**Power BI Project Summary**

**1. Sources Used**

1. **Web Source**
   * **HTML table from *Worldometers.info* containing country-wise GDP data.**
2. **Folder Source (Sales Datasets)**
   * **Multiple Excel files: *Sales\_Jan.xlsx*, *Sales\_Feb.xlsx*, *Sales\_Mar.xlsx*.**
   * **Loaded using Append Queries from Folder for automated monthly sales consolidation.**
3. **Excel Source (Employee Data)**
   * **Employee dataset containing EmployeeID, Name, Department, Region, Join Date, Birth Date.**

**2. Transformations Applied**

**Basic Data Cleaning (All Queries)**

* **Promoted first row to headers.**
* **Removed blank rows and columns.**
* **Renamed columns for clarity.**
* **Changed data types (using Change Type with Locale for currency & dates).**
* **Removed duplicates and filtered null values.**

**Text Transformations (Sales Dataset)**

* **Standardized text using UPPER(), LOWER(), TRIM(), CLEAN().**
* **Replaced text values where needed using REPLACE().**
* **Split columns using Delimiter (for customer names).**

**Numeric Transformations (Sales Dataset)**

* **Rounded off revenue to 2 decimal places.**
* **Created calculated column: Profit = Revenue – Cost.**

**Date & Time Transformations**

* **Extracted Day, Month, Year, Quarter from Order Date.**
* **Calculated Age from Birthdate in Employee Data.**

**Conditional & Index Columns (Sales Dataset)**

* **Added Sales Category column:**
  + **High (≥ 4000)**
  + **Medium (1,000–4,000)**
  + **Low (< 1,000)**
* **Added index columns (0-based and 1-based).**

**Pivoting & Unpivoting**

* **Pivoted sales data to combine monthly columns into a single column.**
* **Unpivoted data back into normalized format.**

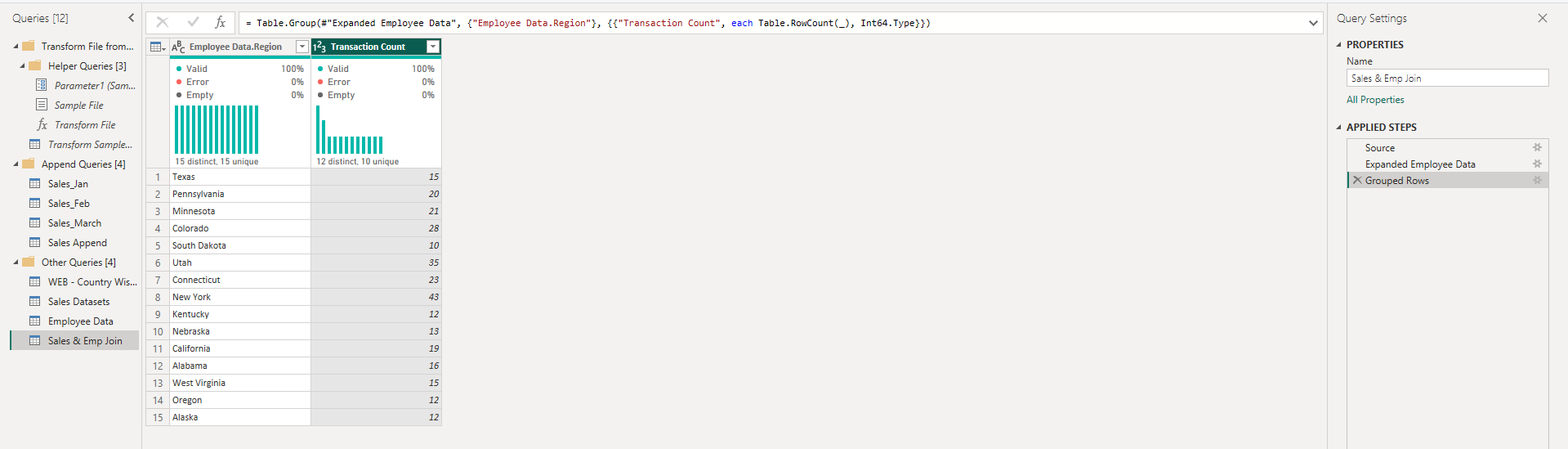
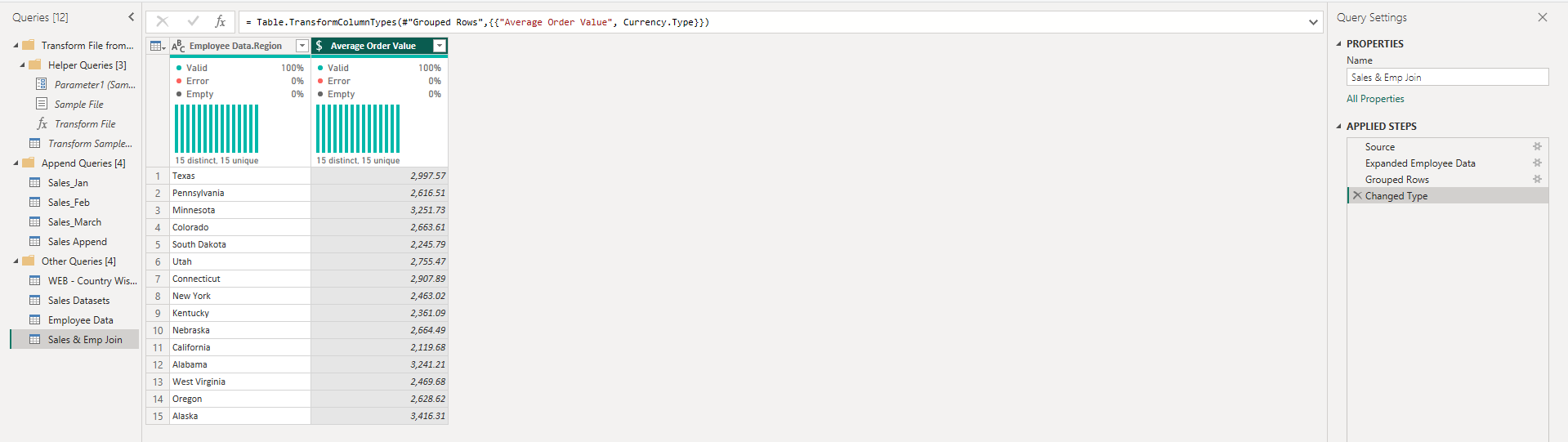
**Merging & Appending**

* **Merged Sales Data with Employee Data on EmployeeID.**
* **Appended Jan–Mar sales datasets into a single query.**

**Grouping & Aggregation**

* **Grouped by Region to calculate:**
  + **Total Sales**
  + **Average Order Value**
  + **Transaction Count**

**The 3 Screenshots of the groupby tables are attached below, please refer them for detailed info on groupby tables.**

****

**Data Profiling & Quality Checks**

* **Used Column Profile, Column Distribution, and Column Quality tools to:**
  + **Identify missing values**
  + **Detect column errors**
  + **View distinct & unique value counts**

**Source Settings & Parameters**

* **Created a dynamic folder path parameter for sales data.**
* **Modified data source credentials in Data Source Settings.**

**Refresh Simulation**

* **Added *Sales\_Apr.xlsx* to folder.**
* **Verified data auto-loaded and transformations applied without errors.**

**Challenges Faced and Solutions**

| **Challenge** | **Solution** |
| --- | --- |
| **Different file formats & sources (Web + Excel + Folder)** | **Used Power Query’s flexible connectors to standardize data loading.** |
| **Currency/date type mismatches** | **Applied Change Type with Locale and also changed computer’s regional settings to UK, so that the mm-dd-yyyy date format appeared consistently.** |
| **Inconsistent customer names & addresses** | **Used TRIM, CLEAN, UPPER/LOWER, REPLACE, SPLIT to standardize text fields.** |
| **Monthly file addition without manual updates** | **Implemented Folder Query with Parameters so new files load automatically.** |
| **Missing/null values in key columns** | **Applied Remove Duplicates & Filter Null Values steps, validated via Column Quality tools.** |
| **Combining datasets with different structures** | **Used Append Queries for monthly files and Merge Queries for Sales & Employee data.** |
| **Folder import failed due to different sheet names in Excel files** | **Renamed all sheets in sales files to have the same sheet name, allowing folder importing and combining to work without errors.** |