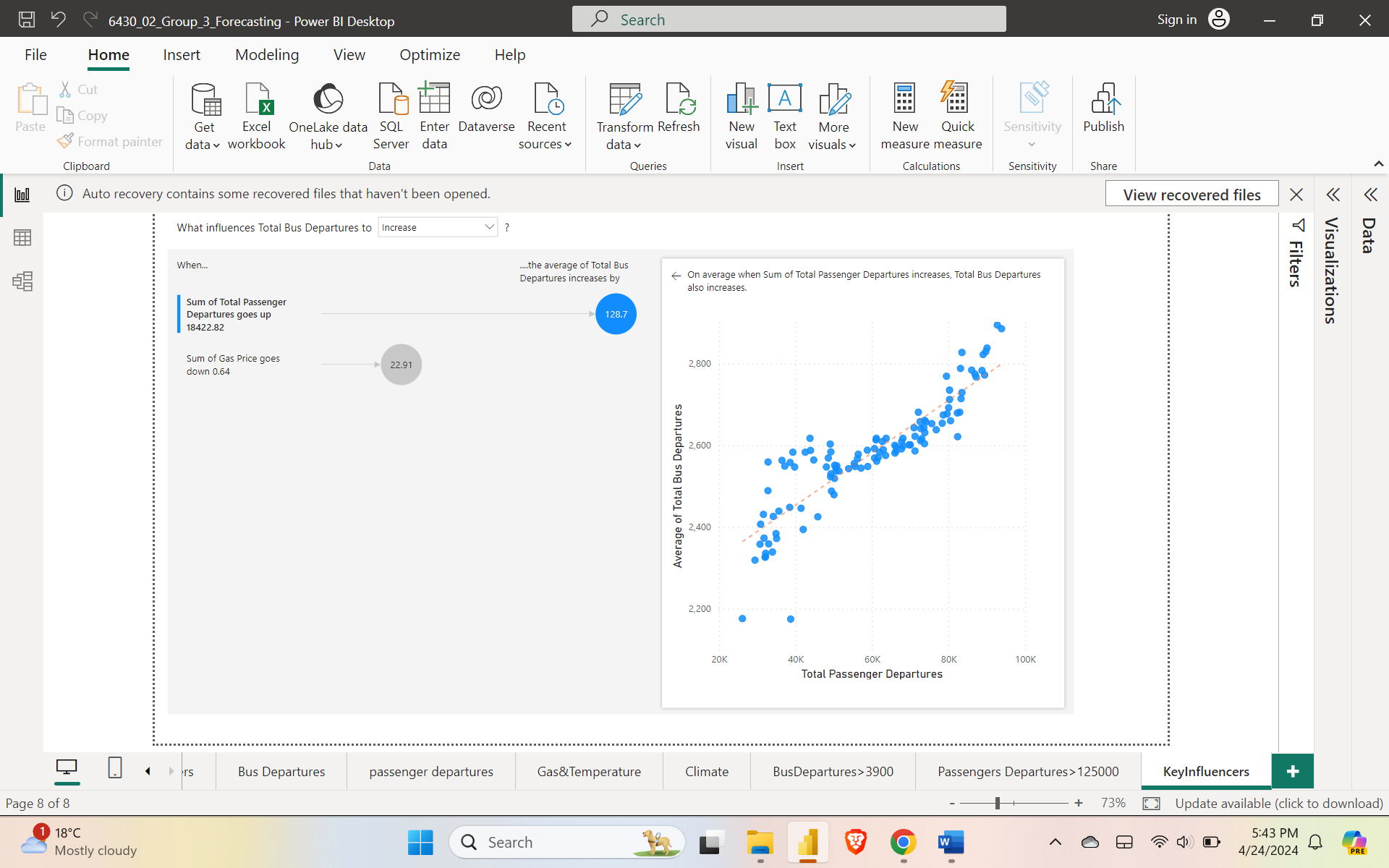
Here we are providing the forecast visuals of the bus and passenger departures as per the requirement of the business need using Microsoft Power BI tool. Visual\_1 represent the forecasting of the bus departures using trend line. In visual we can see that projections indicates anticipated to bus departures to reach 3,900 on 04/30/2029(Quarter 2). Visual\_2 represents the forecasting of passenger’s departures; projections indicate that passenger numbers are anticipated to surpass 125,000 on 06/23/2025(Quarter 2). Visual\_3 represents the key influencer of the bus and passenger departures, here we can see that the external factors like climate, weather, population, public holidays, and gas price. Among these factors gas is the main key influencer for the both bus and passenger departures.



**Visual\_1: Bus Departure Forecast.**



**Visual\_2: Passenger Departures Forecast.**

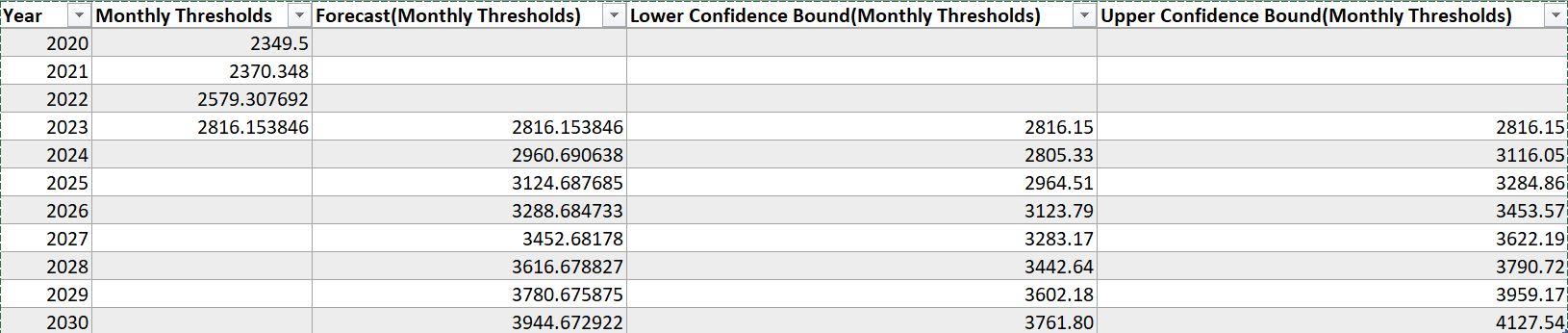


**Visual\_3: Key Influencer.**

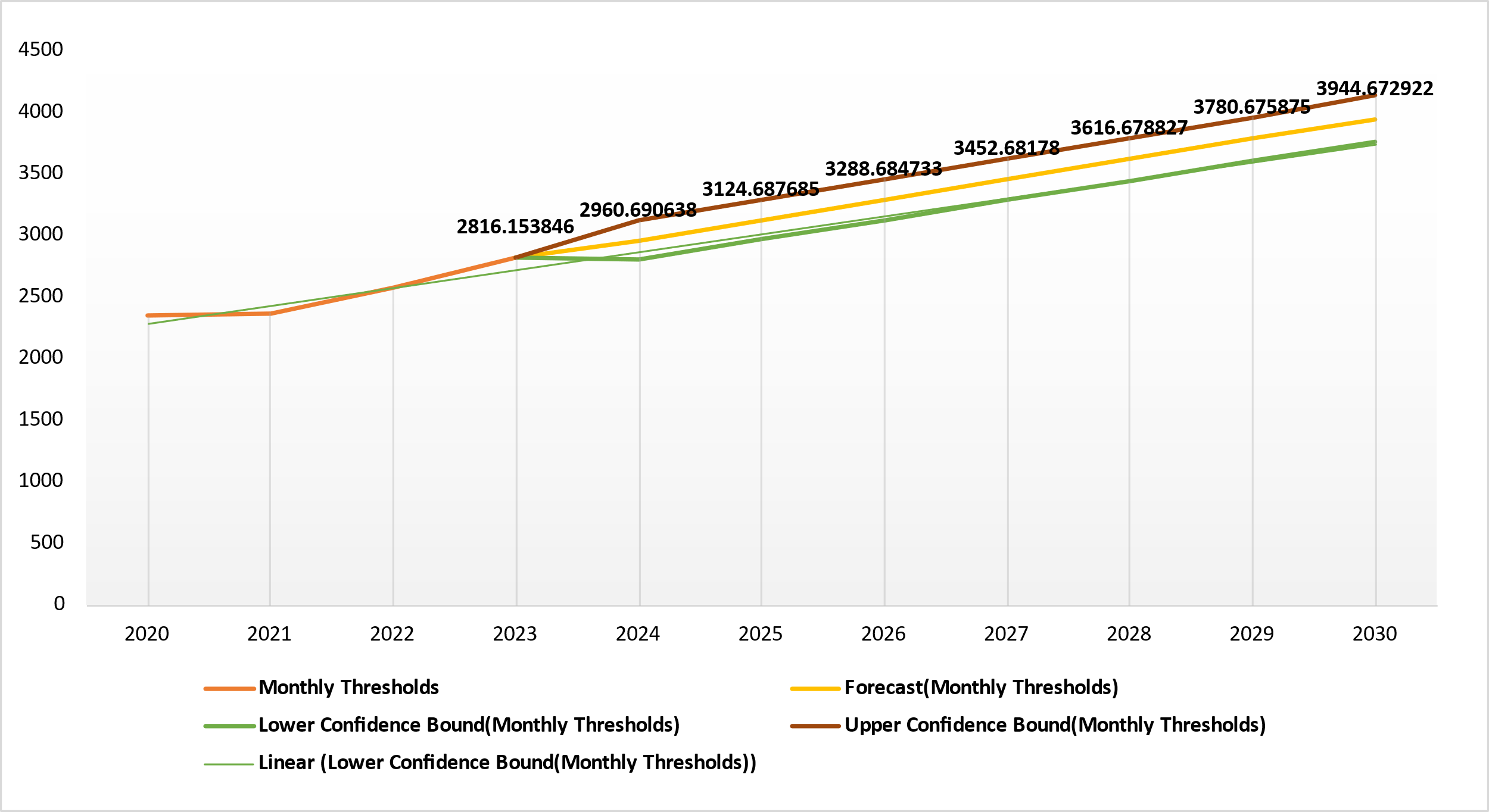
The ACF for the number of bus departures shows a strong positive correlation. This means that the current number of bus departures is highly correlated with the number of bus departures yesterday. The current number of bus departures is also somewhat correlated with the number of bus departures two and three days ago. The ACF for the number of bus departures is similar to the ACF for passenger totals versus bus departures. This suggests that there is a strong positive relationship between the number of bus departures and the number of passengers.

**Exponential Smoothing Output:**

Here we are providing the forecast visuals of the bus and passenger departures as per the requirement of the business need using Exponential Smoothing using Excel. Based on consolidated data prepared we used the FORECAST.ETS function in Microsoft Excel to perform time series forecasting. Forecast Parameters “-1”, this number represents the number of future data points we want to forecast. In this case, we are forecasting just one data point into the future. The FORECAST.ETS function in Excel utilizes historical time series data (x-values and corresponding y-values) to predict future values based on the specified algorithm. The output of this function at a given time point will provide an estimated y-value (forecasted value) corresponding to that x-value. Visuals 4.1 and 4.2 represents the bus departures forecasting, in visual\_4.1 we can observe that by 2030 the bus departures suppress 3900 departures and value is 3944 departures. Visual\_4.2 is the forecasting.

