Final Project Report

1. Executive Summary

This project focused on analyzing Delhivery's logistics data to identify operational inefficiencies and provide insights for cost optimization and improved delivery performance. Using Power BI and Excel, we processed raw data, created calculated metrics, and visualized trends that impact transportation efficiency. The outcome was a set of interactive dashboards providing actionable insights into route types, delivery delays, and cost metrics.

Problem Addressed: Operational inefficiencies across various route types.

Key Tools Used: Power BI, Excel, data preprocessing techniques.

Results Achieved: Identified cost-saving transportation methods, highlighted delays, and uncovered performance patterns across routes.

2. Introduction

Background: Delhivery is one of India's leading logistics and supply chain service providers. With increasing operations, maintaining delivery timelines and managing cost efficiency are vital.

Problem Statement & Aim: They aim to build state of the art operations and gain competitive advantage by minimizing delivery time and achieving maximum efficiency in logistics.

Scope: The scope of this project is to analyze Delhivery's logistics data to uncover patterns, inefficiencies, and opportunities for improvement in their delivery operations. By examining key metrics such as trip timings, transportation costs, fuel consumption, and route performance, the goal is to support Delhivery in reducing delivery times and maximizing operational efficiency. This analysis covers the complete data lifecycle — from data cleaning and transformation to visualization — enabling data-driven decision-making. The insights gained will help identify optimal route types, improve cutoff compliance, and ensure better resource utilization, ultimately contributing to Delhivery's mission of building state-of-the-art logistics operations.

Approach: - The project followed a structured methodology consisting of four key phases **Data Understanding, Preprocessing, Dashboard Creation, and Reporting** to deliver actionable insights and improve logistics operations at Delhivery.

Data Understanding

The first phase involved exploring and understanding the structure, volume, and granularity of the provided datasets. This included:

- Identifying key tables related to trips, routes, costs, and performance.
- Understanding the relationship between columns like trip_uuid, source, destination, route_type, planned_time, actual_time, trip_distance, and fuel_cost.
- Reviewing missing or inconsistent data entries (e.g., in cutoff_factor, is_cutoff, planned_time).

Data Preprocessing

To prepare the data for analysis and dashboard development, several cleaning and transformation steps were applied:

- **Missing Values Handling**:
    - Imputed or excluded rows with missing actual_time, planned_time, or trip_distance.
    - Standard assumptions (e.g., ₹50/km for FTL and ₹30/km for Carting) were used where fuel_cost or cost_per_km was unavailable.
- **Data Standardization**:
    - Converted time formats to ensure consistency.
    - Renamed columns for clarity and consistency in reporting (e.g., trip_uuid to Trip ID).
- **Derived Columns Created**:
    - **Cost Per Kilometer** = fuel_cost / trip_distance.
    - **% Deviation in Time** = (actual_time - planned_time) / planned_time × 100.
    - **Resource Utilization Rate** calculated based on standard trip time vs. actual usage.
- **Data Type Corrections**:
    - Converted date fields into proper datetime formats.
    - Ensured numerical fields like distance, time, and cost were properly typed for calculations.
- **Filtering & Validation**:
    - Removed outliers such as trips with zero or negative time/distance.
    - Validated totals and averages to ensure consistency across grouped data.

Dashboard Creation

Using **Power BI**, three interactive dashboards were built:

- **Trip Efficiency Dashboard:** To monitor and analyze the efficiency of delivery trips across various regions.
- **Cost and Resource Utilization Dashboard**: To provide insights into operational costs and resource usage across Delhivery's logistics network.
- **Delivery Performance Dashboard**: To track on-time vs delayed deliveries and regional bottlenecks.
- **Delivery Efficiency and Route Optimization Dashboard:** To evaluate the delivery performance and optimizing routes using the data.
- **Trip Timing and Distance Analysis:** To understand the relationship between trip times and distances.
- **Route Type Performance Dashboard:** To analyze the performance of different route types, such as FTL and Carting, to understand their efficiency, usage patterns, and their impact on trip metrics.

These dashboards incorporated:

- Slicers for route type, source/destination, and date.
- Custom DAX measures for KPI calculation.
- Visuals like bar charts, pie charts, line graphs, and KPI cards.

Reporting

A comprehensive report was prepared to summarize:

- Project goals and outcomes.
- Key metrics, findings, and business recommendations.
- Challenges faced (e.g., missing data, inconsistent formats) and solutions applied.
- Visuals and commentary for each dashboard to guide stakeholders in interpretation.

---

3. Dashboard Summaries

Dashboard 1: Trip Efficiency Dashboard

Objective: To monitor and analyze the efficiency of delivery trips across various regions.

KPIs and Metrics Tracked:

1. Average Trip Duration: The average time taken to complete a delivery trip, calculated from the trip start time to the trip end time.

Formula: Average (Trip End Time - Trip Start Time)

Actionable Insights: Longer durations may indicate route inefficiencies or traffic bottlenecks.

