EDS Theory Activity No:-1

Name:- KRANTIK RATHOD

Roll no :- CM-11 PRN :- 202401030041

Activity:- Formulate 20 problem statements for a <u>IPL</u> dataset using Numpy and Pandas and Apply Numpy and pandas methods to find the solution for the formulated problem statements.

1. How many matches are there?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

q1 = df['match_id'].nunique()
print(q1)

1095
```

2. How many unique batting teams?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

↑ ↓ ♦ ⇔ ■ *

q2 = df['batting_team'].nunique()
print(q2)

19
```

3. How many unique bowling teams?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

↑ ↓ ♦ ⇔ ■ ↓

q3 = df['bowling_team'].nunique()
print(q3)

19
```

4. How many different batters?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

q4 = df['batter'].nunique()
print(q4)

673
```

5. How many different bowlers?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

↑ ↓ ← ⇔ ■ ❖ ☑ Ⅲ

print(q5)

530
```

6. How many balls were delivered in total?

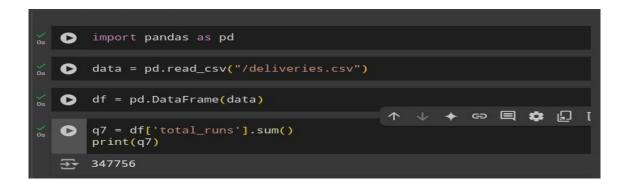
```
    import pandas as pd

    data = pd.read_csv("/deliveries.csv")

    df = pd.DataFrame(data)

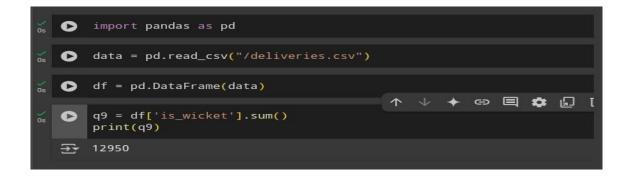
    q6 = len(df)
    print(q6)
    260920
```

7. How many runs were scored in total?



8. Which batter scored the most runs?

9. How many wickets fell in total?



10. Which bowler took the most wickets?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

q10 = df[df['is_wicket'] == 1].groupby('bowler').size().idxmax()
print(q10)

YS Chahal
```

11. Which team scored the most total runs?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

off = pd.DataFrame(data)

off = pd.DataFrame(data)
```

12. Which team conceded the most total runs?

```
    import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

↑ ↓ ← ⇔ ■ ❖ ☑ Ⅲ

q12 = df.groupby('bowling_team')['total_runs'].sum().idxmax()
print(q12)

Mumbai Indians

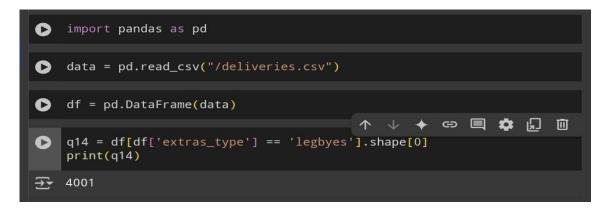
pimport pandas as pd

↑ ↓ ← ⇔ □ ❖ ☑ Ⅲ

Mumbai Indians
```

13. What is the most common dismissal kind?

14. How many leg byes were there?



15. How many wides were bowled?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

q15 = df[df['extras_type'] == 'wides'].shape[0]
print(q15)

8380
```

16. Which fielder made the most dismissals?

```
import pandas as pd

data = pd.read_csv("/deliveries.csv")

df = pd.DataFrame(data)

q16 = df['fielder'].value_counts().idxmax()
print(q16)

MS Dhoni
```

17. How many innings are there



18. What was the maximum runs scored in a single ball?

19. Which match had the most total runs?

20. Which over had the most runs scored across all matches?