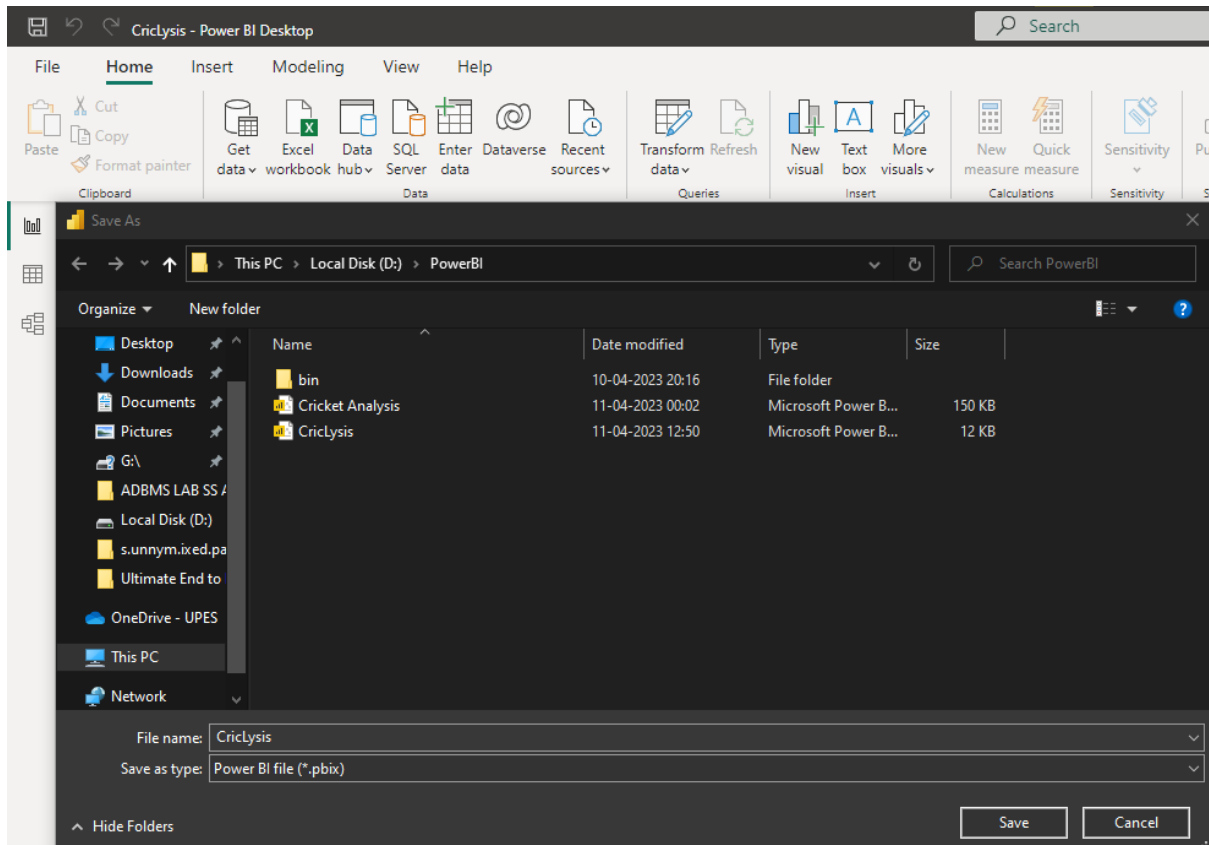
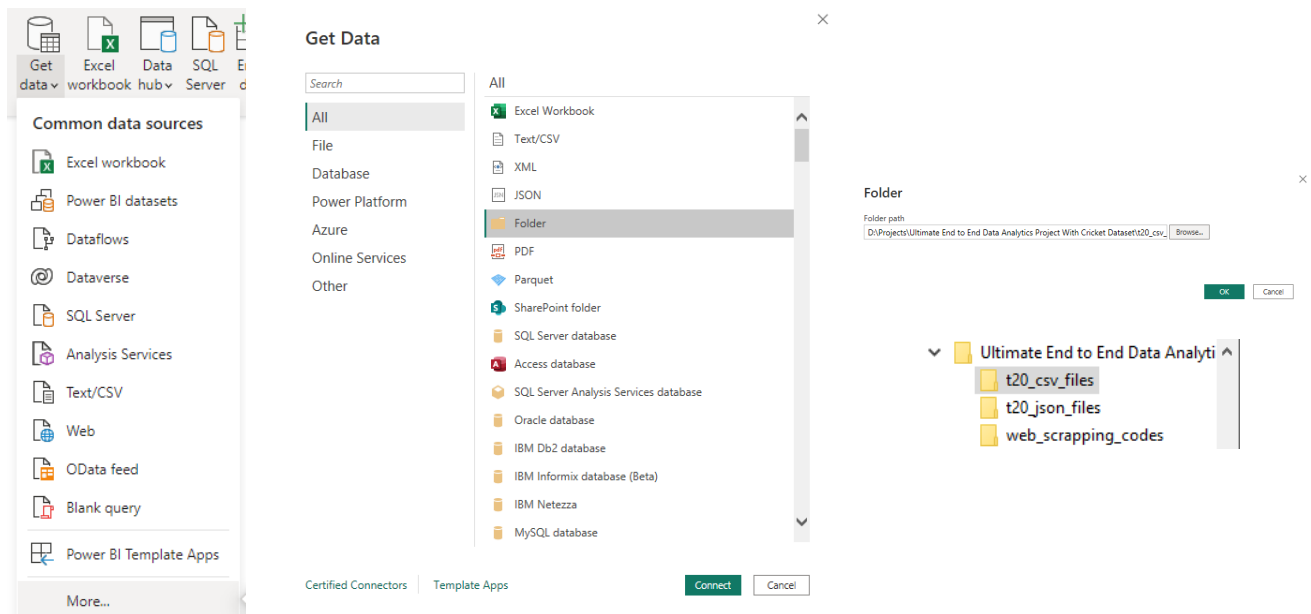


Starting with the PowerBI Desktop for Data Visualization

1. Open PowerBI Desktop and firstly save the untitled project by giving a name and location.

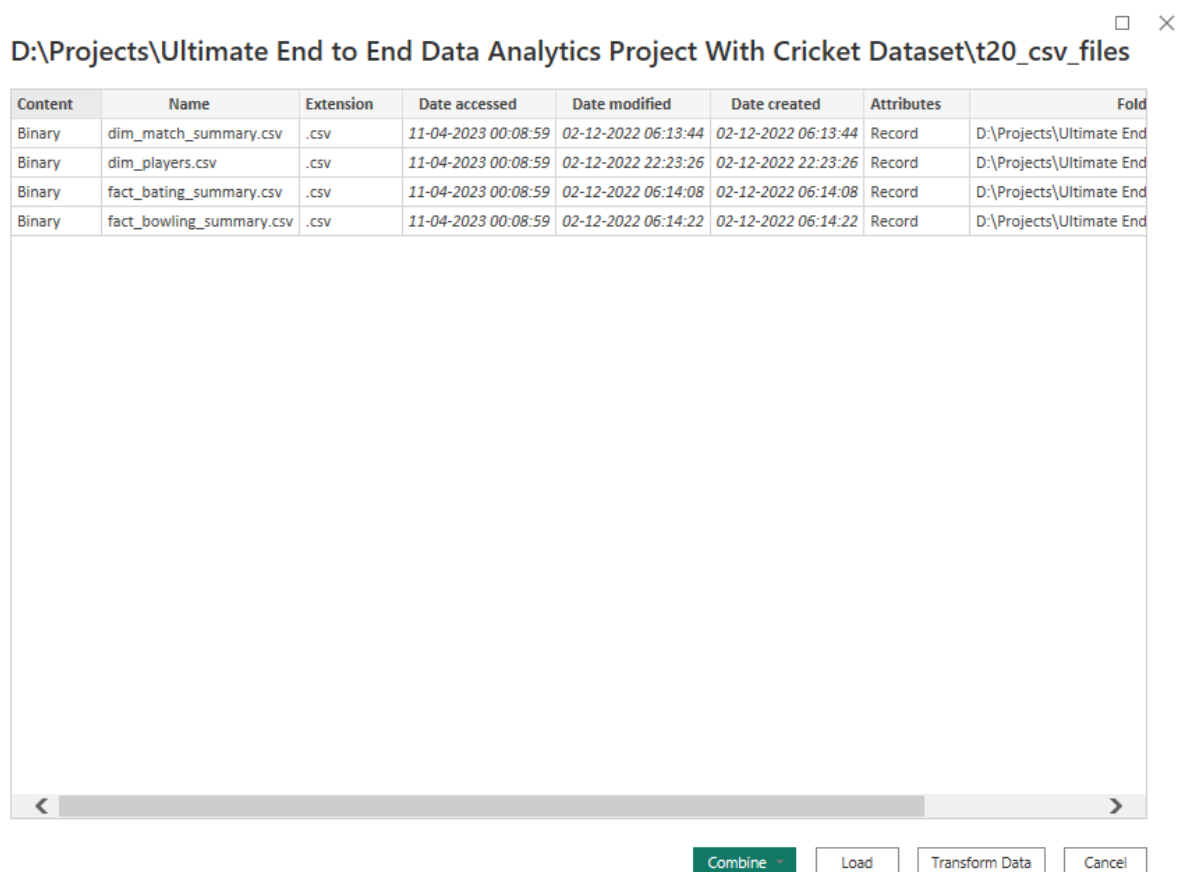


2. Load the data by going to Get Data > More > Folder.

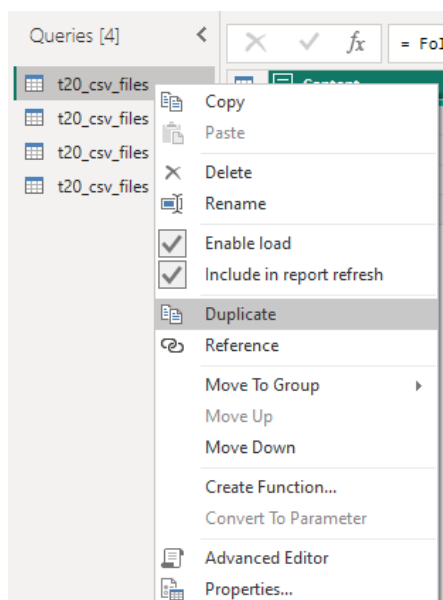


Select the folder with t20_csv_files and click on OK to load the data.

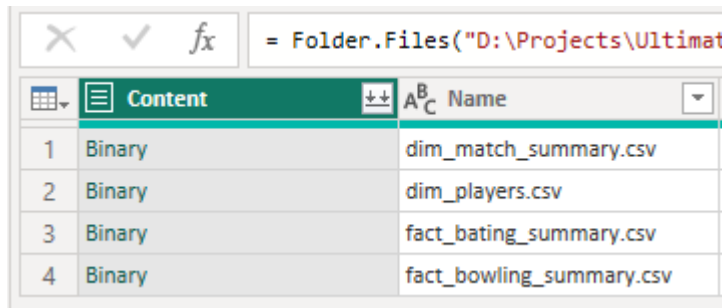
3. A dialog box will appear and click on Transform Data in the lower right corner.



