**Introduction**

The tourism industry in New Zealand significantly influences the country's economy. Over $40.91 billion was brought into New Zealand by the tourist industry, making it the country's single highest export earner (Munasinghe et al., 2022). Approximately 8.4 percent of New Zealand's population, or 229,566 individuals, are employed directly in the tourism industry, demonstrating the industry's enormous positive impact on regional economies and employment (Munasinghe et al., 2022). Before the outbreak of the COVID-19 pandemic, the tourism industry in New Zealand was a thriving and crucial element of the country's economy. New Zealand has firmly established itself as a global tourist destination due to its breathtaking natural beauty and diverse outdoor activities (Cloke & Perkins, 2002). The Southern Alps, verdant woods, and clean beaches drew tourists worldwide. Moreover, the adventure tourism industry proved highly appealing due to its wide range of thrilling activities, such as bungee jumping, skydiving, and skiing.

The tourism industry, nevertheless, experienced an abrupt cessation due to the COVID-19 pandemic. The tourism sector was affected by health and safety concerns, international travel restrictions, and border closures (Henrickson, 2020). As the number of foreign visitors decreased, New Zealand was compelled to make a temporary transition in emphasis towards domestic tourism (McNeill & Asquith, 2022). The tourism industry was severely impacted economically as a result of the stringent restrictions implemented by the government to manage the pandemic (Gössling et al., [2020](https://www.tandfonline.com/doi/full/10.1080/13683500.2021.1874311)). Before COVID-19, New Zealand's tourism industry was substantial and diversified. However, the pandemic brought problems that had never been seen before, forcing tactics to be reevaluated and the need to be flexible in a constantly changing world.

The goal of this analysis is to estimate monthly tourist arrivals in New Zealand without the impact of COVID-19 on tourism revenue and compare the performance of forecasts that include and exclude the COVID-19 pandemic period to predict the financial loss in tourism revenue in New Zealand since the start of COVID-19.

In this analysis, forecasting monthly tourist arrivals in New Zealand with or without COVID-19 is essential to minimize financial losses in the tourism revenue of the country. It is crucial to forecast this financial loss in tourism revenue in New Zealand so that government bodies, businesses, and authorities can develop strategies to mitigate the impact of COVID-19 on the tourism sector and allocate resources efficiently. Policies implemented and developed by the government may effectively reduce the financial losses due to the COVID-19 pandemic in New Zealand. The impact assessment should encourage stakeholders, such as the government and corporations, to allocate more resources. By illustrating financial losses and the potential advantages of specific measures, decision-makers may be motivated to address financial losses, explore potential solutions, and implement effective strategies for the recovery of the tourism industry. Therefore, a descriptive goal is chosen to assess the impact of COVID-19 on financial losses in tourism revenue based on monthly tourist arrivals in New Zealand."

The New Zealand dataset from the Excel file 'Monthly tourist arrivals.xlsx' was utilised for the analysis. Subsequently, the file format was converted to .csv for import into RStudio for further analysis. Please refer to the uploaded R script and the Excel file for more details.

**Visualisation**

In this stage, GgAcf and Ggseasonplot were chosen to visualise the series.

A graph of a line graph

Description automatically generated

Figure 1.3a

From the ACF plot, several observations can be made:

* **Exceeding Critical Values:** All autocorrelation values exceeding the critical threshold indicate that these data points do not adhere to the characteristics of white noise.
* **Positive Autocorrelation at Small Lags:** The autocorrelation values for small lags are notably large and positive. Additionally, the ACF exhibits a scalloping pattern with a slow decay, indicating the presence of a trend in the data.
* **Seasonal Autocorrelations:** Significant autocorrelations are observed at lags 12 and 24, which are consistent with seasonal lags that have higher autocorrelation. These values surpass the corresponding proximity, suggesting monthly data with a clear seasonality pattern. This finding confirms the presence of seasonality within the dataset.

A graph of different colored lines

Description automatically generated

Figure 1.3b

From the seasonal plot, we can observe that:

* The presence of trends is evident. The consistent, ascending pattern of lines over successive years serves as a clear indicator of a positive trend within the data. Nonetheless, it is important to note that this upward trend experienced an interruption in January 2020, as it was abruptly disrupted by the structural shift induced by the COVID-19 pandemic.
* Seasonality is present. The nearly identical lines show a consistent progression in tourist arrivals throughout the months across the years, except for the years 2020 and 2021, which correspond to the COVID-19 pandemic period.
* The graph pattern reveals a notable trend with a high influx of tourists in January and February. This trend continues to rise during the summer months of December to February, which many young people and the general population prefer. It is important to note that the summer season coincides with the global holiday season, which includes various celebrations from different cultures and countries.
* The highest peak in tourist arrivals is observed in December. This surge in arrivals during December can be attributed to the presence of worldwide holidays, such as Christmas and New Year's Eve. Additionally, December is a time when many schools have their breaks in various countries. This alignment with the holiday season makes December a popular period for families to plan vacations and embark on travels. The combination of holiday festivities and school breaks creates a significant incentive for people to explore new destinations during this time of the year.
* There is a significant decline in tourism arrivals during the autumn season (March-May). This is primarily due to the school year being in full swing, limiting family travel opportunities. Additionally, New Zealand experiences cooler temperatures and shorter daylight hours during the autumn season, which may not be favourable for tourism.
* Notably, starting from 2015, there is an interesting pattern, as indicated by the Ministry, showing a slight growth in visitors during the autumn season. This pattern may be attributed to an increase in older people visiting New Zealand who prefer the autumn seasons. This pattern, while subtle in the seasonal plot, becomes more noticeable when compared with previous years.
* As reported by the Ministry of Business, Innovation, and Employment, visitors during the summer period account for approximately twice as many visitors during the winter season (June-August) and the preceding months. This is consistent with the seasonal plots, which indicate the lowest tourism arrivals in June. The low temperature during winter might be less appealing to tourists for outdoor activities, and accommodations may be closed due to snow and ice. Winter sports can also be more expensive during this season.
* There is a notable increase in tourism arrivals during the spring season between September and November. This is mainly due to milder temperatures and longer daylight hours, making it an ideal time for outdoor activities. Spring is less crowded and more affordable for travelling compared to the peak summer season. Additionally, spring marks the beginning of the wine harvest season in New Zealand, attracting wine enthusiasts who visit vineyards, participate in tastings, and experience the wine culture.
* A sudden peak in tourism arrivals in New Zealand occurred in September 2011. This was attributed to the Rugby World Cup, one of the world's most prestigious rugby tournaments, attracting a large number of international visitors and fans.
* During the COVID-19 pandemic, starting in 2020, the pattern followed expectations until February, when COVID-19 outbreaks led New Zealand to restrict international arrivals. This resulted in a sharp drop in historically low tourist arrivals.
* The restrictions persisted until March 2021 when international borders were reinstated, elucidating the abrupt surge in tourist arrivals recorded during the period spanning March to May 2021.
* However, another significant drop from May to August 2021 could be attributed to news of a new variant of Omicron (Delta) originating in India in late 2021, causing COVID-19. This variant was reported to be twice as infectious, leading to stricter lockdown measures and a further halt in international travel until the end of 2021.

A graph showing the number of tourist arrivals

Description automatically generated

Figure 1.3c

A graph showing the growth of tourism arrivals

Description automatically generated

Figure 1.3d

From the STL Decomposition and Seasonal Adjusted data, we can observe that:

* Seasonal adjusted data involves the removal of the seasonal component, leaving behind the trend cycle and remainder components. Consequently, the plot of seasonal adjusted data does not exhibit any seasonality.
* The presence of trends and structural breaks can be further substantiated by both the STL Decomposition and the plot of seasonal adjusted data. These visualisations offer a clearer insight into the trends present in the trend cycle of STL decomposition and the plot of seasonal adjusted data.
* Although there is a discernible positive increasing trend, it is evident that this trend is not linear.
* The remainder component initially appears constant, but exhibits increased variability towards the end, showing significant random variation during the period spanning 2020-2021.

**Data Preprocessing**

The time span selected for the data visualisation begins in January 1980. This is due to its extensive historical perspective, which spans over four decades. The availability of large datasets facilitates the identification of the tourism industry's responses to various challenges over time. It permits the evaluation of the industry's vulnerabilities and resiliency, shedding light on its capacity to recover from crises. This extensive time frame permits a thorough examination of tourism industry trends, patterns, and cycles. By including data from January 1980, deeper understanding of the tourism industry's baseline performance can be obtained. This baseline is essential for distinguishing pandemic-related losses from normal fluctuations and long-term growth trends in the industry.

Next, larger datasets such as spans over four decades enable models to generalise to unknown or future data more effectively. When more examples are available, the model is better equipped to manage a broader range of scenarios, and the risk of overfitting the model to the training dataset and improving the predictability of future outcomes can be reduced. Over the years, the tourism industry has been influenced by policy shifts, economic vicissitudes, and technological advancements. A dataset from January 1980 can aid in tracing the impact of such changes on tourism revenue and provide valuable insight into how external factors impact the industry. Having data from before the COVID-19 pandemic provides a baseline for estimating the extent of the financial loss caused by the outbreak. This historical context is essential for comprehending the scale of the disruption and developing recovery strategies.

