CHANDIGARH
ENGINEERING COLLEGE
CGC, LANDRAN, MOHALI
Building Careers. Transforming Lives.
CEC
LANDRAN

TITLE : Olympic Data Analyzer

Submitted To

Name of CC : Harpreet kaur

Submitted by: Abhishek Kumar

Roll no: 2237786

Abstract

This project explores the evolution of the Olympics from 1896 to 2016, focusing on nations' performance and participation. It analyzes dominant sports contributions, compares performance across sports, and suggests ways to enhance future participation. Factors including increased nation and athlete participation, event numbers, costs, performance improvements, gender ratios, medical facilities, and pandemic impacts are examined to understand Olympic evolution and make future predictions.

Introduction

- The Olympics, starting in 1896, occur every four years and provide a global stage for athletes.
- This project aims to analyze nations' performance and participation from 1896 to 2016.
- It identifies each country's dominant sports contributions over time.
- Sports performances are compared, highlighting areas with potential for increased participation.
- Suggestions for enhancing future Olympic contributions are proposed.
- The Olympics encompass both summer and winter events, with over 200 participating nations.
- Various factors in Olympic evolution are examined, such as increased nation and athlete participation, event numbers, costs, performance improvements, gender ratios, medical facilities, and pandemic impacts.
- Analyzing these factors aids in understanding Olympic evolution and facilitates future predictions.

Objectives of Project

The main objective of this project was to analyze and visualize the factors contributing to the evolution of the Olympic Games over time. By employing Exploratory Data Analysis, the study aimed to provide insights that could help countries and athletes enhance their performance by adapting strategies. Through Python implementation, the analysis revealed significant changes in the Olympics, including the introduction of the Winter Games, increased participation of countries and female athletes, shifts in player demographics, and medal tallies. Visualization techniques such as line graphs, scatter plots, bar graphs, and distribution plots were utilized to validate and present the findings effectively.

Methodology

- Data Collection:
  - Gathered datasets containing player details (e.g., gender, height, weight, country, medals), country information (e.g., participation, medals won), and country codes.
  - Ensured a diverse and extensive dataset for comprehensive analysis.
- Data Pre-Processing:
  - Processed raw data to make it usable for analysis.
  - Employed Deterministic Imputation technique to handle null values by replacing them with appropriate substitutes (e.g., mean, median, similar records).
- Exploratory Data Analysis (EDA):
  - Utilized EDA techniques to analyze and visualize data.
  - Conducted comparative studies on various factors influencing the evolution of the Olympic Games over time.
  - Employed visualization tools such as histograms, bar graphs, box plots, scatter plots, etc., to interpret and present findings effectively.

Work Done

- Data collection from various sources related to Olympic Games, including player details, country information, and country codes.
- Data pre-processing to clean and prepare the datasets for analysis, including handling null values using Deterministic Imputation.
- Exploratory Data Analysis (EDA) conducted using Python, focusing on visualizing and analyzing factors contributing to the evolution of the Olympics.
- Identified trends such as changes in participant demographics, country-wise performance, and the impact of various factors over time.
- Utilized visualization techniques such as histograms, bar graphs, box plots, scatter plots, etc., to present findings effectively.

Work to be Done

- Advanced Data Analysis Techniques: Explore predictive modeling, clustering, or other advanced techniques to gain deeper insights into Olympic Games evolution.
- Temporal Analysis: Conduct a detailed temporal analysis to understand how specific factors have evolved over time and their influence on the Olympics.
- Comparative Studies: Perform comparative studies between different editions of the Olympics or between different countries to identify patterns and trends.
- Performance Prediction: Develop models to predict future performance trends in the Olympics based on historical data and identified factors.
- Integration of Additional Data Sources: Incorporate additional datasets such as economic indicators or technological advancements to provide a holistic understanding of Olympic Games evolution.
- Interactive Visualization: Create interactive visualization tools to allow stakeholders to explore and interact with the data more effectively.

Timelines

May–June
- Data collection and initial preprocessing.
- Commencement of exploratory data analysis (EDA).
- Continued EDA and refinement of analysis techniques.
- Integration of additional datasets if needed.
- Development of initial visualizations.

June–July:
- Further analysis and visualization refinement.
- Comparative studies and exploration of advanced techniques.
- Finalization of analysis and report preparation.
- Presentation of findings and potential recommendations.

References

- Text size: 20 Calibri
- Wikipedia contributors: https://en.m.wikipedia.org/wiki/Olympic Games, last accessed 2020/11/02.
- Dey S K, Rahman M M, Siddiqi U R and Howlader A 2020 Analyzing the epidemiological outbreak of COVID–19: A visual exploratory data analysis
- approach J. Med. Virol. 92 632–8
- Bondu R, Cloutier V, Rosa E and Roy M 2020 An exploratory data analysis approach for assessing the sources and distribution of naturally occurring
- contaminants (F, Ba, Mn, As) in groundwater from southern Quebec (Canada) Appl. Geochem. 114 104500
- Cutait, M.: Management performance of the Rio 2016 Summer Olympic Games. Research Paper submitted and approved to obtain the Master's degree in
- Sports Administration at AISTS in Lausanne, Switzerland.
- Moreno A, Moragas M and Paningua R 1999 The evolution of volunteers at the Olympic Games Proceedings of Symposium on Volunteers (Lausanne, Switzerland: Global Society and the Olympic Movement) pp 1–18

References

- Abeza G, Braunstein-Minkove J R, S′eguin B, O'Reilly N, Kim A and Abdourazakou Y 2020 Ambush marketing via social media: The case of the three most recent Olympic Games Int. J. Sport Communication 1–25
- Yamunathangam D, Kirthicka G and Shahanas P 2018 Performance Analysis in Olympic Games using Exploratory Data Analysis Techniques International
- Journal of Recent Technology and Engineering (IJRTE) 7 251–3
- Wikipedia contributors: Exploratory data analysis, https://en.wikipedia.org/wiki/Exploratory data analysis, last accessed 2020/11/11.
- Ramachandran K. M. and Tsokos C P 2020 Mathematical statistics with applications in R (Academic Press)
- Lange D Summer Olympics: number of participating countries 1896–2016 Statista.com

Thank You
Queries?