

Importing Libraries in Python

```
In [1]: import pandas as pd
import numpy as np
```

Reading DataSet

```
In [2]: df = pd.read_csv("Match.csv")
df.head()
```

```
Out[2]:
```

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman_runs	extra_runs	total_runs	player_dismissed	dismissal_kind
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	0	0	0	0	NaN	NaN
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	0	0	0	0	NaN	NaN
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	0	4	0	4	NaN	NaN
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	0	0	0	0	NaN	NaN
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	0	0	2	2	NaN	NaN

5 rows × 21 columns

```
In [3]: df.shape
```

```
Out[3]: (69761, 21)
```

In [4]: df.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 69761 entries, 0 to 69760
Data columns (total 21 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   match_id              69761 non-null  int64  
1   inning               69761 non-null  int64  
2   batting_team         69761 non-null  object  
3   bowling_team         69761 non-null  object  
4   over                 69761 non-null  int64  
5   ball                 69761 non-null  int64  
6   batsman              69761 non-null  object  
7   non_striker          69761 non-null  object  
8   bowler               69761 non-null  object  
9   is_super_over        69761 non-null  int64  
10  wide_runs            69761 non-null  int64  
11  bye_runs             69761 non-null  int64  
12  legbye_runs          69761 non-null  int64  
13  noball_runs          69761 non-null  int64  
14  penalty_runs         69761 non-null  int64  
15  batsman_runs         69761 non-null  int64  
16  extra_runs           69761 non-null  int64  
17  total_runs           69761 non-null  int64  
18  player_dismissed     3516 non-null   object  
19  dismissal_kind       3516 non-null   object  
20  fielder              2529 non-null   object  
dtypes: int64(13), object(8)
memory usage: 11.2+ MB

```

In [5]: df[["player_dismissed"]].value_counts()

```

Out[5]: player_dismissed
RG Sharma      59
SK Raina       59
AC Gilchrist   57
G Gambhir      56
Yuvraj Singh   55
..
A Choudhary     1
BAW Mendis      1
H Das           1
GR Napier       1
Jaskaran Singh  1
Length: 354, dtype: int64

```

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

```
Out[6]: fielder
RV Uthappa      61
KD Karthik      57
MS Dhoni        55
AC Gilchrist    54
KC Sangakkara   46
..
RA Jadeja (sub)  1
Misbah-ul-Haq   1
AD Mascarenhas  1
M Kaif (sub)    1
KS Williamson   1
Length: 352, dtype: int64
```

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

```
Out[7]: dismissal_kind
caught          2029
bowled          650
run out         379
lbw             223
stumped         121
caught and bowled 106
hit wicket       4
retired hurt     4
dtype: int64
```

Find Team wise runs scored.

```
In [8]: df.groupby(["batting_team"])["total_runs"].sum().sort_values(ascending=False)
```

```
Out[8]: batting_team
Mumbai Indians      11202
Kings XI Punjab     10763
Delhi Daredevils    10654
Royal Challengers Bangalore 10221
Kolkata Knight Riders  9924
Chennai Super Kings  9253
Deccan Chargers     8931
Rajasthan Royals    8039
Rising Pune Supergiant 2470
Gujarat Lions       2412
Sunrisers Hyderabad 2349
Kochi Tuskers Kerala 1760
Pune Warriors       1657
Name: total_runs, dtype: int64
```

Find Total Runs of Virat Kohli.

```
In [9]: df[df.batsman.str.contains("V Kohli")]["batsman_runs"].sum()
```

```
Out[9]: 1428
```

Find Total Runs of Your Fav Player.

```
In [10]: batsman=df.groupby(['batsman'])['batsman_runs'].sum()
batsman['KS Williamson']
```

```
Out[10]: 256
```

Extra Runs Given By All Bowlers in a League.

```
In [11]: df.groupby(['bowler'])['extra_runs'].sum().sort_values(ascending=False)
```

```
Out[11]: bowler
SL Malinga      158
RP Singh       135
I Sharma       118
P Kumar        111
IK Pathan      101
...
I Malhotra      0
B Chipili       0
Tejas Baroka    0
A Nel           0
TM Head         0
Name: extra_runs, Length: 285, dtype: int64
```

Extra Runs Given By TS Mills.

```
In [12]: df[df['bowler']=='TS Mills']['extra_runs'].sum()
```

```
Out[12]: 13
```

Catches Taken By Mandeep Singh.

```
In [13]: df[(df['dismissal_kind']=='caught') & (df['fielder']=='Mandeep Singh)][['batsman','dismissal_kind']]
```

```
Out[13]:
```

	batsman	dismissal_kind
11	DA Warner	caught
2728	MJ McClenaghan	caught
3922	DT Christian	caught
4691	DR Smith	caught
6355	CR Woakes	caught
9751	SE Marsh	caught
69340	IK Pathan	caught

How Many Times DA Warner is dismissed by Mohammed Shami.

```
In [14]: df[(df['player_dismissed']=='DA Warner') & (df['bowler']=='Mohammed Shami)][['bowler']].count()
```

```
Out[14]: 1
```

Total Number of Wickets Taken by YS Chahal.

```
In [15]: df[(df['bowler']=='YS Chahal') & (df['dismissal_kind'].isna() != True)][['dismissal_kind']].count()
```

```
Out[15]: 14
```

Find Number of Stump Out done by MS Dhoni.

```
In [16]: df[(df['fielder']=='MS Dhoni') & (df['dismissal_kind']=='stumped)][['fielder']].count()
```

```
Out[16]: 17
```

How Many Times S Dhawan Got Out by lbw.

```
In [17]: df[(df['player_dismissed']=='S Dhawan') & (df['dismissal_kind']=='lbw)][['dismissal_kind']].count()
```

```
Out[17]: 2
```

Total number sixes hit by MS Dhoni.

```
In [18]: df[(df['batsman']=='MS Dhoni') & (df['batsman_runs']==6)][['batsman']].count()
```

```
Out[18]: 62
```

Total numbers of balls faced by MI Players.

```
In [19]: df[(df['batting_team']=='Mumbai Indians)][['ball']].count()
```

```
Out[19]: 8662
```

Total numbers of balls faced by KKR Players.

```
In [20]: df[(df['batting_team']=='Kings XI Punjab)][['ball']].count()
```

```
Out[20]: 8208
```

Find the total number of runs scored by Sachin Tendulkar.

```
In [21]: df[(df['batsman']=='SR Tendulkar')]['batsman_runs'].sum()
```

```
Out[21]: 1578
```

Find the total number of balls faced by Schin Tendulkar.

```
In [22]: df[(df['batsman']=='SR Tendulkar')]['ball'].count()
```

```
Out[22]: 1388
```

Find Sachin Tendulkar's dismissal types in IPL.

```
In [23]: df[(df['batsman']=='SR Tendulkar')]['dismissal_kind'].value_counts()
```

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

Find the bowlers who dismissed SR Tendulkar.

```
In [24]: df[(df['player_dismissed']=='SR Tendulkar')]['bowler']
```

```
Out[24]: 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
```

How many times was Suresh Raina dismissed.

```
In [25]: df[(df['player_dismissed']=='SK Raina')]['dismissal_kind'].count()
```

```
Out[25]: 59
```

Display player details who hit the maximum no of sixes in the current IPL tournament.

```
In [26]: df[df['batsman_runs']==6]["batsman"].value_counts()
```

```
Out[26]: SK Raina      83
         YK Pathan    80
         Yuvraj Singh  75
         CH Gayle     72
         RG Sharma    72
         ..
         RT Ponting    1
         SD Chitnis    1
         Sohail Tanvir  1
         SP Goswami    1
         WD Parnell    1
         Name: batsman, Length: 234, dtype: int64
```

What is the Overall No Balls Runs for each team in the dataset?

```
In [27]: df.groupby(["batting_team"])['noball_runs'].sum().sort_values(ascending=False)
```

```
Out[27]: batting_team
         Deccan Chargers      45
         Rajasthan Royals    45
         Delhi Daredevils    42
         Mumbai Indians     42
         Kolkata Knight Riders 39
         Kings XI Punjab     36
         Royal Challengers Bangalore 34
         Chennai Super Kings  32
         Kochi Tuskers Kerala  11
         Gujarat Lions       10
         Pune Warriors       9
         Sunrisers Hyderabad  6
         Rising Pune Supergiant 1
         Name: noball_runs, dtype: int64
```

Which player scored the most runs in the dataset?

```
In [28]: df.groupby(['batsman'])['batsman_runs'].sum().sort_values(ascending=False).head(1)
```

```
Out[28]: batsman
         SK Raina    2159
         Name: batsman_runs, dtype: int64
```

which bowler take the most wicket

```
In [29]: wkt= df[df['dismissal_kind'].isin(['caught','bowled','lbw','caught and bowled','stumped','hit wicket'])]
         wicket=wkt.groupby(['bowler'])['dismissal_kind'].count().sort_values(ascending=False).head(1)
         wicket
```

```
Out[29]: bowler
         SL Malinga    71
         Name: dismissal_kind, dtype: int64
```


which bowler take the least wicket

```
In [30]: wkt.groupby(['bowler'])['dismissal_kind'].count().sort_values(ascending=False).tail(1)
```

```
Out[30]: bowler  
MJ Henry    1  
Name: dismissal_kind, dtype: int64
```

Number of sixes in whole IPL dataset

```
In [31]: df[df['batsman_runs']==6]['batsman_runs'].count()
```

```
Out[31]: 2939
```

Number of fours in whole IPL dataset

```
In [32]: df[df['batsman_runs']==4]['batsman_runs'].count()
```

```
Out[32]: 7970
```

Top 10 Scoring batsman:

```
In [33]: df.groupby(['batsman'])['batsman_runs'].sum().sort_values(ascending=False).head(10)
```

```
Out[33]: batsman  
SK Raina      2159  
G Gambhir     1907  
RG Sharma     1794  
MS Dhoni      1585  
SR Tendulkar  1578  
RV Uthappa    1509  
JH Kallis     1490  
Yuvraj Singh  1469  
AC Gilchrist  1446  
SE Marsh      1439  
Name: batsman_runs, dtype: int64
```

how many sixes are hit by CH Gayle

```
In [34]: df[(df['batsman']=='CH Gayle') & (df['batsman_runs']==6)]['batsman_runs'].count()
```

```
Out[34]: 72
```

List of Players run out

```
In [35]: runout=df[df["dismissal_kind"]=="run out"][['batting_team','batsman','dismissal_kind']]
runout
```

```
Out[35]:
```

	batting_team	batsman	dismissal_kind
199	Royal Challengers Bangalore	KM Jadhav	run out
247	Royal Challengers Bangalore	A Choudhary	run out
371	Mumbai Indians	HH Pandya	run out
1063	Royal Challengers Bangalore	Vishnu Vinod	run out
1313	Gujarat Lions	DS Kulkarni	run out
...
68598	Kochi Tuskers Kerala	R Vinay Kumar	run out
68874	Royal Challengers Bangalore	SS Tiwary	run out
69123	Mumbai Indians	Harbhajan Singh	run out
69459	Rajasthan Royals	AL Menaria	run out
69658	Pune Warriors	WD Parnell	run out

379 rows × 3 columns

The top two players are frequently run out.

```
In [36]: runout.value_counts().head(2)
```

```
Out[36]: batting_team    batsman    dismissal_kind
Delhi Daredevils    G Gambhir    run out          7
Deccan Chargers    Y Venugopal Rao    run out          7
dtype: int64
```

List of team showing an extra run in decending order

```
In [37]: exrun=df.groupby('bowling_team')['extra_runs'].sum().sort_values(ascending=False)
exrun
```

```
Out[37]: bowling_team
Mumbai Indians    711
Royal Challengers Bangalore    634
Delhi Daredevils    622
Kings XI Punjab    593
Kolkata Knight Riders    592
Deccan Chargers    541
Chennai Super Kings    490
Rajasthan Royals    452
Rising Pune Supergiant    111
Kochi Tuskers Kerala    106
Pune Warriors    95
Gujarat Lions    90
Sunrisers Hyderabad    84
Name: extra_runs, dtype: int64
```

which team has the least amount of extra run

In [38]: `exrun.tail(1)`

Out[38]: bowling_team
Sunrisers Hyderabad 84
Name: extra_runs, dtype: int64

top 5 wicket taking bolwer

In [39]: `wic=df.groupby(['bowler'])['player_dismissed']
wic.count().sort_values(ascending=False).head(5)`

Out[39]: bowler
SL Malinga 77
IK Pathan 72
RP Singh 71
A Mishra 70
PP Ojha 65
Name: player_dismissed, dtype: int64

Gujrat Lions batsman Score Board

In [40]: `df[(df['batting_team']=='Gujarat Lions')].groupby('batsman')['batsman_runs'].sum().sort_values(ascending=False)`

Out[40]: batsman
SK Raina 442
KD Karthik 361
BB McCullum 320
AJ Finch 300
Ishan Kishan 277
DR Smith 239
RA Jadeja 158
JJ Roy 59
JP Faulkner 54
AJ Tye 53
Basil Thampi 28
P Kumar 8
Ankit Soni 7
IK Pathan 2
PJ Sangwan 2
DS Kulkarni 1
MM Patel 0
AD Nath 0
Name: batsman_runs, dtype: int64

How many wickets are taken By Bowler Zampa

In [41]: `df[(df['bowler']=='A Zampa')]['player_dismissed'].count()`

Out[41]: 7

Which bowler has given up the most extra runs

```
In [42]: df.groupby(['bowler'])['extra_runs'].sum().sort_values(ascending=False).head(1)
```

```
Out[42]: bowler  
SL Malinga    158  
Name: extra_runs, dtype: int64
```

Wicket Taken by Rashid Khan

```
In [43]: wic.get_group('Rashid Khan').count()
```

```
Out[43]: 18
```

List of team Who played the Super Over

```
In [49]: df[(df['is_super_over']==1)][['batting_team','bowling_team']].value_counts()
```

```
Out[49]: batting_team    bowling_team  
Gujarat Lions           Mumbai Indians      8  
Kolkata Knight Riders   Rajasthan Royals    7  
Chennai Super Kings     Kings XI Punjab    5  
Mumbai Indians          Gujarat Lions      5  
Kings XI Punjab         Chennai Super Kings  4  
Rajasthan Royals        Kolkata Knight Riders  4  
dtype: int64
```

list of Players who get Retired Hurt`

```
In [52]: df[(df['dismissal_kind']=='retired hurt')]['batsman'].value_counts()
```

```
Out[52]: KH Pandya      1  
DJ Thorneley           1  
Harbhajan Singh        1  
VVS Laxman              1  
Name: batsman, dtype: int64
```

```
In [ ]:
```