

ipl-match-copy1-checkpoint

June 27, 2023

```
[72]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import warnings
warnings.simplefilter(action = "ignore", category = FutureWarning)
```

```
[73]: data = pd.read_csv("data.csv")
```

```
[74]: data.head(100)
```

```
[74]:
```

	id	season	city	date	team1 \
0	1	2008	Bangalore	2008-04-18	Kolkata Knight Riders
1	2	2008	Chandigarh	2008-04-19	Chennai Super Kings
2	3	2008	Delhi	2008-04-19	Rajasthan Royals
3	4	2008	Mumbai	2008-04-20	Mumbai Indians
4	5	2008	Kolkata	2008-04-20	Deccan Chargers
..
95	96	2009	Kimberley	2009-05-11	Deccan Chargers
96	97	2009	Centurion	2009-05-12	Kolkata Knight Riders
97	98	2009	Centurion	2009-05-12	Kings XI Punjab
98	99	2009	Durban	2009-05-13	Delhi Daredevils
99	100	2009	Durban	2009-05-14	Chennai Super Kings

		team2	toss_winner	toss_decision \
0	Royal Challengers Bangalore	Royal Challengers Bangalore	field	
1		Kings XI Punjab	Chennai Super Kings	bat
2		Delhi Daredevils	Rajasthan Royals	bat
3	Royal Challengers Bangalore		Mumbai Indians	bat
4		Kolkata Knight Riders	Deccan Chargers	bat
..	
95		Rajasthan Royals	Deccan Chargers	bat
96	Royal Challengers Bangalore	Royal Challengers Bangalore	field	
97		Mumbai Indians	Kings XI Punjab	bat
98		Deccan Chargers	Deccan Chargers	field
99	Royal Challengers Bangalore		Chennai Super Kings	bat

	result	dl_applied	winner	win_by_runs	\
0	normal	0	Kolkata Knight Riders	140	
1	normal	0	Chennai Super Kings	33	
2	normal	0	Delhi Daredevils	0	
3	normal	0	Royal Challengers Bangalore	0	
4	normal	0	Kolkata Knight Riders	0	
..	
95	normal	0	Deccan Chargers	53	
96	normal	0	Royal Challengers Bangalore	0	
97	normal	0	Mumbai Indians	0	
98	normal	0	Delhi Daredevils	12	
99	normal	0	Royal Challengers Bangalore	0	

	win_by_wickets	player_of_match	\
0	0	BB McCullum	
1	0	MEK Hussey	
2	9	MF Maharoo	
3	5	MV Boucher	
4	5	DJ Hussey	
..	
95	0	DR Smith	
96	6	LRPL Taylor	
97	8	Harbhajan Singh	
98	0	R Bhatia	
99	2	LRPL Taylor	

	venue	umpire1	umpire2	\
0	M Chinnaswamy Stadium	Asad Rauf	RE Koertzen	
1	Punjab Cricket Association Stadium, Mohali	MR Benson	SL Shastri	
2	Feroz Shah Kotla	Aleem Dar	GA Pratapkumar	
3	Wankhede Stadium	SJ Davis	DJ Harper	
4	Eden Gardens	BF Bowden	K Hariharan	
..	
95	De Beers Diamond Oval	GAV Baxter	HDPK Dharmasena	
96	SuperSport Park	M Erasmus	SS Hazare	
97	SuperSport Park	SS Hazare	RE Koertzen	
98	Kingsmead	DJ Harper	SL Shastri	
99	Kingsmead	BR Doctrove	DJ Harper	

	umpire3
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN
..	...

```

95      NaN
96      NaN
97      NaN
98      NaN
99      NaN

```

```
[100 rows x 18 columns]
```

```
[75]: data.columns
```

```
[75]: Index(['id', 'season', 'city', 'date', 'team1', 'team2', 'toss_winner',
          'toss_decision', 'result', 'dl_applied', 'winner', 'win_by_runs',
          'win_by_wickets', 'player_of_match', 'venue', 'umpire1', 'umpire2',
          'umpire3'],
          dtype='object')
```

```
[76]: data_indexing = data.set_index(['id','season'])
```

```
[77]: data_indexing
```

```
[77]:
```

		city	date	team1 \
id	season			
1	2008	Bangalore	2008-04-18	Kolkata Knight Riders
2	2008	Chandigarh	2008-04-19	Chennai Super Kings
3	2008	Delhi	2008-04-19	Rajasthan Royals
4	2008	Mumbai	2008-04-20	Mumbai Indians
5	2008	Kolkata	2008-04-20	Deccan Chargers
...	
573	2016	Raipur	2016-05-22	Delhi Daredevils
574	2016	Bangalore	2016-05-24	Gujarat Lions
575	2016	Delhi	2016-05-25	Sunrisers Hyderabad
576	2016	Delhi	2016-05-27	Gujarat Lions
577	2016	Bangalore	2016-05-29	Sunrisers Hyderabad

			team2	toss_winner \
id	season			
1	2008	Royal Challengers Bangalore	Royal Challengers Bangalore	
2	2008	Kings XI Punjab	Chennai Super Kings	
3	2008	Delhi Daredevils	Rajasthan Royals	
4	2008	Royal Challengers Bangalore	Mumbai Indians	
5	2008	Kolkata Knight Riders	Deccan Chargers	
...	
573	2016	Royal Challengers Bangalore	Royal Challengers Bangalore	
574	2016	Royal Challengers Bangalore	Royal Challengers Bangalore	
575	2016	Kolkata Knight Riders	Kolkata Knight Riders	
576	2016	Sunrisers Hyderabad	Sunrisers Hyderabad	
577	2016	Royal Challengers Bangalore	Sunrisers Hyderabad	

id	season	toss_decision	result	dl_applied	winner \
1	2008	field	normal	0	Kolkata Knight Riders
2	2008	bat	normal	0	Chennai Super Kings
3	2008	bat	normal	0	Delhi Daredevils
4	2008	bat	normal	0	Royal Challengers Bangalore
5	2008	bat	normal	0	Kolkata Knight Riders
...	
573	2016	field	normal	0	Royal Challengers Bangalore
574	2016	field	normal	0	Royal Challengers Bangalore
575	2016	field	normal	0	Sunrisers Hyderabad
576	2016	field	normal	0	Sunrisers Hyderabad
577	2016	bat	normal	0	Sunrisers Hyderabad

id	season	win_by_runs	win_by_wickets	player_of_match \
1	2008	140	0	BB McCullum
2	2008	33	0	MEK Hussey
3	2008	0	9	MF Maharoo
4	2008	0	5	MV Boucher
5	2008	0	5	DJ Hussey
...	
573	2016	0	6	V Kohli
574	2016	0	4	AB de Villiers
575	2016	22	0	MC Henriques
576	2016	0	4	DA Warner
577	2016	8	0	BCJ Cutting

id	season	venue	umpire1 \
1	2008	M Chinnaswamy Stadium	Asad Rauf
2	2008	Punjab Cricket Association Stadium, Mohali	MR Benson
3	2008	Feroz Shah Kotla	Aleem Dar
4	2008	Wankhede Stadium	SJ Davis
5	2008	Eden Gardens	BF Bowden
...	
573	2016	Shaheed Veer Narayan Singh International Stadium	A Nand Kishore
574	2016	M Chinnaswamy Stadium	AK Chaudhary
575	2016	Feroz Shah Kotla	M Erasmus
576	2016	Feroz Shah Kotla	M Erasmus
577	2016	M Chinnaswamy Stadium	HDPK Dharmasena

id	season	umpire2	umpire3
1	2008	RE Koertzen	NaN
2	2008	SL Shastri	NaN

3	2008	GA Pratapkumar	NaN
4	2008	DJ Harper	NaN
5	2008	K Hariharan	NaN
...	
573	2016	BNJ Oxenford	NaN
574	2016	HDPK Dharmasena	NaN
575	2016	C Shamshuddin	NaN
576	2016	CK Nandan	NaN
577	2016	BNJ Oxenford	NaN

