In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
match_data = pd.read_csv("C:/MY WORK/ABES ENG/dataset/IPL Matches 2008-2020.csv")
ball_data = pd.read_csv("C:/MY WORK/ABES ENG/dataset/IPL Ball-by-Ball 2008-2020.csv")
```

In [3]:
```
match_data.head()
```

Out[3]:

	id	city	date	player_of_match	venue	neutral_venue	team1	team2
0	335982	Bangalore	18-04-2008	BB McCullum	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kolkata Knight Riders
1	335983	Chandigarh	19-04-2008	MEK Hussey	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Chennai Super Kings
2	335984	Delhi	19-04-2008	MF Maharoof	Feroz Shah Kotla	0	Delhi Daredevils	Rajasthan Royals
3	335985	Mumbai	20-04-2008	MV Boucher	Wankhede Stadium	0	Mumbai Indians	Royal Challengers Bangalore
4	335986	Kolkata	20-04-2008	DJ Hussey	Eden Gardens	0	Kolkata Knight Riders	Deccan Chargers

In [4]:
```
ball_data.head()
```

Out[4]:

	id	inning	over	ball	batsman	non_striker	bowler	batsman_runs	extra_runs	total_runs
0	335982	1	6	5	RT Ponting	BB McCullum	AA Noffke	1	0	1
1	335982	1	6	6	BB McCullum	RT Ponting	AA Noffke	1	0	1
2	335982	1	7	1	BB McCullum	RT Ponting	Z Khan	0	0	0
3	335982	1	7	2	BB McCullum	RT Ponting	Z Khan	1	0	1
4	335982	1	7	3	RT Ponting	BB McCullum	Z Khan	1	0	1

In [5]:
```
print('Matches played so far:', match_data.shape[0])
print('\n Cities played at:', match_data['city'].unique())
print('\n Teams participated:', match_data['team1'].unique())
```

```
Matches played so far: 816

 Cities played at: ['Bangalore' 'Chandigarh' 'Delhi' 'Mumbai' 'Kolkata' 'Jaipur' 'Hyd
erabad'
 'Chennai' 'Cape Town' 'Port Elizabeth' 'Durban' 'Centurion' 'East London'
 'Johannesburg' 'Kimberley' 'Bloemfontein' 'Ahmedabad' 'Cuttack' 'Nagpur'
 'Dharamsala' 'Kochi' 'Indore' 'Visakhapatnam' 'Pune' 'Raipur' 'Ranchi'
 'Abu Dhabi' nan 'Rajkot' 'Kanpur' 'Bengaluru' 'Dubai' 'Sharjah']

 Teams participated: ['Royal Challengers Bangalore' 'Kings XI Punjab' 'Delhi Daredevi
ls'
 'Mumbai Indians' 'Kolkata Knight Riders' 'Rajasthan Royals'
 'Deccan Chargers' 'Chennai Super Kings' 'Kochi Tuskers Kerala'
 'Pune Warriors' 'Sunrisers Hyderabad' 'Gujarat Lions'
 'Rising Pune Supergiants' 'Rising Pune Supergiant' 'Delhi Capitals']
```

In [8]:
```
match_per_season = match_data.groupby(['Season'])['id'].count().reset_index().rename(c
match_per_season
```

Out[8]:

	Season	matches
0	2008	58
1	2009	57
2	2010	60
3	2011	73
4	2012	74
5	2013	76
6	2014	60
7	2015	59
8	2016	60
9	2017	59
10	2018	60
11	2019	60
12	2020	60

In [10]:
```
player = (ball_data['batsman']=='SK Raina')
df_raina=ball_data[player]
df_raina.head()
```

Out[10]:

	id	inning	over	ball	batsman	non_striker	bowler	batsman_runs	extra_runs	total_runs
246	335983	1	10	3	SK Raina	MEK Hussey	PP Chawla	2	0	2
247	335983	1	10	4	SK Raina	MEK Hussey	PP Chawla	0	0	0
248	335983	1	10	5	SK Raina	MEK Hussey	PP Chawla	6	0	6
249	335983	1	10	6	SK Raina	MEK Hussey	PP Chawla	4	0	4
253	335983	1	11	4	SK Raina	MEK Hussey	K Goel	6	0	6

In [11]:
```
runs = ball_data.groupby(['batsman'])['batsman_runs'].sum().reset_index()
runs.columns = ['Batsman', 'runs']
y = runs.sort_values(by='runs', ascending = False).head(10).reset_index().drop('index'
y
```

Out[11]:

	Batsman	runs
0	V Kohli	5878
1	SK Raina	5368
2	DA Warner	5254
3	RG Sharma	5230
4	S Dhawan	5197
5	AB de Villiers	4849
6	CH Gayle	4772
7	MS Dhoni	4632
8	RV Uthappa	4607
9	G Gambhir	4217

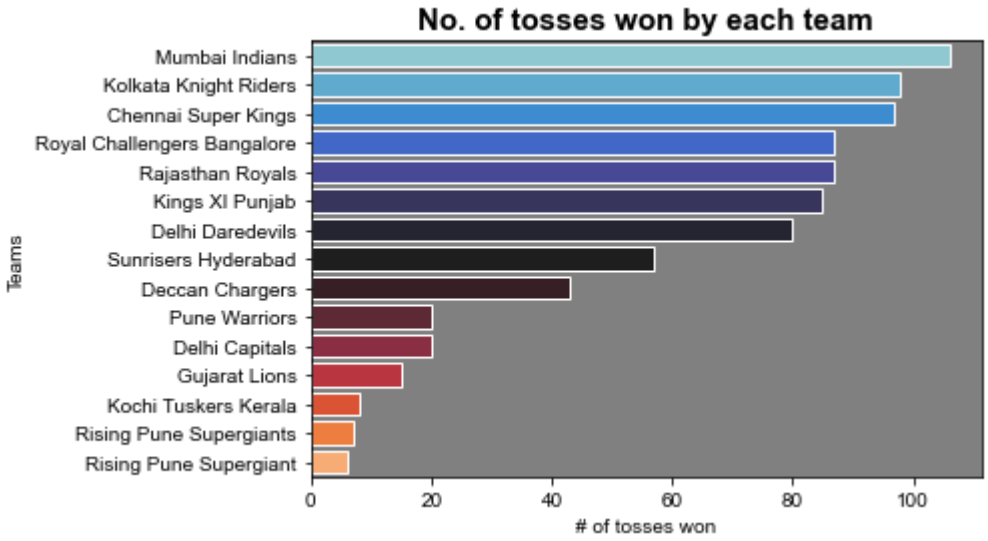
In [12]:
```
def count(df_raina,runs):
    return len(df_raina[df_raina['batsman_runs']==runs])*runs
```

In [13]:
```
print("Runs scored from 1's :",count(df_raina,1))
print("Runs scored from 2's :",count(df_raina,2))
print("Runs scored from 3's :",count(df_raina,3))
print("Runs scored from 4's :",count(df_raina,4))
print("Runs scored from 6's :",count(df_raina,6))
```

```
Runs scored from 1's : 1666
Runs scored from 2's : 528
Runs scored from 3's : 33
Runs scored from 4's : 1972
Runs scored from 6's : 1164
```

In [14]:
```
toss=match_data['toss_winner'].value_counts()
```

```
ax = plt.axes()
ax.set(facecolor = "grey")
sns.set(rc={'figure.figsize':(15,10)},style='darkgrid')
ax.set_title('No. of tosses won by each team',fontsize=15,fontweight="bold")
sns.barplot(y=toss.index, x=toss, orient='h',palette="icefire",saturation=1)
plt.xlabel('# of tosses won')
plt.ylabel('Teams')
plt.show()
```



In [ ]: