**Project Title: Sateesh Store-Insights**

**Project Overview:**

This project aimed to create a unified dashboard utilizing two datasets—Orders.csv and Details.csv—to provide insightful analysis on market-Insight operations, evaluating revenue trends, category performance, and customer behavior.

**Step 1: Data Collection and Understanding**

1. **Datasets Collected:**
   * **Orders.csv:** Contains order-level details such as:
     + OrderDate
     + CustomerName
     + City
   * **Details.csv:** Contains item-level details such as:
     + Category
     + Amount
     + Profit
2. **Data Exploration:**
   * Examined the structure of both datasets:
     + Checked columns and data types for accuracy.
   * Identified primary keys:
     + **Primary Key:** OrderID (assumed to be present in both datasets).
3. **Noted Data Quality Issues:**
   * Missing values in crucial columns.
   * Duplicates identified across the datasets.
   * Inconsistencies in categories and date formats.

**Step 2: Data Cleaning and Transformation**

1. **Data Cleaning:**
   * **Removed Duplicates:**
   * DELETE FROM Orders
   * WHERE OrderID IN (SELECT OrderID FROM Orders GROUP BY OrderID HAVING COUNT(\*) > 1);
   * **Handled Missing Data:**
     + **Filled missing numerical values (e.g., Profit) with the average:**
     + UPDATE Details
     + SET Profit = (SELECT AVG(Profit) FROM Details WHERE Profit IS NOT NULL)
     + WHERE Profit IS NULL;
     + **Removed rows with missing OrderID:**
     + DELETE FROM Orders WHERE OrderID IS NULL;
2. **Data Standardization:**
   * Standardized date formats and categorical values for consistency:
   * UPDATE Orders SET OrderDate = CONVERT(DATE, OrderDate, 101);
3. **Data Merging:**
   * **SQL Query to Join Datasets:**
   * SELECT o.OrderID, o.OrderDate, o.CustomerName, o.City,
   * d.Category, d.Amount, d.Profit
   * FROM Orders o
   * JOIN Details d ON o.OrderID = d.OrderID;
4. **New Calculated Fields:**
   * **Profit Margin:**
   * ALTER TABLE Details ADD ProfitMargin AS (Profit / Amount \* 100);
   * **Year and Month from OrderDate:**
   * ALTER TABLE Orders ADD Year INT, Month INT;
   * UPDATE Orders SET Year = YEAR(OrderDate), Month = MONTH(OrderDate);

**Step 3: Data Analysis and Insights**

1. **Aggregated Metrics Using SQL:**
   * **Total Revenue:**
   * SELECT SUM(Amount) AS TotalRevenue FROM Details;
   * **Total Profit:**
   * SELECT SUM(Profit) AS TotalProfit FROM Details;
   * **Top-Selling Product Categories:**
   * SELECT Category, SUM(Amount) AS TotalSales
   * FROM Details
   * GROUP BY Category
   * ORDER BY TotalSales DESC;
   * **Payment Trends:**
   * SELECT PaymentMethod, COUNT(\*) AS NumberOfTransactions
   * FROM Orders
   * GROUP BY PaymentMethod;
2. **Key Insights Identified:**
   * Derived metrics reflected that:
     + **Electronics** was the most profitable category.
     + **Sales peaks** occurred during specific holiday seasons.
     + **Top payment methods**: Credit Card and EMI.
     + **Highest revenue-generating cities**: Gudur and Kavali.

**SQL:**

1. Retrieve all orders with their amount, profit, and payment mode.

2. Find the total sales for each product category.

3. Calculate the average profit per order.

4. Get the top 5 subcategories by sales amount.

5. Count the number of orders placed in each city.

6. Find the payment mode with the highest frequency.

7. Identify orders with a profit margin greater than 50%.

8. Calculate the total quantity sold for each subcategory.

9. Retrieve all orders placed in 2018.

10. Find the total profit grouped by city.

**Power BI:**

1. What is the trend of total sales over time?

2. How is the profit distributed across different categories?

3. Which are the top 5 cities by total sales?

4. What is the distribution of each payment mode used?

5. How do sales vary across different localities?

6. What is the average order value per category?

7. Which subcategory contributes most to profit?

8. What is the distribution of orders by date?

9. How many orders are placed in each quarter of 2018?

10. What is the breakdown of orders by customer names?

**Data Columns**

**Details.csv Orders.csv**

Order ID Order ID

Amount Order Date

Profit Customer Name

Quantity City

Category Locality

Sub-Category

PaymentMode

**SQL Queries:**

**1. Retrieve all orders with their amount, profit, and payment mode:**

SELECT `Order ID`, Amount, Profit, PaymentMode FROM Details;