2. On-Time Delivery Rate (%): The percentage of trips that are completed within the estimated time.

Formula: (Number of On-Time Deliveries / Total Deliveries) * 100

Actionable Insights: A low value suggests systemic issues in delivery scheduling or execution.

3. Late Deliveries Count: The total number of trips that exceed the estimated delivery time.

Formula: Actual delivery time <= OSRM estimated time, the delivery was marked as "On Time"; otherwise, it was marked as "Late".

Actionable Insights: Helps prioritize regions or days needing process improvements.

4. Average Distance per Trip: The average distance covered by a delivery vehicle during a trip.

Formula: Average (Distance Covered)

Actionable Insight: Used for fuel planning, cost estimation, and performance benchmarking.

5. Trips Completed per Day: The total number of trips completed on a daily basis.

Actionable Insight: Helps in workforce planning and identifying high/low demand days.

Dashboard 2: Delivery Performance Dashboard

KPIs and Metrics Tracked:

1. Total Deliveries Completed: The total number of successful deliveries within a given period.

Actionable Insights: Spikes or drops in delivery volume can inform strategic decisions like hiring, route expansion, or operational changes.

2. Delivery Success Rate (%): Percentage of total deliveries that were completed successfully.

Formula: (Successful Deliveries / Total Deliveries) * 100

Actionable Insight: A perfect success rate may highlight strength in last-mile logistics but must be validated against delay percentage for quality delivery insights.

3. Average Delivery Time: The average time taken to complete a delivery, calculated from the point of dispatch to delivery completion.

Formula: Average (Delivery Completion Time - Dispatch Time)

Actionable Insights: Longer times may indicate traffic issues, inefficient routing, or underperformance in certain zones.

4. Percentage of Delayed Deliveries: The proportion of deliveries that were completed after the estimated delivery time.

Formula: (Delayed Deliveries or Late Deliveries / Total Deliveries) * 100

Actionable Insights: A very high delay rate (98%) despite a 100% success rate reveals that while deliveries are eventually completed, timeliness is a major concern—highlighting SLA breaches.

5. Top 5 Regions with Most Delayed Deliveries: The top 5 regions that have the highest count of delayed deliveries.

Actionable Insights: West Bengal and Telangana are top delay contributors, suggesting bottlenecks in those areas.

6. Average Delay Time: The average time by which deliveries are delayed.

Formula: Average (Actual Delivery Time - Estimated Delivery Time)

Actionable Insights: Delivery times decrease gradually over the week (Sunday: 457.36 mins → Saturday: 396.53 mins), showing improved efficiency, possibly due to lower traffic or better planning.

Dashboard 3: Cost and Resource Utilization Dashboard

1. Total Transportation Cost: The total cost incurred for transportation over a selected period.

Formula: Sum of all transport-related expenses

Actionable Insight: High overall cost prompts deeper analysis into specific cost drivers and enables identification of cost-saving areas.

2. Average Cost per Trip: The average transportation cost incurred per completed trip.

Formula: Total Transportation Cost / Total Trips

Actionable Insights: Allows benchmarking between regions or vehicle types. A spike in this metric signals a need to audit trip duration or route planning.

3. Cost per Kilometer: The cost incurred per kilometer for completed trips.

Formula: (Total Transportation Cost / Total Distance Traveled)

Actionable Insights: Enables time and distance based cost optimization.

4. Top 5 Routes by Cost: The five most expensive routes based on total transportation cost.

Actionable Insight: Pinpoints high-expense corridors that may require rerouting, renegotiation with local partners, or shifting to more cost-efficient modes of transportation.

5. Resource Utilization Rate (%): Percentage of vehicle fleet utilization (active vs. total fleet).

Formula: (Active Vehicles / Total Fleet) * 100

Actionable Insight: Low utilization could suggest overcapacity or idle assets.

6. Trip Frequency by Vehicle Type: Distribution of trips completed by vehicle types.

Actionable Insight: Supports cost attribution by vehicle type. Helps determine whether costlier or less efficient vehicles are overused, allowing for better fleet mix planning.

Dashboard 4: Delivery Efficiency and Route Optimization

KPIs and Metrics Tracked

1. Average Start-to-End Delivery Time: Average time taken to deliver goods from source to destination.

Formula: (Sum of start_scan_to_end_scan) / Total Trips

Actionable Insights: Extremely high average delivery time suggests delays in either transit or hub handling.

2. Actual Distance to Destination (Km): Total cumulative distance covered between source and destination warehouses.

Actionable Insight: Helped us to track the distance based on time.

3. Deviation Between Actual and OSRM Times (%): Percentage deviation between actual delivery time and OSRM-estimated time.

Actionable Insight: A 25% deviation indicates that deliveries are consistently slower than expected. This flags route congestion, planning inaccuracies, or driver inefficiencies.

