ICPC WF Ranking Results (1999 - Present)

Description of the Dataset

The dataset being used for the final project contains information about the ICPC World Finals, which is the international collegiate programming contest. The dataset spans from 1999 to 2023 and includes data on participating teams. There are a total of 2562 rows of data. The columns included in the dataset are:

1. Year: Year of the competition.
2. Date: Specific date of the competition.
3. Host: Host country of the event.
4. City: City where the event was held.
5. Venue: Venue of the competition.
6. Rank: Rank achieved by the team.
7. University: Name of the university.
8. Country: Country of the university.
9. Team: Team name.
10. Contestant 1: Name of the first contestant.
11. Contestant 2: Name of the second contestant.
12. Contestant 3: Name of the third contestant.
13. Gold: Boolean indicating if the team won a gold medal.
14. Silver: Boolean indicating if the team won a silver medal.
15. Bronze: Boolean indicating if the team won a bronze medal.
16. Honorable: Boolean indicating if the team received an honorable mention.
17. Score: Total problems solved by the team.
18. Total: Total number of problems in the contest.
19. Score Percentage: Percentage of problems solved by the team.
20. Penalty: Total penalty points.
21. Prize: Prize received by the team.

Analytical Intentions

The main goal of this analysis is to identify trends and patterns in the ICPC World Finals over the years. Specifically, the analysis will focus on:

1. Performance Trends: Analyzing how the performance of teams (in terms of rank, score, and medal distribution) has changed over time.

2. Geographical Analysis: Understanding the distribution of top-performing universities and countries.

3. Medal Distribution: Examining the distribution of gold, silver, and bronze medals among countries and universities.

4. Team Composition: Analyzing the composition of teams and any observable trends in the contestants' data.

Analytical Strategy

The strategy for analyzing the data will involve the following steps:

1. Data Cleaning: Ensuring the data is clean and free from errors. This includes handling missing values, verifying data types, and standardizing formats.

2. Descriptive Statistics: Calculating basic statistics to summarize the data and provide an overview of the dataset.

3. Trend Analysis: Using time series analysis to identify trends and patterns over the years.

4. Geographical Analysis: Mapping the performance of countries and universities using geographical data visualization tools.

5. Correlation Analysis: Examining the relationships between different variables, such as the relationship between penalty points and rank.

Planned Visualizations

1. Line Charts: To show trends over time for variables such as the number of teams, scores, and penalties.

2. Bar Charts: To compare the number of medals won by different countries and universities.

3. Pie Charts: To show the distribution of medal types (gold, silver, bronze) among the teams.

4. Histograms: To analyze the distribution of scores and penalties.

5. Scatter Plots: To investigate the relationship between score percentage and penalty points.

Planned Maps

1. World Map: Highlighting the countries that have participated in the ICPC World Finals and their respective performances.

2. Country-Specific Maps: For detailed analysis of top-performing countries, showing the distribution of universities and their performance within those countries.

3. Host Cities Map: A map showing the cities that have hosted the ICPC World Finals over the years.

4. University Distribution Map: Mapping the geographic distribution of all participating universities.

5. Medal Distribution Map: Showing the geographic distribution of gold, silver, and bronze medals across different countries.

6. Performance Heatmap: A heatmap indicating the intensity of participation and performance in different regions over the years.

7. Team Origin Map: Mapping the origins of the top-ranked teams in different years.

By following this structured approach, the analysis aims to provide insightful findings about the ICPC World Finals, contributing to a better understanding of the factors influencing team performances and trends in competitive programming at the collegiate level.