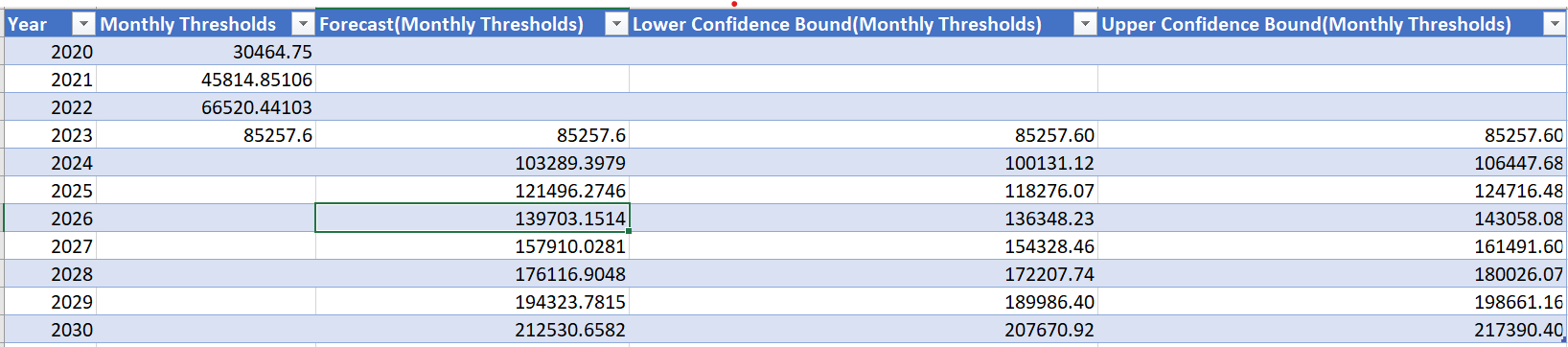


**Visual\_4.1: ES\_Bus\_Departures\_Forecasting\_Values.**

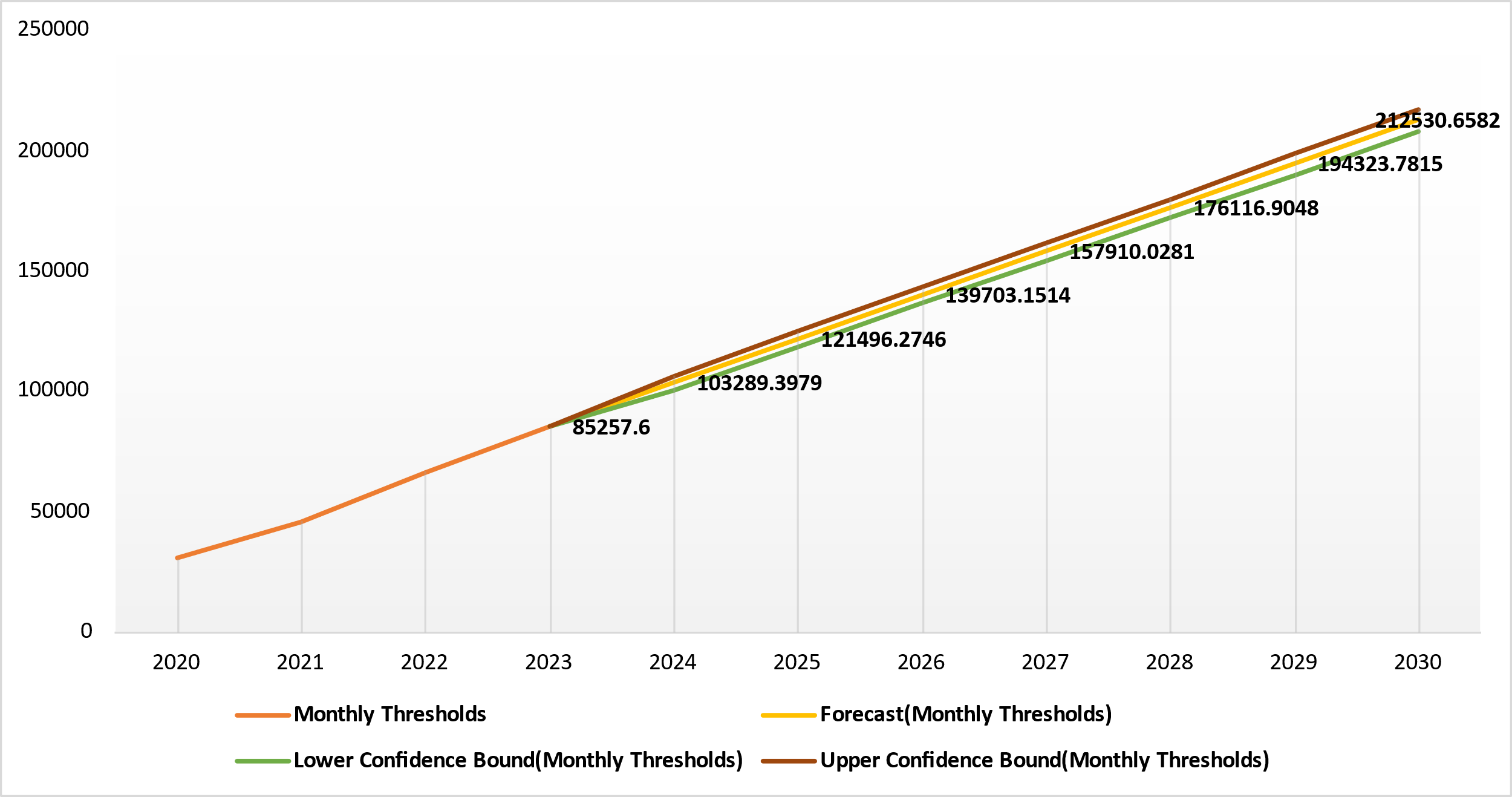


**Visual\_4.2: ES\_Bus\_Departures\_Forecasting.**

In visual\_5.1 we can observe the forecasting values of passenger departures. As per the business requirement question by year 2026 passengers’ departures suppress 125000. The forecasted value is 139703 passengers’ departures. Visual\_5.2 is forecasting of these values using timeseries plot.



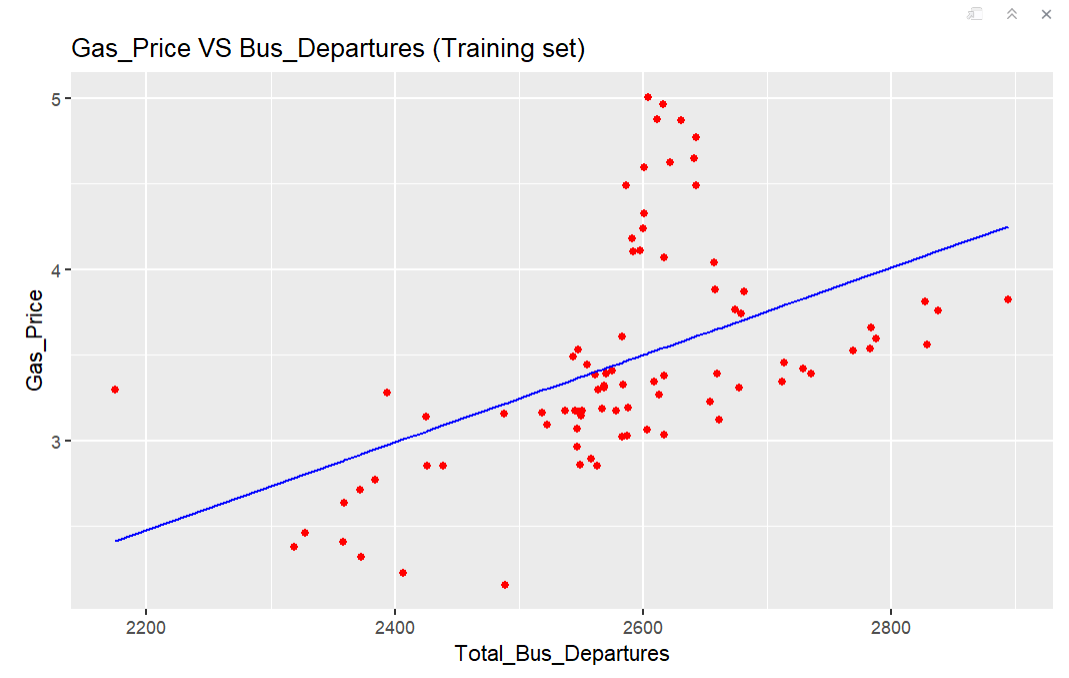
**Visual\_5.1: ES\_Passengers\_Departures\_Forecasting\_Values**.

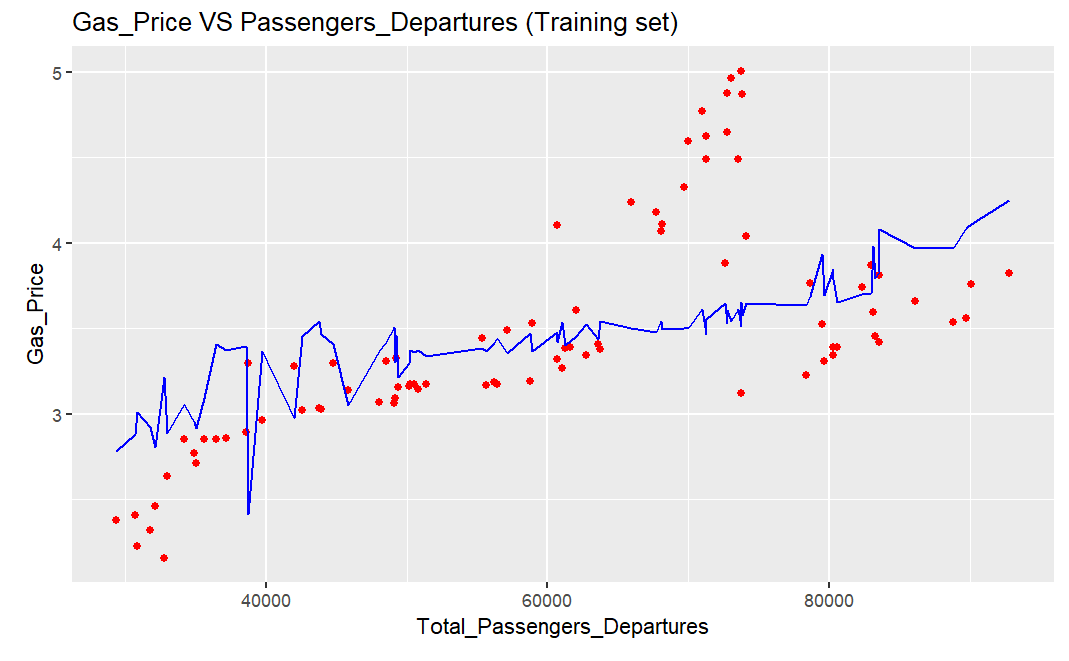


**Visual\_5.2: ES\_Passengers\_Departures\_Forecasting**

**ARIMA Model Outputs:**

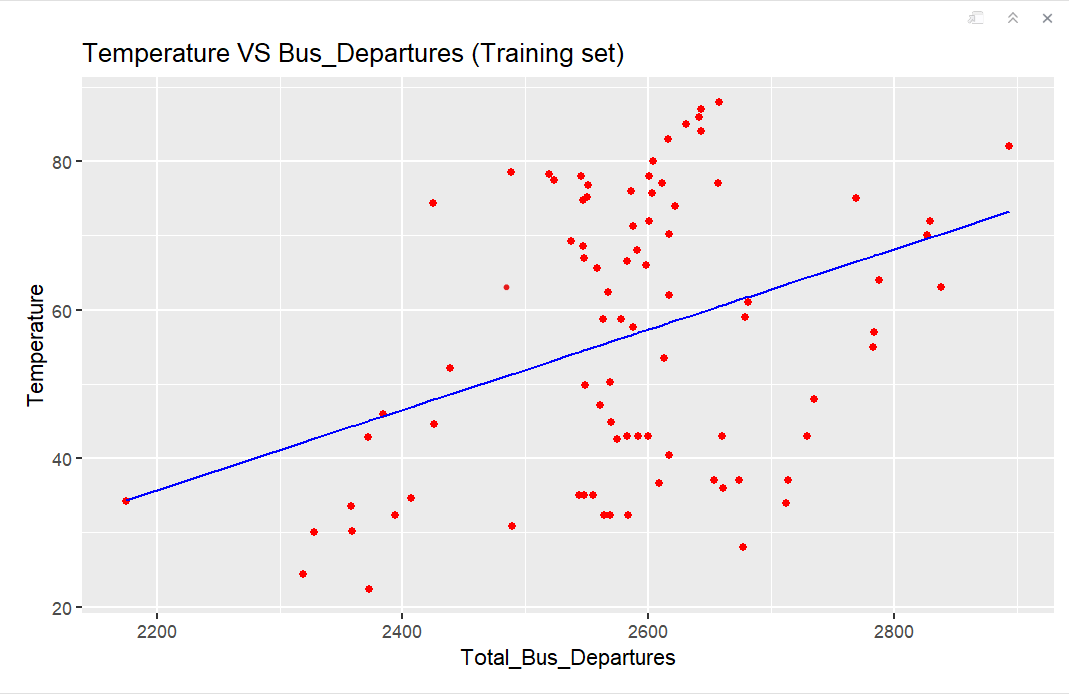
Here we are providing the output of the ARIMA Model R programming output. First, we compared the bus and passenger’s departures with gas price and temperature. We observed that the bus departures and passenger’s departures are more when gas price is between $4 - $5.

**Visual\_6.1: Gas\_Price VS Bus\_Departures**

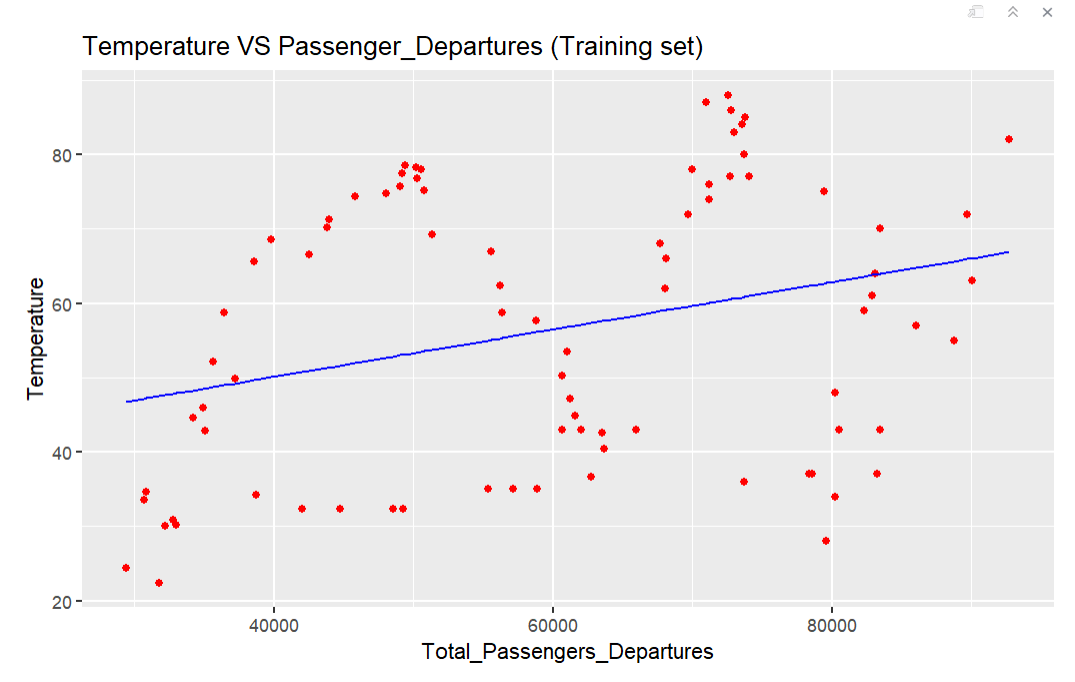


**Visual\_6.2: Gas\_Price VS Passengers\_Departures**

We also compared bus and passenger’s departures with temperature, we observed that when temperature is between 50F – 80F there were more departures of bus and passengers.



**Visual\_7.1: Temperature VS Bus\_Departures**



**Visual\_7.2: Temperature VS Passenger\_Departures**

**Conclusion:**

We conclude that based on rigorous analysis utilizing predictive modeling, our projections indicate that passenger numbers are anticipated to surpass 125,000 on 06/23/2025 and bus departures to reach 3,900 on 04/30/2029 using Microsoft Power BI. When compared with external factors like gas price and temperature we observed that the bus departures and passenger’s departures are more when gas price is between $4 - $5, and when temperature is between 50F – 80F there were more departures of bus and passengers. Also, we observed that the number of passengers and the number of bus departures have a significant positive correlation.

**6430\_02\_Group\_3 Members: Group of 7 Members.**

1. Sravani Kandhukuri, skand36@unh.newhaven.edu
2. Madhulika Ganapathireddy, mgana4@unh.newhaven.edu
3. Abhilash Goud Rampelly, aramp1@unh.newhaven.edu
4. Santhosh Kumar Alakurthi, salak2@unh.newhaven.edu
5. Sai Rahul Chirra: schir8@unh.newhaven.edu
6. Anand Thorpunuri, athor3@unh.newhaven.edu
7. Pavan Kumar Nagam, pnaga7@unh.newhaven.edu