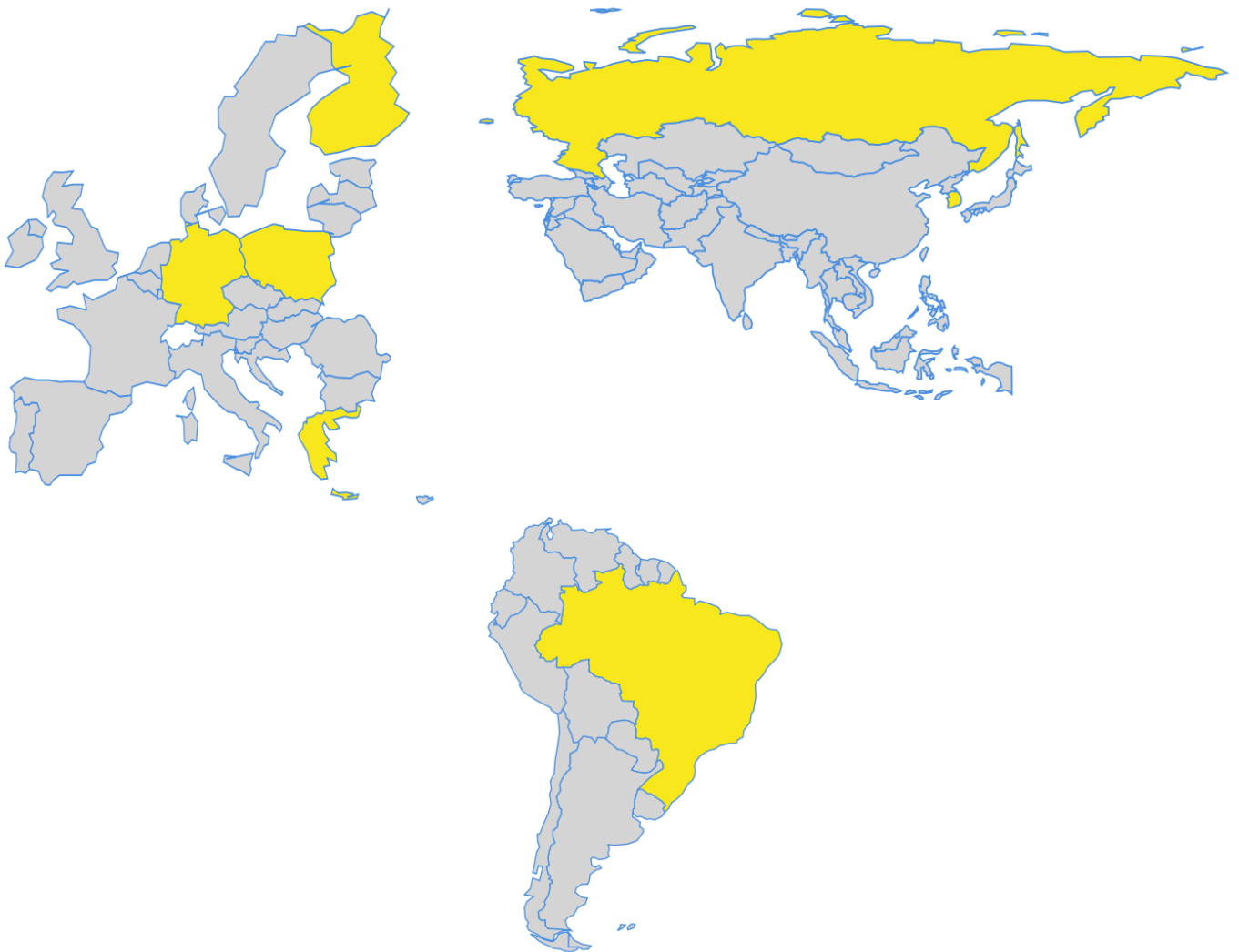# PISA 2015

From Data Scientists' Point of View

Brazil, Germany, Finland, Greece, South Korea, Poland and Russia students comparison

Katarzyna Zdon, Dorota Łępicka, Mateusz Kałamarz

Faculty of Mathematics and Information Science, Warsaw University of Technology

Abstract

The Programme for International Student Assessment (PISA) is a worldwide study by the Organisation for Economic Co-operation and Development (OECD) in member and non-member nations of 15-year-old school pupils' scholastic performance on mathematics, science, and reading. In our research we analysed data of PISA 2015.

After data processing only chosen countries were taken into consideration, these were: Brazil, Germany, Finland, Greece, South Korea, Poland and Russia. Thus analysed data included 41 041 students.

The question posed in our research was: What if these young individuals were (more or less consciously) using different strategies during the test? To test this hypothesis five strategies were proposed.

As a result we determined the strategies used by most of each countries' students.

In the final chapter we showed that if all the students were using single strategy final ranking would stay the same but the best strategy to chose is to prepare for the test.

PISA

The Programme for International Student Assessment (PISA) is a worldwide study by the Organisation for Economic Co-operation and Development (OECD) in member and non-member nations of 15-year-old school pupils' scholastic performance on mathematics, science, and reading. It was first performed in 2000 and then repeated every three years. Its aim is to provide comparable data with a view to enabling countries to improve their education policies and outcomes. It measures problem solving and cognition in daily life.

The 2012 version of the test involved 34 OECD countries and 31 partner countries, with a total of 510,000 participating students. The 2015 version of the test was published on 6 December 2016.

The Data

In our research we analysed data of Programme for International Student Assessment (PISA) 2015. Two datasets were downloaded from PISA website - student-questionnaire data file (which also includes estimates of student performance and parent-questionnaire data) and the cognitive item data file. These files included information concerning 519 334 participants. Data was cleaned from missing values and records not meeting crucial criteria e.g. overall time exceeding time of the test (2h). Finally only chosen countries were taken into consideration, these were: Brazil, Germany, Finland, Greece, South Korea, Poland and Russia. Thus analysed data included 41 041 students.

Analysis
Young strategists?

Many reports and analyses based on PISA data were presented, that is why we would like to take a different approach on the matter.

What if these young individuals were (more or less consciously) using different strategies during the test? To test this hypothesis five strategies were proposed.

This late already?!

First we checked how many students start slowly at the very beginning but later, possibly realizing how much time they have got left, decide to hurry up with last parts of the test.

**Figure 1: Percentage of students in each country using Strategy A**
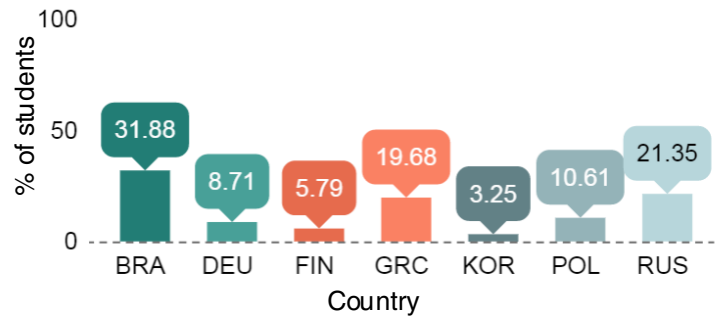
The above plot presents percentage of students using first strategy in each country. As we can see this is one-third of students in Brazil and approximately 20% of students in Greece and Russia.

Five more minutes, please?

Two hours for average of 51 questions (median 53). Is that enough time to fill the whole test?

This time we were looking for students who did not plan their time well enough compared to overall performance. What for sure is worth highlighting is the fact that, according to our analysis, to be included in this group, one needed to lack solution to any single question.

**Figure 2: Percentage of students in each country using Strategy B**
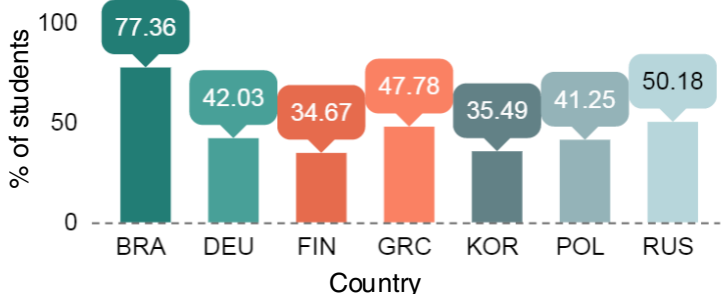
This strategy seems to be equally distributed amongst countries. Biggest proportion is observed for Brazil (77.36%) and the opposite for Finland (34.67%).

Know it or skip it

In this section we were expecting students to spend much more time on the questions that they knew the answer for while skipping other ones.

Figure 3: Percentage of students in each country using Strategy C
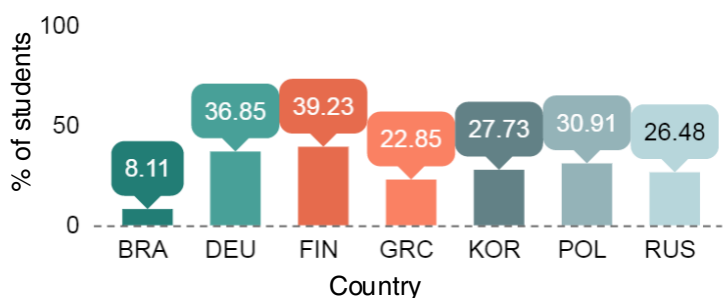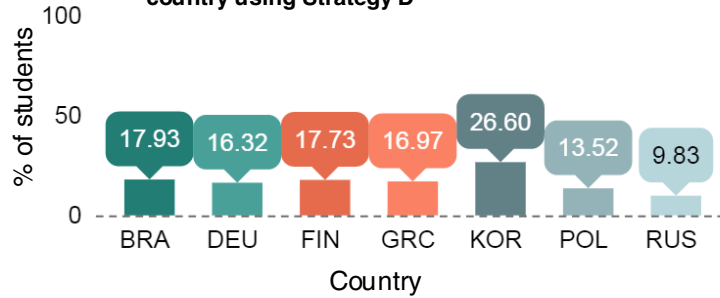

This strategy is mostly the domain of Finnish (39.23%) and German (36.85%) students. The least proportion of students using it were the Brazilian (8.11%) ones.

Make haste!

For the fourth strategy, students willing to finish the test as fast as possible, no matter the results, were considered.

**Figure 4: Percentage of students in each country using Strategy D**

This strategy was most likely to be found in South Korea (26.20%). The least rushing students were found in Russia (9.83%).

Brace Yourselves, PISA is Coming

As Alexander Graham Bell once said - "Before anything else, preparation is the key to success".

Last group considered in our analysis were the students that prepared for the test. These were the individuals having low number of actions, short overal time and highest score.

**Figure 5: Percentage of students in each country using Strategy E**
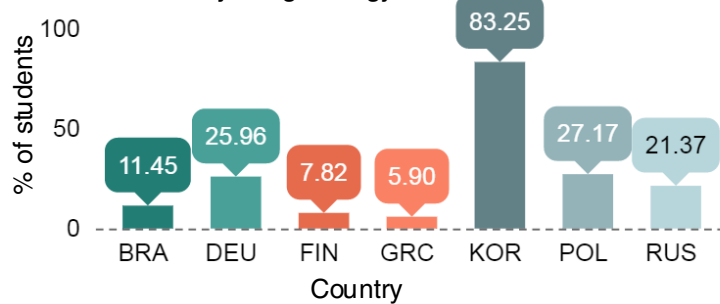
This strategy is the korean domain (83.25%). The least preparing countries are Finland (7.82%) and Greece (5.92%). This of course impacts the final score of South Korea in PISA ranking.

The one has no strategy

Now the question is - did we find a strategy of every single individual? The answer is - no.

Figure 6: Percentage of students in each country using no strategy
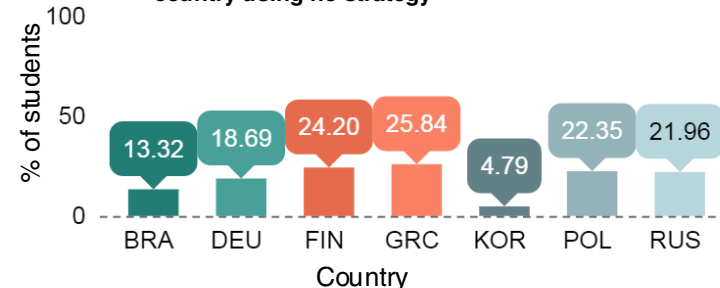

We did not assign any strategy to one-fourth of Greeks (25.84%) and Fins (24.20%). On the other side we well explained strategies of Brazilians (86.68%) and Koreans (95.21%). The average of students with assigned strategy was 81,26%.

The reason for this might be either fact that we considered only five strategies and did not come up with the right ones for these students. Another plausible explanation is that these students who do not follow any obvious pattern. Both answers make a room for further analysis. Either more strategies need to be developed or more complex analysis of considered data should be held.

Summary of strategies

Figure 7: Percentage of students using certain Strategy grouped by countries
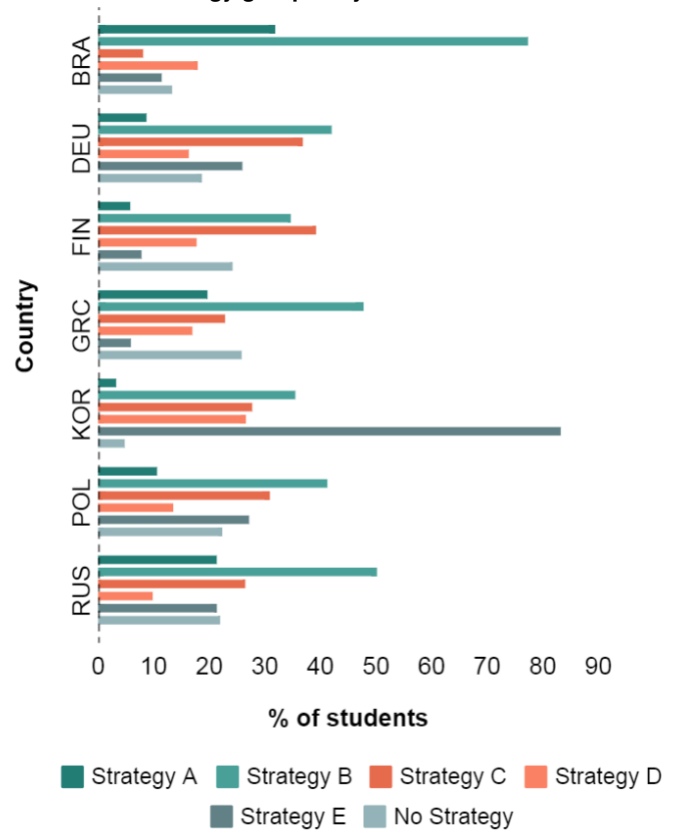
The above plot shows that there were many individuals who were assigned to more than one strategy (summary percentages exceeding 100%). This might be caused by smaller dataset (just seven countries) or by similarity in some strategies.

The most commonly used strategies were:
Strategy B for Brasil, Germany, Greece, Poland and Russia;
Strategy C for Finland;
Strategy D for Korea.

What if?

As the overall mean scores for these countries vary and in the previous chapter we have shown that they use different strategies a new question arises - what if we omitted strategies' impact on the scores? Thus we will analyse what would happen if everyone used just one, chosen strategy.

Only Strategy A

Strategy A - group of students who spent more time at the beginning of the test and less at the end, answering the questions, compared to their colleagues

Figure 8: Comparison of mean score of each country for original data compared to situation of using only Strategy A
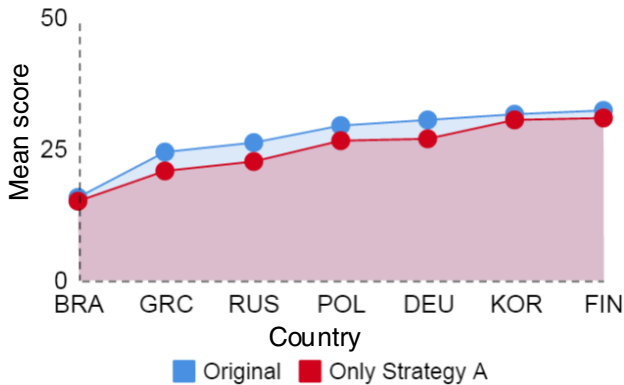
Using only first strategy would not have much impact on the scores. All the results would be slightly worse than the original one. Also the final ranking would not change at all.

Only Strategy B

Strategy B - students who did not plan their time well thus filling smaller part of a test compare to the overall performance. We showed that these are students that lack answer to at least one question.

Figure 9: Comparison of mean score of each country for original data compared to situation of using only Strategy B
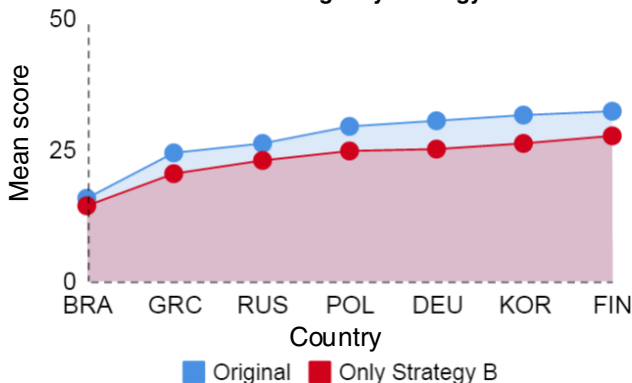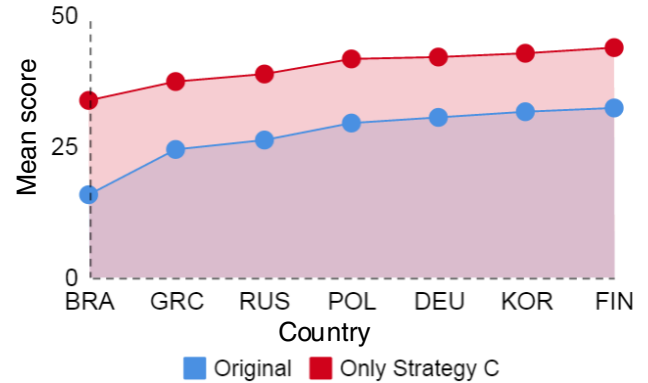
If students used only second strategy the ranking would not change. All of the countries' results would be slightly lower compared to the original data.

Only Strategy C

Strategy C - considered individuals who spent more time on questions that they answered correctly while not spending time on those they answered wrongfuly.

Figure 10: Comparison of mean score of each country for original data compared to situation of using only Strategy C

If Brazilian students used only the third strategy their overall average score would be twice as big (33.76% compared to 16.66%). Using this strategy every single country would have performed much better.

Only Strategy D

Strategy D - group of students who finished the test much faster than others, no matter the result.

Figure 11: Comparison of mean score of each country for original data compared to situation of using only Strategy D
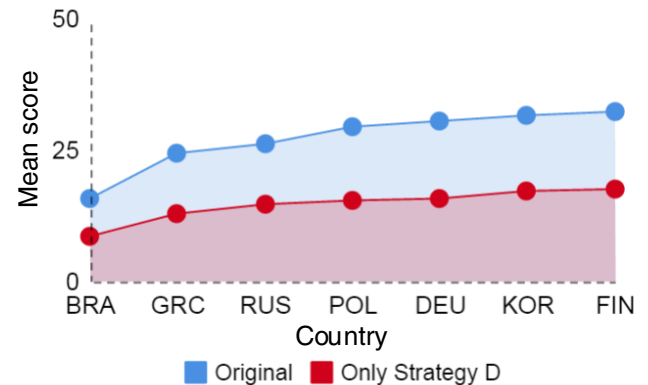
Using only fourth strategy would lower all countries' scores. This strategy would impact Poland, Germany, Korea and Finland the most (more than 14 points lost).

Only Strategy E

Strategy E - students who prepared for the test, having lower time of filling the whole test, better total score and fewer actions (clicks) taken.

Figure 12: Comparison of final score of each country for original data compared to situation of using only Strategy B
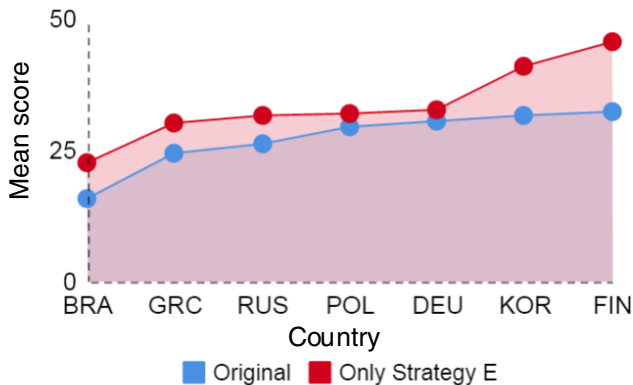
Korea and Finland would benefit the most from using only the fifth strategy compared to the other countries (+9,24 and 13.16 points respectively). This strategy would have the lowest impact on Poland and Germany scores.

Summary

Figure 13: Comparison of final score of each country for original data compared to situation of using only certain Strategy
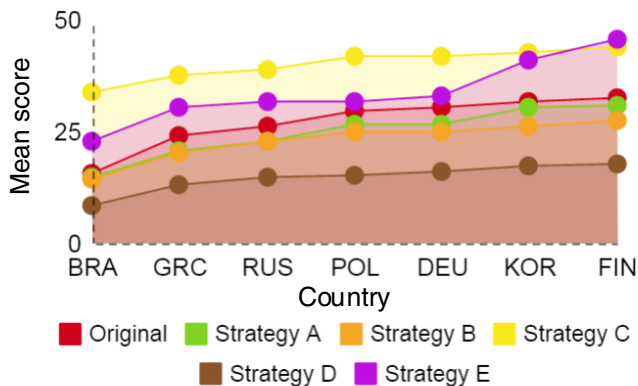
The best strategy for all the countries would be to focus on the tasks that they know the best while skipping the other ones (Strategy C). Only the Finland would benefit more from preparing to the test (Strategy E). The worst strategy for all the countries is to rush through the test willing to finish as fast as possible (Strategy D).

Results

Our analysis showed that all countries' students use different strategies.

Most commonly used strategies were:
Strategy B for Brazil, Germany, Greece, Poland and Russia;
Strategy C for Finland;
Strategy D for Korea.

Strategy was assigned to the average of 81,26% of students ranging from 75,80% for Finland to 95,21% for Korea.

If only one strategy was used:
Strategies C and E Strategy would rise all countries' scores;
Strategies A, B and D would decrease the scores.

Strategies C and E are the best to chose from and the Strategy D would decrease overall scores of all the countries the most.

The ranking would not change no matter the strategy used by countries.

Source

If you want to see the whole process of preparing, clustering, analyzing and visualizing the data take a look at that page, all the files are in HTML format.

Source on GitHub