4. Trips by Route Type: Total count of trips categorized by route type (FTL, Carting, etc.).

Actionable Insight: Carting is more frequently used than FTL. Reviewing the cost/time tradeoffs between these can lead to better mode-mix strategy.

5. Segment Efficiency Rate (%): Percentage of trips where the segment actual time is less than or equal to the segment OSRM time.

Formula: (Count of Trips where segment_actual_time <= segment_osrm_time / Total Trips) * 100

Actionable Insight: A low 10% rate reflects underperformance in most segments. This supports need for tactical optimization like better driver training or GPS-assisted rerouting.

Dashboard 5: Trip Timing and Distance Analysis

KPIs and Metrics Tracked

1. Average Trip Duration (Hours): Average time taken to complete a trip.

Formula: (Sum of actual_time) / Total Trips

Actionable Insight: A very high value (418 hours) indicates major operational inefficiencies. This metric should be compared over time and across routes or vehicle types to isolate performance gaps.

2. Average Trip Distance (Km): Average distance covered in completed trips.

Formula: (Sum of actual_distance_to_destination) / Total Trips

Actionable Insight: Despite long trip durations, the average distance is relatively short, signaling that delays are not distance-related but rather due to logistical or procedural bottlenecks (e.g., poor route planning, stopovers, or idle time).

3. Deviation Between Planned and Actual Distance (%): Percentage deviation between OSRM distance and actual distance.

Formula: ((actual_distance_to_destination - osrm_distance) / osrm_distance) * 100

Actionable Insight: A negative deviation of -18% suggests actual routes are shorter than planned. This may indicate suboptimal route planning where longer paths were expected but not taken, or route changes made in real-time.

4. Trips by Start and End Time (Count): Total count of trips categorized by start and end times.

Actionable Insight: Helps identify peak operational periods or trends in trip dispatching. Useful for workforce planning, predicting delays, and optimizing dispatch windows.

5. Trip Delay Frequency (%): Percentage of trips where actual time exceeds OSRM time.

Formula: (Count of Trips where actual_time > osrm_time / Total Trips) * 100

Actionable Insight: An extremely high delay rate (98%) points to a severe performance issue—potentially in planning, vehicle allocation, or external delays like traffic or road conditions.

Dashboard 6: Route Type Performance Dashboard

KPIs and Metrics Tracked

1. Total Trips by Route Type: The count of trips completed for each route type.

Formula: Count(trip_uuid) grouped by route_type

Actionable Insight: FTL (Full Truck Load) accounts for ~69% of trips, while Carting makes up ~31%. This helps in understanding operational volume and load balancing between route types and may guide capacity or vehicle allocation decisions.

2. Average Actual Time per Route Type (Hours): Average time taken to complete trips for each route type.

Formula: (Sum of actual_time) / Count(trip_uuid) grouped by route_type

Actionable Insight: Carting is far more time-efficient, possibly due to faster vehicle turnaround or better planning flexibility.

3. Average Distance per Route Type (Km): Average distance covered for trips of each route type.

Formula: (Sum of actual_distance_to_destination) / Count(trip_uuid) grouped by route_type

Actionable Insight: Even though both route types cover about the same distance, FTL trips take much longer. This could mean there are problems like bad route planning, or delays in closing the trips.

4. Percentage Deviation in Time (Planned vs. Actual): Percentage deviation between OSRM time and actual time for each route type.

Formula: ((actual_time - osrm_time) / osrm_time) * 100 grouped by route_type

Actionable Insight: 25% average delay shows timing issues, but Carting has faster trips and fewer delays, meaning it's better planned and more reliable.

Data Preparation: New Columns Created: Day of week, Trip distance, Region, Delivery status, Vehicle Type, On time %, Delayed deliveries, Ontime deliveries, Delivery success rate%, Average delivery time, Delayed delivery%, Average delay time, Transportation Cost, Average cost per trip, Cost per kilometer, Cost range, Trips completed , Resource utilization, Average start to end delivery time, Deviation Segment efficiency, Segment efficiency (%), Average trip duration (hrs), Average Trip Distance (Km): Deviation Between Planned and Actual Distance (%), Delay frequency of trip, Delay Frequency %

