SportsStats Olympic Athletes Analysis: 120 Years of Trends & Insights

CLIENT: SPORTSSTATS

BIBI HENA YAQUBI

2025/10/24

Client & Project Overview

Client: SportsStats — provides insights for sports journalists, trainers, and historians.

Goal: Analyze 120 years of Olympic data to discover athlete trends, country performance, physical traits, and event competitiveness.

Data: Athlete events & NOC regions, cleaned & joined.

Initial Hypotheses

- Large countries dominate total medal counts.

- Female athlete participation increased over time.

- Certain sports skew younger (gymnastics) or older (equestrian).

Approach

Cleaned & joined multiple datasets using Python.

Ran SQL queries with DuckDB for data exploration and analysis.
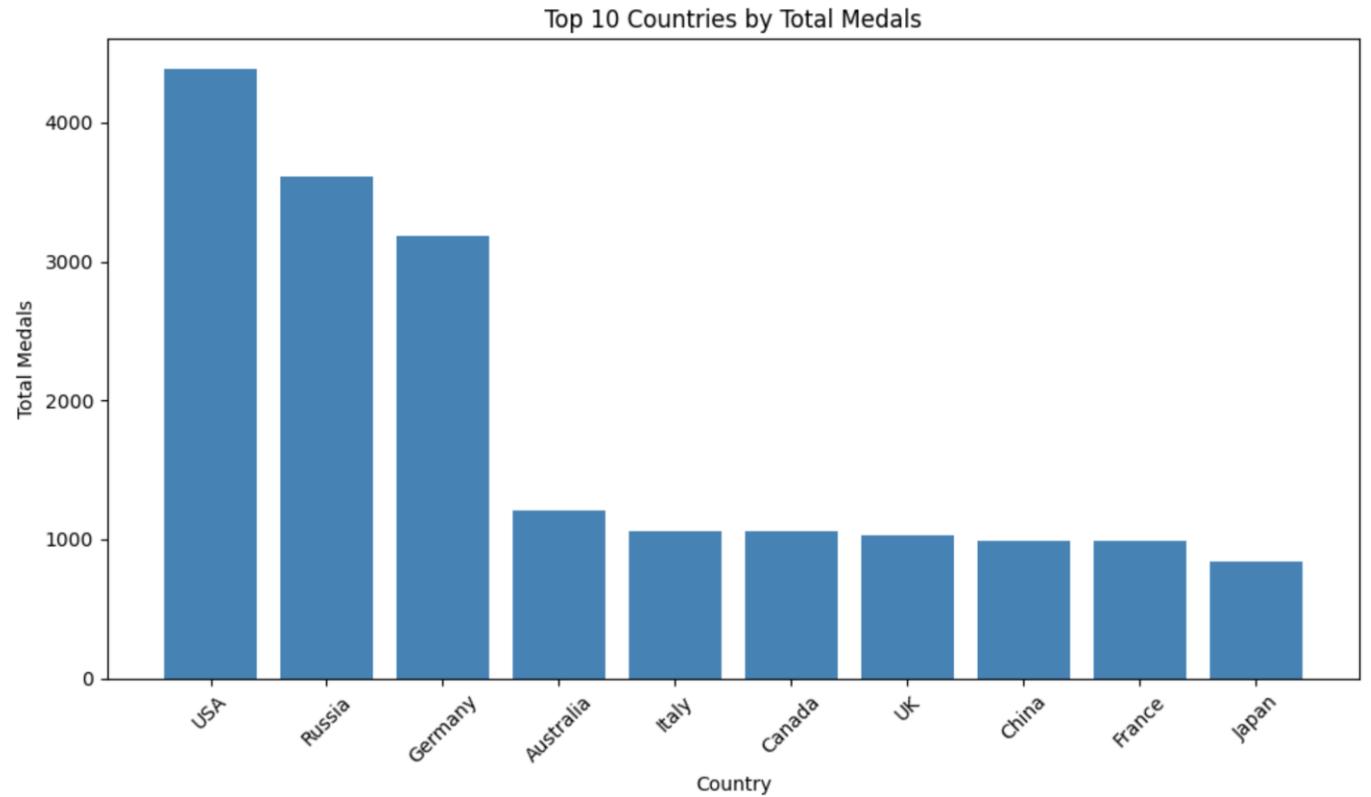
Used descriptive stats, correlations.