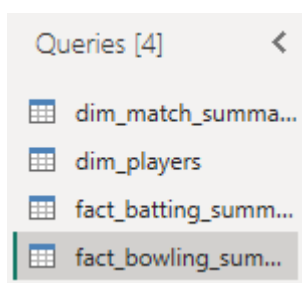
4. Duplicate the t20 csv files in the left side panel into 4 copies like this.



5. For each file on the left side panel, click on the binary tab and name the files according like dim_players, dim_match_summary like this.



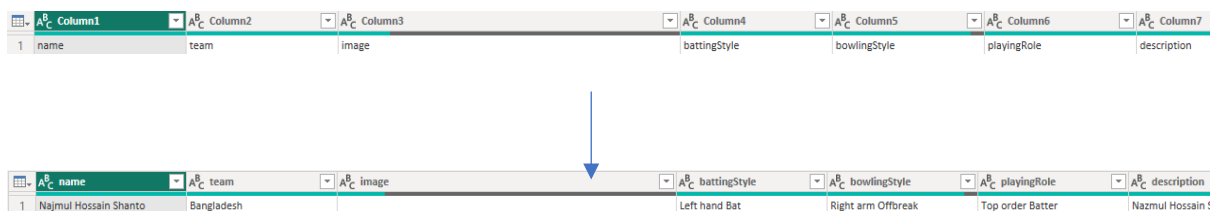
	Content	Name
1	Binary	dim_match_summary.csv
2	Binary	dim_players.csv
3	Binary	fact_bating_summary.csv
4	Binary	fact_bowling_summary.csv



Queries [4]
dim_match_summa...
dim_players
fact_bating_summ...
fact_bowling_sum...

