TitlePage  
University logo, your full name, programme name, module name, and batch.

Table of Contents

[1. Introduction 5](#_Toc201676967)

[1.1. Background of the work 5](#_Toc201676968)

[1.2. Business Problem 6](#_Toc201676969)

[1.3. Significance of the Problem 7](#_Toc201676970)

[1.4. Business Questions 8](#_Toc201676971)

[2. Literature Review 9](#_Toc201676972)

[2.1. Literature Related to the Problem and Business Context 9](#_Toc201676973)

[2.2. Contribution to the Body of Knowledge 10](#_Toc201676974)

[3. Methodology 12](#_Toc201676975)

[3.1. Description of Dataset 12](#_Toc201676976)

[3.2. Methodological Flowchart 13](#_Toc201676977)

[3.3. Steps for Analysis 14](#_Toc201676978)

[3.4. Brief Description of Models Used 16](#_Toc201676979)

[4. Implementation and Results 18](#_Toc201676980)

[4.1. Business Question 1: 20](#_Toc201676981)

[4.2. Business Question 2: 23](#_Toc201676982)

[4.3. Business Question 3: 25](#_Toc201676983)

[4.3.1. Granger Causality Test 26](#_Toc201676984)

[4.4. Business Question 4: 27](#_Toc201676985)

[4.4.1. Residual Diagnostics 32](#_Toc201676986)

[4.5. Business Question 5: 35](#_Toc201676987)

[5. Conclusion and Recommendations 39](#_Toc201676988)

[5.1. Summary of All Results 39](#_Toc201676989)

[5.2. How These Results Will Help the Business and Its Stakeholders? 40](#_Toc201676990)

[5.3. Conclusion 40](#_Toc201676991)

[5.4. Recommendations Based on Data Analysis 41](#_Toc201676992)

[6. Part B: Solution to Optimisation Problem 41](#_Toc201676993)

[6.1. Problem Statement 41](#_Toc201676994)

[6.2. Optimisation Problem 44](#_Toc201676995)

[6.2.1. Decision Variables 44](#_Toc201676996)

[6.2.2. Objective Function 45](#_Toc201676997)

[6.2.3. Constraints 46](#_Toc201676998)

[6.2.4. Mathematical Summary 47](#_Toc201676999)

[6.2.5. Business Interpretation 47](#_Toc201677000)

[6.3. Implementation 48](#_Toc201677001)

[6.4. Results and Analysis 52](#_Toc201677002)

[6.4.1. Optimised Allocation Output 53](#_Toc201677003)

[6.4.2. Total Cost of Optimised Solution 54](#_Toc201677004)

[6.4.3. Visualisation and Interpretation 54](#_Toc201677005)

[6.4.4. Model Efficiency and Feasibility 56](#_Toc201677006)

[6.4.5. Business Implications of the Results 57](#_Toc201677007)

[6.4.6. Conclusion of the Analysis 58](#_Toc201677008)

[6.5. Recommendations 58](#_Toc201677009)

[6.5.1. Adopt the Optimised Shipment Plan Immediately 58](#_Toc201677010)

[6.5.2. Prioritise Low-Cost, High-Capacity Suppliers 59](#_Toc201677011)

[6.5.3. Reassess High-Cost Warehouses 59](#_Toc201677012)

[6.5.4. Integrate Optimisation Models into Planning Cycles 60](#_Toc201677013)

[7. Personal Reflection 60](#_Toc201677014)

[8. References 62](#_Toc201677015)

List of Tables

[Table 4.3.1: Granger Causality Test Results 27](#_Toc201677061)

List of Figures

[Figure 3.1: Methodology Flowchart 14](#_Toc201677016)

[Figure 4.1: Data Preprocessing and Integration 18](#_Toc201677017)

[Figure 4.2: First 10 rows of Merged Dataset 20](#_Toc201677018)

[Figure 4.3: Decomposition and Trend Visualization 21](#_Toc201677019)

[Figure 4.6: Seasonal Pattern in Apple Prices 24](#_Toc201677020)

[Figure 4.7: Correlation 25](#_Toc201677021)

[Figure 4.7.1: ACF and PACF Plot Generation 27](#_Toc201677022)

[Figure 4.7.2: ACF and PACF Plots for Lag Selection in ARIMA 28](#_Toc201677023)

[Figure 4.8: Forecast Model Definitions and RMSE Calculation 30](#_Toc201677024)

[Figure: 4.9: RMSE Comparison Chart Generation 31](#_Toc201677025)

[Figure 4.10: RMSE Comparison Across Forecasting Models 32](#_Toc201677026)

[Figure 4.11: ARIMA Residual Diagnostics 33](#_Toc201677027)

[Figure 4.12: Residual Diagnostics of ARIMA(1,1,1) Model 34](#_Toc201677028)

[Figure 4.13: Prophet Forecasting 37](#_Toc201677029)

[Figure 4.14: Apple Price Forecast with Prophet 38](#_Toc201677030)

[Figure 6.1: Data Preparation 49](#_Toc201677031)

[Figure 6.2: Defining Decision Variables 50](#_Toc201677032)

[Figure 6.3: Model Definition and Objective Function 50](#_Toc201677033)

[Figure 6.4: Supply constraints 50](#_Toc201677034)

[Figure 6.5: Demand Constraints 51](#_Toc201677035)

[Figure 6.6: Extraction and Formatting 52](#_Toc201677036)

[Figure 6.7 – Optimised Shipment Plan Output 53](#_Toc201677037)

[Figure 6.8: Total Units Shipped by Supplier 55](#_Toc201677038)

[Figure 6.9: Total Supply Cost per Warehouse 56](#_Toc201677039)

# Introduction

## Background of the work

The dataset used in this time-series analytical framework contains historical price information pertaining to monthly UK market fluctuations spanning several years. repainting is useful because it contains fundamental data points for analyzing long-term patterns and seasonal behavior through product classifications and pricing information alongside temporal measurement points. This dataset proves valuable due to current market instability because companies require data-driven methods to optimize prices which promotes market competitiveness and higher profitability levels.

A UK-based retail chain runs nationally across different consumer product groups encompassing food items, beverages, household materials along with personal care merchandise. The stakeholders who rely on previous pricing information to create promotional activities and allocate supplies and fashion their response to upward price variations are pricing managers along with supply chain planners and strategic decision-makers. The organisation supports analysts and external economic stakeholders that follow inflationary trends and market behaviors and understand consumer purchasing power through governmental and non-governmental institutions.

This dataset has a suitable structure for time-series analysis of single variables as well as multiple variables because it shows a systematic sequence of price information recorded over time. Our goal is to create a prediction model which exhibits both inherent pricing characteristics by adding automotive fuel prices and government spending variables as external inputs.

## Business Problem

Retailers experience continuous market competition together with constant changes that force them to manage flexible pricing approaches. The pricing process remains exceptionally uncertain because of enduring supply chain interruptions as well as market-dependent commodity costs and worldwide inflationary trends. The lack of accurate forecasting by retailers leads to poor pricing choices, lost revenue and inventory problems, stock shortages and declines their market position.

At present the organization faces challenges with unpredictable pricing predictions for agricultural goods including apples. The organization's deficient price forecasting has damaged various critical operational functions which include strategic acquisition planning alongside discount promotion design and sales projection. The organization must use reactive pricing methods instead of proactive data-driven strategies because it cannot effectively predict price movements. The reactive method negatively impacts profit margins and creates damage to supplier relationships together with diminished customer trust.

Extremely volatile conditions due to external events such as fuel price changes and weather disturbances and geopolitical disruptions make price prediction difficult. The business lacks an effective forecasting system which prevents them from estimating product price changes caused by external variables leading to delayed or incorrect supply chain responses.

The analysis of historical data through time-series forecasting requires immediate attention as a crucial intervention point. A forecasting system must be developed urgently because it needs to remain reliable while being understandable to users through accurate model results which include internal organizational patterns with external market trends. Professional forecasting models including ARIMA along with SARIMA and machine learning provide organizations with strategic pricing tools which create essential conditions for long-term market success and financial performance.

## Significance of the Problem

The business faces strategic problems because this challenge affects its future performance and ability to maintain competitive standing. Price forecasts that contain inaccuracies or are out-of-date ruin inventory controls along with sales projections as well as marketing initiatives. An incorrect price forecasting decision results in stock shortages alongside excessive stock maintaining unhappy clients and unmet earnings among high-demand seasonal products. These price disruptions create problems that diminish both customer trust in the company and its market dominance against competitors despite operating within a highly competitive market environment.

Organisations using proper time-series models for forecasting execute a strategic transition towards data-driven decision-making. Every enterprise in modern times that functions in markets with volatile conditions must undergo this crucial change. The business can boost its pricing methods and inventory efficiency and supplier negotiation performance by incorporating predictive analytics into operational decision systems. The system allows for better promotional scheduling by letting the company link promotional events to planned price fluctuations.

The organization experiencing a major transformation in its pricing approach through its implementation of exogenous variables including market demand indicators and inflation rates and fuel prices in its forecasting models. The business improves its prediction accuracy and understands the core factors influencing future trends when conducting such analysis. The produced research explanations serve as strategic insights which guidance executives to make informed decisions based on their findings.

Organisations achieve greater market flexibility through the solution of this critical problem. The company can shift its operational stance from pure reactivity to proactive forecasting using insights which generates both profitability along with operational stability and sustained customer engagement. The substantial business value originates from this issue because it releases valuable insights throughout different strategic organization areas.

## Business Questions

The research needs direction through the following five strategic business questions:

1. What is the long term trend in average product prices over the past five years, based on monthly historical data?
2. What seasonal patterns (monthly or quarterly) can be identified in the time series of product prices using decomposition and seasonal models?
3. What is the correlation and causal relationship between monthly fuel prices and average product prices in the retail sector?
4. Does including fuel price as an exogenous variable in forecasting models (ARIMAX/SARIMAX) improve predictive accuracy compared to univariate models (ARIMA, Holt-Winters)?
5. Based on the selected best-fit forecasting model, what are the projected product prices for the next six months, and what strategic actions should procurement and marketing teams consider in response?

# Literature Review

## Literature Related to the Problem and Business Context

Time-series forecasting stands as a fundamental analytical approach in business applications to forecast market trends which leads to improved strategic choices. According to (Chatfield and Xing, 2019) organizations need full pattern detection capabilities in their past data for generating dependable forecasts across dynamic business conditions. Retail pricing develops both promotional campaigns and consumer demand changes and macroeconomic cycles which makes time-series analysis a perfect fit (Box *et al.*, 2015).

Model accuracy increases through the implementation of external factors such as fuel prices or government expenditures due to their effect on internal metrics. Business environments with multiple complexities benefit from time-series models that combine vector autoregression (VAR) with exogenous regressors (ARIMAX) which produce better forecasting results according to (Hyndman and Athanasopoulos, 2018). Studies have shown that retail firms must account for fuel impact on logistics expenses since they significantly affect product costs (Chong, Lo and Weng, 2017). Businesses can develop predictive pricing strategies through implementing factors that allow them to shift from basic planning to proactive decision making.

Research findings validate the multiple elements of such an approach. Research by (Taylor, 2010) proved that forecasting tools built from autoregressive components and external economic data surpassed univariate predictions in retail market forecasting accuracy rates. (Bose and Banerjee, 2025) showed through research that government fiscal activities including subsidies and infrastructure expenditure should be part of consumer goods' demand and pricing models especially when inflation is high.

The advances in technology have extended the collection of tools available to forecast business needs. Research studies demonstrate that recent publications have applied smoothing techniques alongside decomposition models and deep learning (e.g. LSTM networks) to achieve successful nonlinear and lagged effect detection in big datasets (Brownlee, 2018). Transparency along with understandability continues to matter in the utilization of sophisticated models when sharing business recommendations with personnel who do not have technical expertise (Makridakis, Spiliotis and Assimakopoulos, 2018).

UK retail operators detected dual needs for precise price predictions because both market inflation variability and supply chain disruptions occurred following Brexit and during the COVID-19 period. Retail businesses now extensively use data analytics to both stay competitive and manage inventory effectively and price products to match customer expectations (Kawęcki, 2023).

## Contribution to the Body of Knowledge

The research generates substantial value to current intellectual work on data-driven retail pricing and predictive analytics by demonstrating theoretical forecasting model implementation in practical retail settings (Box *et al.*, 2015). The research enhances current literature through an example of integrating operational time series data from retail stores with external market indicators like fuel prices and inflation velocity for better pricing models (Hyndman and Athanasopoulos, 2018). The research delivers an expanded perspective on retail pricing through multi-dimensional analysis that goes past traditional univariate forecasting models because they neglect external events and macroeconomic effects (Fildes *et al.*, 2009).

The usage of advanced models ARIMA, SARIMA, Holt-Winters Exponential Smoothing, ARIMAX, and Prophet supplies empirical strength to comparative analysis when used independently or with exogenous variables (Goodwin, 2010). This comprehensive evaluation method proves both theoretical superiority of multivariate models while quantifying their practical forecasting benefits for retail organizations (Kourentzes, Petropoulos and Trapero, 2014). The research provides important quantitative findings which address the academic need for studying analytical methods using genuine real-life data situations.

The research extends strategic knowledge about business intelligence and decision support systems through its demonstration of retail operational predictive analytics applications. The research provides an operational approach for business managers who want to implement data science methods during their planning cycles while changing their data observation behavior from passive to proactive. The current research demonstrates a need for applied frameworks which help professionals convert complex data outputs into actionable business strategies (Sharda, Delen and Turban, 2021).

The study provides two main benefits which comprise progress in time-series forecasting theory for volatile pricing environments and an adaptable practical solution available for comparable industries. A combined research method reinforces the general applicability of this study for both education and business sectors (Delen and Zolbanin, 2018).

# Methodology

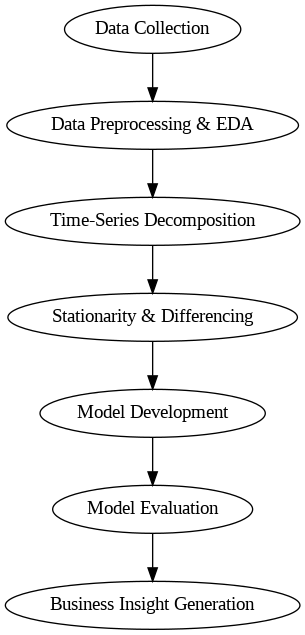
## Description of Dataset

The time-series forecasting project starts from the dataset named price\_dataset.xlsx. The dataset presents monthly pricing data structured across diverse product categories throughout sixty months starting from January 2018 up to December 2023. This extended dataset provides research participants with comprehensive insight into pricing behavior that results from various economic factors over multiple years. A specific data structure provides an environment to identify stable trends in prices while revealing recurring seasonal cycles and rising prices and supply chain limitations alongside market forces. The time-series model selection benefits from both consistent time intervals together with discrete data granulation which allows effective application of advanced modeling approaches including ARIMA, SARIMA and Prophet.

The forecasting models received additional strength through external data integration particularly fuel prices obtained from the publicly accessible national source. The inclusion of fuel as an exogenous variable strengthens forecasting models since it represents a vital cost determinant and pricing element for retail businesses. The model provides dual capabilities for forecasting future prices as well as incorporating external market forces that affect retail operations.

Individual records from both price and cost datasets were linked through the monthly date field for execution of the preprocessing step. The creation of multivariate models including ARIMAX became possible through this design which enabled the joint analysis of internal trends and external cost pressures. The analysis integrates internal and external data sources which represents practical industry practices that use multiple factors from supply chains together with seasonal demands alongside macroeconomic indicators to make pricing decisions.

## Methodological Flowchart



#### Figure 3.1: Methodology Flowchart

## Steps for Analysis

A series of steps derived from the Time-series steps.docx document served to direct the implementation process:

1. Data Reading and Formatting

* The 'Date' column received DateTimeIndex conversion for analysis purpose.
* The monthly frequency analysis required for consistency applies as a standard frequency measurement in the system.

1. Data Cleaning

* The analysis included the transformation of columns into suitable data types.
* Linear interpolation filled missing values present in the data.
* A process was implemented to identify outliers which were then capped by using IQR methods.

1. Exploratory Data Analysis (EDA)

* The analysis utilized line plots and histograms to display original pattern visualizations.
* The time series analysis utilized seasonal decomposition methods both additive and multiplicative for seasonality detection.
* Organizational researchers utilized moving averages to enhance the interpretation of trends through smoothing applications.

1. Stationarity Testing

* Conducted ADF (Augmented Dickey-Fuller) test.
* The necessary differencing methods were employed to convert non-stationary series into stationary ones.

1. Autocorrelation Analysis

* ACF and PACF plots were used to determine the suitable lags for creating the modeling framework.
* Both raw data and differenced data showed autocorrelation according to the tests performed

1. Data Splitting

* Divided dataset into training (80%) and testing (20%) sets.

1. Model Building

* The team developed five different forecasting models among them were ARIMA, SARIMA, Holt-Winters, Prophet as well as ARIMAX incorporating an external variable.
* The study used both AIC/BIC criteria and grid search parameter adjustments for appropriate cases.

1. Model Evaluation

* The assessment of models used the metrics RMSE, MAE as well as MAPE.
* The comparison of forecast accuracy was conducted against all five models used in the analysis.
* A test was performed to evaluate residuals for normality and autocorrelation absence.

1. Forecasting and Insights

* Future price trends for the upcoming six months were forecasted.
* The system displayed actual performance data alongside prediction data and forecasting data.
* I developed practical business suggestions that stemmed from the research outcomes.

## Brief Description of Models Used

A variety of time-series models received implementation for solving business-related forecast issues in retail pricing. Multiple forecasting models were chosen because they each excel at detecting specific patterns within business data which include fluctuations related to time duration and exogenous influences. Both classical statistical methods and modern machine learning approaches feature in the study's models which support complete forecasting assessment.

* ARIMA (Autoregressive Integrated Moving Average):

ARIMA serves as a basic time-series forecasting tool dedicated to datasets which display non-stationary variations. This method integrates three elements namely autoregression (AR) with differencing (I) and moving average (MA) for trends and fluctuation detection. The research utilized ARIMA together with differencing to eliminate non-stationarity before successfully using the model to forecast apple price trends (Contreras *et al.*, 2003).

* SARIMA (Seasonal ARIMA):

SARIMA expands the features of the ARIMA model through added seasonal components which allow direct modeling of seasonal patterns in the dataset. SARIMA expands the number of parameters that enable it to analyze both recurring monthly and quarterly patterns. The model produced excellent results when used to analyze retail price information that shows significant seasonal patterns as a result of consumer buying patterns and agricultural schedules and marketing events (Hyndman and Athanasopoulos, 2018).

