## In[1]:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

## In[2]:

sns.set\_style("darkgrid")

plt.rcParams['figure.figsize'] = (14, 8)

## In[3]:

file\_path='matches1.csv'

matches = pd.read\_csv(file\_path)

## In[4]:

matches.shape

out[4]:

(756, 18)

## In[5]:

matches.info()

out[5]:

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 756 entries, 0 to 755

Data columns (total 18 columns):

id 756 non-null int64

season 756 non-null int64

city 749 non-null object

date 756 non-null object

team1 756 non-null object

team2 756 non-null object

toss\_winner 756 non-null object

toss\_decision 756 non-null object

result 756 non-null object

dl\_applied 756 non-null int64

winner 752 non-null object

win\_by\_runs 756 non-null int64

win\_by\_wickets 756 non-null int64

player\_of\_match 752 non-null object

venue 756 non-null object

umpire1 754 non-null object

umpire2 754 non-null object

umpire3 119 non-null object

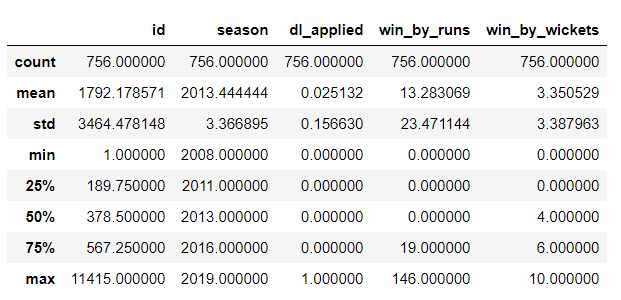
dtypes: int64(5), object(13)

memory usage: 106.4+ KB

## In[6]:

matches.describe()

out[6]:



## In[7]:

matches.head(2)

out[7]:

| **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** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | 2017 | Hyderabad | 05-04-2017 | Sunrisers Hyderabad | Royal Challengers Bangalore | Royal Challengers Bangalore | field | normal | 0 | Sunrisers Hyderabad | 35 | 0 | Yuvraj Singh | Rajiv Gandhi International Stadium, Uppal | AY Dandekar | NJ Llong | NaN |
| **1** | 2 | 2017 | Pune | 06-04-2017 | Mumbai Indians | Rising Pune Supergiant | Rising Pune Supergiant | field | normal | 0 | Rising Pune Supergiant | 0 | 7 | SPD Smith | Maharashtra Cricket Association Stadium | A Nand Kishore | S Ravi | NaN |

## In[8]:

matches['id'].max()

out[8]:

11415

## In[9]:

matches['season'].unique()

out[9]:

array([2017, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2018,

2019], dtype=int64)

## In[10]:

len(matches['season'].unique())

out[10]:

12

## In[11]:

matches.iloc[matches['win\_by\_runs'].idxmax()]

out[11]:

id 44

season 2017

city Delhi

date 06-05-2017

team1 Mumbai Indians

team2 Delhi Daredevils

toss\_winner Delhi Daredevils

toss\_decision field

result normal

dl\_applied 0

winner Mumbai Indians

win\_by\_runs 146

win\_by\_wickets 0

player\_of\_match LMP Simmons

venue Feroz Shah Kotla

umpire1 Nitin Menon

umpire2 CK Nandan

umpire3 NaN

Name: 43, dtype: object

## In[12]:

**matches.iloc[matches['win\_by\_runs'].idxmax()]['winner']**

out[12]:

'Mumbai Indians'

## In[13]:

**matches.iloc[matches['win\_by\_wickets'].idxmax()]['winner']**

out[13]:

'Kolkata Knight Riders'

## In[14]:

**matches.iloc[matches[matches['win\_by\_runs'].ge(1)].win\_by\_runs.idxmin()]['winner']**

out1[4]:

'Mumbai Indians'

## In[15]:

**matches.iloc[matches[matches['win\_by\_wickets'].ge(1)].win\_by\_wickets.idxmin()]**

out[15]:

id 560

season 2015

city Kolkata

date 09-05-2015

team1 Kings XI Punjab

team2 Kolkata Knight Riders

toss\_winner Kings XI Punjab

toss\_decision bat

result normal

dl\_applied 0

winner Kolkata Knight Riders

win\_by\_runs 0

win\_by\_wickets 1

player\_of\_match AD Russell

venue Eden Gardens

umpire1 AK Chaudhary

umpire2 HDPK Dharmasena

umpire3 NaN

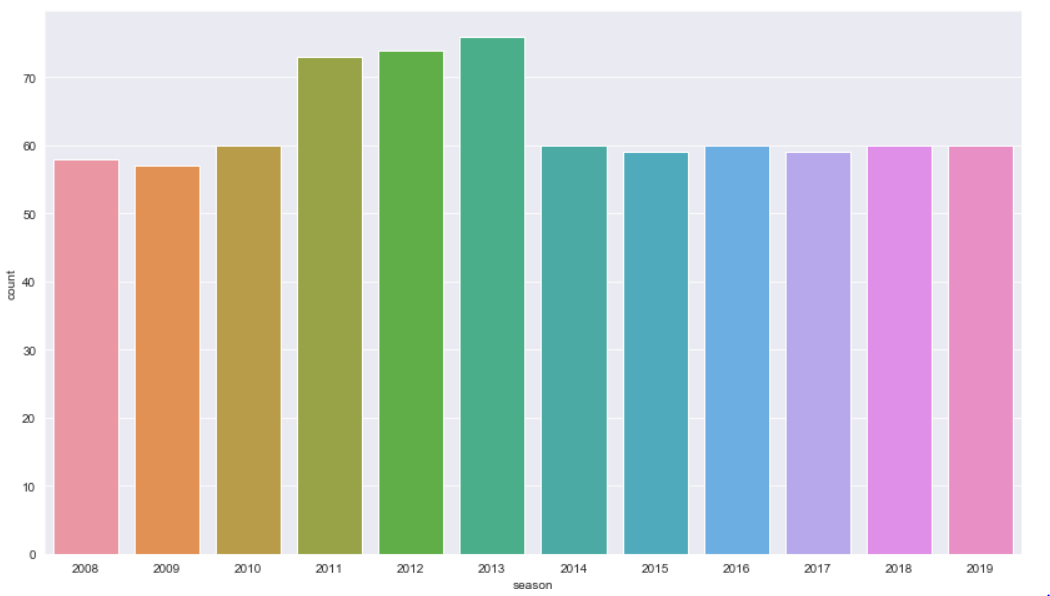
Name: 559, dtype: object

## In[16]:

**sns.countplot(x='season', data=matches)**

**plt.show()**

out[16]:

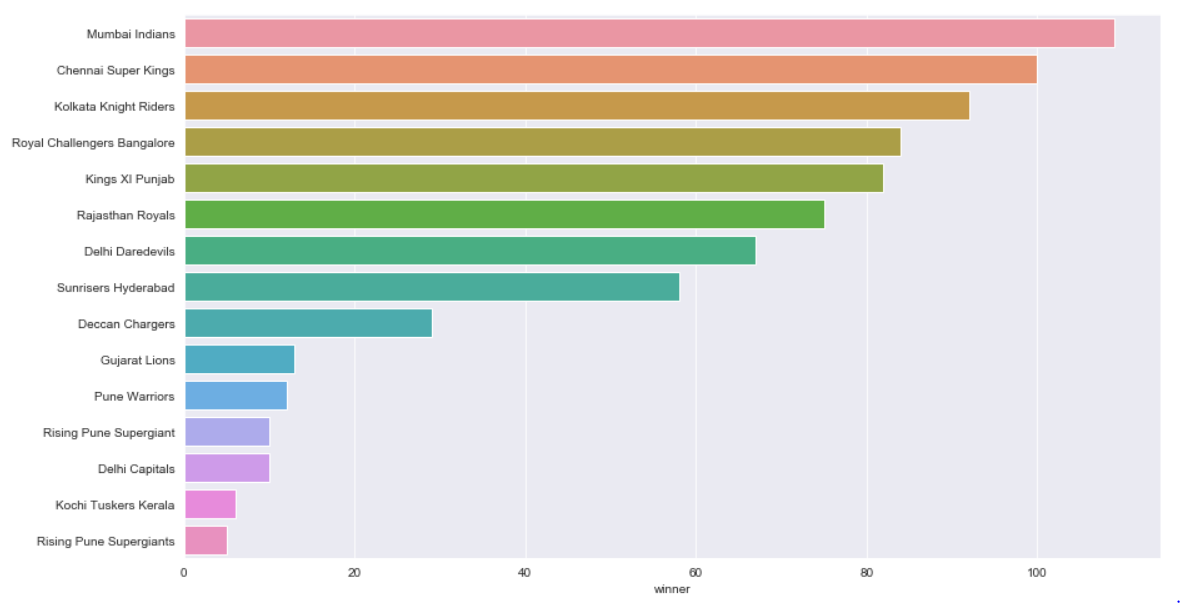
****

## In[17]:

**data = matches.winner.value\_counts()**

**sns.barplot(y = data.index, x = data, orient='h');**

out[17]:

****

## In[18]:

**top\_players = matches.player\_of\_match.value\_counts()[:10]**

**#sns.barplot(x="day", y="total\_bill", data=tips)**

**fig, ax = plt.subplots()**

**ax.set\_ylim([0,20])**

**ax.set\_ylabel("Count")**

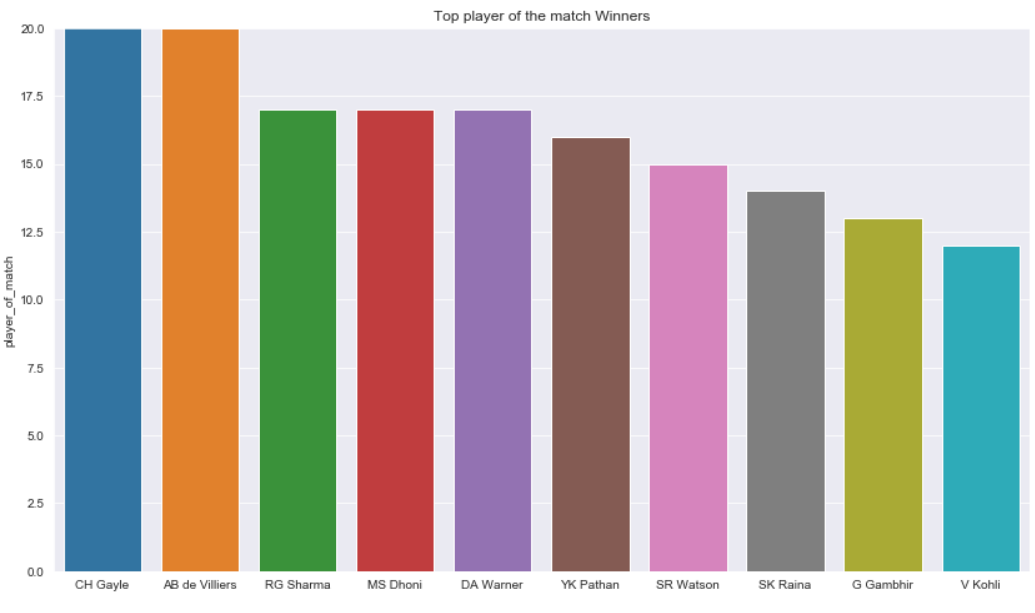
**ax.set\_title("Top player of the match Winners")**

**#top\_players.plot.bar()**

**sns.barplot(x = top\_players.index, y = top\_players, orient='v'); #palette="Blues");**

**plt.show()**

out[18]:

****

## In[19]:

**ss = matches['toss\_winner'] == matches['winner']**

**ss.groupby(ss).size()**

out[19]:

False 363

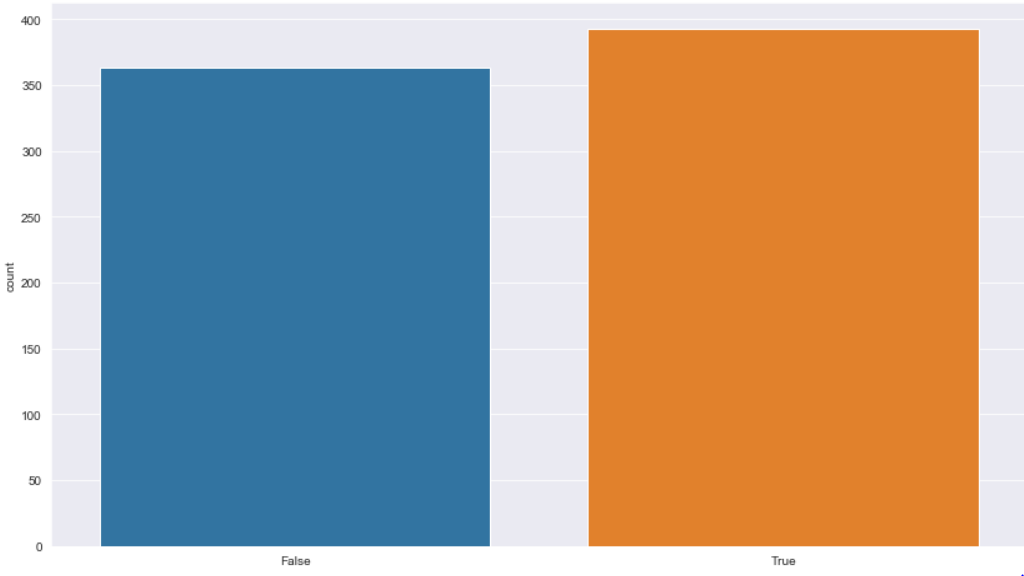
True 393

dtype: int64

## In[20]:

**sns.countplot(ss);**

out[20]:

****