TIME SERIES FORECASTING OF ALASKA AIRLINES REVENUE & PANDEMIC IMPACT

Project Report

TABLE OF CONTENTS

SUMMARY

This project aimed to assess the financial impact of the COVID-19 pandemic on Alaska Airlines' revenue using time series analysis techniques. Quarterly revenue data from 2009 to 2023, sourced from https://www.macrotrends.net/stocks/charts/ALK/alaska-air/revenue, underwent preprocessing to ensure compatibility with R programming for analysis.

The predictability of the dataset was evaluated using two approaches: Approach 1 involved hypothesis testing with an AR(1) model, indicating randomness and unpredictability. Conversely, Approach 2 examined the autocorrelation function (ACF) for the first differenced series, revealing statistical significance and predictability.

Subsequently, the revenue dataset was partitioned into Training (2009 Q1 – 2016 Q4) and Validation Datasets (2017 Q1 – 2019 Q4). Various time series models were constructed and compared using distinct accuracy metrics to forecast revenue for the COVID (2020 Q1 – 2021 Q4) and post-COVID (2022 Q1 – 2023 Q4) periods. Models included Naïve, Seasonal Naïve, Two-Level Forecasting with Linear Trend & Trailing MA for residuals, Two-Level Forecasting with Quadratic Trend & Trailing MA for residuals, Automated Holt-Winter's Model, Regression Models with Linear Trend, Quadratic Trend, Seasonality, and combinations thereof, Two-Level Forecasting with Automated Holt-Winter's Model & AR(1) Model, and an Automated ARIMA Model.

After thorough analysis, two models emerged as top contenders based on accuracy metrics: (1) Two-Level Forecasting with Linear Trend & Trailing MA for residuals and (2) Two-Level

Forecasting with Quadratic Trend & Trailing MA for residuals. However, the model that stood out as the most parsimonious and accurate was the Two-Level Trailing MA Model with Linear Trend and Seasonality, exhibiting the lowest Mean Absolute Percentage Error (MAPE) compared to other models, albeit with a slightly higher Root Mean Square Error (RMSE) than its quadratic counterpart.

Forecast Results:

Utilizing the top two models, the projected loss of revenue for Alaska Airlines was estimated. For the linear model, the total loss of revenue amounted to $24,870.24, representing a percentage loss of 153.4631% compared to the forecasted revenue. Meanwhile, for the quadratic model, the total loss of revenue amounted to $25,259.7, representing a percentage loss of 155.8664% compared to the forecasted revenue.

CHAPTER – I

GENERAL INTRODUCTION OF THE PROJECT

1.1 Introduction

The COVID-19 pandemic has had a profound impact on various industries worldwide. Among the hardest-hit sectors airlines, hospitality, tourism, retail, entertainment etc. which faced significant revenue declines due to travel restrictions, lockdown measures, and changes in consumer behavior. This project aimed to analyze the impact of the pandemic on revenue for "Alaska Airlines". By leveraging time series analysis techniques using historical revenue data spanning from 2009 to 2023, we seek to understand the extent of revenue fluctuations during the pandemic period and evaluate the recovery trends post-COVID-19 crisis.

1.2 Objectives

1.2.1 Determination of the best Time Series Analysis Model

Apply various Time Series Analysis techniques on the historical revenue data for Alaska Airlines considering the dataset from 2009-2016 as training data and the dataset from 2017-2019 as validation data, and then determining the best model based on accuracy metrics.

1.2.2 Forecast Revenue for Covid and Post-Covid Period

Utilize the chosen best Time Series Forecasting Model to predict revenue for the future period, including the COVID (2020-2021) and post-COVID (2022-2023) periods.

1.2.3 Comparison of Forecasted Revenue vs. Actual Revenue

Compare the forecasted revenue with the actual revenue data to quantify the impact of the pandemic on revenue generation.

1.3 Project Scope

Time Period: The dataset represents quarterly revenues (in $million) in Alaska Airlines from the first quarter of 2009 through the fourth quarter of 2023, with a focus on the COVID period (2020-2021) and post-COVID period (2022-2023).

Data Sources: This quarterly revenue data of Alaska Airlines is collected from https://www.macrotrends.net/stocks/charts/ALK/alaska-air/revenue

Methodology: Time series analysis techniques, including identifying time series components, model selection based on accuracy metrics, forecasting, and comparison with actual data, will be employed to achieve the project objectives.

Industry Focus: The project will primarily focus on Alaska Airlines to assess the impact of the pandemic on revenue generation.

Through this analysis, insights into the financial implications of the COVID-19 pandemic on the revenue for Alaska Airlines can be obtained. This will facilitate informed decision-making and strategic planning in response to future crises.

CHAPTER – II

METHODOLOGY

2.1 Data Preparation

The initial dataset comprised "Time" data in Date format and "Revenue" in Currency format (with $ as prefix). To conduct analysis, the dataset was transformed into a CSV file format with Time data represented in Quarter format (e.g., 3/31/2009 converted to Q1-09) and Revenue data in numerical format (e.g., $1,110 converted to 1110).

For the training and validation phases, the pre-COVID-19 period was segmented. Specifically, data from 2009 to 2016 was allocated for training, while data from 2017 to 2019 was designated for validation. Subsequently, the COVID-19 and post-COVID-19 periods, spanning from 2020 to 2023, were reserved for future period analysis.

CHAPTER – III

FORECASTING PROCESS & ANALYSIS

3.1 Forecasting Process

The model development and forecasting process followed an eight-step approach as depicted in the Fig 3.1. This approach ensures robustness and accuracy.

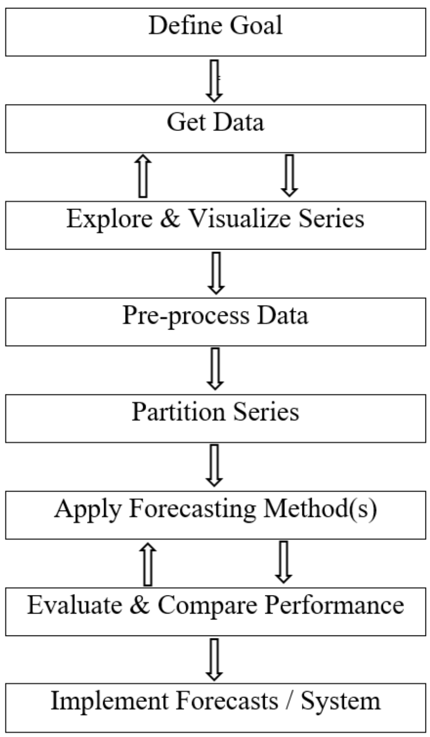

Fig 3.1 8-Step Forecasting Process Flow Chart

3.1.1 Step 1: Define Goal

Goals : The goals encompassed descriptive analysis, including autocorrelation examination, time series component identification, and data visualization. Additionally, predictive analysis aimed to forecast revenue for 2020-2023 under pre-COVID conditions and assess revenue loss by comparing forecasted and actual revenue data.

Forecast Horizon : The forecast horizon extends to the years 2020 Q1 – 2023 Q4 and involves single time point analysis.

Forecast Usage : Stakeholders include executives, financial analysts, and operational managers within Alaska Airlines. The forecast will be numerical, aiming to predict revenue fluctuations.

Forecast Expertise & Automation : Forecasting will be conducted as a one-time task to assess the impact of the pandemic on revenue. Historical revenue records and software like R or Python for time series analysis will be utilized. Implications include selecting appropriate software and ensuring sufficient historical data; given the quarterly dataset, at least 5-6 years of revenue records are needed, totaling 20-24 datapoints. The process involves model development, selecting the best model based on accuracy metrics and simplicity, and deploying the chosen model effectively to forecast revenue.

3.1.2 Step 2: Data Collection

Data quality: The accuracy and reliability of quarterly revenue data of Alaska Airlines from www.macrotrends.net was crucial for robust forecasting. Transformation of data from date to quarterly format was essential to maintain the reliability of revenue figures, particularly given the financial nature of the dataset.

Temporal frequency: Quarterly revenue data provides a balanced view of trends and fluctuations, but adjusting to monthly or weekly data may offer more detailed insights at the cost of increased complexity.

Series granularity: Being a financial forecast, the quarterly granularity for revenue aligns with the forecasting goal.

Domain expertise: Integrating expert knowledge of the aviation industry enriches forecasting by capturing nuanced factors influencing revenue dynamics during the COVID and post-COVID periods (2020-2023) for Alaska Airlines.

3.1.3 Step 3: Explore & Visualize Series

3.1.3.1 Various Time Series Datasets

Name of the series	Duration	Remarks
allrevenue.ts	2009 Q1 – 2023 Q4	Entire Dataset which includes training, validation & future (COVID, post-COVID periods) partitions
revenue.ts	2009 Q1 – 2019 Q4	Pre-COVID period dataset which includes training and validation partitions
future.ts	2020 Q1 – 2023 Q4	Period of interest for which forecasting is done i.e., COVID & post-COVID periods

3.1.3.2 Time Series Dataset *allrevenue.ts* in R

```
> allrevenue.ts <- ts(Alaska.data$Revenue, start = c(2009,1), end = c(2023,4), freq = 4)
> allrevenue.ts
     Qtr1 Qtr2 Qtr3 Qtr4
2009  742  844  967  846
2010  830  976 1068  958
2011  965 1110 1198 1045
2012 1039 1214 1272 1132
2013 1133 1256 1557 1210
2014 1222 1375 1465 1306
2015 1269 1437 1515 1377
2016 1347 1494 1566 1518
2017 1740 2102 2110 1942
2018 1832 2156 2212 2064
2019 1876 2288 2389 2228
2020 1636  421  701  808
2021  797 1527 1953 1899
2022 1681 2658 2828 2479
2023 2196 2838 2839 2553
> |
```

3.1.3.3 Plot of Historical Data & Data Patterns

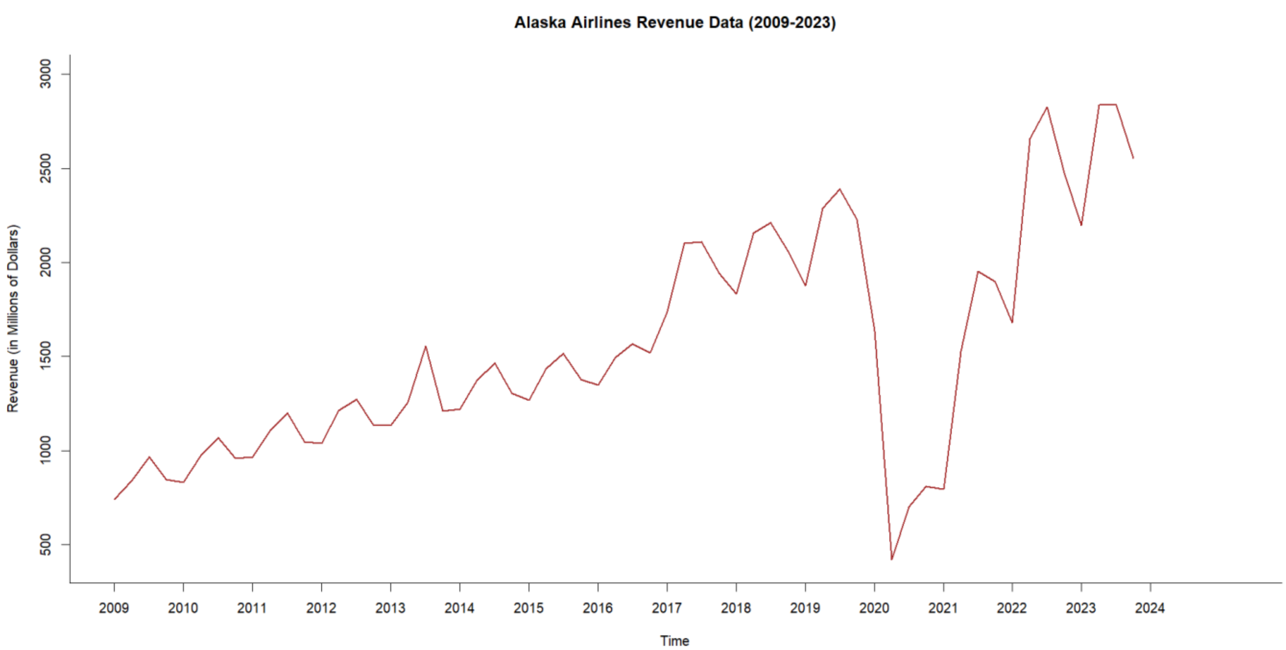

Fig 3.2 Plot of Time Series Data

The visualization in Fig 3.2 illustrates a notable decline in quarterly revenue for Alaska Airlines from 2009 to 2023, particularly evident during the COVID-19 pandemic (2020-2021). This decline is in line with expectations due to travel restrictions and lockdown measures aimed at mitigating the spread of the virus. The clear drop in revenue underscores the importance of this project in deploying time series analysis techniques to quantify its effects on Alaska Airlines' revenue.

3.1.3.4 Visualization of Time Series Components

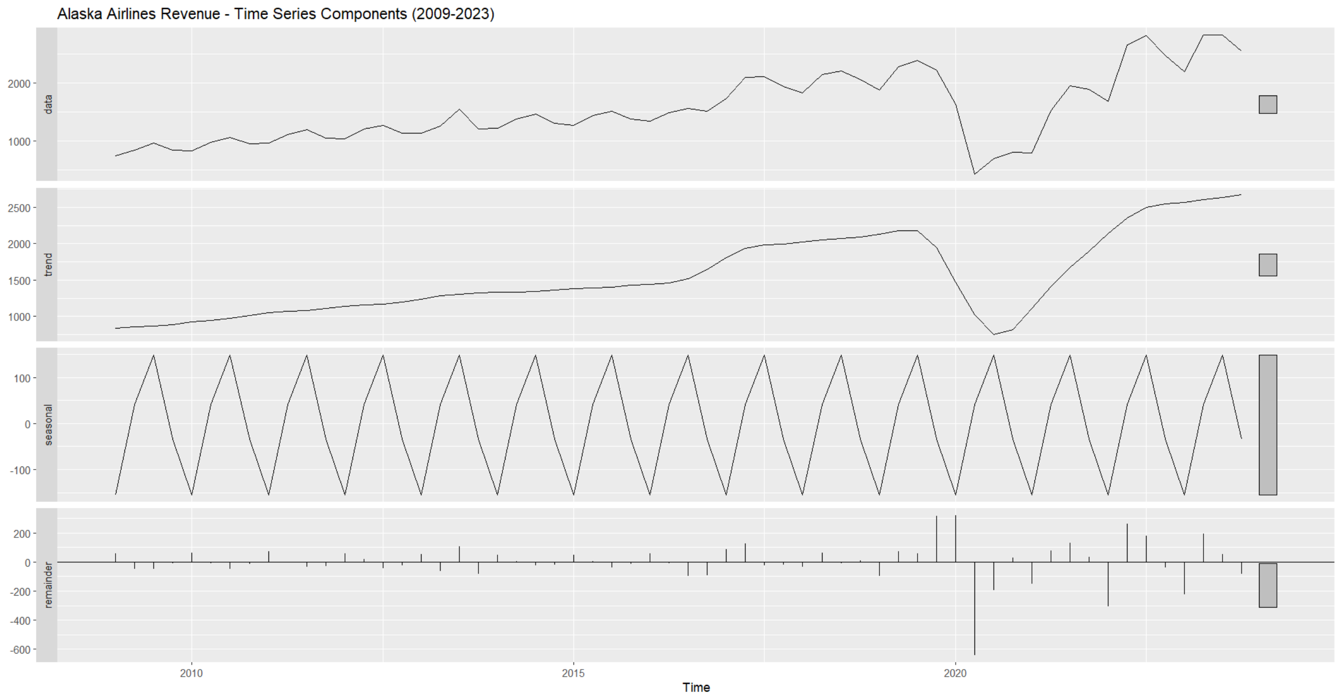

Fig 3.3 Plot of Time Series Components

From Fig 3.3, below inferences are obtained about the time series components for Revenue data of Alaska Airlines

Component	Interpretation
Level	The data in the given time series dataset depicts increasing revenue over time (pre-COVID period) Note : All time series have level
Trend	Since the data exhibits positive growth in the pre-COVID period and post-COVID period, the trend is **upward**
Seasonality	Sales appear to have **cyclical pattern** throughout the year; hence it has **seasonality component**. Pattern observed is lower revenue in the beginning and the end of the year; and comparatively higher revenue in the middle quarters of the year.

3.1.3.5 Interpretation of Autocorrelation of Revenue for pre-COVID period

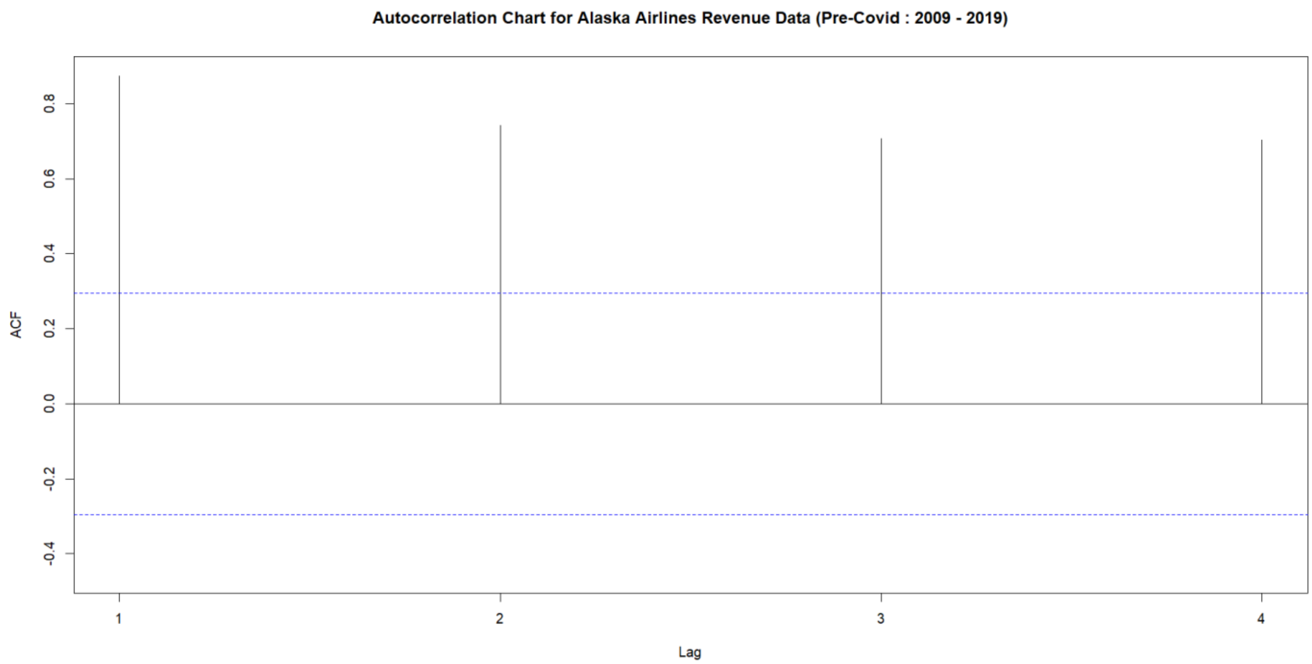

Fig 3.4 Autocorrelation Chart for Alaska Airlines Revenue Data in the pre-COVID period

```
  lag   ACF
1   0 1.000
2   1 0.873
3   2 0.743
4   3 0.707
5   4 0.703
>
```

Component	Lag	ACF	Inference
Trend	1	0.873	Strong positive correlation between consecutive quarters, suggesting a significant trend in revenue
Seasonality	4	0.703	Though ACF is slightly lower than at lag1, it indicates moderate positive correlation, implying seasonality with revenue patterns repeating every four quarters

Hence, the series *revenue.ts* is a series with trend and seasonality.

3.1.4 Step 4: Pre-Process Data

Given the equally spaced nature of the revenue data in this dataset and the absence of extreme or missing values, there was no need for imputation or omission procedures. Yet, to evaluate the predictability of the dataset, two tests were conducted.

Test 1 : Hypothesis Testing using AR(1) model

Summary

```
Series: revenue.ts
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1       mean
      0.9580  1461.2319
s.e.  0.0417   408.7861

sigma^2 = 27617:  log likelihood = -287.64
AIC=581.27   AICc=581.87   BIC=586.62

Training set error measures:
                   ME     RMSE      MAE       MPE    MAPE      MASE         ACF1
Training set 27.30426 162.3618 130.4274 0.7879967 9.31114 0.9373148 -0.04659338
>
```

Structure	Autoregressive Model of Order 1
Parameters	Intercept (α) = 1461.2319 Coefficient of Y_{t-1} (β_1) = 0.9580
Model Equation	*Output Variable,* $Y_t = 1461.2319 + 0.9580\ Y_{t-1}$
Inferences	• Coefficient indicates strong positive autocorrelation; significant • Non-zero mean; indicating the expected value when lagged value is zero.

Hypothesis Testing

```
> # APPLY Z-TEST TO TEST THE NULL HYPOTHESIS THAT BETA COEFFICIENT OF AR(1) = 1
> ar1 <- 0.9580
> s.e. <- 0.0417
> null_mean <- 1
> alpha <- 0.05
> z.stat <- (ar1-null_mean)/s.e.
> z.stat
[1] -1.007194
> p.value <- pnorm(z.stat)
> p.value
[1] 0.1569207
> if (p.value<alpha) {
+   "Reject null hypothesis"
+ } else {
+   "Accept null hypothesis"
+ }
[1] "Accept null hypothesis"
>
```

z-stat = -1.007194	p value = 0.1569207
alpha = 0.05	p value > alpha
Hence, Accept Null Hypothesis	Historical data could be random walk and could be hard to predict

Test 2 : Examination of Auto Correlation Function for First Differenced Series

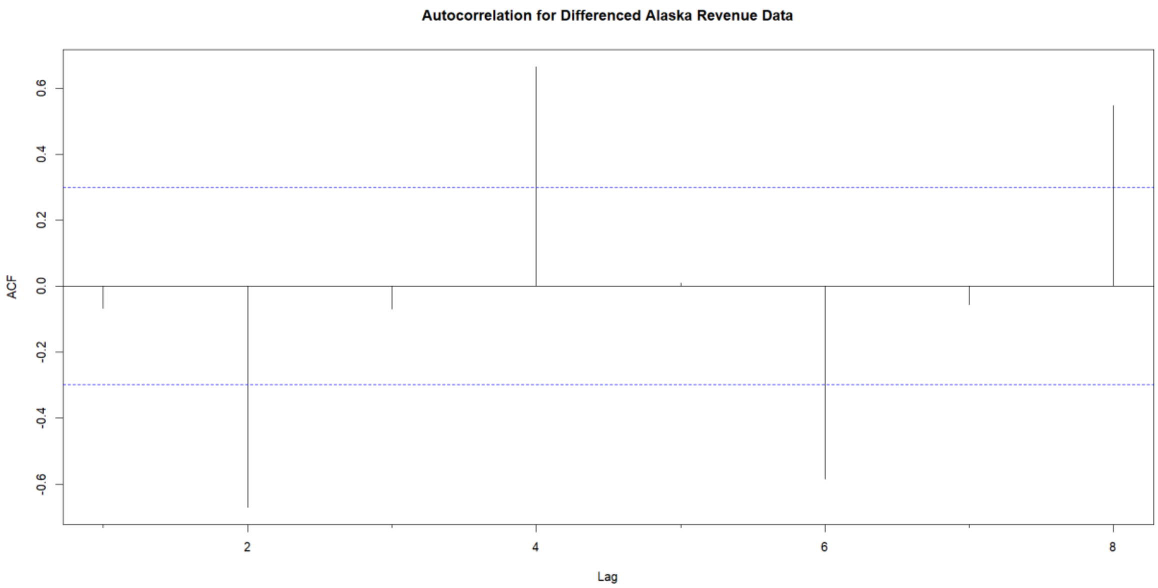

Fig 3.5 Autocorrelation Chart for First Differenced Revenue Data of Alaska Airlines in the pre-COVID period

Since at all levels of lag, the legs are beyond the threshold limit, indicating statistical significance, the **dataset is predictable**.

