In [2]:

```python
#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 [4]:

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

Out[4]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_o
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuv
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	S
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	G
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KI

In [6]:

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

Out[6]:

	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	N
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 × 18 columns

In [7]:

```
df.shape
```

Out[7]:

```
(636, 18)
```

In [8]:

```
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 [ ]:

In [4]:

```
# find null values
```

In [5]:

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

In [6]:

```
#drop null values of city , winner ,player_of_match,umpire1,umpire2 column
```

In [7]:

```
#which city hosted most number of matches?
#draw bar plot and write down your insights
```

In [8]:

```
#find all venue of mumbai city
```

In [9]:

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

In [10]:

```
#what is the preferred choice after winning a toss in mumbai
```

In [11]:
```
#graphical representation of above question
```

In [12]:
```
#which team won most number of toss :-Mumbai Indians
```

In [13]:
```
#show graphical representation of above question
```

In [14]:
```
#find what mumbai indians preferred after winning a toss?
```

In [15]:
```
#head to head winning count of Mumbai Indians vs Chennai Super Kings
```

In [16]:
```
#Which team won most of the matches in mumbai?
```

In [17]:
```
#how many times each team won the toss and won the match in mumbai
```

In [18]:
```
#which venue hosted most number of matches
```

In [19]:
```
#find how many matches chennai super kings played at M Chinnaswamy stadium?
```

In [20]:
```
#who won most matches at M Chinnaswamy stadium?--Royal Challengers Bangalore
```

In [21]:
```
#matches played in each year
```

In [22]:
```
#which city hosted most number of matches in 2013
```

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 [ ]: