Olympiads in Informatics, 2024, Vol. 18, 33–50

DOI: 10.15388/ioi.2024.03

Trends on Returning Contestants and Geography at the International Olympiad in Informatics

Ethan LEE[1], Tymofii REIZIN[2], Farrell Eldrian WU[3], Filbert Ephraim WU[4]
[1]*Stanford University, Stanford, California, United States of America*
[2]*Charles University, Prague, Czechia*
[3]*Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America*
[4]*Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America*
e-mail: eth127@stanford.edu, timreizin@gmail.com, farrellw@mit.edu, filbertw@mit.edu

Abstract. In this paper, we conduct several statistical analyses of IOI 2011 to 2023 performance data, with a focus on tracking returning contestants and identifying geographic trends. This paper identifies several properties of IOI performance data, such that it has strong internal validity while still being subject to random noise. Visualizations are presented throughout to aid the IOI community's understanding of students' competitive programming progress. Afterwards, the geographical analysis shows that countries' IOI performance is correlated to demographic indicators such as population and the Human Development Index. It is, however, more strongly related to competitive programming interest in the country.

Keywords: global talent development, computer science education, academic competitions.

1. Introduction and Background

A. *Background on the International Olympiad in Informatics*

The International Olympiad in Informatics (IOI) is an annual algorithmic programming competition over eighty countries involved around the world. Based on a local selection process (often a series of competitions and training camps), each country sends a team of up to four pre-collegiate students to the IOI. At the competition, contestants are to attempt six algorithmic tasks, presented in the format of two competition days featuring three tasks each day, typically with one excursion day in between, and with each competition day lasting five hours. The current format, consistent from IOI 2011 until now, weighs each problem equally with a full score of 100 points, for a total of 600 points. Medals are then awarded to approximately one-half of all contestants, with gold, silver, and bronze medals awarded in an approximate ratio of 1:2:3. (IOI Regulations)

As a long-running international event, the IOI has substantial recognition as a venue to identify and develop pre-collegiate students' skills in computer science, with a focus on programming skills and algorithmic problem-solving. (IOI Syllabus) In many countries, the local competitions and training programs leading up to the IOI is seen as a major, and sometimes, the primary opportunity for high school students to explore computer science and related subjects. Numerous universities around the world also consider IOI results in their admissions, scholarship, and course placement processes. Therefore, it is of substantial interest to improve the IOI community's collective understanding of the dynamics of IOI results on a global scale.

The "International Olympiad in Informatics—Statistics" website stores IOI competition data ever since the contest's inception in 1989. The available data is extensive: for competitions starting 2010, the website provides the score for each contestant for each competition task, with contestants' results over multiple IOIs are linked through a personal profile page and a time-indexed "Past Participations" graphic showing progress over the past IOIs of all the contestants at a particular IOI. Other performance-related data such as a contestant's total score, medal, rank, and percentile are also provided, together with the gold/silver/bronze medal cutoffs at each year. Altogether, the historical IOI results provide a starting point for statistical analysis and data visualizations, towards yielding insights on contestants' IOI performance trajectories and countries' aggregate IOI performance.

B. *Research Questions*

This paper focuses on exploratory data analysis, particularly on returning contestants' performance over time and aggregate per-country performance.

For the returning contestant analysis, the objective is to broadly seek an understanding of the overall "IOI experience" of a multiple-time contestant, which in turn uncovers salient insights on the educational role of the IOI in developing talent in computer science. The primary statistical instrument in this analysis is the difference between a single contestant's performance in two consecutive IOIs, aggregated in different ways. This statistic is then analyzed using statistical techniques such as regression analysis to investigate and visualize different trends.

For the country comparison analysis, the objective is to showcase geographic trends in IOI results by identifying a range of indicators that correlate to a country's IOI performance, with the goal of better understanding the global landscape of advanced computer science education. In this analysis, we focus on the aggregate performance of a country over a decade-long period spanning 2011 to 2023 which is then regressed on a specific set of indicators and categorized by geographic region.

C. *Data Processing Procedures*

In this section, we describe the data processing produces used in the paper's analysis, which are aimed to reduce statistical noise caused by inherent variations in the IOI while maintaining having sufficiently relevant data for a meaningful analysis. There are two aspects to this: first, subsetting the data, and second, calculating a standardized measure of contestant performance.

Our analysis uses IOI result data starting from 2011, for data reliability purposes. This is shortly after the IOI adapted its current “subtask” partial credit structure for most tasks. Before then, there were concerns regarding the reliability of the IOI’s partial credit system of the IOI (Verhoeff, 2006); the change in partial credit format can have effects on the overall distribution of IOI results due to different contestant approaches. The 2011 competition is also right after the IOI started publishing complete scores for all contestants rather than only the medalists. (The 2009 and 2010 IOIs also had an experimental 8-problem rather than 6-problem format.) For the country analysis, specifically, we only analyze with at least half participation over the IOIs from 2011 to 2023, as average performance metrics can be skewed by incomplete IOI teams.

As for contestant performance, we develop a scaled metric designed for consistency in interpretation across years, due to limitations in the comparability of the numerical IOI score (out of 600) across years. Specifically, using the numerical IOI score is extremely susceptible to changes in the contest’s difficulty and subtask structure. While the percentile ranking properly adjusts for difficulty, this metric is still sensitive to changes in the shape of the distribution of scores and does not distinguish between performance at the highest levels. (In 2023, for example, the entire gold medalist range, spanning over 200 points, spans fewer than ten percentile points.)

To alleviate the above concerns, we calculate a “scaled score” that is a piecewise linear interpolation. In our scale, we use a piecewise linear mapping where the lowest score is mapped to 0, the highest score is mapped to 6, and the bronze, silver, and gold cutoffs are mapped to 2, 3, and 4, respectively. This mapping can be interpreted as a reference to an “ideal IOI” where the bronze, silver, and gold cutoffs are at 200, 300 and 400 points, respectively, which is, after rounding to the nearest hundreds, the average cutoffs from 2011–2023. In addition, we assign the 25^{th}-percentile (middle of the “no medal” group) score to a scaled score of 1, and the 96.67^{th}-percentile (middle of the “gold medal” group) score to a scaled score of 5. This calculated statistic will hereafter be called the “scaled score,” which is different from either the percentile (out of 100%) or the raw score (out of 600).

2. Returning Contestant Analysis

In this section, we present the analysis of returning contestants at the IOI, along two aspects. First, we visualize the joint distribution of a returning contestant’s performance and improvement across multiple IOIs; from here, we draw inferences that characterize the nature of improvement in competitive programming at different skill levels. Second, we describe progress of contestants with multiple IOIs to better understand the nature of long-term learning in competitive programming. In doing so, we also demonstrate the “reversion to the mean” phenomenon common in statistics (Barnett et. al, 2005) as seen in IOI contestant data.

In total, there are 1038 participants with more than one IOI from 2011 to 2023, which provides ample data for the analysis in this section. The trends shown are mostly rather strong and unlikely to be affected by random noise in the data. The selection of this set of recurring contestants indeed may be biased due to different team selection policies across countries and other confounding factors, though we believe that statistical properties depicted are still meaningful. The analysis presented in Figures 4–6 also quantify and contextualize such bias.

A. *Cross-distribution of Performance Across Multiple IOIs*

As an initial analysis, we visualize the scatterplot (Fig. 1) of a multiple-IOI contestant's performance over consecutive IOIs. For this analysis, we use the "scaled score" from 0 to 6, as described previously. So that the datapoints within each scatterplot are comparable to each other, we show separate plots depending on whether we are comparing the first/second or second/third attempts, and whether the contestant had two or 3+ IOIs. We do not analyze IOIs after the third due to small sample sizes in this group. These data divisions were chosen to balance each plot having sufficient data and a useful interpretation.

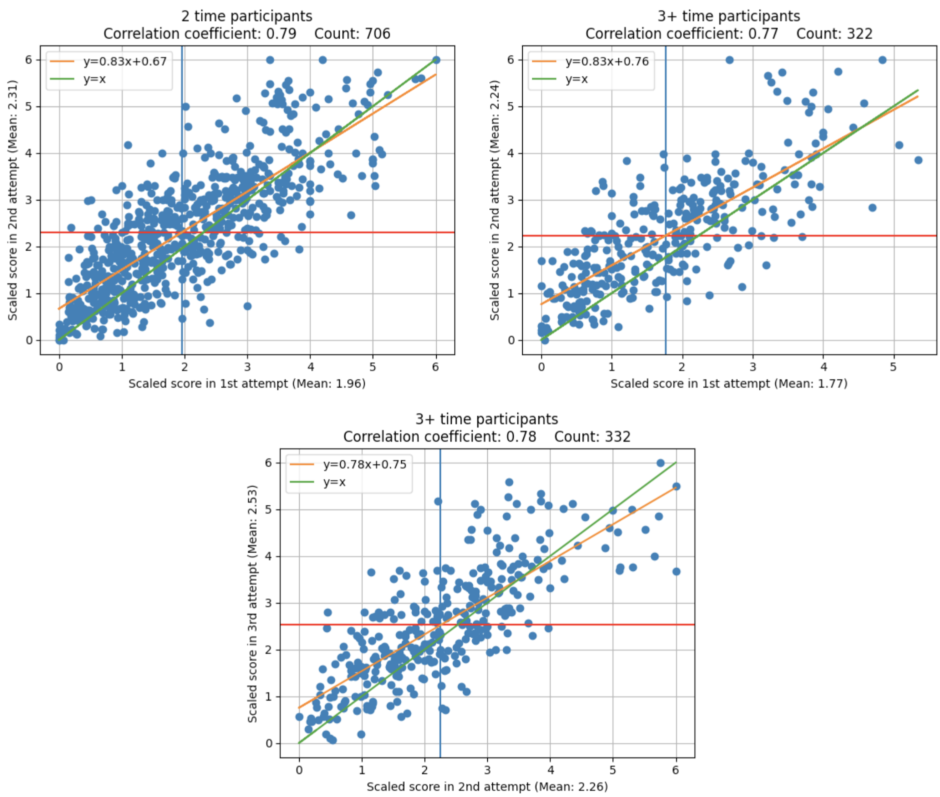


Fig. 1. Scatterplot of returning contestants' performance in consecutive IOIs.

As seen in Fig. 1, a strong correlation is seen across the three scatterplots. The correlation coefficients, a measure of statistical association between the two variables, are calculated to range from 0.77 to 0.79. It is worth nothing, however, that the data for three-or-more-IOI contestants have a greater proportion of outliers at the upper-right corner of the scatterplot, indicating less predictability for high-performers in this group.

The bivariate regression line (a least-squares-optimal prediction of the later attempt's scaled score based on the previous attempt's scaled score) in each scatterplot is plotted in yellow, while the 45-degree y=x line is plotted in green. The regression coefficients are shown to range from 0.78 to 0.83, again indicating strong dependence across IOIs participated. Two features of the scatterplots indicate a general trend of improvement over consecutive IOIs. First, the regression line, for most of the corresponding datapoint x-values, is above the 45-degree line, and second, most of the datapoints in the scatterplot are above the 45-degree line.

The strong correlations in Fig. 1 imply that a contestant's performance at one IOI has substantial predictive power towards performance at the next IOI. It is unclear, however, from the scatterplot whether a single linear trend is appropriate for all IOI performance levels, given the vastly different nature of crucial tasks encountered at different levels. Therefore, we plot a linear spline regression (Marsh and Cormier, 2001) with knots at the integer scaled scores. This method allows flexibility in the regression plot while ensuring continuity of the predicted value and avoiding overfitting from higher order regressors. The results are shown in Fig. 2, where the centroids of the datapoints in each interval are also corresponded to both axes.

From Fig. 2, a positive slope within each scaled score interval is consistently seen, showing that variations of performance, even those within a medal category, yield predictive power towards future IOI performance. The slopes, however, vary substantially. For example, the slope in the [0, 1) scaled score interval (corresponding to a previous IOI performance lower than the 25th percentile of that year) tends to be substantially larger than all other slopes in the same spline. This observation may be interpreted as that having a baseline level of programming and algorithmic proficiency, up to the point of solving the easiest set of IOI subtasks, is crucial towards improving further towards the IOI medal level.

Further examining the scaled score range of [0, 2) in the previous IOI, corresponding to participants who fell short of winning a medal, we note that those with a scaled score in [0, 1), corresponding to a performance below the 25th percentile, are very unlikely (under 10% chance) to win a medal at the following IOI. Meanwhile, those with a scaled score [1, 2), corresponding to a performance at least the 25th percentile, have roughly even odds to do so. Specifically, the regression spline consistently predicts a scaled score of 1.50 (typically between the 35th to 40th percentiles) to correspond to a roughly a bronze medal cutoff performance at the following IOI.

The improvement in the 25th to 50th percentile bucket of the previous IOI, however, is not typical among medalists. Around 40% of bronze medalists achieve a silver medal or better at the next IOI, with a regression spline predicting a middling bronze medalist

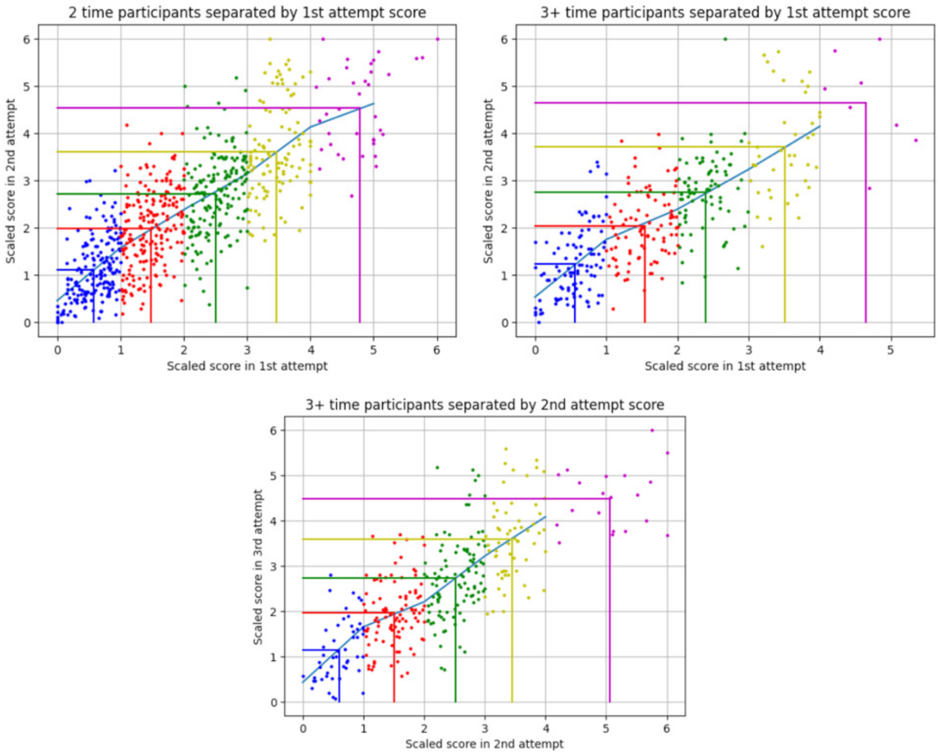


Fig. 2: Regression splines of scaled scores in consecutive IOI attempts

(with a scaled score of 2.50) to achieve an expected scaled score of around 2.80 at the next IOI, short of the silver cutoff of 3.00. Among silver medalists, around 35% manage to improve to a gold medal at the following IOI, while around 20% drop to bronze; meanwhile, very few bronze medalists (under 5%) improve to a gold medal, indicating a large skill gap between bronze and gold medalists.

Next, we consider the distribution of improvement across consecutive IOIs. In this analysis, we divide the dataset in scaled score intervals of size 0.50 for all contestants up to the silver medal level (scaled score of 4.00), plus a single interval for the gold medalists. This data division is chosen to balance having an ample number of datapoints in each bucket while having sufficiently many buckets to demonstrate overall trends. For each bucket, we produce a quantile plot of the scaled score improvement, featuring endpoints indicating the 10th and 90th percentiles, a box spanning the 25th to 75th percentiles, and a red line at the median. The plot is shown in Fig. 3, and a summary of the percentiles is provided in Table 1.

The quantile plot and Table 1 show a general trend of the average improvement decreasing as the initial IOI performance increases, which may be due to the "regression to the mean" phenomenon that is elaborated in the next subsection. A notable exception, however, is the lowest performance interval [0, 0.5), corresponding to contestants

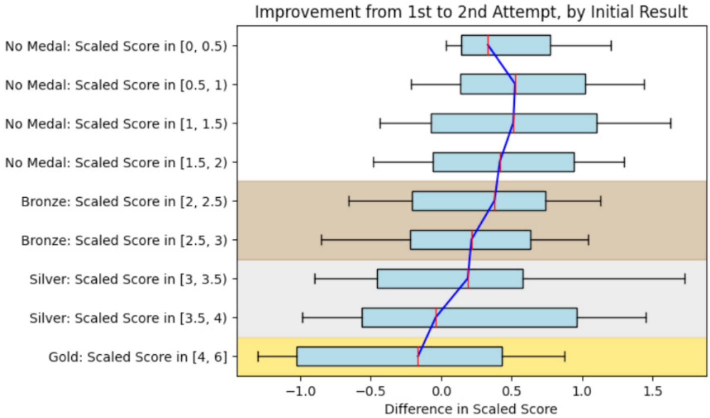


Fig. 3. Quantile plot of improvement across consecutive IOIs, separated by initial result bucket.

Table 1

Summary statistics of scaled score improvement across consecutive IOIs

1st Attempt Result	count	mean	std	10%	25%	50%	75%	90%
No Medal: Scaled Score in [0, 0.5)	112	0.52	0.53	0.04	0.14	0.33	0.77	1.21
No Medal: Scaled Score in [0.5, 1)	171	0.61	0.67	−0.22	0.13	0.52	1.02	1.44
No Medal: Scaled Score in [1, 1.5)	151	0.57	0.8	−0.44	−0.07	0.51	1.1	1.62
No Medal: Scaled Score in [1.5, 2)	145	0.45	0.71	−0.49	−0.06	0.42	0.94	1.31
Bronze: Scaled Score in [2, 2.5)	143	0.3	0.79	−0.68	−0.21	0.38	0.74	1.13
Bronze: Scaled Score in [2.5, 3)	112	0.21	0.82	−0.9	−0.22	0.22	0.63	1.05
Silver: Scaled Score in [3, 3.5)	78	0.24	1.01	−0.9	−0.46	0.19	0.58	1.76
Silver: Scaled Score in [3.5, 4)	61	0.12	0.95	−0.99	−0.56	−0.04	0.96	1.45
Gold: Scaled Score in [4, 6]	54	−0.22	0.92	−1.3	−1.03	−0.16	0.43	0.89

with less than half the score of the 25th percentile contestant. Contestants in this bucket have a lower median scaled score improvement (0.33) than all other no-medal buckets (ranging from 0.42 to 0.51), again indicating the difficulty of improvement among the lowest-performing IOI contestants which may be due to lacking rudimentary programming skills to fully engage in IOI preparation.

Besides median performance, quantile plots allow analysis of variance and skew. Overall, the variance in improvement tends to be large, with the 10th and 90th percentiles spanning more than one, and often two, scaled score points, which correspond to a medal range.) We also notice that the amount of improvement tends to be more variable among medalists than non-medalists at the previous IOI, with silver and gold medalists having the largest variance. Among non-medalists, those above the 25th percentile have a higher variance than those below the 25th percentile due to being more equipped to achieve higher results after a year of preparation. As for skew, there is no discernible trend that is not attributable to noise in the data or the scaled score having a zero lower bound, such as the lowest performance bucket being skewed to the right.

B. *Progress of Multiple-IOI Contestants Over Time*

In this subsection, our analysis concerns understanding the progress of contestants over multiple (two or more) IOIs. As a first step, we visualize the starting profile of IOI contestants with at least two IOIs. Fig. 4 plots the distribution of first-IOI results of these contestants, separated by the number of IOIs attended and capped at four. The distribution is shown as a cumulative histogram on proportion of contestants in the bottom quartile, second quartile, bronze medal, silver medal, and gold medal ranges. These cumulative proportions are shown relative to the actual proportions over all contestants in a typical IOI.

The visualization in Fig. 4 shows that the starting result of a two-time IOI contestant has roughly similar, though slightly worse, performance distribution as the typical IOI contestant. For contestants with more than two IOIs attended, however, the starting results become increasingly skewed towards lower result, with the proportion in the silver and gold ranges combined being under two-thirds the usual proportion. This observation is likely due to the country-based qualification process of the IOI, as it's easier for a contestant to qualify when faced with a weaker pool of contestants in higher cohorts. Note, however, that the set of 4+-time participants is too small to draw statistically-valid comparisons.

Next, we consider tracking the distribution of a contestant's progress over time. To do so, we consider two plots. In Fig. 5, the contestants are grouped by total number of attempts (capped again at four), then the mean scaled score of the contestant over multiple attempts are tracked on a path with arrows to compare the magnitude of progress over time. Fig. 6 then shows quantile plots separated by IOI order (first, second, or IOI order), with a similar $10^{th}/25^{th}/50^{th}/75^{th}/90^{th}$-percentile scheme used as in Fig. 3, allowing for comparison relative to the uniform distribution of percentiles when considering all IOI contestants. The raw data in Fig. 5 is also presented in Table 2 for ease of reference.

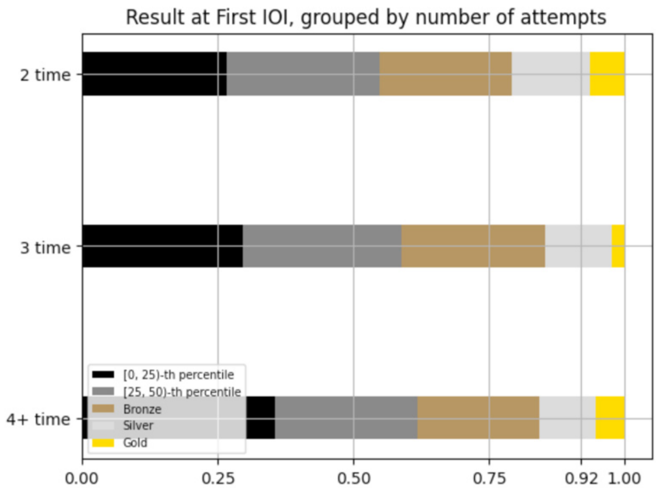


Fig. 4. Histogram of first-IOI result for multiple-IOI contestants, grouped by number of IOIs.

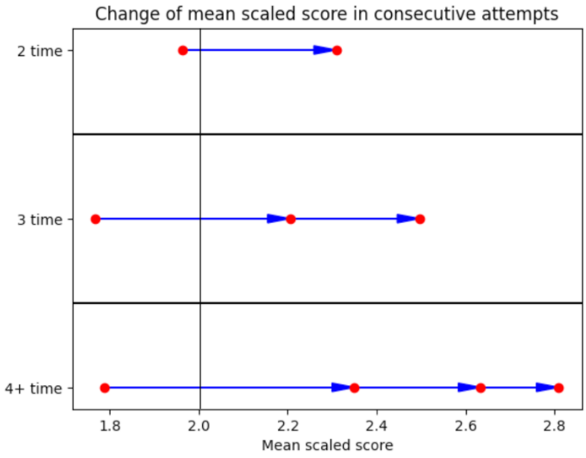


Fig. 5. Progression of mean scaled score in consecutive IOIs, separated by number of IOIs.

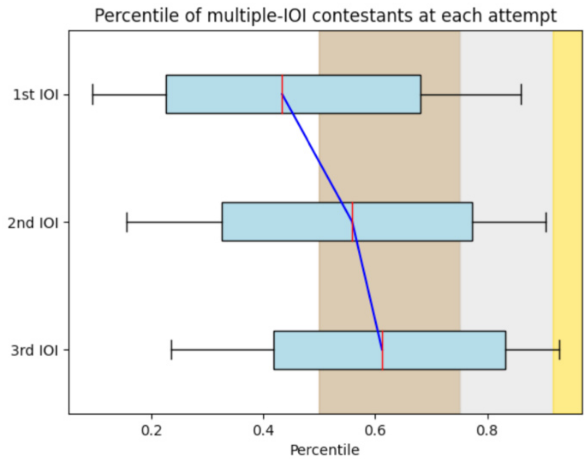


Fig. 6. Quantile plot of percentile distribution separated by ordering of IOI attempt.

Table 2

Summary statistics of percentile distribution separated by IOI order

Percentile in:	count	mean	std	10%	25%	50%	75%	90%
1st IOI	1028	0.46	0.27	0.1	0.23	0.43	0.68	0.86
2nd IOI	1038	0.54	0.27	0.16	0.33	0.56	0.77	0.9
3rd IOI	334	0.6	0.25	0.23	0.42	0.61	0.83	0.93

Fig. 4 shows that on average, contestants tend to improve over multiple IOIs, though the improvement is typically small and amounts to less than a single medal category over consecutive IOIs. Within a contestant's trajectory, improvements tend to decrease over time; for example, for a contestant with four or more IOIs, the improvement from the

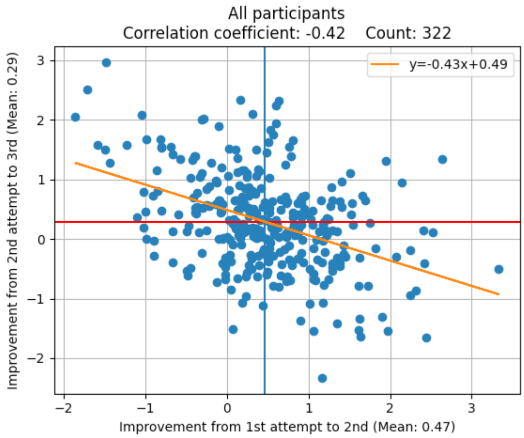


Fig. 7. Scatterplot of performance changes in consecutive IOIs, aggregated data.

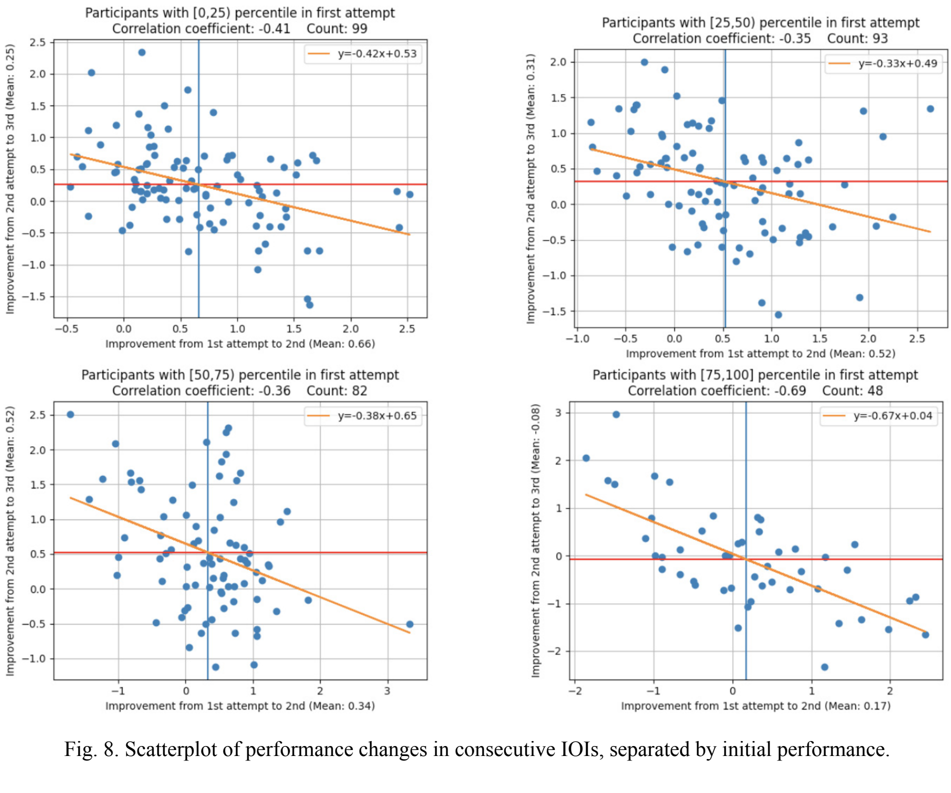


Fig. 8. Scatterplot of performance changes in consecutive IOIs, separated by initial performance.

first to the second IOI is approximately the same as the combined improvements from the second to third and the third to the fourth IOs. This phenomenon shows diminishing marginal returns of more IOI experience, as the easier areas for improvement are likely acted upon by a contestant's second IOI. Contestants with more total IOIs, while starting out at a lower performance level, on average also tend to improve faster and reach higher levels of final IOI performance.

Fig. 5 and Table 2 show that while multiple-IOI contestants start out slightly worse than the average contestant, they eventually improve to be stronger. This observation is true across all quantiles and all performance levels. For example, the median at a multiple-IOI contestant's first IOI is at the 43rd percentile, which improves to the 56th percentile in the second IOI then to the 61st percentile in the third IOI. Diminishing improvements are seen at the higher quantiles (50th and above) but not at the lower quantiles (10th and 25th), which suggests different learning and improvement dynamics at lower as compared to higher performance levels.

Finally, for the set of contestants with at least three IOIs, we compare, under the scaled score metric, the change in performance from the first to second IOI with the change in performance from the second to third IOI. Fig. 7 shows a single scatterplot with a negative correlation, while Fig. 8 separates the data based on the performance at the first IOI. Both plots consistently show the "regression to the mean" phenomenon, indicating that there is a substantial amount of idiosyncratic variance in IOI performance at all levels of competition.

3. Country Classification and Analysis

In this section, we present a cross-sectional analysis of the aggregate performance of countries participating at the IOI. For this analysis, we compare the average percentile of the contestants from a country from IOI 2011 through 2023; to avoid including data with too much noise, we subset our analysis to countries with at least 26 contestants (half of the maximum 52) in this time. We first compare this aggregate performance metric against two predictors, population and Human Development Index (HDI), then afterwards consider differences between geographic regions as well as measures of interest in competitive programming.

While the IOI is officially an individual competition, preparation and selection for this competition is typically done on the country level, so it is still of interest to consider the aggregate performance for a country; such data is also subject to less variance than individual participant or anecdotal data. We also use the participant-to-country correspondence provided at the time of IOI registration and consistent with IOI regulations, with no attempt to distinguish country of origin in the case of immigration or foreign diasporas. While these directions may be interesting for future research, we believe these are less necessary for an initial analysis.

For the country classification into regions, we use the United Nations Geoscheme, which divides the countries in the world into sub-divisions of continents. However, we make the following changes to balance the number of IOI-participating countries in each category:

- Only one category is provided for Africa, as well as for Latin America and the Caribbean
- Central Asia and Southern Asia are combined into a single category

- The United States, Canada, Australia, New Zealand, and the United Kingdom are assigned the "Anglosphere" category due to their cultural similarities and to avoid having multiple small categories (such as Northern America and Oceania)

Thus, our geographical analysis involves eleven regions, with seven to thirteen IOI-participating countries in each:

- 1–4: Asia: Southeastern Asia, Eastern Asia, Central and Southern Asia, Western Asia
- 5–8: Europe: Eastern Europe, Southern Europe, Northern Europe, Western Europe
- 9: Africa, 10: Latin America, 11: Anglosphere

A. *Comparison of Average IOI Percentile to Population and Human Development Index*

As a first analysis, we generate scatterplots to compare the average IOI percentile of each country's contestants with the population and HDI. Our prior is that having a large population means having a larger pool of talent to draw from, irrespective as to how well this talent is nurtured in the country. Meanwhile, the HDI is a widely used metric for development which correlates to the availability of educational resources and strength of institutions. Therefore, these two metrics are a good start towards uncovering the causes that explain the variation between different countries' average IOI performance.

For population, we use a log scale as is standard in economic studies, typically done to avoid distortion from the roughly log-normal distribution of country populations. For HDI, we use a linear scale due to the approximately linear distribution among IOI-participating countries. We also plot a trendline on each scatterplot to evaluate the strength of the correlation with these two indicators. The points in the scatterplot are also colored by region for ease of reference and for a preliminary display of regional trends, for example, as to which regions are mostly above or below each trendline. The results are shown in Fig. 9.

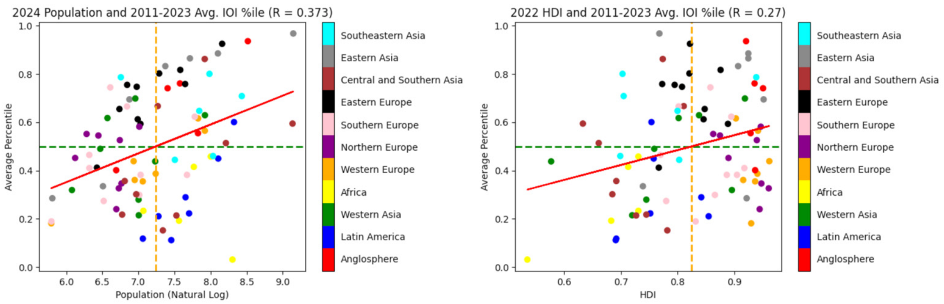


Fig. 9. Scatterplot of average IOI percentile by country, with log Population and HDI.

In Fig. 9, a moderate correlation of 0.373 is seen for the log population variable, while a slightly weaker correlation of 0.270 is seen for the HDI variable. The correlations, and particularly the population correlation, also appear to be a trend for the dataset as a whole and are not driven by a small number of outliers in the data. These correlation coefficients show that both explanatory variables have some correlation with a country's IOI performance, but a lot of the variation is unexplained by these two indicators.

B. *Regression Analysis and Residual Grouping by Region*

Next, we consider a regression analysis using the same data as the previous subsection, the results of which are shown in Table 3. The regression results show that both the log population and the HDI metric are significant in their predictive power towards a country's IOI performance, even when the other is considered. The overall R-squared, however, is only 0.312, meaning that most (around 70%, and likely higher due to possible overfitting) of the variance is unexplained by these two predictors.

The regression analysis yields a fitted model, so we can take calculate the residual for each country datapoint, representing the variance in country average IOI percentiles unexplained by the predictors. We then calculate the mean and standard deviation of the residual by geographical region, a summary of which is presented in Table 4. The residual mean represents each region's overall strength at the IOI beyond what is predicted by population size and HDI, while the residual standard deviation is a measure of heterogeneity of country strengths.

Due to the small bucket sizes, testing for statistical significance will be difficult. The residual means column in Table 4, however, still provides a broad picture as to the strengths of countries in each geographical region after controlling for population and HDI. Three regions, Eastern Europe, Eastern Asia, and Southeastern Asia, have high residual means (at least ten percentile points), with Eastern Europe having by far the

Table 3

Regression output of average IOI percentile on HDI and log population

OLS Regression Results

Dep. Variable:	Average Percentile	R-squared:	0.312
Model:	OLS	Adj. R-squared:	0.294
Method:	Least Squares	F-statistic:	17.88
No. Observations:	82	AIC:	-35.68
Df Residuals:	79	BIC:	-28.46
Df Model:	2		

	coef	std err	t	P>|t|	[0.025	0.975]
constant	-1.5287	0.340	-4.495	0.000	-2.206	-0.852
2022 HDI	1.0051	0.226	4.451	0.000	0.556	1.455
2024 Population (ln)	0.1657	0.032	5.237	0.000	0.103	0.229

Table 4
Residual summary statistics by geographical region

Region	Residual Mean	Residual Stdev.	Count
Eastern Europe	0.220	0.094	10
Eastern Asia	0.122	0.106	7
Southeastern Asia	0.105	0.153	6
Western Asia	0.059	0.151	8
Anglosphere	0.007	0.131	5
Southern Europe	-0.009	0.195	10
Central & Southern Asia	-0.018	0.193	9
Northern Europe	-0.044	0.173	8
Africa	-0.170	0.120	5
Western Europe	-0.171	0.057	7
Latin America	-0.221	0.115	7

highest, consistent with common knowledge on these countries' scientific education traditions. (Lovheim, 2021) Meanwhile, the three regions with the lowest residual means are Africa, Western Europe, and Latin America, likely related to the lack of prominence of the scientific Olympiads in these regions, and in the case of Western Europe, relative to other educational opportunities.

C. *Incorporating Codeforces Registration Data*

In this final analysis, we consider a country-level metric of student interest in competitive programming, the number of active participants on the CodeForces (CF) competitive programming platform. Fig. 10 shows a scatterplot, where we note that the CF participant count measures both a country's population and level of interest in competitive programming.

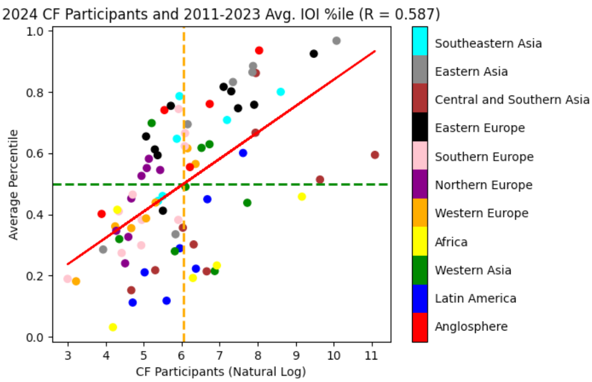


Fig. 10. Scatterplot of CodeForces participant count and country average percentile.

Table 5

Regression output of average IOI percentile on HDI, log population, and log CF density

```
                                OLS Regression Results
==============================================================================
Dep. Variable:     Average Percentile   R-squared:                      0.602
Model:                            OLS   Adj. R-squared:                 0.587
Method:                 Least Squares   F-statistic:                    39.32
No. Observations:                  82   AIC:                           -78.59
Df Residuals:                      78   BIC:                           -68.96
Df Model:                           3
==============================================================================
                           coef    stderr       t      P>|t|     [0.025   0.975]
------------------------------------------------------------------------------
constant                -1.1855     0.264   -4.487   0.000     -1.711   -0.660
2022 HDI                 1.2334     0.175    7.030   0.000      0.884    1.583
2024 Population (ln)     0.2608     0.027    9.553   0.000      0.206    0.315
2024 CF Density (ln)     0.2645     0.035    7.542   0.000      0.195    0.334
```

The correlation in Fig. 10 is noticeably stronger than that in either scatterplot of Fig. 9, showing the drastic importance of the level of student interest and engagement in a practice platform towards achieving strong IOI results. To place this correlation in context, we run a second regression, similar to the regression in the previous section, where in addition to HDI and population we include a "CF Density" metric which is the ratio of CF participants in a country to its total population. The regression results in Table 4 show a much higher R-squared metric of 0.602 compared to 0.312 the previous regression, meaning that the CF density explains nearly half of the residual variance not predicted by population size and HDI.

4. Conclusion and Further Work

This paper appears to be the first large-scale analysis of IOI data, particularly with the focus on returning contestants and geography. Overall, the analyses provide statistical justification for several trends that are likely to be known in an approximate sense within the IOI community, while providing convincing evidence for interested parties outside the IOI community.

A. *Summary of Findings*

From the analysis on returning contestants' IOI performance trajectories, the overarching observation is that IOI performance, as a metric, has good internal validity, given the large correlation between consecutive IOI results. (Figures 1 and 2) There is, however, still substantial variance in skill acquisition over time and unpredictability in the results of a single IOI. These are seen by the large spreads in improvement over consecutive IOIs (Fig. 3) and the reversion to the mean phenomenon (Fig. 7). We have also seen diminishing marginal improvements over multiple consecutive IOIs. (Figures 4 and 5)

Meanwhile, from the geographical analysis, we demonstrated that while broad demographic statistics such as population and Human Development Index (HDI) are significantly correlated with a country's aggregate IOI performance. (Fig. 8 and Table 3) A more important factor, however, is the level of competitive programming interest in the country. (Tables 5) We have also identified regions that are substantially stronger or weaker than what would have been predicted from population and HDI alone, while classifying some regions as more heterogenous or homogenous than others in terms of their countries' IOI performance.

B. *Recommendations for Further Research*

There are many directions in which the analyses in this paper can be extended to yield an even deeper understanding of the educational nature of the IOI in its global setting.

As a first step, it may be interesting to combine the two directions considered in this paper and analyze their interactions. This paper considers the trajectory of returning contestants and the aggregate performance of countries separately, though there can be some insight in analyzing the distribution of improvement in returning contestants from different categories of countries. Analysis, however, might be more difficult due to having few datapoints at this proposed level of granularity.

The concern of having few datapoints made conducting statistical tests difficult particularly for the returning contestant analysis, as the sheer variance in differences in performances causes standard errors to be large. Therefore, only the less surprising and insightful trends could be tested under this framework. Eventually, as the IOI continues to run over the years and generates more high-quality data, it might be feasible to revisit this approach. For example, after ten more IOIs, the amount of usable returning contestant data will nearly double.

Finally, the IOI statistics can be combined with a qualitative analysis and substantive interpretation of the nature of the IOI. There are many interesting directions, such as considering which tasks may be favorable to different populations; for example, identifying tasks that are more approachable by first-time, less experienced contestants or contestants from weaker countries. This direction can also be applied understand the role of the IOI in global talent development and identification, such as for example, in what ways are strong contestants from weaker countries tend to be different from a typical strong contestant.

Acknowledgements

We thank Jay Leeds, Gerard Francis Ortega, and Payton Yao for discussion and insights.

References

Barnett, A.G., van der Pols, J.C., Dobson, A.J. (2004). Regression to the mean: What it is and how to deal with it. International Journal of Epidemiology, 34(1), 215–220. https://doi.org/10.1093/ije/dyh299

Codeforces Country Ratings. Retrieved from https://codeforces.com/ratings/countries

International Olympiad in Informatics International Committee (2023). IOI Regulations. Retrieved from https://ioinformatics.org/files/regulations23.pdf

International Olympiad in Informatics International Scientific Committee (2024). IOI Syllabus. Retrieved from https://ioinformatics.org/files/ioi-syllabus-2024.pdf

International Olympiad in Informatics – Statistics. Retrieved from https://stats.ioinformatics.org/

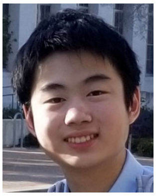
Lövheim, D. (2021). Cold War fostering of scientific elites: International youth olympiads in chemistry and physics 1967–1984. History of Education, 50(5), 685–703. https://doi.org/10.1080/0046760x.2021.1890239

Marsh, L.C., Cormier, D.R. (2005). Spline regression models. Sage Publications.

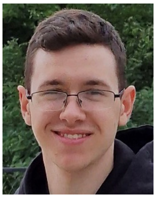
Verhoeff, T. (2006). The IOI is (not) a science olympiad. *Informatics in Education*, *5*(1), 147–159. https://doi.org/10.15388/infedu.2006.25

United Nations Department of Economic and Social Affairs. (n.d.). United Nations geoscheme. Retrieved from https://esa.un.org/MigFlows/Definition%20of%20regions.pdf

United Nations Development Programme. (n.d.). Human Development Index (HDI) Dataset.
Retrieved from https://hdr.undp.org/data-center/human-development-index



E. Lee – is an incoming first-year undergraduate at Stanford University. He plans to pursue a major in the area of applied mathematics. He is an avid enjoyer of mathematics and algorithmic competitions, and his accomplishments in these areas include earning a USA Math Olympiad bronze medal as well as qualifying for the Platinum Division of the USA Computing Olympiad. He shares his interests by educating others about problem-solving and organizing his own contests as well.



T. Reizin – is studying his Bachelors in Computer Science at Charles University in Prague. During his high school years he actively participated in the Ukrainian Olympiads in Informatics. After graduating he continued his involvement with the Ukrainian olympiads, by teaching in preparatory camps and helping with organization of team selection. He also managed to advance the ICPC European Championship. His research interests lie in combinatorics, in particular extremal combinatorics.

F.E. Wu – is a PhD student in Operations Research at the Massachusetts Institute of Technology, studying information-theoretic algorithms for reinforcement learning and decision making. In high school, he won a gold medal at the IMO and a bronze medal at the IOI. Prior to starting the PhD, he worked in a quantitative research position in industry and as an undergraduate was extensively involved in teaching probability and statistics. He also coaches the Philippines' IOI 2024 team.

F.E. Wu is an incoming first-year undergraduate at the Massachusetts Institute of Technology. He represented the Philippines at the International Mathematical Olympiad (IMO) and the IOI, winning a silver medal and bronze medal, respectively, in 2023. Within competitive programming, besides competing he is interested in building learning communities and making the learning process more enjoyable. He plans to continue studying mathematics and computer science in college.