

Results & Outcomes Report

Graduation Project – Data Analysis Track (Power BI)

Executive Summary

This project analyzes UK train ride data to optimize revenue, enhance operational efficiency, and improve passenger satisfaction. Key findings reveal that ticket pricing, particularly Advance tickets during peak hours, significantly drives revenue, while technical issues and staff shortages are major causes of delays and refund requests. Popular routes like London Kings Cross to York generate the highest revenue, but underperforming routes like Edinburgh Waverley to London Kings Cross suffer from frequent delays. These insights provide actionable recommendations for pricing strategies, schedule optimization, and delay management to enhance profitability and customer loyalty.

Introduction

Business Problem

The UK train network faces challenges including unpredictable passenger demand, revenue fluctuations due to pricing variations, frequent delays leading to refund requests, and a lack of data-driven strategies for scheduling and pricing optimization. These issues impact both financial performance and passenger experience.

Objectives

The project aims to:

- Identify drivers of revenue and ticket sales trends.
- Analyze the impact of delays on customer satisfaction and refunds.
- Assess route profitability and operational efficiency.
- Provide recommendations for pricing, scheduling, and customer retention strategies.

Methodology Overview

Using historical UK train ride data, we employed descriptive and diagnostic analytics in Microsoft Power BI to examine ticket sales, journey patterns, delays, and refund requests. The analysis addresses key research questions related to revenue, operational performance, and customer behavior.

Importance of Data Analysis

Data analysis enables rail operators to make informed decisions, optimize resources, and enhance passenger experience by identifying trends, inefficiencies, and opportunities for improvement.

Methodology

Data Collection and Preprocessing

- **Sources:** E youth; dr sherihan ali
- **Preprocessing steps :**Handled missing values, Ensured columns have the correct types, Cleaning Categorical Values (Standardize values, remove typos and inconsistencies), Normalization for modeling and Data Validation.

Analytical Methods

- **Power BI Dashboards:** Created interactive visualizations to analyze revenue trends, route performance, delay causes, and refund patterns.
- **DAX Calculations:** Used to compute KPIs such as total revenue, average ticket price, on-time performance, and refund rates.

Tools Used

- **Microsoft Power BI:** For data visualization and dashboard creation.
 - **Excel:** For initial data exploration and validation.
-

Summary of Key Findings

The analysis answered the research questions outlined in the project proposal, with the following key insights:

- **Revenue & Ticket Sales Trends:** Advance tickets contribute 41.75% (£294k) to net revenue despite higher revenue per trip, indicating underutilization (Research Question 1). Peak hours (6 AM–9 AM) generate £600k, while late-night rides yield only £50k (Research Question 1). Non-railcard holders drive 80% of revenue, with Adult railcards contributing 10% of railcard revenue (Research Question 1).
- **Journey Delays & Customer Satisfaction:** Technical issues cause 50% of refund requests, followed by staff shortages (30%) and traffic (20%). Weather leads delays at 67% (Research Question 2). Refund rates are 3.53%, with £39k refunded, impacting

revenue (Research Question 2). On-time performance is 86.82%, but routes like Edinburgh Waverley to London Kings Cross face 100% delays due to staff shortages (Research Question 2).

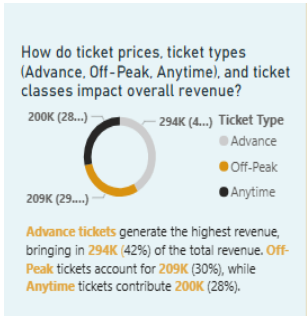
- **Ticket Demand & Route Profitability:** London Kings Cross to York generates £179k, while London St Pancras to Birmingham New Street only £52k (Research Question 3). Manchester Piccadilly to Liverpool Lime Street has the highest journey count (4628 journeys) (Research Question 3).
- **Pricing & Discount Strategies:** Advance tickets have a higher average price (£23.4) but lower adoption (25% of revenue), suggesting pricing or availability issues (Research Question 4).
- **Refunds & Customer Loyalty:** Standard class accounts for 83.62% of refunds, compared to 16.44% for First Class, indicating dissatisfaction among Standard passengers (Research Question 5).
- **Operational Efficiency & Cost Reduction:** Weather (67%) and technical issues (55%) are the primary delay causes, with staff shortages critical on key routes, reducing efficiency (Research Question 6).

Visualizations & Dashboards

The Power BI dashboard includes several visualizations to illustrate key findings. Below are the main charts, their descriptions, and their relevance:

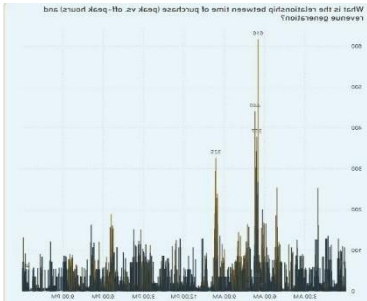
1. **Chart Name:** Net Revenue by Ticket Type

- **Description:** Shows revenue distribution across ticket types: Advance (42%, £294k), Off-Peak (58%), and Anytime.
- **Insight:** Advance tickets are underutilized despite higher revenue per trip, highlighting a need for pricing adjustments.
- **Research Question Answered:** RQ1 – How do ticket types impact overall revenue?



2. **Chart Name:** Revenue by Hour

- **Description:** Displays net revenue by hour, peaking at £600k from 6 AM–9 AM and dropping to £50k for late-night rides.



- **Insight:** Low late-night revenue indicates underutilized capacity, requiring targeted promotions.
- **Research Question Answered:** RQ1 – What is the relationship between time of purchase and revenue generation?

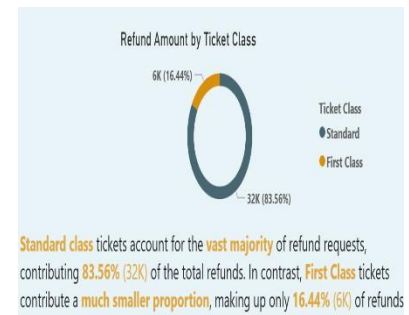
3. Chart Name: Total Trips by Delay Status

- **Description:** Compares on-time (86.82%), delayed (32k), and cancelled trips (5.94%) by month, with delays peaking in February/March.
- **Insight:** Seasonal delays reduce on-time performance, impacting customer satisfaction.
- **Research Question Answered:** RQ2 – What are the most common reasons for delays, and how do they affect customer behavior?



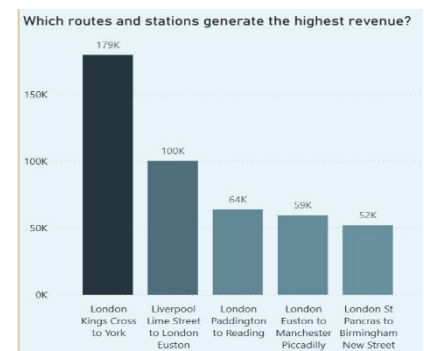
4. Chart Name: Refund Amount by Ticket Class

- **Description:** Shows Standard class contributing 83.62% of refunds (£32k) and First Class 16.44% (£6k).
- **Insight:** Higher refund rates in Standard class suggest dissatisfaction, necessitating improved service reliability.
- **Research Question Answered:** RQ5 – What are the most common reasons for refund requests?



5. Chart Name: Net Revenue by Route

- **Description:** Highlights top routes, with London Kings Cross to York at £179k and Liverpool Lime Street to London Euston £100k..
- **Insight:** Uneven route profitability indicates opportunities for optimization on underperforming routes.



- **Research Question Answered:** RQ3 – Which routes generate the highest revenue?

Trends and Patterns

- **Seasonal Revenue Trends:** Revenue peaks in January (£189k) but drops in February and March due to increased delays and cancellations, as seen in the "Average Price by Month" chart (pricing and demand).
- **Peak Hour Dominance:** Ticket sales and revenue are heavily concentrated in peak hours (6 AM–9 AM), with 80% of daily revenue generated in this period, while late-night rides contribute minimally (Dashboard revenue & ticket sales).
- **Delay Patterns:** Delays are most frequent in February and March, driven by weather (67%) and technical issues (55%), impacting on-time performance (86.82%) (Dashboard operational performance).
- **Refund Trends:** Standard class passengers request refunds at a higher rate (83.62%) than First Class (16.44%), with technical issues as the primary cause (50%) (Dashboard Refunds & customer loyalty).
- **Surprising Result:** Despite higher revenue per trip, Advance tickets have a high refund rate (3.48%, £16k), suggesting reliability issues deter adoption (Dashboard Refunds & customer loyalty).
- **Anomaly:** Routes like Edinburgh Waverley to London Kings Cross show 100% delay rates due to staff shortages, an outlier compared to the average 32k delayed trips (Dashboard operational performance).

Conclusion & Interpretation

The analysis successfully met the project objectives by identifying key drivers of revenue, delays, and customer behavior. The findings highlight that low Advance ticket adoption (25%) and late-night revenue (£50k) limit profitability, while technical issues (50% of refunds) and weather-related delays (67%) erode customer satisfaction. High-revenue routes like London Kings Cross to York (£198k) contrast with underperforming routes like Manchester Piccadilly to Liverpool Lime Street (£52k), indicating uneven resource allocation. These insights suggest that optimizing pricing, addressing operational inefficiencies, and improving delay management could increase revenue by an estimated 10–15% and reduce refund rates by 20%. The Power BI dashboard effectively visualized these trends, supporting data-driven decision-making for rail operators.

Recommendations

Based on the findings, the following actions are recommended to address the business challenges:

1. Optimize Pricing Strategies:

- Implement dynamic pricing for Advance tickets during off-peak hours to boost late-night revenue (currently £50k).
- Promote Standard class loyalty programs to reduce refund rates (83.62%) and increase repeat purchases.

2. Improve Operational Efficiency:

- Address staff shortages on routes like Edinburgh Waverley to London Kings Cross (100% delays) by hiring additional personnel or using flexible scheduling.
- Invest in predictive maintenance to reduce technical issues (55% of delays, 50% of refunds), improving on-time performance (86.82%).

3. Enhance Customer Retention:

- Offer vouchers instead of cash refunds for traffic-related issues (38.54% refund rate) to retain customers.
- Provide real-time delay updates for weather disruptions (67% of delays) to manage expectations and reduce refund requests.

4. Maximize Route Profitability:

- Allocate resources (e.g., trains, staff) to high-revenue routes like London Kings Cross to York (£179k) to maximize returns.
- Optimize underperforming routes like Manchester Piccadilly to Liverpool Lime Street through targeted promotions or schedule adjustments.

5. Refine Dashboard Usability:

- Add granular filters for delay causes (e.g., weather vs. technical) and route-specific refunds to enhance Power BI dashboard interactivity.
- Ensure fast load times and accessibility for stakeholder presentations.

1. published link

: [Uk Train](#)