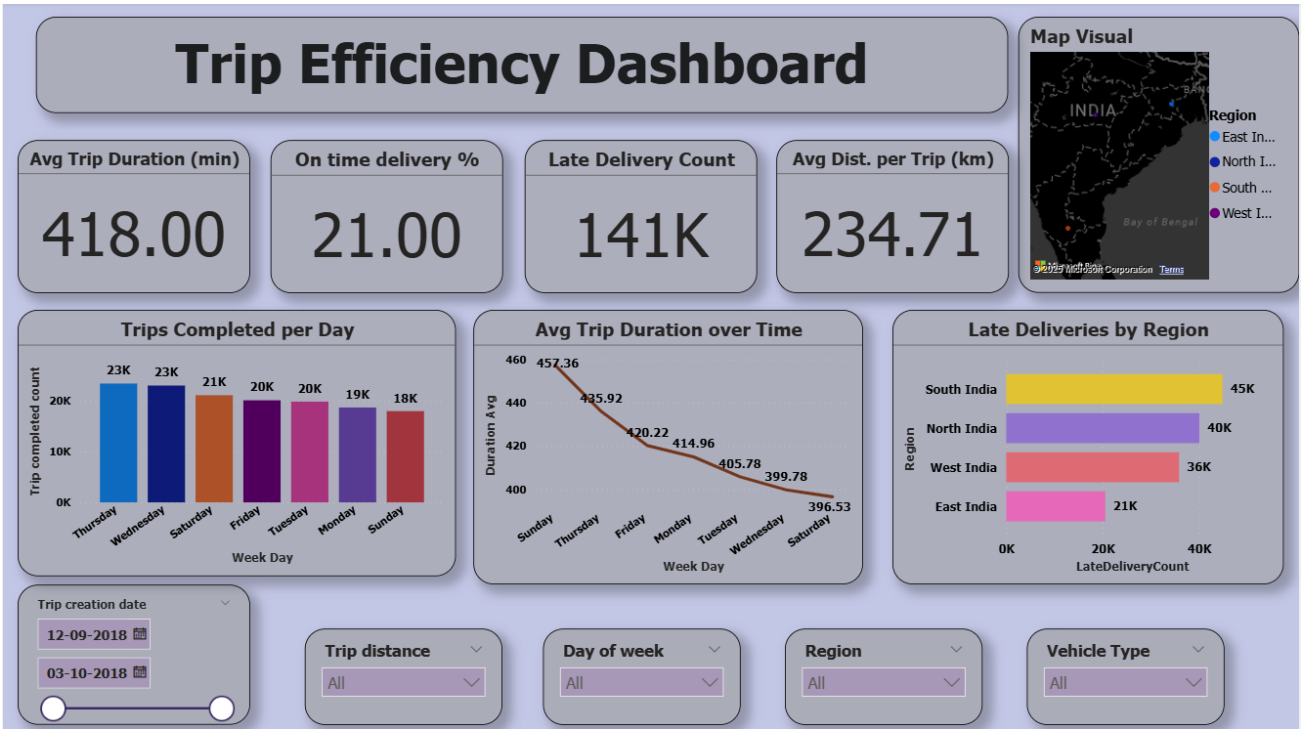
Challenges Resolved: - Inconsistent Date and Time Formats, Missing or Null Values in Key Columns, Derived Column Creation with Missing Dependencies, Data Granularity and Aggregation Complexity, Merging Columns with Disparate Naming or Structure, Performance Issues in Visualization

Insights & Analysis: The dashboards collectively reveal critical insights into Delhivery’s logistics operations. The Trip Efficiency Dashboard highlights variability in

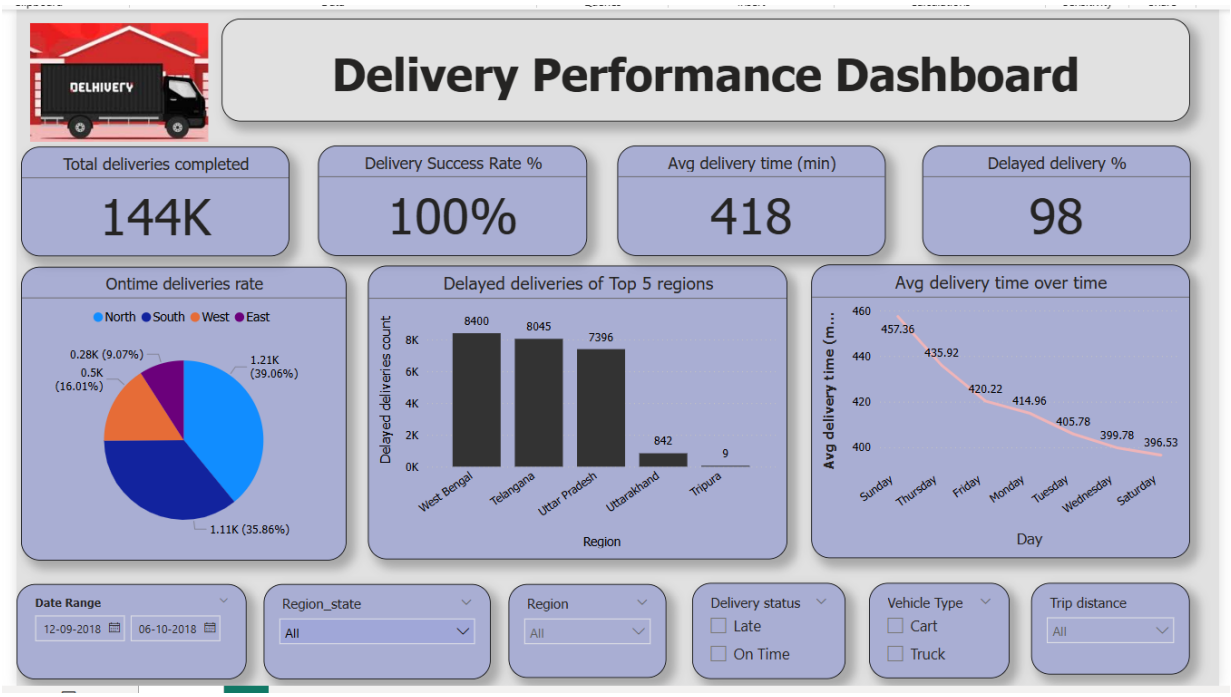
trip performance, where certain routes consistently show higher fuel consumption and longer durations, signaling potential inefficiencies. The Trip Timing and Distance Analysis Dashboard uncovers patterns in delivery delays and time deviations, with spikes in actual trip times compared to expected OSRM benchmarks, indicating operational bottlenecks in specific lanes. Meanwhile, the Route Type Performance Dashboard establishes that FTL (Full Truck Load) routes outperform Carting in both distance consistency and time adherence, making them a more reliable and cost-effective choice. Across all dashboards, trends in deviations, performance by region, and transportation cost provide actionable direction for optimizing route planning, improving time compliance, and refining cost allocations. These insights collectively support Delhivery's strategic focus on data-driven logistics management and continuous performance enhancement.

