20-01-2025

P1 - Exhaustive Analysis Of Indian Agriculture Sector Using Power BI

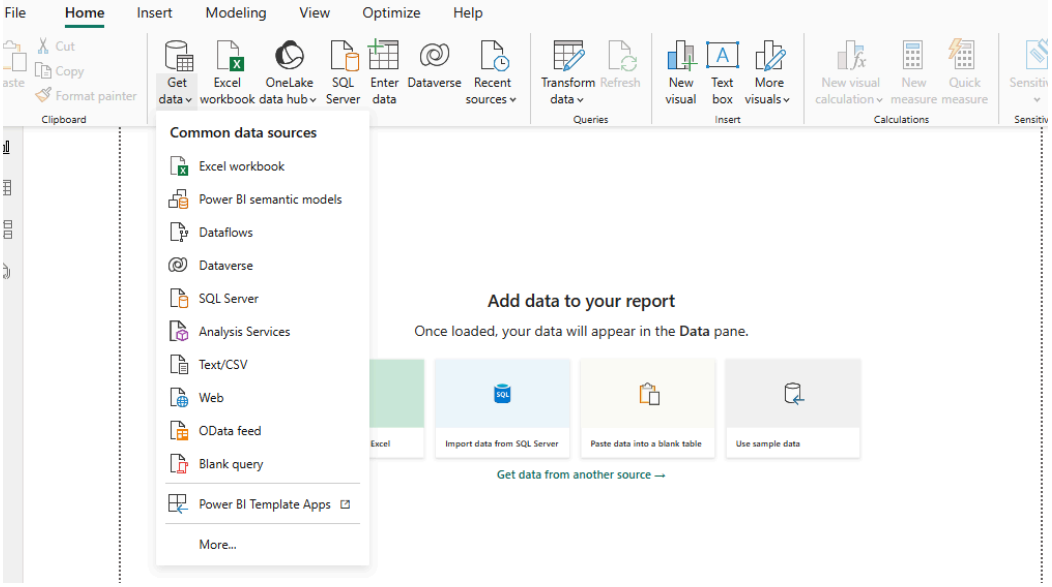
Session 1 - Importing and Preprocessing of data

1. Importing Data

Step 1: Open Power BI Desktop

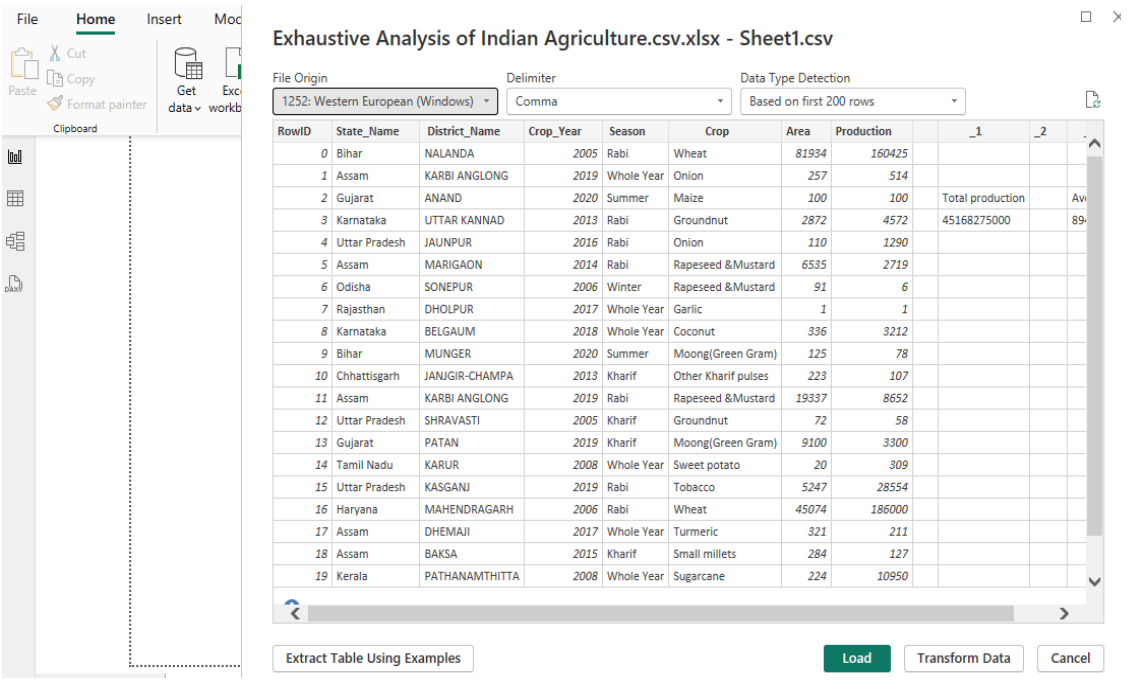
Step 2: In the Home ribbon at the top of the window, click on the "Get Data" button.

In the drop-down menu, select "Text/CSV".



Step 3: Locate Your CSV File

Step 4: Preview data

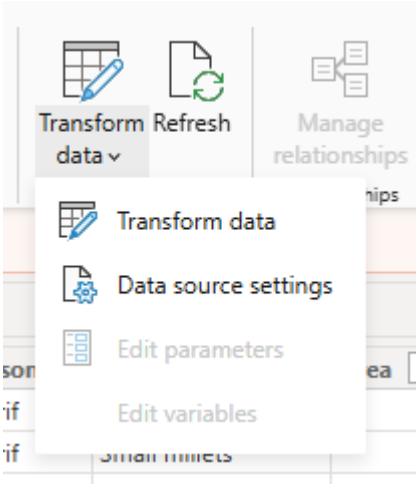


RowID	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production		_1	_2
0	Bihar	NALANDA	2005	Rabi	Wheat	81934	160425			
1	Assam	KARBI ANGLONG	2019	Whole Year	Onion	257	514			
2	Gujarat	ANAND	2020	Summer	Maize	100	100		Total production	
3	Karnataka	UTTAR KANNAD	2013	Rabi	Groundnut	2872	4572		45168275000	
4	Uttar Pradesh	JAUNPUR	2016	Rabi	Onion	110	1290			
5	Assam	MARIGAON	2014	Rabi	Rapeseed &Mustard	6535	2719			
6	Odisha	SONEPUR	2006	Winter	Rapeseed &Mustard	91	6			
7	Rajasthan	DHOLPUR	2017	Whole Year	Garlic	1	1			
8	Karnataka	BELGAUM	2018	Whole Year	Coconut	336	3212			
9	Bihar	MUNGER	2020	Summer	Moong(Green Gram)	125	78			
10	Chhattisgarh	JANJGIR-CHAMPA	2013	Kharif	Other Kharif pulses	223	107			
11	Assam	KARBI ANGLONG	2019	Rabi	Rapeseed &Mustard	19337	8652			
12	Uttar Pradesh	SHRAVASTI	2005	Kharif	Groundnut	72	58			
13	Gujarat	PATAN	2019	Kharif	Moong(Green Gram)	9100	3300			
14	Tamil Nadu	KARUR	2008	Whole Year	Sweet potato	20	309			
15	Uttar Pradesh	KASGANJ	2019	Rabi	Tobacco	5247	28554			
16	Haryana	MAHENDRAGARH	2006	Rabi	Wheat	45074	186000			
17	Assam	DHEMAJI	2017	Whole Year	Turmeric	321	211			
18	Assam	BAKSA	2015	Kharif	Small millets	284	127			
19	Kerala	PATHANAMTHITTA	2008	Whole Year	Sugarcane	224	10950			

Step 5: Load data

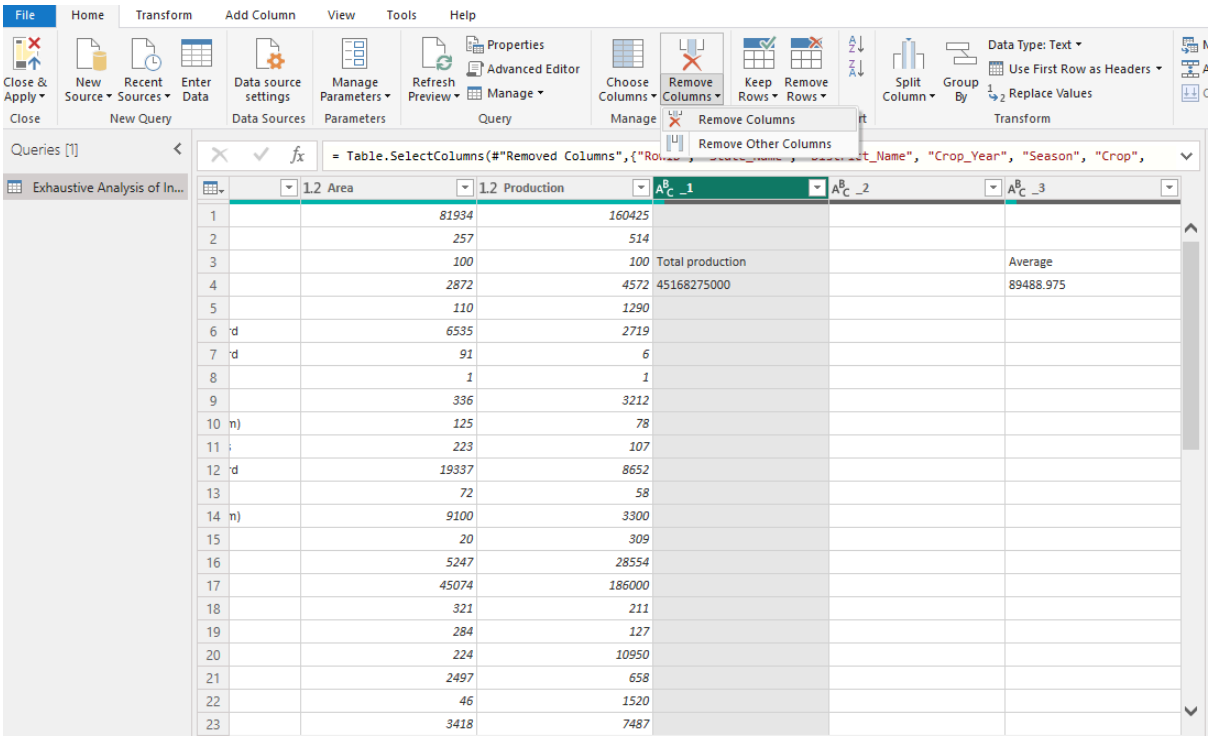
2. Data Preprocessing

Step 1: After importing your data, if you need to transform it, click Transform Data.

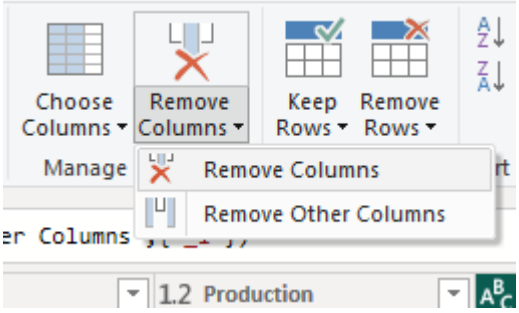


Step 2: Open Power Query Editor

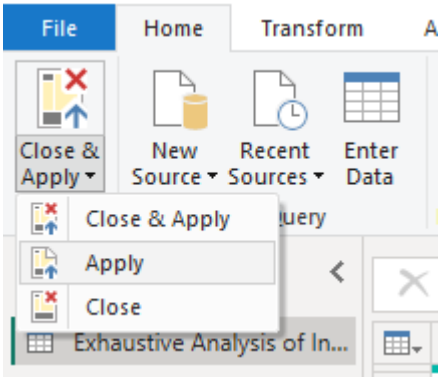
Step 3: Remove Unnecessary Columns.



Step 4: Right-click the column header and select Remove or select the column and click Remove Columns in the ribbon.



Step 5: Click Close & Apply to save and apply your changes to Power BI.

