```
In [1]:   #import all libraries

          import numpy as np
          import pandas as pd
          import matplotlib.pyplot as plt
          %matplotlib inline
          import seaborn as sns
          import warnings
          warnings.filterwarnings('ignore')
```

```
In [2]:   #read ipl.csv file and print first 5 records
          ipl= pd.read_csv('ipl data.csv')
          ipl.head()
```

Out[2]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_a
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	

```
In [3]:   ipl.shape
```

Out[3]:   (636, 18)

```
In [4]:  #check total number of columns,entries note down your findings
         ipl.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 [5]:  # find null values
         ipl.isnull().sum()
```

```
Out[5]:  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
```

drop umpire3 column as it contains more than 75% of null values

In [6]:
```python
ipl.dropna(axis=1,how='all',inplace=True)
ipl
```

Out[6]:

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

636 rows × 17 columns
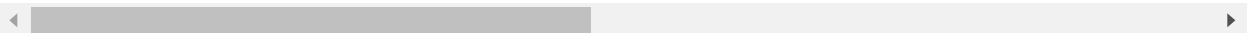
In [7]:
```python
#drop null values of city , winner ,player_of_match,umpire1,umpire2 column
ipl.dropna(inplace=True)
```

```
In [8]: ipl.isnull().sum()
```

```
Out[8]: id                  0
        season              0
        city                0
        date                0
        team1               0
        team2               0
        toss_winner         0
        toss_decision       0
        result              0
        dl_applied          0
        winner              0
        win_by_runs         0
        win_by_wickets      0
        player_of_match     0
        venue               0
        umpire1             0
        umpire2             0
        dtype: int64
```

```
In [9]: ipl.head()
```

Out[9]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_ap
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	

```
In [10]: ipl.shape
```

```
Out[10]: (625, 17)
```

#which city hosted most number of matches?

#draw bar plot and write down your insights

In [11]:
```python
cities = ipl.groupby('city').id.count()
cities =cities.sort_values(ascending=False)
cities
```

Out[11]:
```
city
Mumbai            85
Bangalore         63
Kolkata           61
Delhi             59
Hyderabad         49
Chennai           48
Chandigarh        46
Jaipur            33
Pune              32
Durban            15
Centurion         12
Ahmedabad         12
Visakhapatnam     11
Rajkot            10
Dharamsala         9
Johannesburg       8
Cape Town          7
Cuttack            7
Ranchi             7
Port Elizabeth     7
Abu Dhabi          7
Raipur             6
Sharjah            6
Indore             5
Kochi              5
Kanpur             4
Nagpur             3
Kimberley          3
East London        3
Bloemfontein       2
Name: id, dtype: int64
```

In [ ]:

```
In [12]: plt.figure(figsize=(12,6))
         plt.xticks(rotation=75)
         city_host = sns.barplot(x = cities.index, y = cities)
         city_host.set(xlabel = 'Seasons', ylabel = 'No. of Matches')
```

