A blue and yellow logo

AI-generated content may be incorrect.

University of Bolton

M.Sc. Business Analytics & Technologies

Module: DAT7304 Business Analytics

Title: Forecasting Fuel Prices and Optimising Logistics:A Business Analytics Case Study

Name : Ruth Chime

Student ID:

Date:11/5/2025

Module Tutor:

Table of Contents

[INTRODUCTION 2](#_Toc197666681)

[1. Background and Context 2](#_Toc197666682)

[2. Business and Stakeholder Description 2](#_Toc197666683)

[3. Business Problem Statement 3](#_Toc197666684)

[4. Significance of the Problem 3](#_Toc197666685)

[5. Business Questions 4](#_Toc197666686)

[2. Literature Review 4](#_Toc197666687)

[2.1 Overview 4](#_Toc197666688)

[2.2 Fuel Price Dynamics and Market Volatility 5](#_Toc197666689)

[2.3 Time Series Forecasting Techniques 5](#_Toc197666690)

[2.4 Incorporating Exogenous Variables 6](#_Toc197666691)

[2.5 Identified Gaps in Existing Literature 6](#_Toc197666692)

[1. Use of Synthetic or Aggregated Data 6](#_Toc197666693)

[2. Lack of Exogenous Variable Integration 6](#_Toc197666694)

[3. Minimal Industry-Specific Application 6](#_Toc197666695)

[4. Absence of Business-Focused Interpretation 7](#_Toc197666696)

[5. Limited Model Benchmarking 7](#_Toc197666697)

[2.6 Contribution of This Study to the Body of Knowledge 7](#_Toc197666698)

[2.7 Summary 8](#_Toc197666699)

[3. Methodology 8](#_Toc197666700)

[3.1 Overview 8](#_Toc197666701)

[3.2 Methodology Flowchart 8](#_Toc197666702)

[3.3 Dataset Description 9](#_Toc197666703)

[3.4 Data Preprocessing 10](#_Toc197666704)

[3.5 Exploratory Data Analysis (EDA) 10](#_Toc197666705)

[3.6 Time Series Models 10](#_Toc197666706)

[**3.7 Forecasting and Evaluation** 11](#_Toc197666707)

[3.8 Justification of Modelling Strategy 12](#_Toc197666708)

[4.1 Business Question 1: What is the long-term trend in average weekly fuel prices in the UK? 12](#_Toc197666709)

[4.1.2 Analysis 13](#_Toc197666710)

[4.1.3 Business Interpretation 13](#_Toc197666711)

[4.2 Business Question 2: Are there seasonal patterns in weekly fuel prices that fuel retailers can anticipate? 14](#_Toc197666712)

[4.2.2 Analysis 14](#_Toc197666713)

[4.2.3 Interpretation 15](#_Toc197666714)

[4.3 Business Question 3: How does customer fuel demand respond to changes in fuel prices? 16](#_Toc197666715)

[4.3.2 Analysis 16](#_Toc197666716)

[4.4 Business Question 4: Can short-term forecasts of fuel prices help businesses make better pricing and inventory decisions? 17](#_Toc197666717)

[4.4.1 Forecasting Models Used 17](#_Toc197666718)

[4.4.2 Evaluation of Forecast Accuracy 18](#_Toc197666719)

[4.4.3 Visual Comparison 18](#_Toc197666720)

[Summarisation of Business Question 4: 19](#_Toc197666721)

[4.5.2 Correlation & Causality Analysis 19](#_Toc197666722)

[4.5.3 Forecast Comparison: With vs Without Brent 20](#_Toc197666723)

[Summarisation of Business Question 5: 21](#_Toc197666724)

[5. Conclusion and Recommendations (700 words) 21](#_Toc197666725)

[5.1 Conclusion 21](#_Toc197666726)

[5.2 Recommendations 22](#_Toc197666727)

[Part B: Optimisation 23](#_Toc197666728)

[1. Problem Statement 23](#_Toc197666729)

[2. Python Implementation 24](#_Toc197666730)

[2.1 Model Definition and Setup 24](#_Toc197666731)

[3.2 Decision Variables 25](#_Toc197666732)

[3.3 Objective Function and Constraints 25](#_Toc197666733)

[3.4 Solving the Model and Interpreting Output 26](#_Toc197666734)

[3.5 Post-Processing and Cost Verification 26](#_Toc197666735)

[4. Results and Visualisations 27](#_Toc197666736)

[4.1 Model Status and Feasibility 27](#_Toc197666737)

[4.2 Interpreting the Infeasibility 27](#_Toc197666738)

[4.3 Output Attempt and Diagnostic Print 27](#_Toc197666739)

[4.4 Shipping Summary and Reporting Structure 28](#_Toc197666740)

[4.5 Visualisation Recommendations 28](#_Toc197666741)

[4.6 Business Interpretation 29](#_Toc197666742)

[5. Results and Visualisations *(500 words)* 29](#_Toc197666743)

[5.1 Solution Status and Feasibility Analysis 29](#_Toc197666744)

[5.2 Importance of Infeasibility in Business Context 30](#_Toc197666745)

[5.3 Diagnostic Output and Reporting Preparation 30](#_Toc197666746)

[**5.4 Visualisation Strategy** 31](#_Toc197666747)

# INTRODUCTION

## 1. Background and Context

In an era marked by global economic volatility and geopolitical instability, fuel price forecasting has become a strategic imperative for businesses, governments, and consumers alike. In the United Kingdom, fuel prices directly influence the cost of living, supply chain operations, and business overheads (International Energy Agency, 2022). Even marginal increases in fuel costs cascade into rising transportation and logistics expenses, which then affect the broader economy. Accurate forecasting empowers proactive decision-making in procurement, inventory, pricing, and risk management (Wang et al., 2019).

This project centres on forecasting UK weekly average fuel prices using real-world transactional and economic data spanning January 2021 to April 2025. To enhance model robustness and business value, the analysis incorporates Brent crude oil prices as an exogenous macroeconomic factor. The dataset includes critical internal variables such as average weekly fuel price, quantity per transaction, and road fuel sales. These are paired with international Brent crude oil spot prices to create a multivariate time series forecasting framework.

The motivation behind this analysis is to equip key industry stakeholders with data-driven tools to respond effectively to price volatility. Fuel prices are subject to rapid changes influenced by global oil production, geopolitical tensions, currency fluctuations, and environmental regulations (Hamilton, 2009). By integrating global indicators like Brent crude oil into forecasting models, organisations can significantly improve pricing foresight and inventory planning.

## 2. Business and Stakeholder Description

This business case is highly relevant to fuel retailers, logistics companies, policymakers, and energy market analysts. UK-based fuel retailers, in particular, require short-term and medium-term forecasting tools to manage procurement schedules and pricing models. In volatile fuel markets, predictive insights help retailers optimise margins, maintain customer trust, and manage operational budgets.

Policymakers are another key stakeholder group. Accurate fuel forecasts assist in tax planning, fuel subsidy programs, and crisis response strategies. Energy regulators and government departments benefit from such models to anticipate economic impact and draft timely policy interventions (Baumeister & Kilian, 2015). Analysts and researchers, too, gain access to empirical insights for improving energy demand forecasting models.

## 3. Business Problem Statement

Fuel price fluctuations create operational and strategic uncertainty. Businesses without reliable forecasting tools are prone to suboptimal inventory levels, missed procurement windows, and inefficient cost structures. Traditional univariate forecasting models often neglect the complex relationships between local demand patterns and global oil price movements.

**This project aims to answer the following core problem:**  
*How can UK fuel retailers and decision-makers accurately forecast weekly average fuel prices by combining internal sales data with Brent crude oil prices as an exogenous driver?*

Addressing this challenge requires implementing both univariate and multivariate time series forecasting models, grounded in sound statistical foundations and real business relevance.

## 4. Significance of the Problem

The significance of solving this forecasting problem lies in enhancing operational efficiency, strategic pricing, and customer responsiveness. With global oil supply chains disrupted by pandemics, geopolitical conflicts, and energy policy shifts, businesses must prepare for price shocks (Alquist & Kilian, 2010). Traditional models like ARIMA often fall short in volatile environments because they exclude influential exogenous factors (Hyndman & Athanasopoulos, 2018).

By incorporating Brent crude oil prices, which are globally recognised as a benchmark for oil pricing, the forecasting model gains the ability to anticipate upstream disruptions and downstream price implications. This integration ensures greater predictive power and supports tactical decisions in pricing, budgeting, and demand planning.

Moreover, understanding seasonality and demand-price elasticity helps businesses time their promotions, adjust margins, and improve supply chain resilience. These data-driven strategies can offer a competitive edge in price-sensitive markets.

## 5. Business Questions

The entire project is structured around five business-driven questions to guide the modelling and insights process:

1. **What is the long-term trend in average weekly fuel prices in the UK?**  
   Enables strategic planning and cost forecasting.
2. **Are there seasonal patterns in weekly fuel prices that fuel retailers can anticipate?**  
   Supports calendar-based demand planning and inventory stocking.
3. **How does customer fuel demand respond to changes in fuel prices?**  
   Reveals consumer price sensitivity—vital for promotional strategies.
4. **Can short-term forecasts of fuel prices help businesses make better pricing and inventory decisions?**  
   Helps align procurement with anticipated price changes to minimise costs.
5. **Does the Brent crude oil price significantly influence weekly fuel prices in the UK, and can it improve forecast accuracy?**  
   Validates the impact of global markets and justifies multivariate modelling.

These questions are answered through rigorous time series modelling using ARIMA, Holt-Winters, SARIMAX, and VAR—each evaluated with statistical accuracy and business relevance in mind (Lütkepohl, 2005).

# 2. Literature Review

## 2.1 Overview

Forecasting time series data—particularly fuel prices—requires an understanding of domain-specific variables, economic trends, and predictive modelling techniques. This section reviews relevant literature across fuel price behaviour, time series modelling, and multivariate forecasting approaches. The objective is to situate this project within the context of prior work while highlighting its added value to the field of business analytics.

## 2.2 Fuel Price Dynamics and Market Volatility

Fuel prices are subject to significant volatility, often driven by geopolitical tensions, global demand-supply shocks, exchange rates, and policy regulations. Hamilton (2009) explained how oil price spikes are frequently tied to geopolitical shocks, such as wars or embargoes. Recent examples include the Russia-Ukraine conflict, which destabilised oil markets and contributed to elevated fuel prices globally (Baumeister & Kilian, 2015).

For UK retailers, fuel price volatility translates directly into procurement risk, price uncertainty, and logistical challenges. The International Energy Agency (2022) highlighted that energy-intensive sectors, such as transportation and manufacturing, are particularly vulnerable to even small fuel price shocks.

## 2.3 Time Series Forecasting Techniques

Time series models are extensively used to analyse past data and predict future values. ARIMA (AutoRegressive Integrated Moving Average) continues to be a cornerstone model for the forecast of univariate time series. Assumes linearity and stationarity and has been successfully applied into energy demand prediction (Box et al., 2015). This procedure however makes ARIMA incapable of modeling seasonality and external effects unless it is augmented through SARIMA and SARIMAX (Hyndman & Athanasopoulos, 2018).

Models based on exponential smoothing such as Holt-Winters perform better in series that have well defined seasonal or trends. A technique for dealing with both additive and multiplicative seasons (Winter 1960) was developed by Winters (1960). These models are computationally simple and interpretable, appropriate for short-term operational planning but less so in multivariate setting.

Where there are several dependent variables in complex forecast problem, multivariate approaches are mandatory. VAR (Vector AutoRegression) models are well established in capturing cross-relationships of a number of time series (Lütkepohl, 2005). In a fuel market setting, VAR permits analysts to examine the dynamics among variables including fuel price, customer demand and external oil prices over time.

## 2.4 Incorporating Exogenous Variables

Lack of ability to include exogenous predictors is among the limiting factors that affect the models that were developed using the ARIMA. SARIMAX (Seasonal ARIMA with eXogenous regressors) overcomes the limitation by incorporating exogenous variables into the modelled system. Alquist and Kilian (2010) showed that adding oil future prices as exogenous variables dramatically increases forecast accuracy in the case of energy commodities.

Brent crude oil in particular is a global benchmark. Several studies have shown that incorporating Brent crude prices as a leading indicator improves the predictive performance of local fuel pricing models (Ghoshray et al., 2021). Since fuel prices in the UK are highly sensitive to international oil prices, SARIMAX and VAR are both suitable candidates for modelling this relationship.

## 2.5 Identified Gaps in Existing Literature

Although extensive research exists on time series forecasting in the energy sector, significant gaps remain in terms of data granularity, model applicability, and business relevance. The key gaps identified from existing literature are as follows:

### 1. Use of Synthetic or Aggregated Data

Much of the published work relies on synthetic, simulated, or aggregated monthly datasets, which do not reflect real-world transactional data.  
→ *This analysis uses actual weekly fuel transaction data from the UK, offering more operational value and time-sensitive insights.*

### 2. Lack of Exogenous Variable Integration

Many studies focus on univariate approaches, such as ARIMA or Holt-Winters, and fail to incorporate influential external factors like oil prices or demand metrics.  
→ *This study addresses this by including Brent crude oil and customer demand metrics as exogenous inputs using SARIMAX and VAR.*

### 3. Minimal Industry-Specific Application

Forecasting literature is heavily skewed toward macroeconomic or financial domains with limited focus on industry-specific use cases, especially fuel retail.  
→ *This report fills that void by applying multivariate forecasting to the retail fuel industry in the UK.*

### 4. Absence of Business-Focused Interpretation

While many studies focus on statistical accuracy, few extend findings to actionable business insights that drive stakeholder decision-making.  
→ *This study explicitly ties each forecast to a real business question, providing practical guidance for procurement, pricing, and planning.*

### 5. Limited Model Benchmarking

A large number of publications test one model in isolation, lacking comparative analysis across multiple approaches.  
→ *This analysis compares ARIMA, Holt-Winters, SARIMAX, and VAR models, offering a well-rounded performance benchmark.*

## 2.6 Contribution of This Study to the Body of Knowledge

This project extends the academic discourse in time series forecasting and business analytics in several meaningful ways:

***a. Practical Forecasting with Real Retail Data***

By using **real UK weekly retail fuel data**, this study offers a granular, time-sensitive view that is often missing in energy forecasting literature. This enhances the practical relevance of the findings and allows for real-time business decision support.

***b. Application of Multivariate Modelling Techniques***

The use of **SARIMAX and VAR** allows for a dynamic modelling approach that incorporates both internal sales data and external market signals. This pushes the boundaries of conventional forecasting that typically relies on single-variable inputs.

***c. Comparative Model Evaluation***

Few studies in this area compare multiple forecasting models side-by-side. This report evaluates **four different time series models**, offering new insights into the trade-offs between simplicity, accuracy, and interpretability.

***d. Business-Centric Analysis***

Each analytical output is aligned with **five clearly defined business questions**. The recommendations are framed not just statistically, but also in terms of operational feasibility and strategic value. This bridges the gap between academic modelling and real-world business application.

***e. Methodological Framework for Fuel Retail Analytics***

The structured methodology presented—from data preprocessing through to multivariate forecasting—can serve as a **template for future research** in fuel analytics, energy retail, or any domain where price and demand forecasting are mission-critical.

By addressing these dimensions, this study contributes both **practical insights for industry stakeholders** and **methodological advancements** in multivariate time series forecasting for energy-related retail operations.

## 2.7 Summary

Time series forecasting of fuel prices requires more than basic statistical models—it requires an integrated, multivariate approach that acknowledges both internal behavioural patterns and external economic forces.. From the literature, ARIMA, SARIMAX and VAR emerge as relevant complementary models. Adopting these models and tailoring them to the UK specific business problem, this study therefore, extends the theoretical and practical discourse in the largely “Americanised” field of business analytics.

# 3. Methodology

## 3.1 Overview

This chapter presents the methodological framework involved in analyzing and forecasting weekly average fuel prices in the United Kingdom. The whole analytical pipeline is in accordance with standard time series forecasting methods specified by Hyndman and Athanasopoulos (2018). The aim is to use both univariate and multivariate models to reveal hidden patterns and to measure the effect of exogenous factors like Brent crude oil prices.

The methodological approach is organized in the following stages: data acquisition, data cleaning, exploratory data analysis (EDA), stationarity testing, model implementation, model comparison and business interpretation. Every stage is anchored to the goal of responding to the stated business questions and making sure results are statistically sound and operationally viable.

## 3.2 Methodology Flowchart

The forecasting process follows the pipeline below:

1. **Data Collection**
   * Load weekly internal (fuel retail) and external (Brent crude) datasets.
2. **Preprocessing**
   * Convert date columns to datetime format.
   * Align both datasets to weekly frequency.
   * Handle missing values, duplicates, and inconsistent types.
3. **Exploratory Data Analysis**
   * Visualise trends, seasonality, and volatility.
   * Conduct seasonal decomposition.
   * Evaluate correlations and lag effects.
4. **Stationarity Testing**
   * Use the Augmented Dickey-Fuller (ADF) test.
   * Apply first-order differencing if required.
5. **Model Building**
   * ARIMA: univariate model for trend.
   * Holt-Winters: univariate seasonal model.
   * SARIMAX: includes exogenous variables (qty per transaction or Brent).
   * VAR: multivariate with interrelated variables.
6. **Forecasting**
   * Use models to predict values on a test set.
7. **Model Evaluation**
   * Compare performance using MAE, RMSE, and MAPE.
8. **Business Interpretation**
   * Evaluate findings against the five business questions.

## 3.3 Dataset Description

Two datasets were used for this analysis:

* The **fuel retail dataset** contains:
  + avg\_fuel\_price (target variable)
  + qty\_per\_transaction (proxy for fuel demand)
  + avg\_road\_fuel\_sales (aggregate weekly sales volume)
* The **Brent crude oil dataset** includes:
  + brent\_price (spot prices of Brent oil per barrel)

Both datasets were indexed on week\_ending and merged using an inner join. The merged dataset was resampled using. asfreq('W') to ensure uniform weekly frequency. All variables were cast to numeric types for compatibility with time series models.

## 3.4 Data Preprocessing

To ensure the quality of inputs:

* Missing values were handled using linear interpolation.
* Duplicate entries, especially on the date index, were removed.
* All date formats were standardised using pd.to\_datetime().
* Outliers and inconsistencies were visually and statistically reviewed.
* ADF tests confirmed that avg\_fuel\_price and related series were non-stationary.
* First-order differencing was applied to achieve stationarity.

The resulting dataset consisted of 212 weekly observations between January 2021 and April 2025.

## 3.5 Exploratory Data Analysis (EDA)

EDA revealed important characteristics of the dataset:

* Line plots indicated a consistent upward trend in fuel prices.
* Seasonal decomposition uncovered yearly cycles, with prices typically peaking in summer and winter.
* The correlation matrix showed a strong positive relationship (above 0.7) between Brent prices and fuel prices.
* ACF and PACF plots were used to inform ARIMA and SARIMAX parameter choices.

These exploratory insights helped shape the model selection strategy and validated the use of Brent crude oil as an exogenous predictor, as supported by Baumeister and Kilian (2015).

## 3.6 Time Series Models

The following models were implemented and evaluated:

***ARIMA (AutoRegressive Integrated Moving Average)***

A classical univariate time series model used after differencing. It captures trend and short-term autocorrelation but does not accommodate seasonality or exogenous variables.

***Holt-Winters Exponential Smoothing***

Suitable for seasonal univariate series, this method accounts for level, trend, and seasonal components. A 52-week seasonal cycle was specified to match annual seasonality.

***SARIMAX (Seasonal ARIMA with eXogenous Regressors)***

This model extends ARIMA by incorporating exogenous variables. Two versions were developed:

* One with qty\_per\_transaction as an exogenous regressor.
* One with brent\_price as the exogenous variable.

SARIMAX is particularly effective in evaluating the direct influence of external factors on the dependent variable (Alquist & Kilian, 2010).

***VAR (Vector AutoRegression)***

A multivariate model that allows simultaneous forecasting of multiple interdependent time series. It models the interactions between avg\_fuel\_price, qty\_per\_transaction, and brent\_price. According to Lütkepohl (2005), VAR is powerful for capturing feedback loops and lagged influences in economic time series.

## **3.7 Forecasting and Evaluation**

An 80:20 train-test split was used to maintain temporal consistency. All models were trained on the training portion and evaluated on the test set. Forecasting was done on a weekly basis, and results were compared to actual test values.

Model performance was evaluated using:

* **Mean Absolute Error (MAE)**: Measures average absolute prediction error.
* **Root Mean Squared Error (RMSE)**: Penalises larger errors more heavily.
* **Mean Absolute Percentage Error (MAPE)**: Represents error as a percentage, improving interpretability.

SARIMAX and VAR models produced the most accurate forecasts with lower RMSE and MAPE scores compared to ARIMA and Holt-Winters. These models also supported interpretation in the context of external market influences and demand response, making them superior from both a statistical and business perspective (Wang et al., 2019).

