UK Train Railway Operations and Delay Analysis

Team Members:

- *Ahmed Mohamed Ahmed Fouad Abdelhamid*
- *Jana Reda Mohamed*
- *Khaled Mustafa El-Sayed*
- *Rawan Tarek Gamal*
- *Seif Abdulaziz El-Sayed*
- *Toka Khaled Mostafa*

Instructor: Eng. Osama Adel

Overview of the UK Railway System

Understanding the Scale and Importance of Railway Data in the UK

The UK railway system is a vital component of national infrastructure, connecting millions of passengers and facilitating commerce. Analyzing railway data plays a crucial role in improving efficiency, and enriching the passenger experience. Effective use of this data can lead to significant advancements in transportation planning and operations.

Project Idea

The Railway Operations and Delay Analysis project aims to design a data warehouse that integrates railway transaction and journey data to analyze performance, customer experience, and operational delays.

The dataset captures information about each ticket purchase, including transaction details, routes, payment methods, journey times, and delay reasons.

This analysis enables:

- Tracking journey punctuality and identifying delay causes.
- Monitoring sales by purchase type, payment method, and ticket class.
- Evaluating refund requests linked to delay incidents.
- Building dashboards for performance analytics and decision-making.

Tools Used in The Project

Microsoft Excel

Used for saving the dataset as an excel file

SQL Server

Used for cleaning, extracting dimensions and facts table from the dataset and answering questions

Python

Used in predicting

Power Bi

Used in visualization

Railway Flat Datset Columns

Transaction_ID
Date_of_Purchase
Time_of_Purchase
Departure_Station
Arrival_Destination
Date_of_Journey
Departure_Time
Arrival_Time
Actual_Arrival_Time
Journey_Status
Reason_for_Delay
Refund_Request
Purchase_Type
Payment_Method
Railcard
Ticket_Class
Ticket_Type
Price

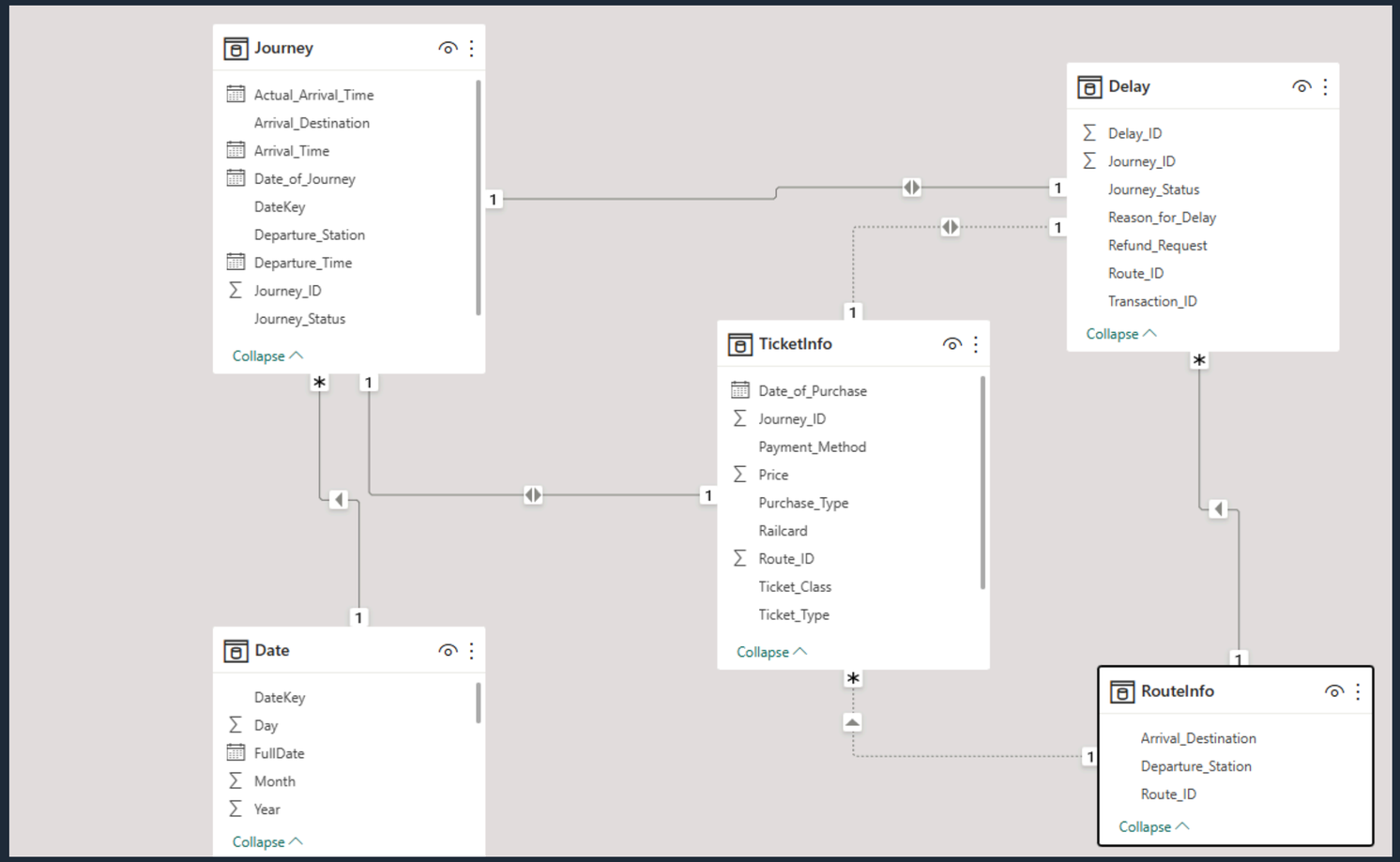
Dataset Dimensions

Dimension	Attributes	Description
Route	Route ID, Departure Station, Arrival Destination	Provides time-based context for transactions and journeys.
Journey	Journey ID, Transaction ID, Departure Station, Arrival Destination, Date of Journey, Departure Time, Arrival Time, Actual Arrival Time, Journey Status, Datekey	Contains scheduled and actual journey times with performance status.
Date	Date Key, Full Date, Day, Month, Year	Provides time-based context for transactions and journeys.

Dataset Facts

Name	Attributes	Description
Ticket Info	Transaction ID, Date of Purchase, Time of Purchase, Purchase Type, Payment Method, Railcard, Ticket Class, Ticket Type, Price, Route_ID, Journey_ID	Details about the type and pricing of the ticket purchased and the purchase info.
Delay	Delay ID, Transaction ID, Journey ID, Route ID, Journey Status, Reason for Delay, Refund Request	Provides information on the cause of delays when applicable.

Dataset Schema



Analysis Scope

Journey Performance

Delay & Cancelled Causes

Stations & Routes Performance

Revenue Tracking

Journey Performance

Total Journeys

31653

On - Time Journeys

27,481

Delayed Journeys

2292

Cancelled Journeys

1880

Delay & Cancelled Causes

Top Delay Cause

Weather

Top Cancelled Cause

Signal Failure

Top Month with Delay & Cancel

March

1137

Avg for Delayed Journeys

42 mins

Important Routes Index

Route 1: York to Liverpool Lime Street

Route 4: Manchester Piccadilly to Liverpool Lime Street

Route 9: London St Pancras to Wolverhampton

Route 21: Liverpool Lime Street to London Euston

Route 29: London Kings Cross to Liverpool Lime Street

Route 48: Manchester Piccadilly to Sheffield

Route 58: London Kings Cross to York

Route 61: Liverpool Lime Street to Sheffield

Stations & Routes Tracking

Best Route Performance (Least Delays):

London Kings Cross → Liverpool Lime Street (Route 29)

London St Pancras → Wolverhampton (Route 9)

York → Liverpool Lime Street (Route 1)

Liverpool Lime Street → Sheffield (Route 61)

Manchester Piccadilly → Sheffield (Route 48)

Route with Most Delays and Cancels respectively:

Liverpool Lime Street → London Euston (Route 21) (Most Delays)

Manchester Piccadilly → Liverpool Lime Street (Route 4) (Most Cancels)

Stations & Routes Tracking

Departure Station with Most Delays:

Liverpool Lime Street

Departure Station with Most Cancels:

London Paddington

Route with Highest Volume:

Manchester Piccadilly → Liverpool Lime Street (Route 4)

Revenue Tracking

Total Revenue

741,921

Total Net Revenue

703,219

Lost Revenue from Refund

38,702

5%

Top Route with Revenue

Route 58

179498

Top Station with Revenue

London Kings Cross

195770

Prediction Analysis

April 2024
Avg Price

24.13

Ticket Model
Accuracy = 63.19%

May 2024
Avg Price

23.92

April 2024
Total Journeys

7781

Journey Model
Accuracy = 68.90%

May 2024
Total Journeys

7817

Prediction Analysis

April 2024 Revenue

187,755

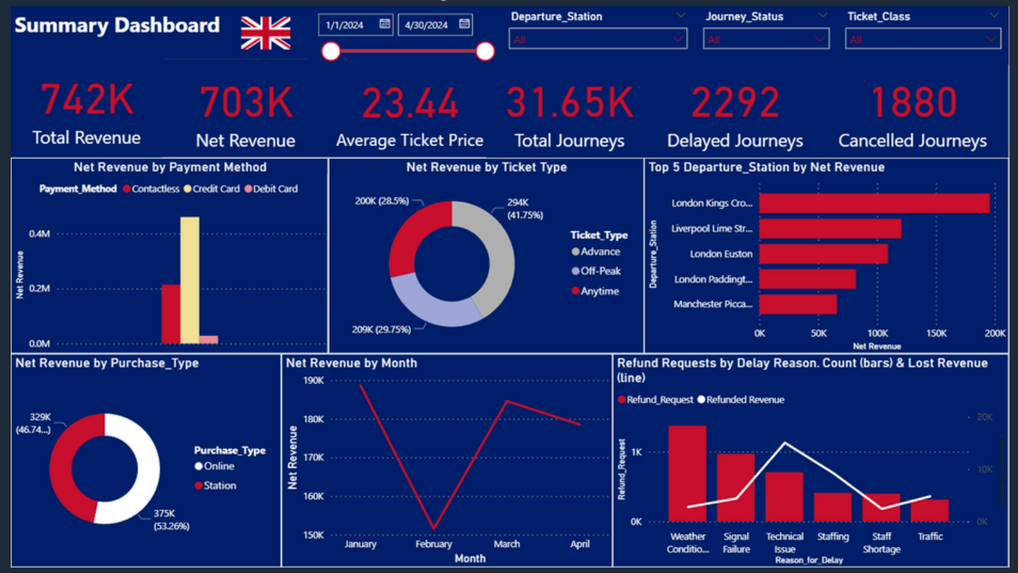
Revenue without Subtracting Refunded Request

May 2024 Expected Revenue

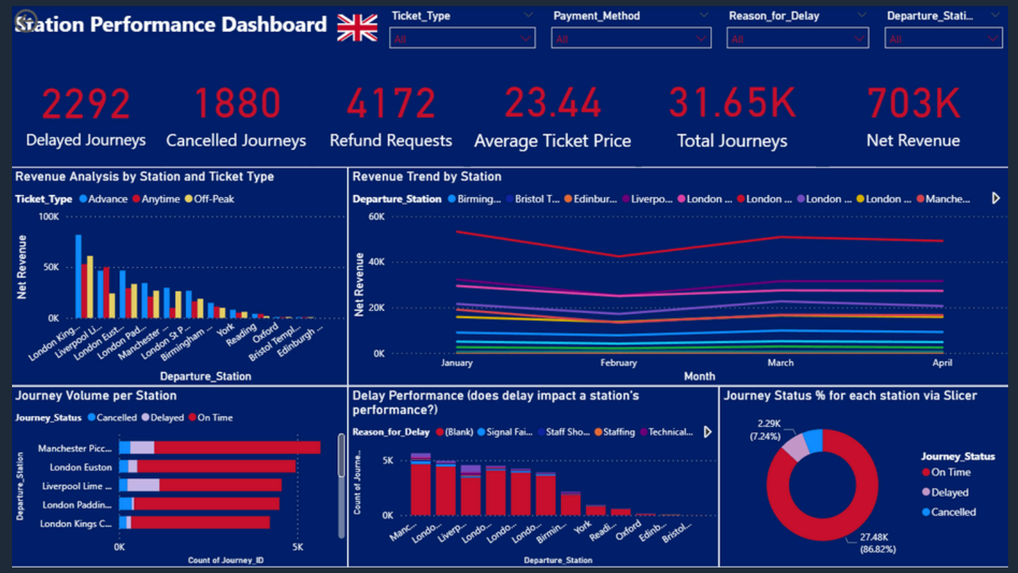
186,983

Dashboards

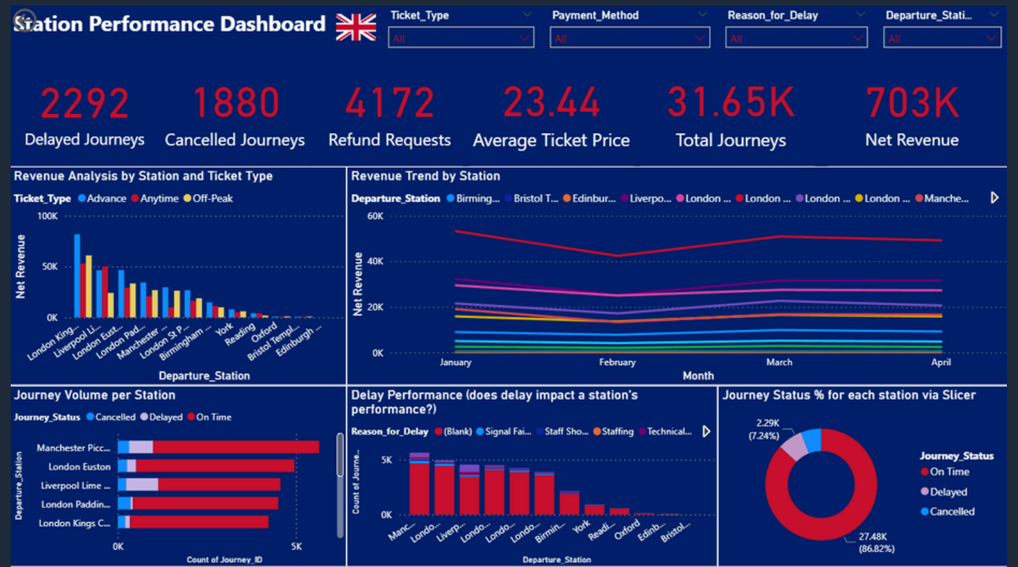
Summary Dashboard



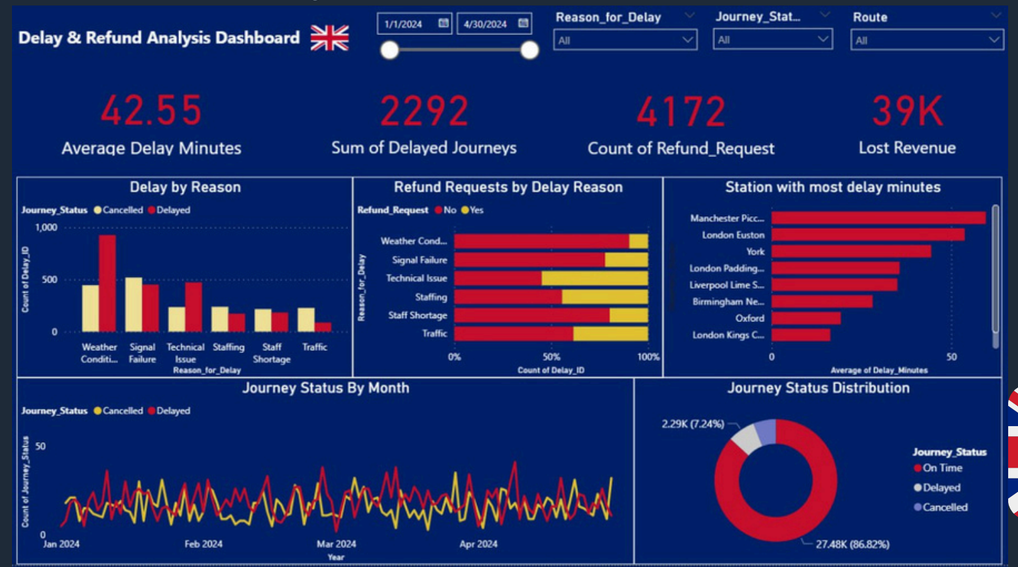
Stations Performance Dashboard



Routes Performance Dashboard



Delay & Refund Dashboard



Insights & Suggested Recommendations

💡Weather alongside signal failure are the biggest cause of refund-driving delays
🎯So improving weather resilience and fixing signal issues can protect revenue.

💡High delay count due to weather conditions and technical issues at Liverpool Lime Street and staff shortages at Manchester Piccadilly are the key issues
🎯The solutions are infrastructure checks and hiring more staff respectively.

💡Most revenue comes from five main stations
🎯So maintain them carefully and use marketing to boost smaller stations.

💡Credit cards dominate payments
🎯So ensure other payment methods work reliably.

💡February’s revenue drop is seasonal
🎯So promotions during low-demand months could help.

Insights & Suggested Recommendations

💡Address technical issues, which cause the highest refunds (54.88%)
🎯By implementing IoT monitoring and proactive maintenance.

💡Manage high traffic on Manchester Piccadilly to Liverpool and reduce cancellations there
🎯By adding backup trains, improving maintenance, and weather-proofing infrastructure.

💡Focus on top routes like Manchester Piccadilly to Liverpool and London Euston to Birmingham to
🎯Increase capacity and reduce delays.

🎯Increase marketing and improve service on
💡Low-volume routes such as Birmingham New Street to London St Pancras and London Paddington to Oxford.

💡Take preventive action for weather conditions, the main cause of delays with 927 incidents
🎯By installing de-icing systems, improving drainage, and using weather sensors to detect adverse conditions early.

Thank You

Questions?