SWIGGY SALES ANALYSIS

BUSINESS REQUIREMENTS

Project Overview

This project aims to analyze Swiggy food delivery sales data using SQL. The objective is to clean raw transactional data, model it into an analytics-friendly structure, and derive meaningful business insights through KPIs and deep-dive analysis.

Data Source

The data source consists of historical Swiggy food delivery order records, with each record representing a single customer order.

Data Cleaning & Validation Requirements

The raw table (swiggy_data) is not analysis-ready and must undergo the following checks:

Null Checks

Identify missing values in:

- State
- City
- Order_Date
- Restaurant_Name
- Location
- Category
- Dish_Name
- Price_INR
- Rating
- Rating_Count

Blank / Empty String Checks

Detect text fields containing blank or empty values that may distort analysis results.

Duplicate Detection

Identify duplicate transactional records by grouping on all business-critical columns.

Duplicate Removal

Use ROW_NUMBER() to remove surplus duplicate rows while retaining one clean record per unique order.

Data Modelling Requirements (Star Schema)

To support efficient analytics and reporting, the cleaned data must be modelled using a Star Schema.

Dimension Tables

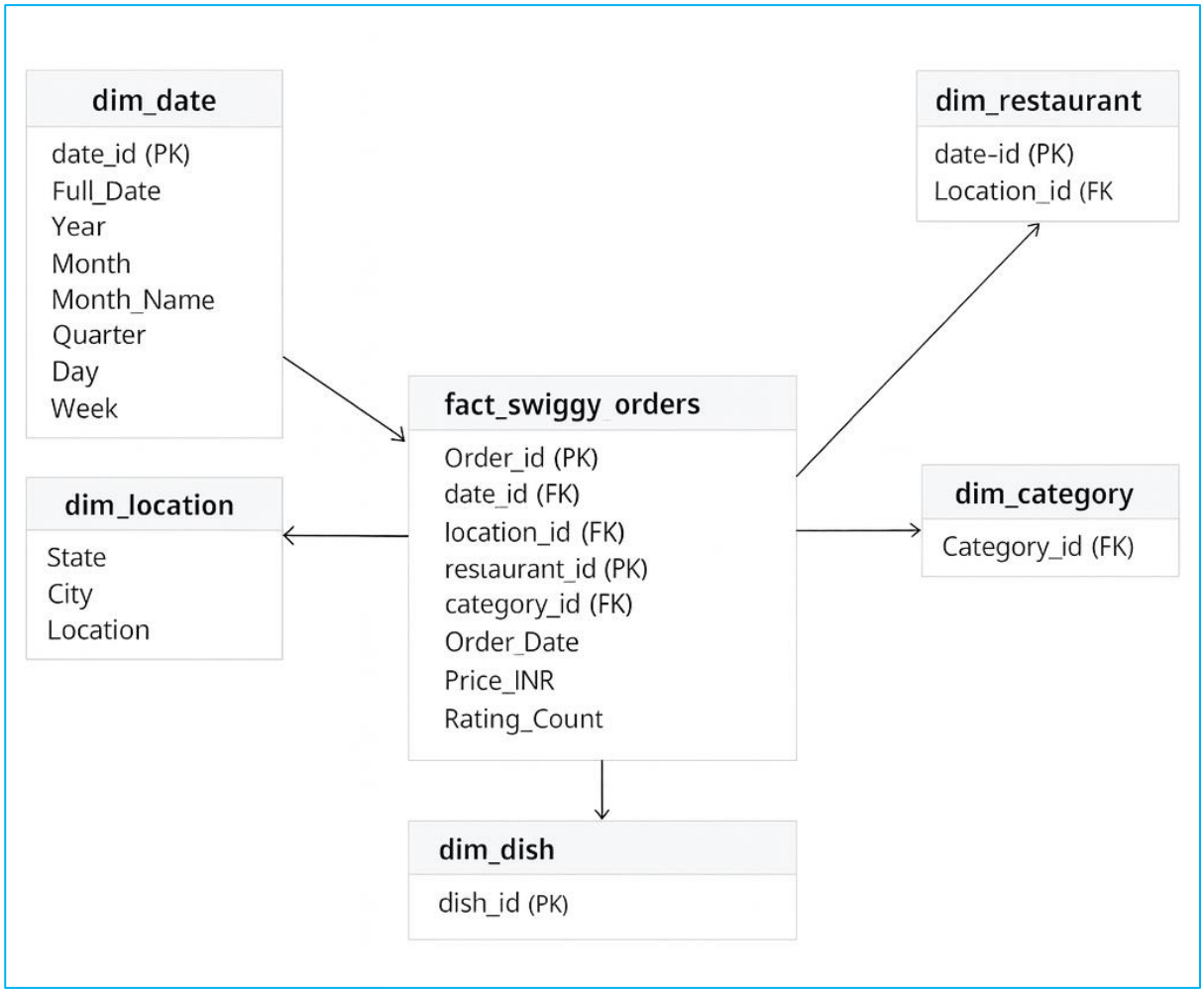
- dim_date: Year, Month, Month_Name, Quarter, Day, Week
- dim_location: State, City, Location
- dim_restaurant: Restaurant_Name
- dim_category: Category
- dim_dish: Dish_Name

Central Fact Table:

- fact_swiggy_orders: Price_INR, Rating, Rating_Count, foreign keys to all dimensions

Each dimension must be populated with distinct values from the cleaned dataset, and the fact table must resolve all foreign keys correctly.

ERD DIAGRAM: STAR SCHEMA



KPI Requirements

- Total Orders
- Total Revenue (reported in INR Millions)
- Average Dish Price
- Average Rating

Business Analysis Requirements

Date-Based Analysis

- Monthly order and revenue trends
- Quarterly order trends
- Year-wise growth analysis
- Day-of-week order patterns

Location-Based Analysis

- Top 10 cities by order volume
- Revenue contribution by states

Food & Cuisine Performance

- Top 10 restaurants by orders
- Top food categories
- Most ordered dishes
- Cuisine performance based on total orders and average rating

Customer Spending Insights

Analyze total order distribution across price ranges:

- Under 100
- 100–199
- 200–299
- 300–499
- 500+.

Ratings Analysis

Analyze distribution of dish ratings from 1 to 5.

Expected Outcome

The final solution should deliver a clean, structured dataset, reliable KPIs, and actionable insights into customer behaviour, food performance, and sales trends. This foundation should be suitable for SQL-based reporting and future BI dashboard development.