

1. Do not change the name of the data file that is shared with the problem statement.
2. If intermediate data files are created, retain in the present working directory and attach them during submission.
3. Retain the data file in the same directory as that of this workbook.
4. Retain the Visualizations that is produced in the file. Don't clear them away.
5. Submit only the .IPYNB file. Intermediate files to be attached as mentioned in (2).
6. All the visuals should adhere to the visualization principles learnt in the Course and must be presentation ready. Most effective visuals would fetch maximum credits
7. Submissions done via means other than CANVAS will strictly be NOT graded.

Objective

The objective is to analyze cricket teams and players performance accross the seasons to identify the outstanding performers.

Both input data files should be considered

Group No: (mention your group number here)

Full Name	BITS ID
Name 1	BITS ID 1
Name 2	BITS ID 2
Name 3	BITS ID 3

▼ Data Preparation : 1 Mark

Import the libraries needed

```
[ ]
```

```
[ ]
```

Load data and store in dataframe

```
[ ]
```

```
[ ] # Display First ten items from both the input files
```

```
[ ] # Check for NULL values and Replace Nan with 0 (For both files)
```

```
[ ] #Display an overview - The participating teams and the unique venues used for matches, and the number of matches played
```

```
[ ]
```

▼ Visualisation Questions - 2 X 5 = 10 Marks

▼ Question 1

Identify the top 10 batsmen of all time by considering how does these players score runs.

Write the python code in the below cell to create appropriate visual to perform the above task.

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 4.Mention the gestalt principles used.(atleast 2)

```
[ ]
```

▼ Question 2

The general notion in cricket is, Win the toss to win the match . To verify the general notion, perform a match winning analysis based on the toss winners.

Write the python code in the below cell to create the appropriate visual to perform the above task .

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 4.Mention the gestalt principles used.(atleast 2)

```
[ ]
```

▼ Question 3

Are T20 formats batsmen friendly? To identify this, Do an analysis to find out the average runs scored per match across the seasons.

Write the python code in the below cell to create the appropriate visual to perform the above task .

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 4.Mention the gestalt principles used.(atleast 2)

```
[ ]
```

▼ Question 4

Create an interactive visual to identify the most player-of-the-matches award.

Write the python code in the below cell to create the appropriate visual to perform the above task .

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 2) 4.Mention the gestalt principles used.(atleast 2

```
[ ]
```

▼ Question 5

Prepare a tabular visual to identify teams scoring more than 225 runs.

Write the python code in the below cell to create the appropriate visual to perform the above task .

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 4.Mention the gestalt principles used.(atleast 2)

```
[ ]
```

▼ Group's choice-2 Marks

Frame 1 (more) question which will help in the EDA(Exploratory Data Analysis) of the given data set and answer the same using the best visual.

1. Write the question in a markdown cell
2. Below the question,in a coding cell,write the python code to create the visual to answer the question

Answer in markdown cells below the visual

- 1.Summarise your findings from the visual.
- 2.The reason for selecting the chart type you did
- 3.Mention the pre-attentive attributes used.(atleast 2)
- 4.Mention the gestalt principles used.(atleast 2)

```
[ ]
```

***** END OF ASSIGNMENT *****