M R Kuladeep

In [204]:
```
import pandas as pd
import numpy as np
```

Reading the file

In [18]:
```
data = pd.read_csv('Desktop\matches.csv')
```

In [296]:
```
data.tail()
```

Out[296]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	playe
751	11347	2019	Mumbai	05/05/19	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	9	
752	11412	2019	Chennai	07/05/19	Chennai Super Kings	Mumbai Indians	Chennai Super Kings	bat	normal	0	Mumbai Indians	0	6	
753	11413	2019	Visakhapatnam	08/05/19	Sunrisers Hyderabad	Delhi Capitals	Delhi Capitals	field	normal	0	Delhi Capitals	0	2	
754	11414	2019	Visakhapatnam	10/05/19	Delhi Capitals	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	6	
755	11415	2019	Hyderabad	12/05/19	Mumbai Indians	Chennai Super Kings	Mumbai Indians	bat	normal	0	Mumbai Indians	1	0	

Most successful Teams

In [22]:
```
success = data['winner'].groupby(data['winner']).count()
```

In [23]:
```
success.sort_values(ascending=False)
```

Out[23]:
```
winner
Mumbai Indians                109
Chennai Super Kings           100
Kolkata Knight Riders          92
Royal Challengers Bangalore    84
Kings XI Punjab                82
Rajasthan Royals               75
Delhi Daredevils               67
Sunrisers Hyderabad            58
Deccan Chargers                29
Gujarat Lions                  13
Pune Warriors                  12
Rising Pune Supergiant         10
Delhi Capitals                 10
Kochi Tuskers Kerala            6
Rising Pune Supergiants         5
Name: winner, dtype: int64
```

Successful team in 2017

In [24]:
```
s17 = data[data['season']==2017]
```

In [25]:
```
s17.head()
```

Out[25]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvra
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPI
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	C
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ M
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM

In [26]:
```
ss17=s17['winner'].groupby(s17['winner']).count()
```

In [27]:
```
ss17.sort_values(ascending=False)
```

Out[27]:
```
winner
Mumbai Indians                12
Rising Pune Supergiant        10
Kolkata Knight Riders          9
Sunrisers Hyderabad            8
Kings XI Punjab                7
Delhi Daredevils               6
Gujarat Lions                  4
Royal Challengers Bangalore    3
Name: winner, dtype: int64
```

Players with most man of the matches

In [28]:
```
M=data['player_of_match'].groupby(data['player_of_match']).count()
```

In [29]:
```
M.sort_values(ascending=False).head(10)
```

Out[29]:
```
player_of_match
CH Gayle          21
AB de Villiers    20
MS Dhoni          17
DA Warner         17
RG Sharma         17
YK Pathan         16
SR Watson         15
SK Raina          14
G Gambhir         13
V Kohli           12
Name: player_of_match, dtype: int64
```

Loading file

In [30]:
```
dely = pd.read_csv('Desktop\deliveries.csv')
```

In [31]:
```
dely.head()
```

Out[31]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs	penalty_ru
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

Top Run scorers

In [329]:
```
delys=dely.groupby(['batsman'])['match_id'].nunique()
```

In [330]:
```
delysort= dely['batsman_runs'].groupby(dely['batsman']).sum()
```

In [331]:
```
ds=pd.merge(delys,delysort,on='batsman')
```

In [333]:
```
dss=ds.sort_values(by='batsman_runs',ascending=False)
```

In [338]:
```
dss.columns = ['Matches','Runs']
```

In [339]:
```
dss.head(10)
```

Out[339]:

batsman	Matches	Runs
V Kohli	169	5434
SK Raina	189	5415
RG Sharma	182	4914
DA Warner	126	4741
S Dhawan	158	4632
CH Gayle	124	4560
MS Dhoni	170	4477
RV Uthappa	170	4446
AB de Villiers	142	4428
G Gambhir	151	4223

Highest wicket takers

In [36]:
```
m2=dely.dropna()
```

In [37]:
```
m12=m2[m2['dismissal_kind']!= 'run out']
```

In [38]:
```
m123=m12['player_dismissed'].groupby(m12['bowler']).count()
```

In [39]:
```
m123.sort_values(ascending=False).head(10)
```

Out[39]:
```
bowler
A Mishra           114
DJ Bravo           113
Harbhajan Singh    101
SL Malinga          91
UT Yadav            89
PP Chawla           86
R Ashwin            86
B Kumar             85
R Vinay Kumar       82
A Nehra             76
Name: player_dismissed, dtype: int64
```

Most no. of stumpings

In [40]:
```
m1234=m12[m12['dismissal_kind']=='stumped']
```

In [41]:
```
kula=m1234['player_dismissed'].groupby(m1234['fielder']).count()
```

In [42]:
```
kula.sort_values(ascending=False).head(10)
```

Out[42]:
```
fielder
MS Dhoni         38
RV Uthappa       32
KD Karthik       29
WP Saha          18
AC Gilchrist     16
PA Patel         16
NV Ojha          10
KC Sangakkara     9
RR Pant           9
Q de Kock         8
Name: player_dismissed, dtype: int64
```

No. of matches played

In [348]:
```
data['id'].count()
```

Out[348]:
```
756
```

No. of times a team won when it didn't win the toss

In [44]:
```
d1=data[data['toss_winner']!=data['winner']]
```

In [45]:
```
d12=d1['winner'].groupby(d1['winner']).count()
```

In [46]:
```
d12.sort_values(ascending=False)
```

Out[46]:
```
winner
Mumbai Indians                 53
Kings XI Punjab                47
Royal Challengers Bangalore    43
Chennai Super Kings            43
Kolkata Knight Riders          39
Sunrisers Hyderabad            35
Rajasthan Royals               33
Delhi Daredevils               32
Deccan Chargers                10
Pune Warriors                   9
Rising Pune Supergiant          5
Gujarat Lions                   3
Delhi Capitals                  3
Rising Pune Supergiants         2
Kochi Tuskers Kerala            2
Name: winner, dtype: int64
```

Most extra runs conceded by a bowler

In [47]:
```
delyy=dely['extra_runs'].groupby(dely['bowler']).sum()
```

In [48]:
```
delyy.sort_values(ascending=False).head(10)
```

Out[48]:
```
bowler
SL Malinga         293
P Kumar            236
UT Yadav           219
DJ Bravo           201
B Kumar            197
I Sharma           194
RP Singh           181
SR Watson          171
DW Steyn           171
Harbhajan Singh    170
Name: extra_runs, dtype: int64
```

No of times the toss winner chose to field

In [349]:
```
d1=data[data['toss_decision']=='field']
d1['toss_decision'].count()
```

Out[349]:
```
463
```

No. of times the toss winner chose to field & won the match

In [352]:
```
d1[d1['toss_winner']==d1['winner']]['id'].count()
```

Out[352]:
```
259
```

No. of times toss winner chose to bat

In [51]:
```
d2=data[data['toss_decision']=='bat']
```

In [353]:
```
d2['toss_winner'].count()
```

Out[353]:
```
293
```

No. of times toss winner chose to bat and won the match

In [354]:
```
d2[d2['toss_winner']==d2['winner']]['id'].count()
```

Out[354]:
```
134
```

No. of big margin wins(30 runs) by a team batting first

In [54]:
```
dat1=data[data['win_by_runs']!=0]
```

In [55]:
```
dat12=dat1[dat1['win_by_runs']>=30]
```

In [56]:
```
dat123=dat12['win_by_runs'].groupby(dat12['winner']).count()
```

In [57]:
```
dat123.sort_values(ascending=False)
```

Out[57]:
```
winner
Mumbai Indians                 24
Chennai Super Kings            22
Kolkata Knight Riders          15
Royal Challengers Bangalore    13
Rajasthan Royals               13
Delhi Daredevils               10
Kings XI Punjab                 9
Sunrisers Hyderabad             8
Deccan Chargers                 5
Delhi Capitals                  2
Rising Pune Supergiants         1
Rising Pune Supergiant          1
Pune Warriors                   1
Name: win_by_runs, dtype: int64
```

No. of big margin wins (>4 wickets) by a team bowling first

In [58]:
```
dat4=data[data['win_by_wickets']!=0]
```

In [59]:
```
dat45=dat4[dat4['win_by_wickets']>=4]
```

In [60]:
```
dat456=dat45['win_by_wickets'].groupby(dat45['winner']).count()
```

In [61]:
```
dat456.sort_values(ascending = False)
```

Out[61]:
```
winner
Kolkata Knight Riders          52
Mumbai Indians                 48
Royal Challengers Bangalore    45
Chennai Super Kings            45
Kings XI Punjab                41
Delhi Daredevils               40
Rajasthan Royals               39
Sunrisers Hyderabad            25
Deccan Chargers                11
Gujarat Lions                  10
Pune Warriors                   6
Rising Pune Supergiant          5
Delhi Capitals                  5
Kochi Tuskers Kerala            4
Rising Pune Supergiants         3
Name: win_by_wickets, dtype: int64
```

Most no. of catches

In [373]:
```
m2
```

Out[373]:

bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman_runs	extra_runs	total_runs	player_dismissed	dismissal_kind	fielder
A ɔudhary	0	...	0	0	0	0	0	0	0	DA Warner	caught	Mandeep Singh
R Binny	0	...	0	0	0	0	0	0	0	S Dhawan	caught	Sachin Baby
Chahal	0	...	0	0	0	0	0	0	0	MC Henriques	caught	Sachin Baby
I Hooda	0	...	0	0	0	0	0	0	0	CH Gayle	caught	DA Warner
MC nriques	0	...	0	0	0	0	1	0	1	KM Jadhav	run out	BCJ Cutting
...	...	...	...	...	...	...	...	...	...	...	...	...
KH Pandya	0	...	0	0	0	0	0	0	0	F du Plessis	stumped	Q de Kock
Bumrah	0	...	0	0	0	0	0	0	0	AT Rayudu	caught	Q de Kock
HH Pandya	0	...	0	0	0	0	1	0	1	MS Dhoni	run out	Ishan Kishan
Bumrah	0	...	0	0	0	0	0	0	0	DJ Bravo	caught	Q de Kock
SL Malinga	0	...	0	0	0	0	1	0	1	SR Watson	run out	KH Pandya

In [63]:
```
m3=m2[m2['dismissal_kind']=='caught']
```

In [74]:
```
m4=m3.groupby(m3['fielder'])['fielder'].count()
```

In [73]:
```
m4.sort_values(ascending=False).head(10)
```

Out[73]:
```
fielder
KD Karthik        109
SK Raina           99
MS Dhoni           98
AB de Villiers     93
RV Uthappa         84
RG Sharma          82
KA Pollard         76
V Kohli            73
PA Patel           69
S Dhawan           68
Name: fielder, dtype: int64
```

Matches where dl was applied

In [78]:
```
data[data['dl_applied']!=0]['id'].count()
```

Out[78]:
```
19
```

Bowlers conceding most no balls & penalties

In [88]:
```
dely['exr']=dely['noball_runs']+dely['penalty_runs']
```

In [91]:
```
y1=dely['exr'].groupby(dely['bowler']).sum()
```

In [93]:
```
y1.sort_values(ascending=False).head(10)
```

Out[93]:
```
bowler
S Sreesanth    27
SL Malinga     25
JJ Bumrah      23
I Sharma       21
A Mishra       20
UT Yadav       19
JA Morkel      18
AB Dinda       14
SW Tait        14
SR Watson      13
Name: exr, dtype: int64
```

No. of matches where match is a tie & super over was played

In [97]:
```
dely[dely['is_super_over']!=0]['match_id'].nunique()
```

Out[97]:
```
7
```

In [110]:
```
d10=dely[dely['dismissal_kind']=='run out']
```

In [115]:
```
d11= d10['fielder'].groupby(d10['fielder']).count()
```

Most runouts

In [116]:
```
d11.sort_values(ascending=False).head(10)
```

Out[116]:
```
fielder
MS Dhoni          23
RA Jadeja         20
V Kohli           17
SK Raina          16
MK Pandey         14
AB de Villiers    14
KD Karthik        14
DJ Bravo          12
PA Patel          12
YK Pathan         10
Name: fielder, dtype: int64
```

Best strike rate

In [128]:
```
d7=dely.groupby(dely['batsman'])['ball'].count()
```

In [129]:
```
d8=dely.groupby(dely['batsman'])['batsman_runs'].sum()
```

In [189]:
```
d9=pd.merge(d7,d8, on = 'batsman')
```

In [190]:
```
d9['strike_rate']=round(d9['batsman_runs']*100/d9['ball'],2)
```

In [191]:
```
d10=d9.sort_values(by='strike_rate',ascending=False)
```

batsman playing <500 balls not to be considered in top strike rated players as it is not enough to judge

In [192]:
```
d10[d10['ball']>=500].head(10)
```

Out[192]:

batsman	ball	batsman_runs	strike_rate
AD Russell	803	1445	179.95
RR Pant	1104	1792	162.32
GJ Maxwell	902	1403	155.54
HH Pandya	736	1118	151.90
JC Buttler	954	1431	150.00
V Sehwag	1833	2728	148.83
AB de Villiers	2977	4428	148.74
CH Gayle	3131	4560	145.64

Most economical bowlers (per over)

In [162]:
```
e9=dely.groupby(dely['bowler'])['total_runs'].sum()
```

In [169]:
```
e10=dely.groupby(dely['bowler'])['ball'].count()
```

In [197]:
```
e11=pd.merge(e9,e10 , on='bowler')
```

In [201]:
```
e11['economy']=round(e11['total_runs']*6/e11['ball'],2)
```

In [321]:
```
e12=e11[e11['ball']>=72]
e12.columns = ['total_runs','no. of balls','economy']
```

In [322]:
```
e12.sort_values(by='economy',ascending=True).head(10)
```

Out[322]:

bowler	total_runs	no. of balls	economy
Sohail Tanvir	275	265	6.23
A Chandila	245	234	6.28
FH Edwards	160	150	6.40
L Ngidi	175	163	6.44
SMSM Senanayake	211	195	6.49
J Yadav	248	226	6.58
SM Pollock	307	280	6.58
LH Ferguson	96	87	6.62
A Kumble	1089	983	6.65
GD McGrath	366	329	6.67

Most no. of sixes

```
In [182]: dely1=dely[dely['batsman_runs']==6]

In [183]: dely12=dely1.groupby(dely1['batsman'])['batsman_runs'].count()

In [184]: dely12.sort_values(ascending=False).head(10)
```

```
Out[184]: batsman
          CH Gayle          327
          AB de Villiers    214
          MS Dhoni          207
          SK Raina          195
          RG Sharma         194
          V Kohli           191
          DA Warner         181
          SR Watson         177
          KA Pollard        175
          YK Pathan         161
          Name: batsman_runs, dtype: int64
```

Best Average

```
In [311]: f1=dely.groupby(dely['batsman'])['match_id'].nunique()

In [312]: f2=dely.groupby(dely['batsman'])['batsman_runs'].sum()

In [313]: f3=pd.merge(f1,f2,on='batsman')

In [315]: f3['average']=round(f3['batsman_runs']/f3['match_id'],2)

In [316]: f4=f3.sort_values(by='average',ascending=False)

In [318]: f4.columns = ['no. of matches','total runs','average']

In [320]: f4[f4['no. of matches']>=20].head(10)
```

Out[320]:

batsman	no. of matches	total runs	average
DA Warner	126	4741	37.63
LMP Simmons	29	1079	37.21
CH Gayle	124	4560	36.77
SE Marsh	69	2489	36.07
KL Rahul	58	2013	34.71
ML Hayden	32	1107	34.59
MEK Hussey	58	1977	34.09
RR Pant	54	1792	33.19
KS Williamson	41	1319	32.17
V Kohli	169	5434	32.15

Caught & bowled

```
In [303]: d=dely[dely['bowler']==dely['fielder']]

In [304]: d1=d[d['dismissal_kind']!='run out']

In [323]: d11=d1.groupby(d1['bowler'])['bowler'].count()
          d11.sort_values(ascending=False).head()
```

```
Out[323]: bowler
          S Gopal           5
          Kuldeep Yadav     5
          Imran Tahir       3
          RA Jadeja         2
          Sandeep Sharma    2
          Name: bowler, dtype: int64
```

Average score of a team in an innings

In [340]:
```
av=dely.groupby(dely['batting_team'])['match_id'].nunique()
av1=dely.groupby(dely['batting_team'])['total_runs'].sum()
a=pd.merge(av,av1,on='batting_team')
```

In [347]:
```
a.sort_values(by='Team avg score',ascending=False)[['Matches','Team avg score']]
```

Out[347]:

batting_team	Matches	Team avg score
Delhi Capitals	16	164.0
Gujarat Lions	30	162.0
Chennai Super Kings	164	161.0
Mumbai Indians	187	159.0
Kings XI Punjab	176	158.0
Sunrisers Hyderabad	108	158.0
Royal Challengers Bangalore	180	156.0
Kolkata Knight Riders	178	154.0
Rajasthan Royals	146	154.0
Rising Pune Supergiant	16	154.0
Deccan Chargers	75	153.0
Delhi Daredevils	161	151.0
Rising Pune Supergiants	14	147.0
Pune Warriors	45	141.0
Kochi Tuskers Kerala	14	136.0