Therefore, a large dataset of monthly tourist arrivals can assist in predicting the tourism industry's adaptability to future challenges, such as new pandemics, global events, and consumer behaviour shifts. Understanding past observations and trends may provide strategies for minimising financial losses and promoting resilience.

**Series Partitioning**

A graph showing the number of datasets

Description automatically generated

Figure 1.5

A black background with a black square

Description automatically generated with medium confidence

Table 1.5

In Table 1.5, the chosen full dataset spans from January 1980 to December 2021, encompassing a total of 504 observations. The Pre-Covid dataset, spanning from January 1980 to December 2019, is divided into a training set and a test set. The training set covers the period from January 1980 to December 2011, while the test set spans from January 2012 to December 2019. This partitioning adheres to an 80/20 split, with approximately 80% of the 480 observations allocated to the training set (384 observations) and the remaining 20% assigned to the test set (96 observations). The Covid dataset, which encompasses the period from January 2020 to December 2021, comprises 24 observations.

**Modelling**

**Simple Forecasting Method: Seasonal Naive**

The seasonal naive method involves using the value from the same season in the previous year as the forecast for the current year's season. The drift method, on the other hand, is only suitable for data with a linear trend, which may not be appropriate in this case due to the data's nonlinear trend. This is because it only captures the trend component but ignores seasonality, leading to inaccurate forecasts. The average and naive methods are designed for data with no trend or seasonality. The average method calculates the historical data's average for forecasting, while the naive method uses the most recent data as the forecast. Consequently, both methods are inappropriate for forecasting data with both trend and seasonality. As a result, the seasonal naive forecast became the only option when dealing with data that are very heavily influenced by seasonality.

**Forecasting Model with STL Decomposition: ETS (M,Ad,N) + STL Decomposition**

The forecasting model with STL decomposition, utilising the ETS(M,Ad,N) model with an additive trend, was chosen due to the data's error distribution, which is multiplicative. This choice aligns with the training data containing only positive values and no zeros or negative values. Additionally, the use of an additive damped trend is preferred to prevent over-forecasting, especially when dealing with a long forecast horizon of 96 data points. The stlf() function forecasts the seasonal component using the seasonal naive method, while the seasonally adjusted data is forecast using the selected method or model. Therefore, it is unnecessary to include seasonality in the model.

**ETS Exponential Smoothing Model: ETS (M,Ad,M )**

The choice of an ETS (M,Ad,M) model is driven by the absence of zero or negative values in the data, making models with multiplicative errors suitable for strictly positive data. To mitigate over-forecasting concerns associated with a long forecast horizon of 96 data points, a model with an additive damped trend is selected. The data displays seasonality with increasing variation, making a model with multiplicative seasonality a logical choice.

**R Generated ETS Model: ETS(M,A,M)**

By utilising the ets() function in R, an ETS (M,A,M) model was generated for the dataset. The smoothing parameters are as follows: α=0.2866, β=0.0031, and γ=0.4134. Here, alpha controls level smoothing, beta controls trend smoothing, and gamma controls seasonality smoothing. The small beta value indicates that the dataset's trend remains relatively stable over time. The sigma value of 0.0583 represents the estimated standard deviation of the model's error term, quantifying the variability or dispersion of the residuals, which are the differences between actual and predicted values. Additional information values returned by the ets() function in R are detailed in the summary below.

**Parameter Estimates, Initial Values and Measures of the Within-Sample Forecast Residuals**

**A black background with a black square

Description automatically generated with medium confidence**

Table 1.6a

The estimated smoothing parameters of the forecast model (M,Ad,N) with STL decomposition consist of α, β, and ϕ. A relatively high value of α, specifically 0.4646, signifies that the model assigns considerable weight to the most recent observations within the time series. Conversely, the value of β is notably low at 0.0112, indicating that historical data primarily influence the trends within the data and are less responsive to recent fluctuations. Despite the low value of β in the forecast model (M,Ad,N) with STL decomposition, it remains the highest among the three models, implying that the forecasts generated are the most uncertain among the three models. The absence of a gamma (γ) value in this model can be attributed to the non-seasonality assumption, as seasonal components are separately forecasted using the seasonal naive method via the stlf() function. Additionally, the value of ϕ, closely approaching 1 at 0.98, indicates a strong damping effect, which is the outcome of the "damping = TRUE'' setting.

The estimated smoothing parameters of the forecast model (M,Ad,M) include α, β, γ, and ϕ. A lower value of α compared to the (M,Ad,N) model with STL decomposition, specifically 0.2931, suggests that lesser weight is assigned to the most recent observations. Similarly, the value of β is lower at 0.0098, signifying that changes in trends are even more influenced by historical data and less responsive to recent changes compared to the previous model. On the other hand, a relatively higher value of γ, 0.4026, indicates that the recent seasonality components are given more weight than past seasonality components. This could result in the forecasted data having a comparable pattern, seasonal cycle duration, and rate of increase in the level of variability in the seasonal cycle, as seen in recent historical data. Similar to the previous model, the value of ϕ remains constant at 0.98, denoting a strong damping effect.

The estimated smoothing parameters for the forecast model (M,A,M) encompass α, β, and γ. The lowest value of α among the three models, at 0.2866, suggests that the most recent observations receive the least emphasis compared to the other models. A low β value of 0.0031 also indicates that trends are most influenced by historical data and minimally responsive to recent changes when compared to the other forecast models. The higher value of γ  at 0.4134, signifies a greater weighting of the most recent seasonal components in the data. This may lead to the forecasted data displaying a similar pattern, seasonal cycle duration, and an increase in the level of variability in the seasonal cycle, much like what has been observed in recent historical data. The absence of a ϕ value in this model is attributed to the lack of damping applied in this particular modelling approach.

**Separate Plots of the Training Set Data with the Fitted Values from each of the Method and Models.**

**A graph showing a number of different weather forecasts

Description automatically generated with medium confidence**

Figure 1.6a

**A graph showing a number of times

Description automatically generated with medium confidence**

Figure 1.6b

**A graph showing a graph of a number of tourists

Description automatically generated with medium confidence**

Figure 1.6c

**A graph showing a graph of a number of tourists

Description automatically generated with medium confidence**

Figure 1.6d

In Figure 1.6a, there are 12 missing values at the beginning. This occurs because the seasonal naive forecast method predicts the current season's value based on the most recent observation from the previous season. Consequently, they require at least one complete seasonal cycle of data as a reference before they can start forecasting. As a result, they cannot generate predictions for the initial data points of the series due to the absence of previous observations from the same season. In this case, a complete seasonal cycle includes 12 months of data, accounting for the 12 missing values.

In contrast, STL decomposition with ETS(M,Ad,N), ETS(M,Ad,M), and ETS(M,A,M) can handle initial data elements more flexibly. These models utilise the available data to estimate the initial state, which includes the initial level, trend, and seasonality. To address missing or incomplete historical data, ETS models initialise the state variables in a way that allows them to make forecasts despite the limited historical data at the beginning of the time series. Consequently, the plots in Figures 1.6b, 1.6c, and 1.6d show that the variance magnitude on the training dataset is now relatively comparable. However, there are still regions of underfitting that may not be visually identifiable.

The four plots exhibit the absence of a cyclic pattern and the presence of positive trends and seasonality. This is evident from the sustained long-term increase and the consistent patterns of peaks and troughs occurring at regular time intervals. However, the seasonal naive forecast method plot displays these characteristics at a different magnitude. Apart from these distinctions, distinguishing between the plots, particularly among the three models, poses a visual challenge.

**Residual Diagnostic Checks**

A graph of different types of graphs

Description automatically generated with medium confidence

Figure 1.6e

H0∶ ρ1= ρ2= ρ3=⋯= ρn

Ha∶ at least one of the  ρi ≠ 0 , for i = 1,2,...,n

p-value < 2.2e-16

Since the p-value is less than the 5% level of significance, H0 is rejected. There is autocorrelation within the forecast residuals. The forecast can be further improved.

A graph of a graph of a graph

Description automatically generated with medium confidence

Figure 1.6f

H0∶ ρ1= ρ2= ρ3=⋯= ρn

Ha∶ at least one of the  ρi ≠ 0 , for i = 1,2,...,n

p-value = 0.002656

Since the p-value is less than the 5% level of significance, H0 is rejected. There is autocorrelation within the forecast residuals. The forecast can be further improved.

A graph of a graph of a graph

Description automatically generated with medium confidence

Figure 1.6g

H0∶ ρ1= ρ2= ρ3=⋯= ρn

Ha∶ at least one of the  ρi ≠ 0 , for i = 1,2,...,n

p-value = 0.03459

Since the p-value is less than the 5% level of significance, H0 is rejected. There is autocorrelation within the forecast residuals. The forecast can be further improved.

A graph of a graph of a graph

Description automatically generated with medium confidence

Figure 1.6h

H0∶ ρ1= ρ2= ρ3=⋯= ρn

Ha∶ at least one of the  ρi ≠ 0 , for i = 1,2,...,n

p-value = 0.05393

Since the p-value is more than the 5% level of significance, H0 is rejected. There is no autocorrelation within the forecast residuals. The residuals are white noise.

The residuals should adhere to the characteristics of white noise, which include a zero mean, no autocorrelation, and constant variance with an additional component of a normal distribution required. Through the utilisation of the checkresiduals() function in R, valuable information was obtained, including a time plot, an ACF plot, a histogram of the residuals, and the p-value resulting from the Ljung-Box test. The null hypothesis for this test assumes the absence of serial correlation in the residuals.

Upon examining the histograms of the four sets of forecast residuals, it is evident that these sets closely approximate a normal distribution. However, the mean of the histogram for the seasonal naive method differs from zero, unlike the other three models where the mean is indeed zero. This observation was further confirmed by the time plots for these four sets of forecasts, indicating that all forecast residuals have a mean of zero, except for the seasonal naive method plot. This suggests that the seasonal naive method for forecasting may not be suitable, as it introduces bias in the forecasted data. The time plots also reveal that only the ETS(M,Ad,M) and ETS(M,A,M) models maintain constant variance, while both the ETS(M,Ad,N) + STL decomposition model and the seasonal naive method do not, violating one of the properties of residuals in forecasting.

An analysis of the ACF plot for each set raised concerns. Among the 36 lags, 10 lags from the seasonal naive method, 6 lags from the ETS(M,Ad,N) + STL decomposition model, 4 lags from the ETS(M,Ad,M) model, and 3 lags from the ETS(M,A,M) model fell outside the critical value. This suggests the potential presence of correlation in these lags.

This judgement is further supported by the p-values obtained from the Ljung-Box test for the first three sets of residuals. In all cases, the p-values are lower than 0.05, leading to the rejection of the null hypothesis and providing strong evidence of serial correlation in the residuals. This violet the assumption of residuals as it indicates the presence of valuable information being left out that should be considered when computing the forecast. However, in the case of the ETS(M,A,M) forecast, the p-value from the Ljung-Box test exceeded 0.05. This suggests that the data may not exhibit correlations.

In conclusion, the forecast residuals obtained from the seasonal naive method do not exhibit the characteristics of white noise. They are neither constant, with a mean of zero, nor independent. Similarly, the forecast residuals from the STL decomposition + ETS(M,Ad,M) model do not resemble white noise, as they lack constancy and show autocorrelation. The forecast residuals of the ETS(M,Ad,M) model also do not meet the criteria for white noise, primarily due to their lack of independence. It is only the forecast residuals of the ETS(M,A,M) model that satisfy all three white noise requirements and exhibit a normal distribution. These residuals align with all four properties and assumptions of residuals, which is the desired outcome.

**Best goodness fit**

Based on the plots for the method and model as presented earlier, the fitted value plot for the seasonal naive forecast method in Figure 1.6a is eliminated from consideration due to its poor goodness of fit. This is evident from the visual inspection, where the actual training set plot and the fitted values exhibit substantial disparities in the magnitude of peaks and troughs of variances when compared to the other three models. Conversely, the models, namely STL decomposition with ETS(M,Ad,N), the exponential smoothing model of ETS(M,Ad,N), and ETS(M,A,M) generated by R, exhibit no significant differences in their goodness of fit when assessed visually. Consequently, all three models are subjected to further evaluation.

In order to make an informed selection, we refer to the information criteria presented in Table 1.6a. The AIC is used to assess the quality of model selection and goodness of fit, where a lower value corresponds to a better fit to the data while considering the complexity of the model. The ETS(M,Ad,N) model in combination with STL decomposition, yields the lowest AIC value of 8830.757, signifying a model with fewer initial states and estimated parameters, thus indicating its superiority in terms of goodness of fit.

This preference is further affirmed by the BIC, which also favours lower values to indicate a better fit but is generally more conservative regarding model complexity. BIC penalises the inclusion of additional parameters in the model. However, by prioritising the evaluation of the model's true fit, BIC sometimes compromises some predictive capability in favour of a model that more closely approximates actuality. The model with the lowest BIC value aligns with the model having the lowest AIC value, which, in this case, is the STL decomposition with ETS(M,Ad,N) with a BIC value of 8854.461.

In conclusion, both AIC and BIC criteria consistently point to the superiority of the STL decomposition with the ETS(M,Ad,N) model, which exhibits the best goodness of fit among the models considered in this analysis. This selection is supported by its lower AIC and BIC values, indicating a strong fit to the data while maintaining relative model simplicity.

By looking at the set error measures from Table 1.6a of the training set on RMSE, MAE, MAPE, and MASE, the forecasting model with STL decomposition with ETS(M,Ad,N) has the lowest error measures, which are 5842.2222, 4003.649, 3.7391 and 0.4849 respectively. Although a measure of forecast accuracy is not necessary for this case, it is a good practice.

**Plot of “Pre-Covid” data set with forecasts for the test set period**

**A graph showing the number of covid-19

Description automatically generated**

Figure 1.6i

**A graph of a graph showing the number of covid-19

Description automatically generated**

Figure 1.6j

Before describing and comparing forecast plots, it is essential to consider the factors influencing the data, such as the assumption that the future will resemble the past and the common belief that ongoing environmental changes will persist even if they don't. These constraints can lead forecasting methods or models to overlook events or circumstances occurring outside the training set period. Conversely, events that took place, particularly in the more recent training set data, may have enduring effects. If such events did occur, there would be significantly impacted forecasted data and cause the anticipated values to deviate from the actual values. Unfortunately, this happened in both the test and training sets. For instance, the Christchurch and Canterbury earthquakes, which occurred only during the training set (September 2010–December 2011), had a considerable impact on tourism arrivals in New Zealand due to safety concerns. These incidents led to a decrease in visitor arrivals during the training set period but not so much during the test set period (2012–2019). Additionally, the release of "The Hobbit: An Unexpected Journey" movie that was filmed in New Zealand, attracted "Lord of the Rings" fans, particularly from China, in November 2012. Despite its significant attraction to tourism, this event was not included in the training data.

Upon visual inspection of Figure 1.6j, it becomes evident that all four forecasting methods and models exhibit similar patterns of peaks and troughs at regular intervals when compared to the actual test set data. The influence of the smoothing parameter, gamma (γ), on this aspect, was minimal, as the seasonal pattern remained consistent between past and recent seasons. This suggests that all the models and methods effectively capture and replicate the seasonal fluctuations, regardless of the gamma value.

However, there are variations in the level of variability across the forecasts. The seasonal naive method and STL decomposition with the ETS(M,Ad,N) model exhibit constant variability, while the ETS(M,Ad,M) model, ETS(M,A,M) model, and the test set data display increasing variability. This divergence is attributed to the seasonal naive forecast method, which solely considers the most recent seasonal pattern and variability, repeating it in forecasts. Similar to the STL decomposition with ETS(M,Ad,N) model, seasonal components are forecasted using the seasonal naive method and seasonally adjusted data. The multiplicative seasonal model in ETS(M,Ad,M) and ETS(M,A,M) results in increasing variability. However, these levels of variability still differ from the actual values, primarily due to forecasting limitations in the assumption that the future will resemble the past. New Zealand's crisis management and response policies, particularly after the Canterbury earthquakes, only took effect after the training set period (2011) and significantly impacted tourist arrivals.

All forecast plots indicate positive trends except for the seasonal naive method, which does not account for trends when forecast. The STL decomposition + ETS(M,Ad,N) model exhibits the slowest trend increase due to its higher beta (β), influenced by the flatter trend in the most recent training data. The ETS(M,Ad,M) model displays a faster increase in trend than the previous model, driven by its lower β value, making it more responsive to the high rate of increase in past data trends. Damping in the forecast model reduces trend growth, explaining the flatter trend in both models. In contrast, the ETS(M,A,M) model has the highest trend increase rate due to its lowest β value, which is influenced more by historical data trends rather than the recent data's flatter trend and the absence of use of damping in the model.

However, even when compared to the ETS(M,A,M) model, the trend in the test set is increasing at an even higher rate. This outcome can be primarily attributed to forecasting limitations and the nature of the exponential smoothing model, which places greater weight on recent data. In this case, the decrease in tourism arrivals in New Zealand in 2011, primarily caused by the Canterbury earthquakes, coincided with the most recent year in the training data. Consequently, the forecasted data was heavily influenced by the post-earthquake data, displaying a flatter trend compared to normal circumstances. Conversely, the limitations of forecasting also stem from two significant factors contributing to the rapid growth in tourism since 2012. These factors were not considered during the forecast using the training data, which only extended up to 2011. They include the release of "The Hobbit: An Unexpected Journey" and the recovery efforts undertaken by the New Zealand government to rebuild infrastructure that was destroyed during the earthquakes, causing novelty and attracting tourism into the landscape.

**Evaluate & Compare Forecasting Performance**

**Traditional ApproachA black background with a black square

Description automatically generated with medium confidence**

Table 1.7a

The out-of-sample forecast accuracy for the forecasting methods or models is determined by calculating the forecast error, which represents the difference between each forecasted value and its corresponding observed value. These values are then subjected to various error measures. Some commonly used error measures include RMSE, which penalises larger errors more, MAE, which assigns equal weight to all errors regardless of their size or direction, and MAPE, which is scale-independent but meaningful only when all observed values are greater than zero. All of this data can be obtained using the "accuracy()" function in R Studio. A single forecast with a significant error can have a more substantial impact on the tourism sector than multiple forecasts with small errors. For example, forecasting and preparing for 10 tourists when the actual count is 100 tourists visiting in a day will have a more substantial impact than forecasting and preparing for 100 tourists when the actual count is 105 tourists visiting over multiple days. Therefore, while all error measures will be considered, RMSE will receive greater emphasis due to its more severe penalty for larger forecast errors. A smaller value in error measures typically indicates a more accurate forecast. Based on the error measures generated by R Studio, the ETS(M,A,M) forecast model provides the smallest value among all error measures. Therefore, the ETS(M,A,M) model forecasts best among the four models in terms of forecast accuracy.

**Time Series Cross-Validation**

**A black background with a black square

Description automatically generated with medium confidence**

Table 1.7b

Another reliable method for evaluating the best forecasting performance model is by using the forecast error generated by time series cross-validation. This forecast error can be obtained using the "tsCV()" function in R. As we are particularly concerned with larger forecast errors, the forecast error will be converted into RMSE again. The model with the smallest RMSE will indicate the most accurate forecast. According to Table 1.7b, the ETS(M,A,M) model forecasts best among the four models in terms of its forecast accuracy, as it produces the smallest RMSE value. This aligns with the conclusion obtained from the traditional approach.

However, having the most accurate forecast on the test set does not necessarily mean it is the most suitable forecast model. It simply implies that it was better in terms of accuracy. It is also essential to assess the overall context and the specific requirements of the area of application. As mentioned in Stage 6, there are significant factors contributing to the rapid growth in tourism that were not included in the training set. This can cause the out-of-sample forecast data to be closer to the actual value and misguide us on the accuracy of the forecast. Damping should be applied to the model if we are trying to compute longer-term forecasts. Therefore, the ETS(M,Ad,M) forecast model should be used for longer-term forecasts (h>75) to prevent over-forecasting, and the ETS(M,A,M) forecast model will be chosen for shorter-term forecasts (h<76) as it forecast best in terms of forecast accuracy.

**Implement Forecasts**

Since forecasting only 24 months ahead(shorter-term forecasts),  ETS(M,A,M) forecast Model will be used.

A black background with a black square

Description automatically generated with medium confidence

Table 1.8

A graph showing a number of tourists

Description automatically generated

Figure 1.8

The trend and seasonality in the forecasted plot generated by the ETS(M,A,M) model during the Covid period in Figure 1.8 align with the historical data. The low β value of 0.0033 indicates that trends are significantly influenced by historical data and less responsive to recent changes compared to the other forecast models. This explains the forecasted positive trends despite a slight decrease in the recent trend of historical data.

The moderate γ value, at 0.3489, implies that the forecasting model gives moderate weight to recent changes in the seasonal component of the data. This results in the forecasted values exhibiting a similar seasonal pattern as the historical data, seasonal duration and increase in the level of variability, which is also consistent with the chosen seasonal multiplicative model. The relatively low α value, at 0.2743, suggests that recent data are not heavily weighted in the forecast.

The narrow but high-confidence forecast intervals indicate that many uncertainties are ignored, making the interval reliable. The lower bounds of the interval suggest the possibility of a slight decrease, aligning with occasional dips observed in historical data. Additionally, the forecast indicates a potential extreme increase in the upper bound, consistent with the historical upward trend. Notably, there are no obvious outliers or unusual data points in the forecasts.

However, the forecast outcome differs significantly from the actual tourist arrivals during the "Covid" period, highlighting once again the limitations of forecasting. Historical data indicated a consistent increase in tourist arrivals, leading to the assumption that this trend would continue. However, the emergence of COVID-19 in December 2019 introduced a new and critical factor that significantly impacted the number of tourists visiting New Zealand. This unexpected event was not accounted for in the historical data, resulting in a future that deviated from the past and leading to a substantial disparity between forecasts and actual values. As this causes the time series to change at a point in time abruptly, we can conclude that a structural break has occurred.

**Stage 9: Quantifying Forecasted Loss in Tourism**

A screenshot of a report

Description automatically generated

Table1.9a

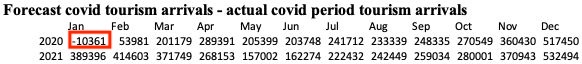


Table1.9b

A table of numbers and a few months

Description automatically generated with medium confidence

Table1.9c

Average visitor revenue :

total tourism expenditure 2018 + total tourism expenditure 2019 Total tourism arrival  from April 2017 to March 2019

Equation 1.9a

Average loss per month:

⨊(forecast covid tourism arrivals - actual covid period tourism arrivals)\*average cost per visitors  Total month with forecasted lost

Equation 1.9b

Upper value of range:

Max(Upper boundaries of prediction interval - actual covid tourism arrivals)\*average revenue per person

Equation 1.9c

Lower value of range:

Min(lower boundaries of prediction interval - actual covid tourism arrivals)\*average revenue per person

Equation 1.9c

Table 1.9a presents annual tourism expenditure by component from April 1, 2017, to March 31, 2021. This dataset is highly reliable due to the government's commitment to transparency and data quality assurance. Over the years, tourism revenue exhibited annual increases, but a significant drop occurred in 2021, coinciding with the COVID-19 pandemic. This adds credibility to the data, aligning with the trends and structural breaks observed in the tourism arrivals data we are analysing.

The data follows a fiscal year ending in March and starting in April of the previous corresponding year. Consequently, we could not use 2020 data as it extended until March 2022, which falls within our 'COVID period.' Thus, we will work with the data for 2018 and 2019, covering the period from April 2017 to March 2019, as the most recent data not affected by the 'COVID period.' All forecasted values for tourism arrivals will be rounded to integers using the 'as.integer()' function since tourism arrivals are discrete data. By computing Equation 1.9a in R, we find the average revenue per visitor arrival to be 10,883.43 NZD. Table 1.9b displays the differences between each month's forecasted COVID tourism arrivals and the corresponding actual tourism arrivals during the COVID period. From Table 1.9b, we observe that the actual values exceed our forecasts for January. Therefore, we will only consider the remaining 23 months where actual tourism arrivals are less than our forecasts. Applying the appropriate values for Equation 1.9b, we determine an average forecasted loss in tourism revenue incurred by New Zealand since the start of the pandemic, amounting to 3,073.881 million NZD per month.

The range of forecasted losses in tourism revenue is calculated using Equation 1.9c to obtain the upper value and Equation 1.9d to obtain the lower value. Since R provides two prediction intervals, we will conduct calculations for both. Again, all forecasted values for tourism arrivals will be rounded to integers with the 'as.integer()' function. Table 1.9c illustrates the differences between the boundaries and the actual tourism arrivals during the COVID period. In Table 1.9c, we observe that the actual values are greater than both the lower boundaries of our prediction interval for January. Therefore, we will only consider the remaining 23 months where actual tourism arrivals are lower than our forecasts. Applying the relevant values twice (for 80% PI and 95% PI), we determine the range of forecasted losses in tourism revenue using a 95% PI to be between 7,001.281 million NZD and 80.05854 million NZD, while the range using an 80% PI falls between 6,583.868 million NZD and 255.7063 million NZD.

Reference List

Cloke, P., & Perkins, H. C. (2002). Commodification and adventure in New Zealand tourism. Current issues in tourism, 5(6), 521-54.

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. Journal of sustainable tourism, 29(1), 1-20.

Henrickson, M. (2020). Kiwis and COVID-19: The Aotearoa New Zealand response to the global pandemic. The International Journal of Community and Social Development, 2(2), 121-133.

McNeill, J., & Asquith, A. (2022). End of the World: New Zealand’s Local Government and COVID-19 Je. In Local government and the COVID-19 pandemic: A global perspective (pp. 757-781). Cham: Springer International Publishing.

Munasinghe, S., Powell, L., Herath, H. M., & Frey, R. A. (2022). Lessons for resilience building in tourism and hospitality: the case of New Zealand. Worldwide Hospitality and Tourism Themes, 14(6), 604-609.

Tourism and the economy. Ministry of Business, Innovation and Employment. (n.d.). [https://www.mbie.govt.nz/immigration-and-tourism/tourism-research-and-data/tourism-data-releases/tourism-and-the-econ](https://www.mbie.govt.nz/immigration-and-tourism/tourism-research-and-data/tourism-data-releases/tourism-and-the-economy/)