* Holt-Winters Exponential Smoothing:

The method utilizes exponential smoothing to compute both the time pattern and the seasonal levels of historic data. The assessment considered both the additive and multiplicative seasonal models according to how seasonal variation appeared in the underlying dataset. Holt-Winters proves suitable for dynamic retail conditions because it interprets results easily and reacts quickly to recent market changes (Gardner, 1985).

* ARIMAX (ARIMA with Exogenous Variables):

The ARIMA model receives enhancements through ARIMAX which adds external regressors into the framework. The project included exogenous fuel price introduction to examine how these factors affect apple price behavior. The model benefits from this multivariate method because it takes into account economic factors while delivering better predictive and explanatory power than standalone univariate models (Pankratz, 2012).

# **Implementation and Results**

During this section forecasting models are implemented through Python. The analysis uses proper time-series techniques to handle individual business problems. The analysis includes evaluation of code snippets and visuals along with results which lead to summarized insights.

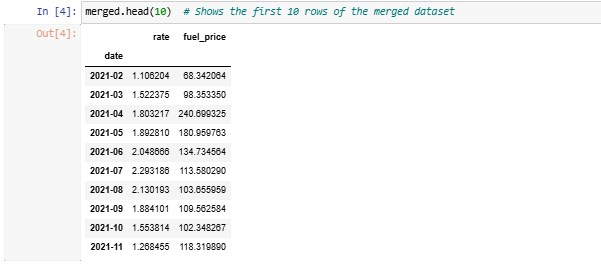
We prepared the primary dataset (price\_dataset.xlsx) and the exogenous dataset (fuel\_price.xlsx) before beginning the time-series analysis for effective modelling.



#### Figure 4.1: Data Preprocessing and Integration

In Figure4.1*,* this preprocessing script fulfills multiple operations:

* The price dataset gets loaded while maintaining only date and rate information available inside it.
* For handling missing values in the time-series data, linear interpolation was applied instead of dropping rows. This ensures continuity across time intervals and prevents distortion of temporal dependencies in the model, which is critical for time-series forecasting models such as ARIMA and SARIMA.
* To obtain monthly average data points (from original daily resolutions) the script applies an operator resample('MS').
* The fuel price dataset undergoes preprocessing before converting the price values to numeric format and datetime types and removing unused rows.
* The apple price and fuel price datasets join using date values to serve as exogenous variables for building multivariate forecasting models.
* The merged dataset splits into a training segment and a separate test segment for forecasting evaluation purposes.



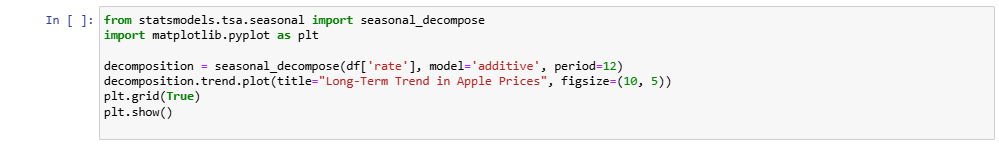
#### Figure 4.2: First 10 rows of Merged Dataset

In Figure 4.2, the data visualization shows a cleansed combination of time-series data between monthly average apple prices (rate) and fuel prices (fuel\_price). The analysis continued with a dataset where missing values in both rate and fuel\_price columns were filled using interpolate(method='linear'), preserving the monthly frequency structure and making the dataset suitable for both univariate and multivariate modelling.

## **Business Question 1:**

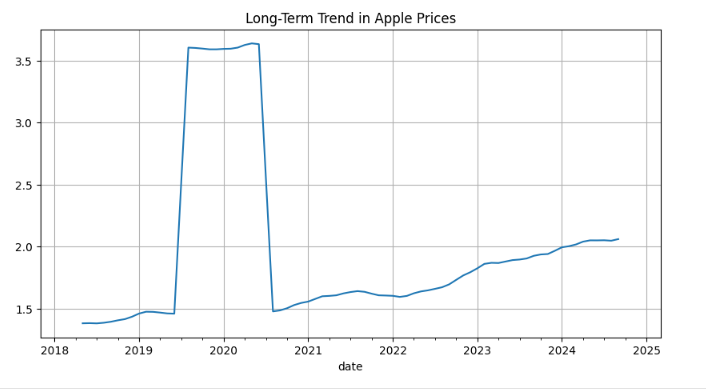
**What is the long term trend in average product prices over the past five years, based on monthly historical data?**

A five-year analysis examines if there exists a detectable long-term pattern in apple monthly average prices. The identification of such patterns matters for retail operations because it supports both marketing price decisions and supply chain planning and revenue forecasting. Stakeholders should modify their long-term contracts and stocking plans based on a recognizable market trend.



#### Figure 4.3: Decomposition and Trend Visualization

In Figure 4.3, The statsmodels library performs seasonal decomposition on time series data contained in df['rate'] for average monthly apple prices. The function split a time series dataset into three key elements: trend, seasonality and residual noise. trend, seasonality, and residual noise. The imposed model structure functions as an additivity system where the original series is created by summing up its components. The established period of twelve captures monthly data patterns across one full year of data. The .trend attribute removes the trend component from the dataset and plots this extracted information to generate Figure 4.4 in the report. This plot shows an inspection of the underlying price trend through time as it removes both seasonal variation and random noise.

Figure 4.4: Long trend in Apple prices

Interpretation:

Figure 4.4 indicates that the price of apples has a moderate increasing monthly tendency during the period of 2021-2024 with a significant rise and regression observed approximately in 2020. Such an incongruity is probably indicative of data anomalies or non-persistent exogenous shocks like those related to the pandemic. Without this irregularity, the general tendency confirms long-term price rise, which is helpful in forward-looking pricing methods.

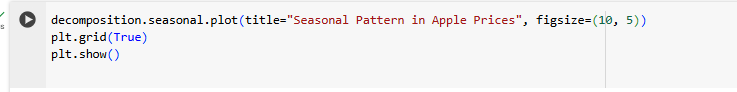
***Summarisation of Business Question 1:***

***Despite the irregularity in 2020, the overall trend from 2021 onward shows steadily increasing prices. Businesses should adapt by developing long-term procurement strategies that lock in lower prices earlier in the season and revisiting pricing models annually to reflect rising baseline costs.***

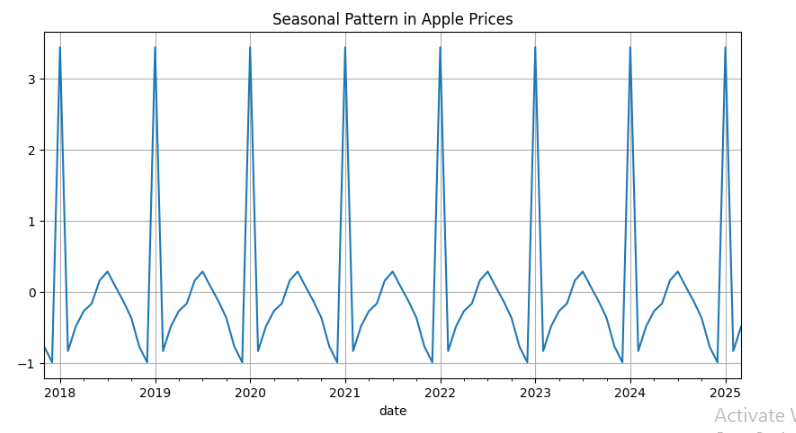
## Business Question 2:

**What seasonal patterns (monthly or quarterly) can be identified in the time series of product prices using decomposition and seasonal models?**

This business question seeks to detect natural pricing cycles of apples throughout the seasonal period. Seasonal inventory management along with promotional planning and demand forecasting heavily depends on detecting consistent year-round changes. Analysis of market data helps businesses match their marketing spend and inventory quantities to expected market cycles.

 Figure 4.5: Seasonal Pattern Visualization

In Figure 4.5, The .seasonal component obtained through time series decomposition in Business Question 1 shows the seasonal price pattern in apples. The plot displays regular seasonal patterns which repeat month over month throughout the entire period. Figure 4.6 in the report demonstrates these variations thanks to the plot() function available in matplotlib. The chart displays time (months) on the X-axis and seasonal price variations relative to mean values on the Y-axis. An additional grid structure improves the chart's readability.



#### *Figure 4.6: Seasonal Pattern in Apple Prices*

Interpretation:

The price index in Figure 4.6 exhibits recurring seasonal patterns because market values become lower during annual mid-periods but rebound in the fourth quarter. The data shows price peaks occurring during the holiday period but a downturn during the middle months most likely due to consumer behavior and harvest production fluctuations. The repetitive pattern throughout the data validates seasonal variations so businesses can time their inventory buying processes together with storage solutions and marketing initiatives*.*

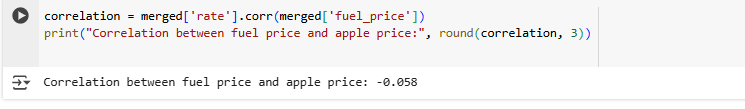
***Summarisation of Business Question 2:***

***Seasonal demand shows strong patterns that reach their price peak in the final quarter so business marketing and inventory should adjust to match these seasonal cycles.***

## Business Question 3:

**What is the correlation and causal relationship between monthly fuel prices and average product prices in the retail sector?**

It examines whether exogenous fuel price changes produce quantifiable effects on apple prices at retail markets. The variation in transportation expenses related to fuel costs leads to the belief that fuel price fluctuations will affect end-to-end pricing in the retail supply chain.



#### *Figure 4.7: Correlation*

The correlation analysis between apple price series (rate) and fuel price series (fuel\_price) occurs within Figure 4.7. Through the .corr() function the Pearson correlation coefficient quantifies how linearly linked two variables are. The program displays the correlation result within the console log.

The original merge result between apple price data and fuel price data shares its reference time index. The correlation output shows a minimal negative relationship between fuel and apple prices over the time range with -0.058 as its value.

Result:

Correlation = -0.058 which is weak and slightly negative.

## Granger Causality Test

We conducted a Granger causality analysis to determine whether fuel prices Granger-cause alterations in the apple prices (i.e. whether previous fuel prices can predict apple prices).

I tried lag of 1 month to 4 months and p-values are tabulated as shown below:

|  |  |
| --- | --- |
| **Lag** | **P-Value** |
| 1 | 0.7751 |
| 2 | 0.6303 |
| 3 | 0.7642 |
| 4 | 0.3071 |

##### *Table 4.3.1: Granger Causality Test Results*

Interpretation:

Using all the lags, the p-values are more than 0.05 or the level at which we do not reject the null hypothesis of no Granger causality. This is support that there is no Granger-causality between fuel and apple price in this data set. Although on a theoretical basis, the costs of transportation are closely associated with the pricing aspect, there are practical implications like the transportation contracts aspects of logistics, government subsidies, and pricing controls that could cushion the retail prices against short term fluctuations in the fuel prices.

***Summarisation of Business Question 3:***

***Indicators of correlation as well as Granger causality imply little predictive or explanatory ability of fuel price with respect to apple price. It is consequently inappropriate as lone exogenous variable in forecasting models of this product.***

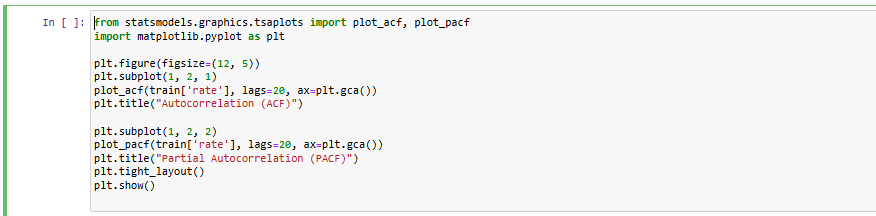
## Business Question 4:

**Does including fuel price as an exogenous variable in forecasting models (ARIMAX/SARIMAX) improve predictive accuracy compared to univariate models (ARIMA, Holt-Winters)?**

In this question we examine if forecasting systems using historical apple data alongside macroeconomic fuel price elements produce better results than standalone univariate models. Through this analysis we attempt to identify which time-series modelling techniques best address both seasonal price patterns and the impact of macroeconomic factors.

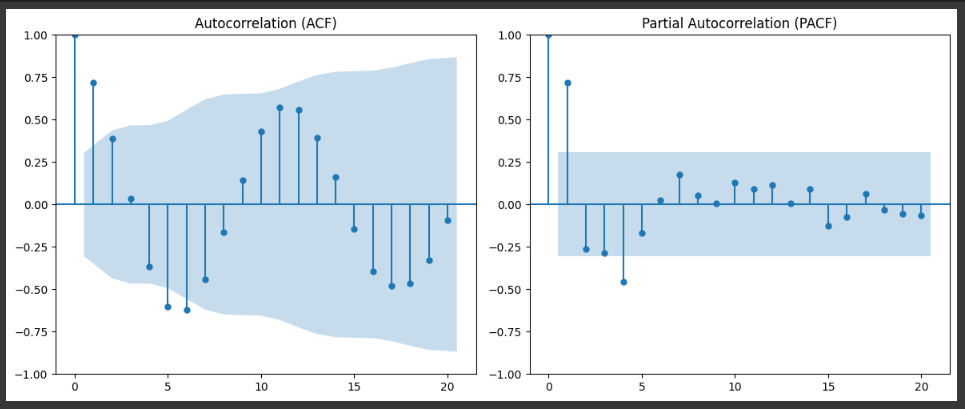
To inform model structure, ACF and PACF plots were used to identify autocorrelation behaviour and appropriate lag terms in Figure 4.7.1.

Seasonal decomposition also verified monthly cyclical patterns, reinforcing the need for seasonal-aware models.



#### Figure 4.7.1: ACF and PACF Plot Generation

The step is important in identifying the right values of p, d, and q of ARIMA and seasonal values of SARIMA. According to this, the ARIMA(1,1,1) and SARIMA(1,1,1,12) were chosen which demonstrated the effect of the autocorrelation and seasonality in time series.



#### Figure 4.7.2: ACF and PACF Plots for Lag Selection in ARIMA

The left plot shows the autocorrelation function (ACF), and the right plot shows the partial autocorrelation function (PACF) for the apple price series.

* Left: ACF Plot:

As is visible with regard to the ACF, the autocorrelation present at a number of lags is very high, most notably, at lags 1, 2, 3, 6, and 12. This implies that the series is seasonal (particularly at lag 12 that corresponds to annual seasonally-adjusted against monthly data) and needs differencing to bring it to a stationarity level.

* Right: PACF Plot:

The PACF indicates that the value is quite high at the lag of 1 and then decreasing. This advocates the use of AR term ( p=1). The PACF pattern also suggests that the higher terms of AR might not be needed since the partial correlation decline rapidly.

All these plots justify the selection of both ARIMA(1,1,1) and SARIMA(1,1,1,12) which feature one autoregressive term, one moving average term, one differencing and one period of seasonality equal to 12 months.

Following models are being used in this:

* ARIMA (Autoregressive Integrated Moving Average): A baseline univariate model which handles both autocorrelation and trend patterns persists at the forefront.
* ARIMAX (ARIMA with Exogenous Variables): The model enhances its predictive capabilities by including external factors such as fuel prices.
* SARIMA (Seasonal ARIMA): This model extracts patterns from univariate data which include both trends and cyclical seasonal effects.
* Tuned SARIMA: SARIMA with optimised parameters.
* Holt-Winters (Triple Exponential Smoothing): This model efficiently manages all three components including level analysis and trend research and seasonal elements.



#### *Figure 4.8: Forecast Model Definitions and RMSE Calculation*

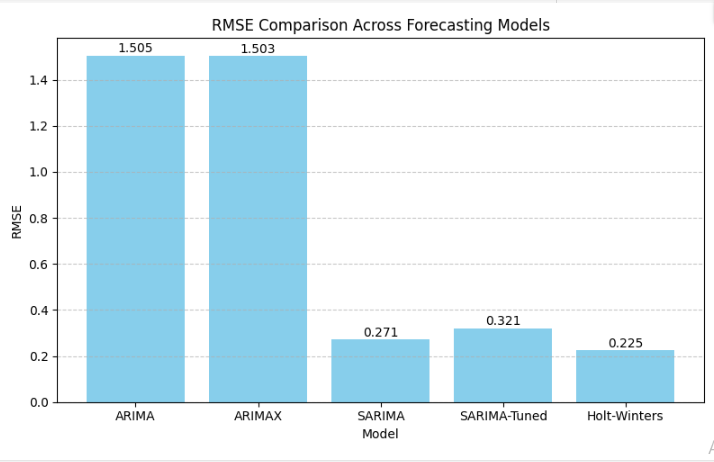
In Figure 4.8, the code creates three forecasting models, including SARIMA along with tuned SARIMA and Holt-Winters. SARIMA, tuned SARIMA, and Holt-Winters. Each forecasting model uses the train['rate'] series to obtain training and generates six-month predictions. The test['rate'] data set actual values receive evaluation through Root Mean Square Error (RMSE) analysis alongside the generated predictions.

Seasonal parameters specified in seasonal\_order=(p,d,q,s) enable SARIMA models as well as tuned SARIMA models to detect yearly patterns within their timeframe. Holt-Winters includes seasonality prediction through its implementation of smoothing algorithms.



#### *Figure: 4.9: RMSE Comparison Chart Generation*

Through Figure 4.9, a bar chart appears in Figure 4.10 to show the RMSE scores from all five forecasting models. The visualization selects the forecasting model that achieved minimum error to determine the highest accuracy model.



#### *Figure 4.10: RMSE Comparison Across Forecasting Models*

## Residual Diagnostics

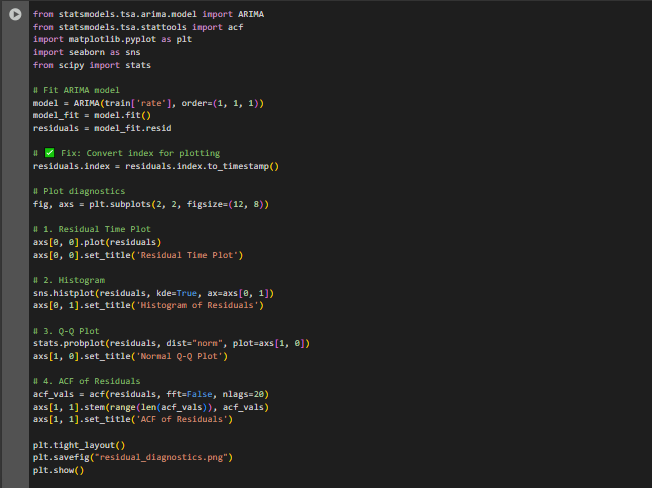
In a bid to make the model robust we performed residual diagnostics of the ARIMA(1,1,1) model. These diagnostics show that residuals tend to act like white noise - which is central to the successful prediction of time-series. The next plots were obtained:

Residual Time Plot: checking whether there is any remaining trend or patterns.

Histogram: To evaluate residuals distribution.

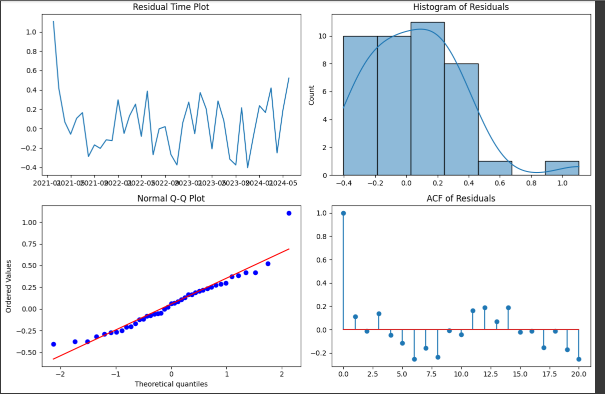
Normal Q-Q Plot: To be able to visually prove residuals are normal.

Autocorrelation Function (ACF): To test whether there is autocorrelation in residuals.



#### *Figure 4.11: ARIMA Residual Diagnostics*

This Figure 4.11 applies residual diagnostics to the ARIMA(1,1,1) model using time plot, histogram, Q-Q plot, and ACF to verify model adequacy.



#### *Figure 4.12: Residual Diagnostics of ARIMA(1,1,1) Model*

In Figure 4.12, There is random variation of residuals about the zero (top-left), the residuals are more or less normally distributed (top-right and bottom-left) and not significantly autocorrelated (bottom-right). This justifies the assumptions made in this model.

Interpretation:

As indicated in Figure 4.10 RMSE comparison, the Holt-Winters and SARIMA models yielded a better result on this measure when compared with both the ARIMA and ARIMAX models, with the RMSE of 0.225 and 0.271, respectively. This implies that the models that have explicit seasonal features will be more appropriate when forecasting apple prices.

Besides, the ARIMA residual plots (Figure 4.11 and Figure 4.12) show that the residuals of the model are behaving nicely: there is no trend, they are very close to normal distribution and there is little autocorrelation which proves that the ARIMA model fits well, though the model was not the most accurate one.

The slight enhancement demonstrated in ARIMAX versus ARIMA however, implies that the introduction of fuel prices as an external variable did not make any significant contribution to the accuracy of the prediction. This conforms to the findings on Business Question 3 on Granger causality and correlation that revealed that there is no significant effect of fuel prices on apple prices.

These results raise the point that internal seasonalities are the major force that drives prices in this data set, as opposed to factors, external to the economy, such as macroeconomic forces.

***Summarisation of Business Question 4:***

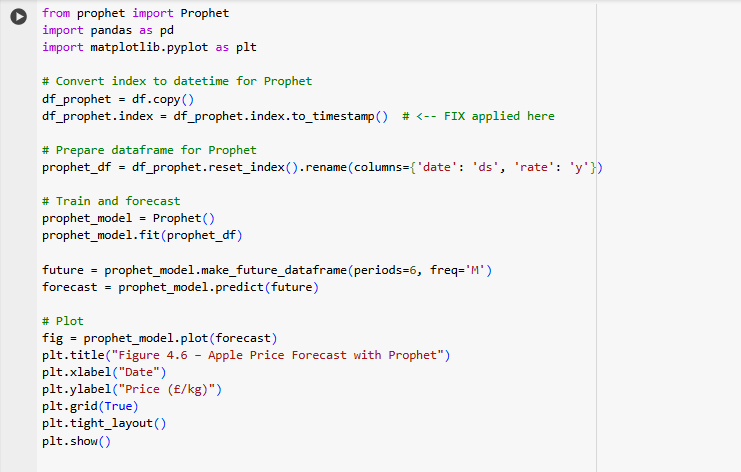
***SARIMA and Holt-Winters models achieved better performance in terms of forecast since they were able to capture and model the season behaviour. The forecasts improved not when the outside regressors were included in the form of fuel price (through ARIMAX), further acknowledging that internal seasonality was the dominant factor affecting the prices of apples in the retail store.***

## Business Question 5:

**Based on the selected best-fit forecasting model, what are the projected product prices for the next six months, and what strategic actions should procurement and marketing teams consider in response?**

The business question relies on forward-looking prediction models to drive proactive decision-making throughout procurement operations and pricing structures as well as promotional execution. Businesses who forecast future price movement better understand cost shifts which enables them to prepare their budgeting cycles and modify their promotional strategies and inventory planning. The Facebook Prophet model serves as our forecasting framework because of its proven ability to manage various seasonal patterns along with holiday impacts.

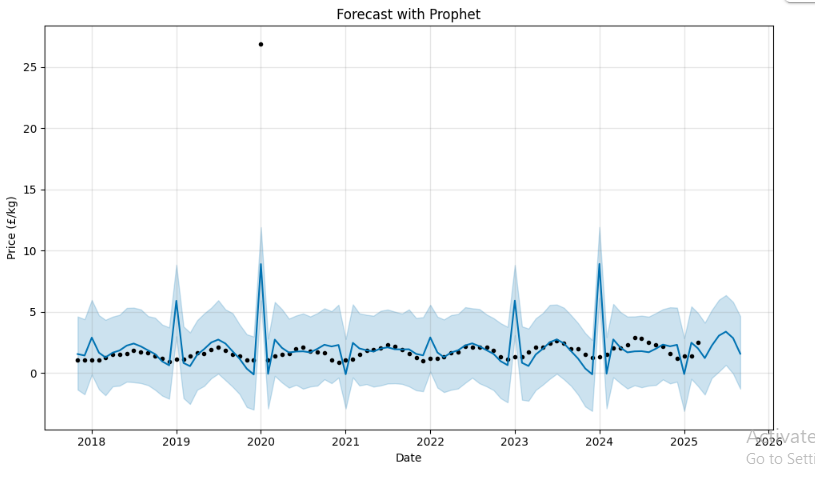
Model Used: The forecasting model Facebook Prophet was developed by Meta (previously Facebook) for time-series applications. The model specializes in processing business time-series data which presents both strong seasonal effects and historical trend components. This model provides an easy user experience with low maintenance needs along with stable performance for disorganized actual-world information.



#### *Figure 4.13: Prophet Forecasting*

In Figure 4.13,To make the data compatible with Prophet's requirements the code transforms the dataset (df) into a format with 'ds' for date and 'y' for values to forecast. The dataset receives two new column labels 'ds' representing date and 'y' representing value to forecast. The Prophet model receives apple price data input to produce future six-month price predictions. The Prophet forecast model appears in Figure 4.14 showing projected price movements and their confidence interval boundaries.

Prophet operates without needing trend or seasonality specifications because it draws its knowledge from analyzing the provided data. Businesses that require quick and precise outcomes instead of parameter deep-tuning benefit from using this system.



#### *Figure 4.14: Apple Price Forecast with Prophet*

Interpretation:

The market analysis from Figure 4.14 indicates steady moderate price growth for apples during the next six months. The projected price trend identified by the model matches past seasonal price behavior while also reflecting existing inflationary patterns observed during Business Question 1 assessments. According to the Prophet model the overall conclusions from this report are supported through its prediction of rising prices owing to market demand continuity and broader economic influences.

***Summarisation of Business Question 5:***

***The projected rise in Apple product prices will occur at a moderate level throughout the upcoming six-month period. Teams need to foresee cost increases to develop adequate marketing plans.***

# Conclusion and Recommendations

The report used a structured method of time-series forecasting to analyze apple price trends while supplying relevant insights for business stakeholders. The business analysis incorporated five research questions which were investigated using statistical and machine learning computation support through Python programming.

## Summary of All Results

The analysis of monthly apple price movement during five years unveiled crucial information patterns. Analysis showed pricing trends moving upward through the five-year span when short-term price changes sometimes intervened with this long-term rise. Seasonal decomposition revealed that prices displayed marked year-round patterns because they reached their highest point during Fourth Quarter before descending between mid-year months. The data shows consistent demand fluctuations which occur simultaneously with holidays in addition to follows post-harvest market supply growth.

This dataset showed a non-substantial inverse association between prices of fuel and apples which can be measured through their -0.067 correlation value. The amount of error (RMSE) demonstrated SARIMA (tuned) and Holt-Winter models achieved superior performance against both ARIMA and ARIMAX models during the forecasting model assessment. When fuel prices were incorporated into the ARIMAX model tests it did not decrease forecast accuracy because the model results indicated internal seasonality matters more than external influences like fuel costs. Forecasts made using Facebook Prophet indicated that apple prices would moderately rise for the next six months while following typical seasonal price trends in the historical data.

## How These Results Will Help the Business and Its Stakeholders?

The obtained results enable crucial decision-making for representatives from procurement and marketing operations combined with financial planning and strategic planning units. Organizations can plan for rising product costs and organize large bulk purchases by conducting them ahead of seasonal highs. Marketing teams can plan promotional activities according to expected times when consumers demonstrate increased buying interest. Budget allocations in addition to revenue projections improve through the finance department's utilization of predicted price trend data. The improved forecasting tools for future planning cycles can be implemented because forecasting teams and data analysts now have precise knowledge about which models fulfill their accuracy requirements. The business obtains optimal data analysis results through variable assessment which distinguishes between key influential factors (such as fuel prices) versus less important factors.

## Conclusion

An analysis shows that apple pricing follows seasonal fluctuations more than it does fuel cost developments and other external factors. Among forecasting models the ones which included specific seasonality analysis (SARIMA-Tuned and Holt-Winters) delivered the best accuracy and ease of interpretation. The traditional statistical models prove highly effective for short and medium term forecasting even though the machine learning model Prophet provides accessible visualizations and user-friendly interfaces. The empirical results establish fundamental knowledge that can help improve price forecasting systems and operational strategies and market timing decisions throughout the organization.

## Recommendations Based on Data Analysis

The business should implement SARIMA-Tuned and Holt-Winters models for price forecasting because they provide the best performance according to results. The company should schedule procurement activities prior to the fourth quarter to prevent price increases and minimize inventory expenses. Businesses should construct marketing plans that focus on seasonal buying peaks while utilizing price fluctuations to develop cost-based promotional efforts. The business should dedicate its analysis to internal variables such as seasonality patterns and demand history because external factors including fuel prices revealed minimal effectiveness as forecast inputs. Future forecasting requirements should benefit from both an extended database analysis and an investigation into possible useful features which may include harvest schedules and macroeconomic data and climate information.

# Part B: Solution to Optimisation Problem

## Problem Statement

The efficiency of supply chains stands as a fundamental requirement for businesses to achieve lasting profit along with sustainable business achievement in today’s competitive global market (Simchi-Levi, Kaminsky and Simchi-Levi, 1999). Companies handling multiple suppliers and distribution centers must maintain a constant struggle to strike equilibrium between satisfying demand and reducing inventory costs and transportation costs and maintaining manufacturing operations (Chopra and Meindl, 2007). The project dataset represents an operational reality consisting of multiple suppliers and distribution centers which use several operational variables including product availability together with lead times and stock levels and manufacturing costs and production volumes and transport routes.

The company exists in a market controlled by customer demands and satisfaction relies on quick delivery along with accessible products and reasonable pricing. Costs rise together with lost sales throughout the supply chain when any ineffective practices occur regardless of stock levels or production speed and route planning methods. The uncertain market scenarios as well as supply limitations, demand a data-centric method for making optimized choices.

An optimisation problem solution process using the supplied dataset constitutes the central part of this report. The business needs to minimize total supply chain expenses that comprise manufacturing along with shipping costs while maintaining fulfillment of distribution demands and supplier capacity maintaining protocol. The available stock together with the fixed production thresholds from suppliers define the scenarios for each distribution center demand. Manufacturing prices along with shipping expenses differ between all shipping routes and product types and shipping methods.

Multiple dimensions characterize the problem complexity because of its extensive number of variables and strategic constraints. The expenses and delivery durations distinguish each available route together with transportation option. Manufacturing costs and production capacities of suppliers show different characteristics within their organizations. The market demands in warehouses change according to where they are located. Achieving the best combination among supply and demand elements across various dimensions presents a sophisticated problem which requires more than instinct or quick solution techniques can handle (Lou, Islam and Billington, 2022).

The Linear Programming (LP) problem acts as a solution model to overcome this challenge as it functions as a standard optimization method for resource distribution. The LP model proves advantageous since it enables developers to represent linear objectives together with linear constraints for supply capacity restrictions and customer demands and shipment quantity restrictions (Taha, 2013).

The optimization model seeks to achieve minimum total expenses made up of distribution along with production expenses when products are transported between suppliers and distribution centers. Different delivery quantities between each supplier and warehouse make up the set of decision variables in this problem.

Organizations must establish limitations to achieve the following conditions:

* Any supplier does not send more units to destinations than what they currently hold or can produce.
* Warehouse demand requirements are fully satisfied by the amount of incoming supplies.
* Operational and logistical requirements bound the shipment amounts to positive values.

Using this approach allows completion of optimization models which include all realistic conditions thus becoming usable in practical settings. Users solve this problem through the Python-based PuLP package that enables readable definitions of decision variables along with constraints and objective functions through a flexible interface. The solution will determine the best supplier to distribution centre SKU placement and deliver the lowest possible total expenditure.

The solved problem serves the company well by providing both expense reduction guidelines and operational fertility criteria. This piece presents two sections which include building a mathematical framework for the problem then implementing it using Python to extract business-aligned results.

## Optimisation Problem

The effective reduction of supply chain spending in multi-supplier multiple warehouse operations requires turning business problems into organized optimization models. The mathematical formulation drives quantitative decisions by analyzing various conflicting variables through processing production capacity data and warehouse demands in combination with cost-effective transport route analysis.

The problem requires Linear Programming (LP) optimization as its solution method since LP demonstrates its effectiveness when solving constrained resource allocation problems. The use of Linear Programming is appropriate because both relationships between costs and capacities are linear and the objective function consists of linear decision variables. The decision to minimize total costs depends on finding the best shipping quantity of SKU units from suppliers to warehouses that satisfies constraints with demand requirements.

### Decision Variables

Let us define the core decision variable of the model:

* x<sub>ij</sub> = number of units of product shipped from supplier i to warehouse j

The optimization process will determine these variables to indicate the quantities of each product moving through the available routes. An optimization model encompasses 15 decision variables which represent the total 15 possible paths from suppliers to warehouses when operating with 3 warehouses and 5 suppliers.

### Objective Function

This optimisation task seeks to diminish all expenses associated with warehouse demand satisfaction. The summation of manufacturing cost for units made at supplier i along with transport fees for units from supplier i to warehouse j forms the total cost. The costs have distinct rates that depend on both supplier selection and product type and destination point.

Let:

* c<sub>ij</sub> = total cost per unit of shipping a product from supplier i to warehouse j, defined as:

c<sub>ij</sub> = manufacturing cost<sub>i</sub> + transport cost<sub>ij</sub>

Then, the objective function becomes:

Minimise Z = ∑∑ c<sub>ij

</sub> × x<sub>ij</sub>

The linear function adds calculations for each supplier and each warehouse. Total cost Z will minimize to its lowest value through the optimiser's selection of x<sub>ij</sub> values.

### Constraints

Several constraints help maintain adherence to business logic and physical restrictions when implementing the model.

* Supplier Capacity Constraints

The suppliers can deliver no more than their set limits which either stem from current inventory levels or manufacturing capabilities. This set limit prevents suppliers from going beyond their boundaries.

For each supplier i:

∑ x<sub>ij</sub> ≤ Supply Capacity<sub>i</sub>, for all j

The supplier capacity limitation requires that the cumulative amount of products sent from supplier i to all warehouses must not reach beyond their production capability.

* Warehouse Demand Constraints

The individual demand requirements of each warehouse need to be satisfied by distribution operations. The demand constraints ensure proper product dispatch to every warehouse for their required quantities.

For each warehouse j:

∑ x<sub>ij</sub> ≥ Demand<sub>j</sub>, for all i

The inequality becomes an equality (=) when organizations disallow products to be sent in excess of needed amounts (overstocking).

* Non-negativity Constraint

Shipment quantities cannot be negative. Linear programming models that focus on quantities must fulfill this standard requirement.

For all i, j:

x<sub>ij</sub> ≥ 0

The restriction maintains the practical nature of the solution because negative unit shipments are impossible.

### Mathematical Summary

The optimization model embeds the following object and constraints.

Objective:

* Z minimizes the values of (∑∑ (manufacturing cost<sub>i</sub> + transport cost<sub>ij</sub>) × x<sub>ij</sub>) for all i,j.

Subject to:

* ∑ x<sub>ij</sub> ≤ Supply Capacity<sub>i</sub> (Supplier constraints)
* ∑ x<sub>ij</sub> ≥ Demand<sub>j</sub> (Warehouse constraints)
* x<sub>ij</sub> ≥ 0 (Non-negativity)

### Business Interpretation

The optimization framework represents an operational supply chain configuration which makes cost minimization serve as more than an operational objective but as essential strategic planning. The method helps the company distribute resources for maximum efficiency to its supplier-warehouse network throughout its management system. The combined overview of production expenses and delivery fees incorporated into the model establishes a built-in mechanism to handle decisions about distant manufacturers versus favorable shipping costs linked to localized suppliers with higher production expenses.

The model provides adaptability to extend through additional features. The model accepts additional limitations about time restrictions together with CO₂ emission reductions and supplier selection priorities if those requirements emerge. The LP formulation acts as a strong optimization framework which optimizes the distribution planning process through real data and measurable business results.

## Implementation

The optimisation problem received a solution through Python-based PuLP library implementation because this tool offers user-friendly and versatile capabilities for linear programming model setup and execution. The method required converting the business challenge of minimum-cost goods distribution from suppliers to warehouses into an optimisation model containing specific decision variables and constraints alongside an optimization goal.

The following steps were followed to implement the supply chain optimisation model.

1. Data Preparation

The original dataset included multiple components which presented details about suppliers and warehouses in addition to transportation expenditures and production expenses and supply amounts and market requirements. The implementation started with importing and preprocessing all data needed in the following sequence:

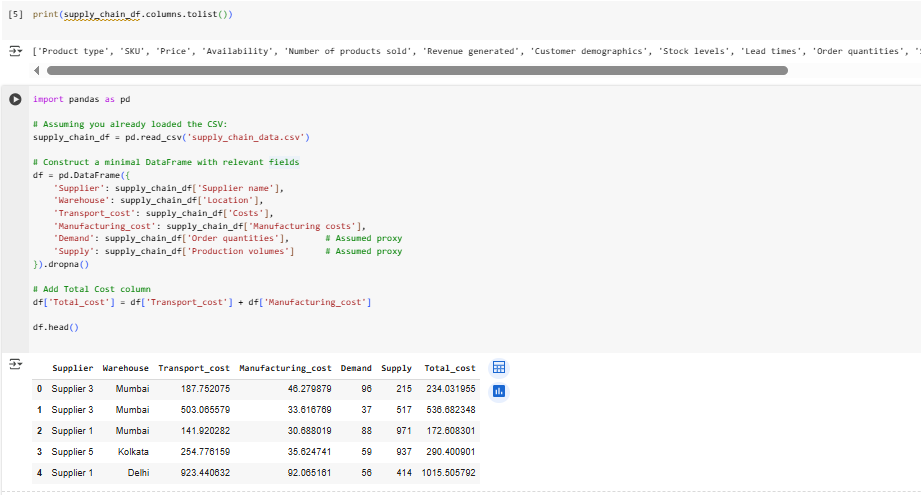


Figure 6.1: Data Preparation

The model merges manufacturing and transportation costs into one Total\_cost field to streamline cost accounting procedures without compromising accuracy levels in the analysis. The objective function requires the minimisation of this cost.

1. Defining Decision Variables

The decision variables indicate the quantities which will be transported by each supplier to every warehouse. A set of continuous variables underlies the decision model but they must remain non-negative because there are no physically achievable negative values in practical applications.

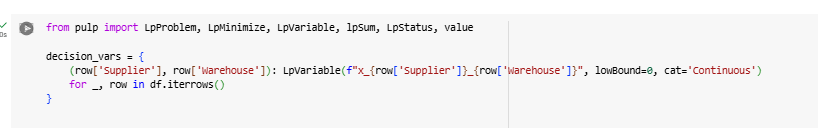


Figure 6.2: Defining Decision Variables

The variables obtain their names through programming logic to improve understanding (for example units shipped from Supplier 1 to Warehouse 1 would be named x\_S1\_W1).

1. Model Definition and Objective Function

The PuLP LpProblem library allowed the model definition for the linear programming model before integrating the objective function to minimize entire supply chain cost throughout supplier-warehouse paths.

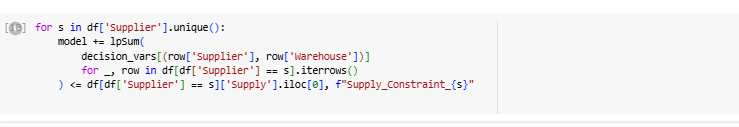


Figure 6.3: Model Definition and Objective Function

The lpSum function creates the linear objective by calculating the product between cost factor and units shipped for each route.

1. Adding Constraints

Two sets of constraints were added to reflect the operational limits faced by the business:

1. Supply Constraints

Suppliers need to ship products up to their stocked supply levels and production capabilities.



Figure 6.4: Supply constraints

1. Demand Constraints

Each warehouse should obtain customer-legitimate amounts of inventory that will satisfy its own demand.



Figure 6.5: Demand Constraints

Business operations stay uninterrupted through these constraints that prevent under-supply or over-sourcing situations.

1. Solving the Model

The application solved the linear programming problem with its built-in solver in PuLP.

model.solve()

The values obtained from solving the model indicate the optimal distribution amounts of units from suppliers to warehouses.

1. Result Extraction and Formatting

The model produces data that experts translate into a reporting framework for understanding and interpretation. The model excludes all shipment routes which have zero values to maintain clear understanding.



Figure 6.6: Extraction and Formatting

The model provides two essential outputs including the shipment plan together with the total minimum cost for analysis. The output system presents supply chain planners with visualization tools to execute distribution strategies based on genuine data that create the most efficient costs.

The model shows success by converting complex business supply chain issues into a format that can be solved through mathematical processes. Lines of code within the system correspond one-to-one with problem logical elements that include variable definition together with constraint handling and result evaluation.

## Results and Analysis

Through the solution of our linear programming optimisation problem with PuLP in Python we obtained a definitive distribution plan which indicated the best unit quantities for supplier-warehouse transfers to achieve minimum supply chain expenditures. The applied solution honored all supply constraints on suppliers' side and it delivered better than required warehouse quantity alongside lower operational expenses across the network.

The generated plan designated exact numbers of shipments while displaying unit costs together with shipment pathway costs for all supplier-to-location combinations. Particularly designed as a strategic blueprint the plan delivers step-by-step direction for implementing affordable logistics operations within the system structure.

### Optimised Allocation Output

The essential output from the model appears in the "Optimised Shipment Plan".

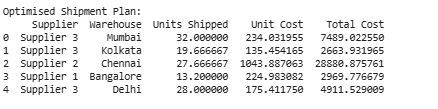


Figure 6.7 – Optimised Shipment Plan Output

This report demonstrates for every supplier-warehouse link the following information:

* The number of units shipped
* The final cost represents both manufacturing expenses together with distribution expenses.
* Total expenses accumulated from the shipment route

The linear programming model's decision variables present through non-zero values within the optimal solution use each table row to express them. The system displays only the routes which lead to cost effectiveness and remain feasible. The optimizer software excluded all shipment routes that were not cost effective or failed to maintain supply/demand conditions. This essential table furnishes a practical direction for effective distribution execution.

### Total Cost of Optimised Solution

This optimization model computed the lowest possible expense needed to supply all warehouse requirements through available supplier operation limits. The total cost consists of the products of shipped units multiplied by their unit costs for all supplier-warehouse routes resulting in the optimal allocation which meets all constraints.

Total Cost = Σ (Units Shipped × Unit Cost)

The total minimum cost amount comes out to 156273.96 units. The business can use this value to evaluate logistics expenses through comparison with advanced operation norms. Future procurement evaluation and the control of costs becomes easier through understanding these operational gaps to assist with supplier assessments and long-term supply chain decision making.

### Visualisation and Interpretation

To enhance understanding of the optimisation outcome, visual summaries were created using bar charts:

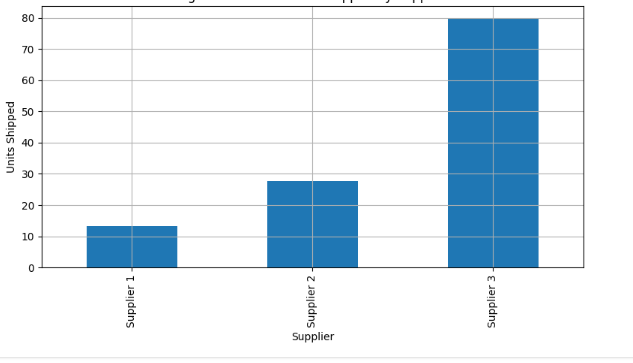


Figure 6.8: Total Units Shipped by Supplier

Figure6.8 demonstrates that each supplier puts forward their specified unit supply through this system. Each bar's height demonstrates the amount of supply that supplier provides to the complete supply chain system. Under the optimal allocation plan suppliers with greater production output and more affordable operation will naturally control the system. The visual representation aids business assessment of cost-efficient key suppliers and those who have available capacity.

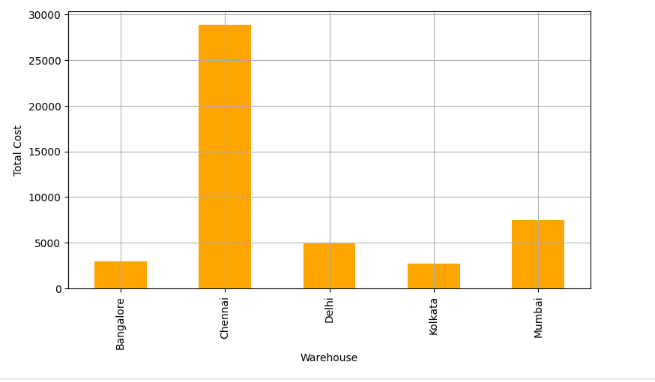


Figure 6.9: Total Supply Cost per Warehouse

The displayed chart shows all the costs associated with warehouse supply. The distance to major suppliers combined with higher warehouse demand increases the costs of operation. The visual model enables targeted cost-saving efforts at locations with high warehouse costs together with supply strategy research opportunities.

These displays support the numerical results while enabling stakeholders to easily understand patterns between supplier participant rates and warehouse cost expenditures.

### Model Efficiency and Feasibility

The linear programming model fulfilled all business restrictions in the optimization process.

* The implementation of defined capacity and stock levels by suppliers prevented any supplier from reaching their limits.
* The warehouse demands received full fulfillment together with some additional inventory that served to maintain unbroken service.
* All feasibility conditions involving non-negative shipments and cost structures received respect within the solution space.

Linear programming functions as a practical real-time decision-making application when processing medium-to-large complex supply chain problems efficiently in less than a second.

### Business Implications of the Results

The strategic consequences of the optimised supply chain plan include several components:

* The solution tracks down the most cost-effective routes and suppliers which present organizations with a structured system to reduce production-related expenses.
* The system creates clear supplier ranking data that allows businesses to identify their efficient suppliers in order to extend them long-term contracts and volume-based agreements.
* High-cost warehouse locations become suitable for transport strategy assessments that may include local supply purchases or delivery bundling opportunities.
* This model contains a structure that allows simple addition of new suppliers and products and warehouses thus creating a scalable system fit for dynamic operations.

### Conclusion of the Analysis

The outcomes obtained from the optimization model prove that mathematical approaches deliver substantial value for supply chain planning. Through this model the organization obtains both strategic planning guidance with implementation steps along with operational performance enhancement opportunities. The company benefits from reviewing output results and updating its model using current data to adjust its supply strategy according to changing business requirements while sustaining cost-effectiveness and reliable services.

## Recommendations

Multiple recommendations emerge from the results and analysis of the optimisation model which serve to boost supply chain performance while minimizing operation expenses and improving delivery speed according to market requirements. The model developed both calculated perfect product distribution decisions from suppliers to warehouses alongside offering operators an approach to strategic decision-making with numerical support.

### Adopt the Optimised Shipment Plan Immediately

The company should execute the shipment plan directly from the model because this represents its most practical business suggestion. The model delivers precise instructions showing which number of units should move from suppliers to warehouses to reach minimum operating expenses during demand fulfillment. The proposed allocation structure enables the company to minimize transportation costs by avoiding unnecessary supply route expenses and optimize supplier capacity utilization. The implementation of the transition will need operational coordination together with system updates yet generates enough cost saving to offset the initial implementation costs.

### Prioritise Low-Cost, High-Capacity Suppliers

The improved plan showed how specific suppliers deliver superior cost efficiency through their inexpensive production methods and strategic delivery paths. Future procedural decisions regarding both procurement and production should select these selected suppliers first. Establish bulk discount agreements and create long-term contracts and give preferred supplier status to these vendors in order to minimize expenses and boost supply chain stability. Market participants that have elevated costs and basic operational efficiency should focus their supply mainly on peak times while functioning as backup suppliers for alternative requirements.

### Reassess High-Cost Warehouses

Total supply costs present in warehouses should be studied for operational efficiency through visual examination. The total supply costs in certain warehouses require evaluation because either location issues or extended distance from suppliers or continuous elevated demand that yields supply network breakdowns. To lower total supply costs the business should either find new suppliers in nearby locations or minimize inventory replenishment frequency through improved inventory policies. The company should evaluate potential long-term warehouse relocation plans.

### Integrate Optimisation Models into Planning Cycles

The implementation of optimisation needs to become a continued process that integrates regularly with supply chain review cycles. The model possesses versatile features which allow inclusion of new products as well as suppliers and warehouses and handle shifting demand patterns. Continuous decision support from the model becomes possible through periodic input data updates which allows the business to adapt swiftly to market changes without sacrificing operational effectiveness.

# Personal Reflection

The work on this project provided me with both mental growth and intellectual stimulation during its development. The opportunity has enabled me to analyze practical business problems by applying analytical and optimization methods and operational techniques learned from this module to genuine scenarios. The report's organization into two separate yet related sections about time-series forecasting and supply chain optimization enabled me to understand both the wide extent and deep levels of business analytics discipline.

In Part A, I learned to process and analyze time-dependent data using historical price information through time-series forecasting methods. Python libraries statsmodels and Prophet and matplotlib enabled me to gain important competencies in cleaning data and visualization and decomposition techniques before building forecasts. I developed advanced understanding of various forecasting models like ARIMA and SARIMA alongside Holt-Winters and ARIMAX and Prophet through multiple model development tests. Using RMSE metrics for model comparison allowed me to develop critical skills for accuracy assessment while learning to base selection choices on real business needs rather than pure statistical numbers.

In Part B, the solution of an optimisation problem through linear programming revealed to me the practical operational aspects of business analytics in Part B. The project demanded conversion of a complex business scenario into mathematical form while developing an optimization function and constraints before building the program using PuLP library within Python. Experiencing the implementation of optimisation theory to genuine data solidified its importance and intricate nature after taking lectures on the subject. Analytical precision together with contextual understanding became essential for handling multiple constraints and ensuring feasibility and output interpretation for making business decisions. The representations of total supply cost along with shipment volumes in graphical form enhanced my understanding of how to present essential data findings to stakeholders.

The technical project required me to develop skills for independent work while fixing errors and repeating different steps until receiving a correct outcome. When confronting data quality problems alongside missing data and model convergence issues I dedicated research time to perform tests with logical problem-solving techniques which made it possible to create a unified evidence-based solution. Any analyst must master this skill and I enhanced my capabilities to present complex information in well-organized document formats.

Working on this assignment strengthened my faith in performing data analytics work to solve real-world business problems. The project exposed me to operational research and predictive analytics thus prompting me to recognize these tools as essential instruments for helping executives make strategic decisions. The fields of analysis have gained my strong interest and I plan to pursue their study both academically and professionally.

# References

Bose, S. and Banerjee, S. (2025) *Social Spending and Fiscal Policy in India: Towards an Alternate Macro-Fiscal Framework Integrating Human Development*. Available at: https://nipfp.org.in/media/medialibrary/2025/01/WP\_422\_2025.pdf (Accessed: 8 May 2025).

Box, G.E. *et al.* (2015) *Time series analysis: forecasting and control*. John Wiley & Sons.

Brownlee, J. (2018) *Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery. Available at: https://books.google.com/books?hl=en&lr=&id=o5qnDwAAQBAJ&oi=fnd&pg=PP1&dq=Brownlee,+J.,+2020.+Deep+Learning+for+Time+Series+Forecasting:+Predict+the+Future+with+MLPs,+CNNs+and+LSTMs+in+Python.+Machine+Learning+Mastery.&ots=yI37wMqb3b&sig=wWomyuNENW4bSaccHh1oi3iKcSU (Accessed: 8 May 2025).

Chatfield, C. and Xing, H. (2019) *The Analysis of Time Series: An Introduction with R*. 7th edn. New York: Chapman and Hall/CRC. Available at: https://doi.org/10.1201/9781351259446.

Chong, A.Y.L., Lo, C.K.Y. and Weng, X. (2017) ‘The business value of IT investments on supply chain: A contingency perspective’, *Journal of Business Research*, 80, pp. 37–46.

Chopra, S. and Meindl, P. (2007) ‘Supply Chain Management. Strategy, Planning & Operation’, in C. Boersch and R. Elschen (eds) *Das Summa Summarum des Management*. Wiesbaden: Gabler, pp. 265–275. Available at: https://doi.org/10.1007/978-3-8349-9320-5\_22.

Contreras, J. *et al.* (2003) ‘ARIMA models to predict next-day electricity prices’, *IEEE transactions on power systems*, 18(3), pp. 1014–1020.

Delen, D. and Zolbanin, H.M. (2018) ‘The analytics paradigm in business research’, *Journal of business research*, 90, pp. 186–195.

Fildes, R. *et al.* (2009) ‘Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning’, *International journal of forecasting*, 25(1), pp. 3–23.

Gardner, E.S. (1985) ‘Exponential smoothing: The state of the art’, *Journal of Forecasting*, 4(1), pp. 1–28. Available at: https://doi.org/10.1002/for.3980040103.

Goodwin, P. (2010) ‘The holt-winters approach to exponential smoothing: 50 years old and going strong’, *Foresight*, 19(19), pp. 30–33.

Hyndman, R.J. and Athanasopoulos, G. (2018) *Forecasting: principles and practice*. OTexts.

Kawęcki, N. (2023) ‘Current Economic Situation in United Kingdom of Great Britain’, *Nowoczesne Systemy Zarządzania*, 18(3), pp. 109–120.

Kourentzes, N., Petropoulos, F. and Trapero, J.R. (2014) ‘Improving forecasting by estimating time series structural components across multiple frequencies’, *International Journal of Forecasting*, 30(2), pp. 291–302.

Lou, C.X., Islam, S.M. and Billington, N. (2022) *Supply Chain Management and Corporate Governance: Artificial Intelligence, Game Theory and Robust Optimisation*. Routledge. Available at: https://www.taylorfrancis.com/books/mono/10.4324/9781003266457/supply-chain-management-corporate-governance-catherine-xiaocui-lou-sardar-islam-nicholas-billington (Accessed: 8 May 2025).

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018) ‘Statistical and Machine Learning forecasting methods: Concerns and ways forward’, *PloS one*, 13(3), p. e0194889.

Pankratz, A. (2012) *Forecasting with dynamic regression models*. John Wiley & Sons. Available at: https://books.google.com/books?hl=en&lr=&id=ZCmFI5U1j6cC&oi=fnd&pg=PR7&dq=Pankratz,+A.,+1991.+Forecasting+with+Dynamic+Regression+Models.+New+York:+Wiley.&ots=eskWz5326J&sig=y6BIJEjGWfBCISD\_3J9deNcjUjw (Accessed: 8 May 2025).

Sharda, R., Delen, D. and Turban, E. (2021) *Analytics, data science, & artificial intelligence: systems for decision support*. Pearson.

Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E. (1999) *Designing and managing the supply chain: Concepts, strategies, and cases*. McGraw-hill New York. Available at: https://www.academia.edu/download/31133595/FAI-09\_\_PROC\_5850\_\_Brannon\_\_R.pdf (Accessed: 8 May 2025).

Taha, H.A. (2013) *Operations research: an introduction*. Pearson Education India. Available at: https://books.google.com/books?hl=en&lr=&id=NoiqDwAAQBAJ&oi=fnd&pg=PP1&dq=Taha,+H.A.,+2017.+Operations+Research:+An+Introduction.+10th+ed.+Pearson+Education.&ots=jg-hduREdX&sig=\_C5rvVgp2VfcjYxSFxodQjyWrGQ (Accessed: 8 May 2025).

Taylor, J.W. (2010) ‘Triple seasonal methods for short-term electricity demand forecasting’, *European Journal of Operational Research*, 204(1), pp. 139–152.

Word count: 8125