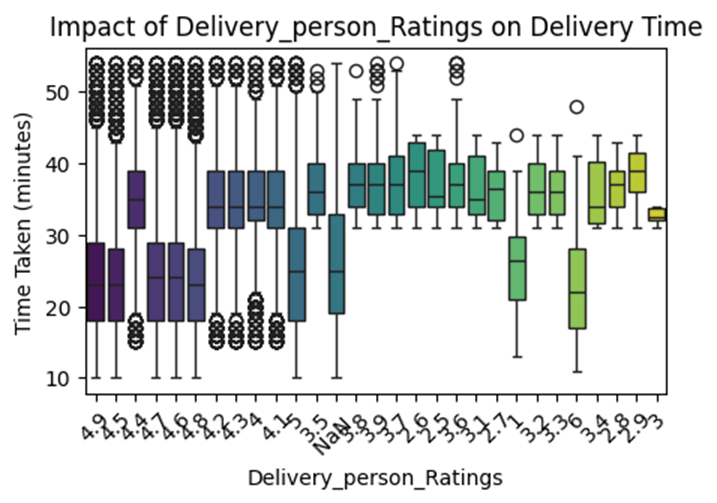
OUTPUTS & OBSERVATIONS

A chart of a number of boxes

AI-generated content may be incorrect.

A chart of different colored boxes

AI-generated content may be incorrect.A diagram of a box diagram

AI-generated content may be incorrect.A diagram of multiple delivery boxes

AI-generated content may be incorrect.A diagram of a vehicle condition

AI-generated content may be incorrect.A diagram of a road traffic density

AI-generated content may be incorrect.A chart of weather conditions

AI-generated content may be incorrect.A chart of a number of boxes

AI-generated content may be incorrect.

Some important insights for box plots

Type of vehicle, Weather conditions, Road traffic density, Vehicle condition, Multiple deliveries, Festival, City type, Delivery person ratings

1. Delivery people with higher ratings (e.g., 4.8, 4.9) generally complete deliveries faster, as shown by lower median times.
2. The NaN group shows a wide box and many outliers, suggesting inconsistent and possibly longer delivery times for delivery persons with missing ratings.
3. Deliveries in urban areas are the fastest, while deliveries in Semi-Urban areas are the slowest. Metropolitan areas have high variability, which may be due to factors like traffic density.
4. Festivals significantly increase delivery time due to higher demand and traffic congestion.
5. Traffic jams significantly increase delivery time, while low-traffic conditions enable faster and more predictable deliveries.
6. Adverse weather conditions like storms and sandstorms increase delivery time and variability.

A diagram of a number of numbers

AI-generated content may be incorrect.

The correlation heatmap shows that Delivery Person Age has a moderate positive correlation (0.30) with Time Taken. In contrast, Delivery Person Ratings have a negative correlation (-0.34), indicating that higher-rated delivery persons tend to complete deliveries faster, whereas distance has minimal impact on delivery time.

A group of purple and white bars

AI-generated content may be incorrect.

From, the frequency distribution charts:

Most deliveries occur in metropolitan areas using motorcycles under low or jammed traffic conditions.

Festivals and extreme weather conditions have fewer recorded deliveries, which may impact predictive modeling.

After building the predictive model, I got the Random Forest Regressor as the best fit with an R square of 0.82 and Root Mean Squared Error (RMSE): 3.9920

Random Forest R2 Score: 0.8182

Mean Absolute Error (MAE): 3.1756

Mean Squared Error (MSE): 15.9360

Root Mean Squared Error (RMSE): 3.9920

Using the Select K Best method, I analyzed the top features that impact the delivery time:

Multiple deliveries  
Delivery person Ratings  
Festival  
Delivery person Age  
Road traffic density

**Conclusion:**

Multiple deliveries, Road traffic density, and festivals introduce unavoidable delays, whereas Delivery person Ratings and Age can indicate efficiency and experience. Younger delivery people (20s to early 30s) tend to deliver faster, while older delivery people take slightly longer, possibly due to careful driving. Also, from the Box Plot: Delivery people with ratings above 4.5 have lower median delivery times, whereas those with lower ratings tend to take longer, likely due to inexperience or inefficiency.

All these observations align well with the dataset’s correlation, box plots, and frequency graph outputs as well.