**2. Find the total sales for each product category:**

SELECT Category, SUM(Amount) AS Total\_Sales FROM Details GROUP BY Category;

**3. Calculate the average profit per order:**

SELECT AVG(Profit) AS Average\_Profit FROM Details;

**4. Get the top 5 subcategories by sales amount:**

SELECT Sub-Category, SUM(Amount) AS Total\_Sales FROM Details GROUP BY Sub-Category ORDER BY Total\_Sales DESC LIMIT 5;

**5. Count the number of orders placed in each city:**

SELECT City, COUNT(`Order ID`) AS Order\_Count FROM Orders GROUP BY City;

**6. Find the payment mode with the highest frequency:**

SELECT PaymentMode, COUNT(\*) AS Frequency FROM Details GROUP BY PaymentMode ORDER BY Frequency DESC LIMIT 1;

**7. Identify orders with a profit margin greater than 50%:**

SELECT \* FROM Details WHERE Profit > Amount \* 0.5;

**8. Calculate the total quantity sold for each subcategory:**

SELECT `Sub-Category`, SUM(Quantity) AS Total\_Quantity FROM Details GROUP BY `Sub-Category`;

**9. Retrieve all orders placed in 2018:**

SELECT \* FROM Orders WHERE YEAR(STR\_TO\_DATE(`Order Date`, '%d-%m-%Y')) = 2018;

**10. Find the total profit grouped by city:**

SELECT o. City, SUM(d.Profit) AS Total\_Profit

FROM Details d

JOIN Orders o ON d.`Order ID` = o.`Order ID`

GROUP BY o.City;

**POWER BI:**

1. **Trend of Sales:** A line chart showing total sales over time.
2. **Profit Distribution:** A bar chart displaying profit across categories.
3. **Top Cities:** A column chart showing the top 5 cities by sales.
4. **Payment Modes:** A pie chart depicting the proportion of each payment mode.
5. **Sales by Locality:** A map visualization displaying sales density by locality.
6. **Average Order Value:**  A table showing average order value per category.
7. **Top Subcategories:** A bar chart highlighting subcategories contributing most to profit.
8. **Orders Over Time:**  A line graph showing order counts by date.
9. **Quarterly Orders:**  A stacked bar chart illustrating order volume per quarter in 2018.
10. **Customer Breakdown:** A heatmap showing orders grouped by customer names.

**Step 4: Dashboard Development and Presentation**

1. **Importing the Dataset into Power BI:**
   * Loaded the cleaned and processed data into Power BI ensuring data integrity.
2. **Developed Visualizations Using DAX:**
   * **Key Performance Indicators (KPIs):**
     + **Total Revenue:**
     + TotalRevenue = SUM(Details[Amount])
     + **Average Profit:**
     + AverageProfit = AVERAGE(Details[Profit])
     + **Year-over-Year Growth:**
     + YoYGrowth =
     + VAR PreviousYearRevenue = CALCULATE(SUM(Details[Amount]),
     + SAMEPERIODLASTYEAR('Date'[Date]))
     + RETURN
     + DIVIDE(SUM(Details[Amount]) - PreviousYearRevenue, PreviousYearRevenue, 0)
3. **Interactive Filters:**
   * Implemented slicers for Category, City, and Order Date to allow users to drill down.
4. **Visualizations Created:**
   * **Line Charts:** Showcased sales and profit trends over time.
   * **Bar Charts:** Compared revenues across different categories.
   * **Heatmaps:** Highlighted top-performing geographical regions.
5. **Final Presentation:**
   * Delivered a comprehensive, interactive dashboard, highlighting key trends and their implications for decision-making.