

```
In [2]: import pandas as pd
import numpy as np
```

```
In [3]: df = pd.read_csv("Match.csv")
df.head()
```

Out[3]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills	0	...
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills	0	...
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills	0	...
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills	0	...
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills	0	...

5 rows × 21 columns

```
In [4]: df.shape
```

Out[4]: (69761, 21)

```
In [5]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 69761 entries, 0 to 69760
Data columns (total 21 columns):
match_id          69761 non-null int64
inning            69761 non-null int64
batting_team      69761 non-null object
bowling_team      69761 non-null object
over              69761 non-null int64
ball              69761 non-null int64
batsman           69761 non-null object
non_striker       69761 non-null object
bowler            69761 non-null object
is_super_over     69761 non-null int64
wide_runs         69761 non-null int64
bye_runs          69761 non-null int64
legbye_runs       69761 non-null int64
noball_runs       69761 non-null int64
penalty_runs      69761 non-null int64
batsman_runs      69761 non-null int64
extra_runs        69761 non-null int64
total_runs        69761 non-null int64
player_dismissed  3516 non-null object
dismissal_kind     3516 non-null object
fielder           2529 non-null object
dtypes: int64(13), object(8)
memory usage: 11.2+ MB
```

```
In [6]: df["player_dismissed"].value_counts()
```

```
Out[6]: RG Sharma          59
        SK Raina           59
        AC Gilchrist       57
        G Gambhir          56
        Yuvraj Singh       55
        RV Uthappa         54
        YK Pathan          53
        V Kohli            51
        PA Patel           49
        S Dhawan           48
        KD Karthik         47
        V Sehwag           47
        JH Kallis          46
        KC Sangakkara      45
        R Dravid           45
        MS Dhoni           44
        BB McCullum        42
        DA Warner          40
        SR Tendulkar       39
        SC Ganguly         37
        DPMD Jayawardene   37
        NV Ojha            35
        Y Venugopal Rao    35
        AB de Villiers     34
        IK Pathan          33
        MK Pandey          32
        KA Pollard         32
        SR Watson          31
        ML Hayden          31
        SE Marsh           31
        ..
        GD McGrath         1
        SS Shaikh          1
        Harpreet Singh     1
        C Madan            1
        Jaskaran Singh     1
        S Ladda            1
        JEC Franklin       1
        AA Chavan          1
        B Kumar            1
        D Kalyankrishna    1
        JJ van der Wath     1
        Washington Sundar  1
        VS Malik           1
        AS Rajpoot         1
        RS Gavaskar        1
        CJ Jordan          1
        DL Chahar          1
        AD Nath            1
        D du Preez         1
        H Das              1
        BAW Mendis         1
        Anureet Singh      1
        Mohammad Ashraful  1
        SE Bond            1
        Shoaib Akhtar      1
        Kamran Khan        1
        VS Yeligati        1
        TM Srivastava      1
```

DH Yagnik	1
FH Edwards	1

Name: player_dismissed, Length: 354, dtype: int64

```
In [7]: df["fielder"].value_counts()
```

```
Out[7]: RV Uthappa          61
        KD Karthik         57
        MS Dhoni           55
        AC Gilchrist       54
        KC Sangakkara      46
        AB de Villiers     44
        SK Raina           40
        RG Sharma          39
        WP Saha            34
        KA Pollard         32
        PA Patel           31
        DA Warner          30
        NV Ojha            28
        V Kohli            28
        DPMD Jayawardene   27
        RA Jadeja          27
        S Dhawan           27
        HH Gibbs          26
        Yuvraj Singh       26
        MK Pandey          25
        JH Kallis          24
        V Sehwag           23
        RP Singh           23
        YK Pathan          22
        BB McCullum        21
        PP Chawla          21
        A Symonds          21
        AT Rayudu          20
        SR Tendulkar       20
        DJ Bravo           20
        ..
        AL Menaria         1
        Mohammed Shami     1
        Mashrafe Mortaza   1
        LH Ferguson        1
        AD Mascarenhas     1
        AS Rajpoot         1
        LR Shukla (sub)    1
        A Mukund (sub)     1
        KS Williamson      1
        MD Mishra          1
        AA Chavan          1
        Umar Gul           1
        R Bishnoi          1
        R Ninan            1
        Misbah-ul-Haq      1
        VS Malik           1
        JJ Roy             1
        IR Jaggi           1
        DA Miller (sub)    1
        AA Bilakhia        1
        AN Ghosh           1
        MR Marsh           1
        MG Johnson         1
        Gurkeerat Singh    1
        RD Chahar          1
        B Chipli (sub)     1
        GJ Bailey (sub)    1
        S Aravind          1
```

```
R Powell (sub)      1
JEC Franklin        1
Name: fielder, Length: 352, dtype: int64
```

```
In [8]: df["dismissal_kind"].value_counts()
```

```
Out[8]: caught          2029
        bowled          650
        run out         379
        lbw             223
        stumped         121
        caught and bowled 106
        retired hurt      4
        hit wicket        4
        Name: dismissal_kind, dtype: int64
```

```
In [9]: #Find Team wise runs scored.
        k=df.groupby(['batting_team'])['total_runs'].sum()
        k
```

```
Out[9]: batting_team
Chennai Super Kings      9253
Deccan Chargers          8931
Delhi Daredevils         10654
Gujarat Lions            2412
Kings XI Punjab          10763
Kochi Tuskers Kerala     1760
Kolkata Knight Riders     9924
Mumbai Indians           11202
Pune Warriors            1657
Rajasthan Royals         8039
Rising Pune Supergiant    2470
Royal Challengers Bangalore 10221
Sunrisers Hyderabad      2349
        Name: total_runs, dtype: int64
```

```
In [10]: # Find Total Runs of Virat Kohli.
        df[df['batsman']=='V Kohli']['batsman_runs'].sum()
```

```
Out[10]: 1428
```

```
In [11]: # Find Total Runs of Your Fav Player.

        df[df['batsman']=='Yuvraj Singh']['batsman_runs'].sum()
```

```
Out[11]: batsman_runs      1469
        dtype: int64
```

```
In [12]: # Extra Runs Given By All Bowlers in a League.  
df.groupby(['bowler'])[['extra_runs']].sum()
```

Out[12]:

extra_runs	
bowler	
A Choudhary	7
A Flintoff	1
A Kumble	62
A Mishra	60
A Mithun	19
A Nehra	62
A Nel	0
A Singh	18
A Symonds	38
A Uniyal	3
A Zampa	4
AA Chavan	4
AA Jhunjunwala	3
AA Kazi	1
AA Noffke	6
AB Agarkar	54
AB Dinda	60
AB McDonald	1
AC Thomas	9
AC Voges	2
AD Mascarenhas	6
AD Mathews	12
AF Milne	13
AG Murtaza	14
AJ Finch	1
AJ Tye	7
AL Menaria	4
AM Nayar	9
AM Rahane	0
AM Salvi	15
...	...
T Natarajan	7
T Thushara	12

	extra_runs
bowler	
TA Boult	11
TG Southee	13
TL Suman	2
TM Dilshan	4
TM Head	0
TS Mills	13
Tejas Baroka	0
UT Yadav	64
Umar Gul	32
V Kohli	11
V Sehwag	3
VR Aaron	22
VRV Singh	18
VS Malik	12
VS Yeligati	8
VY Mahesh	26
WA Mota	4
WD Parnell	27
WPUJC Vaas	19
Washington Sundar	6
Y Ganeswara Rao	0
Y Nagar	6
Y Venugopal Rao	9
YA Abdulla	17
YK Pathan	50
YS Chahal	30
Yuvraj Singh	11
Z Khan	89

285 rows × 1 columns

```
In [13]: # Extra Runs Given By TS Mills.
df[df['bowler']=='TS Mills']['extra_runs'].sum()
```

```
Out[13]: extra_runs    13
dtype: int64
```



```
In [14]: # Catches Taken By Mandeep Singh.
j=(df['fielder']=='Mandeep Singh') & (df['dismissal_kind']=='caught')
j.value_counts()
df[j]
```

Out[14]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_supe
11	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	2	5	DA Warner	S Dhawan	A Choudhary	
2728	12	2	Mumbai Indians	Royal Challengers Bangalore	3	3	MJ McClenaghan	RG Sharma	S Badree	
3922	17	1	Rising Pune Supergiant	Royal Challengers Bangalore	17	5	DT Christian	BA Stokes	S Aravind	
4691	20	2	Gujarat Lions	Royal Challengers Bangalore	2	2	DR Smith	BB McCullum	YS Chahal	
6355	27	1	Kolkata Knight Riders	Royal Challengers Bangalore	17	3	CR Woakes	SA Yadav	TS Mills	
9751	42	1	Kings XI Punjab	Royal Challengers Bangalore	7	6	SE Marsh	M Vohra	P Negi	
69340	293	2	Delhi Daredevils	Kings XI Punjab	15	2	IK Pathan	JR Hopes	PC Valthathy	

7 rows × 21 columns

```
In [15]: # How Many Times DA Warner is dismissed by Mohammed Shami.

x=(df['bowler']=='Mohammed Shami') & (df['player_dismissed']=='DA Warner')
x.value_counts()
df[x]
```

Out[15]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_ov
9029	39	1	Sunrisers Hyderabad	Delhi Daredevils	6	2	DA Warner	S Dhawan	Mohammed Shami	

1 rows × 21 columns

In [16]: # Total Number of Wickets Taken by YS Chahal.

```
y=df[df['bowler']=='YS Chahal']['player_dismissed'].value_counts()
y.count()
```

Out[16]: 13

In [17]: # Find Number of Stump Out done by MS Dhoni.
s=df[(df['fielder']=='MS Dhoni') & (df['dismissal_kind']=='stumped')]
s

Out[17]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super
3024	13	2	Gujarat Lions	Rising Pune Supergiant	12	1	BB McCullum	SK Raina	RD Chahar	
4006	17	2	Royal Challengers Bangalore	Rising Pune Supergiant	11	3	AB de Villiers	KM Jadhav	Imran Tahir	
11944	51	1	Delhi Daredevils	Rising Pune Supergiant	16	2	CJ Anderson	KK Nair	Washington Sundar	
28296	122	2	Royal Challengers Bangalore	Chennai Super Kings	8	5	RV Uthappa	R Dravid	M Muralitharan	
32944	141	2	Delhi Daredevils	Chennai Super Kings	16	5	DA Warner	KD Karthik	SB Jakati	
35415	152	1	Rajasthan Royals	Chennai Super Kings	13	2	GC Smith	RA Jadeja	M Muralitharan	

In [18]: # How Many Times S Dhawan Got Out by Lbw.
df[(df['player_dismissed']=='S Dhawan') & (df['dismissal_kind']=='lbw')]

Out[18]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super	over
31390	135	2	Mumbai Indians	Kings XI Punjab	5	1	S Dhawan	JP Duminy	YA Abdulla		
46812	199	1	Mumbai Indians	Deccan Chargers	4	3	S Dhawan	SR Tendulkar	Jaskaran Singh		

2 rows × 21 columns

```
In [19]: # Total number sixes hit by MS Dhoni.
js=df[(df['batsman']=='MS Dhoni') & (df['total_runs']==6)]
js.total_runs.value_counts()
```

```
Out[19]: 6      62
         Name: total_runs, dtype: int64
```

```
In [20]: # Total numbers of balls faced by MI Players.

df[df['batting_team']=='Mumbai Indians']['ball'].sum()
```

```
Out[20]: 31672
```

```
In [21]: # Total numbers of balls faced by KKR Players.

df[df['batting_team']=='Kolkata Knight Riders']['ball'].sum()
```

```
Out[21]: 28638
```

```
In [22]: # Find the total number of runs scored by Sachin Tendulkar.
df[df['batsman']=='SR Tendulkar']['total_runs'].sum()
```

```
Out[22]: 1730
```

```
In [23]: # Find the total number of balls faced by Schin Tendulkar.
df[df['batsman']=='SR Tendulkar']['ball'].sum()
```

```
Out[23]: 4996
```

```
In [24]: # Find Sachin Tendulkar's dismissal types in IPL.
df[df['batsman']=='SR Tendulkar']['dismissal_kind'].value_counts()
```

```
Out[24]: caught          23
         bowled          7
         lbw            5
         run out         3
         caught and bowled 1
         stumped         1
         Name: dismissal_kind, dtype: int64
```

```
In [25]: # Find the bowlers who dismissed SR Tendulkar.
```

```
df[df['player_dismissed']=='SR Tendulkar']['bowler']
```

```
Out[25]: 22501    Joginder Sharma
23139    Shoaib Akhtar
23597    Shahid Afridi
24663      IK Pathan
25219      VY Mahesh
26239      SK Trivedi
29727      PP Ojha
30608      LR Shukla
31373      VS Malik
32028      AB Agarkar
33240      D du Preez
34298      RP Singh
34887      PJ Sangwan
35588      R Vinay Kumar
37502      SK Warne
37846      SK Raina
39836      R Bhatia
41253      AD Mascarenhas
42994      A Mishra
44154      R Vinay Kumar
45810      L Balaji
46897      RP Singh
47440      PP Chawla
48791      R Sharma
49947      T Thushara
50753      PP Chawla
52211      PJ Sangwan
53353      JH Kallis
54523      DW Steyn
55419      SB Jakati
60095      R Sharma
60650      DE Bollinger
61427      A Mishra
63078      AL Menaria
64590      RJ Harris
65497      Yuvraj Singh
66654      IK Pathan
67929      P Kumar
69098      Anand Rajan
Name: bowler, dtype: object
```

In [26]: *# How many times was Suresh Raina dismissed.*

```
jn=df[df['player_dismissed']=='SK Raina']
```

```
jn
```

1261	6	1	Gujarat Lions	Sunrisers Hyderabad	9	6	SK Raina	KD Karthik	Rashid Khan
3634	16	1	Gujarat Lions	Mumbai Indians	12	1	SK Raina	BB McCullum	Harbhajan Singh
4708	20	2	Gujarat Lions	Royal Challengers Bangalore	4	7	SK Raina	BB McCullum	YS Chahal
5508	23	2	Gujarat Lions	Kolkata Knight Riders	18	5	SK Raina	RA Jadeja	Kuldeep Yadav
6183	26	2	Gujarat Lions	Kings XI Punjab	9	6	SK Raina	KD Karthik	AR Patel
7870	34	1	Gujarat Lions	Mumbai Indians	5	2	SK Raina	Ishan Kishan	JJ Bumrah
8802	38	1	Gujarat Lions	Rising Pune Supergiant	8	6	SK Raina	BB McCullum	Imran Tahir
9557	41	1	Gujarat Lions	Delhi Daredevils	14	2	KD Karthik	SK Raina	Mohammed Shami

In [27]: *# Display player details who hit the maximum no of sixes in the current IPL tournament.*

```
x=df[df['batsman_runs']==6]
```

```
x.groupby('batsman')['batsman_runs'].sum().sort_values(ascending=False).head(10)
```

Out[27]:

batsman	
SK Raina	498
YK Pathan	480
Yuvraj Singh	450
CH Gayle	432
RG Sharma	432
AC Gilchrist	426
RV Uthappa	420
BB McCullum	378
MS Dhoni	372
V Sehwag	360

Name: batsman_runs, dtype: int64

```
In [28]: # What is the Overall No Balls Runs for each team in the dataset?
df.groupby('batting_team')['ball','total_runs'].sum()
```

Out[28]:

	ball	total_runs
batting_team		
Chennai Super Kings	25363	9253
Deccan Chargers	25563	8931
Delhi Daredevils	29614	10654
Gujarat Lions	6044	2412
Kings XI Punjab	29795	10763
Kochi Tuskers Kerala	5330	1760
Kolkata Knight Riders	28638	9924
Mumbai Indians	31672	11202
Pune Warriors	5083	1657
Rajasthan Royals	23278	8039
Rising Pune Supergiant	6845	2470
Royal Challengers Bangalore	29562	10221
Sunrisers Hyderabad	6037	2349

```
In [31]: # Which player scored the most runs in the dataset?
df.groupby('batsman')[['batsman_runs']].sum().max()
```

Out[31]: batsman_runs 2159
dtype: int64

```
In [36]: # Create Your Own Question Atleast 20 To 30 Questions.
```

ALL THE BEST

```
In [32]: # 1- how many match played by Mumbai Indians
MI=df[df['batting_team']=='Mumbai Indians']
MI.batting_team.value_counts()
```

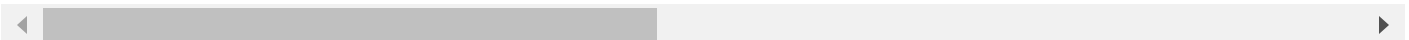
```
Out[32]: Mumbai Indians      8662
Name: batting_team, dtype: int64
```

```
In [33]: # 2- total run of S Dhawan against Sunrisers Hyderabad
x=df[df['bowling_team']=='Sunrisers Hyderabad']
x[x['batsman']=='S Dhawan']
```

```
Out[33]:
```

match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...	b
----------	--------	--------------	--------------	------	------	---------	-------------	--------	---------------	-----	---

0 rows × 21 columns



```
In [35]: # Most wicket tacker bowler
x=df.groupby('bowler')['player_dismissed'].count().sort_values(ascending=False).head(10)
x
```

```
Out[35]: bowler
SL Malinga      77
IK Pathan      72
RP Singh       71
A Mishra       70
PP Ojha        65
JA Morkel      64
Z Khan         62
PP Chawla      62
Harbhajan Singh 62
SK Warne       59
Name: player_dismissed, dtype: int64
```

```
In [36]: #Which team scored the most runs in the first innings?
tf=df[df['inning']==1]
af=tf.groupby('batting_team')['total_runs'].count().sort_values(ascending=False)
af.head(1)
```

```
Out[36]: batting_team
Mumbai Indians      4764
Name: total_runs, dtype: int64
```

```
In [ ]: #which team scored most run in Last 5 over
```