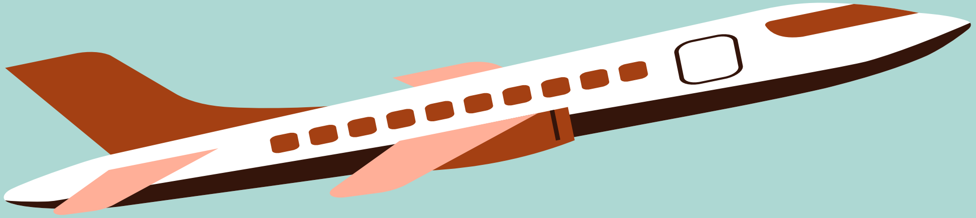
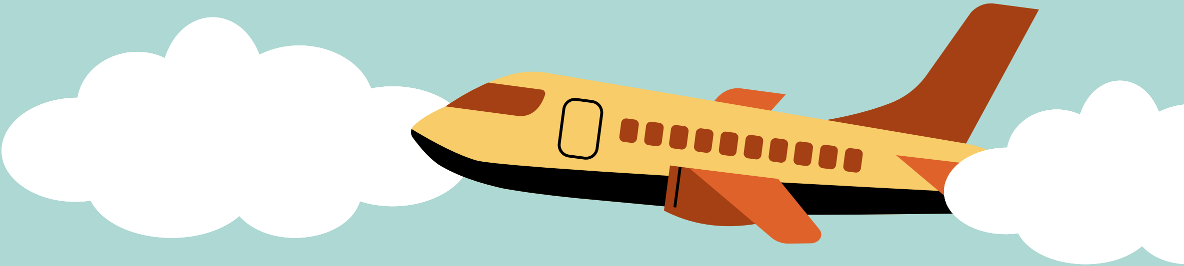


U.S. Air Enplanements Time Series Analysis

Team B: Adriana Choy Sampei, Yiyuan Cui, Yabi Yirga, Jing Zhang

Table of Contents:



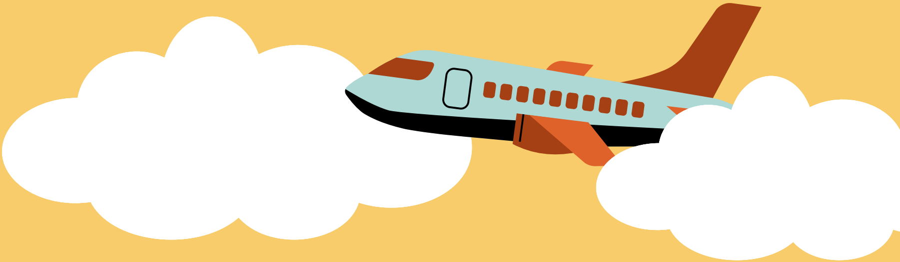
1. Introduction

Background

This analysis serves in many different aspects:

- Strategic decision-making
- Allowing for better planning
- Resource allocation
- Service optimization
- Consumer Behavior and Preferences

Dataset



The dataset contains information from all scheduled passenger flights.

It includes both domestic and international flights operated by U.S. airlines.

The data is monthly, covering the period from January 2000 to July 2023.

Convert the data to average number of enplanements per day.

