**PROJECT TITLE: "** **Optimizing Superstore Operations: Empowering Demand Forecasting And Supply Chain Efficiency Through Data Analytics"**

# CHAPTER 1: INTRODUCTION

## Project Overview:

The objective of this project is to leverage data analytics techniques to enhance demand forecasting and optimize supply chain management for a fictional Superstore – ***ABC Superstore***. This analysis was conducted by utilizing three distinct datasets, each containing valuable insights into sales transactions, inventory management, and supply chain operations.

## Goal:

* To optimize operational efficiency and inventory management.
* To improve demand forecasting accuracy and supply chain performance.
* To enhance customer satisfaction and loyalty through improved product availability and service quality.

## Rationale:

Accurate demand forecasting and efficient supply chain management are crucial for suppliers to meet customer demand effectively and maintain competitive advantage in the market. By leveraging data-driven approaches, retailers can gain insights into sales trends, customer segmentations, and product categories, enabling them to make informed decisions to optimize operations and maximize profitability.

## Expected outcomes:

* Conduct descriptive and predictive analytics on each dataset independently to gain insights into specific aspects of demand forecasting and supply chain management.
* Build predictive models to forecast sales, optimize inventory levels, evaluate supplier performance, and improve transportation and logistics efficiency.
* Implement data-driven strategies and recommendations based on the analysis to enhance operational efficiency and drive business growth.
* Detailed analysis reports for each dataset, including findings, insights, and recommendations.
* Predictive models for demand forecasting, inventory optimization, supplier performance evaluation, and transportation logistics.
* Actionable insights and recommendations for improving demand forecasting accuracy, inventory management efficiency, and supply chain performance.

# CHAPTER 2: DATA IDENTIFICATION AND EXPLORATION

## 2.1 Data Sources:

### Sales Transaction Dataset:

This dataset provides detailed information on sales transactions from various shopping malls, including customer demographics, product categories, quantities sold, and revenue generated. The data was sourced from: <https://github.com/souravpatro26/Retail_Data_Analysis.git>

### Inventory Management Dataset:

The inventory management dataset offers insights into inventory levels, stock movements, and order quantities for different products, enabling analysis of inventory turnover analysis, stockout rates, and lead times. The data was sourced from: https://github.com/sjpradhan/Inventory-ABC-Analysis.git

### Supply Chain Operations Dataset:

This dataset contains information on supply chain operations, such as supplier performance, transportation modes, production volumes, and manufacturing lead times, enabling analysis of supply chain costs, efficiency, and quality control. The data was sourced from: <https://www.kaggle.com/datasets/amirmotefaker/supply-chain-dataset>

## 2.2 Dataset key variables

### Sales Transaction Dataset:

* Product category information table: prod\_cat, prod\_cat\_code, prod\_subcat\_code, prod\_subcat
* Transaction table: transaction\_id, cust\_id, tran\_date prod\_subcat\_code, prod\_cat\_code, Qty, Rate, Tax, total\_amt, Store\_type
* Customer demographics table: customer\_Id, DOB, Gender, city\_code.

### Inventory Management Dataset:

Order history table: Order Date, SKU ID, Order Quantity

Stock table: SKU ID, Current Stock Quantity, Units (Nos/Kg), Average Lead Time (days), Maximum Lead Time (days), Unit Price

### Supply Chain Operations Dataset:

Identify key variables: Product type, SKU, Price, Availability, Number of products sold, Revenue generated, Customer demographics, Stock levels, Lead times, Order quantities, Shipping times, Shipping carriers,Shipping costs, Supplier name, Location, Lead time, Production volumes, Manufacturing lead time, Manufacturing costs, Inspection results, Defect rates, Transportation modes, Routes, Costs

## 2.3 Data Exploration:

* Conduct exploratory data analysis (EDA) to understand sales trends over time.
* Analyze customer demographics to identify patterns in purchasing behavior.
* Calculate aggregate sales metrics such as total revenue and average transaction value.
* Explore inventory levels and stock movements to assess inventory turnover rates.
* Analyze order quantities and lead times to optimize inventory replenishment strategies.
* Identify instances of stockouts and assess their impact on sales and customer satisfaction.
* Investigate supplier performance metrics such as lead times and order accuracy.
* Analyze transportation modes and shipping costs to optimize logistics operations.
* Assess production volumes and manufacturing lead times to optimize production scheduling.

## 2.4 Data Validation and Cleansing:

* Cleanse, preprocess/transform and validate the datasets to ensure data integrity and consistency.
* Handle missing values, outliers, and data formatting issues.
* Transform categorical variables and engineer new features where necessary.
* Standardize units and transform variables as needed.

# CHAPTER 3: ANALYSIS AND MODELING - SALES TRANSACTION DATASET

## Sales Trend Analysis

1. Descriptive Analysis:

* Analyze transaction data over time to identify sales trends and seasonality.
* Calculate aggregate sales metrics such as total revenue, average transaction value, and number of transactions per period.
* Visualize sales trends using time series plots and identify patterns or fluctuations in sales volume.

1. Predictive Analysis:
   1. Time Series Forecasting Model (ARIMA & Holt-Winters): Use AutoRegressive Integrated Moving Average (ARIMA) to model sales over time. Incorporate seasonality and trend components into the forecast using Holt-Winters.

* Machine Learning Models (Linear Regression): Predict future sales based on key features and relevant variables.

## Customer Segmentation

1. Descriptive Analysis:

* Segment customers based on demographics (e.g., total sales by age group and gender)
* Segment customers based on RFM Analysis rule.

1. Predictive Analysis:

* (Classification Models (Logistic Regression)): Predict customer segments based on demographic variables and transactional data.

## Product Category Analysis:

Descriptive:

* Use Market Basket Analysis to determine products are are most likely to be sold together. Analyze sales performance by product category to identify top-selling categories and assess product demand.
* Determine which product sale seasonality.

Predictive:

* ARIMA or Holt-Winters: Predicts sales for each product category over time.
* Linear Regression: Analyzes the relationship between product categories and sales revenue, considering factors like price and seasonality.

## Store Performance Analysis

Descriptive:

* Evaluate sales performance across different store types, product category and subcategories.
* Identify high-performing stores and assess factors contributing to their success.

Predictive:

* Multiple Linear Regression: Predicts store performance based on factors like store type, product category and subcategories.

Chapter 4: Analysis and Modeling - Inventory Management Dataset

1. **Inventory Turnover Analysis:**

Descriptive Analysis:

* Calculate inventory turnover ratios to assess how quickly inventory is sold and replenished.
* Identify slow-moving or obsolete inventory items that may require liquidation or markdowns.

**Predictive**

* Linear Regression: Predicts inventory turnover rates based on factors like sales volume, order frequency, and lead time.

1. **Stockout Analysis:**

Descriptive Analysis:

* Analyze instances of stockouts and assess their impact on sales revenue and customer satisfaction.
* Identify trends in stockouts and determine potential causes, such as insufficient inventory levels or supply chain disruptions.

**Predictive**

* Logistic Regression: Predicts the probability of stockouts based on factors like demand variability and lead time.

1. **Lead Time Analysis:**

Descriptive Analysis:

* Analyze lead time data to assess the time it takes to replenish inventory after placing an order.
* Evaluate the reliability and consistency of lead times across different suppliers or product categories.

**Predictive:**

* Identifies trends and seasonality in lead times to optimize inventory management.

**Inventory-ABC-Analysis**

Descriptive Analysis:

* Conduct ABC analysis on the inventory management dataset to categorize products based on their value and usage.
* Assess the impact of ABC categorization on overall inventory performance and supply chain efficiency.

# Chapter 5: Analysis and Modeling - Supply Chain Operations Dataset

Descriptive Analysis:

1. Supply Chain Cost Analysis:

* Analyze supply chain costs, including transportation costs, manufacturing costs, and inspection costs.
* Identify cost drivers and areas for cost reduction or optimization within the supply chain.

Predictive:

* Cost Estimation Models: Predicts supply chain costs based on factors like transportation expenses, manufacturing costs, and inspection costs.
* Activity-Based Costing (ABC) Models: Allocates costs to different activities within the supply chain to identify cost drivers.

1. Production Volume Forecasting:

* Forecast production volumes based on historical data and demand forecasts.
* Optimize production schedules and resource allocation to meet demand while minimizing costs.

Predictive:

* Exponential Smoothing: Forecasts production volumes while accounting for trend and seasonality.

1. Manufacturing Lead Time Analysis:

* Analyze manufacturing lead times to identify bottlenecks and inefficiencies in the production process.
* Implement process improvements to reduce lead times and improve production efficiency.

1. Quality Control Analysis:

* Analyze inspection results and defect rates to assess product quality and compliance with quality standards.
* Implement corrective actions to reduce defects and improve product quality assurance processes.

Predictive:

* Defect Prediction Models: Predicts the likelihood of defects or quality issues based on historical inspection results and process parameters.

Validation using SQL and Python

|  |  |  |
| --- | --- | --- |
| Objectives |  |  |
| 1 | Analyze transaction data over time to identify sales trends and seasonality.  SELECT COUNT(\*) AS "Number of Transactions"  FROM sales\_transaction\_transactions\_tbl  WHERE tran\_date = '09-03-2013'; | C:\Users\admin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Screenshot (156).png |
| 2 |  |  |

# Chapter 6: Key Insights