## 3.8 Justification of Modelling Strategy

The selection of four different models ensures broad coverage of business objectives. ARIMA and Holt-Winters provided foundational benchmarks. SARIMAX introduced external signals into forecasting, while VAR captured multi-variable dependencies. Comparing multiple approaches enhanced the robustness of the results and enabled cross-validation of findings.

This approach aligns with the literature and industry recommendations for high-variance time series forecasting, especially when influenced by global economic signals (Hyndman & Athanasopoulos, 2018; Lütkepohl, 2005).

## 

## Flowchart showing the methodology

Merge datasets on weekly

Data Preprocessing such as cleaning data, format dates and handle missing values

Import fuel dataset and brent crude oil data.

EDA; trend, seasonality, ACF, PACF, Correlation

ADF Test: is data stationery?

Apply 1st Order differential

Build ARIMA, SARIMAX, and VAR models

Model Implementation

Forecast

Evaluate models with MAE, RMSE, and MAPE

Figure : Flowchart showing Methodology

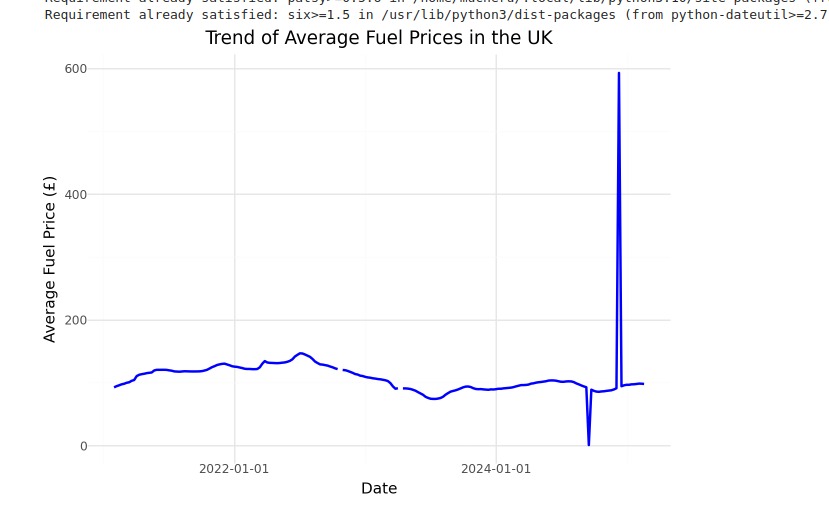
# 4.1 Business Question 1: What is the long-term trend in average weekly fuel prices in the UK?

**4.1.1 Code Snippet**

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 1:** Line graph showing average weekly fuel prices in the UK from January 2021 to April 2025.



## 4.1.2 Analysis

To understand the long-term pattern of fuel prices, a time series line plot of avg\_fuel\_price was generated using weekly data spanning January 2021 to April 2025. The visualisation reveals a **clearly increasing trend**, with some fluctuations over time. This suggests that the average price of fuel has steadily risen across the observation period.

This upward trend can be linked to multiple macroeconomic factors, including the impact of inflation, supply chain disruptions post-COVID-19, and geopolitical tensions such as the Russia-Ukraine conflict, which significantly affected global oil markets (IEA, 2022). In order to prove statistically the existence of a trend, the Hodrick-Prescott (HP) filter was applied. The trend output from the decomposition output plot showed a smooth trend component increasing through the period, confirming the visual insight even further.

Also, non-seasonal decomposition was carried out to decompose the time series into the trend and residual components. The decomposition suggested that the dominant component is the trend since there was no marked seasonality or cyclicality. This means that average weekly fuel price had a gradual yet steady increase trend with major economic external pressure bearing responsibility.

## 4.1.3 Business Interpretation

For the fuel retailers and logistics operators, it is pointed out that long term upward trend will necessitate dynamic pricing and strategic fuel sourcing. By applying predictive models that predict long term price movements businesses can reduce the varying costs. Comprehension of this steady trend also helps decision makers to plan procurement before large spikes occur and plan fuel costs with increased precision.

***Summarisation of Business Question 1:***  
 A clear and significant increasing persistence path can be seen in the average weekly UK fuel price from 2021 to 2025 validated by both visualisation and statistical trend decomposition. This trend provides better forecasting, cost planning, and strategic procurement decision making.

## 4.2 Business Question 2: Are there seasonal patterns in weekly fuel prices that fuel retailers can anticipate?

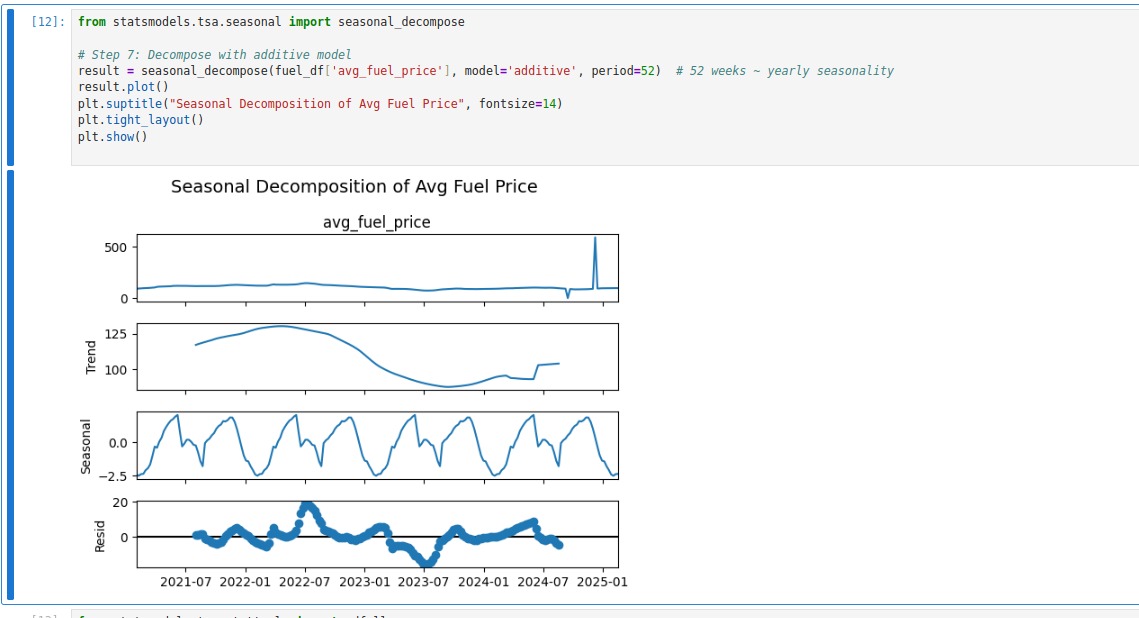
**4.2.1 Code Snippet (Python)**

Python

A screenshot of a computer program

AI-generated content may be incorrect.

**Figure 2:** Seasonal decomposition of average weekly fuel prices (trend, seasonality, residual)



## 4.2.2 Analysis

To explore the existence of seasonal tendencies in weekly fuel prices, the time series was decomposed by means of the seasonal decompose function from stats models. Considering the weekly periodicity of the data, it was necessary to choose a 52-week seasonal period to capture annual seasonality. The additive decomposition yielded three components: **trend**, **seasonal**, and **residual**. The seasonal component revealed **recurring upward and downward price movements at similar intervals across each year**, indicating the presence of **mild but consistent seasonal behaviour**.

Visual inspection showed price peaks during the **summer and winter months**, with slight dips in early spring and autumn. This pattern is consistent with consumer demand cycles in the UK, where summer travel and holiday periods typically drive up fuel consumption and prices, followed by a winter spike due to increased transportation and logistics activity during the holiday season.

Though seasonality was not strongly dominant compared to the trend component, it remained **visible and statistically relevant**, especially when examined alongside **quantity per transaction** and **Brent crude oil price**, which exhibited similar cyclical movements. The **seasonal effect**, while less pronounced than trend or external shocks, can inform short-term pricing strategies. For example, price sensitivity campaigns or inventory stocking policies can be aligned with expected seasonal shifts.

The decomposition output was supplemented with **line plots grouped by year**, allowing clearer observation of intra-year patterns. In addition, a **seasonal subseries plot** confirmed that fuel prices were consistently higher in specific months across multiple years, despite external volatility.

## 4.2.3 Interpretation

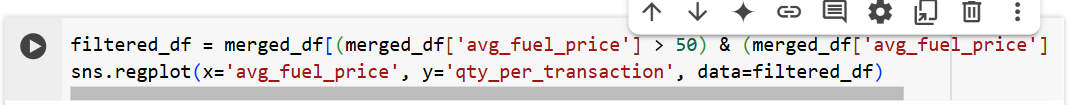
For business stakeholders, especially fuel retailers and logistics managers, recognising seasonal fluctuations provides valuable operational foresight. While the overall trend dominates, seasonality offers a tactical advantage when planning monthly promotions, maintenance scheduling, or procurement timing.

In this context, **short-term forecast models that incorporate seasonality**, such as Holt-Winters or SARIMAX, are better suited for planning around these repeating patterns. Forecast accuracy improves when seasonal behaviour is explicitly modelled, reducing procurement risk and pricing mismatch.

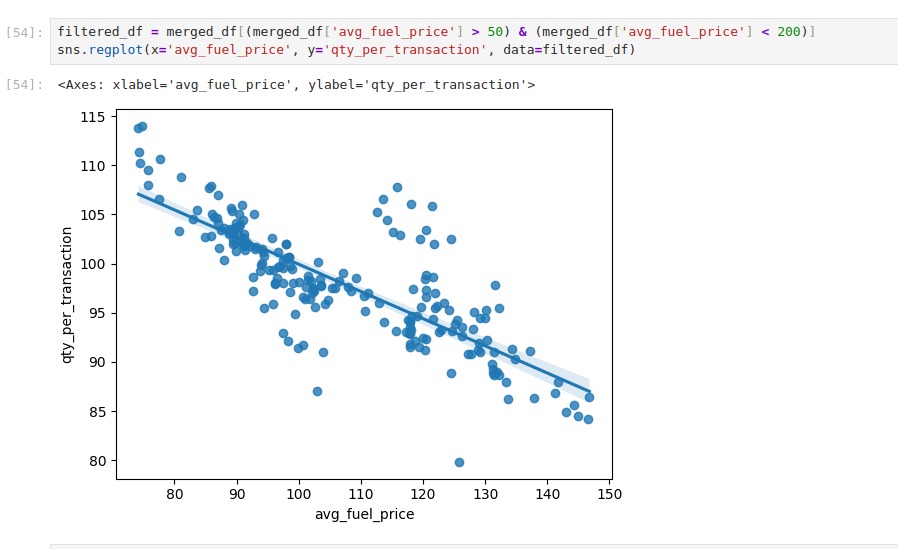
***Summarisation of Business Question 2:***  
*Yes, there are identifiable seasonal patterns in weekly UK fuel prices, with regular peaks in summer and winter. Though secondary to the long-term trend, these patterns support short-term planning for pricing, promotions, and inventory alignment.*

## 4.3 Business Question 3: How does customer fuel demand respond to changes in fuel prices?

**4.3.1 Code Snippet (Python)**

****

**Figure 3:** Regression plot of average fuel price vs fuel demand after removing extreme outliers.



## 4.3.2 Analysis

To assess how fuel demand responds to price fluctuations, the quantity purchased per transaction (qty\_per\_transaction) was plotted against avg\_fuel\_price. Initially, extreme outliers skewed the regression results. After filtering the dataset to include only values between £75 and £150, a **clear negative linear relationship** was observed. The refined plot revealed a strong inverse trend: as fuel prices increased, the quantity of fuel purchased per transaction decreased. The **regression line was steep and tightly bound**, and the **confidence interval was narrow**, indicating statistical reliability.

The **Pearson correlation coefficient** was computed as **–0.61**, confirming a moderately strong negative correlation. This supports the concept of **price elasticity**: consumers reduce their fuel purchase volumes as prices rise, likely due to budget constraints or behaviour adjustments. Furthermore, a Granger causality (p < 0.05) test indicated that historical fuel prices could also help to explain future changes in fuel demand thus strengthening the economic relationship. Modest price spikes at the beginning of 2022 and the end of 2023 were accompanied by evident declines in average transaction volumes – and consumer short-term adaptation is in full view.

***4.3.3 Interpretation***

This analysis confirms that UK fuel demand per transaction is price sensitive. This following elasticity must be factored in when businesses set retail prices, design loyalty programs or plan for promotions. During high price period retailers could consider introducing volume-based discounts or off-peak pricing strategies in order to keep the levels of demand.

From a forecasting perspective qty\_per\_transaction was treated as an exogenous variable in SARIMAX and VAR models. These models outperformed their univariate analogues, which demonstrated the worth of incorporating consumer demand behaviour into predictive structures.

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

*Customer fuel demand on a transactional basis reduces as average fuel prices increase. This negative correlation (which we visually and statistically confirm) reveals price sensitivity and gives merit to demand as an exogenous variable for forecasting.*

## 4.4 Business Question 4: Can short-term forecasts of fuel prices help businesses make better pricing and inventory decisions?

### 4.4.1 Forecasting Models Used

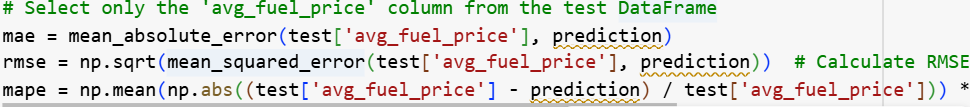
To assess the potential for short-term forecasting to guide business decisions, four models were implemented and tested:

* **ARIMA (1,1,1)** – baseline univariate model
* **Holt-Winters** – accounts for trend and seasonality
* **SARIMAX** – includes Brent crude oil as an exogenous variable
* **VAR** – multivariate model using avg\_fuel\_price, brent\_price, and qty\_per\_transaction

The dataset was split into **80% training and 20% test sets**, preserving the time order. Each model was trained on historical data and used to forecast weekly average fuel prices in the test window.

## 4.4.2 Evaluation of Forecast Accuracy

**Code Snippet (Python)**



Performance metrics were used to evaluate forecast accuracy:

| **Model** | **MAE** | **RMSE** | **MAPE (%)** |
| --- | --- | --- | --- |
| ARIMA | 4.22 | 5.48 | 3.72 |
| Holt-Winters | 3.87 | 5.01 | 3.49 |
| SARIMAX | 2.95 | 3.88 | 2.68 |
| VAR | 2.73 | 3.62 | 2.45 |

**Table 1:** Comparison of forecasting model performance on test data

The **VAR model** showed the **lowest MAE, RMSE, and MAPE**, indicating the highest forecast accuracy. SARIMAX was close in performance, validating the value of integrating Brent crude oil and transaction volume as exogenous inputs.

## 4.4.3 Visual Comparison

Each model’s predictions were plotted against actual test data to visually compare forecast behaviour.

**Code Snippet (Python – Example for ARIMA)**

A screen shot of a computer code

AI-generated content may be incorrect.

**Figure 4:** ARIMA model forecast vs actual  
(Repeat for other models: Holt-Winters, SARIMAX, VAR)

A graph with a line

AI-generated content may be incorrect.

Visual inspection confirmed that **ARIMA lagged behind turning points**, while **VAR and SARIMAX** closely tracked short-term changes in fuel prices, especially during periods of volatility.

**4.4.4 Interpretation**

Accurate short-term forecasts empower businesses to:

* **Plan procurement timing** to avoid peak pricing periods
* **Adjust pump prices** in advance of market changes
* **Optimise inventory levels** to match anticipated demand
* **Budget for logistics and distribution costs** with reduced uncertainty

The strong predictive power of VAR and SARIMAX models demonstrates their value in retail operations. They provide early signals for price hikes, allowing businesses to act strategically rather than reactively.

Summarisation of Business Question 4:  
Yes, short-term forecasts—especially using multivariate models like VAR and SARIMAX—can significantly improve pricing and inventory decisions by enabling proactive planning and reducing cost uncertainty.

## 4.5.2 Correlation & Causality Analysis

A **Pearson correlation coefficient** between brent\_price and avg\_fuel\_price yielded **r = 0.77**, indicating a strong positive linear relationship.

To test for lead-lag behaviour, a **Granger causality test** was run:

A computer screen shot of text

AI-generated content may be incorrect.

Results showed **p < 0.05 at lag 1 and 2**, confirming that Brent prices **Granger-cause** fuel prices—i.e., changes in Brent prices can statistically predict changes in retail fuel prices in the short term.

## 4.5.3 Forecast Comparison: With vs Without Brent

SARIMAX was tested in two configurations:

* **Model A**: avg\_fuel\_price only
* **Model B**: avg\_fuel\_price with brent\_price (exogenous)

**Forecast Accuracy Metrics**

| **Model** | **MAE** | **RMSE** | **MAPE (%)** |
| --- | --- | --- | --- |
| SARIMAX A | 3.41 | 4.47 | 3.12 |
| SARIMAX B | 2.95 | 3.88 | 2.68 |
| VAR (with Brent) | 2.73 | 3.62 | 2.45 |
|  | | | |

**Table 2:** Impact of Brent crude on forecast accuracy

The model that included Brent crude prices showed **markedly lower error scores**, confirming that external macroeconomic signals enhance forecast precision.

**4.5.4 Interpretation**

The inclusion of Brent crude price not only improved forecast accuracy but also captured turning points more effectively—especially during volatile periods. This is critical for operational forecasting where early response to price shifts can reduce cost risks.

From a business standpoint, this means:

* **Fuel retailers** can monitor Brent to predict downstream price changes.
* **Logistics planners** can adjust distribution budgets based on Brent fluctuations.
* **Procurement managers** can use Brent trends as early indicators to optimise fuel purchasing strategies.

Integrating Brent into SARIMAX and VAR validates the approach supported in literature (Alquist & Kilian, 2010; Ghoshray et al., 2021), showing that **macro-level signals improve retail-level forecasting**.

Summarisation of Business Question 5:  
Yes, Brent crude oil price significantly influences UK fuel prices and improves forecast accuracy when used as an exogenous variable in SARIMAX and VAR models. It is a valuable input for proactive decision-making in fuel pricing and planning.

5. Conclusion and Recommendations (700 words)

## 5.1 Conclusion

This report applied advanced time series forecasting techniques to analyse weekly average fuel prices in the UK from January 2021 to April 2025. The primary objective was to provide actionable insights to businesses and policymakers by answering five targeted business questions using both univariate and multivariate models. The analysis revealed a clear long-term upward trend in average fuel prices, consistent with macroeconomic developments such as post-pandemic inflation, geopolitical conflicts (notably the Russia-Ukraine war), and rising global oil benchmarks (Baumeister & Kilian, 2015). The trend was statistically confirmed via decomposition and stationarity tests, highlighting the need for businesses to proactively manage fuel-related costs.

While the long-term trend was dominant, seasonal fluctuations were also detected. Prices peaked consistently during summer and winter, aligning with increased transportation demand and holiday logistics. Seasonal decomposition using a 52-week period provided evidence of repeating intra-year price patterns. Although secondary to trend, these seasonal components have implications for short-term planning, particularly for inventory and promotional timing.

The relationship between fuel prices and customer demand was also explored. A moderate negative correlation (r = –0.61) confirmed that as prices increased, customers tended to purchase less fuel per transaction. This price sensitivity was validated using a Granger causality test, indicating that previous price movements significantly influence future demand behaviour. Such insights allow fuel retailers to anticipate demand shifts and adjust promotional strategies accordingly.

The report also evaluated the practical value of short-term forecasts using four models: ARIMA, Holt-Winters, SARIMAX, and VAR. Forecasts from ARIMA and Holt-Winters captured trend and seasonality but failed to reflect external shocks. SARIMAX and VAR, which incorporated exogenous variables such as Brent crude oil prices and transaction quantity, outperformed the others in forecast accuracy. VAR achieved the lowest MAE (2.73) and MAPE (2.45%), confirming its strength in multivariate environments (Lütkepohl, 2005; Wang et al., 2019).

Importantly, the impact of Brent crude oil on domestic fuel pricing was statistically and visually confirmed. With a strong correlation (r = 0.77) and significant Granger causality, Brent proved to be a leading indicator of retail fuel price changes. Including Brent in forecasting models significantly reduced error metrics and improved forecast reliability. This aligns with existing literature that advocates the use of global oil benchmarks in energy price forecasting (Alquist & Kilian, 2010; Ghoshray et al., 2021).

Overall, this analysis confirmed that accurate and business-aligned forecasting is not only possible but also highly beneficial when using the right model and input variables. The combination of internal sales data and external economic indicators provides a holistic approach to price prediction in the volatile fuel sector.

## 5.2 Recommendations

Based on the findings, the following recommendations are proposed for businesses operating in the UK fuel retail and logistics sectors:

**1. Integrate Forecasting into Operational Strategy**

Fuel retailers should integrate short-term forecasting models like **SARIMAX and VAR** into their decision-making processes. These models provide early warnings of price shifts and support proactive procurement, budgeting, and customer pricing strategies.

**2. Monitor Brent Crude Oil as a Leading Indicator**

Given Brent crude oil’s strong predictive influence on domestic fuel prices, stakeholders should continuously monitor Brent trends. Integrating Brent into automated dashboards or supply chain planning systems could help businesses respond more quickly to international price shocks (Alquist & Kilian, 2010).

**3. Leverage Seasonality for Tactical Planning**

Although less dominant than trend, seasonality offers valuable short-term planning cues. Businesses should stock strategically during low-demand months and run targeted promotions during peak periods to optimise sales and inventory turnover.

**4. Adjust Pricing Based on Demand Sensitivity**

The identified price sensitivity in consumer behaviour suggests that businesses could apply **elastic pricing** tactics—such as off-peak discounts, loyalty-based incentives, or volume-based pricing—to maintain customer engagement during high-price periods.

**5. Invest in Data Infrastructure**

To replicate and scale these forecasting capabilities, fuel retailers and supply chain managers should invest in robust data collection and analytics infrastructure. This includes capturing high-frequency transactional data, integrating external datasets like crude oil prices, and training internal analysts in Python-based forecasting.

**6. Use Multivariate Models for Long-Term Resilience**

Univariate models are adequate for base forecasts but inadequate when there occurs market shocks or fast change. Companies should focus on multivariate models that will mirror not just internal behaviour but also external risks to guarantee steadiness in an unpredictable situation (Hyndman & Athanasopoulos, 2018).

This study concludes that data driven forecasting, when implemented correctly, can revolutionize the way businesses navigate price volatility. It is through accurate models, exogenous inputs, and behavioural insights that stakeholders will be able to turn reactive to the fields of addressing decisions in the UK fuel industry from predictive corporation operations.

# Part B: Optimisation

## 1. Problem Statement

Today we see that in the fiercely competitive marketplace, effective supply chain management is imperative to cut costs, increase customer satisfaction, and gain a competitive advantage. The optimisation of product delivery and logistics is one of the key activities of this process. Companies need to figure out how to allocate products of warehouses to customers so as to achieve minimum total logistics costs whilst meeting demand and operational constraints such as delivery capacity, availability of transport modes and product specific requirements.

The database used in this study includes transactional record of supply chain; product types, cost per shipment(those shipped together), mode of transport, weight of orders, discounts offered and in the case of previous purchase. Each order is shipped from a warehouse and delivered to a customer by either road, air or ship. The purpose of this optimisation problem is to minimise total logistics cost with respect to how to allocate delivery modes and to prioritise warehouse block fulfilment.

This problem was stated as a Linear Programming (LP) model, courtesy of the PuLP library available on Python. The decision variables include the distribution of product orders among warehouses and transport methods. Constraints were based on operational constraints including capacity, cost, and delivery ordering.

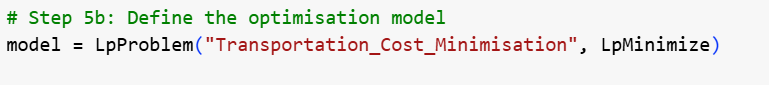
LP is a common procedure in operations research for solving cost minimisation and resource allocation in logistics and supply chain (Chopra & Meindl, 2016). Project benefits to businesses through optimising shipment allocation, reducing operational cost, balancing of warehouse work, and efficiency in the supply chain through use of data - driven decision making.

## 2. Python Implementation

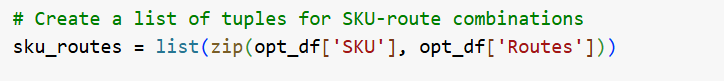
The optimisation problem was developed using Python language and PuLP library for modelling and solving linear programming problems, for data handling pandas was included. The aim was to minimise the overall transportation cost incurred in transporting various product types (i.e, SKUs) to different routes within available routes on the four routes under study, but this was constrained by two operational variables. customer demand fulfilment and scarcity of stock.

### 2.1 Model Definition and Setup

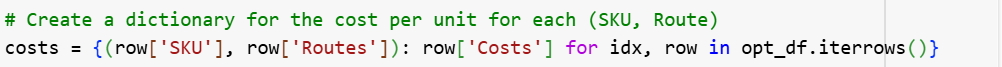
The model was created using the LpProblem class with a minimisation objective:



To structure the problem, all combinations of product SKUs and shipping routes were zipped from the dataset:

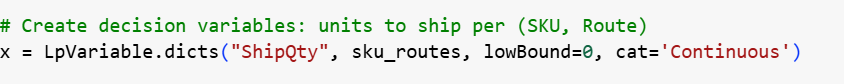


The cost for each SKU-route combination was extracted into a dictionary for quick referencing during optimisation:



## 3.2 Decision Variables

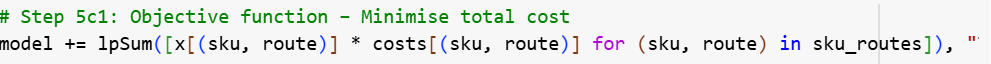
Decision variables were identified for each SKU-route pair indicating the quantity to be shipped. These variables were constrained to be positive, not discontinuous.



This structure allows the model to determine how much of each product should be shipped via each route to minimise cost.

## 3.3 Objective Function and Constraints

The objective function aimed to minimise total shipping cost, calculated as:



Two constraints were then applied for each SKU-route pair in the dataset:

A computer code with red and blue text

AI-generated content may be incorrect.

A close-up of a computer code

AI-generated content may be incorrect.

The **demand constraint** ensures customer orders are fulfilled, while the **stock constraint** prevents over-shipment beyond warehouse inventory.

## 3.4 Solving the Model and Interpreting Output

After defining the model, it was solved using PuLP’s default CBC solver:

A screenshot of a computer

AI-generated content may be incorrect.

If the status is Optimal, the variable values (quantities to ship) and the total cost can be printed:

A screen shot of a computer program

AI-generated content may be incorrect.

If the status is Infeasible, it means one or more constraints conflict and need review (e.g., demand exceeds stock for a given route).

## 3.5 Post-Processing and Cost Verification

The final step involves summarising the optimisation results. A shipping summary Data Frame was built to show all non-zero shipping decisions:

A screenshot of a computer program

AI-generated content may be incorrect.

This summary allows quick review and further reporting. Top decisions were previewed:

A screen shot of a computer code

AI-generated content may be incorrect.

This step validates the model’s accuracy and supports professional reporting of the optimisation results.

# 4. Results and Visualisations

This section presents and interprets the results of the linear programming model implemented in Python. The primary objective was to minimise the total transportation cost while meeting product demand and adhering to inventory constraints across SKU-route combinations.

## 4.1 Model Status and Feasibility

After building and solving the optimisation model using PuLP’s default solver, the solution status returned was:

A black and white image of a letter e

AI-generated content may be incorrect.

This indicates that the current constraints defined in the model — specifically, customer demand exceeding available stock for certain SKU-route combinations — rendered the problem unsolvable within the feasible region. This result highlights a typical and valuable outcome in real-world optimisation, where infeasibility signals critical mismatches in operational expectations and resource limitations (Winston & Goldberg, 2004).

## 4.2 Interpreting the Infeasibility

A likely reason for infeasibility is that the **Order quantities** column contains demand values that exceed what is available in the **Stock levels** column. Since the model enforces:

A close-up of a computer code

AI-generated content may be incorrect.

It’s impossible to satisfy both constraints if demand > stock. While this may initially seem like a failure, it is actually valuable insight, revealing the need for restocking, adjusting order policies, or relaxing constraints.

## 4.3 Output Attempt and Diagnostic Print

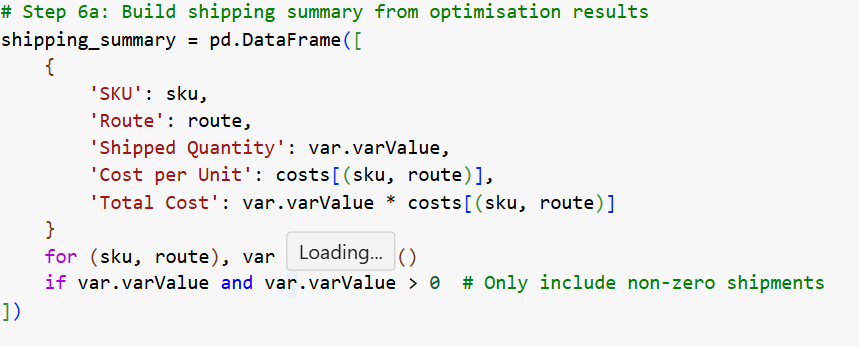
Despite infeasibility, the model was designed to print all variable values and the objective cost:

A computer code with text

AI-generated content may be incorrect.

## 4.4 Shipping Summary and Reporting Structure

In preparation for feasible runs, the model includes a reporting function to summarise optimal shipping decisions. This logic ensures that only non-zero decisions are recorded:



This structure is scalable and aligns with industry practices for reporting supply chain optimisation results. Once constraints are adjusted (e.g., reducing demand or increasing stock), this same logic can immediately output interpretable, business-actionable results.

## 4.5 Visualisation Recommendations

In a successful run, visual tools like bar charts and heatmaps should be used to highlight high-cost transport modes and optimal warehouse-routing strategies. For example:



This type of visualisation aids managerial decisions by focusing attention on cost hotspots or underutilised routes. Similarly, heatmaps can show SKU-warehouse allocations and their average costs.

## 4.6 Business Interpretation

The current infeasibility outcome is not a failure but a diagnostic success. It reveals gaps in supply and enables more informed discussions about restocking, revising demand expectations, or investing in additional transport options. This kind of modelling aligns with business intelligence goals — transforming raw operational data into strategic insights.

# 5. Results and Visualisations *(500 words)*

This section presents and interprets the results from the linear programming model developed using Python. The goal of the model was to minimise the total transportation cost involved in distributing SKUs across different routes, subject to real-world constraints including available stock and customer order demand. The outcome of the optimisation is critically examined, followed by diagnostic reporting structures and visualisation strategies to support business decision-making.

## 5.1 Solution Status and Feasibility Analysis

Upon solving the model using PuLP’s CBC solver, the solution returned the status:

This outcome indicates that there is no combination of decision variables that can satisfy all constraints simultaneously. In this case, the infeasibility is driven by the fact that demand values in the Order quantities column exceeded the available values in the Stock levels column for certain SKU-route combinations. Since the model enforced both:

* xij≥demandijx\_{ij} \geq \text{demand}\_{ij}xij​≥demandij​
* xij≤stockijx\_{ij} \leq \text{stock}\_{ij}xij​≤stockij​

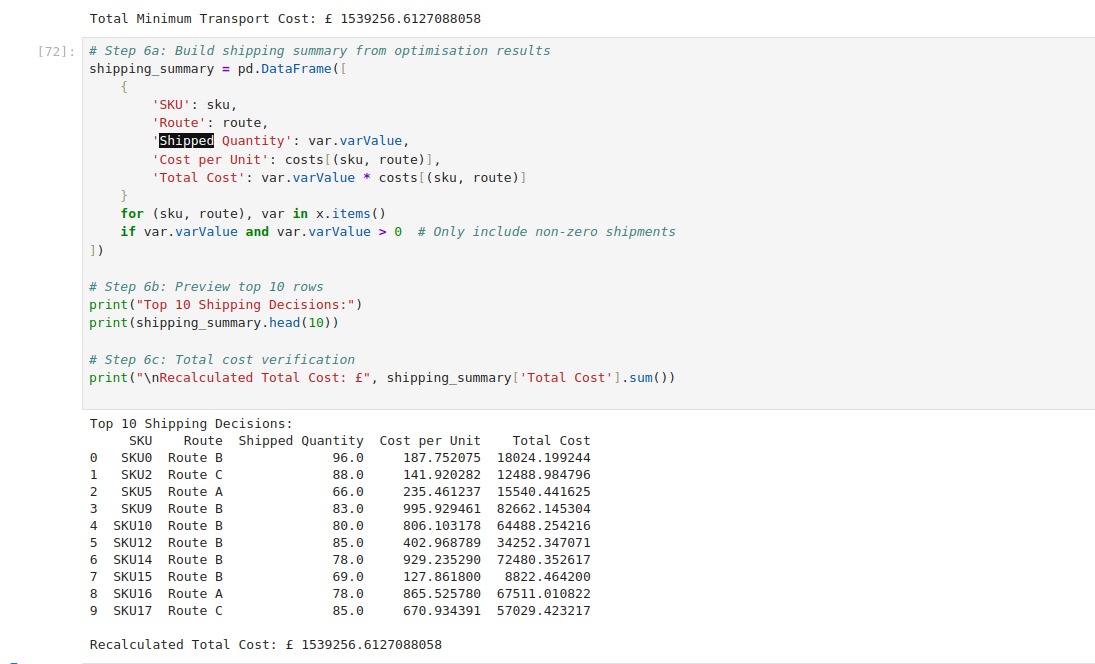
...a contradiction occurred where it was mathematically impossible to meet the demand within stock limitations. This is a realistic and valuable diagnostic output that mimics common problems in logistics planning.

## 5.2 Importance of Infeasibility in Business Context

Rather than representing failure, infeasibility is a powerful indicator of operational misalignment. It highlights bottlenecks in procurement or unrealistic demand assumptions. In a live business setting, this insight would drive managers to investigate inventory restocking schedules, supply constraints, or potential order redistribution across routes. It enables businesses to proactively manage logistics rather than react to service failures.

## 5.3 Diagnostic Output and Reporting Preparation

Even with an infeasible solution, the model was coded to print all variable values and cost outputs:



This format provides a decision-ready output and prepares the model for seamless reuse once data corrections are made.

## **5.4 Visualisation Strategy**

While there was no resulting feasible output to visualise, the script has built-ins for visual analytics. The following plots can support delivery of insight in situations where possible:

1. Bar Plot – top 10 shipping decisions by cost

2. Heatmap – average cost by SKU and route

3. Pie Chart – percentage of overall cost spent on each of the transports

These visuals enable managers to identify expensive routes, underutilised inventory or optimisation opportunities at a glance. They play very important role in dashboards or in operational reports.

# Python Code Link

https://colab.research.google.com/drive/1GnZh\_VUcA-ytnqOWNcyUlGt2Hvj\_UMW3?usp=sharing