PES UNIVERSITY
100 feet Ring Road, BSK 3rd Stage
Bengaluru 560085

Department of Computer Science and Engineering
B. Tech. CSE - 6th Semester
Jan – May 2025

UE22CS343BA1
SUPPLY CHAIN MANAGEMENT (SCM)
For Engineers

PROJECT REPORT
on

Delivery Visibility and Performance Dashboard

Submitted by: Team #: 032-045-172-248

PES1UG22CS045	Aditya Shankar	CSE-6A
PES1UG22CS032	Aditi Vadmal	CSE-6A

PES1UG22EC248	Saanvi Shetty	ECE-6E
PES1UG22EC172	Narreddy Pedda Vemanna Gari Bavana	ECE-6C

Class of _Prof. Raghu B. A._

Table of Contents

Introduction

With the fast-paced nature of today's supply chain, immediate visibility into delivery operations is critical to on-time delivery of orders and minimizing logistics-related inefficiencies. Delayed delivery not only impacts customer satisfaction but also incurs additional transportation and fuel costs, leading to total supply chain disruption.

As a response, our project will develop an interactive dashboard offering detailed insight into delivery performance. This entails tracking average time of delivery, on-time delivery rates, fuel spending per vehicle category, and delivery timescales distribution.

From a synthetic dataset simulating 500 delivery records alongside an Amazon Delivery dataset from Kaggle, we graphed and examined key delivery metrics by vehicle type (Truck, Van, Bike) and date. Supply chain managers are empowered with data-driven decision making by exposing real-time delivery inefficiencies and cost chokepoints.

Scope & Objective:

Scope:

- Build an interactive Power BI dashboard to visualize delivery metrics using a synthetic dataset.
- Analyse delivery time, fuel usage, cost, and performance.
- Apply machine learning to predict late deliveries based on delivery characteristics.

Objectives:

- Design KPI visuals and filterable charts for operational delivery analysis.
- Identify inefficiencies in delivery performance across vehicle types and time periods.
- Train and apply a logistic regression model to classify potential late deliveries.
- Integrate ML predictions into the Power BI dashboard for decision-making support.

Data Inputs & Datasets Used :

This project utilizes **two datasets** — one synthetic and one real-world — each serving a distinct purpose in achieving the project's KPIs and ML objectives.

1. Synthetic Delivery Dataset

- **Source:**
  Custom-generated using Python.
- **Purpose:**
  Used as the primary dataset for both Power BI visualizations and machine learning prediction.
- **Why Synthetic?**
  While the Amazon dataset provided real-world delivery conditions, it lacked certain operational KPIs critical to our problem scope — such as **fuel consumption**, **cost per delivery**, and **precise SLA time tracking**. Hence, a synthetic dataset was created to simulate realistic delivery operations while including these key performance metrics.
- **Key Fields:**

Column	Description
delivery_id	*Unique delivery identifier*
delivery_time_min	*Actual time taken for delivery*
expected_time_min	*SLA time allowed for delivery*
distance_km	*Distance travelled*
fuel_consumed_litres	*Fuel consumed during the trip*
delivery_cost	*Cost incurred*
driver_rating	*Rating of driver performance*
vehicle_type	*Truck, Van, or Bike*
is_late	*1 = Late delivery, 0 = On-time*
predicted_late	*ML prediction of delivery outcome*

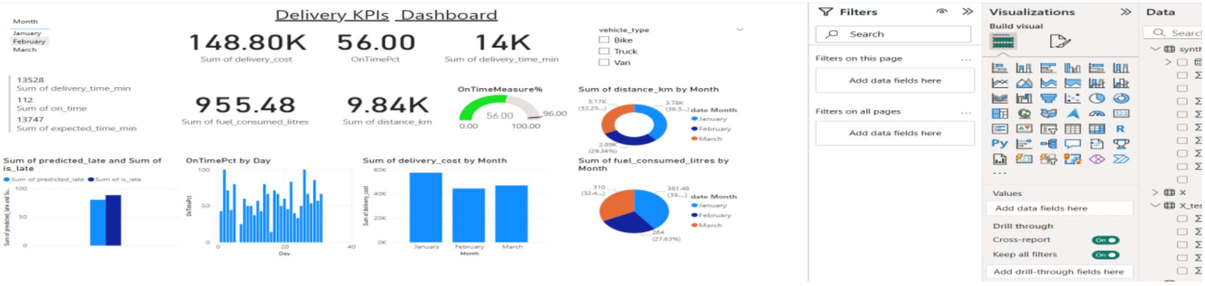
2. Amazon Delivery Dataset (Kaggle)

- **Source:**
  Kaggle (Amazon Last Mile Delivery dataset)
- **Purpose:**
  It was used for exploratory analysis and to understand real-world delivery variables such as agent ratings, traffic, and weather. While it offers rich real-life scenarios, it did not include essential KPIs like fuel usage or delivery cost, and hence was not sufficient for end-to-end KPI analysis or ML implementation in our use case. Therefore, we combined it with another dataset that was synthesized using Python's Scikit Learn ML library for better visualization of the identified important delivery KPIs.
- **Relevant Fields Explored:**

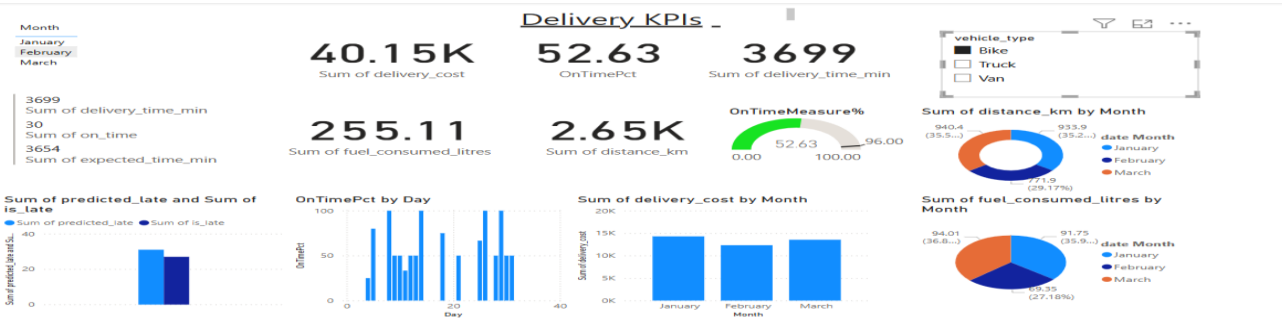
Column	Description
Delivery_person_Age	Age of delivery agent
Delivery_person_Ratings	Average rating of agent
Order_Date, Time_Orderd, Time_Order_picked	Timing details
Weatherconditions	Weather during delivery
Road_traffic_density	Traffic levels
Type_of_order, Type_of_vehicle	Order/vehicle info

Conclusion:
The synthetic dataset was crafted to align with the exact KPIs needed for delivery performance monitoring and ML-based delay prediction, while the Amazon dataset served to validate feature design and to offer real-world context.

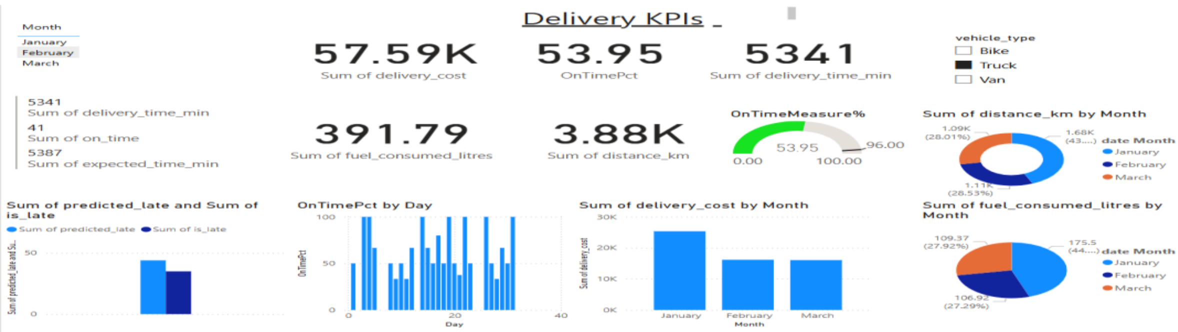
Dashboard & Graph Images :



