

Swiggy Food Delivery Data Analysis

End-to-End SQL Analytics Project

Comprehensive Business Intelligence & Data Analytics Report
transforming raw transactional data into actionable insights for food-tech decision-making.



Business Objectives



Data Standardization

Clean and prepare raw Swiggy order data for analytical readiness



Scalable Design

Build star schema for reporting and BI integration



KPI Tracking

Monitor orders, revenue, ratings, and pricing trends



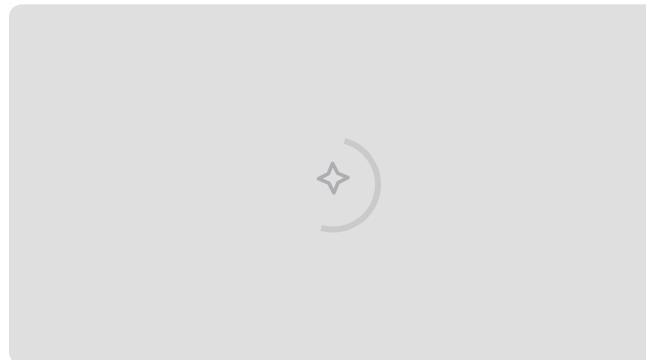
Strategic Insights

Enable decision-making for growth, operations, and marketing teams

Data Overview & Key Attributes

Dataset Source

Swiggy transactional dataset with one row per dish order, providing granular insights into customer behavior and restaurant performance.



Core Data Fields

- Location details (State, City, Area)
- Restaurant and cuisine information
- Dish-level pricing
- Ratings and rating counts
- Order date timestamps

Data Preparation & Cleaning

01

Column Standardization

Converted all column names to snake_case (e.g., Order Date → order_date) for SQL best practices and BI tool compatibility.

02

Null Value Analysis

Comprehensive audit using conditional aggregation to identify missing values across business-critical fields like location, pricing, and ratings.

03

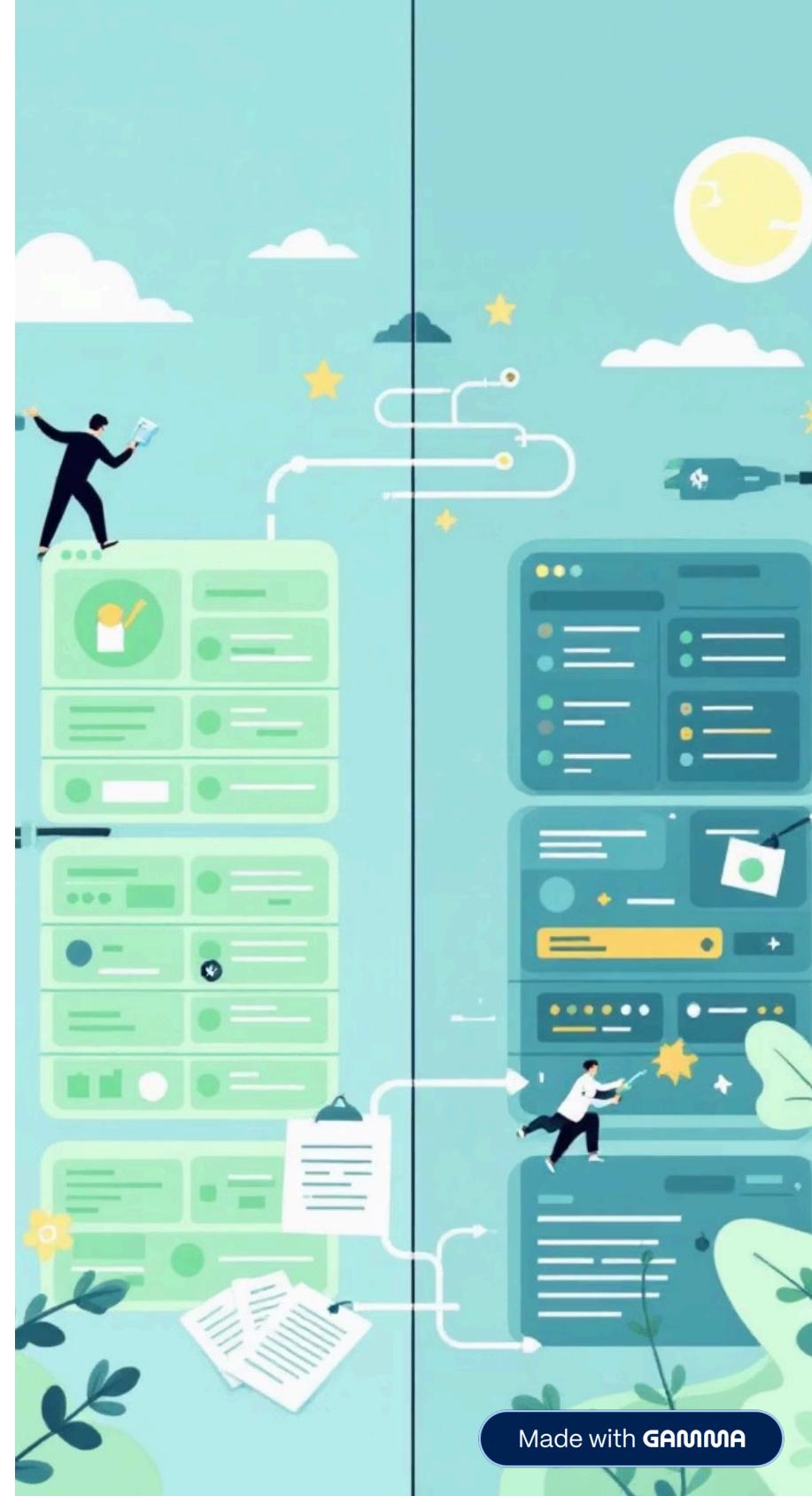
Duplicate Detection

Identified duplicates using composite keys across location, restaurant, dish, date, price, and rating fields.

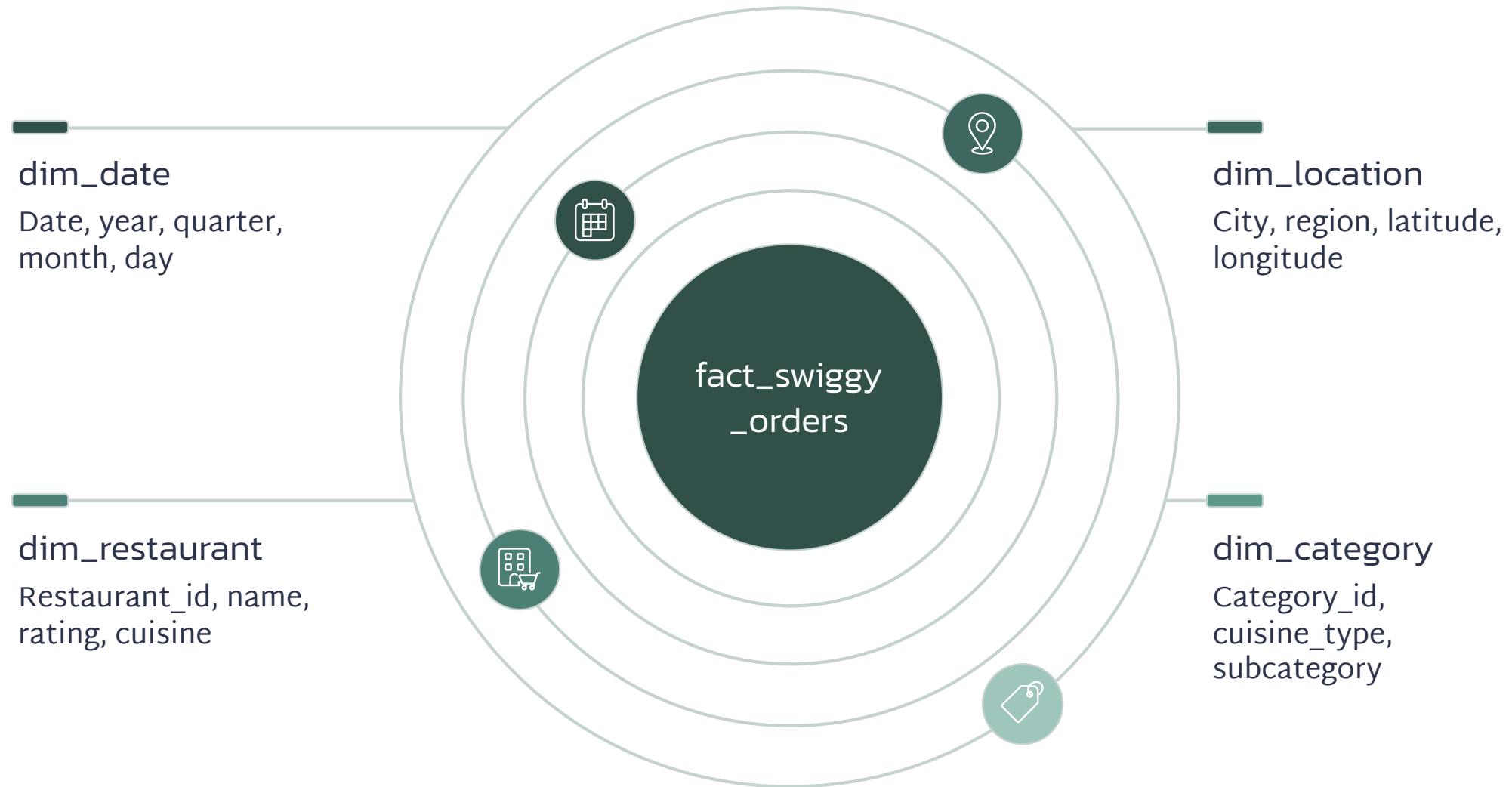
04

Data Removal

Removed duplicates using ROW_NUMBER() window functions and Python execution, ensuring one unique order per transaction.



Star Schema Design



Transformed dataset into a star schema supporting fast aggregations, clean joins, and Power BI/Tableau readiness for analytical queries and BI dashboards.

Key Performance Indicators

1.2M

₹450M

₹215

Total Orders

Platform demand and transaction volume

Total Revenue

Topline performance at scale

Avg Dish Price

Customer spending behavior

4.2

Avg Rating

Overall customer satisfaction

These executive-level KPIs provide a comprehensive performance snapshot for stakeholders to track business health and growth.



Time-Based Analysis

- 1 Monthly Trends
Order volume and revenue analyzed month-wise to identify seasonal demand patterns for campaign planning and capacity forecasting.
- 2 Quarterly Trends
Business growth phases highlighted to support strategic quarterly planning and identify slowdown periods.
- 3 Day-of-Week
Peak ordering days identified (weekends vs weekdays) for promotions, discounts, and staffing decisions.



Location & Restaurant Performance

Geographic Insights

- **Top Cities by Orders:** Reveals demand concentration in high-volume urban markets
- **Top Cities by Revenue:** Differentiates volume vs value cities for premium expansion
- **State-Level Revenue:** Macro-level performance for regional strategy



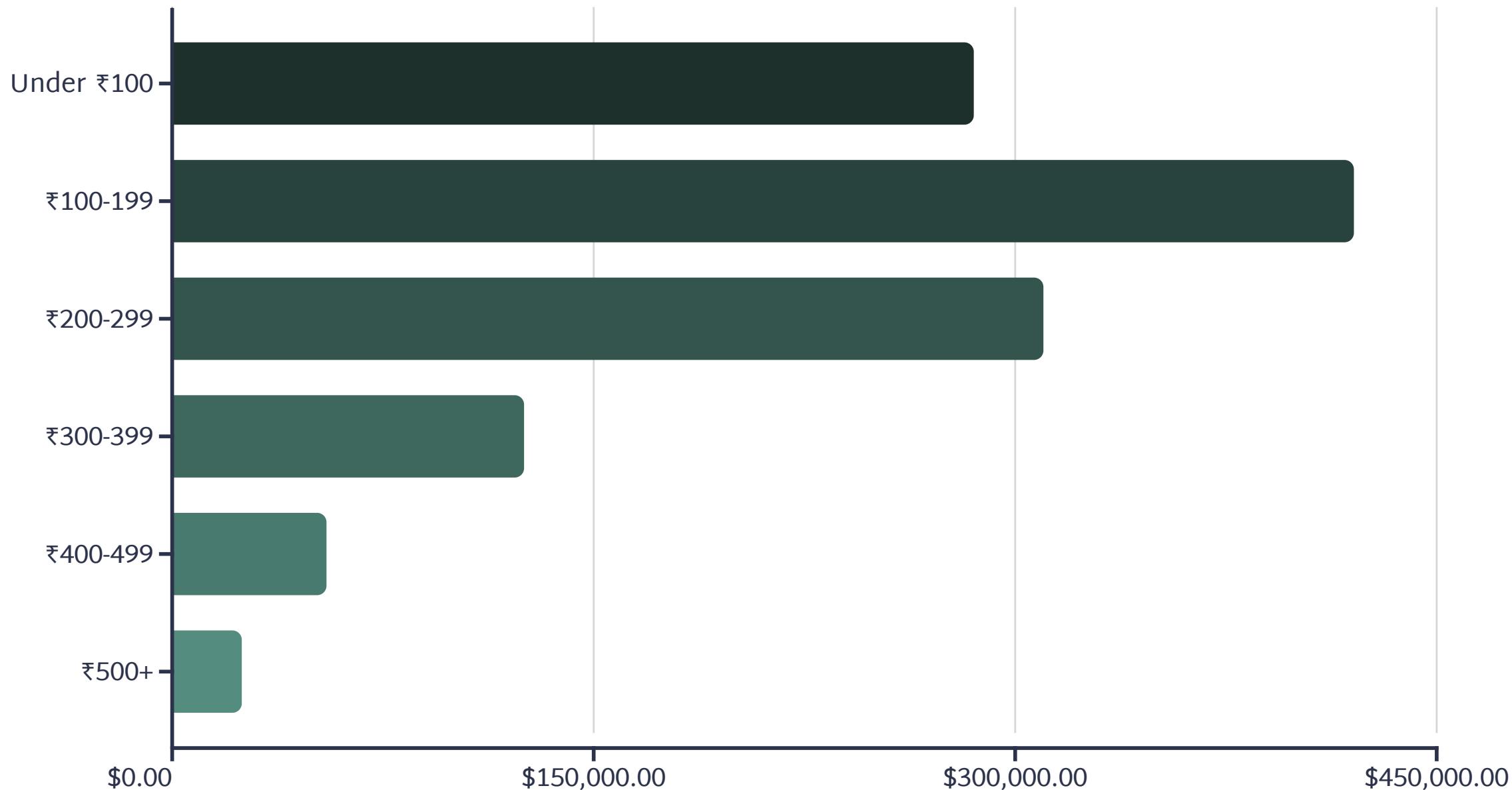
Restaurant Analysis

Top restaurants ranked by order volume and total revenue to identify high-performing partners.

Category Performance: Top cuisines by order count with combined volume and average rating analysis for cuisine-level promotions.

Most Ordered Dishes: Customer favorites pinpointed for menu optimization and bundling strategies.

Customer Spending Insights



Price segmentation reveals mass-market vs premium customer segments and optimal pricing bands for maximum volume.
Rating distribution analysis supports restaurant quality monitoring.

Tools & Key Takeaways



SQL (MySQL)

Data cleaning, modeling, and comprehensive analysis



JupyterLab

Duplicate handling at scale with Python



Star Schema

Dimensional data modeling for BI readiness

This project demonstrates strong proficiency in SQL querying, data quality management, dimensional modeling, and business-oriented analytical thinking—mirroring real-world analytics workflows in food-tech and e-commerce organizations.

