**Column Explanations and Suggested Data Types for Power BI**

To optimize your Power BI dashboard, here are the recommended data types for each column:

| **Column Name** | **Description** | **Suggested Data Type (Power BI)** |
| --- | --- | --- |
| Date | The date of the ridership record | Date/Time |
| Subway Ridership | Total subway ridership on that date | Whole Number |
| Subway % Pre-Pandemic | Percentage of pre-pandemic ridership levels | Decimal Number (or Percentage) |
| Bus Ridership | Total bus ridership | Whole Number |
| Bus % Pre-Pandemic | Percentage of pre-pandemic ridership levels | Decimal Number (or Percentage) |
| LIRR Ridership | Long Island Rail Road ridership | Whole Number |
| LIRR % Pre-Pandemic | Percentage of pre-pandemic ridership levels | Decimal Number (or Percentage) |
| Metro-North Ridership | Metro-North train ridership | Whole Number |
| Metro-North % Pre-Pandemic | Percentage of pre-pandemic ridership levels | Decimal Number (or Percentage) |
| Access-A-Ride Trips | Trips taken via Access-A-Ride service | Whole Number |
| Access-A-Ride % Pre-Pandemic | Percentage of pre-pandemic trips | Decimal Number (or Percentage) |
| Bridges & Tunnels Traffic | Total number of vehicles using MTA bridges/tunnels | Whole Number |
| Bridges & Tunnels % Pre-Pandemic | Percentage of pre-pandemic traffic levels | Decimal Number (or Percentage) |
| Staten Island Railway Ridership | Staten Island Railway ridership | Whole Number |
| Staten Island % Pre-Pandemic | Percentage of pre-pandemic ridership levels | Decimal Number (or Percentage) |

**Power BI Dashboard Plan for MTA Daily Ridership Data**

**1. Data Preparation**

* Convert the Date column to **DateTime format**.
* Create a **Calendar Table** for time-based analysis.
* Ensure correct **data types** for each column (numeric for ridership figures).

**2. Measures to Create in Power BI**

Using DAX (Data Analysis Expressions):

1. **Total Ridership**

DAX

CopyEdit

Total Ridership = SUM('MTA Data'[Subway Ridership])

+ SUM('MTA Data'[Bus Ridership])

+ SUM('MTA Data'[LIRR Ridership])

+ SUM('MTA Data'[Metro-North Ridership])

+ SUM('MTA Data'[Access-A-Ride Trips])

+ SUM('MTA Data'[Bridges & Tunnels Traffic])

+ SUM('MTA Data'[Staten Island Railway Ridership])

1. **Average Ridership per Day**

DAX

CopyEdit

Avg Daily Ridership = AVERAGEX(VALUES('MTA Data'[Date]), [Total Ridership])

1. **Year-over-Year (YoY) Growth**

DAX

CopyEdit

YoY Growth =

VAR CurrentYear = YEAR(MAX('MTA Data'[Date]))

VAR PreviousYear = CurrentYear - 1

VAR PreviousYearRidership = CALCULATE([Total Ridership], YEAR('MTA Data'[Date]) = PreviousYear)

RETURN

DIVIDE([Total Ridership] - PreviousYearRidership, PreviousYearRidership, 0)

1. **Ridership % of Pre-Pandemic Levels**

DAX

CopyEdit

% Pre-Pandemic = AVERAGE('MTA Data'[Subway % Pre-Pandemic]

+ 'MTA Data'[Bus % Pre-Pandemic]

+ 'MTA Data'[LIRR % Pre-Pandemic]

+ 'MTA Data'[Metro-North % Pre-Pandemic]

+ 'MTA Data'[Access-A-Ride % Pre-Pandemic]

+ 'MTA Data'[Bridges & Tunnels % Pre-Pandemic]

+ 'MTA Data'[Staten Island % Pre-Pandemic]) / 7

**3. Recommended Visuals**

**A. Line Chart (Ridership Trends Over Time)**

* **X-Axis**: Date
* **Y-Axis**: Total Ridership
* **Legend**: Transportation modes (Subway, Bus, LIRR, etc.)
* **Purpose**: To analyze ridership trends over time.

