**Supply Chain Data Analysis Report**

**Description**

The dataset used in this analysis is from the DataCo Supply Chain. It contains various attributes related to orders, customers, products, shipping, and financial performance. The data includes customer details, product pricing, order statuses, shipping times, and financial metrics such as discounts, profit margins, and sales figures. The goal of the analysis was to clean, explore, and derive key performance indicators (KPIs) that can help in understanding the overall business operations and performance.

**Objective**

The primary objective of this analysis was to:

1. Clean and preprocess the dataset to handle missing or inconsistent data.
2. Calculate key performance indicators (KPIs) to measure the efficiency of the supply chain, financial performance, and customer satisfaction.
3. Provide actionable insights based on the KPIs to guide strategic business decisions.

**Steps Taken**

* **Database Setup**

Created a new database and set it up for data analysis.

* **Data Exploration**

We explored the data by selecting all the columns to understand its structure and identify any missing or null values.

* **Data Cleaning and Preprocessing**

1. **Checking for duplicate values**
2. **Handling Missing Values**

* Checked for null values in critical columns (e.g., Shipping\_Mode and others) and then filtered those values
* **Customer\_Lname** column: Replaced missing last names with the customer's first name.
* **Customer\_Zipcode** column: Updated missing zip codes based on Customer\_State and Customer\_City:
* **Order\_Zipcode** column: Missing values were replaced using an inference strategy based on matching Order\_City and Order\_State

1. **Dropping Useless Columns**

* Dropped columns with irrelevant or redundant data, such as Customer\_Email, Customer\_Password, Product\_Description, and Product\_Image.
* Dropped columns which have the same values:

(Benefit\_per\_order (dropped) - Order\_Profit\_Per\_Order (kept)) and

(Sales\_per\_customer (dropped) - Order\_Item\_Total (kept))

* **KPI Calculation**

1. **Order Accuracy Rate**

To measure the percentage of orders fulfilled without errors (order delivery).

**Result**: 44%

1. **On-time Delivery Rate (OTD)**

To measure the percentage of orders delivered on or before the scheduled shipping date.

**Result**: 42%

1. **Perfect Order Rate**

To measure the percentage of orders that were accurate, delivered on time, and without any issues.

**Result**: 18%

1. **Order Lead Time**

To measure the average time it takes to fulfill an order from the time of order placement to shipment (Internal process).

**Formula**: Shipping\_Date - Order\_Date

**Result**: average 3 days

1. **Order Cycle Time**

To measure the average time it takes to fulfill an order from the time of order placement to delivery.

**Formula**: (Shipping\_Date + Shipping\_Days)- Order\_Date

**Result**: average 6 days

* 1. **Cycle time with schedualed date**
  2. **Real vs Scheduled Shipping Days per Shipping Mode**
  3. **Real vs Scheduled Cycle time per Shipping mode**
  4. **Real vs Scheduled Cycle time per Customer City**
  5. **Late Orders**

1. **Important KPIs** like: Total Orders, Sales, Profit
   1. **KPIs per Type**
   2. **KPIs per Customer Segment**
   3. **KPIs per Department**
   4. **KPIs per Category Name**
   5. **KPIs per Region**
2. **Average Order Value (AOV)** To calculate the average revenue generated per order.

**Result:** $183.11

1. **Lost Sales:** the orders that have been cancelled.

**Result:** $668,244.99

1. **Return Rate** measures the percentage of orders that were returned.

**Result:** 35%

* 1. **Return Rate per Category Name**
  2. **Return Rate per Shipping Mode**
* **Dashboard design**
* **Results and Insights**
* By cleaning and preparing the data, we improved its usability for analysis and decision-making.
* KPIs like Order Accuracy Rate, On-time Delivery Rate, and Average Order Value provide a clear picture of supply chain efficiency.
* Handling missing data and removing redundant columns optimized the dataset.
* **Order Accuracy:** The high order accuracy rate reflects the efficiency of the company's order processing system. A high rate shows that the majority of orders are fulfilled without issues.
* **On-time Delivery:** While the on-time delivery rate was moderate, there is room for improvement. Delays in shipping impact customer satisfaction and can lead to returns or cancellations.
* **Order Cycle Time:** The average time to complete an order provides insight into the operational efficiency of the supply chain. Shorter cycle times are indicative of a smooth and efficient order process.
* **Return Rate:** The analysis revealed a notable return rate, especially in certain categories. This could be due to product quality issues or mismatches between customer expectations and delivered products.
* **Lost Sales:** Significant sales were lost due to order cancellations, representing an opportunity for improvement. Understanding why orders are canceled can help reduce these losses.
* **Recommendations**
* **Optimize Shipping Processes:**

Improve the on-time delivery rate by addressing bottlenecks in the shipping process. This can involve better supplier coordination or improving inventory management to prevent stockouts.

* **Reduce Order Cancellations:**

Investigate the reasons behind order cancellations and take steps to prevent them, such as improving customer communication and addressing common issues that lead to cancellations.

* **Improve Product Quality:**

Analyze the return rate by product category and take action on products with high return rates. This could involve working with suppliers to improve product quality or enhancing product descriptions to set clearer customer expectations.

* **Enhance Customer Segmentation Strategies:**

Use the insights from customer segments and regions to target high-performing segments with tailored marketing campaigns or special offers.