Analyzing the International Collegiate Programming Contest (ICPC) from 1999 to Present

Brayan Duran Medina

Bachelor of Computer Science


icpc
International Collegiate
Programming Contest
icpc.foundation
advancing the art and sport
of competitive programming

What is the International Collegiate Programming Contest (ICPC)?

- Global competitive programming contest for university students.
- Founded in 1970 at Texas A&M University.
- Teams of three solve complex problems in five hours.
- Languages used: C, C++, Java, Ada, Python, Kotlin.
- Regional contests lead to the annual World Finals.
- Participants from over 3,000 universities in 111 countries.

Motivation for Choosing This Project:

- Competitive programmer since 2016.
- Passionate about problem-solving and programming contests.
- Achievements: 16th place in 2022 ICPC SWERC, Italy; Gold Medal in 2020 ICPC Latin American Regional Contest.
- Participation: IOI 2019 in Azerbaijan; IOI 2018 in Japan.
- Multiple medals in national and regional programming contests since 2017.

Objectives for the Project:

1. Analyze the growth in the number of participating teams and countries over the years.

Objectives for the Project:

1. Analyze the growth in the number of participating teams and countries over the years.
2. Rank the performance of universities, countries, and continents in the ICPC.

Objectives for the Project:

1. Analyze the growth in the number of participating teams and countries over the years.
2. Rank the performance of universities, countries, and continents in the ICPC.
3. Use linear regression to examine performance trends of countries.

Objectives for the Project:

1. Analyze the growth in the number of participating teams and countries over the years.
2. Rank the performance of universities, countries, and continents in the ICPC.
3. Use linear regression to examine performance trends of countries.
4. Analyze the correlation between competition hosting countries and the average ranking of participating countries.

Data Source:

- **Dataset Origin:**
  - Kaggle: "ICPC World Finals Ranking Since 1999"
  - Kaggle Dataset Link: https://www.kaggle.com/datasets/justinianus/icpc-world-finals-ranking-since-1999

Data Quality:

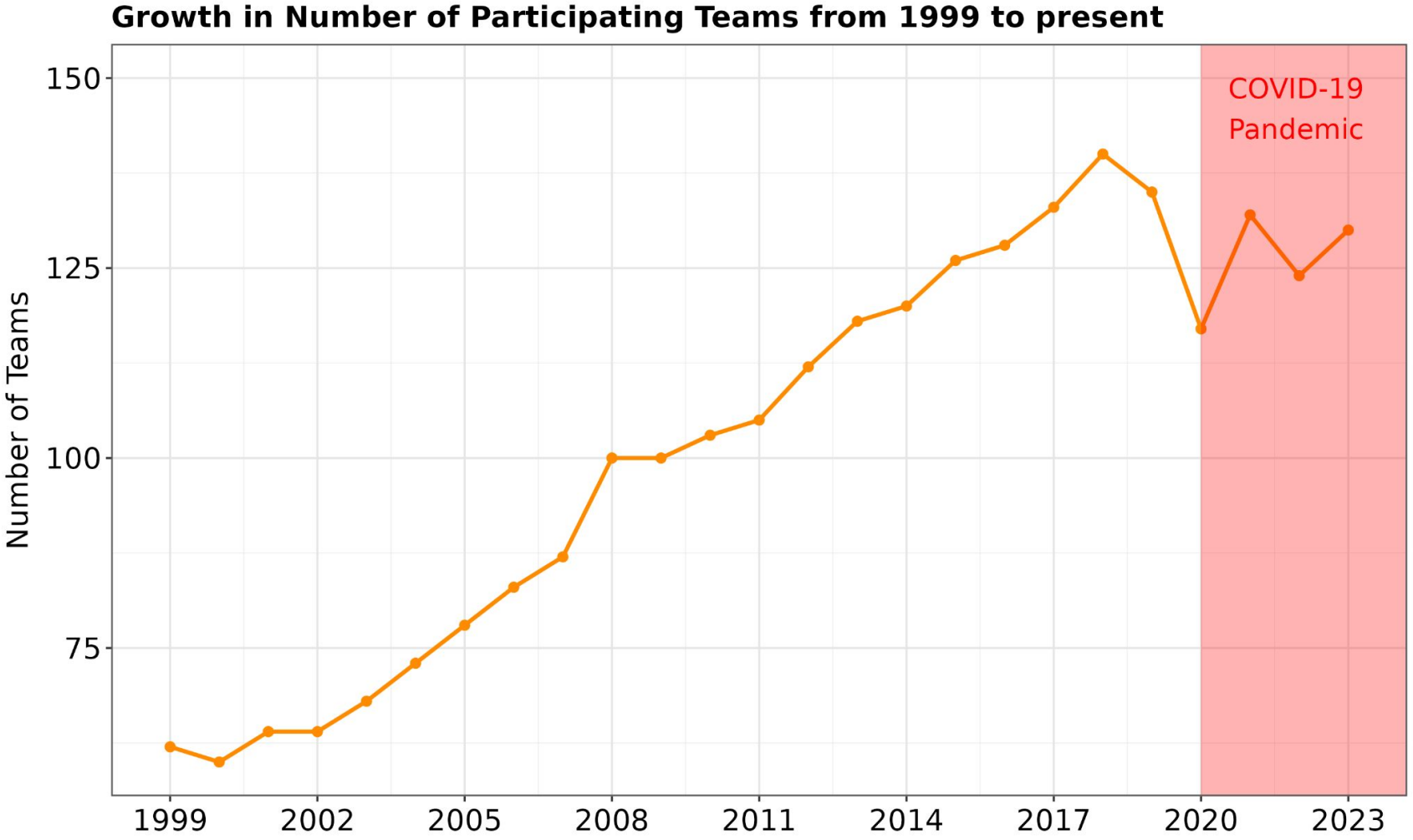
- **Missing Data:**
  - Some missing values in earlier years.
  - Missing fields: team members' names, scores, percentages, penalties.
- **Coverage:**
  - Data from 1999 to present.
  - Comprehensive coverage for available years.
- **Country Name Mismatches:**
  - Some country names did not match plotting library conventions, requiring manual correction.

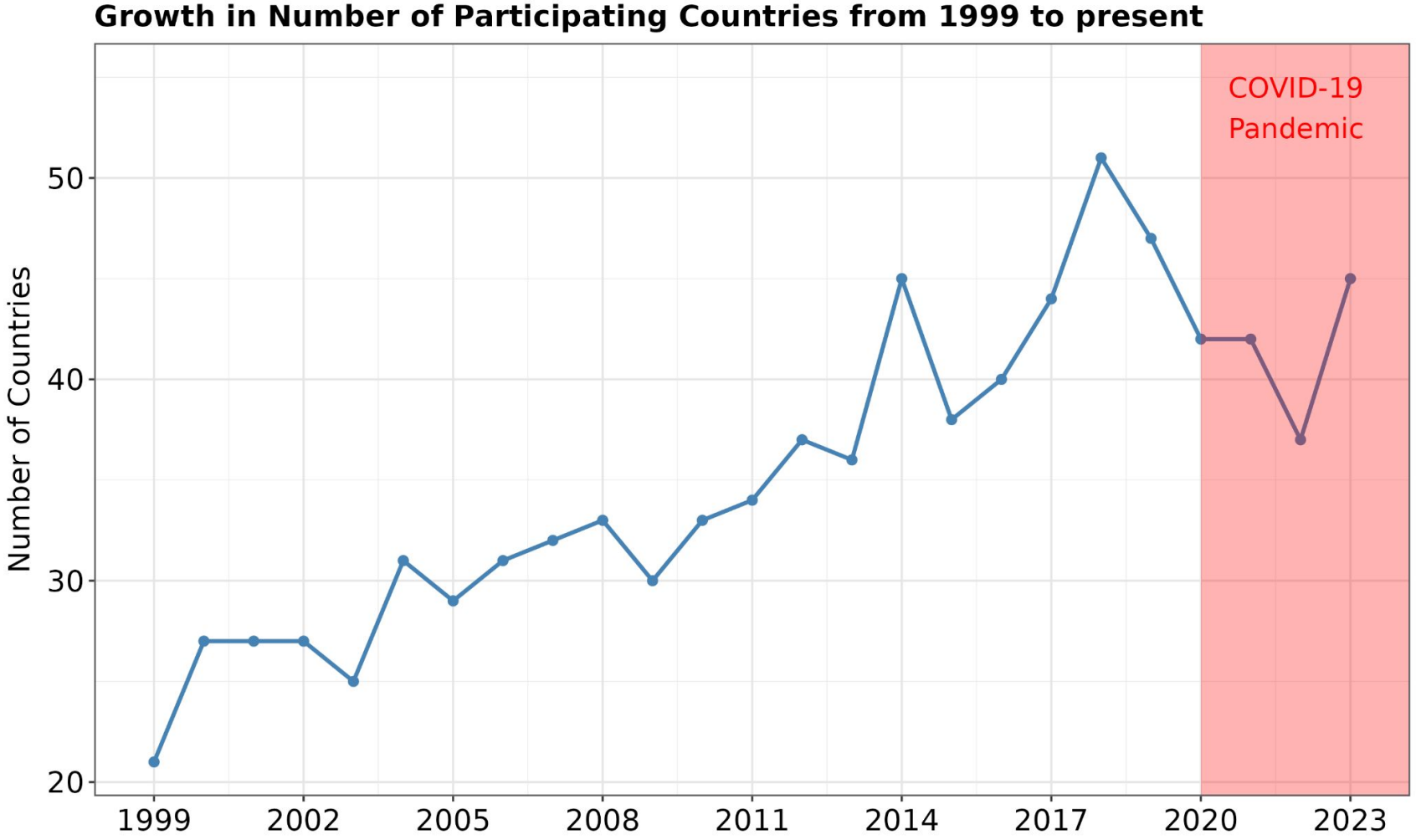
Handling Data Quality Issues:

- **Data Handling and Exclusion Strategy:**
  - Excluded records with missing scores and non-essential fields.
  - Standardized formats; focused on essential metrics.
  - Grouped by universities, countries, continents.
  - Corrected mismatched country names.
- **Assessment:**
  - Data is robust and suitable for trend and performance analysis.
  - Provides a substantial view of ICPC World Finals history.
  - Despite some missing data, offers meaningful insights.

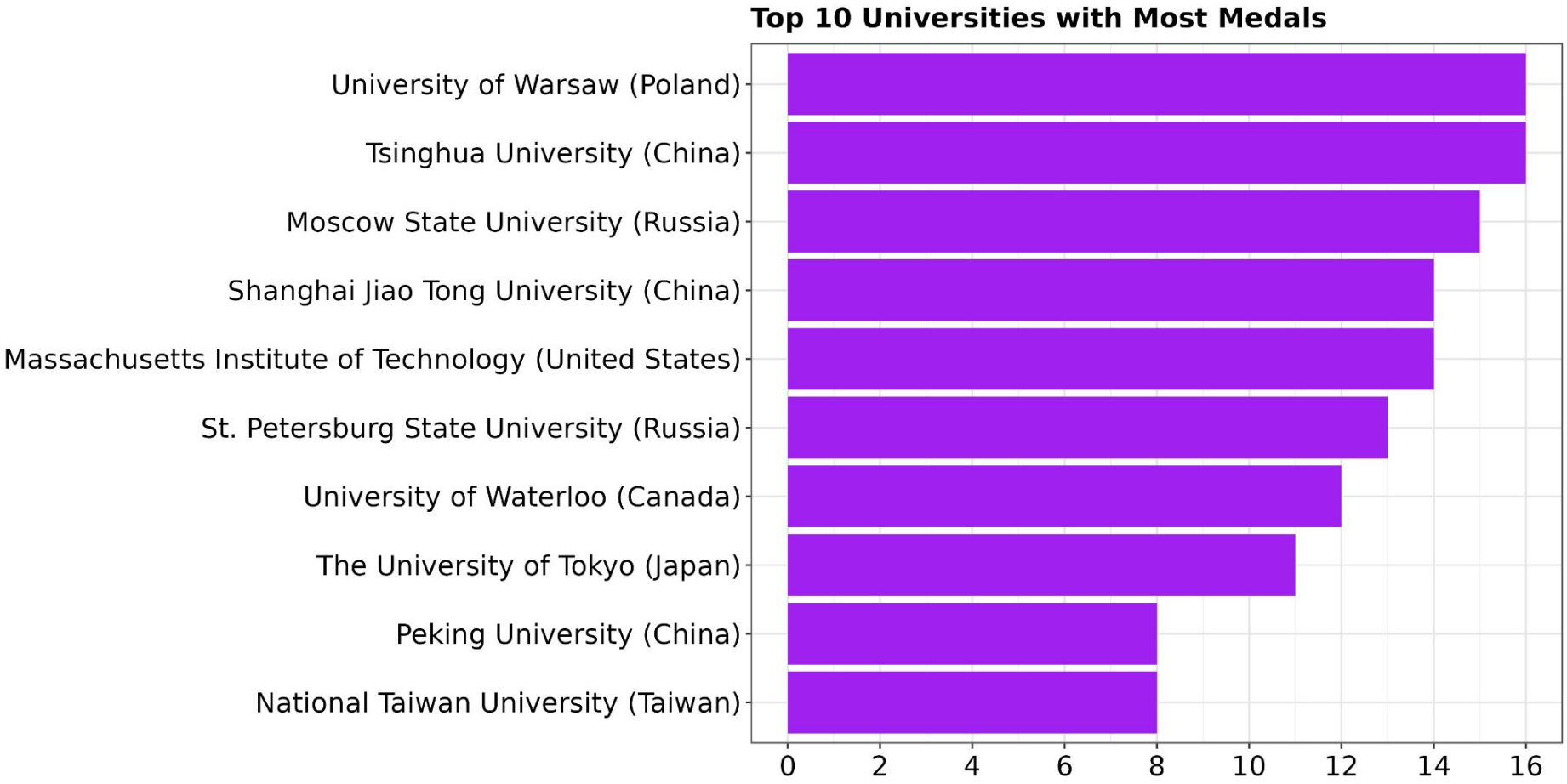
Results

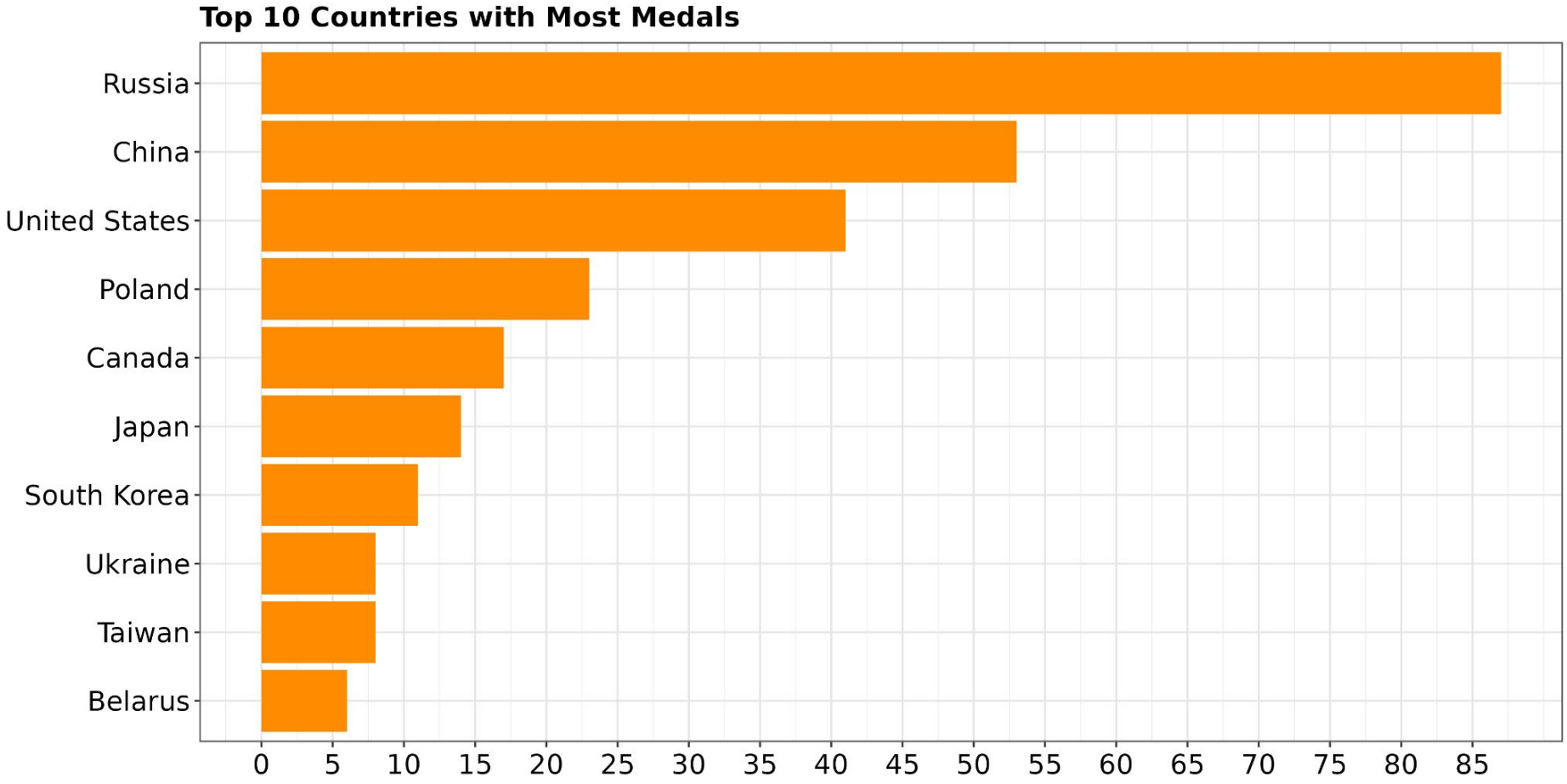
Growth of teams and countries over the years

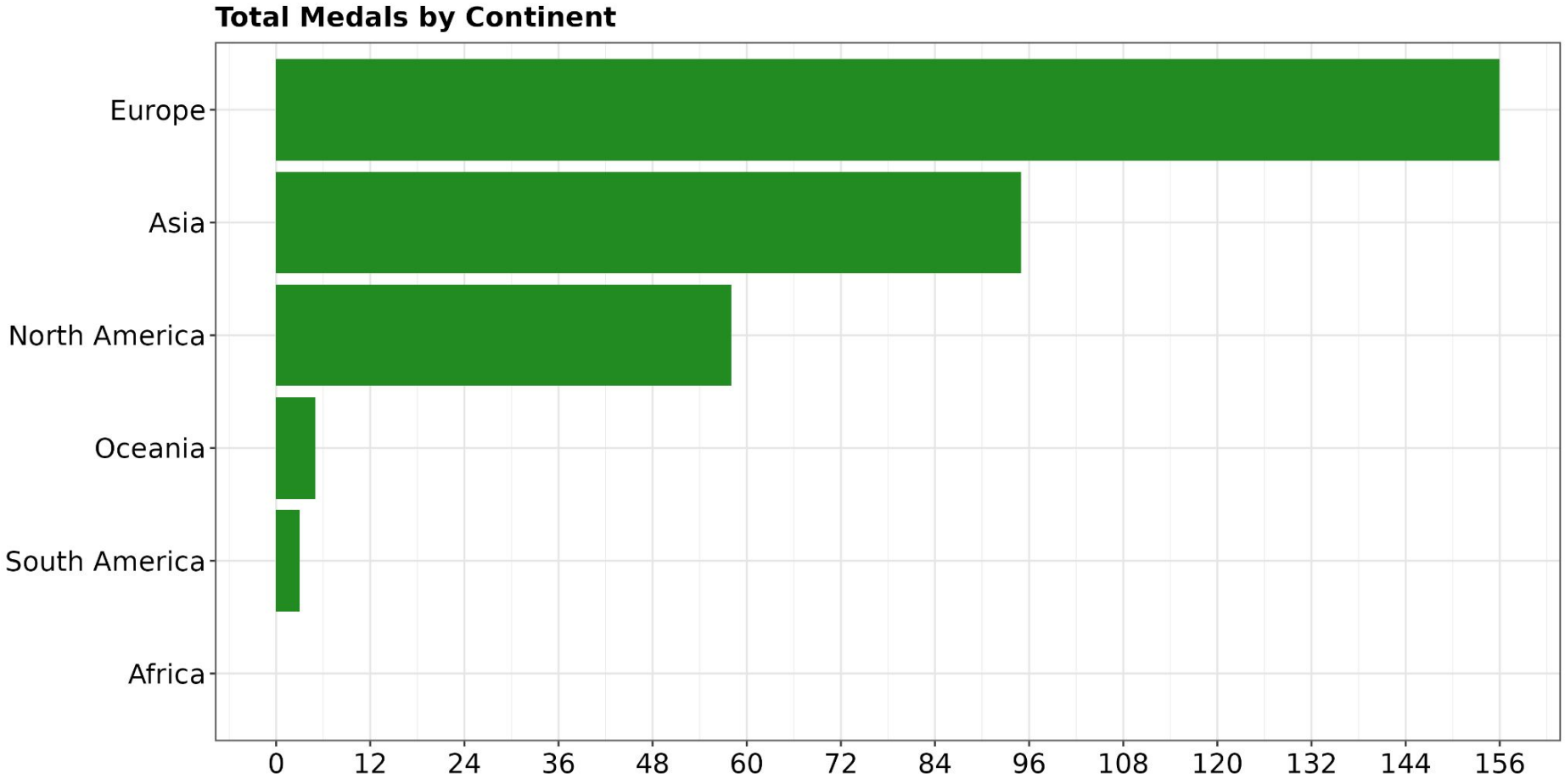

Growth in Number of Participating Teams from 1999 to present
Number of Teams
150
125
100
75
1999
2002
2005
2008
2011
2014
2017
2020
2023
COVID-19
Pandemic


Growth in Number of Participating Countries from 1999 to present
COVID-19
Pandemic
Number of Countries
50
40
30
20
1999
2002
2005
2008
2011
2014
2017
2020
2023

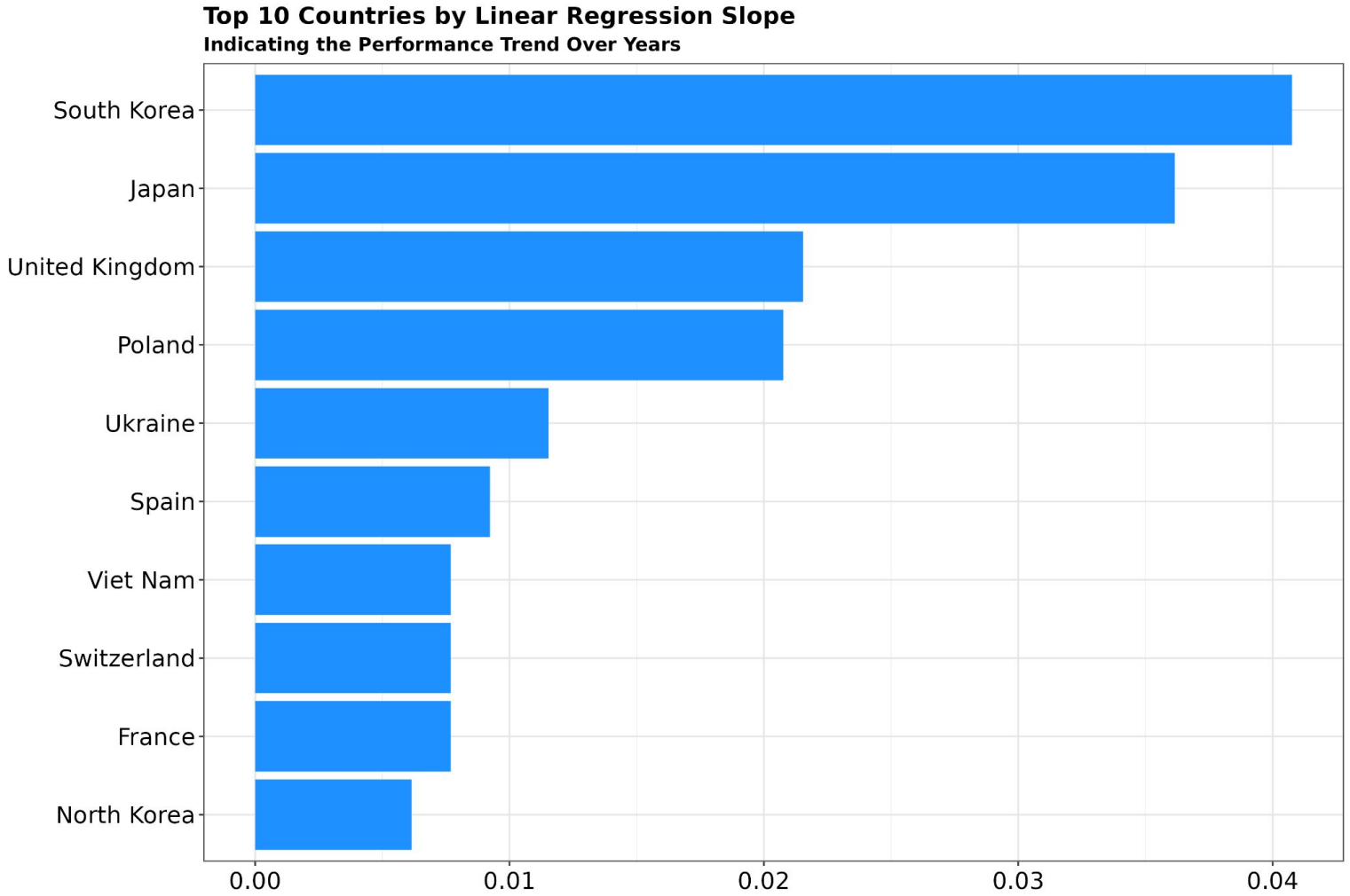
Ranking of universities, countries and continents by medals

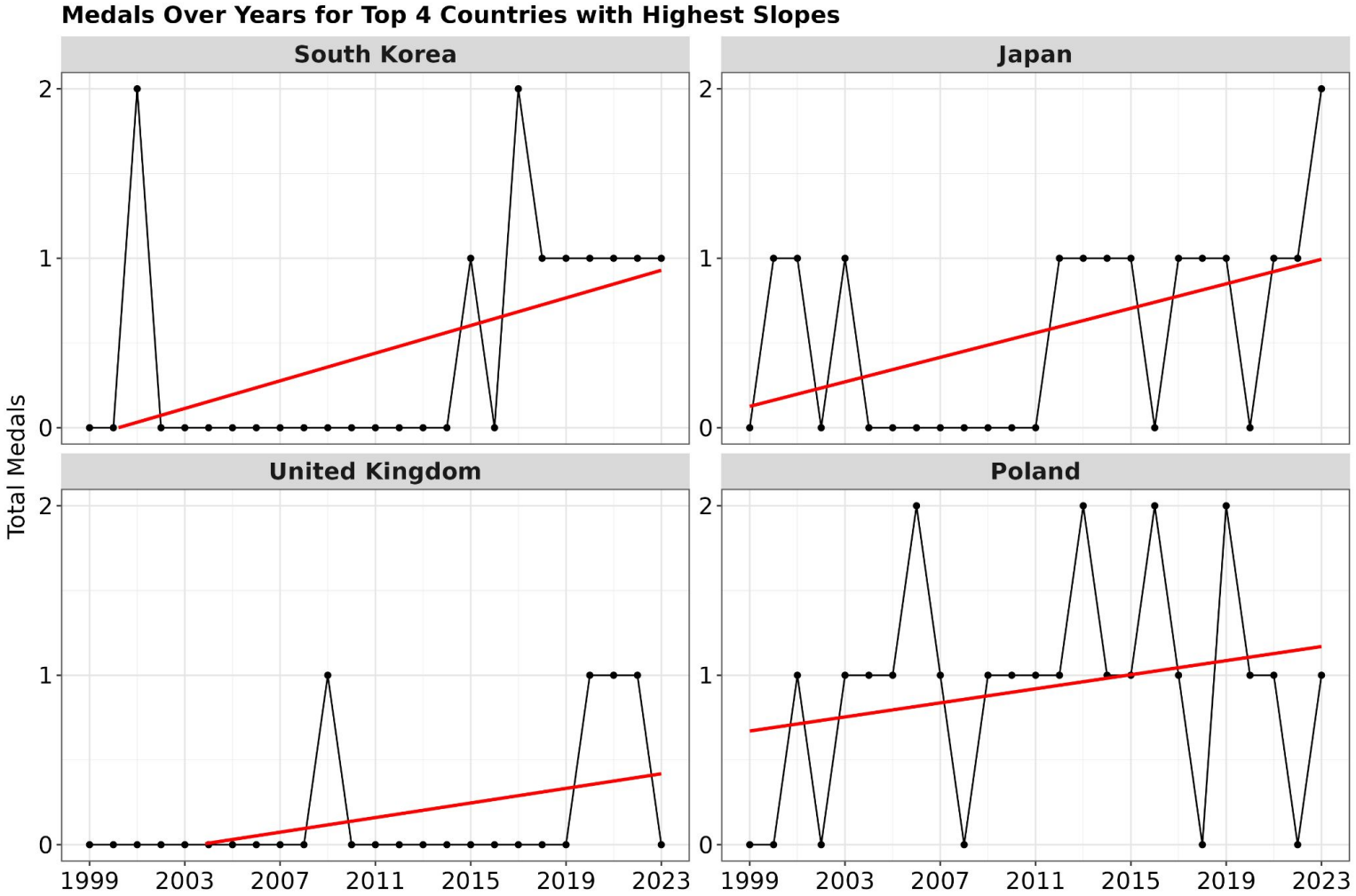

Top 10 Universities with Most Medals
University of Warsaw (Poland)
Tsinghua University (China)
Moscow State University (Russia)
Shanghai Jiao Tong University (China)
Massachusetts Institute of Technology (United States)
St. Petersburg State University (Russia)
University of Waterloo (Canada)
The University of Tokyo (Japan)
Peking University (China)
National Taiwan University (Taiwan)
0
2
4
6
8
10
12
14
16


Top 10 Countries with Most Medals
Russia
China
United States
Poland
Canada
Japan
South Korea
Ukraine
Taiwan
Belarus
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

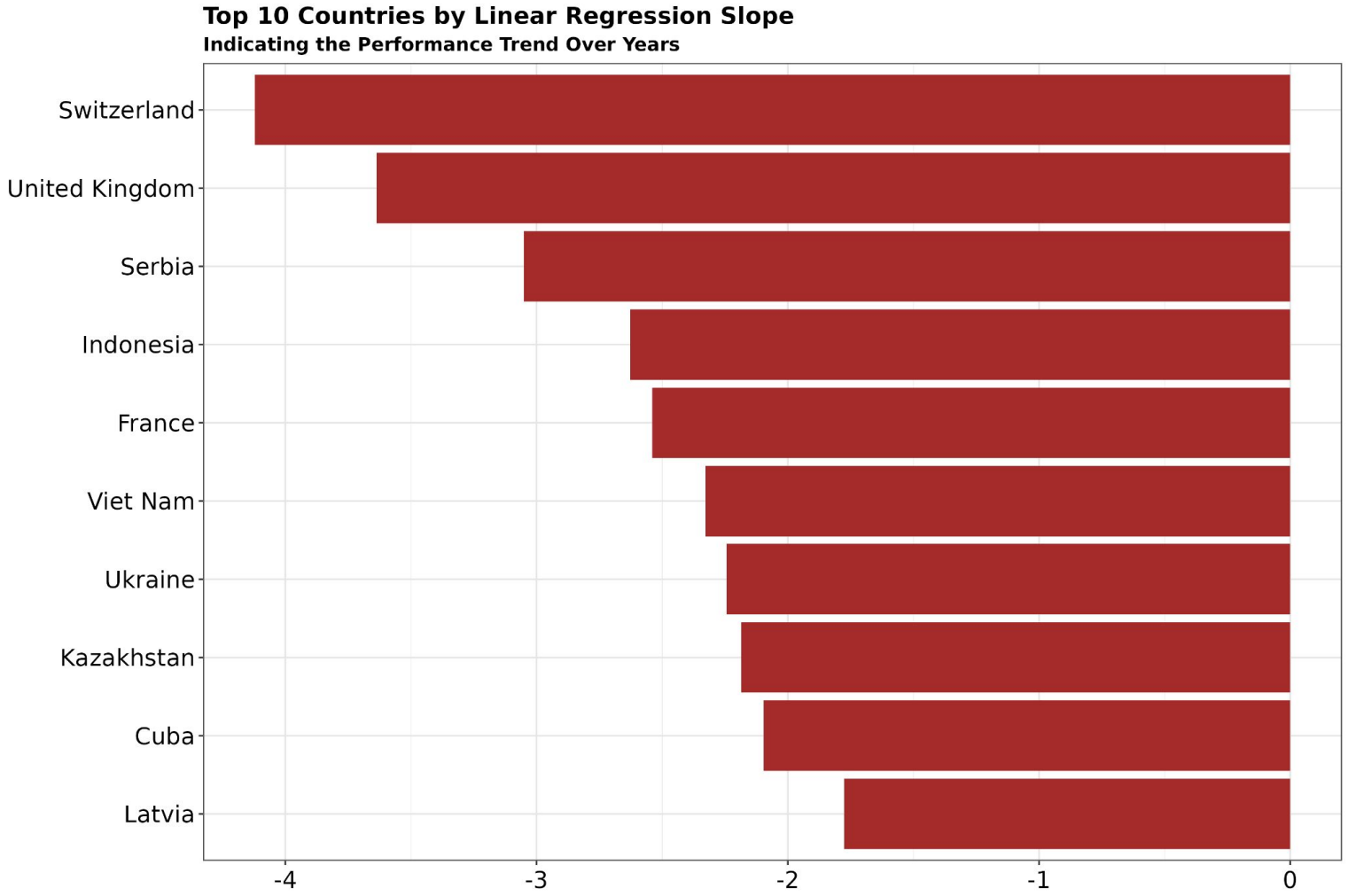

Total Medals by Continent
Europe
Asia
North America
Oceania
South America
Africa
0
12
24
36
48
60
72
84
96
108
120
132
144
156

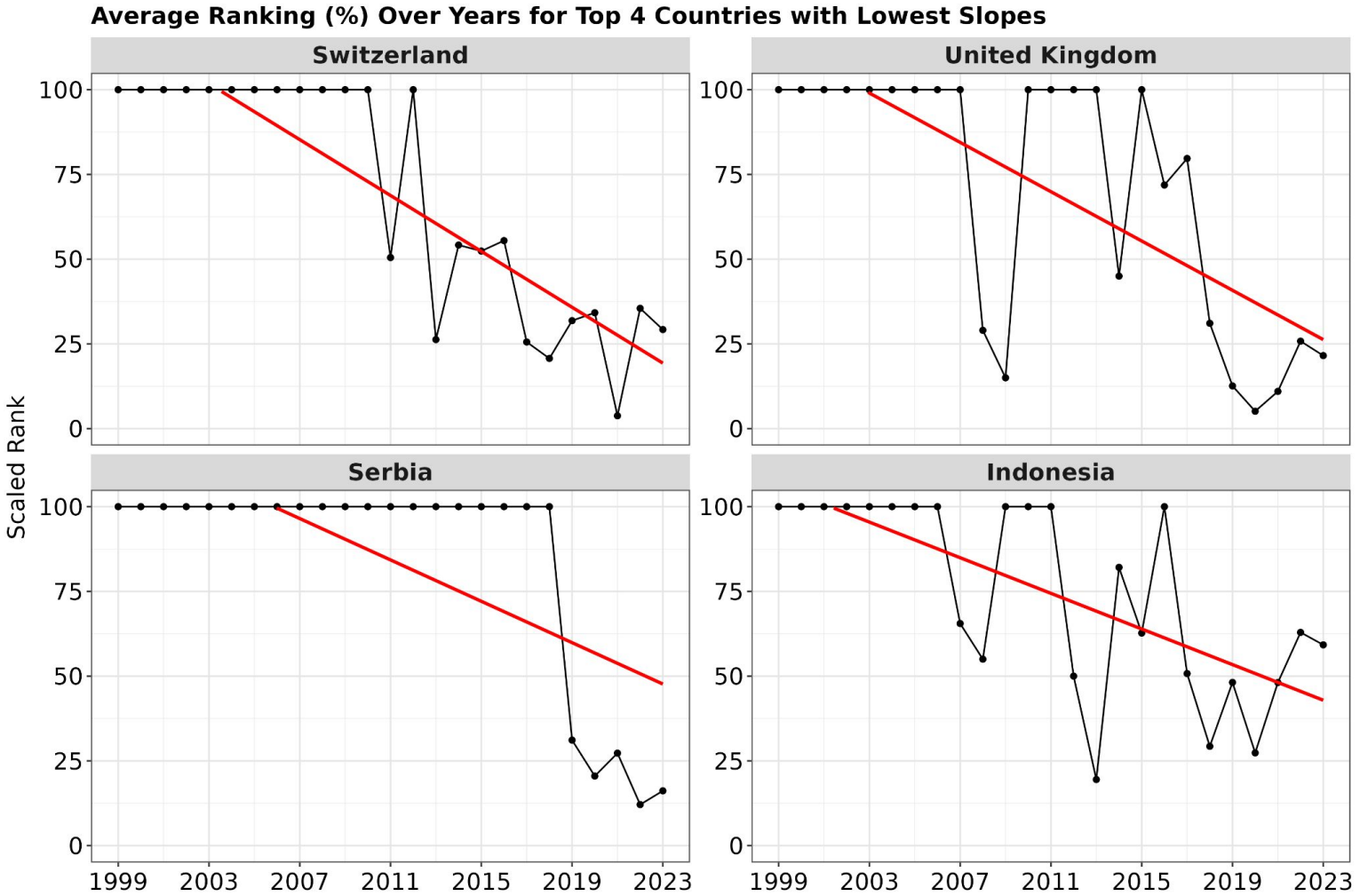
Performance trends of countries by medals

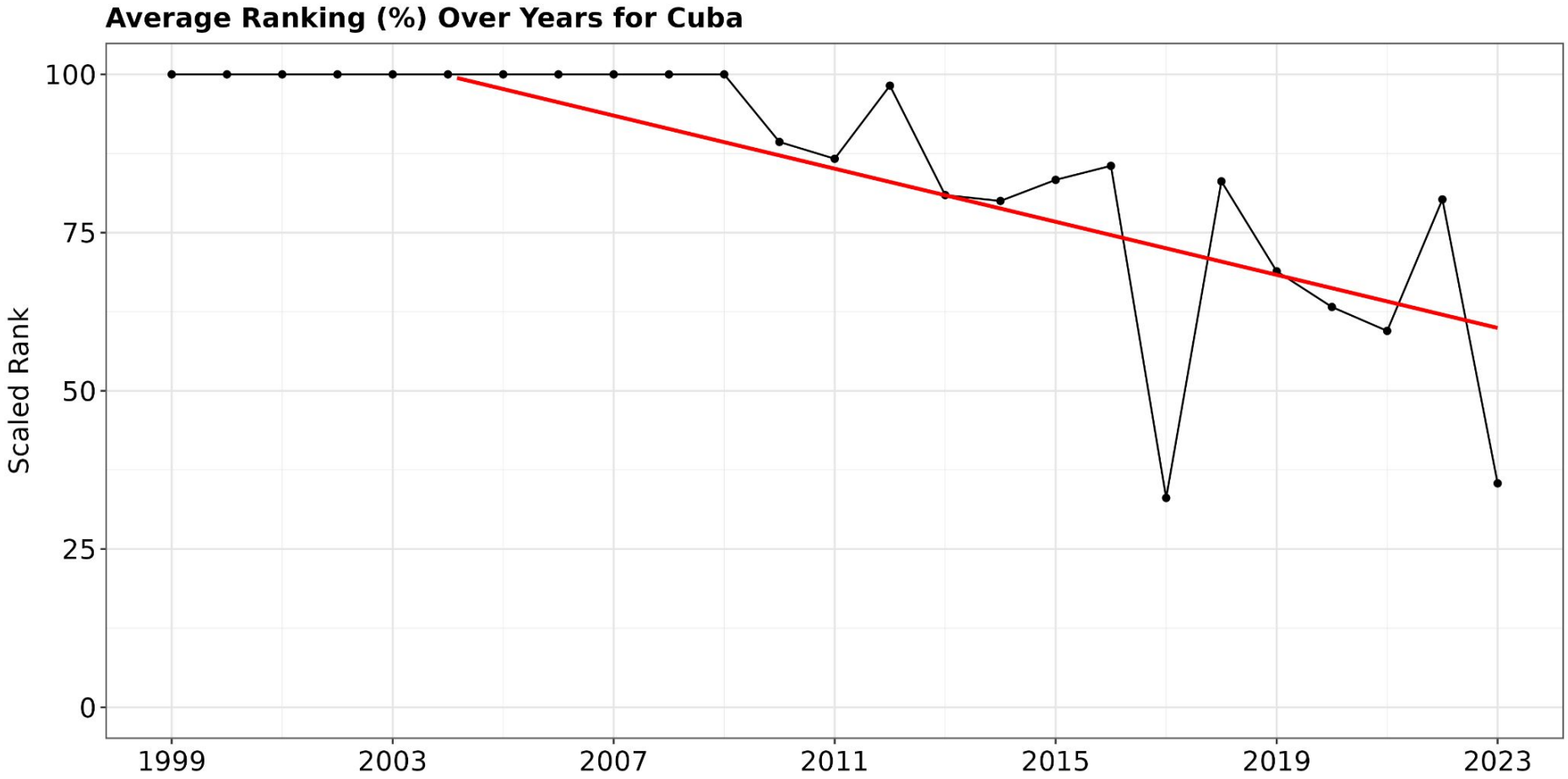

Top 10 Countries by Linear Regression Slope
Indicating the Performance Trend Over Years
South Korea
Japan
United Kingdom
Poland
Ukraine
Spain
Viet Nam
Switzerland
France
North Korea
0.00
0.01
0.02
0.03
0.04


Medals Over Years for Top 4 Countries with Highest Slopes
South Korea
Japan
United Kingdom
Poland
Total Medals
0
1
2
1999
2003
2007
2011
2015
2019
2023

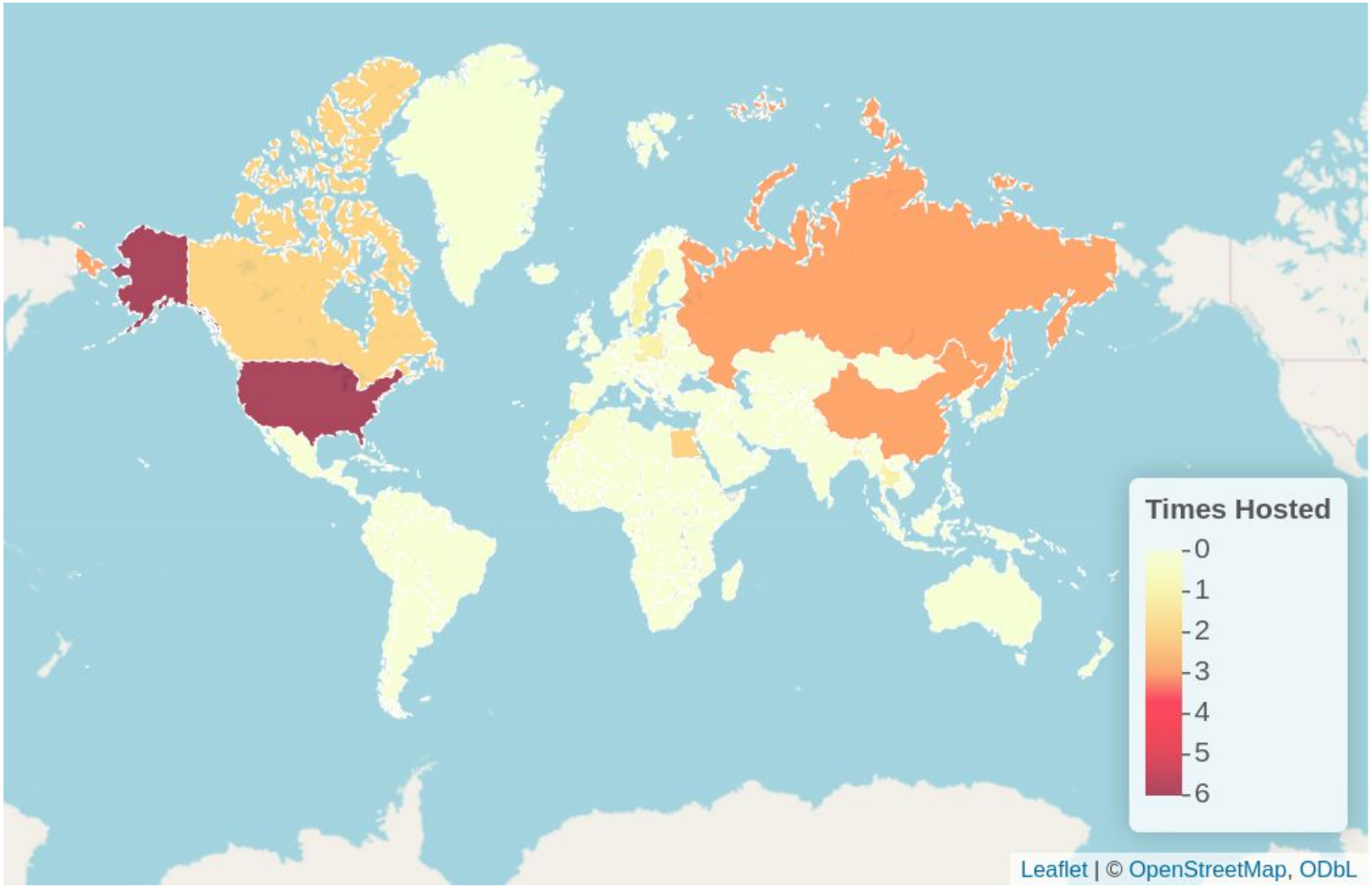
Performance trends of countries by ranking average in %

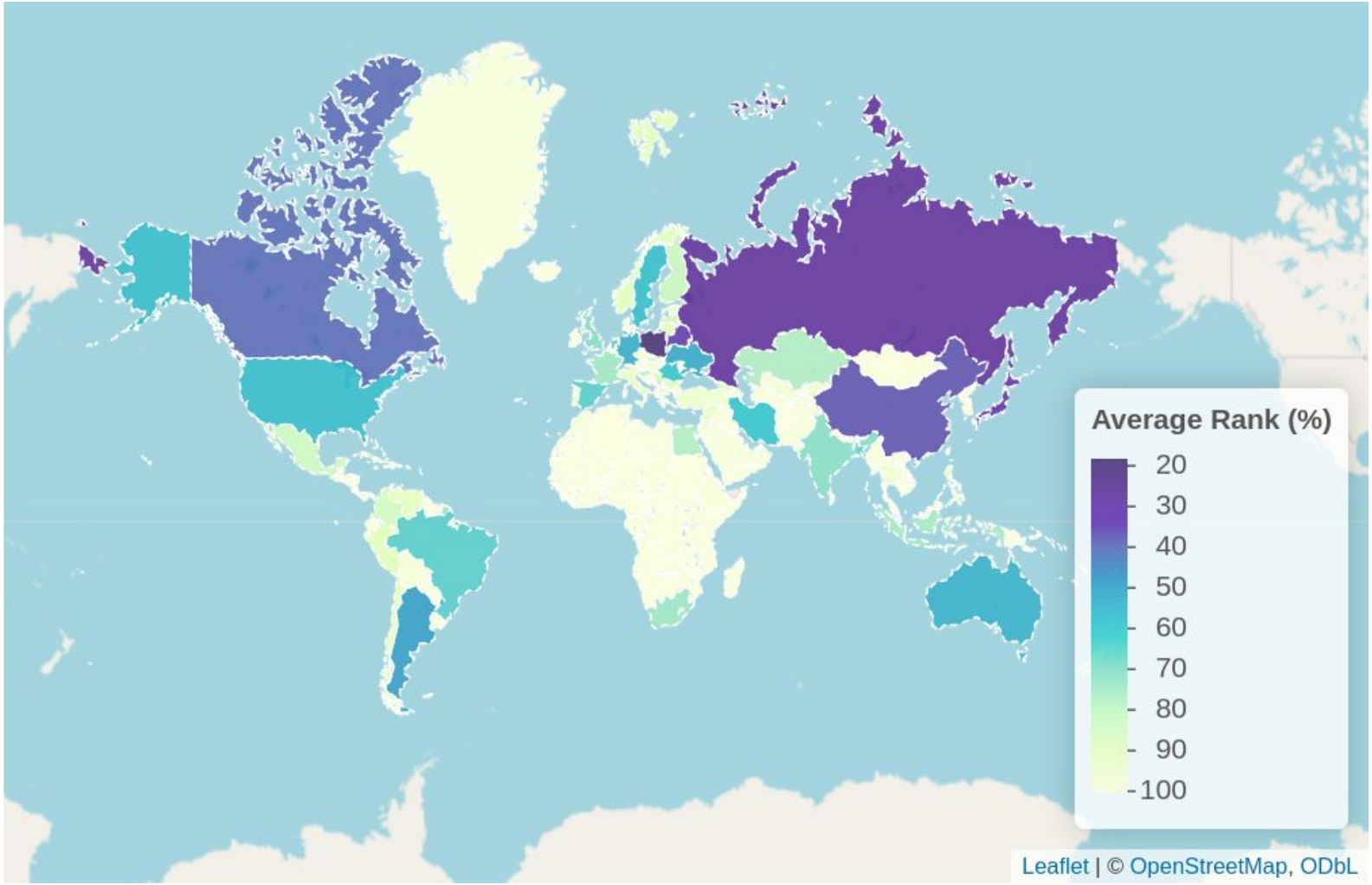

Top 10 Countries by Linear Regression Slope
Indicating the Performance Trend Over Years
Switzerland
United Kingdom
Serbia
Indonesia
France
Viet Nam
Ukraine
Kazakhstan
Cuba
Latvia
-4
-3
-2
-1
0


Average Ranking (%) Over Years for Top 4 Countries with Lowest Slopes
Switzerland
United Kingdom
Serbia
Indonesia
Scaled Rank
100
75
50
25
0
1999
2003
2007
2011
2015
2019
2023


Average Ranking (%) Over Years for Cuba
Scaled Rank
100
75
50
25
0
1999
2003
2007
2011
2015
2019
2023

Correlation between competition hosting and average country ranking in %


Times Hosted
0
1
2
3
4
5
6
Leaflet | © OpenStreetMap, ODbL


Average Rank (%)
20
30
40
50
60
70
80
90
100
Leaflet | © OpenStreetMap, ODbL

Correlation Analysis Between Competition Hosting and Average Country Ranking in %:

- **Correlation Coefficient:**
  - -0.4647881
- **Interpretation:**
  - Indicates a moderate negative correlation.
  - Countries with better average rankings tend to host more competitions.
- **Implications:**
  - Countries with good rankings are more likely to be selected or have the resources to host competitions.
  - There is a noticeable inverse relationship, suggesting that better-ranked countries tend to host more often.

Conclusion

Conclusion:

- **Growth in Participation:**
  - Increasing number of countries and teams over time, with minor declines.
  - Significant drop in participation post-COVID-19.
- **Medal Analysis:**
  - Top 10 universities from top 10 countries.
  - Leading countries are in Europe, Asia, and North America.
  - Europe has the most medals, with Russia leading.

Conclusion:

- **Performance Analysis:**
  - South Korea showed the best improvement in medals.
  - Switzerland had the greatest improvement in rankings.
- **Hosting Correlation:**
  - Moderate negative correlation (-0.46) between hosting frequency and average ranking.

Next Steps:

- **Expand Data Coverage:**
  - Gather and integrate data from earlier years (1970-1998) to provide a more comprehensive analysis.
- **Deeper Regional Analysis:**
  - Conduct detailed studies on performance trends in underrepresented regions.
- **Advanced Statistical Methods:**
  - Use more sophisticated models to explore performance factors and predict future trends.

Thank you!

ICPC International Collegiate Programming Contest

icpc.foundation

advancing the art and sport of competitive programming