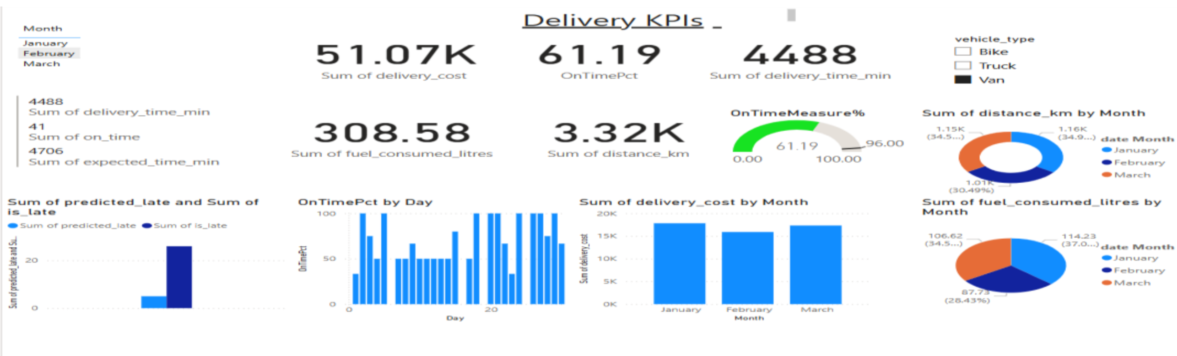
Bike :

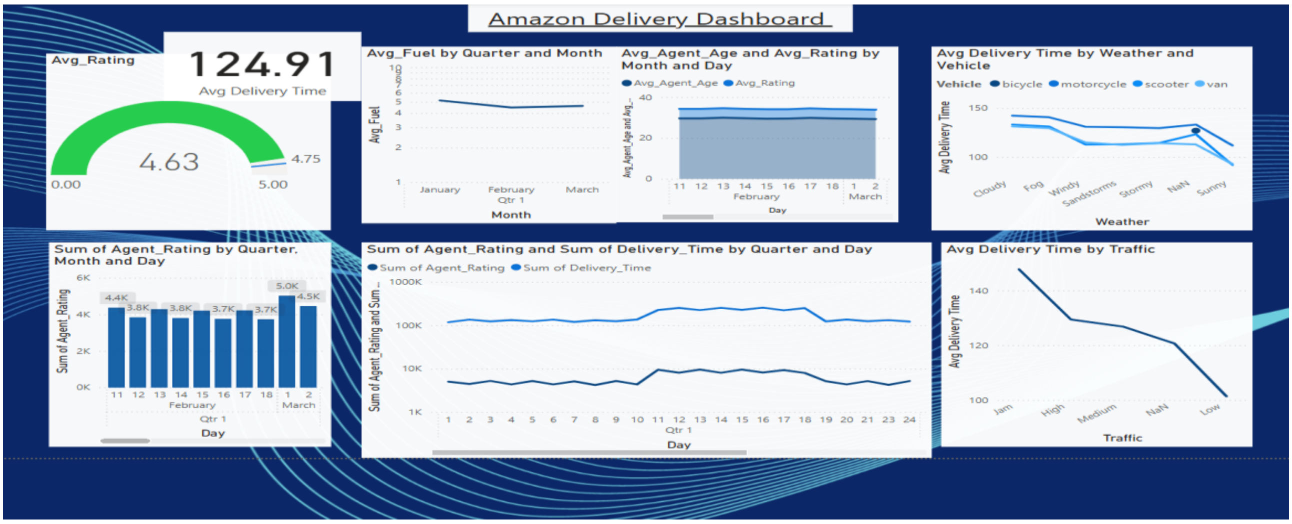


Truck:



Van:





Project Report:

Our analysis reveals the following insights regarding fuel consumption, on-time delivery percentage, and delivery costs for the Bike, Truck, and Van:

1. **Fuel Consumption**:
    - The Bike consumed significantly less fuel on average compared to the Truck (**34.87% lower**) and the Van (**20.97% lower**).
2. **On-Time Percentage**:
    - The Bike has a lower On-Time Percentage, being **16.25% less** than the Van and **2.5% less** than the Truck.
3. **Delivery Costs**:
    - The total delivery cost for the Bike was **44.31% cheaper** than Truck and **27.23% cheaper** than Van.

Conclusion:

From our analysis, we conclude that the Truck mode of delivery is the least efficient overall. In contrast, the Bike (lowest fuel consumption) and the Van (highest On-Time Percentage), are better options for delivery efficiency. The effect of traffic on delivery time is also noticeable, and weather too plays an important role in delivery time (minimum in sunny weather, maximum during fog and cloudy weather). In the case of unfavourable weather or when carrying fragile shipments, it's better to utilize Van mode of delivery, and in high traffic it's ideal to use bike as two wheelers can swiftly navigate through narrow shortcuts helping save time and boost the supply chain's responsiveness and customer satisfaction.

List of ML Techniques Used and Purpose:

Algorithm / Technique	Purpose
Logistic Regression	*To classify deliveries as on-time or late based on features like distance, fuel consumed, driver rating, and vehicle type.*
Train-Test Split (70:30)	*To evaluate model accuracy and generalization by training on 70% of the data and testing on 30%.*
One-Hot Encoding	*To convert the categorical variable vehicle_type into numeric format so that it can be used in the ML model.*
Feature Selection	*To choose the most relevant features (distance_km, fuel_consumed_litres, driver_rating, etc.) for the model.*
Confusion Matrix & Accuracy Score	*To assess model performance in terms of correctly predicted late vs on-time deliveries.*

ML Tools Used & Installation of Software:

📦 ML Tools Used

- Python 3.10: Used as the primary language for machine learning model development.
- scikit-learn: Used for implementing logistic regression, model training, and evaluation.
- pandas: For data preprocessing and manipulation (handling datasets, creating features, etc.).
- Jupyter Notebook: Used for initial experimentation and testing of the ML model before integrating into Power BI.
- Power BI: Used to run Python scripts for ML and display prediction results inside the dashboard.

🛠 Installation of Software

- Power BI Desktop was installed from the official Microsoft Power BI website.
- Python 3.10 was installed from https://www.python.org.
- Required Python packages were installed using pip:
  pip install pandas scikit-learn

- In Power BI:
- Python scripting was enabled via **File → Options → Python scripting** in Power Bi.
- The local Python installation path was configured inside the Power BI settings menu.
- Visuals like predicted late deliveries were generated using Power BI's built-in support for Python

Conclusion:

The project successfully demonstrated how delivery visibility can be enhanced through interactive dashboards and predictive analytics. Using a custom synthetic dataset, a Power BI dashboard was developed to track key performance indicators such as delivery time, fuel consumption, cost, and on-time percentage. Additionally, a logistic regression model was integrated to predict the likelihood of late deliveries based on operational features.

To strengthen the model design and feature selection, a real-world Amazon delivery dataset from Kaggle was explored, though machine learning was applied only on the synthetic data due to missing operational KPIs such as fuel consumption and delivery cost.

Overall, the system empowers supply chain (logistics) managers to make data-driven decisions by highlighting cost chokepoints and inefficiencies in delivery operations. The project demonstrates how combining visualization and machine learning can provide powerful insights for logistics performance monitoring. Future improvements could include real-time data streaming and route optimization to further enhance delivery efficiency.

References:

1. https://www.thoughtspot.com/data-trends/dashboard/supply-chain-kpis-metrics-for-dashboard

2. https://supplychaindigital.com/logistics/seven-key-features-effective-supply-chain-dashboards-take-supply-chain

3. https://www.bizinfograph.com/blog/supply-chain-dashboard/

4. https://inforiver.com/blog/general/supply-chain-kpis-inforiver-charts-powerbi/

5. https://www.slideteam.net/blog/top-10-supply-chain-dashboard-examples-with-templates-and-samples

6. https://cashflowinventory.com/blog/demand-and-supply-planning-kpis/

7. https://www.linkedin.com/pulse/7-mini-case-studies-successful-supply-chain-cost-rob-o-byrne/

8. https://www.data.gov.in/search?title=transport&type=resources&sortby=_score

9. https://www.kaggle.com/datasets/sujalsuthar/amazon-delivery-dataset

10. https://github.com/austinlasseter/datasets-shipping-logistics