UK TRAIN RIDES REPORT

REPRESENTED BY:

ZIAD SHERIF 234232

MALAK AMR 234139

MARIAM TALAAT 234222

BASMALA KHALED 234246

JAN MAGED 234023

1. Introduction

This report provides comprehensive analysis of the UK train
The data was provided by [Kaggle] and covers journeys between [2024-01-01 → 2024-04-30].

Objective

The primary objective of this analysis is to evaluate the performance of train services across the United Kingdom. This includes analyzing patterns in delays, identifying causes of cancellations, and understanding passenger behavior related to ticket purchases. The insights derived aim to enhance service quality, improve operational efficiency, and inform strategic decision-making.

2. Overview

The dataset comprises 31,653 rows and 18 columns, capturing detailed information about train ticket purchases in the United Kingdom. Key attributes include:

- **Transaction Details**: Payment method (online or at the station), ticket price, departure and arrival stations.
- **Journey Information**: Departure time, actual arrival time, journey status (e.g., on time, delayed, cancelled), and delay reasons.
- **Passenger Behavior**: Refund requests and purchase methods.

This dataset supports the analysis of train service performance over time, identification of delay causes, and exploration of passenger preferences based on ticket purchase patterns.

3. Business Question

What is the overall passenger satisfaction score for the period?

Which stations or routes receive the most customer complaints or praise?

What are the most common reasons for passenger dissatisfaction?

What are the most common reasons for passenger dissatisfaction?

How does train occupancy compare across different times of day or routes?

What percentage of trains arrive on time compared to delayed and cancelled trains?

What are the key causes of service delays and cancellations, and how can we prioritize efforts to reduce them?

What are the train routes with the most delays?

What is the average delay time per route?

4. Key Findings

Route Activity:

Manchester Piccadilly has the most active route with 4.6K trips, making it the busiest among the top 7 routes.
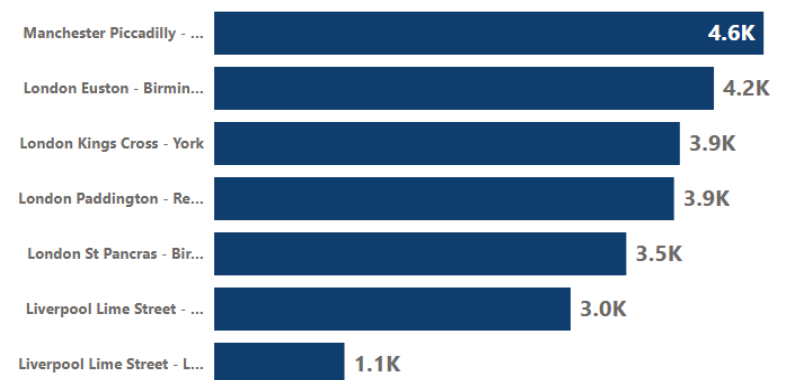
London Euston - Birmingham follows with 4.2K trips, also indicating high travel demand.

London Kings Cross - York and London Paddington each have 3.9K trips, showing equal popularity on these routes.

London St Pancras - Birmingham has a slightly lower figure at 3.5K.

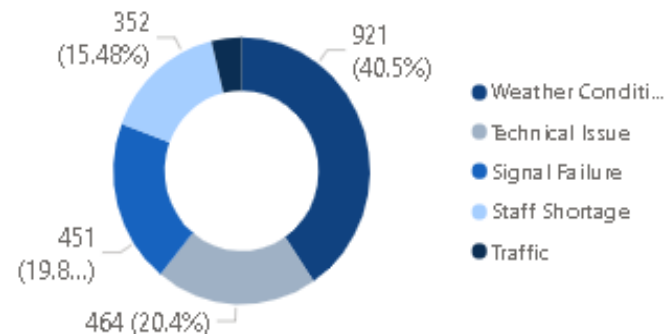Liverpool Lime Street appears twice in the ranking:

Top 7 Active Route

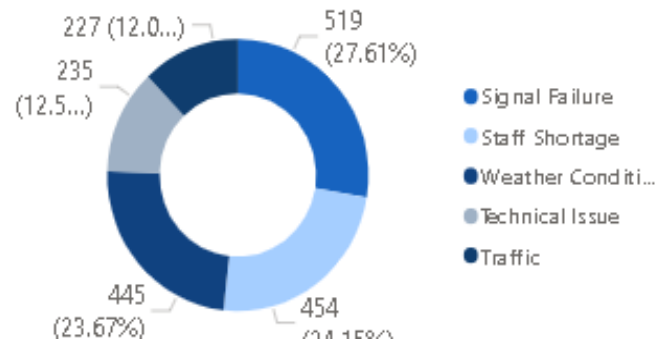| Route	Trips
Manchester Piccadilly - ...	4.6K
London Euston - Birmin...	4.2K
London Kings Cross - York	3.9K
London Paddington - Re...	3.9K
London St Pancras - Bir...	3.5K
Liverpool Lime Street - ...	3.0K
Liverpool Lime Street - L...	1.1K

Once with 3.0K trips,

And again, with a lower 1.1K trips,

indicating possibly two different destination routes or different directions on the same route.

Overall, London appears frequently as an origin or destination, highlighting its role as a central hub in the rail network.

There is a noticeable drop-off in volume from the top route (4.6K) to the lowest (1.1K), showing variability in route popularity.

Delay & Cancellation

Delay Reasons Breakdown

352 (15.48%)
921 (40.5%)
451 (19.8...)
464 (20.4%)

● Weather Conditi...
● Technical Issue
● Signal Failure
● Staff Shortage
● Traffic

Cancelled Reasons Breakdown

227 (12.0...)
519 (27.61%)
235 (12.5...)
445 (23.67%)
454 (24.15%)

● Signal Failure
● Staff Shortage
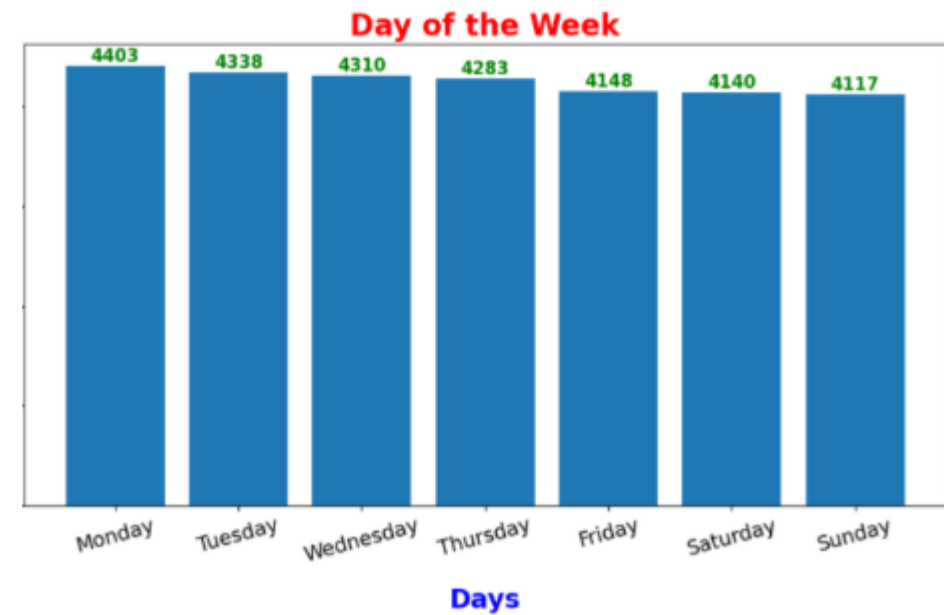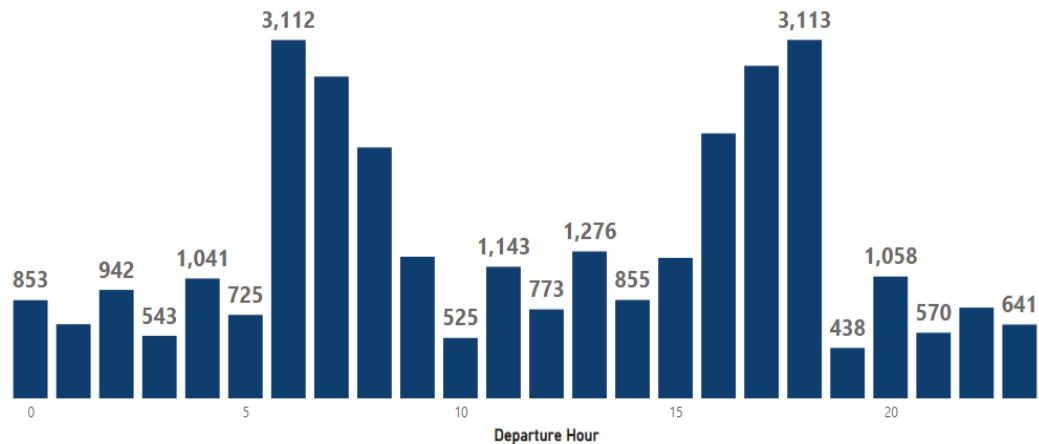● Weather Conditi...
● Technical Issue
● Traffic

- Weather Conditions are the major cause for delays and the major cause for cancellations is signal failure.

- Technical Issues are consistently impactful across both metrics.

- Traffic and Staff Shortage have moderate effects on both delays and cancellations, but never the highest.

Organizations should focus on improving weather resilience and system maintenance to reduce both delays and cancellations.

Temporal Trends

Trips by Departure Hour

A. Daily Ridership:

Highest on Mondays (4,403 passengers), gradually declining to Sundays (4,117).

B. Hourly Activity:

Peaks at 6:00 AM (3,112) and 6:00 PM (3,113) — likely due to work commutes.

Lowest activity overnight (9:00 PM–4:00 AM).

Insight: Demand aligns with typical work schedules.

5. Summary

A. Passenger Patterns:

Over 90% choose standard class, showing strong preference for budget travel.

 Manchester Piccadilly and London stations handle 10x more passengers than smaller stations like Coventry.

B. Delay Trends:

Weather alone causes 35,480 delay minutes – more than staff + technical issues combined.

Weather is the fastest-growing delay reason (33% of incidents).

C. Peak Times:

Rush hours match typical work hours ([8 AM → 8] and [5 PM → 17])

Mondays are the busiest (4,403 passengers), Sundays are the least busy (4,117).

Some Dataset limitations

We don't know why people are traveling (work/leisure)

Missing passenger details (age, frequent travelers, etc.)