

How to use R to hack the publicly available data about skills of 2M+ worldwide students?

Przemysław Biecek
University of Warsaw
Warsaw University of Technology



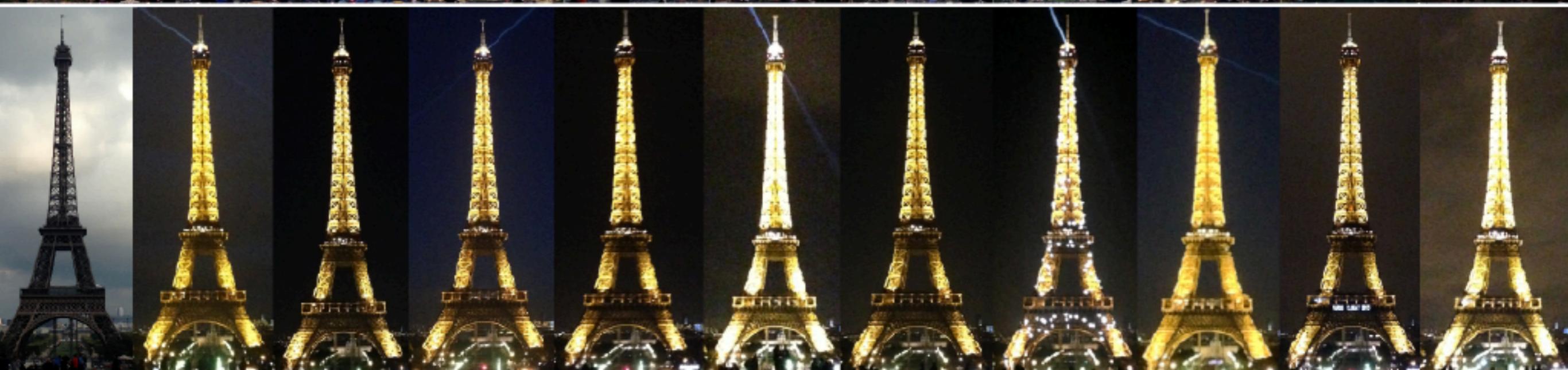
A statistician with research interests:

Applications of data visualisation + machine learning + $p \gg n$ modelling

in molecular human genetics, especially oncology.

Sabbatical in 2013:

Thomas J. Alexander Fellowship - Edu @ OECD



PISA





Programme for International Student Assessment (PISA)

[Home](#)[About PISA](#)[PISA products](#)[Key findings](#)[Contacts](#)[PISA FAQ](#)

Programme for International Student Assessment [PISA]

- Every 3 years hundreds of thousands 15-years old kids from different countries take part in the assessment in three areas: reading, math and science
- Along with their academic performance also some additional data is gathered. Mostly about their environment, attitude, family and school.



Programme for International Student Assessment (PISA)

[Home](#)[About PISA](#)[PISA products](#)[Key findings](#)[Contacts](#)[PISA FAQ](#)

PISA in numbers:

- Around 2 000 features for each student
- Over 2 000 000 students [4x the population of Poznań]

Year	Main subject	Countries	Students
2000	Reading	43	265000
2003	Mathematics	41	275000
2006	Science	57	400000
2009	Reading	85	470000
2012	Mathematics	65	510000
2015	Science	69	475000



OECD

1. WHAT STUDENTS KNOW AND CAN

How do countries/economies perform in reading over

Figure 1.2. Comparing performance in reading

- Statistically significantly above the OECD average
- Not statistically significantly different from the OECD average
- Statistically significantly below the OECD average



Student questionnaire

Q31 Thinking about your views on mathematics: to what extent do you agree with the following statements?

(Please darken only one circle in each row.)

	<i>Strongly agree</i>	<i>Agree</i>	<i>Disagree</i>
a) I enjoy reading about mathematics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) I look forward to my mathematics lessons.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) I do mathematics because I enjoy it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Learning mathematics is worthwhile for me because it will improve my career prospects.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) I am interested in the things I learn in mathematics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g) Mathematics is an important subject for me because I need it for what I want to study later on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q45 Thinking about mathematical concepts: how familiar are you with the following terms?

(Please tick only one box in each row.)

	<i>Never heard of it</i>	<i>Heard of it once or twice</i>	<i>Heard of it a few times</i>	<i>Heard of it often</i>	<i>Know it well, understand the concept</i>
a) Exponential Function	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Divisor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Quadratic Function	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) <Proper Number>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Linear Equation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Vectors	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Complex Number	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h) Rational Number	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
i) Radicals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
j) <Subjunctive Scaling>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
k) Polygon	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
l) <Declarative Fraction>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
m) Congruent Figure	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
n) Cosine	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
o) Arithmetic Mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Student questionnaire

Q31 Thinking about your views on mathematics: to what extent do you agree with the following statements?

(Please darken only one circle in each row.)

	<i>Strongly agree</i>	<i>Agree</i>	<i>Disagree</i>
a) I enjoy reading about mathematics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) I look forward to my mathematics lessons.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) I do mathematics because I enjoy it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Learning mathematics is worthwhile for me because it will improve my career prospects.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) I am interested in the things I learn in mathematics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g) Mathematics is an important subject for me because I need it for what I want to study later on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q45 Thinking about mathematical concepts: how familiar are you with the following terms?

(Please tick only one box in each row.)

	<i>Never heard of it</i>	<i>Heard of it once or twice</i>	<i>Heard of it a few times</i>	<i>Heard of it often</i>	<i>Know it well, understand the concept</i>
a) Exponential Function	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Divisor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Factor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) <Proper Number>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Linear Equation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Vectors	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Complex Number	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h) Rational Number	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
i) Declarative Sentence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
j) <Subjunctive Scaling>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
k) Congruent Figure	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
l) <Declarative Fraction>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
m) Cosine	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
n) Arithmetic Mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Parent questionnaire

Q How important are the following reasons for choosing a school for your child?

(Please tick only one box in each row.)

	<i>Not important</i>	<i>Somewhat important</i>	<i>Important</i>	<i>Very important</i>
--	----------------------	---------------------------	------------------	-----------------------

a) The school is at a short distance to home. ₁ ₂ ₃ ₄

b) The school has a good reputation. ₁ ₂ ₃ ₄

c) The school offers particular courses or school subjects. ₁ ₂ ₃ ₄

d) The school adheres to a particular <religious philosophy>. ₁ ₂ ₃ ₄

e) The school has a particular approach to <pedagogy/didactics, e.g. example>. ₁ ₂ ₃ ₄

f) Other family members attended the school. ₁ ₂ ₃ ₄

g) <Expenses are low> (e.g. tuition, books, room and board). ₁ ₂ ₃ ₄

h) The school has <financial aid> available, such as a school loan, scholarship, or grant. ₁ ₂ ₃ ₄

i) The school has an active and pleasant school climate. ₁ ₂ ₃ ₄

Q We are interested in what you think about your child's school.

How much do you agree or disagree with the following statements?

(Please tick only one box in each row.)

	<i>Strongly agree</i>	<i>Agree</i>	<i>Disagree</i>	<i>Strongly disagree</i>
--	-----------------------	--------------	-----------------	--------------------------

a) Most of my child's school teachers seem competent and dedicated. ₁ ₂ ₃ ₄

b) Standards of achievement are high in my child's school. ₁ ₂ ₃ ₄

c) I am happy with the content taught and the instructional methods used in my child's school. ₁ ₂ ₃ ₄

d) I am satisfied with the disciplinary atmosphere in my child's school. ₁ ₂ ₃ ₄

e) My child's progress is carefully monitored by the school. ₁ ₂ ₃ ₄

f) My child's school provides regular and useful information on my child's progress. ₁ ₂ ₃ ₄

g) My child's school does a good job in educating students. ₁ ₂ ₃ ₄

School questionnaire

Q

In your school, to what extent is the learning of students hindered by the following phenomena?

(Please tick one box in each row.)

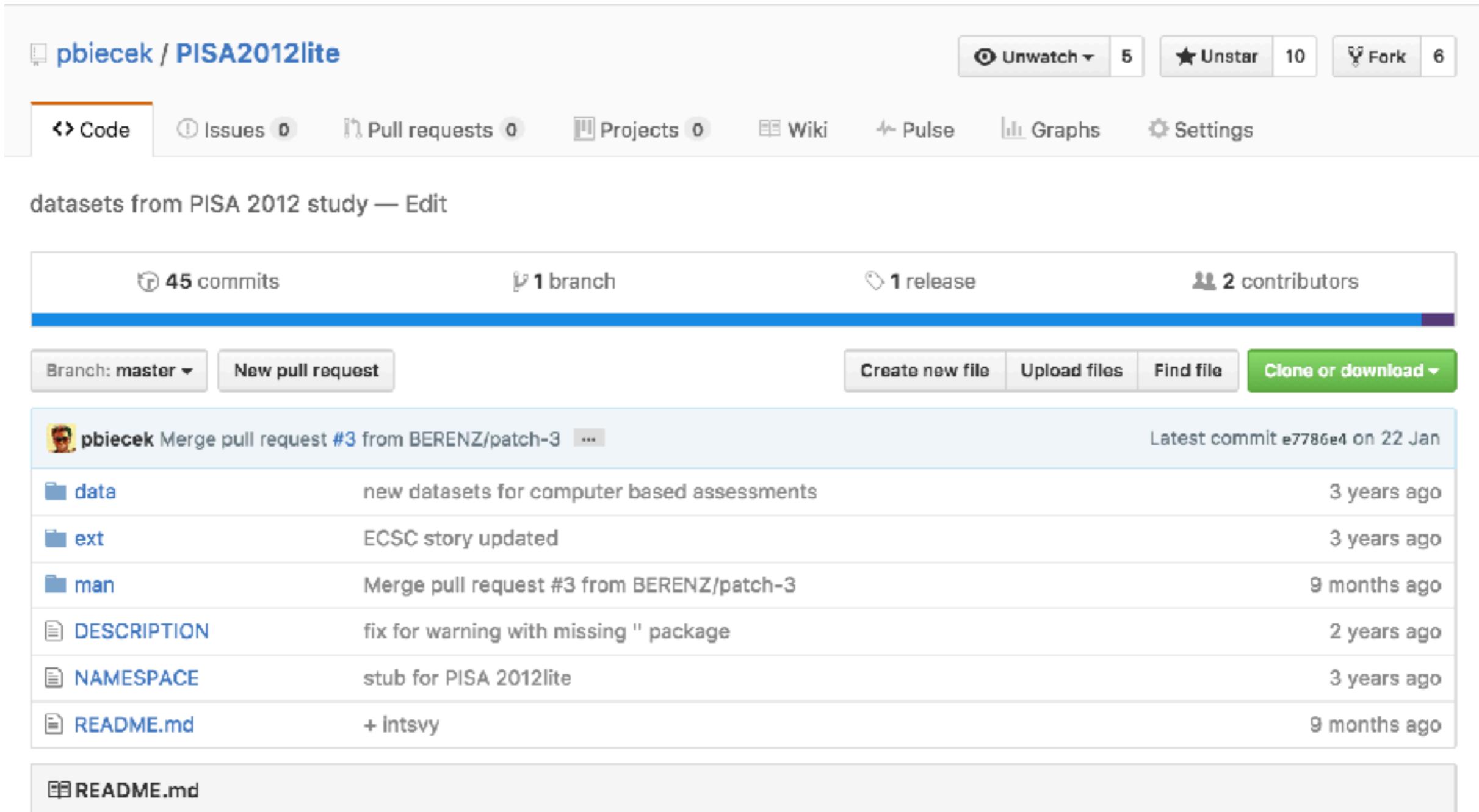
	<i>Not at all</i>	<i>Very little</i>	<i>To some extent</i>
a) Student truancy	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
b) Students skipping classes	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
c) Students arriving late for school	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
d) Students not attending compulsory school events (e.g. sports day) or excursions	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
e) Students lacking respect for teachers	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
f) Disruption of classes by students	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
g) Student use of alcohol or illegal drugs	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
h) Students intimidating or bullying other students	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
i) Students not being encouraged to achieve their full potential	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
j) Poor student-teacher relations	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
k) Teachers having to teach students of			

To what extent have appraisals of and/or feedback to teachers directly led to the following?

(Please tick one box in each row.)

	<i>No change</i>	<i>A small change</i>	<i>A moderate change</i>
a) A change in salary	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
b) A financial bonus or another kind of monetary reward	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
c) Opportunities for professional development activities	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
d) A change in the likelihood of career advancement	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
e) Public recognition from you	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
f) Changes in work responsibilities that make the job more attractive	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
g) A role in school development initiatives (e.g. curriculum development group, development of school objectives)	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3

Flat text files are on the OECD website
R packages with PISA data are on GitHub

A screenshot of a GitHub repository page for the user pbiecek's repository "PISA2012lite". The page shows basic repository statistics: 45 commits, 1 branch, 1 release, and 2 contributors. It includes navigation links for Code, Issues (0), Pull requests (0), Projects (0), Wiki, Pulse, Graphs, and Settings. Below the stats, there are buttons for Branch: master (dropdown), New pull request, Create new file, Upload files, Find file, and Clone or download (green button). A list of recent commits is displayed, showing changes to data, ext, man, DESCRIPTION, NAMESPACE, README.md, and README.md. The latest commit was made on Jan 22, 2018.

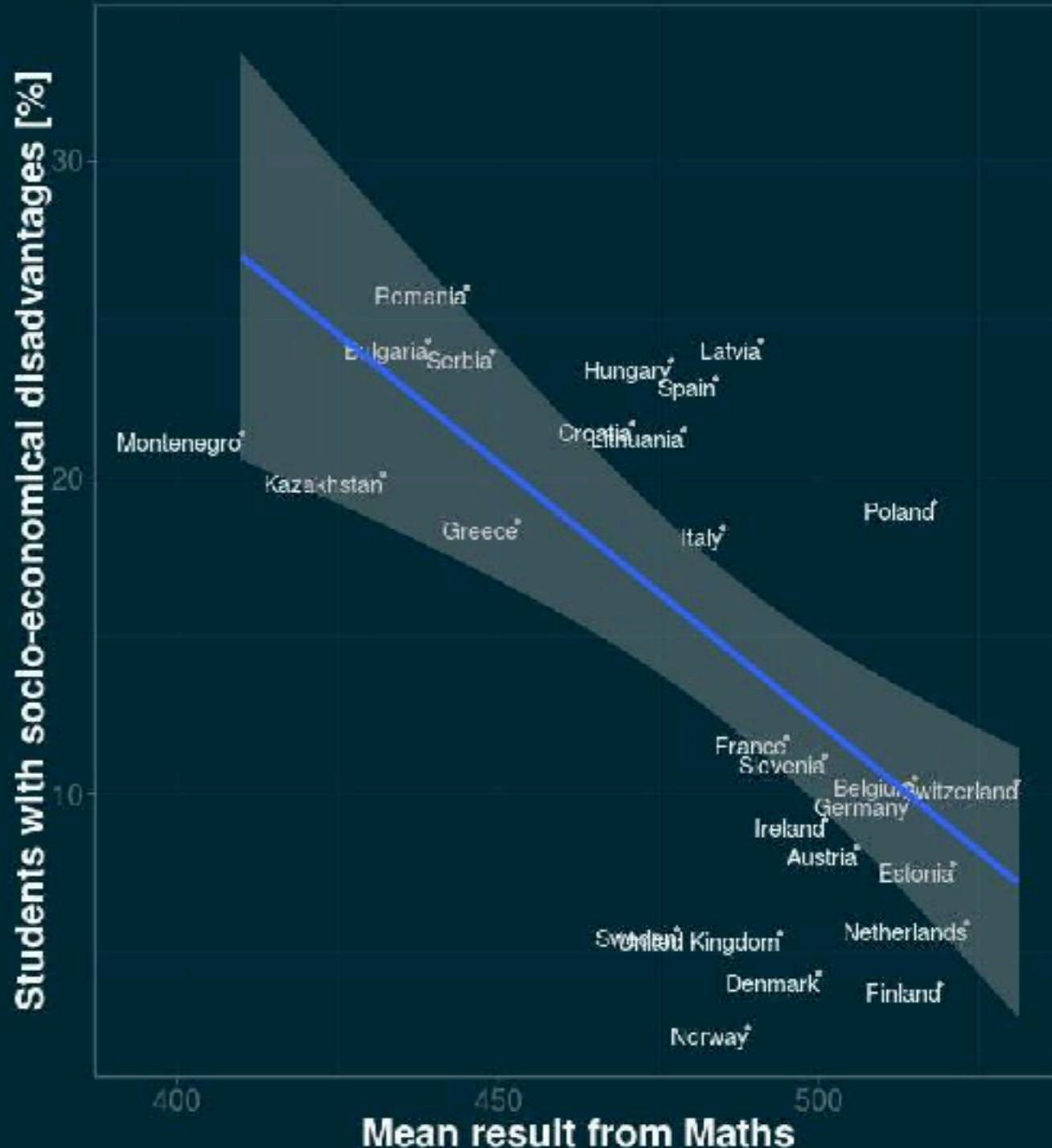
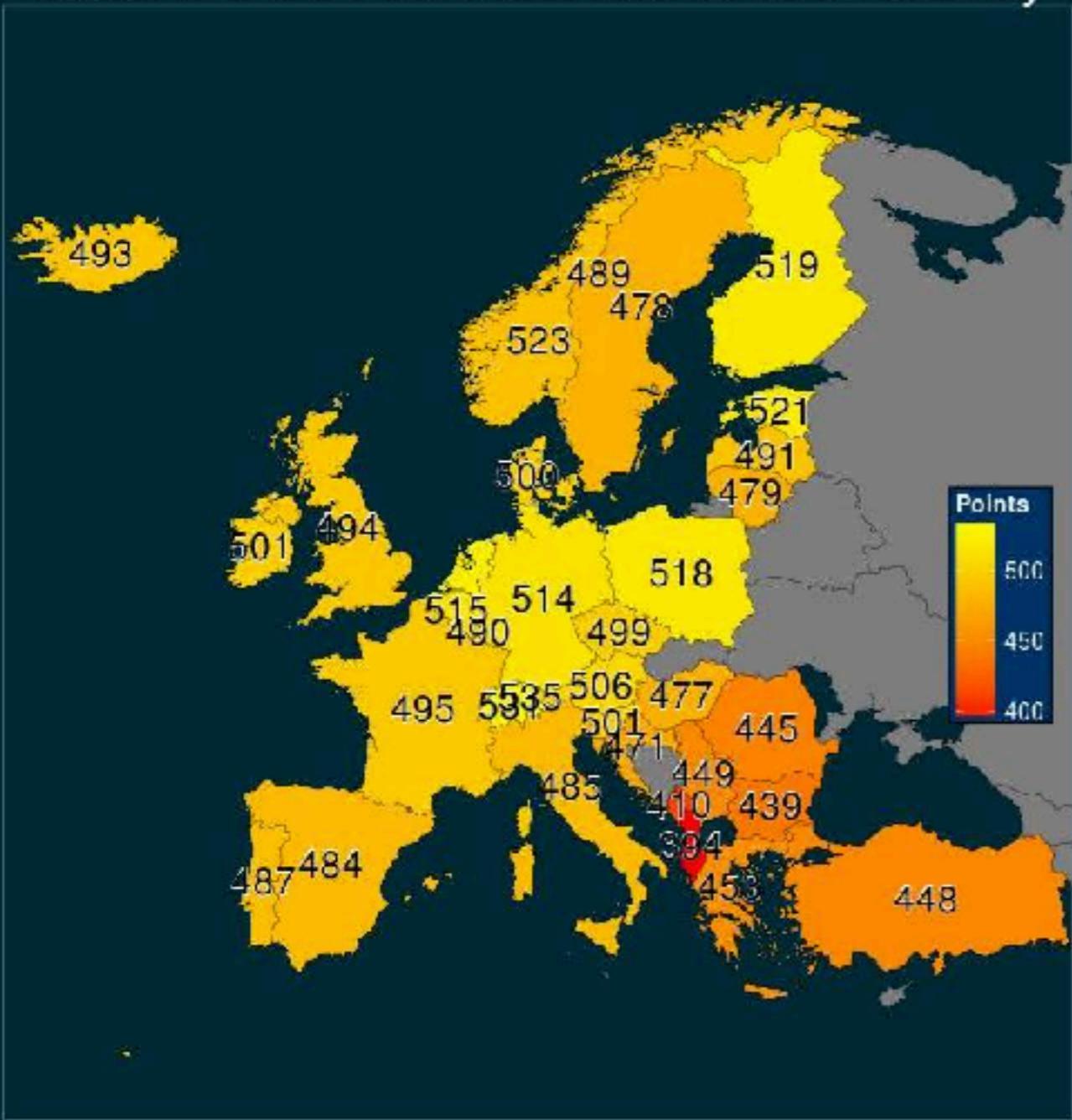
File / Commit Type	Description	Date
data	new datasets for computer based assessments	3 years ago
ext	ECSC story updated	3 years ago
man	Merge pull request #3 from BERENZ/patch-3	9 months ago
DESCRIPTION	fix for warning with missing " package	2 years ago
NAMESPACE	stub for PISA 2012lite	3 years ago
README.md	+ intsvy	9 months ago
README.md		

Programme for International Student Assessment 2012

PISA in teaching: It's a great blasting site for students' projects

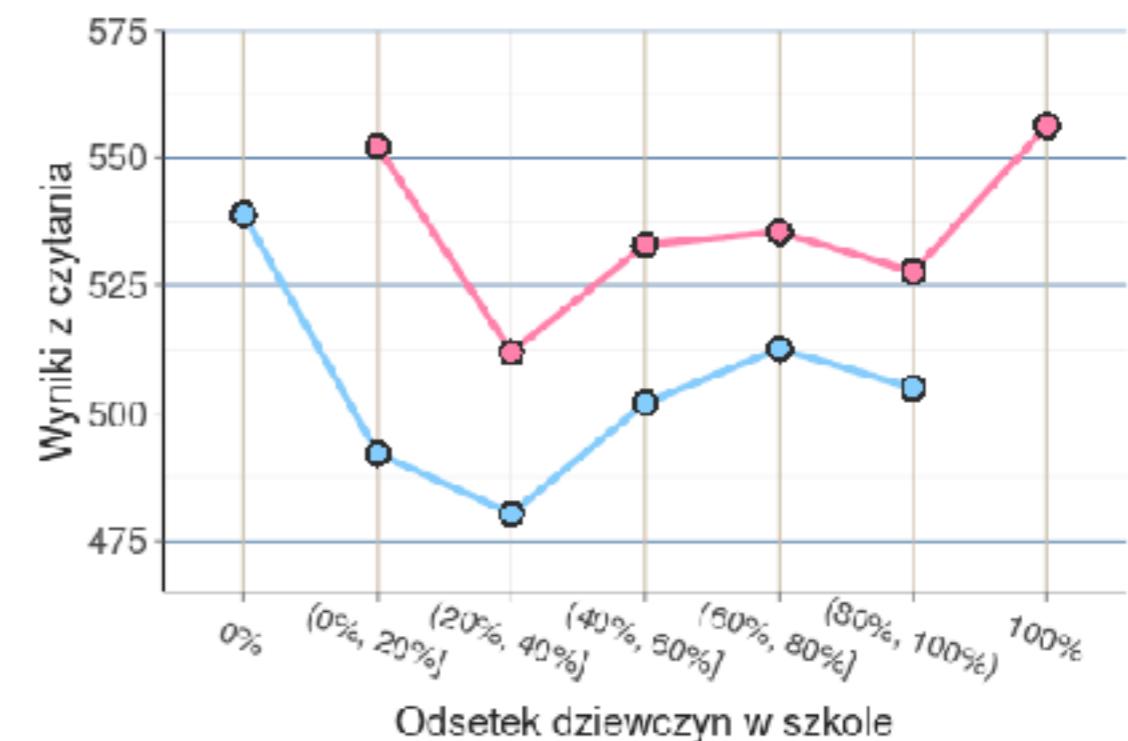
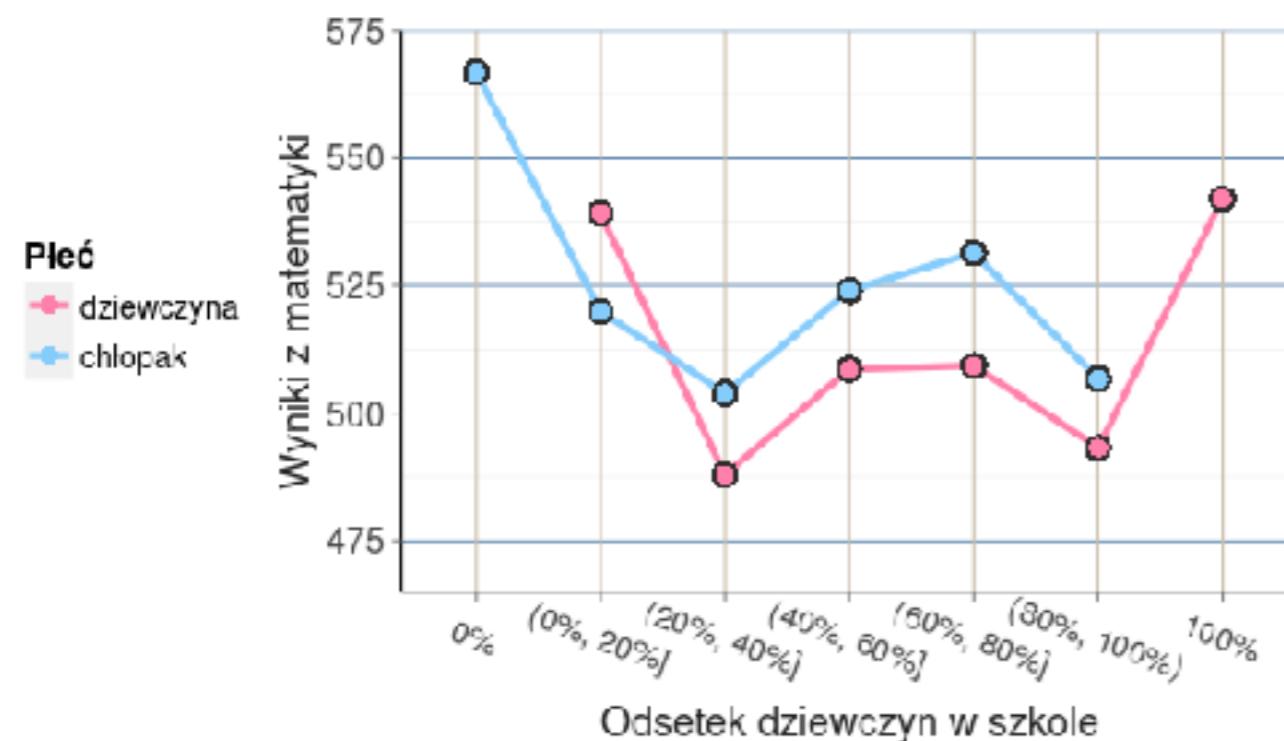
Socio-economic background **DOES** matter in results of PISA math Test

Mean score on the math test for each country

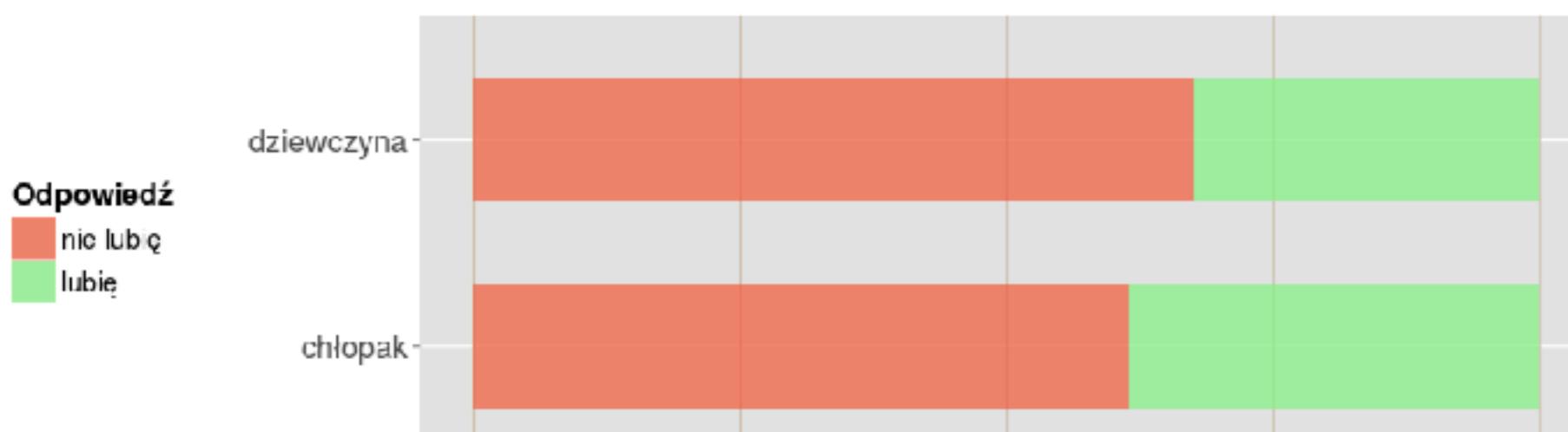


PISA in teaching: It's a great blasting site for students' projects

Czy obecność płci przeciwej w szkole obniża wyniki w nauce ?



Czy uczniowie lubią matematykę ?





education data education evidence education policy education analysis education statistics education data education evidence education policy

Do parents' occupations have an impact on student performance?

- Students whose parents work in professional occupations generally outperform other students in mathematics, while students whose parents work in elementary occupations tend to underachieve compared to their peers.
- The strength of the relationship between parents' occupations and student performance varies considerably across countries: for example, when it comes to mathematics performance, the children of cleaners in Shanghai-China outperform the children of professionals in the United States, and the children of professionals in Germany outperform the children of professionals in Finland, on average.
- Finland and Japan achieve high levels of performance by ensuring that the children of parents who work in elementary occupations are given the same education opportunities and the same encouragement as the children of professionals.

**Q20 What is your father's main job?
(e.g., school teacher, cook, sales manager)**

(If he is not working now, please tell us his last main job.)

Please write in the job title. _____

**Q21 What does your father do in his main job?
(e.g., teaches high school students, helps prepare meals in a restaurant, manages a sales team)**

Please use a sentence to describe the kind of work he does or did in that job.

ISCO classification for Occupations <http://bit.ly/2dKE8K4>

Element

Example

ISCO level 1: major group

2 Professionals

ISCO level 2: sub-major group

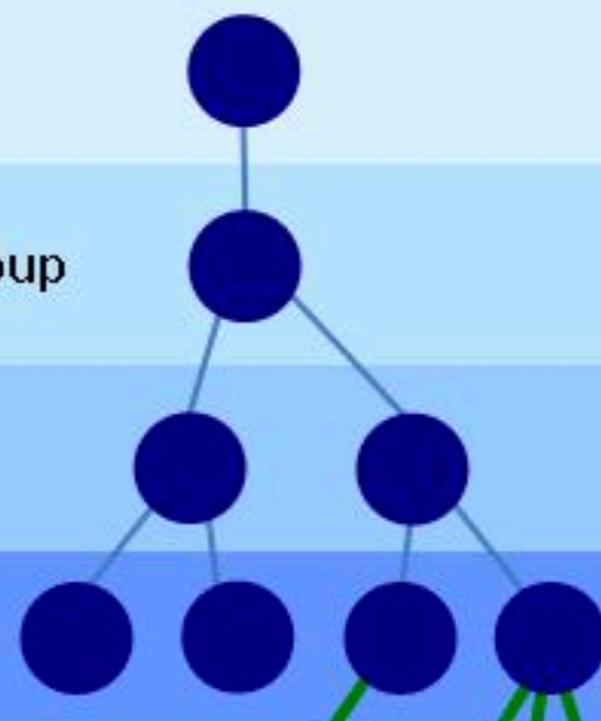
21 Computing, engineering and science professionals

ISCO level 3: minor group

213 Computing professionals

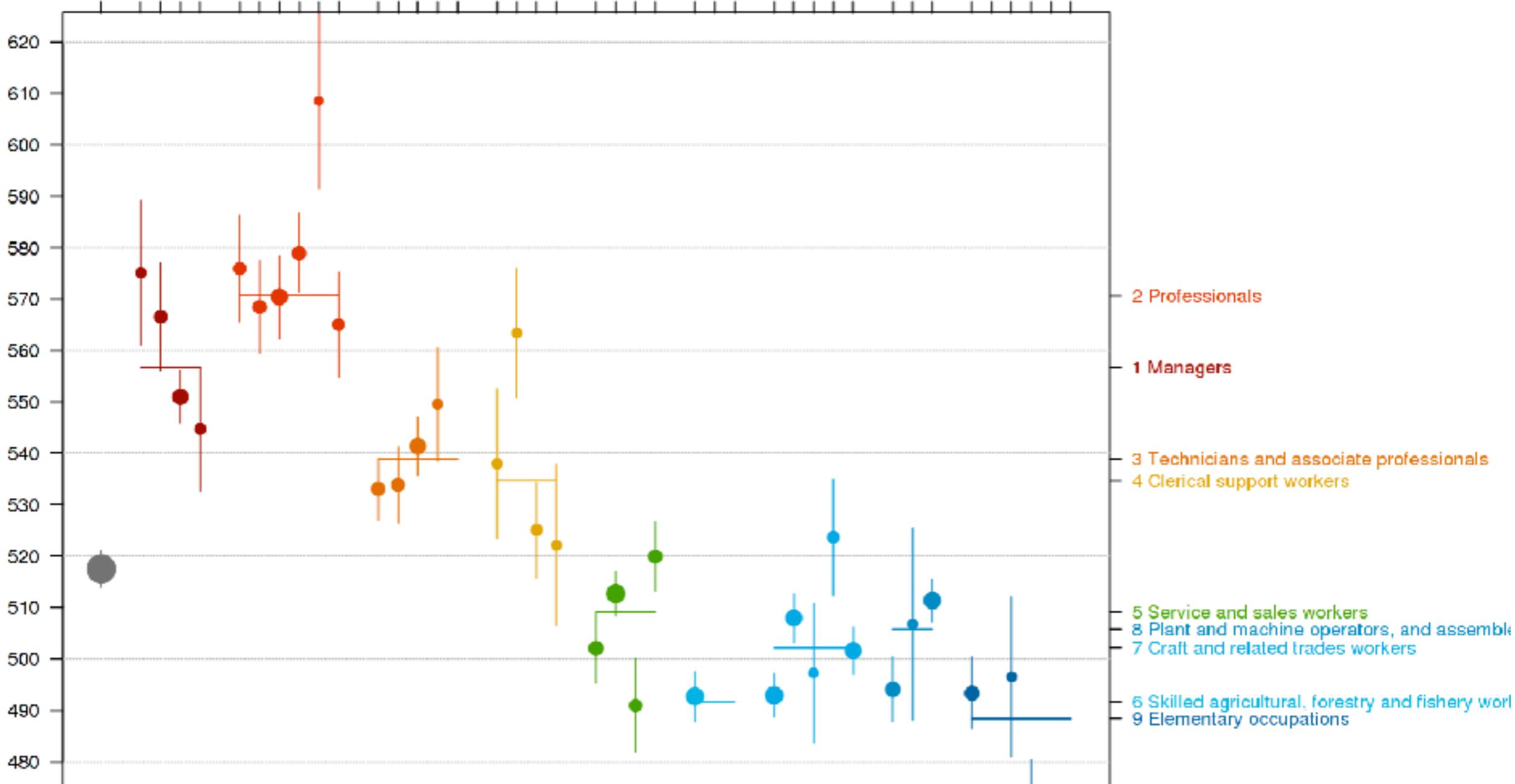
ISCO level 4: unit group

2132 Computer programmers



ISCO

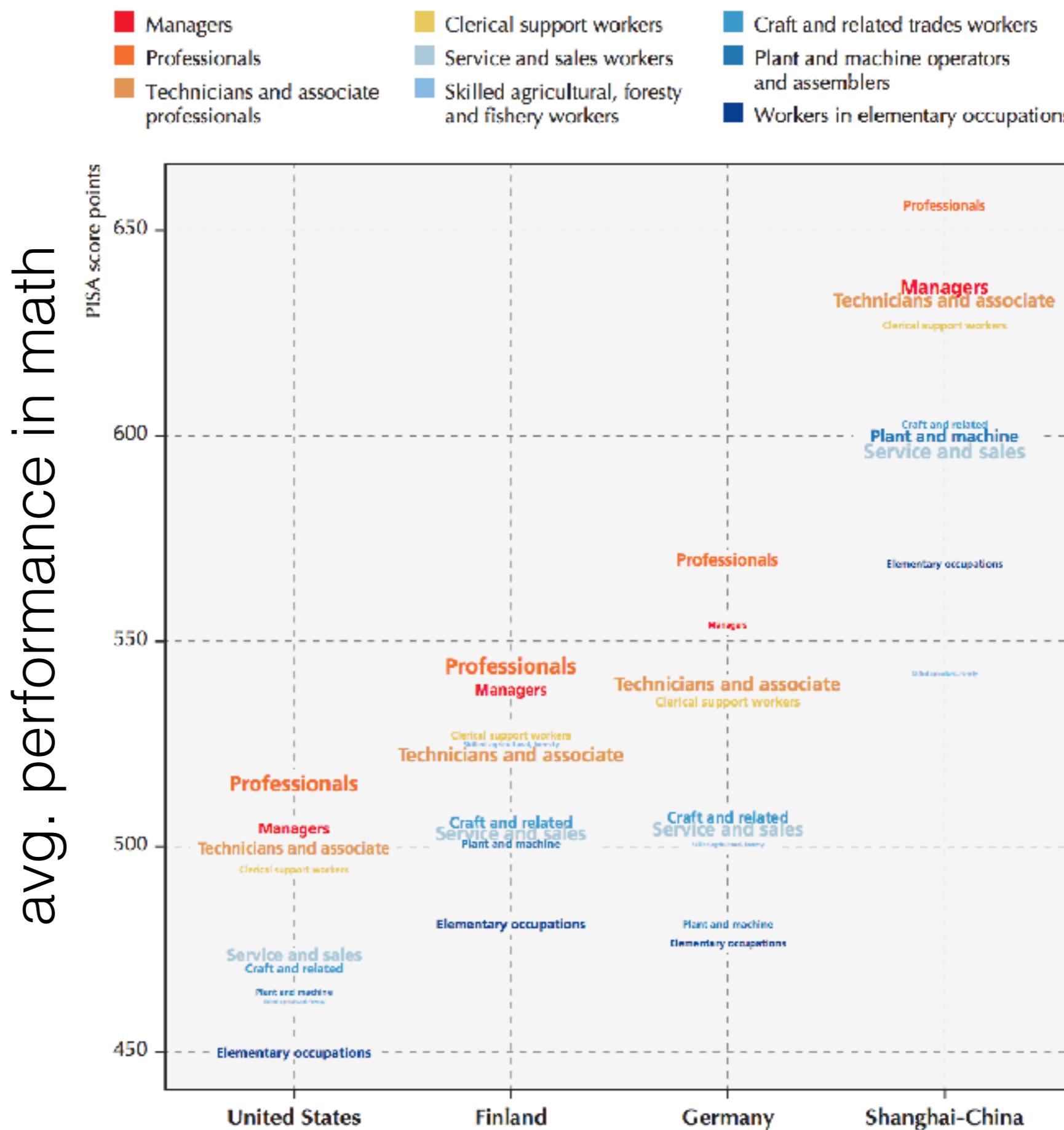
avg. performance in math



Student performance in mathematics in Germany and Finland, by parents' occupations



Student performance in mathematics in the United States, Finland, Germany and Shanghai-China, by parents' occupations



Explore the story for your country with the shiny app <http://bit.ly/1W2TTKf>

Occupations@PISA2012

How much can we infer about a student's performance in school by looking at what his or her parents do for a living? To find out, PISA 2012 asked participating students their parents' occupations. **Occupations@PISA2012** is a web-based application that allows you to explore the relationship between parents' occupations and their children's performance in mathematics, reading and science - in your own country and in other countries.

