POWER BI – TRANSFORMATION & MODELING

Ivy Professional School

Top Ranked **Data Science Education Provider since 2007**

OUTLINE

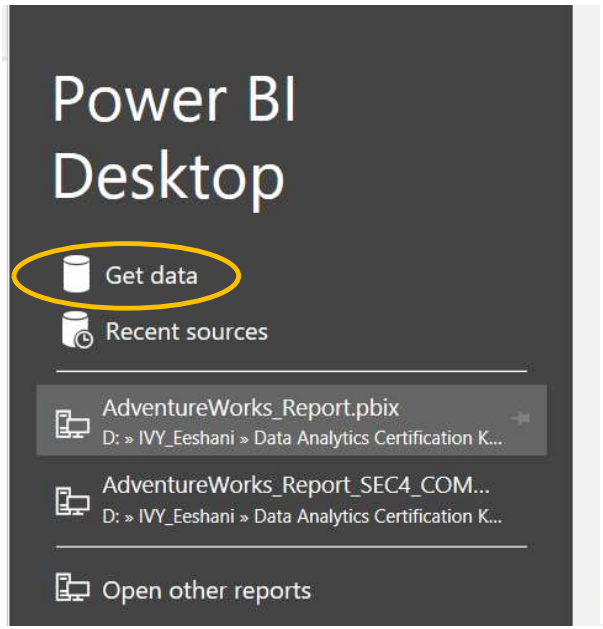
OUTLINE

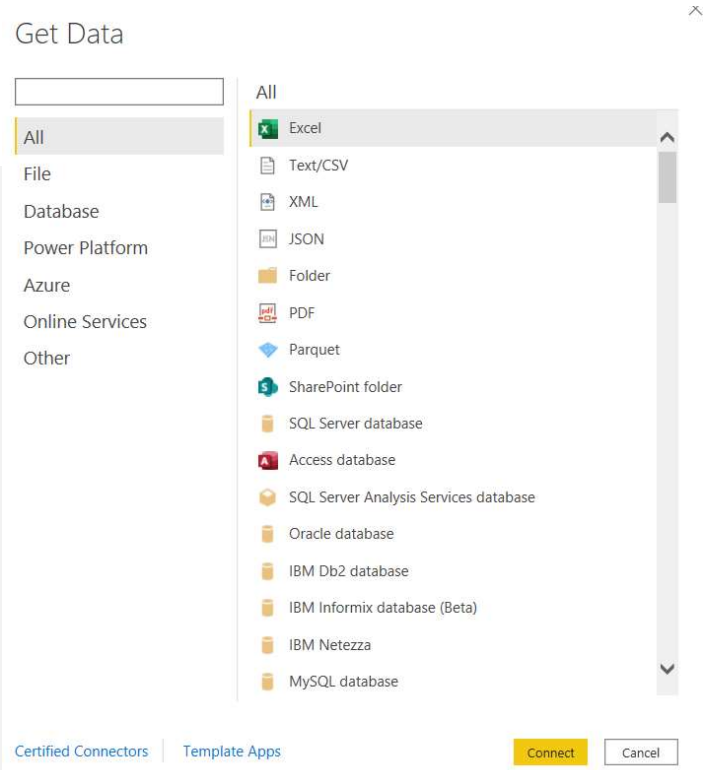
- Connecting with Data
- Data Query Editor
- Data Transformation
- Data Modelling

CONNECTING WITH DATA

Connecting with Data

- Power BI can connect with variety of Data
- Open Power BI
- Starting Screen
- Get Data
- Connect to data from file also using Get Data button in Home Tab

