



ZOMATO DATA ANALYSIS USING PYTHON

With the aim of strengthening my data analysis skills and gaining hands-on experience, I worked on the project 'Zomato Data Analysis Using Python.' This project allowed me to explore real-world data, apply analytical techniques, and demonstrate my proficiency in Python programming and data visualization, which are essential for a career in data analytics.

DATA CLEANING & PREPROCESSING :

In this project, I followed essential data preprocessing steps to ensure the dataset was clean and ready for analysis. The first step involved loading the dataset into a Python environment using pandas. Then, I performed data cleaning to remove inconsistencies and errors within the data. Missing values were handled through appropriate techniques, including deletion, to maintain the integrity of the analysis. These preprocessing steps were crucial for ensuring the quality and accuracy of the data, which laid the foundation for meaningful insights in the analysis.

EXPLORATORY DATA ANALYSIS

1. Univariate Analysis:

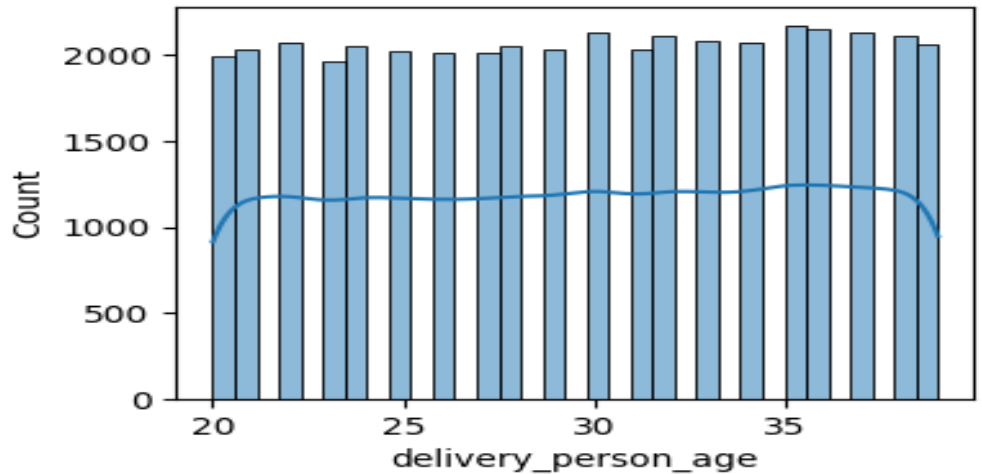
Delivery Person Age Distribution:

The first histogram showcases the distribution of delivery persons' ages. The bars indicate the count of delivery persons in each age range. Observations

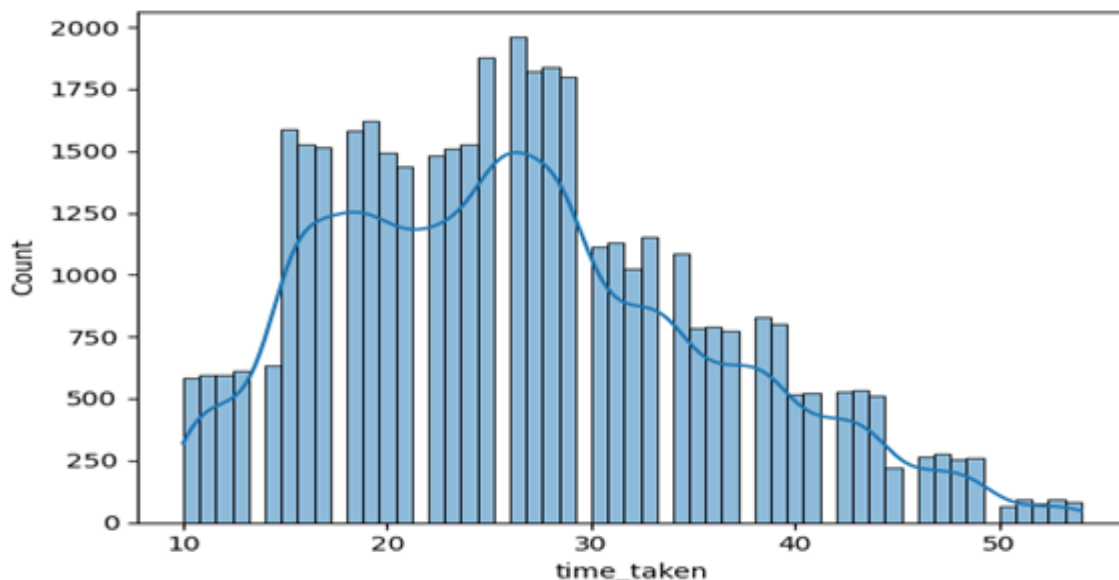
- The ages are evenly distributed across the range, with no significant peaks or troughs.

- Most delivery persons fall within the age group of 20 to 40 years.

- This indicates that the workforce in food delivery is relatively young and fairly consistent in terms of age.



Delivery Time Distribution:



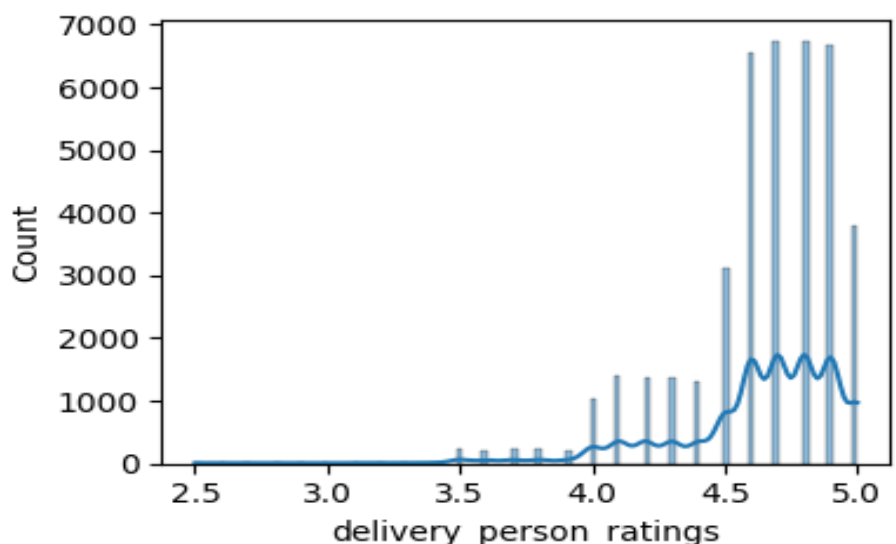
The second histogram depicts the time taken for deliveries. Key points:

- The data is right-skewed, with most deliveries occurring within 20 to 30 minutes.
- A noticeable decline in counts is observed as the delivery time increases beyond 30 minutes, indicating fewer long-duration deliveries.
- The peak around the 20–30 minute mark aligns with typical delivery service expectations, reflecting operational efficiency.

Delivery Person Ratings:

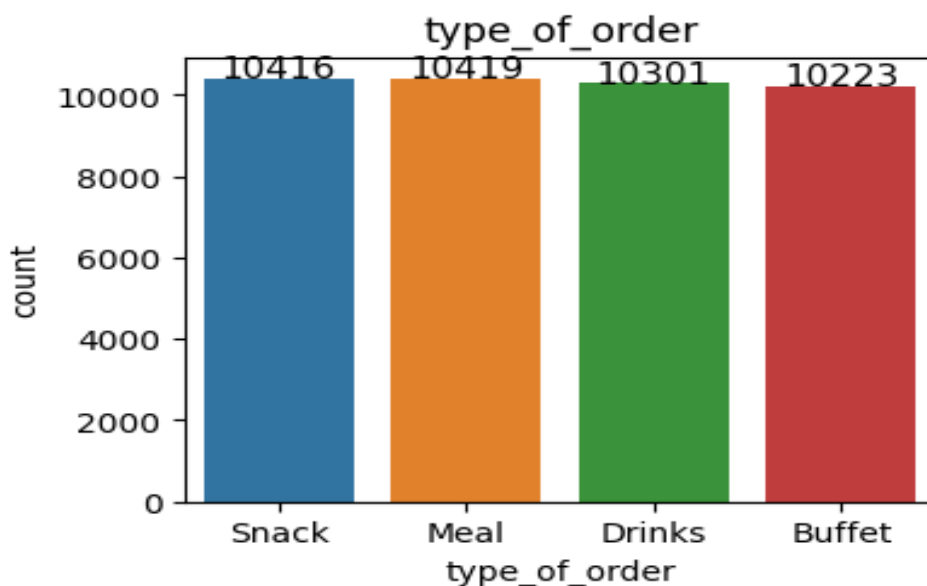
The third histogram shows the distribution of ratings for delivery personnel. Insights include:

- The majority of ratings are concentrated between 4.5 and 5, indicating high customer satisfaction.
- A smaller number of ratings fall below 4, with minimal data points for ratings lower than 3.
- This suggests that Zomato customers generally have positive experiences with delivery personnel.



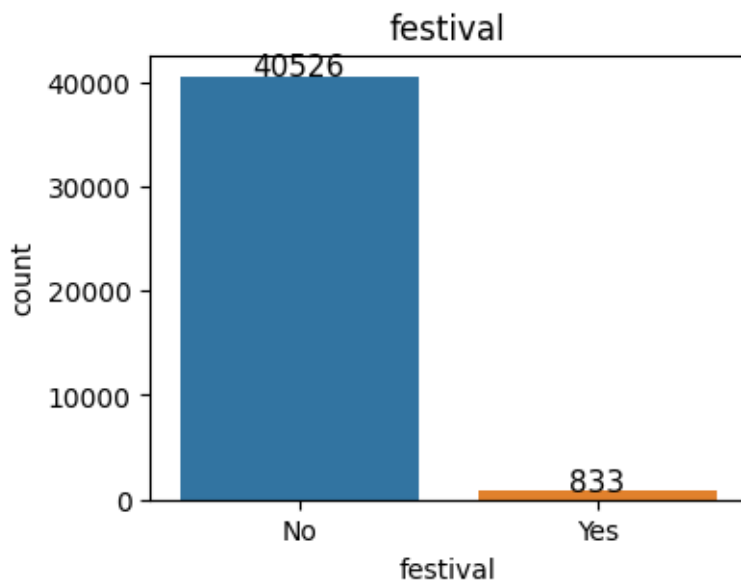
These histograms provide valuable insights into Zomato's operations, including workforce demographics, delivery efficiency, and customer satisfaction levels. Such analyses help identify strengths and areas for improvement, enhancing decision-making in operations and customer service.

Type of Orders :



- The data shows that Snacks, Meals, Drinks, and Buffets are nearly equally preferred, with order counts ranging between 10,000 to 10,419.
- **Insight:** This balanced distribution indicates that Zomato caters effectively to diverse customer needs across categories.
- **Actionable Suggestion:** Further analyze customer feedback to identify trends and enhance popular offerings.

Festival delivery's :

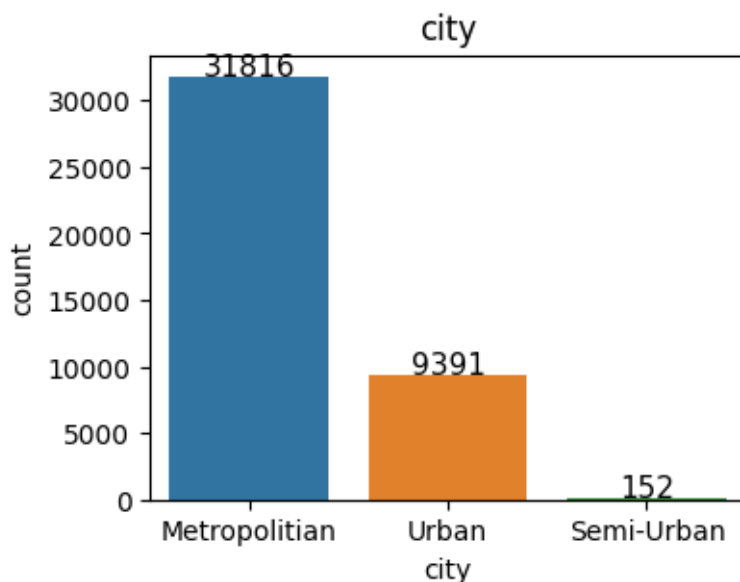


- Non-festival periods dominate order volumes (40,526), while festival periods record significantly lower orders (833).

- **Insight:** Low orders during festivals may indicate a preference for dine-in experiences or special home-cooked meals.

- **Actionable Suggestion:** Introduce festive promotions, exclusive deals, and personalized recommendations to boost festival-season engagement.

City delivery's :

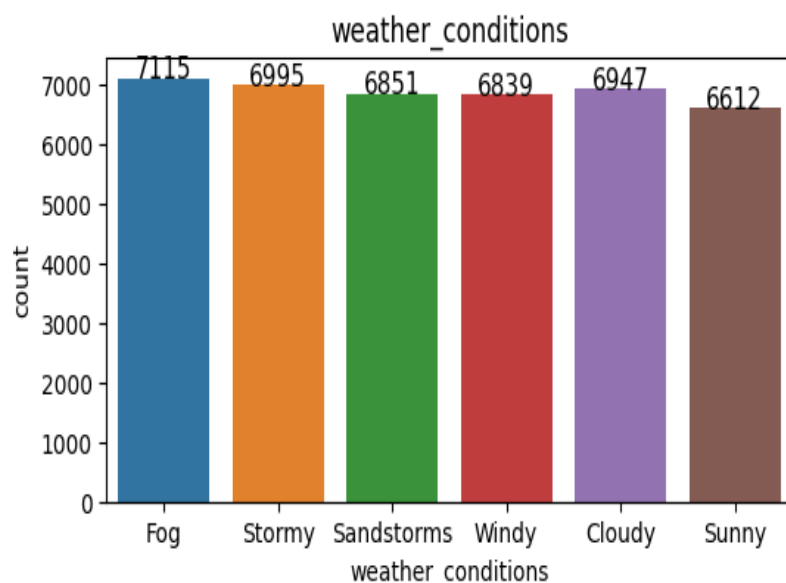


- Metropolitan cities account for the highest orders (31,816), followed by Urban areas (9,391), while Semi-Urban areas contribute only 152 orders.

- **Insight:** The vast majority of orders come from well-connected urban regions, with Semi-Urban areas remaining an underutilized market.

- **Actionable Suggestion:** Evaluate delivery feasibility and market potential in Semi-Urban areas to expand Zomato's reach.

Weather Conditions :

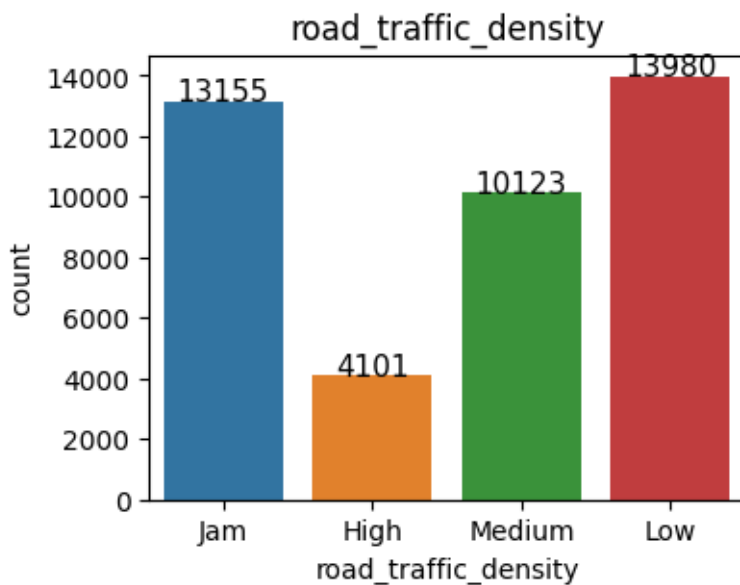


- Order counts remain relatively consistent across weather types: Fog (7,015), Stormy (6,995), Sandstorms (6,851), Windy (6,839), Cloudy (6,947), and Sunny (6,612).

- **Insight:** Weather does not significantly impact customer orders, demonstrating reliability in customer demand despite external conditions.

- **Actionable Suggestion:** Ensure operational resilience during extreme weather by preparing delivery teams with protective measures and robust logistics.

Road Traffic Density :



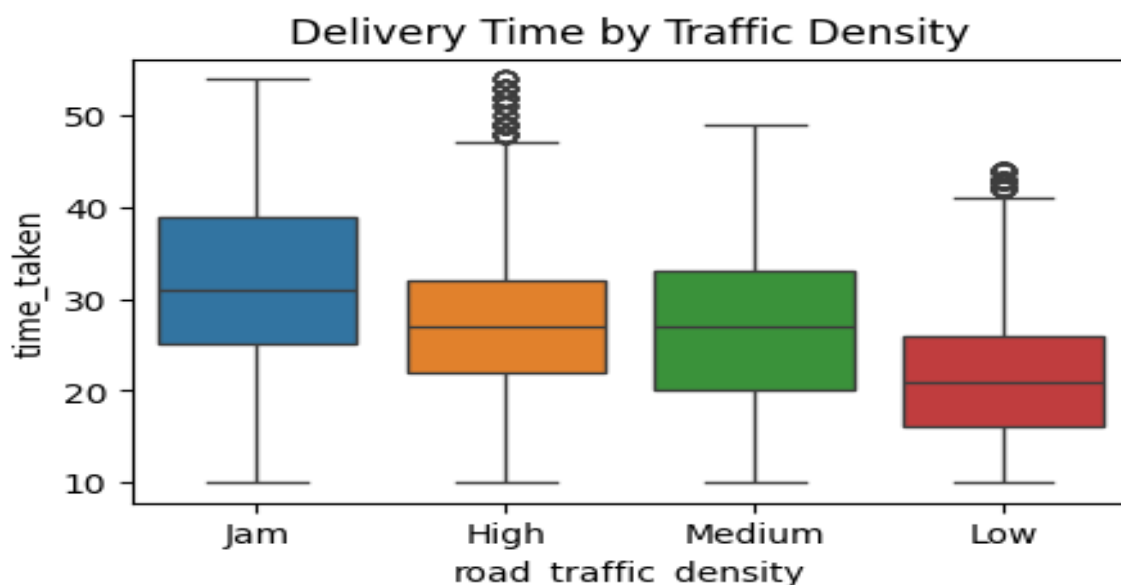
- Traffic conditions influence order volumes: Low (13,980) and Jam (13,155) have the highest orders, while High traffic conditions see only 4,101 orders.

- **Insight:** Customers may avoid ordering during heavy traffic due to anticipated delays.

- **Actionable Suggestion:** Utilize advanced traffic data to optimize routes and maintain timely deliveries during peak traffic hours.

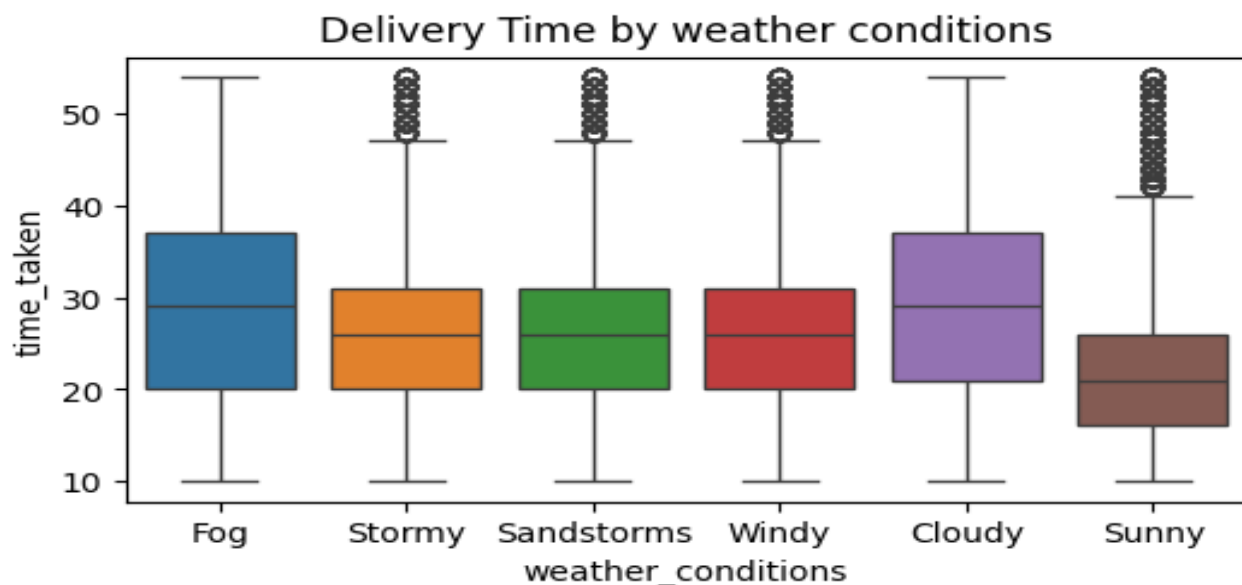
2.Bivariate Analysis :

Delivery Time by Traffic Density :



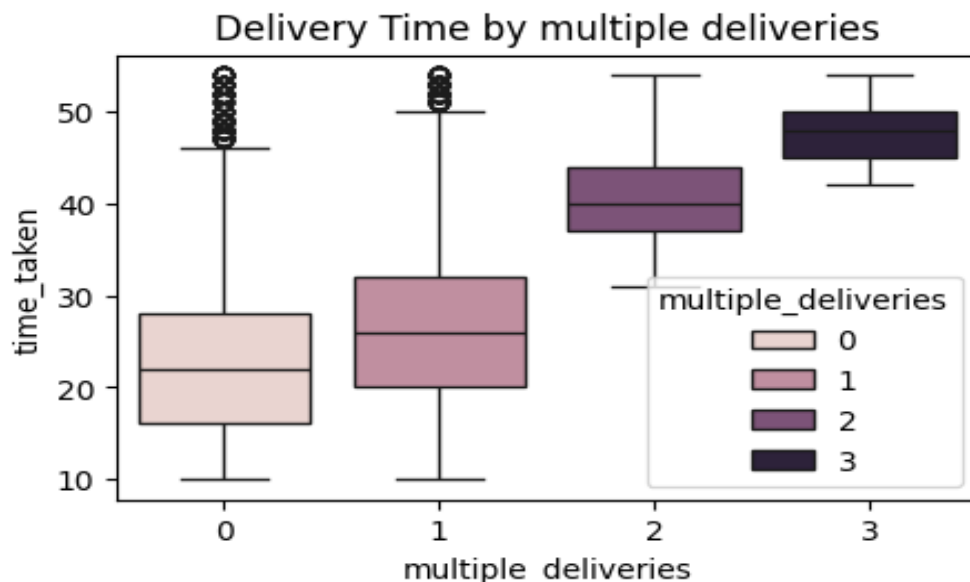
- Delivery time is longest under Jam and High traffic densities, with noticeable delays and variability.
- Medium and Low traffic densities show significantly reduced delivery times.
- **Insight:** Traffic conditions are a key determinant of delivery efficiency.
- **Actionable Suggestion:** Implement real-time traffic monitoring tools and prioritize traffic-aware route planning.

Delivery Time by Weather Conditions :



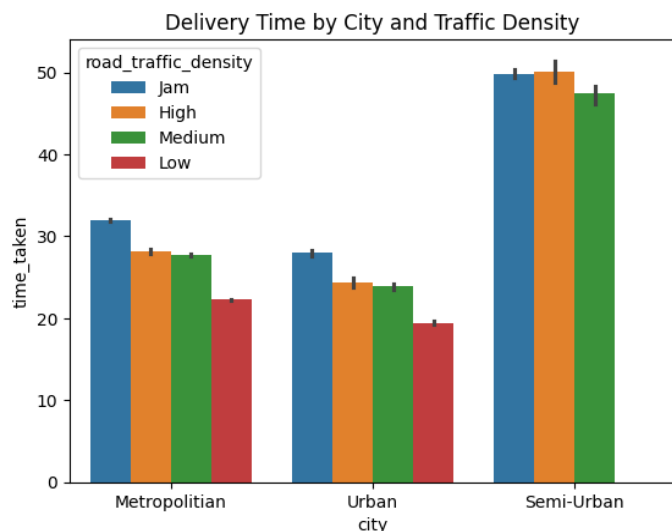
- Stormy and Cloudy conditions lead to longer delivery times, while Sunny weather enables quicker deliveries.
- Insight:** Weather conditions impact rider efficiency and road safety.
- Actionable Suggestion:** Adapt delivery strategies to include allowances for delays in adverse weather and incentivize riders during such conditions.

Delivery Time by Multiple Deliveries :



- Single deliveries are the quickest, with delivery times increasing as the number of simultaneous deliveries increases.
- Insight:** Overloading delivery personnel affects speed and customer satisfaction.
- Actionable Suggestion:** Optimize task allocation by limiting the number of simultaneous deliveries per rider based on real-time conditions.

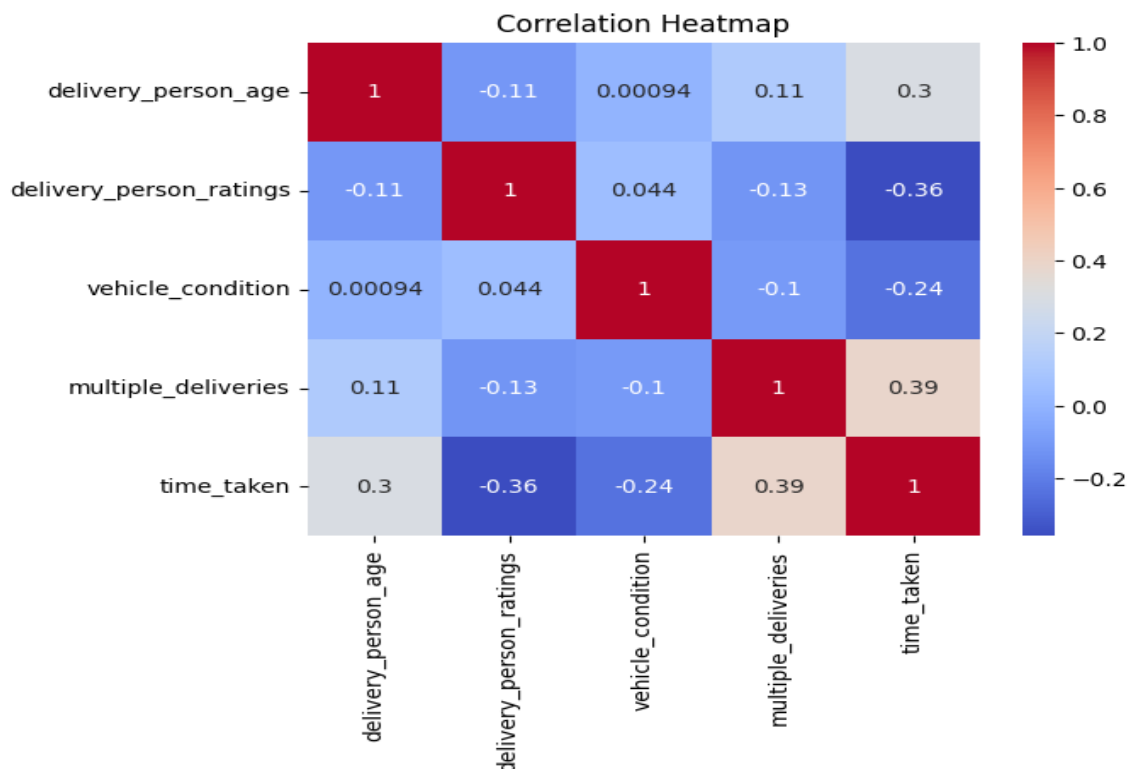
Delivery Time by City and Traffic Density :



- Semi-Urban areas experience the longest delivery delays across all traffic densities, while Metropolitan cities demonstrate better time management, even during high traffic.
- **Insight:** Infrastructure and delivery processes are more streamlined in urban centers compared to semi-urban regions.
- **Actionable Suggestion:** Improve logistics and delivery processes in Semi-Urban areas by deploying additional delivery personnel and vehicles.

3.Multivariate Analysis :

Correlation Heatmap :



- **Positive Correlations:**
 - Multiple deliveries positively correlate with delivery time, indicating delays increase with the workload.
- **Negative Correlations:**
 - Delivery person ratings drop with increased delivery time (-0.36), emphasizing the importance of timely service.
 - Vehicle condition has a minor but noticeable impact on efficiency, showing that poor maintenance can slow down deliveries.
- **Actionable Suggestion:** Invest in fleet maintenance, streamline multiple delivery assignments, and monitor delivery times closely to enhance customer experience.