Zomato Delivery Operations Analysis Report

# 1. Introduction

With the rising demand for online food delivery services, companies like Zomato face increasing pressure to optimize logistics while maintaining high service standards. This project analyzes Zomato's delivery dataset, which contains detailed operational data including delivery timings, road and weather conditions, and performance metrics. The goal is to uncover operational inefficiencies, enhance route planning, and improve overall delivery performance through data analysis and machine learning techniques.

# 2. Purpose

The purpose of this project is to evaluate the key factors affecting delivery time and customer experience in Zomato’s delivery operations. By applying data analytics and predictive modeling, the project seeks to understand how internal and external variables—such as traffic, vehicle condition, and delivery personnel ratings—influence delivery outcomes. These insights will support operational decisions and strategic improvements in efficiency and service quality.

# 3. Expected Outcomes

This project aims to deliver a comprehensive, data-driven assessment of Zomato's delivery operations. Key outcomes include:

- Identifying variables most strongly correlated with delivery delays.

- Formulating strategies for peak hours, festivals, multiple deliveries, and delivery staff performance.

- Recommendations to improve delivery quality, customer experience, and logistics efficiency.

# 4. Dataset Information

Source: Zomato Delivery Dataset (Kaggle)

Link: https://www.kaggle.com/datasets/saurabhbadole/zomato-delivery-operations-analytics-dataset

The dataset includes 56,634 delivery records with attributes categorized into five key areas:

4.1 Delivery Staff

- Delivery\_person\_ID: Unique identifier of the delivery personnel.

- Delivery\_person\_Age: Age of the delivery personnel.

- Delivery\_person\_Ratings: Average rating given to the delivery person.

4.2 Vehicles

- Vehicle\_condition: Condition score of the delivery vehicle (0 to 2).

- Type\_of\_vehicle: Type of vehicle used (motorcycle, scooter, etc.).

- Multiple\_deliveries: Number of deliveries made in the same trip.

4.3 Order and Timing

- Order\_Date, Time\_Ordered, Time\_Order\_picked: Key timestamps.

- Type\_of\_order: Type of food order (Meal, Snack, Drinks, Buffet).

- Festival: Indicates whether the delivery was made during a festival.

4.4 Weather and Traffic Conditions

- Weather\_conditions: Weather at the time of delivery (Sunny, Fog, Stormy, etc.).

- Road\_traffic\_density: Traffic levels (Low, Medium, High, Jam).

4.5 Area

- Restaurant\_latitude/longitude, Delivery\_location\_latitude/longitude: GPS coordinates.

- City: Area type (Metropolitan, Urban, Semi-Urban).

# 5. Tools & Technologies

- Python: Pandas, Matplotlib, Seaborn for data processing and analysis.

- Jupyter Notebook: For scripting and prototyping.

- Power BI: For interactive data visualization and dashboard reporting.

# 6. Analysis Plan

6.1 Data Preparation & Cleaning

- Fix data types (e.g., convert dates and times properly).

- Handle missing values via imputation or removal.

- Clean inconsistencies in categorical variables and string formatting.

6.2 Exploratory Data Analysis (EDA)

- Measures of Central Tendency: Custom functions used to analyze mean, median, mode, min, and max delivery times by different dimensions.

- Histogram: Plotted distribution of delivery times by number of deliveries.

- Density Plot: Analyzed distribution for age, rating, and time ordered.

- Box Plot: Compared delivery time by festival and type of order.

- Regression: Identify factors that increase delivery time.

6.3 Visualization and Reporting

- Key metrics visualized using Power BI dashboards.

- Interactive filters for city, delivery type, traffic, and time slots.

- Custom icons used for visual storytelling (e.g., ⏱, ⭐).

# 7. Conclusions and Insights

Delivery Staff

- Age: Delivery personnel aged 28 and below tend to deliver faster.

- Rating: Higher ratings are associated with shorter delivery times.

Vehicles

- Condition: Poor vehicle condition slows delivery. Regular maintenance is recommended.

- Multiple Deliveries: Increases delivery time; routing must be adjusted accordingly.

Order & Timing

- Time Ordered: Orders placed between 5 PM–10 PM experience significant delays. More delivery staff should be allocated in this window.

- Festival: Deliveries during festivals take longer. Festival-specific planning is essential.

Weather & Traffic

- Weather: Best delivery times observed in clear/sunny weather.

- Traffic: High traffic or jams increase time taken. Routing algorithms should avoid high-density areas.

Area

- City: Semi-Urban areas show the longest delivery times. Operational or logistic improvements are needed for these locations.

**Top yếu tố làm tăng thời gian giao hàng (dựa theo hệ số hồi quy):**

| **Biến** | **Hệ số ảnh hưởng** | **Giải thích** |
| --- | --- | --- |
| Delivery\_location\_longitude | **48.46** | Khoảng cách địa lý xa hơn (toạ độ longitude lớn hơn) → tăng thời gian giao hàng. |
| City\_Semi-Urban | **11.31** | Giao hàng tại vùng bán đô thị khiến thời gian giao hàng cao hơn đáng kể. |
| Festival\_Yes | **10.60** | Giao hàng trong dịp lễ tốn thời gian hơn do tắc nghẽn, thiếu nhân lực. |
| multiple\_deliveries | **3.14** | Mỗi lần thêm 1 đơn hàng trong cùng chuyến đi làm tăng thời gian. |
| Delivery\_person\_Age | **0.38** | Tuổi lớn hơn liên quan đến tốc độ giao hàng chậm hơn một chút. |
| Road\_traffic\_density\_Jam | **0.26** | Khi giao hàng trong điều kiện "Jam" (kẹt xe), thời gian tăng lên đáng kể. |