Insights:

- In the dataset, there are 14817 unique trips between September, 2018 to October, 2018
- 1504 unique delivery routes available in data. 1508 unique source centres and 1481 unique destination centres are there.
- 72% training data and 28% test data are available in dataset
- There are 60% of route types are carting means small vehicles are used in 60% of route types and 40% of route type are FTL means full truck loads are used in 40% of route types
- Most of the trips (18.5%) in the data are created on Wednesday followed by Saturday
- Almost all states and union territories (total 31) are covered in India by Delhivery
- 1240 cities are covered in India by Delhivery
- Maximum trips are starting from Maharashtra state closely followed by Karnataka but in the city, Bengaluru is the source of highest no of trips followed by Gurgaon
- Most orders are coming from Maharashtra state again followed by Karnataka and in the city, most orders are coming from Bengaluru followed by Mumbai
- Delhi to Haryana and Haryana to Delhi are the busiest routes between states
- Most trips are happening in Bengaluru only followed by Mumbai and Bhiwani to Mumbai
- Bhiwandi to Mumbai and Mumbai to Bhiwandi are the busiest routes between cities followed by Delhi-Gurgaon
- Longest average distance covered by Delhivery is Guwahati to Bhiwandi followed by Chandigarh to Bengaluru
- Among the longest routes, most trips are happening between Gurgaon to Bengaluru

- **Statistical Test Results:**
    1. od_time_difference and start_scan_to_end_scan is similar. The result was expected because start_scan_to_end_scan is the time taken to deliver from source to destination and od_time_difference is the time difference between trip start time and trip end time, both are supposed to be almost same
    2. There is statistically significant difference between actual time and osrm time. It means actual time taken to complete the delivery and times provided by routing engine are different
    3. Actual time and segment actual time are similar. The result time was expected because the segment actual time is the sum of time taken in each segment of the delivery so it is supposed to be almost same to the actual time taken to complete the delivery
    4. There is statistically significant difference between osrm distance and segment osrm distance
    5. There is statistically significant difference between osrm time and segment osrm time

Recommendations:

- As we can see that there is difference between actual time and osrm time so optimization of scanning time at the start and the end can help reduce the actual time to complete the delivery
- As per analysis, it is better to use Carting (small vehicles) for delivery with in the city and Full Truck load (heavy vehicles) for long distance trips in order to reduce the time taken to complete the delivery
- As we can see that most trips are generating on Wednesday followed by Saturday so no of vehicles can be increased on these days to complete the delivery on time
- Similarly, no of vehicles can also be optimised according to busiest routes. For example, Maharashtra and Karnataka are the states having most trips so more no of vehicles can be made available in these states