NAME: GAURAV DHADGE

ROLL NO: CS6-07

PRN NO: 202401100139

EDS THEORY ACTIVITY 1:

DATA SET LINK: https://www.kaggle.com/datasets/enggbilalalikhan/odi-world-cup-2023-complete-dataset

DATA SET : CRICKET WORLD CUP

Connect the set:

```
[13]  import pandas as pd
      import numpy as np

      # Load datasets
      batting = pd.read_csv('/content/batting_summary.csv')
      bowling = pd.read_csv('/content/bowling_summary.csv')
      matches = pd.read_csv('/content/match_schedule_results.csv')
      players = pd.read_csv('/content/world_cup_players_info.csv')
```

```
#1 Batsman with the highest strike rate
max_sr_index = np.argmax(batting['Strike_Rate'].values)
batting.iloc[max_sr_index][['Batsman_Name', 'Strike_Rate']]
```

	379
Batsman_Name	Imam-ul-Haq
Strike_Rate	98.592

dtype: object

```
[10]  #2 Players who hit more than 10 sixes
      batting[batting['6s'] > 10][['Batsman_Name', '6s']]
```

	Batsman_Name	6s
693	Fakhar Zaman	11

```
[14]  #3 Filter players from India
      india_players = players[players['team_name'].str.lower() == 'india']
      print(india_players[['player_name', 'description']])
```

```
          player_name                                 description
28       Rohit Sharma   Rohit Sharma is an Indian cricketer known for ...
29       Ishan Kishan   Ishan Kishan is an Indian cricketer known for ...
30       Shreyas Iyer   Shreyas Iyer is an Indian cricketer known for ...
47       Shubman Gill   Shubman Gill is an Indian cricketer known for ...
54     Ravindra Jadeja  Ravindra Jadeja is an Indian cricketer known f...
58     Jasprit Bumrah   Jasprit Bumrah is an Indian cricketer known fo...
59      Kuldeep Yadav   Kuldeep Yadav is an Indian cricketer known for...
65     Mohammed Siraj   Mohammed Siraj is an Indian cricketer known fo...
66      Shardul Thakur  Shardul Thakur is an Indian cricketer known fo...
123       Virat Kohli   India has given to the world many a great cric...
124         KL Rahul    A tall, elegant right-hand batsman who can kee...
125      Hardik Pandya  Hardik Pandya swears by living life king size ...
143    Suryakumar Yadav  Hard-hitting 360-degree batter Suryakumar Yada...
144     Mohammed Shami  Mohammed Shami was India's leading fast bowler...
149  Ravichandran Ashwin  R Ashwin took the tricks and skills he learned...
```

```
#4 Right-handed batsmen
right_hand_batsmen = players[players['battingStyle'].str.contains('right hand Bat', case=False, na=False)]
num_right_handers = right_hand_batsmen.shape[0]
print(num_right_handers)
```
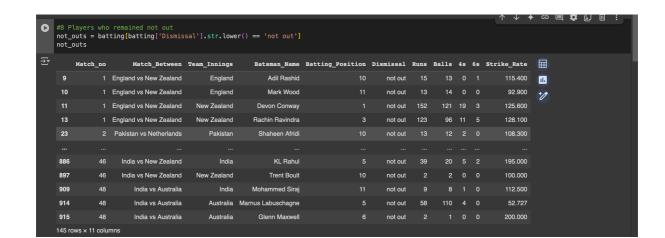
```
101
```

```
#5 Batsman who faced the most balls
balls_faced_by_batsman = batting.groupby('Batsman_Name')['Balls'].sum()
top_batsman = balls_faced_by_batsman.idxmax()
max_balls = balls_faced_by_batsman.max()
print(max_balls)
```

```
847
```

```
#6 Left-handed batsmen
left_hand_batsmen = players[players['battingStyle'].str.contains('Left hand Bat', case=False, na=False)]
num_left_handers = left_hand_batsmen.shape[0]
print(num_left_handers)
```

```
49
```

```
#7 Matches in Pune
matches_in_pune = matches[matches['Venue'].str.contains('Pune', case=False, na=False)]
matches_in_pune
```

	Match_no	Date	Venue	Team1	Team2	Winner	Scorecard URL
16	17	October 19	Pune	India	Bangladesh	India	https://www.cricketwa.com/scorecard/23022/indi...
29	30	October 30	Pune	Afghanistan	Sri Lanka	Afghanistan	https://www.cricketwa.com/scorecard/23035/afgh...
31	32	November 1	Pune	New Zealand	South Africa	South Africa	https://www.cricketwa.com/scorecard/23037/new-...
39	40	November 8	Pune	England	Netherlands	England	https://www.cricketwa.com/scorecard/23045/engl...
42	43	November 11	Pune	Australia	Bangladesh	Australia	https://www.cricketwa.com/scorecard/23049/aust...

```
#8 Players who remained not out
not_outs = batting[batting['Dismissal'].str.lower() == 'not out']
not_outs
```

	Match_no	Match_Between	Team_Innings	Batsman_Name	Batting_Position	Dismissal	Runs	Balls	4s	6s	Strike_Rate
9	1	England vs New Zealand	England	Adil Rashid	10	not out	15	13	0	1	115.400
10	1	England vs New Zealand	England	Mark Wood	11	not out	13	14	0	0	92.900
11	1	England vs New Zealand	New Zealand	Devon Conway	1	not out	152	121	19	3	125.600
13	1	England vs New Zealand	New Zealand	Rachin Ravindra	3	not out	123	96	11	5	128.100
23	2	Pakistan vs Netherlands	Pakistan	Shaheen Afridi	10	not out	13	12	2	0	108.300
...	...	...	...	...	...	...	...	...	...	...	...
886	46	India vs New Zealand	India	KL Rahul	5	not out	39	20	5	2	195.000
897	46	India vs New Zealand	New Zealand	Trent Boult	10	not out	2	2	0	0	100.000
909	48	India vs Australia	India	Mohammed Siraj	11	not out	9	8	1	0	112.500
914	48	India vs Australia	Australia	Marnus Labuschagne	5	not out	58	110	4	0	52.727
915	48	India vs Australia	Australia	Glenn Maxwell	6	not out	2	1	0	0	200.000

145 rows × 11 columns

```python
[26] #9 Most played venue
     venue_matches = matches['Venue'].value_counts().head(1)
     venue_matches
```

```
              count
     Venue
     Ahmedabad     5

     dtype: int64
```

```python
#10 Bowler who bowled most overs
bowling.groupby('Bowler_Name')['Overs'].sum().sort_values(ascending=False).head(5)
```

```
                   Overs
     Bowler_Name
     Adam Zampa       96.0
     Kuldeep Yadav    95.1
     Josh Hazlewood   93.1
     Ravindra Jadeja  92.9
     Mitchell Santner 92.4

     dtype: float64
```

```python
[28] #11 Player with most boundaries
     batting['Boundaries'] = batting['4s'] + batting['6s']
     batting[['Batsman_Name', 'Boundaries']].sort_values('Boundaries', ascending=False).head(1)
```

```
         Batsman_Name  Boundaries
     767  Glenn Maxwell         31
```

```python
[48] #12 Filter rows where players scored more than 30 runs
     batting_above_30 = batting[batting['Runs'] > 30]  # Fix: Corrected spelling of 'batting'

     # Count occurrences of each batting position
     most_players_30 = batting_above_30['Batting_Position'].value_counts()  # Fix: Corrected column name typo

     # Display results
     print(most_players_30)
```

```
     Batting_Position
     4     52
     3     48
     1     47
     2     37
     5     30
     6     26
     7     20
     8      7
     9      4
     10     1
     11     1
     Name: count, dtype: int64
```

```python
[31] #13 Average runs
     average_runs = np.mean(batting['Runs'].values)
     average_runs
```

```
     np.float64(26.510917030567686)
```

```python
#14 Top 5 bowlers with most wickets
bowling[['Bowler_Name', 'Wickets']].groupby('Bowler_Name').sum().sort_values('Wickets', ascending=False).head(5)
```

```
                      Wickets
     Bowler_Name
     Mohammed Shami        24
     Adam Zampa            23
     Dilshan Madushanka    21
     Gerald Coetzee        20
     Jasprit Bumrah        20
```

```
[36]  #15 Total runs
      total_runs = np.sum(batting['Runs'].values)
      total_runs
```

np.int64(24284)

```
[37]  #16 Team with most wins
      matches['Winner'].value_counts().head(2)
```

	count
Winner	
Australia	9
India	9

dtype: int64

```
#17 Total matches
total_matches = matches.shape[0]
total_matches
```

48

```
#19 Total runs by player
total_runs_team = batting.groupby('Batsman_Name')['Runs'].sum().sort_values(ascending=False)
total_runs_team
```

	Runs
Batsman_Name	
Virat Kohli	765
Quinton de Kock	706
Rohit Sharma	597
Rachin Ravindra	578
David Warner	577
...	...
Dushan Hemantha	4
Lahiru Kumara	4
Fazalhaq Farooqi	2
Alex Carey	0
Usama Mir	0

146 rows × 1 columns

dtype: int64

```
#20 Top 5 batsmen who scored the most runs
batting[['Batsman_Name', 'Runs']].groupby('Batsman_Name').sum.sort_values('Runs', ascending=False).head(5)
```