ZOMATO DELIVERY OPERATION ANALYTICS

Optimizing Delivery Efficiency Through Data-Driven Insights and Predictive Analysis

DEEPAK BEHERA

ABSTRACT

This project analyzes Zomato's delivery operations with the goal of optimizing delivery times, reducing costs, and enhancing customer satisfaction. By leveraging historical data on delivery patterns, customer locations, restaurant preparation times, and delivery partner availability, the study identifies key performance indicators (KPIs) such as average delivery time, order cancellation rates, and customer feedback scores.

A combination of descriptive and predictive analytics is used to uncover trends and bottlenecks. The project explores the efficiency of Zomato's algorithms in assigning delivery partners, minimizing delays, and maximizing the utilization of resources. Furthermore, it evaluates the impact of peak hours, traffic conditions, and delivery distance on operational efficiency.

The project proposes actionable insights such as dynamic partner allocation, optimizing delivery zones, and predictive demand forecasting to improve the delivery process. Machine learning models are applied to predict delays and suggest operational improvements.

The results offer Zomato a data-driven strategy for enhancing delivery performance, reducing costs, and improving the overall customer experience. This project ultimately contributes to the broader goal of building a more efficient and scalable delivery system.

TABLE OF CONTENTS

INTRODUCTION

1.1 Project Overview :-

In today's fast-paced food delivery industry, efficiency and customer satisfaction are critical to staying competitive. Zomato, being one of the largest food delivery platforms, faces the challenge of managing timely deliveries, optimizing resources, and ensuring customer satisfaction across multiple regions. To address these challenges, the **Zomato Delivery Operations Analytics Project** was initiated with the goal of analyzing and improving the efficiency of Zomato's delivery operations.

Using data analytics, this project aims to provide insights into bottlenecks in the delivery pipeline and offer data-driven solutions to improve Zomato's overall operational efficiency, reduce delays, and enhance customer satisfaction.

1.2 Objective :-

- Reduce the time taken for food delivery by optimizing rider allocation, routing, and delivery zones.
- Improve customer feedback and ratings through timely deliveries and better service.
- Track critical KPIs like delivery time, order acceptance rates, rider availability, and customer satisfaction scores.

1.3 Scope :-

- Analyze the flow of orders from restaurants to customers, identifying inefficiencies and delays.
- Study the effect of traffic patterns, time of day, and other external factors (e.g., weather) on delivery times.
- Analyze customer behavior trends like order frequency, preferred restaurants, and peak ordering times.

COMPANY BACKGROUND

2.1 Zomato Business Model :-

a. Market Place Model

- Zomato's initial business model was a **restaurant discovery and review platform** where users could browse restaurant menus, check ratings, and read reviews. This model continues to serve as a foundational element of Zomato's platform.
- Revenue Scource :-
  Restaurants pay Zomato for sponsored listings, which increases their visibility on the platform.

b. Food Delivery Services

  - Users place food orders from partnered restaurants via the app, and Zomato's delivery fleet or third-party riders deliver the food to the customer's location.
  - Zomato charges a commission on every order fulfilled through its platform, usually between 20-30% of the order value.

c. Advertising

  - Restaurants and food brands can advertise on Zomato's platform to increase visibility, attract customers, and promote offers.
  - Zomato charges restaurants for sponsored listings, banner ads, and other in-app promotional tools.

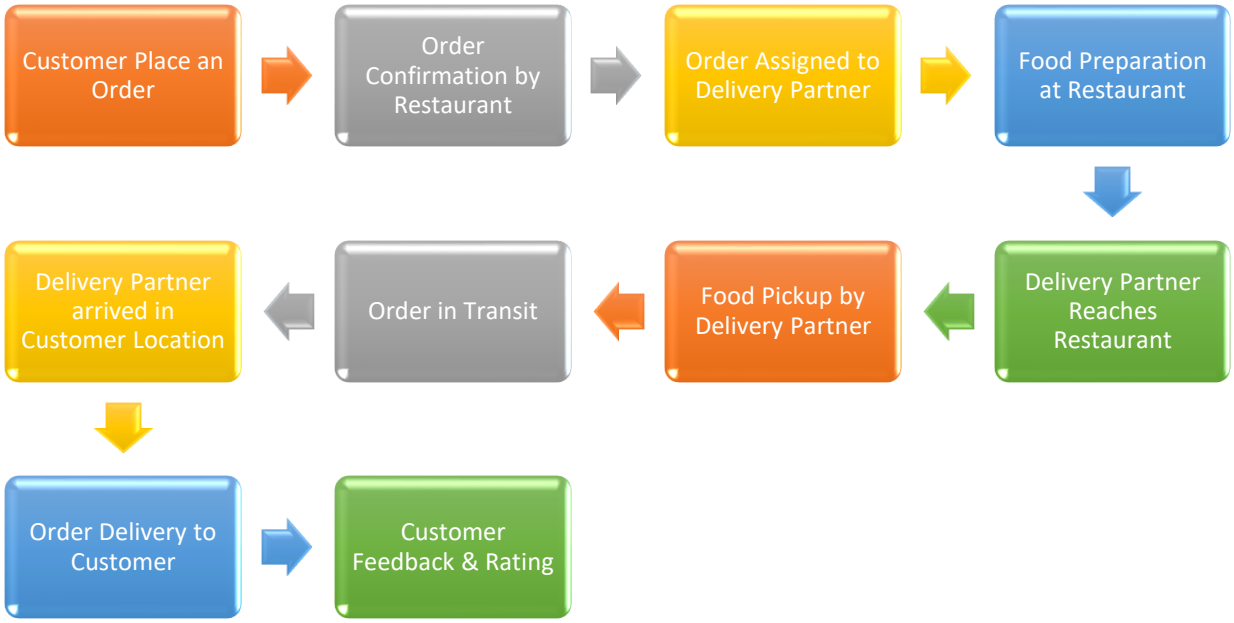
d. Restaurant Analytics

- Restaurants use Zomato's dashboard to analyze order patterns, customer demographics, reviews, and optimize their offerings.

e. Cloud Kitchen
   - Zomato partners with restaurants and helps them set up cloud kitchens, enabling them to expand delivery coverage without investing in physical dining spaces.

f. Zomato Delivery Partner
   - Delivery partners are paid per delivery or based on incentives, but Zomato may take a percentage of the delivery charge or use variable pricing based on demand (surge pricing).

2.2 Delivery Operation Structure :-



PROBLEM STATEMENT

3.1 Key Challenges in Delivery Operation :-

Delivery operations face several challenges that can affect efficiency, customer satisfaction, and profitability. Here are some of the **key challenges** in delivery operations, especially for platforms like Zomato or similar food delivery services:

- Logistics and Route Optimization :- Ensuring that delivery routes are optimized for time and fuel efficiency is critical, especially when handling multiple deliveries. Poor route planning can lead to delays and higher operational costs.
- Delivery Partner Availability and Retentation :- Maintaining an adequate and reliable pool of delivery partners, especially in high-demand areas and peak times, is difficult. Retaining delivery personnel in a competitive market is also a significant concern.
- Weather and Traffic Condition :- Adverse weather conditions like rain, snow, or extreme heat, as well as unpredictable traffic patterns, can significantly delay deliveries.
- Customer Expectation and Satisfaction :- Customers expect fast, accurate, and high-quality service, making it difficult to balance cost-efficiency with speed.
- Food Safty and Hygiene Concern :- Ensuring that the food is handled safely during delivery, especially with the increased focus on hygiene post-pandemic.

Delivery operations, particularly in the food delivery industry, must continuously adapt to a dynamic environment of logistical, technological, and customer challenges. Success depends on balancing efficiency, cost, and customer satisfaction while managing external factors like traffic, weather, and restaurant performance.

3.2 **Importance of Operational Efficiency :-**

- Fast Delivery Time :- Enhancing operational efficiency helps reduce delivery times. This leads to higher customer satisfaction, which is critical for maintaining user loyalty in a competitive food delivery market.
- Resource Optimization :- Through better demand forecasting and real-time data analysis, Zomato can ensure optimal allocation of delivery personnel, ensuring that neither under-utilization nor over-utilization occurs.
- Improved Customer Experiences :- Operational efficiency helps in delivering a seamless user experience. Accurate ETAs, fast delivery, and fewer errors during the ordering process contribute to higher user satisfaction and retention.
- Data Driven Decision Making :- nalyzing operational data helps Zomato identify bottlenecks in the delivery process, predict peak demand times, and determine which areas need more delivery partners or restaurants, thereby continually refining their service.

DATA COLLECTION AND METHODOLOGY

4.1 **Data Sources :-**

- Customer Data :- Order history, preferences, location, delivery feedback.
- Restaurant Data :- Menu, order preparation time, location.
- Delivery Partner Data :- Route, travel time, time at restaurant, delivery time.
- Geographical Data :- Route, travel time, time at restaurant, delivery time.

4.2 **Data Cleaning and Preprocesssing :-**

❖ Data Cleaning :-

- Handelling Missing Value
- Removing Duplicate data
- Handelling Outliers data

❖ Data Preprocessing :-

- Normalization the Data
- Minmax Scaler
- Standred Scaler
- One-hot Encoding
- Label Encoding
- Feature Engineering
- Handelling Imbalanced Data
- Data Aggregation

4.3 **Analytical Tools and Techniques :-**

In a Zomato delivery operation analytics project, various **analytical tools and techniques** are employed to derive actionable insights and optimize the delivery process. These tools and techniques fall under several categories such as descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics.

❖ Analytical Tools :-

- Python (Pandas, Numpy) : For data cleaning, exploration, and manipulation.
- MYSQL : To query and manage large datasets stored in relational databases.
- Power BI : Another popular BI tool for creating dynamic reports.
- Excel : Simple statistical operations and data analysis for quick insights.
- Sciket-Learn : For building predictive models like regression, classification, clustering, etc.
- XGBoost : For high-performance gradient boosting algorithms used in prediction tasks.
- Geopandas : For spatial analysis and mapping deliveries based on geographic data.

❖ Analytical Techniques :-

- Descriptive Analytics : This involves summarizing and interpreting historical data to understand the performance of delivery operations.
  - ✓ Mean, median, mode, standard deviation to summarize delivery times, distances.
  - ✓ Bar charts, histograms, pie charts, and line graphs to understand the distribution of orders, delays, peak times, and locations.
  - ✓ Mapping delivery locations to identify high-demand areas, delivery zones, or bottlenecks in delivery routes.
- Diagonistic Analytics : Diagnostic analytics helps in understanding the "why" behind certain patterns or anomalies.
  - ✓ Root Cause Analysis
  - ✓ Correlation Analysis
  - ✓ Hypothesis Testing
- Predictive Analytics : Predictive analytics uses historical data to forecast future outcomes, helping to predict potential delivery delays or customer demand patterns.
  - ✓ Regression Analysis : Predicting delivery times based on factors like distance, time of day, traffic, and weather conditions.
- Prescriptive Analytics : Prescriptive analytics recommends actions to optimize delivery operations based on predictions and data-driven insights.
  - ✓ Optimization Algorithms : Using algorithms like the Traveling Salesman Problem (TSP) or vehicle routing to determine the most efficient delivery routes for drivers.

❖ Performance Monitoring & KPI Analysis :-

- KPI : racking metrics like average delivery time, on-time delivery rate, customer satisfaction scores, and average delivery cost.
- Dashboarding : Creating real-time dashboards to monitor performance metrics and delivery health using Tableau or Power BI.

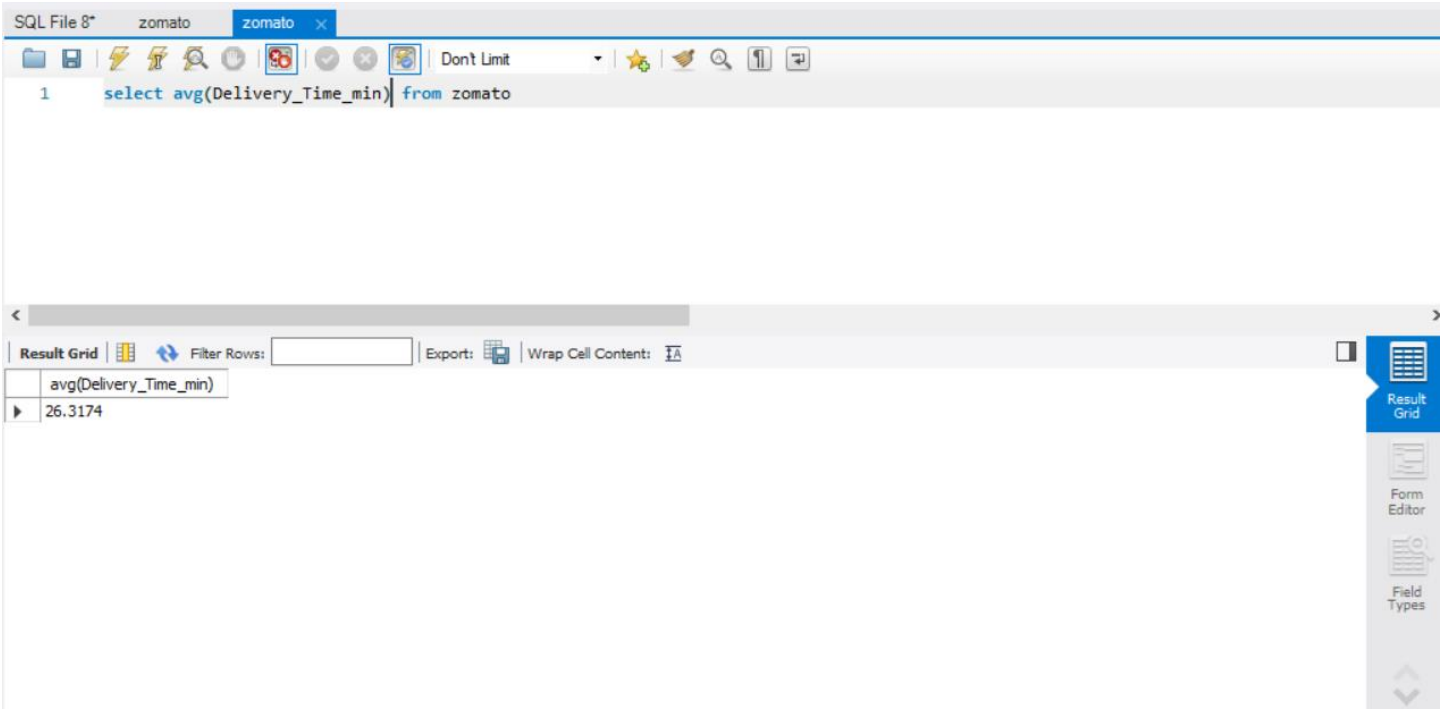
DESCRIPTIVE ANALYTICS

5.1 KPIs

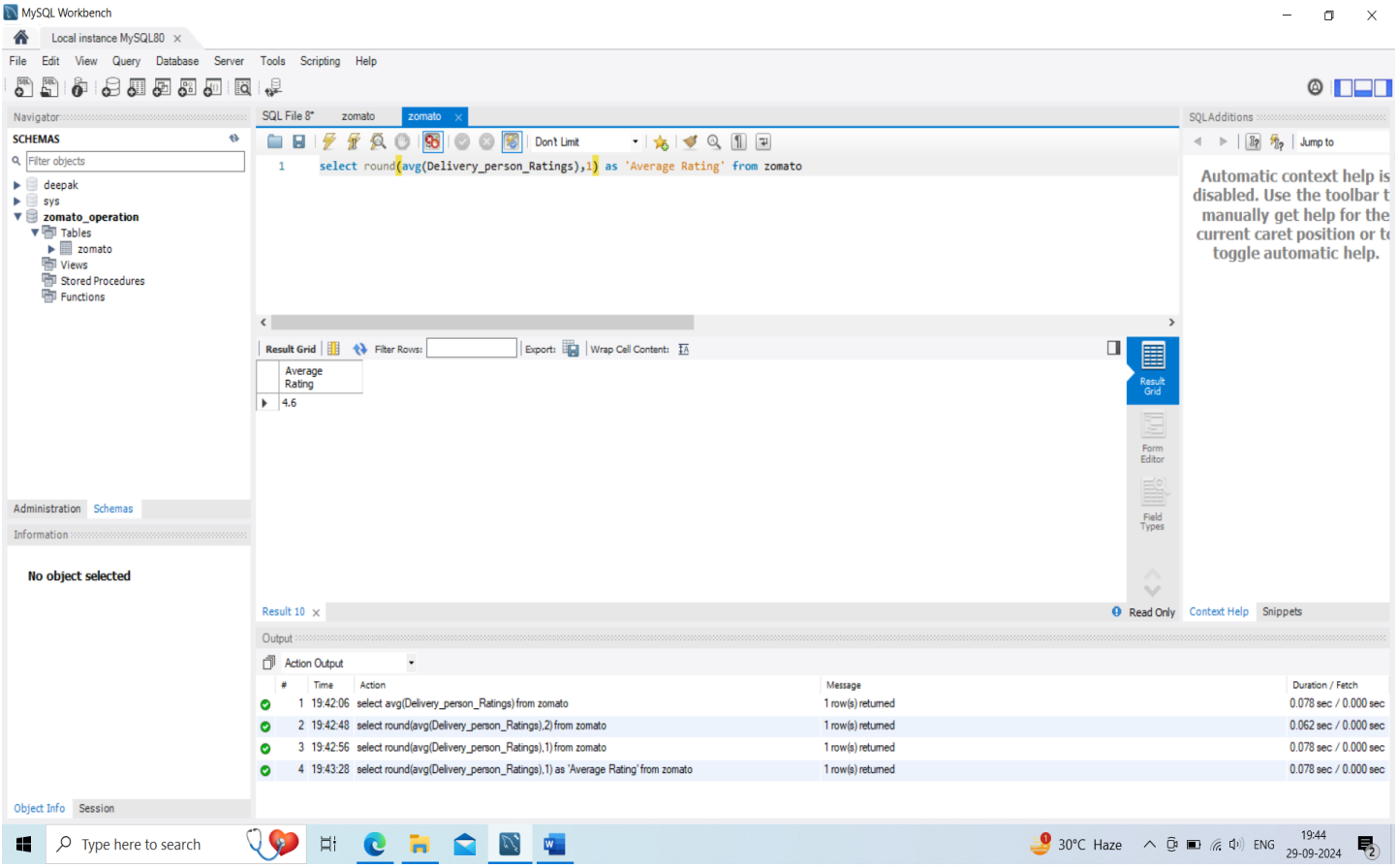
- Average Delivery Time : The average time taken from the moment an order is placed to when it is delivered. Lower values indicate faster delivery.
- Average Rating Per Delivery : Based on customer reviews and ratings for each delivery, often on a 1-5 scale.
- Order Volume : The total number of deliveries completed within a given time frame (daily, weekly, monthly).
- Order Growth Rate : The percentage increase or decrease in order volume over time, indicating market expansion or contraction.
- Average Order Value : The average monetary value of each order, useful for tracking customer spending behavior.
- Average Distance Per Delivery : The average distance that delivery personnel must travel to fulfill an order, which affects delivery time and cost.

- Delivery Personnel Rating : Ratings given by customers based on their experience with the delivery person's behavior and professionalism.
- Heat Map of Delivery Demand : Identifying areas with the highest concentration of orders, which can help optimize resource allocation and delivery routes.
- Order Fulfillment Time : The total time taken from order placement to restaurant acceptance, cooking, and handoff to the delivery person.

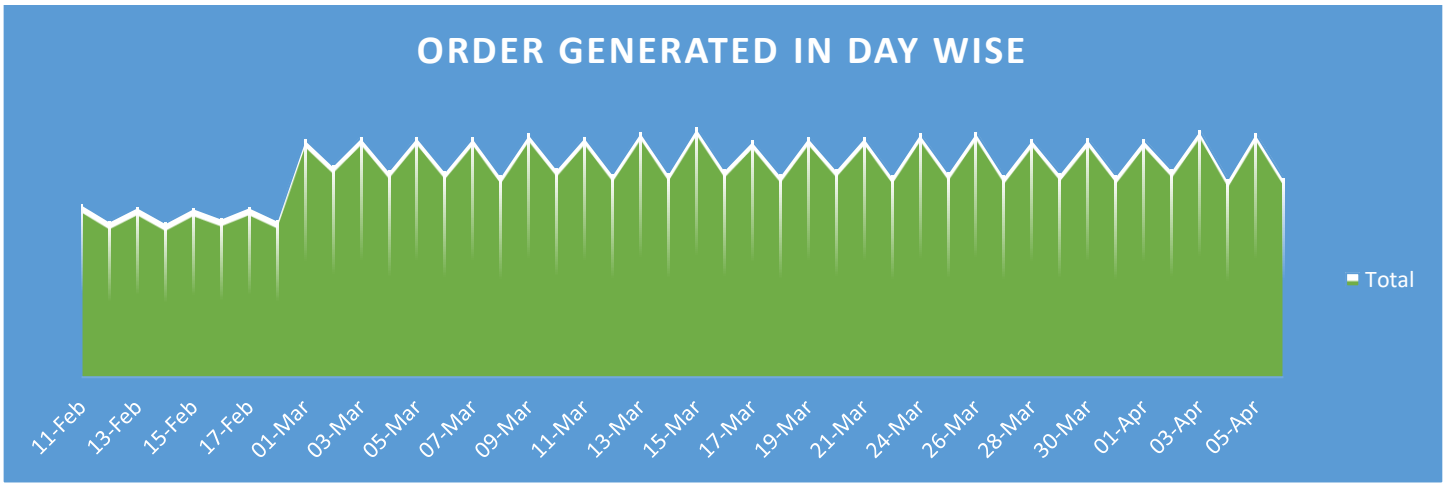
5.2 Average Delivery Time



5.3 Customer Feedback



5.4 Analysis of Delivery Pattern

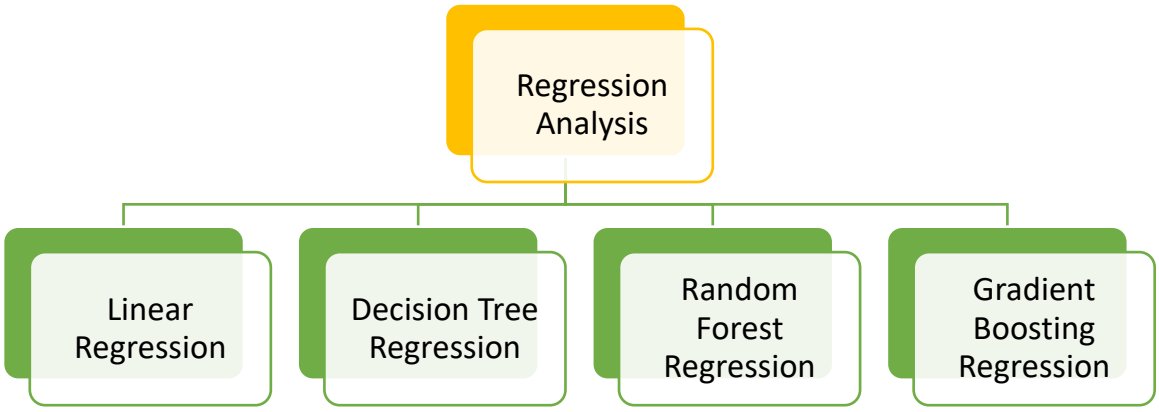


PREDICTIVE ANALYTICS

6.1 Machine Learning Model

In a **Zomato delivery analytics project**, machine learning models can be used to improve various aspects of the delivery process, such as predicting late deliveries, optimizing delivery routes, or forecasting demand. Below are examples of machine learning models that could be applied in different areas of the delivery ecosystem.

- ❖ Delivery Time Estimation : Accurate estimation of delivery time can greatly improve customer satisfaction by setting the right expectations.
  - Problem Type : Regression

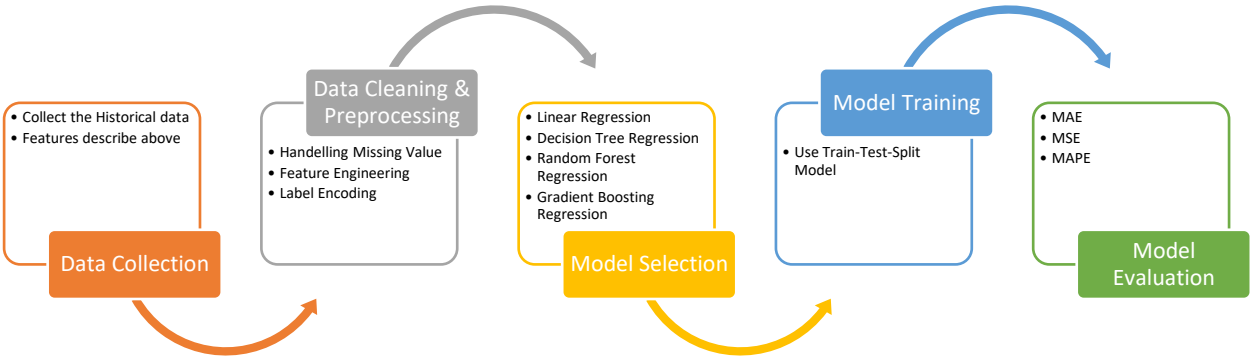


6.2 Delivery Time Prediction

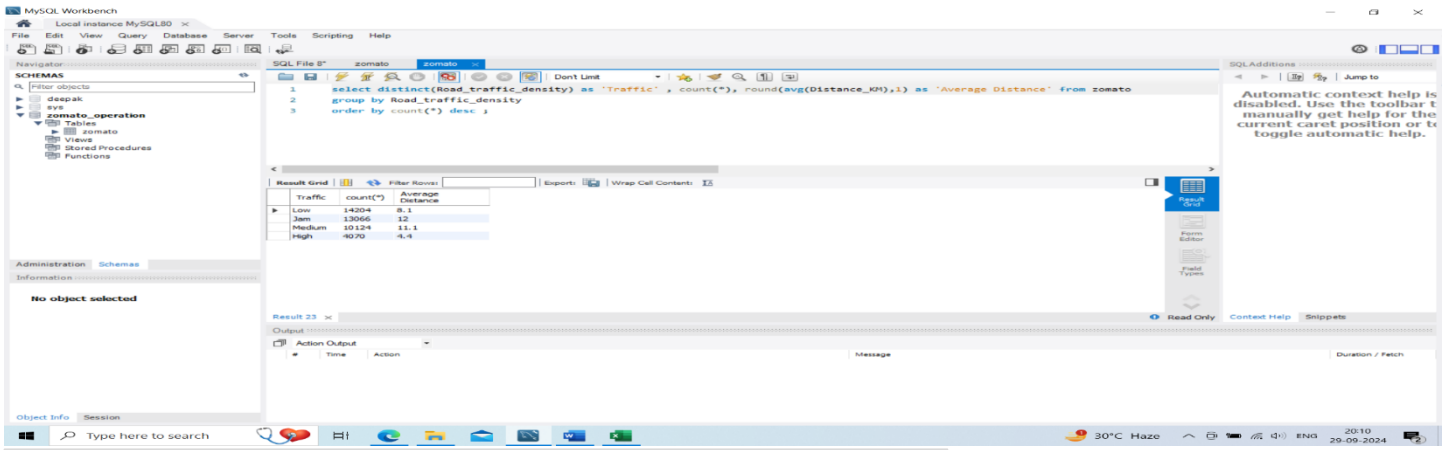
To build a machine learning model that predicts the estimated delivery time from the moment an order is placed to its successful completion, taking into account factors such as distance, traffic conditions, restaurant preparation time, and weather.

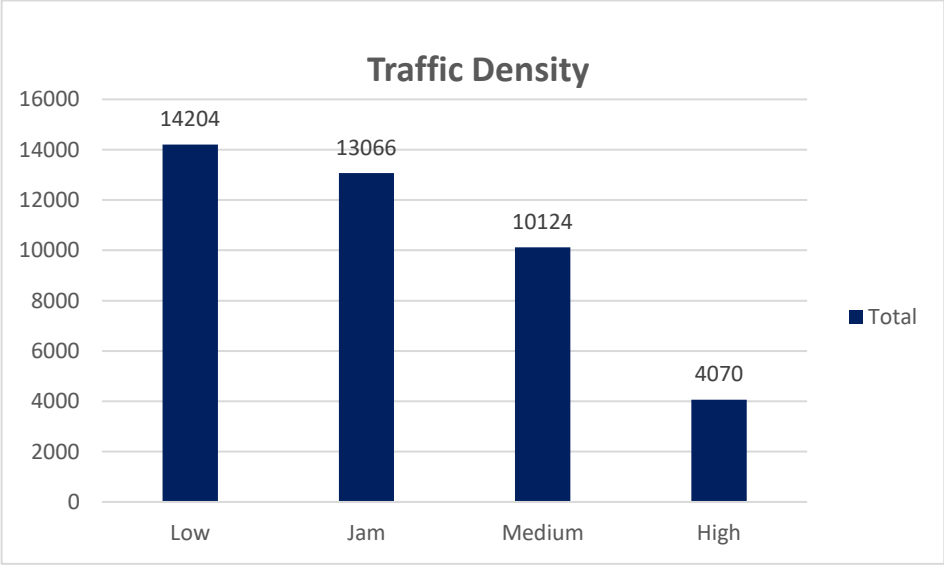
- Problem Type :-
  - ✓ Regression : This is a **supervised learning** task where the model predicts a continuous variable (i.e., the delivery time in minutes).

- Data Features for Delivery Time Pediction :-
  - ✓ Order Placement Time
  - ✓ Order Preparation Time
  - ✓ Delivery Distance
  - ✓ Traffic Condition
  - ✓ Weather Condition
  - ✓ Geographical Location
  - ✓ Age of Delivery Person
  - ✓ Vehicle Condition
- Step to Build a Delivery Time Prediction



6.3 Traffic and Distance Impact Analysis





OPTIMIZATION TECHNIQUES

7.1 Route Optimization using Traffic and Distance Data

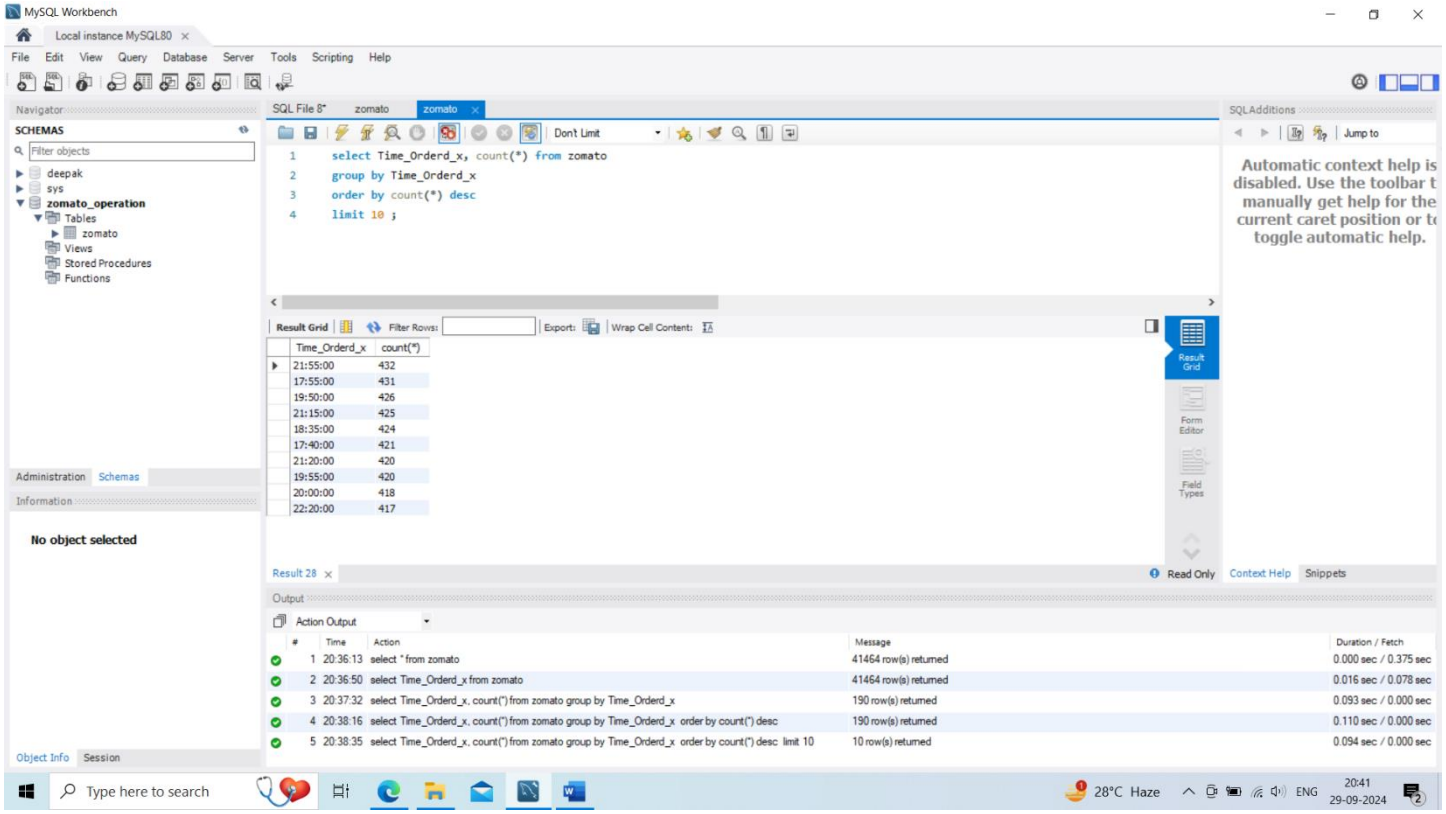
- Optimal Route Generation :- Based on the scenario, generate the optimal delivery route by balancing traffic conditions, shortest distance, and delivery priorities.
- Real Time Adjustment :- Continuously monitor traffic changes, order cancellations, and new orders, adjusting routes on the fly.

7.2 Reducing Restaurant Preparation Time

- Task Sequencing :- Organize the sequence of food preparation steps so that tasks are completed efficiently and in parallel where possible
- Pre-Preparation :- Prepare ingredients (chopping, marinating, etc.) ahead of time, especially for commonly ordered dishes, so chefs can assemble meals quickly during peak hours.
- Cross- Utilize Gradient :- Design dishes that use overlapping ingredients to reduce preparation complexity and time.
- Cross Training Staff :- Train staff to handle multiple roles, allowing flexibility in the kitchen, especially during peak times when bottlenecks occur.
- Order Management System :- Integrate order management systems that streamline the flow of orders between front-of-house and the kitchen, reducing miscommunication and delays.

OPERATIONAL INSIGHT

8.1 Peak Hours and Traffic Considerations

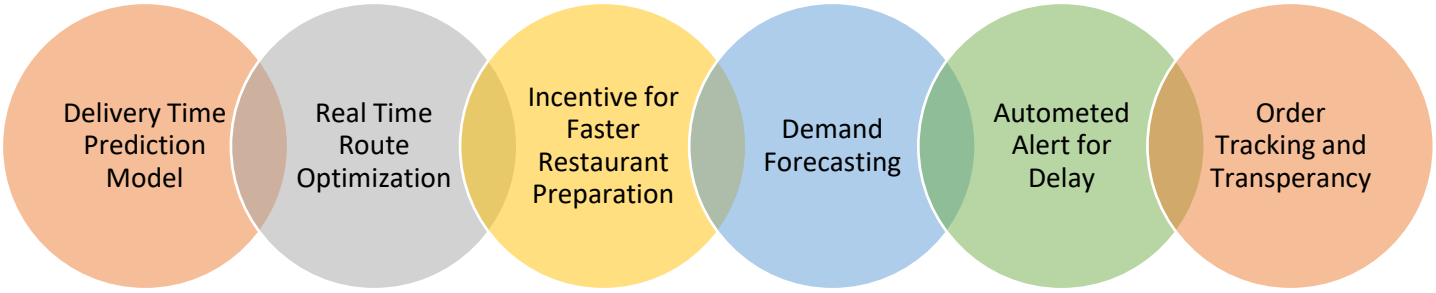


8.2 Bottleneck Identification

- Restaurant Preparation Bottlenecks
- Order Dispatch and Routing Bottlenecks
- Traffic and Weather Bottlenecks

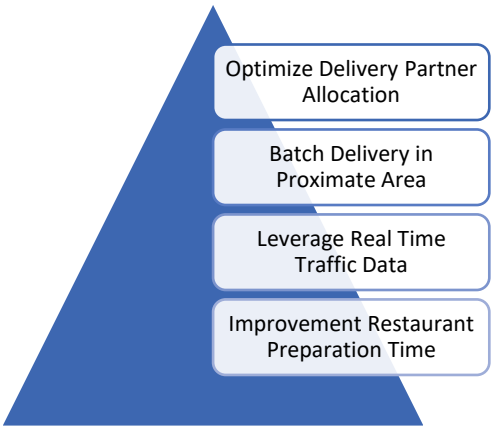
8.3 Strategies to Reduce Delays

To reduce delays in Zomato's delivery operations, especially during peak hours and traffic congestion, several strategies can be employed. These strategies focus on streamlining the order preparation, optimizing delivery routes, improving coordination between restaurants, delivery partners, and customers, and using data-driven insights to enhance the overall delivery process.



RECOMMENDATION

9.1 Improving Delivery Times

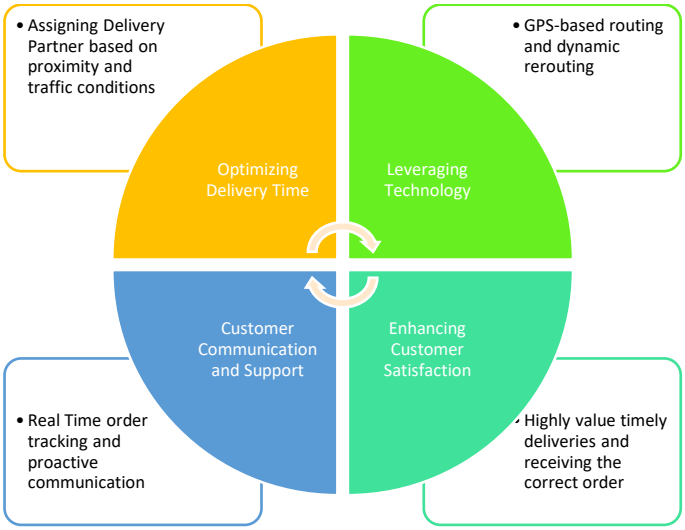


9.2 Enhancing Customer Satisfaction

CONCLUSION

10.1 Summary of Findings

- This project focuses on improving delivery operations and enhancing customer satisfaction for a food delivery platform like Zomato.
- The main goals are to reduce delivery times, optimize operational efficiency, and increase customer loyalty through a combination of advanced technologies, data-driven insights, and operational enhancements.
- By leveraging AI, machine learning, real-time data, and customer-centric strategies, the project aims to deliver tangible improvements in the speed and quality of service.



- The project's findings underscore the importance of technology-driven optimization, real-time data integration, and customer-centric strategies to improve delivery operations and enhance customer satisfaction. Implementing these solutions can lead to faster deliveries, better service, and higher customer retention, making the platform more competitive and efficient.

10.2 Future Scope and Research Opportunities

❖ AI and Machine Learning for Deeper Optimization

- Advanced AI models can be further developed to continuously learn from new data to improve delivery route planning, demand forecasting, and partner allocation.
- Explore reinforcement learning algorithms that enable adaptive decision-making based on real-time scenarios (traffic changes, weather, etc.).

❖ Advanced Customer Personalizatin and Predictive Behavior Analysis

- Predicting customer behavior more accurately (e.g., when and what they'll order) could lead to even better operational planning, like pre-positioning delivery partners.
- Research ways to improve customer experience using predictive analytics to offer personalized menus, time-slot-based offers, and pre-ordering suggestions.

❖ Real Time Customer Feedback Loop

- Using AI to analyze real-time customer feedback and adjust operations dynamically could greatly improve service and satisfaction.
- Explore how sentiment analysis from customer reviews and feedback can be integrated into delivery operations to predict and preempt potential issues.

- ✓ By exploring these future areas, food delivery platforms like Zomato can stay at the forefront of innovation, continually improving their efficiency, sustainability, and customer satisfaction. Each of these areas presents unique research opportunities that can shape the next generation of food delivery services.

REFERENCES

11.1 Sources and Literature Review

- Exploratory Data Analysis of Zomato Data in Python: Unveiling Insights for Restaurant Success | by Vijay Sundaram Mohana | Medium
- (PDF) Zomato Data Analysis | IJRASET Publication - Academia.edu
- (PDF) An Analysis of Online Food Ordering Applications in India: Zomato and Swiggy (researchgate.net)
- End-to-End Predictive Analysis on Zomato - Analytics Vidhya
- Zomato Data Analysis Using Python - GeeksforGeeks

DASHBOARD BUILDING POWER BI

KEY METRICS FOR STAKE HOLDERS

1. Numer of Delivery Per Delivery Person
2. Multiple Deliveries :- Number of cases where delivery personnel handled multiple orders at once
3. Rating Scale Distribution
4. Rating by Weather Condition
5. Impact on Traffic and Weather on Delivery Time
6. Vehicle Condition on Delivery Time
7. Impact of Type of vehicle on Delivery Time
8. Daily / Weekly Delivery Trends
9. Seasonal and Weather Delivery Trends
10. Impact of Delivery Person age Range on Delivery Time
11. Percentage of Rating Scale

SQL QUERIES 100 Query for Praticing

1. Delivery Person age (Average, min, max)
2. Value counts of Type of Order columns
3. What is average food processing time
4. What is Distance (Average, min, max)
5. Delivery state value counts
6. Type of vehicle value counts
7. Weather conditions value counts
8. Road taffic density value counts
9. Delivery Time (Average, min, max)
10. Delivery person rating (Average, min, max)
11. Rating value count
12. Age below 25 and order snack
13. How many delivery persons are between 25-30

14. Retrive the data where delivery person id start from (D, K, P)
15. Average delivery time of each age group
16. Average delivery time of each state category
17. Average delivery time in vehicle
18. Average delivery time in weather conditions
19. Average delivery time in road traffic density
20. Write a query where rating is excellent and very good

21. Order Generated in each date
22. Ascending and Descending order in date
23. Average delivery time of each delivery person
24. Identify top 5 delivery persons who's delivery order in highest number
25. How many customer are from maharashtra or Uttarkhand
26. 5^{th} highest number of order generated state
27. Average distance of each state
28. Average distance travel of each delivery persons
29. Which time is generated highest number of order
30. Average delivery time of the order in Order generated time

31. Average delivery time in each vehicle conditions
32. Average delivery time of each multiple deliveries
33. Peak time when the highest number of order are generated
34. How many delivery persons are assigned in Kerala state
35. Find the data which Is longer than 30 min delivery time
36. Find out the highest excellent rating in each state
37. Average delivery time of each type of order (food)
38. List the top 5 delivery persons based on their rating
39. Which delivery person has the highest average delivery rating
40. Average food processing time of each type of order

41. Average rating of each state
42. Average rating on type of food order
43. Averge rating on vehicle conditions
44. Which vehicle is beter delivery rating
45. The maximum distance travel of each state
46. Find out the data where the average delivery time is greather tha 30 and the distance is grather than average distance
47. Group the orders by Delivery_State and count the number of deliveries for each state. Only return states with more than 5 deliveries.
48. Group the orders by Weather Condition and count the number of deliveries for each state. Only return states with more than 5 deliveries.
49. Find the average Delivery_person_Ratings grouped by Type_of_vehicle. Return only those vehicle types that have an average rating of 4.5 or higher.
50. Count the number of deliveries made by each delivery person and filter to show only those who have made more than 3 deliveries.

51. Group the results by Weather_conditions and calculate the average Time_taken (min)_x. Return only those weather conditions that have an average time greater than the overall average.
52. Create age groups for Delivery_person_Age (e.g., 20-30, 31-40, etc.) and calculate the average Delivery_person_Ratings for each age group. Return groups with an average rating above 4.0.
53. Group by Type_of_order and find the total distance traveled (Distance_KM) for each type. Only return types where the total distance is greater than 15 KM.
54. Calculate the average Food Processing (min) for each Type_of_order, but only include types where the average processing time exceeds 10 minutes.

55. Group by Rating Scale and count how many deliveries fall into each rating category. Only show categories with more than 5 entries.
56. Find the average Time_taken (min)_x for each Type_of_order and return only those where the average time taken is significantly longer (e.g., greater than 5 minutes) than the overall average delivery time
57. Group the orders by month (extract the month from Order_Date) and count the total number of deliveries for each month. Only return months with more than 10 deliveries.
58. Calculate the average Distance_KM for each Type_of_order. Return only those order types that have an average distance greater than the overall average distance.
59. Group the deliveries by Delivery_State and calculate the average Delivery_person_Ratings for each state. Filter the results to show only those states with an average rating below 4.0.
60. Find the average Time_taken (min)_x for each Vehicle_condition. Return only those conditions where the average time taken exceeds 30 minutes.

61. Group the data by both Type_of_order and Weather_conditions, and calculate the count of deliveries for each combination. Return combinations where the count is greater than 3.
62. Group by Type_of_order and calculate the total Time_taken (min)_x for each type. Filter to show only those types where the total time taken is more than 100 minutes.
63. Group the results by Delivery_State and calculate the average Time_taken (min)_x for each state. Return only states with an average delivery time greater than the overall average.
64. Group the deliveries by Type_of_vehicle and calculate the average delivery time (Time_taken (min)_x). Return only those vehicle types where the average delivery time is less than the overall average delivery time.
65. Find the delivery persons who have a rating above 4.5 and calculate the total number of deliveries they have made. Return only those who have made more than 5 deliveries.
66. Group by Vehicle_condition and calculate the count of deliveries for each condition. Return only conditions that have fewer than 5 deliveries.
67. Write a query to find all delivery persons who have a Delivery_person_Ratings above the average rating of all delivery persons.
68. Find the details of all orders where the Time_taken (min)_x exceeds the average delivery time for all orders.
69. Select the delivery persons who have a rating greater than the average rating of delivery persons from a specific state.
70. Write a query to find delivery persons who have made more deliveries than the average number of deliveries per delivery person.

71. Find orders where the Food Processing (min) is greater than the average food processing time for all orders.
72. Identify the states where the number of deliveries is greater than the average number of deliveries across all states.
73. Write a query to find orders with a Distance_KM greater than the average distance for all orders.
74. Find the delivery persons whose ratings are higher than the average rating of all delivery persons in the same vehicle type.
75. Retrieve all details of orders that have a Delivery_person_Ratings greater than the average rating for orders of the same type.
76. Write a query to find states where the average Time_taken (min)_x is less than the average time across all states.
77. Identify all Type_of_order that have an average Delivery_person_Ratings greater than 4.0.
78. Find all deliveries made in weather conditions where the average delivery time exceeds the average delivery time for all conditions.
79. List all delivery persons who are older than the average age of all delivery persons.
80. Write a query to find the order types that have more deliveries than the average number of deliveries for all order types.

81. Identify the Type_of_vehicle that has a higher average delivery rating than the overall average delivery rating.
82. Write a query that retrieves the Delivery_person_ID along with their total number of deliveries as a separate column.
83. Find all weather conditions where the average delivery time is greater than the overall average delivery time, and list the number of deliveries for each condition.

84. Identify the delivery persons whose ratings are higher than the average rating of delivery persons in their age group.
85. List orders that have a Food Processing (min) longer than the average processing time of all orders of the same type.
86. Write a query to find states with a number of deliveries greater than the average number of deliveries per state.
87. Write a query to find the average Time_taken (min)_x for deliveries made in fog, grouped by Delivery_State. Return only states with an average time above 35 minutes.
88. Write a query to find delivery persons whose average Delivery_person_Ratings is greater than the average rating for their Type_of_vehicle.
89. For each state, find the delivery location (Delivery_location_latitude, Delivery_location_longitude) with the highest number of deliveries.
90. Find the top 10% of fastest deliveries (Time_taken (min)_x) made by delivery persons in each age group (Delivery_person_Age), sorted by speed.

91. Use a subquery to find which Type_of_order has a higher average Delivery_person_Ratings compared to the overall average rating of all orders.
92. Which type of order are maximum in between 17.00.00 to 21.00.00 PM
93. Write a query to find all orders where the Food Processing (min) is greater than 30 **AND** the Weather_conditions are either "Stormy" **OR** "Fog."
94. Write a query to find all orders that were delivered by vehicles **NOT** classified as a "motorcycle" (Type_of_vehicle).
95. Write a query to find all orders where the Weather_conditions are "Clear" **AND** the Road_traffic_density is "Low."
96. Write a query to retrieve all deliveries made in either "Uttarakhand," "Kerala," or "Punjab" using the IN operator.
97. Write a query to find all deliveries where the Delivery_person_Ratings is **BETWEEN** 4 and 5.
98. Write a query to find all orders where the Road_traffic_density is either "Low" **OR** "Medium" **AND** the Delivery_person_Ratings is greater than the average rating of all orders.
99. Write a query to find all orders where the Time_taken (min)_x is greater than 40 **AND** either the Road_traffic_density is "High" **OR** there were multiple deliveries (multiple_deliveries > 1).
100. Write a query to find all orders where the Delivery_person_Ratings is **BETWEEN** 3 and 4.5 **AND** the Road_traffic_density is "High."