Olympics Medals Analysis Project

Project under Unified Mentor Data Science Internship

1. Project Overview

This project explores data from the Summer Olympics (1976–2008) to understand patterns in medal wins, top-performing countries, gender participation, and how performance changed over the years.

The main goal is to find out which countries dominate, how participation evolved, and what trends shaped Olympic history during this period**.**

2. Data Cleaning and Preparation

The data came from a file named *"Summer-Olympic-medals-1976-to-2008.csv."*
Before starting the analysis, the data was cleaned properly:

- Empty rows and missing values were removed.
- The "Year" column was converted into a number for easier use.
- The dataset was checked to make sure there were no errors.

Insight: After cleaning, the data became neat and ready for visualization and analysis.

3. Top 10 Countries by Total Medals

A bar chart was created showing which countries won the most medals from 1976 to 2008.
Top Performers:
US United States, RU Russia, DE Germany, CN China, and a few European countries like Great Britain and France.

Insight:
These countries dominate because of their **strong sports infrastructure**, **training programs**, and **consistent participation**.
They also invest heavily in athlete development and host major sporting events more often.

4. Medals Over the Years

A line graph showed how the total number of medals changed from 1976 to 2008.
There was a **slight increase** in the number of medals over time.

Insight: The Olympics have grown bigger over the years, with **more countries participating** and **new sports being added**.
It shows how global interest in sports has expanded.

5. Gender Distribution

A pie chart showed the distribution of male and female athletes.
It was found that **men still had a higher share** of participation and medals.
But over time, **women's participation increased a lot**.

Insight: The gap between male and female athletes is closing.
This is a positive sign that shows **more equality and inclusiveness** in the Olympics today compared to the past.

6. Popular Sports and Events

Even though the project focused more on medals and countries, the data also hinted that **Athletics, Swimming, and Gymnastics** are the most medal-rich sports.

Insight: Countries that focus on these sports tend to win more medals.
For example, the USA performs really well in swimming and track events, which helps it stay at the top.

7. Machine Learning (Prediction Part)

A Logistic Regression model setup was included in the notebook, which indicates a plan to predict outcomes like:

- Which country might win a medal, or
- What factors (e.g., gender, event type) influence medal wins.

Insight:
Using predictive models can help analyze which characteristics or strategies increase a country's chances of winning medals.

8. Key Takeaways

Here's what we learned from this project:

- The **USA** has been the most successful Olympic country overall.
- **More people and countries** are taking part each year.
- **Women's participation** has grown, showing progress toward equality.
- **Certain sports** like swimming and athletics dominate the medal table.
- Data visualization makes it easy to understand Olympic history and trends.

9. Conclusion

This project gives a clear picture of how the Olympics have evolved over time — not just in sports, but also in global participation and gender equality.

It shows that success in the Olympics isn't just about talent — it's about **preparation, opportunity, and investment in athletes**.

"The Olympics tell the story of how nations grow through sports — celebrating diversity, hard work, and unity across the world."