

Example - 2

Web Data Scraapping

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 organization's 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~~ ~~was~~ ↓

Unpivot this data



and we want 3 columns

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

How to Load such Data into
Your Excel

select the website link



Data: Ribbon



Get 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 : Ribbon



Remove Rows



Remove Top rows



No. of rows = 1



OK

(2) Remove Bottom (Grand Total) Rows

(i) Home : Ribbon



Remove Rows



Remove Bottom rows



No. of rows = 1



OK

[iii]

Go to Rank column filter ☒



Uncheck Total ☐ OK

(3.) Remove Rank Column

select Rank Column



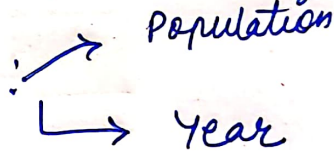
Right Click



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

Name,

Add Column: Ribbon



Extract ▼



Text B/W delimiters



Start: C

End: space



OK

Now,

- * Remove Population census Column
- * Rename "Year" Column
- * Rename "Population" Column

Now,

- * We'll see many states have N/A values

Change data type

- * ~~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.