Name-Omkar Arun Medankar

Roll no-675

Batch-F4

Prn-202201030042

EDS ASSIGNMENT

Problem statement-Develop a cricket match outcome prediction model that accurately predicts the result of a cricket match based on team performance metrics, player statistics, and match conditions. The objective is to assist cricket analysts and fans in making accurate predictions and strategic decisions, leading to improved team performance and match outcomes

Data Set-

Year	Format	Host	Final Venue	Winner	Margin	Runner up
1984	ODI	UAE	Sharjah	India	54 runs	Pakistan
1986	ODI	Sri Lanka	Colombo (SSC)	Sri Lanka	5 wickets	Pakistan

1988	ODI	Bangladesh	Dhaka	India	6 wickets	Sri Lanka
1991	ODI	India	Kolkata	India	7 wickets	Sri Lanka
1995	T20I	UAE	Sharjah	India	8 wickets	Sri Lanka
1997	ODI	Sri Lanka	Colombo (RPS)	Sri Lanka	8 wickets	India
2000	ODI	Bangladesh	Dhaka	Pakistan	39 runs	Sri Lanka
2004	ODI	Sri Lanka	Colombo (RPS)	Sri Lanka	25 runs	India
2008	T20I	Pakistan	Karachi	Sri Lanka	100 runs	India
2010	ODI	Sri Lanka	Dambulla	India	81 runs	Sri Lanka
2012	ODI	Bangladesh	Mirpur	Pakistan	2 runs	Bangladesh
2014	ODI	Bangladesh	Mirpur	Sri Lanka	5 wickets	Pakistan
2016	T20I	Bangladesh	Mirpur	India	8 wickets	Bangladesh

Year	Format	Host	Final Venue	Winner	Margin	Runner up
1984	ODI	UAE	Sharjah	India	54 runs	Pakistan
1986	ODI	Sri Lanka	Colombo (SSC	Sri Lanka	5 wickets	Pakistan
1988	ODI	Bangladesh	Dhaka	India	6 wickets	Sri Lanka
1991	ODI	India	Kolkata	India	7 wickets	Sri Lanka
1995	T20I	UAE	Sharjah	India	8 wickets	Sri Lanka
1997	ODI	Sri Lanka	Colombo (RPS	Sri Lanka	8 wickets	India
2000	ODI	Bangladesh	Dhaka	Pakistan	39 runs	Sri Lanka
2004	ODI	Sri Lanka	Colombo (RPS	Sri Lanka	25 runs	India
2008	T20I	Pakistan	Karachi	Sri Lanka	100 runs	India
2010	ODI	Sri Lanka	Dambulla	India	81 runs	Sri Lanka
2012	ODI	Bangladesh	Mirpur	Pakistan	2 runs	Bangladesh
2014	ODI	Bangladesh	Mirpur	Sri Lanka	5 wickets	Pakistan
2016	T20I	Bangladesh	Mirpur	India	8 wickets	Bangladesh

1.Find the number of matches where the winning team had a margin of victory greater than 50 runs.

answer-INPUT

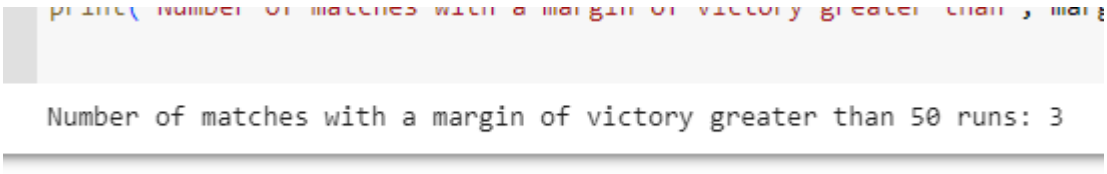
```
import csv

def count_matches_with_margin(datacsv, margin_threshold):
    count = 0

    with open(datacsv, 'r') as file:
        reader = csv.DictReader(file)
        if 'Margin' not in reader.fieldnames:
            print("Error: 'Margin' column not found in the CSV file.")
            return count

        for row in reader:
            margin_value = row['Margin']
            if margin_value.strip() == '':
                continue    # Skip rows with empty margin values

            margin = int(margin_value.split()[0])
            if margin > margin_threshold:
                count += 1

    return count

datacsv = '/content/drive/MyDrive/Colab Notebooks/dataset1.csv'
```

margin_threshold = 50

num_matches = count_matches_with_margin(datacsv, margin_threshold)

print("Number of matches with a margin of victory greater than", margin_threshold, "runs:", num_matches)

OUTPUT-Number of matches with a margin of victory greater than 50 runs: 3



2.Find the matches where the winning team won by a wicket margin of less than 10.

Answer-Input

```
import csv


def find_matches_with_wicket_margin(datacsv, margin_threshold):
    matches = []

    with open(datacsv, 'r') as file:
        reader = csv.DictReader(file)
        if 'Margin' not in reader.fieldnames:
            print("Error: 'Margin' column not found in the CSV file.")
            return matches

        for row in reader:
```

```
            margin_value = row['Margin']
            if 'wickets' in margin_value:
                margin = int(margin_value.split()[0])
                if margin < margin_threshold:
                    matches.append(row)

    return matches

datacsv = '/content/drive/MyDrive/Colab Notebooks/dataset1.csv'
margin_threshold = 10

matches = find_matches_with_wicket_margin(datacsv, margin_threshold)

print("Matches where the winning team won by a wicket margin of less than", margin_threshold, ":")
for match in matches:
    print(match)
```

OUTPUT-Matches where the winning team won by a wicket margin of less than 10 :

{'Year': '1986', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (SSC)', 'Winner': 'Sri Lanka', 'Margin': '5 wickets', 'Runner up': 'Pakistan'}

{'Year': '1988', 'Format': 'ODI', 'Host': 'Bangladesh', 'Final Venue': 'Dhaka', 'Winner': 'India', 'Margin': '6 wickets', 'Runner up': 'Sri Lanka'}

{'Year': '1991', 'Format': 'ODI', 'Host': 'India', 'Final Venue': 'Kolkata', 'Winner': 'India', 'Margin': '7 wickets', 'Runner up': 'Sri Lanka'}

{'Year': '1995', 'Format': 'T20I', 'Host': 'UAE', 'Final Venue': 'Sharjah', 'Winner':

'India', 'Margin': '8 wickets', 'Runner up': 'Sri Lanka'}

{'Year': '1997', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '8 wickets', 'Runner up': 'India'}

{'Year': '2014', 'Format': 'ODI', 'Host': 'Bangladesh', 'Final Venue': 'Mirpur', 'Winner': 'Sri Lanka', 'Margin': '5 wickets', 'Runner up': 'Pakistan'}

{'Year': '2016', 'Format': 'T20I', 'Host': 'Bangladesh', 'Final Venue': 'Mirpur', 'Winner': 'India', 'Margin': '8 wickets', 'Runner up': 'Bangladesh'}

```
Matches where the winning team won by a wicket margin of less than 10 :
{'Year': '1986', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (SSC)', 'Winner': 'Sri Lanka', 'Margin': '5 wickets', 'Runner up': 'Pakistan'}
{'Year': '1988', 'Format': 'ODI', 'Host': 'Bangladesh', 'Final Venue': 'Dhaka', 'Winner': 'India', 'Margin': '6 wickets', 'Runner up': 'Sri Lanka'}
{'Year': '1991', 'Format': 'ODI', 'Host': 'India', 'Final Venue': 'Kolkata', 'Winner': 'India', 'Margin': '7 wickets', 'Runner up': 'Sri Lanka'}
{'Year': '1995', 'Format': 'T20I', 'Host': 'UAE', 'Final Venue': 'Sharjah', 'Winner': 'India', 'Margin': '8 wickets', 'Runner up': 'Sri Lanka'}
{'Year': '1997', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '8 wickets', 'Runner up': 'India'}
{'Year': '2014', 'Format': 'ODI', 'Host': 'Bangladesh', 'Final Venue': 'Mirpur', 'Winner': 'Sri Lanka', 'Margin': '5 wickets', 'Runner up': 'Pakistan'}
{'Year': '2016', 'Format': 'T20I', 'Host': 'Bangladesh', 'Final Venue': 'Mirpur', 'Winner': 'India', 'Margin': '8 wickets', 'Runner up': 'Bangladesh'}
```

3.Find the matches hosted in Sri Lanka where India was the runner-up.

Answer-INPUT-

```
import csv


def find_matches_in_sri_lanka_with_india_runner_up(datacsv, host_country, runner_up):
    matches = []


    with open(datacsv, 'r') as file:
        reader = csv.DictReader(file)
        if 'Host' not in reader.fieldnames or 'Runner up' not in reader.fieldnames:
            print("Error: 'Host' or 'Runner up' column not found in the CSV
```

```
file.")
                return matches

        for row in reader:
                if row['Host'] == host_country and row['Runner up'] ==
runner_up:
                        matches.append(row)

    return matches

datacsv = '/content/drive/MyDrive/Colab Notebooks/dataset1.csv'
host_country = 'Sri Lanka'
runner_up = 'India'

matches = find_matches_in_sri_lanka_with_india_runner_up(datacsv,
host_country, runner_up)

print("Matches hosted in", host_country, "where India was the runner-up:")
for match in matches:
    print(match)
```

OUTPUT-Matches hosted in Sri Lanka where India was the runner-up:

{'Year': '1997', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '8 wickets', 'Runner up': 'India'}

{'Year': '2004', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '25 runs', 'Runner up': 'India'}

```
Matches hosted in Sri Lanka where India was the runner-up:
{'Year': '1997', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '8 wickets', 'Runner up': 'India'}
{'Year': '2004', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Colombo (RPS)', 'Winner': 'Sri Lanka', 'Margin': '25 runs', 'Runner up': 'India'}
```

4.Find the matches that were played in ODI format and had a winning margin of more than 50 runs.

ANSWER-INPUT-

```
import csv


def find_matches_with_odi_format_and_margin(datacsv, format_type, margin_threshold):
    matches = []


    with open(datacsv, 'r') as file:
        reader = csv.DictReader(file)
        if 'Format' not in reader.fieldnames or 'Margin' not in reader.fieldnames:
            print("Error: 'Format' or 'Margin' column not found in the CSV file.")
            return matches


        for row in reader:
            if row['Format'] == format_type:
                margin_value = row['Margin']
                if 'runs' in margin_value:
```

```
                margin = int(margin_value.split()[0])
                if margin > margin_threshold:
                    matches.append(row)

    return matches

datacsv = '/content/drive/MyDrive/Colab Notebooks/dataset1.csv'
format_type = 'ODI'
margin_threshold = 50

matches = find_matches_with_odi_format_and_margin(datacsv, format_type, margin_threshold)

print("Matches played in", format_type, "format with a winning margin of more than", margin_threshold, "runs:")
for match in matches:
    print(match)
```

OUTPUT-

```
Matches played in ODI format with a winning margin of more than 50 runs:
{'Year': '1984', 'Format': 'ODI', 'Host': 'UAE', 'Final Venue': 'Sharjah', 'Winner': 'India', 'Margin': '54 runs', 'Runner up': 'Pakistan'}
{'Year': '2010', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Dambulla', 'Winner': 'India', 'Margin': '81 runs', 'Runner up': 'Sri Lanka'}
```

```
Matches played in ODI format with a winning margin of more than 50 runs:
{'Year': '1984', 'Format': 'ODI', 'Host': 'UAE', 'Final Venue': 'Sharjah', 'Winner': 'India', 'Margin': '54 runs', 'Runner up': 'Pakistan'}
{'Year': '2010', 'Format': 'ODI', 'Host': 'Sri Lanka', 'Final Venue': 'Dambulla', 'Winner': 'India', 'Margin': '81 runs', 'Runner up': 'Sri Lanka'}
```

5.Find the matches where the winning team was Pakistanand the runner-up was Sri Lanka.

ANWER-INPUT

```
import csv


def find_matches_with_teams(datacsv, winner_team, runner_up_team):
    matches = []

    with open(datacsv, 'r') as file:
        reader = csv.DictReader(file)
        if 'Winner' not in reader.fieldnames or 'Runner up' not in reader.fieldnames:
            print("Error: 'Winner' or 'Runner up' column not found in the CSV file.")
            return matches

        for row in reader:
            if row['Winner'] == winner_team and row['Runner up'] == runner_up_team:
                matches.append(row)

    return matches

datacsv = '/content/drive/MyDrive/Colab Notebooks/dataset1.csv'
```

```
winner_team = 'Pakistan'

runner_up_team = 'Sri Lanka'


matches = find_matches_with_teams(datacsv, winner_team, runner_up_team)


print("Matches where the winning team was", winner_team, "and the runner-up was", runner_up_team, ":")

for match in matches:

    print(match)
```

OUTPUT-Matches where the winning team was Pakistan and the runner-up was Sri Lanka :

{'Year': '2000', 'Format': 'ODI', 'Host': 'Bangladesh', 'Final Venue': 'Dhaka', 'Winner': 'Pakistan', 'Margin': '39 runs', 'Runner up': 'Sri Lanka'}