Task 1 – ETL using Power BI

Author – S. Prajanya

# Getting Started

Our task of the day would be:

* Learning what is Power BI
* Performing ETL

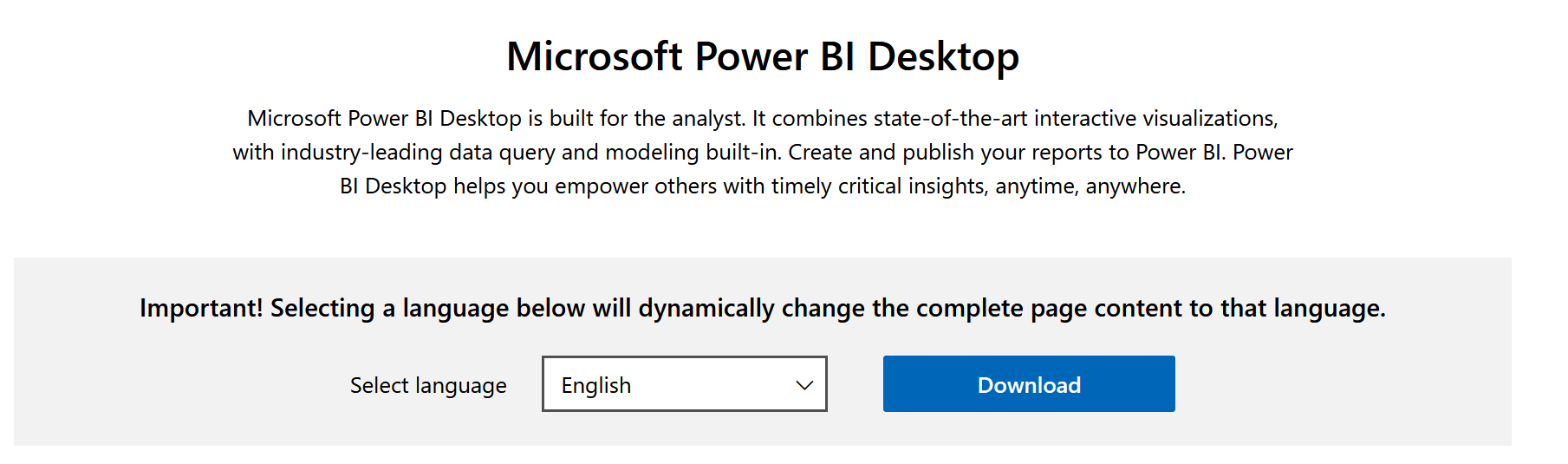
Power BI is a business analytics tool from Microsoft used to visualise data from various data sources.

ETL is to :

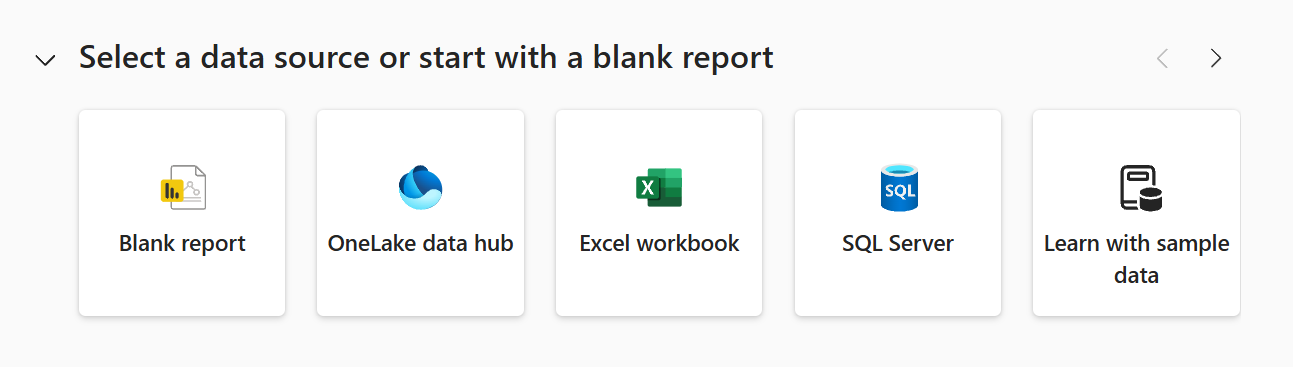
* Extract data : Pull data from data sources
* Transform : Processing and cleaning data
* Loading : Analysing data using dashboards

# Power BI

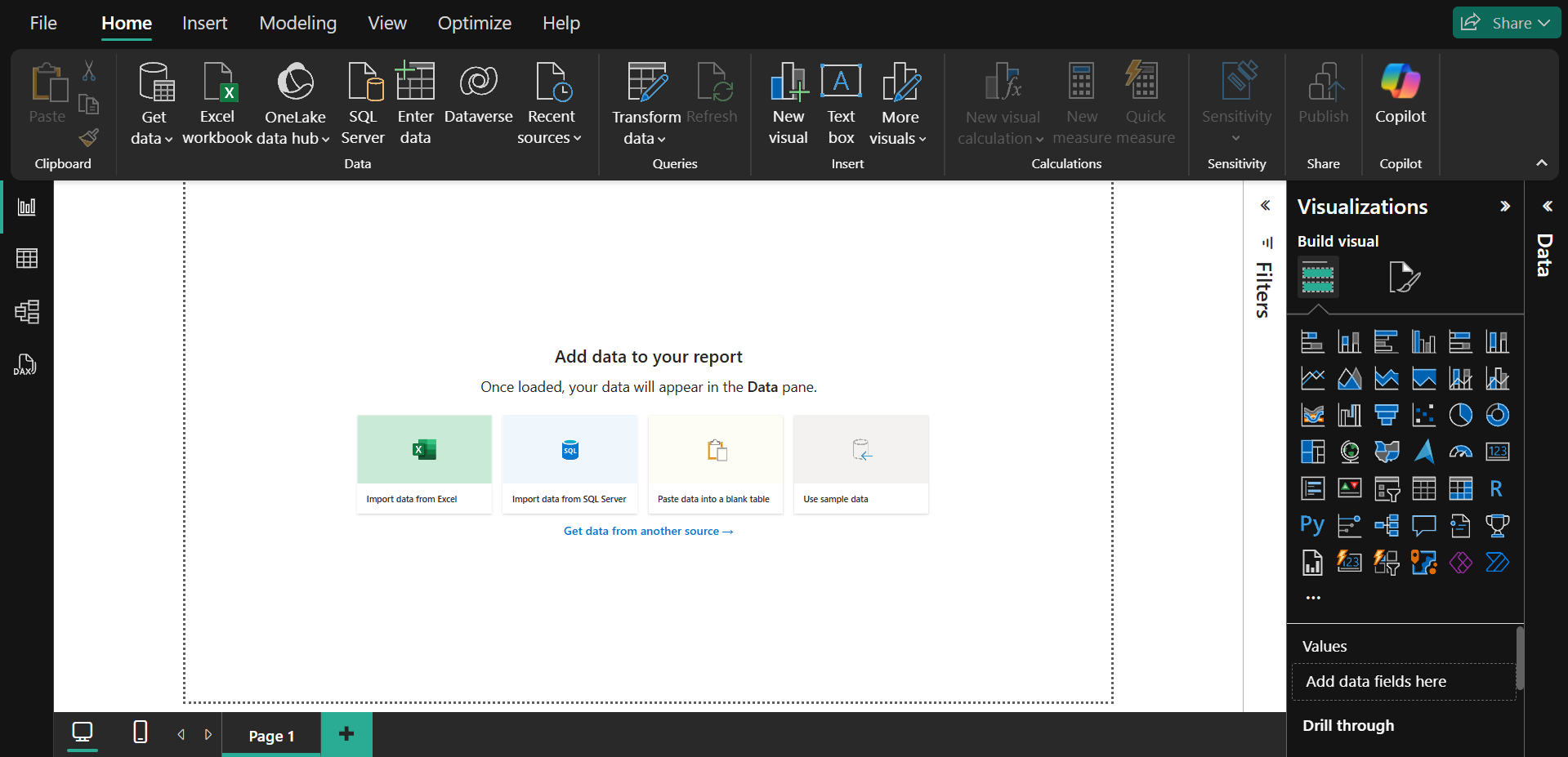
Download Power BI from Microsoft’s own site.



After downloading Power BI, open it and create a blank report.



This is the Power BI dashboard. On the right-hand side, we’ve got the Visualisation tools. On top, we have tabs like Home, Insert, Modelling, etc.

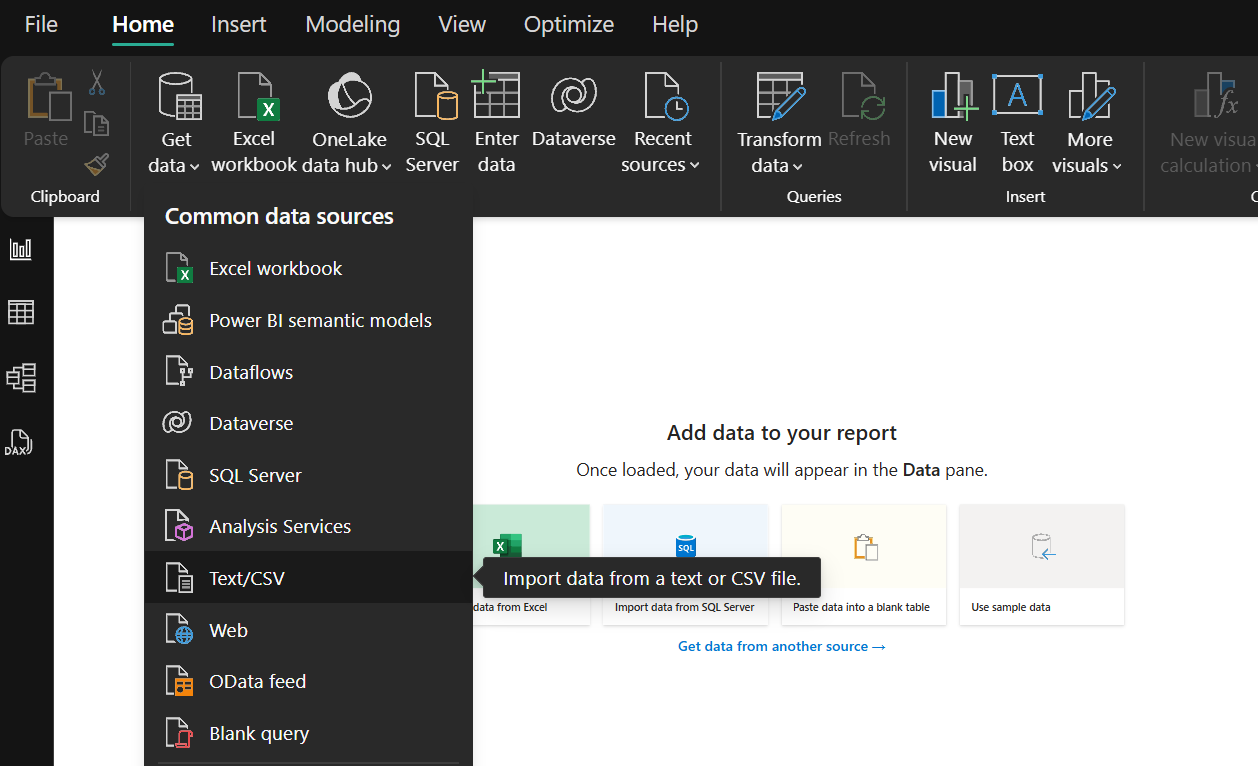


On the left-hand side, we’ve 4 tabs :

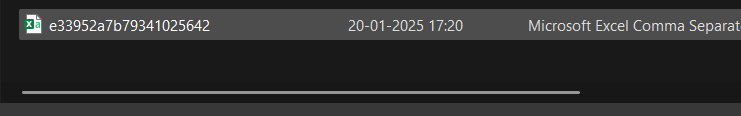
* Report view : Place to perform all visualisations.
* Table view : Place to deal with imported datasets.
* Model view : Place to merge different datasets.
* DAX Query view : Place to run DAX queries.

# Extract

Go to Home tab, then click on Get Data. Something like this would appear. Click on Text/CSV option.

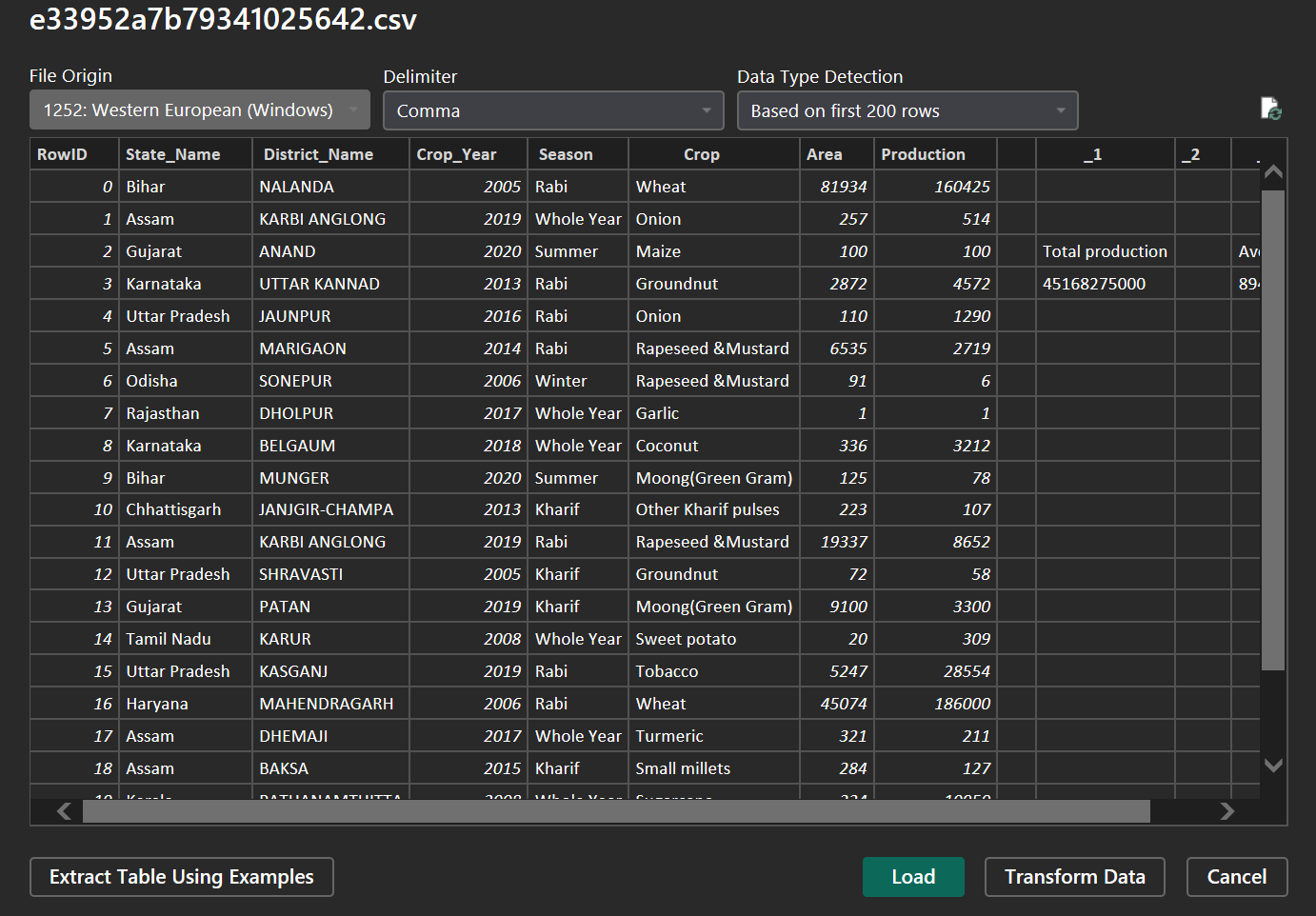


Select this particular file [the dataset] which you’d have downloaded from the LMS website.

