

In [2]:

```
#import all libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
```

In [3]:

```
#read ipl.csv file and print first 5 records
df=pd.read_csv("ipl data.csv")
df.head()
```

Out[3]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_the_match
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	Suresh Raina
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	Andre Russell
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	Gautam Gambhir
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	Kieron Pollard

In [4]:

```
#check total number of columns,entries note down your findings
df
#Findings:-
# There are total 636 enteries with 18 columns
#
```

Out[4]:

am2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	umpire2	ump
oyal gers alore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh	Rajiv Gandhi International Stadium, Uppal	AY Dandekar	NJ Llong	
sing une jiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD Smith	Maharashtra Cricket Association Stadium	A Nand Kishore	S Ravi	
kata ight ders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	CA Lynn	Saurashtra Cricket Association Stadium	Nitin Menon	CK Nandan	
js XI njab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ Maxwell	Holkar Cricket Stadium	AK Chaudhary	C Shamshuddin	
Delhi evils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav	M Chinnaswamy Stadium	NaN	NaN	
...	
oyal gers alore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	6	V Kohli	Shaheed Veer Narayan Singh International Stadium	A Nand Kishore	BNJ Oxenford	
oyal gers alore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	4	AB de Villiers	M Chinnaswamy Stadium	AK Chaudhary	HDPK Dharmasena	
kata ight ders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	22	0	MC Henriques	Feroz Shah Kotla	M Erasmus	C Shamshuddin	
sers ibad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	4	DA Warner	Feroz Shah Kotla	M Erasmus	CK Nandan	
oyal gers alore	Sunrisers Hyderabad	bat	normal	0	Sunrisers Hyderabad	8	0	BCJ Cutting	M Chinnaswamy Stadium	HDPK Dharmasena	BNJ Oxenford	

In [5]:

```
df.shape
```

Out[5]:

(636, 18)

In [6]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 636 entries, 0 to 635
Data columns (total 18 columns):
 #   Column                Non-Null Count  Dtype  
---  --
 0   id                    636 non-null   int64  
 1   season               636 non-null   int64  
 2   city                 629 non-null   object  
 3   date                 636 non-null   object  
 4   team1                636 non-null   object  
 5   team2                636 non-null   object  
 6   toss_winner          636 non-null   object  
 7   toss_decision        636 non-null   object  
 8   result               636 non-null   object  
 9   dl_applied           636 non-null   int64  
10  winner               633 non-null   object  
11  win_by_runs          636 non-null   int64  
12  win_by_wickets       636 non-null   int64  
13  player_of_match      633 non-null   object  
14  venue                636 non-null   object  
15  umpire1              635 non-null   object  
16  umpire2              635 non-null   object  
17  umpire3              0 non-null     float64
dtypes: float64(1), int64(5), object(12)
memory usage: 89.6+ KB
```

In []:

In [7]:

```
df.isnull().sum()
```

Out[7]:

```
id                0
season            0
city              7
date             0
team1            0
team2            0
toss_winner       0
toss_decision     0
result           0
dl_applied        0
winner           3
win_by_runs       0
win_by_wickets    0
player_of_match   3
venue            0
umpire1           1
umpire2           1
umpire3          636
dtype: int64
```

In [8]:

```
df[df.isnull()]
```

Out[8]:

ity	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	umpire2	umj
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
iN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

S

In [10]:

```
#drop umpire3 column as it contains more than 75% of null values
df.drop(columns=['umpire3'])
```

Out[10]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	play	
	0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	
	1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	
	2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	
	3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	
	4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	
	
	631	632	2016	Raipur	2016-05-22	Delhi Daredevils	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	6	
	632	633	2016	Bangalore	2016-05-24	Gujarat Lions	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	4	
	633	634	2016	Delhi	2016-05-25	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	22	0	M
	634	635	2016	Delhi	2016-05-27	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	4	
	635	636	2016	Bangalore	2016-05-29	Sunrisers Hyderabad	Royal Challengers Bangalore	Sunrisers Hyderabad	bat	normal	0	Sunrisers Hyderabad	8	0	

636 rows x 17 columns

