



Project Work

Informatics Practices Class XII Project (2020–21)

Front-end: Pandas

Back-end: csv

Plots: matplotlib

#Project on Cricket 20-20 Analysis

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt #visualization
plt.rcParams['figure.figsize'] = (14, 8)

df = pd.read_csv("E:\\t20.csv")
print('-----')
print('-----')
print(df.info())
print()
print('-----')
print('-----')
print('Total Matches are:::',df['id'].max())
print()
print('-----')
print('-----')
print('How many seasons data we\'ve got in the dataset?')
print(df['season'].unique())
print()
```

```

print('-----')
print('-----')
print('Which Team had won by maximum runs?')
print(df.iloc[df['win_by_runs'].idxmax()])
print()
print('-----')
print('-----')
print('Which Team had won by maximum wickets?')
print(df.iloc[df['win_by_wickets'].idxmax()][['winner']])
print()
print('-----')
print('-----')
print('Which Team had won by (closest margin) minimum runs?')
print(df.iloc[df[df['win_by_runs'].ge(1)].win_by_runs.idxmin()][
['winner']])
print()
print('-----')
print('-----')
print('Which Team had won by minimum wickets?')
print(df.iloc[df[df['win_by_wickets'].ge(1)].win_by_wickets.idxmin()])

print()
print('-----')
print('-----')
print('Which season had most number of matches?')
plt.bar(x='season',data=df,height=df['season'])
plt.show()
print()
print('-----')
print('-----')
print('The Most Successful IPL Team is:::')
data = df.winner.value_counts()
print(data)
print()
print('-----')
print('-----')
print('The Players who got maximum times Man of the Match are:::')
top_players = df.player_of_match.value_counts()[:10]

```

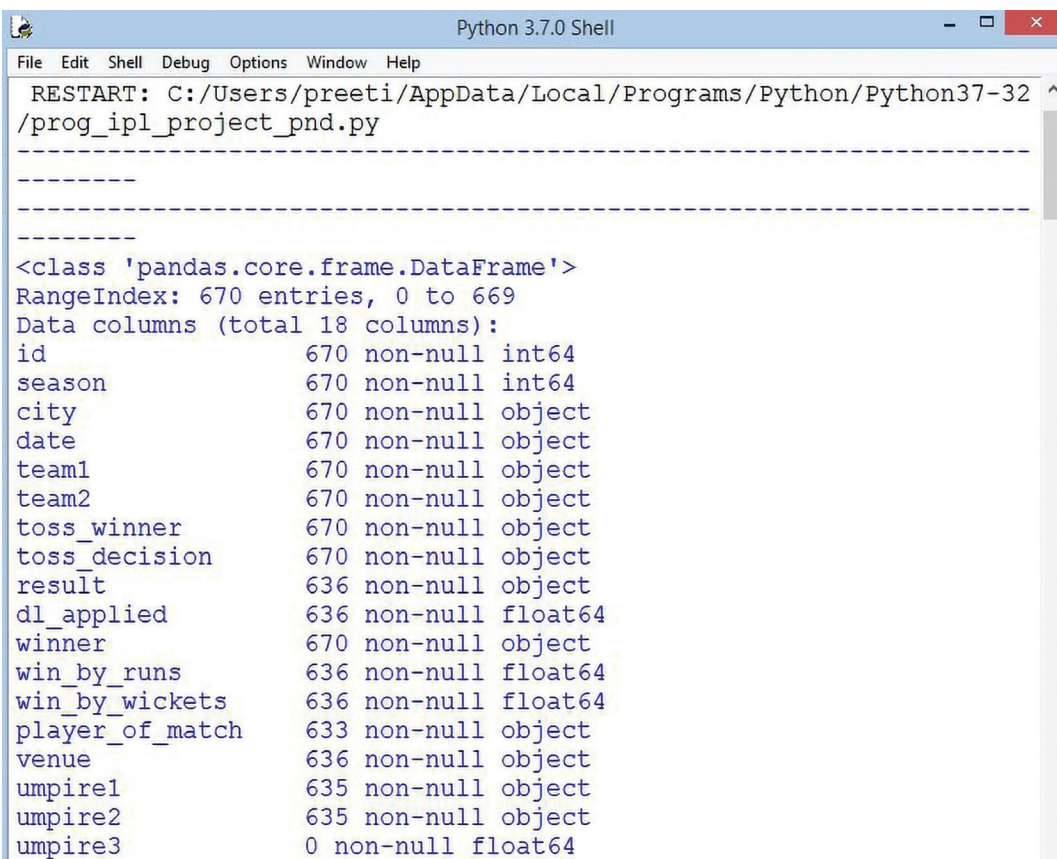
```

print(top_players)
print("Data Frame Analysis")
menu='' 1. Top record of the Players
\n 2. Bottom Records of the Players
\n 3. To print particular column
\n 4. To print multiple columns
\n 5. To display complete statistics of the Matches
\n Press enter to go back ''
print(menu)
ch=int(input("Enter your choice"))
if ch==1:
    n=int(input("Enter the number of records to be displayed"))
    print("Top ", n, " records from the dataframe")
    print(df.head(n))
elif ch==2:
    n=int(input("Enter the number of records to be displayed"))
    print("Bottom ", n, " records from the dataframe")
    print(df.tail(n))
elif ch==3:
    print("Name of the columns\n",df.columns)
    columns\n",df.columns)
    co=input("Enter the column name to be displayed")
    print(df[[co]])
elif ch==4:
    print("Name of the columns\n",df.columns)
    co=eval(input("Enter the column names as list in square bracket"))
    print(df[co])
elif ch==5:
    print("The statistics of the dataframe is:", df.describe(include='all'))

```

OUTPUT SCREENSHOTS

```
-----
Which Team had won by maximum runs?
id                44
season            2017
city              Delhi
date              05-06-2017
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
-----
```

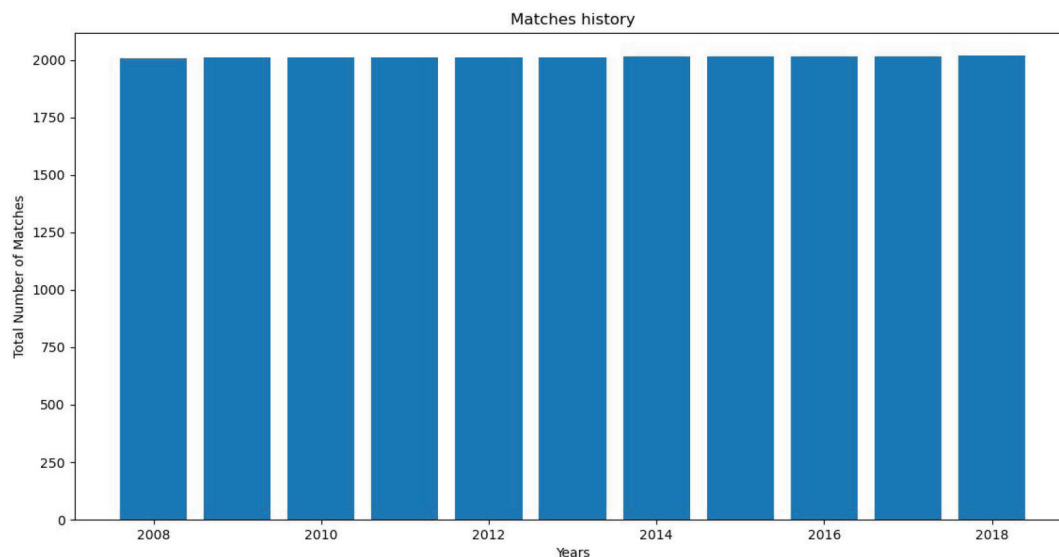


The screenshot shows a Python 3.7.0 Shell window with a menu bar (File, Edit, Shell, Debug, Options, Window, Help) and a title bar. The command prompt shows the execution of a script: `RESTART: C:/Users/preeti/AppData/Local/Programs/Python/Python37-32 /prog_ipl_project_pnd.py`. The output displays a pandas DataFrame with 670 entries and 18 columns. The columns and their data types are listed as follows:

Column	Count	Null Status	Data Type
id	670	non-null	int64
season	670	non-null	int64
city	670	non-null	object
date	670	non-null	object
team1	670	non-null	object
team2	670	non-null	object
toss_winner	670	non-null	object
toss_decision	670	non-null	object
result	636	non-null	object
dl_applied	636	non-null	float64
winner	670	non-null	object
win_by_runs	636	non-null	float64
win_by_wickets	636	non-null	float64
player_of_match	633	non-null	object
venue	636	non-null	object
umpire1	635	non-null	object
umpire2	635	non-null	object
umpire3	0	non-null	float64

```
-----  
-----  
Total Matches are:::: 670  
-----  
-----  
-----
```

```
How many seasons data we've got in the dataset?  
[2017 2008 2009 2010 2011 2012 2013 2014 2015 2016 2018]
```



```
-----  
-----  
The Most Successful IPL Team is:::
```

Mumbai Indians	95
Chennai Super Kings	85
Kolkata Knight Riders	82
Royal Challengers Bangalore	76
Kings XI Punjab	75
Rajasthan Royals	67
Delhi Daredevils	65
Sunrisers Hyderabad	47
Deccan Chargers	29
Gujarat Lions	13
Pune Warriors	12
Rising Pune Supergiant	10
Kochi Tuskers Kerala	6
Rising Pune Supergiants	5
no result	3

```
Name: winner, dtype: int64  
-----  
-----  
-----
```

```
The Players who got maximum times Man of the Match are:::
```

CH Gayle	18
YK Pathan	16
DA Warner	15
AB de Villiers	15
SK Raina	14

```

646 Bangalore
647 Mohali
648 Kolkata
649 Mumbai
650 Jaipur
651 Mohali
652 Pune
653 Kolkata
654 Bangalore
655 Hyderabad
656 Jaipur
657 Delhi
658 Mumbai
659 Bangalore
660 Hyderabad
661 Delhi
662 Pune
663 Jaipur
664 Bangalore
665 Pune
666 Bangalore
667 Delhi
668 Kolkata
669 Indore

[670 rows x 1 columns]
>>>

```

Ln: 221 Col: 4

Data Frame Analysis

1. Top record of the Players
2. Bottom Records of the Players
3. To print particular column
4. To print multiple columns
5. To display complete statistics of the Matches

Press enter to go back

Enter your choice3

Name of the columns

```

Index(['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'],
      dtype='object')

```

Enter the column name to be displayedcity

```

city
0 Hyderabad
1 Pune
2 Rajkot
3 Indore
4 Bangalore

```