

PROJECT UK TRAIN RIDES REPORT

PRES E N T A T I O N





Team Members



Ahmed Zain

Ahmed Korany

Ahmed Kaoud

Marina Isaac

Mohanad Omar

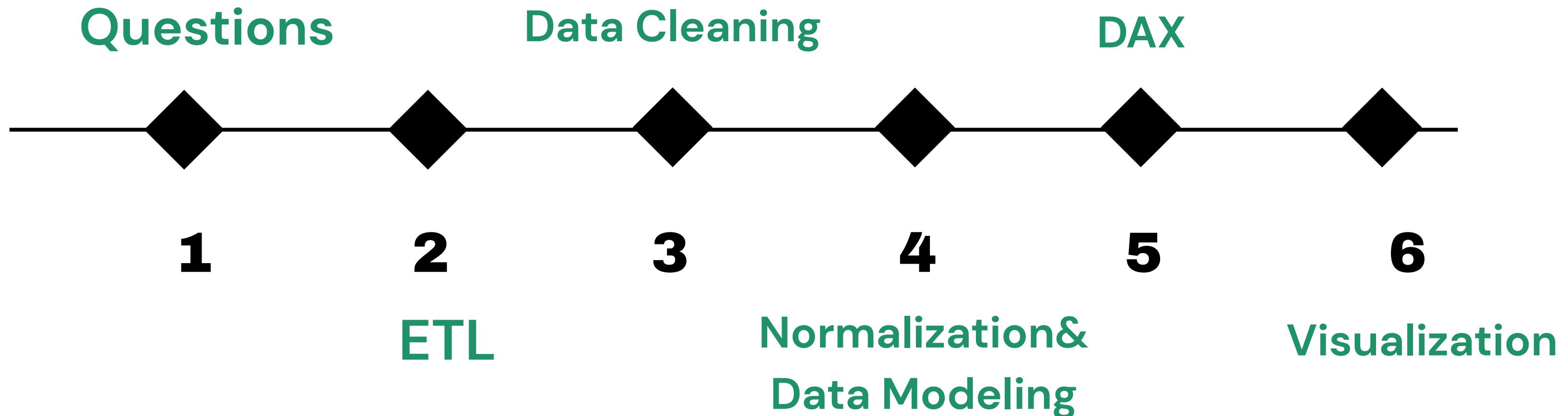


Introduction

The "UK Train Rides Report" project was developed to provide a comprehensive and detailed analysis of railway operations in the United Kingdom. This report is the result of a collaborative effort focused on identifying and resolving critical issues in train operations, as well as optimizing the passenger experience.



Project Phases



Data Analysis Questions



ETL

Initiated - POWER BI DESKTOP

Select a data source or start with a blank report

- Blank report
- OneLake catalog
- Excel workbook

Recommended

Getting started

Intro—What is Power BI?

File Origin: 1252: Western European (Windows) Delimiter: Comma Data Type Detection: Based on first 200 rows

railway.csv

Transaction ID	Date of Purchase	Time of Purchase	Purchase Type	Payment Method	Railcard	Ticket Class	Ticket Type
da8a6ba8-b3dc-4677-b176	12/8/2023	12:41:11 PM	Online	Contactless	Adult	Standard	Advance
b0cdd1b0-f214-4197-be53	12/16/2023	11:23:01 AM	Station	Credit Card	Adult	Standard	Advance
f3ba7a96-f713-40d9-9629	12/19/2023	7:51:27 PM	Online	Credit Card	None	Standard	Advance
b2471f11-4fe7-4c87-8ab4	12/20/2023	11:00:36 PM	Station	Credit Card	None	Standard	Advance
2be00b45-0762-485e-a7a3	12/27/2023	6:22:56 PM	Online	Contactless	None	Standard	Advance
4e1dcdb8-3d95-44ef-99fa	12/30/2023	7:56:06 AM	Online	Credit Card	None	Standard	Advance
1c74479d-85a4-4ba1-a607	12/31/2023	12:02:01 AM	Station	Credit Card	Adult	Standard	Advance
febfb8dab-f808-46fa-bf2b	12/31/2023	1:35:18 AM	Station	Contactless	Disabled	Standard	Advance
01df916f-4291-41ec-a37d	12/31/2023	1:43:09 AM	Station	Credit Card	None	Standard	Advance
a8cedba7-1923-459d-b046	12/31/2023	3:05:52 AM	Online	Credit Card	None	Standard	Advance
b3e5ca7d-e76c-49f2-b49f	12/31/2023	3:26:37 AM	Online	Contactless	None	Standard	Advance
6c63f7ac-d590-4356-9eaa	12/31/2023	3:52:11 AM	Online	Contactless	Adult	Standard	Advance
2e7add75-566a-41aa-9468	12/31/2023	5:55:22 AM	Online	Contactless	None	Standard	Advance
7ed9b545-eb6f-49b2-9b5a	12/31/2023	6:44:35 AM	Online	Contactless	None	Standard	Advance
2e05e2a6-88a8-40fb-bacc	12/31/2023	8:05:50 AM	Online	Credit Card	Disabled	Standard	Advance
8a18d3b4-995e-49bf-93a3	12/31/2023	8:16:53 AM	Online	Credit Card	None	Standard	Advance
7493a611-342a-4b17-90dc	12/31/2023	8:23:15 AM	Online	Credit Card	None	Standard	Advance
054676ac-a976-4909-a26e	12/31/2023	9:09:20 AM	Online	Credit Card	None	Standard	Advance
6b62b452-c491-468d-b39c	12/31/2023	9:12:21 AM	Online	Credit Card	None	Standard	Advance
85e38992-6c6c-4569-914e	12/31/2023	10:42:22 AM	Online	Credit Card	None	Standard	Advance

Extract Table Using Examples

Visualizations

Values

Drill through

Cross-report

Keep all filters

Add drill-through fields

Data Cleaning

naa15

Home Transform Add Column View Tools Help

New Recent Enter Data source settings Manage Parameters Refresh Preview Advanced Editor Manage Choose Columns Remove Columns Keep Rows Remove Rows Reduce Rows Sort Split Column Group By Replace Values Data Type: Text Merge Queries Append Queries Combine Files Combine

Trips [4] X ✓ fx = Table.RemoveColumns(#"Renamed Columns7", {"Station.Station"})

Query Settings

PROPERTIES

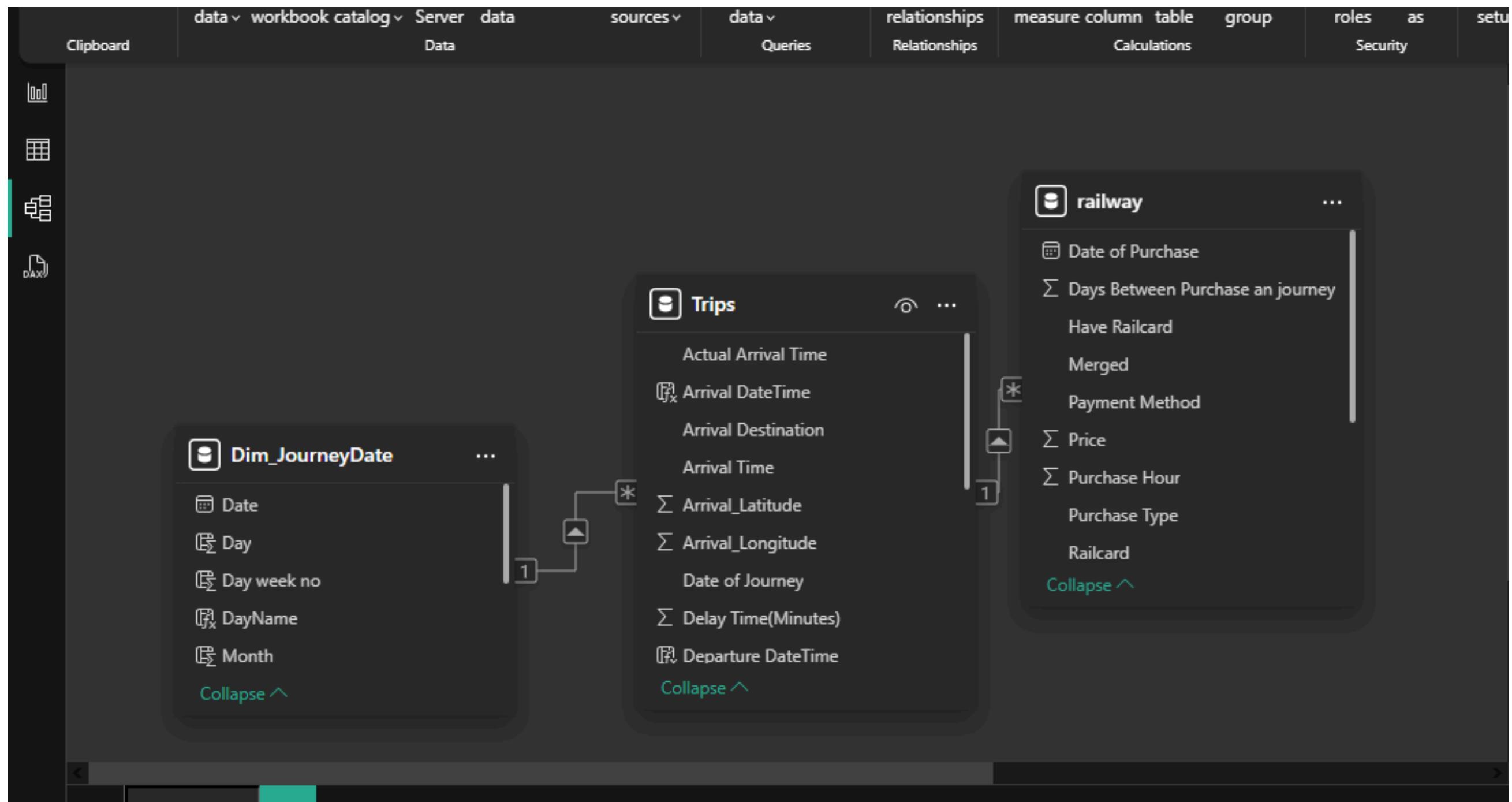
Name: Trips
All Properties

APPLIED STEPS

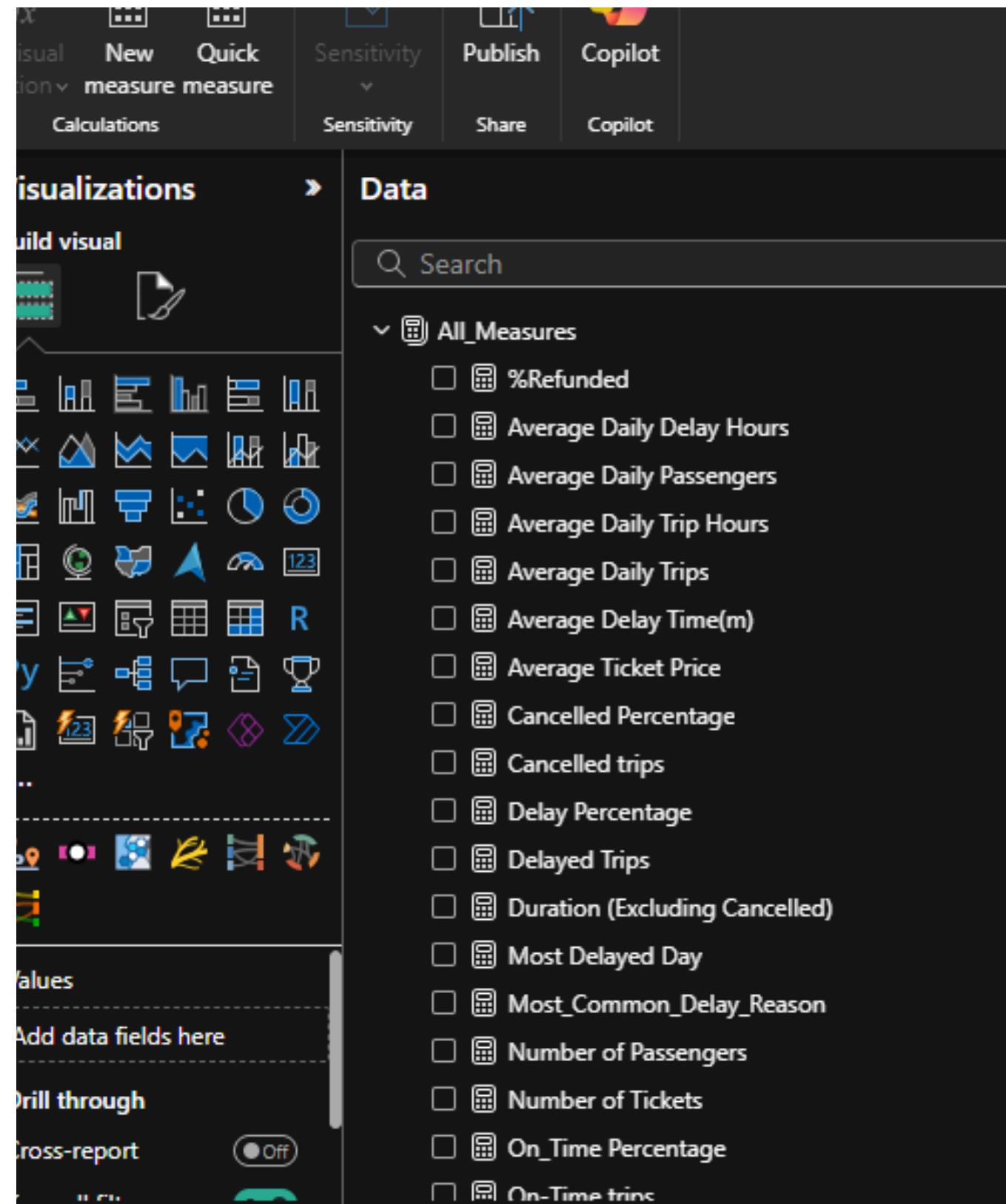
- Removed Columns7
- Renamed Columns4
- Inserted Merged Column
- Removed Duplicates
- Removed Columns8
- Added Index
- Reordered Columns4
- Renamed Columns5
- Added Prefix
- Merged Queries
- Expanded Station
- Renamed Columns6
- Removed Columns9
- Merged Queries1
- Expanded Station1
- Renamed Columns7
- Removed Columns10

	Trip ID	Departure Station	Arrival Destination	Date of Journey	Journey Day	Departure
1	Tr-1	London Paddington	Liverpool Lime Street	1/1/2024	Monday	
2	Tr-13	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	Monday	
3	Tr-14	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	Monday	
4	Tr-23	London Euston	Birmingham New Street	1/1/2024	Monday	
5	Tr-24	London Euston	Birmingham New Street	1/1/2024	Monday	
6	Tr-28	London Euston	Birmingham New Street	1/1/2024	Monday	
7	Tr-17	London St Pancras	Birmingham New Street	1/1/2024	Monday	
8	Tr-30	London St Pancras	Birmingham New Street	1/1/2024	Monday	
9	Tr-4	London Paddington	Reading	1/1/2024	Monday	
10	Tr-10	London Paddington	Reading	1/1/2024	Monday	
11	Tr-12	London Paddington	Reading	1/1/2024	Monday	
12	Tr-16	London Paddington	Reading	1/1/2024	Monday	
13	Tr-27	London Paddington	Reading	1/1/2024	Monday	
14	Tr-31	London Paddington	Reading	1/1/2024	Monday	
15	Tr-26	Oxford	Bristol Temple Meads	1/1/2024	Monday	
16						

Normalization & Data Modeling



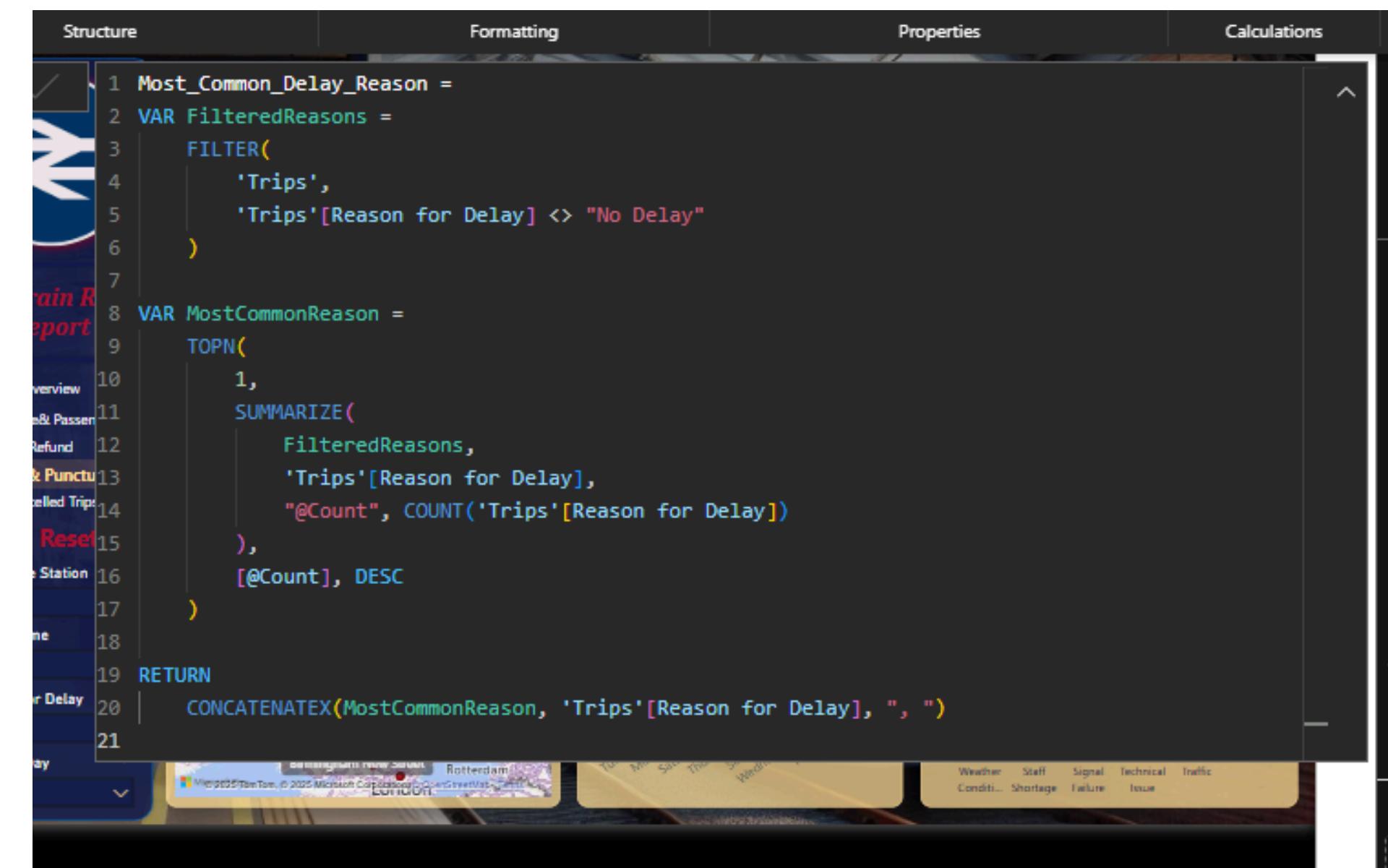
DAX



The screenshot shows the Microsoft Power BI DAX editor interface. The top navigation bar includes 'Visual' (highlighted), 'New', 'Quick', 'Sensitivity', 'Publish', and 'Copilot'. Below the navigation bar, there are tabs for 'measure measure' (selected), 'Calculations', 'Sensitivity', 'Share', and 'Copilot'. On the left, a sidebar titled 'Visualizations' contains icons for various chart types like bar, line, pie, and map. A 'Build visual' section is also present. The main area is titled 'Data' with a search bar. Under 'All_Measures', a list of measures is displayed:

- %Refunded
- Average Daily Delay Hours
- Average Daily Passengers
- Average Daily Trip Hours
- Average Daily Trips
- Average Delay Time(m)
- Average Ticket Price
- Cancelled Percentage
- Cancelled trips
- Delay Percentage
- Delayed Trips
- Duration (Excluding Cancelled)
- Most Delayed Day
- Most_Common_Delay_Reason
- Number of Passengers
- Number of Tickets
- On_Time Percentage
- On-Time trips

At the bottom, there are sections for 'Add data fields here', 'Drill through', 'Cross-report' (set to Off), and a 'DAX' button.



The screenshot shows the Microsoft Power BI DAX editor interface with the 'Structure' tab selected. The code is as follows:

```

1 Most_Common_Delay_Reason =
2 VAR FilteredReasons =
3     FILTER(
4         'Trips',
5         'Trips'[Reason for Delay] <> "No Delay"
6     )
7
8 VAR MostCommonReason =
9     TOPN(
10        1,
11        SUMMARIZE(
12            FilteredReasons,
13            'Trips'[Reason for Delay],
14            "@Count", COUNT('Trips'[Reason for Delay])
15        ),
16        [@Count], DESC
17    )
18
19 RETURN
20 CONCATENATEX(MostCommonReason, 'Trips'[Reason for Delay], ", ")
21

```

The code defines a measure named 'Most_Common_Delay_Reason' that uses a variable 'FilteredReasons' to filter trips where the reason for delay is not 'No Delay'. It then uses the 'TOPN' function to find the top reason by count. Finally, it returns a concatenated string of the reason names separated by commas.

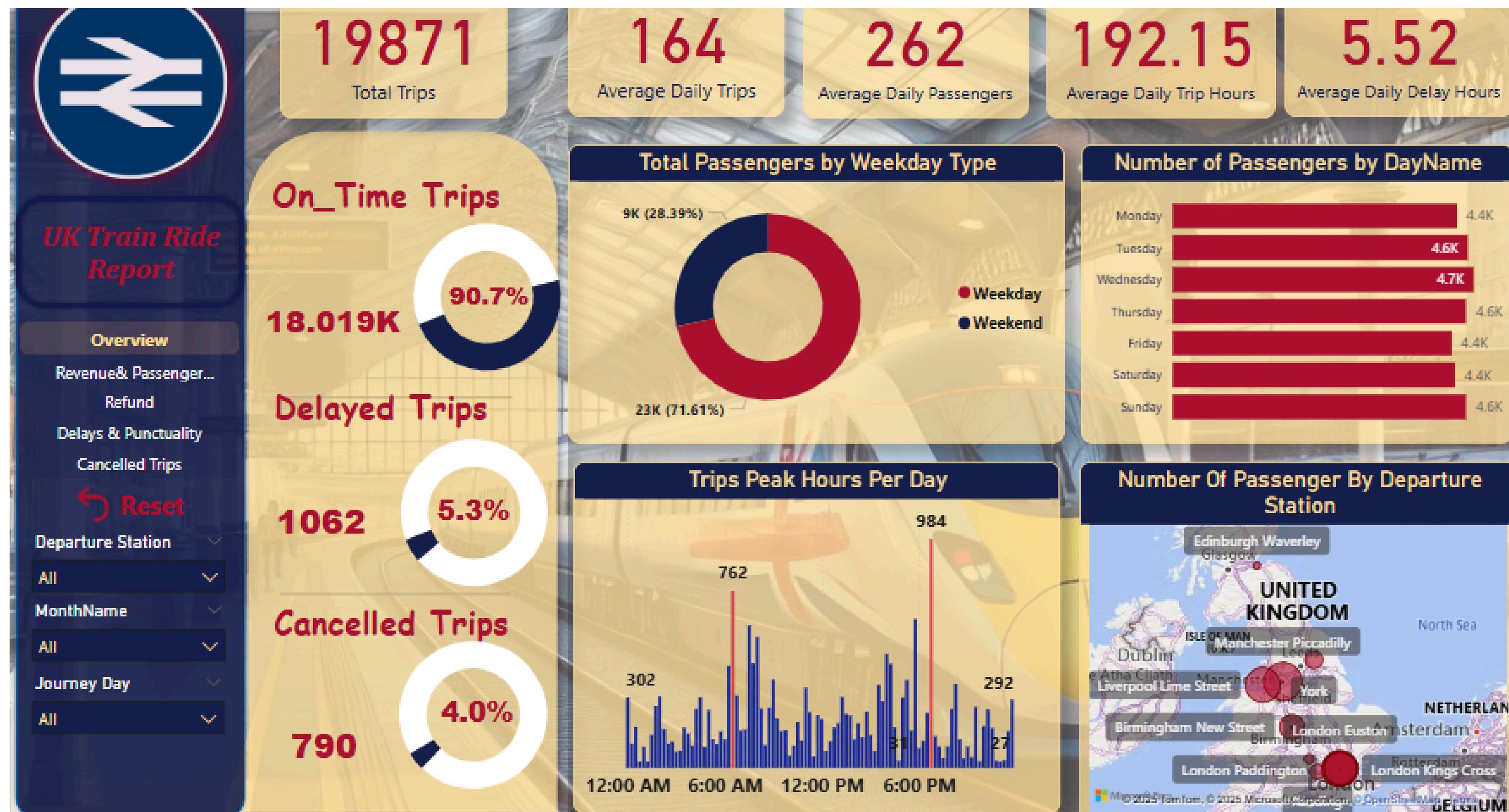
Visualization

Cover



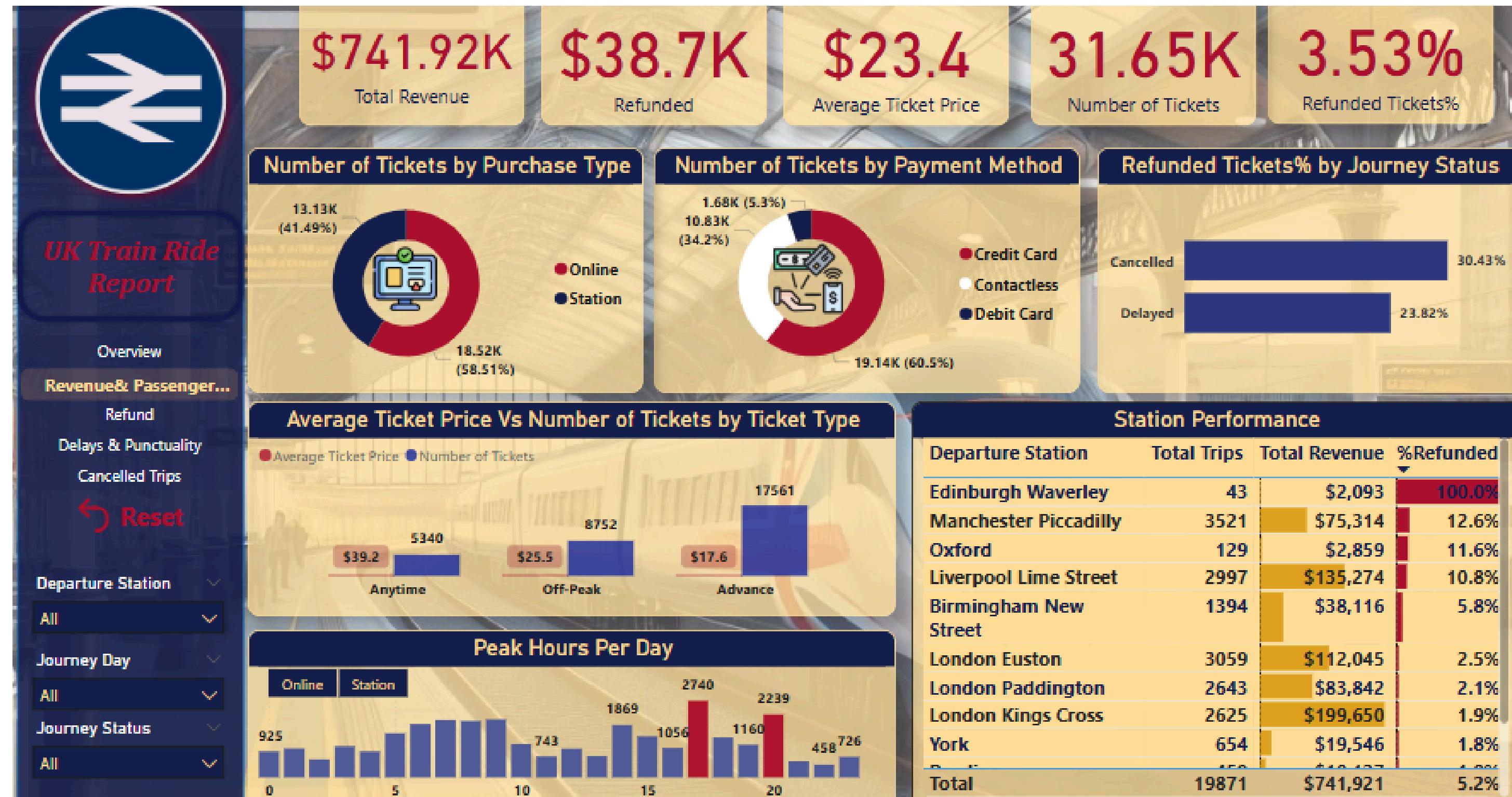
Visualization

Overview



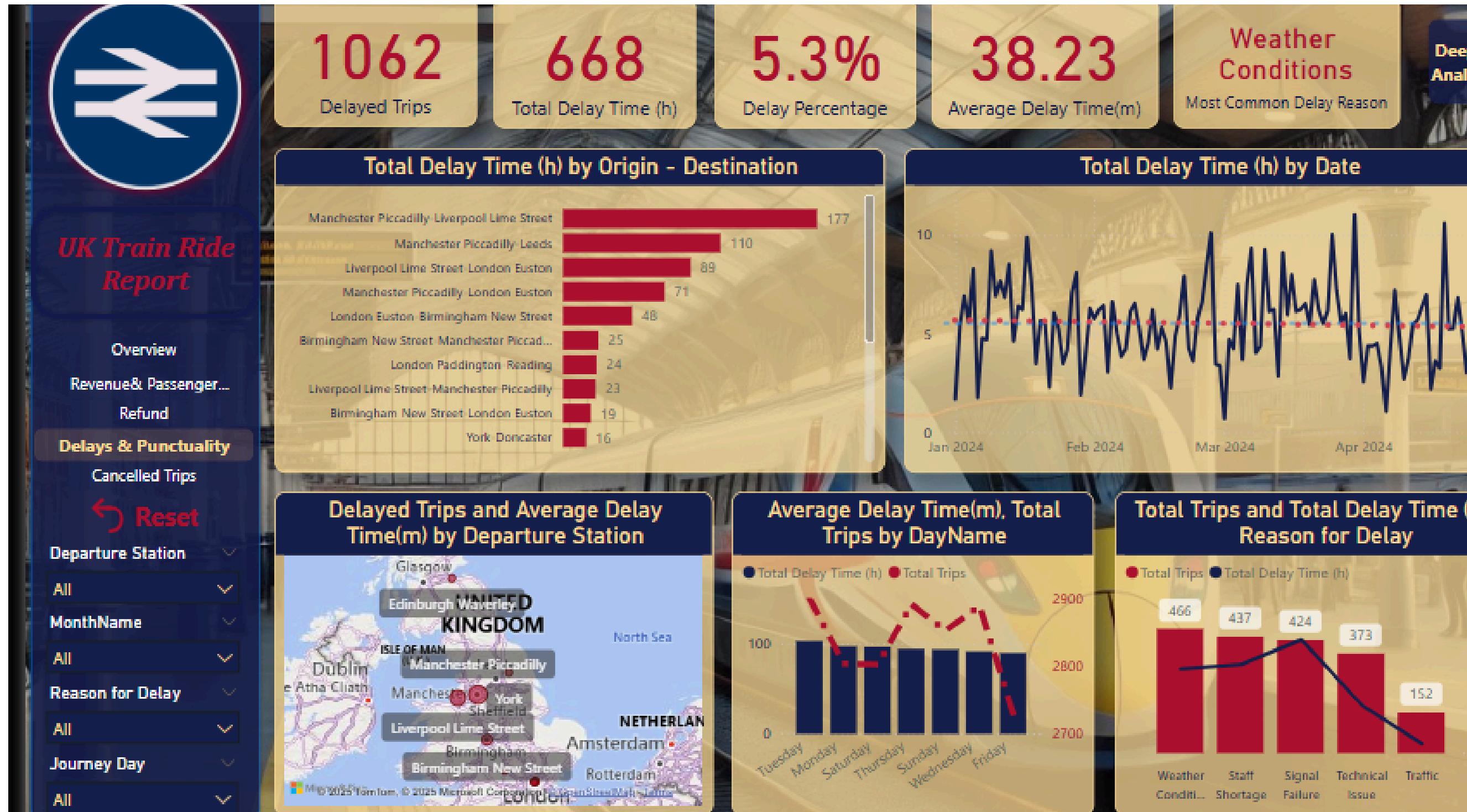
Visualization

Revenue & Passenger Behaviour



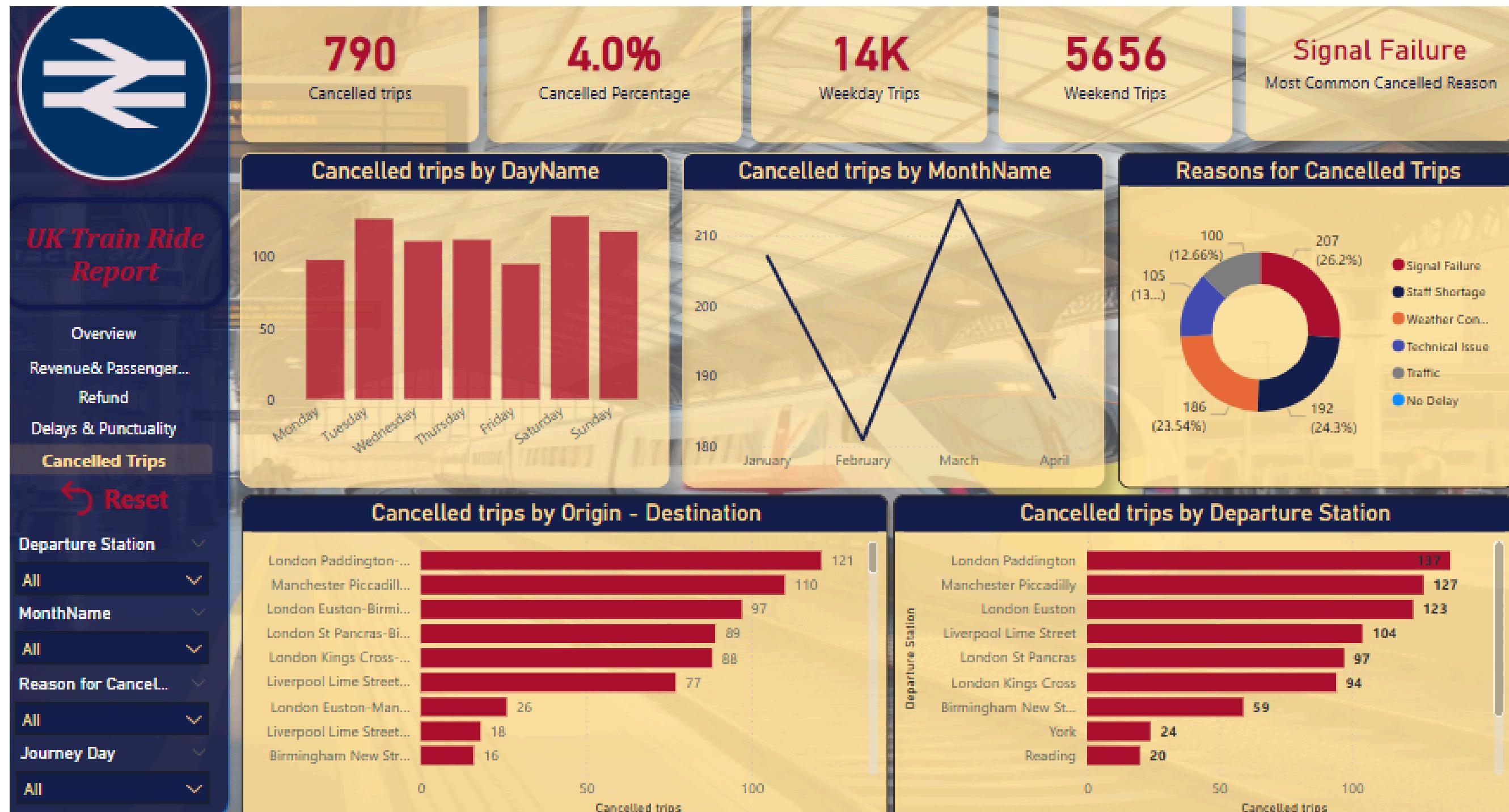
Visualization

Delays & Punctuality



Visualization

Cancelled Trips



Conclusion & Recommendations

- Priority Actions: Resolve signal failures and staff shortages to cut cancellations by 30% in 6 months.
- Quick Wins: Introduce flexible pricing for off-peak travel and improve weather-related contingency plans.
- Long-Term Strategy: Invest in infrastructure and workforce stability to boost punctuality beyond 90.7%.



Thank You