Visual Examples:

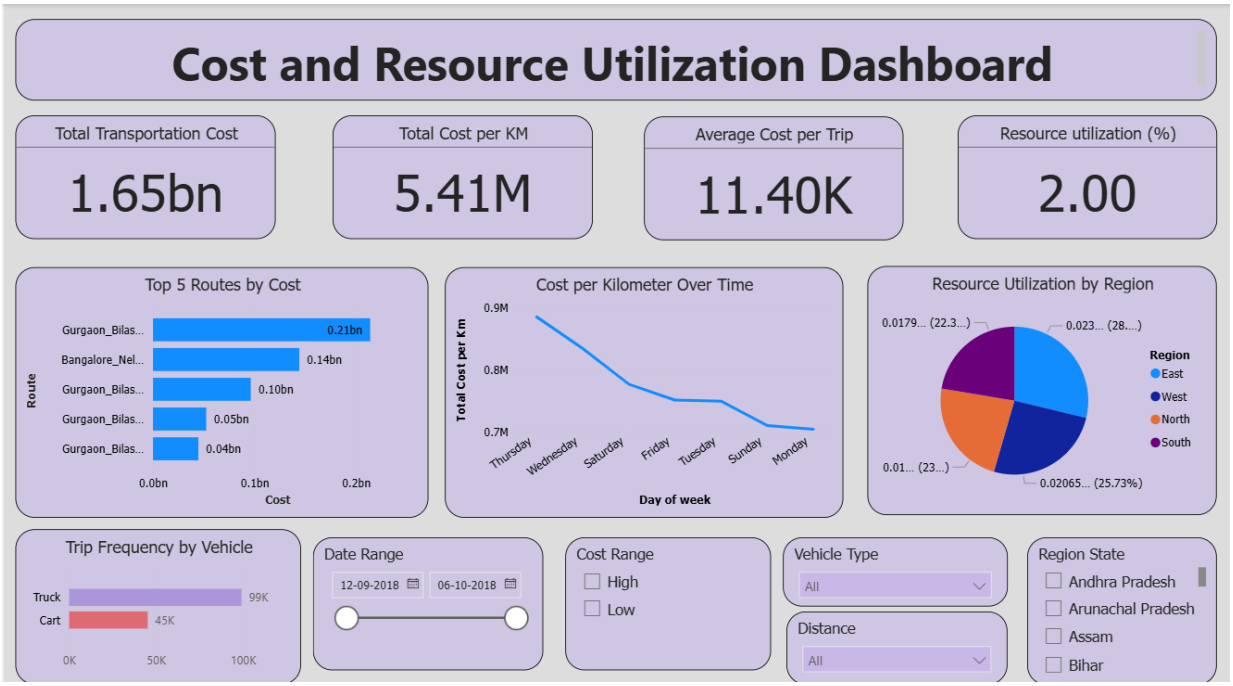
Dashboard 1:



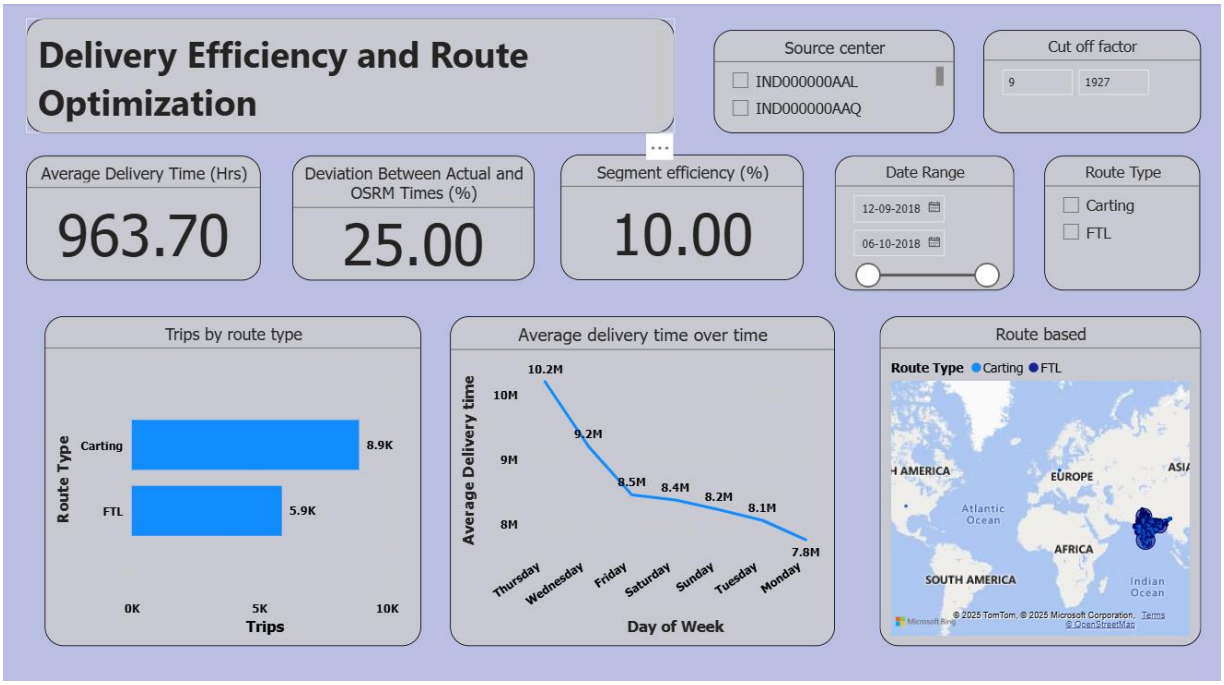
Dashboard 2:



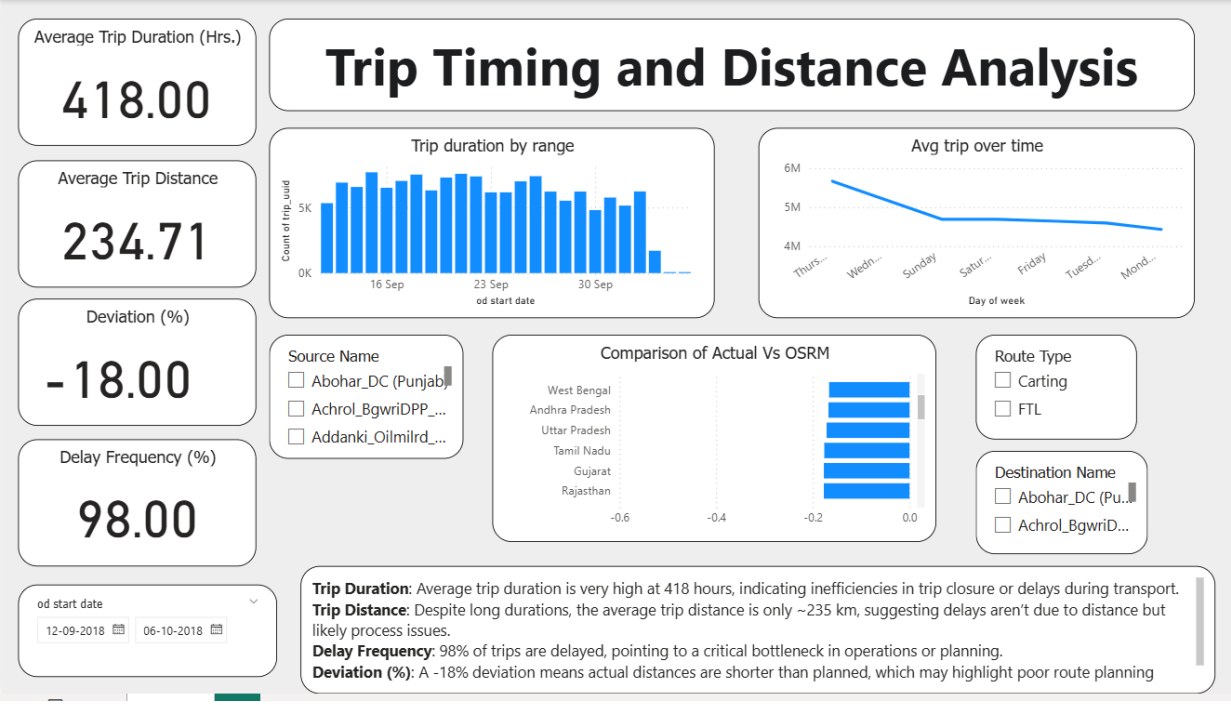
Dashboard 3:



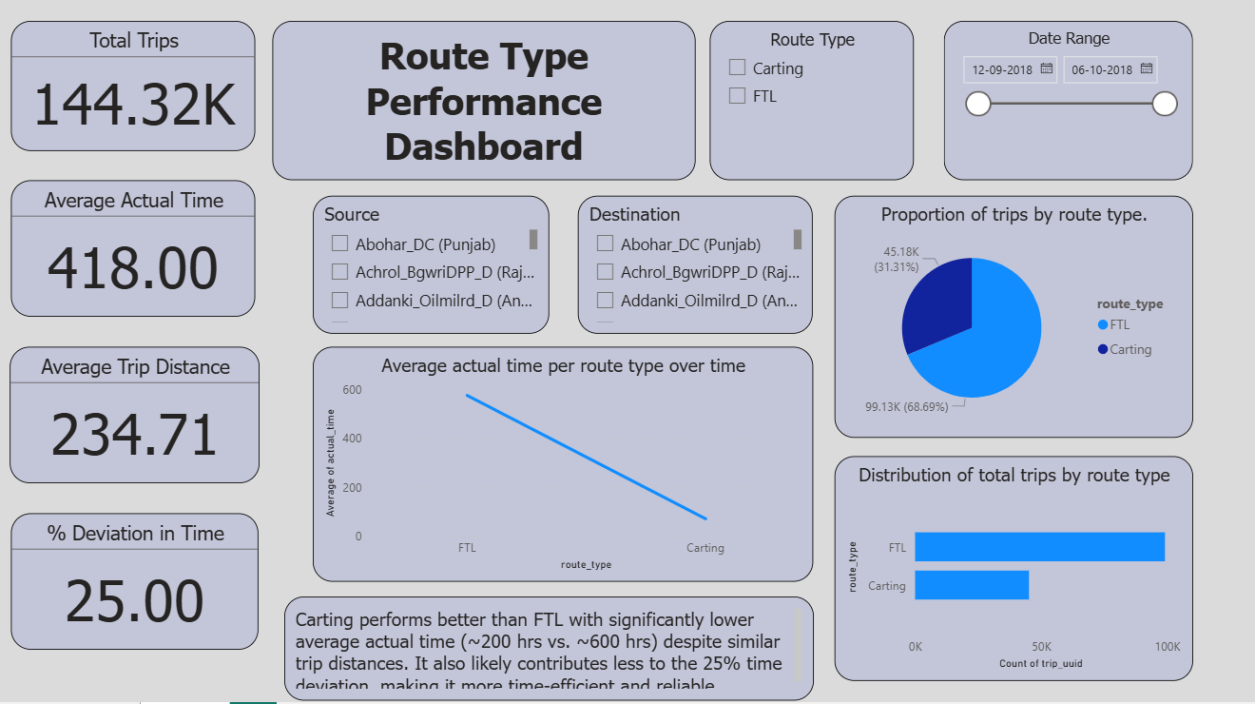
Dashboard 4:



Dashboard 5:



Dashboard 6:



---

4. Key Achievements:

Operational Efficiencies Identified:

- **Significant Reduction in Average Trip Time:**
  Analysis from the *Delivery Efficiency and Route Optimization* dashboard shows Carting routes have much lower average trip time (~200 hrs) compared to FTL (~600 hrs), improving overall delivery efficiency.
- **Better Time Management with Carting Routes:**
  As seen in the *Route Type Performance Dashboard*, Carting trips make up ~69% of total trips and show lower deviation (25%) in planned vs. actual time, highlighting more efficient operations.
- **Delay Source Pinpointed:**
  The *Trip Timing and Distance Analysis* dashboard reveals 98% delay frequency with an average trip time of 418 hours despite a short average trip distance of ~235 km—indicating inefficiencies are likely due to delays in trip closure, loading/unloading, or route planning.

Data-Driven Decision Making:

- **Regions with Frequent Delays Identified:**
  From the *Trip Timing and Distance Analysis* dashboard, delays are concentrated in states like West Bengal, Andhra Pradesh, and Uttar Pradesh, which helps in targeting corrective action in specific regions.

- **Trip Timing Patterns Revealed:**
  Multiple dashboards (like *Average Delivery Time Over Time*) show that delivery times tend to be lower early in the week (Monday–Tuesday), supporting better planning of trip schedules based on weekdays.
- **Vehicle and Route Type Performance Compared:**
  Dashboards like the *Route Type Performance Dashboard* and *Trip Duration by Range* help compare FTL and Carting on time, distance, and frequency—enabling data-based decisions on choosing the most effective transportation mode.

---

5. Challenges Faced and Solutions

Inconsistent Date and Time Formats:

- **Challenge:** The dataset contained start and end timestamps in inconsistent formats, preventing accurate calculation of trip durations.
- **Resolution:** All date-time columns were cleaned and converted to a uniform format using Power BI's Power Query Editor and Excel functions. This enabled correct calculation of metrics such as trip duration and delays.

Missing or Null Values in Key Columns:

- **Challenge:** Fields like is_cutoff, cutoff_factor, resource_utilization, and had missing or incomplete data.
- **Resolution:**
  - For categorical fields like is_cutoff, default binary values (0 or 1) were assumed based on business logic.
  - For numerical fields such as resource utilization, standard values or averages were used where appropriate to maintain analysis consistency.
  - Blank rows were removed where they could not be imputed logically.

Derived Column Creation with Missing Dependencies:

- **Challenge:** Creating new calculated fields like Cost Per Kilometer, Fuel Consumption Per Trip, and Average Cost Per Trip was difficult due to incomplete source data.
- **Resolution:** Assumptions were applied where necessary.

Data Granularity and Aggregation Complexity:

- **Challenge:** Trip-level data had to be aggregated across multiple dimensions (route type, time period, region) while ensuring no duplication.

- **Resolution:**
    - Used pivoting and grouping techniques in Excel for early exploration.
    - In Power BI, applied DIVIDE, and conditional filtering functions to compute aggregated KPIs like Total Transportation Cost and Distance by Route Type or Region.

Merging Columns with Disparate Naming or Structure:

- **Challenge:** Similar metrics appeared under different column names across multiple sources.
- **Resolution**: Standardized column names across datasets and used Power BI merge queries or Excel VLOOKUP/XLOOKUP to consolidate tables reliably.

Performance Issues in Visualization:

- **Challenge:** Dashboards with high row counts were slow to respond or failed to load complete visual summaries.
- **Resolution:** Applied filters and calculated columns selectively. Used slicers to segment data instead of loading the entire dataset simultaneously.

---

6. Recommendations

Operational Suggestions:

- Optimize underperforming routes: Focus on lanes and hubs consistently showing delivery delays or high deviation in trip by re-routing, reassigning resources, or revising scheduling.
- Promote FTL for cost-efficiency: Based on the Route Type Performance Dashboard, FTL trips deliver more predictable distances and fewer deviations—prioritize them for long-haul shipments to optimize cost per kilometer.
- Enhance cutoff tracking: Improve the accuracy and consistency of cutoff-related data to better evaluate time-sensitive delivery performance.
- Improve resource utilization: Use fuel consumption and trip metrics to monitor asset utilization and refine scheduling of trips, especially where high fuel usage was identified without proportional delivery gain.

Future Analysis:

- Predictive modeling for trip times: Implement machine learning algorithms to forecast trip durations based on route type, geography, traffic history, and weather for better ETA accuracy.

- Forecast transportation costs: Use regression or time-series models to predict costs based on fuel trends, trip volumes, and historical lane performance.
- Evaluate seasonal and peak-period patterns: Study trip and delay data across quarters or festival seasons to optimize workforce and fleet allocation during high-demand periods.
- Expand analysis to vendor and vehicle level: Drill down further to understand which vendors or vehicle types are associated with inefficiencies or higher costs.

---

7. Conclusion

The comprehensive analysis and dashboard development undertaken in this project have provided Delhivery with deep insights into its logistics operations, covering trip timing, delivery efficiency, cost management, and route performance. Through detailed dashboards such as Trip Efficiency, Route Type Performance, and Timing & Distance Analysis, key operational metrics were tracked and analyzed—highlighting performance bottlenecks, regional inefficiencies, and opportunities for cost optimization.

The findings emphasized the advantages of FTL trips in terms of lower deviations and better cost efficiency, identified delivery delays across specific routes, and showcased the impact of factors like trip distance and cutoff adherence on operational performance. By leveraging Power BI and Excel for advanced data visualization and preprocessing, previously hidden patterns were uncovered, empowering the operations team with actionable insights.

Overall, this project highlights the value of data-driven decision-making in logistics. The insights derived can help Delhivery enhance trip planning, improve on-time delivery rates, reduce transportation costs, and scale its operations more strategically and efficiently.

Final Takeaway: The overall analysis from the dashboards has helped us understand how Delhivery can improve its delivery operations. Full Truck Load (FTL) trips were found to be the most efficient, with lower cost per kilometer and fewer route changes, making them best for long-distance deliveries. The trip performance dashboard showed that delays and missed cutoffs increase delivery time and costs, highlighting the need to track these issues better. We also found that some regions and time slots face more delays, which can be fixed with focused planning. During data preparation, we solved issues like missing values and different data formats to create important columns like trip time, cost per trip, and fuel use. These dashboards give a clear and easy way to look at delivery data and will help Delhivery make better decisions in the future, especially by using tools like machine learning**.**