

```
[In (1)]: import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
data=pd.read_csv('IPL_Matches 2008-2020.csv')
data.head()
```

Out (1):

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminator	method	umpire1	umpire2
0	335982	Bangalore	19-04-2008	BB McCullum	Chinnaswamy Stadium	M	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs	140.0	N	NaN	Asad Rauf	RE Kovi
1	335993	Chandigarh	19-04-2008	MEK Hussey	Punjab Cricket Association Stadium, Mohali		Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	33.0	N	NaN	MR Benson	SL Srinivasan
2	335984	Delhi	19-04-2008	MF Maharoof	Feroz Shah Kotla		Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	9.0	N	NaN	Aleem Dar	Pratap Kumar
3	335985	Mumbai	20-04-2008	MV Boucher	Wankhede Stadium		Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	5.0	N	NaN	SJ Davis	DJ Harris
4	335986	Kolkata	20-04-2008	DJ Hussey	Eden Gardens		Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	5.0	N	NaN	BF Bowden	K Hariharan

In (2):

#counting row and column of dataset

Out (2):

(16, 17)

In (4):

data['player\_of\_match'].value\_counts()[0:5] #counting the value of player of the match top 5 (0to4)

Out (4):

AB de Villiers	23
CH Gayle	22
RG Sharma	18
DA Warner	17
MS Dhoni	17

Name: player\_of\_match, dtype: int64

In (35):

plt.figure(figsize=(8,5)) #plt.bar(categorical values(x), numerical values(y))
plt.bar(list(data['player\_of\_match'].value\_counts()[0:5].keys()),list(data['player\_of\_match'].value\_counts()[0:5]),color='skyblue')
plt.title('ipl stats')
plt.xlabel('player of the match')
plt.ylabel('runs')
plt.show()

ipl stats

In (34):

list(data['result'].value\_counts().keys())

Out (34):

['wickets', 'runs', 'tie']

In (27):

list(data['result'].value\_counts().keys())

Out (27):

['wickets', 'runs', 'tie']

In (37):

data['toss\_winner'].value\_counts()[0:5]

Out (37):

Mumbai Indians	196
Kolkata Knight Riders	98
Chennai Super Kings	97
Royal Challengers Bangalore	87
Rajasthan Royals	87

Name: toss\_winner, dtype: int64

In (61):

batting\_first=data[data['result']=='0'] #condition where number assigning a variable

Out (61):

batting\_first.head()

Out (42):

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminator	method	umpire1	umpire2
0	335982	Bangalore	19-04-2008	BB McCullum	Chinnaswamy Stadium	M	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs	140.0	N	NaN	Asad Rauf	RE Kovi
1	335993	Chandigarh	19-04-2008	MEK Hussey	Punjab Cricket Association Stadium, Mohali		Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	33.0	N	NaN	MR Benson	SL Srinivasan
2	335984	Delhi	19-04-2008	MF Maharoof	Feroz Shah Kotla		Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	9.0	N	NaN	Aleem Dar	Pratap Kumar
3	335985	Mumbai	20-04-2008	MV Boucher	Wankhede Stadium		Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	5.0	N	NaN	SJ Davis	DJ Harris
4	335986	Kolkata	20-04-2008	DJ Hussey	Eden Gardens		Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	5.0	N	NaN	BF Bowden	K Hariharan

In (68):

#making histogram use for visualization of the numerical data
plt.figure(figsize=(7,5))
plt.hist(batting\_first['result\_margin'])
plt.title('distribution of runs')
plt.show()

distribution of runs

In (42):

#finding out the number of wins
batting\_first['winner'].value\_counts()

Out (42):

Mumbai Indians	120
Chennai Super Kings	106
Kolkata Knight Riders	99
Royal Challengers Bangalore	91
Kings XI Punjab	88
Rajasthan Royals	81
Delhi Daredevils	67
Sunrisers Hyderabad	66
Deccan Chargers	29
Delhi Capitals	19
Gujarat Lions	13
Pune Warriors	12
Rising Pune Supergiant	10
Kochi Tuskers Kerala	6
Rising Pune Supergiants	5

Name: winner, dtype: int64

In (87):

#making bar plot for batting first team
plt.figure(figsize=(10,5))
plt.bar(list(batting\_first['winner'].value\_counts()[0:5].keys()),list(batting\_first['winner'].value\_counts()[0:5]),color=['blue','yellow','purple','red','pink'])
plt.title('IPL')
plt.xlabel('teams')
plt.ylabel('no of wins')
plt.show()

IPL

In (51):

#making pie(parameter(numerical),parameter(categorical))
plt.pie(list(batting\_first['winner'].value\_counts()),labels=list(batting\_first['winner'].value\_counts().keys()),autopct='%0.1f%%')
plt.show()

Chennai Super Kings

In (73):

#extracting those records where a team has won after batting second
batting\_second=data[data['result']!='runs'] #making of variable dataframe
batting\_second.head()

Out (73):

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminator	method	umpire1	umpire2
2	335984	Delhi	19-04-2008	MF Maharoof	Feroz Shah Kotla		Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	9.0	N	NaN	Aleem Dar	Pratap Kumar
3	335985	Mumbai	20-04-2008	MV Boucher	Wankhede Stadium		Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	5.0	N	NaN	SJ Davis	DJ Harris
4	335986	Kolkata	20-04-2008	DJ Hussey	Eden Gardens		Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	5.0	N	NaN	BF Bowden	K Hariharan
5	335987	Japur	21-04-2008	SR Watson	Sawai Mansingh Stadium		Rajasthan Royals	Kings XI Punjab	Kings XI Punjab	bat	Rajasthan Royals	wickets	6.0	N	NaN	Aleem Dar	RB Tripathi
6	335988	Hyderabad	22-04-2008	V Sehwag	Rajiv Gandhi International Stadium, Uppal		Deccan Chargers	Delhi Daredevils	Deccan Chargers	bat	Delhi Daredevils	wickets	9.0	N	NaN	IL Howell	AM Sahebi

In (74):

batting\_second.value\_counts()

#making barplot of team batting\_second
plt.figure(figsize=(11,5))
plt.bar(list(batting\_second['winner'].value\_counts()[0:5].keys()),list(batting\_second['winner'].value\_counts()[0:5]),color=['purple','blue','red','pink','yellow'])
plt.title('IPL')
plt.xlabel('Teams win by wickets')
plt.ylabel('no of matches')
plt.show()

IPL

In (86):

#extracting in which stadium most matches played
data['venue'].value\_counts()[0:5]

Out (86):

Eden Gardens	77
Feroz Shah Kotla	74
Wankhede Stadium	73
Chinnaswamy Stadium	65
Rajiv Gandhi International Stadium, Uppal	64

Name: venue, dtype: int64

In (81):

#identifying team wins toss also wins the match
import numpy as np
import pandas as pd
np.sum(data['toss\_winner']==data['winner'])

Out (81):

458

Out (82):

np.sum(data['toss\_winner']==data['winner'])

Out (82):

398

In (83):

#extraction no of eliminator
data['eliminator'].value\_counts()

Out (83):

N	799
Y	13

Name: eliminator, dtype: int64

In (85):

#extraction where matches played most in which city
data['city'].value\_counts()

Out (85):

Mumbai	120
Kolkata	77
Delhi	74
Bangalore	65
Hyderabad	64
Chennai	57
Chandigarh	56
Jalpur	47
Pune	38
Abu Dhabi	29
Dubai	26
Bengaluru	15
Durban	15
Vizagapatnam	13
Ahmedabad	12
Sharjah	12
Cuttack	12
Rajkot	10
Dharamsala	9
Indore	9
Johannesburg	8
Gatwick	7
Cape Town	7
Port Elizabeth	7
Ranchi	7
Raipur	6
Kochi	5
Kanpur	4
East London	3
Kiaerley	3
Nagpur	3
Bloufontein	2
Name: city, dtype: int64	

In (86):

#extraction of umpire
data['umpire1'].value\_counts()

Out (86):

HDPK Dharmasena	78
AK Chaudhary	56
Asad Rauf	51
H Erasmus	47
Aleem Dar	38
BF Bowden	38
S Ravi	37
BR Doctrove	34
CG Gaffaney	34
AV Jadesha	33
AV Jadesha	33
NJ Llong	28
Mtin Menon	20
Clarence L. T. Southey	19
S Asnani	19
VK Kulkarni	19
BH Owenford	16
JD Cloete	16
YC Barde	14
A Nandkishore	14
BC Jerrling	13
PK Tillingworth	13
SS Hazare	11
S Das	10
MR Benson	10
K Hariharan	10
GJ Ward	9
UV Gandhi	8
CK Nandan	8
IL Howell	7
GAV Baxter	7
PH Dooland	7
SD Prisy	6
VK Sharma	6
SD Prisy	5
RJ Tucker	5
A Becham	5
PR Reiffel	5
KL Nandan	5
PG Pathak	4
JJ Gould	3
Bharatan	3
AM Sahebi	3
AV Jadesha	2
K Srikanth	2
SJA Taufel	1
SL Shastri	1
Name: umpire1, dtype: int64	

In (2):

Out (12):

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	1	Jos Buttler	17	17	2	863	116	57.53	579	149.05	4	4	83	45
1	2	K L Rahul	15	15	3	616	103*	51.33	455	135.38	2	4	45	30
2	3	Quinton De Kock	15	15	1	508	140*	36.29	341	148.97	1	3	47	23
3	4	Hardik Pandya	15	15	4	487	87*	44.27	371	131.26	0	4	49	12
4	5	Shubman Gill	16	16	2	483	96	34.5	365	132.32	0	4	51	11
5	6	David Miller	16	16	9	481	94*	68.71	337	142.72	0	2	32	23
6	7	Faf Du Plessis	16	16	1	468	96	31.2	367	127.52	0	3	49	13
7	8	Shikhar Dhawan	14	14	2	460	88*	38.33	375	122.66	0	3	47	12
8	9	Saruj Samson	17	17	1	458	55	28.63	312	146.79	0	2	43	26
9	10	Deepak Hooda	15	14	0	451	59	32.21	330	136.66	0	4	36	18

In (16):

Out (10):

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
152	153	Shivam Mavi	6	3	1	3	2	1.5	7	42.85	0	0	0	0
153	154	Harshit Rana	2	2	1	2	2	2	2	100.00	0	0	0	0
154	155	Naveed Sani	2	1	0	2	2	2	2	100.00	0	0	0	0
155	156	Tristan Stubbs	2	2	0	2	2	1	4	50.00	0	0	0	0
156	157	Pradeep Sangwan	3	1	0	2	2	2	5	40.00	0	0	0	0
157	158	Fazalhaq Farooqi	3	1	1	2	2*	-	8	25.00	0	0	0	0
158	159	Jagadeesha Suchith	5	2	0	2	2	1	8	25.00	0	0	0	0
159	160	Tim Southey	9	5	1	2	1*	0.5	12	16.66	0	0	0	0
160	161	Nathan Coulter-Nile	1	1	1	1	1*	-	2	50.00	0	0	0	0
161	162	Aurich Nortje	6	1	1	1	1*	-	6	16.66	0	0	0	0

In (13):

Out (13):

Jos Buttler	1
Simran Singh	1
Mohamed Siraj	1
Tim Seifert	1
Bhuvneshwar Kumar	1
Rinku Singh	1
San Billings	1
Oswald Brevis	1
Ayush Badoni	1
Aurich Nortje	1

Name: Player, Length: 162, dtype: int64

In (19):

import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
data=pd.read\_csv('Most Runs - 2022.csv')
data.head(10)

Out (19):

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	1	Jos Buttler	17	17	2	863	116	57.53	579	149.05	4	4	83	45
1	2	K L Rahul	15	15	3	616	103*	51.33	455	135.38	2	4	45	30
2	3	Quinton De Kock	15	15	1	508	140*	36.29	341	148.97	1	3	47	23
3	4	Hardik Pandya	15	15	4	487	87*	44.27	371	131.26	0	4	49	12
4	5	Shubman Gill	16	16	2	483	96	34.5	365	132.32	0	4	51	11
5	6	David Miller	16	16	9	481	94*	68.71	337	142.72	0	2	32	23
6	7	Faf Du Plessis	16	16	1	468	96	31.2	367	127.52	0	3	49	13
7	8	Shikhar Dhawan	14	14	2	460	88*	38.33	375	122.66	0	3	47	12
8	9	Saruj Samson	17	17	1	458	55	28.63	312	146.79	0	2	43	26
9	10	Deepak Hooda	15	14	0	451	59	32.21	330	136.66	0	4	36	18

In (24):

data['SR'].max()

Out (24):

488.0

In (23):

data['Player'].value\_counts()[0:5]

Out (23):

Jos Buttler	1
Simran Singh	1
Mohamed Siraj	1
Tim Seifert	1
Bhuvneshwar Kumar	1

Name: Player, dtype: int64

In (26):

data['HS'].value\_counts()

Out (26):

8	6
2	6
4	4
3*	3
18	3
13	1
67*	1
89	1
67	1
35	1

Name: HS, Length: 180, dtype: int64

In (27):

data['HS'].min() #minimum run

Out (27):

1\*

In (28):

data['HS'].max()

Out (28):

99\*

In (29):

data['Runs'].max() #max runs

Out (29):

863

In (30):

data['6s'].max() #most sixes

Out (30):

45

In (32):

#most runs
plt.figure(figsize=(8,5))
plt.hist(data['Runs'],color='orange')
plt.title('most runs')
plt.show()

most runs

In (61):

plt.figure(figsize=(7,7))
sns.barplot(x=data['Player'][(0:5)],y=data['Runs'][(0:5)],color='green')
plt.title('most runs')
plt.xlabel('players name')
plt.ylabel('players runs')
plt.show()

most runs

In (61):