

Examples:

- (1) Employee Data set:- Load & Clean
- (2) web scrapping
- (3) consolidating folder of files
- (4) Merge & Append Tables

What is Power Query?

1) # 

(Extract Transform Load)

→ process normally in data analysis and reporting lines of work where in order for you to build any kind of reporting.

→ 1st: Extract the data & store it

2nd: Transform: change nature of data
(edit them)

3rd: Load (store it from where data analysis can happen)

2) Data Clean

3) Connect (any kind of file:- Sql Excel)

4) get the data for analysis

• (4.) Consolidator → Merge
→ Append

• (5.) Data Generator → calendar Tables
→ list for Automation

• (6.) works in Excel & Power B.I

User Interface of Power-Query:

(U.I)

(1) ~~Ribbon~~ → Access various features of P.Q

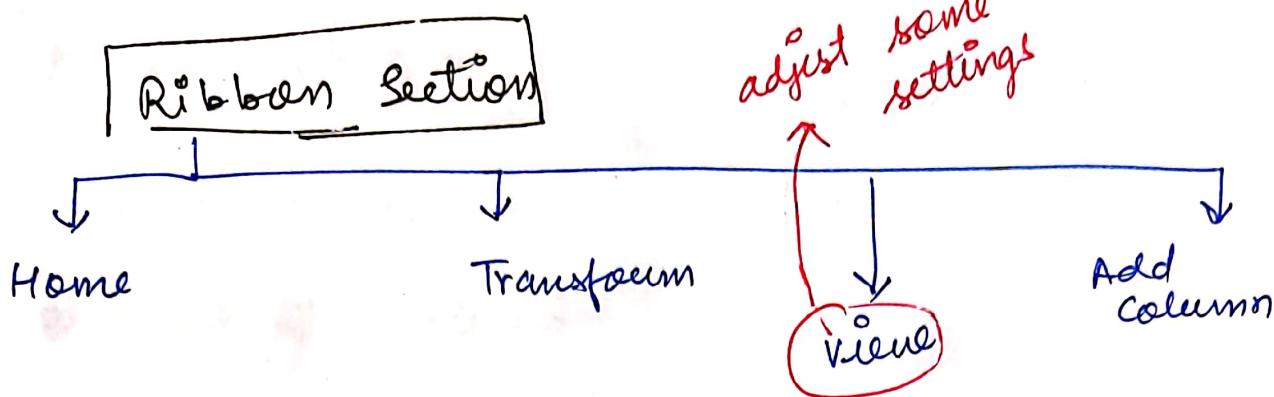
(2) ~~Preview/Data view Grid~~
→ Data being shown

(3) Queries span
→ if you've multiple queries or multiple connections within power query they will all be listed; you can select one at a time to view & work on that

(4) Steps/Query settings
→ Right corner: we can see editing done currently or done on that Query

group Holders

hidden



Home :

→ where common operations & frequently used items are usually bunched together and lot of things about query management

Transform :

→ Any operation on the data where things are happening within place so that means you don't want to add anything you just want to change the nature of data itself.

Add Column :

→ Add new column to our data
→ Specially date related things

Preview Section

- Data Grid → column profile
- Preview of selection

Things to know About Power Query

- Power Query uses its own language
 ⇒ "M"
- ↳ we can use 'm' language to develop advanced things within P.Q.
- we don't need to learn 'm' - language
- P.Q. is case sensitive
- Step - Recorder
 - (when we do any operation ; P.Q. records those steps)
- Data Process Automation Tool
- Each step in P.Q. depends on previous steps

for ex:

1

2

(3) → error if comes

then

it will not execute

4

step 4 & 5

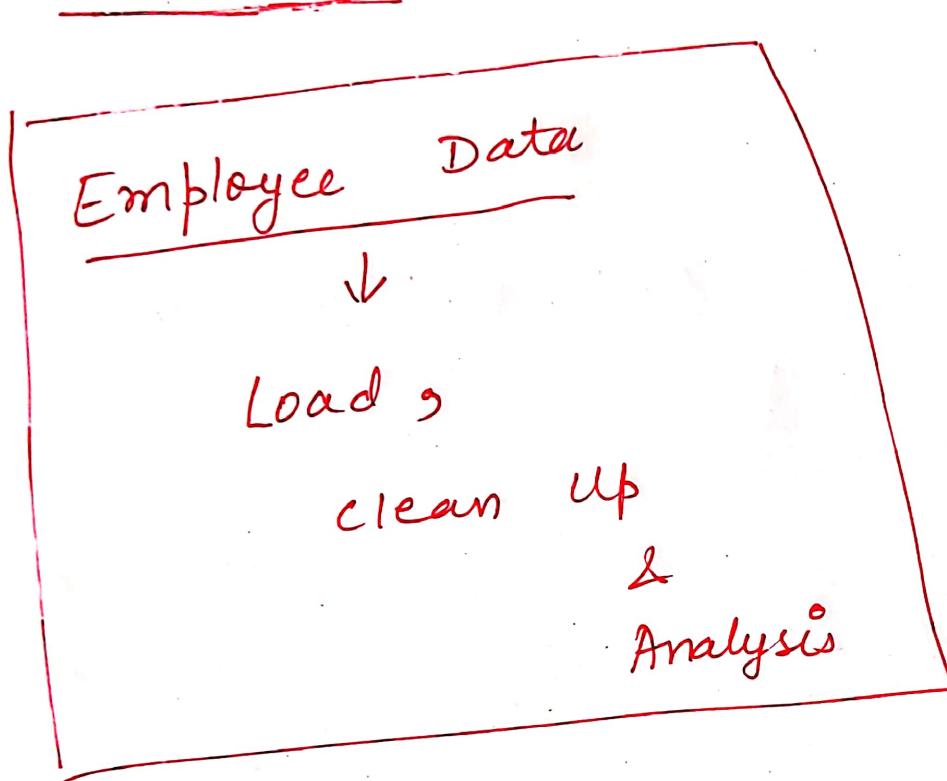
5

as they depend on step 3.

Data - Sets that can be connected
to Power Query

- (1) Excel
- (2) CSV
- (3) web connection
- (4) XML
- (5) API
- (6) SQL server
- (7) Data Bases (oracle)
- (8) Cloud sources
- (9) Folders
- (10) PDF

EXAMPLE - I :



Using data :-

Practical - 6 Data

↓
"mol" (Employee Data)

↓
Excel file Name

open your excel file: "M01"



(Name gender Dept. salary location start Date)

Column (6 columns of Employee Data)



Check the data carefully



This data is not adequate for
doing Analysis

Note: (in the data table)

- Some salary information missing
- Some Dept. are coded as "NULL"
- Some gender values are missing
- Inconsistent pattern while writing the address

Opening Your File/Data in PoQ.

Now, open a blank Excel workbook file



Go to → Data : Ribbon



Get & Transform Data :



Get Data



click here



From file



From Xcel Workbook



Select our file



Open

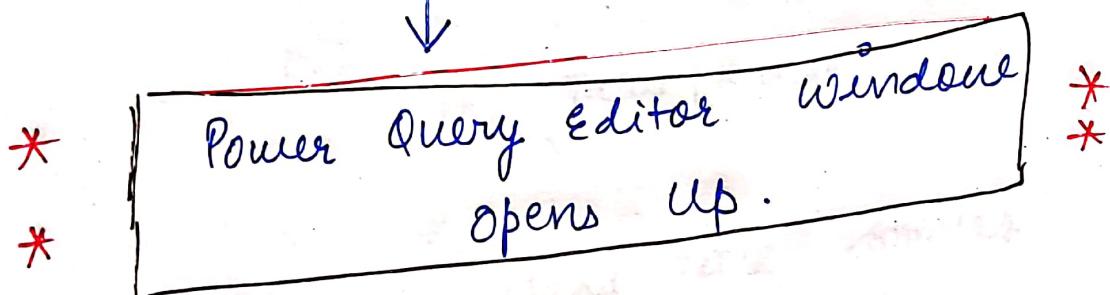


A navigator screen dialogue
Box opens up

↓
Here, we can preview our data
and select particular worksheet
or table

↓
select it

↓
click : Transform
[change / format our data]



Note :-

- * when our data is present
in the same Excel file



Data : Ribbon



Get & Transform Data :



(from Table | Range)



Other steps are similar
as prev. wala

Now,

Power Query Editor opens up

(Resize it to full screen)



We'll see :⇒

- file had data from rev no. = 3 & i-2 rev has some additional information

User Interface - Power Query

⇒ window called : "P.Q. Editor"

⇒ Access this from both

: Excel & Power BI

by using Queries button

⇒ Top : Ribbons

⇒ 3 main areas

1. Left

Queries Area

(expand & collapsed it)

Data set

Preview Window

Right

2. Middle

3. Top

query setting
Area

Preview Window (Middle)

it will show preview of the data

it will show top 1000 rows
of the data here

shown by rows & columns

we can't edit in the box/cell
(i.e., can't update them individually)

Query Settings Area (Right)

→ it has 2 imp. BOX

Properties

(Name of the query)

Applied steps

(No. of steps used / applied)

→ All operations performed on the data
are listed in this area

Applied Steps

1. source
2. navigation
3. Promoted Headers
4. changed Type

⇒ It has already performed these 4 steps on the data.

⇒ It went to source (Excel file selected)

↓
then, navigated to data table

↓
Automatic put row 1 + 2 as headers

↓
changed data type for each column.

Note:-

→ If not happy/want any step, just select and delete it (X)

⇒ P.Q. is step-by-step language
so, every subsequent steps by default depends on prev. step.

↓
So, good way to delete the steps is from bottom.

Employee :- Data Clean Up

Keep only these steps

- 1) Source
- 2) Navigation

1) Delete 1st two Rows

Home : Ribbon



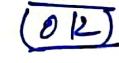
Remove rows  click here



Remove top rows



A dialogue box opens up

No. of Rows = 2  

Note:- your original excel file is not changing but when it is brought over into power query

⇒ Power Query makes a note that bring

the table, being the data worksheet
and delete the first two rows
and use the remaining data to
process.

→ we don't have to redo these steps
P.Q will apply these steps automatically

(2) Use 1st Row As Headers

Home : Ribbon



Use 1st Row as Header

(3) if Extra Column Appears, just remove them

select those extra column



right click



Remove Columns

Applied Steps :-

1. Source ✓
2. Navigation ✓
3. Remove Top Rows ✓
4. Promoted Headers ✓
5. Changed Type → Remove it as it is
6. Removed column ✓ automatically applied by P.Q.

Now,

Our Task to do in the
Power Query Editor :

1. Load the Data
2. Power Query window & buttons
3. Blanks (Gender + Salary column)
4. NULLS (Dept. column)
5. Missing Salary - filter
6. Extract Country
7. Publish
8. Pivot

9. > Update the file

10. > Refresh

(4.) Rename NULL in Gender column
as Others

Select Gender column

↓
right click

↓
Replace values

↓
A dialogue Box opens up

value To find : null

replace with : Others

↓
OK

*+ null ≠ NULL (In PQ)

* By each step, you can see its applied formula in the formula bar at the top.

If formula bar is not visible

↓
View : Ribbon

↓
 enable formula Bar

(5) Rename NULL in Dept. column
as Engineering

Select Department Column

↓
Right click

↓
Replace values

value to find :> NULL

Replace :> Engineering

↓
OK

* it means initial employee had been joined but later on they moved to engineering section.

(6) Remove person whose salary is missing

↓
It means they are now not a part of company

Go to salary  click



Uncheck (null)

↓
OK

(7) Extract last word of location column and then remove location column

Method - I

Home : Add Column



Extract



Last character



No. of last character = 3 

Method - II

select location Column



Add Column : Ribbon



Extract



Text After Delimiter

Method - III

select location Column



Add Column : Ribbon



Column from Examples

▼ click here



From selection



A new window opens up



Type : In

Column 1



whatever we type here

it will define the logic so far

↓
fill 3-4 boxes

↓
it will automatically fill up
rest

↓
click : ok

↓
Now , it will be added in
P.C table

↓
this table will be named = "custom"

↓
Delete click on it : rename it
↓
Country

Name,

remove location ~~place~~ column

↓
select location column

↓
Right click : delete / remove



close & load



Now, it will come up and sit
as a nice table within excel
with all the data.

Data Analysis & Refresh

Once the data is in Excel



~~Close & Reset~~



Table Design



summarize with Pivot Table

* Put : Dept → Rows
Name & Salary → Values

* Apply formatting to Numbers

Gender as Slicer

* Add

Now,
open over original Excel file

↓
go to end of the table

↓
paste some extra data

↓
(ctrl + S) this excel file

Now,
come to P.Q wall excel sheet

↓
open Pivot sheet

↓
Data : Ribbon

↓
click: Refresh All

↓
Now, automatically, everything
will get updated

Editing the Query &

Adding one More Step

Go to Data ribbon

Click: Queries & Connections

It will open Query span
And show you data added

Right click on data: on Query Span

edit

You'll be back to P.C editor

now

Select Start Date column

right click (heading)

Change Type

Date

Now, this is converted to Date data type

Select start date column



Add column : Ribbons



Date : ↑
Year



Close & Load

Now, Year column will be added to
P.O table

Again,
Go to Pivot sheet.



Data : Ribbons



Refresh

Now,
Add : Year to columns in
pivot field

Example - 2

Web Data Scrapping with Power Query

Using Data

Practice - 6 Data

↓
"PQT-sample"

↓
web data (sheet 1)

we will be looking at how we can connect to a website and fetch the data from the websites into Excel now from a business data analysis point of view connecting to website is not something that you would do often but from an organizational data analysis point of view many times you may have to connect to data sources that are like websites.

Data

open that excel sheet



go web data → click on that link



Indian state population from wikipedia

Transform this website table

↓
Data years

↓
Unpivot this data

↓
and we want 3 columns

- [i) state
ii) year
iii) Population]

Select them

How to Load such Data into
Your Excel

select the website link

↓

Data: Rilon

↓

Create Data  click here

↓

From other Source

↓

From web

↓

A dialogue Box opens up

Paste URL
↓
OK

Now, navigator dialogue BOX
opens

↓
Select them which has table
format

↓
Transform Data
↓
It will bring your data to
Power Query Editor

* Remove steps
changed Type: from Applied
in P.Q. Editor

(1) Remove top Rows

Home : Ribbons



Remove Rows



Remove Top rows



No. of rows = 1

↓
OK

(2) Remove Bottom (Grand Total) Rows

(i) Home : Ribbons



Remove Rows



Remove Bottom rows



No. of rows = 1



OK

(ii) Go to Rank Column filter Y



Uncheck Total J
OK

(3.) Remove Rank Column

Select Rank Column



Right click



Remove

(4.) Transform on Pivot

→ Other than state/UT column we don't need them. We just want to

have format like pivot table
so unpivot these data

i.e., for every column we want to split into 2 :
→ population
→ year

1) This form of operation is called

"Transform on Pivot" bcz we are changing the nature of the data

→ we want to keep state/UT and unpivot all others.

Steps:

Select State/UT Column



Transform: Ribbon



Unpivot Columns ↑ click here



Unpivot other columns

To get only Year

Now,

Add column : ribbon



extract



Text B/W Delimiters



Start : (

End : ↓ space



OK

Note,

* Remove Population census column

* Rename "Year" column

* Rename "Population" column

Note, many states have N/A values

* ~~Convert~~ select Population column

↓
Right click

↓
change type

↓
whole NO.