Created new metrics: medals per athlete & weight-to-height ratio.

Visualized trends using matplotlib.
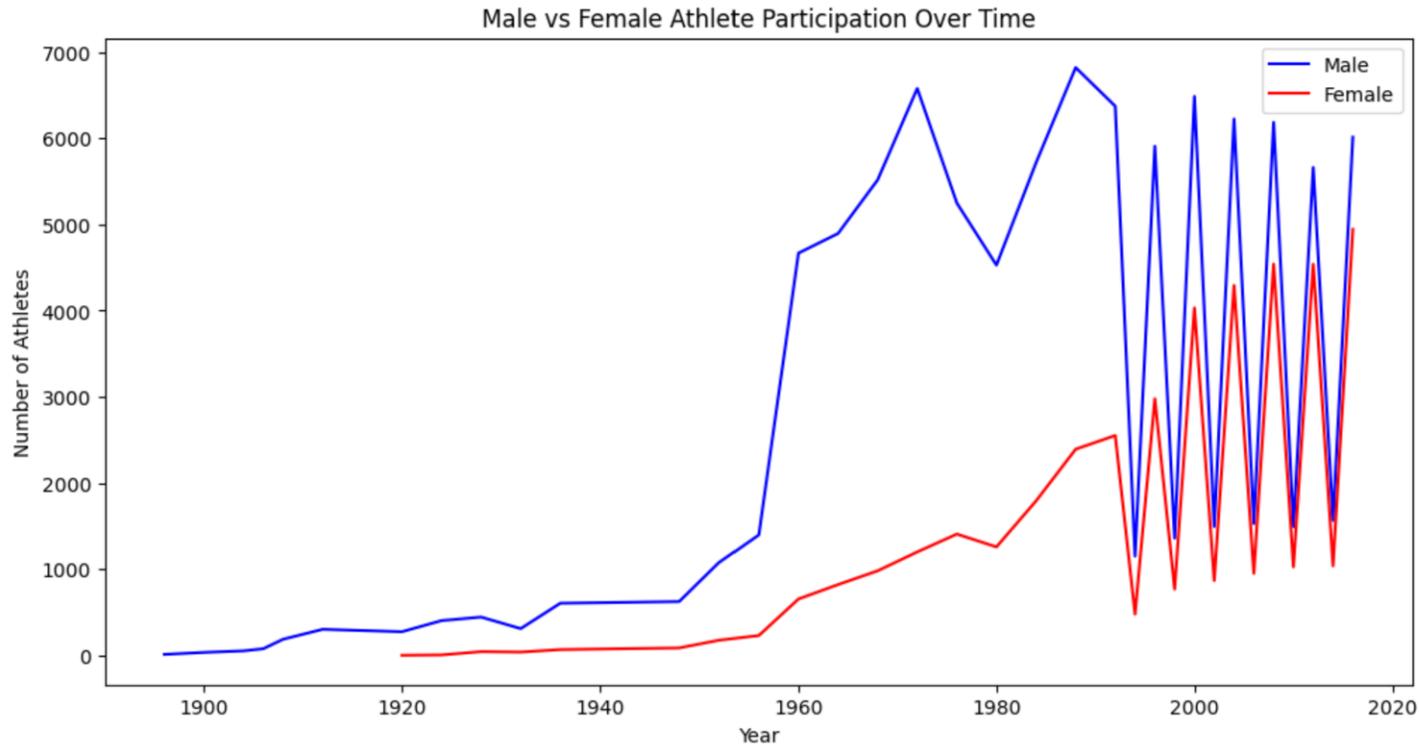
Descriptive Statistics

Key Stats:

◦ Total Records: 200k+ athletes.

◦ Top Countries: USA, Russia, Germany.

◦ Top Sports: Athletics, Swimming, Gymnastics.

Top 10 Countries by Total Medals

Gender Participation Trends

Clear increase in female athletes, especially post-1980.

Male vs Female Athlete Participation Over Time

Physical Profiles by Sport

- Average age, height, weight calculated for all medalists.
- *Rhythmic Gymnastics*: youngest average age (19).
- *Equestrianism*: oldest average age (35).

Average Age of Medalists
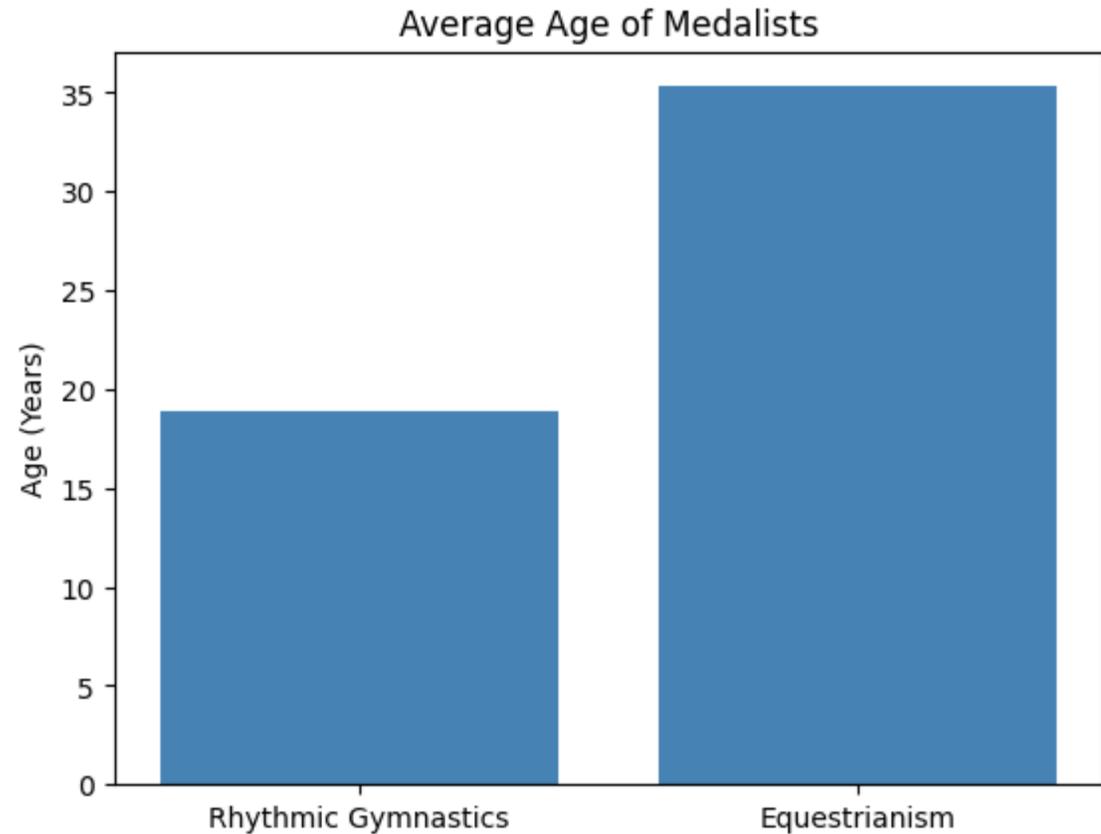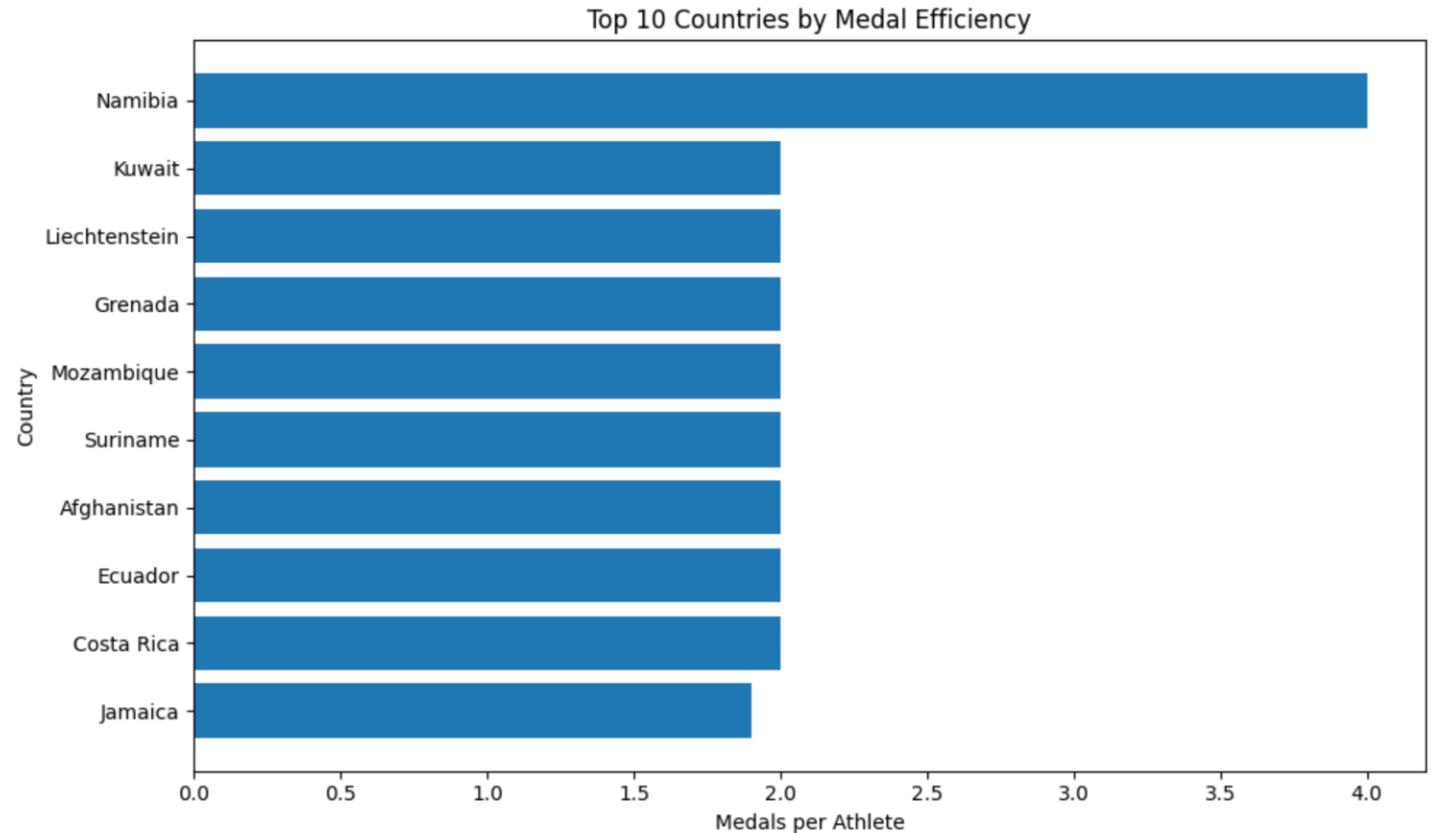
Correlations

Age vs Height:

Pearson correlation: 0.14 → weak positive.

Age vs Weight:

Pearson correlation: 0.21 → weak positive.

New Discoveries & Aha! Moments

• Smaller countries like Namibia and Kuwait stand out for medal efficiency - earning multiple medals per athlete despite limited teams.

Top 10 Countries by Medal Efficiency

Insights & Outcomes

Hypotheses 1–3 all supported.

Small countries can be medal efficient.

Gender participation rising → more balanced representation.

Recommendations & Next Steps

• Develop data-driven stories that highlight medal efficiency and standout countries with high medals per athlete — e.g., Namibia.

• Create athlete training benchmarks based on physical profiles — use age, weight, and height trends to tailor coaching by sport.

• Build dynamic dashboards (e.g., Tableau, DuckDB dashboards) so sports journalists can track gender trends, event competitiveness, and country performance in real time.

Thank You