6. Now we have to transform the data and clean it.

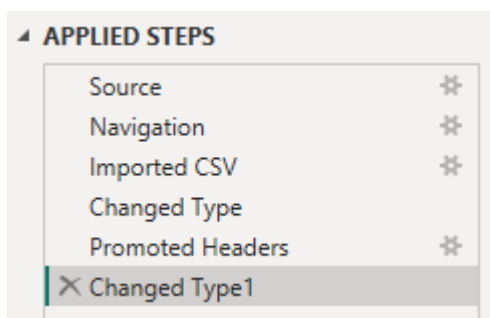
Firstly we go to the dim_players table and under the Transform tab, we select the Use First Row as Headers.



A _C Column1	A _C Column2	A _C Column3	A _C Column4	A _C Column5	A _C Column6	A _C Column7
1 name	team	image	battingStyle	bowlingStyle	playingRole	description

A _C name	A _C team	A _C image	A _C battingStyle	A _C bowlingStyle	A _C playingRole	A _C description
1 Najmul Hossain Shanto	Bangladesh		Left hand Bat	Right arm Offbreak	Top order Batter	Nazmul Hossain S

We can also view the changed log and applied steps in this view in the right side panel.

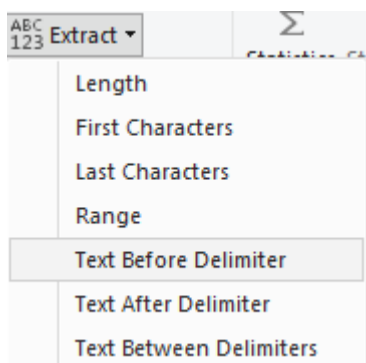


APPLIED STEPS	
Source	✱
Navigation	✱
Imported CSV	✱
Changed Type	
Promoted Headers	✱
✕ Changed Type1	

7. Next, under the name column, we can see that © after some names as they are the captain of respective teams. So we need to delete those as these are not useful.

	A ^B _C name		A ^B _C name
1	Najmul Hossain Shanto		1 Najmul Hossain Shanto
2	Soumya Sarkar		2 Soumya Sarkar
3	Litton Das		3 Litton Das
4	Shakib Al Hasan(c)		4 Shakib Al Hasan
5	Afif Hossain		5 Afif Hossain
6	Mosaddek Hossain		6 Mosaddek Hossain
7	Nurul Hasan		7 Nurul Hasan
8	Yasir Ali		8 Yasir Ali
9	Wessly Madhevere		9 Wessly Madhevere
10	Craig Ervine(c)		10 Craig Ervine

The process for removing is Select Extract > Text Before Delimiter > Write ' (' in the box of Delimiter and OK. It will remove all the unwanted © after the names.



Text Before Delimiter

Enter the delimiter that marks the end of what you would like to extract.

Delimiter

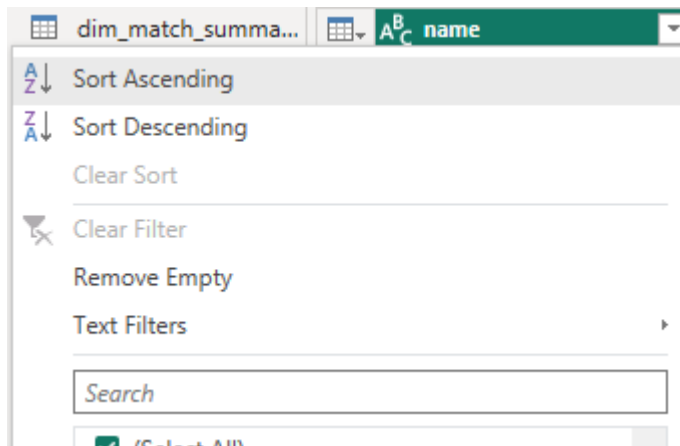
▸ Advanced options

OK

Cancel

8. Now, we have to remove all the duplicate values from the table.

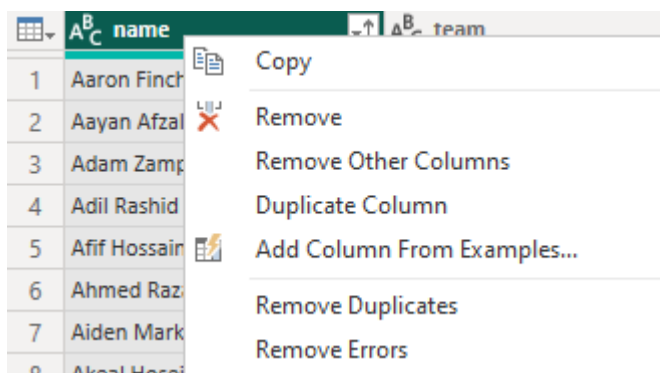
We can check the duplicate values by sorting the columns in ascending order.



Duplicate values like:

124	Matthew Wade	137	Mohammad Nabi
125	Matthew Wade	138	Mohammad Nabi

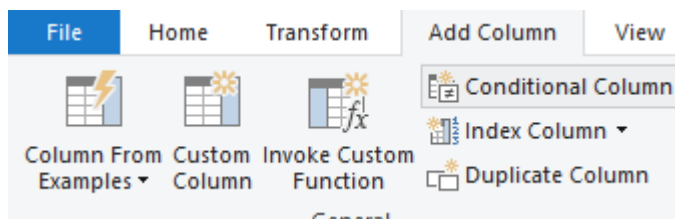
For removing duplicate values, right click on the name column, and select remove duplicate values.



So here, we can see that the rows gets decreased because of the removal of duplicate values.

9. So the dim players data cleaning is over now. We will now proceed towards the dim match summary table.

Here, we will first need to add a conditional column. We can do that by selecting conditional column under the Add column tab.



Our condition for this is like

If matchdate < 22-10-2022, then the column will have 'Qualifier' else if it is after 22-10-2022, then it will have 'Super 12'.

So we need to fill the data like this:

×

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

Stage

	Column Name	Operator	Value		Output
If	matchDate	is before or equal...	22-10-2022	Then	Qualifier

Add Clause

Else

Super 12

OK

Cancel

The output will be like this:

A ^B _C ground	matchDate	A ^B _C match_id	ABC 123 Stage
Geelong	16-10-2022	T20I # 1823	Qualifier
Geelong	16-10-2022	T20I # 1825	Qualifier
Hobart	17-10-2022	T20I # 1826	Qualifier
Hobart	17-10-2022	T20I # 1828	Qualifier
Geelong	18-10-2022	T20I # 1830	Qualifier
Geelong	18-10-2022	T20I # 1832	Qualifier
Hobart	19-10-2022	T20I # 1833	Qualifier
Hobart	19-10-2022	T20I # 1834	Qualifier
Geelong	20-10-2022	T20I # 1835	Qualifier
Geelong	20-10-2022	T20I # 1836	Qualifier
Hobart	21-10-2022	T20I # 1837	Qualifier
Hobart	21-10-2022	T20I # 1838	Qualifier
Sydney	22-10-2022	T20I # 1839	Super 12
Perth	22-10-2022	T20I # 1840	Super 12
Hobart	23-10-2022	T20I # 1841	Super 12
Melbourne	23-10-2022	T20I # 1842	Super 12
Hobart	24-10-2022	T20I # 1843	Super 12
Hobart	24-10-2022	T20I # 1844	Super 12
Perth	25-10-2022	T20I # 1845	Super 12
Melbourne	26-10-2022	T20I # 1846	Super 12
Melbourne	26-10-2022	T20I # 1846a	Super 12
Sydney	27-10-2022	T20I # 1847	Super 12
Sydney	27-10-2022	T20I # 1848	Super 12
Perth	27-10-2022	T20I # 1849	Super 12
Melbourne	28-10-2022	T20I # 1849a	Super 12
Melbourne	28-10-2022	T20I # 1849b	Super 12
Sydney	29-10-2022	T20I # 1850	Super 12
Brisbane	30-10-2022	T20I # 1851	Super 12


10. We have done the data cleaning in the dim_match_summary table. Now, we will move on to the fact_bowling_summary table.

Firstly we will need to change the column names of 4's and 6's to Words.

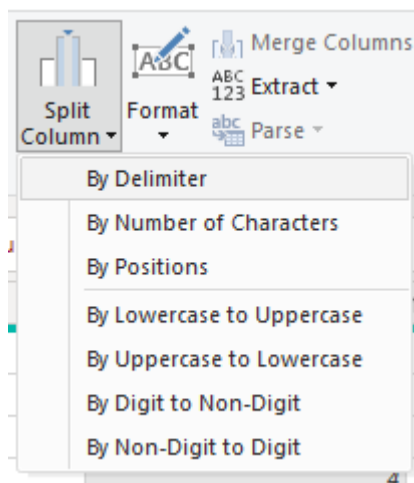
1 ² ₃ 4s	1 ² ₃ 6s
1 ² ₃ Fours	1 ² ₃ Sixes

11. Then we will need to split the column 'overs' into overs and balls.

1.2 overs
4
4
4
4
4
1
4
4
3
3
4
4
2
3
4
3
4
1
3
4
4
2.5
1
2
4



Select the column > Under transform tab, select Split Column > By Delimiter



Fill the dialog box like this :

Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter

--Custom--
.

Split at

- ☐ Left-most delimiter
☐ Right-most delimiter
☒ Each occurrence of the delimiter

▸ Advanced options

Quote Character

"

☐ Split using special characters

Insert special character

OK

Cancel

12. The table will be looking like this:

t23 overs	t23 balls
4	null
4	null
4	null
4	null
4	null
1	null
4	null
4	null
3	null
3	null
4	null
4	null
4	null
2	null
3	null
4	null
3	null
4	null
1	null
3	null
4	null
4	null
2	5
1	null
2	null
4	null
4	null
3	null
3	null
4	null
4	null
4	null
3	null
3	3
4	null

We need to replace all null values with ‘0’ in the ‘balls’ column.

Replace Values

Replace one value with another in the selected columns.

Value To Find

null

Replace With

0

Advanced options

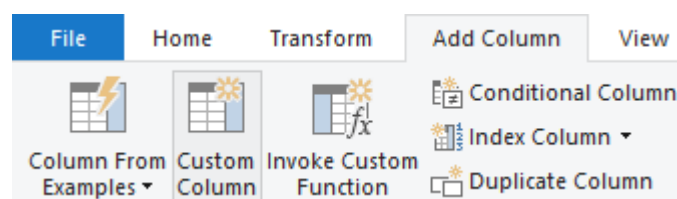
OK

Cancel

The column will look like this :

0
5
0
0
0
0
0
0
0
0
0
0
0
0
0
3

**13. We need to add a column ‘total balls’ by the condition:
Overs * 6 + balls.**



Custom Column

Add a column that is computed from the other columns.

New column name

Total Balls

Custom column formula ⓘ

= [overs]*6 + [balls]

Available columns

match
bowlingTeam
bowlerName
overs
balls
maiden
runs

<< Insert

[Learn about Power Query formulas](#)

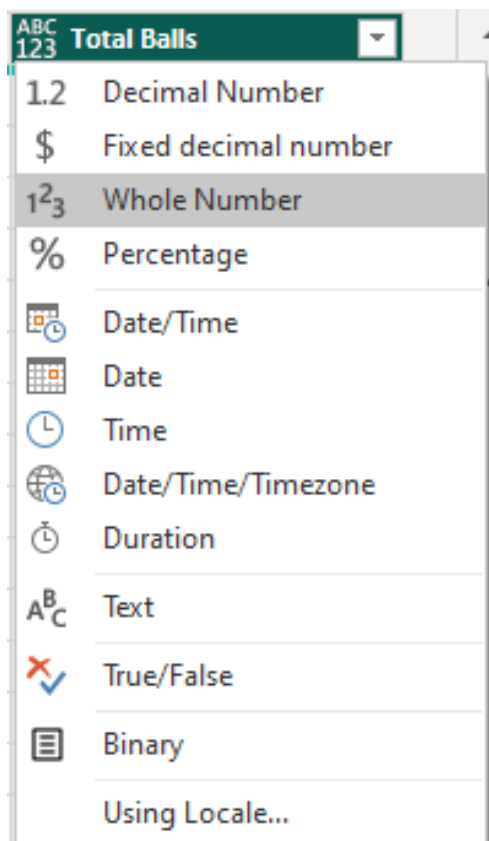
✓ No syntax errors have been detected.

OK

Cancel

ABC 123	Total Balls	
	24	
	24	
	24	
	24	
	24	
	6	
	24	
	24	
	18	
	18	
	24	
	24	
	24	
	12	
	18	
	24	
	18	
	24	
	6	

14. We need to change the datatype of the column to Whole Number.



15. Next, we move to our last table fact batting summary.
We need to do the Step.10 here again on the columns ‘4’
and ‘6’.

1^2_3 Fours	1^2_3 Sixes
0	0
1	0
1	2
2	0
0	0

16. We now need to change the column of ‘out/not out’ to whole number by substituting 0 for ‘out’ and 1 for ‘not out’.

Select the column and select Replace Values. Also change the datatype of the column to Whole Number.

Replace Values

Replace one value with another in the selected columns.

Value To Find

out

Replace With

0


▷ Advanced options

OK

Cancel

[illegible]

17. Repeat Step.7 again on this table for removing © from the names table.



A ^B _C batsmanName
Michael van Lingen
Divan la Cock
Jan Nicol Loftie-Eaton
Stephan Baard
Gerhard Erasmus(c)
Jan Frylinck
David Wiese
JJ Smit
Pathum Nissanka
Kusal Mendis
Dhananjaya de Silva
Danushka Gunathilaka
Bhanuka Rajapaksa
Dasun Shanaka(c)
Wanindu Hasaranga de Silva
Chamika Karunaratne
Pramod Madushan

A ^B _C batsmanName
Michael van Lingen
Divan la Cock
Jan Nicol Loftie-Eaton
Stephan Baard
Gerhard Erasmus
Jan Frylinck
David Wiese
JJ Smit
Pathum Nissanka
Kusal Mendis
Dhananjaya de Silva
Danushka Gunathilaka
Bhanuka Rajapaksa
Dasun Shanaka
Wanindu Hasaranga de Silva
Chamika Karunaratne
Pramod Madushan

18. Add a custom column by the condition with name = Boundary Runs.

Boundary runs = fours*4 + sixes*6

Custom Column

Add a column that is computed from the other columns.

New column name

Boundary Runs

Custom column formula ⓘ

= [Fours]*4 + [Sixes]*6

[Learn about Power Query formulas](#)

Available columns

battingPos
batsmanName
runs
balls
Fours
Sixes
SR

<< Insert

✓ No syntax errors have been detected.

OK

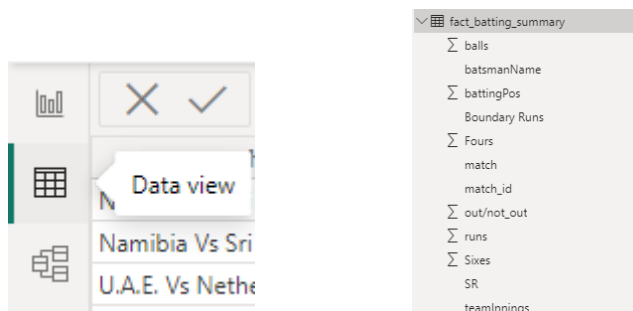
Cancel

19. We are now over with the cleaning and transformation of the data. Next, we will move to creating the DAX Measures. The excel file is provided with the project so that you can refer to the file and use it as per your standards.

So I will create a measure for the first one and you can create the rest as per the same steps.

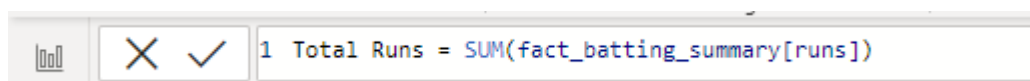
Measures	Description / Purpose	DAX FORMULA	TABLE
1 Total Runs	Total number of runs scored by the batsman	Total Runs = SUM(fact_batting_summary[runs])	fact_batting_summary

Go to the data view and select the table mentioned in the sheet.

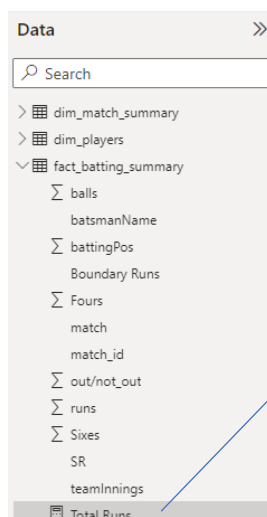


Then select New Measure under Table tools.

Paste the measure given in the sheet on the tab like this.



Hit Enter and that's it. In this way, create all the measures.



The measure created is shown here.

20. Next, you need to apply your PowerBI Design knowledge and create your own dashboard using various tools available.

Thank You