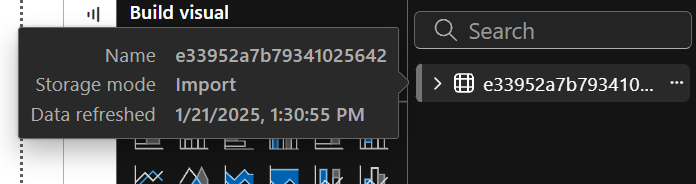


Something like this would appear. This is the overview of your dataset. It has 3 buttons :

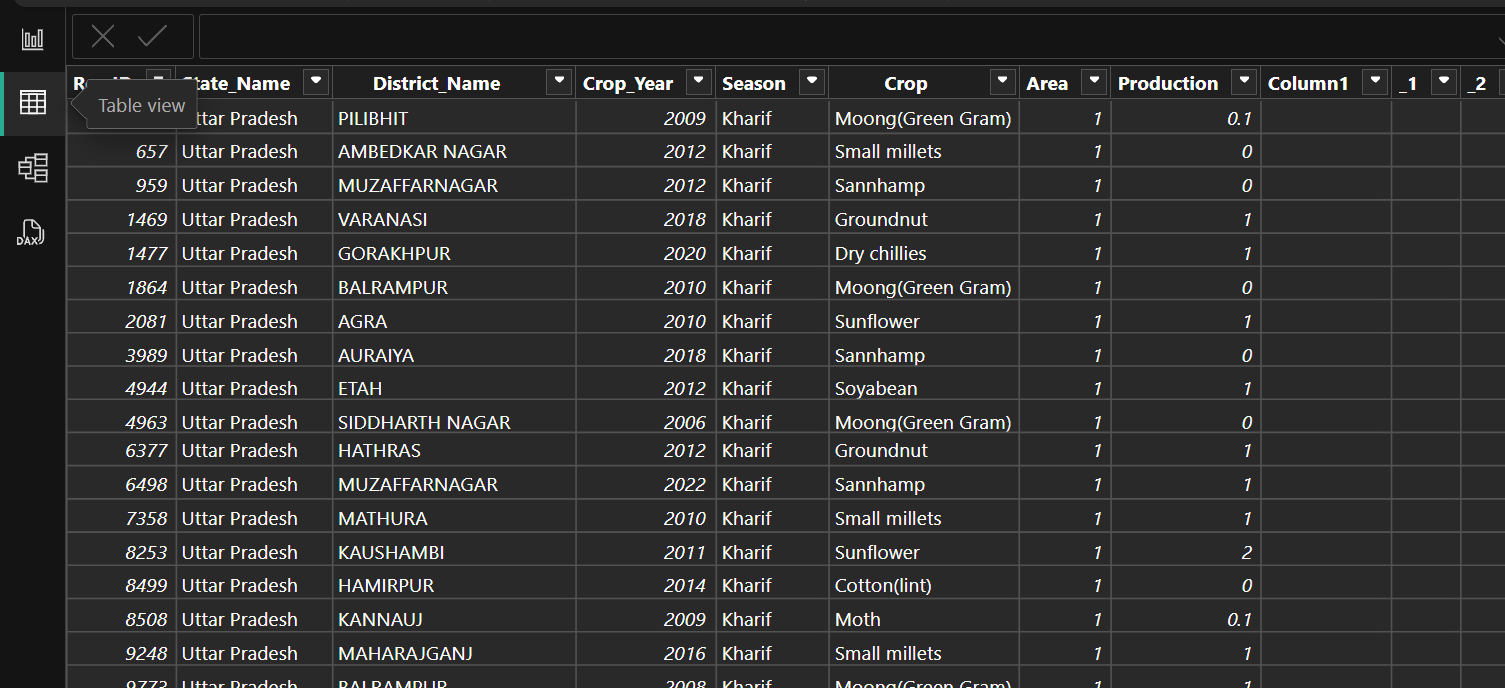
* Load : Click if your data is clean.
* Transform : Click if your data needs some correction / cleaning
* Cancel : Undo / Terminate this process



Click on Load. Something like this would’ve appeared.

