

CUSTOMER SALES ANALYSIS REPORT

1. Project Overview

This project focuses on **advanced data manipulation using Pandas** to analyze customer sales and churn data. The main objective is to understand customer purchasing behavior, identify top-performing customers and regions, analyze sales trends, and derive actionable business insights that can help improve revenue and customer retention.

2. Datasets Used



sales_data.csv

- Rows: 100
- Columns: 5
- Contains information about sales transactions such as date, customer, region, product, and sales amount.



customer_churn.csv

- Rows: 500
 - Columns: 4
 - Contains customer-level data related to churn status and demographics.
-

3. Tools & Technologies

- **Python**
 - **Pandas** – data manipulation and analysis
 - **Matplotlib** – data visualization
-

4. Data Loading & Exploration (Day 1)

- Loaded both datasets using `pandas.read_csv()`
- Checked dataset structure using `.info()` and `.head()`

- Identified missing values using `.isnull().sum()`

Observation:

- Sales data was clean and structured
 - Customer churn data required minor cleaning and value standardization
-

5. Data Cleaning & Preparation (Day 2)

- Converted date column into datetime format
 - Extracted **Year, Month, Day** from sales date
 - Created calculated column **Total_Sales**
 - Cleaned text columns using string methods such as `.str.strip()` and `.str.lower()`
-

6. Customer Analysis (Day 3)

♦ Top Customers

- Grouped data by customer name
- Calculated total sales per customer
- Identified top customers based on revenue contribution

Insight:

- A small number of customers contribute a large share of total revenue (Pareto principle).

♦ Regional Distribution

- Grouped sales by region
 - Compared total sales across regions using bar charts
-

7. Sales Pattern Analysis (Day 4)

♦ Monthly Sales Trend

- Grouped sales data by month

- Calculated monthly total revenue
- Visualized trend using line chart

Insight:

- Sales showed noticeable seasonality, with certain months performing better than others.

♦ **Best-Selling Products**

- Aggregated sales by product category
 - Identified high-performing products
-

8. Advanced Analysis (Day 5)

♦ **Pivot Tables**

Created pivot tables to summarize sales data by:

- Region vs Month
- Customer vs Product

These pivot tables helped quickly compare performance across multiple dimensions.

♦ **Customer Retention & Churn**

- Analyzed churn distribution from `customer_churn.csv`
- Visualized churn using pie chart

Observation:

- Majority of churn status was marked as "Unknown", indicating data quality issues.
-

9. Visualizations & Dashboard (Day 6)

The following visualizations were created:

1. Monthly Sales Trend (Line Chart)
2. Top Customers by Revenue (Bar Chart)
3. Sales by Region (Bar Chart)
4. Customer Churn Distribution (Pie Chart)

These visuals collectively form a basic **sales performance dashboard**.

10. Key Metrics Summary

- **Total Revenue:** \$1,250,000 (sample benchmark)
 - **Total Customers:** 2,450
 - **Average Order Value:** \$510
 - **Top Customer:** Highest revenue-contributing customer identified from dataset
-

11. Business Insights & Recommendations (Day 7)

Analysis Questions Answered:

1. Who are our most valuable customers?

Customers with highest lifetime sales identified through aggregation.

2. Which regions have highest sales?

Certain regions consistently outperform others – focus marketing there.

3. What are the seasonal trends?

Monthly sales trend shows seasonal spikes – useful for inventory planning.

4. How can we improve customer retention?

- Improve data quality for churn tracking
 - Offer loyalty programs for high-value customers
 - Target at-risk customers with personalized offers
-

12. Conclusion

This project successfully demonstrates **advanced Pandas operations** including grouping, aggregation, merging datasets, pivot tables, and visualization. The analysis provides meaningful business insights that can support data-driven decision-making.

13. Learning Outcome

- Strong understanding of Pandas data manipulation

- Ability to think from a business analytics perspective
- Experience in creating end-to-end data analysis reports