

BBL(BIG BASH LEAGUE) Analysis using python EDA

```
In [1]: #BBL matches analysis
import numpy as np
import pandas as pd
import os
from matplotlib import pyplot as plt
```

```
In [2]: os.chdir(r"C:\Users\hp\Desktop\csv files")
```

```
In [3]: bbl=pd.read_csv("BBL matches.csv")
```

```
In [4]: bbl.head()
```

```
Out[4]:
```

	id	city	date	player_of_match	venue	neutral_venue	team
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	Sydney Sixers
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Melbourne Stars
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Adelaide Strikers
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Perth Scorchers
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Brisbane Heat

```
In [5]: #data cleaning
```

```
In [6]: #check null value presenet or not
bbl.isnull()
```

Out[6]:

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...
348	False	False	False	False	False	False	False	False	False
349	False	False	False	False	False	False	False	False	False
350	False	False	False	False	False	False	False	False	False
351	False	False	False	False	False	False	False	False	False
352	False	False	False	False	False	False	False	False	False

353 rows × 17 columns

In [7]: `#info of all column`
`bb1.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 353 entries, 0 to 352
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                     353 non-null   int64
1   city                   353 non-null   object
2   date                   353 non-null   object
3   player_of_match       348 non-null   object
4   venue                  353 non-null   object
5   neutral_venue         353 non-null   int64
6   team1                  353 non-null   object
7   team2                  353 non-null   object
8   toss_winner           353 non-null   object
9   toss_decision         353 non-null   object
10  winner                 348 non-null   object
11  result                 348 non-null   object
12  result_margin         344 non-null   float64
13  eliminator            348 non-null   object
14  method                 18 non-null    object
15  umpire1                353 non-null   object
16  umpire2                353 non-null   object
dtypes: float64(1), int64(2), object(14)
memory usage: 47.0+ KB
```

In [8]: `#drop null values`
`bb1.dropna()`

Out[8]:

	id	city	date	player_of_match	venue	neutral_venue	
15	524933	Sydney	1/8/2012	MA Starc	Stadium Australia	0	T
31	571236	Perth	12/12/2012	SL Malinga	Western Australia Cricket Association Ground	0	Sc
36	571241	Brisbane	12/18/2012	NM Coulter-Nile	Brisbane Cricket Ground, Woolloongabba	0	B
53	571265	Perth	2/5/2013	SE Marsh	Western Australia Cricket Association Ground	0	Sc
87	654097	Sydney	2/5/2014	CJ Simmons	Sydney Cricket Ground	0	
166	1023603	Melbourne	1/1/2017	GB Hogg	Melbourne Cricket Ground	0	Mel
197	1114863	Brisbane	12/27/2017	AI Ross	Brisbane Cricket Ground, Woolloongabba	0	B
235	1152513	Canberra	12/21/2018	DR Sams	Manuka Oval	0	T
256	1152534	Sydney	1/8/2019	CA Lynn	Sydney Showground Stadium	0	T
281	1152559	Sydney	2/2/2019	DP Hughes	Sydney Cricket Ground	0	
287	1152565	Canberra	2/9/2019	PJ Cummins	Manuka Oval	0	T
301	1195582	Adelaide	12/23/2019	J Weatherald	Adelaide Oval	0	A
307	1195588	Launceston	12/30/2019	NM Coulter-Nile	Aurora Stadium	0	Hur
316	1195597	Sydney	1/6/2020	T Banton	Sydney Showground Stadium	0	T
328	1195609	Canberra	1/15/2020	SB Harper	Manuka Oval	0	T
333	1195614	Sydney	1/18/2020	CH Morris	Sydney Showground	0	T

	id	city	date	player_of_match	venue	neutral_venue	
					Stadium		
336	1195617	Sydney	1/20/2020	MC Henriques	Sydney Cricket Ground	0	
345	1195626	Sydney	1/26/2020	AD Hales	Sydney Showground Stadium	0	T

In [9]: `bbl.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 353 entries, 0 to 352
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                     353 non-null   int64
1   city                   353 non-null   object
2   date                   353 non-null   object
3   player_of_match       348 non-null   object
4   venue                  353 non-null   object
5   neutral_venue         353 non-null   int64
6   team1                  353 non-null   object
7   team2                  353 non-null   object
8   toss_winner           353 non-null   object
9   toss_decision         353 non-null   object
10  winner                 348 non-null   object
11  result                 348 non-null   object
12  result_margin          344 non-null   float64
13  eliminator             348 non-null   object
14  method                 18 non-null    object
15  umpire1                353 non-null   object
16  umpire2                353 non-null   object
dtypes: float64(1), int64(2), object(14)
memory usage: 47.0+ KB
```

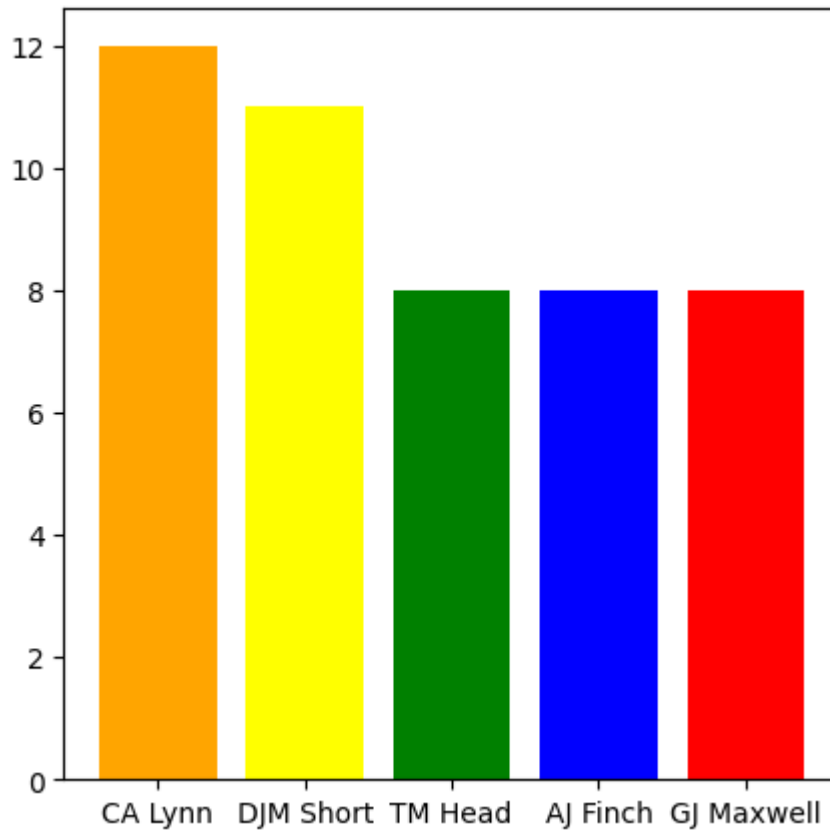
In [10]: `# frist five records of dataset`
`bbl.head()`

Out[10]:	id	city	date	player_of_match	venue	neutral_venue	team
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	Sydney Sixers
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Melbourne Stars
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Adelaide Strikers
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Perth Scorchers
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Brisbane Heat

```
In [11]: #most time player of the match
bbl["player_of_match"].value_counts()[0:5]
```

```
Out[11]: player_of_match
CA Lynn      12
DJM Short    11
TM Head      8
AJ Finch     8
GJ Maxwell   8
Name: count, dtype: int64
```

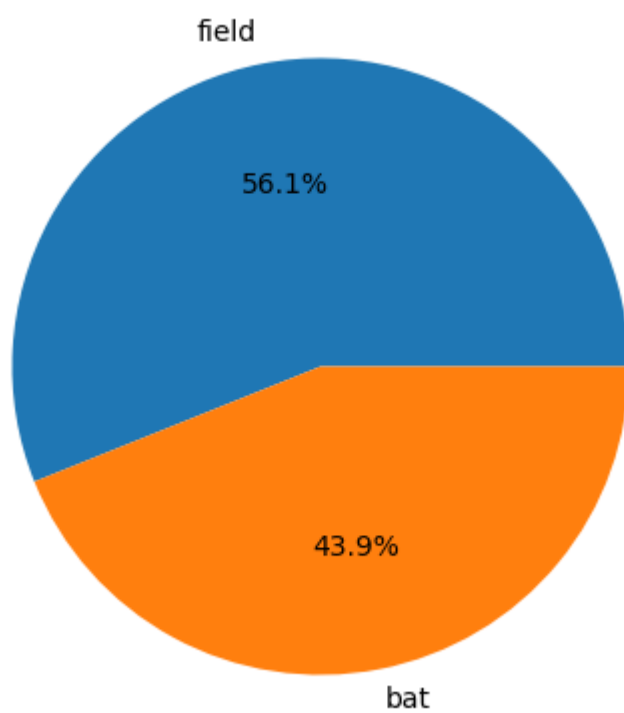
```
In [12]: #most time player of the match bar plot
plt.figure(figsize=(5,5))
plt.bar(list(bbl["player_of_match"].value_counts()[0:5].keys()),list(bbl["player_of
plt.show()
```



```
In [13]: #toss decision after win toss  
bbl["toss_decision"].value_counts()
```

```
Out[13]: toss_decision  
field    198  
bat      155  
Name: count, dtype: int64
```

```
In [14]: plt.figure(figsize=(5,5))  
plt.pie(list(bbl["toss_decision"].value_counts()),labels=list(bbl["toss_decision"]  
plt.show()
```



In []:

In [15]: **bb1**

Out[15]:

	id	city	date	player_of_match	venue	neutral_venue	
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Mell
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Ac
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Sc
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Br
...	
348	1195629	Hobart	1/30/2020	JD Cook	Bellerive Oval	0	Hur
349	1195630	Melbourne	1/31/2020	SNJ O'Keefe	Melbourne Cricket Ground	0	Mell
350	1195631	Adelaide	2/1/2020	AD Hales	Adelaide Oval	0	Ac
351	1195632	Melbourne	2/6/2020	NCR Larkin	Melbourne Cricket Ground	0	Mell
352	1195633	Sydney	2/8/2020	JR Philippe	Sydney Cricket Ground	0	

353 rows × 17 columns

In [32]: `bbl["toss_winner"].value_counts()`

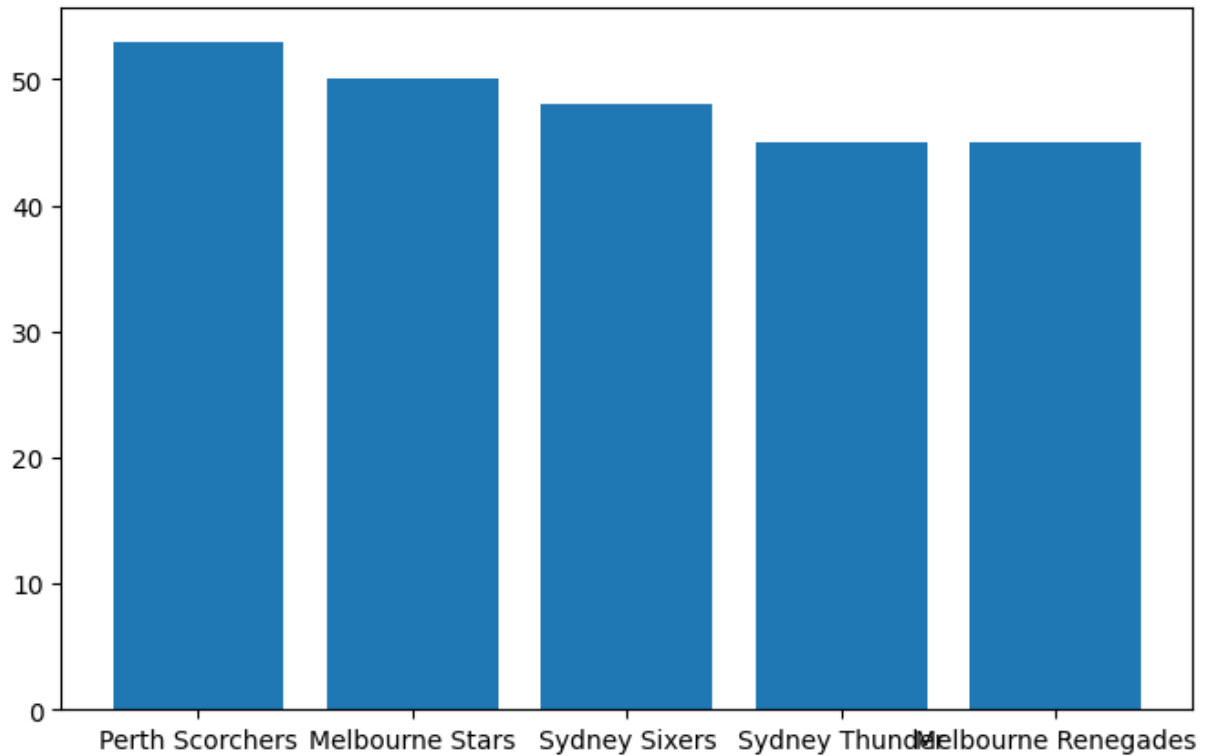
Out[32]:

toss_winner	
Perth Scorchers	53
Melbourne Stars	50
Sydney Sixers	48
Sydney Thunder	45
Melbourne Renegades	45
Brisbane Heat	39
Hobart Hurricanes	37
Adelaide Strikers	36

Name: count, dtype: int64

In [44]:

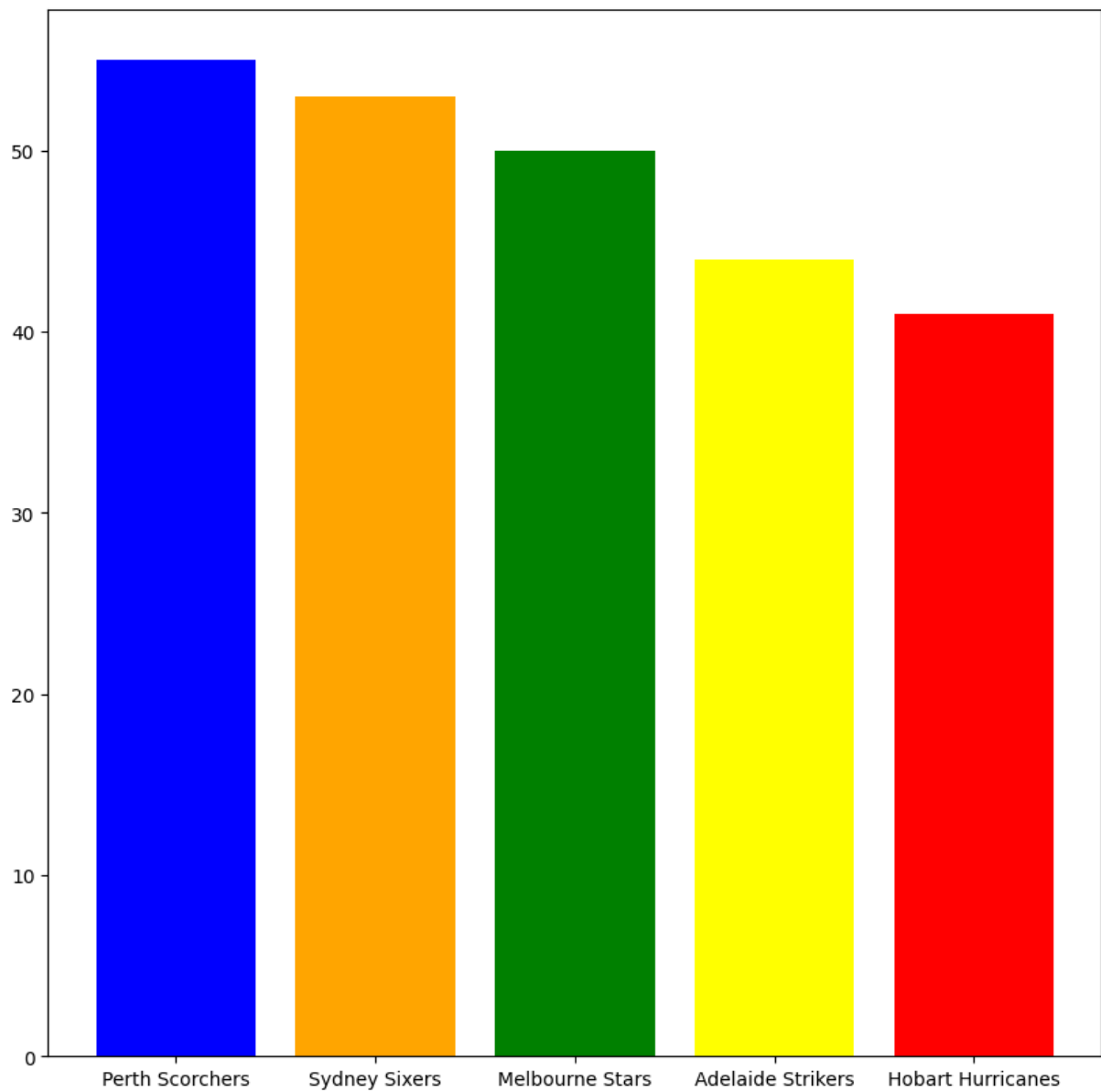
```
plt.figure(figsize=(8,5))
plt.bar(list(bbl["toss_winner"].value_counts()[0:5].keys()),list((bbl["toss_winner"]
plt.show()
```