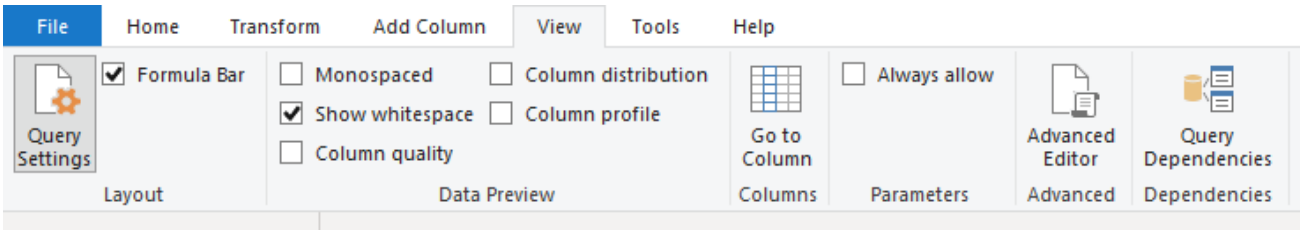


3. Assessing the Quality of our column

Step 1: Open Power Query Editor

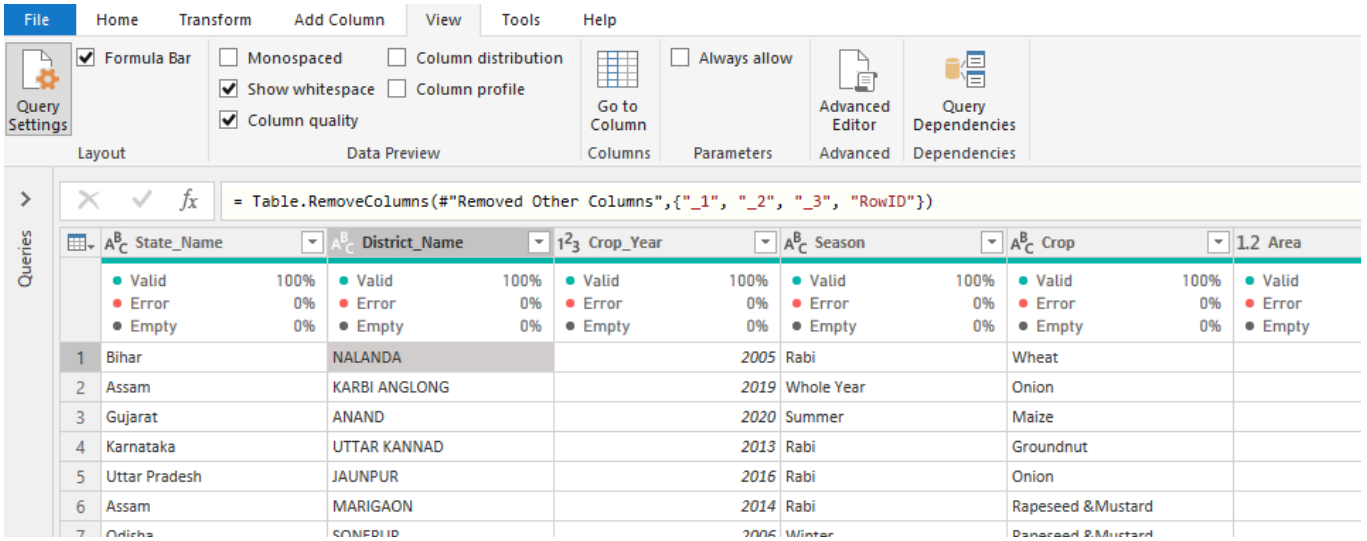
Step 2: Click on any column header to select the column you want to analyze.

Step 3: In the View tab of the ribbon (located at the top of the Power Query Editor), check the option called Column Quality.

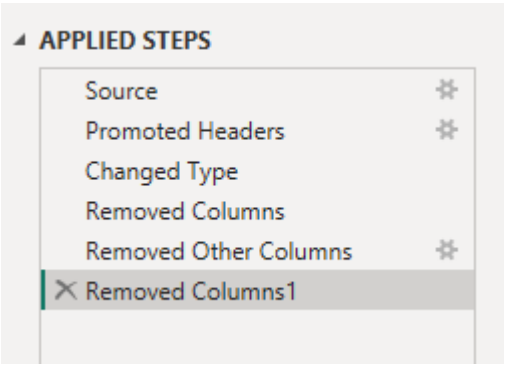


Step 4: Once enabled, you'll see a small bar chart above each column header, representing the distribution of valid, empty, and error values:

- **Green bars** represent the valid values in the column.
- **Yellow bars** represent empty (null) values.
- **Red bars** represent error values.



4. Notes

i) All the changes that we have made so far in the document will appear in the Applied steps bar.

ii) As we apply transformations like removing columns, filtering rows, changing data types, Power BI automatically records each transformation in the Applied Steps pane on the right side of the screen.