Date:11-11-25

Power Query - Step-by-step Analysis of Transformations

This document describes, step-by-step, the transformations applied in the Power Query shown in the provided image. It explains what each applied step does, what changes were likely made to the dataset, and provides short recommendations for next steps.

Dataset Overview

From the preview grid we can observe these columns: Project Name, Sectors, Name, Job Role, Total Project Revenue. There are about 52 rows (preview shows 52 rows) and the Total Project Revenue column contains numeric values. Most other columns are textual. The Query Settings pane shows a sequence of Applied Steps which reveal the cleaning and transformation history.

Applied Steps — Explanation (in order shown)

1. Source

The data source was connected (for example Excel, CSV, database or other). This step brings the raw table into Power Query.

2. Navigation

A specific worksheet, table or node inside the source was selected to load into the query.

3. Promoted Headers

The first row of the data was promoted to column headers. Raw column names (if present in row1) became the table headers.

4. Changed Type

Power Query automatically detected and set data types for the columns (e.g., text, number, date). This ensures numeric columns (like revenue) are treated as numbers.

5.   Split Column by Delimiter

A column was split by a delimiter (for example splitting a combined column such as 'Name - Role' or splitting a compound field into parts).

6.   Changed Type1

After the split, data types for the new columns were adjusted (for example converting split pieces to text or numbers).

7.   Split Column by Delimiter1

A second split operation was applied to another column, again breaking one column into multiple columns using a delimiter.

8.   Changed Type2

Data types were changed again to reflect the new columns created by the second split.

9.   Renamed Columns

One or more columns were renamed to clearer or standardized names (for example renaming 'Column1' to 'Project Name').

10. Lowercased Text

Text values in one or more columns were converted to lowercase, helping standardize text for comparison and grouping.

11. Trimmed Text

Leading and trailing spaces were removed from text fields to avoid mismatches (important before deduplication or joins).

12. Replaced Value

Specific text replacements were applied (there are multiple 'Replaced Value' steps). These commonly fix misspellings, replace shorthand with full names, or standardize values (for example 'IT ' → 'IT').

13. Replaced Value1

A second replace operation — further standardization of values.

14. Replaced Value2

Additional replace operation(s) — possibly correcting names, sectors or roles.

15. Replaced Value3

More replacements for standardization.

16. Replaced Value4

Another replacement pass — keeps the dataset consistent.

17. Replaced Value5

Final replacement pass seen in the applied steps.

18. Lowercased Text1

A second lowercasing operation — possibly applied to another column or re-applied after replacements.

19. Filtered Rows

Rows were filtered to keep only relevant records. The applied filter might remove blanks, remove header rows that reappeared, or filter by a condition (for example only include rows where Total Project Revenue > 0 or where Project Name is not null).

Additional Observations from the Preview
- Column order appears to be: Project Name, Sectors, Name, Job Role, Total Project Revenue.
- Total Project Revenue cells show values like 136400, 120800, 299100 — these appear numeric and probably formatted as numbers.
- Names and roles are in lowercase in the preview (consistent with the Lowercased steps).
- Several project names repeat (e.g., 'New Satellite launch', 'Invest India', 'Expand Reach'), indicating a transactional dataset where each row is a person-role for a project.
- The Query Settings show many incremental cleaning steps — typical for iterative cleaning in Power Query.

Likely Purpose of Transformations
- Standardize text values (case, spacing, spelling) so grouping and joins are reliable.
- Split combined fields into atomic columns to allow analysis by each component.
- Ensure numeric columns have correct data type for aggregations (sum, avg).
- Filter to relevant rows to remove noise before downstream analysis or export.

Recommended Next Steps
1. Validate data types: confirm 'Total Project Revenue' is numeric and Dates (if any) are Date type.
2. Remove duplicates if required: use 'Remove Duplicates' on combination of relevant columns.
3. Add a column quality check: count nulls per column to identify missing data.
4. Add descriptive column names if any are still generic.
5. Add documentation: create a small Notes column or a separate documentation sheet explaining replacements performed (helps reproducibility).
6. If you plan to pivot or aggregate, ensure 'Project Name' and 'Sectors' values are fully standardized.