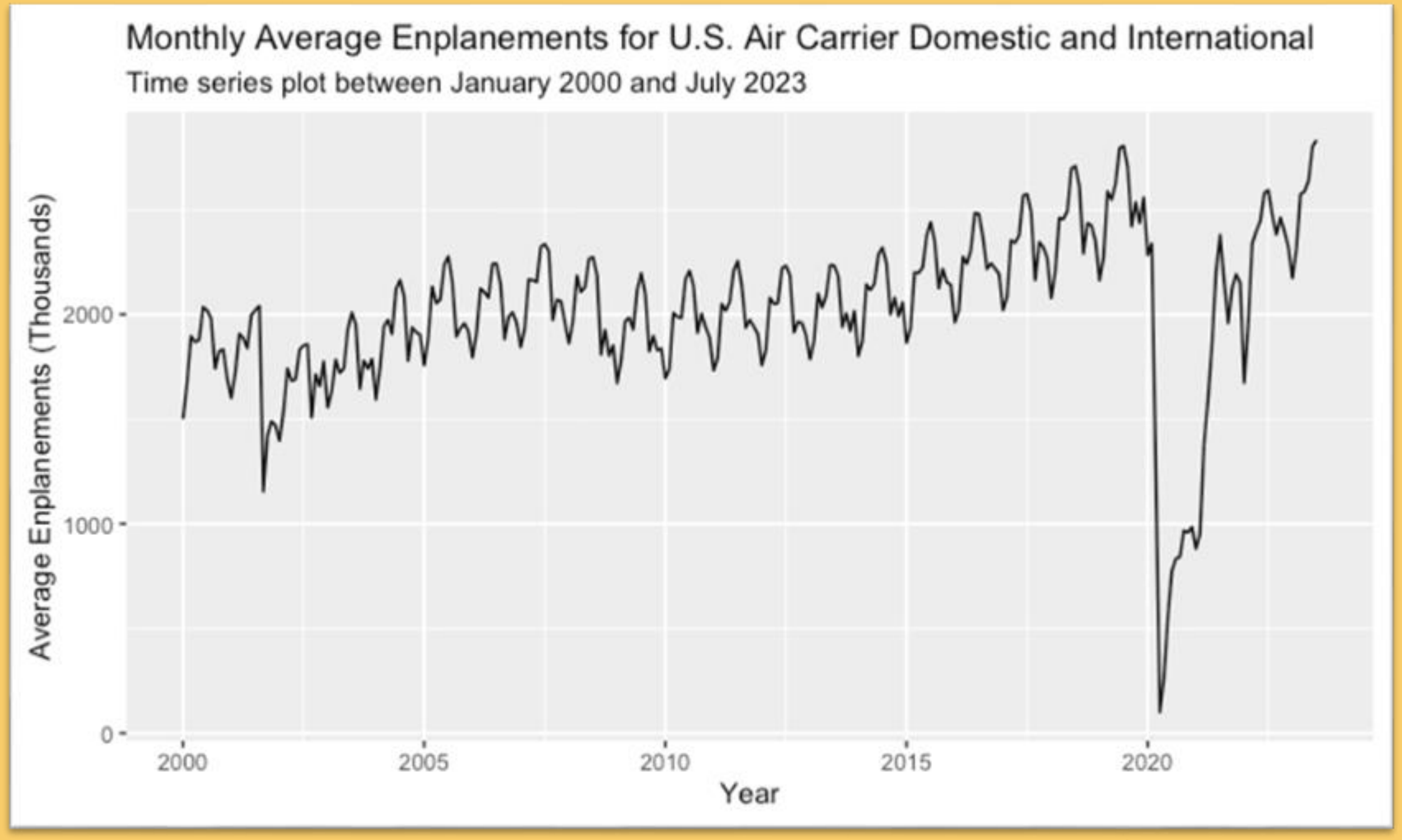
Notable features of the dataset include evident seasonality and Visible trend.

Questions we will be exploring

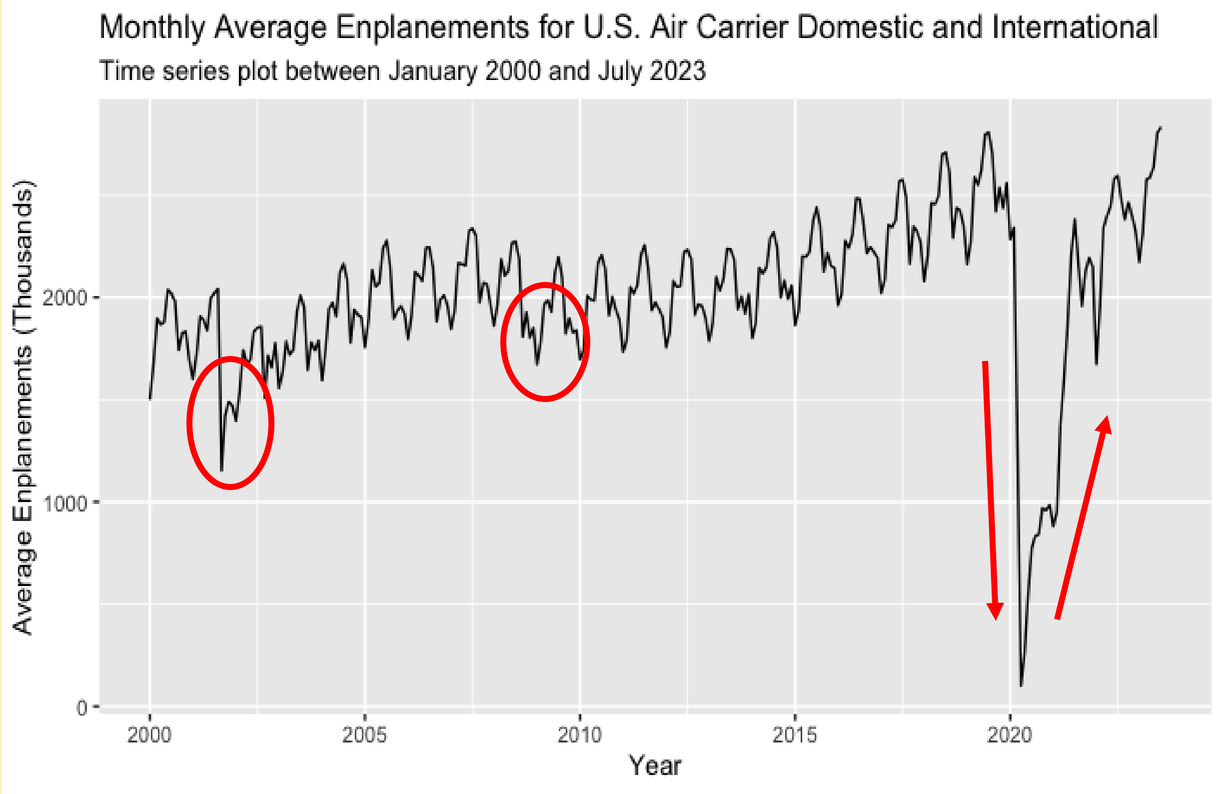
- What is the impact of major events on the number of passenger flights?
- What drives the seasonality of the time series?
- What time of the year should airlines expect a larger number of passenger flights?
- What is the best model to forecast the number of passenger flights for the following year?

2. Qualitative Description

Time Series Plot



Qualitative Description

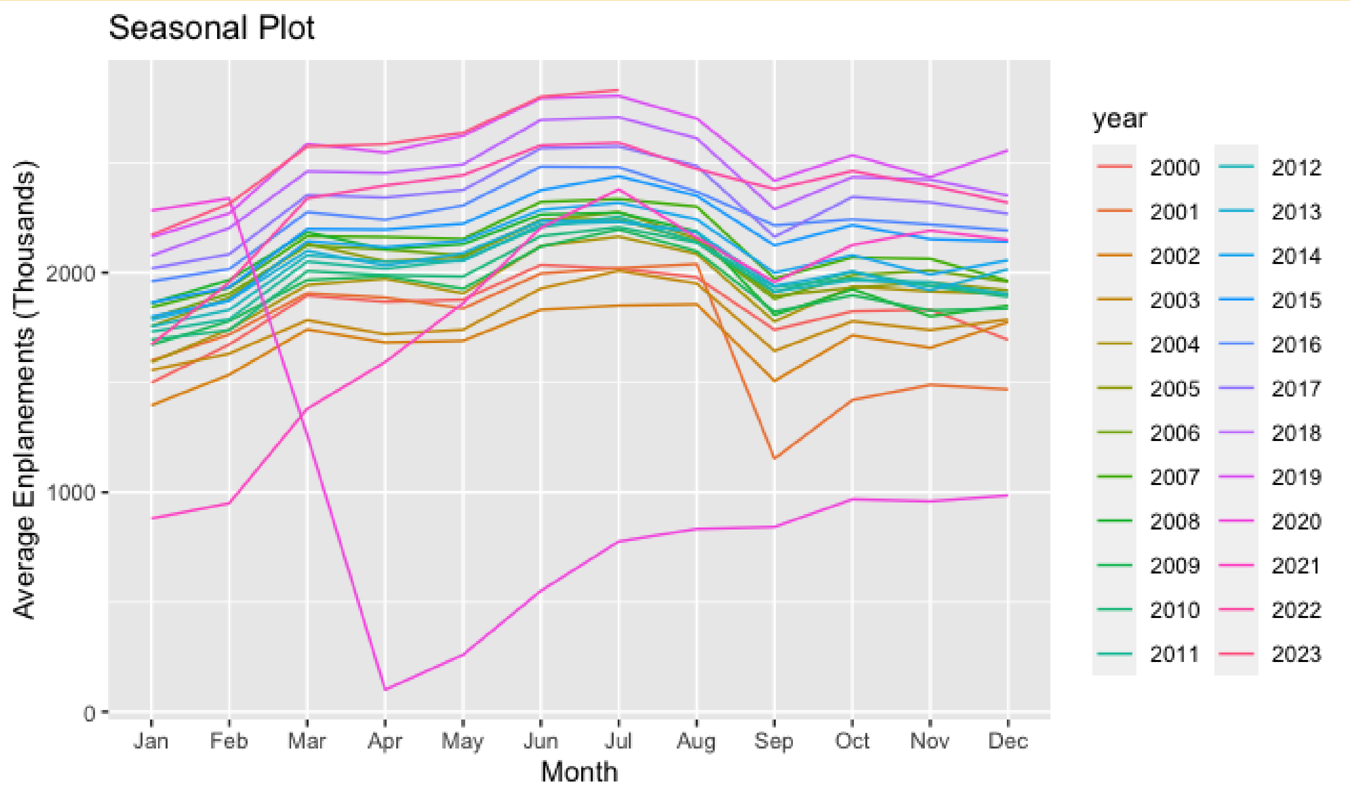


Trend:
An overall upward trend from 2000 to 2019;
Declines in September 2001 and at the end of 2008;
A sharp drop at early 2020, recovering from March 2021, then gradually increasing.

Cycles:
September 11 attacks (9/11/2001);
2008 financial crisis;
COVID-19 pandemic (lockdowns beginning from March 2020).

Seasonality:
Annual seasonality.

Seasonality



Peak seasons: From May to August;
A significant drop in September;
Winter months tend to have lower enplanements.

Decomposition

- Strength of trend: 0. 82
- Strength of seasonality: 0.53

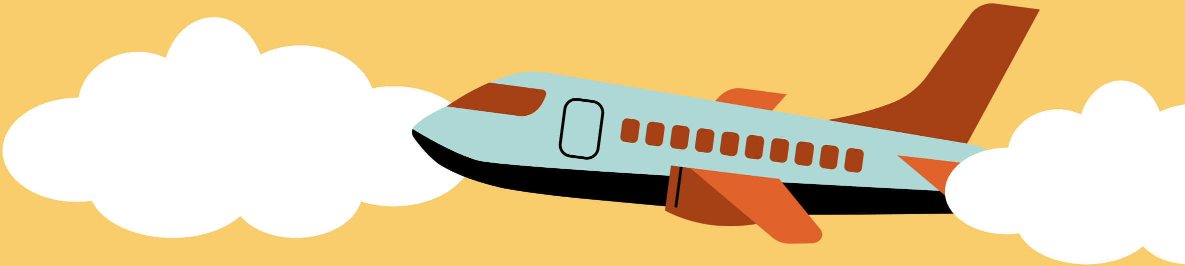
Trend and seasonality do exist in the time series.

Related Time Series: Rail Passenger Miles

- Similar trends and seasonality;
- No significant drop in September 2001 for rail transportation;
- Correlation: 0.76. Rail is not a good predictor for flights.

3. Model Creation

Appropriate Models



Seasonal Naïve

RMSE = 488.5077

Holt-Winters

RMSE = 125.3203

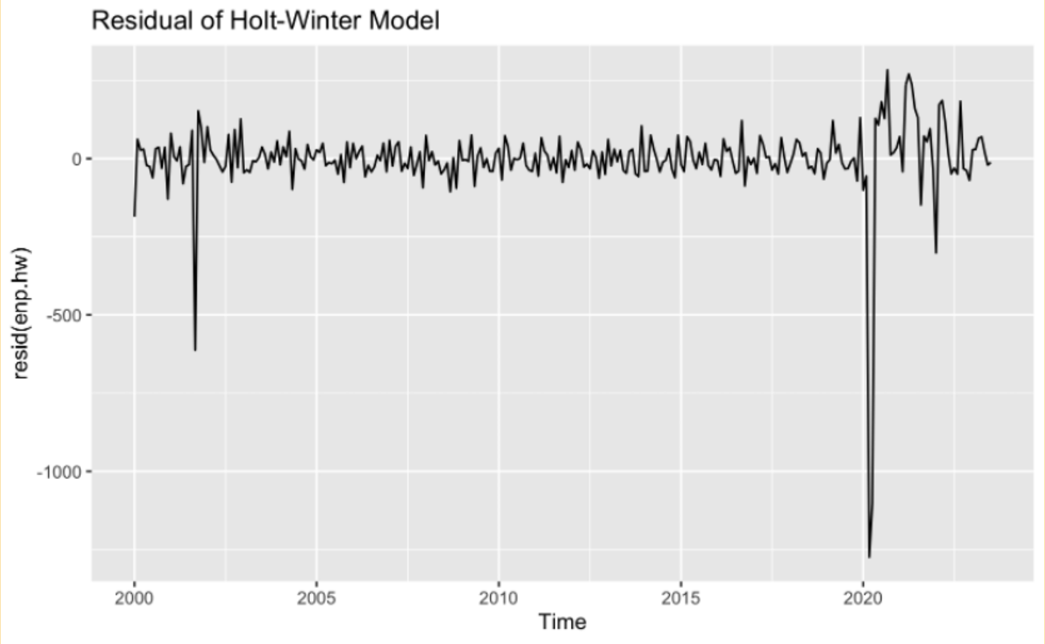
ARIMA(2,1,0)(2,0,0)[12]

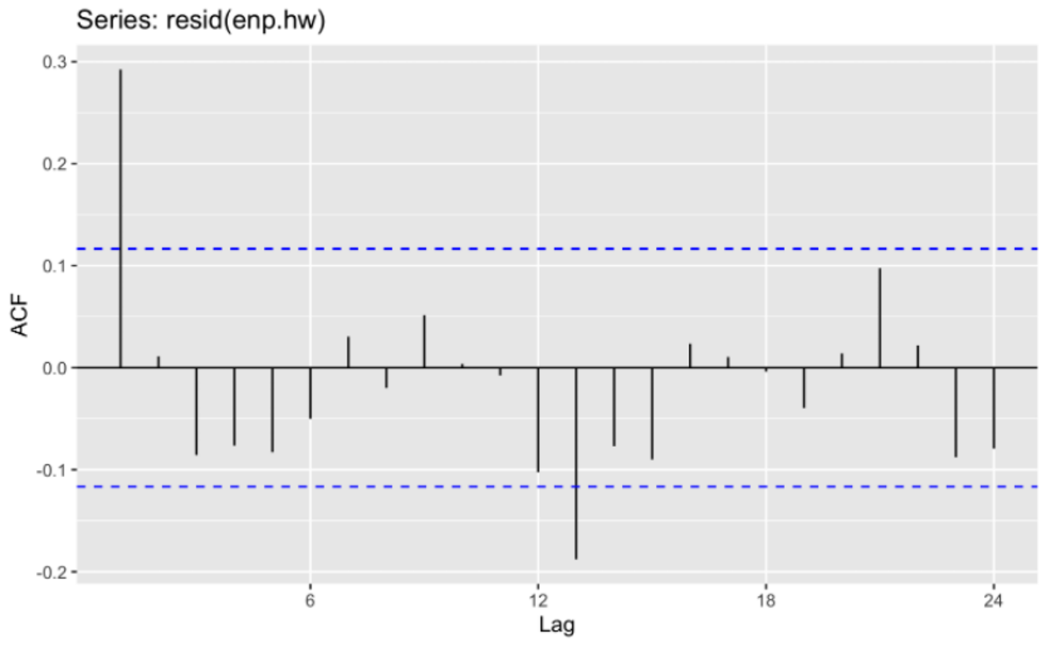
RMSE = 140.6094

Linear (season + poly(trend,2)

RMSE = 341.2709

Residual of the Holt-Winter Analysis

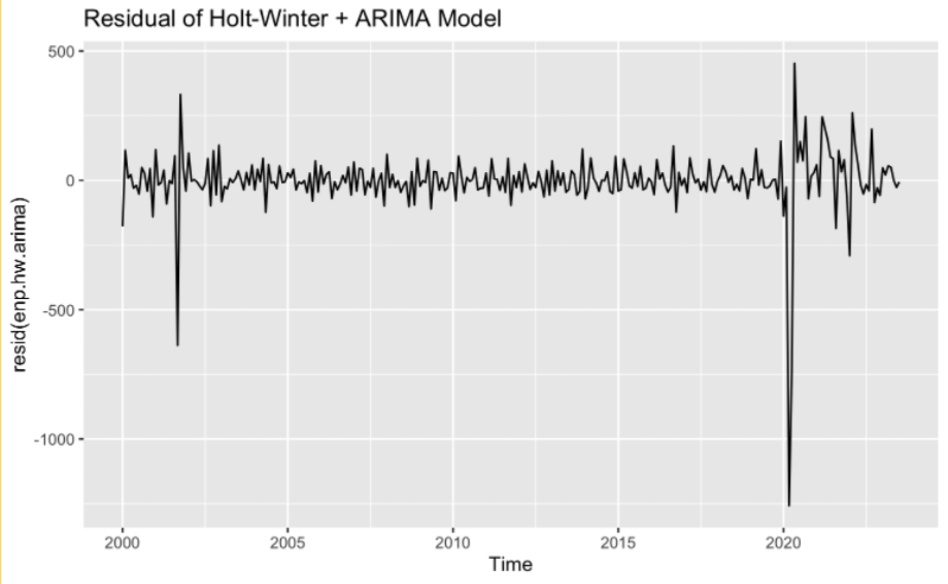


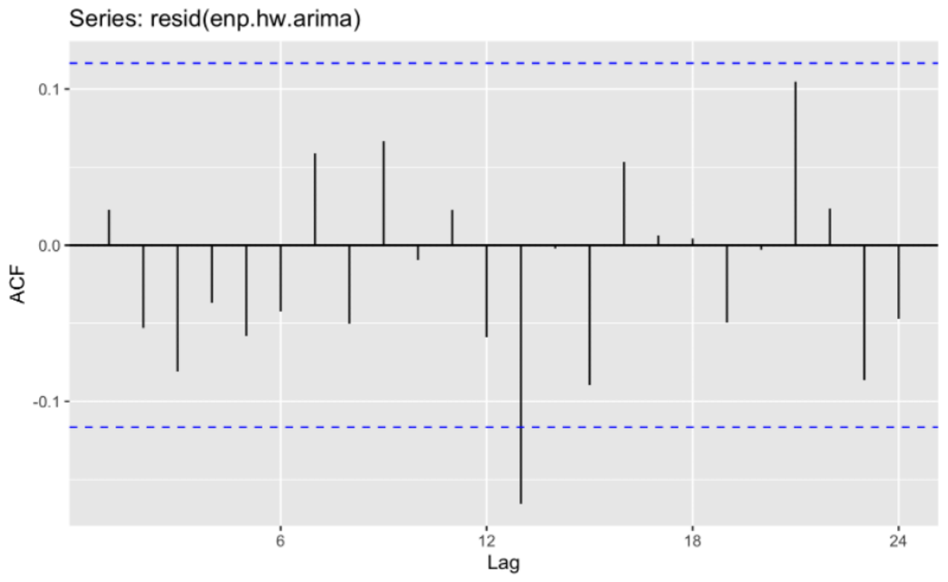


Two-stage Model

Fit an ARIMA model to the residuals from the Holt-Winters model.

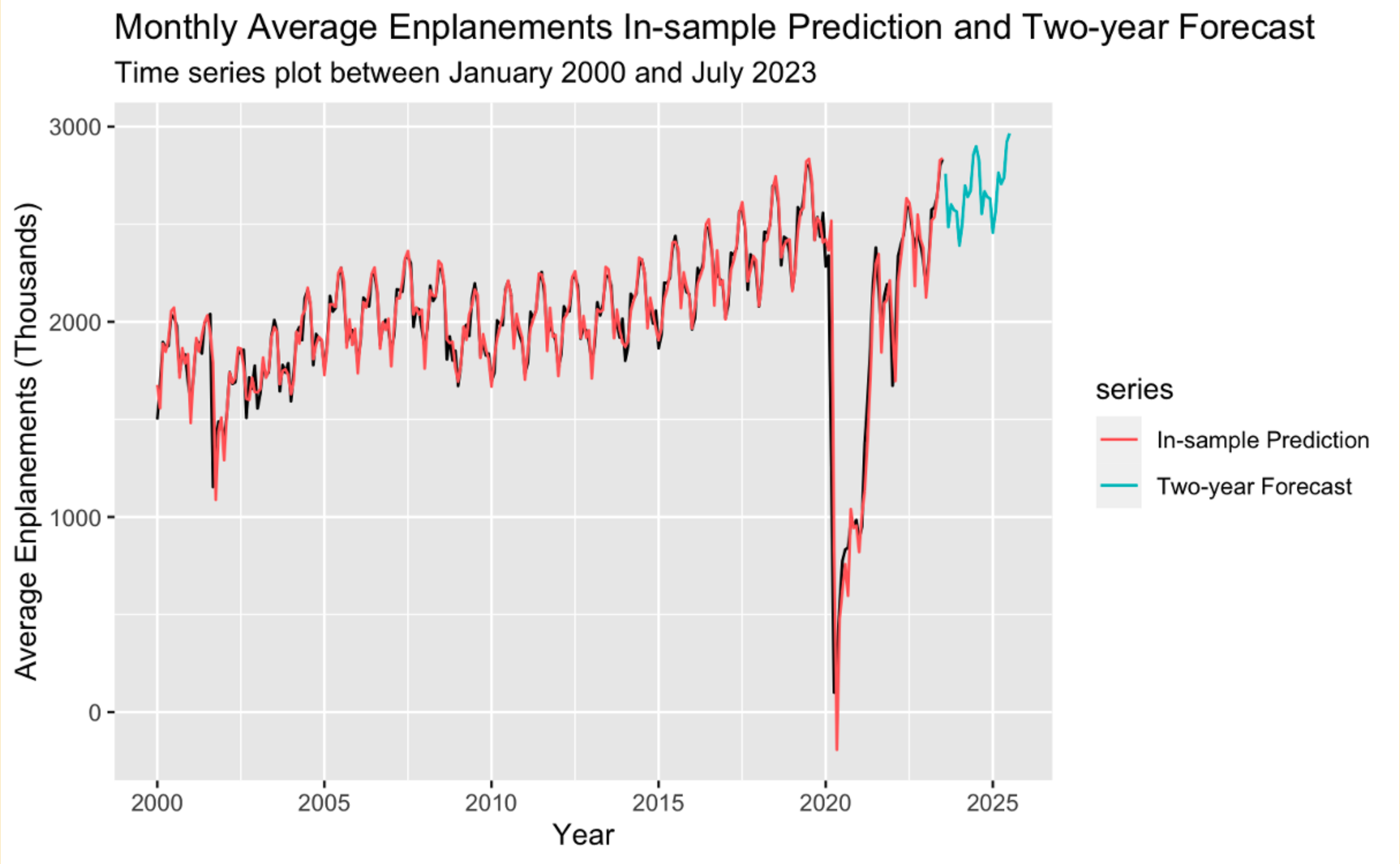
RMSE = 119.7914





In-sample Prediction & Two-year Forecast

Two-stage model (Holt-winters + ARIMA(1,0,0))



4. Conclusion

Conclusion & Takeaways

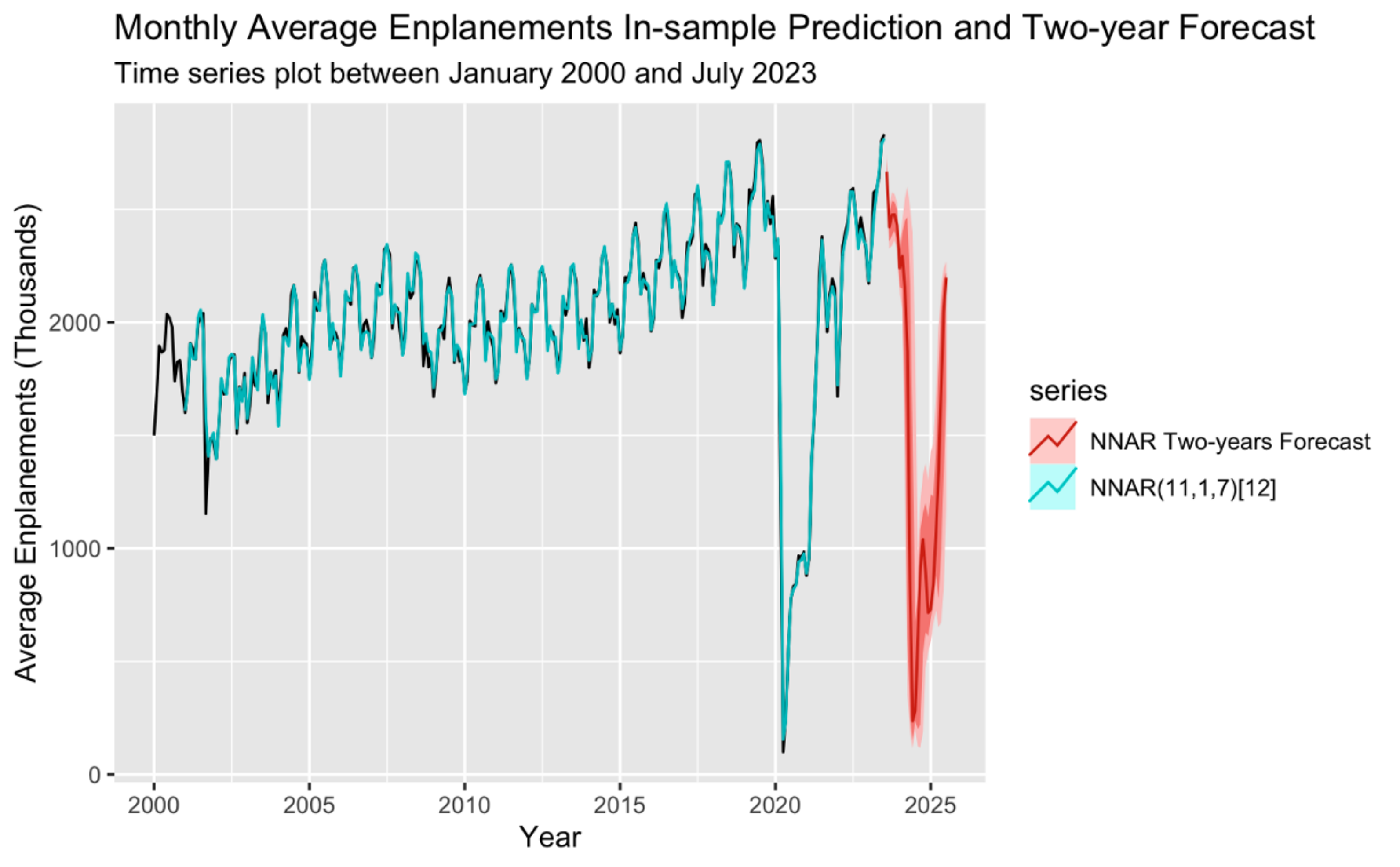
1. Events that affect the trust in the mode of transportation or the economy have a significant impact on the number of passenger flights.

2. Number of flights has a yearly seasonality that peaks during warmer months, which is likely associated with better weather and more available leisure time.

Seasonal Naïve	**Holt-Winters**	SARIMA	Linear Regression
488.5077	**125.3203**	140.6094	341.2709