* Right click on Pop. column

↓
Remove errors

Go to Query setting at Right

Name : India Population



close & Load



This will generate a table in
our excel sheet in this
pattern.

Example - 3

Consolidating a folder
of files with
Power Query

⇒ How to take a folder of all
files & consolidate them into
one big table using
Power - B.I

Data



practise 6 Data



csv folder

we can also use this technique
to combine data that is in
text files ; csv files ; even pdfs
through Power Query.



make a note/copy that folder path

(ctrl+c) : its path

Now,

Data : ribbon



Get Data ▾



From File



From Folder



just paste that path (ctrl+v)



OK.

Now, it will open the Navigator screen

Click : Combine ▾

Combine & Transform Data

Now, Combine files screen opens up

just: 4 files in the folder you want me
to combine but what do you
want in each file , what is
the process you want to run on
each file teach me for one and
then I'll do the same for all
of them

Sample file : (pick any file)
eg: 1st file

Sheet 1 select Bottom

OK

Now, P·Q will open the editor
and load up query here

Now you'll see a bunch of Queries
before csv query
(there are things that power
query needs in order to generate
this)

Understanding

Folder

Combine

Operation

final Query \Rightarrow CSV

Note:-

so, we taught it what we want to do which is select the sheet 1 in the 1st file and then it will do the same for all of that so that kind of process is usually done through functions:

we loaded : sample file

so, based on the sample file P.Q builds a small internal Query called = "Transform sample file" where operations we want to do is defined.

Once these are done, based on it; P.Q. will create a

"Transform file fx" "f"



is for defining the process
that we want to run on all the
files in the folder so once
this process is defined then P.Q.
will go to the folder for each
file in the folder; it will
simply run that process that
function & get the output &
put everything together here
in I go.

Output :

SourceName

Month

Document

New,

Formatting :→

(1.) Rename source column as
Project Name and remove
from its name
·xlsx

⇒ * Select source name column

↓
Right click

↓
Replace values

↓
To : ·xlsx

with : —
↓
space

↓
OK

* Double click on source name
and rename "Project Name"

Adding
Extra
Operations
to
Transform
&
functions
Sample file

we don't want Month No.
but instead want name of
the Month #

→ so, these things will happen
on all the files so any changes
to that need to happen at the
functions level but we don't
know to edit functions -

Then,
simply we'll make changes
to "Transform Sample file"

Now,

click :  Transform Sample File



Add column : Ribbon



Click : Custom Column

↓
Custom Column dialogue Box Appears

↓
New Column Name: Date

formula: ~~=~~ #date(2020, [Month], 1)

Available Column: select:- Month

↓
OK

↓
then date column will appear

Remove Month column

↓

Go to CSV

all the same will be updated
there (for all files)

Note :- if we don't want this function on all the files on the folder, let's say it has some additional file like - excel, ppt, word, pdf, text file but we want only excel file

↓
To do this → steps :



Applied steps



Click: Source
(listing all items in the folder)



** P.Q. ignores hidden files in the folder

↓
Go to Extension column



↑ click here



↓
Text filters

↓
Equals

↓
equals = .xlsx (extension of
excel file)

↓
Select .xlsx
OK

↓
Now, close & load it

↓
All the data will come in
your Excel sheet

↓
Note: Date is shown as NO.

↓
Select Date column

↓
Home: Ribbon of Excel

↓
General: Short Date

Note:-

Add a new file to this folder



just open your excel sheet



Select any cell



Right : Queries & Connections



Right click : CSV



Refresh



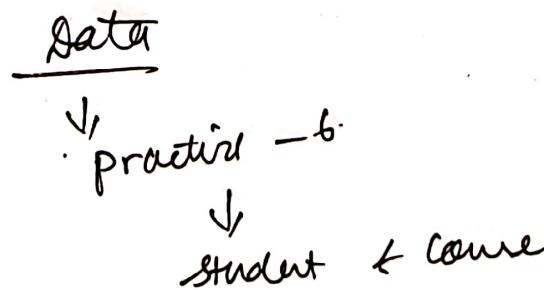
Now, you'll see new file
will be updated here.

EXAMPLE - 4

Merging & Appending Queries

Join 2 or more tables Using
Merge & Append Operations

(Take 2 sets of data: either join or
combine them) to get 1 full set
of data



We have:

Year 11 :

Year 12 :

Courses :

Task to do :

- ① Append both sets of student
(Year¹¹ + Year¹²)
- ② Combine course & student Data
- ③ Filter female students on 3 or
more credit courses.

Open Excel blank sheet



Data : Ribbon



Get Data ▾



From File



From Xcel Workbook



select your file



open (import)



navigator screen opens up



or select multiple items



Course

Year 11

Year 12



Transform Data



Power Query Editor opens up



We'll see in Query section

at left side



Year 11

Year 12

Courses

⇒ we'll also have our previous
queries listed if any

Appending

Year 11

&

Year 12

Data

open them in P.Q



Select any either Year 11 or
Year 12

from left (Queries)



Name: Ribbons



Append Queries

↑ click here



Append Queries as new

(appended wala separate rahaega)

or
just Append



② Two Tables

Tables to append: other than
current wala

↓
OK



Now, Both are appended together
and data of both will be shown
altogether

↓
But actually there is 1 problem.
we can't specify whether somebody
is in Year 11 or Year 12

* Undo these steps by deleting from

Applied steps
 $X \rightarrow$ Appended Query

Now,
click on Both by + one:
Year 11 & Year 12

(i) ↓
Home: Ribbon
↓
use 1st row as header

(ii) Go to year 11
↓
Add column: Ribbon
↓
Custom column
↓

new column name : year
custom formula : = "Year 11"
OK ↴

⇒ Repeat the same for Year 12

Now,

Append Year 11 & Year 12 tables
as prev. done.



Select Any



Name : Append Queries

Now,

Rename the appended table as "Student"

from Right side



Query settings



Name : students

Merging Student & Course Data

Here,

Join Students table &
Course Table

This \Downarrow operation = "Merging"

* click on students table

\downarrow
Home : Merge Queries

\downarrow
A dialogue Box opens up

merge

students Table \rightarrow just click here to select

selected : course \rightarrow table which u want to join

* select the column on each basis you
table on which are going these 2 table.

* we can also make them joins
on basis of 2 column

↓
select 2 columns in each
table



Join kind : → just like SQL
select any option



Note,
we are back in P.Q editor

↓
we'll see that for each
course it will add a new table
so it's not directly merging . It is
actually giving you the corresponding
row of the other table

note: → select red & then click on the
table here , you'll see the entire
information for the red course at the
bottom



Now, expand the table



Course



Course



disable course

disable original column name
as prefix



Right Click OK



Now, it will be added

It will create just a copy.

Students table.

The course field: female

(from query setting or view)

Query Referencing

Now, we'll filter female students on 3 or more credits courses

click on student table in
Queries

↓
right click

↓
④ Reference

↓
it will create just a copy of
Students table

↓
Rename this: Female Students
(from Query setting on the
right)

Gender
↑
click here



filter just female

♀ only female



OK



Go to credit 
↓ click here



No. filters



Greater than or equal = 3



OK

Customizing Load Behaviour

(i) Select female students in Queries
↓
Close & load

(ii) At the left side
↓
Queries section
↓
Right click on empty space
in Queries section
↓
New Group

↓
Name: Support Queries

↓
And in this folder
↓
move: Year 11, Year 12, Course &
students table

↓
move them by just dragging
them

↓
now, click on close & load to

↓
or only create connection

↓
OK

↓
At the right side of your Excel
file ; queries & connection opens

↓
right click : Female Students

↓

Load to

↓

or Table

or new worksheet

↓
OK



so, this will bring up all
the female students data table
here in new sheet.