

Business Performance Analysis Through Data Exploration

1. Data Preparation:

- **Loading Data:** You loaded the datasets (Customers.csv, Products.csv, and Transactions.csv) into Pandas DataFrames.
- **Data Preprocessing:** Checked for missing values, explored data types, and reviewed basic statistics to understand the data structure.

2. Exploratory Data Analysis (EDA):

- **Sales Analysis by Region:**
 - Grouped data by Region and calculated the total sales for each region.
 - Visualized the total sales by region using a bar plot and represented the percentage contribution of each region with a pie chart.
- **Product Category Revenue and Order Size:**
 - Grouped data by Category to calculate the total revenue and average order size.
 - Visualized total revenue and average order size by product category using bar plots.
- **Customer Signup Trends:**
 - Converted SignupDate to a datetime format and extracted the year and month for trend analysis.
 - Grouped data by month to analyze the number of unique signups over time and visualized the trends with a line plot.
- **Top Customers by Revenue:**
 - Grouped data by Customer Name and summed the total revenue.
 - Identified the top 10 customers based on their total revenue and visualized this using a bar plot.
- **Repeat Purchases Analysis:**
 - Grouped data by ProductID and counted the number of transactions and unique customers per product.
 - Calculated repeat purchases and identified the top 10 products with the highest repeat purchases. This was visualized using a bar plot.

3. Key Insights:

1. **Sales by Region:** Identified the region with the highest sales and its percentage contribution to the total sales.
2. **Product Category Performance:** Determined which product categories generated the highest revenue and average order size.
3. **Signup Trends:** Analyzed customer signup trends over time to understand growth patterns.
4. **Top Customers:** Found the top 10 customers contributing the most revenue.

5. **Repeat Purchases:** Identified the top 10 products with the highest repeat purchases.

4. Visualizations:

- You used Matplotlib and Seaborn to create various plots, including:
 - Bar plots for total sales by region, product category revenue, and top customers.
 - Pie chart for the percentage contribution by region.
 - Line plot for customer signup trends over time.
 - Bar plot for repeat purchases by product.