```
In [11]:  
  
#drop null values of city , winner ,player_of_match,umpire1,umpire2 column  
df.dropna(subset=['city','winner','player_of_match','umpire1','umpire2'],how='all')
```

Out[11]:

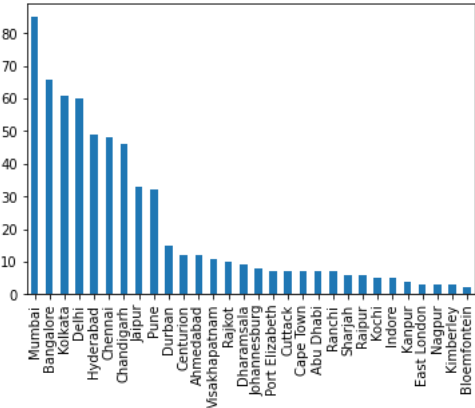
	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Y
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	(
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	9	Ri
...
631	632	2016	Raipur	2016-05-22	Delhi Daredevils	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	6	
632	633	2016	Bangalore	2016-05-24	Gujarat Lions	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	4	AE
633	634	2016	Delhi	2016-05-25	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	22	0	MC
634	635	2016	Delhi	2016-05-27	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	4	
635	636	2016	Bangalore	2016-05-29	Sunrisers Hyderabad	Royal Challengers Bangalore	Sunrisers Hyderabad	bat	normal	0	Sunrisers Hyderabad	8	0	B

325 rows × 18 columns

```
In [14]:  
  
#which city hosted most number of matches?  
#draw bar plot and write down your insights  
df['city'].value_counts().plot(kind='bar')
```

Out[14]:

<AxesSubplot:>



In [17]:

```
#find all venue of mumbai city
df[df['city']=='Mumbai']['venue'].value_counts()
```

Out[17]:

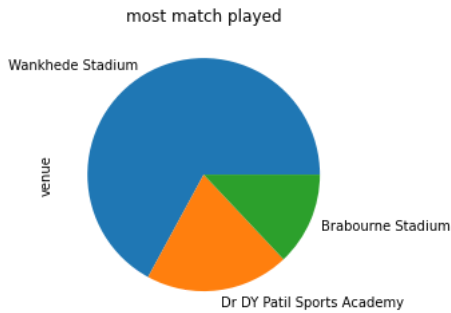
```
Wankhede Stadium      57
Dr DY Patil Sports Academy  17
Brabourne Stadium     11
Name: venue, dtype: int64
```

In [18]:

```
#now compare in which venue of mumbai most number of matches played (draw bar plot and write down insights)
df[df['city']=='Mumbai']['venue'].value_counts().plot(kind='pie',title='most match played')
```

Out[18]:

<AxesSubplot:title={'center':'most match played'}, ylabel='venue'>



In [19]:

```
#what is the preferred choice after winning a toss in mumbai
df[df['city']=='Mumbai']['toss_decision'].value_counts()
```

Out[19]:

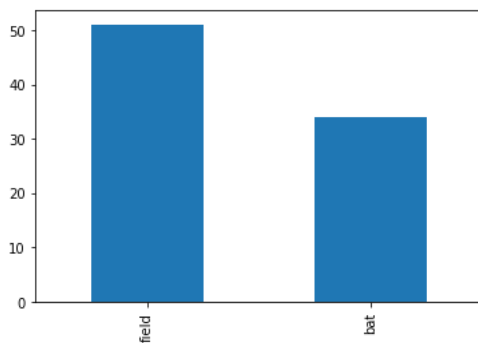
```
field    51
bat      34
Name: toss_decision, dtype: int64
```

In [20]:

