

EXPERIMENT NO: 3(A)

Pandas library -Create own dataset in CSV format

Aim:

To write a Python program to create a dataset using the Pandas library.

Algorithm:

1. Define a dictionary with the relevant fields.
2. Convert the dictionary into a Pandas DataFrame.
3. Display the DataFrame to verify the structure and contents.
4. Perform basic data inspection using functions like .head(), .info(), and .describe().

Program:

```
[1]: import pandas as pd

[2]: data = {
    'Name': ['Virat Kohli', 'Rohit Sharma', 'Ben Stokes', 'Babar Azam', 'Pat Cummins',
            'Steve Smith', 'Kane Williamson', 'Jasprit Bumrah', 'Shakib Al Hasan', 'David Warner'],
    'Age': [35, 37, 34, 30, 32, 35, 34, 31, 36, 33],
    'Country': ['India', 'India', 'England', 'Pakistan', 'Australia',
               'Australia', 'New Zealand', 'India', 'Bangladesh', 'Australia'],
    'Matches_Played': [275, 250, 150, 130, 110, 160, 145, 100, 220, 190],
    'Runs': [12800, 10700, 5500, 5400, 950, 8700, 6500, 400, 7000, 8200],
    'Wickets': [4, 8, 150, 15, 210, 25, 20, 180, 280, 5],
    'Batting_Avg': [58.2, 49.8, 37.5, 56.3, 11.4, 44.1, 47.8, 8.6, 35.6, 45.2],
    'Bowling_Avg': [120.5, 98.2, 29.3, 60.1, 23.5, 42.4, 50.8, 25.1, 31.4, 80.7],
    'Role': ['Batsman', 'Batsman', 'All-Rounder', 'Batsman', 'Bowler',
            'Batsman', 'Batsman', 'All-Rounder', 'Batsman', 'Bowler']
}

[3]: df = pd.DataFrame(data)
print(df)

   Name  Age  Country  Matches_Played  Runs  Wickets  \
0  Virat Kohli   35    India          275  12800         4
1  Rohit Sharma   37    India          250  10700         8
2   Ben Stokes   34  England          150   5500       150
3   Babar Azam   30  Pakistan          130   5400        15
4   Pat Cummins   32  Australia          110    950       210
5   Steve Smith   35  Australia          160   8700        25
6 Kane Williamson   34 New Zealand          145   6500        20
7  Jasprit Bumrah   31    India          100    400       180
8  Shakib Al Hasan   36  Bangladesh          220   7000       280
9   David Warner   33    Australia          190   8200         5

   Batting_Avg  Bowling_Avg  Role
0      58.2      120.5    Batsman
1      49.8      98.2    Batsman
2      37.5      29.3  All-Rounder
3      56.3      60.1    Batsman
4      11.4      23.5    Bowler
5      44.1      42.4    Batsman
6      47.8      50.8    Batsman
7       8.6      25.1    Bowler
8      35.6      31.4  All-Rounder
9      45.2      80.7    Batsman

[4]: df.to_csv('cricketer_stats.csv', index=False)
```

Result:

Thus, the Python program is executed successfully for creating the dataset using Pandas.