

EDS Activity No 1

Name-Sanskruti Santosh Warekar

Roll No. -CS3-10

PRN- 202401040075

Dataset- IPL

Dataset - IPL

Untitled0.ipynb

```
[9] import pandas as pd

file_path = "ipl2019.csv"
ipl_2019_df = pd.read_csv(file_path)
print(ipl_2019_df)
```

	Match_ID	Date	Team_A	Team_B	Winner	Margin	Venue
0	1	2019-03-23	MI	RCB	MI	4	Mumbai
1	2	2019-03-25	RCB	CSK	CSK	6	Bangalore
2	3	2019-03-27	CSK	KKR	KKR	5	Chennai
3	4	2019-03-29	KKR	SRH	SRH	7	Kolkata
4	5	2019-03-31	SRH	PBKS	PBKS	3	Hyderabad
5	6	2019-04-02	PBKS	RR	RR	2	Mohali
6	7	2019-04-04	RR	DC	DC	6	Jaipur
7	8	2019-04-06	DC	MI	MI	8	Delhi
8	9	2019-04-08	MI	KKR	MI	10	Mumbai
9	10	2019-04-10	RCB	SRH	SRH	1	Bangalore
10	11	2019-04-12	CSK	PBKS	PBKS	14	Chennai
11	12	2019-04-14	KKR	RR	RR	5	Kolkata
12	13	2019-04-16	SRH	DC	DC	6	Hyderabad
13	14	2019-04-18	PBKS	MI	KKR	9	Mohali
14	15	2019-04-20	RR	RCB	RCB	4	Jaipur
15	16	2019-04-22	DC	CSK	CSK	7	Delhi
16	17	2019-04-24	MI	PBKS	PBKS	3	Mumbai
17	18	2019-04-26	RCB	RR	RR	6	Bangalore
18	19	2019-04-28	CSK	DC	CSK	10	Chennai
19	20	2019-04-30	KKR	MI	MI	8	Kolkata
20	21	2019-05-02	SRH	RCB	RCB	7	Hyderabad
21	22	2019-05-04	PBKS	CSK	PBKS	11	Mohali
22	23	2019-05-06	RR	KKR	KKR	9	Jaipur
23	24	2019-05-08	DC	SRH	SRH	5	Delhi
24	25	2019-05-10	MI	DC	DC	6	Mumbai
25	26	2019-05-12	RCB	MI	MI	3	Bangalore
26	27	2019-05-14	CSK	RCB	RCB	7	Chennai
27	28	2019-05-16	KKR	CSK	CSK	4	Kolkata
28	29	2019-05-18	SRH	KKR	KKR	2	Hyderabad
29	30	2019-05-20	PBKS	RR	RR	5	Mohali
30	31	2019-05-22	RR	PBKS	PBKS	6	Jaipur

0s completed at 11:50 PM

Untitled0.ipynb

```
29 30 2019-05-20 PBKS RR RR 5 Mohali
30 31 2019-05-22 RR PBKS PBKS 6 Jaipur
31 32 2019-05-24 DC SRH SRH 8 Delhi
32 33 2019-05-26 MI SRH SRH 7 Mumbai
33 34 2019-05-28 RCB PBKS PBKS 3 Bangalore
34 35 2019-05-30 CSK RR RR 5 Chennai
35 36 2019-06-01 KKR DC DC 9 Kolkata
36 37 2019-06-03 SRH MI MI 10 Hyderabad
37 38 2019-06-05 PBKS RCB RCB 11 Mohali
38 39 2019-06-07 RR CSK CSK 6 Jaipur
39 40 2019-06-09 DC KKR KKR 4 Delhi
40 41 2019-06-11 MI RR RR 7 Mumbai
41 42 2019-06-13 RCB DC DC 8 Bangalore
42 43 2019-06-15 CSK MI MI 5 Chennai
43 44 2019-06-17 KKR RCB RCB 6 Kolkata
44 45 2019-06-19 SRH CSK CSK 7 Hyderabad
45 46 2019-06-21 PBKS PBKS PBKS 4 Mohali
46 47 2019-06-23 RR SRH SRH 3 Jaipur
47 48 2019-06-25 DC RR RR 2 Delhi
```

0s completed at 11:50 PM

20 Problem statement based on IPL dataset

Problem 1: List all matches played at "Chennai"

```
mumbai_matches = ipl_df[ipl_df['Venue'] == "Chennai"]
print("1. Matches played at Mumbai:\n", mumbai_matches, "\n")
```

1. Matches played at Mumbai:

	Match_ID	Team_A	Team_B	Winner	Margin	Top_Scorer	Man_of_the_Match	Venue
0	1	SRH	CSK	DC	58	Player1	Player4	Chennai
1	2	KKR	CSK	CSK	51	Player1	Player4	Chennai
3	4	CSK	KKR	DC	10	Player2	Player1	Chennai
5	6	RR	SRH	RCB	46	Player5	Player5	Chennai
8	9	DC	RCB	CSK	70	Player1	Player3	Chennai
17	18	RCB	KKR	RR	61	Player4	Player5	Chennai
31	32	RR	CSK	KKR	86	Player3	Player1	Chennai
36	37	MI	MI	KKR	32	Player5	Player4	Chennai
39	40	PBKS	DC	PBKS	8	Player2	Player5	Chennai
40	41	KKR	MI	CSK	13	Player5	Player1	Chennai
43	44	DC	RR	PBKS	88	Player4	Player5	Chennai
45	46	RR	MI	KKR	48	Player4	Player1	Chennai

0s completed at 11:50 PM

```
Commands + Code + Text

[11] # Problem 2: Find the total number of matches won by "MI"
mi_wins = len(ipl_df[ipl_df['Winner'] == "MI"])
print("2. Total matches won by MI:", mi_wins, "\n")

2. Total matches won by MI: 6

[12] # Problem 3: Identify the team with the maximum wins
max_wins_team = ipl_df['Winner'].value_counts().idxmax()
print("3. Team with maximum wins:", max_wins_team, "\n")

3. Team with maximum wins: KKR

# Problem 4: Calculate the average winning margin for all matches
avg_margin = ipl_df['Margin'].mean()
print("4. Average winning margin:", avg_margin, "\n")

4. Average winning margin: 51.92

[30] # Problem 5: Find top 5 matches where the winning margin was greater than 10 runs/wickets
large_margin_matches = ipl_df[ipl_df['Margin'] > 10]
print("5. Matches with winning margin > 10:\n", large_margin_matches.head(), "\n")

5. Matches with winning margin > 10:
  Match_ID Team_A Team_B Winner Margin Top_Scorer Man_of_the_Match \
0         1   SRH   CSK    DC     58   Player1         Player4
1         2   KKR   CSK    CSK     51   Player1         Player4
2         3    DC   MI    DC     38   Player2         Player4
5         6   RR   SRH   RCB     46   Player5         Player5
6         7   CSK   RCB    DC     89   Player1         Player5

  Venue
0  Chennai
1  Chennai
2  Bangalore
5  Chennai

✓ 0s completed at 11:50 PM
```

```
Untitled0.ipynb ☆
File Edit View Insert Runtime Tools Help

Commands + Code + Text

6 Kolkata

# Problem 6: How many matches did "RCB" play?
rcb_matches = len(ipl_df[(ipl_df['Team_A'] == "RCB") | (ipl_df['Team_B'] == "RCB")])
print("6. Total matches played by RCB:", rcb_matches, "\n")

6. Total matches played by RCB: 11

[16] # Problem 7: List all matches where "CSK" played against "MI"
csk_vs_mi_matches = ipl_df[
    ((ipl_df['Team_A'] == "CSK") & (ipl_df['Team_B'] == "MI")) |
    ((ipl_df['Team_A'] == "MI") & (ipl_df['Team_B'] == "CSK"))
]
print("7. Matches where CSK played against MI:\n", csk_vs_mi_matches, "\n")

7. Matches where CSK played against MI:
  Match_ID Team_A Team_B Winner Margin Top_Scorer Man_of_the_Match Venue
7         8    MI   CSK   KKR     67   Player3         Player3  Mumbai

# Problem 8: Find the venue where "KKR" won most matches
kk_r_venues = ipl_df[ipl_df['Winner'] == "KKR"]['Venue'].value_counts()
most_common_venue = kkr_venues.idxmax()
print("8. Venue where KKR won most matches:", most_common_venue, "\n")

8. Venue where KKR won most matches: Chennai

[18] # Problem 9: How many matches ended with a margin of 6 runs/wickets?
six_margin_matches = len(ipl_df[ipl_df['Margin'] == 6])
print("9. Matches with margin = 6:", six_margin_matches, "\n")

9. Matches with margin = 6: 1

✓ 0s completed at 11:50 PM
```

Untitled0.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

0s

[19] # Problem 10: Determine the winning percentage of "SRH"
srh_matches = len(ipl_df[(ipl_df['Team_A'] == "SRH") | (ipl_df['Team_B'] == "SRH")])
srh_wins = len(ipl_df[ipl_df['winner'] == "SRH"])
win_percentage = (srh_wins / srh_matches) * 100 if srh_matches > 0 else 0
print("10. Winning percentage of SRH:", win_percentage, "%\n")

10. Winning percentage of SRH: 21.428571428571427 %

1s

[20] # Problem 11: Matches where the venue was "Delhi" and the margin was < 5
close_delhi_matches = ipl_df[(ipl_df['Venue'] == "Delhi") & (ipl_df['Margin'] < 5)]
print("11. Close matches at Delhi:\n", close_delhi_matches, "\n")

11. Close matches at Delhi:
Empty DataFrame
Columns: [Match_ID, Team_A, Team_B, Winner, Margin, Top_Scorer, Man_of_the_Match, Venue]
Index: []

0s

[21] # Problem 12: Match with the highest winning margin
max_margin_match = ipl_df[ipl_df['Margin'] == ipl_df['Margin'].max()]
print("12. Match with highest margin:\n", max_margin_match, "\n")

12. Match with highest margin:
Match_ID Team_A Team_B Winner Margin Top_Scorer Man_of_the_Match Venue
11 12 CSK PBKS MI 96 Player1 Player2 Delhi

0s

[22] # Problem 13: Count of matches played in each venue
matches_per_venue = ipl_df['Venue'].value_counts()
print("13. Matches per venue:\n", matches_per_venue, "\n")

13. Matches per venue:
Venue
Mumbai 13
Chennai 12

0s completed at 11:50PM

Untitled0.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

Kolkata 10
Delhi 8
Bangalore 7
Name: count, dtype: int64

0s

[23] # Problem 14: Identify the team that played the most matches
team_counts = pd.concat([ipl_df['Team_A'], ipl_df['Team_B']]).value_counts()
most_matches_team = team_counts.idxmax()
print("14. Team that played the most matches:", most_matches_team, "\n")

14. Team that played the most matches: MI

0s

[24] # Problem 15: Matches where "PBKS" lost
pbks_lost_matches = ipl_df[(ipl_df['winner'] != "PBKS") &
((ipl_df['Team_A'] == "PBKS") | (ipl_df['Team_B'] == "PBKS"))]
print("15. Matches where PBKS lost:\n", pbks_lost_matches, "\n")

15. Matches where PBKS lost:
Match_ID Team_A Team_B Winner Margin Top_Scorer Man_of_the_Match Venue
11 12 CSK PBKS MI 96 Player1 Player2 Delhi
16 17 SRH PBKS SRH 54 Player2 Player1 Kolkata
19 20 SRH PBKS MI 81 Player1 Player1 Mumbai
21 22 RR PBKS KKR 18 Player2 Player5 Delhi
34 35 PBKS MI KKR 64 Player3 Player3 Kolkata

0s

[25] # Problem 16: Average number of matches per venue
avg_matches_per_venue = matches_per_venue.mean()
print("16. Average matches per venue:", avg_matches_per_venue, "\n")

16. Average matches per venue: 10.0

0s

[26] # Problem 17: First match where "RR" won
rr_first_win = ipl_df[ipl_df['winner'] == "RR"].iloc[0]
print("17. First match where RR won:\n", rr_first_win, "\n")

17. First match where RR won:

0s completed at 11:50PM

```
17. First match where RR won:
Match_ID      13
Team_A        MI
Team_B        MI
Winner         RR
Margin        40
Top_Scorer     Player1
Man_of_the_Match Player2
Venue         Bangalore
Name: 12, dtype: object
```

```
[27] # Problem 18: Count of matches won by each team
matches_won_by_team = ipl_df['Winner'].value_counts()
print("18. Matches won by each team:\n", matches_won_by_team, "\n")
```

```
18. Matches won by each team:
Winner
KKR    9
RR     8
PBKS   7
CSK    7
MI     6
DC     6
RCB    4
SRH    3
Name: count, dtype: int64
```

```
[28] # Problem 19: Number of matches "KKR" played at "Kolkata"
kkr_kolkata_matches = ipl_df[((ipl_df['Team_A'] == "KKR") | (ipl_df['Team_B'] == "KKR")) &
                             (ipl_df['Venue'] == "Kolkata")]
print("19. Matches KKR played in Kolkata:", len(kkr_kolkata_matches), "\n")
```

```
19. Matches KKR played in Kolkata: 1
```

```
RR      8
PBKS    7
CSK     7
MI      6
DC      6
RCB     4
SRH     3
Name: count, dtype: int64
```

```
[28] # Problem 19: Number of matches "KKR" played at "Kolkata"
kkr_kolkata_matches = ipl_df[((ipl_df['Team_A'] == "KKR") | (ipl_df['Team_B'] == "KKR")) &
                             (ipl_df['Venue'] == "Kolkata")]
print("19. Matches KKR played in Kolkata:", len(kkr_kolkata_matches), "\n")
```

```
19. Matches KKR played in Kolkata: 1
```

```
[29] # Problem 20: Team with the lowest winning margin overall
min_margin_team = ipl_df[ipl_df['Margin'] == ipl_df['Margin'].min()][['Winner']].iloc[0]
print("20. Team with lowest winning margin:", min_margin_team, "\n")
```

```
20. Team with lowest winning margin: MI
```

[] Start coding or generate with AI.