Out[12]: [Text(0.5, 0, 'Seasons'), Text(0, 0.5, 'No. of Matches')]
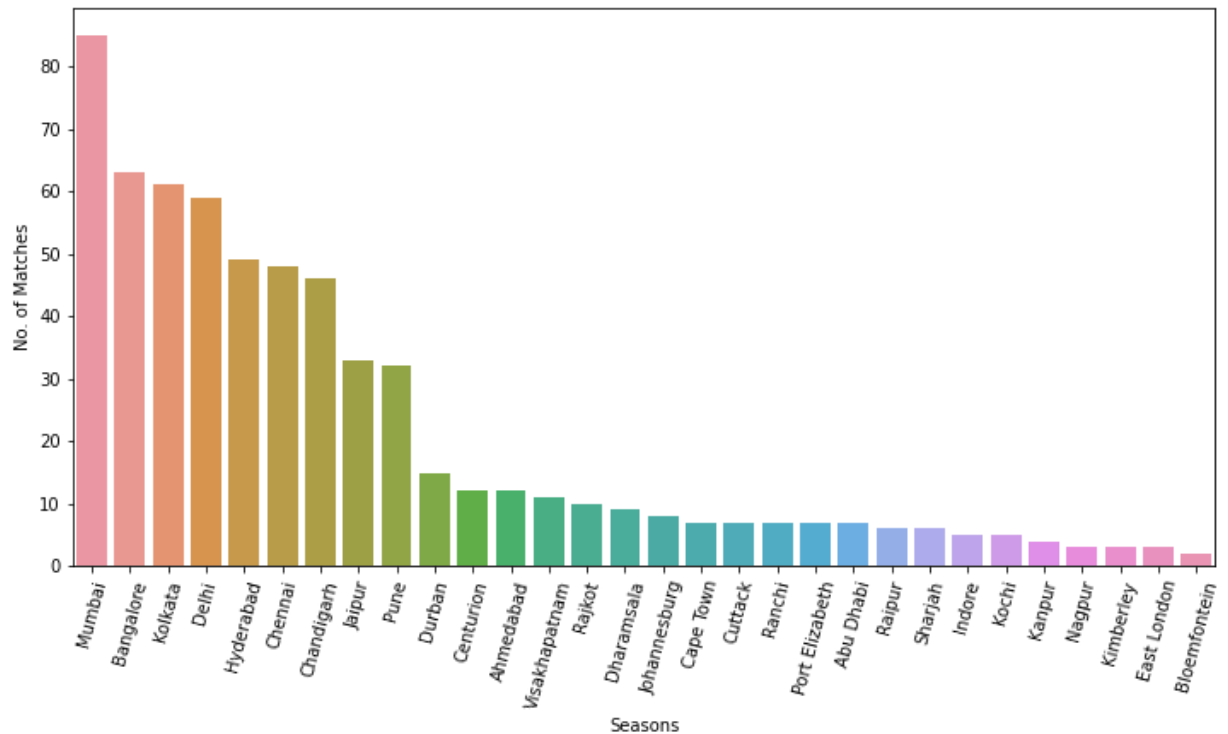
In [ ]:

find all venue of mumbai city

```
In [13]: a=ipl.loc[ipl['city']=='Mumbai'].venue.unique()
         a
```

Out[13]: array(['Wankhede Stadium', 'Dr DY Patil Sports Academy',
                'Brabourne Stadium'], dtype=object)
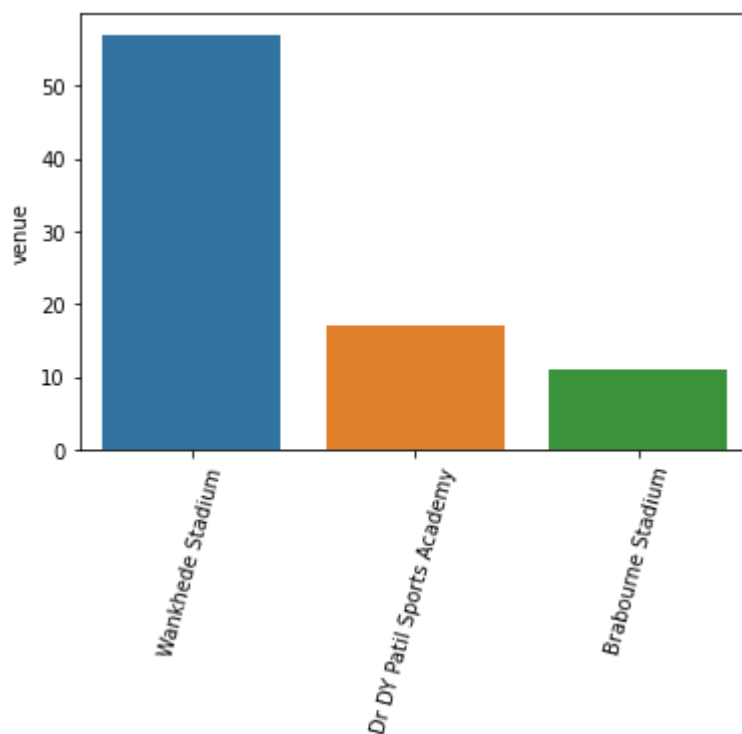
In [ ]:

In [ ]:

#now compare in which venue of mumbai most number of matches played (draw bar plot and write down insights)

In [14]:
```python
venue=ipl.loc[ipl['city']=='Mumbai'].venue.value_counts()
```

In [15]:
```python
sns.barplot(x=venue.index,y=venue)
plt.xticks(rotation=75)
plt.show()
```


1. Wankhede stadium hosted most number of matches in Mumbai the match count(number of matches) goes greter than 50
2. where as dy patil and brabourne stadium are not even half of the number of wandkhede stadium the match count(number of matches) is less than 20
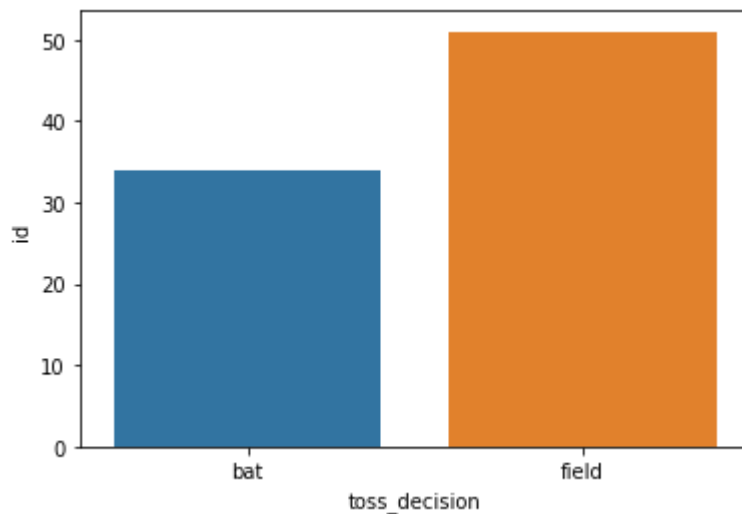
#what is the preferred choice after winning a toss in mumbai

In [16]:
```python
a=ipl.loc[ipl['city']=='Mumbai'].groupby('toss_decision').id.count()
a
```

Out[16]:
```
toss_decision
bat      34
field    51
Name: id, dtype: int64
```

In [17]:
```python
#graphical representation of above question
sns.barplot(x=a.index,y=a)
```

Out[17]: `<AxesSubplot:xlabel='toss_decision', ylabel='id'>`

In [ ]:

which team won most number of toss

In [18]:
```python
a=ipl['toss_winner'].value_counts()
a
```

Out[18]:
```
Mumbai Indians                  83
Kolkata Knight Riders           77
Delhi Daredevils                71
Royal Challengers Bangalore     67
Kings XI Punjab                 67
Chennai Super Kings             66
Rajasthan Royals                61
Deccan Chargers                 43
Sunrisers Hyderabad             34
Pune Warriors                   20
Gujarat Lions                   15
Kochi Tuskers Kerala             8
Rising Pune Supergiants          7
Rising Pune Supergiant           6
Name: toss_winner, dtype: int64
```

In [19]:
```python
#show graphical representation of above question
sns.barplot(x=a.index,y=a)
plt.xticks(rotation=75)
plt.show()
```
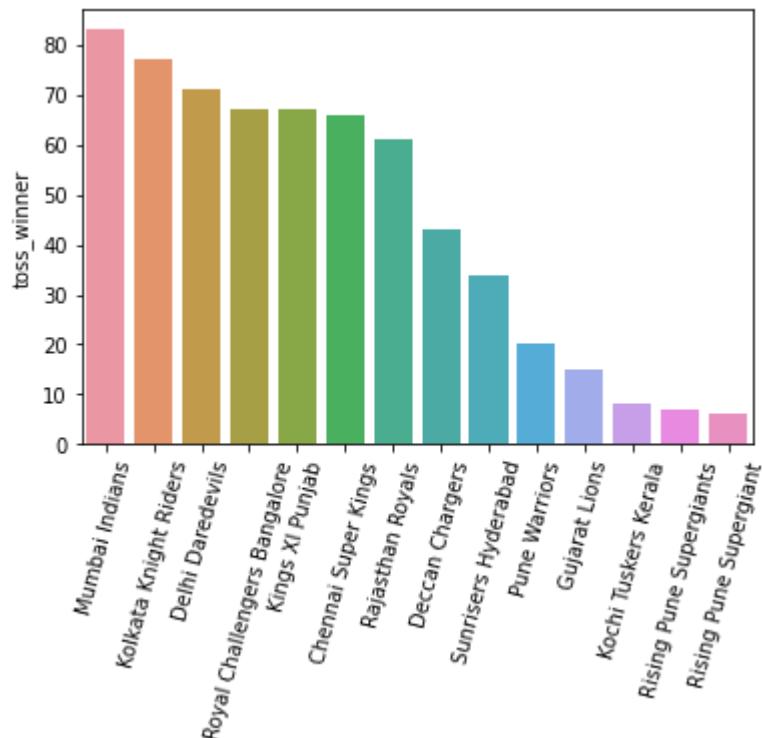
find what mumbai indians preferred after winning a toss?

In [20]:
```python
a=ipl.loc[ipl['toss_winner']=='mumbai indians'.title()].groupby('toss_decision').
a
```

Out[20]:
```
toss_decision
bat      39
field    44
Name: id, dtype: int64
```

In [ ]:

head to head winning count of Mumbai Indians vs Chennai Super Kings

In [21]:
```python
micsk=ipl[np.logical_or(np.logical_and(ipl['team1']=='Mumbai Indians',ipl['team2'
```

In [22]:
```python
sns.set(style='dark')
fig=plt.gcf()
fig.set_size_inches(12,8)
sns.countplot(micsk['winner'],order=micsk['winner'].value_counts().index)
plt.text(-0.1,13,str(micsk['winner'].value_counts()['Mumbai Indians']),size=29,cd
plt.text(0.9,10,str(micsk['winner'].value_counts()['Chennai Super Kings']),size=2
plt.xlabel('Winner',fontsize=15)
plt.ylabel('Count',fontsize=15)
plt.yticks(fontsize=0)
plt.title('MI vs CSK - head to head')
plt.show()
```
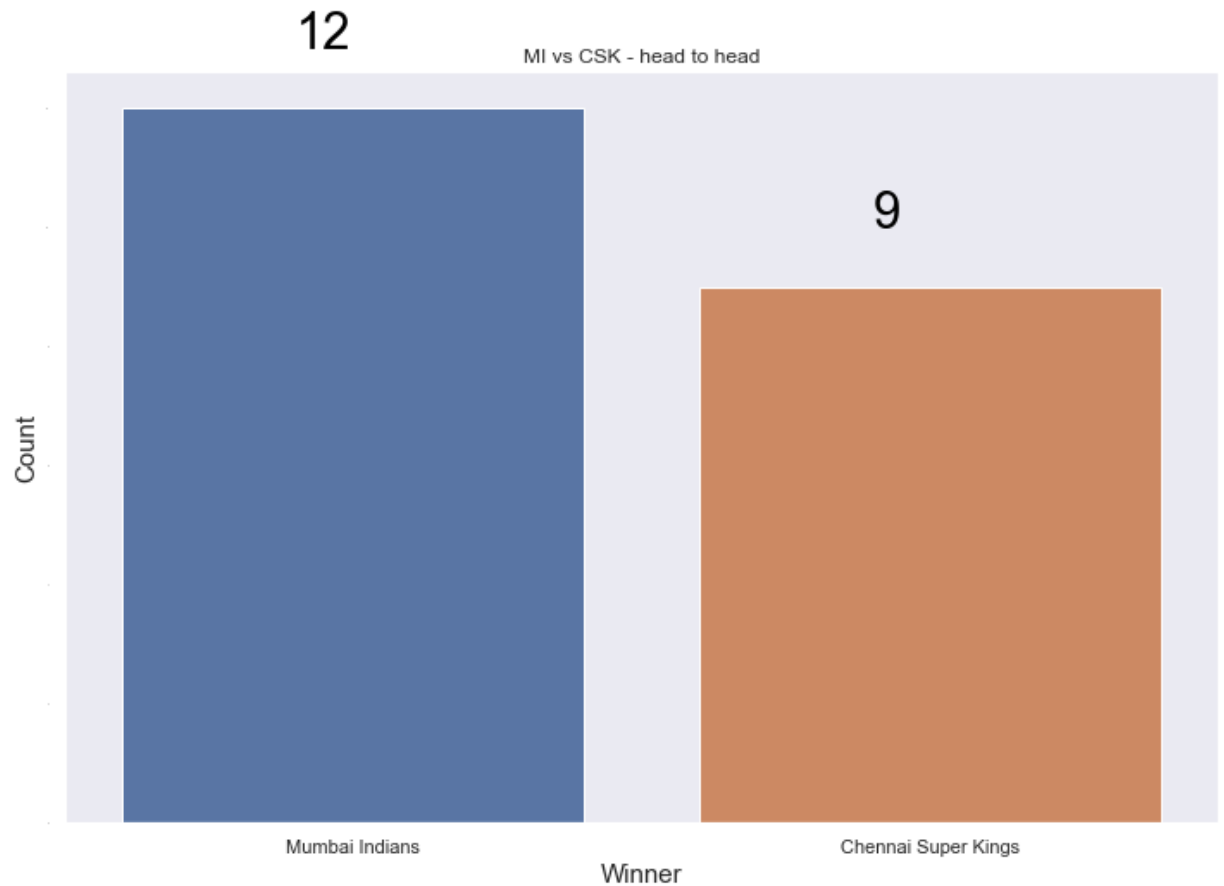

#Which team won most of the matches in mumbai?

```
In [23]: ipl['winner'].value_counts()
         a=ipl.loc[ipl['city']=='Mumbai']
         a['winner'].value_counts()
```

```
Out[23]: Mumbai Indians                45
         Chennai Super Kings            8
         Kings XI Punjab                5
         Royal Challengers Bangalore    5
         Rajasthan Royals               5
         Deccan Chargers                3
         Kolkata Knight Riders          3
         Pune Warriors                  3
         Rising Pune Supergiant         2
         Delhi Daredevils               2
         Kochi Tuskers Kerala           1
         Sunrisers Hyderabad            1
         Rising Pune Supergiants        1
         Gujarat Lions                  1
         Name: winner, dtype: int64
```

In [ ]:

In [ ]:

#how many times each team won the toss and won the match

```
In [24]: a=ipl.loc[ipl['toss_winner']==ipl['winner']]
         a['team1'].value_counts()
```

```
Out[24]: Chennai Super Kings            46
         Kolkata Knight Riders          39
         Mumbai Indians                 38
         Royal Challengers Bangalore    35
         Delhi Daredevils               35
         Kings XI Punjab                30
         Rajasthan Royals               27
         Deccan Chargers                23
         Sunrisers Hyderabad            18
         Gujarat Lions                  12
         Pune Warriors                   9
         Rising Pune Supergiants         4
         Rising Pune Supergiant          3
         Kochi Tuskers Kerala            2
         Name: team1, dtype: int64
```

In [ ]:

#how many times each team won the toss and won the match in mumbai

In [ ]:

In [25]:
```python
x=(ipl.loc[(ipl['city']=='Mumbai') & (ipl['toss_winner']==ipl['winner'])])
x['toss_winner'].value_counts()
```

Out[25]:
```
Mumbai Indians                 26
Chennai Super Kings             5
Deccan Chargers                 3
Rajasthan Royals                2
Delhi Daredevils                2
Royal Challengers Bangalore     2
Kochi Tuskers Kerala            1
Kolkata Knight Riders           1
Gujarat Lions                   1
Name: toss_winner, dtype: int64
```

In [26]:
```python
(ipl['winner']).value_counts()
```

Out[26]:
```
Mumbai Indians                 92
Kolkata Knight Riders          77
Chennai Super Kings            77
Royal Challengers Bangalore    71
Kings XI Punjab                69
Rajasthan Royals               63
Delhi Daredevils               61
Sunrisers Hyderabad            40
Deccan Chargers                29
Gujarat Lions                  13
Pune Warriors                  12
Rising Pune Supergiant         10
Kochi Tuskers Kerala            6
Rising Pune Supergiants         5
Name: winner, dtype: int64
```

In [27]:
```python
ipl['city'].loc[ipl['city']=='Mumbai']
```

Out[27]:
```
6       Mumbai
9       Mumbai
15      Mumbai
23      Mumbai
27      Mumbai
         ...
572     Mumbai
576     Mumbai
584     Mumbai
589     Mumbai
599     Mumbai
Name: city, Length: 85, dtype: object
```

In [28]: `ipl.loc[(ipl['toss_winner']==ipl['winner'])]`

Out[28]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	d
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	
6	7	2017	Mumbai	2017-04-09	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	
...	...	...	...	...	...	...	...	...	...	
629	630	2016	Kanpur	2016-05-21	Mumbai Indians	Gujarat Lions	Gujarat Lions	field	normal	
631	632	2016	Raipur	2016-05-22	Delhi Daredevils	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	
632	633	2016	Bangalore	2016-05-24	Gujarat Lions	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	
634	635	2016	Delhi	2016-05-27	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	
635	636	2016	Bangalore	2016-05-29	Sunrisers Hyderabad	Royal Challengers Bangalore	Sunrisers Hyderabad	bat	normal	

321 rows × 17 columns
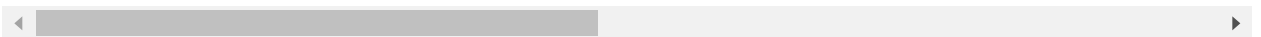
In [ ]:

which venue hosted most number of matches

In [129]: `ipl.head()`

Out[129]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_ap
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	

```
In [133]: venue=ipl['venue'].value_counts()
          venue
```

```
Out[133]: M Chinnaswamy Stadium                                        63
          Eden Gardens                                                 61
          Feroz Shah Kotla                                             59
          Wankhede Stadium                                             57
          Rajiv Gandhi International Stadium, Uppal                     49
          MA Chidambaram Stadium, Chepauk                              48
          Punjab Cricket Association Stadium, Mohali                    35
          Sawai Mansingh Stadium                                       33
          Dr DY Patil Sports Academy                                   17
          Subrata Roy Sahara Stadium                                   17
          Maharashtra Cricket Association Stadium                       15
          Kingsmead                                                    15
          Sardar Patel Stadium, Motera                                 12
          SuperSport Park                                              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
          St George's Park                                              7
          Sheikh Zayed Stadium                                          7
          JSCA International Stadium Complex                             7
          Barabati Stadium                                              7
          Shaheed Veer Narayan Singh International Stadium               6
          Sharjah Cricket Stadium                                        6
          Holkar Cricket Stadium                                         5
          Nehru Stadium                                                 5
          Green Park                                                     4
          Vidarbha Cricket Association Stadium, Jamtha                   3
          De Beers Diamond Oval                                          3
          Buffalo Park                                                   3
          OUTsurance Oval                                                2
          Name: venue, dtype: int64
```

In [142]:
```python
plt.figure(figsize=(12,8))
# plt.xticks(rotation=75)
venues = sns.barplot(x = venue, y =venue.index)
# venue.set(xlabel = 'Seasons', ylabel = 'No. of Matches');
```
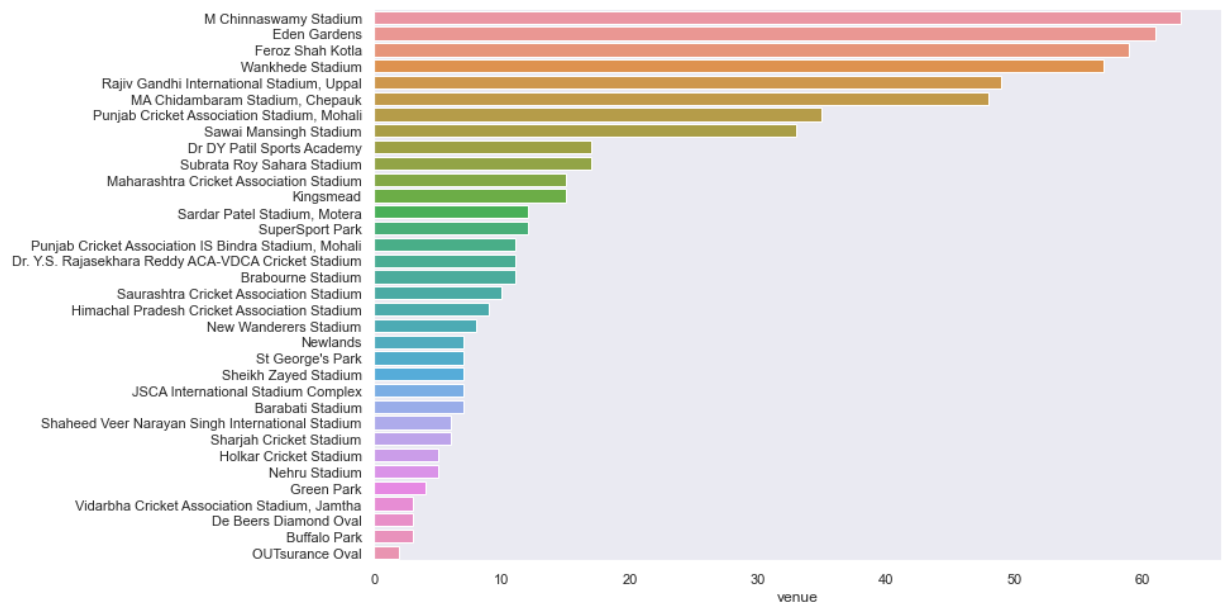

In [ ]:

#find how many matches chennai super kings played at M Chinnaswamy stadium?

In [163]:
```python
t1=ipl.loc[(ipl['team1'] =='Chennai Super Kings') & (ipl['venue']=='M Chinnaswamy
t1
```

. . .

In [165]:
```python
t2=ipl.loc[(ipl['team2'] =='Chennai Super Kings') & (ipl['venue']=='M Chinnaswam
t2
```

. . .

In [167]:
```python
a=pd.concat([t1,t2])
a.shape[0]
```

Out[167]: 7

In [ ]:

#who won most matches at M Chinnaswamy stadium?

In [178]:
```python
x=ipl.loc[ipl['venue']=='M Chinnaswamy Stadium']
x['winner'].value_counts()
```

Out[178]:
```
Royal Challengers Bangalore     29
Mumbai Indians                   8
Kolkata Knight Riders            6
Kings XI Punjab                  5
Chennai Super Kings              4
Rajasthan Royals                 3
Delhi Daredevils                 3
Sunrisers Hyderabad              2
Rising Pune Supergiant           1
Gujarat Lions                    1
Deccan Chargers                  1
Name: winner, dtype: int64
```

In [ ]:

year wise analysis

In [179]:
```python
#matches played in each year
```

In [180]:
```python
#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 [189]:
```python
year=ipl.groupby('season')
year2013=year.get_group(2013)
year2013.head()
```

. . .

In [ ]:

In [ ]:

#which city hosted most number of matches in 2013

In [192]: `year2013['city'].value_counts()`

Out[192]:
```
Kolkata        8
Bangalore      8
Hyderabad      8
Delhi          8
Chennai        8
Pune           8
Jaipur         8
Mumbai         8
Chandigarh     6
Dharamsala     2
Raipur         2
Ranchi         2
Name: city, dtype: int64
```

In [ ]: