STATISTICS APPLIED TO BUSINESS

Time-Series Forecasting and Demand Volatility in Supply Chain Logistics

Submitted by: Roshan

Student ID: PBM26037

Programme/Course: Bachelor of Business Administration and Management

Submission date: 11/10/2026

Table of Contents

Contents

Table of Contents 2

Executive Summary 3

Introduction 4

Exploratory Data Analysis & Methodology 4

Findings & Model Results 6

Discussion & Business Implications 8

Conclusion & Recommendations 9

References 10

Appendix A – Simulated Monthly Demand Dataset 11

Appendix B – Time-Series Decomposition and Seasonality 12

Appendix C – Forecast Model Comparison and Accuracy Metrics 14

Appendix D – 2026 Demand Forecast 15

Appendix E – Safety Stock and Reorder-Point Scenarios 16

Appendix F – Forecasting Roadmap and Model Documentation 17

- Implementation Roadmap 17
- Forecast Model Definitions 17

Appendix G – Complete Raw Monthly Demand Dataset 18

1. Executive Summary

This report introduces the topic of time-series forecasting and how it can enhance demand planning and inventory decisions in supply chain logistics. The main case is Walmart with supporting examples of Amazon and Inditex (Zara). The report introduces the concepts of trend, seasonality and stationarity, considers moving-average, seasonal naive and Holt-Winters and seasonal ARIMA modelling methods, compares the errors of the forecasts, and interprets forecast uncertainty into decisions for safety stock and reorder points. Using real-time AI and automation to forecast demand, detect stock outages and re-stock products from point to point across Walmart's network (Walmart, 2025). It uses a combination of historic sales, seasonality, local demand, weather and logistics restrictions to create predictive systems by 2025, the technology account reads (Walmart Global Tech, 2025).

The practical will employ simulated demand observations for 60 different months (Jan 2021 – Dec 2025) for a hypothetical medium-sized omnichannel retailer. The first 48 months are used for model estimation and the last 12 months are a holdout test. Four forecasting techniques are evaluated and compared with MAE, RMSE, MAPE and WAPE. The holdout WAPE values for Holt-Winters additive smoothing and seasonal ARIMA are very close at 2.62% and 2.66% respectively. A 12-month forecast for the selected model is created by refitting it. It is then used in a two-month lead-time safety-stock illustration at 90%, 95%, 98% and 99% service levels in the holdout forecast-error standard deviation.

The main suggestion is to not choose a forecasting model based solely on in-sample fit. A medium-sized operation should be a mix of clean data, SKU segmentation, rolling-origin back-testing, forecast-bias monitoring, and inventory policies which take into account demand and lead-time uncertainty. All numbers in the observations and models in the workbook are for illustration only and not to be linked to Walmart, Amazon or Inditex.

2. Introduction

Estimates of future demand drive decisions in the supply chain. Under forecasting may lead to stock outages, lost sales, emergency freight costs and poor customer experience. The over-forecasting can lead to a build-up of stock levels, a drop in selling prices, the pressure on working capital of the business and cost of storage. In omnichannel retail, where the same product is being requested both in-stores and on the internet as well as through marketplace channels and fulfilment centres, these costs are particularly significant. Forecasting is not only a statistical exercise but a managerial process which deals with sales, procurement, inventory, warehouse capacity, transport and finance.

A case study that has become relevant is from Walmart, as they continue in their efforts to invest in their technology-driven supply chain operations. “Walmart will have real-time AI and automation in its playbook, and some of the features include self-healing inventory where the system anticipates out-of-stocks and routes products away from stores before they are unavailable,” Walmart said in July 2025. Walmart Global Tech article says: "Predictive systems are used to place inventory at Walmart and make decisions about fulfilment, taking into account history, seasonality, local demand, weather and logistics constraints. The following are some corporate descriptions that illustrate how forecasts can be combined with execution systems. They don't share a full public SKU-by-SKU demand level, so this assignment is not a true representation of Walmart's internal forecast (Walmart, 2025).

3. Exploratory Data Analysis & Methodology

A time series for a business is a series of observations taking place at regular time periods. Characteristic features of its structure are trend, seasonality, cyclical movement and irregular variation. Trend is a steady and continuous rising or falling tendency over a long period of time. Seasonality refers to a recurring pattern that is associated with a given calendar period, such as increased retail sales in November and December. Cycles are longer variations linked to economic or industry conditions which may not occur at a regular time. Random variation includes all of the

unpredictable influences in the short term; such as weather, promotions, supplier disruption and measurement noise. According to NIST, observations over time may exhibit autocorrelation, and it is important to account for this or for trend and seasonal variation when selecting a model (National Institute of Standards and Technology [NIST], n.d.).

If the mean, variance and dependence structure of a series is reasonably stable over time, then the series is stationary. A stationary series is often necessary following appropriate transformations or differencing to identify many ARIMA identification procedures. If there is a visible trend or the variance or seasonal pattern is changing, then the stationarity assumption is likely to be violated. In this case, the diagnostic test we are interested in is the Augmented Dickey-Fuller (ADF) test, with a null hypothesis that a unit root exists. A large p-value does not prove that the null cannot be rejected, it is just to say that the null cannot be rejected. Depending on the type of instability, first differencing or seasonal differencing, log transformation or detrending might be considered.

The illustrative dataset is comprised of the 60 monthly observations January 2021 to December 2025 of units demanded by a hypothetical mid-sized distribution chain of an omnichannel retailer. Has a gradual increasing trend, a calendar-month seasonal profile and a lot of random noise. The dataset is split chronologically: 48 months of training data (January 2021 – December 2024); 12 months of holdout (January – December 2025). The estimation of the candidate models does not take into consideration any actual observation from 2025. This prevents look-ahead leakage and provides a more proper evaluation of forecast performance.

The following four models are compared: The mean of the most recent three observations is used to predict the next month in the first step, which is a simple moving average over 3 months. It is clear and filters out noise, but it is slow to respond to trends and it does not make an explicit model of annual seasonality. Secondly, the actual demand of the same month of the previous year is used in the seasonal-naive method. It can be a useful indicator in an annual seasonal pattern, but cannot show a

trend. Third, Holt-Winters additive exponential smoothing has an estimating component for level, trend and seasonal. It is appropriate if the units of the seasonal change are fairly consistent. Fourth, Seasonal ARIMA (SARIMA) incorporates a seasonal autoregressive term, a seasonal differencing term and a seasonal moving-average term in addition to the terms in ARIMA. The (p,d,q) indicate ARIMA terms and the (P,D,Q)s indicate seasonal terms. Attention should be given to the stationarity and seasonality as well as the behavior of the residuals when identifying a model (NIST, n.d.; Hyndman & Athanasopoulos, 2021).

The models are estimated using Python statistical libraries and the workbook contains the raw observations, model predictions, statistics, future projections and inventory calculations. The accuracy is assessed on the 12-month holdout. Units: MAE is the mean absolute error in units and the RMSE is the square root of the mean squared error, which penalizes large misses; the MAPE is the mean absolute percentage error; and the WAPE is the total absolute error divided by the total actual demand. MAPE can become unstable when values in actual calculations are zero or very small, and WAPE is beneficial for determining the error of the entire portfolio (Hyndman, n.d.; Amazon Web Services, n.d.). The selected model is the one with the smallest holdout WAPE, and MAE, bias and operational judgement are taken into consideration.

4. Findings & Model Results

The features of the design simulated series are indicated on the exploratory charts: an upward trend, a seasonal increase about the year-end peak and irregular month-to-month variation. The demand is greater during the later years than the former years. The p-value of the ADF test on the level series is 0.9215, which is not less than 0.05, indicating that the unit-root null cannot be rejected. The first difference of the series has a p-value less than 0.001, giving evidence that there is no unit root in the first difference. The results justify the need to focus on trend before modeling an ARIMA-family model. These should not be used as a mechanical model-selection rule but rather in conjunction

with the time plot and decomposition. While a first difference can eliminate a stochastic trend, it may not eliminate annual seasonality completely, and additional differencing of the way or explicit seasonal terms might be necessary.

The holdout comparison shows relative performance of the four methods. The seasonal-naive model is the closest to the benchmark WAPE, at 4.57%, with the three-month moving average at 9.85% and the Holt-Winters additive model at 2.62%. The SARIMA model has the best WAPE, which is 2.66%. An MAE of about 336 units and RMSE of about 386 units are also recorded by Holt-Winters. Its forecast bias (the average of the forecasted demand less the average of actual demand) is approximately +79 units, which would mean that forecasts were slightly above the actual demand in this holdout, on average. The small difference between Holt-Winters and SARIMA means that using a simpler exponential-smoothing model may be a suitable model for this data. The result does not imply that Holt-Winters is generally the best model; another SKU, horizon or market might benefit from SARIMA or a more sophisticated model.

The worksheet in the workbook displaying the safety stocks translates misjudgments in forecasting into a stock buffer. The example assumes that the lead time for replenishment is 2 months, and it calculates the standard deviation of the errors for the 12-month holdout period of the model selected. The following equation is used to determine safety stock: SS = Z × standard deviation of the forecast error × the square root of lead time. The reorder point is equal to the sum of the forecast demand during the lead time and safety stock. The illustrative safety stock is about 919 units with a service level of 95% and the reorder point is about 25,598 units. The higher the service-level targets the higher the Z and safety stock. This decreases the risk of running out of stock and increases holding costs and working capital. Where the errors in the demand are correlated in time or the lead time for the supplies varies, the operation should estimate the distribution of the cumulative lead time forecast errors directly, using the formula as a simplified approximation (Gonçalves et al., 2020).

5. Discussion & Business Implications

The corporate examples demonstrate that the value of demand forecasting comes to light when forecasts are linked to inventory execution. Walmart's self-healing inventory initiative is said to be discovering stock imbalances and rerouting inventory throughout the network. Later technology identified in the report is seeking to determine the integration of a decision process taking into account historical sales, seasonal patterns, local demand, weather, store conditions and driver capacity (Walmart, 2025; Walmart Global Tech, 2025). A manager's hint is that he should not regard a forecast as a number of fixed months. It should be updated if there are new information that affects demand and/or fulfilment constraints.

Amazon is another example of Scale and Model variety. Amazon Forecast is a machine-learning service for time-series forecasting that enables retailers to make inventory decisions at the granular level for their products, according to AWS. Amazon Science also released Amazon Web Services' (AWS) 2025 research paper on deep inventory management, which integrates demand learning with supply and capacity risk awareness (Liu et al., 2025). This implies that the next step after the univariate ARIMA is a multi-signal forecasting and optimization system.

The results report for FY25 show sales growth of 3.2% to €39.9 billion and inventories (31 January 2026) 2% lower than a year ago. The company also notes that the integration of sales and distribution channels in stores and online and continued investment in centralized sales and distribution channels (Inditex, 2026). It has reported every year about RFID and an Integrated Stock Management System, which enable staff to find products in stores and online warehouses (Inditex, 2024). This is not proof that Zara uses only a statistical time-series model, but rather an illustration of the fact that inventory records and customer feedback are just as useful in forecasting in a swiftly changing assortment.

There are three lessons to learn from the comparison. Walmart has figured out how to orchestrate inventory across the entire network; Amazon has figured out how to optimize inventory at the granular level; and Inditex has figured out how to synchronize stock at the item level and in an agile fashion. A mid-sized company can do this step-by-step: First fix data quality, then start segmenting products by value and variability, and finally, when there is a business case, move to machine learning.

6. Conclusion & Recommendations

When time series forecasting processes identify trend, seasonality, stationarity and random shocks, it becomes easier to make a proper inventory decision. Holt-Winters model: Represent level, trend, seasonality, explain easily but can be laggy, provide good benchmark for seasonal naive model; SARIMA may take into account more complex autocorrelation / seasonal structure; etc. Chronological holdout and rolling-origin validation should be used to decide on the best model, rather than in-sample fit. A variety of accuracy assessment metrics should be used since MAE, RMSE, MAPE and WAPE place emphasis on different aspects of errors.

For a medium-sized supply chain, a staged forecasting roadmap should be deployed: creating an integrated demand dataset; SKUs classified according to the ABC/XYZ forecasting strategy and then testing several models on each product segment; incorporating the uncertainty in the forecast into the service level and maintaining the safety stocks for that level; binding the forecasts for inventory replenishment; and monitoring forecast bias, WAPE, the risk of a stock out and excess inventory using monthly Sales and Operations Planning. Forecast over-rides should be documented and reviewed. Retraining a model should occur when there is a change in material demands, promotions, new materials and/or a significant drop in model accuracy caused by a supplier's change or disruption or a sustained drop in model accuracy. Central concept: Forecasts are a means to decisions, never mandates.

References

Amazon Web Services. (n.d.). *Amazon Forecast*. https://aws.amazon.com/forecast/

Gonçalves, J. N. C., Carvalho, M. S., & Cortez, P. (2020). Operations research models and methods for safety stock determination: A review. *Operations Research Perspectives, 7*, 100164. https://doi.org/10.1016/j.orp.2020.100164

Hyndman, R. J. (n.d.). *Measuring forecast accuracy*. https://www.robjhyndman.com/papers/forecast-accuracy.pdf

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

Inditex. (2024). *Inditex Group annual report 2023*. https://static.inditex.com/annual_report_2023/en/Inditex_Group_Annual_Accounts_2023.pdf

Inditex. (2026). *FY2025 results*. https://www.inditex.com/itxcomweb/se/en/press/news-detail/b870d5ec-6b7e-491d-b38e-340cd69036df/fy2025-results

Liu, F., Liu, Y. (Bellevue), & Eisenach, C. (2025). *Deep inventory management with supply and capacity risk awareness*. Amazon Science. https://www.amazon.science/publications/deep-inventory-management-with-supply-and-capacity-risk-awareness

National Institute of Standards and Technology. (n.d.). *Introduction to time series analysis*. NIST/SEMATECH e-Handbook of Statistical Methods. https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm

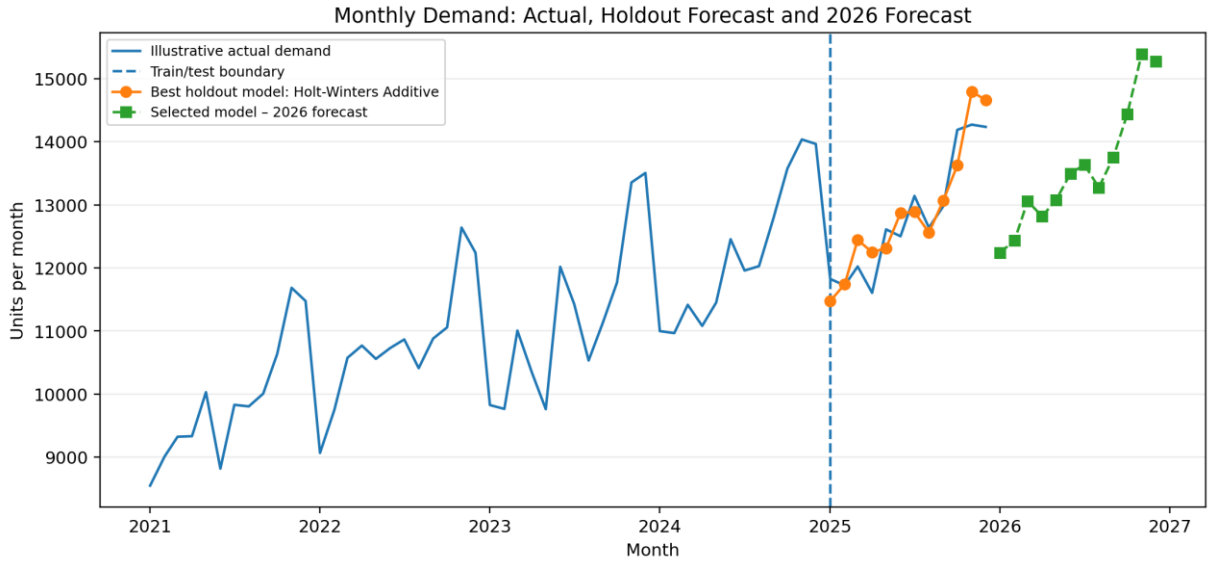
Walmart. (2025, July 17). *Walmart's U.S. supply chain playbook goes global—and it's reinventing retail at scale*. https://corporate.walmart.com/news/2025/07/17/walmarts-us-supply-chain-playbook-goes-global-and-its-reinventing-retail-at-scale

Walmart Global Tech. (2025, December 8). *Inside the AI network orchestrating Walmart's fastest holiday deliveries yet*. https://tech.walmart.com/content/walmart-global-tech/en_us/blog/post/inside-the-ai-network-orchestrating-walmarts-fastest-holiday-deliveries-yet.html

Walmart Inc. (2026). *2026 annual report and proxy statement*. https://corporate.walmart.com/news/2026/04/23/walmart-releases-2026-annual-report-and-proxy-statement

A. Appendix A – Simulated Monthly Demand Dataset

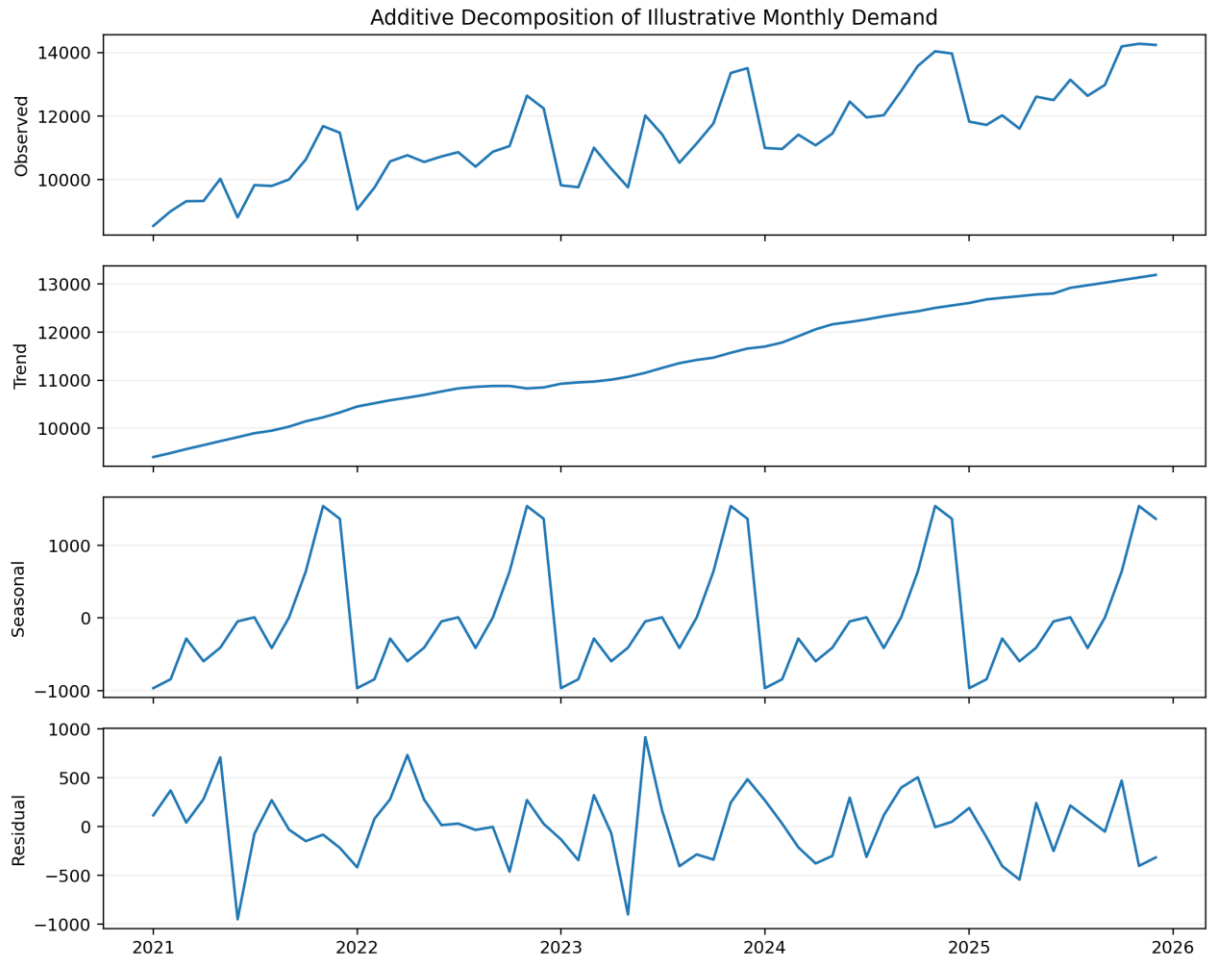
The entire 60 months of row-level data is provided. The chart shows the historical series, holdout forecast and selected 2026 forecast. The observations are emulated.



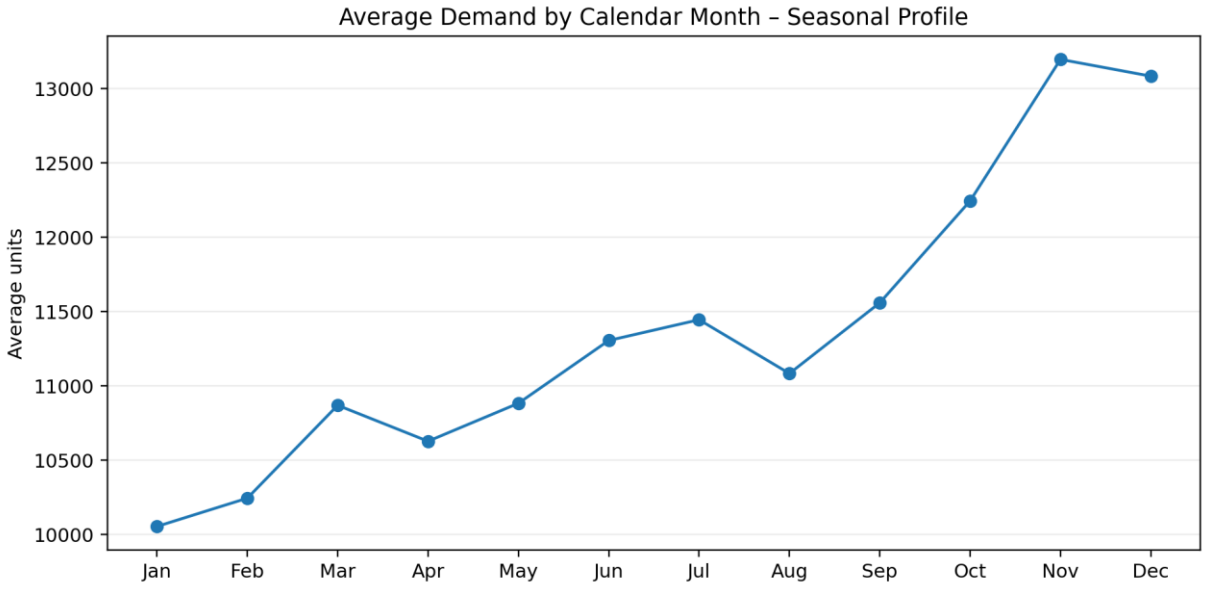
Source: Statistical model and excel workbook were prepared for this assignment. Data that have been synthetically created are marked as such.

B. Appendix B – Time-Series Decomposition and Seasonality

The decomposition is used to split the observed demand into the three components, namely the trend, the seasonal and the residual component. Using the seasonal profile, the average demand is shown by calendar month.



Source: A statistical model and worksheet (Excel) are made-up for present assignment. Synthetic data are clearly labelled.



Source: A statistical model and worksheet (Excel) are made-up for present assignment. Synthetic data are clearly labelled.

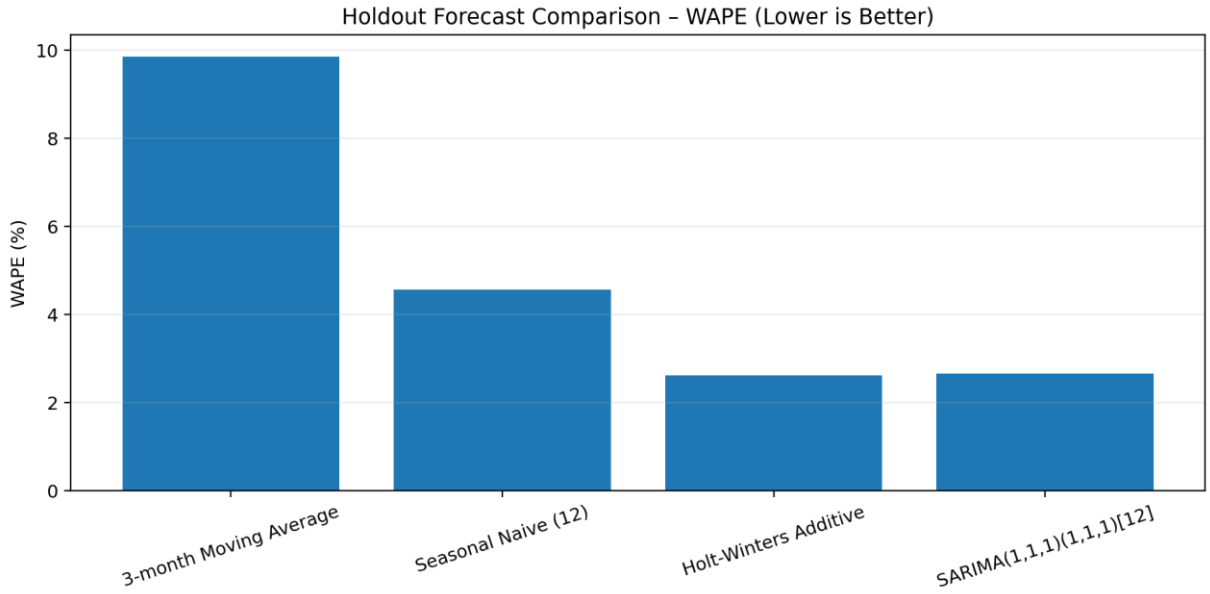
C. Appendix C – Forecast Model Comparison and Accuracy Metrics

The 2025 hold out period is not used in estimation of the models. The lower the error value the more accurate the results will be and these should be read in conjunction with each other.

Excel/Python Output – Forecast Accuracy on 2025 Holdout

Model	MAE (units)	RMSE (units)	MAPE (%)	WAPE (%)	Bias (units)
3-month Moving Average	1,262	1,451	10.30	9.85	1,108
Seasonal Naive (12)	585	680	4.65	4.57	-585
Holt-Winters Additive	336	386	2.61	2.62	79
SARIMA(1,1,1)(1,1,1)[12]	341	394	2.64	2.66	118

Source: This is a table designed with statistics and an excel workbook used for this assignment. Clearly mark and label synthetic data.



Source: Statistical model and Excel workbook prepared for this assignment. Synthetic data are clearly labelled.

D. Appendix D – 2026 Demand Forecast

To compile a 2026 forecast, the selected model is refitted to the entire 60-month run prior to the forecast.

Selected Model – 2026 Monthly Demand Forecast

Month	Forecast units
Jan 2026	12,244
Feb 2026	12,435
Mar 2026	13,060
Apr 2026	12,818
May 2026	13,073
Jun 2026	13,496
Jul 2026	13,636
Aug 2026	13,274
Sep 2026	13,750
Oct 2026	14,435
Nov 2026	15,388
Dec 2026	15,275

Source: Statistical model and excel model prepared on the basis of this assignment. Synthetic data are clearly labelled.

E. Appendix E – Safety Stock and Reorder-Point Scenarios

In some cases, the desired service level will dictate the amount required for safety stock. The calculation is based on two-month venture lead time and autonomy of the error in each forecast month.

Safety Stock Scenarios – Two-Month Lead Time

Service level	Z factor	Safety stock (units)	2-month forecast demand	Reorder point
90%	1.2816	716	24,679	25,395
95%	1.6449	919	24,679	25,598
98%	2.0537	1,147	24,679	25,826
99%	2.3263	1,299	24,679	25,979

Source: Created statistical model and workbook in Excel. Clearly label data collected by means of a synthetic model.

F. Appendix F – Forecasting Roadmap and Model Documentation

I. Implementation Roadmap

Phase	Timing	Deliverable
1. Data foundation	Month 1-2	At least 98% valid SKU-location-date records; documented data dictionary.
2. Baseline and segmentation	Month 2-3	SKU segmentation and baseline forecast report.
3. Model pilot	Month 3-5	WAPE/MAE by SKU segment; bias and forecast-value-added.
4. Inventory policy	Month 5-6	Fill rate, stockout rate, excess stock, inventory turns.
5. Controlled deployment	Month 6-8	Planner adoption; override rate; forecast-to-order traceability.
6. Monitor and retrain	Month 9-12	Monthly dashboard and model change log.

II. Forecast Model Definitions

Method	Formula / definition	Limitation
Simple moving average	F(t+1) = mean of the most recent k observations	Smooths short-term noise; lags behind trend and cannot represent seasonality unless adapted.
Seasonal naive	Forecast for month m = actual from the same month last year	Strong baseline for stable annual seasonality; does not capture trend or structural change.
Holt-Winters additive	Level + trend + additive seasonal component	Useful when seasonal swings are roughly constant in units; multiplicative seasonality may be preferable when amplitude grows with level.
ARIMA(p,d,q)	AR: lagged observations; I: differencing; MA: lagged forecast errors	Requires attention to stationarity, differencing, autocorrelation and residual diagnostics.
SARIMA	ARIMA with seasonal terms (P,D,Q)s	Extends ARIMA to repeating seasonal patterns; order selection should be justified and validated.
MAE	Mean(|actual − forecast|)	Easy to interpret in units; treats each absolute error linearly.
RMSE	sqrt(mean((actual − forecast)^2))	Penalises large misses more strongly; useful when large errors are costly.
MAPE	Mean(|actual − forecast| / actual) × 100	Undefined at zero and unstable for very small actual values.
WAPE	Sum(|actual − forecast|) / Sum(actual) × 100	Weighted overall absolute error; useful across a portfolio with positive demand.
Safety stock	SS = Z × σ(error) × sqrt(L)	Simplified independent-period model; use lead-time forecast-error quantiles or full lead-time error distribution if errors are autocorrelated.

G. Appendix G – Complete Raw Monthly Demand Dataset

All 60 monthly observations in the model exercise are shown in the following table. The data is synthetic representing the hypothetical sales of a medium sized omnichannel retailer, not Walmart, Amazon or Inditex sales data.

Appendix – Complete 60-Month Simulated Demand Dataset

Month	Demand (units)	Month	Demand (units)
Jan 2021	8,550	Jul 2023	11,428
Feb 2021	9,010	Aug 2023	10,533
Mar 2021	9,325	Sep 2023	11,141
Apr 2021	9,334	Oct 2023	11,769
May 2021	10,031	Nov 2023	13,354
Jun 2021	8,820	Dec 2023	13,505
Jul 2021	9,831	Jan 2024	10,999
Aug 2021	9,806	Feb 2024	10,967
Sep 2021	10,008	Mar 2024	11,416
Oct 2021	10,636	Apr 2024	11,082
Nov 2021	11,685	May 2024	11,451
Dec 2021	11,476	Jun 2024	12,455
Jan 2022	9,068	Jul 2024	11,959
Feb 2022	9,755	Aug 2024	12,028
Mar 2022	10,576	Sep 2024	12,787
Apr 2022	10,770	Oct 2024	13,574
May 2022	10,559	Nov 2024	14,036
Jun 2022	10,732	Dec 2024	13,965
Jul 2022	10,867	Jan 2025	11,826
Aug 2022	10,412	Feb 2025	11,723
Sep 2022	10,881	Mar 2025	12,023
Oct 2022	11,058	Apr 2025	11,605
Nov 2022	12,640	May 2025	12,611
Dec 2022	12,241	Jun 2025	12,502
Jan 2023	9,827	Jul 2025	13,141
Feb 2023	9,766	Aug 2025	12,638
Mar 2023	11,007	Sep 2025	12,979
Apr 2023	10,346	Oct 2025	14,188
May 2023	9,762	Nov 2025	14,271
Jun 2023	12,018	Dec 2025	14,235

Source: This is an academic dataset that has been created for the purpose of the assignment.