Project 1: Retail Analytics - Customer Segmentation & Lifetime Value Engine

Project Duration

4 Weeks (End-to-End Analytics Project)

1. Executive Summary

This project focuses on building an end-to-end **Customer Segmentation and Churn Risk Analysis** system using the **RFM (Recency, Frequency, Monetary)** framework. The objective was to convert raw transactional data into actionable business insights, automate the analytics pipeline, and deliver executive-ready dashboards and reports.

The project was executed across four weeks covering data extraction, transformation, analytics, visualization, automation, and final business handoff.

2. Business Problem

Retail businesses often struggle to:

- Identify high-value customers
- Detect customers at risk of churn
- Prioritize marketing and retention efforts
- Automate recurring analytics processes

Without segmentation, marketing spends are inefficient and customer churn goes unnoticed until revenue is lost.

3. Dataset Overview

Source: Retail transactional CSV data

Key Columns Used:

- order_id
- order_date
- customer_name
- product_name
- category
- quantity
- price

The dataset represents historical purchase transactions across multiple customers.

4. Tools & Technologies Used

- **Python:** pandas, datetime
- **SQL:** Conceptual design & queries
- **Power BI:** Dashboard & RLS

- **Automation:** Windows Task Scheduler (ETL scheduling)
- **Version Control:** GitHub

5. Week 1 – Data Understanding & Preparation

Objectives

- Understand raw data
- Clean and prepare dataset
- Design analytical approach

Key Activities

- Imported CSV data into Python
- Converted date columns to datetime format
- Created revenue field (quantity × price)
- Validated data quality

Output

- Cleaned transaction-level dataset
- Ready for analytical processing

6. Week 2 – RFM Analysis & Customer Segmentation

Objectives

- Calculate RFM metrics
- Score customers
- Segment customers based on behavior

RFM Metrics

- **Recency:** Days since last purchase
- **Frequency:** Number of unique orders
- **Monetary:** Total revenue generated

Segmentation Logic

- High Value Customers
- Potential Customers
- Churn Risk Customers

Output Files

- rfm_final_output.csv
- churn_risk_customers.csv

7. Week 3 – Power BI Dashboard & Row Level Security

Objectives

- Build executive dashboards
- Enable secure data access

Dashboard Features

- KPI cards (Revenue, Customers, Avg Recency)
- Segment distribution charts
- RFM customer table
- Interactive slicers

Row Level Security (RLS)

- Implemented role-based data access
- Restricted data by city/region

Output

- Power BI dashboard (.pbix)

8. Week 4 – Automation, Handoff & Executive Reporting

Objectives

- Automate ETL pipeline
- Create churn-focused executive report
- Prepare project for business handoff

Automation

- Finalized Python ETL script
- Scheduled execution using Windows Task Scheduler
- Enabled repeatable analytics refresh

Executive Presentation

- Business problem & solution overview
- Key insights & churn risk findings
- Actionable recommendations
- Automation & scalability

9. Key Business Insights

- A significant portion of customers fall into the **Churn Risk** segment
- High-value customers contribute majority of revenue
- Recent and frequent buyers show strong loyalty potential

- Targeted retention campaigns can significantly reduce revenue loss

10. Business Recommendations

- Launch discount campaigns for churn-risk customers
- Create loyalty programs for high-value customers
- Automate weekly RFM refresh
- Use Power BI dashboards for ongoing monitoring

11. Project Impact

- Improved customer visibility
- Data-driven retention strategy
- Automated analytics pipeline
- Executive-ready reporting

12. Conclusion

This project demonstrates a full lifecycle analytics solution—from raw data to automated insights—using industry-standard tools and best practices. It is scalable, business-ready, and suitable for real-world deployment.