Power BI Case Studies:

Case Study 1: (Power Query)

Objective: To reconcile GSTR 2A Returns with Purchase Journal using Power Query ETL (Extract, Transform and Load) features.

Source Files:

1. Three months GSTR 2A as *.json files
2. Three months Purchase Journal Report as Excel Workbook

Problem Statement:

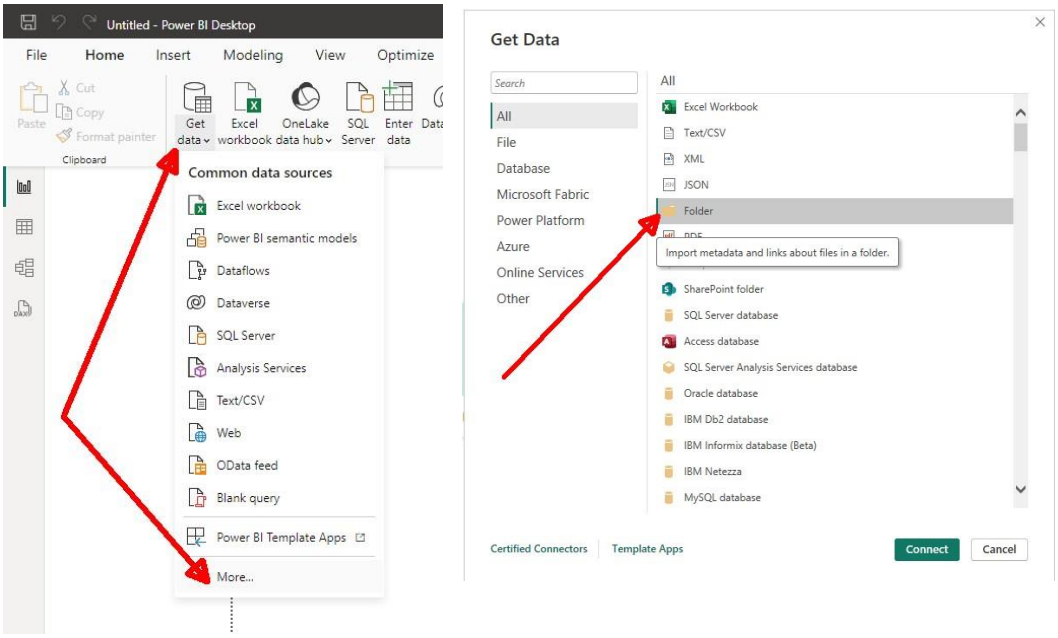
1. The Excel file contains two sheets one for purchase journal with suppliers name and without their GSTIN. Another sheet contains the list of suppliers with their GSTIN.
2. The JSON files have nested tables and we need to reconcile with the B2B table which contains the suppliers data.
3. From the next month onwards, we need to automate the cumbersome process.

Overview of steps:

1. Use extract from folder feature to extract multiple json files and identify B2B table.
2. Keep only the relevant fields for reconciliation from the above table.
3. Extract purchase journal data from Excel file and merge the same with GSTIN table to get the GSTIN details of the suppliers.
4. Reconcile the both JSON and PJ tables by merging with INNER JOIN of common fields.
5. Find unreconciled JSON and PJ by merging with LEFT ANTI and RIGHT ANTI of common fields.

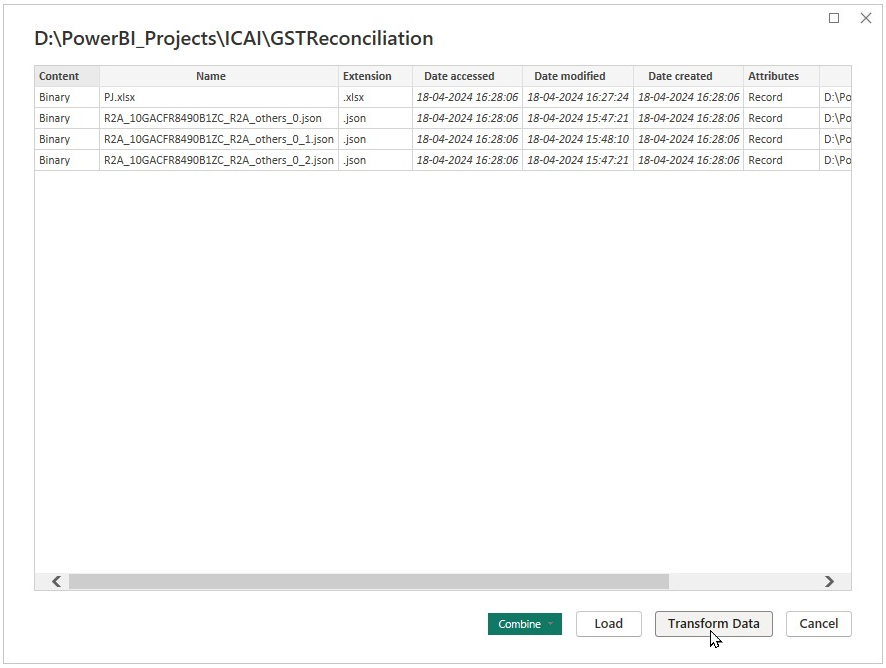
Detailed Steps:

1. Open Power BI and create a new project. Click Home Tab -> Get Data -> More to select Get Data from folder and select the Folder option and click Connect.

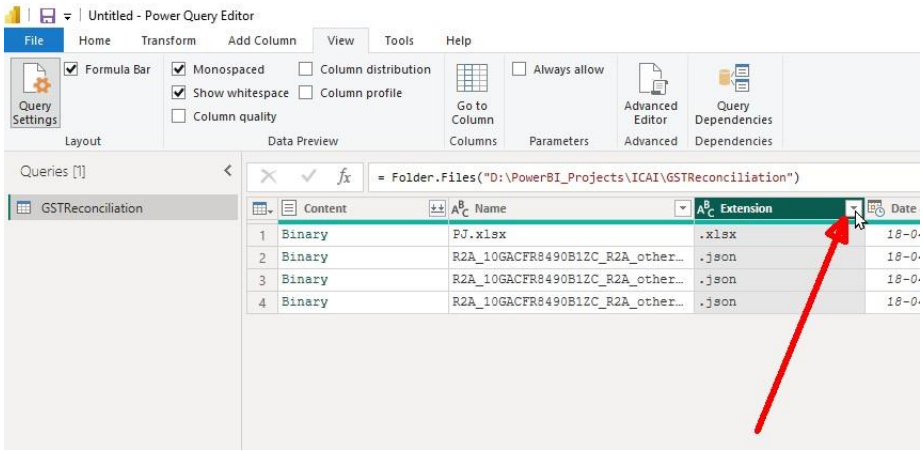


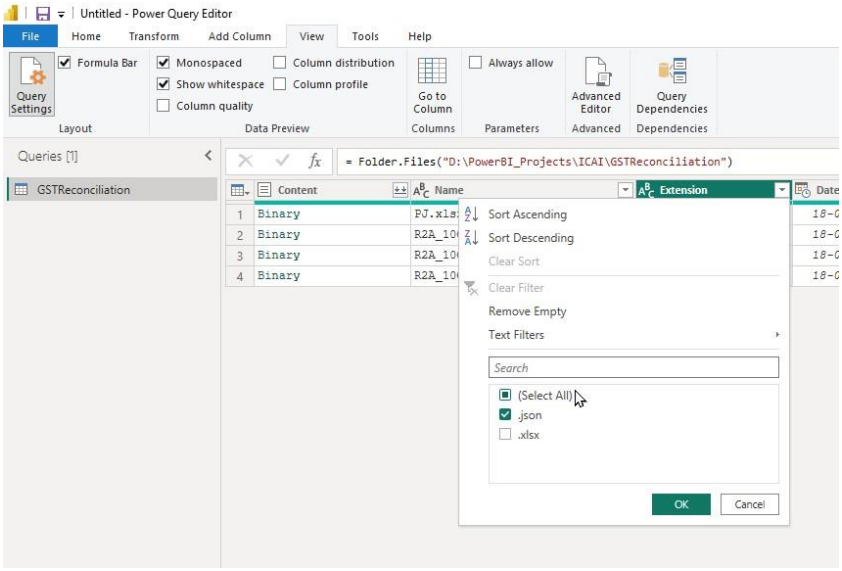
2. Select the folder where JSON files are kept.

3. Click Transform Button. Combine files can be clicked only when all the files are having the same structure and already clean and no transformations required. Since we need to transform the files and having Excel file also in the same directory we need to click Transform.

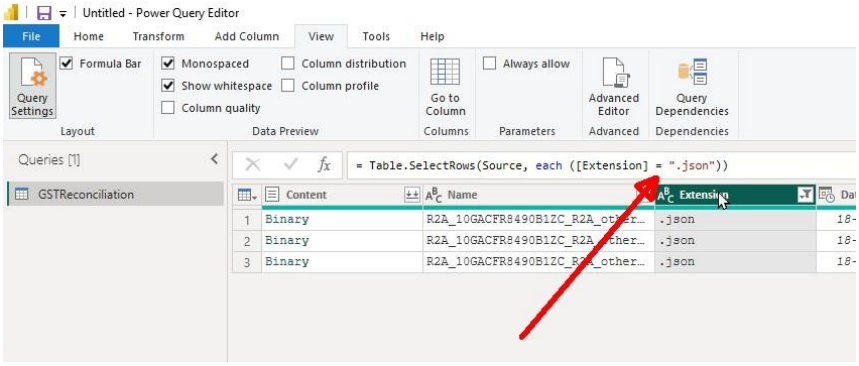


4. Click on filter button of Extension column and select only JSON files as in the image below:

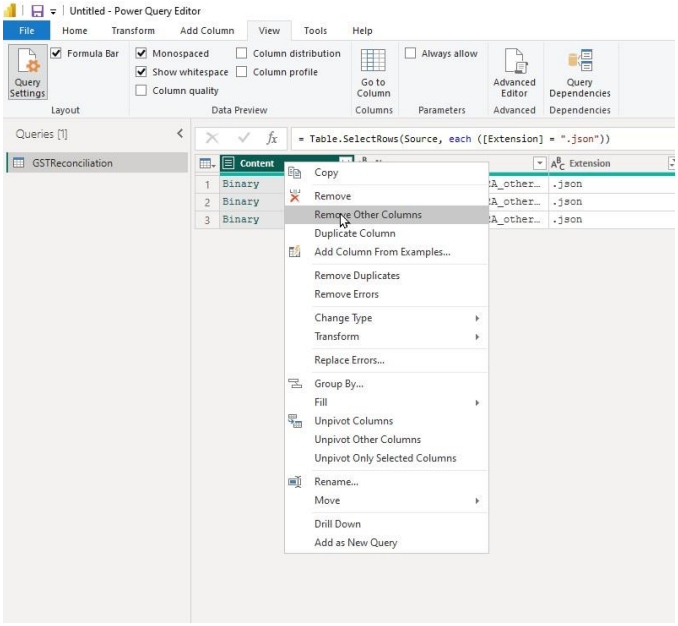




5. Confirm the M Language formula filtering only .json extensions.

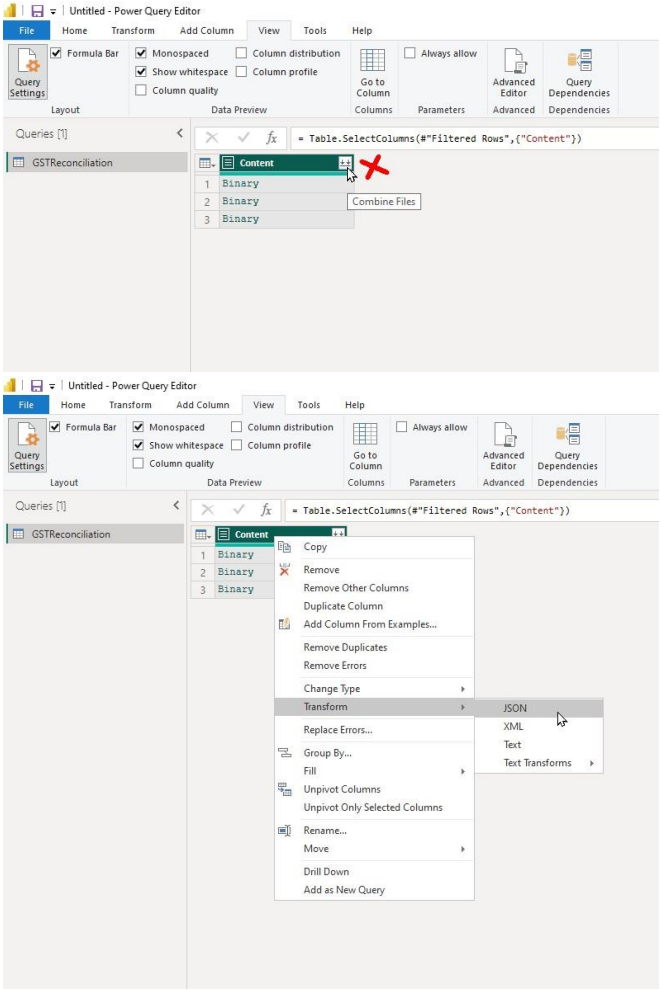


6. Right Click on Content Column header and select Remove Other Columns since we don't require other than the content column:

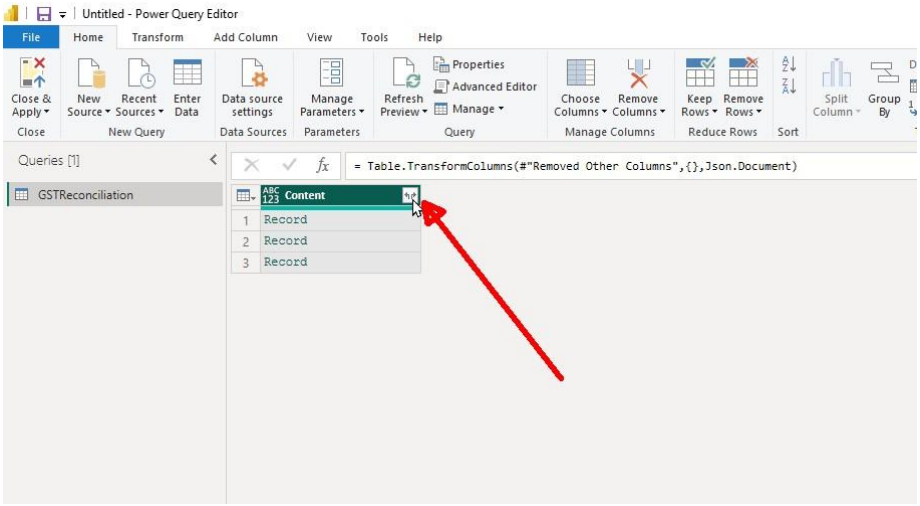


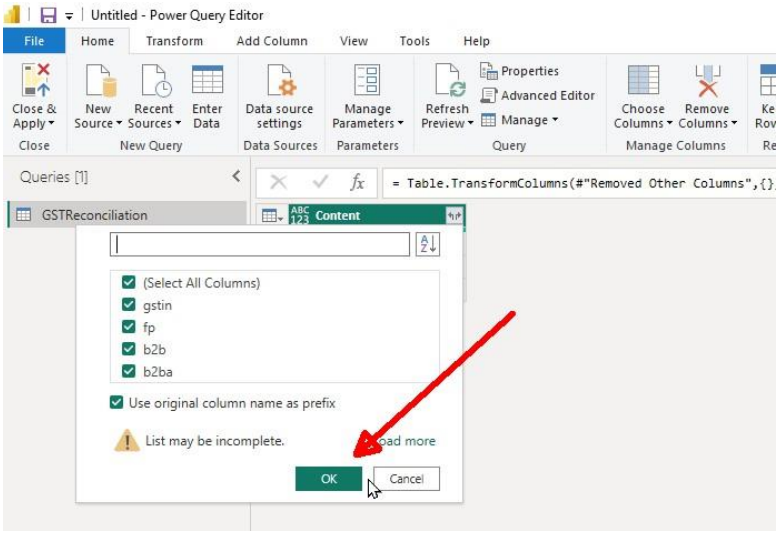
7. Now also we should not click Combine files button. The reason is: GSTR2A JSON files are not having the same structure always. Some may have B2B tables, some may not

have. Hence, we have to Right Click on Content column header, Select Transform -> JSON. By this way we are telling Power Query that, each file should be separately transformedas JSON files.

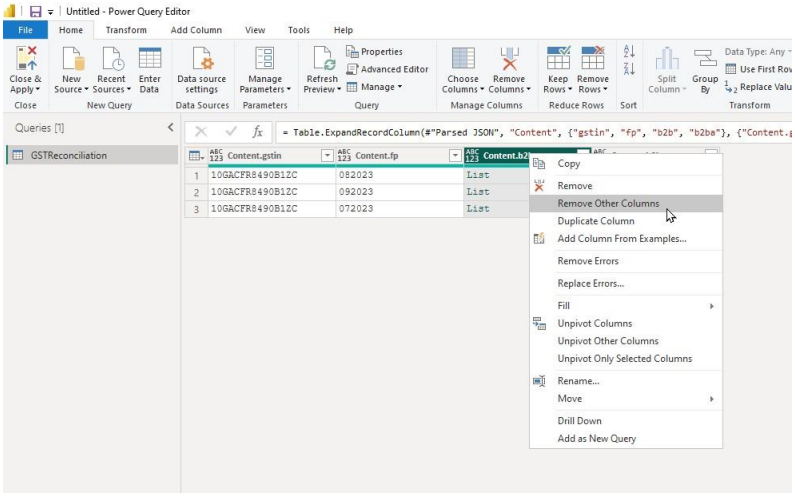


8. We can see that JSON files are transformed from Binary to Records. Click on Expand Table button as shown in the image below. Click Ok.

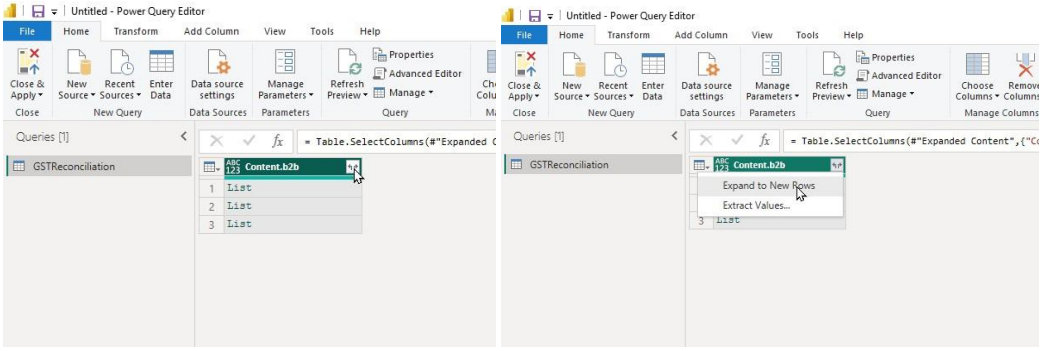




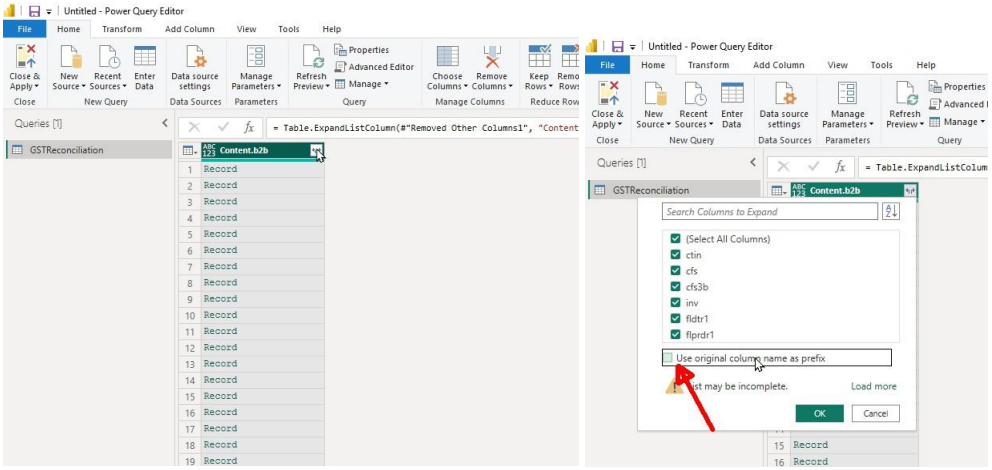
9. Since we need only the B2B List for reconciliation purpose, Right Click on B2B Column and Select Remove Other Columns.



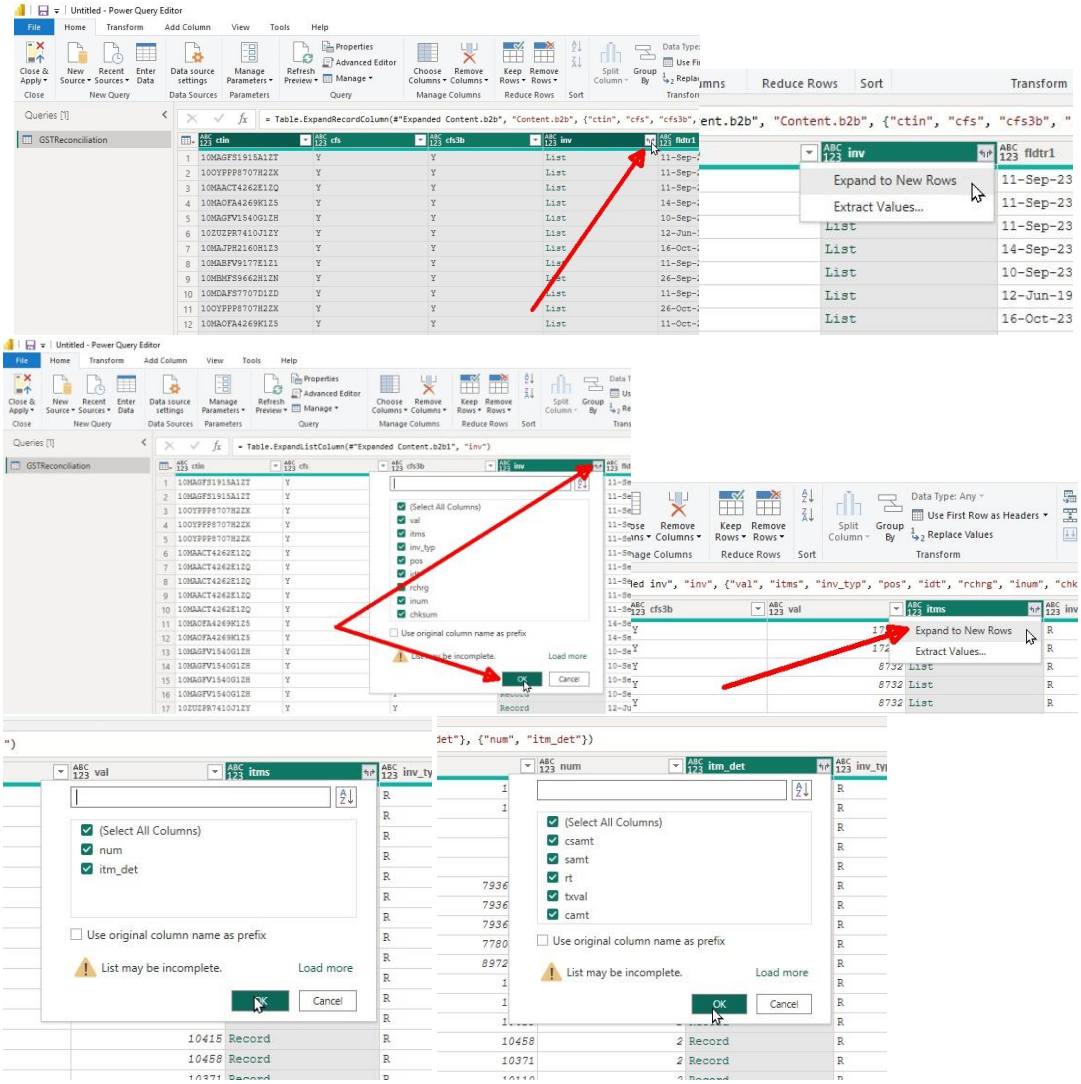
10. Again click Expand Button of B2B Column. Now we need to select Expand to New Rows (since it is a list).



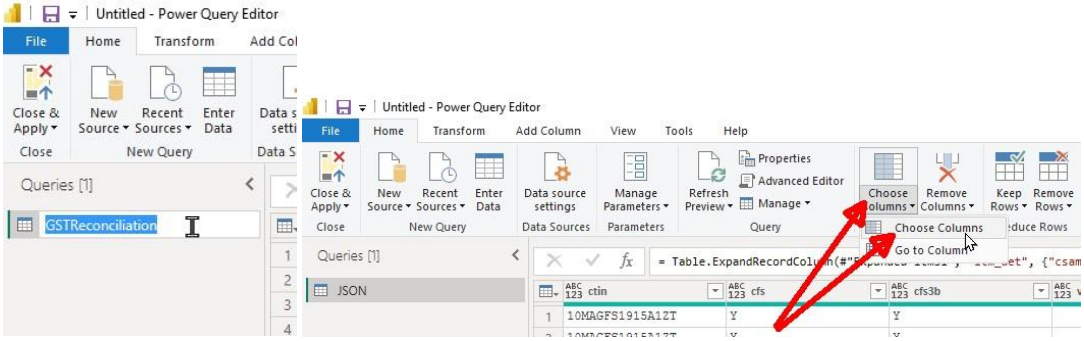
11. Now the lists are expanded to new Records. Again click Expand Button. This time remove the check mark on Use original column name as prefix since we don't require and click Ok.