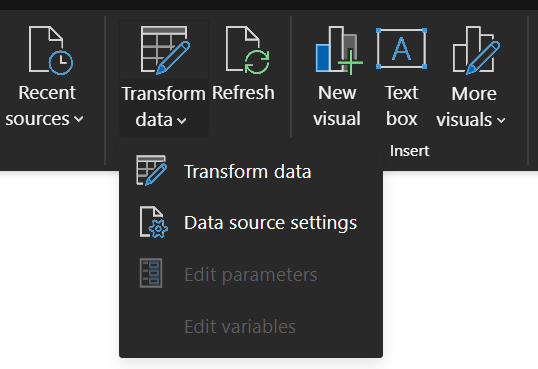


Go to Table view and you’ll get something like this :

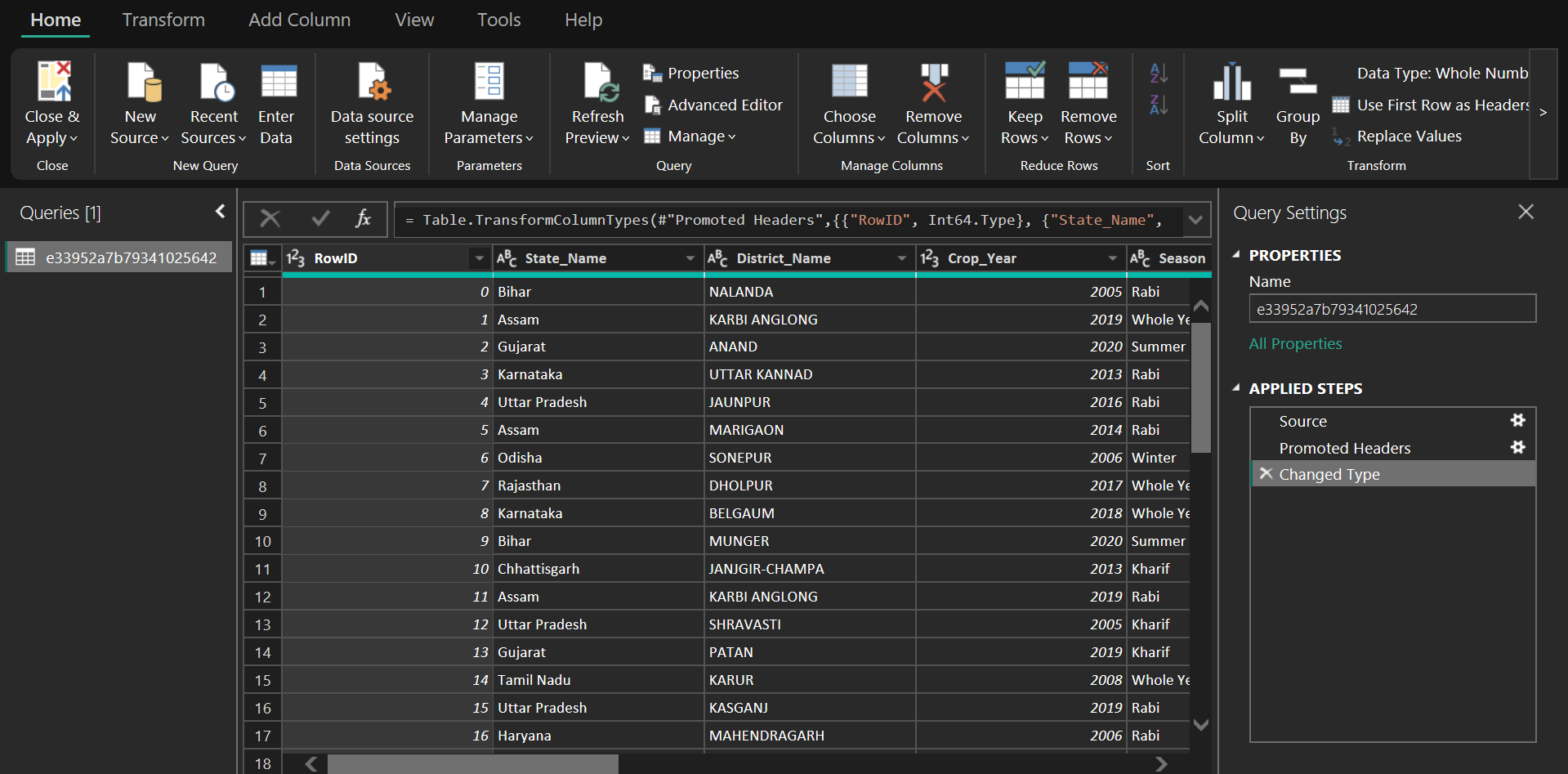


# Transform

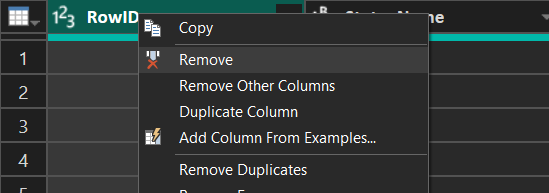
In the Home tab, click on the Transform Data.



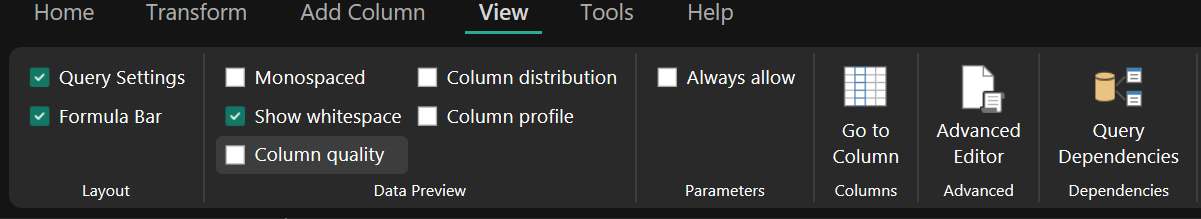
Another window pops up. This is where you perform data-cleaning. This window is called Power Query Editor.



Since RowID column has no use in this dataset, we’re gonna remove it. To do so, right-click on that column and hit Remove.

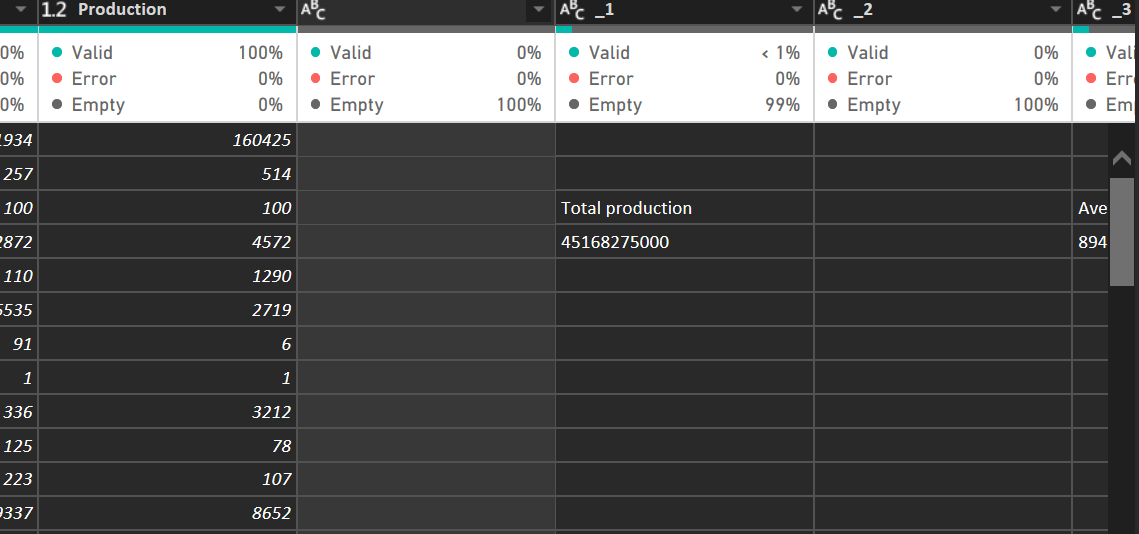


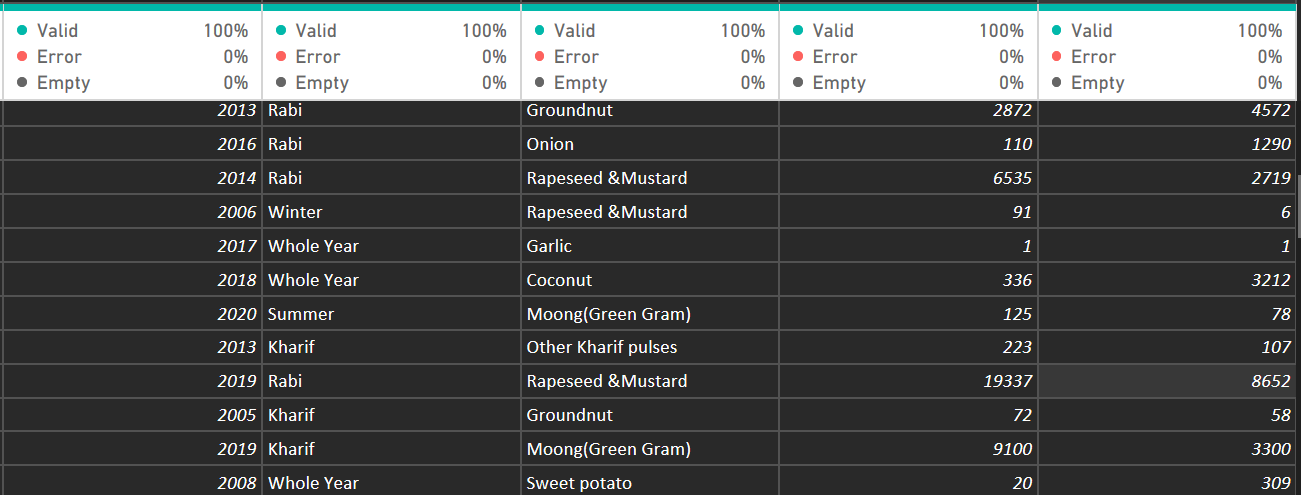
In the same dataset, we’ve some empty columns as well. Delete them all! Now, to know which column is delete-worthy, go to View tab and check Column quality option.



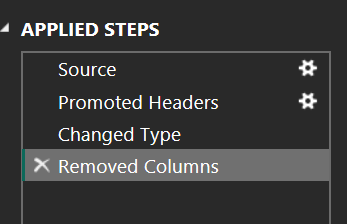
It shows you which column is Empty, Null, etc.

Remove the columns that show less than 100% Valid.



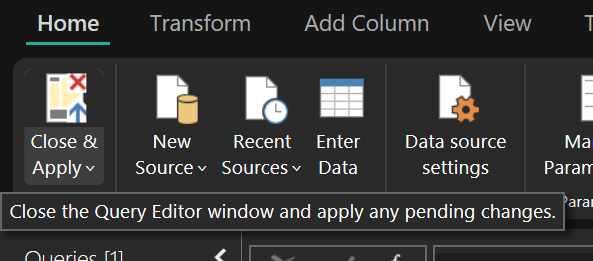


On your right-hand side, there is this column Applied Steps that shows the actions taken in this window. It is useful to roll-back any unwanted changes.



# Load

Finally, after transforming your data, click on Close & Apply.



Your table in the Table View may have some changes like this :

