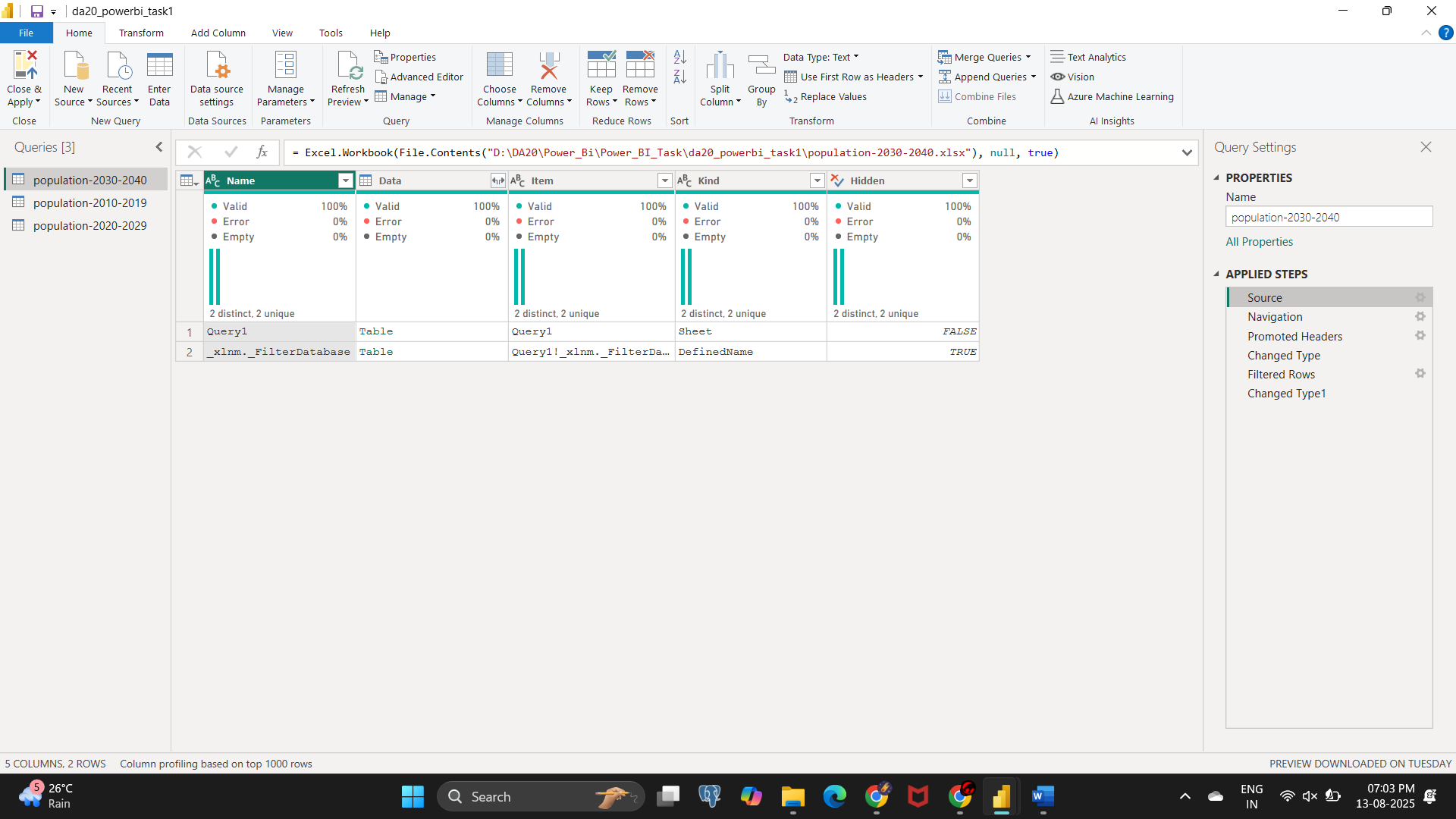
## Power BI Data Preparation

This document provides a comprehensive overview of the data preparation steps for the population-2030-2040 query, as shown in the provided screenshots. The purpose of this documentation is to ensure clarity, reproducibility, and maintainability of the data model. By detailing each step of the Power Query process, we can understand how the raw data was transformed into a clean and structured format suitable for analysis and reporting.

### 1. Data Source and Connection

This initial step establishes the link between Power BI and the raw data source.

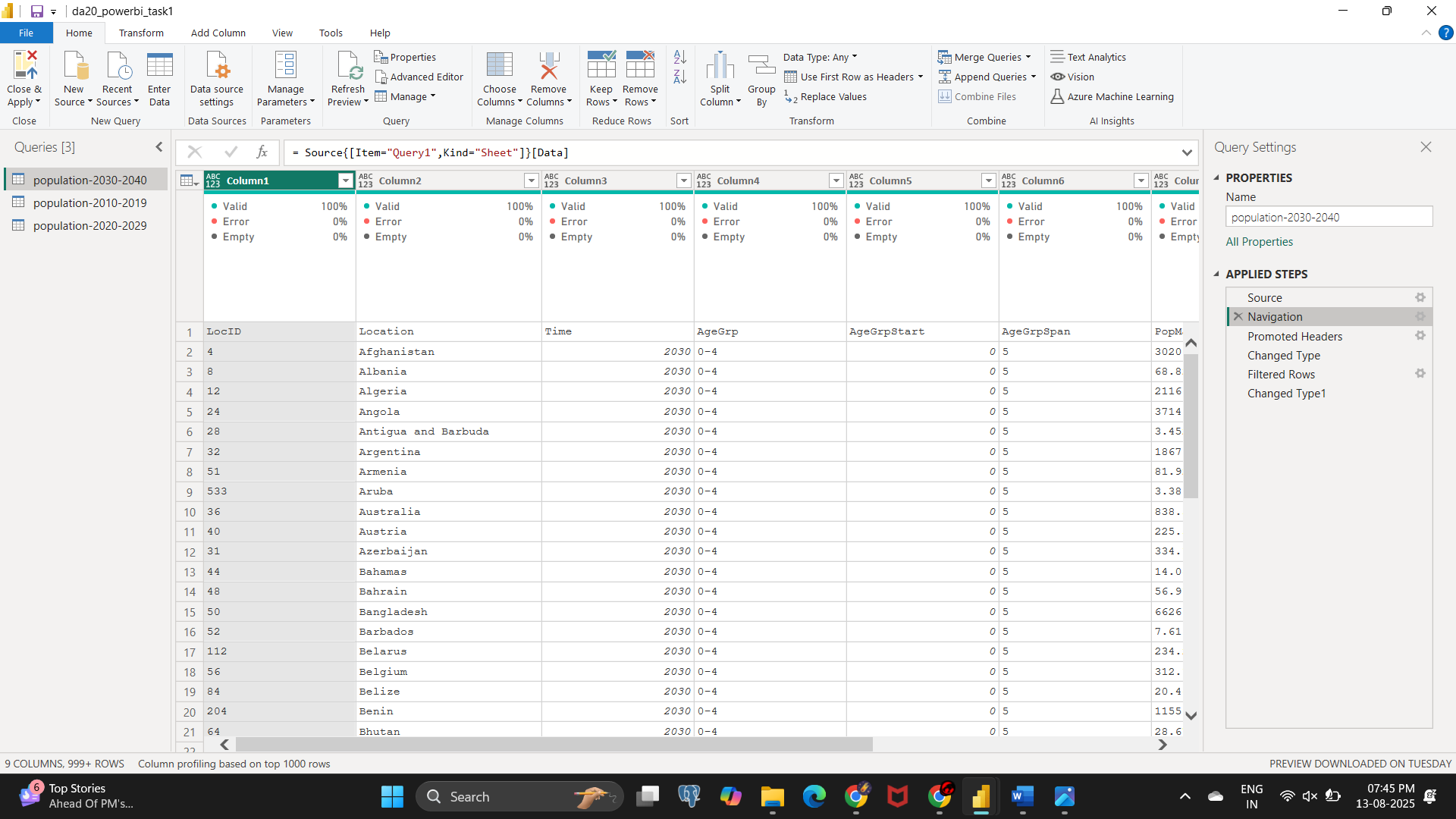
* **Details:** The query connects to an Excel file named population-2030-2040.xlsx located at the file path D:\DA20\powerbi\_task1\.



### 2. Navigation

Once the connection is made, this step selects the specific data to be loaded into the query editor.