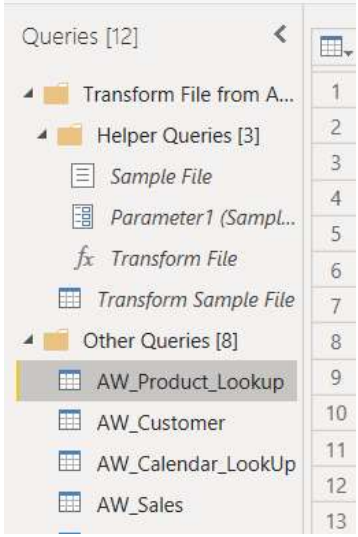




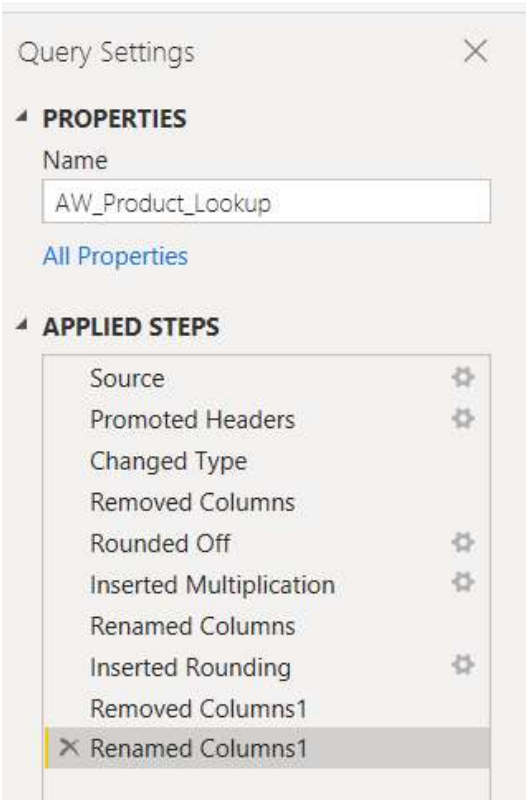
DATA QUERY EDITOR

Query Editor

- It is used for transformation of data
- Three main parts
  - Left – all queries
  - Middle – selected table
  - Right – Properties of the table and Applied Steps

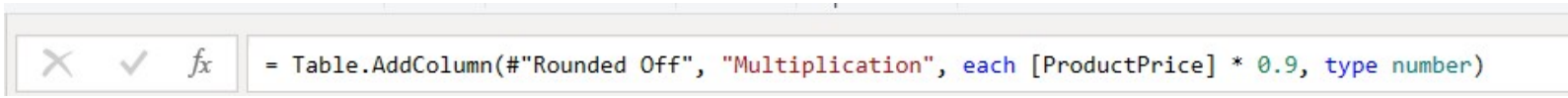


	ProductKey	ProductSubcategoryKey	ProductSKU
1	214	31	HL-U509-R
2	215	31	HL-U509
3	218	23	SO-B909-M
4	219	23	SO-B909-L
5	220	31	HL-U509-B
6	223	19	CA-1098
7	226	21	LJ-0192-S
8	229	21	LJ-0192-M
9	232	21	LJ-0192-L



Applied Steps

- Like Macros of Excel
- Automates the process
- Runs through the same set of steps again when data is refreshed
- We can delete steps and change the sequence
- Mcode shows up in Formula bar is used
- It is Power BI scripting language



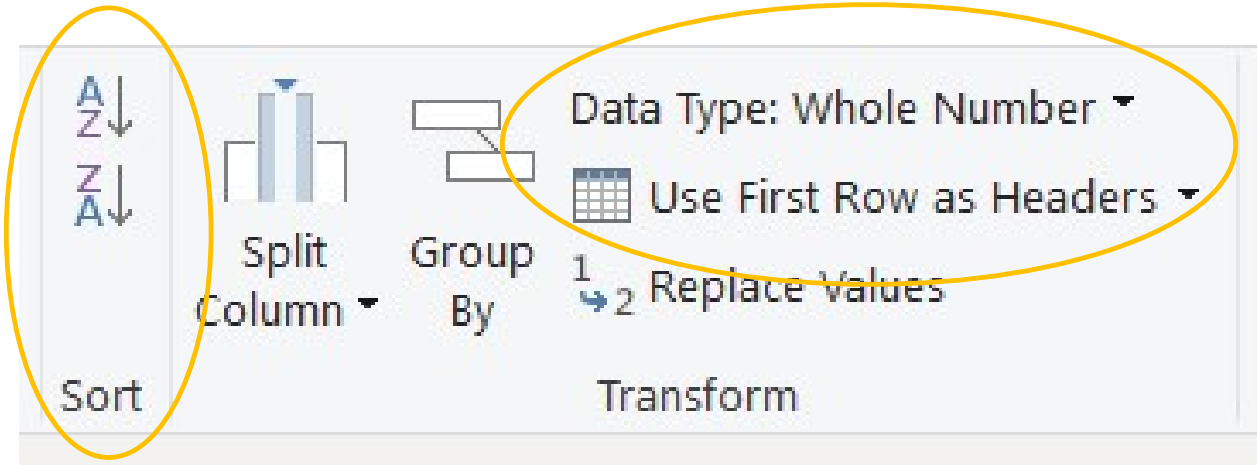
DATA TRANSFORMATION

Column and Row

- Remove unwanted columns
  - Columns not required for analysis
  - Helps to reduce the processing speed
- Remove unwanted row
  - Footers
  - Notes
  - Headings

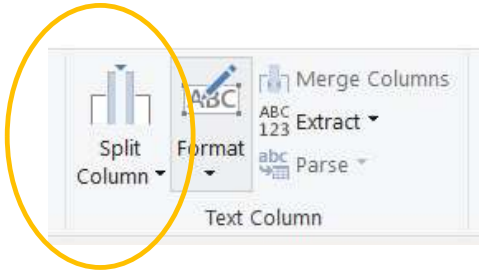
Basic Table Transformation

- Check data types of each field
  - To change data type –
    - Right click → Change Type → Select new data type
    - Or, Transform Tab → Data Type Tool on top
- Sorting columns
- Promote header rows



Text Tools #1

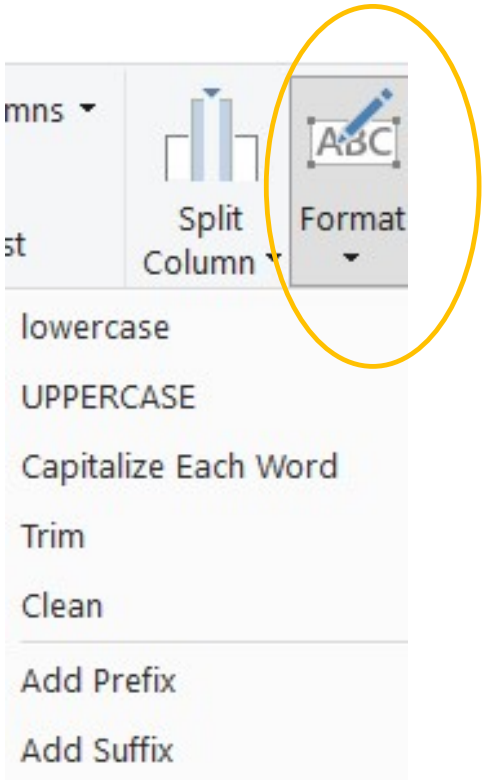
- Under Transformation Tab
- Split



i. Delimiter – based on any character
ii. Position – splits into multiple columns with specified number of char. Mention the position starting from 0 and in asc order like Eeshani → 0,5 →Eesha / ni
iii. Number of character
iv. Upper to Lower Case – Split from where Upper case changes to lower case. Splits into two columns
v. Lower to upper – similar to upper to lower
vi. Digit to Non-digit – splits from where the numbers end and alpha starts. First column will be number and second non digit
vii. Non-digit to Digit – opposite of Digit to Non-Digit

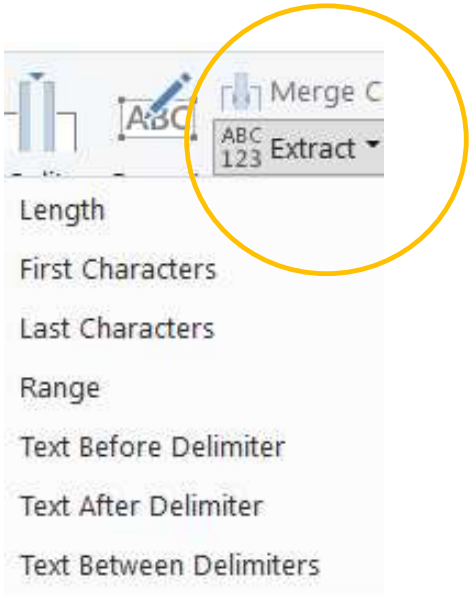
Text Tools #2

- Format
  - Changes cases – lower, upper, Capitalize Each word
  - Trim – remove leading and trailing spaces
  - Clean – removes non-printable characters
  - Add – Prefix and Suffix



