EDS Activity No. 1

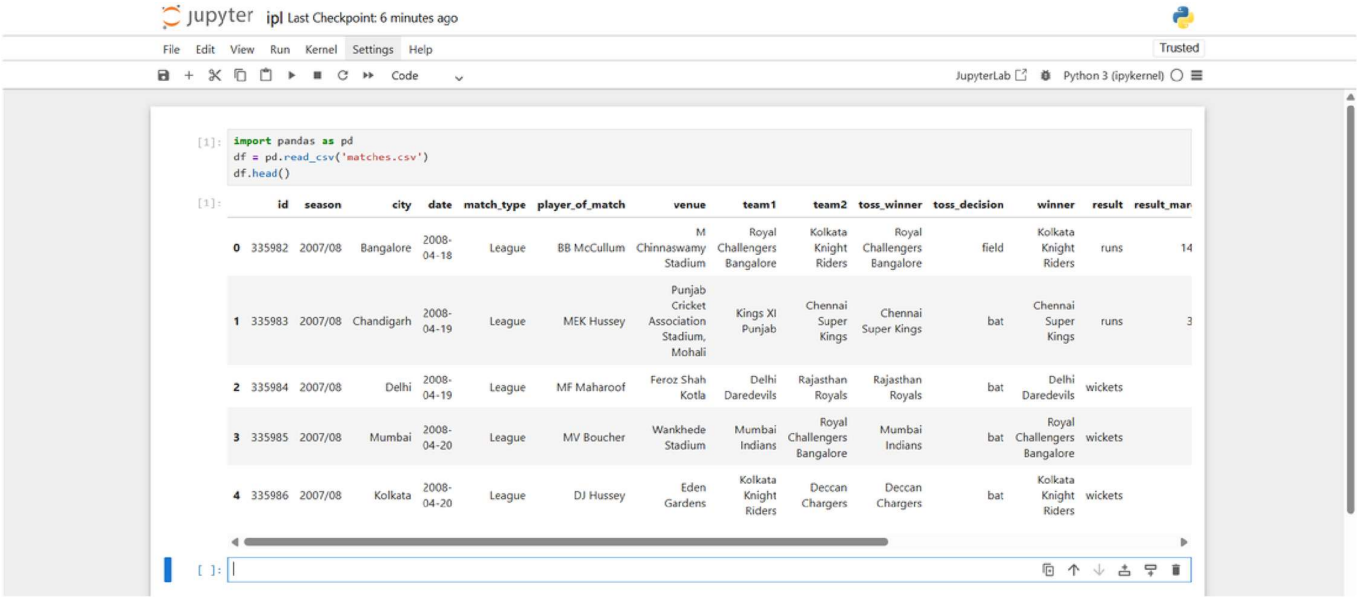
Name : Yash Jambhale

PRN : 202401040254

Roll No. : CS1-09

Dataset : IPL

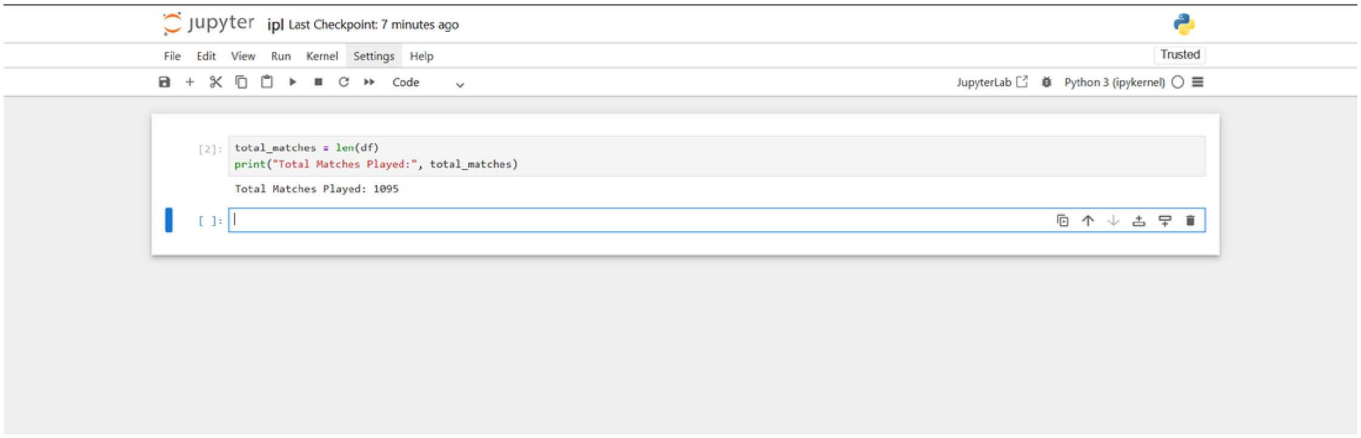
Importing dataset into Jupyter Python Notebook



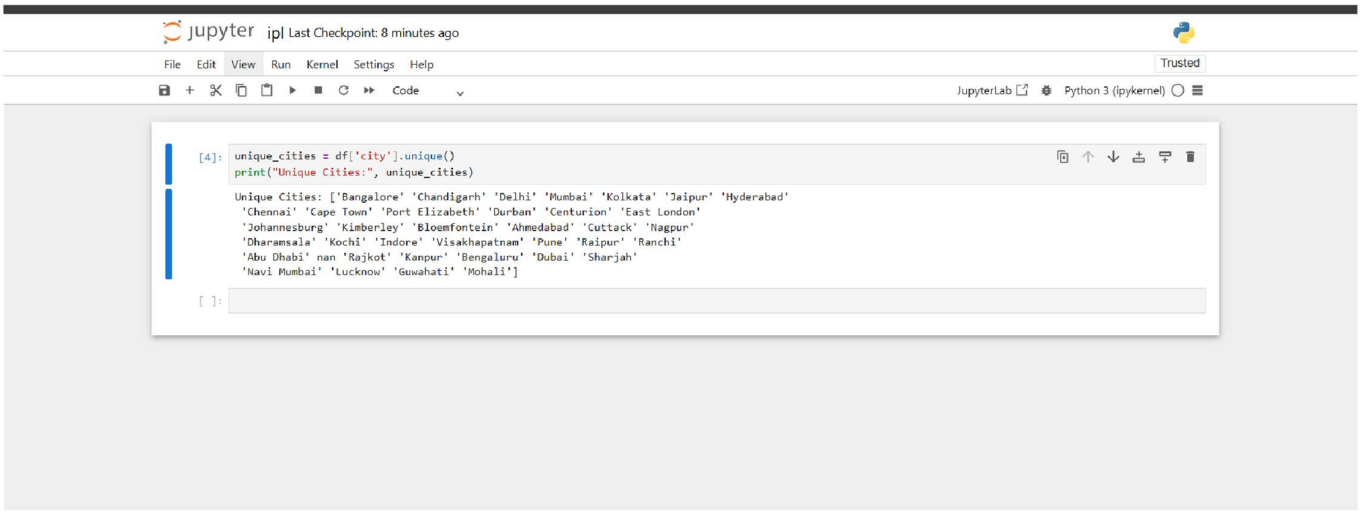
	id	season	city	date	match_type	player_of_match	venue	team1	team2	toss_winner	toss_decision	winner	result	result_mar
0	335982	2007/08	Bangalore	2008-04-18	League	BB McCullum	M Chinnaswamy Stadium	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs	14
1	335983	2007/08	Chandigarh	2008-04-19	League	MEK Hussey	Punjab Cricket Association Stadium, Mohali	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	3
2	335984	2007/08	Delhi	2008-04-19	League	MF Maharoof	Feroz Shah Kotla	Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	
3	335985	2007/08	Mumbai	2008-04-20	League	MV Boucher	Wankhede Stadium	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	
4	335986	2007/08	Kolkata	2008-04-20	League	DJ Hussey	Eden Gardens	Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	

Problem Statements:

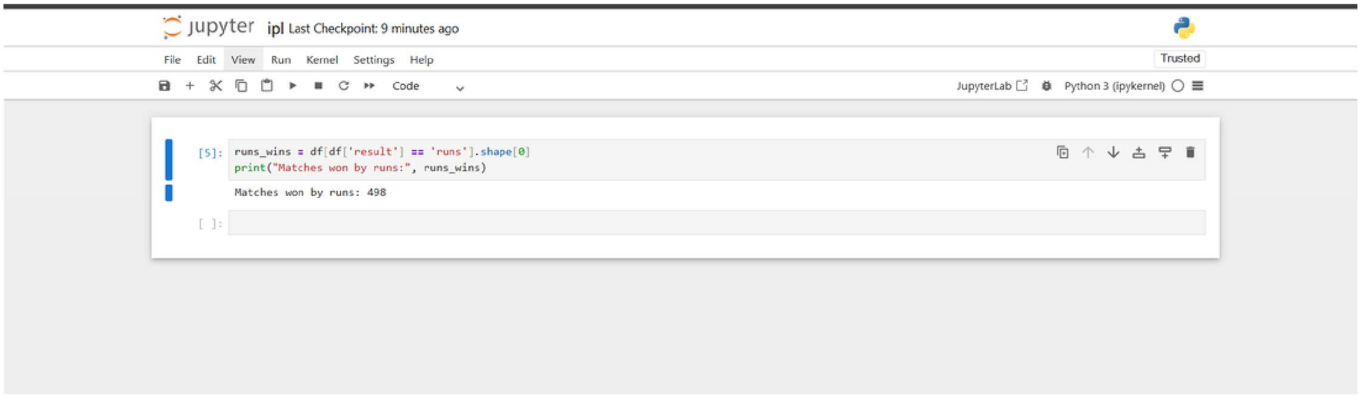
1.Find the total number of matches played.



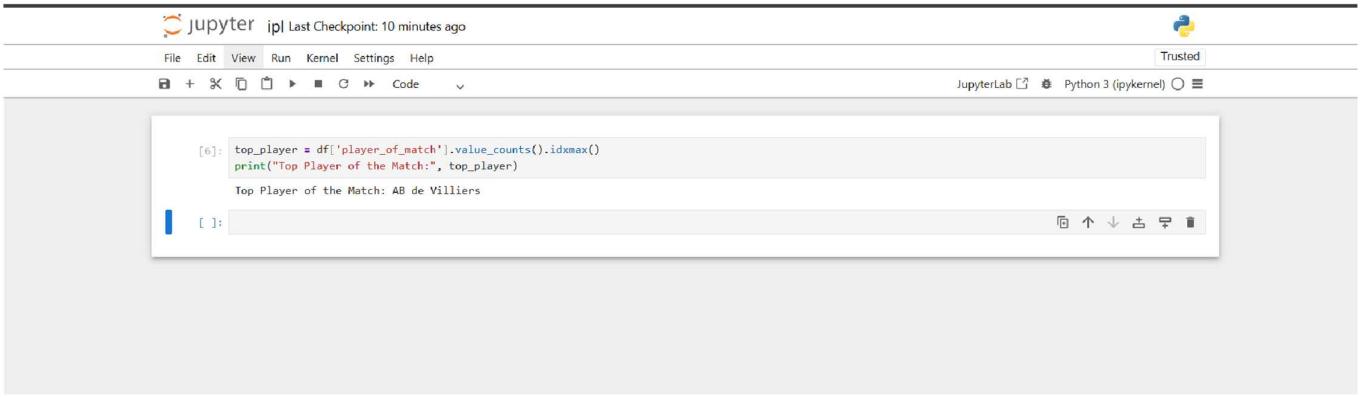
2.Listing all unique cities where matches were played.



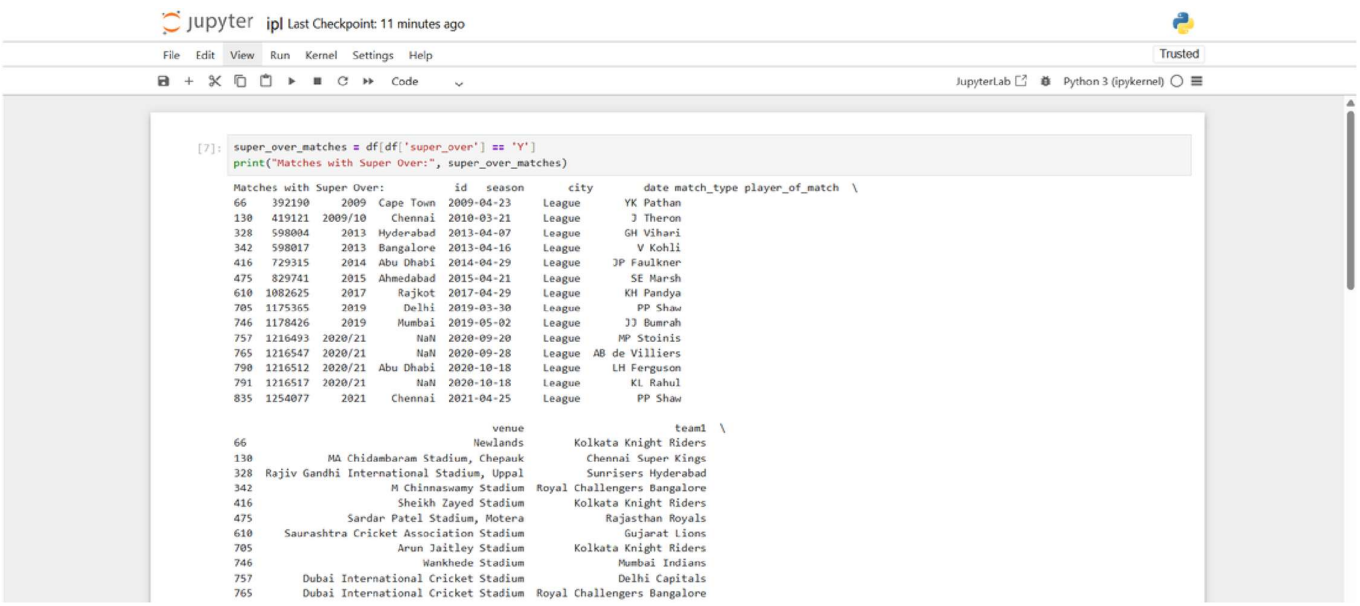
3.Finding out how many matches had a result decided by runs.



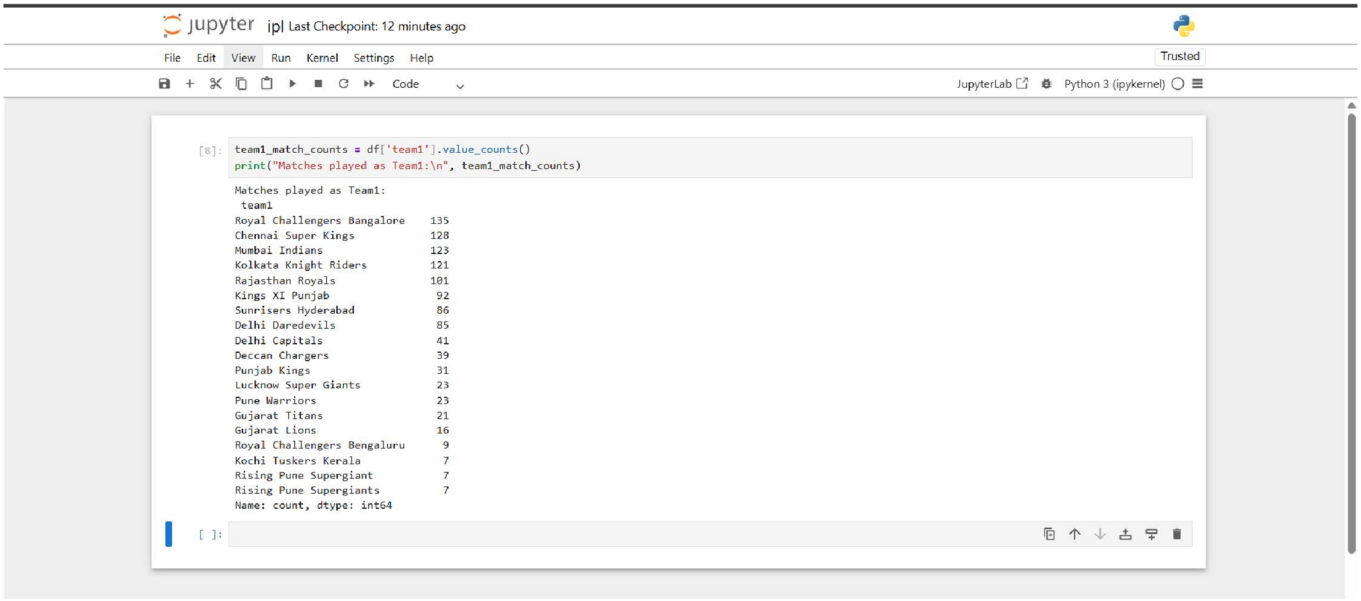
4.Find the player with the most 'Player of the Match' awards.



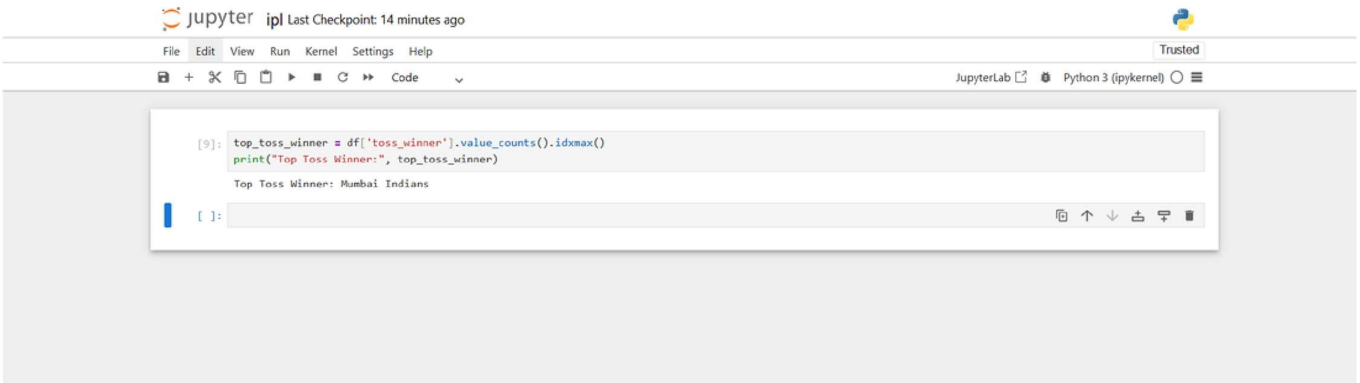
5.Find all matches where a Super Over was played.



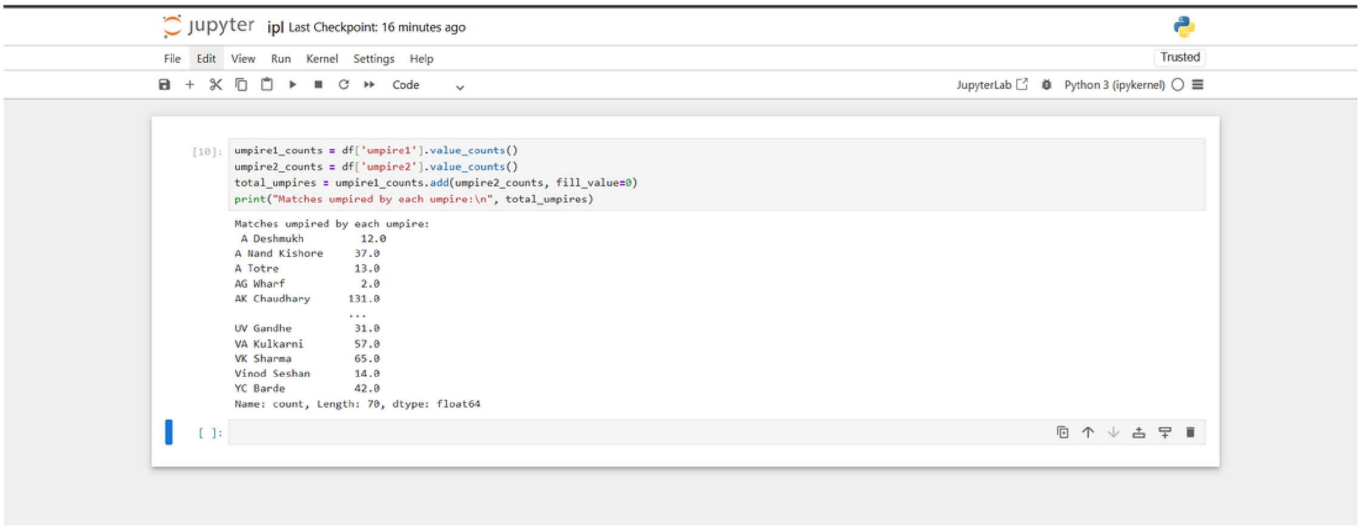
6.Count how many matches each team has played as Team1.



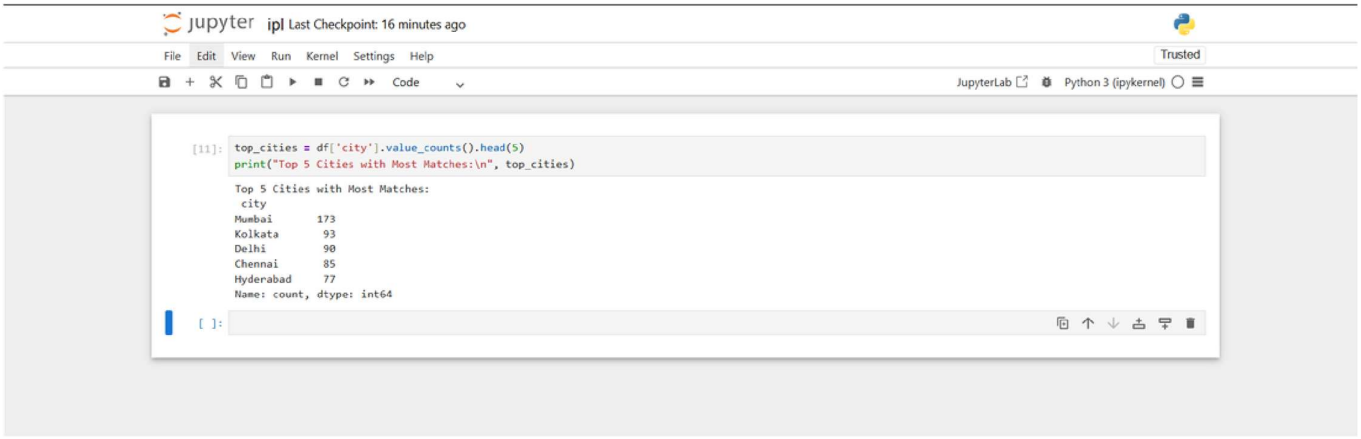
7.Find out which team won the maximum tosses.



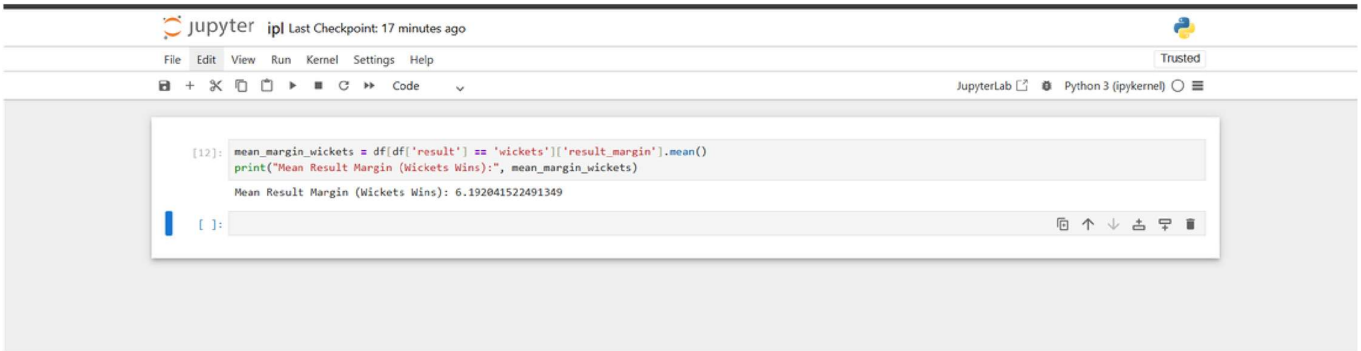
8.Find out the number of matches each umpire officiated.



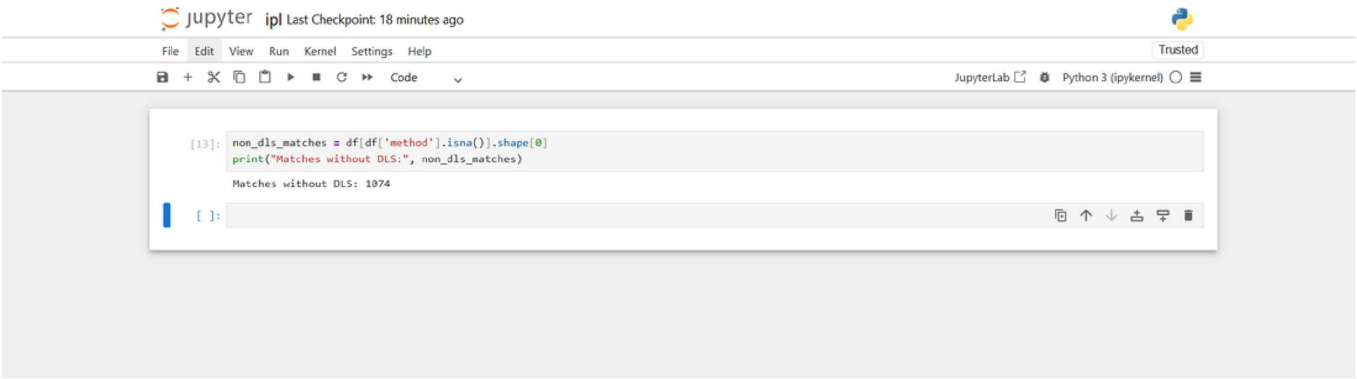
9.List the top 5 cities that hosted the most matches.



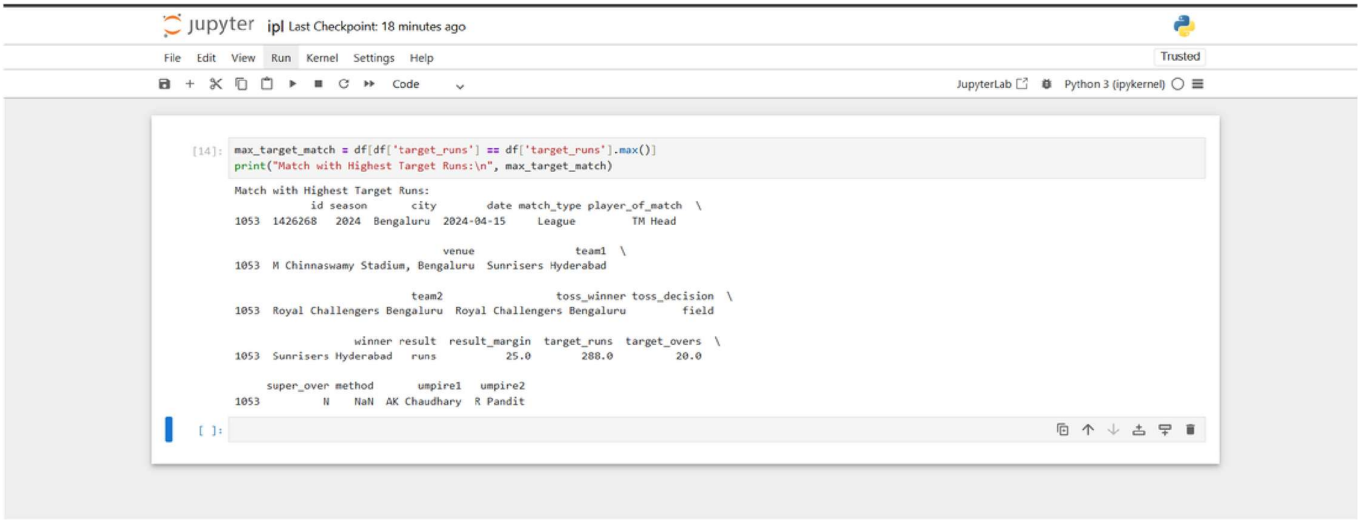
10.Calculate the mean result margin when matches were won by wickets.



11.Find the total number of matches decided without using DLS method.

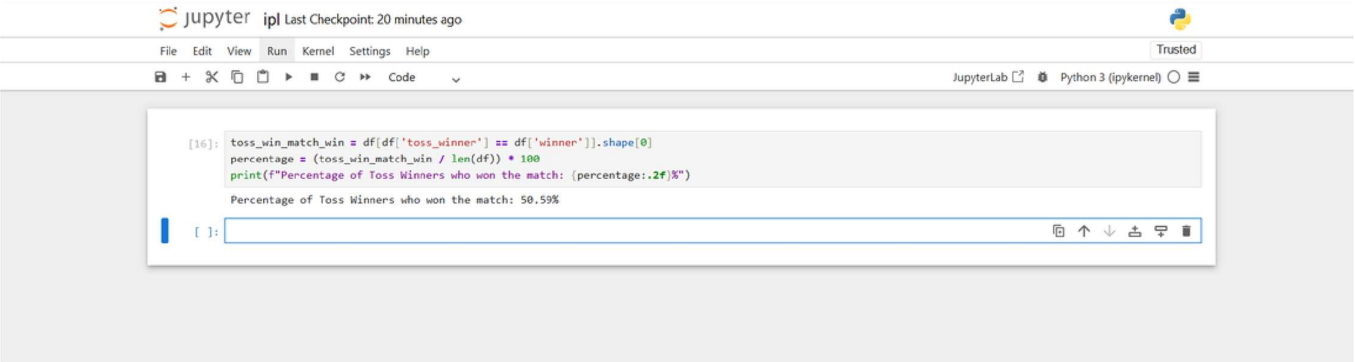


12.Find the match with the highest 'target_runs'.

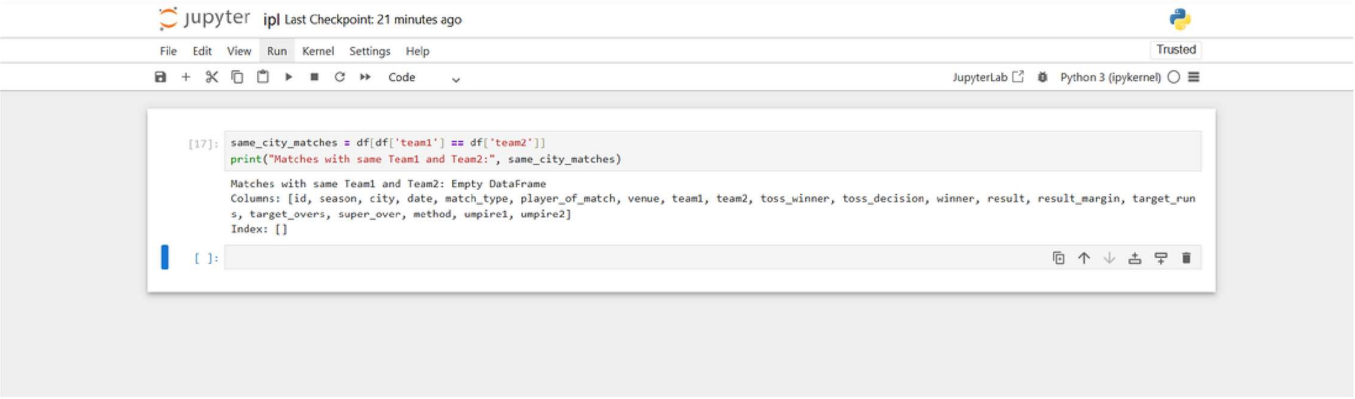


13.Find which season had the most matches played.

14.Find the percentage of matches where toss-winning team won the match.



15.List matches where both teams belonged to the same city (if any)



16.Find top 3 players who won Player of the Match most often in finals only.

17.Group by city and find average 'result_margin' per city.

jupyter ipl Last Checkpoint: 22 minutes ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

```
[19]: avg_margin_city = df.groupby('city')['result_margin'].mean()
print("Average Result Margin per City:\n", avg_margin_city)
```

```
Average Result Margin per City:
 city
Abu Dhabi          17.971429
Ahmedabad          16.914286
Bangalore          23.870968
Bengaluru          12.000000
Bloemfontein       10.000000
Cape Town          27.166667
Centurion          10.833333
Chandigarh         15.213115
Chennai            17.987952
Cuttack            13.285714
Delhi              18.181818
Dharamsala         31.461538
Dubai              12.000000
Durban             12.800000
East London        31.333333
Guwahati           22.333333
Hyderabad          15.592105
Indore              9.888889
Jaipur             16.824561
Johannesburg        7.500000
Kanpur              5.500000
Kimberley          21.000000
Kochi              24.600000
Kolkata            14.946237
Lucknow            20.307692
```

18.Using NumPy, find how many matches had a 'target_overs' of exactly 20.

jupyter ipl Last Checkpoint: 23 minutes ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

```
[20]: import numpy as np
target_20 = np.sum(df['target_overs'] == 20)
print("Matches with 20 overs target:", target_20)
```

```
Matches with 20 overs target: 1062
```

```
[ ]:
```

19.Find the cities which hosted matches in multiple seasons.

jupyter ipl Last Checkpoint: 24 minutes ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

```
[21]: city_season = df.groupby('city')['season'].nunique()
multi_season_cities = city_season[city_season > 1]
print("Cities hosting matches in multiple seasons:\n", multi_season_cities)
```

```
Cities hosting matches in multiple seasons:
 city
Abu Dhabi          3
Ahmedabad          7
Bangalore          9
Bengaluru          5
Chandigarh        12
Chennai           11
Cuttack            3
Delhi             14
Dharamsala         6
Guwahati           2
Hyderabad         12
Indore             3
Jaipur             9
Kanpur             2
Kolkata           14
Lucknow            2
Mumbai            15
Pune               7
Raipur             3
Rajkot             2
Ranchi             3
Visakhapatnam      5
Name: season, dtype: int64
```

```
[ ]:
```

20.List the matches where toss decision was to field but the team lost.

jupyter ipl Last Checkpoint: 24 minutes ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

[22]:
```
field_loss = df[(df['toss_decision'] == 'field') & (df['toss_winner'] != df['winner'])]
print("Matches where toss winner fielded and lost:\n", field_loss)
```

```
Matches where toss winner fielded and lost:
            id   season       city        date   match_type player_of_match  \
0       335982  2007/08  Bangalore  2008-04-18       League      BB McCullum
7       335989  2007/08    Chennai  2008-04-23       League       ML Hayden
9       335991  2007/08  Chandigarh  2008-04-25       League   KC Sangakkara
16      335998  2007/08      Delhi  2008-04-30       League      GD McGrath
21      336034  2007/08  Bangalore  2008-05-03       League         P Kumar
...        ...      ...        ...         ...          ...             ...
1085  1426300     2024  Bengaluru  2024-05-12       League         C Green
1086  1426302     2024      Delhi  2024-05-14       League        I Sharma
1088  1426305     2024     Mumbai  2024-05-17       League        N Pooran
1089  1426306     2024  Bengaluru  2024-05-18       League    F du Plessis
1093  1426311     2024    Chennai  2024-05-24  Qualifier 2   Shahbaz Ahmed

                                      venue                        team1  \
0                     M Chinnaswamy Stadium  Royal Challengers Bangalore
7           MA Chidambaram Stadium, Chepauk          Chennai Super Kings
9     Punjab Cricket Association Stadium, Mohali              Kings XI Punjab
16                         Feroz Shah Kotla             Delhi Daredevils
21                    M Chinnaswamy Stadium  Royal Challengers Bangalore
...                                     ...                          ...
1085          M Chinnaswamy Stadium, Bengaluru  Royal Challengers Bengaluru
1086               Arun Jaitley Stadium, Delhi              Delhi Capitals
1088                   Wankhede Stadium, Mumbai        Lucknow Super Giants
1089          M Chinnaswamy Stadium, Bengaluru  Royal Challengers Bengaluru
1093    MA Chidambaram Stadium, Chepauk, Chennai          Sunrisers Hyderabad

                         team2                toss_winner toss_decision  \
0          Kolkata Knight Riders  Royal Challengers Bangalore         field
```