**Final Report**

**Project Objective**

The objective of this project is to analyze customer trends in an online retail setting to identify factors influencing purchase behavior and recommend strategies for improved sales performance.

This report outlines the steps taken using SQL for data exploration and manipulation, Python for data cleaning, analysis, and visualization, and Power BI for data visualization and storytelling.

**Steps**

1. **SQL: Data Exploration and Manipulation.**

The initial step involves using SQL to explore and manipulate the data to gain basic insights into the dataset.

**Key Insights:**

* The total number of orders, total sales, and average sales per order provide an overview of the sales performance.
* The number of unique customers indicates the customer base size.
* Summarizing sales by product line and country helps identify the best-performing products and key markets.

**2. Python: Data Cleaning, Analysis, and Visualization**

Data cleaning is crucial to ensure the accuracy of subsequent analysis.

* Data Cleaning and Transformation Step

1. Standardize date formats and change Columns Type.

df['ORDERDATE'] = pd.to\_datetime(df['ORDERDATE'], format='%m/%d/%Y %H:%M')

1. Filling NAN Values in Columns and Remove all unwanted columns

columns\_to\_delete = ['PHONE', 'ADDRESSLINE1', 'ADDRESSLINE2', 'POSTALCODE', 'CONTACTLASTNAME', 'CONTACTFIRSTNAME', 'STATE']

1. Mapping dictionary for product name standardization and Mapping Dictionary for Country to Territory (Products and Countries of same description)

**Key Insights:**

* Segment customers based on purchase frequency and total spending amount based on quantiles. (0.25,0.75 quantiles) to High Value Frequent, High Value Occasional, Low Value Frequent, Low Value Frequent.
* Segmenting Customers Based on Purchase Frequency and Total Spending using technique called RFM analysis employed, RFM stands for Recency, Frequency, and Monetary Value. Focusing on Frequency and Monetary Value.

After cleaning and analyzing the data in Python, export the cleaned dataset to a Excel file using Pandas

1. **Power BI: Data Visualization and Storytelling**

Using Power BI, creating interactive dashboards to showcase customer trends and insights.

**Analyzing Customer Trends and Insights**

**1. Sales Overview Page**

**Visualization**: Line charts, Pie Chart, Column Charts showing sales trends over different time periods (monthly, quarterly, yearly), Sales by Product line, Sales and profit over time, Sales by deal size, Map.

**Insights**:

* **Seasonal Trends**: The peaks and dips in sales are in Quarter 4, so the online retailer should adjust inventory and marketing strategies accordingly.

**Geographic Analysis:** Europe with high sales Volume & Asia with the lowest sales volume.

* **Implications**: Plan promotions and campaigns during peak buying periods to maximize revenue and manage inventory levels effectively during slower periods.

Customize marketing campaigns based on regional preferences and optimize shipping and logistics to improve customer satisfaction.

**2. Customer Insights**

**Visualization**: Using bar charts and column charts to segment customers based on key metrics such as Customer Segmentation Vs. Total Sales, Top 5 Customers, Customer name across Territories.

Using Scatter Chart showing purchase behavior (**High Value Frequent, High Value Occasional, Low Value Frequent, Low Value Frequent)** based onhigh CLTV and frequent purchases versus those with lower engagement.

* **Insights**: **High Value Frequent, High Value Occasional, Low Value Frequent, Low Value Frequent.** Showing segments with high CLTV and frequent purchases versus those with lower engagement.
* **Implications**: Tailor marketing efforts differently for each segment. High-value customers may respond well to loyalty programs, while low-value customers may need incentives to increase their spending.

Tailor customer service and marketing strategies to meet the unique needs and preferences of each segment, thereby enhancing customer retention and satisfaction.

**3. Products Performance**

**Visualization**: Tree maps to visualize sales distribution across different product lines (PRODUCTLINE) and territory, Sales Trends Over Time.

* **Insights**:Top-Selling Products: Classic Cars contribute the most to sales.
* **Implications**: Allocate resources to promote top-selling Classing Cars or optimize underperforming product lines which is Trains through targeted marketing or product improvements.

**Recommendations for Improve**

1. **Personalized Marketing Campaigns**:
   * Utilize insights from customer segmentation to create personalized offers and recommendations.
   * Implement targeted email campaigns and promotions based on customer preferences and past purchase behavior according to years (2003-2004-2005).
2. **Enhanced Product Offering**:
   * Focus on promoting best-selling products (Classic Cars) and optimizing inventory management.
   * Introduce new products or variants based on customer feedback and market demand trends.
3. **Optimized Customer Experience**:
   * Improve website usability, checkout process, and customer support based on insights into customer behavior.
4. **Data-Driven Decision Making**:
   * Continuously monitor key metrics and adjust strategies in real-time based on performance data.
   * Invest in analytics tools and training to empower teams to make informed decisions that drive sales growth.

By leveraging these visualizations and insights in Power BI, the online retailer can effectively analyze customer trends, identify factors influencing purchase behavior, and implement targeted strategies to improve sales performance and customer satisfaction. This approach ensures a data-driven approach to decision-making that is responsive to market dynamics and customer needs.