Resources: Spreadsheet Analyses [Sheet](https://docs.google.com/spreadsheets/d/1596h7jRJw2CN2XYImiE0Sgcj3kSIWUR2O61BkhVEJ8o/edit#gid=1528166832)

1. Advantage of Google sheets (sheets.google.com) :
   1. Unlike MS excel , its free
   2. Most excel features are available on sheets as well
   3. Can be used on any system linux, windows etc as long as you have an easy browser like google chrome
2. How many of you are not familiar with working with excel at all?
3. IPL Dataset [link](https://drive.google.com/file/d/14F-pKXkAzYvLqhaS3VnPMawFezxwBeHP/view) (Actual data). This is zip file, download it and uncompress it
   1. This IPL dataset has 2 csv files. CSV = Comma separated values
   2. The biggest challenge with CSV file is, if a cell contains “hello,world” - although its supposed to be in one column, it will split it into two different columns
4. Importing Files: file -> import (ipl matches)
   1. Under import location -> import new sheets
   2. Under separator type -> detect automatically (or you can use comma)
   3. Import another file: file -> import (ball by ball)
      1. Under import location -> import new sheets
      2. Under separator type -> detect automatically (or you can use comma)
5. Explore the attributes of the file
   1. Matches sheet: 817 rows = 817 IPL matches played during the period
   2. Ball by ball sheet: has info for each ball
6. Fixing headers and header aesthetics:
   1. How to self navigate: Go to help on the menu bar
      1. Type freeze -> choose freeze 1 rows
      2. You can undo - ctrl+z (in windows) OR cmd+z on mac
   2. Click on the row ->View -> Freeze -> 1 row
   3. Make the row bold -> Change color of text -> Fill the row with any color you want
7. Partial columns visible (Expand columns):
   1. Opt1:
      1. Select the column you want to expand and drag it right
   2. Opt2:
      1. Select all (cmd+a) -> Format (on menu bar) -> wrapping -> wrap
8. Sorting the data
   1. Ball by ball data
   2. Data is unordered ; Ideally for each match , in each innings, for each over, I want the balls ordered
   3. Data (menu bar) -> Create filter
   4. Click on ball ( from the filtered sheet) -> Sort (A-Z)
   5. Click on over ( from the filtered sheet) -> Sort (A-Z) [sorting goes from inner to outer]
   6. Click on innings ( from the filtered sheet) -> Sort (A-Z)
   7. Click on ID ( from the filtered sheet) -> Sort (A-Z)
   8. Overs should be only till 19 since its starting from 0 ; Then you can see the second innings
9. You want to analyse the data for match ID = 335982 [Filter] All further analyses should be done on the filtered sheet
   1. Filter on match ID -> clear (to clear filter) and select the match ID you want
   2. Copy data -> create new sheet and paste date
10. In the filtered match, How many runs were scored by KKR by the end of first 5 overs?
    1. Look for which innings KKR was batting = innings 1
    2. =sum('Data for 335982'!J2:J33)
11. How many extra runs were there in the first 5 overs
    1. =SUM('Data for 335982'!I2:I33)
12. Translate in Google Sheets
    1. GOOGLETRANSLATE("hello, how are you", "en", "hi") (Other languages = te)
13. Number of legbys in whole match [CountIfs - Conditional Counting]
    1. Countif counts only if the “if” condition is satisfied
    2. =COUNTIF('Data for 335982'!P2:P226,"legbyes")
    3. Teach how to search for syntax using official documentation
    4. Show how you can even just filter the match data for legbyes and then just highlight the rows and see what the count is
14. Whats the run rate for first innings for the filtered match
    1. Run rate = total runs / total overs [for first innings for the filtered match]
    2. Total runs =sum('Data for 335982'!J2:J125)
    3. Total overs =max('Data for 335982'!C2:C125)+1
    4. Run Rate = Cell of [Total runs ] / Cell of [Total overs ]
15. Given the list of names of batsman, find their last names?
    1. Share a demo list of batsmen name

| **Batsman** |
| --- |
| SC Ganguly |
| BB McCullum |
| RT Ponting |
| CL White |
| MV Boucher |
| B Akhil |
| AA Noffke |
| P Kumar |
| Z Khan |
| SB Joshi |
| W Jaffer |
| JH Kallis |
| R Dravid |
| V Kohli |

* 1. First we want to find the index of the space
     + 1. =FIND(" ",B5) -> Drag the formula down to affect all cells
       2. Length of the string: =LEN(B5)
       3. Last name = RIGHT(B5,D5-C5)
       4. First name = =LEFT(B5,FIND(" ",B5)-1)

1. Given the instances of "catch-out", display:"b bowler\_name, c fielder\_name" in a match-card
   1. Filter the match ID for dismissal kind = caught and paste this data in the sheet
   2. For each batsman for the above filtered data, we should show match card
   3. =IF(M5="caught", CONCATENATE("b ",G5," c ",O5 ))
2. Count the number of matches played in May-2009 [IF, AND, SUM
   1. In the original match sheet , create a new column =IF(AND(MONTH(C2)=5, YEAR(C2)=2009),1,0)
   2. Extend to all rows
   3. In Q7 sheet, Formula: =SUM('IPL Matches 2008-2020'!R2:R817)
3. Given a match-ID, determine who is the winner. [VLOOKUP, IFERROR]
   1. Explain vlookup using official [doc](https://support.google.com/docs/answer/3093318?hl=en)
   2. IFERROR(VLOOKUP(F5,'IPL Matches 2008-2020'! $A$2:$Q$817,11,FALSE),"Incorrect Match ID")
   3. For venue, use index as 5
   4. For city, use index as 2
4. Given, match-ID and column-name (pick this column name via drop down boxP, obtain the cell-value. [VLOOKUP, MATCH]
   1. Right click under Column name—> Dropdown—> Criterion(from a range) —> ='IPL Matches 2008-2020'!$A$1:$Q$1—> Done [Imagine this was filled in E5
   2. Under Value =VLOOKUP(D5,'IPL Matches 2008-2020'!$A$2:$Q$817,MATCH(E5,'IPL Matches 2008-2020'!$A$1:$Q$1,0),FALSE)