Using BI to Optimize Last-Mile Logistics

Objective:

This project explores a large dataset of Amazon deliveries to uncover insights that can improve delivery efficiency and optimize operations. The analysis focuses on customer behavior, delivery performance, product trends, and traffic impacts—both overall and by area.

1. Ask the Right Questions (Define the Problem)

Guiding Questions:

- What factors affect delivery efficiency?
- Which product categories are most and least popular?
- How do traffic, delivery area, and vehicle type impact transit time?
- When are orders most frequently placed?
- What patterns exist in agent performance across different conditions?

2. Collect the Data

Dataset Source:

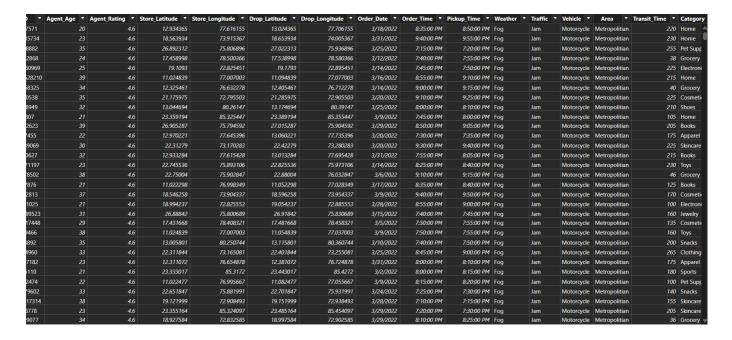
The dataset used for this analysis was downloaded from Kaggle and can be found at Amazon Delivery Dataset

Key Features:

Size: 43,648 rows x 16 columns

Features:

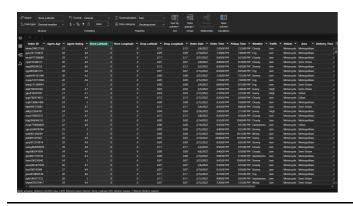
- Order_ID
- Agent_Age (numeric)
- Agent_Rating (decimal)
- Transit Time (in minutes)
- Traffic, Weather, Vehicle (categorical)
- Area, Order_Date, Category

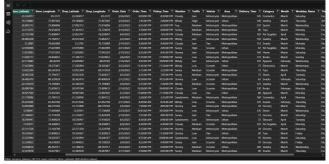


3. Clean & Prepare the Data

Data Cleaning Tasks:

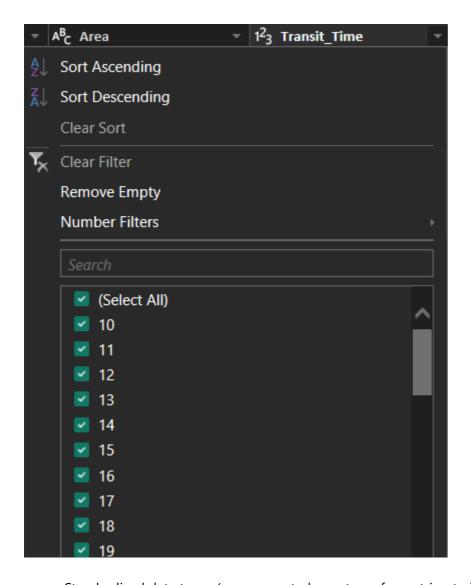
• Removed invalid geolocation values (Store_Latitude/Store_Longitude = 0 - in the Indian Ocean)





Before After

Dropped rows with a blank in the Transit Time column to preserve metric integrity



- Standardized data types (e.g., converted agent age from string to integer)
- Capitalized first letter of Vehicle types for better visuals

Feature Engineering:

• Created Month, Weekday, Weekday_Name columns using Power Query date functions

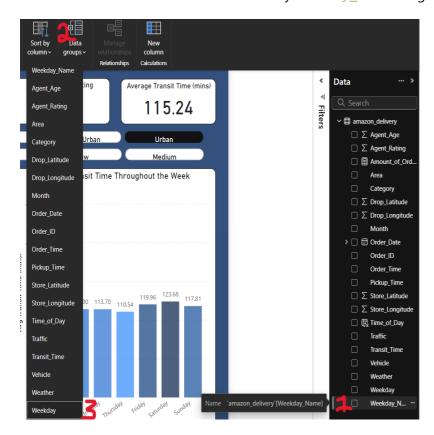






- Built Time_of_Day column using DAX SWITCH() for Morning/Afternoon/Evening/Night
- Created Amount_Of_Orders measure to track volume

Ensured bar chart would be sorted by Weekday_Name using numerical Weekday values



4. Analyze the Data

a. Time-Series Trends:

- Significant order increases from Feb 18 Mar 1 (likely due to end-of-month promotions)
- March had the most complete data; February and April had partial coverage
- Delivery times varied significantly based on traffic and area:
 - o Fastest: low-traffic urban areas
 - Slowest: semi-urban in high traffic

b. Traffic Pattern Insights:

- Traffic Jams = longest average transit times (148 mins), lowest ratings
- Low traffic = quickest deliveries (101 mins), highest ratings

c. Descriptive & Exploratory Analysis:

- Majority of deliveries were made to Metropolitan areas (30K+ orders)
- Motorcycles were the most used delivery vehicle across all areas
- Evenings had the highest order activity, especially on Wednesdays and Sundays
- Electronics, books, and jewelry were the top products purchased

5. Share & Visualize Results

4.a Time-Series Trends Visuals:

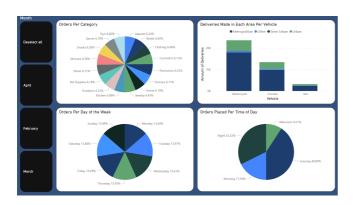


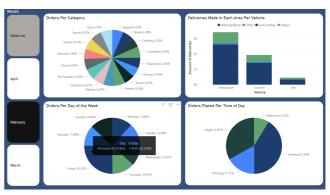
4.b Traffic Pattern Insights Visuals:

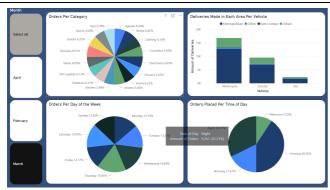


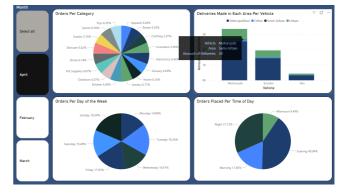


4.c Descriptive & Exploratory Analysis Visuals:









6. Act (Draw Conclusions & Recommend Actions)

Key Takeaways:

- Transit time influences agent ratings the quicker the delivery, the better the customer feedback
- Agents riding motorcycles in low traffic, urban areas are the fastest and get the highest ratings
- Delivery optimization should prioritize time of day, vehicle type, and regional traffic trends
- High traffic times correlate with Afternoon deliveries that consist mostly of skincare, sports, and pet supplies products
- Evening orders dominate—especially on Wednesdays and Sundays

Recommendations:

- Route deliveries to Urban/Other zones during low-traffic times
- Assign motorcycle deliveries for congested routes
- Focus on high-efficiency time slots (Evenings, Weds/Thurs) to deliver packages
- Use agent performance data for training purposes
- Develop demand-based staffing forecasts by day & hour