```
#graphical representation of above question
df[df['city']=='Mumbai']['toss_decision'].value_counts().plot(kind='bar')
```

Out[20]:

<AxesSubplot:>



In [23]:

```
#which team won most number of toss :-Mumbai Indians
df['toss_winner'].value_counts()
```

Out[23]:

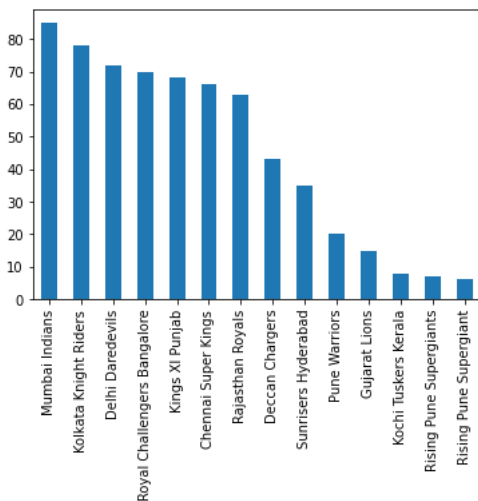
```
Mumbai Indians      85
Kolkata Knight Riders 78
Delhi Daredevils    72
Royal Challengers Bangalore 70
Kings XI Punjab     68
Chennai Super Kings 66
Rajasthan Royals    63
Deccan Chargers     43
Sunrisers Hyderabad 35
Pune Warriors       20
Gujarat Lions       15
Kochi Tuskers Kerala 8
Rising Pune Supergiants 7
Rising Pune Supergiant 6
Name: toss_winner, dtype: int64
```

In [26]:

```
#show graphical representation of above question
df['toss_winner'].value_counts().plot(kind='bar')
```

Out[26]:

<AxesSubplot:>



In [27]:

```
#find what mumbai indians preferred after winning a toss?
df[df['toss_winner']=='Mumbai Indians']['toss_decision'].value_counts()
```

Out[27]:

```
field    45
bat      40
Name: toss_decision, dtype: int64
```

In [30]:

```
#head to head winning count of Mumbai Indians vs Chennai Super Kings
df[(df['team1']=='Mumbai Indians') & (df['team2']=='Chennai Super Kings']]['winner'].value_counts() + df[(df['team1']=='Chennai Super Kings') & (df['team2']=='Mumbai Indians']]['winner'].value_counts()
```

Out[30]:

```
Chennai Super Kings    10
Mumbai Indians         12
Name: winner, dtype: int64
```

In [32]:

```
df[(df['team1']=='Mumbai Indians') & (df['team2']=='Chennai Super Kings']]['winner'].value_counts()
```

Out[32]:

```
Mumbai Indians    7
Chennai Super Kings 5
Name: winner, dtype: int64
```

In [33]:

```
df[(df['team1']=='Chennai Super Kings') & (df['team2']=='Mumbai Indians')]['winner'].value_counts()
```

Out[33]:

```
Chennai Super Kings    5
Mumbai Indians         5
Name: winner, dtype: int64
```

In []:

In []:

In [36]:

```
#Which team won most of the matches in mumbai?
df[(df['city']=='Mumbai')]['winner'].value_counts()
```

Out[36]:

```
Mumbai Indians           45
Chennai Super Kings      8
Rajasthan Royals         5
Royal Challengers Bangalore 5
Kings XI Punjab          5
Pune Warriors            3
Deccan Chargers          3
Kolkata Knight Riders     3
Rising Pune Supergiant    2
Delhi Daredevils          2
Sunrisers Hyderabad      1
Rising Pune Supergiants   1
Kochi Tuskers Kerala      1
Gujarat Lions            1
Name: winner, dtype: int64
```

In [48]:

```
#how many times each team won the toss and won the match in mumbai
df[(df['city']=='Mumbai') & (df['toss_winner']==df['winner'])].shape
```

Out[48]:

```
(43, 18)
```


In [49]:

```
df[df['city']=='Mumbai']
```

Out[49]:

team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	umpire2	umpire3
Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	N Rana	Wankhede Stadium	Nitin Menon	CK Nandan	
Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	JJ Bumrah	Wankhede Stadium	Nitin Menon	CK Nandan	
Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	6	N Rana	Wankhede Stadium	A Nand Kishore	S Ravi	
Delhi Daredevils	Delhi Daredevils	field	normal	0	Mumbai Indians	14	0	MJ McClenaghan	Wankhede Stadium	A Nand Kishore	S Ravi	
Mumbai Indians	Mumbai Indians	field	normal	0	Rising Pune Supergiant	3	0	BA Stokes	Wankhede Stadium	A Nand Kishore	S Ravi	
...
Chennai Super Kings	Mumbai Indians	bat	normal	0	Mumbai Indians	25	0	KA Pollard	Wankhede Stadium	HDPK Dharmasena	RK Illingworth	
Rising Pune Supergiants	Mumbai Indians	bat	normal	0	Rising Pune Supergiants	0	9	AM Rahane	Wankhede Stadium	HDPK Dharmasena	CK Nandan	
Gujarat Lions	Gujarat Lions	field	normal	0	Gujarat Lions	0	3	AJ Finch	Wankhede Stadium	HDPK Dharmasena	VK Sharma	
Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	6	RG Sharma	Wankhede Stadium	AK Chaudhary	CK Nandan	
Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	6	RG Sharma	Wankhede Stadium	Nitin Menon	RJ Tucker	

In []:

In []:

In []:

In []:

```
In [50]:  
  
#which venue hosted most number of matches  
df['venue'].value_counts()
```

```
Out[50]:  
  
M Chinnaswamy Stadium      66  
Eden Gardens               61  
Feroz Shah Kotla           60  
Wankhede Stadium           57  
Rajiv Gandhi International Stadium, Uppal    49  
MA Chidambaram Stadium, Chepauk             48  
Punjab Cricket Association Stadium, Mohali   35  
Sawai Mansingh Stadium      33  
Subrata Roy Sahara Stadium  17  
Dr DY Patil Sports Academy  17  
Maharashtra Cricket Association Stadium     15  
Kingsmead                   15  
SuperSport Park             12  
Sardar Patel Stadium, Motera 12  
Punjab Cricket Association IS Bindra Stadium, Mohali 11  
Dr. Y.S. Rajasekhara Reddy ACA-VDCA Cricket Stadium 11  
Brabourne Stadium           11  
Saurashtra Cricket Association Stadium      10  
Himachal Pradesh Cricket Association Stadium  9  
New Wanderers Stadium        8  
Newlands                     7  
Dubai International Cricket Stadium          7  
JSCA International Stadium Complex           7  
St George's Park                   7  
Barabati Stadium                   7  
Sheikh Zayed Stadium               7  
Sharjah Cricket Stadium            6  
Shaheed Veer Narayan Singh International Stadium 6  
Nehru Stadium                      5  
Holkar Cricket Stadium             5  
Green Park                         4  
De Beers Diamond Oval              3  
Vidarbha Cricket Association Stadium, Jamtha 3  
Buffalo Park                       3  
OUTsurance Oval                   2  
Name: venue, dtype: int64
```

```
In [ ]:  
  
df['venue'].value_counts().plot()
```

```
In [64]:  
  
#find how many matches chennai super kings played at M Chinnaswamy stadium?  
mask1=(df['team1']=='Chennai Super Kings')|(df['team2']=='Chennai Super Kings')  
mask2=df['venue']=='M Chinnaswamy Stadium'  
df[mask1&mask2]
```

```
Out[64]:
```

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_the_match
73	74	2008	Bangalore	2008-04-28	Chennai Super Kings	Royal Challengers Bangalore	Chennai Super Kings	bat	normal	0	Chennai Super Kings	13	0	
191	192	2010	Bangalore	2010-03-23	Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	field	normal	0	Royal Challengers Bangalore	36	0	
301	302	2011	Bangalore	2011-05-22	Chennai Super Kings	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Royal Challengers Bangalore	0	8	
378	379	2012	Bangalore	2012-05-23	Chennai Super Kings	Mumbai Indians	Mumbai Indians	field	normal	0	Chennai Super Kings	38	0	
451	452	2013	Bangalore	2013-05-18	Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	field	normal	0	Royal Challengers Bangalore	24	0	
509	510	2014	Bangalore	2014-05-24	Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	8	
537	538	2015	Bangalore	2015-04-22	Chennai Super Kings	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Chennai Super Kings	27	0	

In [68]:

```
#who won most matches at M Chinnaswamy stadium?--Royal Challengers Bangalore  
df[df['venue']=='M Chinnaswamy Stadium']['winner'].value_counts()
```

Out[68]:

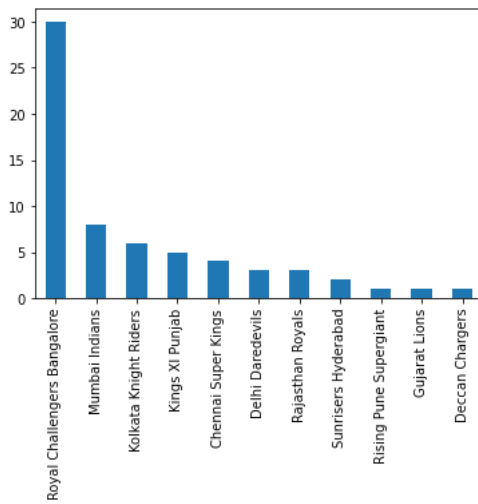
```
Royal Challengers Bangalore    30  
Mumbai Indians                8  
Kolkata Knight Riders         6  
Kings XI Punjab              5  
Chennai Super Kings           4  
Delhi Daredevils              3  
Rajasthan Royals              3  
Sunrisers Hyderabad          2  
Rising Pune Supergiant        1  
Gujarat Lions                 1  
Deccan Chargers               1  
Name: winner, dtype: int64
```

In [69]:

```
df[df['venue']=='M Chinnaswamy Stadium']['winner'].value_counts().plot(kind='bar')
```

Out[69]:

<AxesSubplot:>



In [73]:

```
#matches played in each year  
df['season'].value_counts()
```

Out[73]:

```
2013    76  
2012    74  
2011    73  
2010    60  
2014    60  
2016    60  
2015    59  
2017    59  
2008    58  
2009    57  
Name: season, dtype: int64
```

In [77]:

```
#which city hosted most number of matches in 2013
df[df['season']==2013]['city'].value_counts()
```

Out[77]:

Kolkata	8
Jaipur	8
Mumbai	8
Chennai	8
Delhi	8
Pune	8
Hyderabad	8
Bangalore	8
Chandigarh	6
Dharamsala	2
Ranchi	2
Raipur	2

Name: city, dtype: int64

In [23]:

```
#lets analyse ipl season held in 2013
#extract all the details of 2013 season
#here we will create yearwise groups (hint: use groupby() function)
```

In [80]:

```
df[df['season']==2013]
```

Out[80]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_the_match
381	382	2013	Kolkata	2013-04-03	Delhi Daredevils	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	6	
382	383	2013	Bangalore	2013-04-04	Royal Challengers Bangalore	Mumbai Indians	Mumbai Indians	field	normal	0	Royal Challengers Bangalore	2	0	
383	384	2013	Hyderabad	2013-04-05	Sunrisers Hyderabad	Pune Warriors	Pune Warriors	field	normal	0	Sunrisers Hyderabad	22	0	
384	385	2013	Delhi	2013-04-06	Rajasthan Royals	Delhi Daredevils	Rajasthan Royals	bat	normal	0	Rajasthan Royals	5	0	
385	386	2013	Chennai	2013-04-06	Mumbai Indians	Chennai Super Kings	Mumbai Indians	bat	normal	0	Mumbai Indians	9	0	
...
452	453	2013	Hyderabad	2013-05-19	Kolkata Knight Riders	Sunrisers Hyderabad	Kolkata Knight Riders	bat	normal	0	Sunrisers Hyderabad	0	5	
453	454	2013	Delhi	2013-05-21	Chennai Super Kings	Mumbai Indians	Chennai Super Kings	bat	normal	0	Chennai Super Kings	48	0	
454	455	2013	Delhi	2013-05-22	Sunrisers Hyderabad	Rajasthan Royals	Sunrisers Hyderabad	bat	normal	0	Rajasthan Royals	0	4	
455	456	2013	Kolkata	2013-05-24	Rajasthan Royals	Mumbai Indians	Rajasthan Royals	bat	normal	0	Mumbai Indians	0	4	Harbhajan Singh
456	457	2013	Kolkata	2013-05-26	Mumbai Indians	Chennai Super Kings	Mumbai Indians	bat	normal	0	Mumbai Indians	23	0	

76 rows × 18 columns

In [84]:

```
mask1=(df['team1']=='Chennai Super Kings')|(df['team2']=='Chennai Super Kings')
mask2=df['winner']=='Chennai Super Kings'
df[mask1&mask2]['season'].value_counts()
```

Out[84]:

```
2013    12
2011    11
2012    10
2014    10
2015    10
2008     9
2010     9
2009     8
Name: season, dtype: int64
```

In [85]:

```
df.describe()
```

Out[85]:

	id	season	dl_applied	win_by_runs	win_by_wickets	umpire3
count	636.000000	636.000000	636.000000	636.000000	636.000000	0.0
mean	318.500000	2012.490566	0.025157	13.682390	3.372642	NaN
std	183.741666	2.773026	0.156726	23.908877	3.420338	NaN
min	1.000000	2008.000000	0.000000	0.000000	0.000000	NaN
25%	159.750000	2010.000000	0.000000	0.000000	0.000000	NaN
50%	318.500000	2012.000000	0.000000	0.000000	4.000000	NaN
75%	477.250000	2015.000000	0.000000	20.000000	7.000000	NaN
max	636.000000	2017.000000	1.000000	146.000000	10.000000	NaN

In [86]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 636 entries, 0 to 635
Data columns (total 18 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                    636 non-null   int64
1   season                636 non-null   int64
2   city                 629 non-null   object
3   date                 636 non-null   object
4   team1                636 non-null   object
5   team2                636 non-null   object
6   toss_winner          636 non-null   object
7   toss_decision        636 non-null   object
8   result               636 non-null   object
9   dl_applied           636 non-null   int64
10  winner               633 non-null   object
11  win_by_runs          636 non-null   int64
12  win_by_wickets       636 non-null   int64
13  player_of_match      633 non-null   object
14  venue                636 non-null   object
15  umpire1              635 non-null   object
16  umpire2              635 non-null   object
17  umpire3              0 non-null     float64
dtypes: float64(1), int64(5), object(12)
memory usage: 89.6+ KB
```

In [88]:

```
df.shape
```

Out[88]:

```
(636, 18)
```

In [89]:

```
df.groupby('season').winner.value_counts()
```

Out[89]:

season	winner	
2008	Rajasthan Royals	13
	Kings XI Punjab	10
	Chennai Super Kings	9
	Delhi Daredevils	7
	Mumbai Indians	7
2017	Sunrisers Hyderabad	8
	Kings XI Punjab	7
	Delhi Daredevils	6
	Gujarat Lions	4
	Royal Challengers Bangalore	3

Name: winner, Length: 84, dtype: int64

In []: