120 years of Olympic History Analysis

Lovely Professional University

Requirements Specification

Version 1.0

Developed By

Sendrela Gulzar (12209369)

Ravi Kumar (12216513)

Sarthak Tiwari (12216889)

Table of Contents

1. Executive Summary

1.1 Project Overview

This project is to develop a web application that allows users to explore and analyze the Olympic Games dataset from Kaggle. The application will provide users with information on medals won, athletes, and events.

1.2 Purpose and Scope of this Specification

This specification defines the functional and non-functional requirements for the web application. The functional requirements describe the features and functionality that the application must provide. The non-functional requirements describe the performance, reliability, security, and other aspects of the application.

2. Product/Service Description

2.1 User Characteristics

The target users of the web application are people, who are interested in the Olympic Games. This includes sports fans, researchers, and students.

2.2 Assumptions

The following assumptions are made about the users of the web application:

- Users have access to a computer with an internet connection.
- Users have a basic understanding of how to use a web browser.
- Users are interested in learning about the Olympic Games.

2.3 Constraints

The following constraints apply to the development of the web application:

- The application must be developed within a budget of $10,000.
- The application must be developed within a timeframe of 6 months.
- The application must be compatible with all major web browsers.

2.4 Dependencies

The web application will depend on the following software and hardware:

Software: Python, Streamlit, Pandas, Plotly, Matplotlib, Seaborn
Hardware: Web server, database server

3. Requirements

3.1 Functional Requirements

The web application must provide the following features:

- Allow users to view the medal tally for each Olympic Games.
- Allow users to filter the medal tally by country, year, and sport.
- Allow users to view the top athletes for each Olympic Games.
- Allow users to filter the top athletes by country, year, and sport.
- Allow users to view the top events for each Olympic Games.
- Allow users to filter the top events by country, year, and sport.
- Allow users to create custom visualizations of the Olympic Games data.

Requirement no.	Requirement name	Requirement description
1	Medal Tally	Allow users to view the medal tally for each Olympic Games.
2	Overall Analysis	Allow users to view the top athletes and events for each Olympic Games.
3	Country-wise Analysis	Allow users to filter the medal tally, athletes, and events by country.
4	Athlete wise Analysis	Allow users to filter the medal tally, athletes, and events by athlete.
5	Create Custom Visualizations	Allow users to create custom visualizations of the Olympic Games data.

Non-Functional Requirements:

Hardware Requirements:

Processor: Intel Dual-Core or above
RAM: 4 GB or above
Software Requirements:

Python 3.6 or above
Streamlit 1.2 or above
Pandas 1.2 or above
Plotly 5.1 or above
Matplotlib 3.5 or above
Seaborn 0.11 or above
Web server (e.g., Apache, Nginx)
Database server (e.g., MySQL, PostgreSQL)
3.2 User Interface Requirements

The web application must have a user-friendly interface that is easy to navigate. The interface should be consistent and visually appealing.

3.3 Performance

The web application must be able to handle a large number of concurrent users without sacrificing performance. The application should also be responsive and load quickly.

3.3.1 Capacity

The web application must be able to handle up to 100 concurrent users.

3.3.2 Availability

The web application must be available 99% of the time.

3.3.3 Latency

The web application must respond to user requests within 1 second.

3.4 Manageability/Maintainability

The web application must be easy to manage and maintain. The application should be well-documented and modular in design.

3.4.1 Monitoring

The web application must be monitored for performance and errors.

3.4.2 Maintenance

The web application must be easy to update and maintain. New features and bug fixes should be released on a regular basis.

3.4.3 Operations

The web application should be easy to deploy and operate.

3.5 System Interface/Integration

The application should interface with external data files ('athlete_events.csv' and 'noc_regions.csv').

3.6 Security

3.6.1 Protection: The application should protect user data and prevent unauthorized access.
3.6.2 Authorization and Authentication: User access should be controlled and authenticated.

3.7 Data Management

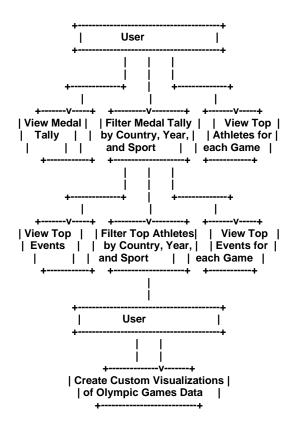The application must efficiently manage and process Olympic data.

3.8 Standards Compliance

The application should adhere to coding and design standards.

4. User Scenarios/Use Cases

Users can select different analysis options (Medal Tally, Overall Analysis, etc.) and specify years and countries for analysis.
Users can view graphical representations of Olympic data.

```
                    +------------------------------------+
                    |              User                  |
                    +------------------------------------+
                        |     |     |
                        |     |     |
                  +-------------+   |   +---------------+
                  |             |   |   |               |
            +-------V-----+ +---------V--------+ +-------V-----+
            | View Medal  | | Filter Medal Tally |  |  View Top  |
            |   Tally     | |  by Country, Year, |  | Athletes for |
            |             | |     and Sport      |  | each Game  |
            +-------------+ +--------------------+ +-------------+
                        |     |     |
                        |     |     |
                  +-------------+   |   +---------------+
                  |             |   |   |               |
            +-------V-----+ +---------V--------+ +-------V-----+
            | View Top    | | Filter Top Athletes|  |  View Top  |
            |  Events     | |  by Country, Year, |  | Events for |
            |             | |     and Sport      |  | each Game  |
            +-------------+ +--------------------+ +-------------+
                                   |
                                   |
                    +------------------------------------+
                    |              User                  |
                    +------------------------------------+
                        |   |
                        |   |
                  +--------------V-------+
                  | Create Custom Visualizations |
                  | of Olympic Games Data    |
                  +--------------------------+
```

In this use case diagram:

- The "User" interacts with the system through various use cases, such as "View Medal Tally," "Filter Medal Tally," "View Top Athletes," "Filter Top Athletes," "View Top Events," and "Create Custom Visualizations."
- Arrows represent the interaction between the user and the system through these use cases.
- The use cases are organized based on their functionalities.
- This diagram provides an overview of how users interact with the system to perform specific tasks related to Olympic Games data analysis. You can create a visual representation of this diagram using various diagramming tools or software like Lucidchart, draw.io, or similar tools.

5. About Dataset

Context
This is a historical dataset on the modern Olympic Games, including all the Games from Athens 1896 to Rio 2016.

Note that the Winter and Summer Games were held in the same year up until 1992. After that, they staggered them such that Winter Games occur on a four year cycle starting with 1994, then Summer in 1996, then Winter in 1998, and so on. A common mistake people make when analyzing this data is to assume that the Summer and Winter Games have always been staggered.

Content

The file athlete_events.csv contains 271116 rows and 15 columns. Each row corresponds to an individual athlete competing in an individual Olympic event (athlete-events). The columns are:

ID - Unique number for each athlete
Name - Athlete's name
Sex - M or F
Age - Integer
Height - In centimeters
Weight - In kilograms
Team - Team name
NOC - National Olympic Committee 3-letter code
Games - Year and season
Year - Integer
Season - Summer or Winter
City - Host city
Sport - Sport
Event - Event
Medal - Gold, Silver, Bronze, or NA

5. Deleted or Deferred Requirements

No requirements have been deleted or deferred at this time.

This Software Requirements Specification (SRS) provides an overview of the "120 Years of Olympic History" project, its functional and non-functional requirements, and user scenarios. It serves as a reference for the development and evaluation of the application.

APPENDIX A. REFERENCES

The Olympic data on www.sports-reference.com is the result of an incredible amount of research by a group of Olympic history enthusiasts and self-proclaimed 'statistorians'. Check out their blog for more information. All I did was consolidated their decades of work into a convenient format for data analysis.

Dataset link: https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-results