COURSE

[redacted] TIME SERIES MODELLING AND FORECASTING

SEMESTER 1 2024/2025

TITLE

ASSIGNMENT 1

LECTURER

PROF. DR. ZAIDI ISA

NAME

AZRUL ZULHILMI BIN AHMAD ROSLI [redacted]

a. Three approaches to forecasting exist. Kindly provide an explanation.

The three approaches to forecasting are using qualitative models, time-series methods and causal methods.

i) Qualitative models: Based on expert opinions and intuition, used when historical data is unavailable or insufficient. Examples include the Delphi method, consumer market research, jury of executive opinion and sales force composite.

ii) Time-series methods: Utilize past data to predict future values by identifying patterns such as trends, seasonality, and cycles. Examples include ARIMA, exponential smoothing, moving averages, time trend projections, decompositions.

iii) Causal methods: Use cause-and-effect relationships between variables to make predictions. Regression analysis is a common causal method, where independent variables explain variations in the dependent variable. If there are more than one independent variables, multiple regression can be use.

b. What characteristics are essential for a reliable forecast?

A good forecast is characterized by its accuracy, reliability, and relevance to the specific context and objectives of the forecasting task. Accuracy is paramount, as a forecast should closely match actual outcomes or trends to enable informed decision-making and planning. This entails capturing the underlying patterns and dynamics within the data, such as trends, seasonality, and irregular fluctuations, with precision.

Reliability is equally essential, as stakeholders rely on forecasts to guide strategic decisions and allocate resources effectively. A reliable forecast should be based on sound methodologies, robust data analysis, and transparent assumptions, instilling confidence in its predictions and projections. Additionally, a good forecast should be relevant to the intended purpose, providing actionable insights and addressing the key questions and uncertainties faced by decision makers. This requires aligning the forecasting approach with the specific needs and requirements of the organization or project, considering factors such as time horizon, level of detail, and degree of uncertainty. Ultimately, a good forecast empowers stakeholders with accurate, reliable, and relevant information to navigate uncertainty, mitigate risks, and seize opportunities effectively.

c. Describe the properties of the data consisting of time series.

Time series data exhibits distinct characteristics that distinguish it from other types of data, primarily due to its sequential nature and temporal dependencies. One key characteristic is temporal ordering, where observations are collected and recorded at regular intervals over time, such as hourly, daily, monthly, or yearly intervals. This temporal structure enables analysts to identify trends, seasonality, and patterns that evolve over time, providing insights into underlying dynamics and relationships. Here the explanations of each characteristics:

i) Trend: The trend component accounts for the gradual shifting of the time series to relatively higher or lower values over a long period of time. Trend is usually the result of long-term factors such as changes in population, demographics, technology, or consumer preferences.

ii) Seasonality: The seasonal component accounts for regular patterns of variability within certain time periods such as a year. The variability does not always correspond with the season of the year. The variability does not always correspond with the seasons of the year such as winter, spring, summer, fall. There can be for example, within-week or within-day "seasonal" behaviour.

iii) Cyclical patterns: Any regular pattern of sequences of values above and below the trend line lasting more than one year can be attributed to the cyclical component. Usually, this component is due to multiyear cyclical movements in the economy.

iv) Irregular effects: Unpredictable and short-term variations due to unseasonable weather/natural disasters, strikes, sampling errors and nonsampling error.

Additionally, time series data often displays autocorrelation, where observations at one time point are correlated with observations at previous time points. This autocorrelation reflects the persistence of patterns and influences the choice of modeling techniques, such as autoregressive models, designed to capture these dependencies.

Moreover, time series data may exhibit non-stationarity, where statistical properties such as mean and variance change over time. Detecting and addressing non-stationarity is crucial for accurate modeling and forecasting. Understanding these characteristics is essential for effectively analyzing and interpreting time series data, enabling analysts to uncover insights, identify trends, and make informed forecasts and decisions based on historical patterns and trends.

d. Provide a comprehensive explanation of the time series normalization methods used in forecasting.

Normalization methods are especially important when the data varies in scale, units, or has different levels of variability, which can affect the performance of forecasting algorithms. The goal of time series normalization is to make the data easier for models like ARIMA to process and generate accurate forecasts.The time series normalization methods used in forecasting are min-max scaling, z-score normalization, logarithmic transformations and differencing.

i) Min-Max Normalization: Min-max normalization scales the data into a fixed range, usually [0, 1], by subtracting the minimum value and dividing by the range which is the difference between the maximum and minimum values.

Formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Where:

- x' is the normalized value.
- x is the original value.
- min(x) and max(x) are the minimum and maximum values of the dataset.

Use case: This method is useful when want to scale the data to a specific range for neural networks, where the model benefits from having inputs in a uniform range. However, this method is sensitive to outliers, as they will affect the min and max values.

Pros:

- Transforms the data into a defined range (e.g., [0, 1]).
- Useful for algorithms that rely on the distance between data points.

Cons:

- Sensitive to outliers, which can distort the normalization.
- Does not handle non-stationary data well.

ii) Z-Score Normalization (Standardization): Z-score normalization standardizes the data by subtracting the mean and dividing by the standard deviation. This centers the data around 0 with a standard deviation of 1.

Formula:

$$z = \frac{x - \mu}{\sigma}$$

Where:

- z is the normalized value (z-score).
- x is the original value..
- μ is the mean of the dataset.
- σ is the standard deviation of the dataset.

Use case: Z-score normalization is often used when the data follows a Gaussian (normal) distribution. It's useful in many machine learning algorithms, such as linear regression, SVM, and neural networks, where data is assumed to have a mean of zero and a standard deviation of one.

