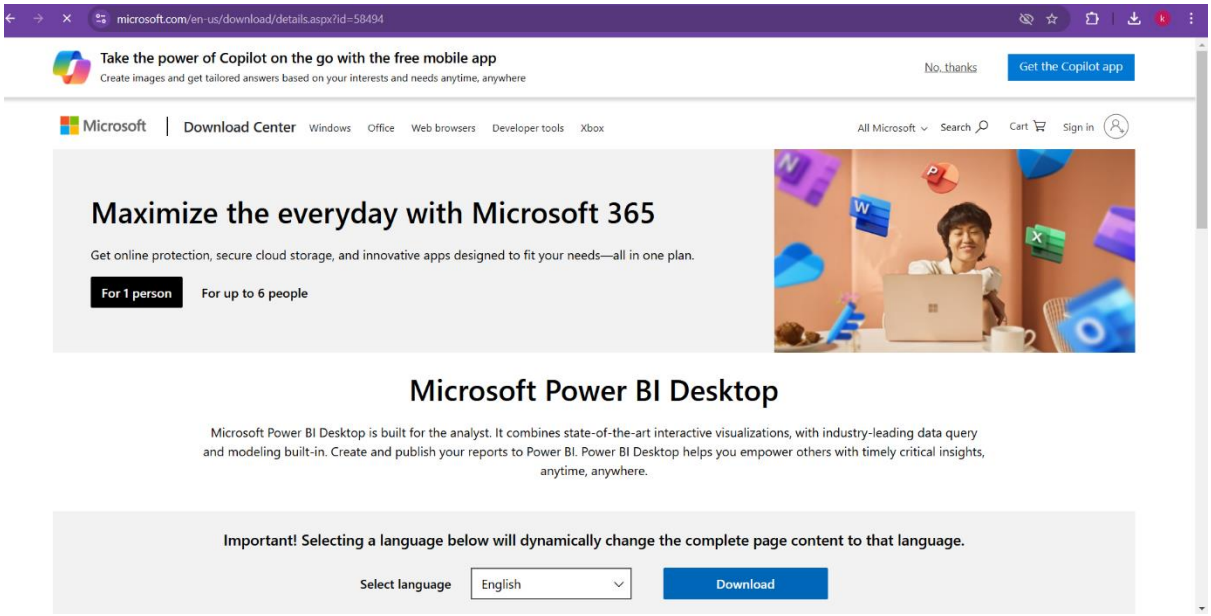
Week 1 Task

Project 1- Exhaustive Analysis of Indian Agriculture using Power BI

1. Installing Power BI and Exploring Its Interface

1. **Download Power BI Desktop:**
    - **Visit the official Power BI website to get the desktop version.**
    - **Complete the installation by following the on-screen instructions.**

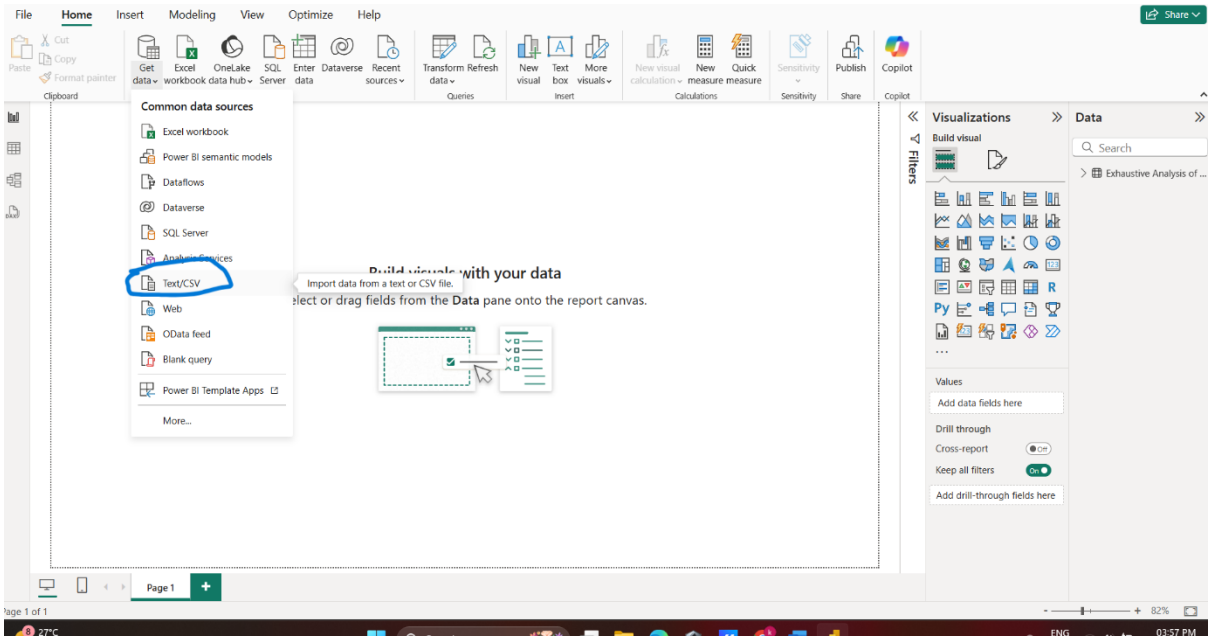
2. **Understanding the Power BI Interface:**
    - **Data Pane: Displays all the data tables you import (located on the right).**
    - **Visualizations Pane: Offers various chart and graph options for your reports (also on the right).**
    - **Main Workspace: The central area where you design and build your reports.**



Name : alok mehta

2. Importing a File into Power BI

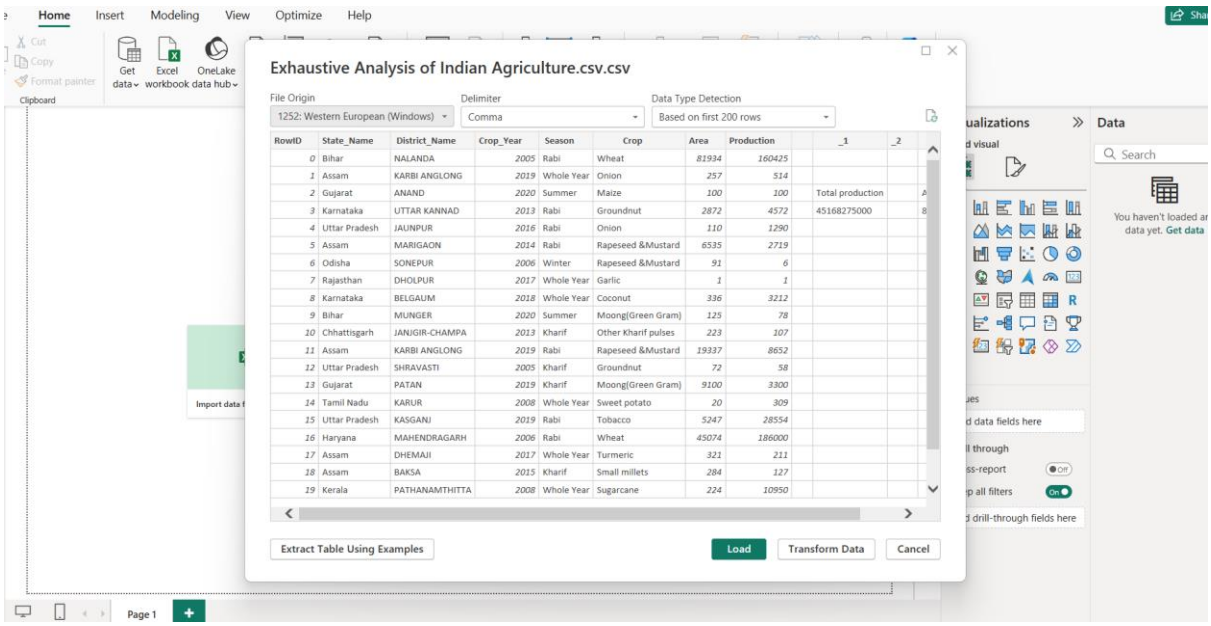
1. Launch Power BI Desktop on your device.
2. Select **"Get Data"** from the Home menu.
3. Pick the type of file you want to upload (e.g., CSV) and click **Connect**.
4. Locate your file on your system, select it, and hit **Open**.



Name : alok mehta

Step 5: File Dialog Box - Choose Your Action

- **Once you select the file, a dialog box will pop up showing you a preview of the data inside. There are three options in the dialog box:**
    - **Load**: Import the file directly without making changes.
    - **Transform**: Make adjustments or clean the data using the Power Query Editor.
    - **Cancel**: Exit the process without importing the file.