Choose perspective ->

Compare results by subject

Mathematics

and country

Germany

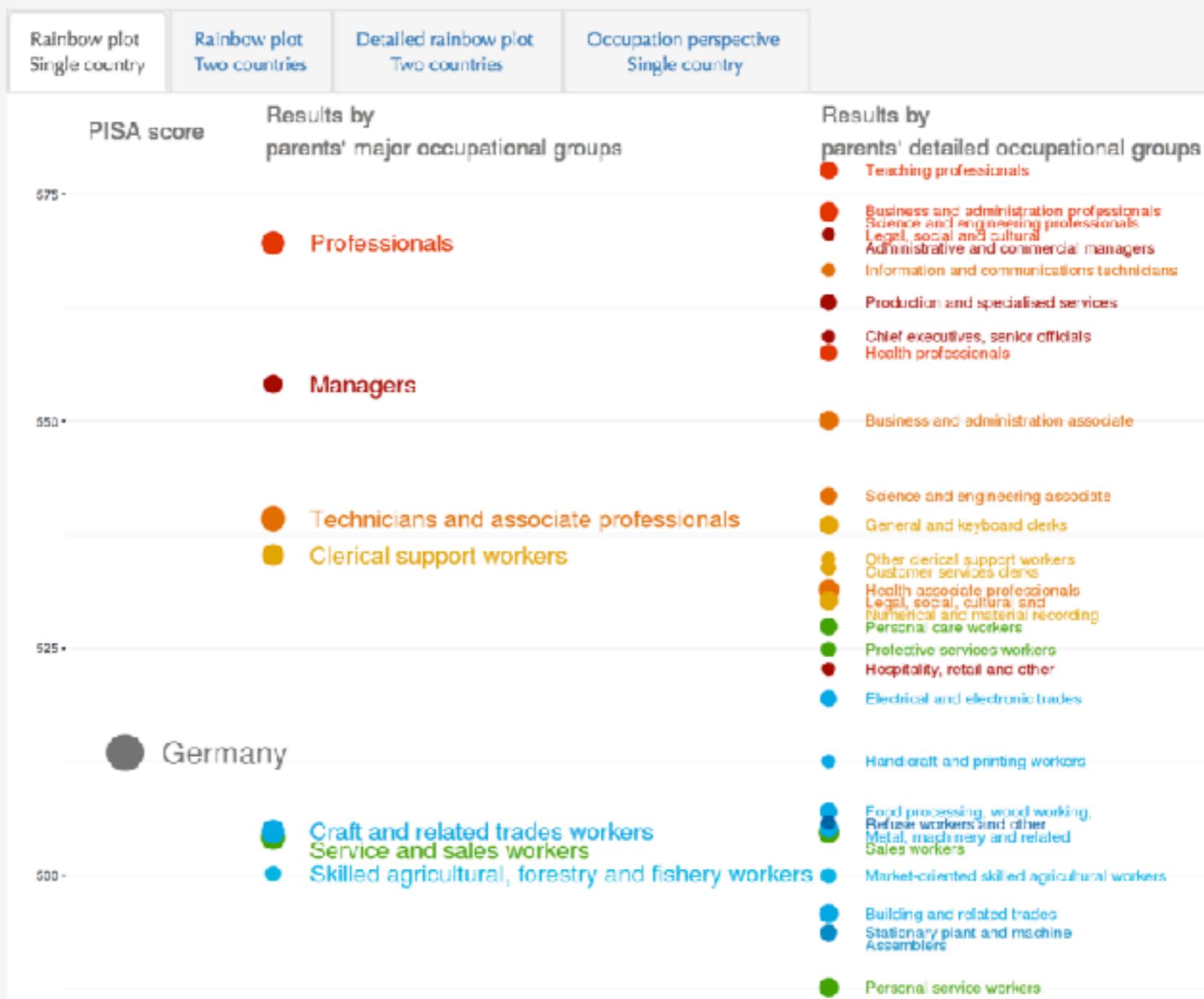
Set range for vertical axis:



You can find more interesting results in our [Pisa In Focus article](#).

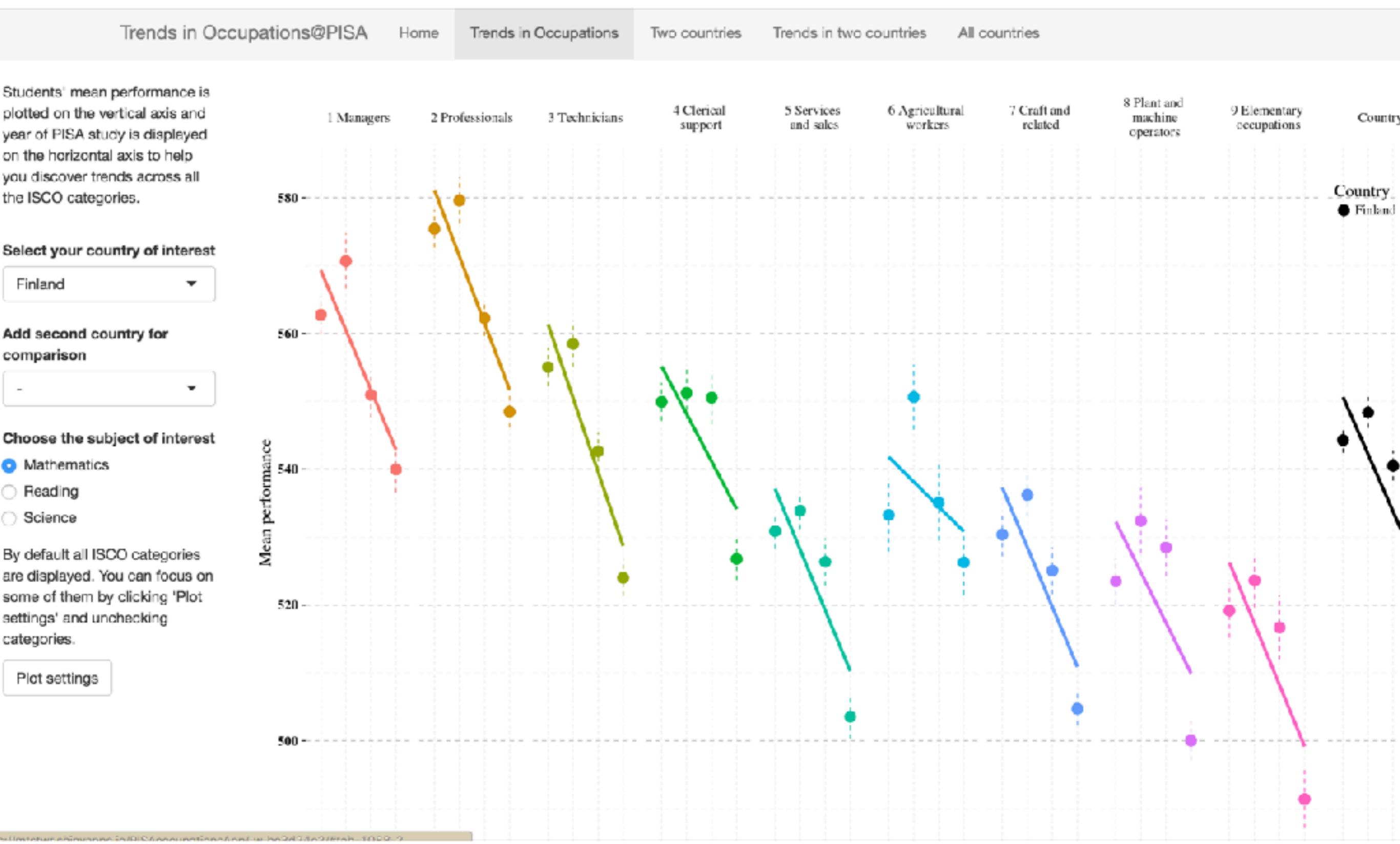
Here are country profiles [as pdf file](#), and raw data [as Excel file](#) and regional data [as Excel file](#).

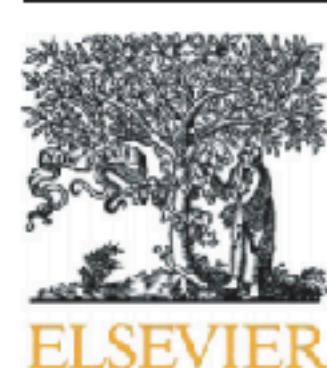
The app is prepared by [Przemyslaw Biecek](#) as a part of IJA fellowship.



An ongoing project with Mateusz Staniak

Trends in Occupations <http://bit.ly/2e60j85>





An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test



Francesca Borgonovi ^{a,*}, Przemyslaw Biecek ^b

^a Organisation for Economic Co-operation and Development, 2 Rue André Pascal, 75116 Paris, France

^b Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw, Poland

ARTICLE INFO

Article history:

Received 20 May 2015

Received in revised form 20 April 2016

Accepted 3 June 2016

Available online xxxx

ABSTRACT

The paper examines academic endurance, defined as students' ability to maintain their baseline rates of successful test completion for the duration of a low-stakes cognitive test. The random allocation of test booklets to students in the Programme for International Student Assessment (15-year-olds, 60+ countries, N > 450,000) offers a natural experiment to examine whether the decline in students' correct responses to test items placed at the beginning vs. at the end of the booklet varies across countries, population subgroups, and assessment domains. Results reveal differences in academic endurance both between countries and within countries across different population subgroups. The decline in students' correct responses tends to be more pronounced among boys, among socio-economically disadvantaged students and when students solve the reading part of the PISA test. In some countries students' endurance in reading, mathematics and science depends on the response format of the assessment questions.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Large scale international assessments are one of the instruments

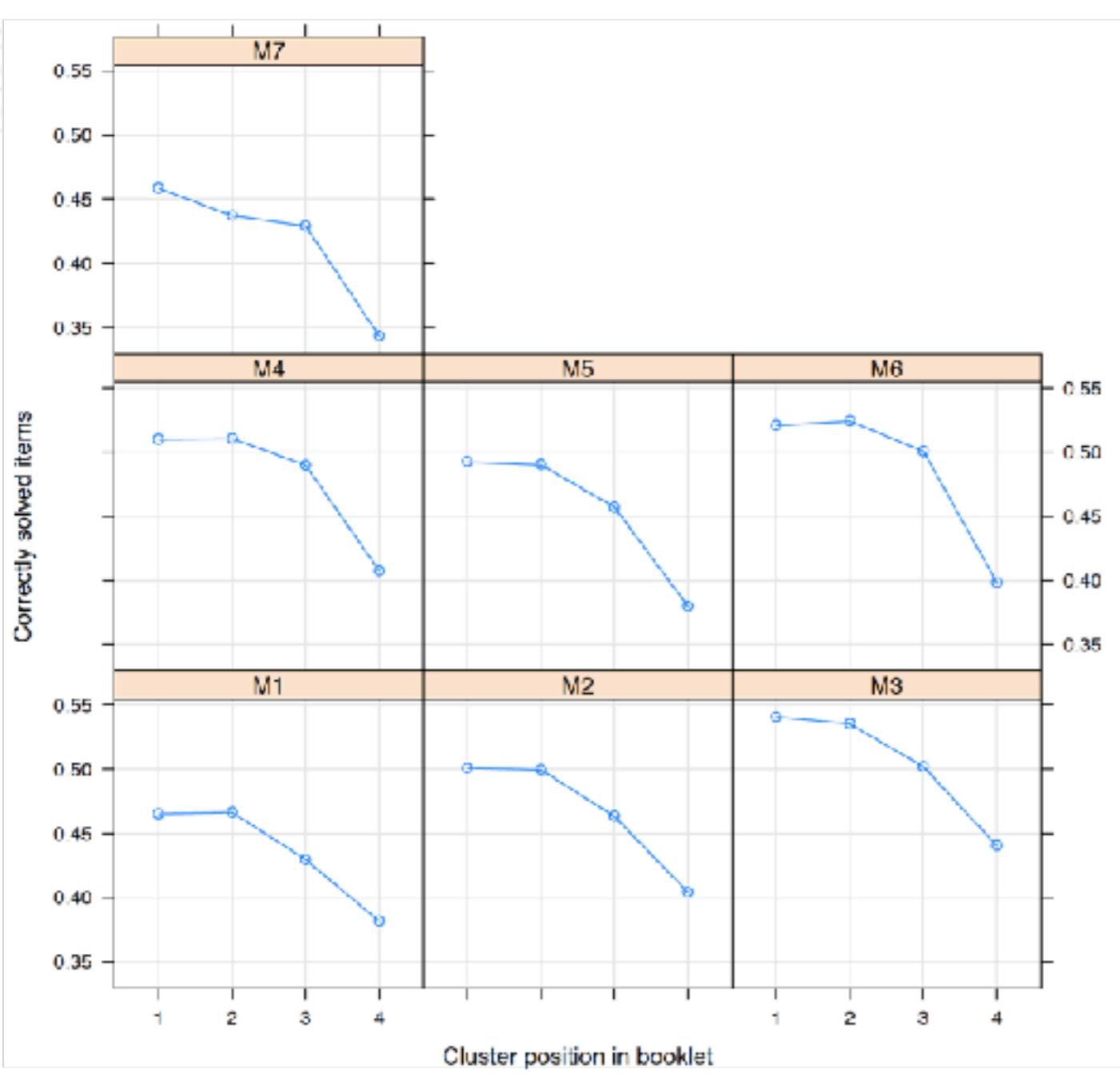
2. Theory and hypotheses

2.1. Theoretical framework

2006			
Position 1	Position 2	Position 3	Position 4
Science 1	Science 2	Science 4	Science 7
Science 2	Science 3	Maths 3	Reading 1
Science 3	Science 4	Maths 4	Maths 1
Science 4	Maths 3	Science 5	Maths 2
Science 5	Science 6	Science 7	Science 3
Science 6	Reading 2	Reading 1	Science 4
Science 7	Reading 1	Maths 2	Maths 4
Maths 1	Maths 2	Science 2	Science 6
Maths 2	Science 1	Science 3	Reading 2
Maths 3	Maths 4	Science 6	Science 1

2009				
	Position 1	Position 2	Position 3	Position 4
Booklet 1	Maths 1	Reading 1	Reading 3A	Maths 3
Booklet 2	Reading 1	Science 1	Reading 4A	Reading 7
Booklet 3	Science 1	Reading 3A	Maths 2	Science 3
Booklet 4	Reading 3A	Reading 4A	Science 2	Reading 2
Booklet 5	Reading 4A	Maths 2	Reading 5	Maths 1
Booklet 6	Reading 5	Reading 6	Reading 7	Reading 3A
Booklet 7	Reading 6	Maths 3	Science 3	Reading 4A
Booklet 8	Reading 2	Maths 1	Science 1	Reading 6
Booklet 9	Maths 2	Science 2	Reading 6	Reading 1
Booklet 10	Science 2	Reading 5	Maths 3	Science 1

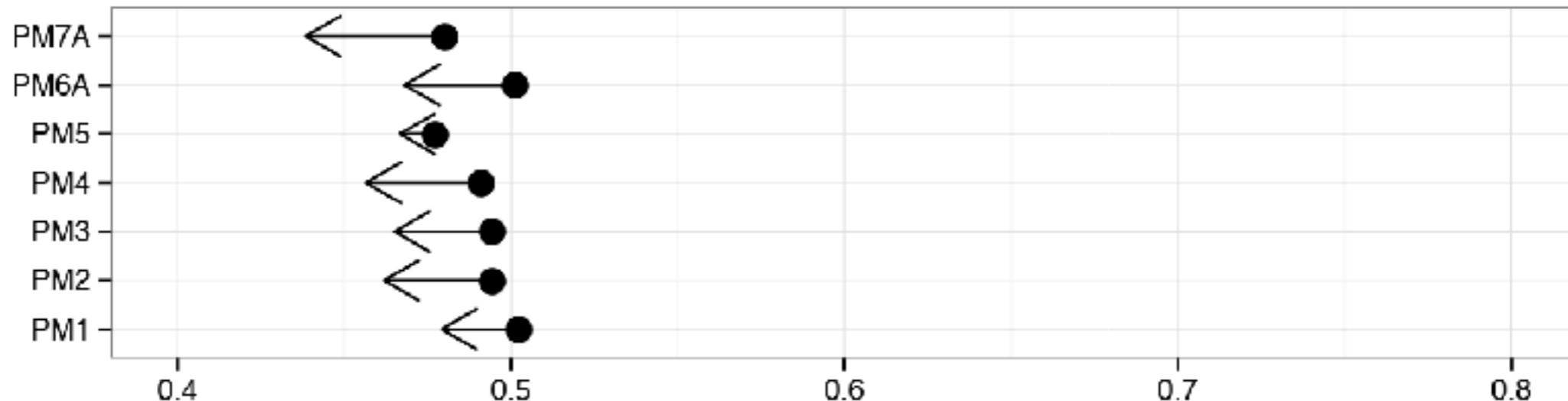
2012				
	Position 1	Position 2	Position 3	Position 4
Booklet 1	Maths 5	Science 3	Maths 6A	Science 2
Booklet 2	Science 3	Reading 3	Maths 7A	Reading 2
Booklet 3	Reading 3	Maths 6A	Science 1	Maths 3
Booklet 4	Maths 6A	Maths 7A	Reading 1	Maths 4
Booklet 5	Maths 7A	Science 1	Maths 1	Maths 5
Booklet 6	Maths 1	Maths 2	Reading 2	Maths 6A
Booklet 7	Maths 2	Science 2	Maths 3	Maths 7A
Booklet 8	Science 2	Reading 2	Maths 4	Science 1
Booklet 9	Reading 2	Maths 3	Maths 5	Reading 1
Booklet 10	Maths 3	Maths 4	Science 3	Maths 1
Booklet 11	Maths 4	Maths 5	Reading 3	Maths 2
Booklet 12	Science 1	Reading 1	Maths 2	Science 3
Booklet 13	Reading 1	Maths 1	Science 2	Reading 3



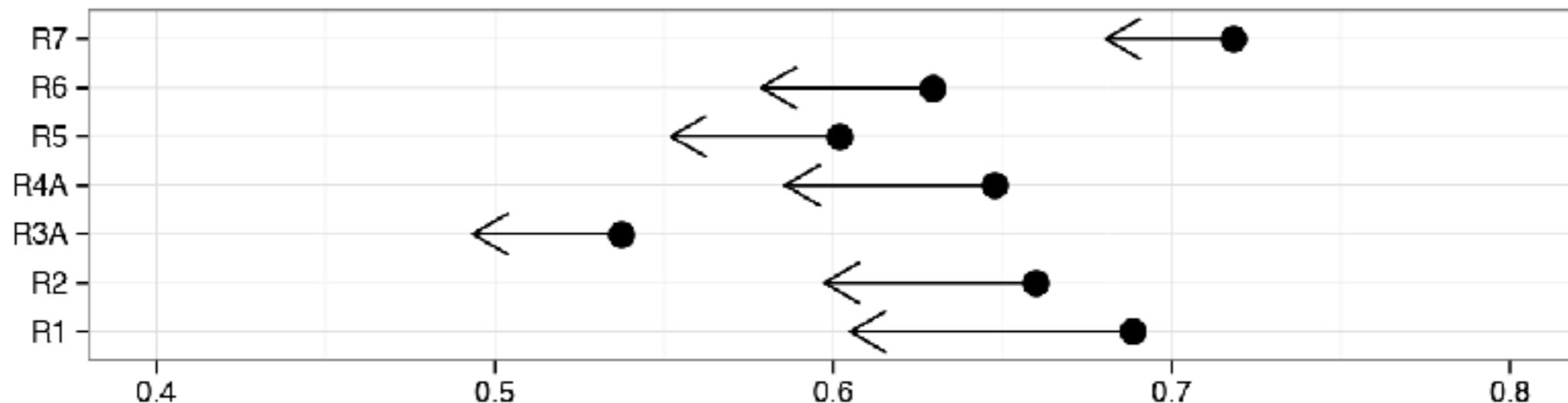
Student's performance declines with time

But some students are more resistant and able to keep high performance over time

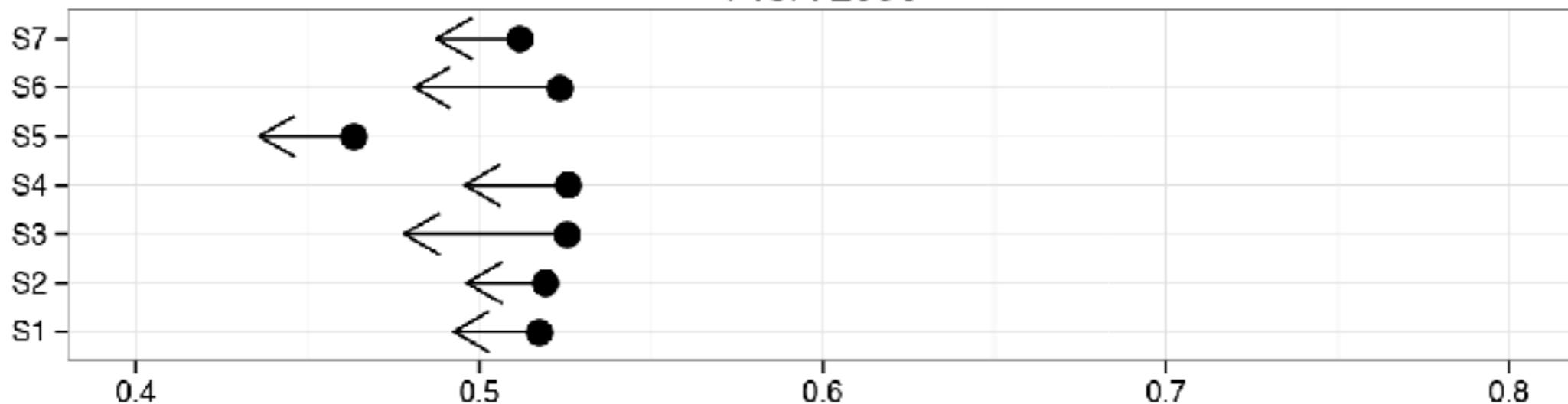
PISA 2012



PISA 2009

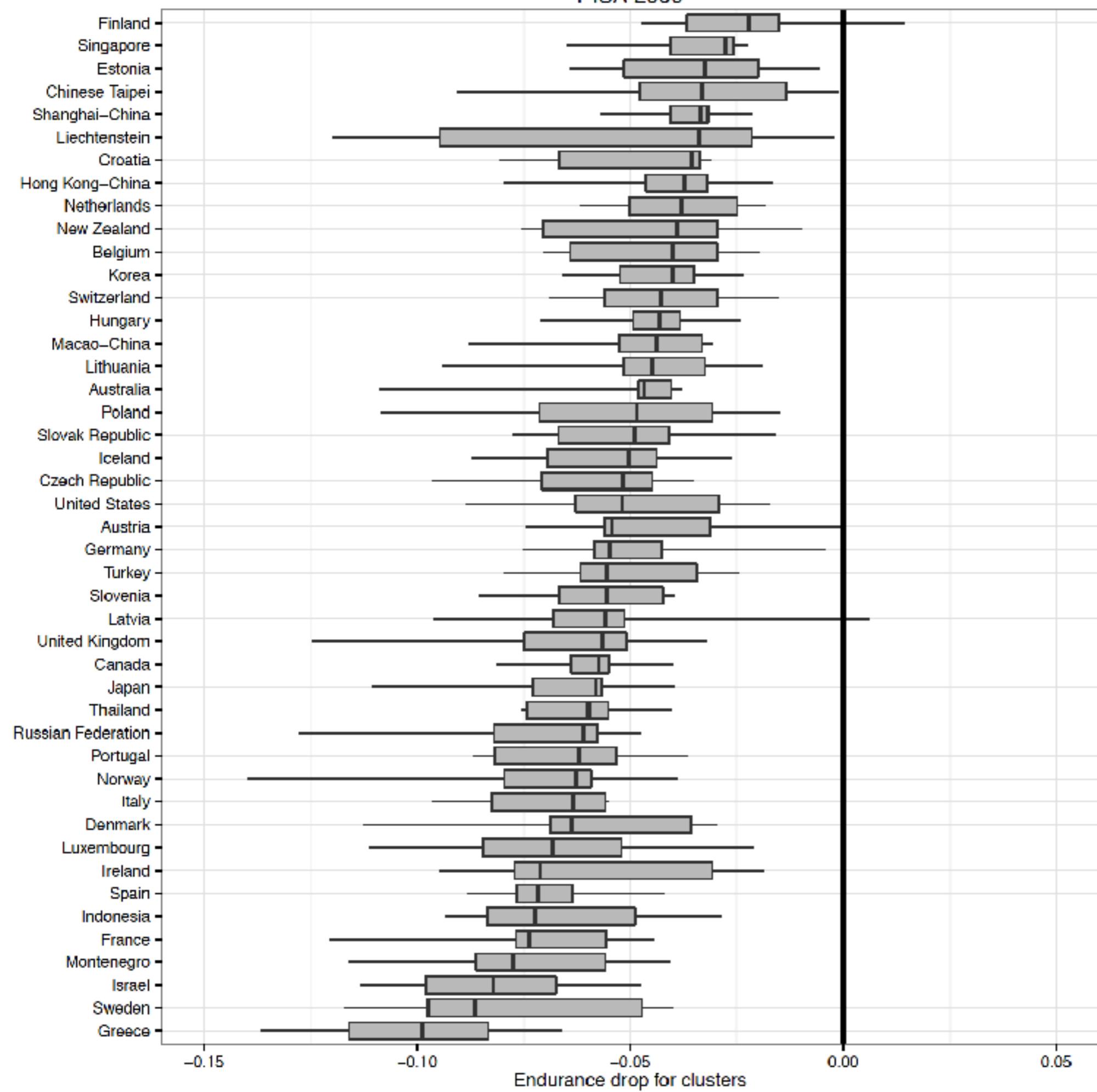


PISA 2006



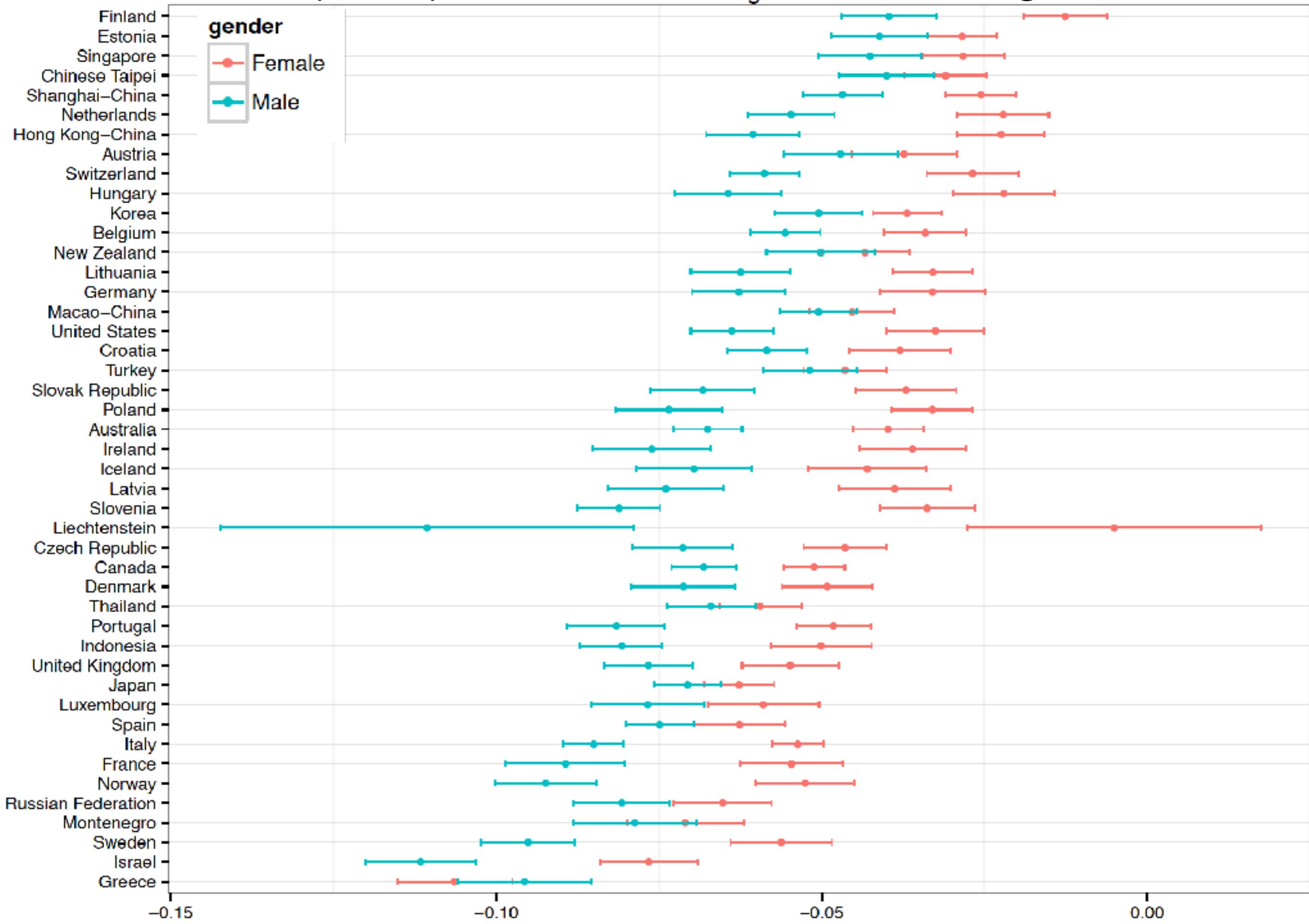
avg. performance

PISA 2009



Drops in performance are different for different countries

Drops in performance differ between genders



PISA Data
Analysis Manual
SPSS® SECOND EDITION

PISA Data
Analysis Manual
SAS® SECOND EDITION



Programme for International Student Assessment



Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II.

<http://www.jstatsoft.org/>

intsvy: An R Package for Analysing International Large-Scale Assessment Data

Daniel H. Caro
University of Oxford

Przemyslaw Biecek
University of Warsaw

Abstract

This paper introduces `intsvy`, an R package for working with international assessment data (e.g., PISA, TIMSS, PIRLS). The package includes functions for importing data, performing data analysis, and visualising results. The paper describes the underlying methodology and provides real data examples. Tools for importing data allow useRs to select variables from student, home, school, and teacher survey instruments as well as for specific countries. Data analysis functions take into account the complex sample design (with replicate weights) and rotated test forms (with plausible values of achievement scores) in the calculation of point estimates and standard errors of means, standard deviations, regression coefficients, correlation coefficients, and frequency tables. Visualisation tools present data aggregates in standardised graphical form.

intsvy: International Assessment Data Manager

intsvy is an R package for working with international assessment data from PISA (Programme for International Student Assessment), TIMSS (Science Study), PIRLS (Progress in International Reading Literacy Study), PIAAC (Programme for the International Assessment of Adult Competencies) and Information Literacy Study (ICILS). The package includes functions for importing data, performing data analysis, and visualising results. It handles complex sample design (with replicate weights) and rotated test forms (with plausible achievement values) in the calculation of point estimates, standard deviations, regression coefficients, correlation coefficients, and frequency tables. Visualization tools present data aggregates in standardised graphical formats.

How to install and load 'intsvy' into R?

Package intsvy is available from [CRAN](#). It can be installed and loaded into R, as follows:

```
> install.packages('intsvy')  
  
> library(intsvy)
```

This is the recommended installation method. Alternatively, users may install and use the [development version](#) of 'intsvy' available in GitHub.

```
> install.packages("devtools")  
  
> library(devtools)  
  
> install_github("cldafani/intsvy")
```

How to use 'intsvy' for the analysis of international assessment data?

Documentation for R intsvy functions can be accessed via `help(package='intsvy')`. More resources to learn how to use 'intsvy' are provided below.

The 'intsvy' tutorial

The intsvy tutorial can be downloaded from [here](#).

Datasets and preliminary results for PISA 2015
will be published on December 6, 2016 at 11h00
<https://www.oecd.org/pisa/pisaproducts/>



<http://www.matholympus.com/>

Educate policy makers! Reach them early!

Introduce the data science to primary and secondary schools

Beta bit



[Amazing Adventures of Beta and Bit](#)



[Stories](#)



[Games](#)



[Gadgets](#)

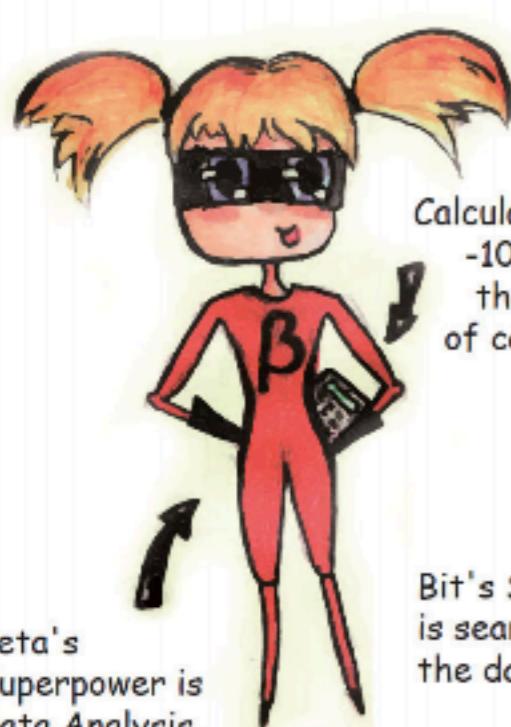
Adventures of Beta and Bit

How to weigh a dog with a ruler?



Beta, who has a passion for mathematics, chess and good books, changes into SuperBeta under the influence of puzzles.

Text: Przemysław Biećk
Illustrations: Klaudia Korniluk



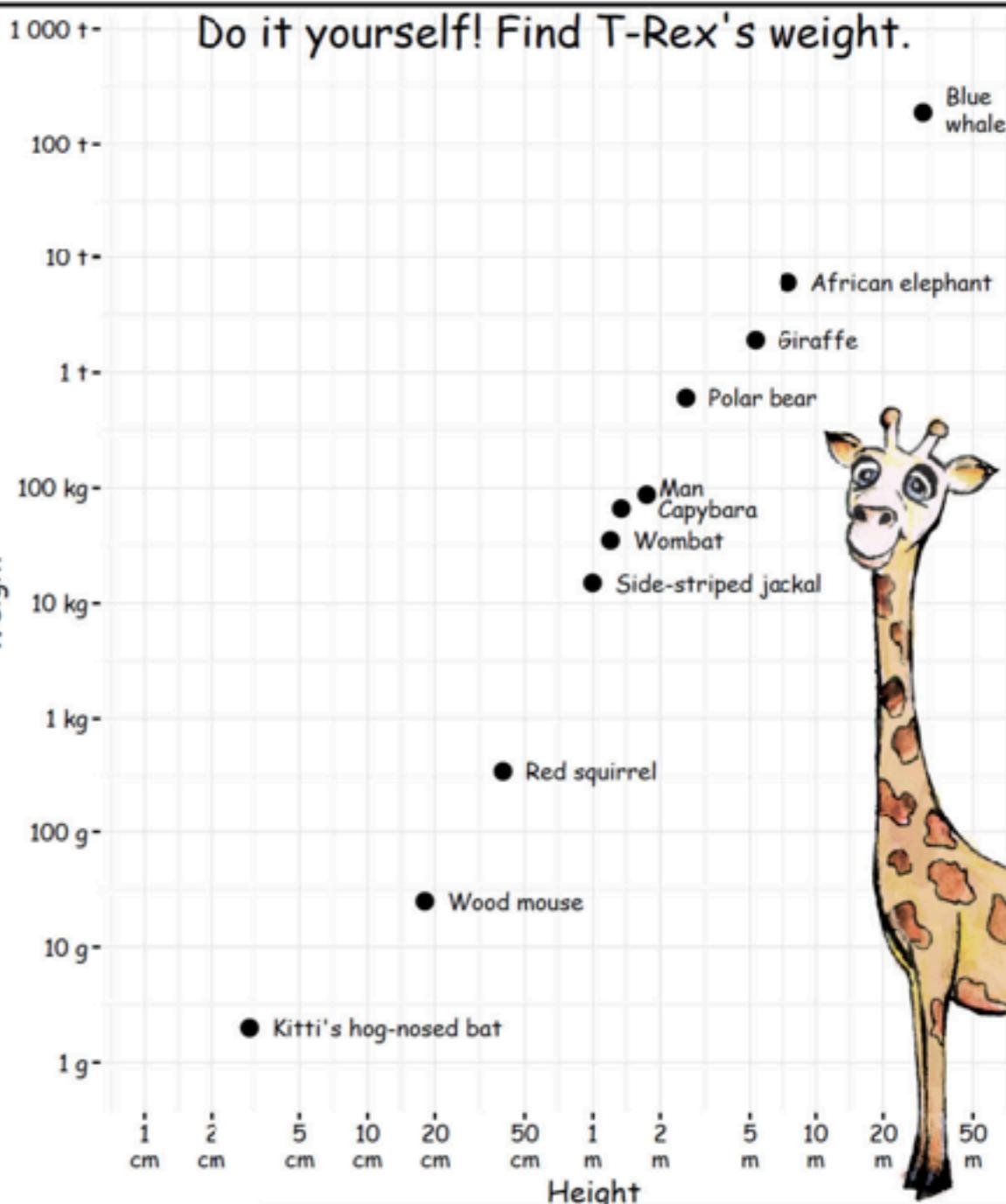
Beta's Superpower is Data Analysis.



Superglasses:
+10 to the speed of browsing the Internet.

Bit, who is a computer, programming as well as robot maniac, changes into SuperBit under the influence of puzzles.

The early spring: Beta and Bit are wandering through the park.



On the above-mentioned chart it was marked the height and weight of different animals. From a small bat to a giant whale. Having such a combination try to estimate how much the 10 meter dinosaur weighed.

(1) Draw a line passing close to the marked points, as you did for the dogs.

(2) Having such a relationship estimate T-Rex's weight (*), find out what weight corresponds to 10 meters of height.

<http://betabit.wiki/>

```
> library(BetaBit)
> proton()
-----  
[1] "Proton"  
-----  
[1] "Proton"  
-----  
[1] "Proton"  
-----  
[1] "Proton"
```

Ready to hack the Pietraszko's password?
The Proton Game is a text based R console game. Just install and load the BetaBit package from CRAN.

```
install.packages("BetaBit")
library("BetaBit")
proton()
```

The Python version.

This time you have to do the frequency analysis of ciphertext. **The Frequon Game** is a text based R console game. Just install and load the Betabit package

```
install.packages("BetaBit")
library("BetaBit")
frequon()
```

ARTSCIENCE MUSEUM™ PRESENTS

BIG BANG DATA

WHAT DOES DATA MEAN TO YOU?



ACTIVITY BOOKLET



A study done in 2015 shows how often people post about their food experience on social media...



Notes on Matters Affecting the Health of Our Troops in the Indian Army - Florence Nightingale 1856

RADAR CHART:
To compare the total values of multiple data series.

In the Neck'n' Cholera Section of Cholera - John Snow 1856

BAR CHART:
To compare and track the changes of different data over a period of time.

Recruiting Area of Subsequent Men Lost in the French Army During the Russian Campaign 1812-1813 - Charles Joseph Minard 1869

LINE CHART:
To illustrate trends from data collected over a long time period.

2.

What have you discovered from the chart you created on the previous page?

Beta



Random thoughts:

- take sabbaticals!
- Paris is a lovely place!
- use shiny!
- play with PISA data!
- there is an R package waiting for you!

Thank you!