Text Tools #3

- Extracts
  - Data extraction from each row of table
  - Len
  - First Character etc
- Merge
  - Creates new column by concatenation
  - Select columns to merge then Merge will activate



Exercise

Customer file

a. Convert the name into proper cases

b. Create a new column with full name

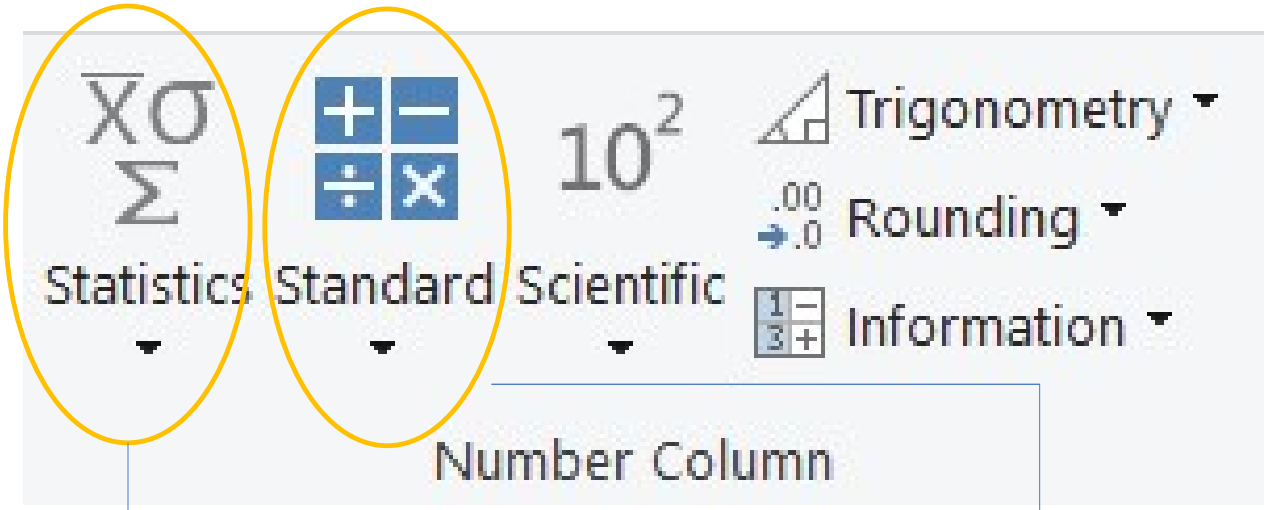
c. Create a new columns with user name and domain name from the email ID

d. In Domainname col replace '–' with space (use Replace Value Tool)

Exercise

- In Order Table – Sales 2015
  - Convert Sales into Whole number
  - Extract company name from the Product Name (first word of the product name)

Number Tools



- Statistics is available in Transform Tab and Not Add Column
- It aggregated values in column and returns one value only
- Not applied on row level but on column level
- Used to get quick summary of data

- Standard is available in Transform Tab and Add Column
- Used to Add, Subtract, Divide etc on row level
- Used to get modify data or create new columns

Date Tools

- Date and time tools will be active when data type of selected column is date/time
- Used to extract
  - Age based on the date and today's date
  - Various components of dates like Year, Quarter, Month etc.
- Earliest and Latest will return one single value
  - It is Transformation Tab only
- Start of the week can be defined using the Mcode

= Table.AddColumn(#"Inserted Day Name", "Start of Week", each Date.StartOfWeek([Date], **1**), type date). Add ,1 since 1 represents Monday. Or Table.AddColumn(#"Inserted Day Name", "Start of Week", each Date.StartOfWeek([Date], **Day.Monday**), type date). Add

Exercise

- Calculate the Age of the Customer

Index and Conditional Column

1. Index Column –

a. It is used to create sequential number column. Used for creating the Unique key for the tables which are used in the relationships.

2. Conditional Columns –

a. Helps to create new columns based on logical rules and conditions (If/Then Statement)

b. Goes through each row