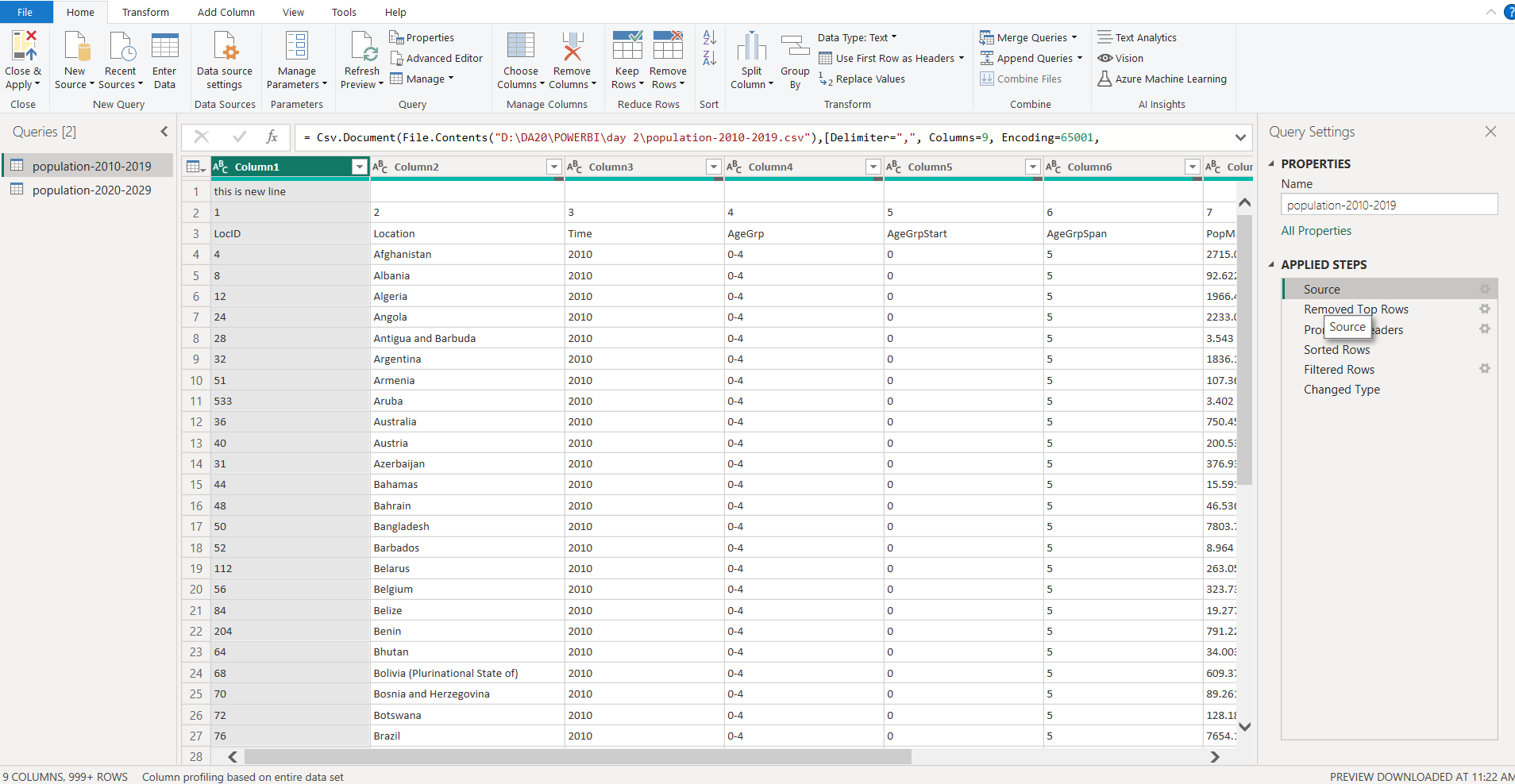
**Data Preparation Explained: Power BI (Population\_2010-2019)**

Here’s a detailed, presentation-style explanation of how the data was prepared in Power BI — going beyond basic steps, focusing on best practices and rationale for each transformation.