Pros:

- Useful for algorithms that assume normality of data (e.g., linear regression, PCA).
- Robust to outliers (though they can still influence the mean and standard deviation).
- Data with different units or scales will be standardized to the same scale.

Cons:

- May not work well with data that is heavily skewed or has extreme outliers.
- Assumes that the data is approximately normally distributed.

iii) Log Transformation: Log transformation is applied to reduce the effect of large outliers or variance in a dataset by compressing the range of values. It takes the natural logarithm of the data values.

Formula:

$$x' = \log(x + 1)$$

Where:

- x' is the transformed value.
- x is the original value.

Use case: This transformation is especially useful for time series data that is positively skewed or has exponential growth patterns (like financial data, population growth, etc.). Log transformation helps stabilize variance and make the data more homoscedastic (constant variance over time).

Pros:

- Reduces skewness and helps stabilize variance.
- Improves performance for models that assume constant variance, such as ARIMA or other linear models.

Cons:

- The method is only applicable for positive data.
- Log transformation can distort relationships when applied indiscriminately.

iv) Differencing: Differencing is a technique used to transform a time series dataset into a stationary series by subtracting the previous value from the current value. This method is frequently used to remove trends and seasonality.

Formula:

$$y'_t = y_t - y_{t-1}$$

Where:

- yt′: the differenced series.
- yt and yt_{-1}: the current and previous values of the time series.

Use case: Differencing is especially used in time series models like ARIMA to make a non-stationary series stationary, which is a prerequisite for many time series forecasting models.

Pros:

- Helps in removing trends and seasonality in the data.
- Makes data stationary, which is essential for many time series forecasting methods.

Cons:

- Increases the complexity of the model by adding lagged terms.
- May lose information about the original series, which could be important for forecasting.

e. What distinguishes smoothing techniques from decomposition approaches in time series forecasting?

Smoothing techniques and decomposition methods are both valuable approaches used in time series forecasting, albeit with distinct methodologies and outcomes. Smoothing techniques aim to reduce noise and smooth out fluctuations in the data to identify the underlying pattern like remove the random variation. Smoothing is typically applied when identifying the general trend or level of the data, often making short-term variations less pronounced. Decomposition approaches break down a time series into its component parts, specifically for the trend, seasonality, and residuals (noise). The goal is to separate these components for better analysis and forecasting. Decomposition provides a clear structure to the time series, making it easier to model each component independently. It is also used for deeper analysis of time series structure and long-term forecasting.

Common smoothing methods are:

- Arithemic Moving Average: Smooths the series by averaging a fixed number of previous data points.
- Smoothing Model: Weights past observations with exponentially decreasing weights.
- Holt-Winters Methods: A type of exponential smoothing method that accounts for both trend and seasonality.

Common decomposition methods are:

- Classical Decomposition: Decomposes a series into components like trend, seasonality, and residuals using moving averages.
- Seasonal and Trend decomposition using Loess (STL): A flexible method for decomposing time series data. It segregates the time series into trend, seasonality, and residual components. By individually analyzing each component, analysts gain insight into the nuanced patterns inherent in the data, including long-term trends and seasonal variations.

In summary, while both smoothing techniques and decomposition methods serve the purpose of uncovering underlying patterns in time series data, they vary in their methodologies and depth of analysis. Smoothing techniques prioritize noise reduction and trend identification through averaging, while decomposition methods provide a more comprehensive breakdown of the time series into its constituent elements, facilitating nuanced insights into the data's structure. The selection between these approaches depends on the specific characteristics of the time series and the level of detail required for accurate forecasting and analysis.

f. Explain the distinction between cyclical and seasonal effects in time series.

In time series analysis, distinguishing between cyclical and seasonal effects is crucial for understanding the underlying patterns and dynamics within the data. Cyclical components in time series data refer to patterns that fluctuate above and below the trend line, typically lasting more than a year. These cycles are irregular in length and can span several years or even decades, making them difficult to predict. They are often driven by macroeconomic factors such as economic growth and recession cycles or external events like political changes or global crises.

In contrast, seasonal components involve predictable, recurring patterns that occur at fixed intervals, usually within a year, but they are not necessarily tied to the traditional seasons (winter, spring, summer, fall). Seasonal patterns can also occur on shorter time scales, such as daily, weekly, or monthly cycles, driven by factors like weather, holidays, or societal behaviors. These effects are highly predictable, as they follow regular intervals, such as increased retail sales during holidays or changes in behavior on weekends.

The key difference lies in their time frame and predictability: cyclical patterns are long-term and irregular, while seasonal patterns are shorter, fixed, and much easier to anticipate. Additionally, cyclical effects are often driven by broader economic or external factors, whereas seasonal effects are primarily influenced by natural, societal, or commercial events.

g. What are the criteria for assessing the forecasting performance of a model?

The forecasting performance of a model can be assessed using several key criteria that evaluate its accuracy, reliability, and predictive power. Some of the most common criteria are:

i. Accuracy Merics:

  a. Mean Absolute Error (MAE): Measures the average magnitude of errors in a set of predictions, without considering their direction. It is calculated as the average of the absolute differences between the predicted and actual values.
  b. Mean Squared Error (MSE): Similar to MAE but penalizes larger errors more heavily by squaring the differences. It is more sensitive to outliers.
  c. Root Mean Squared Error (RMSE): The square root of MSE, which provides a measure of error in the original units of the data, making it easier to interpret than MSE.
  d. Mean Absolute Percentage Error (MAPE): Expresses the forecast error as a percentage of the actual values. It is useful for comparing the relative accuracy of forecasts across different datasets.
  e. Symmetric Mean Absolute Percentage Error (SMAPE): A variation of MAPE that prevents the error from being biased by very small values in the actual data.
  f. Relative Absolute Error (RAE): Compares the forecast errors to the total variation in the actual data.

ii. Forecasting Bias:

  a. Bias or Menad Forecast Error (MFE): Measures the average of the errors. If the average error is consistently positive or negative, the model might be biased such consistently overestimating or underestimating.
  b. Tracking Signal: A statistic used to detect any bias in the forecast model over time. It's the cumulative sum of forecast errors divided by the mean absolute deviation. A tracking signal outside a predefined range suggests the model is biased.

iii. Error Distribution:

  a. Residual Analysis: Examining the distribution of forecast errors (residuals). Ideally, residuals should be randomly distributed which is no patterns. Systematic patterns in residuals could indicate that the model is misspecified.

iv. Complexity and Model Fit:
   a. Overfitting and Underfitting: Evaluating whether the model is too complex (overfitting) or too simple (underfitting) by comparing its performance on both training and validation data.
   b. Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC): These are used to assess model fit while accounting for the number of parameters, with lower values indicating better model performance.

v. Reliability: Assessing the consistency and stability of the forecasting model over time. Reliability metrics consider the variability of forecast errors and the model's ability to provide consistent predictions across different time periods. Consistent and stable forecast performance across various time periods enhances the reliability and trustworthiness of the model.

vi. Ability of the model to capture and adjust to changes in underlying patterns and dynamics within the data is critical. Models should be able to adapt to changes in trends, seasonality, and other factors influencing the time series data. Monitoring the model's performance over time and its ability to adapt to changing conditions is essential for ensuring its long-term effectiveness.

vii. Interpretability and simplicity: A model's interpretability refers to its ease of understanding and the ability to explain how predictions are generated. Simple models that are easy to interpret and implement are often preferred, as they are more transparent and facilitate better understanding and communication of forecasting results.

viii. Robustness to outliers.

h. Discuss the distinctions between additive and multiplicative models, as well as the criteria for selecting the more suitable model.

The differences between additive and multiplicative models lie in how they handle the components of a time series—trend, seasonality, and randomness. In an additive model, these components are combined by adding them together, whereas in a multiplicative model, they are multiplied. The distinction between additive and multiplicative models are:

i. Additive model: When the size of the seasonal component stays about the same as the trend changes, then an additive methos is usually best. Assumes the dependent variable is the sum of contributions from independent variables. This means that the observed values are equal to the sum of the trend, seasonal variation, and random noise at each time point. Additive models are typically used when the magnitude of the seasonal fluctuations does not depend on the level of the time series. For example, if the seasonal variations remain relatively constant over time regardless of the overall level of the series, an additive model may be appropriate.

Formula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

- Relationships between variables are linear and independent.

ii. Multiplicative model: Assumes the dependent variable is the product of contributions from independent variables. In this case, the observed values are equal to the product of the trend, seasonal variation, and random noise at each time point. Multiplicative models are often used when the magnitude of the seasonal fluctuations depends on the level of the time series. For instance, if the seasonal variations grow or shrink proportionally with the overall level of the series, a multiplicative model may be more suitable.

Fomula:

$$Y = \beta_0 \cdot X_1^{\beta_1} \cdot X_2^{\beta_2} \cdot e^{\epsilon}$$

- Typically log-transformed for linearization:

$$\ln(Y) = \ln(\beta_0) + \beta_1 \ln(X_1) + \beta_2 \ln(X_2) + \epsilon$$

iii. Criteria for Model Selection: When size of the seasonal component increases as the trend increases, then a multiplicative method may be better.

- Nature of data: Use additive models if the relationships between variables are linear and use multiplicative models if variables interact or exhibit exponential growth/decay.
- Model fit: Compare models using metrics like Adjusted R-squared, Akaike Information Criterion (AIC), or Bayesian Information Criterion (BIC).
- Residual analysis: Evaluate residual patterns; multiplicative models may correct non-constant variance or nonlinearity.

In summary, the choice between additive and multiplicative models depends on the relationship between the level of the time series and the magnitude of its seasonal fluctuations. Additive models are appropriate when the seasonal variations remain constant regardless of the level of the series, while multiplicative models are preferred when the seasonal fluctuations are proportional to the level of the series. Understanding these differences helps analysts select the most appropriate model for their time series data and improve the accuracy of their forecasts.

i. What distinguishes the moving average model from the exponential smoothing model?

Both moving average and exponential smoothing are time series forecasting methods, but they differ in their approach to weighting historical data, responsiveness to changes, and computational complexity.

i. Moving Average Model: A simple method where forecasts are calculated as the average of the last n observations in the time series.

Fomula:

$$\text{Forecast at time } t = \frac{1}{n}\sum_{i=1}^{n} Y_{t-i}$$

ii. Exponential Smoothing Model: A forecasting method that assigns exponentially decreasing weights to past observations to predict future values, giving more importance to recent data.

Formula:

$$F_t = \alpha Y_{t-1} + (1 - \alpha) F_{t-1}$$

Where:

- F_t: Forecast at time t
- Y_{t-1}: Actual observation at time t-1
- A: Smoothing factor ($0 < \alpha < 1$)

So, exponential smoothing carries all past history forever while moving average model eliminates bad data after N periods.Based on weighting of data, moving average model must have all observations withing the window while exponential smoothing it exponentially decreasing weight and recent data has more influence. Then, adaptability to change for moving average model is slower and less responsive to recent change while exponential smoothing model is more faster and highly responsive to recent changes. The complexity of moving average model is simple to compute and use a fixed number of observations while exponential smoothing is requiring parameter selection (α) and iterative computation. Then, moving average model is poor to handle trends and seasonality while exponential smoothing can be extended to account for trends and seasonality. Lastly, lagging effect for moving average model is significant lag especially for larger windows while exponential smoothing is less lag due to higher weighting of recent observations.

j. What is the key difference between exponential smoothing methods and Holt-Winters models?

Exponential smoothing methods and Holt-Winter models are both techniques used in time series forecasting, but they differ in complexity and the components they incorporate. The key difference between exponential smoothing methods and Holt-Winters models lies in their ability to handle trends and seasonality in time series data.

i. Exponential Smoothing Model: A forecasting method that assigns exponentially decreasing weights to past observations to predict future values, giving more importance to recent data. Focuses on smoothing the time series by prioritizing recent data points, making it suitable for short-term forecasting of stationary data which is no trend or seasonality.

Types:

a. Simple Exponential Smoothing (SES):
   i. Assume no trend or seasonality.
   ii. Formula:

$$F_t = \alpha Y_{t-1} + (1 - \alpha) F_{t-1}$$

   iii. Best for stationary time series which is data with constant mean and variance.

b. Holt's Linear Exponential Smoothing:
   i. Extends simple exponential smoothing to handle trends in the data.
   ii. Adds a trend component to the model:

$$F_t = \alpha Y_{t-1} + (1 - \alpha)(F_{t-1} + T_{t-1})$$

   Where:
   - T_t: Trend factor

c. SES is primarily used for forecasting data without trend or seasonality, while Holt's method incorporates trend by considering both level and trend components in the smoothing process.

ii. Holt-Winters Models: A further extension of exponential smoothing that incorporates both trend and seasonality components, making it suitable for forecasting time series with both patterns. It also known as triple exponential smoothing where comprise three smoothing equations: one for the level, one for the trend, and one for the seasonal component. These models enable analysts to capture and forecast seasonal variations, trends, and level changes in the data, making them more robust and flexible for forecasting tasks compared to simple exponential smoothing methods.

Types of Holt Winter Models:

a. Additive Model: Used when seasonal variations are constant over time.

   i. Formula:

$$F_{t+m} = (L_t + mT_t) + S_{t+m-s}$$

   Where:

   - L_t: Level (baseline value of the time series at time t)
   - T_t: Trend (rate of increase or decrease at time t)
   - S_{t+m-s}: Seasonal component

b. Multiplicative Model: Used when seasonal variations increase or decrease proportionally to the level of the series.

   i. Formula:

$$F_{t+m} = (L_t + mT_t) \times S_{t+m-s}$$

   Where:

   - L_t: Level (baseline value of the time series at time t)
   - T_t: Trend (rate of increase or decrease at time t)
   - S_{t+m-s}: Seasonal component

The key difference between exponential smoothing and Holt-Winters Models are on trend and seasonality handling, use case, complexity and forecasting formula. Basic exponential smoothing model cannot handle trends except Holt's method while Holt-Winter models explicitly incorporate a trend component. Seasonality handling cannot be handled by exponential smoothing while Holt-Winter models can handle both additive (constant seasonality) and multiplicative (proportional seasonality).

Forecasting formula for exponential smoothing is focused on level smoothing while Holt-Winter models include level, trend and seasonality so it is more complex due to trend and seasonality component. Exponential smoothing is suitable for short-term forecasting of stationary data while Holt-Winter models suitable for time series with trend and seasonality.

k. Please provide an explanation of the time trend forecasting model.

Trend projection involves fitting a trend line to historical data points and extending it into the future to forecast medium to long-term values. The time trend forecasting model, a type of regression analysis, predicts future values by identifying the long-term trend in the data, which can be expressed as a mathematical function of time. This model establishes the relationship between a dependent variable (Y) and time (t), assuming the changes in Y follow a predictable linear or nonlinear pattern.

The simplest trend equation is a linear model, derived using regression analysis, which assumes the rate of change in Y is constant over time. More complex models, such as exponential or quadratic equations, capture nonlinear trends. A trend reflects long-term growth or decline in a time series, taking various forms depending on the data's behaviour. In conclusion, time trend is simple to use and interpret, effective for data with clear and predictable trends and sensitive to outliers which can distort the trend estimation. However, it works best when the time series lacks strong seasonal or cyclical fluctuations.

l. Describe the linear and nonlinear time series trend forecasting models in detail.

Trend Time series trend forecasting models are tools used to predict the future behavior of a dataset based on its historical trend. These models can be broadly categorized into linear and nonlinear models, depending on the relationship between the dependent variable (Y) and the independent variable (t, time).

i. Linear Time Series Trend Forecasting Model: The linear trend model assumes a constant rate of change in the dependent variable (Y) over time (t). This means the relationship between Y and t is a straight line.

Model equation:

$$Y_t = a + bt + \epsilon_t$$

Where:

- Y_t: Value of the dependent variable at time *t*.
- *a*: Intercept (value of *Y* when t=0).
- *b*: Slope (rate of change of *Y* per unit of time)
- *t*: Time (independent variable)
- ϵ_t: Error term (captures random variations not explained by the trend).

Characteristics:

- Constant Rate of Change: The slope (b) remains constant, meaning the increase or decrease in Y is uniform over time.
- Straight Line: When plotted, the trend appears as a straight line.
- Simple to Use: Requires fewer computations and is easy to interpret.

ii. Nonlinear Time Series Trend Forecasting Model: The nonlinear trend model assumes the rate of change in Y varies over time. This means the relationship between Y and t is not a straight line but can take other forms, such as exponential, quadratic, or logarithmic.

Type of Nonlinear Model:

- Exponential trend model: Used when the dependent variable grows or decays at an increasing or decreasing rate.
- Quadratic trend model: Used for trends that follow a parabolic shape, such as data that increases initially and then decreases, or vice versa.
- Logarithmic trend model: Used when the rate of change decreases over time.
- Power trend model: Describes situations where the dependent variable grows or declines according to a power function.

Characteristics:

- Varying Rate of Change: The trend curve changes its steepness over time.
- Flexible: Captures complex behaviours in the data that cannot be represented by a straight line.
- More Computationally Intensive: Requires advanced methods to estimate parameters.

The choice between linear and nonlinear models depends on the characteristics of the time series data and the nature of the trends and patterns present. Linear models offer simplicity and ease of interpretation but may be inadequate for capturing nonlinear dynamics. Nonlinear models provide greater flexibility and accuracy in modeling complex time series data but may require more computational resources and expertise for model estimation and interpretation. Ultimately, the selection of the appropriate forecasting model depends on the specific requirements of the forecasting task and the complexity of the underlying time series dynamics.

m. Perform an empirical analysis on forecasting employing smoothing, decomposition, and time trend methodologies. (Any available datasets may be utilized)

CHAPTER I

INTRODUCTION

1.1 TIME SERIES ANALYSIS

Time series analysis is a statistical approach used to examine and interpret data collected over time. It focuses on identifying patterns, trends, and relationships within sequential data points. The key aspects of time series analysis include descriptive analysis, visualization, modeling, and forecasting.

Descriptive analysis involves summarizing the key characteristics of the dataset using statistical measures such as mean, median, and standard deviation. Visualization techniques, such as line plots and histograms, help in detecting trends, patterns, and anomalies within the data.

Modeling methods, including autoregressive integrated moving average (ARIMA) models and exponential smoothing, are used to identify underlying structures and dynamics in the data. These models play a vital role in forecasting future values and analyzing relationships between variables.

Forecasting, an essential component of time series analysis, utilizes historical data to predict future values. Various techniques, including extrapolation, moving averages, and advanced machine learning models, are applied to improve prediction accuracy and support decision-making.

1.2 IMPORTANCE OF TIME SERIES

Time series analysis is highly valuable across multiple fields, as it helps extract meaningful insights from data collected over time. In economics and finance, it is essential for studying stock market trends, predicting economic indicators such as GDP growth and inflation rates, and recognizing financial market patterns to support investment decisions. By analyzing economic cycles, financial analysts can anticipate market fluctuations and minimize risks associated with market volatility.

In business and marketing, time series analysis assists organizations in forecasting sales, monitoring consumer behavior, and detecting seasonal demand patterns for products and services. By leveraging historical sales data, businesses can enhance inventory management, improve marketing strategies, and allocate resources more effectively to meet customer needs. Forecasting methods allow companies to respond to shifting market conditions, predict consumer preferences, and maintain a competitive edge.

In environmental science and climate research, time series analysis is instrumental in tracking and predicting changes in weather patterns, sea levels, and environmental conditions. Researchers use time series data to examine long-term climate trends, evaluate the effects of human activity on the environment, and devise strategies to combat climate change. By identifying trends, anomalies, and cycles in environmental data, time series analysis supports informed policymaking for conservation and sustainability initiatives.

1.3 OBJECTIVE

1. To conduct an empirical study forecasting using smoothing, decomposition, and time trend models.

CHAPTER II

DATA SOURCE

2.1 DATA DESCRIPTION

The dataset focuses on "Advance Retail Sales: Retail Trade" from the U.S. Census Bureau, obtained via Federal Reserve Economic Data (FRED) in the United States. It includes total monthly sales estimates, incorporating e-commerce sales. The selected data spans from 1992 to the end of 2024, covering the retail trade and food services industries. The sales figures are expressed in millions of dollars and are not seasonally adjusted, with data recorded on a monthly basis.

2.2 VARIABLES

The dataset likely contains two variables which are:

1. Period: The date of each data point in monthly.

2. Value: Total sales in month.

The dataset spans from 1992 to the end of 2024 and records sales on a monthly basis. The primary variable in this dataset is Advance Retail Sales, measured in millions of U.S. dollars, which includes both in-store and e-commerce sales. These sales figures are not seasonally adjusted, meaning they reflect raw sales data without adjustments for predictable seasonal fluctuations such as holidays or changing consumer trends. The dataset classifies industries into retail trade and food services, covering businesses like department stores, electronics retailers, restaurants, and cafes. With its monthly frequency, this dataset allows for analyzing sales trends, identifying seasonal patterns, and forecasting future retail performance, making it a valuable tool for understanding consumer spending behavior over time.

CHAPTER III

METHODOLOGY

3.1 METHODS

The simulation is made in Risk Simulator and eight types of models has been used to analyze and predict future outcomes based on historical data. The types of models applied in the software are as below in Table 3.1.

Table 3.1 Type of model

Model	Description
Single Moving Average	Asimple smoothing technique that calculates the average of a fixed number of past observations to forecast future values. It helps in reducing noise and identifying trends but does not adapt well to sudden changes in the data.
Single Exponential Smoothing	Assigns exponentially decreasing weights to past observations, giving more importance to recent data points. It is useful for short-term forecasting when data has no clear trend or seasonality.
Double Moving Average	Applies the moving average technique twice to smooth out fluctuations and capture underlying trends better. It is more effective for time series with trends compared to the Single Moving Average.
Double Exponential Smoothing	Also known as Holt's Method, extends Single Exponential Smoothing by incorporating a trend component. It consists of two equations: one for smoothing the level and another for capturing the trend, making it suitable for data with linear trends but no seasonality.
Seasonal Additive	Assumes that the seasonal effect is constant over time and can be added to the trend. It is suitable when seasonal variations remain relatively stable in magnitude.
Seasonal Multiplicative	Assumes that the seasonal effect varies proportionally to the trend. It is useful when seasonal fluctuations increase or decrease in magnitude over time.

Holt-Winter's Additive	Extends Holt's Method by adding a seasonal component that is added to the trend and level. It works best when seasonal variations are relatively constant over time.
Holt-Winter's Multiplicative	An extension of Holt's Method, where the seasonal component is multiplied by the trend. This model is preferred when seasonal fluctuations change in magnitude over time, making it effective for data with increasing or decreasing seasonal effects.

3.2 EVALUATION OF MODEL

To evaluate the performance of each forecasting model where several metrics can be used for comparison. The metrics is as shown as in Table 3.2.

Table 3.1 Type of model

Evaluation Metrics	Description
Root Mean Squared Error (RMSE)	RMSE measures the average magnitude of the errors between predicted and actual values. Lower RMSE values indicate better accuracy and closer alignment between predicted and actual values.
Mean Absolute Error (MAE)	MAE calculates the average absolute difference between predicted and actual values. It provides a measure of the average magnitude of errors without considering their direction.
Mean Absolute Percentage Error (MAPE)	MAPE measures the percentage difference between predicted and actual values relative to actual values. It provides insight into the relative accuracy of forecasts across different scales of data.
Mean Squared Error (MSE)	MSE calculates the average of the squared differences between predicted and actual values. It penalizes large errors more heavily than smaller ones and is commonly used in conjunction with RMSE.
Theil's U Statistic	Theil's U statistic compares the performance of the forecasting model to a naive forecasting method. A Theil's U value less than 1 indicates that the forecasting method performs better than a simple naive method.
Forecast Bias	Forecast bias measures the tendency of the model to systematically overestimate or underestimate actual values. Positive bias indicates the model consistently overestimates, while negative bias indicates underestimation.
Residual Analysis	Residual analysis involves examining the differences between predicted and actual values to identify patterns or trends in the errors. It

	helps assess whether the model captures all relevant information in the data and if any systematic errors exist.

CHAPTER IV

RESULT AND DISCUSSION

4.1 ANALYSIS

Analysis was done in Risk Simulator in Microsoft Excel and Figure 4.1 shows the best model that was used is Holt-Winter's Multiplicative model.

Methodology

The Best Model: Holt-Winter's Multiplicative	
Alpha	0.2030
Beta	0.0190
Gamma	0.0303
Seasons	4
Root Mean Squared Error (RMSE)	25332.5345
Mean Squared Error (MSE)	0
Mean Absolute Deviation (MAD)	19465.0425
Mean Absolute Percentage Error (MAPE)	5.28%
Theil's (U)	0.6979

Second Best Model: Holt-Winter's Additive	
Alpha	0.1923
Beta	0.0214
Gamma	0.0315
Seasons	4
Root Mean Squared Error (RMSE)	25566.8046
Mean Squared Error (MSE)	653661495.5019
Mean Absolute Deviation (MAD)	18846.5229
Mean Absolute Percentage Error (MAPE)	5.10%
Theil's (U)	0.7045

Third Best Model: Seasonal Multiplicative	
Alpha	0.2824
Gamma	0.0317
Seasons	4
Root Mean Squared Error (RMSE)	25926.9292
Mean Squared Error (MSE)	672205657.4363
Mean Absolute Deviation (MAD)	19754.3189
Mean Absolute Percentage Error (MAPE)	5.31%
Theil's (U)	0.7145

Fourth Best Model: Seasonal Additive	
Alpha	0.2759
Gamma	0.0342
Seasons	4
Root Mean Squared Error (RMSE)	26253.3961
Mean Squared Error (MSE)	689240806.0804
Mean Absolute Deviation (MAD)	19382.4814
Mean Absolute Percentage Error (MAPE)	5.20%
Theil's (U)	0.7235

Fifth Best Model: Double Moving Average	
	12
Root Mean Squared Error (RMSE)	28743.7836
Mean Squared Error (MSE)	826205094.9068
Mean Absolute Deviation (MAD)	20022.4918
Mean Absolute Percentage Error (MAPE)	5.16%
Theil's (U)	0.7616

Sixth Best Model: Double Exponential Smoothing	
Alpha	0.1666
Beta	0.0235
Root Mean Squared Error (RMSE)	29019.7165
Mean Squared Error (MSE)	842143948.2464
Mean Absolute Deviation (MAD)	20761.7237
Mean Absolute Percentage Error (MAPE)	5.62%
Theil's (U)	0.8068

Seventh Best Model: Single Moving Average	
	7
Root Mean Squared Error (RMSE)	29422.1359
Mean Squared Error (MSE)	865662080.7803
Mean Absolute Deviation (MAD)	21991.2332
Mean Absolute Percentage Error (MAPE)	5.80%
Theil's (U)	0.8032

Eighth Best Model: Single Exponential Smoothing	
Alpha	0.2428
Root Mean Squared Error (RMSE)	29845.4658
Mean Squared Error (MSE)	890751826.0357
Mean Absolute Deviation (MAD)	22353.8292
Mean Absolute Percentage Error (MAPE)	5.98%
Theil's (U)	0.8287

Figure 4.1 Best Model in Risk Simulator

In exponential smoothing models, the parameters alpha (α), beta (β), and gamma (γ) play crucial roles in controlling the smoothing of different components: level, trend, and seasonality. The alpha (α) parameter is used in Single Exponential Smoothing (SES), Holt's Method, and Holt-Winters models, where it determines how much weight is assigned to the most recent observation compared to past values. A higher α (close to 1) makes the model more responsive to recent changes, while a lower α (close to 0) results in a smoother but slower adaptation to new data.

The beta (β) parameter, used in Double Exponential Smoothing (Holt's Method) and Holt-Winters models, controls the influence of recent trends in the data. A higher β makes the model adapt quickly to trend changes, but if set too high, it can cause instability. Conversely, a lower β provides a more stable trend estimate but reacts slower to trend shifts.

The gamma (γ) parameter, specific to Holt-Winters Additive and Multiplicative models, governs how the seasonal component is updated over time. A higher γ allows the model to adapt quickly to seasonal changes, making it useful for volatile seasonal patterns, while a lower γ maintains past seasonal patterns for longer, ensuring stability in forecasting.

In Holt-Winters models, these parameters work together to update the level, trend, and seasonal components iteratively. The additive model assumes that seasonal variations remain constant over time, while the multiplicative model accounts for seasonal variations that scale with the trend. These parameters are typically optimized using techniques such as grid search or by minimizing Mean Squared Error (MSE) to ensure the best forecasting performance. Proper tuning of α, β, and γ is essential for achieving accurate forecasts, especially in datasets with complex trends and seasonality.

The Holt-Winters' Multiplicative model emerges as the best forecasting model for the "Advance Retail Sales: Retail Trade" dataset based on a detailed analysis of key performance metrics. This model demonstrates the lowest Root Mean Squared Error (RMSE) of 25,332.53, which penalizes large deviations more heavily, making it the most accurate in minimizing significant forecasting errors. Additionally, it achieves a

Mean Absolute Percentage Error (MAPE) of 5.28%, ranking third lowest among all models. While the Double Moving Average model has a slightly lower MAPE (5.16%), the Holt-Winters Multiplicative model balances MAPE, RMSE, and Theil's U Statistic, making it the most reliable overall. The Theil's U Statistic of 0.6979, the lowest among all models, further confirms that this model significantly outperforms a naive forecast.

The dataset consists of non-seasonally adjusted monthly retail sales data from 1992 to 2024, which typically exhibits growing seasonal trends such as increasing holiday sales spikes over time. The multiplicative model effectively captures this pattern since it allows seasonal variations to scale with the trend, unlike the additive model, which assumes a constant seasonal effect. The smoothing parameters used in this model are alpha (α) at 0.203 for level smoothing, beta (β) at 0.019 for trend smoothing, and gamma (γ) at 0.0303 for seasonal smoothing. It indicates that the model moderately adjusts to recent observations, maintains a stable trend, and responds well to seasonal fluctuations.

Other models underperformed for various reasons. The Holt-Winters Additive model assumes constant seasonal variations, which is less suitable for retail sales where seasonal peaks, such as holiday spending, tend to grow over time. This limitation is reflected in its slightly higher RMSE (25,566.80) and Theil's U (0.7045). Meanwhile, simpler methods like Single and Double Moving Averages, as well as Exponential Smoothing, fail to account for seasonality, resulting in significantly higher RMSE values exceeding 28,000, making them unsuitable for retail data that is heavily influenced by seasonal cycles. Even Seasonal Additive and Seasonal Multiplicative models, while incorporating seasonality, lack a trend component (beta parameter), causing them to fit the data less effectively compared to Holt-Winters.

In conclusion, the Holt-Winters Multiplicative model is the optimal choice for forecasting this dataset. It accurately captures both trend and evolving seasonal patterns, minimizes large forecasting errors with the lowest RMSE, and balances short-term responsiveness with long-term stability, as evidenced by its superior Theil's U score. This makes it particularly well-suited for forecasting non-seasonally adjusted retail data, where trends and seasonal fluctuations, such as holiday sales growth driven by e-commerce, evolve over time.

4.2 TIME SERIES ANALYSIS SUMMARY

Based on the best model which is Holt-Winter's Multiplicative, 12 months forecasted points are generated. Figure 4.2 below shows the time series plot generated by the simulator.

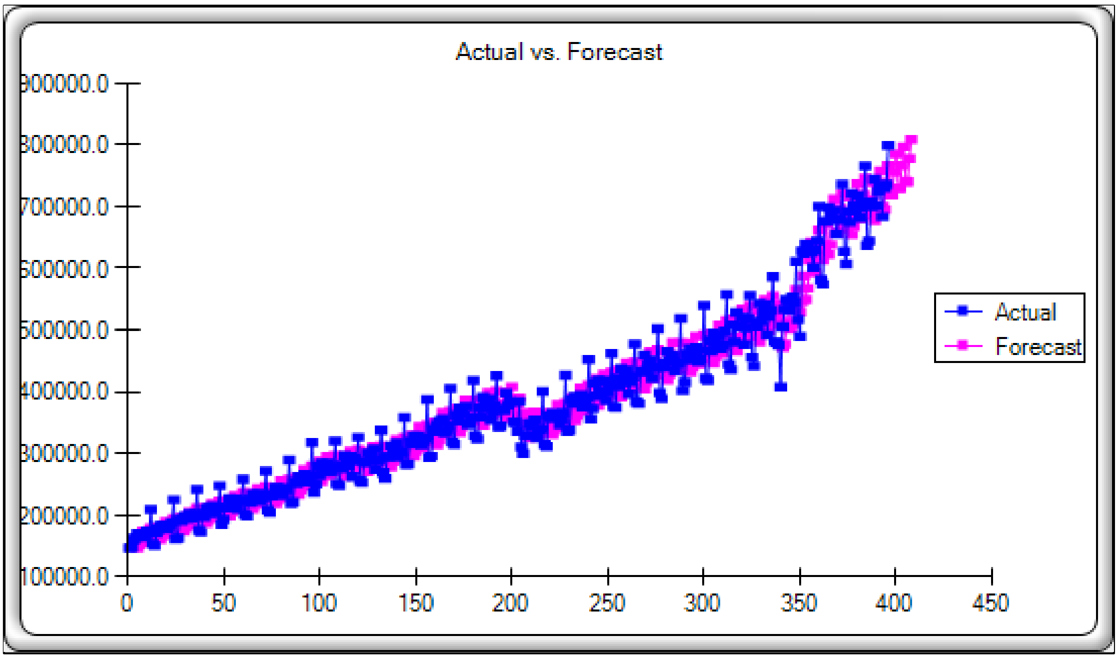

Figure 4.2 Actual vs Forecast graph

Figure 4.3 compares the actual and forecasted values, where can observe that in some cases, the forecasted values are close to the actual values, while in others, there are notable differences. Then, figure 4.4 shows the forecasted values for 2025. The values is keep increasing by the months in year 2025.

Period	Actual	Forecast Fit
1	146376.00	
2	147079.00	
3	159336.00	
4	163669.00	
5	170068.00	146376.00
6	168663.00	152002.51
7	169890.00	168502.37
8	170364.00	173553.47
9	164617.00	155372.76
10	173655.00	158002.46

Figure 4.3 Actual vs Forecast values

395	736461.00	738640.03
396	799769.00	766803.94
Forecast397		719778.69
Forecast398		719366.03
Forecast399		755719.53
Forecast400		785930.30
Forecast401		730558.22
Forecast402		730099.19
Forecast403		766953.20
Forecast404		797569.80
Forecast405		741337.75
Forecast406		740832.36
Forecast407		778186.87
Forecast408		809209.29

Figure 4.4 Forecast values for year 2025.

This analysis was conducted using Holt-Winters' Multiplicative model with alpha (α) at 0.203, beta (β) at 0.019, and gamma (γ) at 0.0303, along with a seasonality period of 4. When both seasonality and trend exist in a dataset, more advanced models are required to decompose the data into three components: the level (L), adjusted by the alpha parameter; the trend (b), adjusted by the beta parameter; and the seasonality (S), adjusted by the gamma parameter. Among the available methods, the two most common are Holt-Winters' Additive and Holt-Winters' Multiplicative models. The additive model assumes seasonal patterns remain constant over time, while the multiplicative model allows seasonal effects to scale with the trend, making it more suitable for datasets where fluctuations increase as the trend grows.

To determine the best-fitting model, multiple error measurement criteria were evaluated. The Root Mean Squared Error (RMSE), which penalizes larger errors more heavily, was 25,332.53, confirming that the Holt-Winters' Multiplicative model minimizes significant deviations from actual sales values. The Mean Squared Error (MSE), which measures the average squared differences between predicted and actual values, was 641,737,305.1, further reinforcing the model’s accuracy in long-term forecasts. The Mean Absolute Deviation (MAD), which quantifies average error magnitude, was 19,465.04, indicating a reasonable level of deviation in predictions.

For relative error evaluation, the Mean Absolute Percentage Error (MAPE) was 5.28%, meaning that, on average, the model's forecasts deviated by 5.28% from actual values. The Median Absolute Percentage Error (MdAPE) was 4.71%, suggesting a lower sensitivity to extreme outliers. Additionally, the Symmetrical Mean Absolute Percentage Error (sMAPE) was 5.3%, further validating the model’s robustness. The Theil’s U Statistic of 0.6979 confirms that the forecast is statistically better than a naive prediction, while the Theil’s U1 Accuracy (0.0619) and Theil’s U2 Quality (0.0310) indicate that the model outperforms simple extrapolation techniques.

Other key error measurements, such as the Root Mean Square Log Error (RMSLE) at 0.0655 and Root Mean Square Percentage Error Loss (RMSPE) at 0.0653, support the model's effectiveness in handling variations in the dataset. The Median Absolute Error (MdAE) of 15,951.95 further demonstrates that the model provides a strong balance between accuracy and consistency.

Overall, based on these performance metrics, the Holt-Winters' Multiplicative model is the best choice for forecasting the "Advance Retail Sales: Retail Trade" dataset. It effectively captures both trends and evolving seasonal patterns, ensuring more accurate forecasts while maintaining stability across different time periods.

REFERENCES

Cowpertwait, P. S. P. (2013). *Introductory time series with R*. Springer.

Cryer, J. D. (2008). *Time series analysis: With applications in R*. Springer.

OSL Risk Management., (2025). Clever Thinking. *Oslriskmanagement.com.* https://oslriskmanagement.com/

U.S. Census Bureau, Advance Retail Sales: Retail Trade [RSXFSN], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/RSXFSN