DATA SCIENCE PROJECT REPORT

(Project Semester August-December 2021)

T20 WORLDCUP 2021 DATA

Submitted by

VANGALA SANDEEP

Registration No: 11911818

Computer Science and Engineering

Section: KM004

Course Code INT217

Under the Guidance of

Sandeep Kaur

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara
<u>CERTIFICATE</u>

This is to certify that Sandeep bearing Registration no. 11911818 has completed Data Science project titled, **"T20 WORLDCUP 2021"** under

my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor

Designation of the Supervisor

School of Computer Science

Lovely Professional University Phagwara,
Punjab.

DECLARATION

I, Vangala Sandeep, student of Lovely Professional University under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this assignment is based onmy own intensive work and is genuine.

Name of the student: Sandeep Vangala

Registration No. 11911818

<u>ACKNOWLEDGEMENT</u>

A project work is a combination of views, ideas, suggestions and contribution of many people. Thus, one of the pleasant parts of writing the report is to thank those who have contributed towards its fulfilment.

I consider it as great privilege to have esteemed Lecturer Ms. Sandeep Kaur as my project guide. I take this opportunity to express my sincere gratitude to her through constant advice and constructive criticism nourished my interest in the subject and provided a free and pleasant atmosphere to work against all odd situations. I avail this opportunity to extend my heart full thanks and deep respect to faculty member for their able guidance during this project.

My gratitude to all those, who **responded to my questionnaire** in a well-defined manner and helped me acquiring knowledge.

I would like to communicate a deep sense of gratitude to all these people without whom my project would not have been such a great learning experience.

Vangala Sandeep
KM004

Reg no: 11911818

Lovely Professional University

Table of Contents

Introduction

The 2021 ICC Men's T20 World Cup was the seventh ICC Men's T20 World Cup tournament, with the matches played in the United Arab Emirates and Oman from 17 October to 14 November 2021. The West Indies were the defending champions, but were eventually eliminated in the Super 12 stage.

Originally the event was scheduled to be held in Australia from 18 October to 15 November 2020, but in July 2020, the International Cricket Council (ICC) confirmed that this tournament had been postponed, due to the COVID-19 pandemic. In August 2020, the ICC confirmed that India would host the 2021 tournament as planned, with Australia being named as the host for the succeeding 2022 tournament. However, in June 2021, the ICC announced that the tournament had been moved to the United Arab Emirates due to growing concerns over the COVID-19 pandemic situation in India, and a possible third wave of the pandemic in the country. The tournament began on 17 October 2021, with the tournament's final taking place on 14 November 2021. The preliminary rounds of the tournament were played in the UAE and Oman.

New Zealand became the first team to reach the final, after they beat England by five wickets in the semi-finals. It was the first time that New Zealand had qualified for a T20 World Cup final. They were joined in the final by Australia, who beat Pakistan by five wickets in the second semi-final. It was the second time that Australia had qualified for the final of the T20 World Cup, after reaching the final of the 2010 tournament. In the final, Australia beat New Zealand by eight wickets to win their first T20 World Cup. Mitchell Marsh was named the Player of the Match with David Warner being named the Player of the Tournament.

Scope of The Analysis

Data Science / Analytics is all about finding valuable insights from the given dataset. In short, Finding answers that could help business.

Since usually such tutorials are based on in-built datasets like iris, It becomes harder for the learner to connect with the analysis and hence learning becomes difficult. To overcome this, The dataset that we use in this notebook is T20 WORLDCUP Dataset posted on Kaggle Datasets sourced from cricsheet. T20 Worldcup is one of the most popular cricket tournaments in the world, thus the problems we try to solve and the questions that we try to answer should be familiar to anyone who knows Cricket.

Questions:

1. What is percentage of the teams toss decision?

2. What is the count of Matches win by each team?

3. What is the Average target set by each team?

4. How many Matches are played in Evening and Afternoon?

5. Who are the best blowers in each team and their most individual wickets?

6. What is the Average temperature played by each team?

7. What is the Average economy of best blowers in each team?

8. What is total count of player of the matches of each team?

9. What is the Percentage of batting style?

10. What is the sum of wickets of different bowling styles?

<u>Existing System</u>

Before existence of Data Science, analyzing data used to be hectic task and existing system didn't used to analyses the data with perfection.

Without existence of current cutting-edge technology of data science, we can get actionable insights in the dataset of the Indian Trade. Following are the benefits which weren't present in the existing system of data analyzing:

1. Making Better Decision with The Help of Data
2. Directing actions based on trends- which later defines the goals required for profit.
3. Doing challenging stuffs with the help of prediction which is done by data.
4. Identifying various opportunities to increase the profit,
5. Making decision with Quantifiable, data driven evidence so that loss doesn't happens.
6. Testing the decisions taken by the data and watching and analyzing the trend.
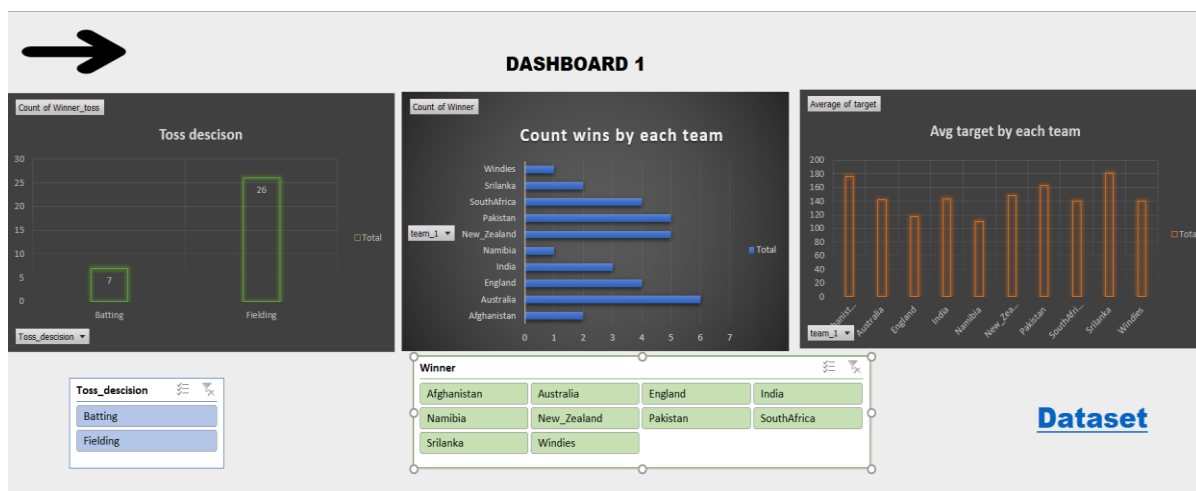
<u>Source of The Dataset</u>

- The dataset is taken from the Kaggle with the name 'T20 WORLDCUP 2021'.

- https:// https://www.kaggle.com/rajsengo/icc-t20-world-cup-2021?select=details.csv

- Author of the Dataset Rajkumar Sengottuvel

- Data last updated
  August 2021
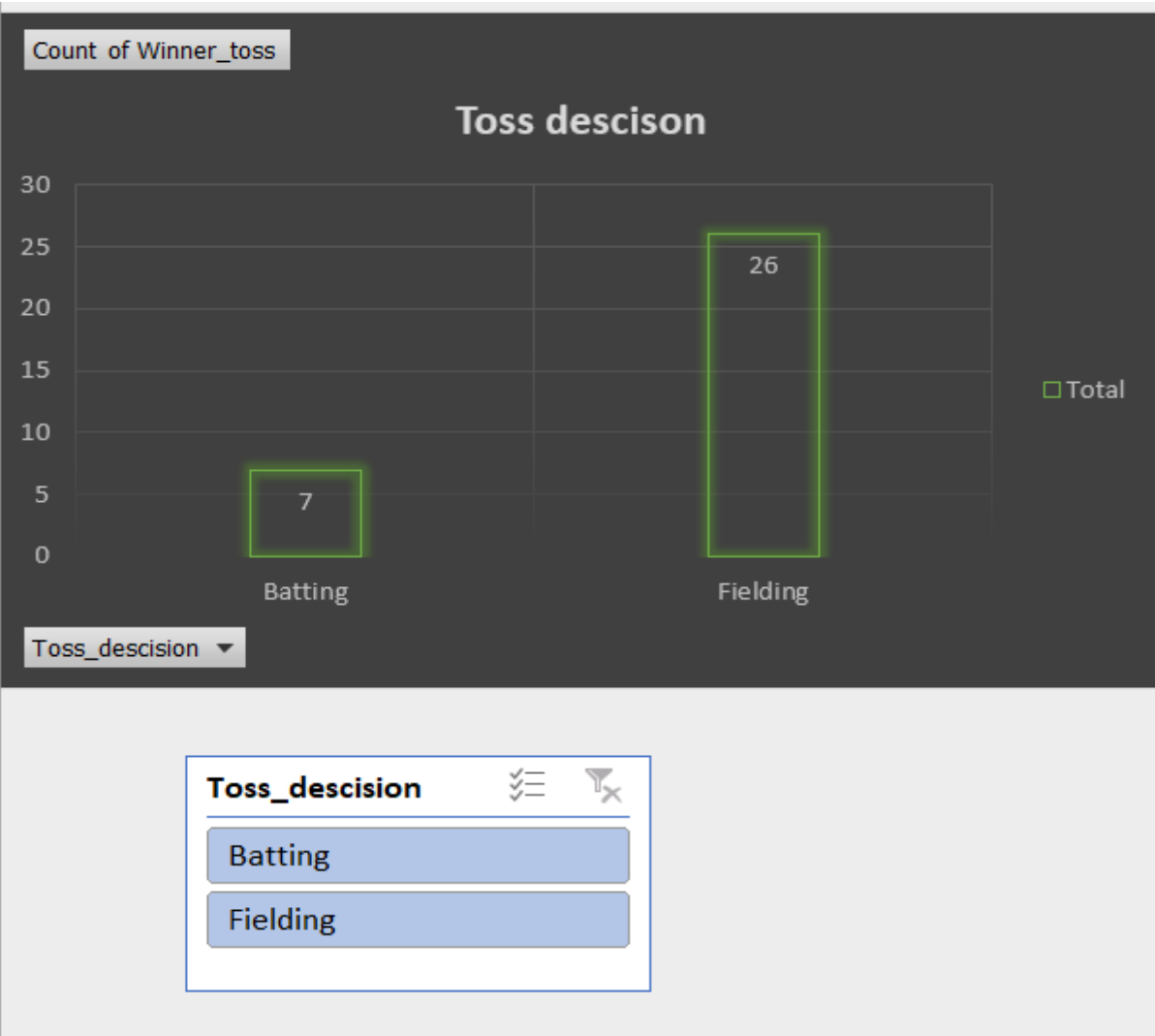
<u>Analysis of Dataset</u>

DASHBOARD OF THE PROJECT:

1. **percentage of the teams toss decision :**
   a) Introduction: The analysis shows the count of winner toss and decision taken by them.
   b) Specific Requirements/Functions and Formulas:
       i)Pivot table of Import Data
      ii)Pivot table of Export Data
      iii)SLICERS AND HYPER
      LINKS ARE USED
      iv)Clustered Column Chart

   c) Analysis Results:

we have observed that most toss winning team has opted to field first.

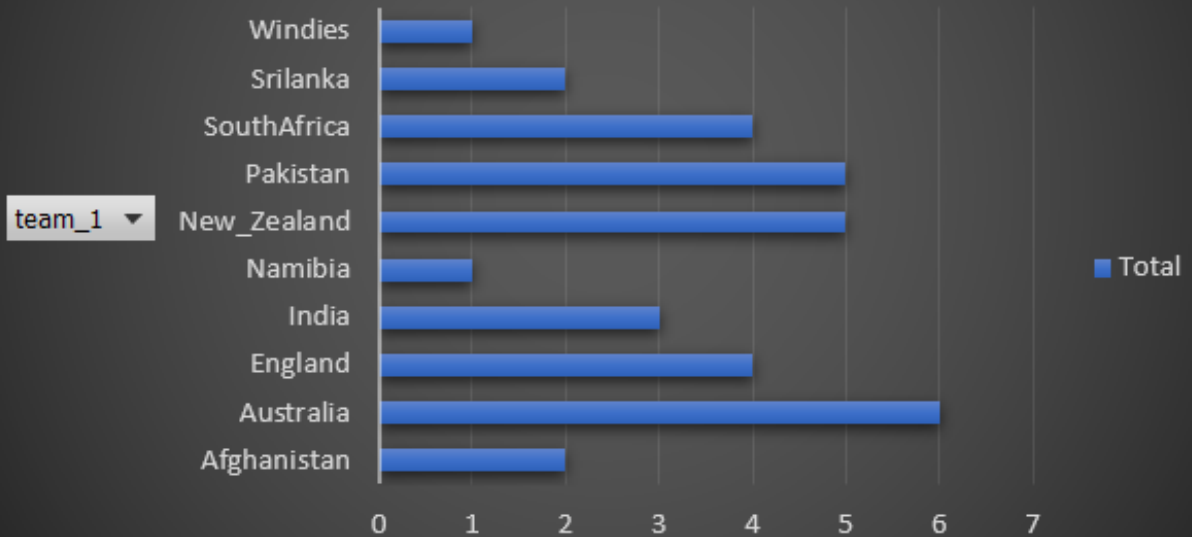Count of Winner_toss

Toss descison

Toss_descision

2.Matches won by each Team overall:
  d) Introduction: The analysis shows the total matches won by each team.
  e) Specific Requirements/Functions and Formulas:
    i)Pivot table of the data of Import ii)With the help of
    data Clustered Bar chart is plotted.
    iii)Slicers are used to change the team.

  f) Analysis Results:
    By change the teams we can observe that Australia has won more matches.

  g) Visualization:

Count of Winner

Count wins by each team

team_1 ▼

Legend: ■ Total

X-axis values: 0 1 2 3 4 5 6 7

Teams (top to bottom):
- Windies
- Srilanka
- SouthAfrica
- Pakistan
- New_Zealand
- Namibia
- India
- England
- Australia
- Afghanistan

Winner

Afghanistan	Australia	England	India
Namibia	New_Zealand	Pakistan	SouthAfrica
Srilanka	Windies		

3.Average target set by each team:

a)Introduction: The analysis shows the Team and the average target by them.
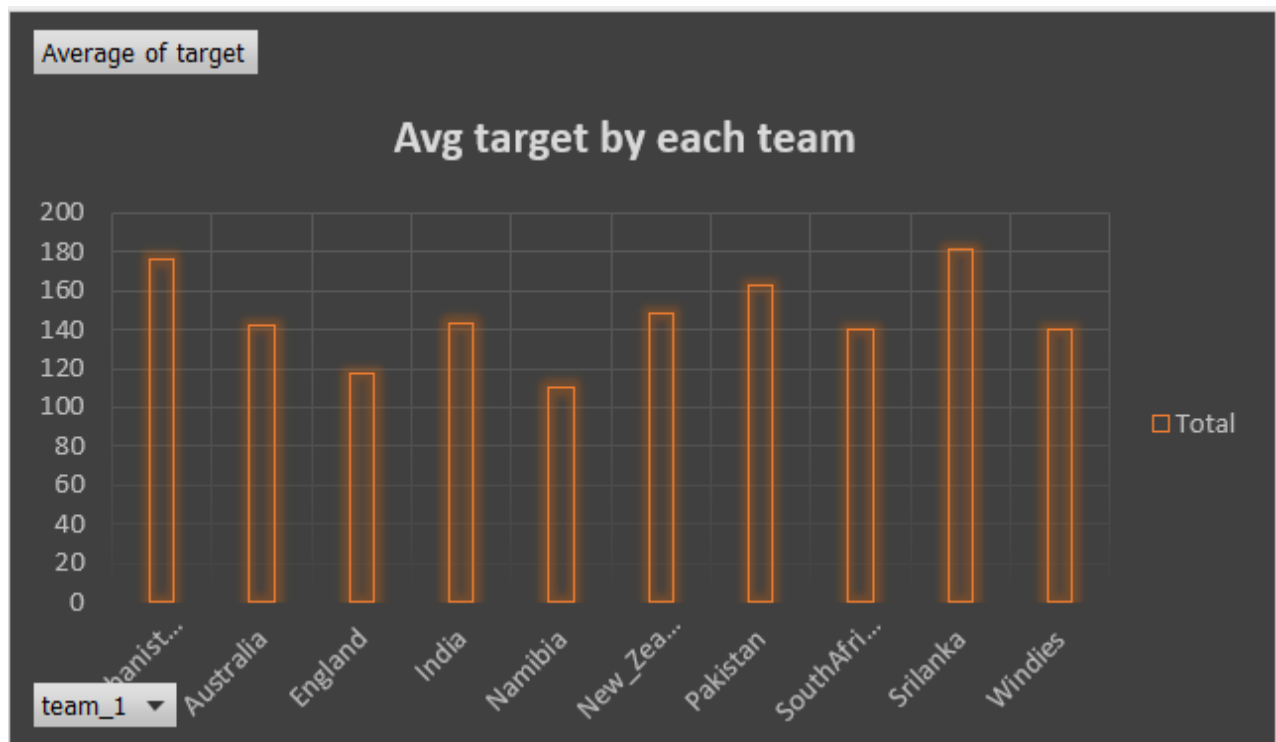
B)Specific Requirements/Functions and Formulas:
i)Pivot table of the data of Import.
ii)With the help of data Clustered Bar chart is plotted.

c)Analysis Results:

From the clustered graph we can observe that

most of teams are having average target 100+

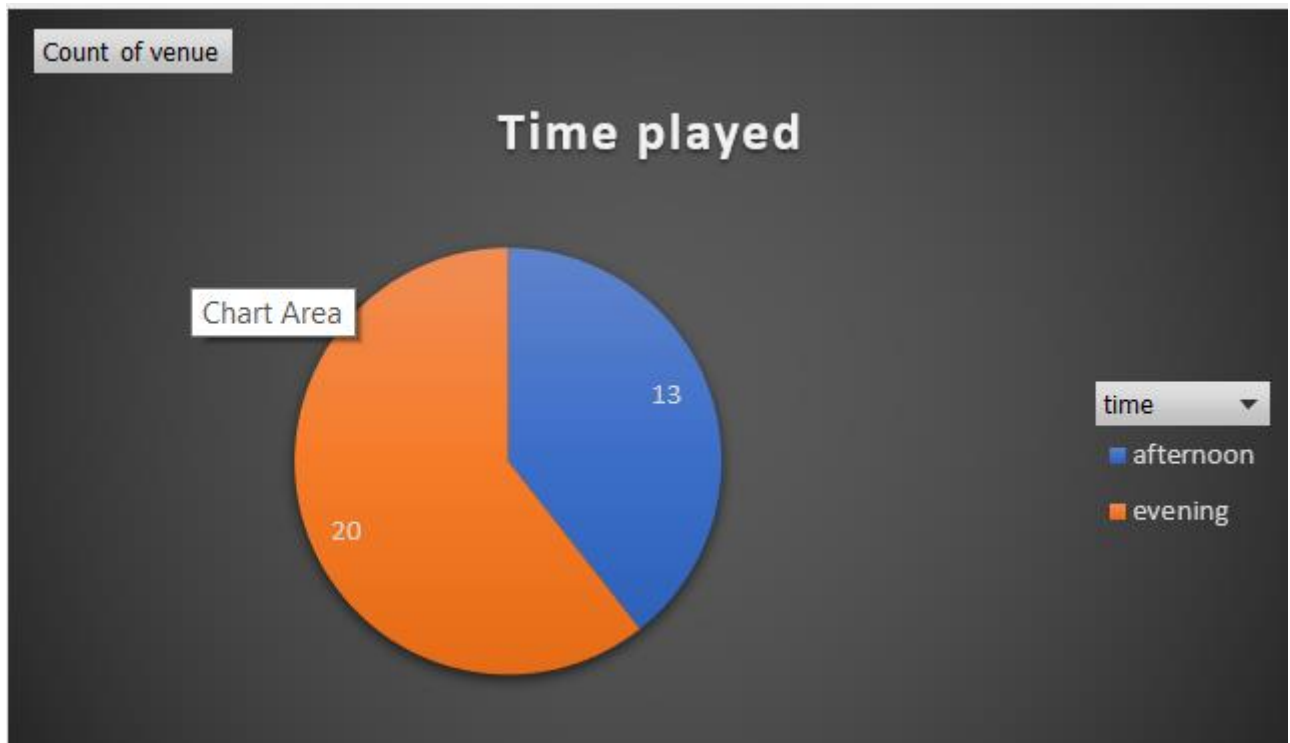### 4.Matches are played in Evening and Afternoon:

a) Introduction: The analysis shows the percentage of matches play and time .

b)Specific Requirements/Functions and Formulas:
i) Pivot table of the data of toss decision ii)With the help of data pie chart is plotted.

c)Analysis Results:

we can observe that most of teams are played in

evening.

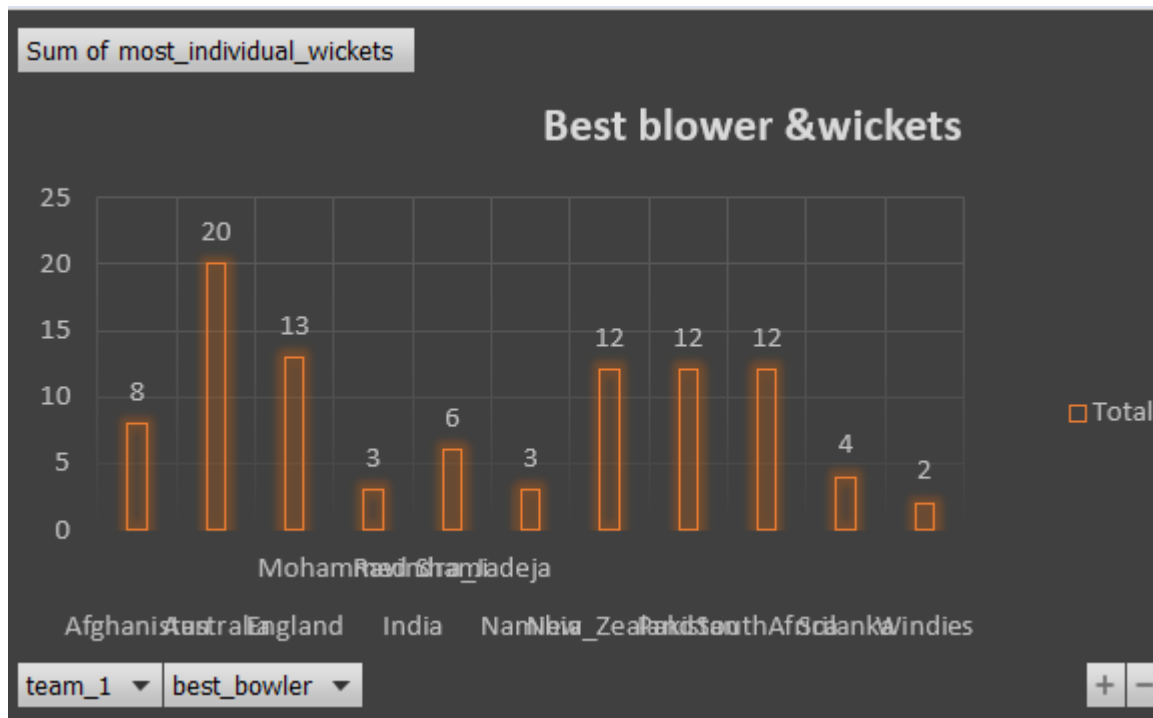5. **best blowers in each team and their most individual wickets:**
a) Introduction: The analysis shows the best bowler in each
team and their individual wickets taken.
b)Specific Requirements/Functions and Formulas:

  i) Pivot table of the data of best
  bowler and wickets.ii)With the
  help of data Clustered Bar chart
  is plotted.

c)Analysis Results:

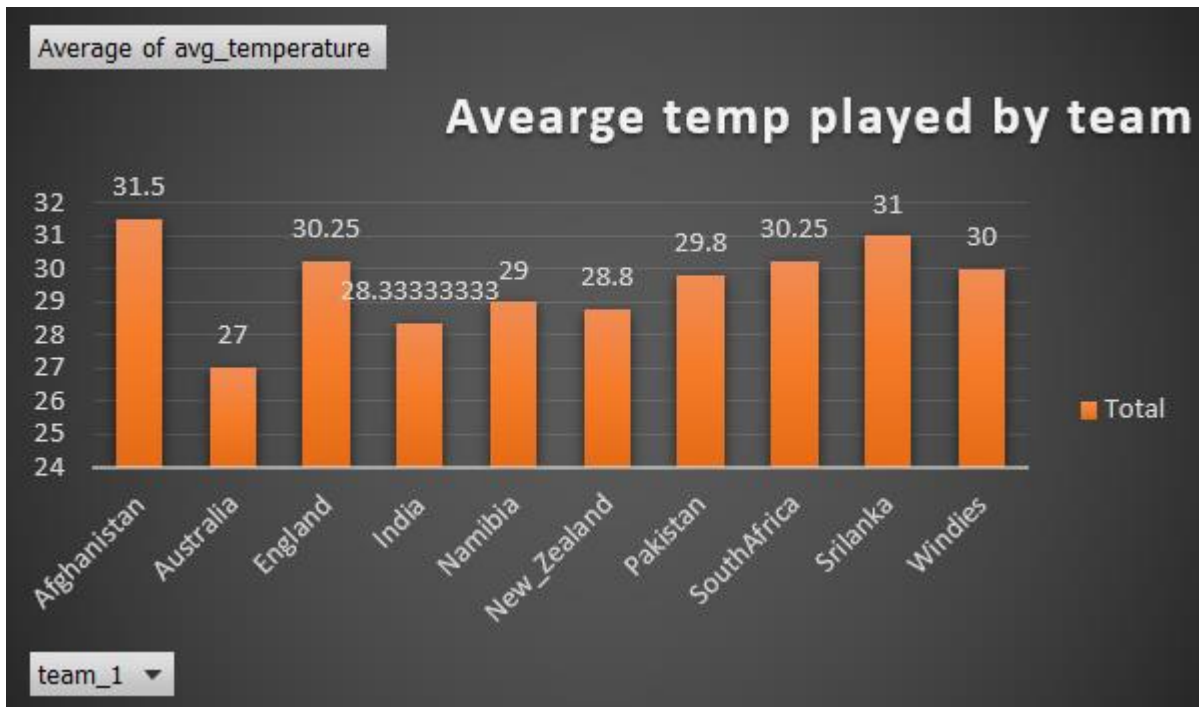we can observe that most of wickets are taken by

Australia.

6. Average temperature played by each team:
 a) Introduction: The analysis shows the Team
 wise average temperature played
 b)Specific Requirements/Functions and Formulas:
  i)Pivot table of the data of
  Import ii)With the clustered
  bar plotted.

c)Analysis Results:
    We can observe that average temperature is
        above 27 degrees.

7. Average economy of best blowers in each team:

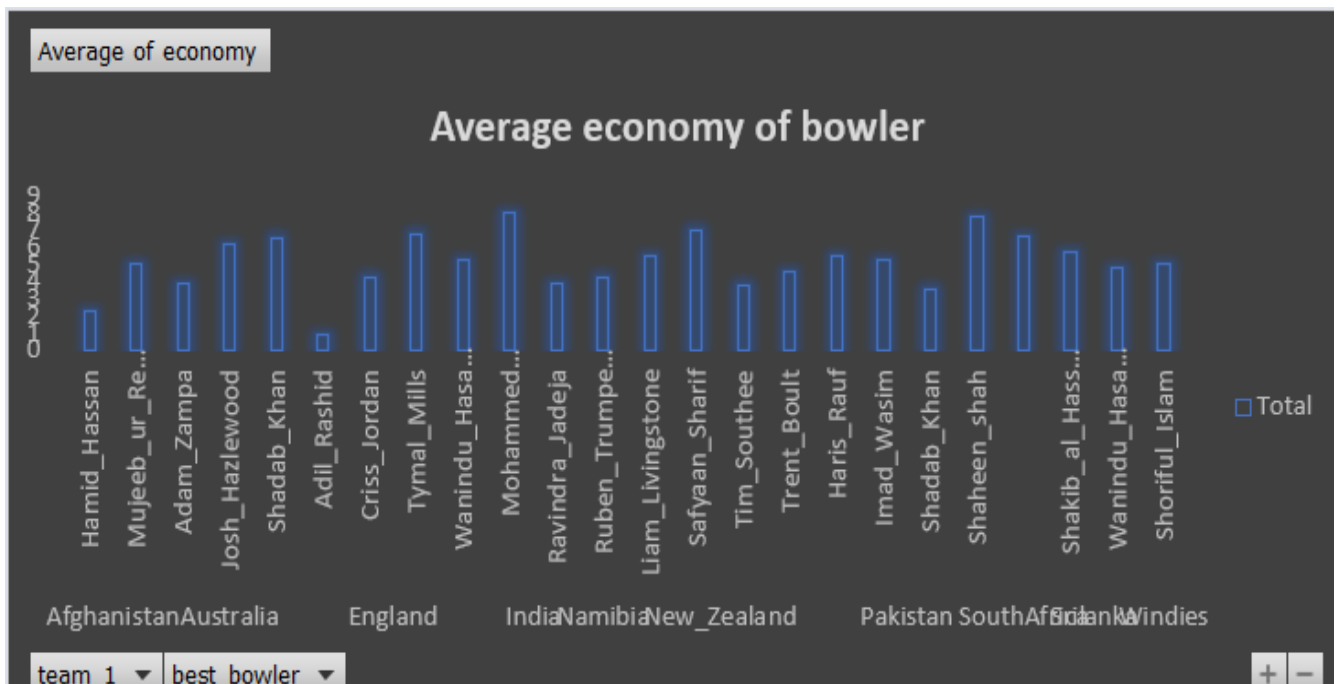a)Introduction: The analysis shows the Team and the average
economy of bolwer.

b)Specific Requirements/Functions and Formulas:
i)Pivot table of the data of Import.
ii)With the help of data Clustered Bar
chart is plotted.

c)Analysis Results:

From the clustered graph we can observe that

having average economy lowest is Adil Rashid

8. Total count of player of the matches of a team:

a)Introduction: The analysis shows the Team and the most man
of the matches each team.

b)Specific Requirements/Functions and Formulas:
i)Pivot table of the data of Import.
ii)With the help of data Clustered Bar
chart is plotted.

Slicer is used to change the team

c)Analysis Results:

From the clustered graph we can observe that having most man of
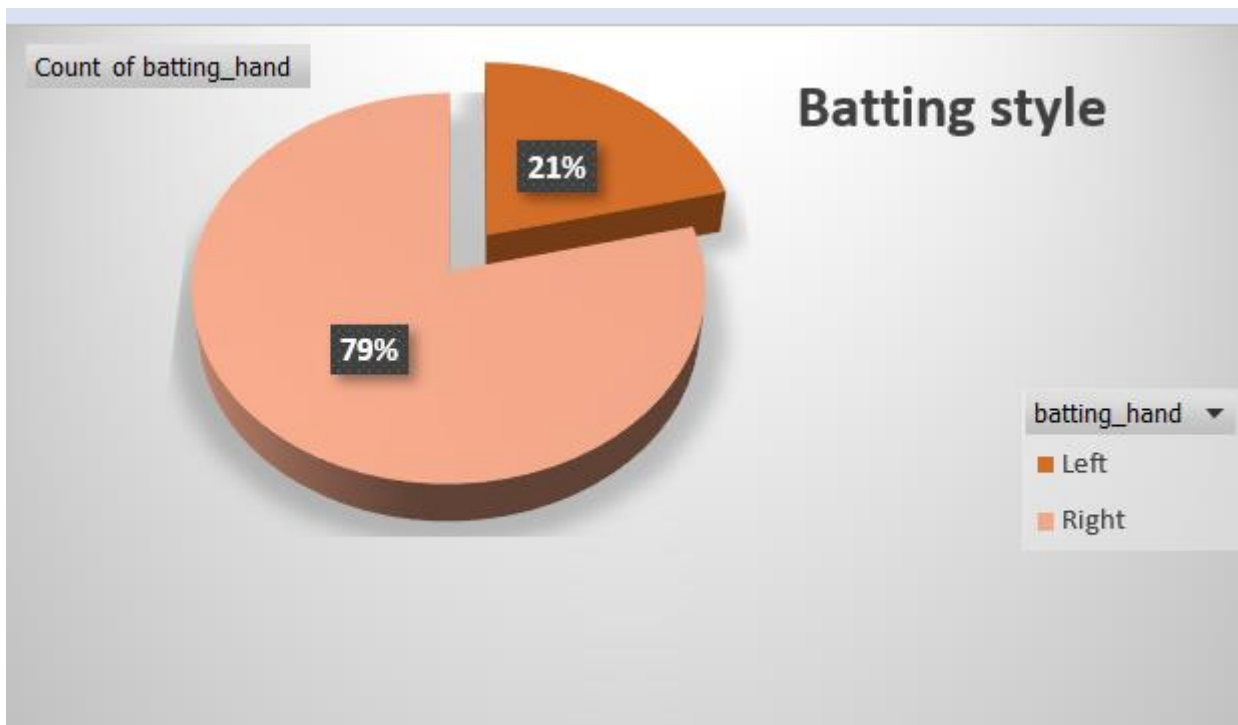the matches are Bangladesh.

Count of Player_of_the_match

Total Player of the matches each team

Chart Title

■ Total

team_2 ▼

team_2		
Afghanistan	Australia	Bangladesh
England	India	Namibia
New_Zealand	Pakistan	Scotland
SouthAfrica	Srilanka	Windies

9. Percentage of batting style:

a) Introduction: The analysis shows the percentage of batting
style.

b)Specific Requirements/Functions and Formulas:
   i) Pivot table of the data of
   batting styleii)With the help of
   data pie chart is plotted.

c)Analysis Results:

we can observe that most of team players are

Right handed.

10. sum of wickets different bowling style:

   a)Introduction: The analysis shows the Team and the most
wickets taken by different style bowling each team.

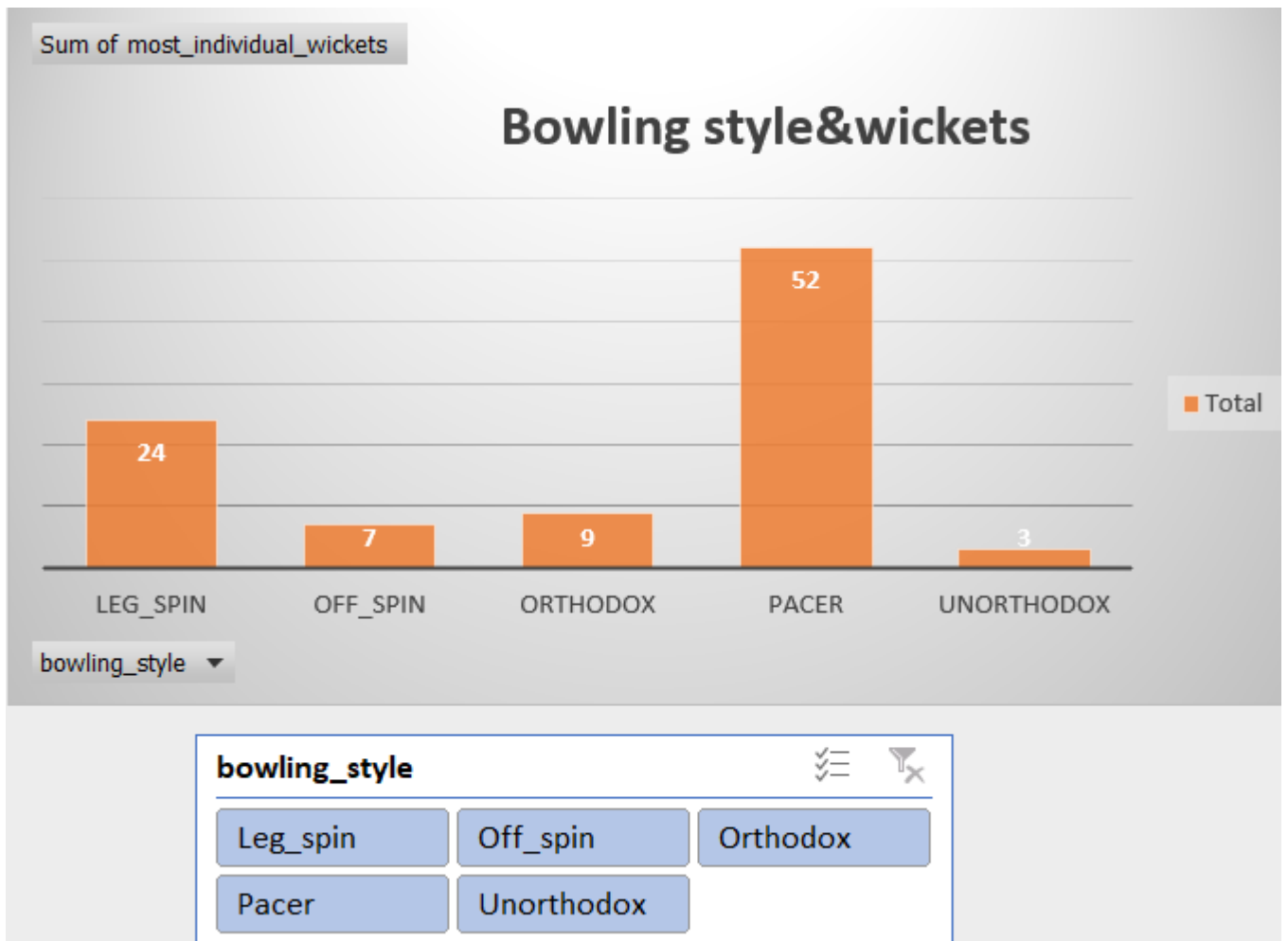   b)Specific Requirements/Functions and Formulas:
i)Pivot table of the data of Import.
ii)With the help of data Clustered Bar
chart is plotted.

   Slicer is used to change the bowling
style .

   c)Analysis Results:

From the clustered graph we can observe that having most wickets of pacer style bowling.

References

1.   www.kaggle.com

2.   www.youtube.com

3.   www.google.com

4.   www.stackoverflow.com

5.   www.github.com

Bibliography

1. Microsoft Excel 2016 Bible: The Comprehensive Tutorial Resource by John Walkenbach, Wiley

2. Fundamentals of Business Analytics by R.N. Prasad, Seema Acharya, Wiley