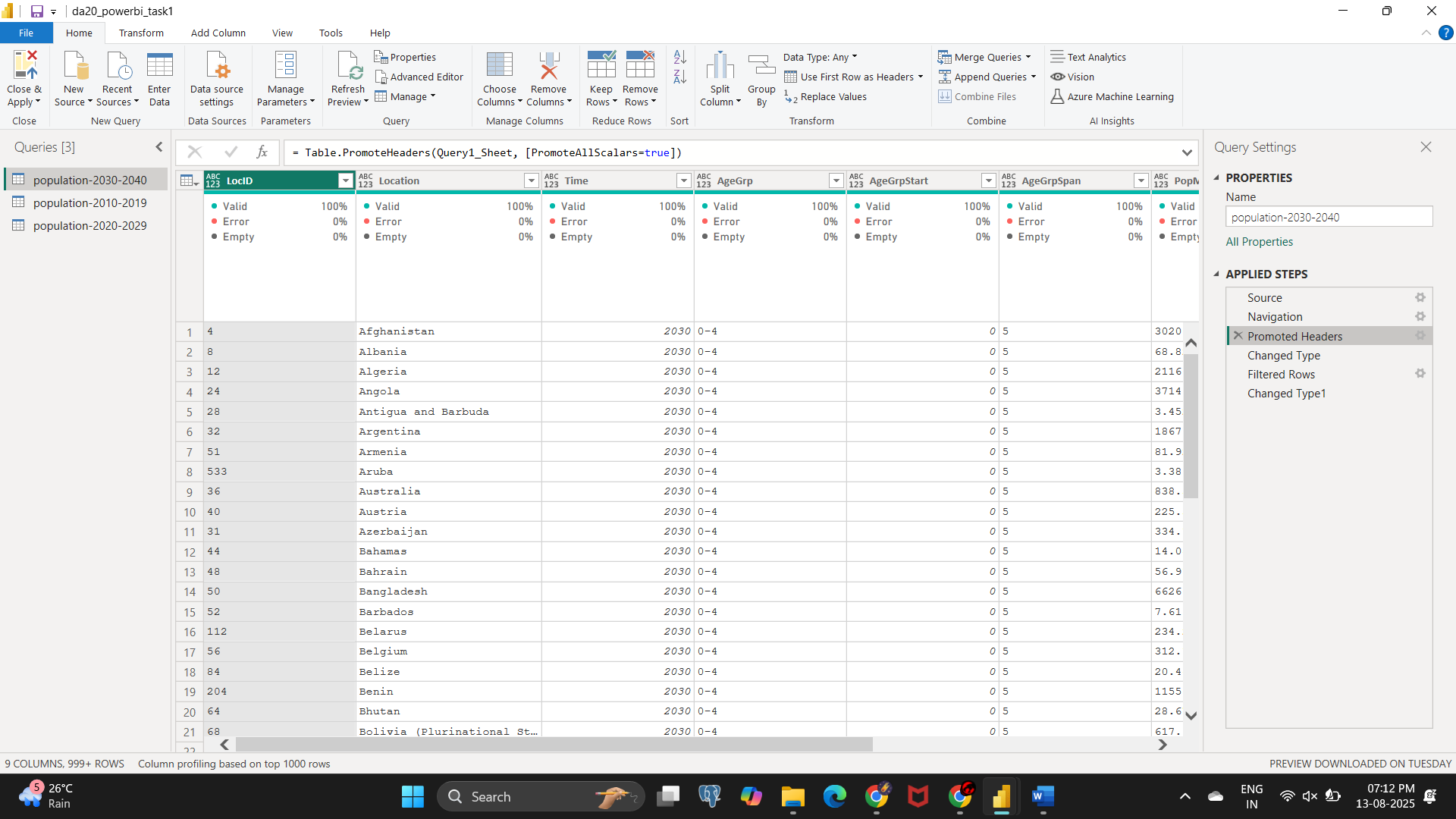
* **Details:** The query navigates to and selects the table named Query1, which contains the population data.



### 3. Promoted Headers

This is a crucial data-shaping step that correctly identifies the column headers.

* **Purpose:** The first row of the data, which contains the column names like LocID, Location, and Time, is promoted to become the official headers for the table. This ensures that Power BI recognizes the correct labels for each column.
* **Tip:** Always check that the first row contains meaningful header names before promoting them. If not, you may need to rename the columns after this step.



### 4. Changed Type

Power Query automatically attempts to determine the appropriate data type for each column.

* **Purpose:** This step automatically assigns a data type to each column (e.g., Text, Number, Date). This initial data typing is essential for preventing errors and ensuring that subsequent transformations work correctly.
* **Details:** Based on the data in the columns, Power Query assigns types. For example, LocID and Time are likely set to a numeric type, while Location is set to Text

A screenshot of a computer

AI-generated content may be incorrect.

### 5. Filtered Rows

This step is critical for data cleaning and quality control.

* **Purpose:** It removes rows that do not meet a specified condition. In this case, rows with null or empty values are removed to prevent incomplete data from affecting the analysis.
* **Details:** The screenshot shows that rows with null values were filtered out, ensuring only complete records are kept.

A screenshot of a computer

AI-generated content may be incorrect.

### 6. Changed Type1

This final step refines the data types, overriding any automatic changes to ensure the data model is perfectly aligned with analytical needs.

* **Purpose:** To manually and explicitly set the data types for each column. This step is a best practice to guarantee that the data is in the correct format for calculations, visualizations, and modeling.
* **Details:** As per the screenshot, the data types were specifically set for the columns:
  + Location: Set to **Text**
  + Time: Set to **Text**
  + AgeGrpStart: Set to **Whole Number**

A screenshot of a computer

AI-generated content may be incorrect.

### Summary

The transformations performed have successfully cleaned and structured the raw Excel data into a usable table. The next steps in a typical Power BI workflow would involve loading this data into the data model, creating relationships with other tables (if applicable), and developing DAX measures for advanced calculations and reporting. Regular checks of the data source and a review of these transformation steps are recommended to maintain data integrity.