[577 rows x 16 columns]

```
[78]: data_plyr_season = data.set_index(['season','player_of_match'])
```

```
[79]: data_plyr_season.head(72)
```

```
[79]:
```

		id	city	date	\
season	player_of_match				
2008	BB McCullum	1	Bangalore	2008-04-18	
	MEK Hussey	2	Chandigarh	2008-04-19	
	MF Maharooof	3	Delhi	2008-04-19	
	MV Boucher	4	Mumbai	2008-04-20	
	DJ Hussey	5	Kolkata	2008-04-20	
...		
2009	RS Bopara	68	Durban	2009-04-24	
	PP Ojha	69	Durban	2009-04-25	
	TM Dilshan	70	Port Elizabeth	2009-04-26	
	KC Sangakkara	71	Cape Town	2009-04-26	
	HH Gibbs	72	Durban	2009-04-27	

		team1	\
season	player_of_match		
2008	BB McCullum	Kolkata Knight Riders	
	MEK Hussey	Chennai Super Kings	
	MF Maharooof	Rajasthan Royals	
	MV Boucher	Mumbai Indians	
	DJ Hussey	Deccan Chargers	
...		...	
2009	RS Bopara	Royal Challengers Bangalore	
	PP Ojha	Deccan Chargers	
	TM Dilshan	Royal Challengers Bangalore	
	KC Sangakkara	Kings XI Punjab	
	HH Gibbs	Chennai Super Kings	

		team2	\
season	player_of_match		
2008	BB McCullum	Royal Challengers Bangalore	

	MEK Hussey	Kings XI Punjab
	MF Maharooof	Delhi Daredevils
	MV Boucher	Royal Challengers Bangalore
	DJ Hussey	Kolkata Knight Riders

...

2009	RS Bopara	Kings XI Punjab
	PP Ojha	Mumbai Indians
	TM Dilshan	Delhi Daredevils
	KC Sangakkara	Rajasthan Royals
	HH Gibbs	Deccan Chargers

		toss_winner	toss_decision	result	\
season	player_of_match				
2008	BB McCullum	Royal Challengers Bangalore	field	normal	
	MEK Hussey	Chennai Super Kings	bat	normal	
	MF Maharooof	Rajasthan Royals	bat	normal	
	MV Boucher	Mumbai Indians	bat	normal	
	DJ Hussey	Deccan Chargers	bat	normal	
...					
2009	RS Bopara	Royal Challengers Bangalore	bat	normal	
	PP Ojha	Deccan Chargers	bat	normal	
	TM Dilshan	Royal Challengers Bangalore	bat	normal	
	KC Sangakkara	Kings XI Punjab	bat	normal	
	HH Gibbs	Deccan Chargers	field	normal	

		dl_applied	winner	win_by_runs	\
season	player_of_match				
2008	BB McCullum	0	Kolkata Knight Riders	140	
	MEK Hussey	0	Chennai Super Kings	33	
	MF Maharooof	0	Delhi Daredevils	0	
	MV Boucher	0	Royal Challengers Bangalore	0	
	DJ Hussey	0	Kolkata Knight Riders	0	
...					
2009	RS Bopara	0	Kings XI Punjab	0	
	PP Ojha	0	Deccan Chargers	12	
	TM Dilshan	0	Delhi Daredevils	0	
	KC Sangakkara	0	Kings XI Punjab	27	
	HH Gibbs	0	Deccan Chargers	0	

		win_by_wickets	\
season	player_of_match		
2008	BB McCullum	0	
	MEK Hussey	0	
	MF Maharooof	9	
	MV Boucher	5	
	DJ Hussey	5	
...			

2009	RS Bopara	7
	PP Ojha	0
	TM Dilshan	6
	KC Sangakkara	0
	HH Gibbs	6

season	player_of_match	venue
2008	BB McCullum	M Chinnaswamy Stadium
	MEK Hussey	Punjab Cricket Association Stadium, Mohali
	MF Maharooof	Feroz Shah Kotla
	MV Boucher	Wankhede Stadium
	DJ Hussey	Eden Gardens
...		...
2009	RS Bopara	Kingsmead
	PP Ojha	Kingsmead
	TM Dilshan	St George's Park
	KC Sangakkara	Newlands
	HH Gibbs	Kingsmead

season	player_of_match	umpire1	umpire2	umpire3
2008	BB McCullum	Asad Rauf	RE Koertzen	NaN
	MEK Hussey	MR Benson	SL Shastri	NaN
	MF Maharooof	Aleem Dar	GA Pratapkumar	NaN
	MV Boucher	SJ Davis	DJ Harper	NaN
	DJ Hussey	BF Bowden	K Hariharan	NaN
...	
2009	RS Bopara	BR Doctrove	TH Wijewardene	NaN
	PP Ojha	HDPK Dharmasena	SJA Taufel	NaN
	TM Dilshan	S Asnani	BG Jerling	NaN
	KC Sangakkara	M Erasmus	K Hariharan	NaN
	HH Gibbs	IL Howell	TH Wijewardene	NaN

[72 rows x 16 columns]

```
[80]: # data1=data.groupby(['season','team1','team2']).mean() mohit
```

```
[81]: # data1.head()
```

```
[82]: data2=data.groupby(['team1','team2']).size()
```

```
[83]: data2
```

team1	team2	
Chennai Super Kings	Deccan Chargers	8
	Delhi Daredevils	11

```

                Kings XI Punjab          9
                Kochi Tuskers Kerala      2
                Kolkata Knight Riders     11
                ..
Sunrisers Hyderabad Mumbai Indians      5
                Pune Warriors            2
                Rajasthan Royals         7
                Rising Pune Supergiants   2
                Royal Challengers Bangalore 5
Length: 113, dtype: int64

```

```
[84]: data5=data.groupby(['season','winner']).sum()[['win_by_runs','win_by_wickets']]
```

```
[85]: data5
```

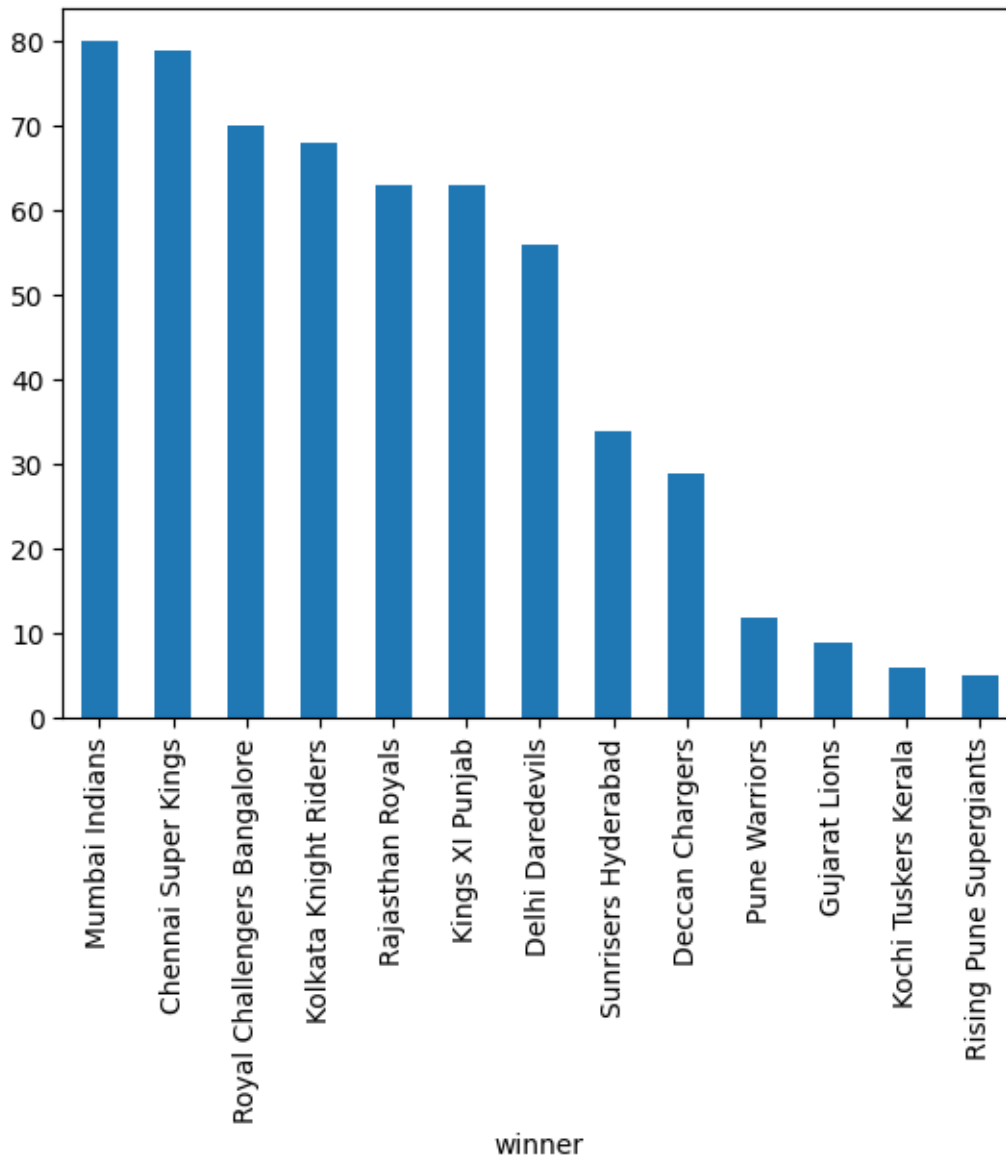
```
[85]:
```

		win_by_runs	win_by_wickets
season	winner		
2008	Chennai Super Kings	73	29
	Deccan Chargers	0	17
	Delhi Daredevils	22	36
	Kings XI Punjab	123	32
	Kolkata Knight Riders	191	8
...	
2016	Kolkata Knight Riders	29	38
	Mumbai Indians	105	32
	Rising Pune Supergiants	53	20
	Royal Challengers Bangalore	285	26
	Sunrisers Hyderabad	134	38

```
[76 rows x 2 columns]
```

```
[86]: data["winner"].value_counts().plot(kind= "bar")
```

```
[86]: <Axes: xlabel='winner'>
```

```
[87]: grouped = data.groupby(["season", "player_of_match"])["winner"]
winner_by_season = grouped.apply(lambda x: x.value_counts())

print(winner_by_season)
```

season	player_of_match	
2008	A Kumble	Royal Challengers Bangalore 1
	A Mishra	Delhi Daredevils 1
	A Nehra	Mumbai Indians 1
	AC Gilchrist	Deccan Chargers 2
	BB McCullum	Kolkata Knight Riders 1

```

2016    SR Watson      Royal Challengers Bangalore    1
        SV Samson      Delhi Daredevils          1
        V Kohli        Royal Challengers Bangalore    4
                        Gujarat Lions            1
        YK Pathan      Kolkata Knight Riders      2
Name: winner, Length: 395, dtype: int64

```

```
[88]: a=data.
      ↪loc[(data["winner"]==data["toss_winner"]),['season','winner','toss_decision']]
```

```
[89]: b=a.groupby(['season', 'winner']).count()
      b
      #a.groupby(['season', 'winner', 'toss_decision']).count()

      #a[a.toss_decision=='bat'].count()
```

```
[89]:                                     toss_decision
season winner
2008  Chennai Super Kings                3
      Deccan Chargers                    2
      Delhi Daredevils                    2
      Kings XI Punjab                     4
      Kolkata Knight Riders               3
...
2016  Kolkata Knight Riders                4
      Mumbai Indians                      5
      Rising Pune Supergiants             3
      Royal Challengers Bangalore         4
      Sunrisers Hyderabad                 7
```

[71 rows x 1 columns]

```
[90]: toss_winner=data[data['winner']==data['toss_winner']].count()
      print(toss_winner['winner'])
```

291

```
[91]: grouped = data.groupby("season")["toss_winner"]
      toss_winner_by_season = grouped.apply(lambda x: x.value_counts())

      print(toss_winner_by_season)
```

```

season
2008  Rajasthan Royals                11
      Deccan Chargers                 9
      Mumbai Indians                  8
      Kings XI Punjab                 8

```

```

        Kolkata Knight Riders        6
        ..
2016   Delhi Daredevils              8
        Rising Pune Supergiants      7
        Kolkata Knight Riders        6
        Royal Challengers Bangalore  6
        Kings XI Punjab              6
Name: toss_winner, Length: 76, dtype: int64

```

```
[92]: # win_loss_ratio
```

```
[93]: csk2008won = len(data[(data.winner=="Chennai Super Kings") & (data.
    ↪season==2008)])

csk2008played = len(data[((data.team1=="Chennai Super Kings") | (data.
    ↪team2=="Chennai Super Kings"))& (data.season==2008)])

csk2008wr = csk2008won/csk2008played

print("csk2008played",csk2008played)
print("csk2008won",csk2008won)
print("csk2008 win loss ratio",csk2008wr)
```

```

csk2008played 16
csk2008won 9
csk2008 win loss ratio 0.5625

```

```
[94]: dc2008won = len(data[(data.winner=="Deccan Chargers") & (data.season==2008)])

dc2008played = len(data[((data.team1=="Deccan Chargers") | (data.team2=="Deccan_
    ↪Chargers"))& (data.season==2008)])

dc2008wr = dc2008won/dc2008played

print("dc2008played",dc2008played)
print("dc2008won",dc2008won)
print("dc2008 win loss ratio",dc2008wr)
```

```

dc2008played 14
dc2008won 2
dc2008 win loss ratio 0.14285714285714285

```

```
[95]: # # This is a for loop that iterates over the range of years from 2008 to 2016.
    ↪Within the loop, it calculates the number of matches won by the Chennai_
    ↪Super Kings (CSK) and the number of matches played by CSK in each season. It_
    ↪then calculates the win-loss ratio of CSK in each season and prints out the_
    ↪results.
```

```
# for i in range(2008,2017):
# #     cskwon=len(data[(data.winner=="Chennai Super Kings") & (data.
# #         ↪season==i)])
# #     cskplayed = len(data[((data.team1=="Chennai Super Kings") | (data.
# #         ↪team2=="Chennai Super Kings"))& (data.season==i)])
# #     cskwr = cskwon/cskplayed
# #     print("csk played in",i,cskplayed)
# #     print("cskwon in",i,cskwon)
# #     print("csk win loss ratio in ",i,cskwr)
```

```
[96]: for i in range(2008,2017):
        dcwon=len(data[(data.winner=="Deccan Chargers") & (data.season==i)])
        dcplayed = len(data[((data.team1=="Deccan Chargers") | (data.team2=="Deccan_
        ↪Chargers"))& (data.season==i)])
        print("dc played in",i,dcplayed)
        print("dcwon in",i,dcwon)
        print("dc win loss ratio in ",i,dcwr)
```

```
dc played in 2008 14
dcwon in 2008 2
dc win loss ratio in 2008 0.26666666666666666
dc played in 2009 16
dcwon in 2009 9
dc win loss ratio in 2009 0.26666666666666666
dc played in 2010 16
dcwon in 2010 8
dc win loss ratio in 2010 0.26666666666666666
dc played in 2011 14
dcwon in 2011 6
dc win loss ratio in 2011 0.26666666666666666
dc played in 2012 15
dcwon in 2012 4
dc win loss ratio in 2012 0.26666666666666666
dc played in 2013 0
dcwon in 2013 0
dc win loss ratio in 2013 0.26666666666666666
dc played in 2014 0
dcwon in 2014 0
dc win loss ratio in 2014 0.26666666666666666
dc played in 2015 0
dcwon in 2015 0
dc win loss ratio in 2015 0.26666666666666666
dc played in 2016 0
dcwon in 2016 0
dc win loss ratio in 2016 0.26666666666666666
```

```
[97]: for i in range(2008,2017):
        ddwon=len(data[(data.winner=="Delhi Daredevils") & (data.season==i)])
        ddplayed = len(data[((data.team1=="Delhi Daredevils") | (data.team2=="Delhi
↳Daredevils"))& (data.season==i)])
        ddwr = ddwon/ddplayed
        print("dd played in",i,ddplayed)
        print("ddwon in",i,ddwon)
        print("dd win loss ratio in ",i,ddwr)
```

```
dd played in 2008 14
ddwon in 2008 7
dd win loss ratio in 2008 0.5
dd played in 2009 15
ddwon in 2009 10
dd win loss ratio in 2009 0.6666666666666666
dd played in 2010 14
ddwon in 2010 7
dd win loss ratio in 2010 0.5
dd played in 2011 14
ddwon in 2011 4
dd win loss ratio in 2011 0.2857142857142857
dd played in 2012 18
ddwon in 2012 11
dd win loss ratio in 2012 0.6111111111111112
dd played in 2013 16
ddwon in 2013 3
dd win loss ratio in 2013 0.1875
dd played in 2014 14
ddwon in 2014 2
dd win loss ratio in 2014 0.14285714285714285
dd played in 2015 14
ddwon in 2015 5
dd win loss ratio in 2015 0.35714285714285715
dd played in 2016 14
ddwon in 2016 7
dd win loss ratio in 2016 0.5
```

```
[98]: for i in range(2008,2017):
        kpwon=len(data[(data.winner=="Kings XI Punjab") & (data.season==i)])
        kpplayed = len(data[((data.team1=="Kings XI Punjab") | (data.team2=="Kings
↳XI Punjab"))& (data.season==i)])
        kpwr = kpwon/kpplayed
        print("kp played in",i,kpplayed)
        print("kpwon in",i,kpwon)
        print("kp win loss ratio in ",i,kpwr)
```

```
kp played in 2008 15
```

```

kpwon in 2008 10
kp win loss ratio in 2008 0.6666666666666666
kp played in 2009 14
kpwon in 2009 7
kp win loss ratio in 2009 0.5
kp played in 2010 14
kpwon in 2010 4
kp win loss ratio in 2010 0.2857142857142857
kp played in 2011 14
kpwon in 2011 7
kp win loss ratio in 2011 0.5
kp played in 2012 16
kpwon in 2012 8
kp win loss ratio in 2012 0.5
kp played in 2013 16
kpwon in 2013 8
kp win loss ratio in 2013 0.5
kp played in 2014 17
kpwon in 2014 12
kp win loss ratio in 2014 0.7058823529411765
kp played in 2015 14
kpwon in 2015 3
kp win loss ratio in 2015 0.21428571428571427
kp played in 2016 14
kpwon in 2016 4
kp win loss ratio in 2016 0.2857142857142857

```

```

[99]: for i in range(2008,2017):
        kkrwon=len(data[(data.winner=="Kolkata Knight Riders") & (data.season==i)])
        kkrplayed = len(data[((data.team1=="Kolkata Knight Riders") | (data.
        team2=="Kolkata Knight Riders"))& (data.season==i)])
        kkrwr = kkrwon/kkrplayed
        print("kkr played in",i,kkrplayed)
        print("kkrwon in",i,kkrwon)
        print("kkr win loss ratio in ",i,kkrwr)

```

```

kkr played in 2008 13
kkrwon in 2008 6
kkr win loss ratio in 2008 0.46153846153846156
kkr played in 2009 13
kkrwon in 2009 3
kkr win loss ratio in 2009 0.23076923076923078
kkr played in 2010 14
kkrwon in 2010 7
kkr win loss ratio in 2010 0.5
kkr played in 2011 15
kkrwon in 2011 8
kkr win loss ratio in 2011 0.5333333333333333

```

```

kkr played in 2012 17
kkrwon in 2012 12
kkr win loss ratio in 2012 0.7058823529411765
kkr played in 2013 16
kkrwon in 2013 6
kkr win loss ratio in 2013 0.375
kkr played in 2014 16
kkrwon in 2014 11
kkr win loss ratio in 2014 0.6875
kkr played in 2015 13
kkrwon in 2015 7
kkr win loss ratio in 2015 0.5384615384615384
kkr played in 2016 15
kkrwon in 2016 8
kkr win loss ratio in 2016 0.5333333333333333

```

```

[100]: for i in range(2008,2017):
        kkrwon=len(data[(data.winner=="Kolkata Knight Riders") & (data.season==i)])
        kkrplayed = len(data[((data.team1=="Kolkata Knight Riders") | (data.
        team2=="Kolkata Knight Riders"))& (data.season==i)])
        kkrwr = kkrwon/kkrplayed
        print("kkr played in",i,kkrplayed)
        print("kkrwon in",i,kkrwon)
        print("kkr win loss ratio in ",i,kkrwr)

```

```

kkr played in 2008 13
kkrwon in 2008 6
kkr win loss ratio in 2008 0.46153846153846156
kkr played in 2009 13
kkrwon in 2009 3
kkr win loss ratio in 2009 0.23076923076923078
kkr played in 2010 14
kkrwon in 2010 7
kkr win loss ratio in 2010 0.5
kkr played in 2011 15
kkrwon in 2011 8
kkr win loss ratio in 2011 0.5333333333333333
kkr played in 2012 17
kkrwon in 2012 12
kkr win loss ratio in 2012 0.7058823529411765
kkr played in 2013 16
kkrwon in 2013 6
kkr win loss ratio in 2013 0.375
kkr played in 2014 16
kkrwon in 2014 11
kkr win loss ratio in 2014 0.6875
kkr played in 2015 13
kkrwon in 2015 7

```

```

kkr win loss ratio in 2015 0.5384615384615384
kkr played in 2016 15
kkrwon in 2016 8
kkr win loss ratio in 2016 0.5333333333333333

```

```
[101]: data.head()
```

```

[101]:   id  season      city      date      team1 \
0    1    2008  Bangalore  2008-04-18  Kolkata Knight Riders
1    2    2008  Chandigarh  2008-04-19    Chennai Super Kings
2    3    2008     Delhi  2008-04-19    Rajasthan Royals
3    4    2008    Mumbai  2008-04-20    Mumbai Indians
4    5    2008    Kolkata  2008-04-20    Deccan Chargers

      team2      toss_winner toss_decision \
0  Royal Challengers Bangalore  Royal Challengers Bangalore      field
1           Kings XI Punjab      Chennai Super Kings      bat
2           Delhi Daredevils      Rajasthan Royals      bat
3  Royal Challengers Bangalore      Mumbai Indians      bat
4           Kolkata Knight Riders      Deccan Chargers      bat

      result  dl_applied      winner  win_by_runs \
0  normal          0      Kolkata Knight Riders      140
1  normal          0      Chennai Super Kings      33
2  normal          0      Delhi Daredevils      0
3  normal          0  Royal Challengers Bangalore      0
4  normal          0      Kolkata Knight Riders      0

      win_by_wickets  player_of_match      venue \
0          0      BB McCullum      M Chinnaswamy Stadium
1          0      MEK Hussey  Punjab Cricket Association Stadium, Mohali
2          9      MF Maharoor      Feroz Shah Kotla
3          5      MV Boucher      Wankhede Stadium
4          5      DJ Hussey      Eden Gardens

      umpire1      umpire2  umpire3
0  Asad Rauf      RE Koertzen      NaN
1  MR Benson      SL Shastri      NaN
2  Aleem Dar  GA Pratapkumar      NaN
3   SJ Davis      DJ Harper      NaN
4  BF Bowden      K Hariharan      NaN

```

```

[102]: man_of_the_match=data['player_of_match'].value_counts()[:20]
man_of_the_match
# man_of_the_match=data['player_of_match'].value_counts()[:20]
# man_of_the_match=pd.DataFrame("player":{man_of_the_match})
# man_of_the_match

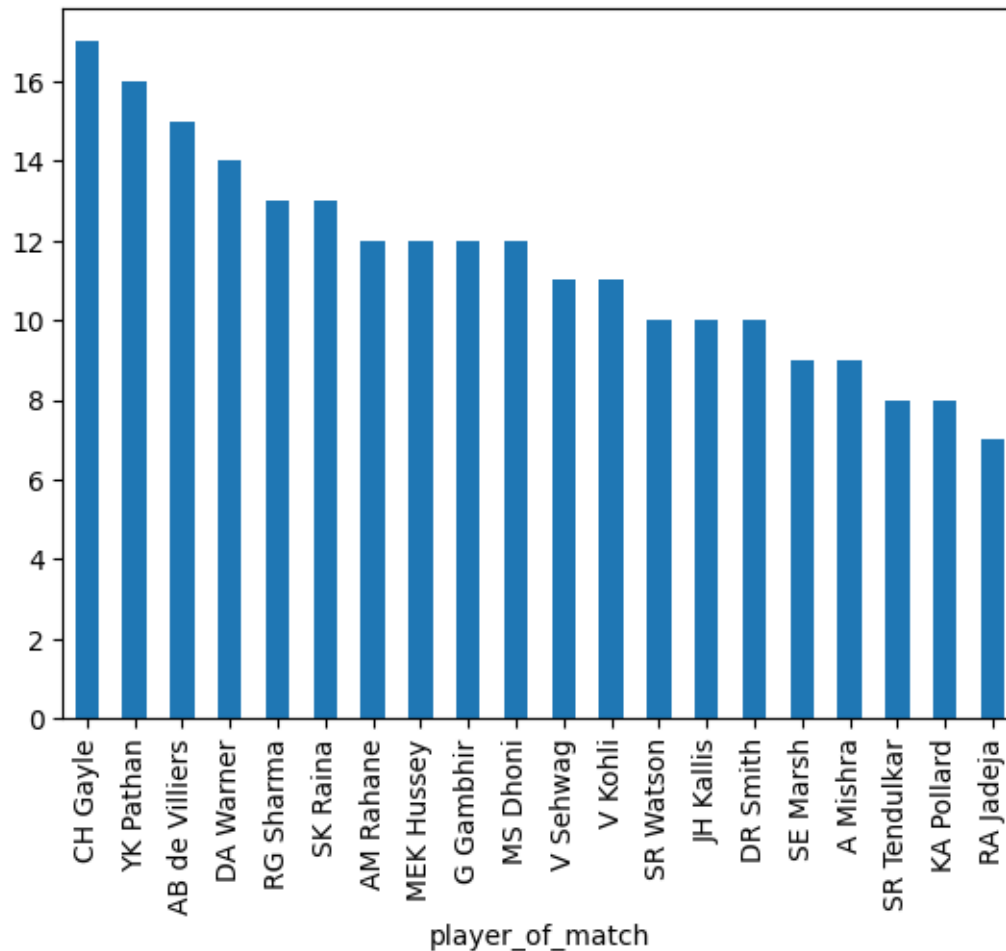
```



```
[102]: player_of_match
      CH Gayle      17
      YK Pathan    16
      AB de Villiers 15
      DA Warner    14
      RG Sharma    13
      SK Raina     13
      AM Rahane    12
      MEK Hussey   12
      G Gambhir    12
      MS Dhoni     12
      V Sehwag     11
      V Kohli      11
      SR Watson    10
      JH Kallis    10
      DR Smith     10
      SE Marsh      9
      A Mishra      9
      SR Tendulkar  8
      KA Pollard    8
      RA Jadeja     7
      Name: count, dtype: int64
```

```
[103]: #
      data['player_of_match'].value_counts()[:20].plot(kind = 'bar')
```

```
[103]: <Axes: xlabel='player_of_match'>
```



```
[104]: team_win_by_runs=data.groupby(by='winner').sum()['win_by_runs']
print(team_win_by_runs)

team_win_by_runs.plot(kind='bar',figsize=(12,12),grid=True,title='Team Score in
all seasons (win by runs)')
```

winner	
Chennai Super Kings	1587
Deccan Chargers	421
Delhi Daredevils	416
Gujarat Lions	1
Kings XI Punjab	796
Kochi Tuskers Kerala	23
Kolkata Knight Riders	798
Mumbai Indians	1362
Pune Warriors	139
Rajasthan Royals	821

```

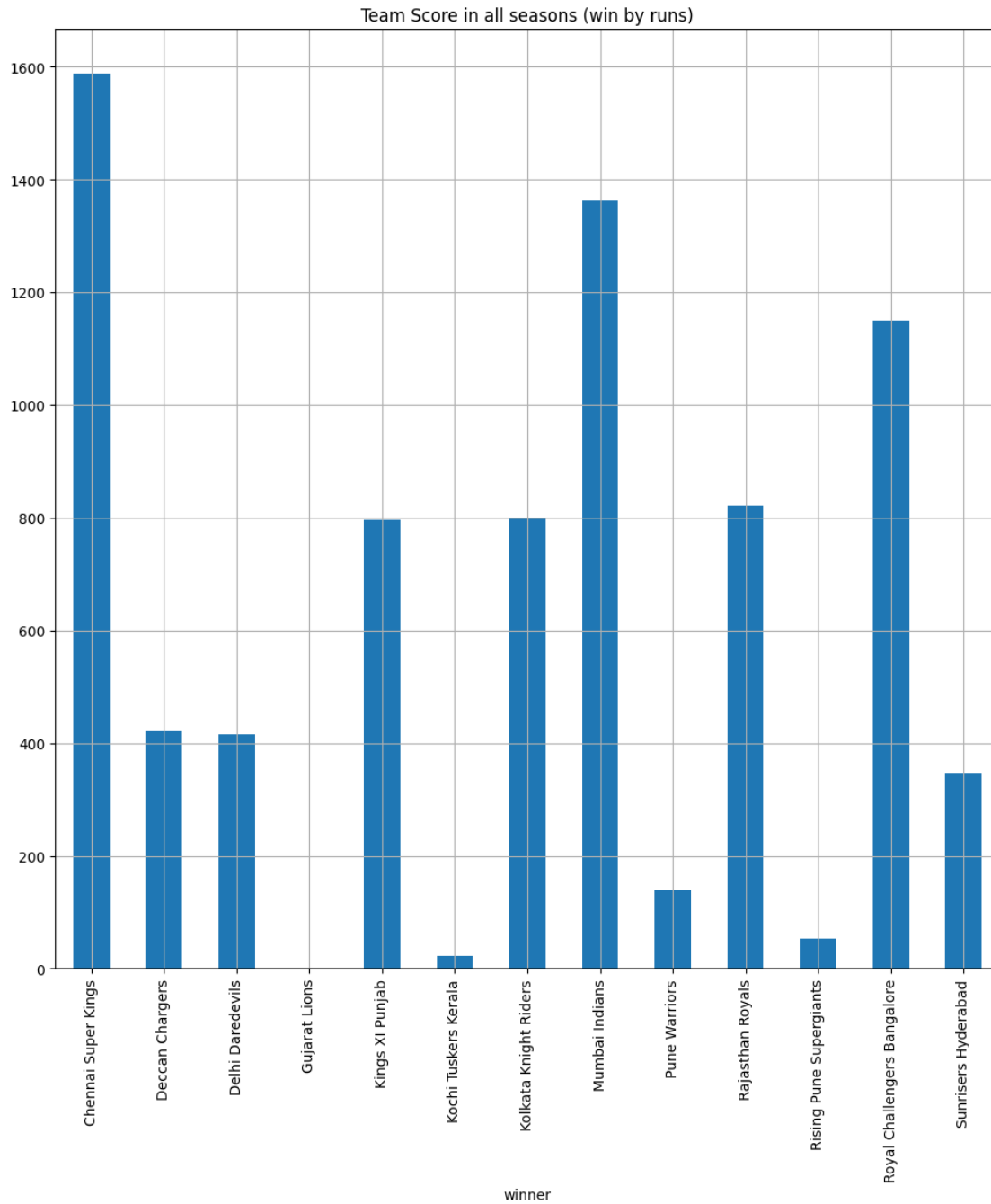
Rising Pune Supergiants      53
Royal Challengers Bangalore  1150
Sunrisers Hyderabad          347
Name: win_by_runs, dtype: int64

```

```

[104]: <Axes: title={'center': 'Team Score in all seasons (win by runs)'},
       xlabel='winner'>

```

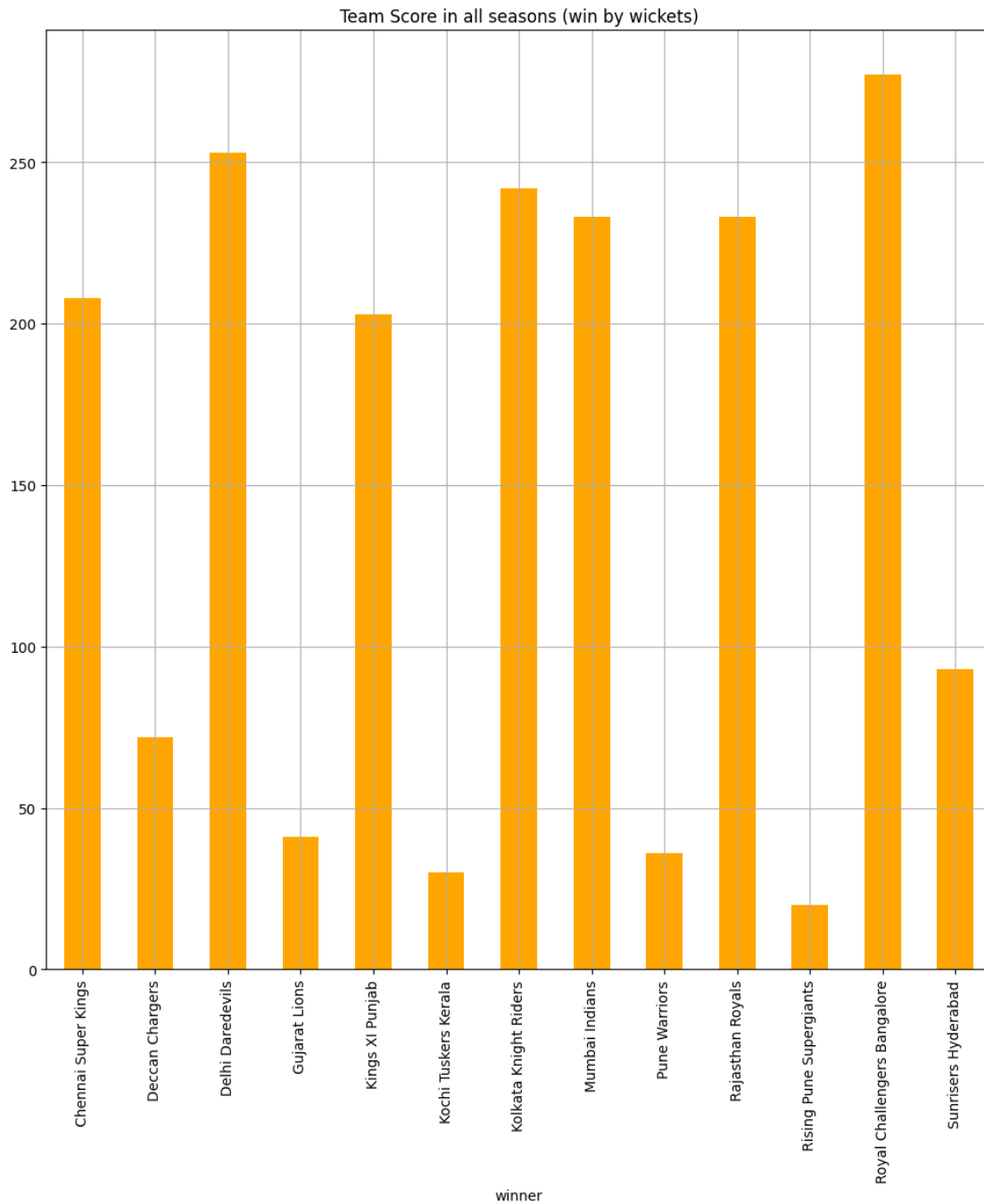


```
[105]: team_win_by_wickets=data.groupby(by='winner').sum()['win_by_wickets']
print(team_win_by_wickets)

team_win_by_wickets.plot(kind='bar',figsize=(12,12),grid=True,title='Team Score_
↳in all seasons (win by wickets)',color="orange")
```

```
winner
Chennai Super Kings      208
Deccan Chargers          72
Delhi Daredevils         253
Gujarat Lions            41
Kings XI Punjab          203
Kochi Tuskers Kerala      30
Kolkata Knight Riders    242
Mumbai Indians           233
Pune Warriors            36
Rajasthan Royals         233
Rising Pune Supergiants   20
Royal Challengers Bangalore 277
Sunrisers Hyderabad      93
Name: win_by_wickets, dtype: int64
```

```
[105]: <Axes: title={'center': 'Team Score in all seasons (win by wickets)'},
xlabel='winner'>
```

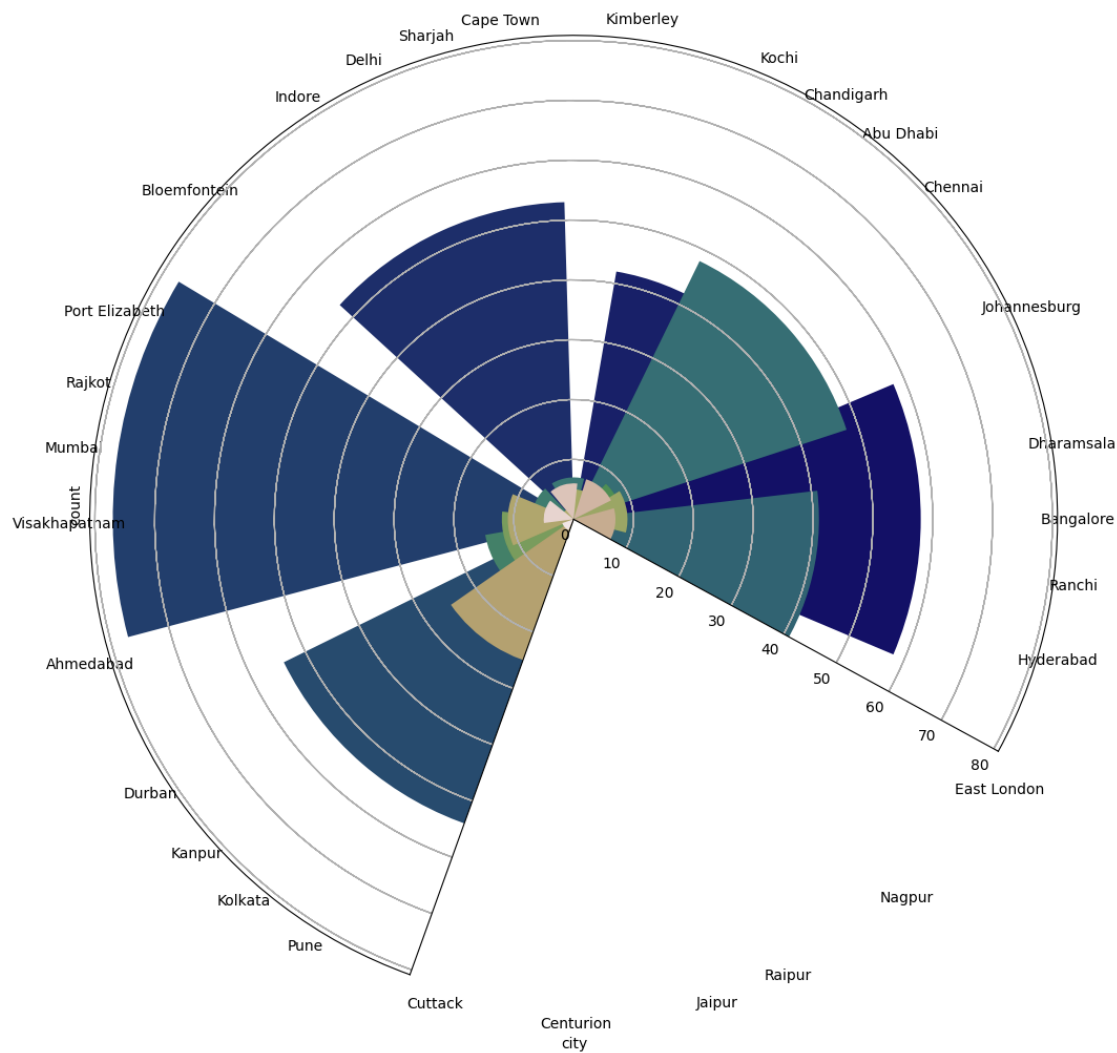


```
[107]: # overall_win_by_runs_and_wickets=data.groupby(by='winner').
        ↪mean()[['win_by_runs','win_by_wickets']]
        # print(overall_win_by_runs_and_wickets)

        # overall_win_by_runs_and_wickets.
        ↪plot(kind='bar',figsize=(12,8),grid=True,title='Team Score in all seasons')
```

```
[108]: fig=plt.figure(figsize=(20,12))
plt.subplot(polar=True)
sns.countplot(x='city',data=data,palette='gist_earth')
```

```
[108]: <PolarAxes: xlabel='city', ylabel='count'>
```



Total Match Played vs Total Match won

```
[110]: team_stats=pd.DataFrame({'TotalMatches': data.team1.value_counts()+data.team2.
↪value_counts()
, 'TotalWin':data.winner.value_counts()})

team_stats=team_stats.reset_index()

team_stats.rename(columns={'index': 'Teams'},inplace=True)
```

```
team_stats.head(13)
```

```
[110]:
```

	Teams	TotalMatches	TotalWin
0	Chennai Super Kings	131	79
1	Deccan Chargers	75	29
2	Delhi Daredevils	133	56
3	Gujarat Lions	16	9
4	Kings XI Punjab	134	63
5	Kochi Tuskers Kerala	14	6
6	Kolkata Knight Riders	132	68
7	Mumbai Indians	140	80
8	Pune Warriors	46	12
9	Rajasthan Royals	118	63
10	Rising Pune Supergiants	14	5
11	Royal Challengers Bangalore	139	70
12	Sunrisers Hyderabad	62	34

1 Comparing man of the match vs Venue(stadium)

```
[111]: man_of_the_match_vs_venue = data.
        ↳groupby(["player_of_match", 'venue', 'city'])['id'].count().
        ↳sort_values(ascending=False)[:20]
        print(man_of_the_match_vs_venue)
```

player_of_match	venue	city	
CH Gayle	M Chinnaswamy Stadium	Bangalore	9
V Kohli	M Chinnaswamy Stadium	Bangalore	6
DA Warner	Rajiv Gandhi International Stadium, Uppal	Hyderabad	6
AB de Villiers	M Chinnaswamy Stadium	Bangalore	6
RG Sharma	Wankhede Stadium	Mumbai	5
SE Marsh	Punjab Cricket Association Stadium, Mohali	Chandigarh	5
MEK Hussey	MA Chidambaram Stadium, Chepauk	Chennai	5
YK Pathan	Eden Gardens	Kolkata	5
SK Raina	MA Chidambaram Stadium, Chepauk	Chennai	4
AM Rahane	Sawai Mansingh Stadium	Jaipur	4
M Vijay	MA Chidambaram Stadium, Chepauk	Chennai	4
AD Russell	Eden Gardens	Kolkata	4
KA Pollard	Wankhede Stadium	Mumbai	4
V Sehwag	Feroz Shah Kotla	Delhi	3
DW Steyn	Rajiv Gandhi International Stadium, Uppal	Hyderabad	3
DA Warner	Feroz Shah Kotla	Delhi	3
SR Watson	Sawai Mansingh Stadium	Jaipur	3
MS Dhoni	MA Chidambaram Stadium, Chepauk	Chennai	3
DR Smith	Wankhede Stadium	Mumbai	3
SC Ganguly	Eden Gardens	Kolkata	3

Name: id, dtype: int64

```
[ ]: # won = data.groupby('winner')['id'].count().sort_values(ascending = False)
# played = []
# for t in won.index:
#     played.append(len(data[(data.team1 == t) | (data.team2 == t)]))

# played = pd.Series(played, index = won.index)
# wonloss_df = pd.DataFrame({'played': played, 'won': won})
# wonloss_df['ratio'] = wonloss_df['won']/wonloss_df['played']
# print(won)
# wonloss_df.sort_values('ratio', ascending = False)
```

2 Comapring winning team vs Venue(stadium)

Insight

-> Consider top 20 teams and bottom 20 teams. In that Chennai super Kings is number one team.

-> By doing this consider Brabourne Stadium,Mumbai.CSK had won one time and Mumbai Indians won 6 times.So that winning probability is very low for CSK to play against Mumbai Indians inBrabourne Stadium,Mumbai.

-> Like this we can do for every teams particularly for young teams.

```
[112]: winner_vs_venue_top_20 = data.groupby(["winner", 'venue', 'city'])['id'].count().
        ↪sort_values(ascending=False)[:20]
print(winner_vs_venue_top_20)
```

winner	venue	city
Kolkata Knight Riders	Eden Gardens	
Kolkata 33		
Chennai Super Kings	MA Chidambaram Stadium, Chepauk	
Chennai 33		
Mumbai Indians	Wankhede Stadium	
Mumbai 29		
Royal Challengers Bangalore	M Chinnaswamy Stadium	
Bangalore 29		
Rajasthan Royals	Sawai Mansingh Stadium	
Jaipur 24		
Delhi Daredevils	Feroz Shah Kotla	Delhi
19		
Kings XI Punjab	Punjab Cricket Association Stadium, Mohali	
Chandigarh 18		
Sunrisers Hyderabad	Rajiv Gandhi International Stadium, Uppal	
Hyderabad 14		
Mumbai Indians	Eden Gardens	
Kolkata 8		
Rajasthan Royals	Sardar Patel Stadium, Motera	

Ahmedabad	7		
Mumbai Indians		M Chinnaswamy Stadium	
Bangalore	6		
		Brabourne Stadium	
Mumbai	6		
Chennai Super Kings		Feroz Shah Kotla	Delhi
5			
Kings XI Punjab		Himachal Pradesh Cricket Association Stadium	
Dharamsala	5		
Mumbai Indians		Dr DY Patil Sports Academy	
Mumbai	5		
Kolkata Knight Riders		Rajiv Gandhi International Stadium, Uppal	
Hyderabad	4		
Royal Challengers Bangalore		Feroz Shah Kotla	Delhi
4			
Kings XI Punjab		Rajiv Gandhi International Stadium, Uppal	
Hyderabad	4		
Delhi Daredevils		Rajiv Gandhi International Stadium, Uppal	
Hyderabad	4		
Kings XI Punjab		Wankhede Stadium	
Mumbai	4		
Name: id, dtype: int64			

```
[113]: winner_vs_venue_bottom_20 = data.groupby(["winner", 'venue', 'city'])['id'].
        ↪count().sort_values(ascending=True)[:20]
        print(winner_vs_venue_bottom_20)
```

winner		venue	
city			
Chennai Super Kings		Brabourne Stadium	
Mumbai	1		
Mumbai Indians		Sardar Patel Stadium, Motera	
Ahmedabad	1		
		Punjab Cricket Association IS Bindra Stadium, Mohali	
Chandigarh	1		
		Newlands	
Cape Town	1		
		Maharashtra Cricket Association Stadium	
Pune	1		
		Buffalo Park	
East London	1		
Kolkata Knight Riders		Wankhede Stadium	
Mumbai	1		
		SuperSport Park	
Centurion	1		
		Sheikh Zayed Stadium	Abu
Dhabi	1		
		Sharjah Cricket Stadium	

Sharjah	1	Sawai Mansingh Stadium
Jaipur	1	Punjab Cricket Association IS Bindra Stadium, Mohali
Chandigarh	1	JSCA International Stadium Complex
Ranchi	1	Shaheed Veer Narayan Singh International Stadium
Sunrisers Hyderabad	1	Wankhede Stadium
Raipur	1	Holkar Cricket Stadium
Kochi Tuskers Kerala	1	Feroz Shah Kotla
Mumbai	1	Eden Gardens
Indore	1	Subrata Roy Sahara Stadium
Delhi	1	St George's Park
Kolkata	1	
Kings XI Punjab	1	
Pune	1	
Port Elizabeth	1	

Name: id, dtype: int64

3 Comapring Umpire vs winning team (overall)

3.1 Umpire1 vs winners

3.2 Umpire2 vs winners

3.3 umpire1 and umpire2 vs winners

```
[ ]: # Insight

# -> umpire1:Mumbai Indians won 80 matches overall in that 17 matches has been
↳umpired by HDPK Dharmasena nearly its 20% in 80

# -> umpire2:Chennai Super Kings won 131 matches overall in that 12 matches
↳has been umpired by SJA Taufel nearly its 10% in 180

# -> umpire1 and umpire2:Mumbai Indians won 80 matches overall in that 4
↳matches has been umpired by Asad Rauf,S Asnani nearly its 3% in 80

[ ]: winner_vs_umpire1 = data.groupby(["winner", 'umpire1'])['id'].count().
↳sort_values(ascending=False)[:20]
print(winner_vs_umpire1)
```

winner	umpire1	
Mumbai Indians	HDPK Dharmasena	17

Kings XI Punjab	HDPK Dharmasena	14
Mumbai Indians	BR Doctrove	10
	Asad Rauf	9
Rajasthan Royals	BF Bowden	8
Royal Challengers Bangalore	Aleem Dar	8
Chennai Super Kings	Asad Rauf	8
Rajasthan Royals	Aleem Dar	8
Royal Challengers Bangalore	HDPK Dharmasena	8
Kolkata Knight Riders	AK Chaudhary	8
	Asad Rauf	7
Chennai Super Kings	HDPK Dharmasena	7
	Aleem Dar	7
Kolkata Knight Riders	S Ravi	6
Delhi Daredevils	HDPK Dharmasena	6
Chennai Super Kings	BR Doctrove	6
Sunrisers Hyderabad	AK Chaudhary	6
Delhi Daredevils	BF Bowden	6
Rajasthan Royals	HDPK Dharmasena	6
Kings XI Punjab	BF Bowden	6

Name: id, dtype: int64

```
[ ]: winner_vs_umpire2 = data.groupby(["winner", 'umpire2'])['id'].count().
    ↪sort_values(ascending=False)[:20]
print(winner_vs_umpire2)
```

winner	umpire2	
Chennai Super Kings	SJA Taufel	12
Mumbai Indians	SJA Taufel	12
Royal Challengers Bangalore	C Shamshuddin	9
Chennai Super Kings	VA Kulkarni	7
Royal Challengers Bangalore	S Ravi	7
Rajasthan Royals	SK Tarapore	7
Chennai Super Kings	C Shamshuddin	7
Mumbai Indians	VA Kulkarni	7
Chennai Super Kings	RB Tiffin	7
Royal Challengers Bangalore	RB Tiffin	6
Delhi Daredevils	S Ravi	6
Rajasthan Royals	RJ Tucker	6
Kings XI Punjab	S Ravi	6
Mumbai Indians	RJ Tucker	6
Kolkata Knight Riders	C Shamshuddin	6
	CK Nandan	6
Kings XI Punjab	AM Saheba	5
Rajasthan Royals	RE Koertzen	5
	SJA Taufel	5
Deccan Chargers	SJA Taufel	5

Name: id, dtype: int64

```
[ ]: winner_vs_umpire1_vs_umpire2 = data.
      ↳groupby(["winner", 'umpire1', 'umpire2'])['id'].count().
      ↳sort_values(ascending=False)[:20]
      print(winner_vs_umpire1_vs_umpire2)
```

winner	umpire1	umpire2	
Mumbai Indians	Asad Rauf	S Asnani	4
	BR Doctrove	SK Tarapore	3
Kings XI Punjab	HDPK Dharmasena	CK Nandan	3
Royal Challengers Bangalore	Aleem Dar	C Shamsuddin	3
Mumbai Indians	HDPK Dharmasena	SJA Taufel	3
	RK Illingworth	VA Kulkarni	3
Chennai Super Kings	Aleem Dar	RB Tiffin	3
Sunrisers Hyderabad	Asad Rauf	S Asnani	3
Royal Challengers Bangalore	AK Chaudhary	HDPK Dharmasena	3
Kings XI Punjab	HDPK Dharmasena	PG Pathak	3
Kolkata Knight Riders	AK Chaudhary	M Erasmus	3
Rajasthan Royals	S Ravi	RJ Tucker	3
Kolkata Knight Riders	S Ravi	C Shamsuddin	3
Deccan Chargers	M Erasmus	AM Saheba	2
	HDPK Dharmasena	SJA Taufel	2
Rajasthan Royals	SD Fry	CB Gaffaney	2
Deccan Chargers	RE Koertzen	S Ravi	2
Mumbai Indians	BR Doctrove	DJ Harper	2
Rajasthan Royals	BF Bowden	SK Tarapore	2
Royal Challengers Bangalore	Aleem Dar	SS Hazare	2

Name: id, dtype: int64