School of Computer Science and Engineering

(Computer Science & Engineering)

Faculty of Engineering & Technology

Jain Global Campus, Kanakapura Taluk - 562112
Ramanagara District, Karnataka, India

Project Outline

Project Title: Olympics Data Analysis

Project Goal: The goal of this data analysis project is to find patterns and information in the Olympics dataset. Through an examination of participation rates, medal distribution, and athlete characteristics in several Olympic Games, the project aims to illuminate factors influencing success, identify areas of growth, and potentially reveal social impacts associated with this global sporting event.

Team Name: Strategy Stars

Team Members and Roles:

- Team Lead: Hitha Choudhary G – 22BTRAD015
- Data Analyst: K Shreeshanth Gouda – 22BTRAD017
- Visualization Specialist: Lakshya Sharma – 22BTRAD021

Resources:

The dataset was collected from a public repository which was a zip file("Olympic Dataset") that included two csv files "athlete_events.csv" and "noc_regions.csv"

- athlete_events.csv contains a total of 271116 rows and 15 columns and the att ributes are as follows:
  ID, Name, Sex, Age, Height, Weight, Team, NOC, Games, Year, Season, City , Sport, Event, Medal
- Noc_regions.csv contains a total of 230 rows and 3 columns and the attributes are as follows:
  NOC, region, notes

Stages with Onus:

Stage 1: Data Acquisition and Exploration

Objective: Acquire dataset and understand its structure and summary statistics.

- Action Item: Team Lead acquires the dataset(example., a public repository)
  Onus: Hitha Choudhary G (Team Lead)
- Action Item: A data analyst examines the data with tools such as pandas to comprehend the data structure, identify inconsistencies or missing values, and obtain summary statistics (numerical columns).
  Onus: K Shreeshanth Gouda (Data Analyst)

Stage 2: Data Cleaning and Analysis

Objective: Clean the data, identify patterns, trends, and relationships in line with project goals.

- Action Item: Assisting with missing numbers and irregularities, the data analyst cleans up the data and does analysis depending on project goals, such as identifying patterns, trends, and relationships in the data.
  Onus: K Shreeshanth Gouda (Data Analyst)
- Action Item: The team lead works with the data analyst to specify certain analysis tasks and guarantee that they are in line with the project objectives.
  Onus: Hitha Choudhary G (Team Lead)

Stage 3: Data Visualization

Objective: Produce clear and informative visualizations to illustrate research results.

- Action Item: Visualization Specialist produces scatter plots, bar charts, and histograms to illustrate research results. While also ensuring that the visualizations are clear and informative.
  Onus: Lakshya Sharma (Visualization Specialist)

Stage 4: Key Inferences

Objective: Collaboratively discuss important conclusions and insights gained from each step.

- Action Item: All the team members collaborate to go over the most important conclusions and practical insights gained from each step and come to a consensus.
- Onus: All Team Members

Deliverables:

- Analysis Methodology:
  Document outlining the approach adopted for data acquisition, cleaning, analysis, visualization, and inference derivation.
- Report on Data Analysis:
  Comprehensive report showcasing significant findings derived from the analysis, including descriptive statistics, patterns, trends, and relationships identified within the dataset.
- Key Findings and Insights:
  Summary document highlighting the most important conclusions and practical insights gained from the project, accompanied by concise justifications and relevant images or visuals.

Project Deadline:

10/04/2024 – Wednesday