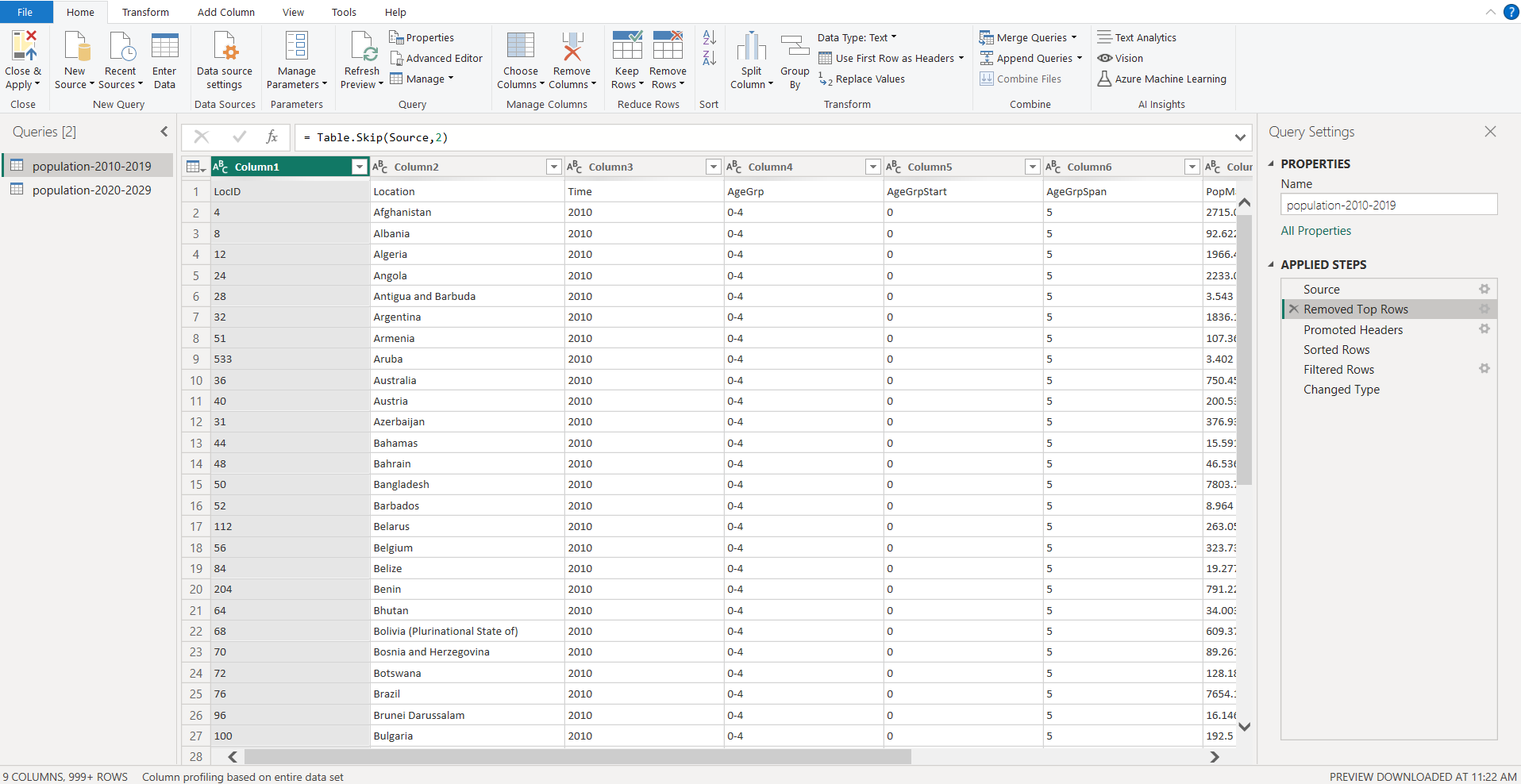
**1. Data Import :**

* **Action:** Loaded raw population CSV/excel data into Power BI.
* **Why:** Provides the starting point; ensures direct access to all source records for cleaning.



