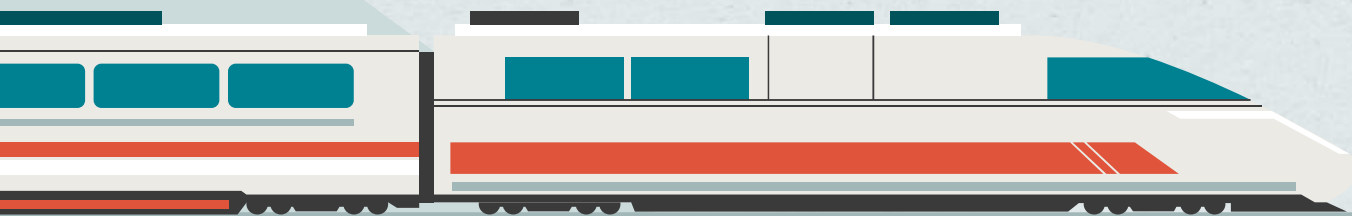# UK TRAINS RIDES DATA ANALYSIS

Microsoft Power BI Engineer - CLS GIZ2_DAT2_G4

رواد مصر الرقمية

UK TRAINS DATA ANALYSIS PROJECT

Project Details

- **Project Name:** UK Train Rides Data Analysis
- **Assigned by:** DEPI Scholarship
- **Duration:** Three Months
- **Tools Used:** Power BI

Our Customer

DEPI Scholarship Management .
CLS Management.

Project Scope

The project involves cleaning and preprocessing UK train ride data using Power BI, identifying key analysis questions, forecasting ride demand and revenue, and visualizing insights through an interactive dashboard. The final phase includes preparing a comprehensive report and presentation to summarize findings and recommendations.

Project Goals

The goal of this project is to analyze UK train rides data using Power BI, enabling data-driven decision-making. The project focuses on data cleaning, analysis, visualization, and forecasting to provide insights into ride demand, ticket class trends, and forecasted revenue, ride delay causes.

Project Time Frame

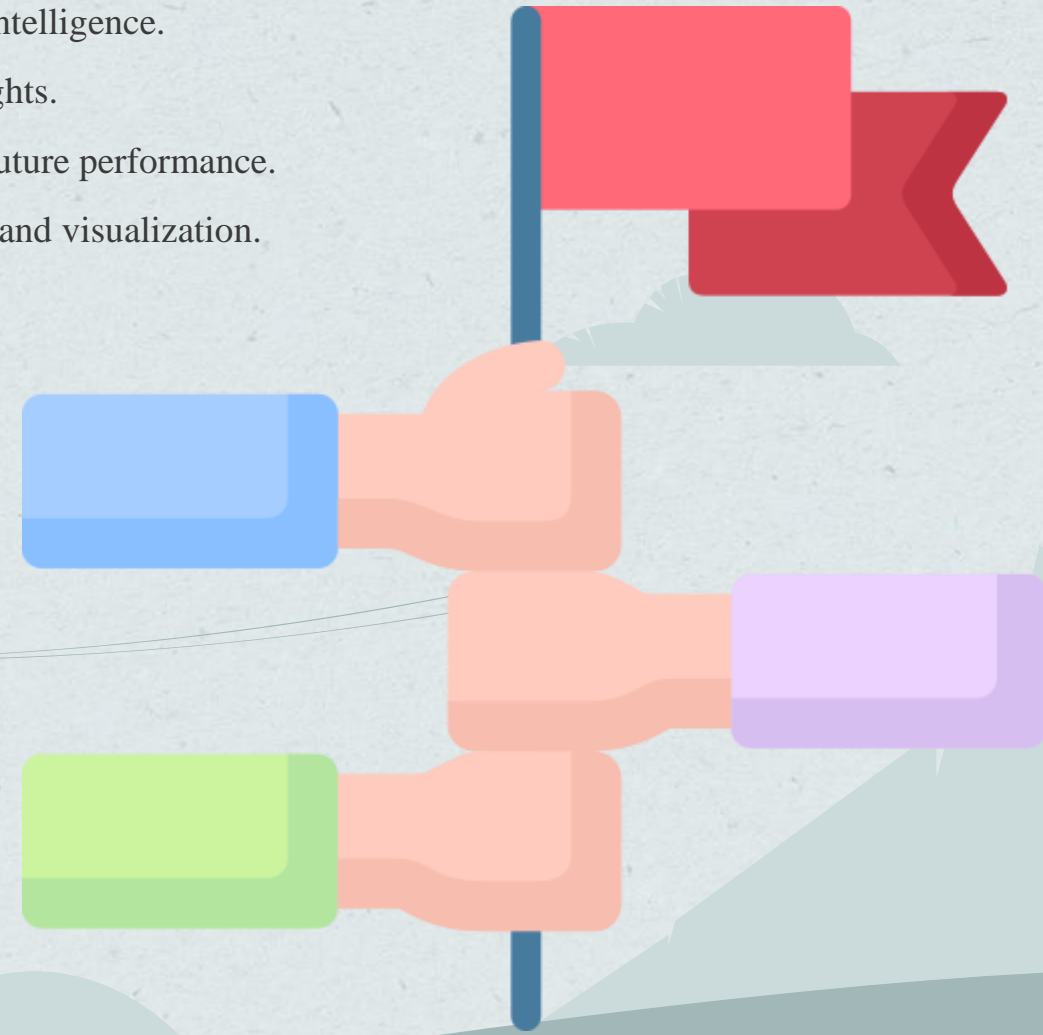|	Start	Deadline
Project Planning & Management	19-Feb-25	24-Feb-25
Literature Review	19-Feb-25	24-Feb-25
Requirements Gathering	19-Feb-25	24-Feb-25
System Analysis & Design	19-Feb-25	24-Feb-25
Implementation	19-Feb-25	11-Apr-25
Final Presentation	19-Feb-25	11-Apr-25

01

Main Objectives

Main Objectives :

- **Unlocking Data Insights:** Transforming raw UK train ride data into actionable business intelligence.

- **Enhancing Decision-Making:** Empowering stakeholders with real-time, data-driven insights.

- **Optimizing Performance:** Identifying demand trends, revenue streams, and forecasting future performance.

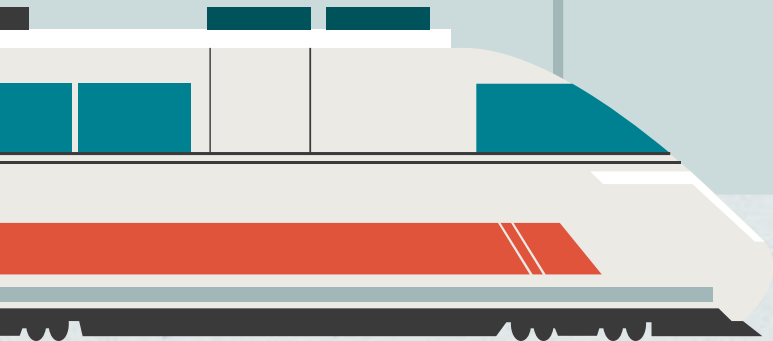- **Interactive Dashboards:** Providing a dynamic Power BI dashboard for in-depth analysis and visualization.

Key Achievements Towards the Goal:

- **Data Cleaning & Processing:** Ensuring accurate and structured data for meaningful analysis.

- **Advanced Data Analytics:** Uncovering hidden trends in train ticket pricing, demand fluctuations, and seasonal patterns.

- **Forecasting Revenue & Demand:** Predicting ticket sales and occupancy rates using Power BI forecasting models.

- **Actionable Recommendations:** Presenting key findings in a comprehensive report for stakeholders.

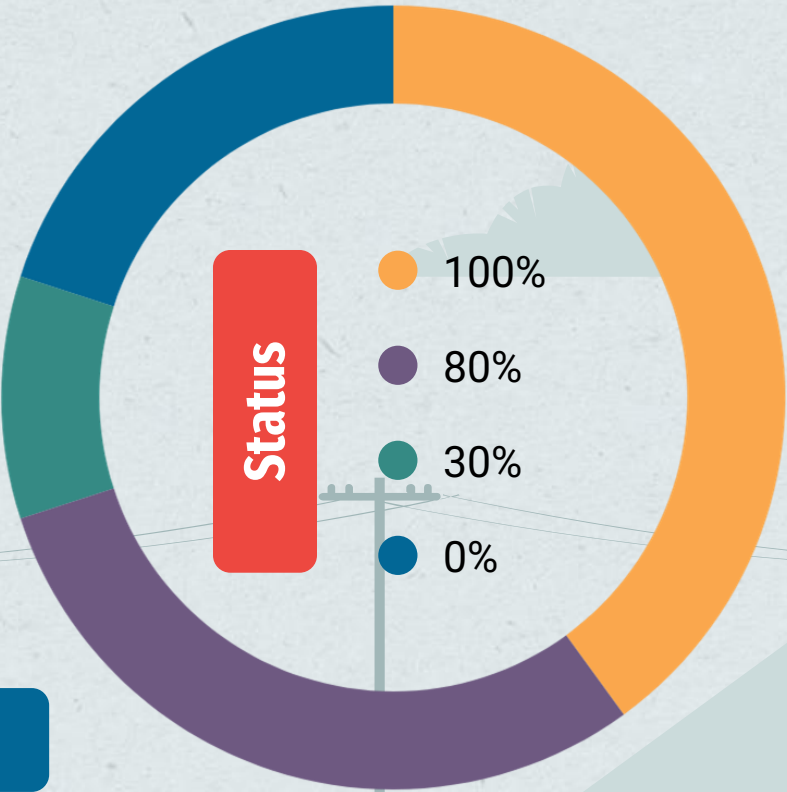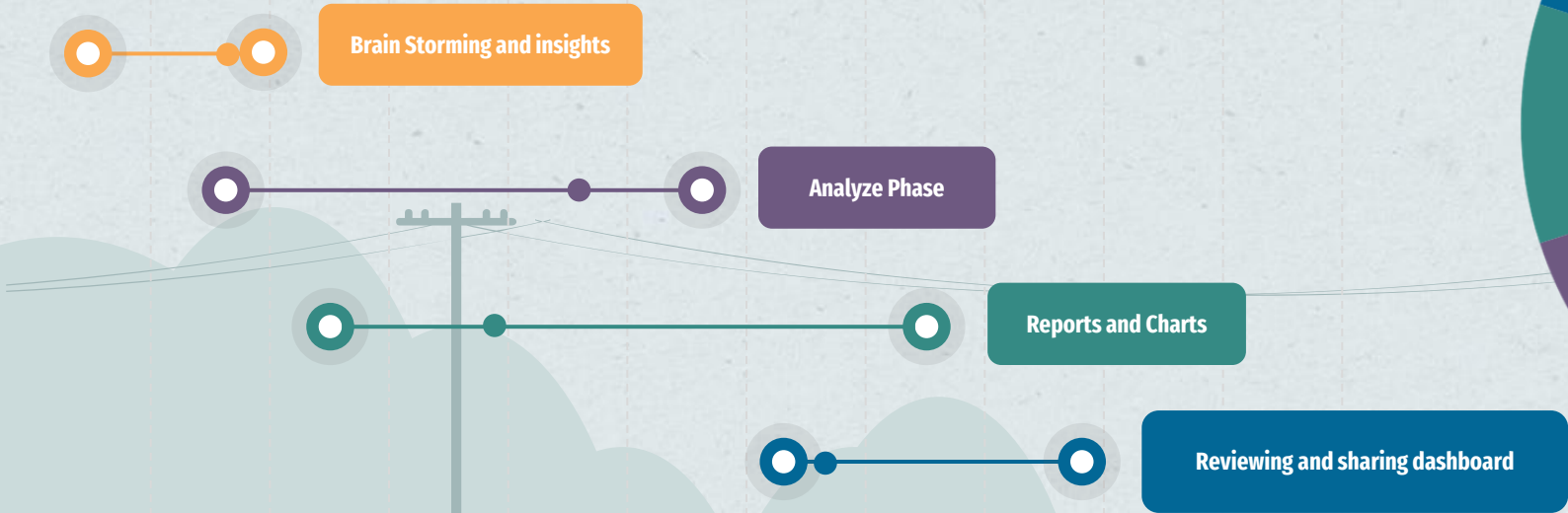# 02

Project Outlines

Project Gantt Chart

	Feb				Mar				Apr			
	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4

Brain Storming and insights

Analyze Phase

Reports and Charts

Reviewing and sharing dashboard

Status

- 100%
- 80%
- 30%
- 0%

03

Risk Assisment

FMEA MODEL

Explanation of the FMEA Model and How It Works:

What is FMEA (Failure Mode and Effects Analysis)?

Failure Mode and Effects Analysis (FMEA) is a structured risk management tool used to identify potential failures in a process, assess their impact, and implement preventive measures. The goal of FMEA is to **prioritize risks** based on their severity, occurrence, and detection difficulty and take proactive steps to mitigate them.

It is widely used in industries such as manufacturing, healthcare, and data analysis to **improve process reliability** and **minimize errors before they cause significant problems**.

Explanation of the FMEA Model and How It Works:

How FMEA Works?

FMEA follows a systematic approach with the following key steps:

1. **Identify Failure Modes** – Recognizing potential issues that could occur in the process.

2. **Determine Effects** – Analyzing the impact of each failure mode.

3. **Find Possible Causes** – Identifying the root causes of failures.

4. **Assign Severity (S), Occurrence (O), and Detection (D) Scores**:

   ❑ **Severity (S):** How serious is the failure? (Scale 1-10, with 10 being the worst)

   ❑ **Occurrence (O):** How often does this failure happen? (Scale 1-10, with 10 being very frequent)

   ❑ **Detection (D):** How difficult is it to detect the failure before it happens? (Scale 1-10, with 10 being very hard to detect)

5. **Calculate Risk Priority Number (RPN)** – Using the formula: $RPN = S \times O \times D$

6. **Prioritize & Implement Solutions** – Addressing failures with the highest RPN values first.

Applying FMEA Failure Mode and Effects Analysis:

Detection Method	S	O	D	RPN	RPN%	Proposed Solution
Power BI data profiling and Power Query validation	9	6	5	270	%19	Apply Power Query transformations: remove duplicates, handle missing data, and standardize formats.
Stakeholder feedback and user testing	8	5	6	240	%17	Simplify dashboards, use clear labels, and implement intuitive slicers and filters.
Model validation and relationship checks	9	6	4	216	%15	Ensure correct table joins, use star schema, and review relationships in Power BI.
Model validation & error measurement (MAPE, RMSE)	9	5	4	180	%13	Compare multiple forecasting methods and validate predictions against past trends.
Regular stakeholder reviews	9	4	5	180	%13	Conduct frequent feedback sessions and align reports with business goals.
Dashboard performance testing	7	5	5	175	%12	Use calculated columns instead of measures where possible, and optimize visual load.
Performance monitoring & DAX optimization	7	6	4	168	%12	Optimize data model, apply aggregations, and reduce unnecessary columns.

FMEA INSIGHTS:

1.**Prioritizing Data Quality**

   1. Power BI **data profiling & Power Query validation** have the highest **Risk Priority Number (RPN: 270, 19%)**, highlighting that **data consistency, missing values, and incorrect formatting** are major concerns.
   2. **Action:** Implement robust **data cleaning & transformation** strategies before analysis.

2.**User-Centric Dashboard Improvements**
   1. **Stakeholder feedback & user testing (RPN: 240, 17%)** is critical to ensure that reports and dashboards **align with business needs** and **improve usability**.
   2. **Action:** Focus on **intuitive design, slicers, filters, and clear labeling** for better stakeholder engagement.

3.**Data Model Integrity & Relationship Checks**
   1. **Incorrect relationships and table joins (RPN: 216, 15%)** could lead to **misleading insights**.
   2. **Action:** Ensure **star schema modeling** and **validate table joins** to prevent reporting errors.

4.**Forecasting & Model Validation Risks**
   1. **Model validation & error measurement (RPN: 180, 13%)** indicate that **MAPE, RMSE, and other error metrics** need continuous monitoring.
   2. **Action:** Compare multiple forecasting methods and **benchmark predictions** against past trends.

5.**Frequent Stakeholder Reviews Enhance Decision-Making**
   1. **Regular stakeholder reviews (RPN: 180, 13%)** ensure that dashboards remain aligned with business needs.
   2. **Action:** Schedule periodic feedback sessions to adjust metrics and reporting structures accordingly.

6.**Performance Optimization for Scalability**
   1. **Dashboard performance testing (RPN: 175, 12%)** and **DAX optimization (RPN: 168, 12%)** show that **slow performance** is a recurring issue.
   2. **Action:** Optimize visuals, use **aggregations, calculated columns instead of measures**, and **remove unnecessary columns** for better efficiency.
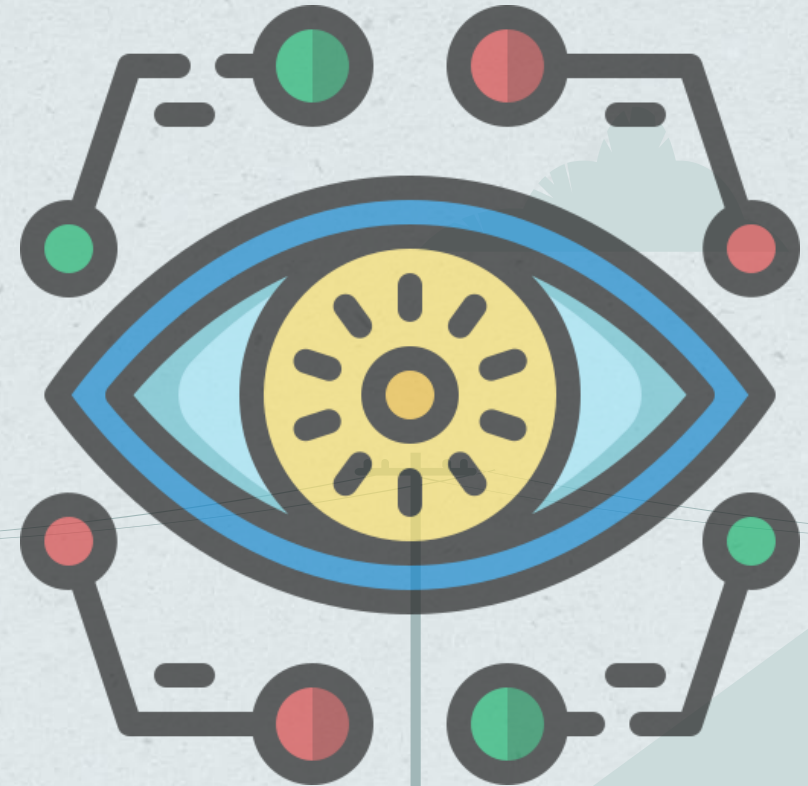
MEET OUR TEAM

Mazen Ramadan
Power BI Engineer

Mahmoud Zakaria
Power BI Engineer

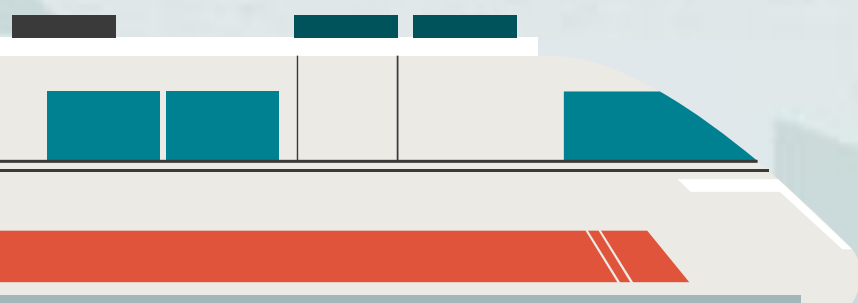**Abdelrahman Elkhateeb**
Power BI Engineer

Iman Fathy
Power BI Engineer

Mahmoud Mostafa
Power BI Engineer

Nada Salem
Power BI Engineer

TASK ASSIGMENT AND RULES

RESPONSABILITES FOR TEAM MEMBERS

Mazen

- Brainstorming.
- Data Visualization.
- Ticketing Pricing Insights.
- DAX Measures

Mahmoud Zakaria

- Brainstorming.
- Data Visualization.
- Trip Insights Analysis.
- Theoretical Part Management.
- PowerPoint Presentation Design.
- Dashboard Design.
- Forecasting.
- FMEA Model

Abdelrahman

- Brainstorming.
- Data Visualization.
- Bookmarks.
- Data Modeling.
- Slicers.
- Trip Insights Analysis.
- Dashboard Design.
- Forecasting.

Iman

- Brainstorming.
- Data Visualization.
- Ticketing Pricing Insights.
- Dax Measures.
- GitHub Management

Mahmoud

- Brainstorming.
- Data Visualization.
- Trip Insights Analysis.
- Theoretical Part Management.

Nada

- Brainstorming.
- Data Visualization.
- Ticketing Pricing Insights,
- Theoretical Part Management.
- GitHub Management

Thanks!

Microsoft Power BI Engineer - CLS GIZ2_DAT2_G4