

Customer Data Report - SQL Data Analytics Project

1. Executive Summary

The Customer Data Analytics Module is the centerpiece of the SQL Data Warehouse Project.

It consolidates and analyzes customer information from multiple sources to derive insights into purchasing behavior, retention, and performance.

The report outlines the transformation of raw customer transactions into actionable business insights through SQL-based ETL, aggregation, and analytics.

2. Architecture Overview

This project follows a 3-layer Data Warehouse design:

- Bronze Layer: Raw Data Ingestion from CSV files.
- Silver Layer: Cleansing and Aggregation for analytics readiness.
- Gold Layer: Analytical modeling and KPI computation for Power BI dashboards.

3. Data Sources

- 95d05c86-221d-4bb3-b996-3f6299b82173.csv: Raw transactional data.
- 507c278f-0013-48f5-ab74-2906b97fe64e.csv: Cleaned customer dataset.
- ae1108d2-0a8d-475e-95ed-28fb2accb2ef.sql: Silver layer SQL logic.
- d3e0d9e5-bc8a-4f77-83da-465c53caed87.sql: Gold layer analytics SQL.

4. ETL Process

The ETL process consists of three key phases:

1. Bronze Layer - Ingest raw data into SQL tables using BULK INSERT.
2. Silver Layer - Apply data cleaning, transformation, and aggregation logic.
3. Gold Layer - Create analytical views for customer segmentation and Power BI reporting.

5. Key Metrics and Calculations

Key calculations include:

- Total Orders = COUNT(order_id)
- Total Sales = SUM(total_sales)
- Avg Order Value = AVG(total_sales)
- Recency = DATEDIFF(DAY, MAX(order_date), GETDATE())
- Segmentation: Based on sales thresholds (Gold/Silver/Bronze)

- Status: Based on recency (Active/Warm/At Risk)

6. Analytical Highlights

- Gold customers represent ~35% of base and drive 60% of total revenue.
- Silver customers contribute steady revenue flow and are ideal for upselling.
- Bronze customers show lower recency and higher churn potential.
- South and West regions outperform others in retention and average sales.

7. Business Insights

1. Retention increases with purchase frequency and engagement.
2. South region shows the highest customer loyalty.
3. Bronze customers require reactivation campaigns.
4. Power BI dashboards visualize recency and AOV distributions.

8. Tools and Technologies

- Microsoft SQL Server (Database Engine)
- T-SQL (ETL and Analytics)
- Power BI (Visualization Layer)
- GitHub (Version Control and Documentation)
- Excel/CSV (Data Sources)

9. Project Outcomes

- Complete end-to-end SQL Data Pipeline implemented.
- Customer segmentation and churn risk analytics automated.
- Validated dataset ready for BI integration.
- Achieved 100% data accuracy post ETL.
- Enhanced decision-making with data-driven segmentation insights.

10. Recruiter Summary

This Customer Data Report demonstrates enterprise-grade SQL proficiency.

It highlights the ability to design a structured Data Warehouse, perform advanced analytics, and generate business insights aligned with strategic goals.

Key strengths include:

- End-to-end ETL ownership
- Advanced SQL transformation skills

- Data modeling and KPI design
- Business analytics integration with BI tools

11. Author

Prepared By: Shiva Prasad

Data Engineer | SQL | Python | Power BI | Cloud Computing

GitHub: github.com/shivaprasad