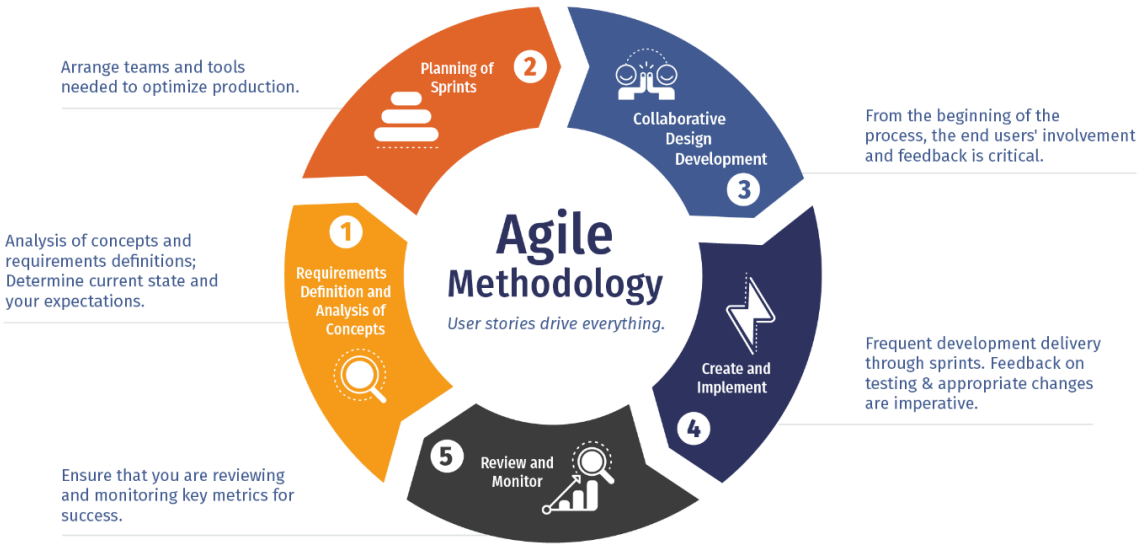
Real-Time Inventory Optimization at Walmart Using Agile Framework

Project Background:

Walmart, the world's largest retailer, is facing operational challenges in managing inventory efficiently across its thousands of stores worldwide. With fluctuating customer demand, promotions, and external factors like weather impacting sales, Walmart's operations team is under pressure to ensure optimal stock levels without overspending on excess inventory or losing sales due to stockouts.

To maintain a competitive edge and meet growing customer demands, Walmart aims to implement **real-time demand forecasting** by analyzing historical sales data. This process will involve understanding the influence of customer purchasing behavior, external factors like weather, and the effectiveness of promotions. The analysis will guide Walmart's inventory decisions, enabling them to streamline operations, reduce costs, and enhance customer satisfaction.

This project follows an **Agile methodology**, broken down into **sprints** for continuous value delivery and adaptation to insights throughout the process. As data analysts, you will be part of the team tasked with analyzing this dataset, discovering insights, and delivering actionable recommendations.



More Information: https://asana.com/resources/agile-methodology

Dataset:

DA01 Retail_Inventory_Optimization_Data.csv

Data Dictionary with Potential Analysis:

1. **transaction_id**:
    - **Description**: Unique identifier for each transaction.
    - **Data Type**: Integer
    - **Example**: 1, 2, 3
    - **Analysis**: This column can be used to count the total number of transactions and track individual transaction patterns or anomalies.
2. **customer_id**:
    - **Description**: Unique identifier for each customer.
    - **Data Type**: Integer
    - **Example**: 2824, 1409, 5506
    - **Analysis**: Can be used for customer segmentation and loyalty analysis. This helps track purchase behavior and identify frequent or high-value customers.
3. **product_id**:
    - **Description**: Unique identifier for each product.
    - **Data Type**: Integer
    - **Example**: 843, 135, 391
    - **Analysis**: Allows for product-level analysis, including sales performance, product popularity, and stockout analysis at the product level.
4. **product_name**:
    - **Description**: Name of the product.
    - **Data Type**: String
    - **Example**: "Fridge", "TV", "Smartphone"
    - **Analysis**: Can be used for product category performance analysis and comparisons across product lines.
5. **category**:
    - **Description**: The product category (e.g., Electronics, Furniture).
    - **Data Type**: String
    - **Example**: "Electronics", "Appliances"
    - **Analysis**: Enables category-level sales analysis and can be used to determine which categories are performing well and which need more attention.
6. **quantity_sold**:
    - **Description**: Number of units sold in each transaction.
    - **Data Type**: Integer
    - **Example**: 3, 4, 5
    - **Analysis**: Essential for volume analysis, demand forecasting, and identifying high-demand products.
7. **unit_price**:
    - **Description**: Price per unit of the product.
    - **Data Type**: Float
    - **Example**: 188.46, 1912.04

    - **Analysis**: Can be used for pricing strategy analysis, identifying the revenue impact of price changes, and calculating total revenue per product.
8. **transaction_date**:
    - **Description**: Date and time of the transaction.
    - **Data Type**: DateTime
    - **Example**: "2024-03-31 21:46:34"
    - **Analysis**: Allows for time series analysis to track sales trends over time, identify peak sales periods, and seasonal patterns.
9. **store_id**:
    - **Description**: Unique identifier for the store where the transaction occurred.
    - **Data Type**: Integer
    - **Example**: 3, 5, 1
    - **Analysis**: Store-level performance analysis can help compare sales across locations, identify high- and low-performing stores, and optimize inventory distribution.
10. **store_location**:
    - **Description**: Geographical location of the store.
    - **Data Type**: String
    - **Example**: "Miami, FL", "Dallas, TX"
    - **Analysis**: Useful for geographical analysis of sales patterns, comparing urban vs. rural performance, and identifying region-specific product demand.
11. **customer_loyalty_level**:
    - **Description**: Loyalty tier of the customer (e.g., Silver, Gold).
    - **Data Type**: String
    - **Example**: "Silver", "Gold", "Platinum"
    - **Analysis**: Helps in customer segmentation analysis, understanding customer lifetime value, and designing targeted loyalty programs.
12. **payment_method**:
    - **Description**: Payment type used for the transaction (e.g., Credit Card, Cash).
    - **Data Type**: String
    - **Example**: "Credit Card", "Cash"
    - **Analysis**: Payment method preference analysis to understand customer preferences and payment trends. Can also highlight potential barriers to transactions (e.g., lack of cashless options).
13. **promotion_applied**:
    - **Description**: Boolean flag indicating whether a promotion was applied.
    - **Data Type**: Boolean
    - **Example**: True, False
    - **Analysis**: Promotion effectiveness analysis to understand how promotions impact sales and demand.
14. **promotion_type**:
    - **Description**: Type of promotion applied (e.g., Percentage Discount).
    - **Data Type**: String
    - **Example**: "Percentage Discount", "None"
    - **Analysis**: Helps analyze which types of promotions are more effective and how they drive sales across different product categories.
15. **weather_conditions**:
    - **Description**: Weather conditions during the transaction (e.g., Sunny, Rainy).

    - **Data Type**: String
    - **Example**: "Stormy", "Sunny"
    - **Analysis**: Correlation analysis between weather conditions and sales, helping to predict demand based on weather forecasts.
16. **holiday_indicator**:
    - **Description**: Boolean flag indicating whether the transaction occurred during a holiday.
    - **Data Type**: Boolean
    - **Example**: True, False
    - **Analysis**: Analyzing sales uplift during holidays can help optimize inventory and promotions during holiday periods.
17. **weekday**:
    - **Description**: The day of the week when the transaction took place.
    - **Data Type**: String
    - **Example**: "Monday", "Friday"
    - **Analysis**: Weekday performance analysis helps in identifying which days of the week generate higher sales, enabling scheduling promotions and inventory optimization.
18. **stockout_indicator**:
    - **Description**: Boolean flag indicating if the product was out of stock at the time of the transaction.
    - **Data Type**: Boolean
    - **Example**: True, False
    - **Analysis**: Critical for understanding stockout frequency, and how it affects sales and customer satisfaction. Can guide inventory planning.
19. **forecasted_demand**:
    - **Description**: Predicted demand for the product.
    - **Data Type**: Integer
    - **Example**: 172, 289
    - **Analysis**: Forecast accuracy analysis. Compare forecasted demand with actual sales to improve future forecasting models and minimize stockouts/overstocking.
20. **actual_demand**:
    - **Description**: Actual demand for the product as recorded in sales.
    - **Data Type**: Integer
    - **Example**: 179, 416
    - **Analysis**: Used to measure forecasting accuracy, sales performance, and make future demand predictions.

Agile Methodology Implementation:

- **Sprint 1**: Data Collection and Cleaning
    - Conduct initial exploratory data analysis (EDA) to identify data patterns, clean missing values, and ensure data consistency.
    - Collaborate with stakeholders (inventory managers and store managers) to define key metrics for the analysis.

- **Sprint 2**: Exploratory Data Analysis and Initial Forecast
    - Analyze customer buying patterns, weather impact, and promotional effectiveness using various data techniques.
    - Present preliminary insights in a sprint review to gather feedback from stakeholders.
- **Sprint 3**: Refining Insights and Developing Forecasts
    - Build a detailed analysis report with advanced forecasting models.
    - Focus on optimizing inventory levels based on projected demand patterns.
- **Sprint 4**: Insights Presentation and Actionable Recommendations
    - Present the final analysis and recommendations to key stakeholders in a sprint review meeting.
    - Document all findings, ensuring alignment with business goals and real-time decision-making processes.

Problem Definition:

Walmart needs a data-driven inventory optimization solution that minimizes both stockouts and overstocking by forecasting demand based on past sales data. The goal is to develop a scalable approach that adjusts inventory levels dynamically using real-time insights from the analysis.

Key Questions for Stakeholders:

1. **Inventory Managers**:
    - How can we improve inventory accuracy and ensure that stock levels match demand?
    - What factors are contributing to stockouts, and how can we address them?
2. **Store Managers**:
    - How can we allocate inventory more efficiently across stores based on local demand patterns?
    - How do promotions and weather influence sales at different store locations?
3. **Marketing and Promotions Team**:
    - What is the return on investment (ROI) for promotions, and how do they affect demand?

Stakeholder Involvement:

- **Internal Stakeholders**:
    - **Inventory Managers**: Will use the demand forecasting insights to ensure optimal stock levels across stores.
    - **Store Managers**: Responsible for managing in-store inventory based on demand.
    - **Marketing Team**: Understands how promotions impact sales and tailors campaigns accordingly.

- **External Stakeholders**:
    - **Customers**: Their purchasing behavior directly influences demand forecasts.
    - **Suppliers**: Will receive more accurate forecasts to ensure timely and sufficient product delivery.

Data Requirements:

- **Sales Transactions**: Data including product, store location, transaction date, and customer info.
- **Weather Conditions**: Weather during the sales period to assess external impact.
- **Promotions Data**: Information on promotional activities and their effect on sales.
- **Inventory and Stockout Data**: Indicators show whether stockouts occurred and whether they were forecasted vs. actual demand.

Metric Development:

1. **Stockout Rate**: Number of stockouts as a percentage of total sales.
2. **Forecast Accuracy**: Difference between forecasted demand and actual demand.
3. **Sales Uplift from Promotions**: Percentage increase in sales during promotions.
4. **Weather Impact on Sales**: Sales variance in different weather conditions.
5. **Customer Loyalty and Purchasing Patterns**: Impact of loyalty tier on purchasing behavior.

Basic-Level Questions (Excel-Oriented):

1. **Q1: What is the total sales volume across all Walmart stores?**
    - **Hint**: Use the `SUM` function in Excel to calculate the total of the `quantity_sold` column.
    - **How it helps**: Provides a bird's-eye view of Walmart's overall product sales volume.
    - **Business Impact**: Helps in determining high-level sales performance and overall inventory demand.
2. **Q2: What is the average unit price of all products sold, by category?**
    - **Hint**: Use `AVERAGE` in combination with `Pivot Table` grouped by `category`.
    - **How it helps**: Identifies product pricing trends across different categories.
    - **Business Impact**: Assists in price optimization for different product lines to maximize profit margins.
3. **Q3: Which stores are underperforming in terms of sales volume?**
    - **Hint**: Use `Pivot Table` with store-wise `SUM` of `quantity_sold`.

    - **How it helps**: Identifies stores that may need additional inventory support or different marketing strategies.
    - **Business Impact**: Allows the operations team to allocate inventory more strategically and support underperforming locations.
4. **Q4: What percentage of transactions involved promotions?**
    - **Hint**: Apply `Filter` and count the number of `promotion_applied` rows divided by total transactions.
    - **How it helps**: Helps assess the reach and usage of promotional campaigns.
    - **Business Impact**: Provides insight into whether promotions are driving enough sales to justify their costs.
5. **Q5: What is the total revenue generated from sales during promotional events?**
    - **Hint**: Create a calculated column for total revenue (`quantity_sold * unit_price`) and filter rows where `promotion_applied` is `True`.
    - **How it helps**: Measures the financial impact of promotional events on overall revenue.
    - **Business Impact**: Helps determine the effectiveness of promotions and refine marketing strategies.

6. **Q6: Which product categories see the largest increase in demand during holidays?**
    - **Hint**: Use `Pivot Table` to analyze `quantity_sold` for `holiday_indicator` `True` and group by `category`.
    - **How it helps**: Understands product demand variation during holidays.
    - **Business Impact**: Enables better inventory planning for peak holiday seasons.
7. **Q7: What is the correlation between weather conditions and sales performance?**
    - **Hint**: Create a summary table comparing sales volumes under different `weather_conditions`.
    - **How it helps**: Reveals how external factors, like weather, influence customer purchasing behavior.
    - **Business Impact**: Assists in demand forecasting based on weather trends, leading to better inventory preparation.
8. **Q8: How does customer loyalty level affect purchasing patterns?**
    - **Hint**: Use a `Pivot Table` to analyze `quantity_sold` by `customer_loyalty_level`.
    - **How it helps**: Helps understand how loyalty programs impact repeat purchases.
    - **Business Impact**: Insights can be used to enhance customer retention strategies.
9. **Q9: What is the forecast accuracy for each store location?**
    - **Hint**: Create a calculated column for `forecasted_demand` vs `actual_demand` and analyze by `store_location`.

- **How it helps**: Measures the accuracy of demand forecasts, which is crucial for inventory management.
- **Business Impact**: Allows Walmart to refine forecasting models, leading to better stock management.

10. **Q10: What is the stockout rate at each store?**
    - **Hint**: Use the `stockout_indicator` to calculate the percentage of stockout incidents per store.
    - **How it helps**: Identifies stores where stock management is failing.
    - **Business Impact**: Reducing stockouts helps prevent lost sales and improves customer satisfaction.

Medium-Level Questions (Excel-Oriented):

1. **Q1: Which products are experiencing the highest variance between forecasted and actual demand, and what are the potential reasons?**
    - **Hint**: Create a calculated column for the variance (`forecasted_demand - actual_demand`) and sort the products with the highest variance.
    - **How it helps**: Helps in understanding the discrepancies in demand forecasts and identifying possible external factors affecting sales (e.g., inaccurate forecasts, unpredicted events).
    - **Business Impact**: Refining the forecasting model to reduce inaccuracies and align actual demand with predicted values can minimize stockouts or excess inventory.
2. **Q2: Which stores have the highest stockout rates during promotional events, and what could be the underlying cause?**
    - **Hint**: Use a `Pivot Table` to filter for promotional events (`promotion_applied` = True) and calculate stockout rates using `stockout_indicator`.
    - **How it helps**: Identifies how promotional events can lead to unexpected demand surges and stockouts.
    - **Business Impact**: Adjusting inventory allocation during promotions can ensure high-demand products are sufficiently stocked to avoid lost sales.
3. **Q3: How does revenue differ between stores in urban vs. rural locations?**
    - **Hint**: Use `Pivot Table` to group by `store_location`, categorize stores based on urban/rural, and sum the revenue (`quantity_sold * unit_price`).
    - **How it helps**: Understanding the revenue gap between urban and rural stores helps in allocating resources based on the needs of different store types.
    - **Business Impact**: Tailoring inventory strategies based on location can improve sales performance and customer satisfaction.
4. **Q4: How does customer loyalty level influence promotion effectiveness?**
    - **Hint**: Create a summary table to compare the `quantity_sold` during promotions across different `customer_loyalty_level` tiers.
    - **How it helps**: Reveals whether loyal customers respond better to promotions than new customers or occasional shoppers.

    - **Business Impact**: Helps Walmart adjust promotion strategies to maximize effectiveness for each customer tier and improve loyalty-based retention.
5. **Q5: What is the impact of weather conditions on high-ticket products (e.g., Electronics)?**
    - **Hint**: Filter for high-ticket products in the `category` column and compare sales across different `weather_conditions` using a `Pivot Table`.
    - **How it helps**: Analyzing the correlation between weather conditions and sales for high-priced items helps identify external drivers for customer purchasing behavior.
    - **Business Impact**: Better planning for weather-driven demand can prevent stockouts or overstocking during specific weather conditions.
6. **Q6: How do holiday promotions compare to regular promotions in terms of sales uplift?**
    - **Hint**: Use the `holiday_indicator` to differentiate sales during holidays and non-holidays, and compare the percentage sales uplift during promotions.
    - **How it helps**: Highlights the effectiveness of holiday-specific promotions versus regular promotional efforts.
    - **Business Impact**: Helps Walmart plan better for holiday-specific promotions and invest in the right types of campaigns.
7. **Q7: What is the percentage of products with forecasted demand higher than actual sales, and what does it indicate?**
    - **Hint**: Calculate the percentage of products where `forecasted_demand` is greater than `actual_demand`.
    - **How it helps**: This metric reveals overestimated demand scenarios and highlights areas for improving the forecasting model.
    - **Business Impact**: Reducing overestimated demand prevents overstocking, which can lead to unnecessary costs.
8. **Q8: Which store locations have the most significant revenue loss due to stockouts?**
    - **Hint**: Calculate the potential lost revenue by filtering `stockout_indicator = True` and summing the `actual_demand * unit_price`.
    - **How it helps**: Identifies store locations where stockouts are causing substantial revenue loss.
    - **Business Impact**: Helps prioritize stores for inventory improvements to reduce stockouts and prevent lost revenue.
9. **Q9: How does promotion type (e.g., percentage discount vs. buy-one-get-one) impact the sales of different product categories?**
    - **Hint**: Use a `Pivot Table` to compare `promotion_type` with `quantity_sold` across various `category`.
    - **How it helps**: Allows for analysis of which promotion types work best for specific product categories.
    - **Business Impact**: Enables Walmart to tailor promotions to product categories for maximum impact on sales.
10. **Q10: Which product categories are most prone to stockouts, and how should Walmart adjust its inventory strategy for these categories?**

- **Hint**: Use a `Pivot Table` to group by `category` and sum the occurrences of `stockout_indicator = True`.
- **How it helps**: Identifies which categories are at higher risk for stockouts.
- **Business Impact**: Helps Walmart adjust inventory levels and stocking policies for vulnerable product categories, reducing the frequency of stockouts.

How These Questions Help the Company:

- **Q1-Q3**: Enable deeper understanding of demand forecasting errors, store performance, and location-based differences in sales, ensuring that resources are allocated efficiently and operations are streamlined.
- **Q4-Q5**: Provide insights into customer behavior and external influences like weather, allowing Walmart to tailor their inventory strategies and improve customer satisfaction.
- **Q6-Q10**: Focus on stockout rates, promotional effectiveness, and category-specific performance, driving actionable improvements in promotions and stock management to minimize lost revenue and overstocking.

Advanced-Level Questions (Excel-Oriented):

1. **Q1: How can Walmart identify which store locations are at the highest risk of stockouts based on sales trends, promotions, and external factors (e.g., weather or holidays)?**
   - **Hint**: Create a weighted scoring model by analyzing historical sales trends, stockout occurrences, promotions, weather conditions, and holiday effects, assigning different weights to each factor. Use a `Pivot Table` and `Weighted Average` to calculate stockout risk for each store.
   - **How it helps**: Helps Walmart identify which store locations are most vulnerable to stockouts under specific conditions.
   - **Business Impact**: Allows Walmart to focus on improving inventory planning for high-risk locations, reducing the likelihood of stockouts and preventing lost sales.
2. **Q2: Can you analyze how forecast accuracy varies across different product categories and suggest improvements in demand forecasting for the worst-performing categories?**
   - **Hint**: Use `Pivot Tables` to group by `category` and calculate the forecast error (`forecasted_demand - actual_demand`). Highlight categories with the largest discrepancies. Suggest which external factors or demand indicators should be included to improve forecasts (e.g., consider seasonality, promotion impact).
   - **How it helps**: Identifies product categories with consistent forecast errors and suggests refinements in the demand forecasting model.

- **Business Impact**: Improving forecast accuracy for underperforming categories will reduce both overstocking and stockouts, optimizing Walmart's inventory management.

3. **Q3: Develop a revenue-maximization strategy by analyzing which combinations of promotions, weather conditions, and customer loyalty levels yield the highest sales uplift.**
   - **Hint**: Use a `Pivot Table` to analyze the interaction between `promotion_type`, `weather_conditions`, and `customer_loyalty_level` and their combined effect on `quantity_sold`. Use correlation analysis in Excel to determine which combination yields the highest sales uplift.
   - **How it helps**: Determines the best strategies to maximize sales by leveraging specific promotion types and understanding how they interact with weather and customer loyalty.
   - **Business Impact**: Walmart can tailor promotions and inventory strategies to match these high-impact combinations, maximizing revenue during peak conditions.
4. **Q4: How can Walmart forecast product demand for new store locations based on the historical performance of similar existing stores?**
   - **Hint**: Identify existing stores with similar demographic, geographic, and economic conditions to the new location. Use historical sales data from these stores to estimate demand at the new store, creating a forecast model using `AVERAGE` and `TREND` functions.
   - **How it helps**: Helps Walmart predict sales for new stores by finding patterns from similar stores.
   - **Business Impact**: Ensures new store locations are stocked appropriately from day one, preventing both overstocking and stockouts.
5. **Q5: How can Walmart develop an inventory management plan that accounts for seasonal sales variations across different regions and product categories?**
   - **Hint**: Analyze the sales data by `store_location` and `season` (derived from `transaction_date`) and group by `product_category`. Use trend analysis to determine seasonal peaks and troughs in sales. Create a region-specific inventory plan based on these insights.
   - **How it helps**: Identifies region-specific and seasonal demand trends, enabling better forecasting for each region's needs during peak seasons.
   - **Business Impact**: Ensures optimized inventory management that aligns with seasonal demand, reducing both stockouts and unnecessary overstocking during off-seasons.
6. **Q6: Which specific products should Walmart prioritize during peak seasons based on past demand, forecast accuracy, and sales volume?**
   - **Hint**: Use `Pivot Tables` to analyze sales volume and forecast accuracy for each `product_name` during peak seasons. Identify the top products with the highest sales volume and most accurate forecasts, and prioritize them for inventory stocking.
   - **How it helps**: Prioritizes high-performing, high-demand products for increased inventory during peak seasons.

    - **Business Impact**: Ensures key products are always available during peak demand periods, driving more revenue and customer satisfaction.
7. **Q7: How can Walmart optimize inventory for low-margin products to ensure profitability while maintaining sufficient stock levels?**
    - **Hint**: Identify low-margin products by calculating gross margin (`unit_price - cost_price`). Use sales data to analyze their performance during promotions and holidays. Create a strategy for balancing inventory levels and promotional pricing to ensure profitability.
    - **How it helps**: Balances inventory costs and promotional pricing for low-margin products, ensuring profitability even with lower profit margins.
    - **Business Impact**: Helps Walmart ensure that low-margin products are stocked and priced efficiently without sacrificing profitability.
8. **Q8: How can Walmart predict future promotions' effectiveness based on historical promotion data and external factors?**
    - **Hint**: Analyze past promotion data to evaluate sales uplift, factoring in external influences like weather, holidays, and customer demographics. Use correlation and trend analysis to predict the likely impact of future promotions.
    - **How it helps**: Forecasts the success of future promotions based on historical trends and external factors.
    - **Business Impact**: Enables Walmart to plan promotions that are more likely to succeed, increasing sales while minimizing inventory strain.
9. **Q9: How can Walmart use cluster analysis (manually in Excel) to segment stores by performance and customize inventory strategies for each segment?**
    - **Hint**: Group stores based on performance metrics like total revenue, stockout rate, and customer demographics. Create custom clusters (e.g., high-revenue, low-revenue, high stockout, etc.) and develop inventory strategies based on the performance of each cluster.
    - **How it helps**: Manual cluster analysis allows Walmart to segment stores and apply tailored inventory strategies for each group.
    - **Business Impact**: By segmenting stores, Walmart can focus on optimizing inventory management for each cluster, ensuring stores with unique challenges receive specific attention.
10. **Q10: How can Walmart optimize store-level inventory allocations during promotional periods by analyzing sales data from past promotional events?**
    - **Hint**: Use `Pivot Tables` to compare sales volume during promotional periods for each store, and identify trends. Use this information to propose store-level inventory allocation strategies during future promotions.
    - **How it helps**: Identifies stores that may require more stock during promotions and helps adjust inventory strategies to meet demand spikes.
    - **Business Impact**: Prevents stockouts at high-performing stores during promotions and ensures that all locations have sufficient inventory to meet promotional demand.

How These Questions Help the Company:

- **Q1-Q3**: Enable Walmart to predict and manage stockouts, optimize promotions, and understand how external factors impact sales. These insights help in better inventory planning and resource allocation.
- **Q4-Q6**: Allow Walmart to optimize inventory for new stores, handle seasonal demand fluctuations, and prioritize high-demand products for peak seasons, ensuring stock availability at the right times.
- **Q7-Q10**: Focus on profitability strategies for low-margin products, predictive modeling for future promotions, store-level customization of inventory, and data-driven strategies for future promotions, maximizing both revenue and customer satisfaction.