3. Lower RMSE models are not necessarily the most ideal ones at the risk of overfitting.

NNAR Model

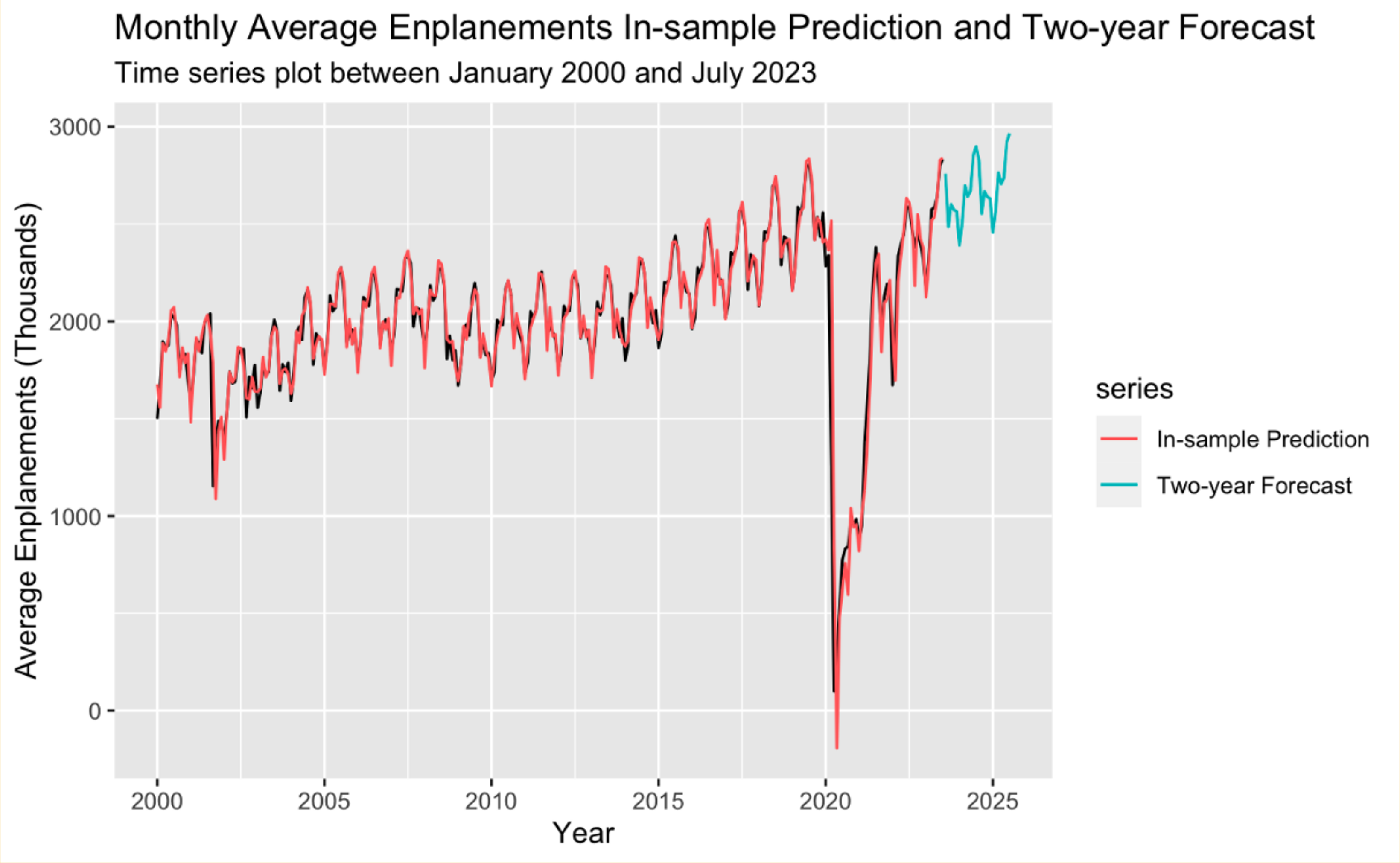


<u>NNAR 2–year forecasting</u>

- RMSE = 41.40211
- Captures the exceptional drop in flights during the COVID-19 lockdowns

Two-stage Model

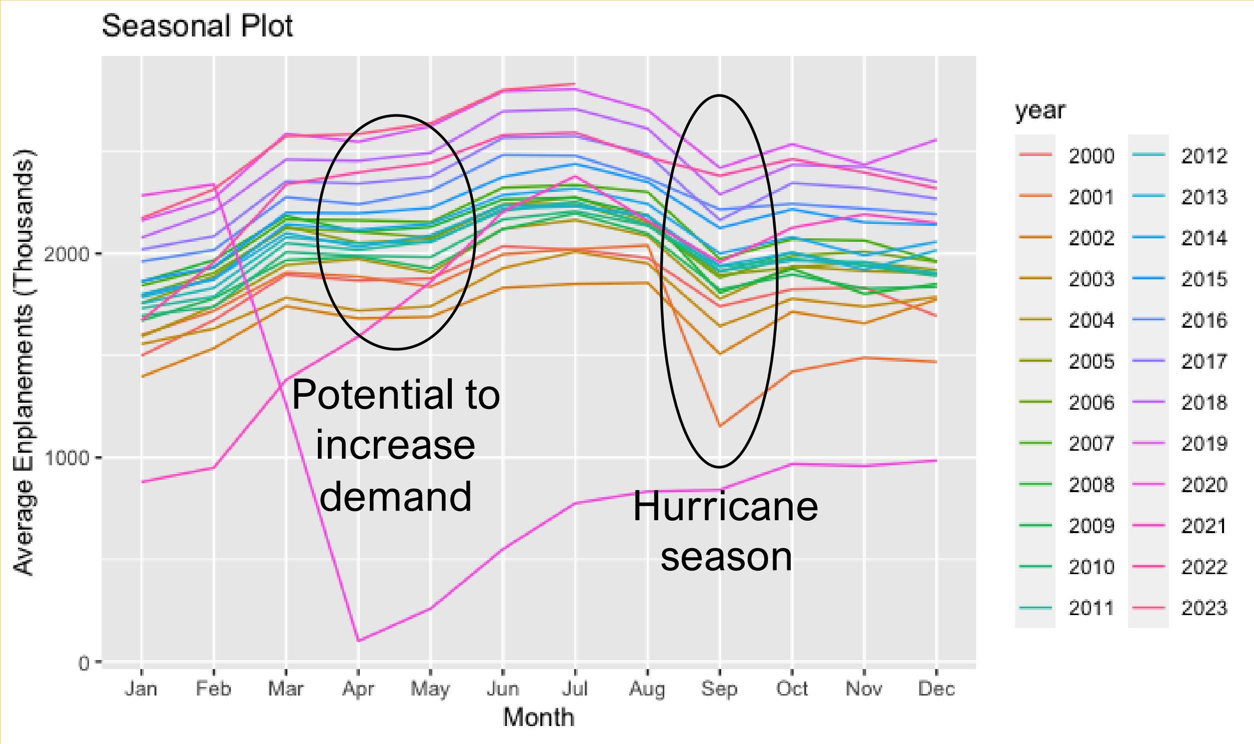
(Holt-Winters + ARIMA)



2–year forecasting

- RMSE = 125.3203
- Captures yearly seasonality, and projects a likely upward trend

Recommendations



Customers

- Implement marketing or promotional tactics to incentivize customer demand on April and May

Resources

- Investment in equipment that might help avoid bad-weather cancellations
- Prepare for a possible all-time peak and what that implies in terms of resource availability and staffing

Thank you!