3.1.5 Step 5: Partition of Time Series

Partition	Series Name	Duration	No. of Data Points
Training	*train.ts*	2009 Q1 – 2016 Q4	8 x 4 = 32 data points
Validation	*valid.ts*	2017 Q1 – 2019 Q4	3 x 4 = 12 data points
Future (COVID, post-COVID periods)	*future.ts*	2020 Q1 – 2023 Q4	4 x 4 = 16 data points

```
> train.ts
     Qtr1 Qtr2 Qtr3 Qtr4
2009  742  844  967  846
2010  830  976 1068  958
2011  965 1110 1198 1045
2012 1039 1214 1272 1132
2013 1133 1256 1557 1210
2014 1222 1375 1465 1306
2015 1269 1437 1515 1377
2016 1347 1494 1566 1518
```

```
> valid.ts
     Qtr1 Qtr2 Qtr3 Qtr4
2017 1740 2102 2110 1942
2018 1832 2156 2212 2064
2019 1876 2288 2389 2228
>
```

```
> future.ts
     Qtr1 Qtr2 Qtr3 Qtr4
2020  742  844  967  846
2021  830  976 1068  958
2022  965 1110 1198 1045
2023 1039 1214 1272 1132
>
```

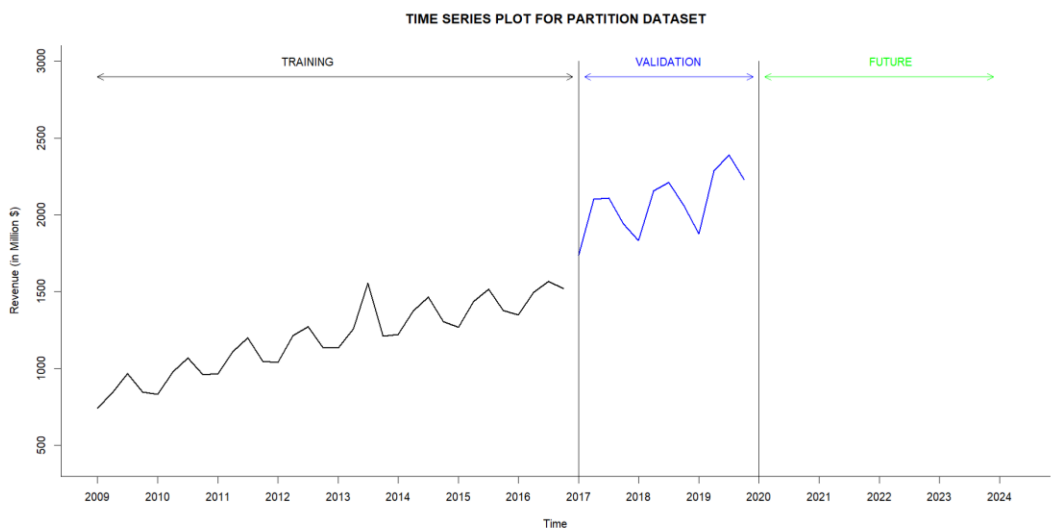

Fig 3.6 Data Partition Graph for Alaska Airlines Revenue Data in the pre-COVID period

3.1.6 Step 6: Apply Forecasting Methods

3.1.6.1 Model 1 : Naïve Model

<u>Forecast of Revenue for the period 2020 - 2023</u>

```
> revenuef.naive.pred$mean
     Qtr1 Qtr2 Qtr3 Qtr4
2020 2228 2228 2228 2228
2021 2228 2228 2228 2228
2022 2228 2228 2228 2228
2023 2228 2228 2228 2228
>
```

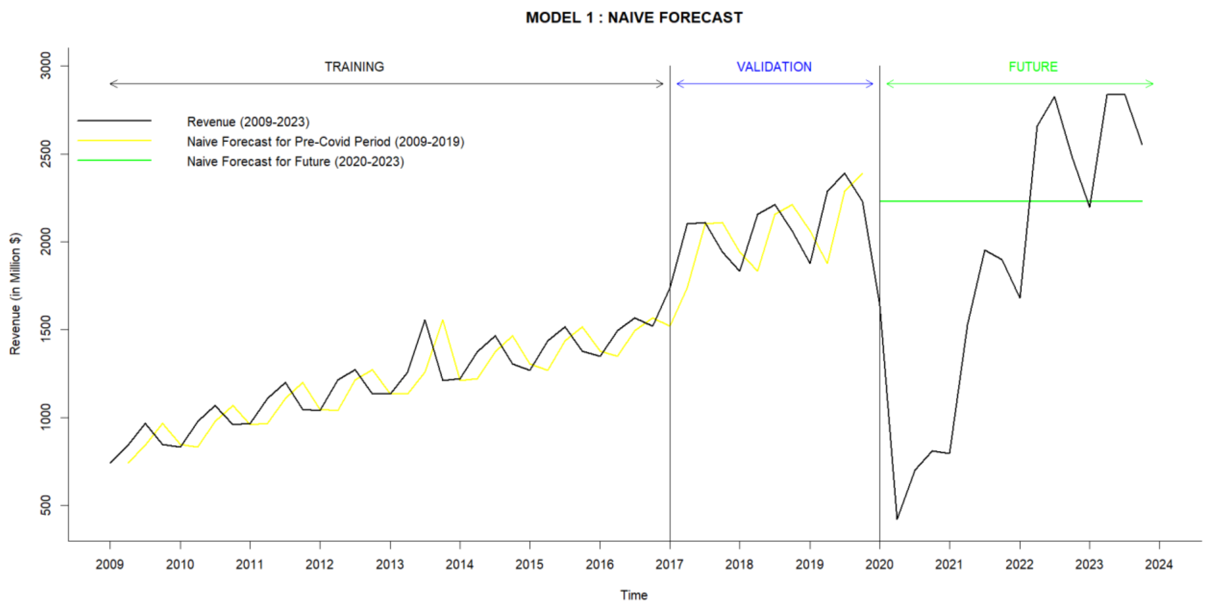

Fig 3.7 Naïve Forecast (green line) of Revenue for Alaska Airlines

3.1.6.2 Model 2 : Seasonal Naïve Model

<u>Forecast of Revenue for the period 2020 – 2023</u>

```
> revenue.snaive.pred$mean
     Qtr1 Qtr2 Qtr3 Qtr4
2017 1347 1494 1566 1518
2018 1347 1494 1566 1518
2019 1347 1494 1566 1518
>
```

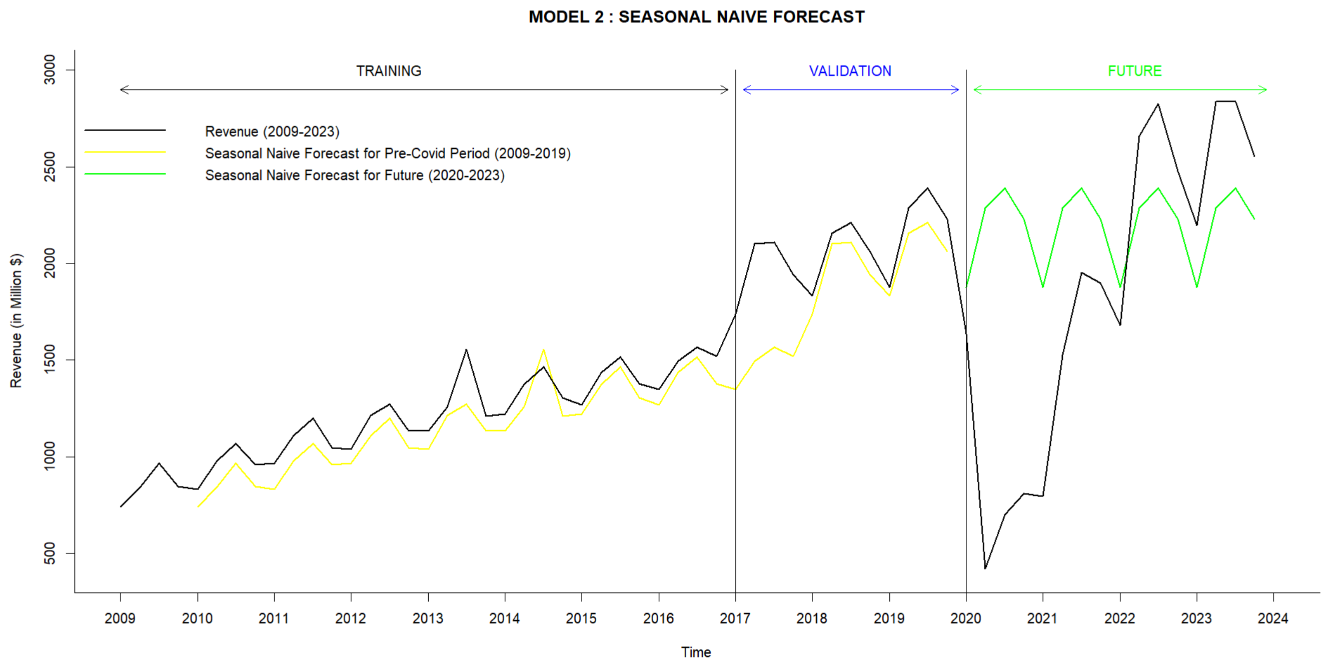

Fig 3.8 Seasonal Naïve Forecast (green line) of Revenue for Alaska Airlines

3.1.6.3 Model 3 : Two-Level Model (Regression + Trailing Moving Average models)

Model 3a. Regression Model with Linear Trend & Seasonality + Trailing MA model

<u>Forecast of Revenue for the period 2020 – 2023</u>

For window width k = 2

```
> future_2.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2067.873              148     2215.873
2        2273.055              148     2421.055
3        2370.055              148     2518.055
4        2216.145              148     2364.146
5        2200.473              148     2348.473
6        2405.655              148     2553.655
7        2502.655              148     2650.655
8        2348.745              148     2496.746
9        2333.073              148     2481.073
10       2538.255              148     2686.255
11       2635.255              148     2783.255
12       2481.345              148     2629.346
13       2465.673              148     2613.673
14       2670.855              148     2818.855
15       2767.855              148     2915.855
16       2613.945              148     2761.946
>
```

For window width k = 3

```
> future_3.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2067.873          147.842     2215.714
2        2273.055          147.842     2420.896
3        2370.055          147.842     2517.896
4        2216.145          147.842     2363.987
5        2200.473          147.842     2348.314
6        2405.655          147.842     2553.496
7        2502.655          147.842     2650.496
8        2348.745          147.842     2496.587
9        2333.073          147.842     2480.914
10       2538.255          147.842     2686.096
11       2635.255          147.842     2783.096
12       2481.345          147.842     2629.187
13       2465.673          147.842     2613.514
14       2670.855          147.842     2818.696
15       2767.855          147.842     2915.696
16       2613.945          147.842     2761.787
>
```

For window width k = 4

```
> future_4.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2067.873          102.349     2170.222
2        2273.055          107.374     2380.429
3        2370.055          111.394     2481.448
4        2216.145          114.610     2330.755
5        2200.473          117.182     2317.655
6        2405.655          119.241     2524.895
7        2502.655          120.887     2623.542
8        2348.745          122.204     2470.950
9        2333.073          123.258     2456.331
10       2538.255          124.101     2662.356
11       2635.255          124.776     2760.030
12       2481.345          125.315     2606.661
13       2465.673          125.747     2591.419
14       2670.855          126.092     2796.947
15       2767.855          126.368     2894.223
16       2613.945          126.589     2740.535
>
```

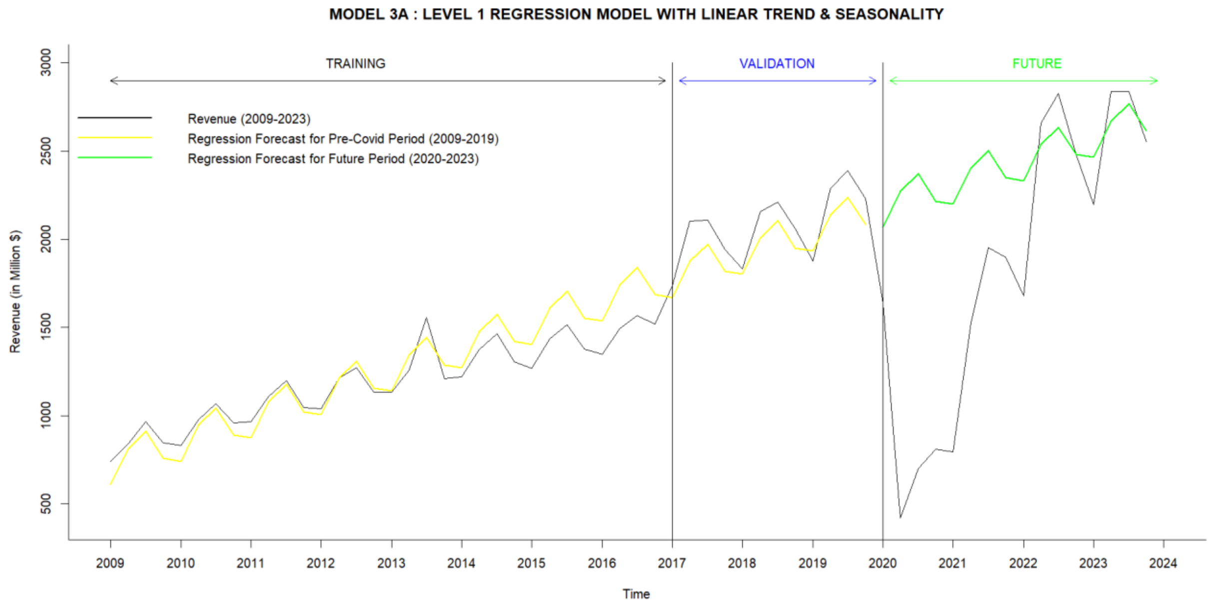

Fig 3.9 Regression (Linear Trend & Seasonality) Model Forecast (green line) of Revenue for Alaska Airlines

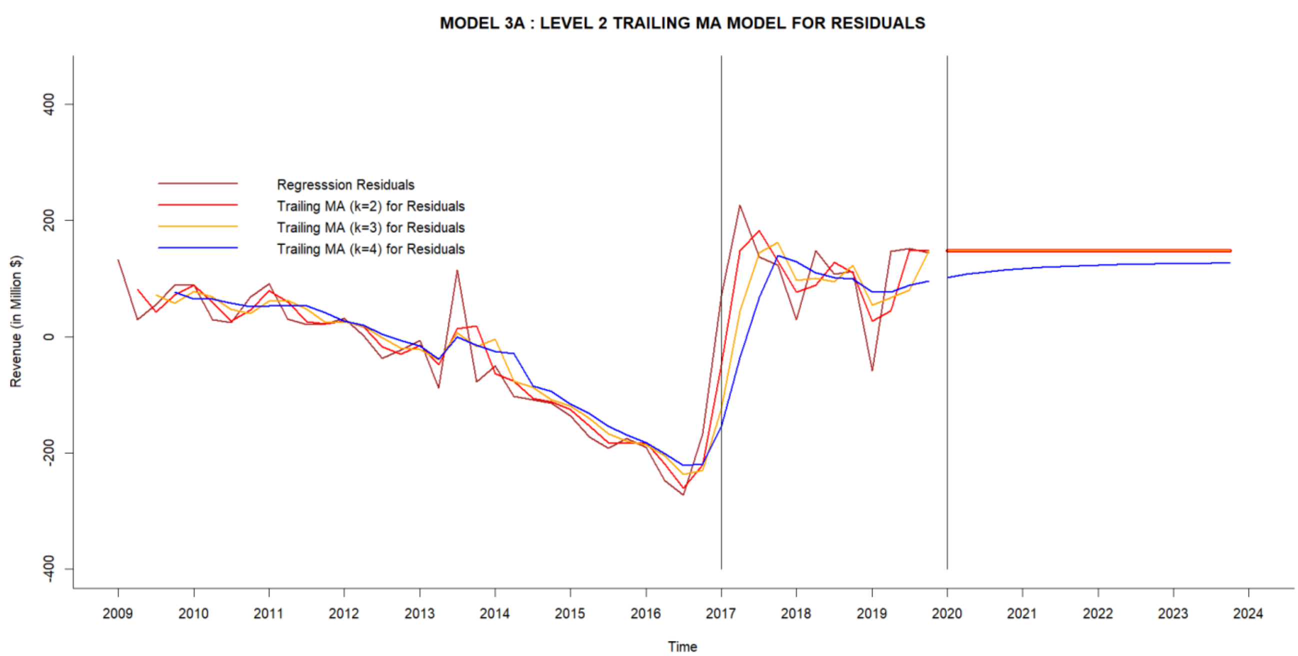

Fig 3.10 Trailing MA Model's Residual Forecast of Revenue for Alaska Airlines

Model 3b. Regression Model with Quadratic Trend & Seasonality + Trailing MA model

<u>Forecast of Revenue for the period 2020 – 2023</u>

For window width k = 2

```
> futureq_2.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2252.127           -1.972     2250.155
2        2483.019           -1.972     2481.047
3        2605.729           -1.972     2603.757
4        2477.529           -1.972     2475.557
5        2489.709           -1.972     2487.737
6        2724.886           -1.972     2722.914
7        2851.881           -1.972     2849.909
8        2727.967           -1.972     2725.994
9        2744.431           -1.972     2742.459
10       2983.893           -1.972     2981.921
11       3115.173           -1.972     3113.201
12       2995.544           -1.972     2993.571
13       3016.293           -1.972     3014.321
14       3260.040           -1.972     3258.068
15       3395.605           -1.972     3393.633
16       3280.260           -1.972     3278.288
>
```

For window width k = 3

```
> futureq_3.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2252.127            8.582     2260.709
2        2483.019            8.582     2491.601
3        2605.729            8.582     2614.311
4        2477.529            8.582     2486.111
5        2489.709            8.582     2498.291
6        2724.886            8.582     2733.468
7        2851.881            8.582     2860.463
8        2727.967            8.582     2736.548
9        2744.431            8.582     2753.013
10       2983.893            8.582     2992.475
11       3115.173            8.582     3123.755
12       2995.544            8.582     3004.125
13       3016.293            8.582     3024.875
14       3260.040            8.582     3268.622
15       3395.605            8.582     3404.187
16       3280.260            8.582     3288.842
>
```

For window width k = 4

```
> futureq_4.df
   Regression.Fst MA.Residuals.Fst Combined.Fst
1        2252.127          -42.911     2209.216
2        2483.019          -51.255     2431.763
3        2605.729          -57.931     2547.798
4        2477.529          -63.271     2414.258
5        2489.709          -67.544     2422.166
6        2724.886          -70.961     2653.924
7        2851.881          -73.696     2778.185
8        2727.967          -75.883     2652.083
9        2744.431          -77.633     2666.798
10       2983.893          -79.033     2904.860
11       3115.173          -80.153     3035.020
12       2995.544          -81.049     2914.495
13       3016.293          -81.766     2934.528
14       3260.040          -82.339     3177.701
15       3395.605          -82.798     3312.807
16       3280.260          -83.165     3197.096
>
```

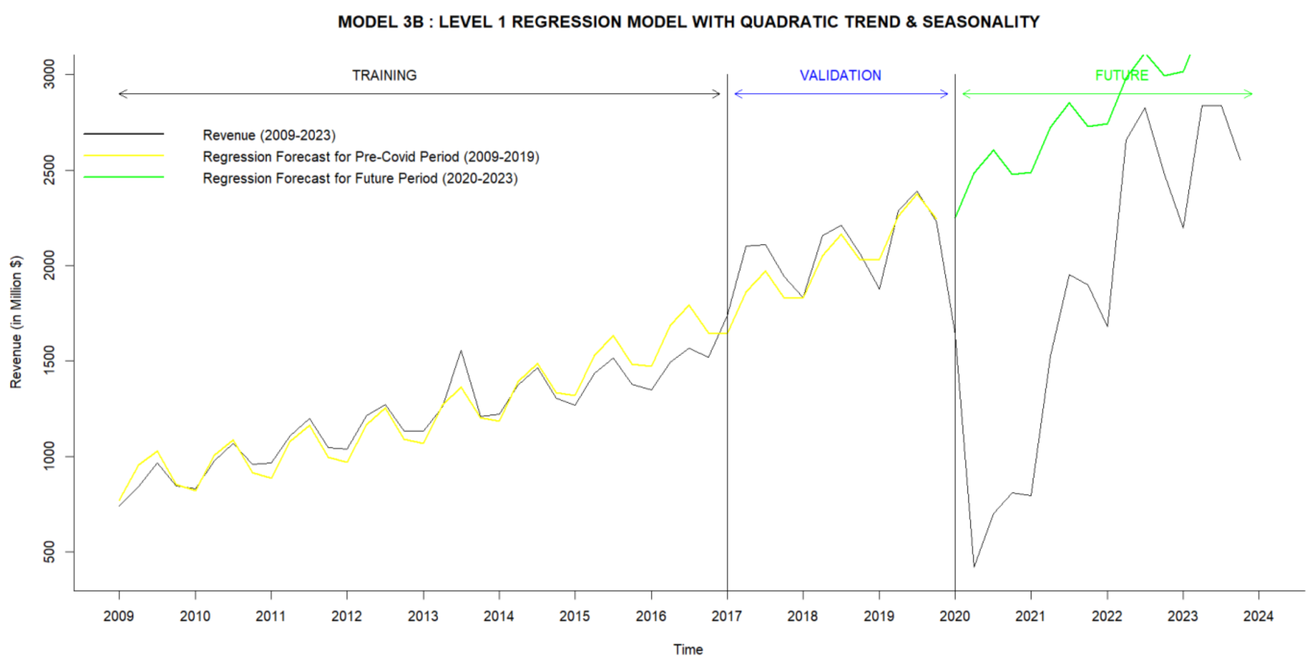

Fig 3.11 Regression (Quadratic Trend & Seasonality) Model Forecast (green line) of Revenue for Alaska Airlines

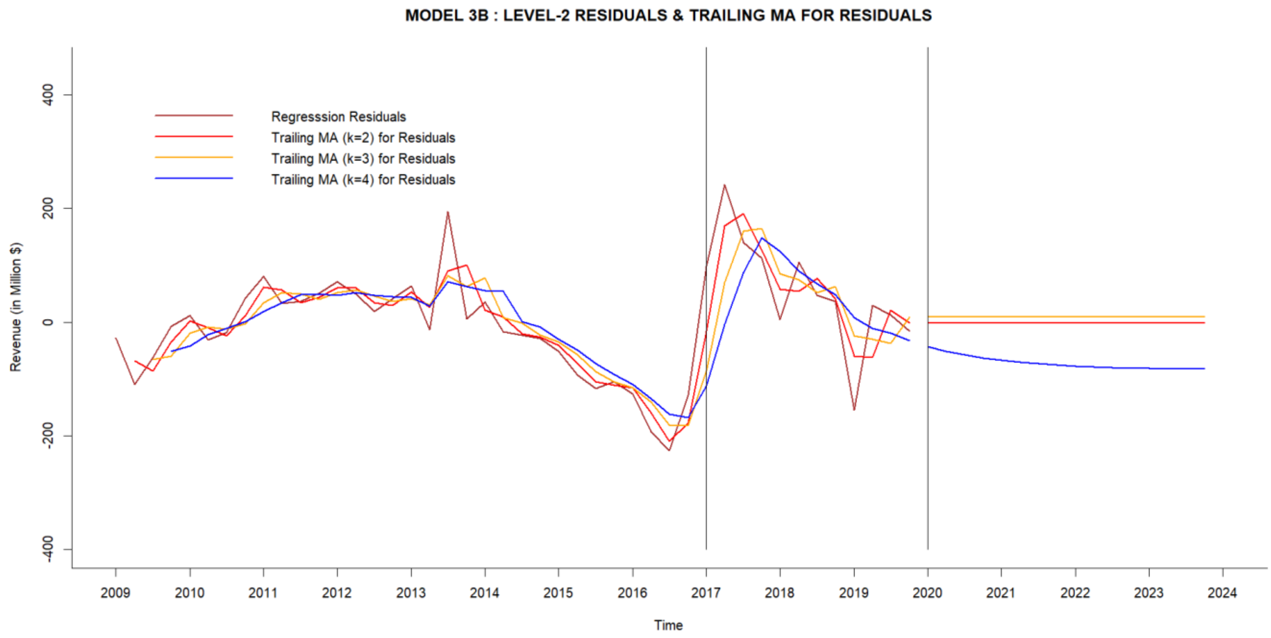

Fig 3.12 Trailing MA Model's Residual Forecast of Revenue for Alaska Airlines

3.1.6.4 Model 4 : Automated Holt-Winter's Model

<u>Summary</u>

```
ETS(M,A,M)

Call:
 ets(y = revenue.ts, model = "ZZZ")

  Smoothing parameters:
    alpha = 0.7178
    beta  = 3e-04
    gamma = 0.0022

  Initial states:
    l = 772.6929
    b = 34.1705
    s = 0.9556 1.0884 1.0346 0.9213

  sigma:  0.0547

     AIC     AICc      BIC
555.6601 560.9542 571.7178
```

Model	**MAM**
Error	Multiplicative
Trend	Additive
Seasonality	Multiplicative

Smoothing parameter for level, Alpha	0.7178	High value Model gives **more weight** to the most recent revenue data while updating the level component
Smoothing parameter for trend, Beta	0.0003	Extremely low value Model gives **very low weight** to the most recent data while updating the trend component Indicates relatively stable trend
Smoothing parameter for seasonality, Gamma	0.0022	Very low value Model gives **low weight** to the most recent data while updating the season component Indicates relatively stable season

<u>Forecast of Revenue for the period 2020 – 2023</u>

```
> HW.ZZZ.pred
        Point Forecast     Lo 0     Hi 0
2020 Q1       2154.844 2154.844 2154.844
2020 Q2       2455.359 2455.359 2455.359
2020 Q3       2619.909 2619.909 2619.909
2020 Q4       2332.804 2332.804 2332.804
2021 Q1       2280.831 2280.831 2280.831
2021 Q2       2596.849 2596.849 2596.849
2021 Q3       2768.737 2768.737 2768.737
2021 Q4       2463.467 2463.467 2463.467
2022 Q1       2406.819 2406.819 2406.819
2022 Q2       2738.339 2738.339 2738.339
2022 Q3       2917.565 2917.565 2917.565
2022 Q4       2594.131 2594.131 2594.131
2023 Q1       2532.808 2532.808 2532.808
2023 Q2       2879.830 2879.830 2879.830
2023 Q3       3066.395 3066.395 3066.395
2023 Q4       2724.796 2724.796 2724.796
>
```

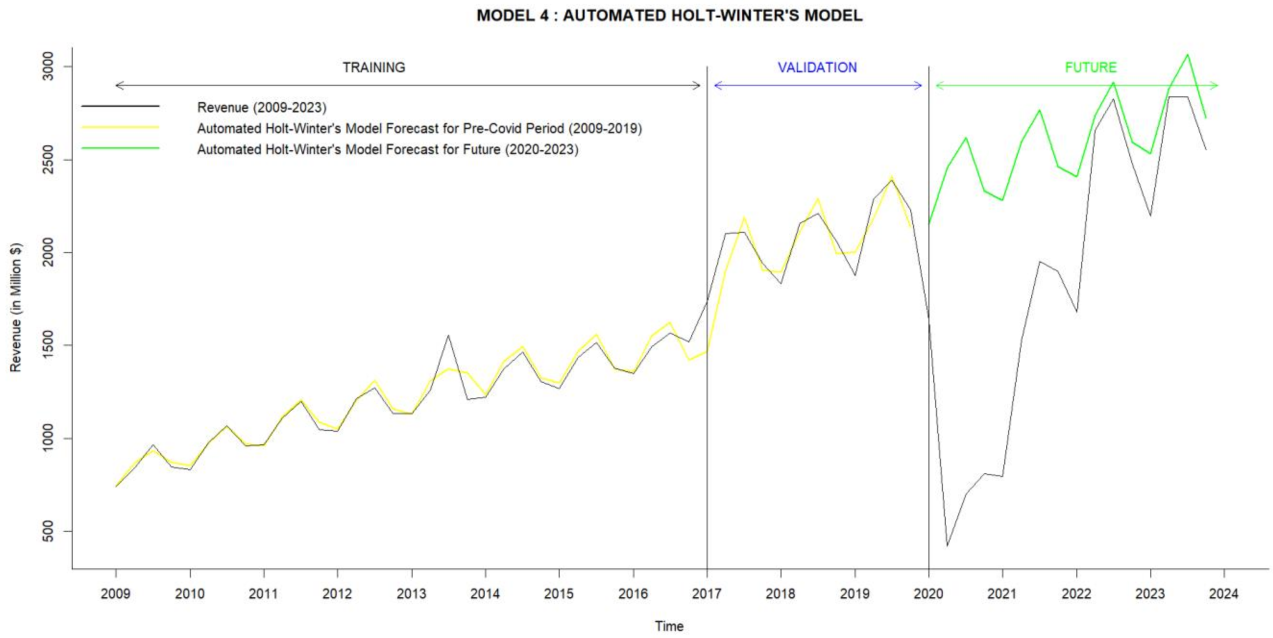

Fig 3.13 Automated Holt-Winter's Model Forecast (green line) of Revenue for Alaska Airlines

3.1.6.5 Model 5 : Regression Models

Model 5a. Regression Model with Linear Trend

<u>Summary</u>

```
Call:
tslm(formula = revenue.ts ~ trend)

Residuals:
     Min       1Q   Median       3Q      Max 
-305.717 -116.823   -8.299   90.357  282.718 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  686.638     48.167   14.26   <2e-16 ***
trend         33.313      1.864   17.87   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 157 on 42 degrees of freedom
Multiple R-squared:  0.8837,	Adjusted R-squared:  0.881 
F-statistic: 319.3 on 1 and 42 DF,  p-value: < 2.2e-16
```

Structure	Regression Model with Linear trend
Parameter	Intercept (β_0) = 686.638 Trend (β_1) = 33.313
Model Equation	*Output Variable,* y_t *= 686.638 + 33.313 t*
Predictors	Time period index, t
Statistical Significance	• Statistically significant with 100% confidence interval indicated by *** • Good fit into training partition data set • Can be used in time series forecasting since R^2 and Adjusted R^2 values are around 88% which is a good indication of the model's ability to explain the variation in the data and make accurate predictions.

Forecast of Revenue for the period 2020 – 2023

```
> lin.trend.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 2185.725 2219.038 2252.351 2285.664
2021 2318.977 2352.290 2385.603 2418.916
2022 2452.229 2485.542 2518.856 2552.169
2023 2585.482 2618.795 2652.108 2685.421
>
```

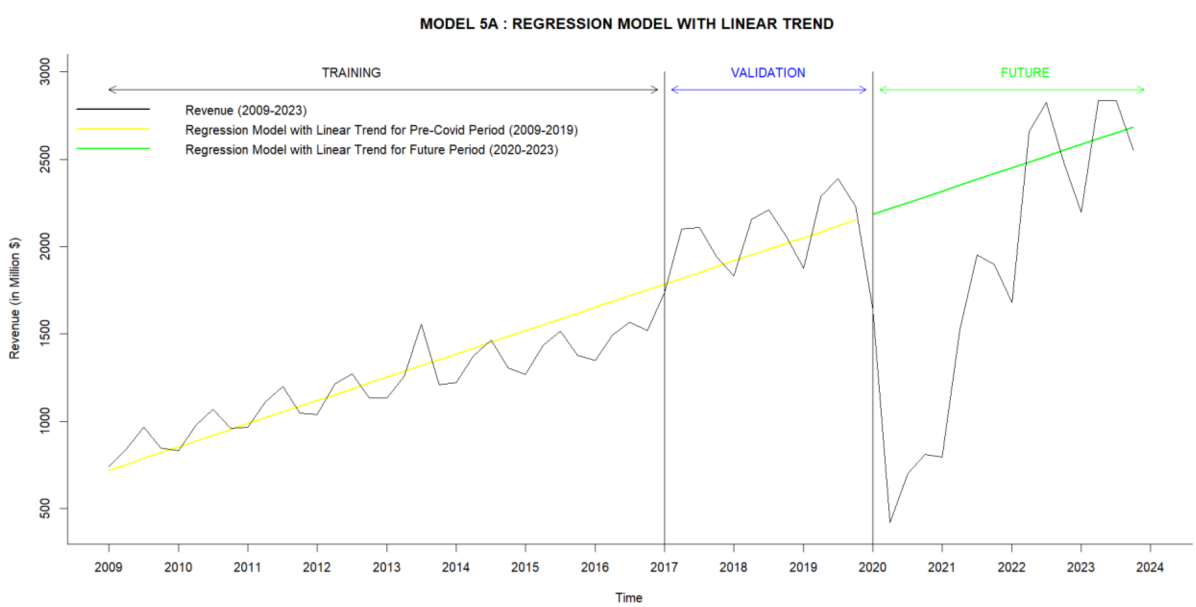

Fig 3.14 Regression Model (Linear Trend) Forecast (green line) of Revenue for Alaska Airlines

Model 5b. Regression Model with Quadratic Trend

<u>Summary</u>

```
Call:
tslm(formula = revenue.ts ~ trend + I(trend^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-272.63  -87.16  -22.80   94.09  316.57 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 869.9283    65.2365  13.335  < 2e-16 ***
trend         9.4057     6.6868   1.407 0.167079    
I(trend^2)    0.5313     0.1441   3.687 0.000658 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 137.7 on 41 degrees of freedom
Multiple R-squared:  0.9127,	Adjusted R-squared:  0.9084 
F-statistic: 214.3 on 2 and 41 DF,  p-value: < 2.2e-16

> 
```

Structure	Regression model with Quadratic trend
Parameter	Intercept (β_0) = 869.9283 Trend (β_1) = 9.4057 I (Trend^2), (β_2) = 0.5313
Model Equation	*Output Variable,* y_t *= 869.9283 + 9.4057 t – 0.5313* t^2
Predictors	Time period indices, t and t^2
Statistical Significance	• Statistically significant with 100% confidence interval indicated by *** • Good fit into training partition data set • Can be used in time series forecasting since R^2 and Adjusted R^2 values are around 91% which is a good indication of the model's ability to explain the variation in the data and make accurate predictions.

<u>Forecast of Revenue for the period 2020 – 2023</u>

```
> quad.trend.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 2369.015 2426.767 2485.581 2545.458
2021 2606.397 2668.399 2731.463 2795.590
2022 2860.780 2927.032 2994.346 3062.724
2023 3132.163 3202.666 3274.230 3346.858
> 
```

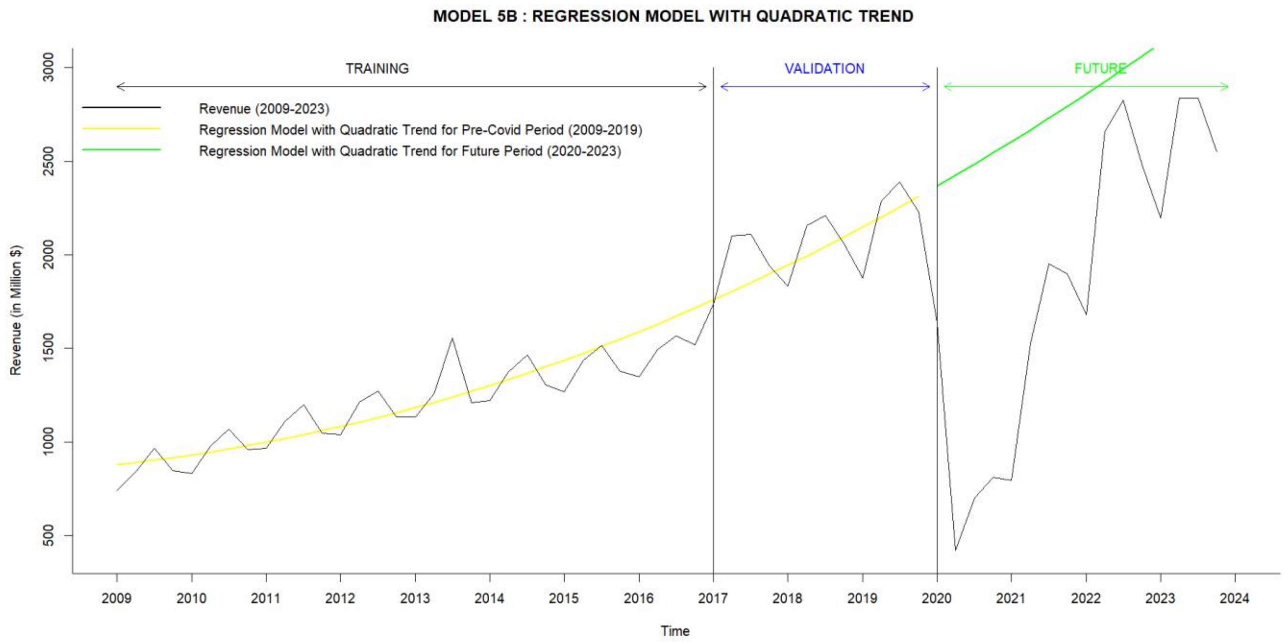

Fig 3.15 Regression Model (Quadratic Trend) Forecast (green line) of Revenue for Alaska Airlines

Model 5c. Regression Model with Seasonality

<u>Summary</u>

```
Call:
tslm(formula = revenue.ts ~ season)

Residuals:
    Min      1Q  Median      3Q     Max 
-633.45 -322.32  -80.95  481.16  814.55 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   1272.3      138.0   9.218 1.92e-11 ***
season2        205.2      195.2   1.051    0.299    
season3        302.2      195.2   1.548    0.129    
season4        148.3      195.2   0.760    0.452    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 457.8 on 40 degrees of freedom
Multiple R-squared:  0.05918,	Adjusted R-squared:  -0.01138 
F-statistic: 0.8387 on 3 and 40 DF,  p-value: 0.4807

>
```

Structure	Regression model with seasonality but without trend
Parameter	Intercept (β_0) = 1272.3 Season2 (β_1) = 205.2 Season3 (β_2) = 302.2 Season4 (β_3) = 148.3
Model Equation	*Outcome Variable,* $y_t = 1272.3 + 205.2\ D_2 + 302.2\ D_3 + 148.3\ D_4$
Predictors	Binary Dummy Variables, D_i
Statistical Significance	• The model is not statistically significant overall, as indicated by the F-statistic (p = 0.4807) • Not a good fit into training partition data set as R^2 and Adjusted R^2 values are too low • Cannot be used in time series forecasting since R^2 and Adjusted R^2 values indicate model's inability to explain the variation in the data and make accurate predictions. Moreover, all the seasonal predictors have insignificant coefficients

Forecast of Revenue for the period 2020 – 2023

```
> revenue.season.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 1272.273 1477.455 1574.455 1420.545
2021 1272.273 1477.455 1574.455 1420.545
2022 1272.273 1477.455 1574.455 1420.545
2023 1272.273 1477.455 1574.455 1420.545
>
```

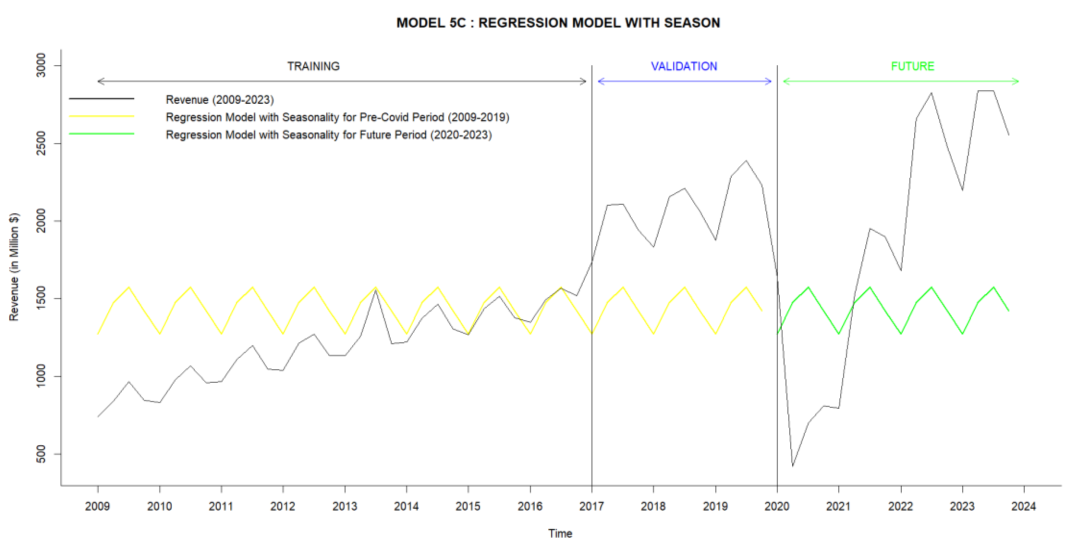

Fig 3.16 Regression Model (only Seasonality) Forecast (green line) of Revenue for Alaska Airlines

Model 5d. Regression Model with Linear Trend & Seasonality

Summary

```
Call:
tslm(formula = revenue.ts ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-273.65  -92.25   26.45   94.68  226.75 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  576.123     50.425  11.425 5.16e-14 ***
trend         33.150      1.534  21.615  < 2e-16 ***
season2      172.032     54.892   3.134 0.003269 ** 
season3      235.882     54.956   4.292 0.000113 ***
season4       48.823     55.063   0.887 0.380692    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 128.7 on 39 degrees of freedom
Multiple R-squared:  0.9275,	Adjusted R-squared:  0.9201 
F-statistic: 124.8 on 4 and 39 DF,  p-value: < 2.2e-16

> |
```

Structure	Regression model with linear trend and seasonality
Parameter	Intercept (β_0) = 576.123 Trend (β_1) = 33.15 Season2 (β_2) = 172.032 Season3 (β_3) = 235.882 Season4 (β_4) = 48.823
Model Equation	*Outcome Variable,* $y_t = 576.123 + 33.15\, t + 172.032\, D_2 - 235.882\, D_3 + 48.823\, D_4$
Predictors	Trend, t Binary Dummy Variables for Season, D_i
Statistical Significance	• The model is statistically significant, indicated by a low p-value for the F-statistic and significant coefficients except for season4 • Good fit into training partition data set indicated by R^2 and Adjusted R^2 values • Can be used in time series forecasting since R^2 and Adjusted R^2 values are around 92% which is a good indication of the model's ability to explain the variation in the data and make accurate predictions.

Forecast of Revenue for the period 2020 – 2023

```
> lin.season.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 2067.873 2273.055 2370.055 2216.145
2021 2200.473 2405.655 2502.655 2348.745
2022 2333.073 2538.255 2635.255 2481.345
2023 2465.673 2670.855 2767.855 2613.945
> 
```

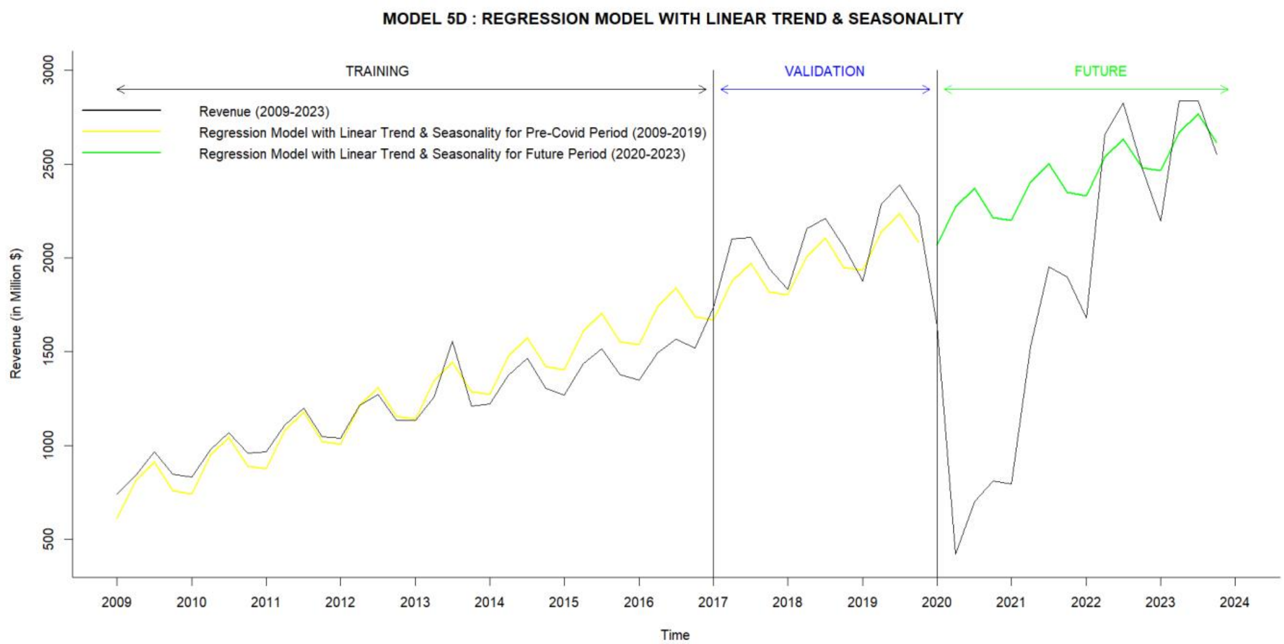

Fig 3.17 Regression Model (Linear Trend & Seasonality) Forecast (green line) of Revenue for Alaska Airlines

Model 5e. Regression Model with Quadratic Trend & Seasonality

Summary

```
Call:
tslm(formula = revenue.ts ~ trend + I(trend^2) + season)

Residuals:
     Min       1Q   Median       3Q      Max 
-226.520  -36.362    8.305   47.587  241.743 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 760.3770    53.4575  14.224  < 2e-16 ***
trend         9.0470     4.8793   1.854 0.071490 .  
I(trend^2)    0.5356     0.1051   5.096 9.84e-06 ***
season2     173.1031    42.8617   4.039 0.000252 ***
season3     236.9531    42.9119   5.522 2.58e-06 ***
season4      48.8227    42.9948   1.136 0.263257    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 100.5 on 38 degrees of freedom
Multiple R-squared:  0.9569,	Adjusted R-squared:  0.9513 
F-statistic: 168.9 on 5 and 38 DF,  p-value: < 2.2e-16

> 
```

Structure	Regression model with quadratic trend and seasonality
Parameter	Intercept (β_0) = 760.377 Trend (β_1) = 9.0470 I (Trend^2) (β_2) = 0.5356 Season2 (β_3) = 173.1031 Season3 (β_4) = 136.9531 Sesaon4 (β_5) = 48.8227
Model Equation	*Outcome Variable,* $y_t = 760.377 + 9.047\, t – 0.5356\, t^2 + 173.1031\, D_2 – 136.9531\, D_3 + 48.8227\, D_4$
Predictors	Trend, t, t^2 Binary Dummy Variables for Season, D_i
Statistical Significance	• The model is statistically significant, indicated by a low p-value for the F-statistic and significant coefficients for all seasons • Good fit into training partition data set indicated by R^2 and Adjusted R^2 values • Can be used in time series forecasting since R^2 and Adjusted R^2 values are around 95% which is a good indication of the model's ability to explain the variation in the data and make accurate predictions.

<u>Forecast of Revenue for the period 2020 – 2023</u>

```
> quad.season.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 2252.127 2483.019 2605.729 2477.529
2021 2489.709 2724.886 2851.881 2727.967
2022 2744.431 2983.893 3115.173 2995.544
2023 3016.293 3260.040 3395.605 3280.260
>
```

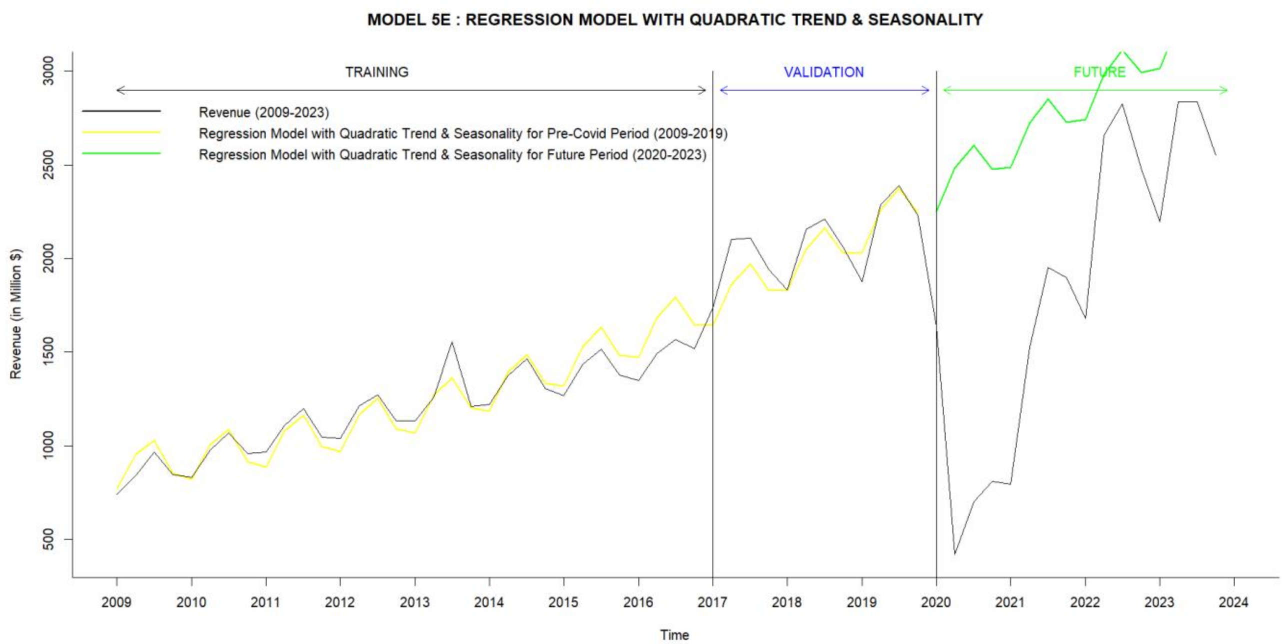

Fig 3.18 Regression Model (Quadratic Trend & Seasonality) Forecast (green line) of Revenue for Alaska Airlines

3.1.6.6 Model 6 : Autocorrelation & Autoregressive Model (Automated HW's Model + AR(1) Model)

Since Automated HW's model had better accuracy compared to Regression Models, the same was chosen for Two-level forecast with AR(1) Model

Intriguingly, the correlogram for the training partition displayed statistical significance, while the validation partition showed a random walk of autocorrelation, indicating statistical insignificance. However, a two-level model was devised to explore potential improvements in autocorrelation. As depicted in Fig 3.21, there's a reduction in autocorrelation, although it was anticipated that this wouldn't yield significant impact, as all autocorrelations were already accounted for prior to the development of the two-level model.

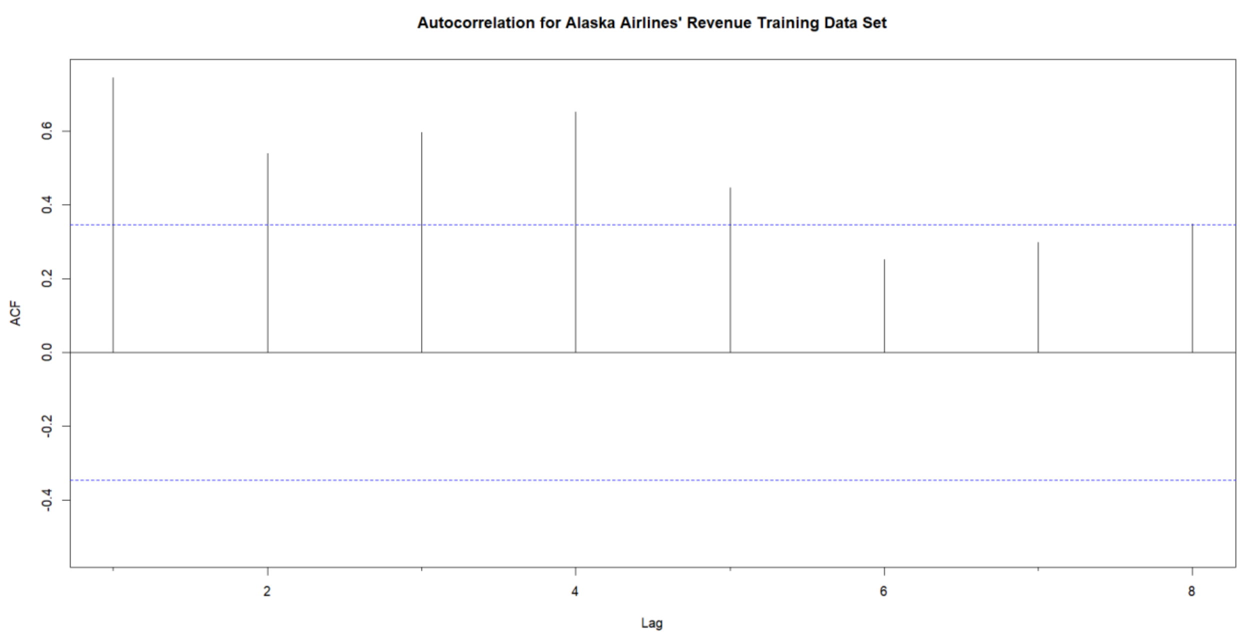

Fig 3.19 Autocorrelation Chart for Training partition of Alaska Airlines Revenue Data

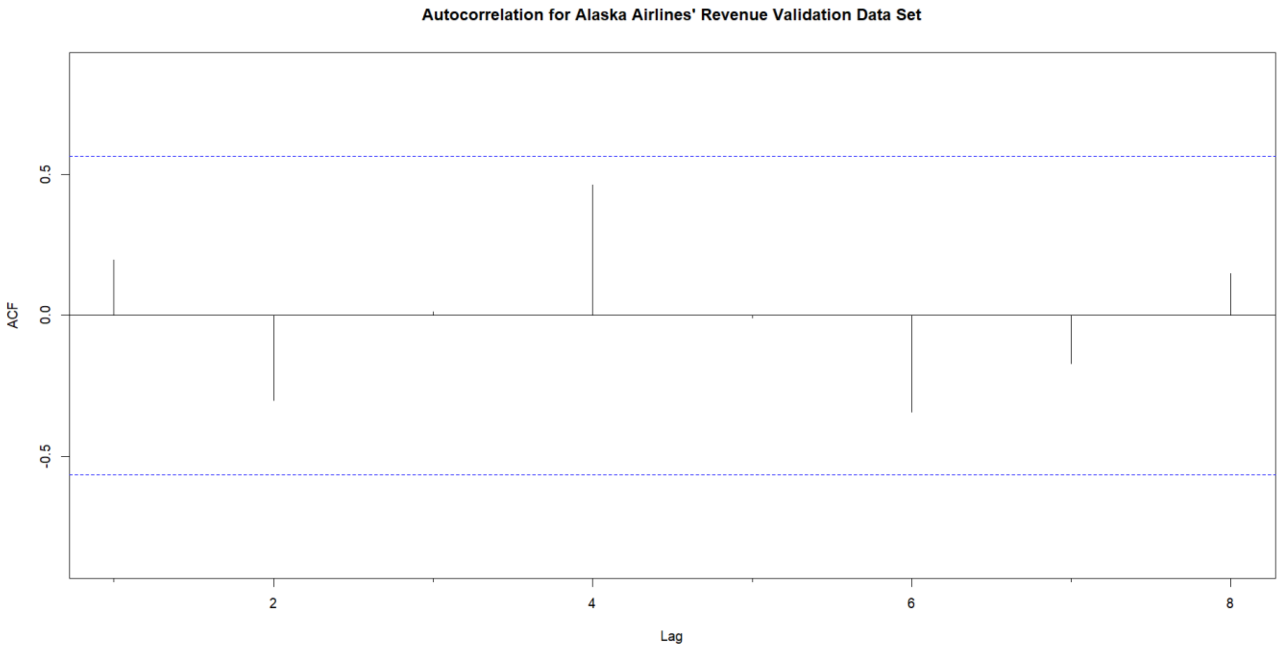

Fig 3.20 Autocorrelation Chart for Validation partition of Alaska Airlines Revenue Data

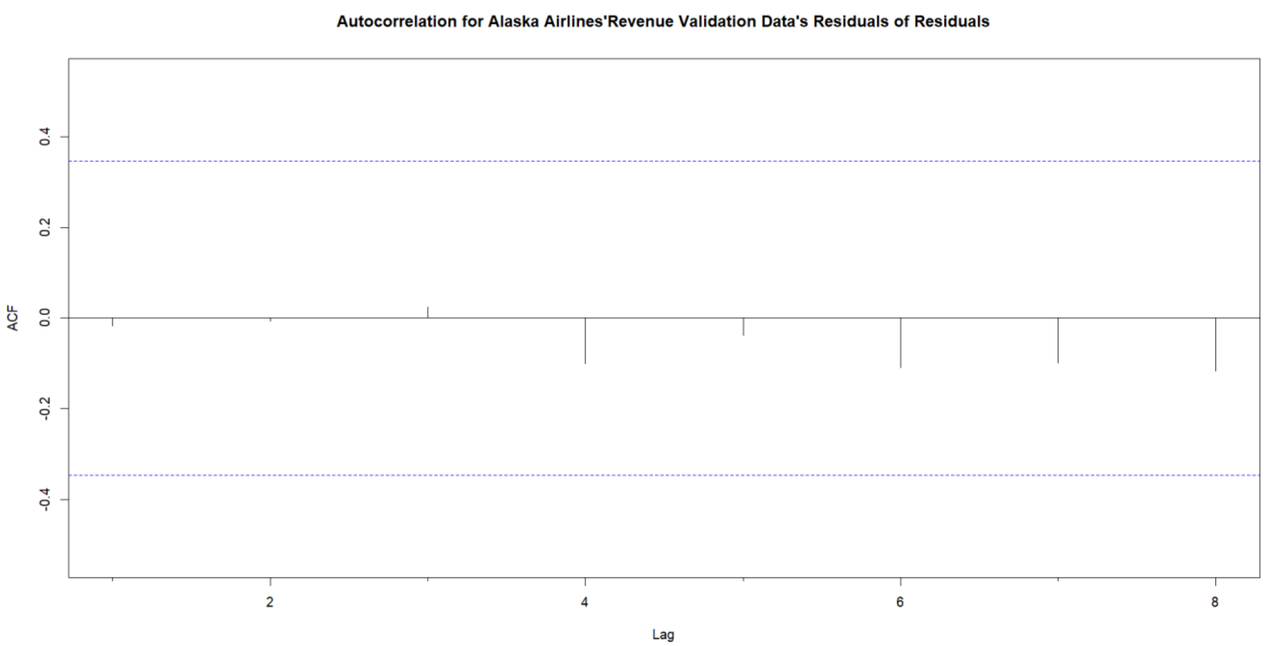

Fig 3.21 Autocorrelation Chart for Validation partition at Level 2 of Alaska Airlines Revenue Data

Summary for Level 2

```
Series: HW.ZZZ$residuals
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1     mean
      0.0061  -0.0011
s.e.  0.1504   0.0075

sigma^2 = 0.002562:  log likelihood = 69.86
AIC=-133.72   AICc=-133.12   BIC=-128.37

Training set error measures:
                       ME       RMSE        MAE      MPE     MAPE      MASE          ACF1
Training set 9.857094e-08 0.04945713 0.03354667 102.2268 102.2268 0.7577864 -0.000525277
```

Structure	Autoregressive Model of Order 1 for Regression Residuals
Parameter	Intercept (α) = -0.0011 Coefficient of e_{t-1} (β_1) = 0.0061
Model Equation	*Outcome Variable,* e_t = -0.0011 + 0.0061 e_{t-1}
Inferences	• Coefficient indicates weak positive autocorrelation implying that there's a slight tendency for the current revenue value to be positively influenced by the previous one, may not be statistically significant • Non-zero mean; indicating the expected value when lagged value is zero.

Forecast of Revenue for the period 2020 – 2023

```
   Reg.Forecast AR(1)Forecast Combined.Forecast
1      2154.844        -0.001          2154.843
2      2455.359        -0.001          2455.358
3      2619.909        -0.001          2619.908
4      2332.804        -0.001          2332.803
5      2280.831        -0.001          2280.830
6      2596.849        -0.001          2596.848
7      2768.737        -0.001          2768.736
8      2463.467        -0.001          2463.466
9      2406.819        -0.001          2406.818
10     2738.339        -0.001          2738.338
11     2917.565        -0.001          2917.564
12     2594.131        -0.001          2594.130
13     2532.808        -0.001          2532.807
14     2879.830        -0.001          2879.829
15     3066.395        -0.001          3066.394
16     2724.796        -0.001          2724.794
> 
```

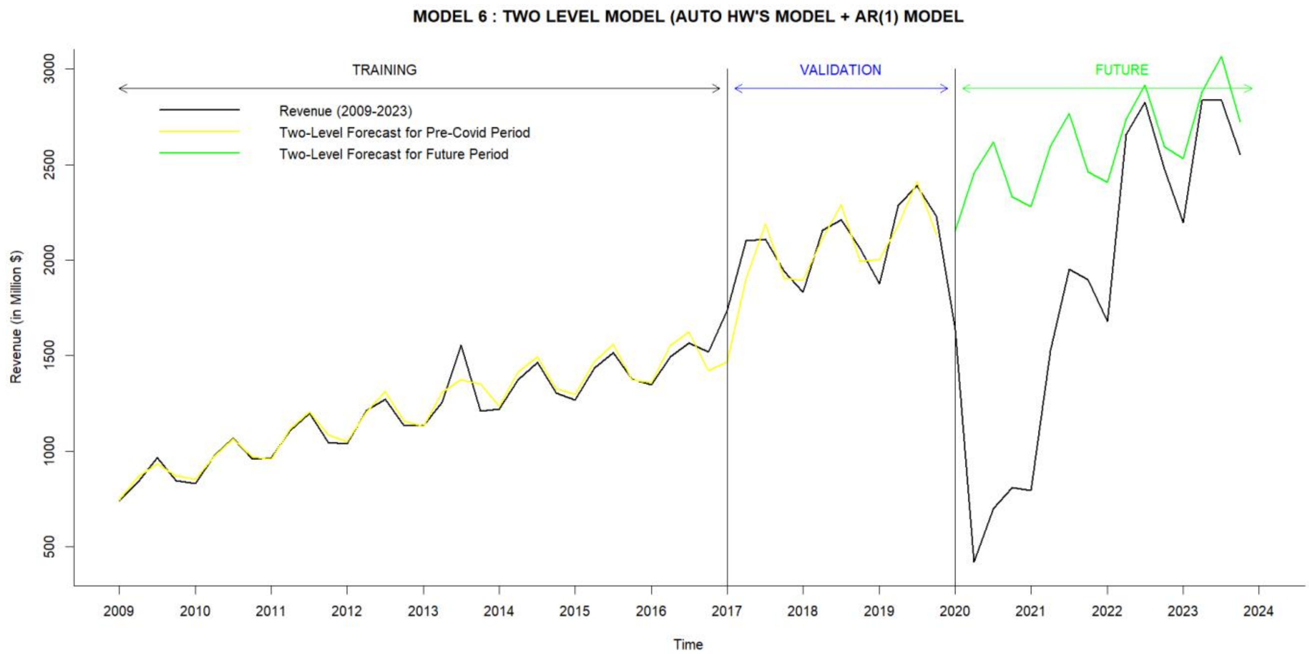

Fig 3.22 Two Level Model (Auto HW Model + AR(1) Model) Forecast (green line) of Revenue for Alaska Airlines

3.1.6.7 Model 7 : Automated Autoregressive Integrated Moving Average (ARIMA) Model

<u>Summary</u>

```
Series: revenue.ts
ARIMA(1,0,0)(0,1,1)[4] with drift

Coefficients:
         ar1     sma1    drift
      0.7454  -0.5369  33.8427
s.e.  0.1102   0.1951   6.6651

sigma^2 = 8496:  log likelihood = -237.05
AIC=482.1   AICc=483.24   BIC=488.86

Training set error measures:
                     ME     RMSE      MAE        MPE     MAPE      MASE       ACF1
Training set -0.1468673 84.52668 56.70033 -0.5052959 3.655566 0.4074763 0.03267285
>
```

Structure	Auto-ARIMA Model
Parameter	p = *1*, Autoregressive model AR(1) d = *0*, No differencing to remove linear trend q = *0*, No moving average model P = *0*, No Autoregressive model for seasonality D = *1*, order 1 differencing to remove linear trend Q = *1*, Moving average model MA(1) for error lags m = *4*, for quarterly seasonality
Model Equation	*Outcome Variable*, $y_t - y_{t-1} = 33.8427 + 0.7454\, y_{t-1} - 0.5369\, \rho_{t-1}$

Forecast of Revenue for the period 2020 – 2023

```
> revenue.auto.arima.pred$mean
         Qtr1     Qtr2     Qtr3     Qtr4
2020 2115.445 2440.432 2503.006 2341.463
2021 2234.487 2563.632 2629.305 2470.072
2022 2364.818 2695.246 2761.876 2603.356
2023 2498.633 2829.458 2896.382 2738.083
>
```

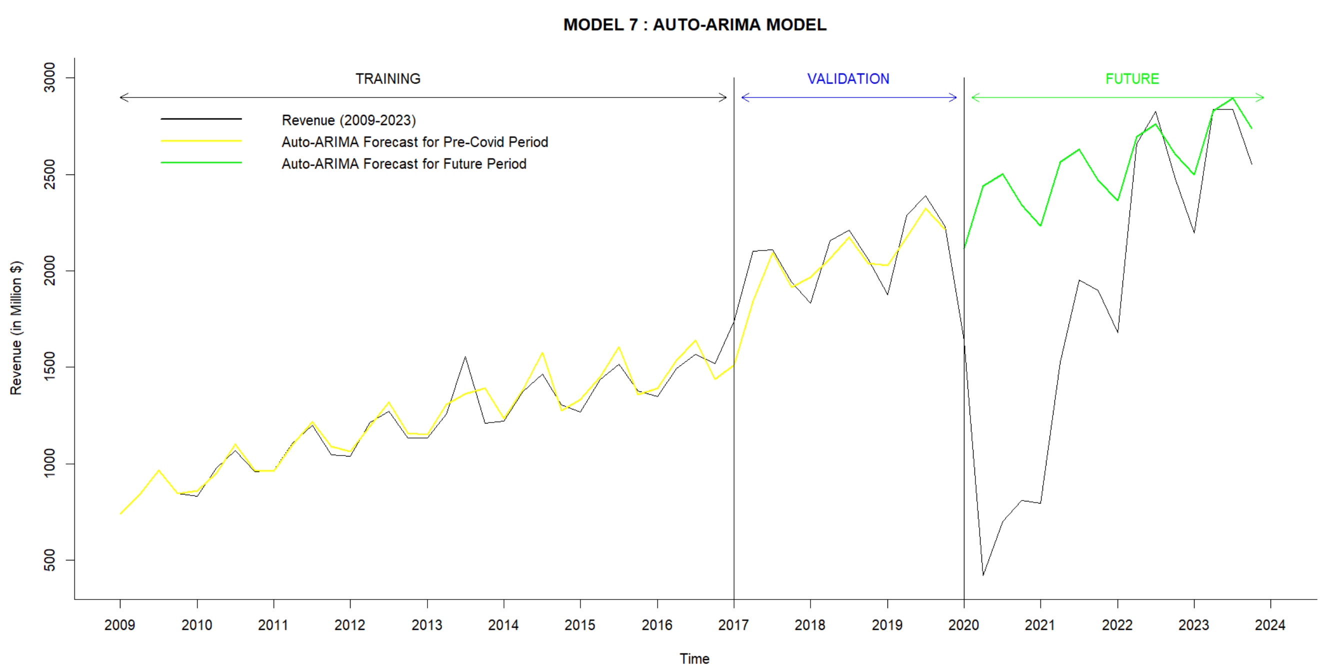

Fig 3.23 Automated ARIMA Model Forecast (green line) of Revenue for Alaska Airlines

3.1.7 Step 7: Evaluate & Compare Performance

Performance measures provide insights into each model's predictive accuracy, assessed through error metrics such as ME, RMSE, MAE, MPE, MAPE, MASE, and ACF1. Lower error values indicate higher accuracy. RMSE and MAPE are key metrics for accuracy evaluation, depicted in the chart below for both the Validation Dataset and the entire dataset. Models highlighted in yellow are the top performers on both datasets.

Model No	**Model Name**	**Valid_RMSE**	**Valid_MAPE**	**Future_RMSE**	**Future_MAPE**
1	Naïve Model	590.888	26.338	163.387	9.091
2	Seasonal Naïve Model	611.04	28.432	189.411	9.109
3A	Two-Level Forecasting Linear Trend + Trailing MA K=2 K=3 K=4	428.308 426.412 427.851	19.974 19.88 19.951	43.042 58.404 72.232	1.954 2.647 3.137
3B	Two-Level Forecasting Quadratic Trend + Trailing MA K=2 K=3 K=4	520.42 516.395 520.155	24.181 23.981 24.167	42.669 57.783 70.969	2.085 2.737 3.111
4	Automated Holt-Winter's Model	(M,Ad,M) 482.079	(M,Ad,M) 22.643	(M,A,M) 77.209	(M,A,M) 3.311
5A	Regression Model : Linear Trend	411.729	17.928	153.427	8.828
5B	Regression Model : Quadratic Trend	532.725	23.749	132.958	7.542

5C	Regression Model : Seasonality	891.752	42.333	436.469	26.64
5D	Regression Model : Linear Trend, Seasonality	399.172	18.526	121.151	7.057
5E	Regression Model : Quadratic Trend, Seasonality	518.751	24.098	93.377	4.955
6	Automated Holt-Winter's Model + AR Model	482.083	22.643	77.209	3.311
7	Automated ARIMA Model	(0,0,0) (1,1,0) [4] 425.072	(0,0,0) (1,1,0) [4] 19.768	(1,0,0) (0,1,1) [4] 84.527	(1,0,0) (0,1,1) [4] 3.656

All the models except the Regression model with seasonality have better accuracy measures compared to Naïve and Seasonal Naïve Forecast.

Given its **parsimony** compared to its quadratic counterpart and its achievement of the **lowest MAPE** among all models, the Two-Level Model combining a Regression Model with **linear trend and seasonality** for level 1, and a **Trailing Moving Average Model** for level 2, is deemed the optimal choice for implementation in the existing enterprise system.

3.1.8 Step 8: Implement Forecasts/ Systems

This selection ensures a reliable forecasting tool for the organization in facing similar pandemic or unforeseen scenarios in the future.

Estimation of Loss of Revenue by the top two models

i. Regression Model with Linear Trend & Seasonality + Trailing MA model

Time Period	Actual Revenue	Forecasted Revenue	Loss of Revenue
2020.00	742	2218.242	1476.242
2020.25	844	2420.119	1576.119
2020.50	967	2515.768	1548.768
2020.75	846	2357.743	1511.743
2021.00	830	2344.440	1514.440
2021.25	976	2546.317	1570.317
2021.50	1068	2641.965	1573.965
2021.75	958	2483.941	1525.941
2022.00	965	2470.637	1505.637
2022.25	1110	2672.515	1562.515
2022.50	1198	2768.163	1570.163
2022.75	1045	2610.139	1565.139
2023.00	1039	2596.835	1557.835
2023.25	1214	2798.713	1584.713
2023.50	1272	2894.361	1622.361
2023.75	1132	2736.337	1604.337

```
> cat("Total Loss of Revenue:", total_loss_LTS, "(in Million $) \n")
Total Loss of Revenue: -24870.24 (in Million $)
> cat("Percentage Loss of Revenue:", percentage_loss_LTS, "%\n")
Percentage Loss of Revenue: -60.54653 %
```

Regression Model with Quadratic Trend & Seasonality + Trailing MA model

Time Period	Actual Revenue	Forecasted Revenue	Loss of Revenue
2020.00	742	2238.302	1496.302
2020.25	844	2434.628	1590.628
2020.50	967	2528.633	1561.633
2020.75	846	2374.059	1528.059
2021.00	830	2370.103	1540.103
2021.25	976	2566.429	1590.429
2021.50	1068	2660.433	1592.433
2021.75	958	2505.859	1547.859
2022.00	965	2501.904	1536.904
2022.25	1110	2698.229	1588.229
2022.50	1198	2792.234	1594.234
2022.75	1045	2637.660	1592.660
2023.00	1039	2633.704	1594.704
2023.25	1214	2830.030	1616.030
2023.50	1272	2924.035	1652.035
2023.75	1132	2769.461	1637.461

```
> cat("Total Loss of Revenue:", total_loss_QTS, "(in Million $)\n")
Total Loss of Revenue: -25259.7 (in Million $)
> cat("Percentage Loss of Revenue:", percentage_loss_QTS, "%\n")
Percentage Loss of Revenue: -60.9171 %
>
```

CHAPTER – IV

CONCLUSION

4.1 Final Recommendations

Based on the analysis conducted, it is recommended to implement the Two-Level Trailing MA Model with Linear Trend and Seasonality for forecasting revenue in Alaska Airlines during similar pandemic or unforeseen scenarios. This model demonstrated the lowest Mean Absolute Percentage Error (MAPE) and exhibited parsimony compared to its quadratic counterpart.

4.2 Remarks on Analysis and Forecasted Results

The analysis revealed significant revenue declines during the COVID-19 pandemic, aligning with expectations due to travel restrictions and lockdown measures. However, it's essential to note that association does not imply causation. While COVID-19 likely played a significant role in the revenue drop, other external factors may have contributed as well.

4.3 Statement on Analysis & Forecast Results

The top-performing models, the Two-Level Trailing MA Model with Linear Trend and Seasonality, and its quadratic counterpart, projected significant losses of revenue. Specifically, the linear model estimated a total loss of revenue amounting to $24,870.24, representing a percentage loss of 60.54653% compared to the forecasted revenue. Similarly, the quadratic model projected a total loss of revenue of $25,259.7, indicating a percentage loss of 60.9171% compared to the forecasted revenue. These findings underscore the substantial financial impact of the pandemic on

Alaska Airlines' revenue generation and emphasize the importance of robust forecasting models in mitigating such risks.

4.4 Benefits and Limitations

4.4.1 Benefits

- The implemented forecasting model provides a reliable tool for predicting revenue during crises, aiding in decision-making and strategic planning.
- The use of time series analysis techniques allows for a systematic evaluation of revenue trends, facilitating informed insights into the financial impact of external events.

4.4.2 Limitations

- The forecasting models are based on historical data and assumptions, which may not fully capture unforeseen events or sudden shifts in consumer behavior.
- External factors beyond the scope of the analysis, such as changes in market dynamics or regulatory policies, could influence revenue outcomes and are not accounted for in the models.