Author - Shubham Agarwal

Case Study 1 – IPL

In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

```
ipl18 = pd.DataFrame({'Team': ['SRH', 'CSK', 'KKR', 'RR', 'MI', 'RCB', 'KXIP', 'DD'],
                      'Matches': [14, 14, 14, 14, 14, 14, 14, 14],
                      'Won': [9, 9, 8, 7, 6, 6, 6, 5],
                      'Lost': [5, 5, 6, 7, 8, 8, 8, 9],
                      'Tied': [0, 0, 0, 0, 0, 0, 0, 0],
                      'N/R': [0, 0, 0, 0, 0, 0, 0, 0],
                      'Points': [18, 18, 16, 14, 12, 12, 12, 10],
                      'NRR': [0.284, 0.253, -0.070, -0.250, 0.317, 0.129, -0.502, -0.222
                      'For': [2230, 2488, 2363, 2130, 2380, 2322, 2210, 2297],
                      'Against': [2193, 2433, 2425, 2141, 2282, 2383, 2259, 2304]},
                      index = range(1,9)
                    )

ipl18
```

Out[2]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	SRH	14	9	5	0	0	18	0.284	2230	2193
2	CSK	14	9	5	0	0	18	0.253	2488	2433
3	KKR	14	8	6	0	0	16	-0.070	2363	2425
4	RR	14	7	7	0	0	14	-0.250	2130	2141
5	MI	14	6	8	0	0	12	0.317	2380	2282
6	RCB	14	6	8	0	0	12	0.129	2322	2383
7	KXIP	14	6	8	0	0	12	-0.502	2210	2259
8	DD	14	5	9	0	0	10	-0.222	2297	2304

In [3]:

```
ipl17 = pd.DataFrame({'Team': ['MI', 'RPS', 'SRH', 'KKR', 'KXIP', 'DD', 'GL', 'RCB'],
                      'Matches': [14, 14, 14, 14, 14, 14, 14, 14],
                      'Won': [10, 9, 8, 8, 7, 6, 4, 3],
                      'Lost': [4, 5, 5, 6, 7, 8, 10, 10],
                      'Tied': [0, 0, 0, 0, 0, 0, 0, 0],
                      'N/R': [0, 0, 1, 0, 0, 0, 0, 1],
                      'Points': [20, 18, 17, 16, 14, 12, 8, 7],
                      'NRR': [0.784, 0.176, 0.469, 0.641, 0.123, -0.512, -0.412, -1.299]
                      'For': [2407, 2180, 2221, 2329, 2207, 2219, 2406, 1845],
                      'Against': [2242, 2165, 2118, 2300, 2229, 2255, 2472, 2033]},
                      index = range(1,9)
                     )

ipl17
```

Out[3]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	MI	14	10	4	0	0	20	0.784	2407	2242
2	RPS	14	9	5	0	0	18	0.176	2180	2165
3	SRH	14	8	5	0	1	17	0.469	2221	2118
4	KKR	14	8	6	0	0	16	0.641	2329	2300
5	KXIP	14	7	7	0	0	14	0.123	2207	2229
6	DD	14	6	8	0	0	12	-0.512	2219	2255
7	GL	14	4	10	0	0	8	-0.412	2406	2472
8	RCB	14	3	10	0	1	7	-1.299	1845	2033

Question-1: Suppose in 'ipl18', you want to filter out the teams that have an NRR greater than zero, and for which the 'For' score exceeds the 'Against' score, i.e. both the conditions should be satisfied. Which teams will be left after you perform the above filtration?

a) CSK, MI b) SRH,CSK, MI c) SRH,CSK, RCB d) SRK,CSK, MI,RCB

In [4]:

```
ipl18.sort_values(['NRR'],ascending=False)
```

Out[4]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
5	MI	14	6	8	0	0	12	0.317	2380	2282
1	SRH	14	9	5	0	0	18	0.284	2230	2193
2	CSK	14	9	5	0	0	18	0.253	2488	2433
6	RCB	14	6	8	0	0	12	0.129	2322	2383
3	KKR	14	8	6	0	0	16	-0.070	2363	2425
8	DD	14	5	9	0	0	10	-0.222	2297	2304
4	RR	14	7	7	0	0	14	-0.250	2130	2141
7	KXIP	14	6	8	0	0	12	-0.502	2210	2259

In [5]:

```
print(ipl18['NRR']>0)
```

```
1     True
2     True
3    False
4    False
5     True
6     True
7    False
8    False
Name: NRR, dtype: bool
```

In [6]:

```
df_res= ipl18[(ipl18['NRR']>0) & (ipl18['For']>ipl18['Against'])]
```

In [7]:

```
print(df_res)
```

```
  Team  Matches  Won  Lost  Tied  N/R  Points    NRR   For  Against
1  SRH       14    9     5     0    0      18  0.284  2230     2193
2  CSK       14    9     5     0    0      18  0.253  2488     2433
5   MI       14    6     8     0    0      12  0.317  2380     2282
```

In [8]:

```
print(df_res['Team'])
```

```
1    SRH
2    CSK
5     MI
Name: Team, dtype: object
```

Answer-1: b)

Question-2: If all the stats are taken for both 'ipl17' and 'ipl18', which team with its total points greater than 25 will have the highest win percentage?

In [9]:

```
ipl17.fillna(0,inplace=True)
ipl18.fillna(0,inplace=True)
```

In [10]:

```
ipl17
```

Out[10]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	MI	14	10	4	0	0	20	0.784	2407	2242
2	RPS	14	9	5	0	0	18	0.176	2180	2165
3	SRH	14	8	5	0	1	17	0.469	2221	2118
4	KKR	14	8	6	0	0	16	0.641	2329	2300
5	KXIP	14	7	7	0	0	14	0.123	2207	2229
6	DD	14	6	8	0	0	12	-0.512	2219	2255
7	GL	14	4	10	0	0	8	-0.412	2406	2472
8	RCB	14	3	10	0	1	7	-1.299	1845	2033

In [11]:

```
rslt_df= ipl17.set_index('Team').add(ipl18.set_index('Team'),fill_value=0)
rslt_df
```

Out[11]:

Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
CSK	14.0	9.0	5.0	0.0	0.0	18.0	0.253	2488.0	2433.0
DD	28.0	11.0	17.0	0.0	0.0	22.0	-0.734	4516.0	4559.0
GL	14.0	4.0	10.0	0.0	0.0	8.0	-0.412	2406.0	2472.0
KKR	28.0	16.0	12.0	0.0	0.0	32.0	0.571	4692.0	4725.0
KXIP	28.0	13.0	15.0	0.0	0.0	26.0	-0.379	4417.0	4488.0
MI	28.0	16.0	12.0	0.0	0.0	32.0	1.101	4787.0	4524.0
RCB	28.0	9.0	18.0	0.0	1.0	19.0	-1.170	4167.0	4416.0
RPS	14.0	9.0	5.0	0.0	0.0	18.0	0.176	2180.0	2165.0
RR	14.0	7.0	7.0	0.0	0.0	14.0	-0.250	2130.0	2141.0
SRH	28.0	17.0	10.0	0.0	1.0	35.0	0.753	4451.0	4311.0

In [12]:

```
rslt_df= ipl17.set_index('Team').add(ipl18.set_index('Team'),fill_value=0).reset_index()
rslt_df
```

Out[12]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
0	CSK	14.0	9.0	5.0	0.0	0.0	18.0	0.253	2488.0	2433.0
1	DD	28.0	11.0	17.0	0.0	0.0	22.0	-0.734	4516.0	4559.0
2	GL	14.0	4.0	10.0	0.0	0.0	8.0	-0.412	2406.0	2472.0
3	KKR	28.0	16.0	12.0	0.0	0.0	32.0	0.571	4692.0	4725.0
4	KXIP	28.0	13.0	15.0	0.0	0.0	26.0	-0.379	4417.0	4488.0
5	MI	28.0	16.0	12.0	0.0	0.0	32.0	1.101	4787.0	4524.0
6	RCB	28.0	9.0	18.0	0.0	1.0	19.0	-1.170	4167.0	4416.0
7	RPS	14.0	9.0	5.0	0.0	0.0	18.0	0.176	2180.0	2165.0
8	RR	14.0	7.0	7.0	0.0	0.0	14.0	-0.250	2130.0	2141.0
9	SRH	28.0	17.0	10.0	0.0	1.0	35.0	0.753	4451.0	4311.0

In [13]:

```
rslt_df.sort_values(['Points'],ascending=False)
```

Out[13]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
9	SRH	28.0	17.0	10.0	0.0	1.0	35.0	0.753	4451.0	4311.0
3	KKR	28.0	16.0	12.0	0.0	0.0	32.0	0.571	4692.0	4725.0
5	MI	28.0	16.0	12.0	0.0	0.0	32.0	1.101	4787.0	4524.0
4	KXIP	28.0	13.0	15.0	0.0	0.0	26.0	-0.379	4417.0	4488.0
1	DD	28.0	11.0	17.0	0.0	0.0	22.0	-0.734	4516.0	4559.0
6	RCB	28.0	9.0	18.0	0.0	1.0	19.0	-1.170	4167.0	4416.0
0	CSK	14.0	9.0	5.0	0.0	0.0	18.0	0.253	2488.0	2433.0
7	RPS	14.0	9.0	5.0	0.0	0.0	18.0	0.176	2180.0	2165.0
8	RR	14.0	7.0	7.0	0.0	0.0	14.0	-0.250	2130.0	2141.0
2	GL	14.0	4.0	10.0	0.0	0.0	8.0	-0.412	2406.0	2472.0

In [14]:

```
maximum=0
max_team=0
for t in rslt_df["Team"]:
    d= rslt_df[rslt_df["Team"]==t]
    if(int(d["Points"])>25):
        win_pct = int((d["Won"]/d["Matches"])*100)

        if(win_pct>maximum):
            maximum=win_pct
            print(int(d["Won"]/d["Matches"])*100)
            print("max" , maximum)
            max_team =t
print(t)
```

```
0
max 57
0
max 60
SRH
```

In [ ]:

```

```

In [15]:

```
maximum=0
max_team=0
for t in rslt_df["Team"]:
    d= rslt_df[rslt_df["Team"]==t]
    if(int(d["Points"])>25):
        win_pct = int((d["Won"]/d["Matches"])*100)
        print(win_pct)
        if(win_pct>maximum):
            maximum=win_pct
            print("Percentage",int(d["Won"]/d["Matches"])*100)
            print("max" , maximum)
            max_team =t
print("Team_Name",t)
```

```
57
Percentage 0
max 57
46
57
60
Percentage 0
max 60
Team_Name SRH
```

Answer = SRH