

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	...	bye_runs	legbye_runs	noball_runs	penalty_r
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills	0	...	0	0	0	
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills	0	...	0	0	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):
#   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 [6]:

```
df[["player_dismissed"]].value_counts()
```

Out[6]:

```
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 [7]:

```
df[["fielder"]].value_counts()
```

Out[7]:

```
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 [8]:

```
df[["dismissal_kind"]].value_counts()
```

Out[8]:

```
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
```

In [10]:

```
#Find Team wise runs scored.
team_wise=df.groupby('batting_team')['total_runs'].sum().sort_values(ascending=False)
team_wise
```

Out[10]:

```
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
```

In [11]:

```
# Find Total Runs of Virat Kohli.
player_run=df.groupby('batsman')['batsman_runs'].sum().sort_values(ascending=False)
player_run['V Kohli']
```

Out[11]:

1428

In [12]:

```
# Find Total Runs of Your Fav Player.
name=input("enter your fav player name :-")
print("Total run of ::", player_run[name])
```

```
enter your fav player name :-DA Warner
Total run of :: 1391
```

In [13]:

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

Out[13]:

```
bowler
A Choudhary      7
A Flintoff       1
A Kumble        62
A Mishra        60
A Mithun        19
..
YA Abdulla      17
YK Pathan       50
YS Chahal       30
Yuvraj Singh    11
Z Khan         89
Name: extra_runs, Length: 285, dtype: int64
```

In [14]:

```
# Extra Runs Given By TS Mills.
bowler=df.groupby('bowler')['extra_runs'].sum().sort_values(ascending=False)
bowler['TS Mills']
```

Out[14]:

13

In [15]:

```
# Catches Taken By Mandeep Singh.
mask=df[df['dismissal_kind']=='caught']
mask[mask['fielder']=='Mandeep Singh']['batsman','dismissal_kind']
```

Out[15]:

	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

```
In [17]:  
  
# How Many Times DA Warner is dismissed by Mohammed Shami.  
msk1=df['player_dismissed']=='DA Warner'  
mask2=df['bowler']=='Mohammed Shami'  
len(df[msk1 & mask2][['batsman','bowler']])  
df[msk1 & mask2][['batsman','bowler']]
```

Out[17]:

	batsman	bowler
9029	DA Warner	Mohammed Shami

```
In [18]:  
  
# Total Number of Wickets Taken by YS Chahal.  
bowl=df[df['bowler']=='YS Chahal']  
bowl['player_dismissed'].count()
```

Out[18]:  
  
14

```
In [19]:  
  
# Find Number of Stump Out done by MS Dhoni.  
m1=df['dismissal_kind']=='stumped'  
m2=df['fielder']=='MS Dhoni'  
df[m1 & m2][['batsman','dismissal_kind','fielder']]
```

Out[19]:

	batsman	dismissal_kind	fielder
3024	BB McCullum	stumped	MS Dhoni
4006	AB de Villiers	stumped	MS Dhoni
11944	CJ Anderson	stumped	MS Dhoni
28296	RV Uthappa	stumped	MS Dhoni
32944	DA Warner	stumped	MS Dhoni
35415	GC Smith	stumped	MS Dhoni
35445	AS Raut	stumped	MS Dhoni
49905	AT Rayudu	stumped	MS Dhoni
51089	MD Mishra	stumped	MS Dhoni
52411	CH Gayle	stumped	MS Dhoni
52422	DJ Hussey	stumped	MS Dhoni
52495	WP Saha	stumped	MS Dhoni
54962	MD Mishra	stumped	MS Dhoni
55630	MS Bisla	stumped	MS Dhoni
55677	EJG Morgan	stumped	MS Dhoni
55692	MK Tiwary	stumped	MS Dhoni
60699	AT Rayudu	stumped	MS Dhoni

```
In [20]:  
  
len(df[m1 & m2 ]) # total 17 time stump out done my ms dhoni
```

Out[20]:  
  
17

```
In [12]:  
  
# How Many Times S Dhawan Got Out by lbw.
```

In [21]:

```
# Total number sixes hit by MS Dhoni.  
  
b1=df['batsman_runs']==6  
b2=df['batsman']=='MS Dhoni'  
len(df[b1 & b2][['batsman','batsman_runs']])
```

Out[21]:

62

In [22]:

```
# Total numbers of balls faced by MI Players.  
df.groupby('batting_team')['ball'].count().sort_values(ascending=False)
```

Out[22]:

batting_team	
Mumbai Indians	8662
Kings XI Punjab	8208
Delhi Daredevils	8197
Royal Challengers Bangalore	8188
Kolkata Knight Riders	7887
Deccan Chargers	7078
Chennai Super Kings	6987
Rajasthan Royals	6443
Rising Pune Supergiant	1900
Sunrisers Hyderabad	1679
Gujarat Lions	1675
Kochi Tuskers Kerala	1455
Pune Warriors	1402

Name: ball, dtype: int64

In [15]:

```
# Total numbers of balls faced by KKR Players.
```

In [23]:

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

Out[23]:

1578

In [24]:

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

Out[24]:

1388

In [25]:

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

Out[25]:

caught	23
bowled	7
lbw	5
run out	2
caught and bowled	1
stumped	1

Name: dismissal\_kind, dtype: int64

In [26]:

```
# Find the bowlers who dismissed SR Tendulkar.
df[df['player_dismissed']=='SR Tendulkar']['bowler'].value_counts()
```

Out[26]:

```
R Sharma      2
PP Chawla     2
RP Singh      2
PJ Sangwan    2
IK Pathan     2
R Vinay Kumar 2
A Mishra      2
T Thushara    1
L Balaji      1
D du Preez    1
DE Bollinger  1
Anand Rajan   1
AB Agarkar    1
P Kumar       1
SB Jakati     1
Shoaib Akhtar 1
SK Warne      1
SK Raina      1
R Bhatia      1
VY Mahesh     1
Shahid Afridi 1
LR Shukla     1
AD Mascarenhas 1
RJ Harris     1
PP Ojha       1
Joginder Sharma 1
JH Kallis     1
Yuvraj Singh  1
DW Steyn      1
VS Malik      1
AL Menaria    1
SK Trivedi    1
Name: bowler, dtype: int64
```

In [27]:

```
# How many times was Suresh Raina dismissed.
df[df['player_dismissed']=='SK Raina']['dismissal_kind'].shape
```

Out[27]:

```
(59,)
```

In [28]:

```
len(df[df['player_dismissed']=='SK Raina']['dismissal_kind'])
```

Out[28]:

```
59
```

In [33]:

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

In [34]:

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

In [29]:

```
# Which player scored the most runs in the dataset?

top_bats=df.groupby("batsman")["batsman_runs"].sum().sort_values(ascending=False)
top_bats.head(1)
```

Out[29]:

```
batsman
SK Raina    2159
Name: batsman_runs, dtype: int64
```

In [36]:

```
# Create Your Own Question Atleast 20 To 30 Questions.
```

ALL THE BEST

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

In [30]:

```
No_ball_runs=df.groupby('batting_team')['noball_runs'].sum().sort_values(ascending=False)
```

In [31]:

```
No_ball_runs
```

Out[31]:

```
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
```

## top ten batsman in term of scored

In [32]:

```
top_bats=df.groupby("batsman")["batsman_runs"].sum().sort_values(ascending=False).head(10)
top_bats
```

Out[32]:

```
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
```

## find the top 10 batsman with max no of six

In [33]:

```
six=df[df['batsman_runs']==6]
six.groupby('batsman')['batsman_runs'].count().sort_values(ascending=False).head(10)
```

Out[33]:

```
batsman
SK Raina      83
YK Pathan     80
Yuvraj Singh  75
CH Gayle      72
RG Sharma     72
AC Gilchrist  71
RV Uthappa    70
BB McCullum   63
MS Dhoni      62
V Sehwag      60
Name: batsman_runs, dtype: int64
```

## find the batsman with most number of 4 and 6 in last 5 overs

In [35]:

```
temp_df=df[df['over']>15]
```

In [36]:

```
temp_df=temp_df[(temp_df['batsman_runs']==4) | (temp_df['batsman_runs']==6)]
```

In [37]:

```
temp_df.groupby('batsman')['batsman_runs'].count().sort_values(ascending=False).head()
```

Out[37]:

```
batsman
MS Dhoni      100
RG Sharma     91
Yuvraj Singh  70
KA Pollard    64
KD Karthik    60
Name: batsman_runs, dtype: int64
```

## find V Kohli record against all teams

In [38]:

```
temp=df[df['batsman']=='V Kohli']
```

In [39]:

```
temp.groupby('bowling_team')['batsman_runs'].sum()
```

Out[39]:

```
bowling_team
Chennai Super Kings    188
Deccan Chargers        255
Delhi Daredevils       188
Gujarat Lions          74
Kings XI Punjab        164
Kochi Tuskers Kerala   50
Kolkata Knight Riders   72
Mumbai Indians         200
Pune Warriors          67
Rajasthan Royals       87
Rising Pune Supergiant  83
Name: batsman_runs, dtype: int64
```

## Find V Kohli four and six against all teams

In [40]:

```
temp=df[(df['batsman']=='V Kohli') & ((df['batsman_runs']==4) | (df['batsman_runs']==6))]
```



In [41]:

```
temp.groupby('bowling_team')['batsman_runs'].count()
```

Out[41]:

```
bowling_team
Chennai Super Kings      25
Deccan Chargers          31
Delhi Daredevils         23
Gujarat Lions            9
Kings XI Punjab          23
Kochi Tuskers Kerala     7
Kolkata Knight Riders    8
Mumbai Indians           27
Pune Warriors            8
Rajasthan Royals         8
Rising Pune Supergiant   9
Name: batsman_runs, dtype: int64
```

## create a function that can retrun the highest score of any batsman

In [42]:

```
def highest(batsman):
    temp=df[df['batsman']==batsman]
    return temp.groupby('match_id')['batsman_runs'].sum().sort_values(ascending=False).head
```

In [43]:

```
highest('MS Dhoni')
```

Out[43]:

```
<bound method NDFrame.head of match_id
182    66
74     65
289    63
25     61
91     60
..
169     2
61     2
83     1
224     0
206     0
Name: batsman_runs, Length: 62, dtype: int64>
```

## Most Wickets taker bowlers

In [44]:

```
df.groupby('bowler')['player_dismissed'].count().sort_values(ascending=False)
```

Out[44]:

```
bowler
SL Malinga      77
IK Pathan       72
RP Singh        71
A Mishra        70
PP Ojha         65
..
M Manhas        0
LRPL Taylor     0
LPC Silva       0
C Ganapathy     0
SR Tendulkar    0
Name: player_dismissed, Length: 285, dtype: int64
```

In [ ]: