⇒ Power BI → Extract data from multiple sources

↳ Modelling (summary based) + ERDs

↳ Shareable | Interactive dashboards.

↳ For statistic based KPIs.

↳ Think of Python Pandas API with filter, group, cols and func as params.

→ Report ; data, Relationships.

↳ Extraction of data from multiple CSV raw files

↳ In premise server.

Power BI → support multiple servers | source data

↳ After securing connection → Query Editor.

→ Transform data → New Pane.

⇒ POWER QUERY

Three major tabs

HOME

Remove duplicates, Header — Choose columns, Remove cols | Row

Reshaping tools ; Text specific tools ← Transform

Add columns

1. Get data

↳ Browse for raw files.

↳ Preview Panel.

↳ If we click transform-data

Name →

Applied steps → source. ↗ define path

promote headers.

Changed datatype.

⇒ ———→ Remove product-sz col.

↳ Apply and Close (Top ~~right~~ left में है)

——————o——————x——————

Text-Specific tools.

Split

Format → case, strip.

Extract → length, ~~fir~~ slicing, split('delimiter') then slice.

Transform → Apply on the same col selected

Add col → New col for the transform.

↳ Customer_data lookup.

↳ Transform → by capitalizing first word of Name-base cols

↳ Add col → Merge Initial, last + first

'' .join (axis=1)' ↳ separator space

extract elements before @ → split("@")[0]

extract element after @ before . → Chain two splits

re.find(r"@(.*)\.", sent)

Transform replace.

⇒ Number Column transformation → Sum, Mean, ~

STATISTIC

↳ Returns a statistic value +Value_counts + unique

↳ We ~~never~~ usually use it in query editor

Standard + Scientific

↳ Numeric transformation - logs, trignometric transformation

↳ Edit Product data

↳ Already uploaded

↳ 0.9 is the discount ratio

↳ Product price × 0.9 → Discount price

Date & Time Columns.

.dt - -

We → Will create date_time columns.

↳ Calendar CSV

↳ Query Editor

Start of the week = Sunday

↳ Can be changed

↳ Has a parameter for this

pd.date_range(start, end, period, freq='D')
↳ Nums of obs

↳ resample(smalle freq).ffill()
↳ पुराना पता है
at resampled time.

⇒ Index columns
↳ Rownums

⇒ Adding conditional columns
↳ If/Else → Apply on a series/col.
Case/when

Transaction table
↳ Sales.
↳ Condition Column.
↳ If, Else

Add index column
From add column
↳ Index column