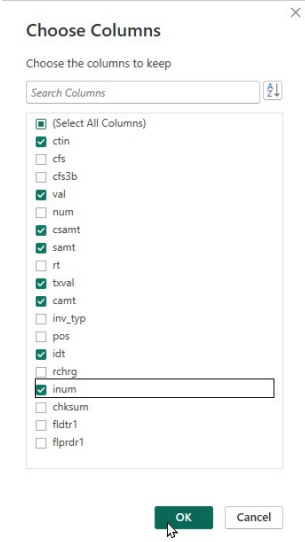
12. The records are expanded to various columns. Remember JSON files can have nested tables. These nested tables are treated as Lists and Records in Power BI. Lists are columns, Records are rows of data in columns. So, our next step is to Expand to new rows if they are lists and Expand if they are records to columns until we don't have any more lists and records. Doing so, we get a flat table of GST B2B Data. Following screenshots show this:



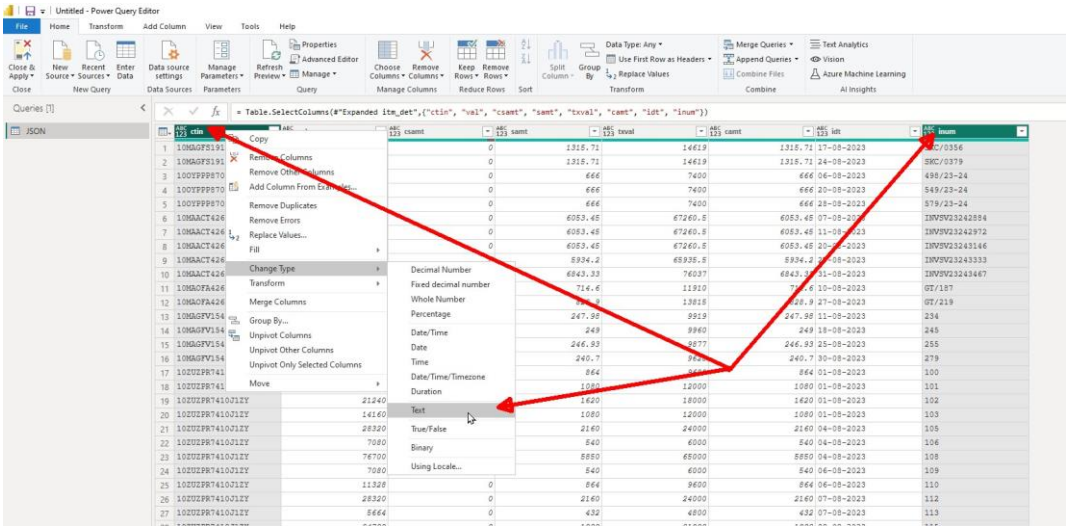
13. Rename the Query name to JSON and Select Choose Columns in the Home Tab to choose only the required columns for reconciliation.



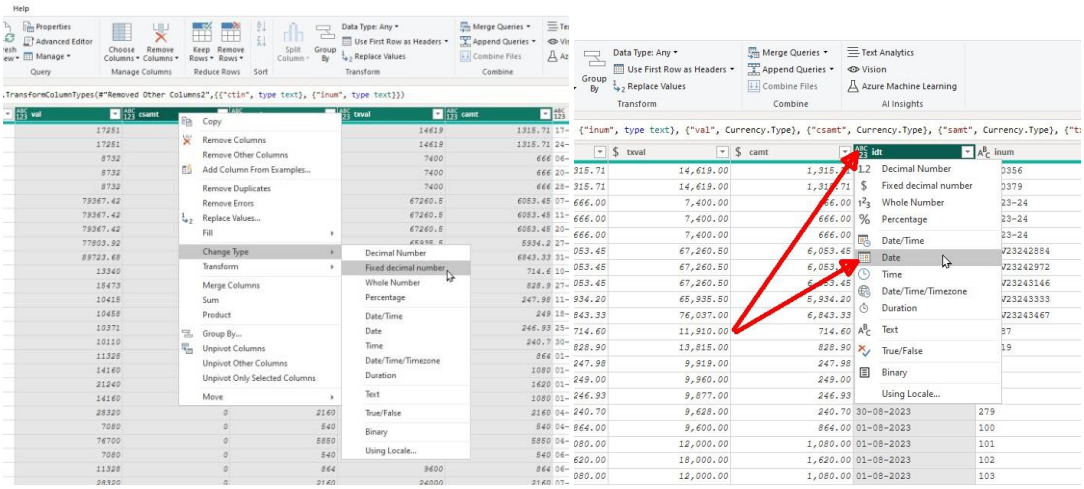
14. Choose only the required columns as shown in the image.



15. Change the type of ctin and inum columns as text by selecting multiple columns (by pressing Ctrl key and clicking on the headers) and right click on any of the column selected headers then select Transform->Change Type-> Text

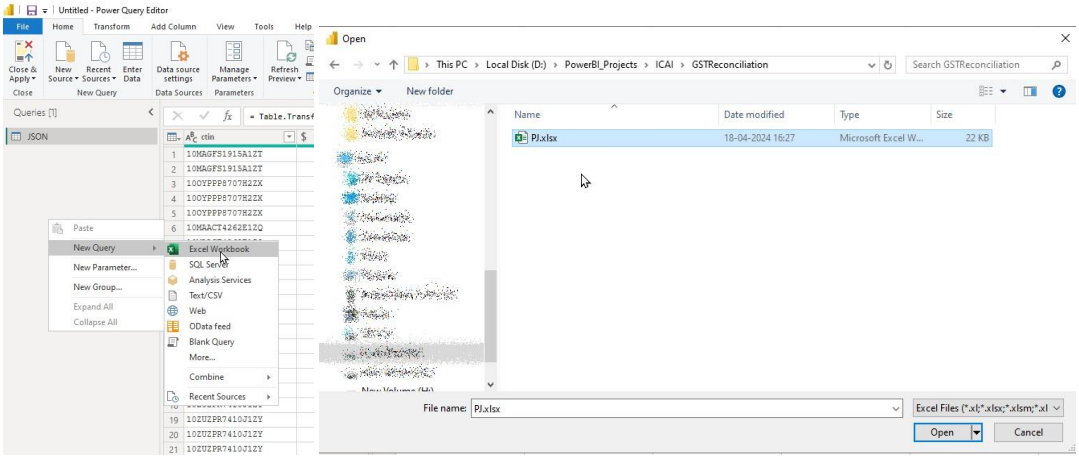


16. Same way select the columns val, csamt, samt, txval, camt and change the type to Fixed Decimal number (since they are currency). Column idt can be changed to date as well.

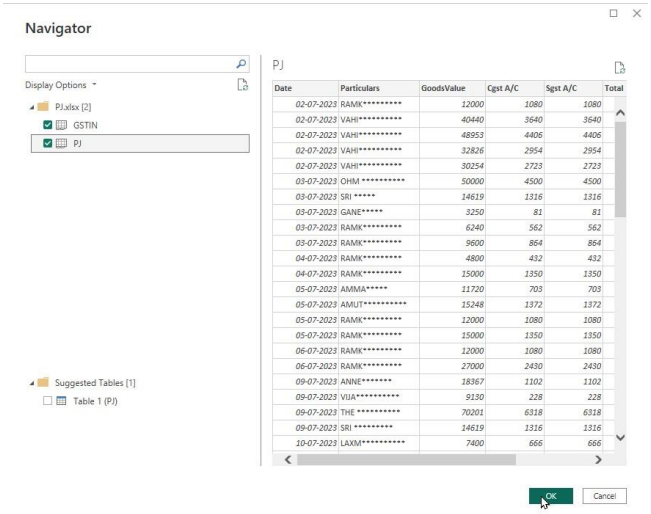


Now we can proceed to extract the Purchase Journal data from Excel file.

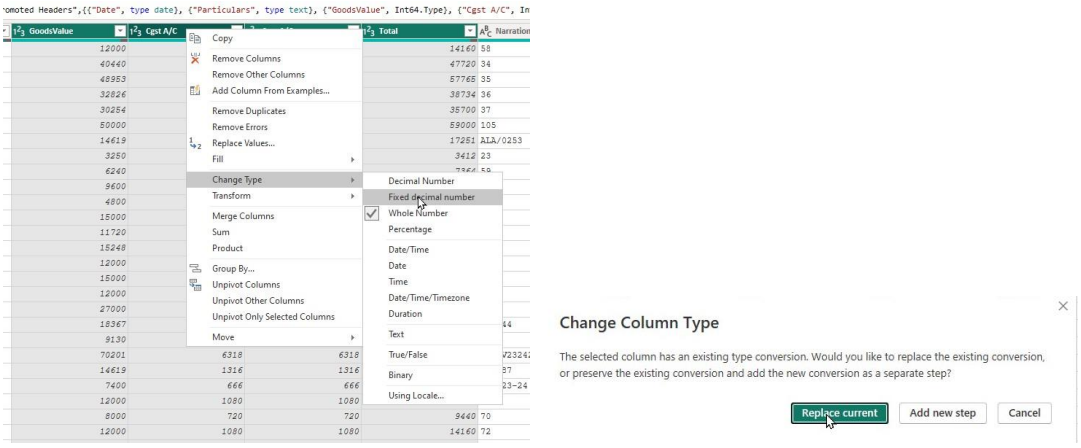
17. Right click on empty area of Queries area and select New Query->Excel Workbook and select PJ.xlsx file.



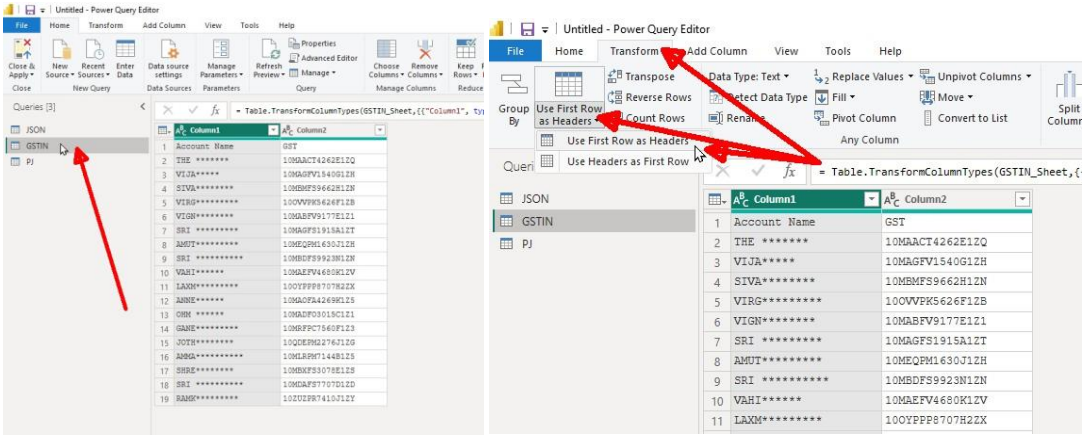
18. Select both the sheets GSTIN and PJ, then click Ok.



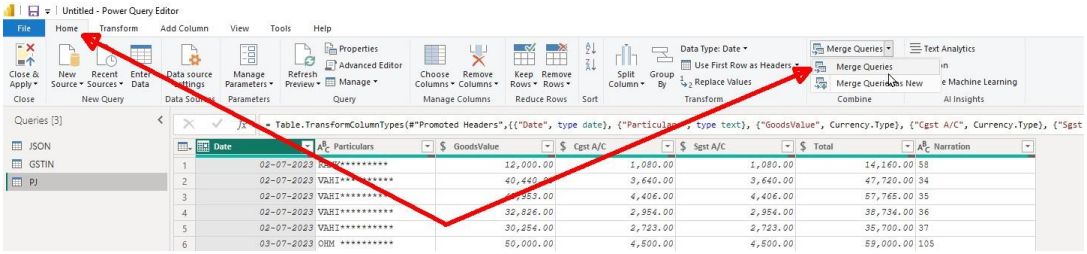
19. As in the previous case, change the type of currency fields Goods Value, Cgst A/c, Sgst A/c, Total to Fixed Decimal number. We can select Replace Current in the window to replace the type.



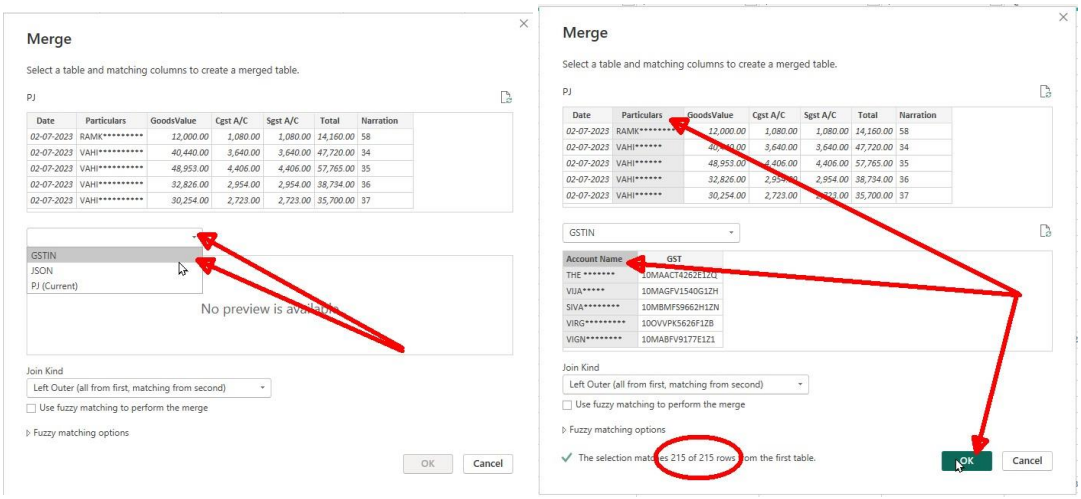
20. In the GSTIN query, we can see that Account Name and GST are not treated as column headers, so let us promote them to headers by clicking Transform -> Use First Row as Headers.



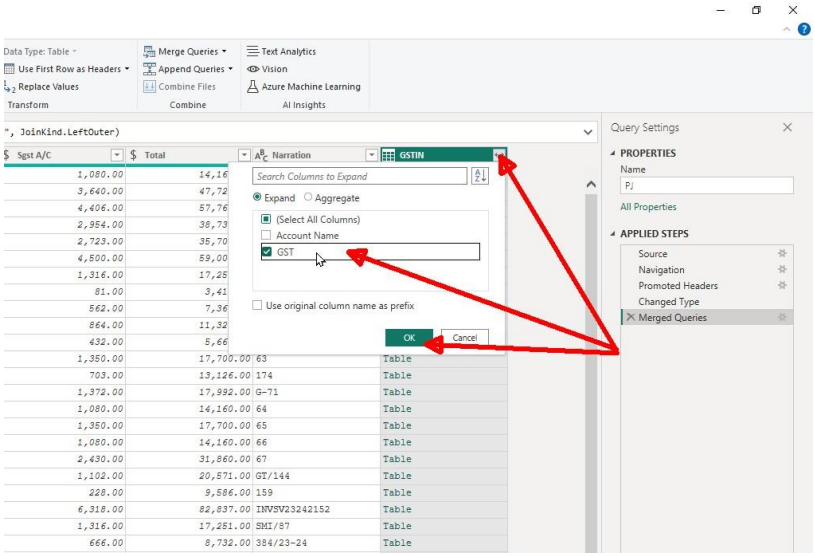
21. Select PJ Query and select Home Tab->Merge Queries since we want GST numbers from GSTIN merged with PJ.



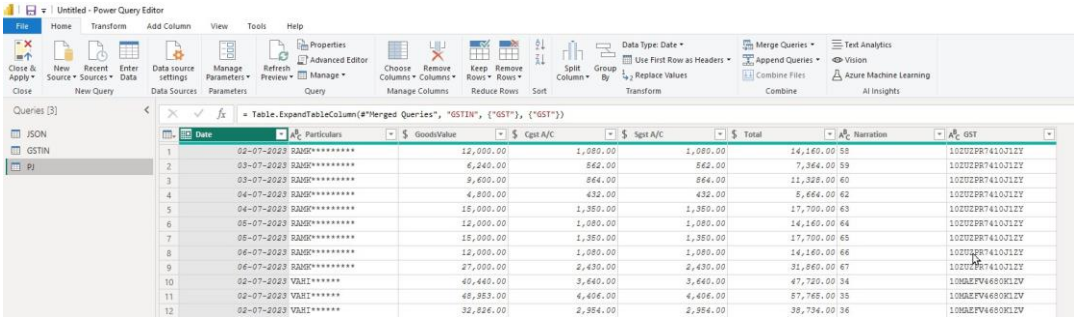
22. Select the GSTIN query in the bottom and match with Particulars of PJ Query with Account name of GSTIN query. This is like VLOOKUP function in Excel where we lookup the field and get the data from another table. Select Left Outer Join (to keep purchase data even if there are no GST numbers). Note that the selection matches all 215 rows.



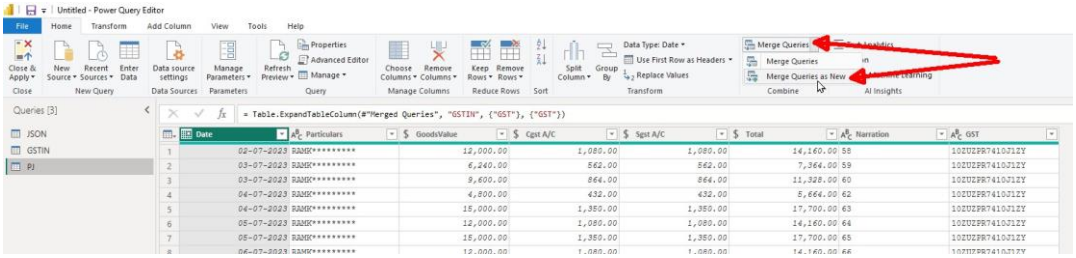
23. Scroll to far right of PJ query to see the merged GSTIN Query and click Expand Button. Then select only GST Column to get the corresponding GST Numbers of the suppliers.



