

```
In [35]: import pandas as pd
```

```
In [36]: delivery=pd.read_csv('deliveries.csv')
```

```
In [37]: match=pd.read_csv('matches.csv')
```

```
In [38]: delivery.head()
```

```
Out[38]:
```

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

5 rows × 21 columns

```
In [39]: runs=delivery.groupby('batsman')
```

```
In [40]: runs.get_group('V Kohli').shape
```

```
Out[40]: (3494, 21)
```

```
In [41]: runs['batsman_runs'].sum().sort_values(ascending=False).head()
```

```
Out[41]:
```

batsman	
SK Raina	4548
V Kohli	4423
RG Sharma	4207
G Gambhir	4132
DA Warner	4014

Name: batsman\_runs, dtype: int64

## Batsman who score maximum number of four

```
In [42]: mask=delivery['batsman_runs']==4
new_delivery=delivery[mask]
```

```
In [43]: new_delivery.shape[0]
```

```
Out[43]: 17033
```

## Batsman who score maximum number of six

```
In [44]: mask=delivery['batsman_runs']==6
new_delivery=delivery[mask]
```

```
In [45]: new_delivery.shape[0]
```

```
Out[45]: 6523
```

```
In [46]: new_delivery.groupby('batsman')['batsman_runs'].count().sort_values(ascending=False)
```

```
Out[46]: batsman
CH Gayle      266
SK Raina      174
RG Sharma     173
V Kohli       160
DA Warner     160
Name: batsman_runs, dtype: int64
```

```
In [47]: vk=delivery[delivery['batsman']=='V Kohli']
```

```
In [48]: vk.groupby('bowling_team')['batsman_runs'].sum().sort_values(ascending=False).head
```

```
Out[48]: bowling_team
Chennai Super Kings      706
Delhi Daredevils         661
Kings XI Punjab          483
Name: batsman_runs, dtype: int64
```

```
In [49]: def run_scored(batsman_name):
vk=delivery[delivery['batsman']==batsman_name]
return vk.groupby('bowling_team')['batsman_runs'].sum().sort_values(ascending=False).head
```

```
In [50]: run_scored('V Kohli')
```

```
Out[50]: 'Chennai Super Kings'
```

```
In [51]: run_scored('G Gambhir')
```

```
Out[51]: 'Kings XI Punjab'
```

```
In [52]: run_scored('MS Dhoni')
```

```
Out[52]: 'Royal Challengers Bangalore'
```

## isin()

Find the most destructive death over batsman in the history of IPL

Strike Rate=(Number of runs/Number of balls)/100

Min batsman 200 balls in over 16-20

```
In [53]: death_over=delivery[delivery['over']>15]
```

```
In [54]: death_over.groupby('batsman')['batsman_runs'].count()>200
```

```
Out[54]: batsman
A Ashish Reddy      False
A Chandila          False
A Chopra            False
A Choudhary         False
A Flintoff          False
...
YS Chahal           False
YV Takawale         False
Yashpal Singh       False
Yuvraj Singh        True
Z Khan              False
Name: batsman_runs, Length: 416, dtype: bool
```

## batsman who plays > 200 balls

```
In [55]: all_batsman=death_over.groupby('batsman')['batsman_runs'].count()
x=all_batsman>200
all_batsman[x]
```

```
Out[55]: batsman
A Mishra      225
AB de Villiers 570
AD Mathews    289
AM Rahane     268
AR Patel      229
AT Rayudu     425
BJ Hodge      385
DA Miller     360
DA Warner     228
DJ Bravo      409
DJ Hussey     234
DPMD Jayawardene 246
Harbhajan Singh 418
IK Pathan     465
JA Morkel     425
JH Kallis     231
JP Duminy     518
JP Faulkner   294
KA Pollard    838
KD Karthik    463
KM Jadhav     338
LRPL Taylor   204
MK Pandey     224
MK Tiwary     423
MS Dhoni      1224
NV Ojha       304
P Kumar       268
PP Chawla     311
R Vinay Kumar 235
RA Jadeja     576
RG Sharma     748
RV Uthappa    275
S Badrinath   283
S Dhawan      243
SK Raina      458
SPD Smith     316
SS Tiwary     300
STR Binny     218
V Kohli       546
WP Saha       295
Y Venugopal Rao 260
YK Pathan     584
Yuvraj Singh  516
Name: batsman_runs, dtype: int64
```

```
In [56]: all_batsman[x].shape
```

```
Out[56]: (43,)
```

```
In [57]: all_batsman[x].index
```

```
Out[57]: Index(['A Mishra', 'AB de Villiers', 'AD Mathews', 'AM Rahane', 'AR Patel',
               'AT Rayudu', 'BJ Hodge', 'DA Miller', 'DA Warner', 'DJ Bravo',
               'DJ Hussey', 'DPMD Jayawardene', 'Harbhajan Singh', 'IK Pathan',
               'JA Morkel', 'JH Kallis', 'JP Duminy', 'JP Faulkner', 'KA Pollard',
               'KD Karthik', 'KM Jadhav', 'LRPL Taylor', 'MK Pandey', 'MK Tiwary',
               'MS Dhoni', 'NV Ojha', 'P Kumar', 'PP Chawla', 'R Vinay Kumar',
               'RA Jadeja', 'RG Sharma', 'RV Uthappa', 'S Badrinath', 'S Dhawan',
               'SK Raina', 'SPD Smith', 'SS Tiwary', 'STR Binny', 'V Kohli', 'WP Saha',
               'Y Venugopal Rao', 'YK Pathan', 'Yuvraj Singh'],
              dtype='object', name='batsman')
```

```
In [58]: batsman_list= all_batsman[x].index.tolist()
```

```
In [59]: # Runs scored by all these 43 batsman
# Balls played by these 43 batsman

final=delivery[delivery['batsman'].isin(batsman_list)]
```

```
In [60]: runs=final.groupby('batsman')['batsman_runs'].sum()
```

```
In [61]: balls=final.groupby('batsman')['batsman_runs'].count()
```

```
In [62]: sr=(runs/balls)*100
```

```
In [63]: sr
```

```
Out[63]: batsman
A Mishra                89.005236
AB de Villiers          145.129059
AD Mathews              120.868114
AM Rahane               117.486549
AR Patel               122.672065
AT Rayudu              123.014257
BJ Hodge               121.422376
DA Miller              137.709251
DA Warner              138.318401
DJ Bravo               122.286822
DJ Hussey              120.072661
DPMD Jayawardene       118.791064
Harbhajan Singh        135.194585
IK Pathan              116.751269
JA Morkel              136.938202
JH Kallis              105.936272
JP Duminy              121.970624
JP Faulkner            129.802956
KA Pollard             140.621266
KD Karthik             123.008475
KM Jadhav              130.555556
LRPL Taylor            120.070838
MK Pandey              116.938453
MK Tiwary              114.127424
MS Dhoni               132.835821
NV Ojha                114.528024
P Kumar                105.263158
PP Chawla              110.278373
R Vinay Kumar          106.666667
RA Jadeja              118.792867
RG Sharma              128.497251
RV Uthappa             127.635135
S Badrinath            115.742972
S Dhawan               118.502496
SK Raina               134.995548
SPD Smith              128.507127
SS Tiwary              114.851485
STR Binny              123.747981
V Kohli                126.588437
WP Saha                127.727646
Y Venugopal Rao        113.872832
YK Pathan              140.751445
Yuvraj Singh           126.390244
Name: batsman_runs, dtype: float64
```

# Merge

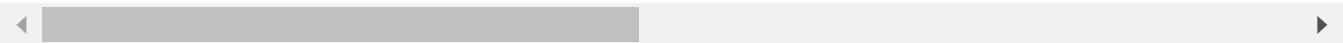
In [64]:

match

Out[64]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal
...	...	...	...	...	...	...	...	...	...
631	632	2016	Raipur	2016-05-22	Delhi Daredevils	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal
632	633	2016	Bangalore	2016-05-24	Gujarat Lions	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal
633	634	2016	Delhi	2016-05-25	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal
634	635	2016	Delhi	2016-05-27	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal
635	636	2016	Bangalore	2016-05-29	Sunrisers Hyderabad	Royal Challengers Bangalore	Sunrisers Hyderabad	bat	normal

636 rows × 18 columns



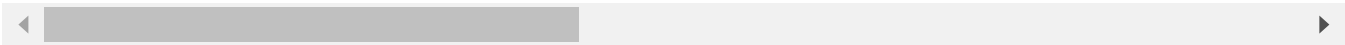
In [65]:

delivery

Out[65]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills
...	...	...	...	...	...	...	...	...	...
150455	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	20	2	Sachin Baby	CJ Jordan	B Kumar
150456	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	20	3	Sachin Baby	CJ Jordan	B Kumar
150457	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	20	4	Iqbal Abdulla	Sachin Baby	B Kumar
150458	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	20	5	Sachin Baby	Iqbal Abdulla	B Kumar
150459	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	20	6	Iqbal Abdulla	Sachin Baby	B Kumar

150460 rows × 21 columns



In [ ]: