PRESIDENCY UNIVERSITY

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

A report on
"Data Visualization on 2021 Tokyo Olympics dataset"

A technical project work submitted in partial fulfillment of requirement for the award of the degree of

Bachelor of Technology
in
Computer Science and Engineering

By

VAISHALI M	20181CSE0764
VANISHREE R	20181CSE0770
VASAVI S M	20181CSE0776

Under the guidance of
Dr. DEEPAK SAKKARI
Assistance professor, Dept of CSE

Table of contents:

1. Introduction to Dataset

2. Python packages and libraries

3. Visualization of graphs

4. Conclusion

Introduction to Dataset:

The dataset used for visualization is "2021 Tokyo Olympics".

This contains the details of over 11,000 athletes, with 47 disciplines, along with 743 Teams taking part in the 2021 Tokyo Olympics.

This dataset contains the details of the Athletes, Coaches, Teams participating as well as the Entries by Gender. It contains their names, countries represented, discipline, gender of competitors, name of the coaches.

2021 Tokyo Olympics

Athletes
Columns: Name, NOC, Discipline
Size: 11085 rows, 3 columns

Coaches
Columns: Name, NOC, Discipline, Event
Size: 394 rows, 4 columns

Entries Gender
Columns: Discipline, Female, Male, Total
Size: 46 rows, 4 columns

Medals
Columns: Rank, Team/NOC, Gold, Silver, Bronze, Total, Rank by Total
Size: 93 rows, 7 columns

Teams
Columns: Name, Discipline, NOC, Event
Size: 743 rows, 4 columns

Python packages and libraries:

To read files/datset

Pandas

For plotting graphs

Matplotlib

Seaborn

Wordcloud

Geopandas

Plotly

** geopandas and plotly to be installed

To ignore warnings during runtime

Warnings

Visualization of graphs:

Using Matplotlib: Horizontal Bar Graph, Pie Chart

Using WordCloud: WordCloud

Using Geopandas: WorldMap

Using Seaborn: Catplot

Using plotly: Bar Graph, Animated Bar Graph, Pie Chart, Sunburst

Conclusion:

Athletes:

From our analysis we can say that United States of America, Japan, Australia, China, Germany are the top 5 countries with most number of participants.

Athletics has the most number of players i.e 2068, followed by swimming with 743 participants and football with 567 participants.

Coaches:

We can depict that Japan has the highest number of coaches i.e 25, followed by Spain and United States of America with 28 coaches and Australia with 22 coaches.

Gender:

We can say that male participants are more than female participants.

The overall male percentage is 52 and female percentage is 48.

Medals:

From the plots we can say that, United States of America has secured highest number of medals i.e 113, followed by China with 88 medals and ROC with 71 medals.

The top contries with gold medal are United States of America, China, Japan, Great Britain and ROC.

Teams:

From our analysis we can see that Japan has the highest number of teams played in 2021 Tokyo Olympics. There were totally 48 teams from Japan, 47 teams from United States of America, 37 teams from Italy and followed by Germany with 36 teams and Australia with 35 teams.