***Milestone 1: Topic Selection and Problem Statement***

CNIT 37200

Samuel Gomez

Noah Zhou

John Hsu

Nicholas Lauw

Submitted To: Dr. Tianyi Li

Date Submitted: 09/24/2024

Date Due: 10/06/2024

### 

## 

### **Domain Selection**

We have selected Option1: Working with a client (Walmart) to improve their data management on their inventory. Our group has decided to work on a retail inventory management problem, specifically focusing on Walmart. Walmart is one of the largest retail companies worldwide and efficiently managing its inventory is important to its operation. The project aims to improve Walmart's data management system for its inventory, ensuring that the company can maintain optimal stock levels and minimize overstock and stockouts.

### **Problem Context**

Walmart's inventory management is an essential aspect of its supply chain. The current system faces challenges in managing large volumes of data across thousands of stores and distribution centers. These challenges can result in inaccurate stock levels, leading to overstock or stockouts, which negatively impacts sales, customer satisfaction, and operational efficiency. Improving the data management system will help Walmart streamline its inventory processes, enabling better decision-making and a more efficient supply chain.

### **Motivation**

The general plan behind this project is to address the needs for a more accurate and efficient inventory management in the retail business. Walmart operates on a global scale, where small inefficiencies can lead to significant losses. By improving the data management system, we want to help Walmart reduce costs, improve customer satisfaction, and ensure that the right products are available in the right quantities at the right time.

### **Data Plan**

We plan to use Walmart's historical inventory and sales data, including stock levels, reorder points, supplier information, and transaction data across their different stores and distribution centers. Since the real data may be sensitive, we will generate synthetic data that reflects these data points while maintaining the integrity of the real-world scenario.

### **Teamwork**

Our group has divided tasks based on individual strengths. One member will focus on database design, another on data analytics, while the rest will handle coding and report writing. We meet regularly to update progress, collaborate on shared tasks, and ensure that everyone contributes equally.

### 

### **Question Development**

* How can Walmart track inventory levels across all stores and distribution centers?
* What are the patterns in product demand across different regions and seasons?
* How can data management help predict stock shortages before they occur?
* What factors contribute to overstock situations, and how can they be minimized?
* How can Walmart optimize reorder points for different product categories?
* What is the impact of inaccurate inventory data on sales performance?
* How can data analysis improve supplier lead time estimates for better stock replenishment?
* What products have the highest return rates, and how does this affect inventory management?
* How can Walmart reduce the time it takes to restock shelves after an item is sold out?
* How can the system identify slow-moving items and adjust stock levels accordingly?

### **Insights Derived from Data**

The insights we aim to derive include:

* Identification of patterns in product demand based on regions and seasons.
* Stock shortage predictions using historical data and trends.
* Identification of overstock and slow-moving items.
* Supplier lead time estimation for efficient replenishment.
* Reorder point optimization for different product categories.

### **Importance of These Insights**

These insights are essential for efficient inventory management and business operations. By optimizing stock levels for each region, businesses can ensure products are readily available where they are most needed, minimizing the risk of stockouts that could lead to lost sales and dissatisfied customers.

* This helps optimize inventory levels by region, ensuring products are available when and where they are needed.
* Preventing stockouts is crucial to maintaining sales and customer satisfaction.
* Reducing excess stock saves costs on storage and decreases the chance of unsold items.

### **Tables Outline**

* **Inventory**: information about products and their stock levels will be stored in each store.
* **Sales Transactions**: record each sale, linking it to the product and the store.
* **Suppliers**: will store supplier details and link to products to track supplier performance.

### **Database Design**

In the following list, here are some of the column names, data types, and constraints that we might use in our database

1. **Inventory Table**:
   * Product\_ID (INT, Primary Key, NOT NULL): Unique identifier for products.
   * Store\_ID (INT, NOT NULL): Unique identifier for stores.
   * Stock\_Level (INT, NOT NULL, CHECK >= 0): Current stock count of the product.
   * Reorder\_Point (INT, NOT NULL, CHECK >= 0): Minimum stock level before reordering.
2. **Sales Transactions Table**:
   * Transaction\_ID (INT, Primary Key, NOT NULL): Unique identifier for each sale.
   * Product\_ID (INT, Foreign Key, NOT NULL): Links to the product in the Inventory table.
   * Store\_ID (INT, Foreign Key, NOT NULL): Links to the store in the Inventory table.
   * Quantity\_Sold (INT, NOT NULL, CHECK > 0): Number of units sold.
   * Sale\_Date (DATE, NOT NULL): Date of the transaction.
3. **Suppliers Table**:
   * Supplier\_ID (INT, Primary Key, NOT NULL): Unique identifier for suppliers.
   * Product\_ID (INT, Foreign Key, NOT NULL): Links to the product in the Inventory table.
   * Lead\_Time (INT, NOT NULL, CHECK >= 0): Days required for delivery.
   * Delivery\_History (TEXT): Delivery performance data.

### **Relationships Between Tables**

Understanding relationships between tables in a database is crucial for managing and analyzing data effectively. These connections, often made through shared fields, allow us to link related information across different tables. Such relationships enable a comprehensive view of the data, supporting better decision-making and operational management.

* The Inventory and SalesTransactions tables are linked via Product\_ID and Store\_ID. This allows tracking of stock levels relative to sales.
* The Inventory and Suppliers tables are linked via Product\_ID, allowing insights into how supplier lead times affect stock levels and reorder efficiency.