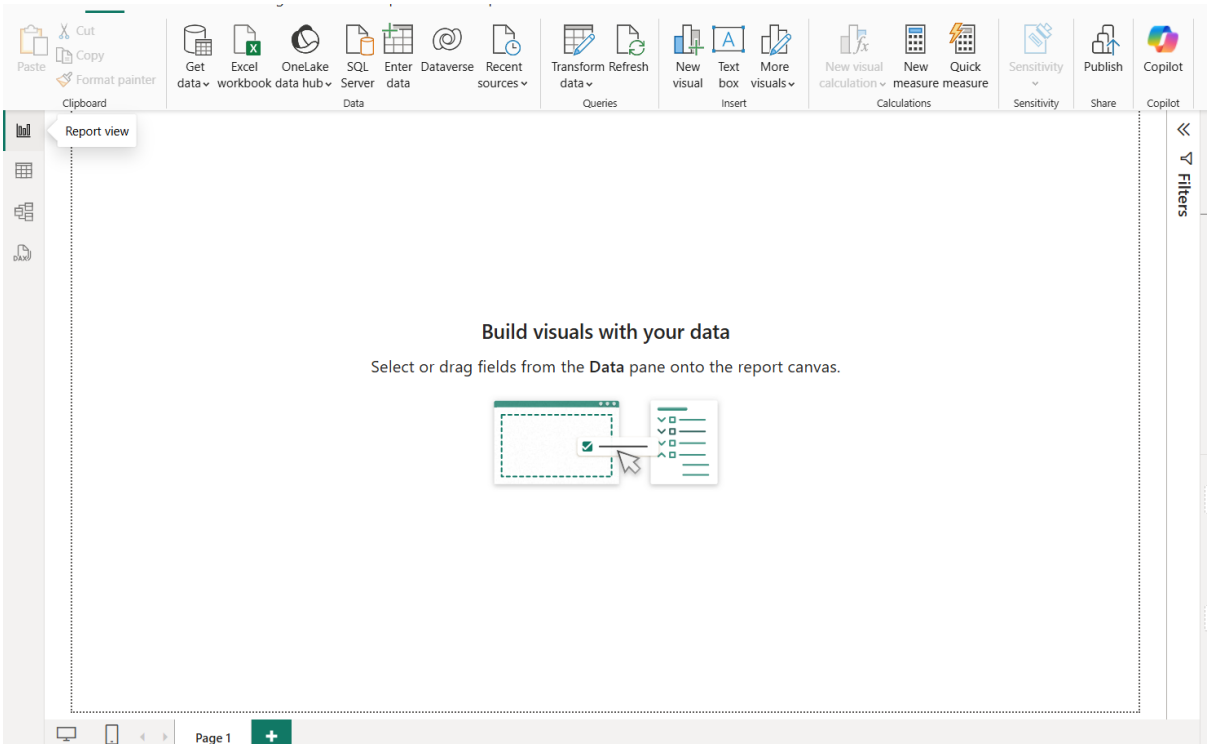
Name : alok mehta

Explaining the Three Buttons on the Left Side

Once your data is loaded into Power BI, you'll notice three buttons on the left side of the screen. These buttons help you navigate through your workspace and interact with your data.

---

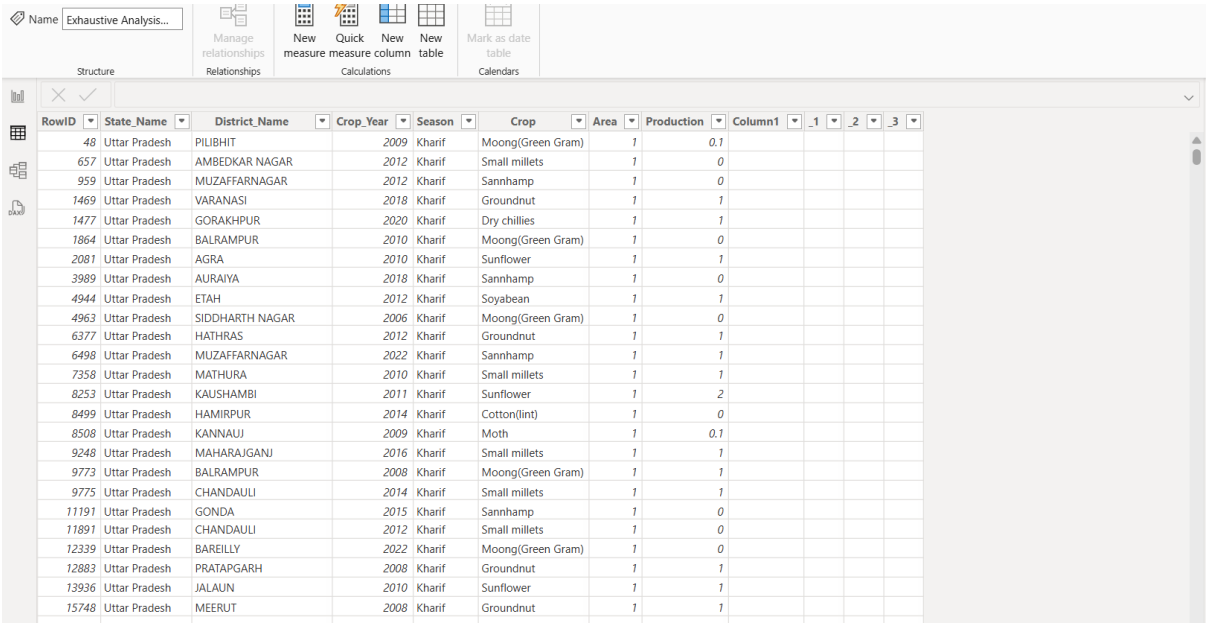
1. **Report View**
    - **Overview**: This section is where you design and create visual reports.
    - **How to Use It**: Drag and drop fields, such as numbers or categories, from your data tables into the main workspace. From there, you can select the type of chart or visual representation you want (e.g., bar chart, pie chart, etc.).
    - **Why It's Useful**: This is the most commonly used view because it allows you to present your data visually, making it easier to analyze and interpret insights.



Name : alok mehta

2. Table View

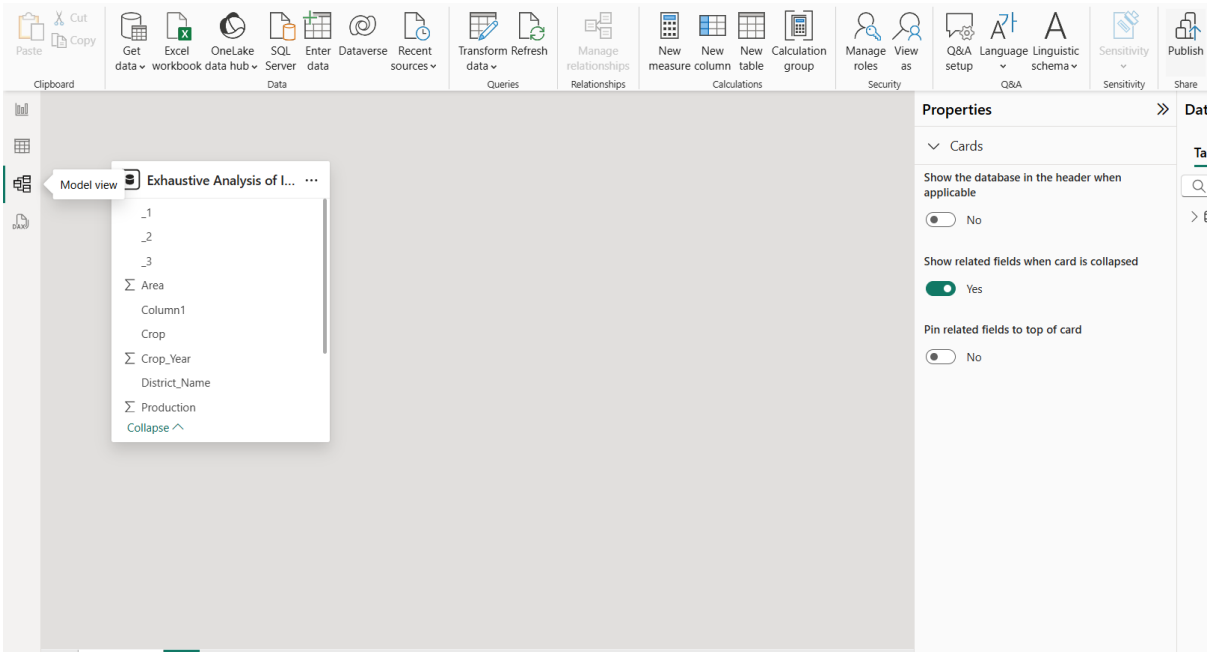
- **Overview**: This view displays your raw data in a structured, table-like format.
- **How to Use It**: View your dataset as rows and columns, similar to a spreadsheet. This is where you can closely examine the data.
- **Why It's Useful**: It's essential for validating your data and ensuring there are no errors or inconsistencies before creating visual reports.



RowID	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production	Column1	_1	_2	_3
48	Uttar Pradesh	PILIBHIT	2009	Kharif	Moong(Green Gram)	1	0.1				
657	Uttar Pradesh	AMBEDKAR NAGAR	2012	Kharif	Small millets	1	0				
959	Uttar Pradesh	MUZAFFARNAGAR	2012	Kharif	Sannhamp	1	0				
1469	Uttar Pradesh	VARANASI	2018	Kharif	Groundnut	1	1				
1477	Uttar Pradesh	GORAKHPUR	2020	Kharif	Dry chillies	1	1				
1864	Uttar Pradesh	BALRAMPUR	2010	Kharif	Moong(Green Gram)	1	0				
2081	Uttar Pradesh	AGRA	2010	Kharif	Sunflower	1	1				
3989	Uttar Pradesh	AURAIYA	2018	Kharif	Sannhamp	1	0				
4944	Uttar Pradesh	ETAH	2012	Kharif	Soyabean	1	1				
4963	Uttar Pradesh	SIDDHARTH NAGAR	2006	Kharif	Moong(Green Gram)	1	0				
6377	Uttar Pradesh	HATHRAS	2012	Kharif	Groundnut	1	1				
6498	Uttar Pradesh	MUZAFFARNAGAR	2022	Kharif	Sannhamp	1	1				
7358	Uttar Pradesh	MATHURA	2010	Kharif	Small millets	1	1				
8253	Uttar Pradesh	KAUSHAMBI	2011	Kharif	Sunflower	1	2				
8499	Uttar Pradesh	HAMIRPUR	2014	Kharif	Cotton(lint)	1	0				
8508	Uttar Pradesh	KANNAUJ	2009	Kharif	Moth	1	0.1				
9248	Uttar Pradesh	MAHARAJGANJ	2016	Kharif	Small millets	1	1				
9773	Uttar Pradesh	BALRAMPUR	2008	Kharif	Moong(Green Gram)	1	1				
9775	Uttar Pradesh	CHANDAULI	2014	Kharif	Small millets	1	1				
11191	Uttar Pradesh	GONDA	2015	Kharif	Sannhamp	1	0				
11891	Uttar Pradesh	CHANDAULI	2012	Kharif	Small millets	1	0				
12339	Uttar Pradesh	BAREILLY	2022	Kharif	Moong(Green Gram)	1	0				
12883	Uttar Pradesh	PRATAPGARH	2008	Kharif	Groundnut	1	1				
13936	Uttar Pradesh	JALAUN	2010	Kharif	Sunflower	1	1				
15748	Uttar Pradesh	MEERUT	2008	Kharif	Groundnut	1	1				

3. Model View

- **Overview**: This view is designed to manage and define relationships between your tables.
- **How to Use It**: Use the Model View to visualize all the tables you've imported into Power BI and establish connections between them. For instance, you can link a Sales table with a Products table to help Power BI understand how the data is interconnected.
- **Why It's Useful**: Establishing relationships between tables enables Power BI to perform advanced and detailed analyses by connecting different data points seamlessly.



Name : alok mehta

Removing and Restoring Empty Columns in Power BI

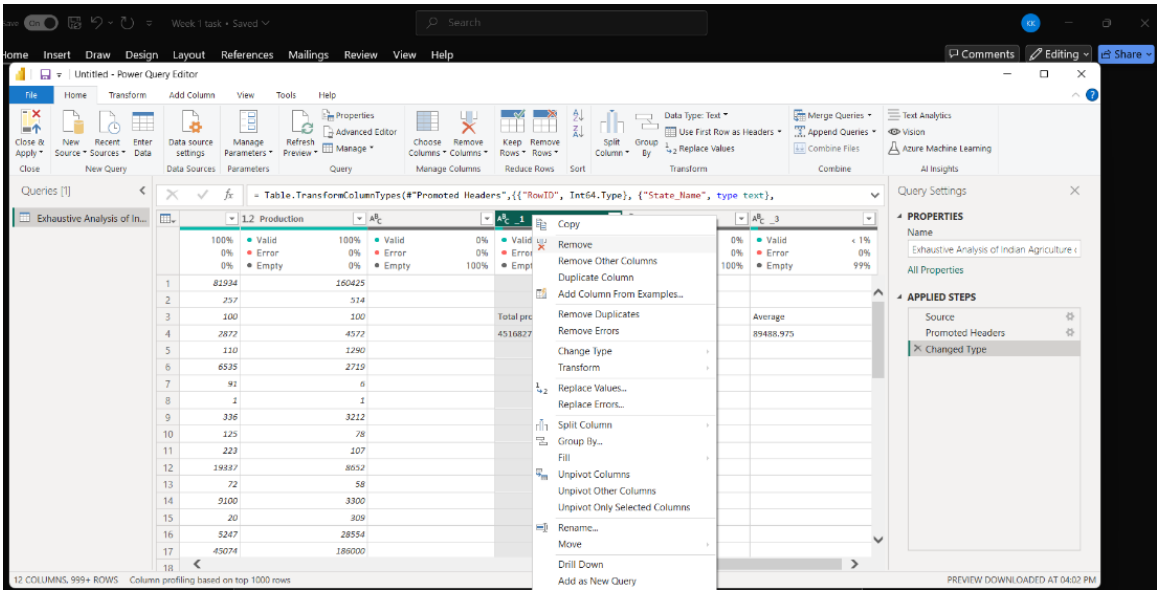
While working with data in Power BI, you might come across columns that have empty or missing values. Here's how you can handle them:

1. **Open the Power Query Editor**
   - After uploading your file (whether you chose Load or Transform), the data should open in the Power Query Editor. If it doesn't, simply click on **Transform Data** to access it.
2. **Review Your Columns**
   - Scroll through your dataset and check for columns with missing values or empty cells. Once identified, you can either remove these columns or restore them if needed.

To Remove a Column

- Right-click on the column header that contains the empty values.
- From the menu that appears, select **Remove** to delete the column from your dataset.

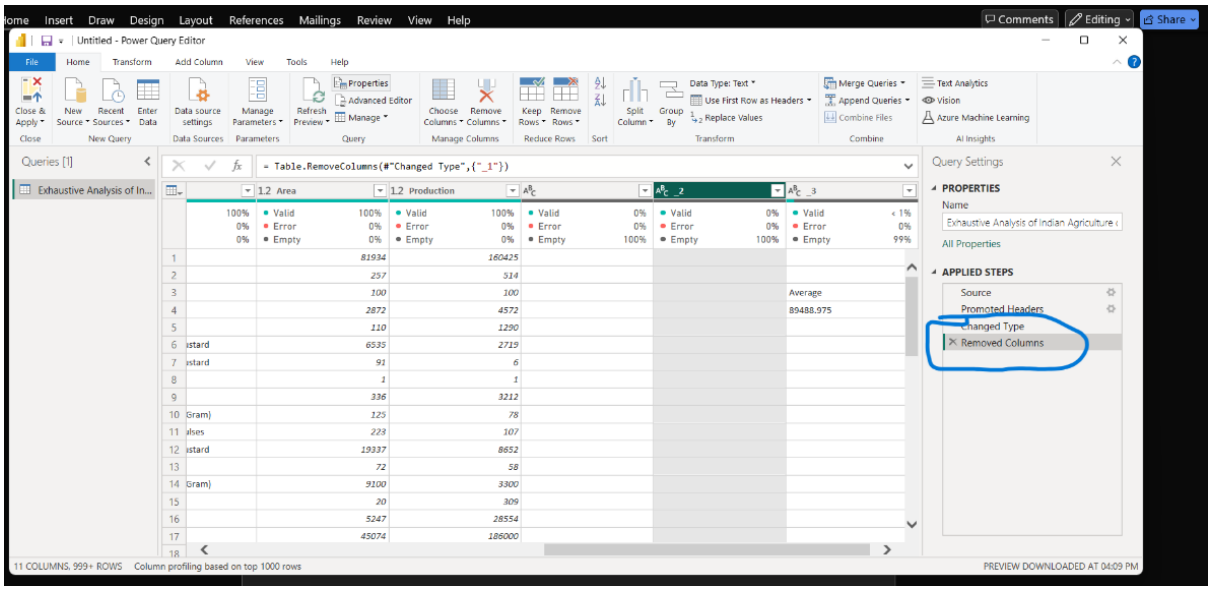
Note: Removing a column in Power BI doesn't erase the data permanently. It only removes it from the dataset you're working with in Power BI.



Name : alok mehta

Restoring a Column

- If you accidentally removed a column or want to bring it back, you can **restore** it by using the **Undo** button in the top-left corner of the Power Query Editor or by going back into the **Applied Steps** pane (on the right) and removing the **Remove Column** step.

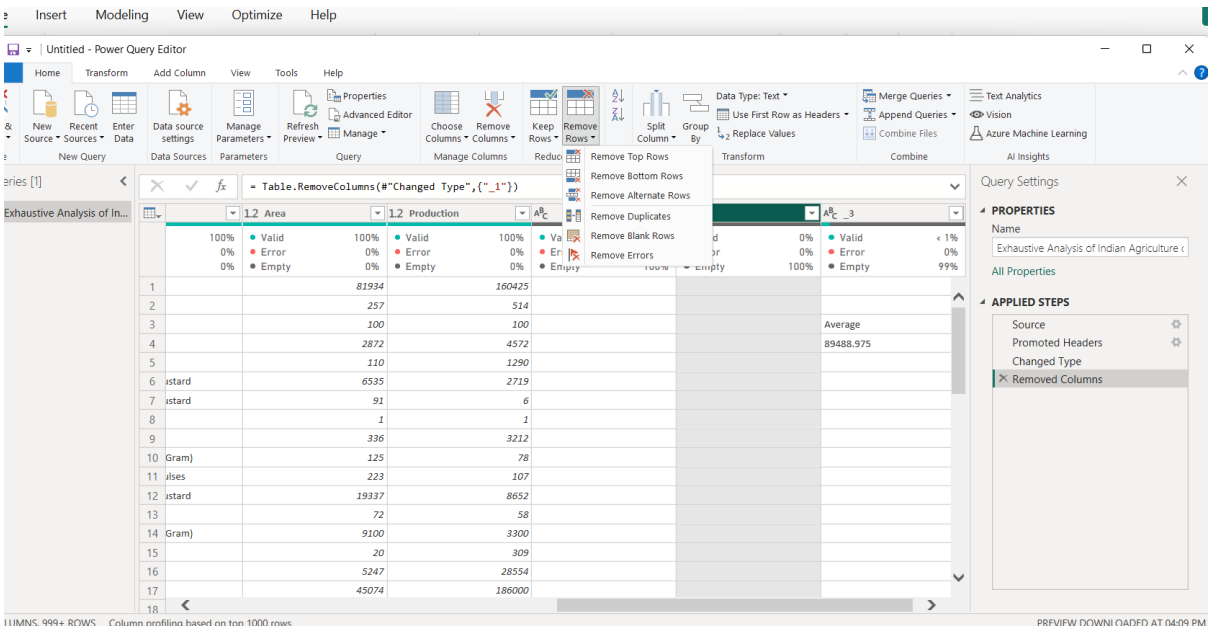


Name : alok mehta

Remove Rows Dropdown

While working in Power BI, you might find rows that are empty or not relevant to your analysis. You can use the **Remove Rows** dropdown to tidy up your data. Here's how:

1. **Open the Remove Rows Menu**
   - In the **Power Query Editor**, navigate to the **Home** tab.
   - Look for the **Reduce Rows** group, where you'll find the **Remove Rows** dropdown.
2. **Choose an Option**
   Here are some of the key options available in the dropdown and how they help:
   - **Remove Top Rows**: Use this to remove a certain number of rows from the start of your dataset. For example, if the first few rows contain headers or irrelevant information, this option lets you specify how many rows to delete.
   - **Remove Bottom Rows**: Similar to the previous option, but this one removes rows from the end of the dataset. You can define how many rows to take out.
   - **Remove Blank Rows**: This is handy for getting rid of rows that are entirely blank (i.e., with no data in any column).
   - **Remove Duplicates**: If your dataset has repeated rows, this option will identify and remove them, ensuring that only unique records remain.
   - **Remove Errors**: Use this to delete rows with errors, such as invalid data types or calculation issues, making your dataset cleaner.



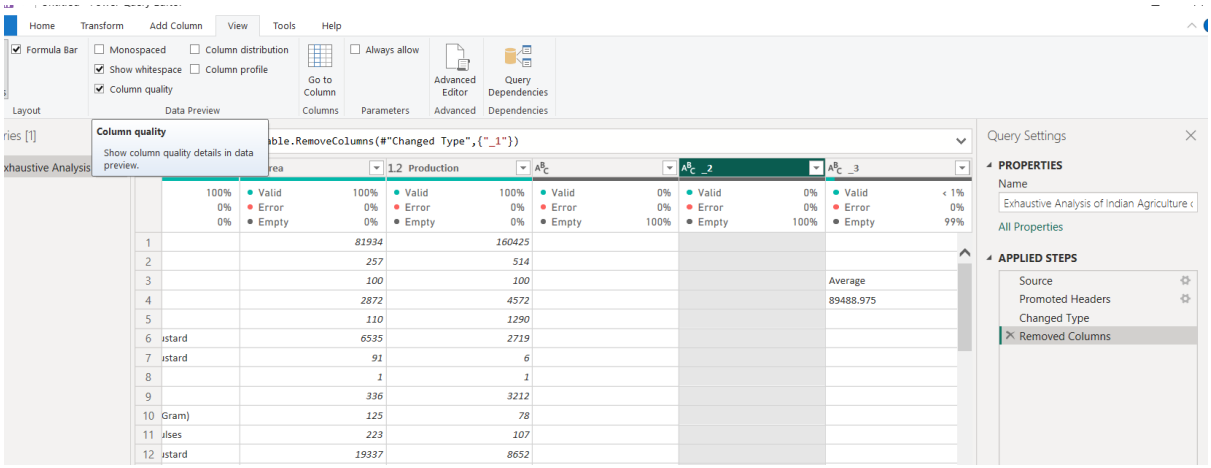
Column Quality in Power BI

Once your data is loaded into Power BI and you're working in the **Power Query Editor**, you can evaluate the quality of each column to make sure your data is clean and accurate. Here's how to use the **Column Quality** feature:

1. **Open the Power Query Editor**
   - After importing your data, click on **Transform Data** to access the Power Query Editor.
2. **Enable Column Quality**
   - In the Power Query Editor, go to the **View** tab at the top menu.
   - Locate the **Column Quality** option in the **Column Quality group** and turn it on.
3. **Interpreting Column Quality Indicators**
   Once enabled, Power BI will provide a visual summary of the data quality for each column in your dataset. You'll see the following metrics:
   - **Valid**: Displays the count or percentage of values in the column that are correct and match the expected data type (e.g., numbers, dates, text).
   - **Empty**: Shows how many rows in the column are blank or contain no data.
   - **Error**: Highlights the number of rows with issues, such as invalid data types or calculation errors.

   Power BI uses color coding for easy identification:

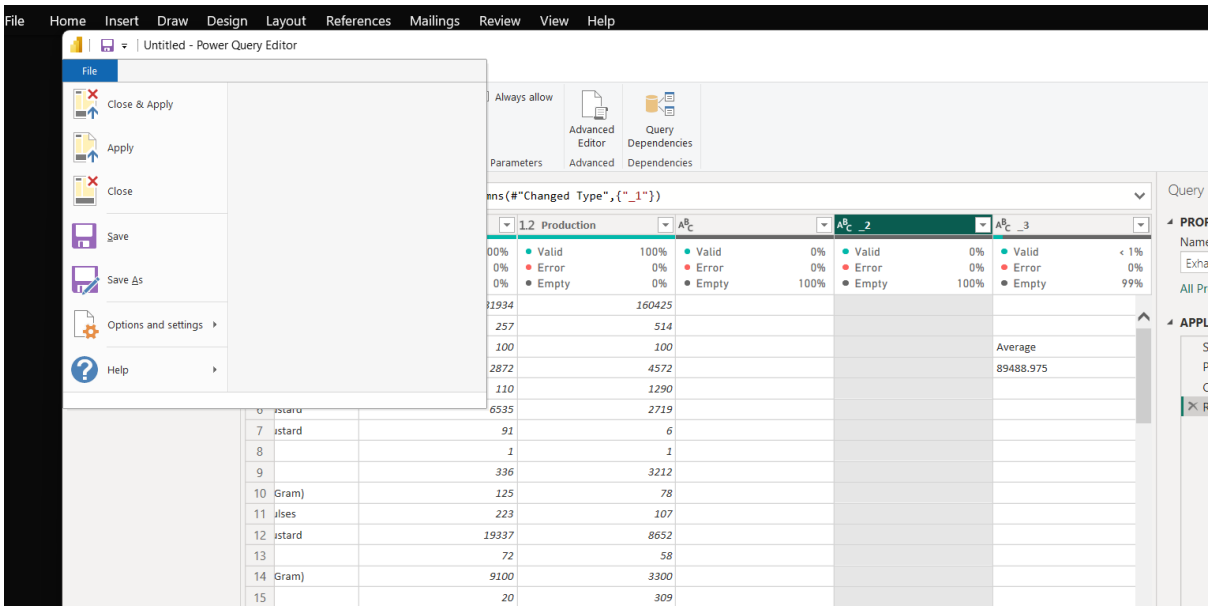
   - **Green**: Indicates a high proportion of valid entries.
   - **Yellow**: Warns of some missing or empty values.
   - **Red**: Flags errors or invalid data in the column.



Name : alok mehta

Saving Your Work in Power BI

- Use the **Close & Apply** button in the Power Query Editor to confirm and apply any changes made to your data.
- Navigate to **File > Save As** to store your Power BI project as a file on your computer.
- Make it a habit to save frequently while working, and when needed, go to **File > Export** to share your reports with others.



Name : alok mehta