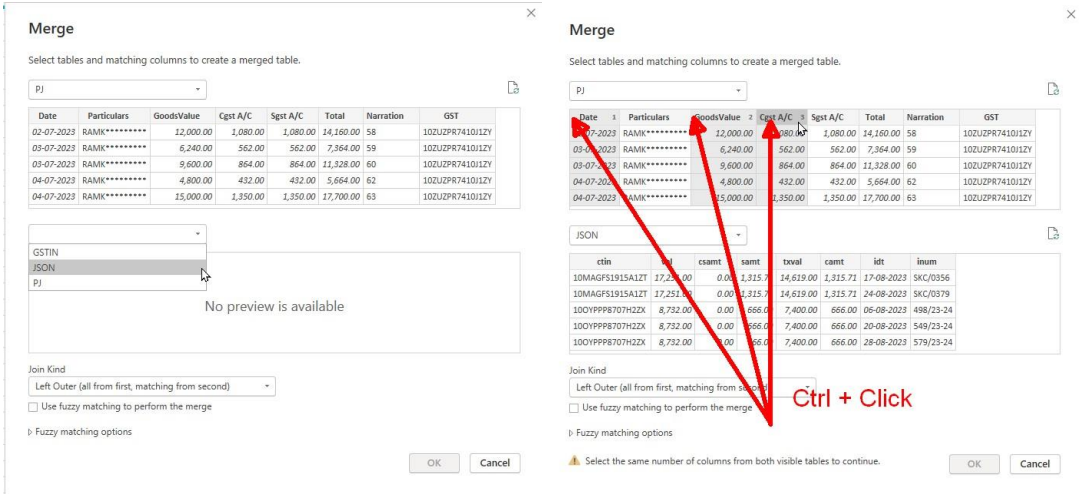
24. Now, the PJ Query will have all the required columns for reconciliation. Both the JSON and PJ queries are ready for reconciliation. You can see that there are 8 columns in both the JSON and PJ Queries. This client is not having IGST suppliers. So, there are only 7 common columns having IGST as extra column in JSON query and Particulars as extra column in PJ Query.



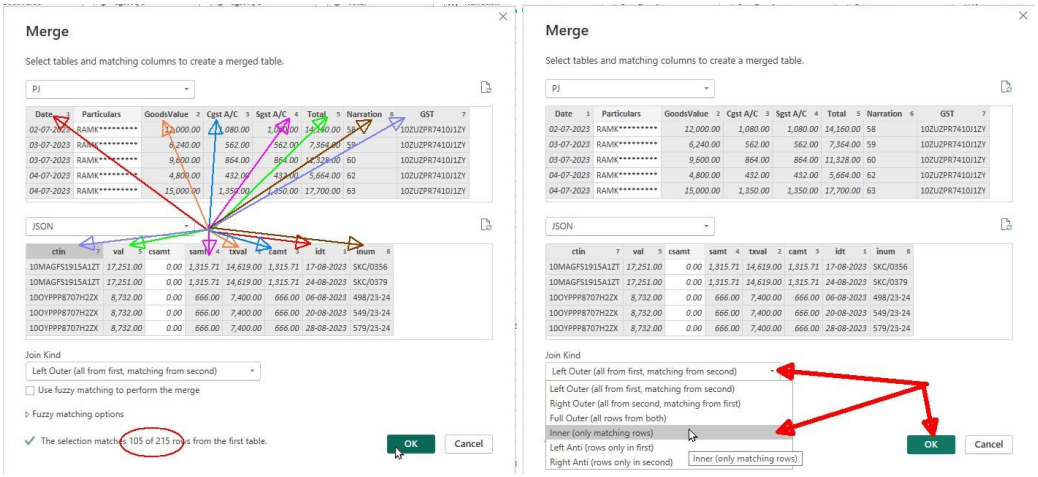
25. Select the PJ Query and Select Home Tab-> Merge Queries-> Merge Queries as New since we want to keep PJ and JSON queries in tact after merging.



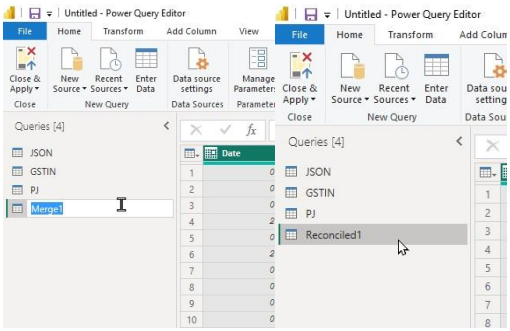
26. In the Merge window, select JSON in the bottom dropdown. We can select multiple common columns to match! It can be done by pressing Ctrl key and Click on the column headers. Note the the sequence numbers 1, 2, 3 … as we select the columns.



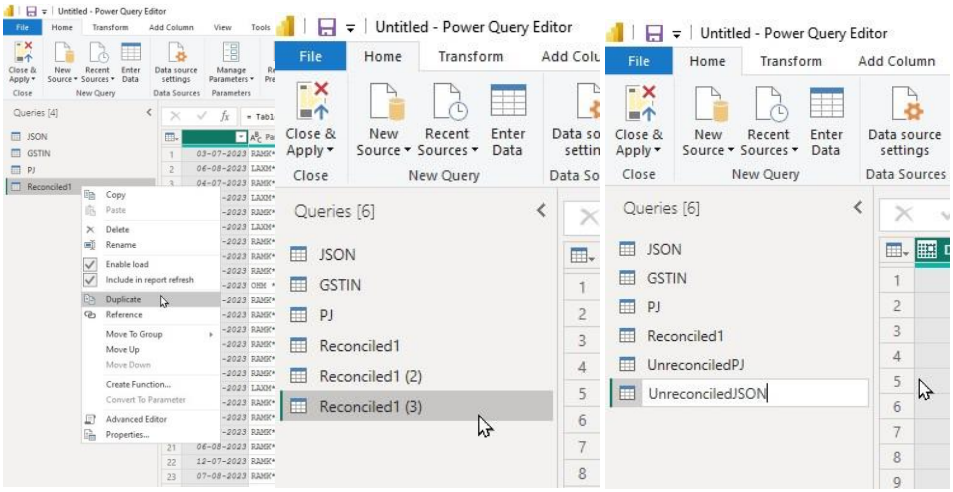
27. Select the relevant 7 common columns from both the queries by matching the selected column serial numbers 1 to 7 as given below. Note that straight away 105 records are matching in both the tables. Now Select Inner Join to have only reconciled records. Click Ok.



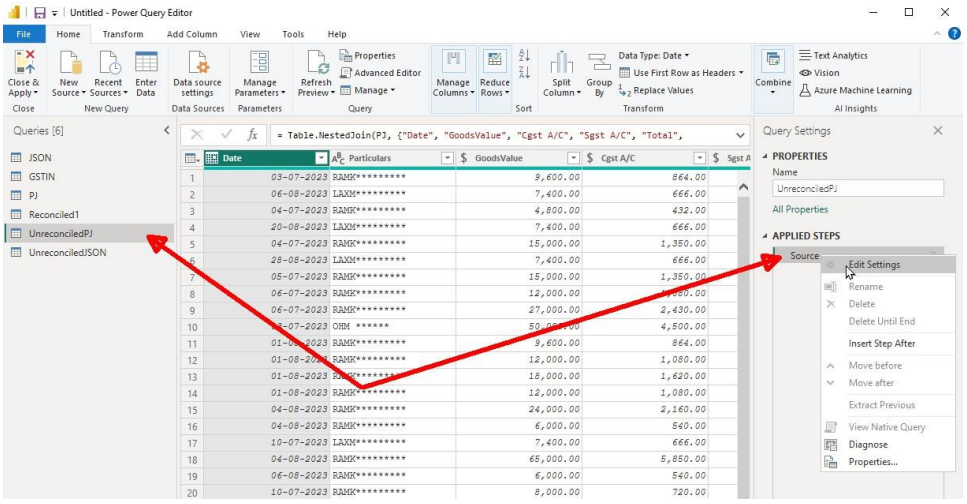
28. Rename the resultant query to Reconciled1



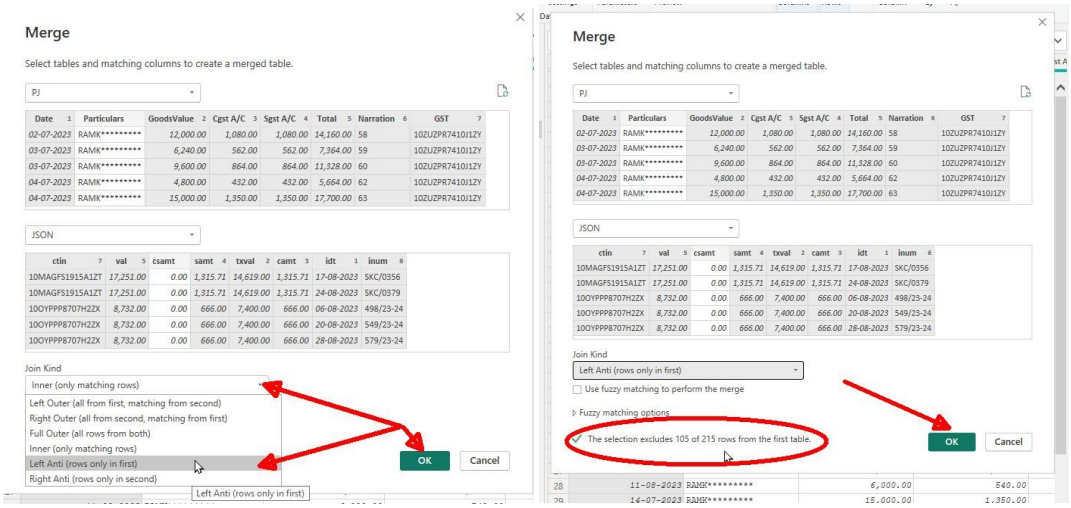
29. Right click on the Reconciled1 query and select Duplicate for two times to create two duplicate queries. This is to make UnreconciledPJ and UnreconciledJSON queries.



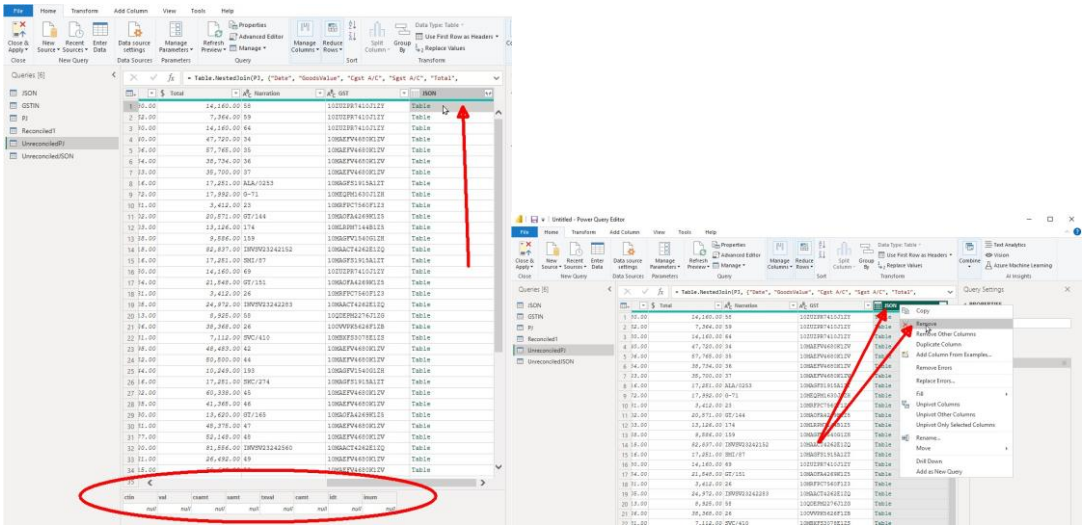
30. Select UnreconciledPJ query and right click on Source Step and select Edit Settings from Applied Steps area.



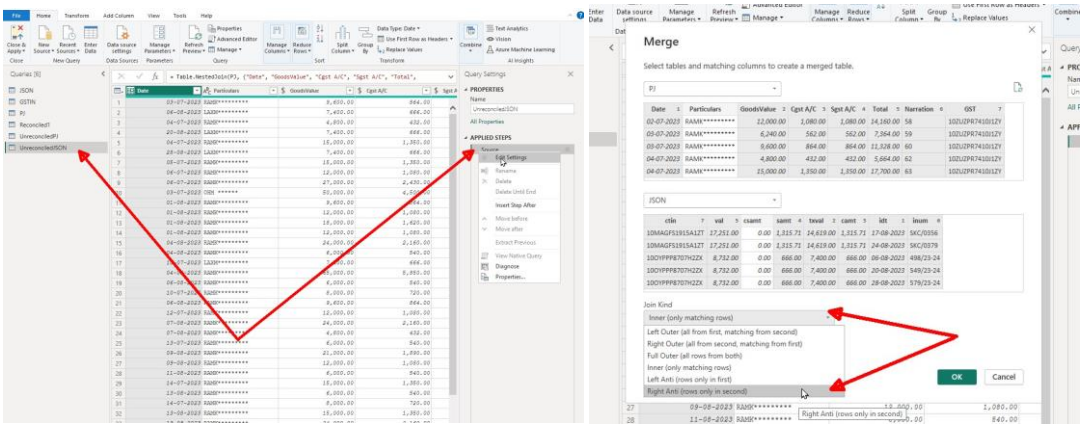
31. Change the Join type from Inner Join to Left Anti (for excluding the common records and keeping only those records available in PJ only)



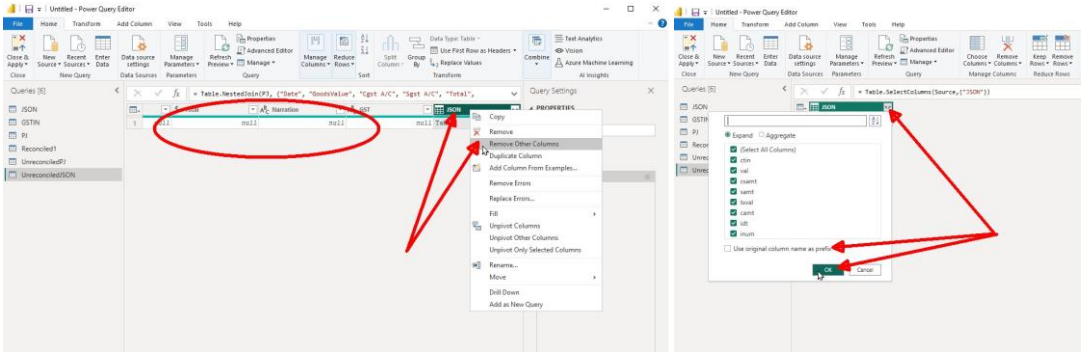
32. In the UnreconciledPJ query, we can see that there is a JSON merged query only with null values. So, right click on the JSON column header and select Remove Column.



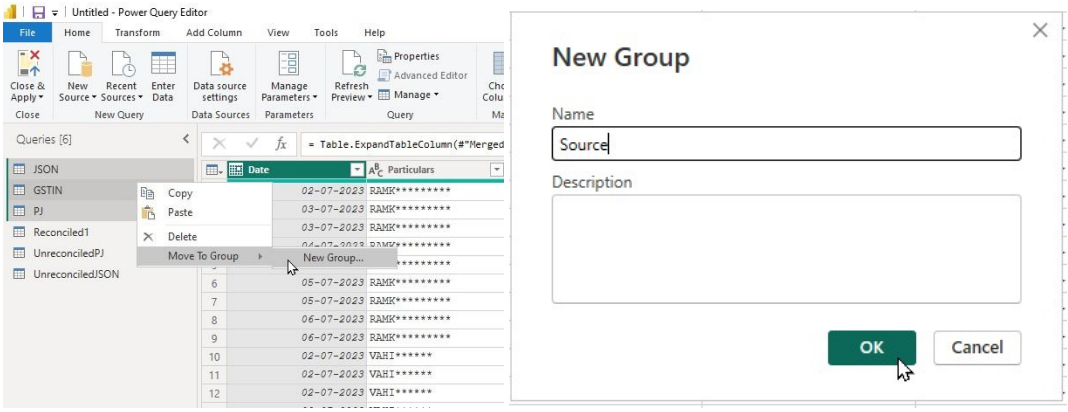
33. Same way, do it for UnreconciledJSON query with Right Anti Join.



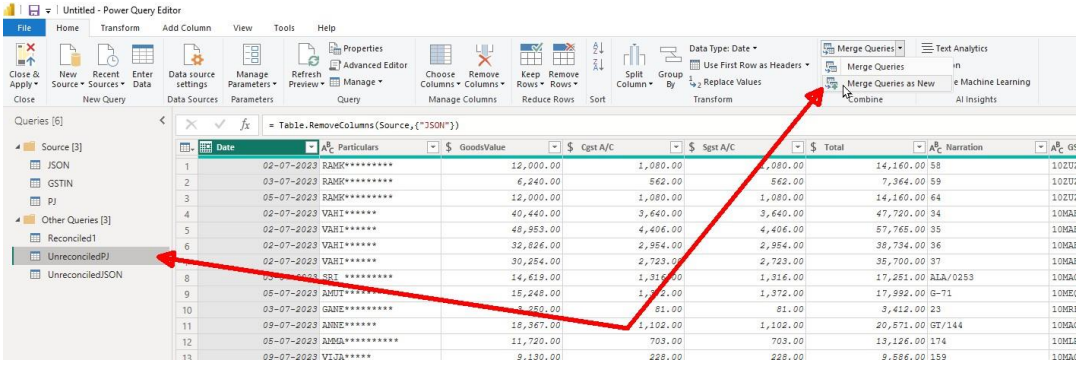
34. Now, in the merged query we can see that all the PJ Query columns are having one row of null values and a JSON column with a table. Right click JSON Column and Remove other Columns since they are related to PJ. Expand the JSON Column to get UnreconciledJSON data back.

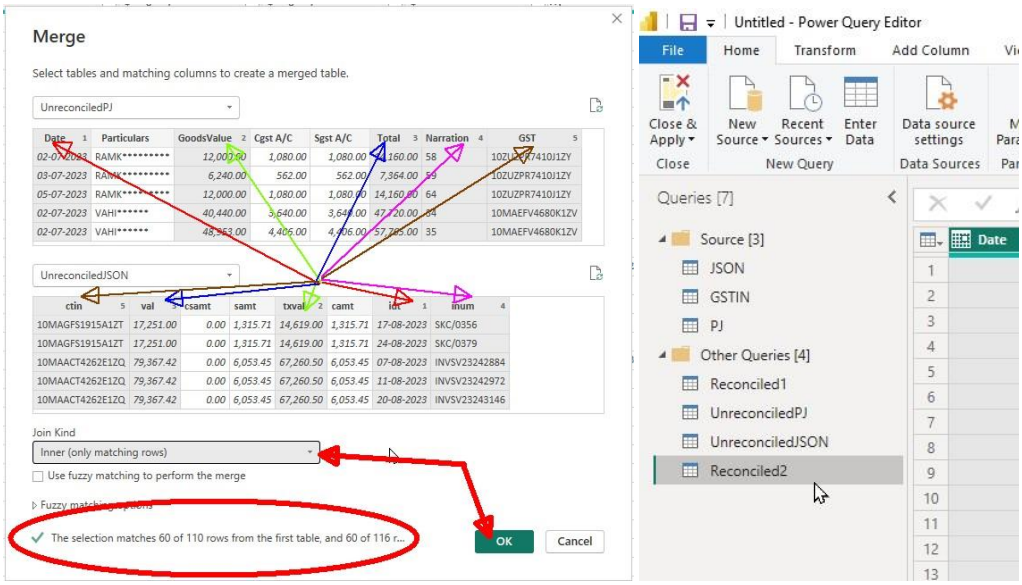


35. Select Source queries JSON, GSTIN and PJ, right click and Move them to a group called Source. This is to organize the queries.

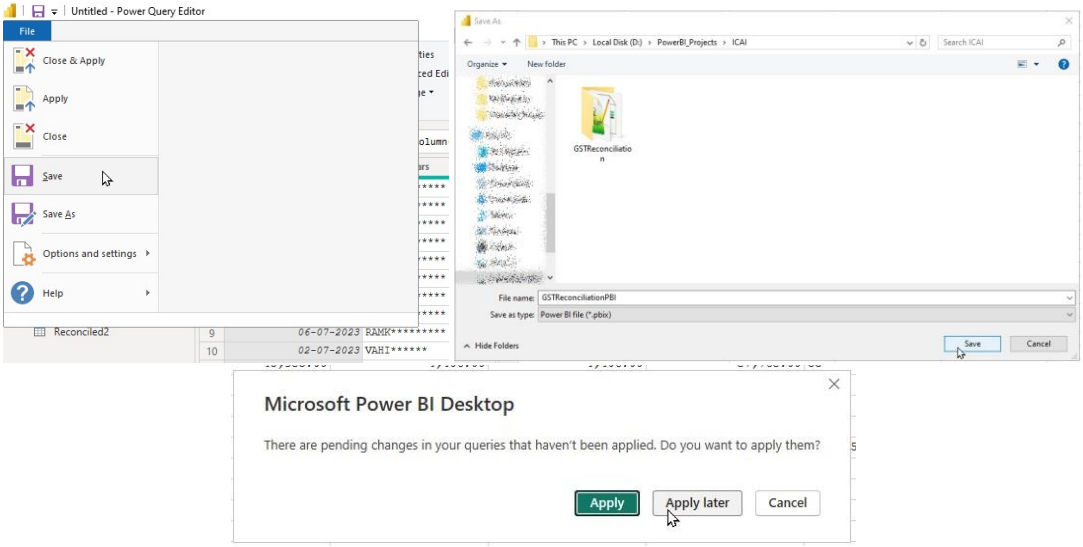


36. Now let us move to reconcile for next stage. We can see that many records are not reconciled because of Rounding Off of GST amounts. Hence, this time we can try with only 5 columns excluding CGST and SGST. Click on UnreconciledPJ and select again Merge Queries as New. Now we can see another 60 records are matching. Rename the query as Reconciled2. Like this we can go for further stages to reconcile.

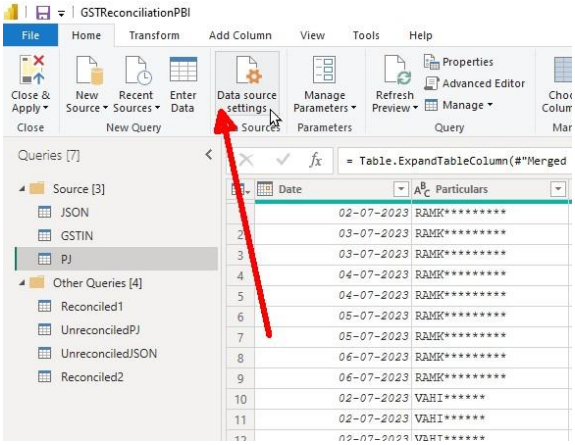




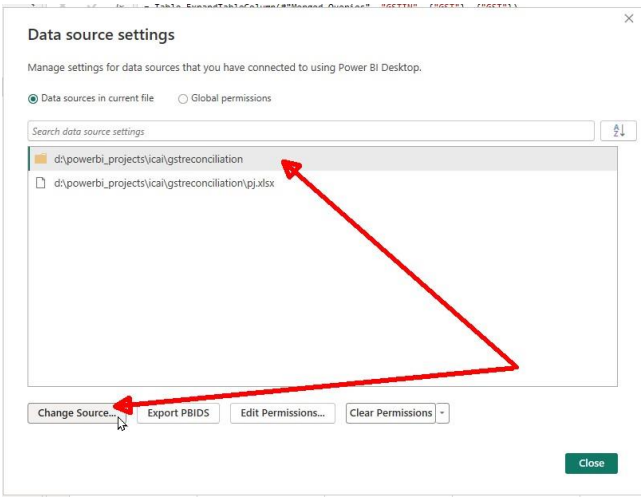
37. Save the file as a Power BI file. Select Apply Later.



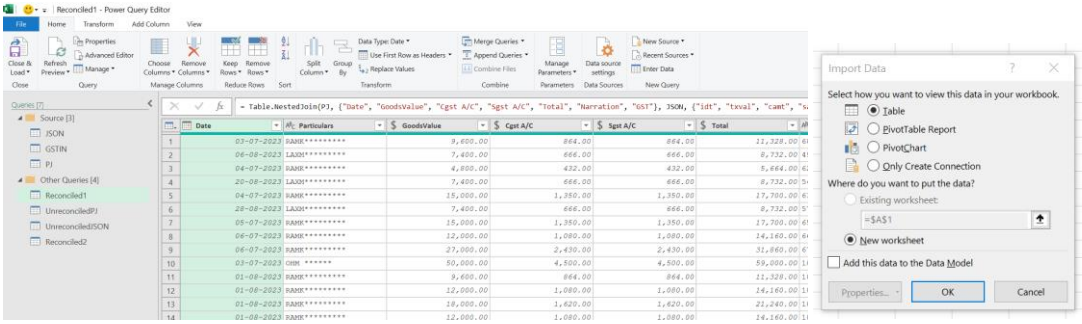
38. From the next month onwards or for another client we can change the data source by selecting the Home Tab -> Data Source Settings.



39. Click Change Source to change the source folder whenever needed.



40. These queries can be executed in Excel Power Query also. In Excel Power Query, we need to select File->Close & Load or Close & Load to for loading the queries as Excel Tables for further analysis. Select File->Close & Load to for loading only the selected queries and not every query.



From the above case study we can learn how Power Query can be used for reconciling GST records. From the next month, it is enough if we put the json files in the folder and update the PJ Excel file, automating the GST Reconciliation.