Written by : Phanuwich Thamjunjuea

Data-Driven Demand Forecasting & Inventory Optimization to Reduce Lost Sales and Holding Cost

Introduction

In today's retail supply chain, effective demand planning and inventory management are critical to maintaining high service levels while controlling operational costs. However, many organizations continue to rely on simplified forecasting methods and uniform inventory policies, which often fail to account for demand variability across products and time periods. As a result, businesses may experience inventory imbalance, characterized by frequent stock-outs for high-demand items and excessive inventory for slow-moving products.

This project utilizes the Superstore Dataset, a widely used retail transaction dataset containing historical sales records, product categories, order dates, and shipment information. By analyzing this dataset, the project aims to simulate a realistic retail supply chain environment and evaluate how demand patterns and inventory decisions affect service performance and financial outcomes.

Exploratory data analysis reveals significant demand variability across product categories and time periods. A small proportion of products contributes to a large share of total demand, while many items exhibit low and intermittent sales. This imbalance suggests that applying a uniform inventory policy across all products may lead to inefficiencies, including lost sales due to stock-outs for high-demand products and unnecessary holding costs for low-demand items.

Further analysis of historical demand trends indicates that simple baseline forecasting methods are insufficient to capture seasonality and demand fluctuations present in the data. These forecasting limitations directly translate into poor inventory decisions, resulting in suboptimal service levels and increased inventory-related costs. The data highlights a clear pain point: inaccurate demand forecasting leads to misaligned inventory levels, negatively impacting both customer satisfaction and working capital utilization.

Based on these observations, this project focuses on applying data-driven demand forecasting and inventory optimization techniques to address inventory imbalance in the retail supply chain. By comparing baseline and improved planning scenarios, the analysis aims to quantify the impact of data-driven decision-making in terms of service level improvement, lost sales reduction, and inventory cost savings.

Ultimately, this project demonstrates how supply chain data analysis can move beyond descriptive reporting to support actionable decisions that improve operational efficiency and deliver measurable business value.

Assumptions

Due to the limitations of publicly available retail transaction data, several assumptions are required to simulate a realistic retail supply chain environment. These assumptions are designed to maintain analytical rigor while reflecting standard practices in supply chain and operations management.

Customer demand is represented by the quantity of products ordered in the Superstore Dataset. All recorded transactions are assumed to reflect actual market demand, with no demand censoring or substitution effects. Since inventory levels are not explicitly available in the dataset, inventory positions are simulated based on demand and predefined replenishment policies.

Inventory is reviewed on a monthly basis, and replenishment decisions are made at the end of each review period. A constant lead time of seven days is assumed for all products, and lead time variability is not considered in order to isolate the impact of demand uncertainty on inventory performance.

Unit cost is estimated as 60% of the selling price, reflecting a simplified and consistent cost structure across products. Inventory holding cost is assumed to be 20% of inventory value per year, incorporating storage, capital, insurance, and obsolescence costs. These values are consistent with commonly used industry benchmarks in inventory management literature.

A stock-out is defined as a situation in which customer demand exceeds available inventory within a review period. Any unfulfilled demand is treated as lost sales and is not backordered. Lost sales are valued at the selling price of the unfulfilled quantity, providing a conservative estimate of revenue impact. Service level is defined as the percentage of customer demand fulfilled immediately without stock-out, with partial fulfillment treated as a stock-out event.

Products are analyzed at the sub-category level to balance demand stability and analytical granularity. Demand is forecasted one period ahead on a monthly basis and is used exclusively to support inventory planning decisions. The same inventory policy structure is applied across both baseline and improved scenarios, ensuring that observed performance differences are attributable to changes in forecasting accuracy and inventory parameter tuning rather than structural policy changes.

The scope of this analysis is limited to demand planning and inventory management. Transportation costs, supplier capacity constraints, and production limitations are excluded from the model. While these assumptions simplify real-world complexity, they provide a transparent and logically consistent framework for evaluating how data-driven demand forecasting can improve service performance and reduce inventory-related costs.

Methodology

This project follows a structured, data-driven methodology to evaluate the impact of demand forecasting accuracy on inventory performance in a retail supply chain. The analysis compares a baseline planning scenario with an improved, data-driven scenario while maintaining a consistent inventory policy framework to ensure fair performance comparison.

Data Preparation and Demand Aggregation

Historical transaction data from the Superstore Dataset were cleaned and aggregated to represent customer demand. Demand was measured as the quantity of products ordered and aggregated on a monthly basis at the sub-category level. This level of aggregation balances analytical tractability with demand stability, reducing short-term noise while preserving meaningful demand patterns relevant to inventory planning decisions.

Outliers and missing values were reviewed to ensure data consistency. All demand observations were assumed to be independent across periods, and no demand substitution or demand censoring effects were considered.

Forecasting Models

This study evaluates two demand forecasting approaches to assess how forecasting accuracy influences inventory performance. Forecasts are generated on a one-period-ahead basis at the monthly level and aggregated by sub-category. The forecasting framework is designed to ensure transparency, reproducibility, and a direct link between forecast values and subsequent inventory decisions.

Baseline Forecasting Model: 3-Year Moving Average

The baseline forecasting model applies a three-year moving average method, representing a long-horizon smoothing approach commonly used to stabilize demand estimates. Forecasted demand for period $t + 1$is calculated as the average of actual demand observed over the previous 36 months.

$$\widehat{D}_{t+1}^{\mathrm{MA}(36)} = \frac{1}{36} \sum_{i=t-35}^{t} D_i$$

where:

- $D_i$denotes actual monthly demand in period i
- The window length is fixed at 36 months.

This method assigns equal weight to all historical observations within the three-year window and implicitly assumes long-term demand stability. While the approach effectively smooths short-term volatility, it reacts slowly to recent demand shifts, seasonal effects, or structural changes in demand patterns. Consequently, the three-year moving average serves as a conservative baseline benchmark for comparison with more adaptive forecasting techniques.

Improved Forecasting Model: Simple Exponential Smoothing

The improved forecasting approach employs Simple Exponential Smoothing (SES) to better capture recent demand dynamics by assigning greater weight to the most recent observations. Unlike the moving average method, SES allows the forecast to respond more quickly to changes in demand levels.

The one-period-ahead forecast for period $t + 1$is calculated as:

$$\widehat{D}_{t+1}^{\text{SES}} = \alpha D_t + (1 - \alpha)\widehat{D}_t$$

where:

- $\widehat{D}_t$is the forecast for period t
- $D_t$is the actual demand observed in period t
- $\alpha$is the smoothing parameter, with values between 0 and 1.

The initial forecast $\widehat{D}_1$is set equal to the first observed demand value to ensure numerical stability. The smoothing parameter $\alpha$is selected through iterative testing to minimize forecast error across sub-categories.

Forecast Accuracy Evaluation

Forecast accuracy is evaluated using Mean Absolute Percentage Error (MAPE), calculated as:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} | \frac{D_t - \widehat{D}_t}{D_t} | \times 100$$

Both forecasting models are evaluated using identical demand data, forecast horizons, and aggregation levels. Performance comparisons are conducted at the sub-category level to ensure consistency and fairness across scenarios.

Integration with Inventory Planning

Forecast outputs from both models are used directly as inputs to inventory planning under an identical periodic review policy. The one-period-ahead demand forecast represents expected demand during the next review period, while demand variability estimated from historical forecast errors is incorporated into safety stock calculations.

By holding the inventory policy structure constant across both forecasting scenarios, this study isolates the impact of forecasting accuracy on inventory outcomes. Observed differences in service level, lost sales, and inventory value can therefore be attributed to changes in forecasting responsiveness rather than structural policy modifications.

Inventory Policy Framework

A periodic review inventory policy was applied consistently across both scenarios. Inventory levels were reviewed monthly, and replenishment quantities were determined at the end of each review period based on forecasted demand and safety stock requirements.

Safety stock levels were calculated using demand variability observed in historical data, while lead time was assumed to be constant across all products. This policy structure ensures that observed performance differences between scenarios are attributable to changes in forecasting accuracy and inventory parameter tuning rather than structural changes in inventory policy design.

Simulation of Inventory Performance

Inventory dynamics were simulated over the analysis horizon to capture the interaction between demand realization, replenishment decisions, and inventory availability. At each review period, forecasted demand and safety stock levels were used to determine order quantities. Actual demand was then realized, and inventory positions were updated accordingly.

Stock-out events occurred when realized demand exceeded available inventory within a review period. Any unmet demand was classified as lost sales and was not backordered. Inventory value was calculated based on simulated inventory positions and estimated unit costs.

Performance Evaluation Metrics

Inventory performance was evaluated using multiple complementary metrics to capture service performance, revenue impact, and inventory efficiency. Service level was measured as the proportion of customer demand fulfilled without stock-out. Lost sales were quantified as the revenue associated with unmet demand.

In addition to service-based metrics, inventory efficiency indicators were calculated to evaluate capital utilization. These include inventory value, inventory turnover, and inventory investment relative to sales. By incorporating efficiency-focused KPIs, the analysis enables a balanced evaluation of trade-offs between customer service and inventory cost.

Comparative Analysis Approach

Performance metrics were compared between the baseline and improved scenarios at the sub-category level. This comparative approach isolates the impact of improved forecasting accuracy on inventory outcomes while holding policy structure constant. Differences in performance were analyzed to identify demand patterns and product categories where data-driven forecasting delivers the greatest operational and financial benefits.

Performance Metrics and Formulations

To ensure transparency and reproducibility, inventory performance is evaluated using clearly defined service, revenue, and efficiency metrics. All metrics are calculated consistently across baseline and improved scenarios.

1. Lost Sales

Lost sales represent unmet customer demand due to insufficient inventory within a review period. Any demand exceeding available inventory is treated as lost and is not backordered.

Formula

$$\text{Lost Sales}_t = \max\left(D_t - I_t, 0\right)$$

where:

- D_t= actual demand in period t
- I_t= available inventory at the beginning of period t

Lost Sales Value

$$\text{Lost Sales Value}_t = \text{Lost Sales}_t \times P$$

where:

- P= selling price per unit

Interpretation

Higher lost sales indicate reduced service performance but may result from intentional inventory reduction aimed at improving capital efficiency.

2. Service Level

Service level measures the proportion of customer demand fulfilled immediately without stock-out.

Formula

$$\text{Service Level} = \frac{\sum_t \min(D_t, I_t)}{\sum_t D_t}$$

where:

- $\min(D_t, I_t)$represents fulfilled demand in period t

Interpretation

A lower service level reflects higher exposure to stock-outs, often associated with lower inventory investment.

3. Inventory Value

Inventory value measures the monetary value of inventory held and represents capital tied up in stock.

Formula

$$\text{Inventory Value}_t = I_t \times C$$

where:

- I_t= inventory on hand in period t
- C= unit cost

Average Inventory Value

$$\text{Average Inventory Value} = \frac{1}{n}\sum_{t=1}^{n} \text{Inventory Value}_t$$

Interpretation

Lower inventory value indicates reduced working capital requirements and holding cost exposure.

4. Inventory Turnover

Inventory turnover evaluates how efficiently inventory is converted into sales.

Formula

$$\text{Inventory Turnover} = \frac{\text{Total Sales}}{\text{Average Inventory Value}}$$

Interpretation

A higher turnover rate indicates faster inventory movement and improved inventory utilization.

5. Inventory Value per Sales

This metric measures inventory investment relative to sales performance.

Formula

$$\text{Inventory per Sales} = \frac{\text{Average Inventory Value}}{\text{Total Sales}}$$

Interpretation

Lower values indicate improved capital efficiency, requiring less inventory to support a given sales level.

6. Capital Efficiency Index

This composite metric integrates service and cost perspectives into a single efficiency indicator.

Formula

$$\text{Capital Efficiency} = \frac{\text{Total Sales} - \text{Lost Sales Value}}{\text{Average Inventory Value}}$$

Interpretation

Higher values indicate that the system generates more realized revenue per unit of inventory investment.

Discussion

The results show that switching from a three-year moving average to exponential smoothing does not automatically improve inventory performance. While the new model is more responsive to recent demand changes, service levels decreased and lost sales increased across most sub-categories when inventory parameters remained unchanged.

Efficiency-focused KPIs also declined in the After scenario. Inventory turnover and capital efficiency decreased, indicating slower inventory movement and weaker utilization of working capital. This suggests that improved forecasting accuracy alone is insufficient to deliver performance gains without aligning inventory policy parameters, particularly safety stock levels.

Overall, the findings highlight that inventory performance is driven by the interaction between forecasting methods and inventory policies rather than forecasting models in isolation.

Managerial Implications

First, changes in forecasting models should be accompanied by recalibration of inventory parameters. Adopting more responsive forecasts without adjusting safety stock exposes the system to higher stock-out risk.

Second, inventory policies should be differentiated by product category, as demand variability and service requirements differ significantly across sub-categories. A uniform policy leads to suboptimal outcomes.

Finally, managers should evaluate inventory decisions using both service-level and efficiency-based KPIs. A balanced view enables better trade-off decisions between customer service and capital efficiency.

Sub-Category	Service Level	Lost Sales ($)	Inventory Value
Accessories	90%	13,629.65	91,328.01
Appliances	89%	9,888.73	40,985.36
Art	89%	1,756.70	9,674.18
Binders	90%	14,879.71	78,770.51
Bookcases	97%	7,543.96	68,928.00
Chairs	91%	19,470.88	262,564.11
Envelopes	90%	1,000.22	14,828.76
Fasteners	90%	1,445.96	7,654.67
Furmishings	90%	6,949.31	44,588.84
Labels	91%	758.10	8,795.71
Paper	91%	5,102.61	35,365.66
Phones	90%	23,579.10	157,367.91
Storage	91%	14,766.97	108,519.49
Tables	88%	18,067.10	129,782.90

Total sales	2,297,200.86
INV TURNOVER	30.36461982
Inv Value per sale	0.032933065
Capital Efficiency Index	28.52943281

Sub-Category	Service Level	Lost Sales ($)	Inventory Value
Accessories	88%	15,933.75	96,838.74
Appliances	86%	13,893.01	37,207.45
Art	87%	2,526.04	9,153.40
Binders	87%	20,946.90	77,132.20
Bookcases	89%	7,874.91	64,041.63
Chairs	88%	28,456.70	252,110.31
Envelopes	88%	1,280.83	17,614.64
Fasteners	98%	276.57	2,161.66
Furmishings	87%	10,132.45	41,091.35
Labels	88%	22,550.99	124,070.69
Paper	88%	22,550.99	124,070.69
Phones	88%	31,597.11	121,912.40
Storage	88%	21,106.47	106,355.99
Tables	86%	22,550.99	124,070.69

Total sales	2,297,200.86
INV TURNOVER	26.8491883
Inv Value per sale	0.037245074
Capital Efficiency Index	24.25826699

Picture indicates Before vs After Forecasting Scenario

Conclusion

This project demonstrates that improving demand forecasting alone does not guarantee better inventory performance. While exponential smoothing provides more responsive demand estimates than a long-horizon moving average, its benefits are not fully realized when inventory policies remain unchanged.

The comparative analysis highlights a fundamental trade-off between service level and inventory efficiency. Changes in forecasting methods must be accompanied by appropriate adjustments in inventory parameters, particularly safety stock, to achieve balanced performance outcomes.

Overall, this study shows how data-driven supply chain analysis can support informed decision-making by revealing system-level interactions and trade-offs. Rather than focusing on a single performance metric, effective inventory management requires an integrated approach that aligns forecasting models, inventory policies, and business objectives.