# Key Findings and Recommendations:

1. **Data Understanding:**
    - The shape of Delhivery data is (144867, 24).
    - Identified 5 unknown fields, which can be dropped.
    - Observed missing area names in both `source_name` and `destination_name`.
    - Categorized columns into numerical and categorical fields.
2. **Categorical Columns:**
    - `data` and `route_type` identified as main categorical columns.
    - Mode of `data` is 'training', with frequent `route_type` as 'FTL'.
    - Top source and destinations are 'Gurgaon_Bilaspur_HB (Haryana)'.
3. **Data Imbalance:**
    - Training data is more than testing data.
    - 60% of transportation type is carting, and 40% is Full Truck Load (FTL).
4. **Outliers:**
    - Identified outliers in numerical columns, but left them untreated.
5. **Correlations:**
    - Positive correlations observed between `actual_time` and `segment_osrm_time_sum`.
    - High correlation (100%) between `actual_time` and `segment_actual_time_sum`.
    - High correlations (98%) between `osrm_distance` and `segment_osrm_distance_sum`, and `osrm_time` and `segment_osrm_time_sum`.
    - Possible glitches in ORSM navigation system.
6. **Temporal Patterns:**
    - More trips created in September.
    - Higher trip creation during night hours.
7. **Geographical Patterns:**
    - Maharashtra is the top state with the highest number of trips.
    - Gurgaon is the top city for trip creation, followed by Mumbai and Delhi.

# Recommendations:

- Address missing values in `source_name` and `destination_name`.
- Consider treating outliers in numerical columns based on the specific context.
- Investigate and improve the accuracy of the ORSM navigation system.
- Monitor and optimize logistics operations during night hours.
- Explore further insights into geographical patterns for targeted marketing and resource allocation.