ESPN Cricket Dataset Analysis

Data source: Web

Link:

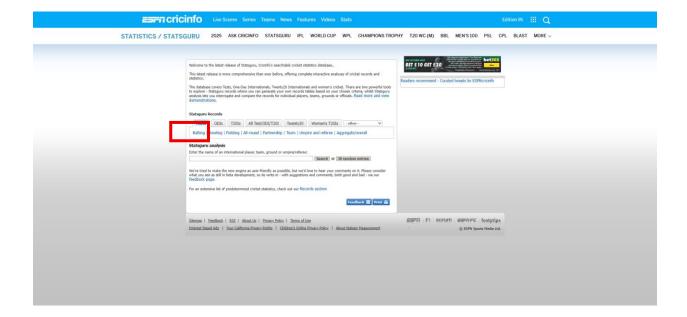
https://stats.espncricinfo.com/ci/engine/stats/index.html

Aim:

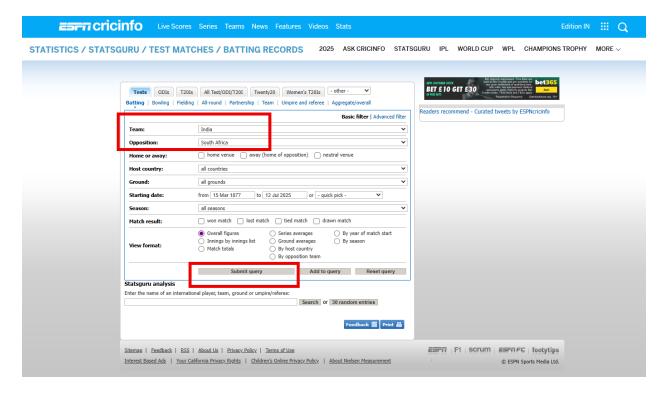
An interactive Power BI dashboard that allows stakeholders to explore **player-level insights** into India vs South Africa matches across batting, bowling, and fielding domains using scraped ESPN Cricinfo data.

Process:

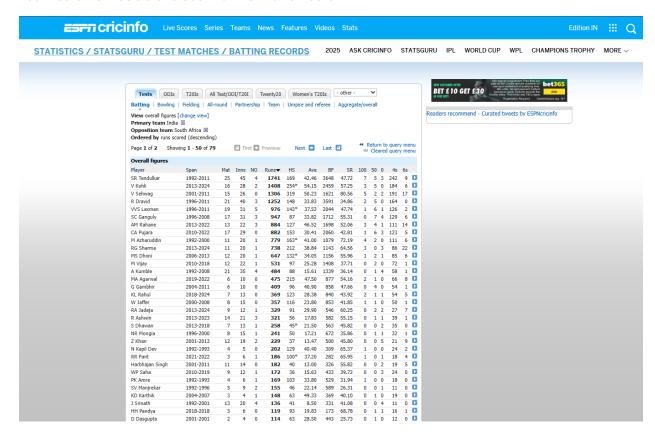
When we opened the link, the landing page would look like -



Click on Batting and enter the details and click on submit query



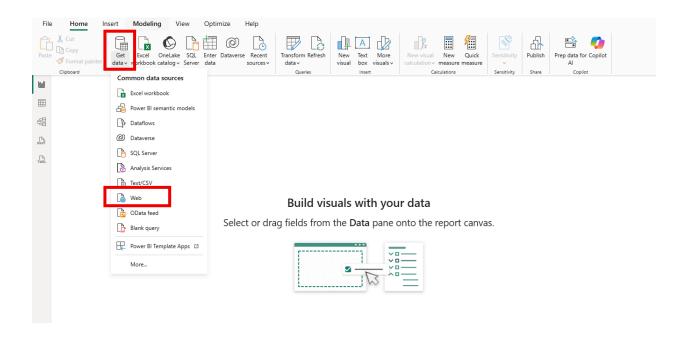
You would now be able to see the information below:



