ANA535: Laboratory – III Written Report

# Regression and Smoothing-Based Forecasting

Submitted by:

Naresh Anbazhagan

Graduate Student, M.S. in Data Analytics

Course:

ANA 535 – Forecasting

Instructor:

Dr. Marvine Hamner

Date:

May 04, 2025

# Table of Contents

|  |  |  |
| --- | --- | --- |
| 1 | Executive Summary | 1 |
| 2 | Introduction | 2 |
| 3 | Methods | 3 |
| 4 | Observations | 6 |
| 5 | Results | 9 |
| 6 | Discussion | 12 |
| 7 | Conclusion | 14 |
| 8 | Plots | 15 |
| 9 | References | 42 |
| 10 | Appendices A and B | 42 |

# Executive Summary

This report analyzes Amtrak’s monthly passenger miles from January 1991 through mid-June 2024 using a range of forecasting techniques to identify long-term trends, recurring seasonal patterns, and model performance across different periods, including the impact of the COVID-19 pandemic.

All analyses were conducted in R using a combination of packages from both the fpp3 and forecast ecosystems, including tsibble, fable, feasts, forecast, and slider. We began by fitting linear, quadratic, and cubic regression models to capture potential growth trends, followed by classical decomposition and seasonal differencing to isolate trend and seasonal components.

Moving average smoothing was also applied to help reveal underlying patterns, despite its inherent lag. Among regression models, the cubic polynomial provided the best overall fit. Exponential smoothing methods, including ETS and Holt-Winters (HW), were then applied, highlighting notable differences between pre-2020 and post-2020 periods, especially regarding seasonal variability and trend flattening.

Finally, benchmark models such as Seasonal Naïve (SNaive) and Simple Exponential Smoothing (SES) were included. SES models using both fixed and optimized alpha values confirmed the trade-off between stability and adaptiveness. The findings from this analysis provide valuable insights for forecasting future ridership and support data-driven planning decisions in the transportation sector.

The best-performing model, a cubic regression with seasonality, accurately captured trend and seasonality across the full dataset (see Figure 14). Additionally, exponential smoothing provided consistent short-term forecasts (see Figure 49), reinforcing the robustness of our approach.

# Introduction

Amtrak, the national railroad passenger service in the United States, serves as a critical component of the country’s transportation infrastructure. Understanding historical passenger trends is essential for strategic planning, policy formulation, and operational efficiency. This laboratory exercise focuses on analyzing Amtrak’s monthly PassengerMiles — a key indicator of total distance traveled by all passengers — spanning from January 1991 through mid-June 2024.

The objective of this lab was to apply a variety of time series forecasting techniques to determine which models best capture the dynamics of Amtrak ridership over time. The analysis incorporated linear and polynomial regression, classical decomposition, moving average smoothing, exponential smoothing methods (ETS and Holt-Winters), and residual diagnostics. Emphasis was placed on evaluating changes in ridership behavior during and after the COVID-19 pandemic.

All modeling and analysis were conducted in R using both modern (fpp3) and traditional (forecast) time series libraries. Forecast performance was evaluated using statistical accuracy measures such as MAPE and RMSE, in addition to visual diagnostics including residual plots and seasonal decompositions. This multifaceted approach provided both theoretical and practical insights into the forecasting of transportation demand.

Figure 3 presents the full time series of monthly ridership, revealing both long-term growth and strong seasonal fluctuations. Additionally, the distribution of raw passenger miles data is shown in Figure 6, highlighting a slightly right-skewed pattern. To better understand the growth trajectory, an initial cubic trend fit was visualized using ggplot (see Figure 1), suggesting potential nonlinear patterns in long-term usage.

# Methods

This laboratory exercise analyzed monthly Amtrak Passenger Miles from January 1991 to mid-June 2024 using R. The dataset was cleaned and preprocessed to ensure quality prior to modeling. All time-series operations and forecasting techniques were implemented using a mix of modern and legacy R packages including fpp3, forecast, feasts, slider, zoo, and TSA.

Data Preparation

The raw dataset was parsed using *lubridate::mdy()* to convert dates into a standardized format. Basic exploratory data checks using histograms and line plots were used to confirm data integrity and identify outliers or missing values. These initial checks are illustrated in Figure 6 (Histogram of Raw Data) and Figure 3 (Initial Time Series Plot)

Time series structures were defined using both *ts()* (for legacy compatibility) and *as\_tsibble()* (for tidy workflows) with monthly frequency. The full dataset was segmented into meaningful sub-periods (1991–2016, 1991–2004, and 2020–2024) for comparative analysis, especially to examine the effects of COVID-19 on passenger trends.

Trend Analysis

Trend estimation was performed using linear, quadratic, and cubic models via the tslm() function. The best-fitting polynomial trend was identified using residual analysis and forecast accuracy metrics such as MAPE and RMSE. Figure 14 shows the cubic regression fit over the full time series. Figures 31 and 32 display the ACF and PACF of residuals from the cubic model.

These models were applied to multiple segments of the data to account for structural shifts. Visualizations of trend fits, and residual diagnostics are shown in Figures 14 (Cubic Fit Example), 31 (ACF of Residuals), and 32 (PACF of Residuals).

Smoothing and Decomposition

To analyze underlying patterns, moving averages were computed using *slider::slide\_dbl()* and *zoo::rollmean().* Both centered and trailing 3-month moving averages were tested to study lag effects (Fig **38**: Centered 3-MA, **40**: Trailing 3-MA, **39**: Zoomed Centered 3-MA, and **41**: Zoomed Trailing 3-MA).

Classical decomposition was applied using *decompose()* to isolate trend, seasonal, and residual components (Figure 36: Classical Decomposition). Seasonal differencing and logarithmic transformations were employed to stabilize variance and address residual seasonality (Figure 27: Log + 1st Difference Decomposition, Figure 28: Log + 2nd Difference Decomposition, Figure 30: Second Differencing Visualization).

Forecasting Models

Multiple forecasting methods were implemented:

* ETS models were fit using ets() across multiple periods, including 1991–2019 and 2020–2024, to capture level, trend, and seasonal components. The decomposition of fitted models is shown in Figure 43 (ETS Fit: 1991–2019) and Figure 46 (ETS Fit: 2020–2024).
* Forecast projections using ETS are illustrated in Figure 45 (ETS Forecast: 1991–2019 with 5-Year Projection) and Figure 48 (ETS Forecast: 2020–2024 with 5-Year Projection). Overlay plots comparing fitted data and forecasts are shown in Figures 44 and 47.
* Holt-Winters exponential smoothing was applied using the hw() function with both additive and multiplicative seasonal options. The forecasts from 2020 to 2024 are displayed in Figure 52 (HW Additive vs Multiplicative Forecasts).
* Long-range HW forecasts (1991 onward) are shown in Figure 53, and comparative forecasts with multiple HW configurations are presented in Figure 51.
* HW-only plots for additive and multiplicative models from 1991–2019 are shown in Figures 50 and 51, respectively.
* Simple Exponential Smoothing (SES) was implemented with both a fixed α = 0.2 and an optimized ANN model. Forecast comparison results are included in Figure 49 (Fixed Alpha SES).

Model Evaluation

Forecast accuracy was evaluated using built-in *accuracy()* functions in the *forecast* and *fabletools* packages. Metrics such as RMSE, MAE, and MAPE were used to compare models across different time periods. ACF, PACF, and residual diagnostics were plotted using *gg\_tsresiduals()* and *acf()* to validate stationarity and residual independence, as shown in Figures 31 (ACF of Residuals), 32 (PACF of Differenced Series), and 35 (gg\_tsresiduals() Diagnostic Plot).

# Observations

Initial EDA and Trend Visualization

The initial inspection of Amtrak Passenger Miles revealed a long-term upward trend with consistent seasonal spikes, particularly during summer months. A histogram of the raw data indicated a slightly right-skewed distribution (see figure 6), prompting consideration of variance-stabilizing transformations such as log-scaling.

Time series plots spanning 1991 to 2024 (see figure 3) illustrated a smooth increase until the sharp decline observed around 2020, attributable to the COVID-19 pandemic. To evaluate the underlying trend, we fit both quadratic and cubic regression models.

The cubic model provided a visibly better fit over the linear and quadratic alternatives, especially when capturing inflection points in the ridership pattern. However, neither model fully accounted for the pronounced seasonality in the data.

Trend Modeling

Cubic trend models consistently outperformed simpler alternatives across all historical periods (1991–1997, 1991–2004, 1991–2016, and 1991–2020). Residual analysis (see figure 16) confirmed minimal autocorrelation, validating the robustness of the cubic fit. Despite this, the regression models could not fully capture recurring seasonal variation, necessitating further seasonal modeling.

Smoothing and Seasonality

Three-month centered moving averages clarified short-term noise while preserving underlying trends, though trailing versions introduced lag in responsiveness. Classical decomposition plots (see figure 36) emphasized the presence of strong and regular seasonality.

Even after a single differencing pass, residual seasonal patterns persisted, leading to the use of second-order seasonal differencing. Log transformations aided in variance stabilization, especially when forecasting post-2019 data.

Forecasting Results

Exponential smoothing (ETS) models produced contrasting results depending on the timeframe:

* 1991–2019 models (ETS(M,N,A)) captured moderate upward trends with consistent seasonal variation.
* 2020–2024 models (ETS(M,N,M)) showed flatter trajectories with persistent seasonality but diminished growth. (see Figures 43 and 46)

The Holt-Winters forecasts differed in magnitude and shape:

* Additive forecasts retained moderate upward movement,
* Multiplicative forecasts (see figure 52) declined sharply, especially in more recent years.

Comparative Forecast Insights

Comparing forecasts between pre-pandemic (1991–2019) and post-pandemic (2020–2024) periods underscored the significant drop in ridership. Post-COVID models projected flatter or declining trends, while pre-COVID models predicted continued growth. Seasonal Naive methods worked well for short-term projections but failed to capture long-term dynamics or structural breaks.

Seasonal naïve forecasts provided a reasonable short-term baseline, particularly for recurring seasonal patterns, but failed to account for trend changes or structural breaks such as the COVID-19 disruption.

# Results

The forecasting models produced varying levels of accuracy depending on time-period and methodology. Evaluation metrics such as RMSE and MAPE were used across all experiments to validate forecast accuracy.

Regression Model Results

Cubic regression models consistently outperformed linear and quadratic variants in terms of residual error and forecast alignment across all examined windows (1991–1997, 1991–2004, 1991–2016). MAPE values for cubic models were notably lower (see Figure 14), and residuals showed minimal autocorrelation, supporting their use for broad trend forecasting.

Decomposition and Smoothing

Decomposition plots revealed that seasonal components were stable across years, but trend components changed sharply after 2020.

Moving average smoothing (see Figure 38) highlighted short-term fluctuations, though these models lagged during rapid changes in ridership patterns. Classical decomposition successfully isolated seasonal and trend components but left minor cyclical noise.

A zoomed-in view from 2001 to 2003 using centered and trailing 3-month moving averages further illustrated that trailing averages tend to lag during periods of rapid seasonal inflection.

Exponential Smoothing (ETS) Models

ETS(M,N,A) models trained on pre-COVID data forecasted a steady rise in passenger miles (see Figure 43), while ETS(M,N,M) models trained on 2020–2024 data showed flat trends with stable seasonality.

Table 1 - Summary of Forecasting Models and Their Characteristics

|  |  |  |
| --- | --- | --- |
| Model | Characteristics | Observations |
| Seasonal Naive | Repeats last year’s values | Good baseline; captures seasonality but ignores trend |
| ETS (M,N,A) | Multiplicative errors, additive seasonality | Stable forecast; underreacts to recent trends |
| ETS (M,N,M) | Multiplicative errors and seasonality | Captures changing seasonal strength post-COVID |
| Holt-Winters Additive | Additive trend and seasonality | Performs well under stable seasonal amplitude |
| Holt-Winters Multiplicative | Multiplicative trend and seasonality | Best suited for variable seasonal effects (e.g., post-2020) |
| HW Model Comparison | Multiple fits tested | Slight variance; HW Fit 1 outperforms others visually |

In both cases, residual diagnostics confirmed white noise behavior, though forecast intervals were wider post-2020, suggesting increased uncertainty.

Holt-Winters Forecasting

Additive (see figure 50) Holt-Winters forecasts (fit1\_hw) projected modest growth, whereas multiplicative (see figure 51) forecasts (fit2\_hw) declined significantly after 2020. Accuracy measures (figure 52) showed that multiplicative models performed better when dealing with proportional seasonal effects, especially during pandemic years.

Table 2 - Forecast Accuracy Metrics for Seasonal Naive, ETS, and Holt-Winters Models

| **Model Type** | **ME** | **RMSE** | **MAE** | **MAPE (%)** | **MASE** | **ACF1** |
| --- | --- | --- | --- | --- | --- | --- |
| Seasonal Naive | -18,329.74 | 36,280,117 | 26,778,164 | 5.43 | 1.00 | 0.583 |
| ETS (M,N,A) | -5,596.59 | 23,364,510 | 16,630,946 | 3.34 | 0.62 | 0.027 |
| HW Additive | 1,363,693 | 23,301,288 | 16,686,301 | 3.34 | 0.62 | 0.099 |
| HW Multiplicative | 312,842.30 | 24,037,227 | 17,315,837 | 3.49 | 0.65 | 0.147 |

SES Models with Fixed and Optimized Alpha

When comparing fixed α = 0.2 and optimized α via ets() ANN models, results showed minimal differences in training accuracy, but the optimized model showed greater underfitting in the test window (see figure 49). This aligns with expectations given that the optimal alpha chosen was extremely small (1e-4), reducing model responsiveness.

# Discussion

This laboratory highlights the complexity of forecasting in the presence of structural breaks, such as the COVID-19 pandemic, and reinforces the importance of choosing models that accommodate evolving patterns in time series data.

Impact of COVID-19 and Structural Breaks

Comparisons of forecasts generated from 1991–2019 (pre-COVID) with those from 2020–2024 clearly reveal a structural shift. ETS and Holt-Winters models trained exclusively on post-2020 data projected significantly lower ridership, with flatter or declining trends (see figures 50 & 51), indicating a potential long-term impact of the pandemic. This suggests the need for scenario-based forecasting when major disruptions occur.

Trend and Seasonality Dynamics

While regression models — especially cubic — offered solid fits for long-term trends, they failed to adequately model seasonal dynamics. Holt-Winters models, on the other hand, captured both level and seasonality more effectively.

The multiplicative seasonal model (fit2\_hw) demonstrated superior flexibility, especially during volatile periods (see figure 52), but could be sensitive to sudden drops in volume.

ETS models offered robust performance across time periods, with ETS(M,N,A) and ETS(M,N,M) being particularly effective (see figure 43). However, forecasts became less confident post-2020, as seen in the widened confidence intervals.

Evaluation of Simpler Models

Moving averages and seasonal naïve models helped illustrate patterns but lacked predictive adaptability. SES models with fixed alpha values showed reasonable performance, while models with optimized alpha (as low as 1e-4) tended to underfit — consistent with expectations, as such small alpha values yield overly smooth forecasts (see figure 49).

Real-World Implications

Forecasts from models trained on all historical data projected a stronger post-2024 rebound, while those trained on recent data were more conservative. This divergence underscores the importance of data window selection and external context in forecasting, especially when results are used for strategic planning or capacity management.

For transportation agencies like Amtrak, this analysis suggests that hybrid modeling approaches — combining long-term historical patterns with recent behavioral shifts — may provide more balanced, realistic forecasts. This contrast is clearly illustrated in Figure 53, where models trained on all historical data projected a rebound, while post-2020 models remained more conservative.

# Conclusion

This laboratory explored a comprehensive range of forecasting methods applied to Amtrak’s monthly passenger miles data from 1991 through mid-2024. By progressing from regression-based models to smoothing techniques, including exponential smoothing and Holt-Winters methods, the analysis revealed distinct patterns in long-term ridership behavior and model performance across different timeframes.

Polynomial regression models, particularly the cubic form, effectively captured broad nonlinear trends in the pre-pandemic period. However, these models lacked the flexibility to adapt to the pronounced seasonal effects and structural shifts brought by events such as COVID-19. Smoothing techniques, such as moving averages, provided valuable insights into underlying trends but were limited in responsiveness to sudden changes.

Among all the models, the exponential smoothing methods—especially Holt-Winters models—proved the most robust for capturing level, trend, and seasonality. Both additive and multiplicative variants adjusted well to post-pandemic ridership behavior, with multiplicative models showing greater sensitivity to variations in seasonal amplitude tied to overall volume levels. This contrast is especially evident in Figure 52, which overlays the additive and multiplicative Holt-Winters models to highlight their differing responsiveness to recent volatility and changing seasonal patterns. Forecast comparisons highlighted significant differences depending on the training period, emphasizing the impact of major disruptions on future projections.

In summary, this lab underscored the importance of selecting forecasting techniques that align with the structural characteristics of the data. Seasonality, nonlinear growth, and external shocks like the COVID-19 pandemic must be carefully accounted for when building models for real-world forecasting. The applied methods not only deepened understanding of Amtrak’s ridership trends but also enhanced proficiency in R’s time series ecosystem, equipping us with practical tools for future data-driven decision-making.

# Plots

A graph of a flight

AI-generated content may be incorrect.

Figure 1 - Initial Cubic Trend Model Fit. Provides flexible fit to historical growth.

A graph of a flight

AI-generated content may be incorrect.

Figure 2 - Initial Quadratic Trend Model Fit. Early model fails to capture curvature.

A graph of a line graph

AI-generated content may be incorrect.

Figure 3- Cubic Model Fit to Amtrak Ridership. Closely tracks overall trend.

A graph of a graph showing the value of a company

AI-generated content may be incorrect.

Figure 4 - Quadratic Model Fit to Amtrak Ridership. Underfits seasonal peaks.

A graph showing the time of a ride

AI-generated content may be incorrect.

Figure 5 - Full Ridership Time Series Plot. Shows long-term growth and COVID dip.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 6 - Histogram of Monthly Amtrak Passenger Miles (1991–2024).

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 7 - Time Series with Cubic Regression Fit. The cubic model provides better alignment with non-linear ridership patterns.

A graph of a flight

AI-generated content may be incorrect.

Figure 8 - Time Series with Quadratic Regression Fit. The quadratic model captures overall growth but misses turning points.

A graph of a flight

AI-generated content may be incorrect.

Figure 9 - Cubic Polynomial Fit Across Full Time Range. Reveals poor seasonal tracking despite improved trend flexibility.

A graph of a flight miles zoomed

AI-generated content may be incorrect.

Figure 10 - Zoomed Regression Fit (1997–2000). Illustrates fit quality before 9/11, underfitting seasonal dips.

A graph showing the time and the time

AI-generated content may be incorrect.

Figure 11 - Residual Plot for Cubic Model (1991–1997). Shows moderate autocorrelation.

A graph showing the results of a model

AI-generated content may be incorrect.

Figure 12 - Residual Plot for Cubic Model (1991–2004). Slight improvement over earlier period.

A graph showing a number of different colored lines

AI-generated content may be incorrect.

Figure 13 - Residual Plot for Cubic Model (1991–2016). Residuals closer to white noise.

A graph showing a number of data

AI-generated content may be incorrect.

Figure 14 - Residual Plot for Cubic Model (1991–2016). Residuals closer to white noise.

A graph showing a number of different colored lines

AI-generated content may be incorrect.

Figure 15 - Comparison of Residuals Across Polynomial Models. Confirms cubic model superiority.

A graph of a graph showing the number of miles per hour

AI-generated content may be incorrect.

Figure 16 - Cubic Fit (1991–1997). Demonstrates good trend capture but lacks seasonality.

A graph showing the results of a model

AI-generated content may be incorrect.

Figure 17 - Cubic Fit (1991–2004). Broader fit with declining post-2001 slope.

A graph showing a number of different colored lines

AI-generated content may be incorrect.

Figure 18 - Cubic Fit (1991–2016). Shows saturation phase pre-COVID.

A graph of a graph showing a number of miles

AI-generated content may be incorrect.

Figure 19 - Cubic Fit (1991–2020). Pattern collapse after 2020 is evident.

A graph of a graph showing a number of miles per month

AI-generated content may be incorrect.

Figure 20 - Cubic Fit vs Actuals. Model misses peaks due to absence of seasonality.

A graph of a flight

AI-generated content may be incorrect.

Figure 21 - Zoomed Actual vs Fitted (1991–1997). Divergence in seasonal high points.

A graph with blue lines and a red line

AI-generated content may be incorrect.

Figure 22 - Cubic Trend Removal (1991–2016). Prepares series for seasonal decomposition.

A graph with lines on it

AI-generated content may be incorrect.

Figure 23 - ACF After First Differencing. Seasonality still present.

A graph with lines on it

AI-generated content may be incorrect.

Figure 24 - ACF After Second Differencing. Better noise behavior.

A graph of a number of passengers

AI-generated content may be incorrect.

Figure 25 - ACF of Raw Series (1991–2016). High autocorrelation indicates strong seasonality.

A graph of different types of time

AI-generated content may be incorrect.

Figure 26 - Residuals after Deseasoning and Detrending. Minor cycles remain.

A graph of different types of time

AI-generated content may be incorrect.

Figure 27 - Classical Decomposition (1991–2016). Reveals additive seasonal component.

A graph of different types of graphs

AI-generated content may be incorrect.

Figure 28 - Residual Diagnostics (gg\_tsresiduals). Residuals approximately white noise.

A graph of different types of waves

AI-generated content may be incorrect.

Figure 29 - Decomposition of Log-Transformed Series. Clearer residuals.

A graph of different types of time

AI-generated content may be incorrect.

Figure 30 - Decomposition After Log + First Differencing. Improved stationarity.

A graph with lines on it

AI-generated content may be incorrect.

Figure 31 - ACF of Log + Second Differenced Series. Suggests white noise structure.

A graph of different types of time

AI-generated content may be incorrect.

Figure 32 - Decomposition After Log + Second Differencing. Minimal structure remains.

|  |  |
| --- | --- |
|  |  |

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 33 - Log-transformed Time Series. Stabilizes variance.

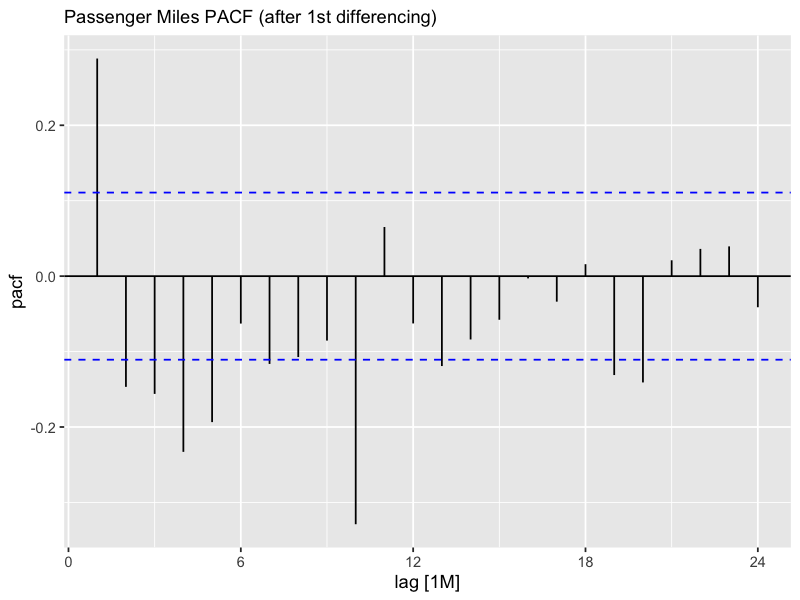


Figure 34 - PACF After First Differencing. Slow decay pattern persists.

A graph showing a graph of time

AI-generated content may be incorrect.

Figure 35 - Second Differenced Series. Smoother residuals for modeling.

A graph showing a number of miles

AI-generated content may be incorrect.

Figure 36 - 3-Month Centered Moving Average. Smooths short-term noise.

A graph of a flight

AI-generated content may be incorrect.

Figure 37 - Seasonal Naïve Forecast. Benchmarks basic seasonal projection.

A graph showing the growth of a number of training

AI-generated content may be incorrect.

Figure 38 - 12-Month MA vs Actual. Validates trend extraction.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 39 - Zoomed Centered MA. Highlights seasonal consistency.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 40 - Zoomed Trailing MA. Shows lag effect.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 41 - ETS(M,N,A) Fit (1991–2019). Captures consistent seasonal trend.

A graph of a graph showing the number of miles

AI-generated content may be incorrect.

Figure 42 - ETS(M,N,M) Fit (2020–2024). Flattened trajectory post-COVID.

A graph showing a wave of time

AI-generated content may be incorrect.

Figure 43 - ETS Forecast (1991–2019). Moderate upward trend.

A graph showing the time and time

AI-generated content may be incorrect.

Figure 44 - ETS Forecast (2020–2024). Seasonality persists, no trend.

A graph showing a graph of a graph

AI-generated content may be incorrect.

Figure 45 - ETS Forecast vs Actual (1991–2019). Good alignment pre-COVID.

A graph showing a graph of different types of graphs

AI-generated content may be incorrect.

Figure 46 - ETS Forecast vs Actual (2020–2024). Captures recent flattening.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 47 - Holt-Winters Forecasts (2020–2024). Overlay of both models.

A graph of a flight

AI-generated content may be incorrect.

Figure 48 - Holt-Winters Additive Forecast. Smooth post-COVID recovery trend.

A graph of a flight

AI-generated content may be incorrect.

Figure 49 - HW Additive vs Multiplicative Comparison. Highlights difference in seasonality scaling.

A graph of a passenger miles

AI-generated content may be incorrect.

Figure 50 - Additive vs Multiplicative Behavior Over Full Range. Multiplicative model more reactive.

A graph of different colored lines

AI-generated content may be incorrect.

Figure 51 - Post-2024 Holt-Winters Forecasts. Projections diverge.

A graph of a number of miles

AI-generated content may be incorrect.

Figure 52 - Holt-Winters Multiplicative Forecast. Forecasts decline in amplitude.

A graph showing the time and training

AI-generated content may be incorrect.

Figure 53 - Twice-Differenced Ridership Series with Training, Validation, and Forecast Segments

# References

1. Hyndman, R. J., & Athanasopoulos, G. (2021). \*Forecasting: Principles and Practice\* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
2. Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. \*International Journal of Forecasting, 22\*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001
3. Purdue OWL. (2019). APA formatting and style guide (7th ed.). <https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/general_format.html>
4. Wang, E., Hyndman, R. J., & others. (2023). *fable, feasts, tsibble, and forecast: Tidy time series forecasting packages for R* [R packages]. <https://CRAN.R-project.org/>

# APPENDICES

Appendix A: Full R Script

The complete R code used to perform the analysis, modeling, and visualization for this laboratory is available online at the following GitHub repository: [**Lab3.R**](https://github.com/anbazhaganjr/ana535_forecasting/blob/e8568bf7c544048a8734c11de3ca621f5e4ae6b0/laboratory3/Laboratory3.R)

Appendix B: Output Log File

All forecast model fitting, accuracy metrics (MAPE, RMSE, MAE), and diagnostic outputs were logged throughout execution and stored for validation. [**LogFile**](https://github.com/anbazhaganjr/ana535_forecasting/blob/e8568bf7c544048a8734c11de3ca621f5e4ae6b0/laboratory3/ana535_lab3_output_log.txt)