`bbl["winner"].value_counts()`

In []:

```
In [17]: plt.figure(figsize=(10,10))
plt.bar(list(bbl["winner"].value_counts()[0:5].keys()),list(bbl["winner"].value_cou
plt.show()
```

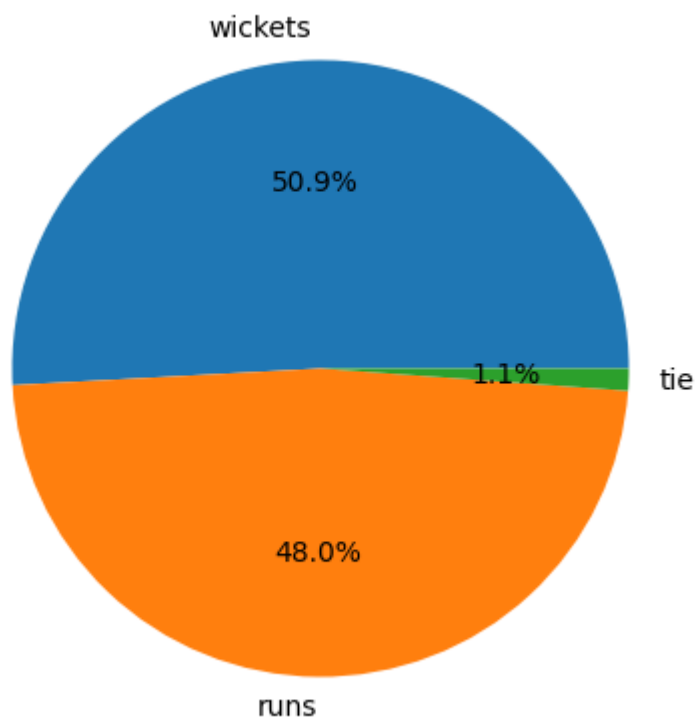


```
In [18]: #result after match  
bbl["result"].value_counts()
```

```
Out[18]: result  
wickets    177  
runs       167  
tie         4  
Name: count, dtype: int64
```

```
In [ ]:
```

```
In [19]: plt.figure(figsize=(5,5))  
plt.pie(list(bbl["result"].value_counts()),labels=list(bbl["result"].value_counts()  
plt.show()
```



In []:

```
In [20]: #teams won by wickets  
bb1[bb1["result"] == "wickets"]
```

Out[20]:

	id	city	date	player_of_match	venue	neutral_venue	
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Mell
6	524921	Melbourne	12/22/2011	SE Marsh	Docklands Stadium	0	Mell Ren
7	524922	Sydney	12/23/2011	CH Gayle	Stadium Australia	0	T
12	524927	Hobart	1/1/2012	Naved-ul-Hasan	Bellerive Oval	0	Hur
...	
337	1195618	Perth	1/20/2020	LS Livingstone	Perth Stadium	0	Sc
340	1195621	Brisbane	1/23/2020	SNJ O'Keefe	Brisbane Cricket Ground, Woolloongabba	0	Br
343	1195624	Sydney	1/25/2020	SPD Smith	Sydney Cricket Ground	0	
345	1195626	Sydney	1/26/2020	AD Hales	Sydney Showground Stadium	0	T
347	1195628	Melbourne	1/27/2020	AJ Finch	Docklands Stadium	0	Mell Ren

177 rows × 17 columns



In [21]:

```
#teams won by runs  
bbl[bbl["result"] == "runs"]
```

Out[21]:

	id	city	date	player_of_match	venue	neutral_venue	
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Ac
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Sc
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Br
5	524920	Hobart	12/21/2011	TR Birt	Bellerive Oval	0	Hur
8	524923	Sydney	12/27/2011	DJ Bravo	Sydney Cricket Ground	0	
...	
348	1195629	Hobart	1/30/2020	JD Cook	Bellerive Oval	0	Hur
349	1195630	Melbourne	1/31/2020	SNJ O'Keefe	Melbourne Cricket Ground	0	Mell
350	1195631	Adelaide	2/1/2020	AD Hales	Adelaide Oval	0	Ac
351	1195632	Melbourne	2/6/2020	NCR Larkin	Melbourne Cricket Ground	0	Mell
352	1195633	Sydney	2/8/2020	JR Philippe	Sydney Cricket Ground	0	

167 rows × 17 columns

In [22]:

```
# toss decision is bat frist
bbl[bbl["toss_decision"] == "bat"]
```

Out[22]:

	id	city	date	player_of_match	venue	neutral_venue	
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Sc
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Br
5	524920	Hobart	12/21/2011	TR Birt	Bellerive Oval	0	Hur
6	524921	Melbourne	12/22/2011	SE Marsh	Docklands Stadium	0	Mell Ren
...	
345	1195626	Sydney	1/26/2020	AD Hales	Sydney Showground Stadium	0	T
346	1195627	Adelaide	1/26/2020	MS Wade	Adelaide Oval	0	Ac s
347	1195628	Melbourne	1/27/2020	AJ Finch	Docklands Stadium	0	Mell Ren
348	1195629	Hobart	1/30/2020	JD Cook	Bellerive Oval	0	Hur
351	1195632	Melbourne	2/6/2020	NCR Larkin	Melbourne Cricket Ground	0	Mell

155 rows × 17 columns

In [23]:

```
#toss decision field frist
bbl[bbl["toss_decision"] == "field"]
```

Out[23]:

	id	city	date	player_of_match	venue	neutral_venue	
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Mell
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Ac
20	524939	Brisbane	1/17/2012	DT Christian	Brisbane Cricket Ground, Woolloongabba	0	Bi
28	571233	Sydney	12/8/2012	BJ Haddin	Sydney Cricket Ground	0	
31	571236	Perth	12/12/2012	SL Malinga	Western Australia Cricket Association Ground	0	Sc
...	
340	1195621	Brisbane	1/23/2020	SNJ O'Keefe	Brisbane Cricket Ground, Woolloongabba	0	Bi
343	1195624	Sydney	1/25/2020	SPD Smith	Sydney Cricket Ground	0	
349	1195630	Melbourne	1/31/2020	SNJ O'Keefe	Melbourne Cricket Ground	0	Mell
350	1195631	Adelaide	2/1/2020	AD Hales	Adelaide Oval	0	Ac
352	1195633	Sydney	2/8/2020	JR Philippe	Sydney Cricket Ground	0	

198 rows × 17 columns

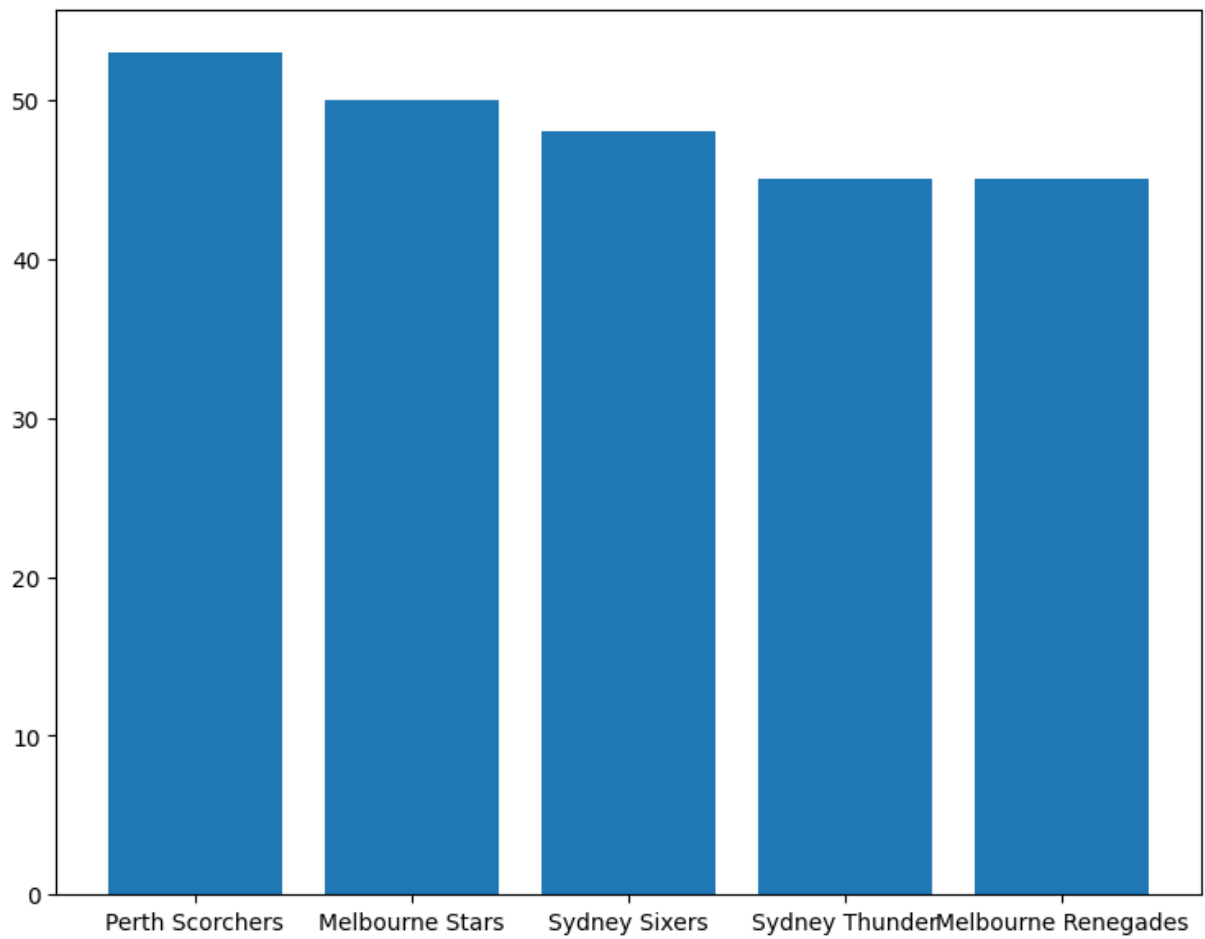
In [24]:

```
#most time toss winner
bbl["toss_winner"].value_counts()
```

Out[24]:

```
toss_winner
Perth Scorchers      53
Melbourne Stars      50
Sydney Sixers        48
Sydney Thunder       45
Melbourne Renegades  45
Brisbane Heat        39
Hobart Hurricanes    37
Adelaide Strikers    36
Name: count, dtype: int64
```

```
In [25]: plt.figure(figsize=(9,7))
plt.bar(list(bbl["toss_winner"].value_counts()[0:5].keys()),list(bbl["toss_winner"]
plt.show()
```



```
In [ ]:
```

```
In [26]: #toss winner == match winner
bbl[bbl["toss_winner"] == bbl["winner"]]
```


Out[26]:

	id	city	date	player_of_match	venue	neutral_venue	
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Mell
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Sc
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Br
5	524920	Hobart	12/21/2011	TR Birt	Bellerive Oval	0	Hur
8	524923	Sydney	12/27/2011	DJ Bravo	Sydney Cricket Ground	0	
...	
343	1195624	Sydney	1/25/2020	SPD Smith	Sydney Cricket Ground	0	
344	1195625	Melbourne	1/25/2020	AB de Villiers	Melbourne Cricket Ground	0	Mell
346	1195627	Adelaide	1/26/2020	MS Wade	Adelaide Oval	0	Ac
348	1195629	Hobart	1/30/2020	JD Cook	Bellerive Oval	0	Hur
351	1195632	Melbourne	2/6/2020	NCR Larkin	Melbourne Cricket Ground	0	Mell

193 rows × 17 columns



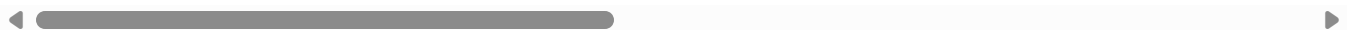
```
In [27]: #teams win with bowling frist
bowl=bbl[bbl["result"] == "wickets"]
```

```
In [28]: bowl
```

Out[28]:

	id	city	date	player_of_match	venue	neutral_venue	
0	524915	Sydney	12/16/2011	BJ Haddin	Sydney Cricket Ground	0	
1	524916	Melbourne	12/17/2011	DA Warner	Melbourne Cricket Ground	0	Mell
6	524921	Melbourne	12/22/2011	SE Marsh	Docklands Stadium	0	Mell Ren
7	524922	Sydney	12/23/2011	CH Gayle	Stadium Australia	0	T
12	524927	Hobart	1/1/2012	Naved-ul-Hasan	Bellerive Oval	0	Hur
...	
337	1195618	Perth	1/20/2020	LS Livingstone	Perth Stadium	0	Sc
340	1195621	Brisbane	1/23/2020	SNJ O'Keefe	Brisbane Cricket Ground, Woolloongabba	0	Br
343	1195624	Sydney	1/25/2020	SPD Smith	Sydney Cricket Ground	0	
345	1195626	Sydney	1/26/2020	AD Hales	Sydney Showground Stadium	0	T
347	1195628	Melbourne	1/27/2020	AJ Finch	Docklands Stadium	0	Mell Ren

177 rows × 17 columns



```
In [29]: #sort values resul margin  
bowl.sort_values(by="result_margin")
```

Out[29]:

	id	city	date	player_of_match	venue	neutral_venue	
93	756743	Perth	12/22/2014	JP Behrendorff	Western Australia Cricket Association Ground	0	Sc
176	1023623	Melbourne	1/10/2017	BW Hilfenhaus	Melbourne Cricket Ground	0	Mell
178	1023627	Melbourne	1/12/2017	BR McDermott	Docklands Stadium	0	Mell Ren
318	1195599	Adelaide	1/8/2020	TK Curran	Adelaide Oval	0	Ac s
154	897763	Melbourne	1/24/2016	UT Khawaja	Melbourne Cricket Ground	0	Mell
...
76	654075	Melbourne	1/14/2014	GS Sandhu	Docklands Stadium	0	Mell Ren
285	1152563	Brisbane	2/8/2019	BCJ Cutting	Brisbane Cricket Ground, Woolloongabba	0	Br
269	1152547	Adelaide	1/21/2019	MS Wade	Adelaide Oval	0	Ac s
133	897719	Melbourne	12/30/2015	M Klinger	Docklands Stadium	0	Mell Ren
331	1195612	Adelaide	1/17/2020	LN O'Connor	Adelaide Oval	0	Ac s

177 rows × 17 columns



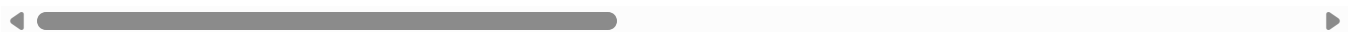
```
In [30]: #teams win by runs  
bat=bbl[bbl["result"] == "runs"]
```

```
In [31]: bat
```

Out[31]:

	id	city	date	player_of_match	venue	neutral_venue	
2	524917	Adelaide	12/18/2011	M Klinger	Adelaide Oval	0	Ac
3	524918	Perth	12/18/2011	BW Hilfenhaus	Western Australia Cricket Association Ground	0	Sc
4	524919	Brisbane	12/20/2011	RJ Quiney	Brisbane Cricket Ground, Woolloongabba	0	Br
5	524920	Hobart	12/21/2011	TR Birt	Bellerive Oval	0	Hur
8	524923	Sydney	12/27/2011	DJ Bravo	Sydney Cricket Ground	0	
...	
348	1195629	Hobart	1/30/2020	JD Cook	Bellerive Oval	0	Hur
349	1195630	Melbourne	1/31/2020	SNJ O'Keefe	Melbourne Cricket Ground	0	Mell
350	1195631	Adelaide	2/1/2020	AD Hales	Adelaide Oval	0	Ac
351	1195632	Melbourne	2/6/2020	NCR Larkin	Melbourne Cricket Ground	0	Mell
352	1195633	Sydney	2/8/2020	JR Philippe	Sydney Cricket Ground	0	

167 rows × 17 columns



In []: