Power Query Processing Summary – Financials_Table

This document summarizes the data■cleaning and type■handling logic applied in **Power Query** for the *Financial Sample* workbook. All steps are written to be reproducible and robust.

Objectives

- Import the true table from the Excel workbook and avoid accidental sheet imports.
- Promote headers only when necessary.
- Normalize column names to the canonical schema.
- Remove the large block of blank rows that caused an “empty” fact table.
- Fix the UnitsID column that turned into zeros.
- Cast types safely using Vietnamese culture (vi-VN) so commas are read as decimals.

1) Import the correct source

Open the workbook with Excel.Workbook(...) and fetch the real table named **Financials_Table** (fallback to **Financials** if needed). This prevents the column drift seen when pulling directly from a sheet.

2) Promote headers only when needed

Detect whether the first row still contains generic headers (e.g., Column1/Column2). Only then call Table.PromoteHeaders to avoid double■promotion.

3) Normalize column names

Trim and remove NBSP (char 160) from all headers so they match the canonical schema below:

```
SegmentID, CountryID, ProductID, DiscountID, UnitsID, Manufacturing Price, Sale Price,
Gross Sales, Discounts, Sales, COGS, Profit, Date, DateID, Month Number, Month Name, Year
```

4) Remove blank/noise rows (root cause of “empty” table)

Keep only records where **Date** and **SegmentID** are not null. This removes ~1,048,575 blank rows (those with only DateID = 011900), which previously made the table look empty in the model.

5) Fix UnitsID becoming 0

Extract the numeric digits from UnitsID and cast to Int64. Remove/avoid any prior steps that force replacements to 0. Result: UnitsID values now reflect the original counts instead of all zeros.

6) Culture■safe type casting (vi-VN)

Apply types using the Vietnamese culture so comma decimals are parsed correctly. Numeric columns → number; Date → date; DateID, Month Number, Year → Int64.

Results

- ~700 valid data rows remain after cleaning; numbers and dates parse correctly.
- UnitsID is fixed (no longer all zeros).
- No phantom rows with DateID 011900.

Applied Steps (as seen in Query Settings)

```
Source → Navigation → NeedPromote → Promoted → CleanNames → Filtered (Keep Rows) → TypePairs
```

Target star schema

Fact: Financials_Table

Dimensions: Segment_Table, Country_Table, Product_Table, Discount_Table, Units_Table, Date_Table (each cleansed to distinct keys, proper types; one■to■many single■direction relationships from dimensions to fact).

Suggested DAX measures (starting set)

```
Total Sales    = SUM(Financials_Table[Sales])
Total Profit   = SUM(Financials_Table[Profit])
Discount %     = DIVIDE(SUM(Financials_Table[Discounts]), SUM(Financials_Table[Gross Sales]))
```

Note: If you also need the M code skeleton, I can export a separate appendix with the parameterized query.