Copy the link to this page:

https://stats.espncricinfo.com/ci/engine/stats/index.html?class=1;opposition=3;team=6;template =results;type=batting

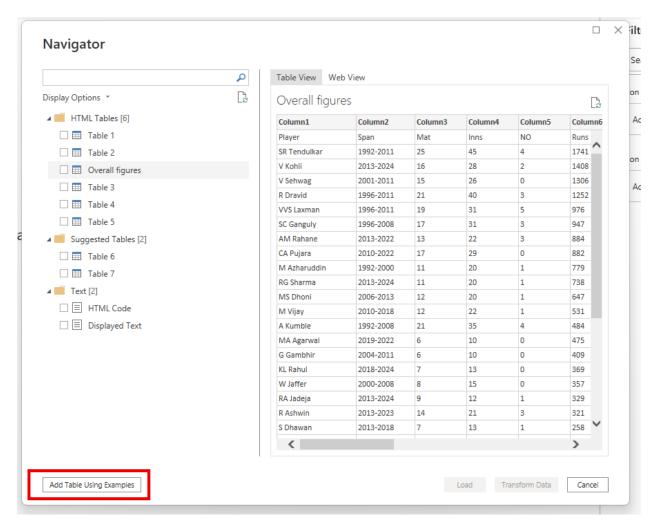
Open PowerBI deskstop and click on get data and then web



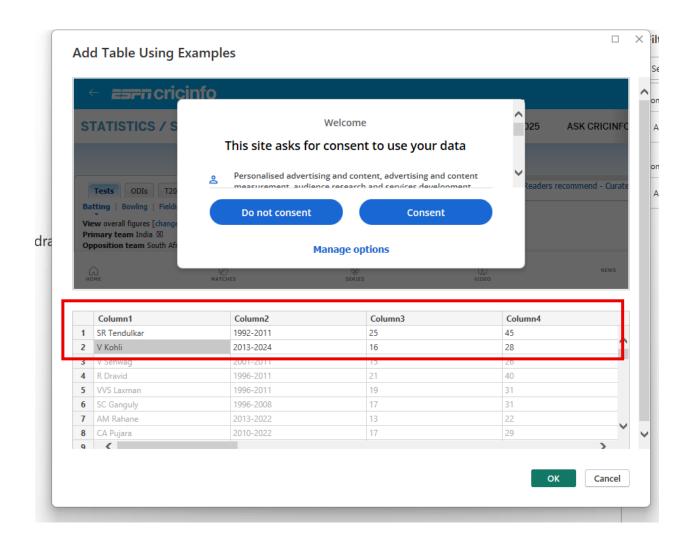
Enter the page URL and just after opposition write 'page=1' and click ok



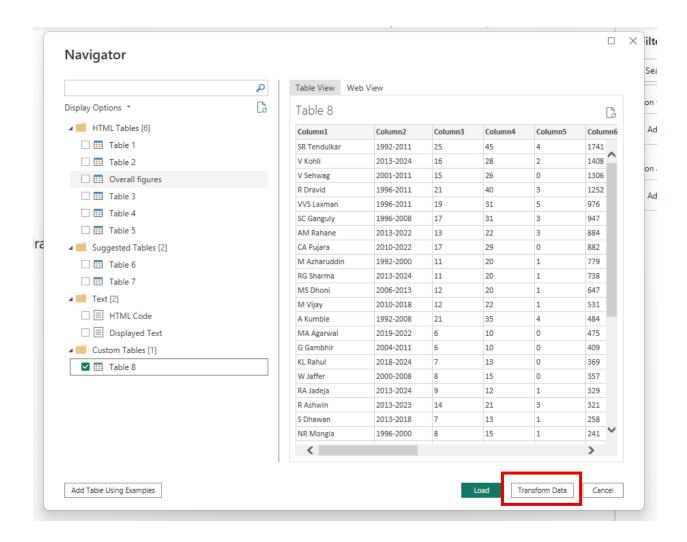
In the Navigator click Add tables using examples



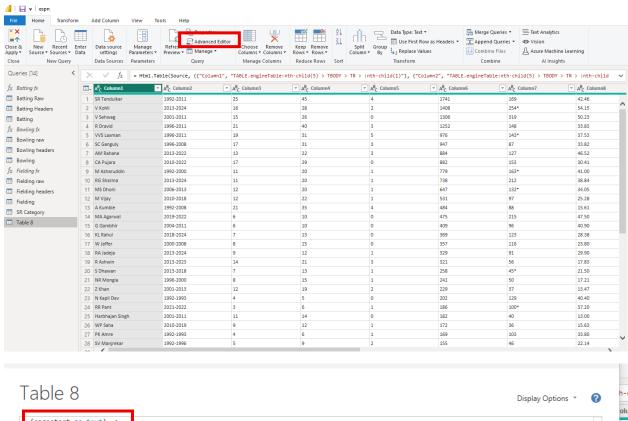
From the main website page, copy the first two rows and paste them into the box below, you would observe that PowerBI is already showing you the suggestions. Click ok to load the table.

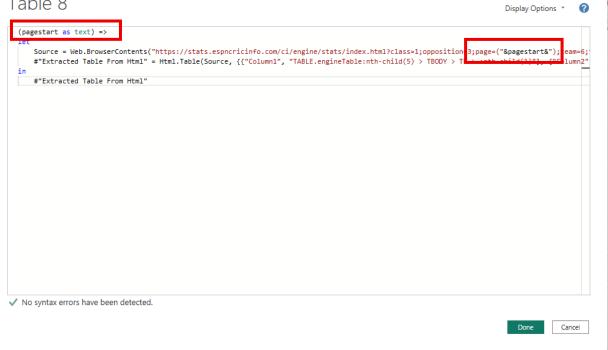


Click on Transform Data to open the power query editor



Since this has only loaded the page 1, and we have 3 pages in the original query we did on the website, therefore we need to convert this table to a function. For this, click on Advanced Editor and enter the pagestart variable .

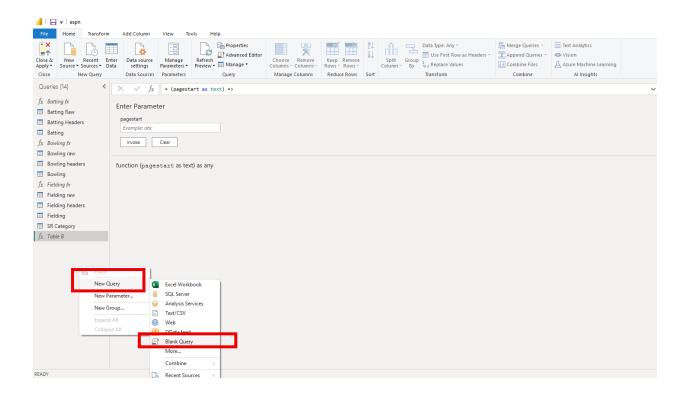




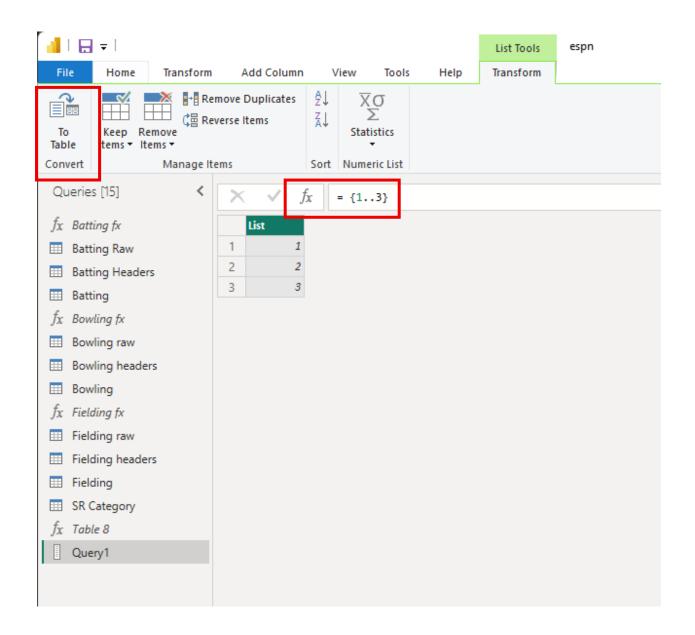
 f_X Table 8

You would now observe that the icon has changed.

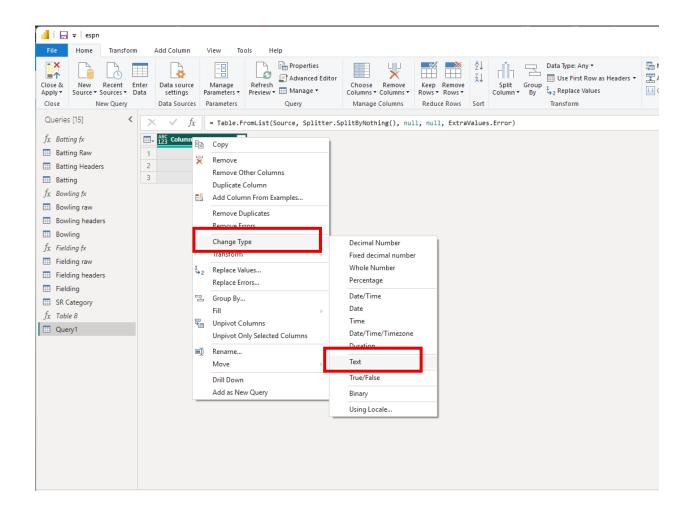
Now we will create a list from 1 to 3 and convert it to a table so that we can use this list to scrap all the pages from 1 to 3. For this right click and click on New query then blank query.



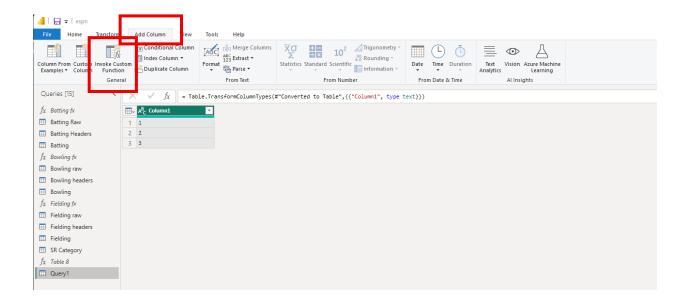
Now click on function tab and write ={1..3} and press enter, this will give you a list. Then on the top left side click on 'to table'. This will give you a table.



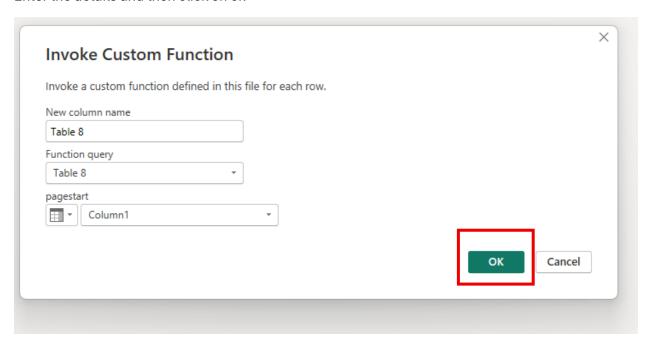
Since our 'pagestart' variable is a text, right click on the column name and change the datatype to text and press enter.



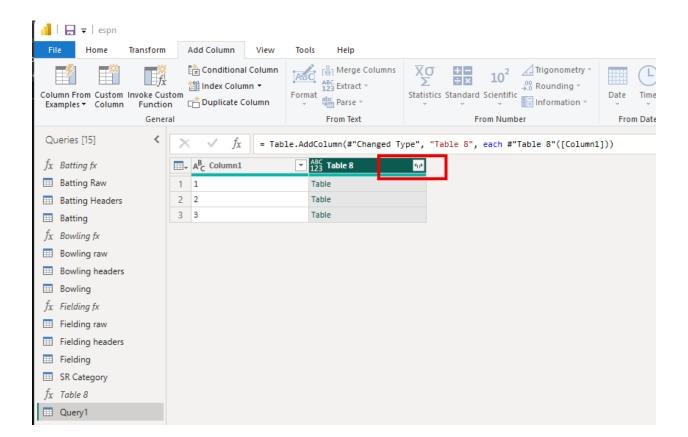
Now next step is to invoke the function. Click on 'Add column' and then 'Invoke Custom Function'



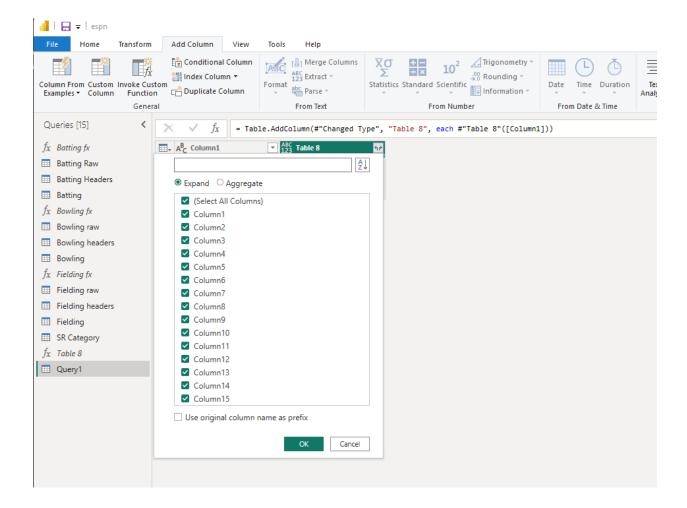
Enter the details and then click on ok



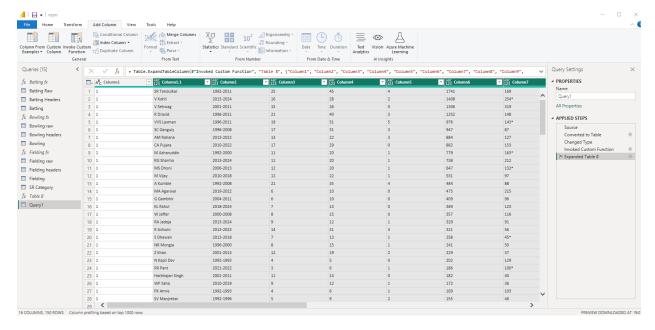
Scrapping the data will take some time. Once it is loaded click on the expand button



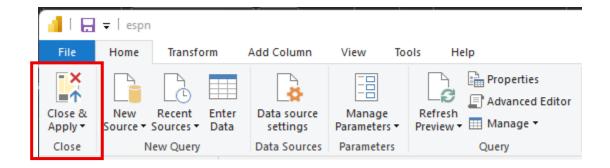
Choose all the columns and click ok



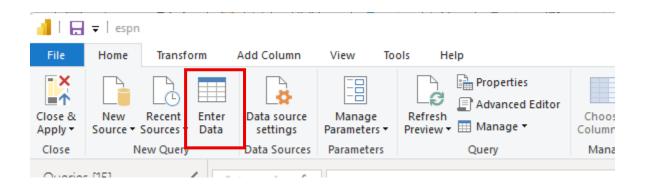
Now remove the Column 1 and rename Column 1.1 as Column 1 (both can be done by right clicking on the column name and choosing the appropriate option).

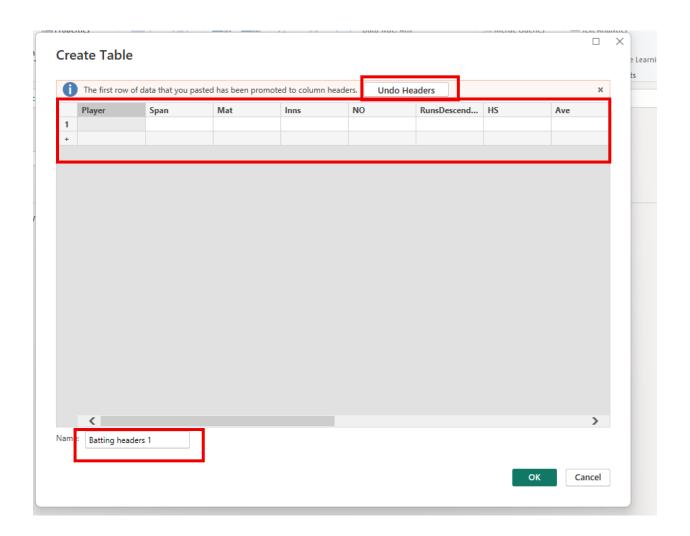


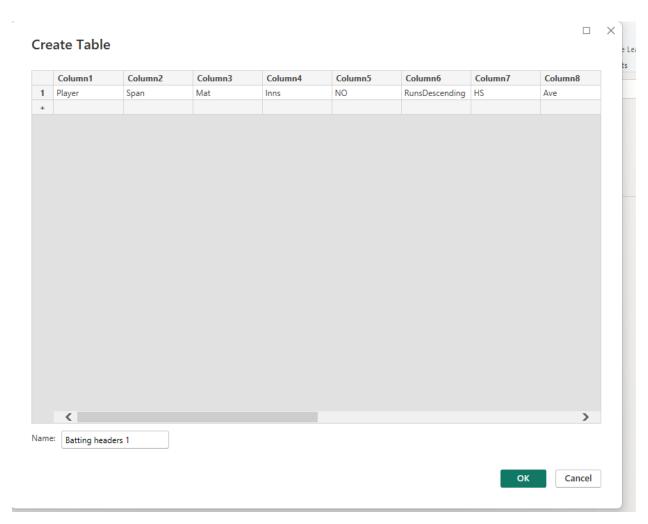
Once the changes are done – Go to Home and click on 'Close and Apply'



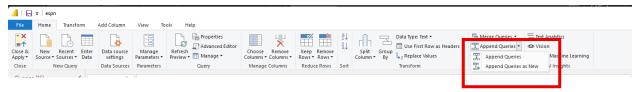
In order to get the headers for the colums, click on enter data from the Home ribbon and then copy the actual headers from the main website page and paste them on the create table page so popped up. You need to click on 'Undo Header Options'. You can also rename this table as 'Batting Headers 1' from the bottom of the page and click ok.



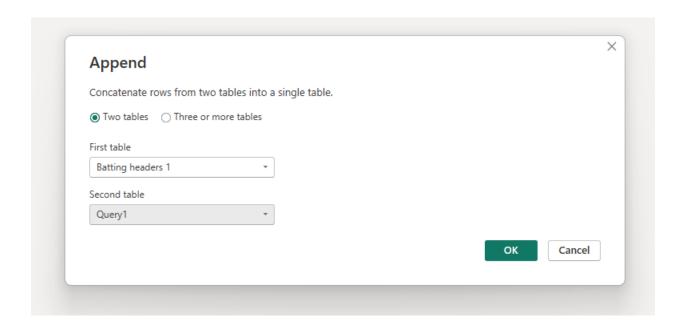




Now, we append the two tables – table 8 and Batting Headers 1 and make the headers. Click on Append queries and then as new.



