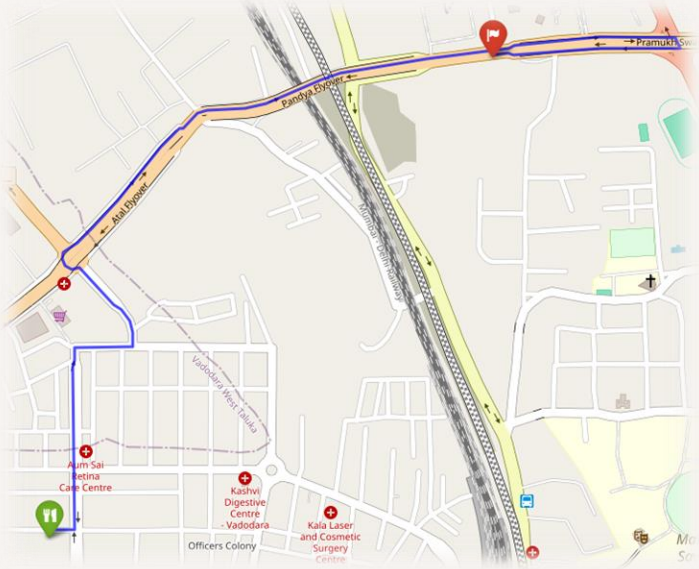
SmartDelivery: Optimizing Zomato's Operations with Predictive Analytics & Route Efficiency

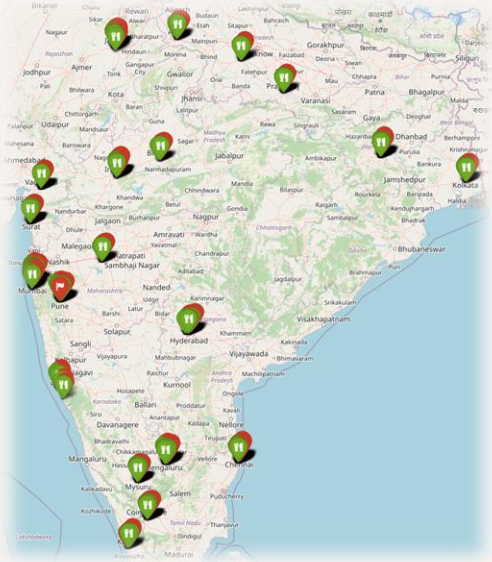


MAIN PROJECT OUTLINE

Project Background

Zomato's delivery arm in India faced rising order volumes, high customer expectations, and growing costs, with inconsistent delivery times and poor routing hurting satisfaction and inflating expenses.

To stay ahead, Zomato needs a holistic, data-driven framework that anticipates demand, measures real-world performance, and plots the most efficient paths for riders.



Disclaimer: *The following story is fictional and created solely to illustrate the challenges addressed by this project.*

Business Objective

📊 Predictive Analytics

- ❑ Develop a machine learning model to estimate delivery times.
- ❑ Identify key factors influencing delivery times using feature importance to provide better operational insights.

🗺️ Route Optimization

- ❑ Utilize the OSRM (Open Source Routing Machine) API to retrieve precise distances, estimated durations, and optimal delivery routes.

📋 Delivery Performance Analysis

- ❑ Identify underperforming drivers by analyzing their actual delivery times, ratings, average speed, vehicle condition, and total completed deliveries.

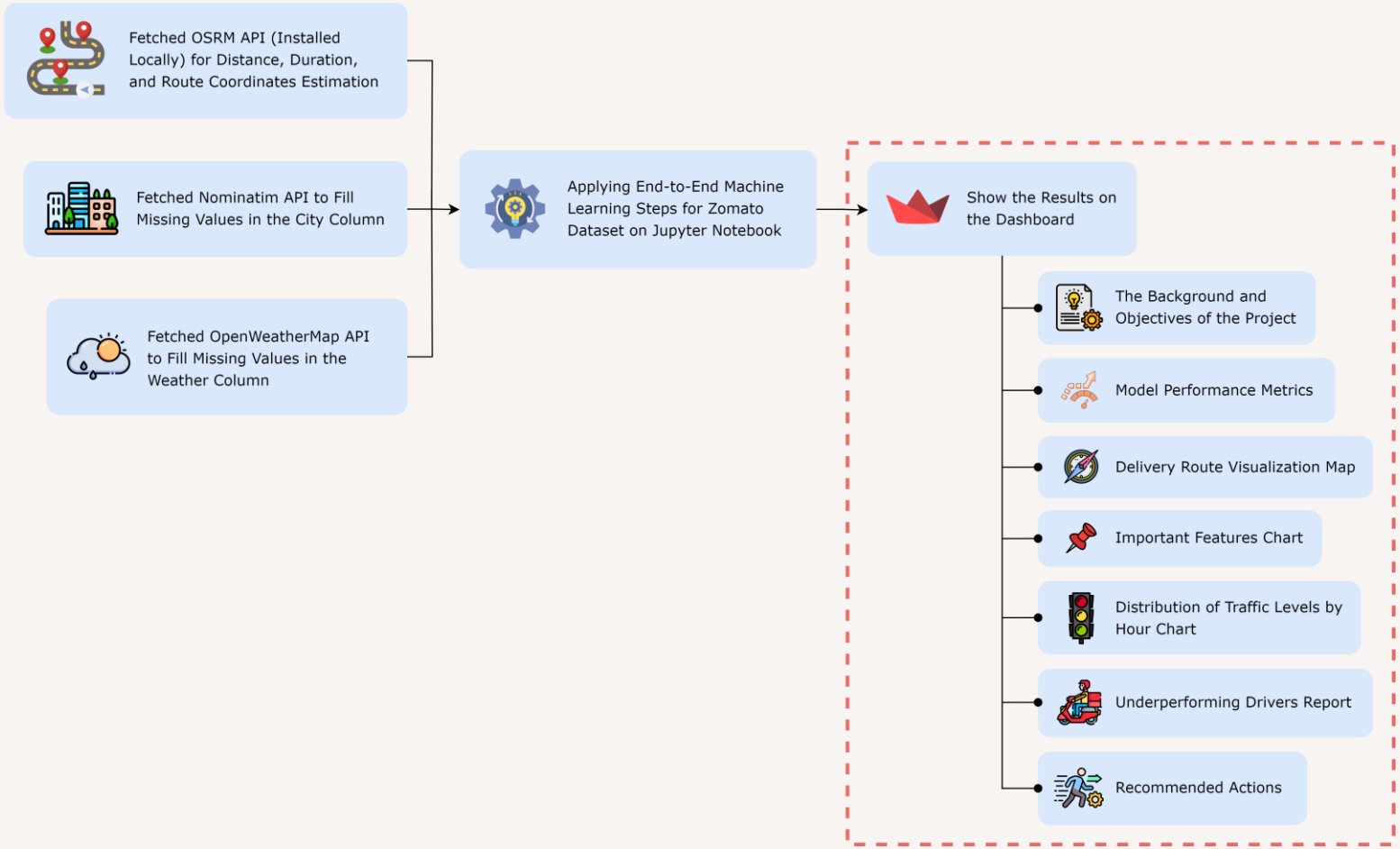
Data Understanding

Collected from Kaggle – Zomato Delivery Operations Analytics Dataset.

Approximately 46,000 orders from Feb-Apr 2022 across India.

Includes order details, driver profile and performance information, as well as delivery and restaurant locations.

Project Overview



Data Preprocessing

Step 1

Fetched OSRM API (Local)

Fetched distance, duration, and route coordinates

Step 2

Fetched OpenStreetMap API

Filled missing values in the City column

Step 3

Fetched Weather API

Filled missing values in the Weather column

Data Preprocessing

Step 4

Remove Anomalous Location

Removed delivery or restaurant locations in the ocean

Step 5

Convert Data Types

Reformatted ‘Order_Date’ to date type

Step 6

Address Outliers

Removed speeds above 150 km/h or below 3 km/h

Data Preprocessing Result

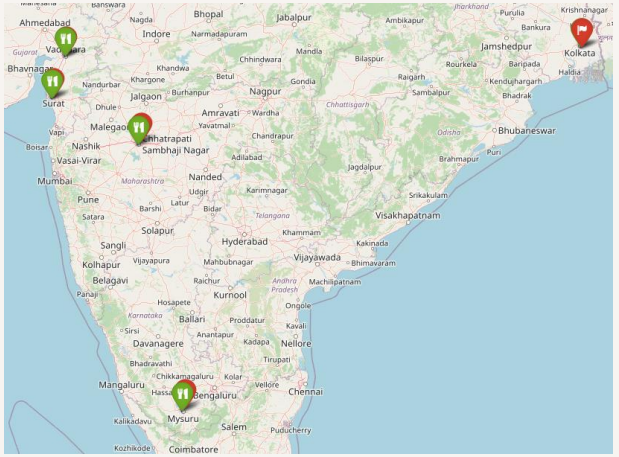
Findings and Results

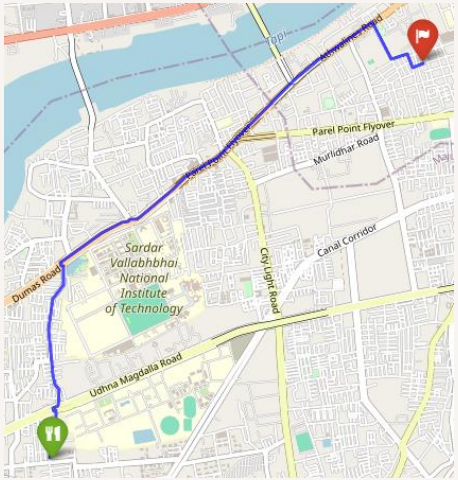
Insight

- The actual time taken by drivers from the restaurant to the delivery location was successfully predicted using an XGBoost machine learning model, achieving an average prediction error (RMSE) of approximately 4.29 minutes. This is a strong result, considering the standard deviation of the actual delivery times is around 9 minutes.

- Additionally, I generated a projected best route using the OSRM API and visualized it on a Folium map. This allows us to better understand the driver's actual route and provides more accurate estimates of both distance and travel time.

Machine Learning Model	Avg. Prediction Error (RMSE)
Random Forest	4.33 minutes
XGBoost	4.29 minutes
LightGBM	4.30 minutes

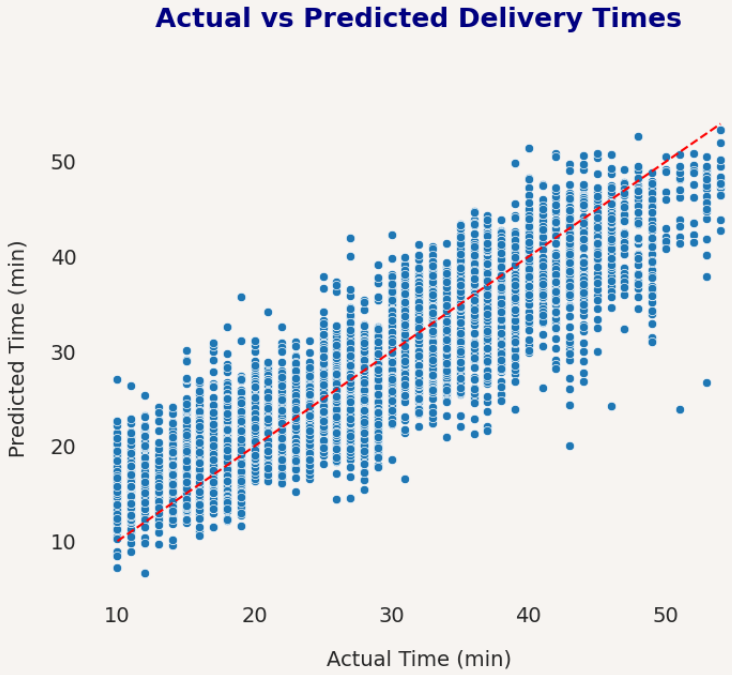




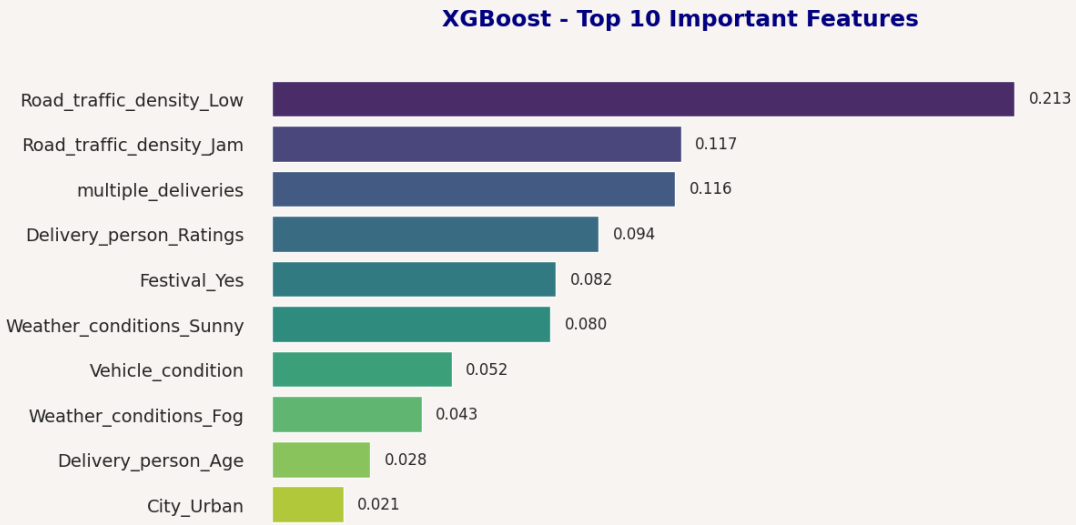
Findings and Results

Insight

- The graph comparing actual and predicted delivery times shows a strong linear relationship, closely following the line y = x.
- The red line represents this ideal relationship, indicating that, on average, the model's predicted delivery times accurately reflect the actual delivery times.



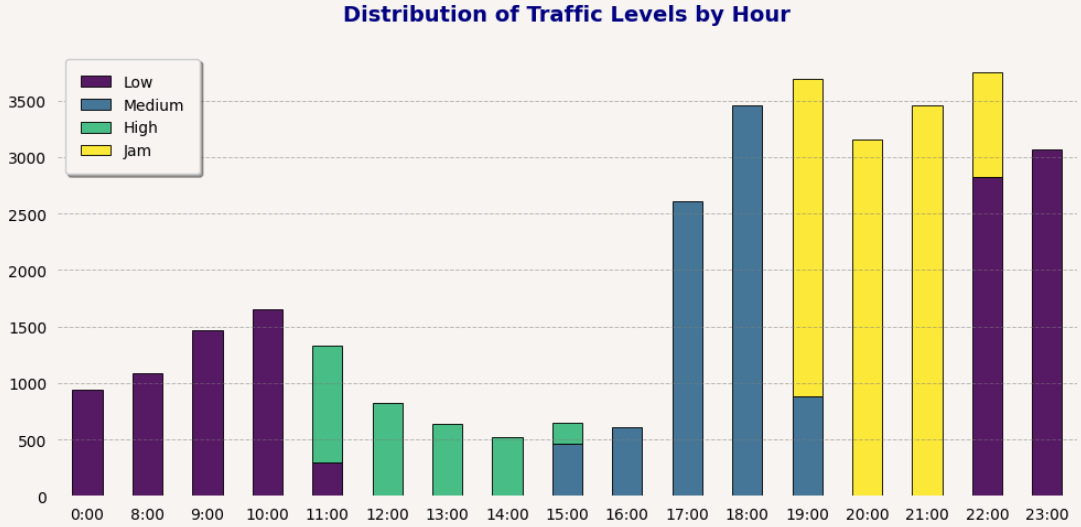
Findings and Results



Insight

- The top three most impactful factors influencing the actual time taken from the driver to the customer are **Road_traffic_density_low**, **Road_traffic_density_jam**, and **Multiple_deliveries**.
- This suggests that low road traffic density helps reduce delivery time, while traffic jams significantly increase it. Additionally, assigning multiple deliveries to a single driver also contributes to longer delivery times, as demonstrated by the **partial dependence plot** in the Jupyter Notebook.

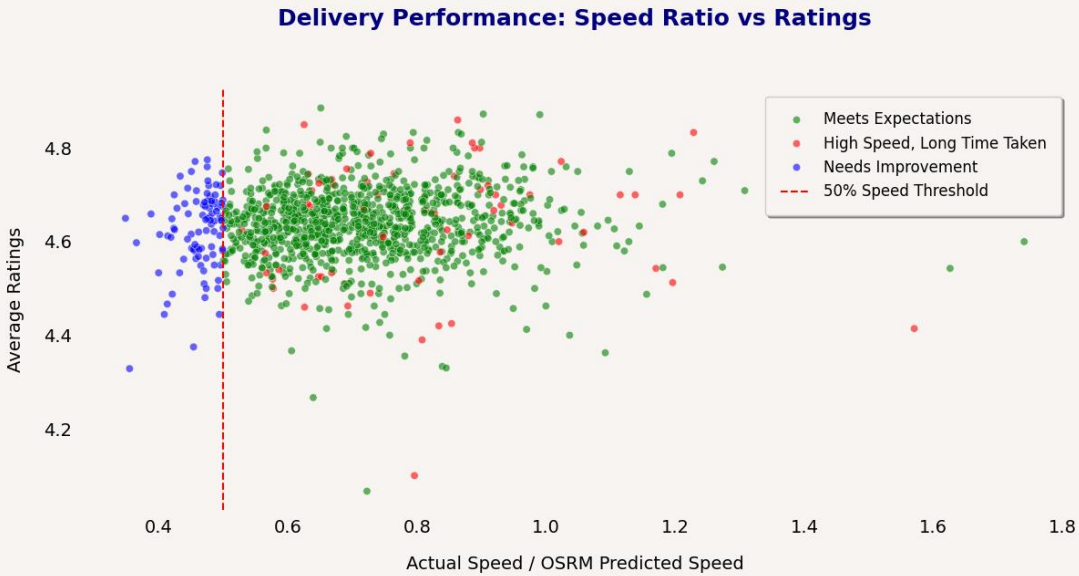
Findings and Results



Insight

- Traffic density trends:
  - Low: Late night to morning (22:16–11:15).
  - High: Midday (11:16–15:15).
  - Medium: Evening rush (15:16–19:15).
  - Jam: Night (19:16–22:15).
- Most orders occur between 17:00–23:00, aligning with peak traffic hours.

Findings and Results



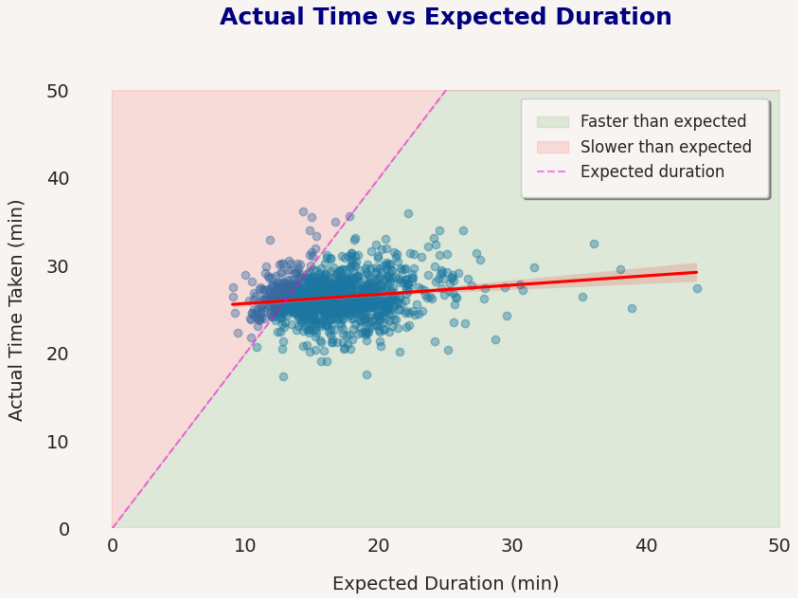
Insight

- OSRM parameters for the car profile were used as the baseline to assess driver performance.
- Since drivers may use a variety of vehicles (Motorcycle, Bicycle, Electric Scooter, and Scooter), and one driver can switch between vehicle types, I used a speed ratio threshold of below 0.5 to identify underperforming drivers.
- Additionally, drivers were also considered underperforming if their actual travel time deviated by more than 2 standard deviations (approximately ±4 minutes) from the mean. For example, if the average actual time is 26 minutes, drivers with times above 30 minutes or below 22 minutes were flagged.

Findings and Results

Insight

- Similar to the previous slide, this shows the expected travel duration from the OSRM API using the car profile as the baseline. I flagged drivers as slower than expected if their actual travel time was less than 50% of the OSRM expected duration.
- Based on this threshold, most drivers (in the green-shaded area) were faster than expected, while a few (in the red-shaded area) were slower than expected.



Findings and Results

Delivery_person_ID	Delivery_person_Ratings	Time_taken (min)	speed_ratio	Vehicle_condition
AGRRES03DEL02	4.70	17.62	1.14	1.25
AGRRES04DEL03	4.59	30.29	0.47	0.86
AGRRES05DEL01	4.67	24.00	0.48	1.14
AGRRES05DEL02	4.58	28.60	0.46	0.70
AGRRES08DEL02	4.67	34.00	0.78	0.44
AGRRES17DEL01	4.75	35.50	0.45	0.30
AGRRES20DEL03	4.54	20.14	1.17	1.00
ALHRES03DEL01	4.51	30.20	0.47	1.10
ALHRES03DEL02	4.74	31.22	0.63	0.44
ALHRES07DEL01	4.55	26.92	0.50	0.92

Summary
Underperforming Drivers (157/1170)
Average Speed Ratio: 0.71
% Drivers Below Speed Threshold (Speed Ratio < 0.5): 7.9%
Average Actual Time Taken: 26 minutes
Average Expected Duration: 17 minutes
Average Vehicle Condition: 1.0 out of 3
Average Rating: 4.6 out of 6

Insight

- This report summarizes the **underperforming drivers**. I identified underperforming drivers based on the following criteria:
  - Rating below 0.35
  - Speed ratio below 0.5
  - Actual travel time deviating by more than 2 standard deviations from the mean
  - Poor vehicle condition (a score greater than 2 on a 0–3 scale, where 0 indicates good condition)

Recommendations

- **Optimize Delivery Routes**: Utilize the OSRM API route projections to design optimized delivery routes, reducing delays caused by traffic and multiple deliveries.
- **Optimize the Schedule**: Schedule deliveries during low-traffic periods (late night to morning) and adjust staffing levels to meet the high demand during peak order times (17:00–23:00).
- **Focus on Driver Performance Improvement**: Provide targeted training or guidance for underperforming drivers.

Dashboard

I have provided a dashboard for relevant stakeholders to monitor delivery time predictions, routes, and driver performance.

Link to the Dashboard: Streamlit

SmartDelivery Analytics

Home

Overview

Dashboard

Delivery Time

Contact

Made with ❤ by Jihad Akbar

SmartDelivery: Optimizing Zomato's Operations with Predictive Analytics & Route Efficiency

Welcome to SmartDelivery!

***Disclaimer**: The following story is fictional and created solely to illustrate the challenges addressed by this project.*

🎬 The Background: Navigating Complexity in Indian Food Delivery

Zomato's delivery network in India faced escalating challenges despite booming order volumes:
⚡ **Skyrocketing Expectations**: Customers demanded faster, more reliable deliveries.
💰 **Rising Costs**: Inefficient routing and delivery delays inflated operational expenses.
🛵 **Inconsistent Performance**: Poor routing led to delays, dissatisfied customers, and mounting inefficiencies.

To maintain its competitive edge, Zomato needed a data-driven transformation—a comprehensive framework to predict delivery times, optimize routes, and analyze delivery performance in order to streamline operations and exceed customer expectations.

Zomato's Driver Illustration, created by ChatGPT

🎯 The Mission: Smarter Delivery, Happier Customers

This project aims to empower Zomato with actionable insights and cutting-edge technology to tackle these challenges head-on. This mission includes:
📊 **Predictive Analytics**: Build a machine learning model to accurately forecast delivery times.
🗺️ **Route Optimization**: Leverage the OSRM (Open Source Routing Machine) API to calculate precise distances, estimated delivery durations, and the most efficient routes for riders.
📋 **Delivery Performance Analysis**: Identify underperforming drivers by analyzing factors such as actual delivery times, ratings, average speeds, vehicle conditions, and total completed deliveries.

This project merges predictive analytics, route optimization, and performance evaluation to enhance Zomato's delivery operations, delighting customers while reducing costs.

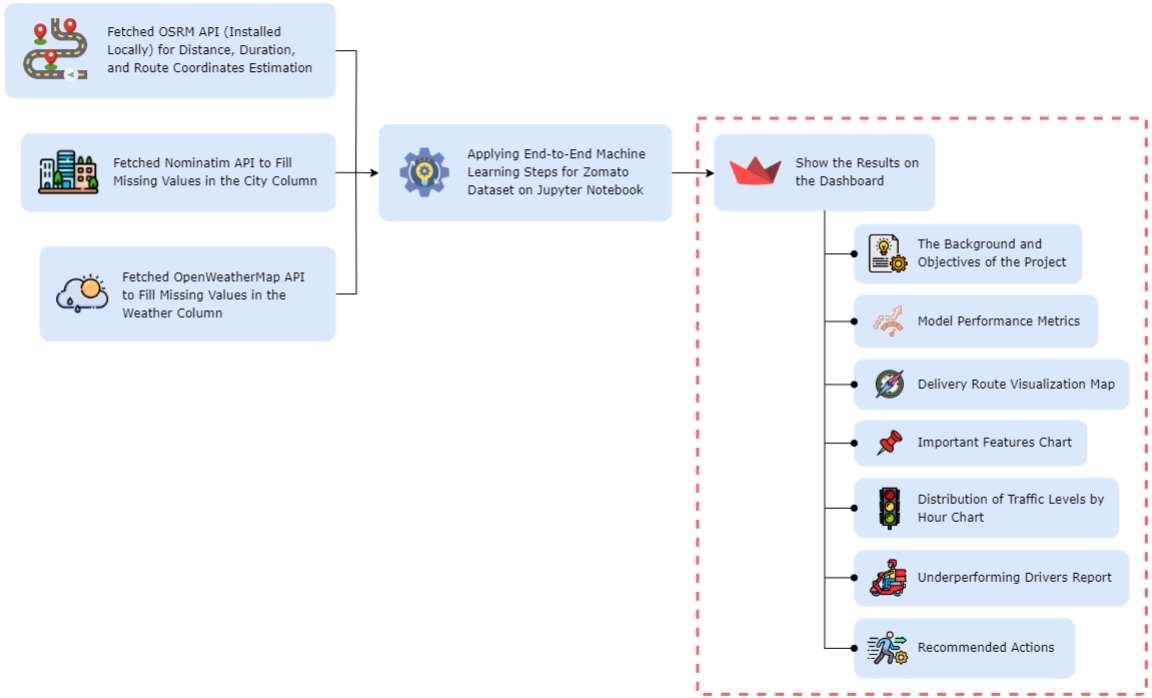
It was powered by the Zomato Delivery Operations Analytics dataset, comprising over 46,000 orders from various regions of India, spanning February to April 2022, sourced from Kaggle.

SmartDelivery Analytics

Home
Overview
Dashboard
Delivery Time
Contact

Made with ❤ by Jihad Akbar

🔍 Project Overview



SmartDelivery Analytics

Home

Overview

Dashboard

Delivery Time

Contact

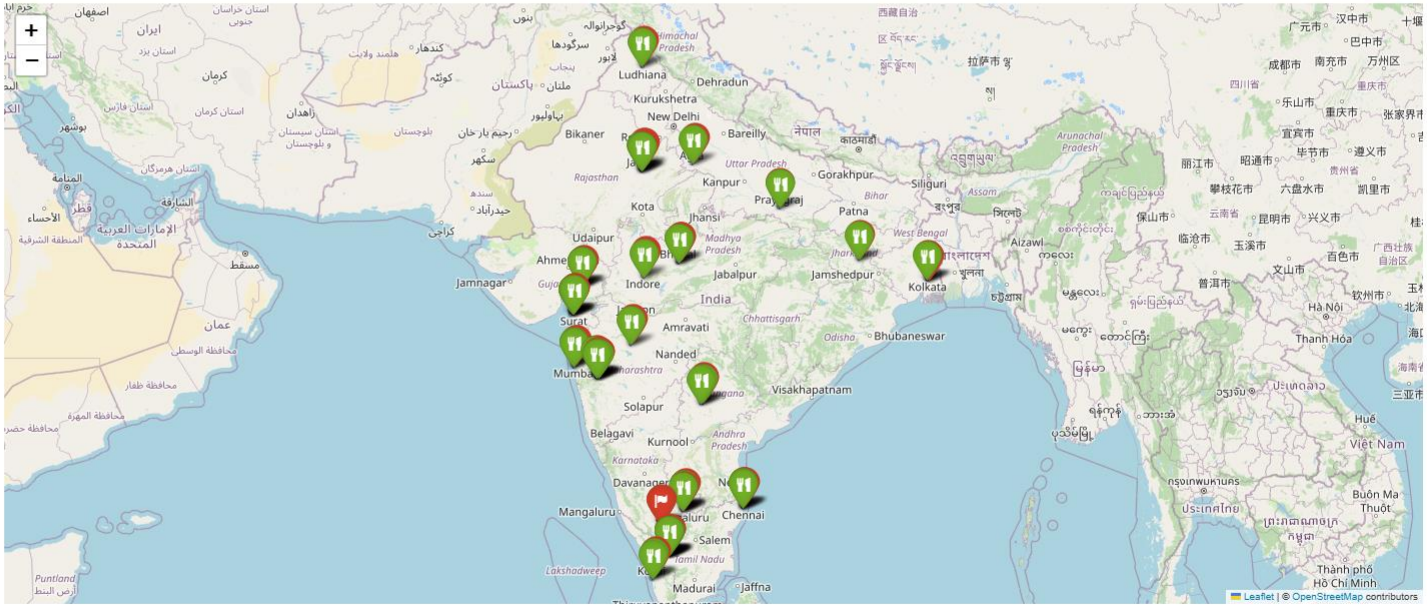
Made with ❤ by Jihad Akbar

SmartDelivery Dashboard

Machine Learning Model

Model	Deviation (RMSE)	Interpretation
XGBoost	±4.29 minutes	On average, predictions are off by 4.29 minutes.

Delivery Routes Visualization



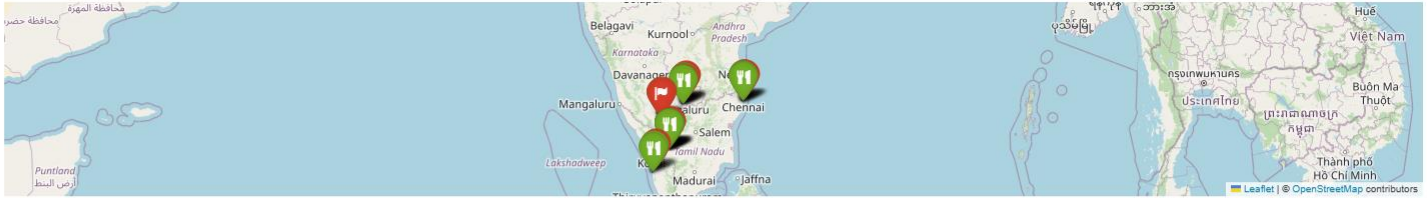
XGBoost - Top 10 Important Features

Share

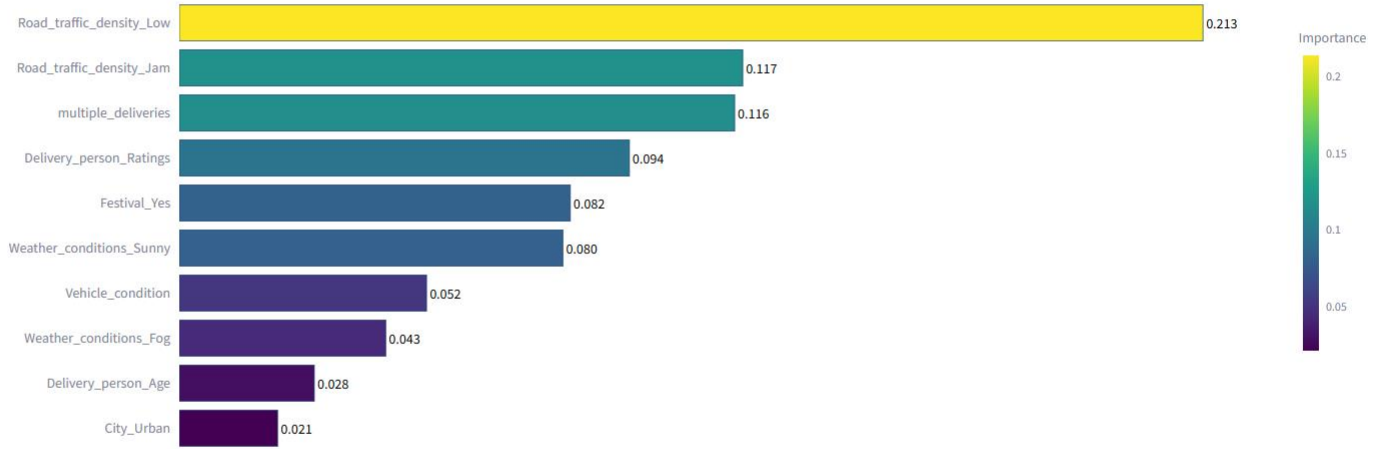
SmartDelivery Analytics

- Home
- Overview
- Dashboard
- Delivery Time
- Contact

Made with ❤ by Jihad Akbar



XGBoost - Top 10 Important Features



Traffic Levels by Hour

Manage app

SmartDelivery Analytics

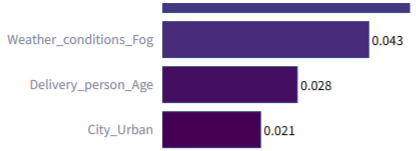
Home

Overview

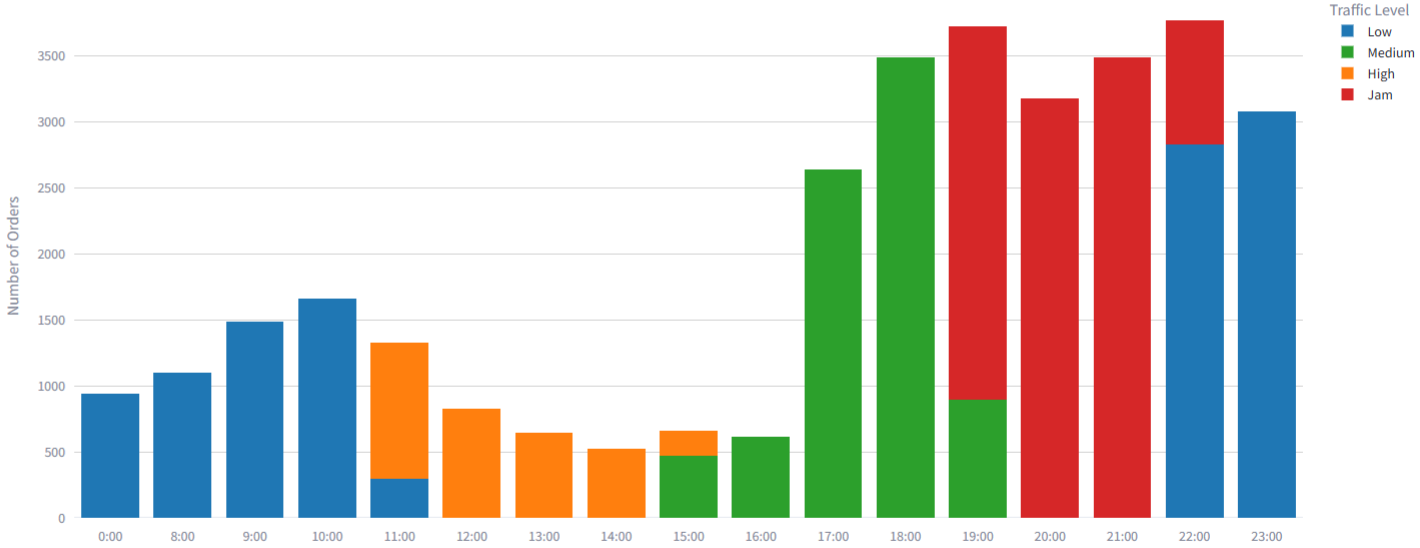
Dashboard

Delivery Time

Contact



Traffic Levels by Hour



Underperforming Drivers (157/1170)

	Delivery_person_ID	Delivery_person_Ratings	Time_taken (min)	Speed Ratio	Vehicle_condition
7	AGRRES03DEL02	4.7	17.62	1.14	

SmartDelivery Analytics

Home
Overview
Dashboard
Delivery Time
Contact

Made with ❤ by Jihad Akbar

Underperforming Drivers (157/1170)

	Delivery_person_ID	Delivery_person_Ratings	Time_taken (min)	Speed Ratio	Vehicle_condition
7	AGRRES03DEL02	4.7	17.62	1.14	1.25
11	AGRRES04DEL03	4.59	30.29	0.47	0.86
12	AGRRES05DEL01	4.67	24	0.48	1.14
13	AGRRES05DEL02	4.58	28.6	0.46	0.7
19	AGRRES08DEL02	4.67	34	0.78	0.44
33	AGRRES17DEL01	4.75	35.5	0.45	0.3
41	AGRRES20DEL03	4.54	20.14	1.17	1
48	ALHRES03DEL01	4.51	30.2	0.47	1.1
49	ALHRES03DEL02	4.74	31.22	0.63	0.44
54	ALHRES07DEL01	4.55	26.92	0.50	0.92

Average Speed Ratio
0.71

Slow Drivers (Speed Ratio < 0.5)
7.9%

Avg Time Taken
26 min

Avg Vehicle Condition
1.0/3.0

Avg Rating
4.6/6.0

Avg Expected Duration
17 min

Recommended Actions

Strategy	Key Actions
Optimize Delivery Routes	Use OSRM API for route planning to avoid traffic and multi-delay bottlenecks
Optimize Delivery Schedule	Shift deliveries to low-traffic hours (night/morning), staff up for peak (17:00-23:00)
Improve Driver Performance	Targeted training for underperforming drivers based on metrics

SmartDelivery Analytics

Home

Overview

Dashboard

Delivery Time

Contact

Made with ❤ by Jihad Akbar

Predict Delivery Time

Welcome to the Delivery Time Prediction Tool!

Eager to know how long a delivery might take? This smart tool helps you estimate delivery times by simply entering a few driver details. It's quick, easy, and designed to help you plan better. Just fill in the form below, and let the tool do the rest!

Driver's Age (Years): 39

Driver's Rating: 4.9

Restaurant Location Latitude: 18.994049

Restaurant Location Longitude: 72.825203

Delivery Location Latitude: 19.074049

Delivery Location Longitude: 72.905203

Vehicle Condition (0=Excellent, 3=Poor): 0

Number of Multiple Deliveries: 1

Weather Conditions: Stormy

Traffic Level: Jam

Order Type: Snack

Vehicle Type: Motorcycle

Is It a Festival?: No

City Type: Metropolitan

Order Day: Monday

Order Month: April

Order Time (Hour): 21

Pick-Up Time (Hour): 21

Pick-Up Time (Minutes): 25

SmartDelivery Analytics

- Home
- Overview
- Dashboard
- Delivery Time
- Contact

Made with ❤ by Jihad Akbar

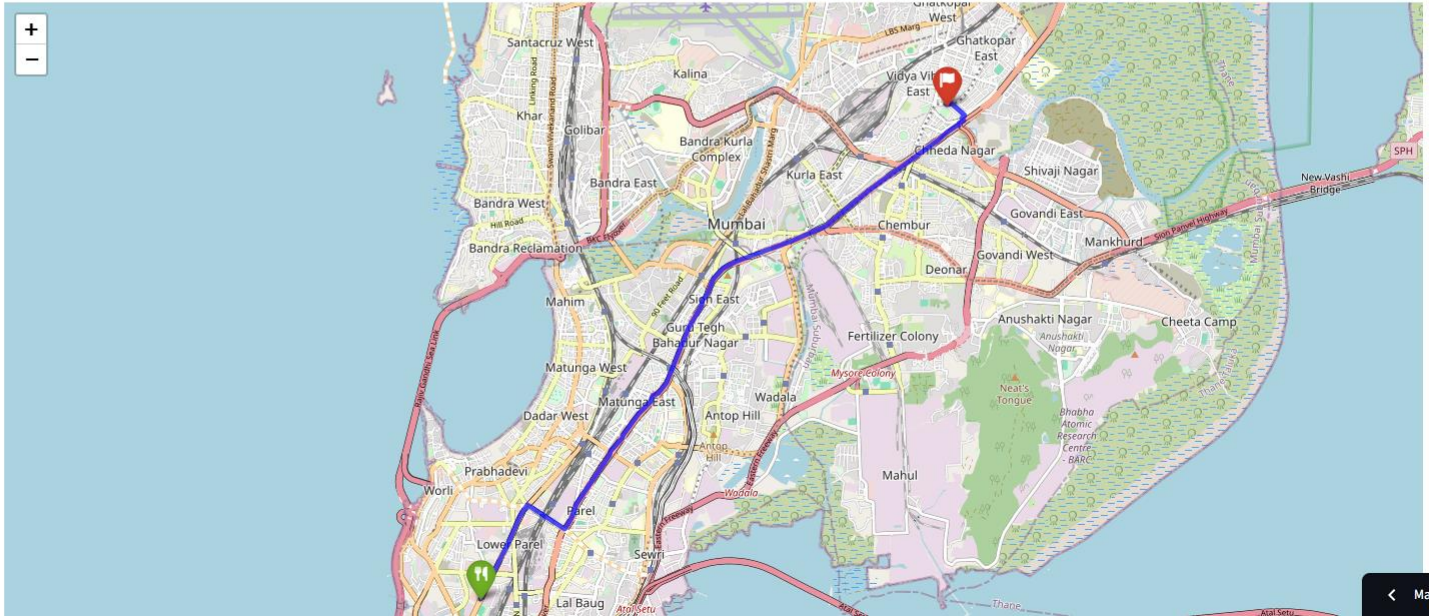
Predict Delivery Time

Delivery ETA

Approximately **37 minutes** (by motorcycle)
Arriving by **22:01** (between 21:57 - 22:05)

Hide Map

Estimated time from OpenStreetMap APIs (by car): 12.9 minutes
Estimated distance from OpenStreetMap APIs (by car): 14.0 km
Estimated speed from OpenStreetMap APIs (by car): 65.1 km/h



Manage app

SmartDelivery Analytics

Home
Overview
Dashboard
Delivery Time
Contact

Made with ❤ by Jihad Akbar

Sample of the Driver

	Features	Driver 1	Driver 2
0	Driver's Age (Years)	39.0	31.0
1	Driver's Ratings	4.9	4.7
2	Restaurant Location Latitude	18.994049	17.430448
3	Restaurant Location Longitude	72.825203	78.418213
4	Delivery Location Latitude	19.074049	17.460448
5	Delivery Location Longitude	72.905203	78.448213
6	Vehicle Condition (0=Excellent, 3=Poor)	0	1
7	Number of Multiple Deliveries	1	1
8	Weather Conditions	Stormy	Cloudy
9	Traffic Level	Jam	Low
10	Order Type	Snack	Meal
11	Vehicle Type	motorcycle	scooter
12	Is It a Festival?	No	No
13	City Type	Metropolitian	Metropolitian
14	Order Day	Monday	Thursday
15	Order Month	April	March
16	Order Time (Hour)	21	23
17	Pick-Up Time (Hour)	21	23
18	Pick-Up Time (Minutes)	25	30
19	Actual Delivery Time (Minutes)	36	18

SmartDelivery Analytics

- Home
- Overview
- Dashboard
- Delivery Time
- Contact

Made with ❤ by Jihad Akbar

Contact Information

Jihad Akbar

Indonesia

(+62) 8133 2326 785

jihadakbr@gmail.com

LinkedIn

GitHub

Feel free to reach out!

04

Appendix

Link

Github: github.com/jihadakbr/time-delivery-route-efficiency-zomato

Streamlit: time-delivery-route-efficiency-zomato.streamlit.app

THANK YOU

Do you have any questions?

Jihad Akbar

+62 813 3232 6785

https://www.linkedin.com/in/jihadakbr

jihadakbr@gmail.com