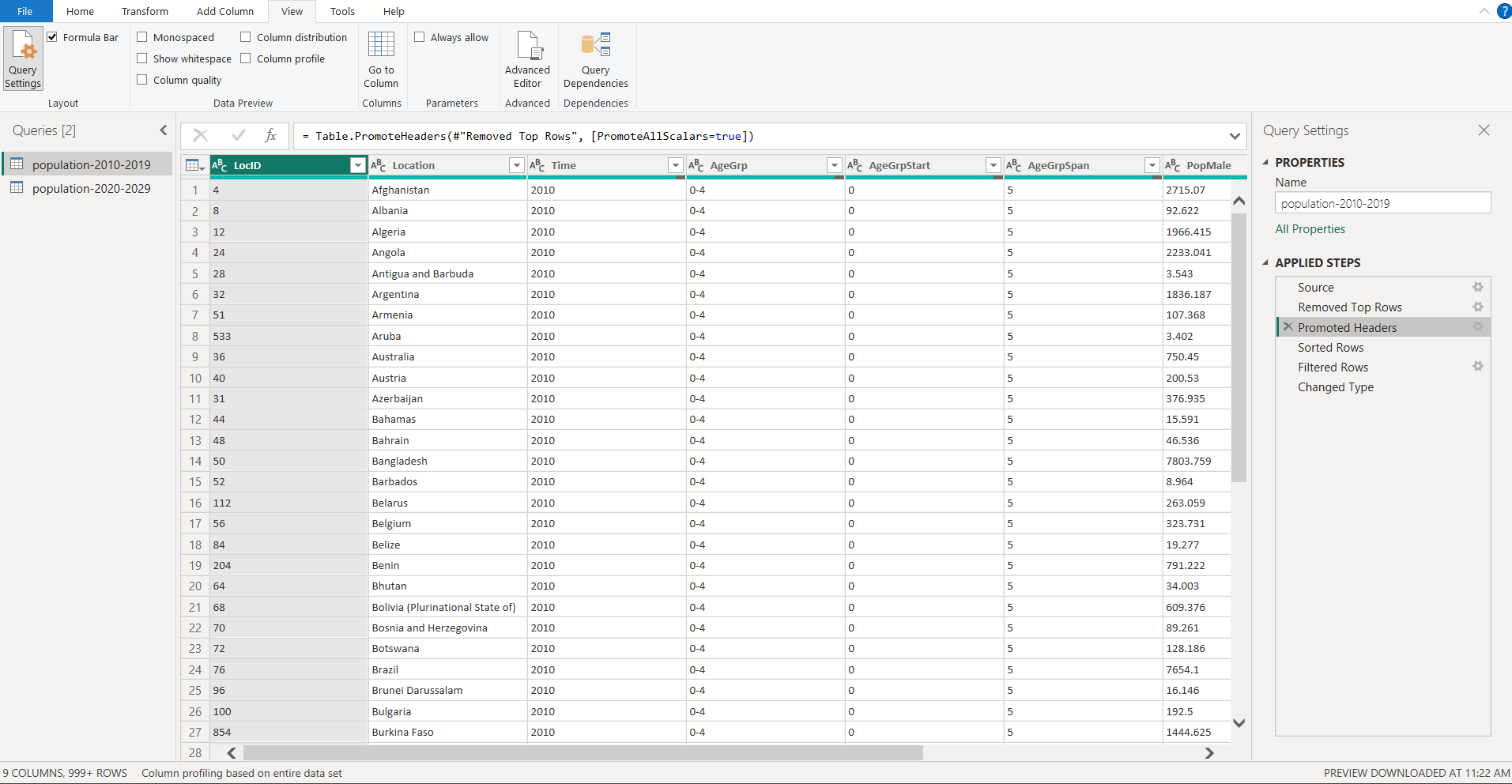
**2. Remove Top Rows :**

* **Action:** Deleted extra rows (e.g., metadata, blank headers) that could interfere with schema recognition.
* **Why:** Cleans dataset for proper header and value alignment.



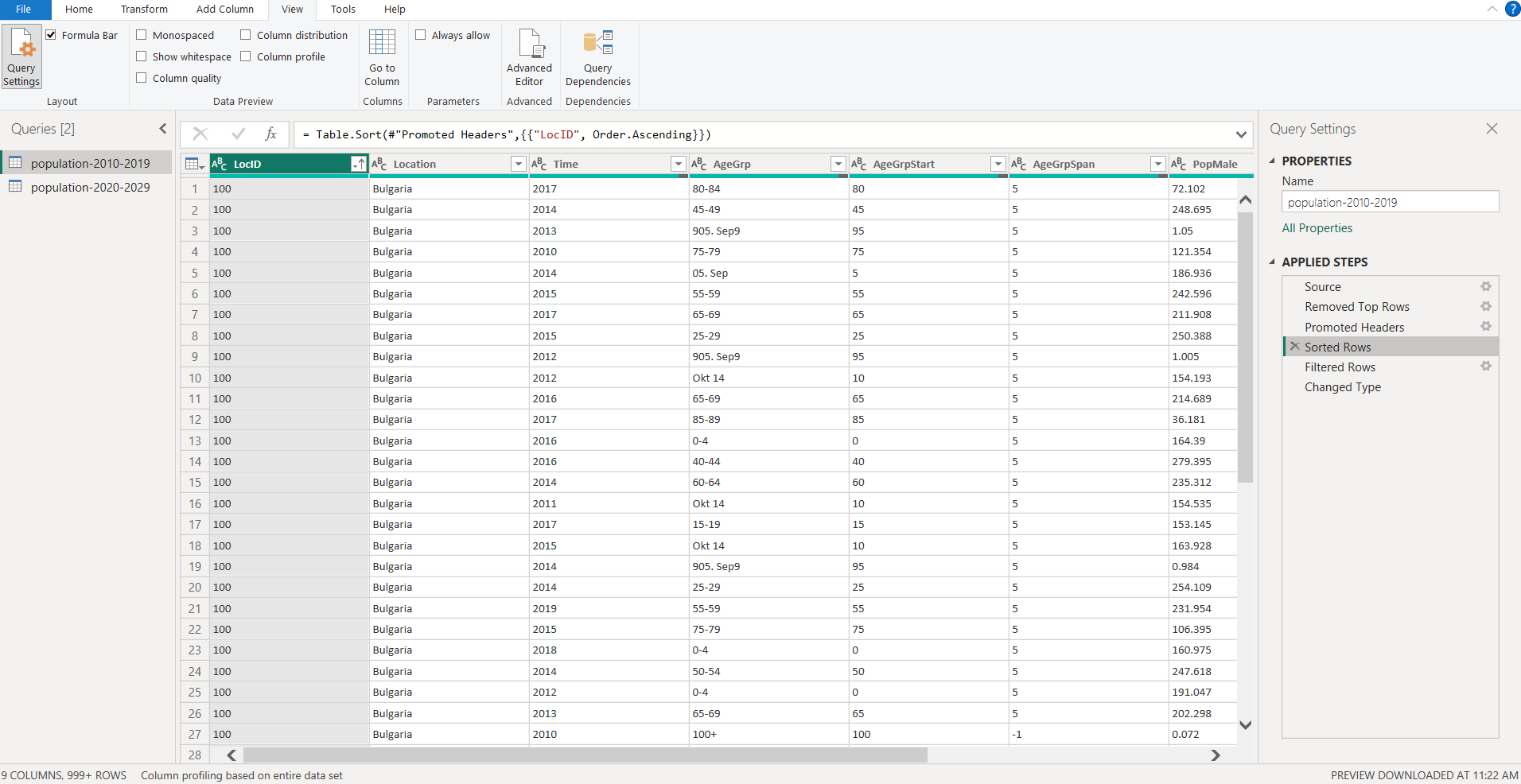
**3. Promote Headers :**

* **Action :** Elevated relevant row as table headers (column names).
* **Why :** Ensures columns like LocID, Location, Time, etc, are correctly identified, essential for downstream transformations.
* **Tip :** Always verify that promoted headers match source documentation.



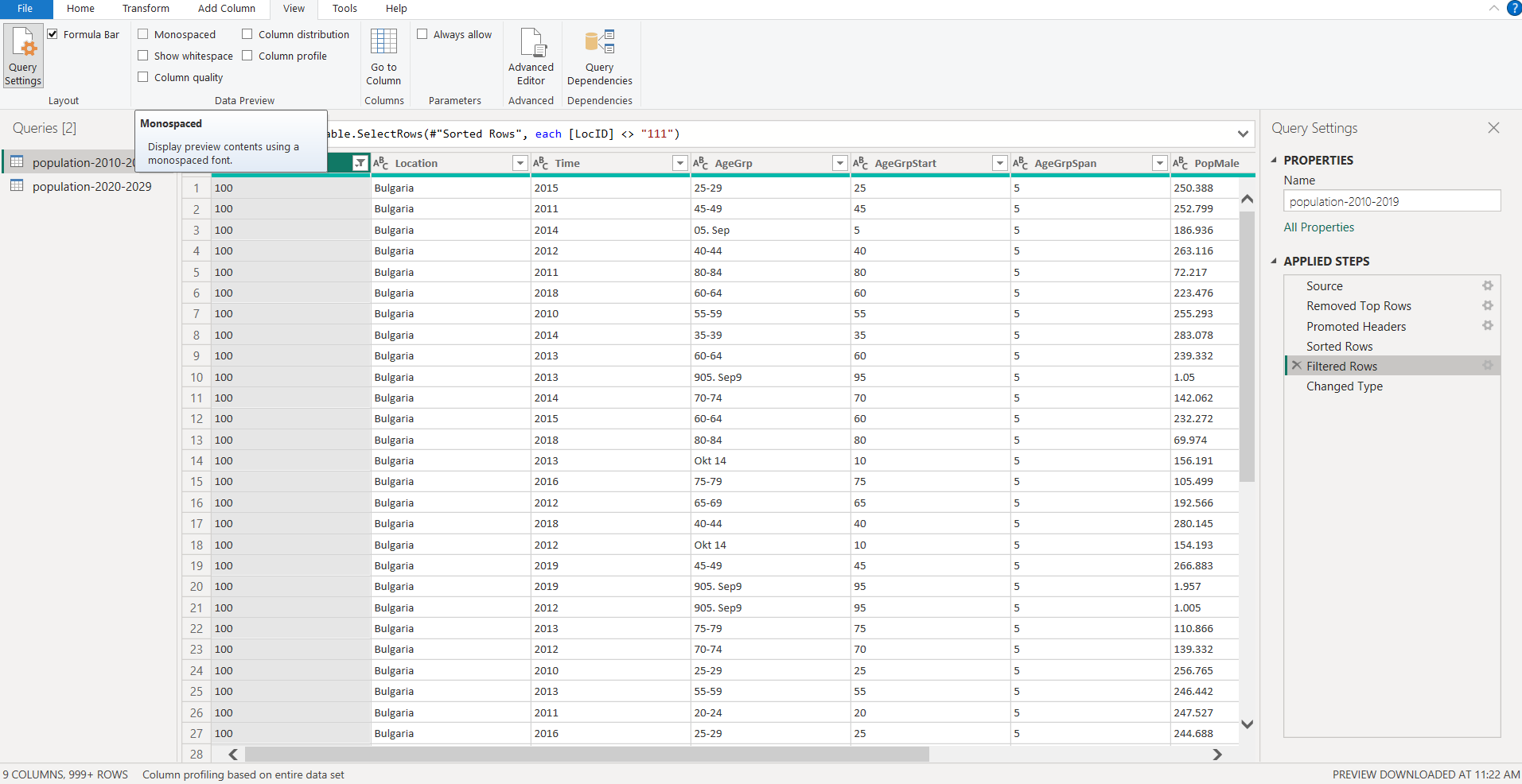
**4. Sort Data :**

* **Action :** Ordered rows by a key column (i.e. , Time or Location).
* **Why :** Facilitates analysis (trends over years, regional comparison), easier data verification, and presentation consistency.



**5. Filter Rows :**

* **Action :** Applied filters to include only useful years or locations (removing errors, blanks, unwanted categories).
* **Why :** Targeted dataset enables accurate analysis, reduces noise, and boosts performance in Power BI visuals.

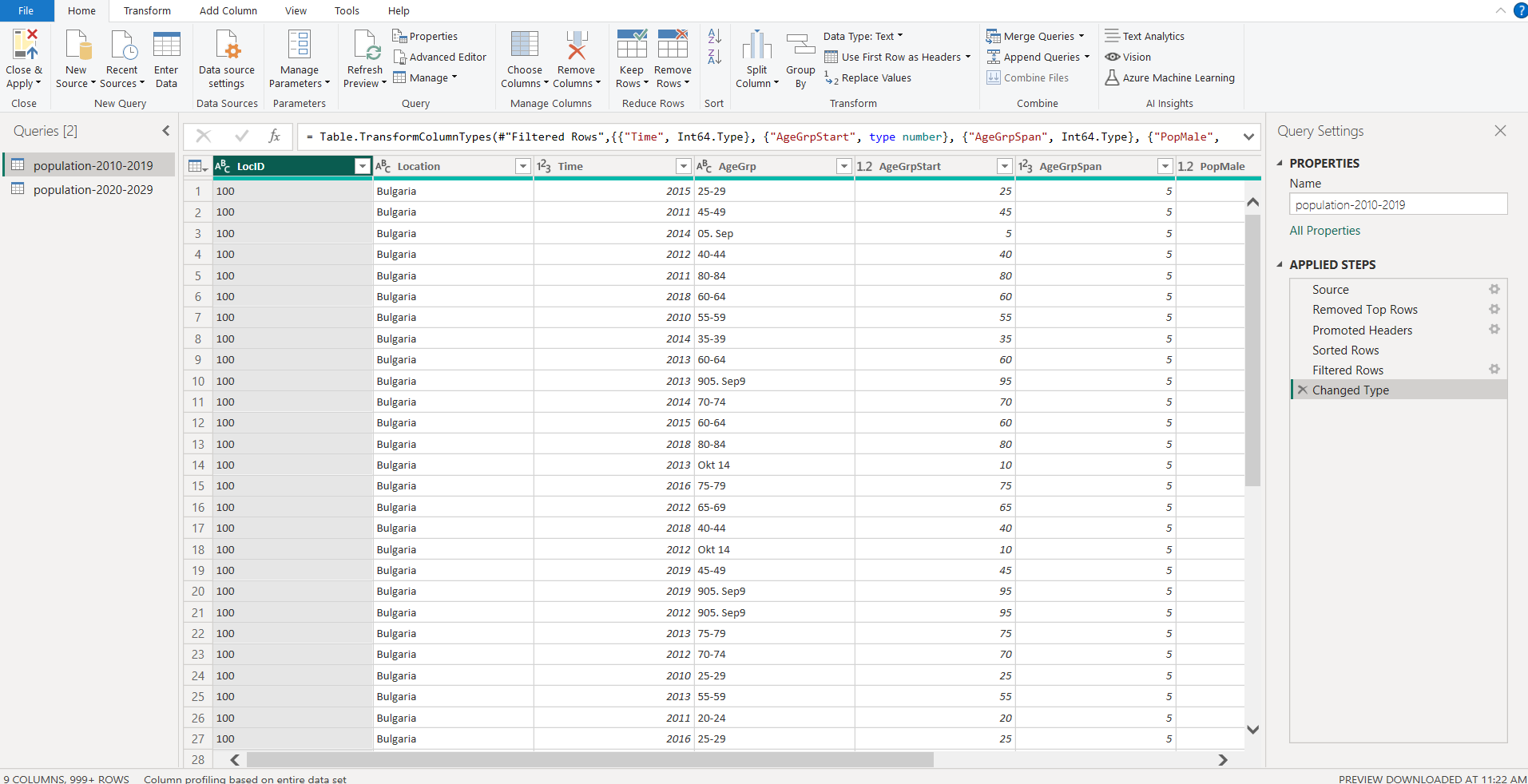
****

**6. Change Data Types :**

* **Action :** Explicitly set correct types:
  + Time (Year): Integer
  + AgeGrpStart : Number/Integer
  + AgeGrpSpan : Number/Integer
  + PopMale : Integer

**Why :** Data types ensure reliable calculations, sorting, and aggregations within Power BI.

**Note :** Mismatched types can break measures (e.g. sum, average).



**7. Review & Validate Columns :**

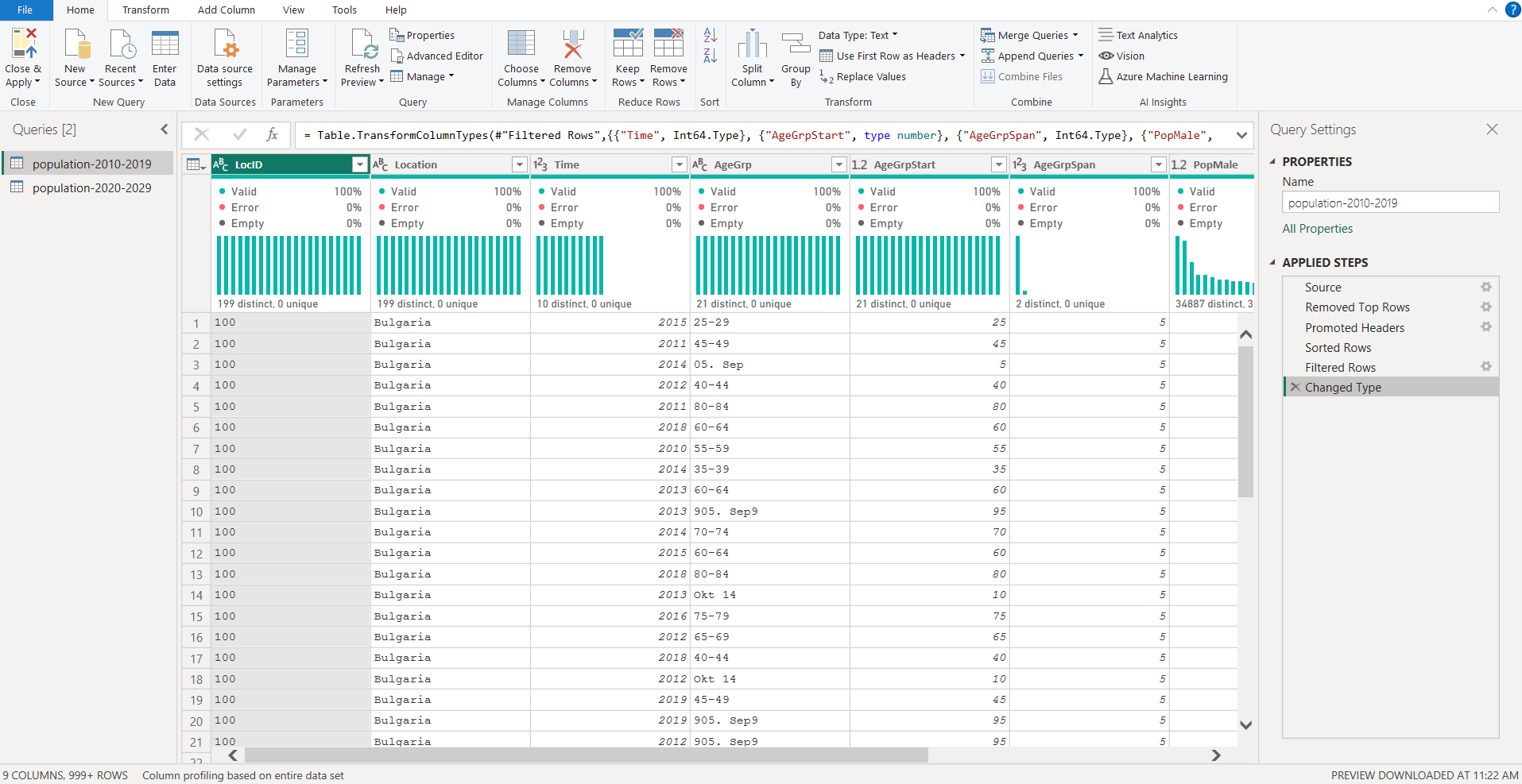
* **Action :** Used column profiling (top of each column) to check validity, detect errors or empty values.
* **Why :** Early detection of data issues saves time, improves trustworthiness.
* **What to look for :** High % valid values, consistent value ranges, logical data (e.g., age spans, population numbers).

**8. Documentation of Applied Steps :**

* **Action :** All actions are recorded in the “Applied Steps” panel.
* **Why :** Allows you to audit, replay, or adjust the pipeline as requirements evolve. Essential for transparency, collaboration, and reproducibility.

**Professional Presentation Tips :**

* **Data Flow Diagrams :** Show the step-by-step journey from raw to clean data.
* **Screenshots :** Capture each transformation phase for easy reference.
* **Annotations :** Highlight why each step was necessary—anticipate audience questions about data integrity.
* **Result :** Clean, analyzable dataset — ready for visualizations, modeling, or reporting.



This process not only readies your data for analysis but establishes a foundation for reliable reporting and deeper insights in any Power BI project.