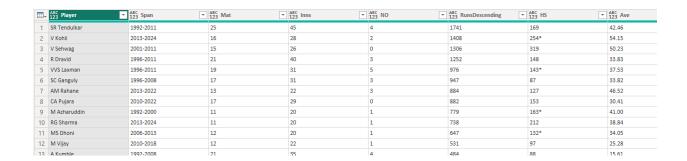
Make the appropriate selection and click ok



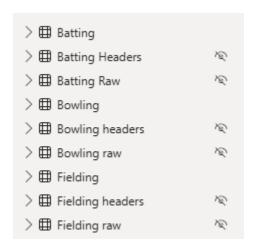
Now click on 'Use first row as header'.



We will now get our final dataset and now you can do the basic data cleaning such as changing data types to whole number and decimals wherever required, replacing '-' with 0 etc.



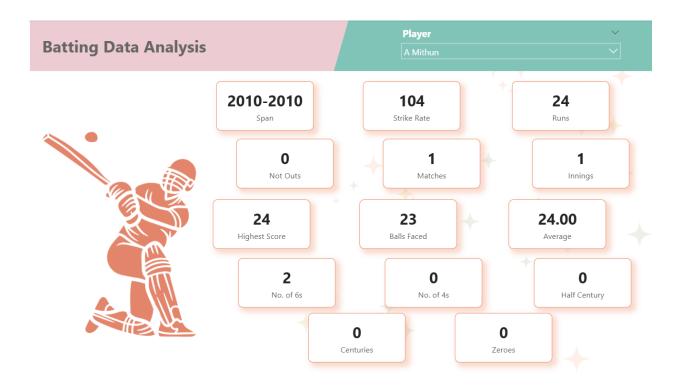
Repeat the same steps to capture data for bowling and fielding as well. You should now have three tables each for batting, bowling and fielding.



Now that our datasets are ready , we will make the reports. Our report will have 3 pages , 1 for each batting, bowling and fielding.

Go to the report page and create the visual of your choice

Report design



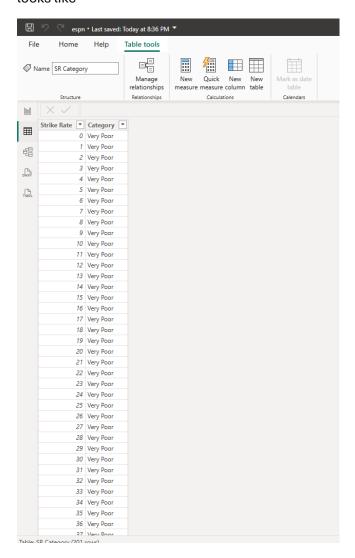




DAX Functions

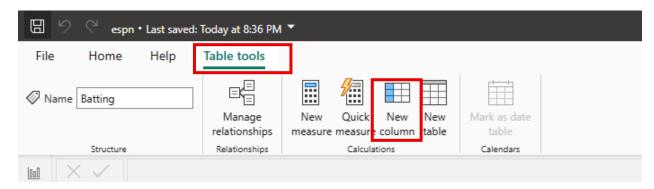
Now we aim to categorise the players on the basis of strike rates. Therefore we made a strike rate category sheet which had categories for each value of strike rate (which is from 0 to 200 and can be

seen as minimum and maximum values of the strike rate column (SR) in batting table). The data type of SR column in batting table should be whole numbers. After loading the SR Category table looks like



Once the table is loaded, now we can use Lookup DAX function.

Click on the batting table and click on 'Table Tools' → 'New column'



In the function bar, enter the following DAX formula:

Lookupvalue

SR Cat = LOOKUPVALUE('SR Category'[Category],'SR Category'[Strike Rate],Batting[SR])



rankX

To enter rank, add new a new column and enter the function as:

Rank = RANKX(ALL(Batting), Batting[SR],, DESC, Dense)

Average

For deviation in runs:

Dev Runs = Batting[RunsDescending]-AVERAGE(Batting[RunsDescending])

Abs

For absolute deviation

abs dev = ABS(Batting[Dev Runs])

Power

For squared deviation

squared dev = POWER(Batting[abs dev],2)

The resultant table would look like:

Player	*	Span 💌	Mat 💌	Inns 💌	NO 💌	RunsDescending •	HS ▼	Ave 💌	BF 💌	SR 💌	100 🔻	50 🔻	0 🔻	4s 🔻	6s 💌	SR Cat	Dev Runs	abs dev ▼	squared dev	rank 💌
KL Rahul		2019-2023	8	7	0	187	56	26.71	264	71	0	2	0	17	0	Poor	21.7478991596639	21.7478991596639	472.971117858908	37
RJ Shastri		1991-1992	6	6	2	179	109	44.75	226	79	1	0	1	17	0	Average	13.7478991596639	13.7478991596639	189.004731304286	31
G Gambhir		2003-2011	10	10	0	155	69	15.5	247	63	0	1	2	21	0	Poor	-10.2521008403361	10.2521008403361	105.105571640421	44
A Kumble		1992-2006	40	28	7	138	16	6.57	229	60	0	0	5	7	0	Poor	-27.2521008403361	27.2521008403361	742.67700021185	47
D Mongia		2003-2006	5	4	1	126	55*	42	182	69	0	1	0	9	0	Poor	-39.2521008403361	39.2521008403361	1540.72742037992	39
K Srikkanth		1991-1992	3	3	0	121	68	40.33	152	80	0	2	1	17	0	Average	-44.2521008403361	44.2521008403361	1958.24842878328	30
NR Mongia		1995-1999	11	10	3	110	24	15.71	150	73	0	0	0	3	0	Poor	-55.2521008403361	55.2521008403361	3052.79464727067	36
M Prabhakar		1991-1995	14	10	2	105	36	13.12	216	49	0	0	0	5	0	Very Poor	-60.2521008403361	60.2521008403361	3630.31565567404	55
MSK Prasad		1999-1999	2	1	0	63	63	63	90	70	0	1	0	6	0	Poor	-102.252100840336	102.252100840336	10455.4921262623	38
B Kumar		2013-2022	16	10	4	60	19*	10	103	58	0	0	2	4	0	Poor	-105.252100840336	105.252100840336	11078.0047313043	49
PA Patel		2003-2011	5	3	0	55	38	18.33	64	86	0	0	0	9	0	Average	-110.252100840336	110.252100840336	12155.5257397076	25
S Ramesh		1999-1999	2	2	0	44	36	22	51	86	0	0	0	7	0	Average	-121.252100840336	121.252100840336	14702.071958195	25
M Vijay		2010-2011	4	4	0	43	25	10.75	59	73	0	0	0	8	0	Poor	-122.252100840336	122.252100840336	14945.5761598757	36
KS More		1991-1992	7	5	2	42	32	14	90	47	0	0	0	4	0	Very Poor	-123.252100840336	123.252100840336	15191.0803615564	56
R Vijay Bharadwaj	j	1999-1999	2	2	1	42	24	42	53	79	0	0	0	6	0	Average	-123.252100840336	123.252100840336	15191.0803615564	31
V Yadav		1992-1993	5	4	1	42	34*	14	38	111	0	0	0	2	0	Good	-123.252100840336	123.252100840336	15191.0803615564	9
BKV Prasad		1995-2001	20	11	9	30	10*	15	43	70	0	0	1	2	0	Poor	-135.252100840336	135.252100840336	18293.1307817245	38
RD Gaikwad		2022-2023	3	3	0	28	19	9.33	54	52	0	0	0	3	0	Poor	-137.252100840336	137.252100840336	18838.1391850858	53
VVS Laxman		2001-2006	3	3	0	27	22	9	44	61	0	0	1	3	0	Poor	-138.252100840336	138.252100840336	19113.6433867665	46
JJ Bumrah		2017-2023	12	2	1	26	14*	26	38	68	0	0	0	3	0	Poor	-139.252100840336	139.252100840336	19391.1475884471	40
SS Dinha		2000-2000	2	9	1	22	17*	22	26	25	0	0	0	0	n	Averane	-1/12 2521008/0226	1/12 2521008/0226	20521 16/2051608	26