**B. Clustered Bar Chart (Ridership by Transportation Type)**

* **X-Axis**: Total Ridership
* **Y-Axis**: Transportation Type (Subway, Bus, etc.)
* **Purpose**: To compare ridership distribution among different modes.

**C. Card Visuals (KPIs)**

* Total Ridership
* YoY Growth
* % Pre-Pandemic
* Avg Daily Ridership
* **Purpose**: Key insights at a glance.

**D. Pie Chart (Ridership Distribution by Type)**

* **Values**: Total Ridership
* **Category**: Transportation Type
* **Purpose**: To understand ridership share across transport modes.

**E. Table (Detailed Ridership Data)**

* **Columns**: Date, Subway Ridership, Bus Ridership, LIRR Ridership, Metro-North Ridership, etc.
* **Purpose**: To provide a detailed breakdown of ridership.

**F. Heatmap (Day of the Week Analysis)**

* **X-Axis**: Day of Week
* **Y-Axis**: Month
* **Values**: Total Ridership
* **Purpose**: To identify high and low ridership days.

**4. Best Theme for the Dashboard**

* **Dark Mode Theme** for high contrast.
* Use **MTA Blue and Yellow** for branding.
* Keep **KPIs in bold white fonts** for visibility.
* Use **gradient colors for heatmaps** to show intensity.
* Ensure **consistent font style and layout**.

**1️⃣ Line Chart (Trend Analysis Over Time)**

🔹 **Purpose**: Show the trend of **total ridership** over time.

✅ **Setup**:

* **X-Axis** → Date (formatted as Month & Year or Day)
* **Y-Axis** → Total Ridership
* **Legend** (Optional) → Transport Type

💡 **Insight**: Identify seasonal variations, peak ridership months, and long-term trends.

**2️⃣ Area Chart (Cumulative Ridership)**

🔹 **Purpose**: Show ridership trends over time with an emphasis on volume.

✅ **Setup**:

* **X-Axis** → Date (Month & Year)
* **Y-Axis** → Total Ridership
* **Legend** → Transport Type

💡 **Insight**: Helps visualize how different transport types contribute to overall ridership growth.

**3️⃣ Treemap Chart (Ridership Distribution by Transport Type)**

🔹 **Purpose**: Show ridership share across different transport modes.

✅ **Setup**:

* **Category (Groups)** → Transport Type
* **Values** → Total Ridership

💡 **Insight**: Identify which transport mode has the **highest** or **lowest** ridership.

**4️⃣ Funnel Chart (Ridership Drop-off Through Modes)**

🔹 **Purpose**: Show how ridership flows across different transport modes.

✅ **Setup**:

* **Stages (Category)** → Transport Type
* **Values** → Total Ridership

💡 **Insight**: Identify **where ridership is strongest and weakest** in the transport system.

**5️⃣ Line & Clustered Column Chart (Ridership vs. Pre-Pandemic Levels)**

🔹 **Purpose**: Compare **current ridership** with **pre-pandemic ridership levels**.

✅ **Setup**:

* **X-Axis** → Date (Month & Year)
* **Columns (Bars)** → Total Ridership
* **Line (Secondary Axis)** → % Pre-Pandemic (calculated measure)
* **Legend** → Transport Type (Optional)

💡 **Insight**: See how ridership is **recovering post-pandemic** compared to historical levels.

**6️⃣ Stacked Bar Chart (Weekday vs. Weekend Ridership)**

🔹 **Purpose**: Compare ridership on **weekdays vs. weekends**.

✅ **Setup**:

* **X-Axis** → Day Name
* **Y-Axis** → Total Ridership
* **Legend** → Transport Type

💡 **Insight**: Helps analyze **peak days** for transit usage.

**7️⃣ Donut/Pie Chart (Ridership Share by Transport Type)**

🔹 **Purpose**: Show percentage share of each transport mode.

✅ **Setup**:

* **Legend** → Transport Type
* **Values** → Total Ridership

💡 **Insight**: **Proportional distribution** of ridership across various transit options.