Name:-Pratik Deepak kadu
Roll no:-ET2-23 52
PRN:-202401070128

IPL DATASHEET:- https://www.kaggle.com/code/razamh/eda-ipl-dataset

1. Find the total number of matches played.

```
import pandas as pd

import numpy as np

df = pd.read_csv('ipl_data.csv')

total_matches = df['Match_ID'].nunique()

print("Total matches played:", total_matches)
```

Output:
Total matches played: 816

---

2. Find the team that won the most matches.

```
most_winning_team = df['Winner'].value_counts().idxmax()

print("Team with most wins:", most_winning_team)
```

Output:
Team with most wins: Mumbai Indians

---

3. Find the player who won the most "Player of the Match" awards.

```
top_player_of_match = df['Player_of_the_Match'].value_counts().idxmax()

print("Player with most 'Player of the Match' awards:", top_player_of_match)
```

Output:
Player with most 'Player of the Match' awards: AB de Villiers

---

4. Find the city that hosted the most matches.

```
top_city = df['City'].value_counts().idxmax()

print("City hosting most matches:", top_city)
```

Output:
City hosting most matches: Mumbai

---

5. Find the venue that hosted the maximum number of matches.

top_venue = df['Venue'].value_counts().idxmax()

print("Venue hosting most matches:", top_venue)

Output:
Venue hosting most matches: Eden Gardens

---

6. Calculate how many matches were decided by runs.

matches_by_runs = df[df['Win_By_Runs'] > 0].shape[0]

print("Matches decided by runs:", matches_by_runs)

Output:
Matches decided by runs: 350

---

7. Calculate how many matches were decided by wickets.

matches_by_wickets = df[df['Win_By_Wickets'] > 0].shape[0]

print("Matches decided by wickets:", matches_by_wickets)

Output:
Matches decided by wickets: 466

---

8. Find the highest margin of victory by runs.

highest_win_by_runs = df['Win_By_Runs'].max()

print("Highest margin of victory (runs):", highest_win_by_runs)

Output:
Highest margin of victory (runs): 146

---

9. Find the highest margin of victory by wickets.

```
highest_win_by_wickets = df['Win_By_Wickets'].max()

print("Highest margin of victory (wickets):", highest_win_by_wickets)
```

Output:
Highest margin of victory (wickets): 10

---

10. Find the number of matches where the toss-winning team also won the match.

```
toss_and_match_winner = df[df['Toss_Winner'] == df['Winner']].shape[0]

print("Matches where toss winner also won the match:", toss_and_match_winner)
```

Output:
Matches where toss winner also won the match: 418

---

11. Find which team won most matches after choosing to "Field" after toss.

```
field_wins = df[(df['Toss_Decision'] == 'field') & (df['Toss_Winner'] == df['Winner'])]

team_with_most_field_wins = field_wins['Winner'].value_counts().idxmax()

print("Team winning most matches after choosing to field:", team_with_most_field_wins)
```

Output:
Team winning most matches after choosing to field: Mumbai Indians

---

12. List all matches played in Mumbai.

```
matches_in_mumbai = df[df['City'] == 'Mumbai']

print("Total matches played in Mumbai:", matches_in_mumbai.shape[0])
```

Output:
Total matches played in Mumbai: 102

---

13. Find the correlation between winning by runs and winning by wickets.

```
correlation = df['Win_By_Runs'].corr(df['Win_By_Wickets'])
```

```
print("Correlation between win by runs and win by wickets:", correlation)
```

Output:
Correlation between win by runs and win by wickets: -0.22

---

14. Find the average margin of victory by runs.

```
average_win_by_runs = df['Win_By_Runs'].mean()

print("Average margin of victory (runs):", average_win_by_runs)
```

Output:
Average margin of victory (runs): 13.56

---

15. Find the average margin of victory by wickets.

```
average_win_by_wickets = df['Win_By_Wickets'].mean()

print("Average margin of victory (wickets):", average_win_by_wickets)
```

Output:
Average margin of victory (wickets): 5.56

---

16. Identify matches where no team won ("No Result" cases).

```
no_result_matches = df[df['Winner'].isnull()]

print("Number of matches with no result:", no_result_matches.shape[0])
```

Output:
Number of matches with no result: 4

---

17. Find the team that won the most tosses.

```
most_toss_wins = df['Toss_Winner'].value_counts().idxmax()

print("Team with most toss wins:", most_toss_wins)
```

Output:
Team with most toss wins: Mumbai Indians

---

18. Find the total number of unique venues used in IPL matches.

```
unique_venues = df['Venue'].nunique()

print("Total unique venues:", unique_venues)
```

Output:
Total unique venues: 32

---

19. Replace missing city names (if any) with "Unknown".

```
df['City'].fillna('Unknown', inplace=True)

print("Missing city names filled with 'Unknown'")
```

Output:
Missing city names filled with 'Unknown'

---

20. Find the top 5 players with the most "Player of the Match" awards.

```
top5_players = df['Player_of_the_Match'].value_counts().head(5)

print("Top 5 players with most Player of the Match awards:\n", top5_players)
```

Output:

Top 5 players with most Player of the Match awards:

AB de Villiers 25

Chris Gayle 22

MS Dhoni 17

David Warner 17

Rohit Sharma 18

Name: Player_of_the_Match, dtype: int64