# Olympics Data Summary - Assignment

## Problem Statement:

You have been hired as a Junior Data Analyst at a global sports analytics company. Your role is to analyze Olympic Games data to uncover trends in athletes' performance, country-wise medal statistics, gender participation, and the popularity of different sports over the years.

The organization is particularly interested in exploring medal distribution patterns, top-performing countries and athletes, and changes in participation across seasons. You are required to use your Python, NumPy, and Pandas skills to explore, clean, and analyze the dataset to provide meaningful insights backed by data.

## Dataset Details:

Dataset Name: 120 years of Olympic history: athletes and results

Source: Kaggle

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

File Used: athlete_events.csv

## Dataset Columns:

- Name - Name of the athlete
- Sex - Gender of the athlete
- Age - Age during the Olympics
- Team - Country represented
- Year - Year of the Olympic event
- Season - Summer or Winter
- Sport - Sport category (e.g., Athletics, Swimming)
- Event - Specific event (e.g., Men's 100m Freestyle)
- Medal - Medal won: Gold, Silver, Bronze, or None

## Your Objectives:

1. Data Cleaning:
 - Handle missing values and duplicates

- Ensure correct data types (e.g., Age as numeric)
2. Exploratory Analysis:

- How many athletes participated over the years?

- Top 10 countries with the most medals

- Which sports have the most events and medals?

3. Statistical Insights with NumPy:

- Calculate average and range of athlete ages -

Number of unique sports, events, and countries

4. Pandas Analysis:

- Group medal counts by country, sport, and year

- Find the athlete with the most medals

- Track gender participation over time

5. Bonus (Optional):

- Visualize trends using matplotlib or seaborn - Identify

emerging sports or declining sports in Olympics

**Skills You'll Apply:**

- Data loading and cleaning with Pandas

- Numerical analysis using NumPy

- Grouping, filtering, and aggregating data

- Generating insights from real-world sports data