SQL_sales_analysis_project

■ Retail Sales SQL Portfolio Project

An end-to-end SQL portfolio project demonstrating how raw retail sales data was cleaned, transformed, and analyzed to generate meaningful business insights.

Author: Pravin Kamble

Project Overview

This project simulates a real-world retail analytics workflow. Starting from raw sales data, SQL techniques were applied to clean, standardize, and prepare the dataset for business analysis.

Dataset Description

Table	Purpose
sales_store	Raw retail sales data
sales_s	Cleaned and analysis-ready data

Data Cleaning & Preparation

• Standardized purchase dates from multiple formats into YYYY-MM-DD.
• Corrected column data types for consistency and accuracy.
• Removed duplicate transactions using window functions.
• Handled missing and blank values by converting them to NULL.
• Standardized categorical values such as gender and payment mode.
• Created a clean analytics table (sales_s).

Business Questions Answered

• Top-selling products by quantity
• Most frequently cancelled products
• Peak purchase time of day
• Top spending customers
• Highest revenue product categories
• Cancellation and return rates
• Preferred payment modes
• Age group purchase behavior
• Monthly sales trends
• Gender-wise product preferences

Business Impact

Insights from this project help retail businesses optimize inventory planning, improve customer targeting, reduce cancellations, and make data-driven decisions.

SQL Skills Demonstrated

• Data Cleaning & Transformation
• CASE WHEN & Conditional Aggregation
• CTEs and Window Functions
• Date & Time Functions
• ROLLUP and Group-Based Analysis
• Business-Oriented SQL Queries

Conclusion

This document combines project overview, execution steps, and business insights into a single GitHub-ready PDF suitable for portfolios and interviews.