



TASK 2

Student Name

Omar Hamed ALi

Course

Power Bi Specialist Final Project

Date of Submission

Tuesday sep 16, 2025

Instructor

Eng. Sherihan Ali

Initiative

DEPI

TemplateLAB



رواد مصر الرقمية

Problem Definition and Research Questions

Title Page:

- Project Title: Analyzing **MTA** Ridership Patterns Post-COVID-19
- Team Members:
 - Omar Hamed Ali
 - Soliman yehia Soliman Mahmoud
 - Mohamed Ahmed Mohamed AlFakharany
 - Samir Mahmoud Samir
 - Ahmed Nasr Abdel Motey Mesalem
- Supervisor: eng. **Sherihan Ali**
- Date: [9/26/2025]

Checking the Data (Assumptions & Data Quality)

What You Checked	What You Found	What You Did
Missing Values	0 blanks	
Duplicates	0 Duplicates	
Outliers	0 indicates no passengers	Removed records with negative sales
Column Types	The columns was text	Changed he columns type from text to whole numbers refer to the number of people and percentage refer to the %

- Steps You Did in Power Query

Checking the Data (Assumptions & Data Quality)

Steps You Did in Power Query

Write step-by-step what you did to clean the data. You can use numbers like a recipe!

Add screenshots if you can.

No Duplicates

The screenshot shows the Power BI Query Editor interface. On the left, there is a preview pane with three columns: Date, Subways: Total Estimated Ridership, and % Subways: % of Comparable Pre-Pandemic Day. A tooltip indicates "This table is empty." On the right, the "APPLIED STEPS" pane is open, showing a step named "Kept Duplicates".

No Blanks

Rows Before blanks = Rows After Remove Blanks

The screenshot shows the Power BI Query Editor interface. On the left, there is a preview pane with four columns: Date, Subways: Total Estimated Ridership, % Subways: % of Comparable Pre-Pandemic Day, and Buses: Total Estimated Ridership. Below the preview are three bar charts showing the distribution of Valid, Error, and Empty values for each column. On the right, the "APPLIED STEPS" pane is open, showing a step named "Removed Blank Rows".

No error

The screenshot shows the Power BI Query Editor interface. On the left, there is a preview pane with five columns: % Staten Island Railway: % of Comparable Pre-Pandemic Day, Year, Month, Day, and Quarter. Below the preview are three bar charts showing the distribution of Valid, Error, and Empty values for each column. On the right, the "APPLIED STEPS" pane is open, showing a step named "Kept Errors".

Added new columns

1 ² ₃ Year	1 ² ₃ Month	1 ² ₃ Day	1 ² ₃ Quarter
<ul style="list-style-type: none"> ● Valid 100% ● Error 0% ● Empty 0% 	<ul style="list-style-type: none"> ● Valid 100% ● Error 0% ● Empty 0% 	<ul style="list-style-type: none"> ● Valid 100% ● Error 0% ● Empty 0% 	<ul style="list-style-type: none"> ● Valid 100% ● Error 0% ● Empty 0% 
5 distinct, 0 unique	12 distinct, 0 unique	31 distinct, 0 unique	4 distinct, 0 unique
2020		3	1
2020		3	2
2020		3	3
2020		3	4
2020		3	5
2020		3	6
2020		3	7
2020		3	8
2020		3	9
2020		3	10
2020		3	11

Mention if you created new columns (e.g., "Year" from "Order Date").

Final Clean Dataset (Before & After)

Before Cleaning	After Cleaning
1706 rows	1706 rows
0 missing customer names	0 missing names (removed them)