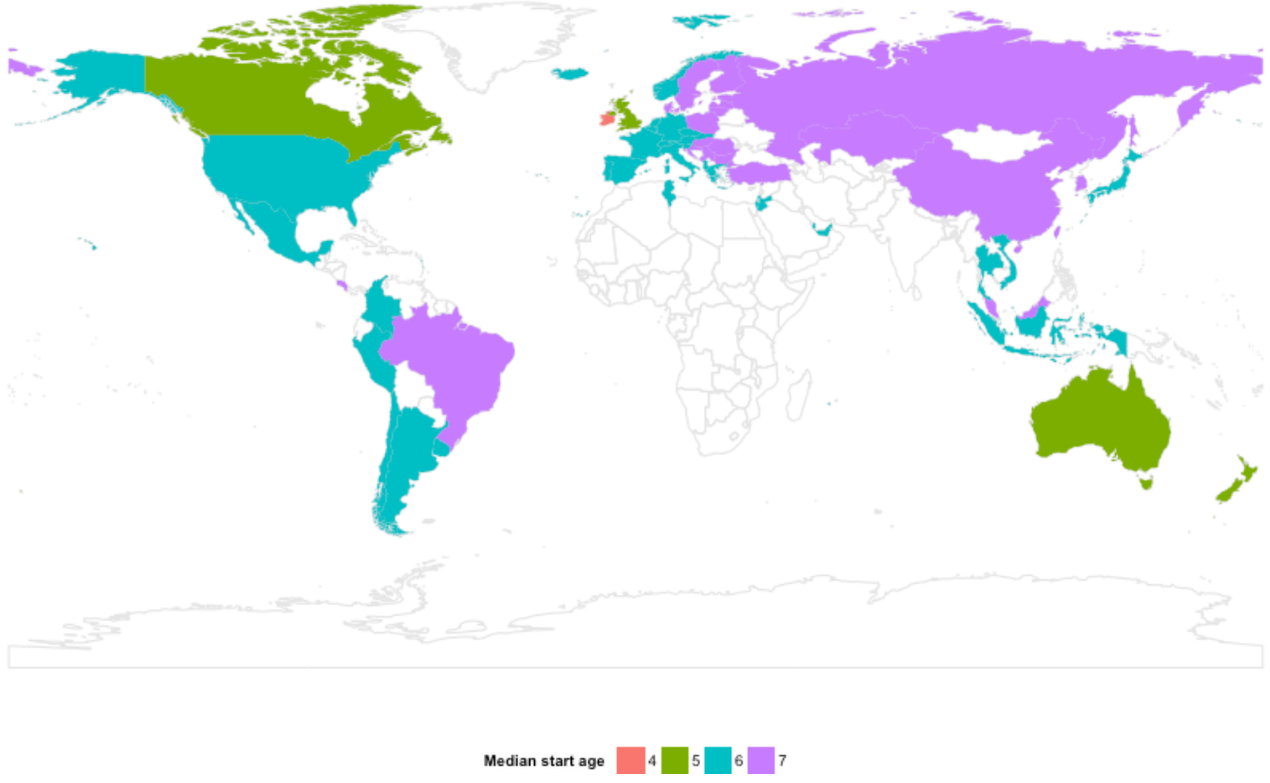
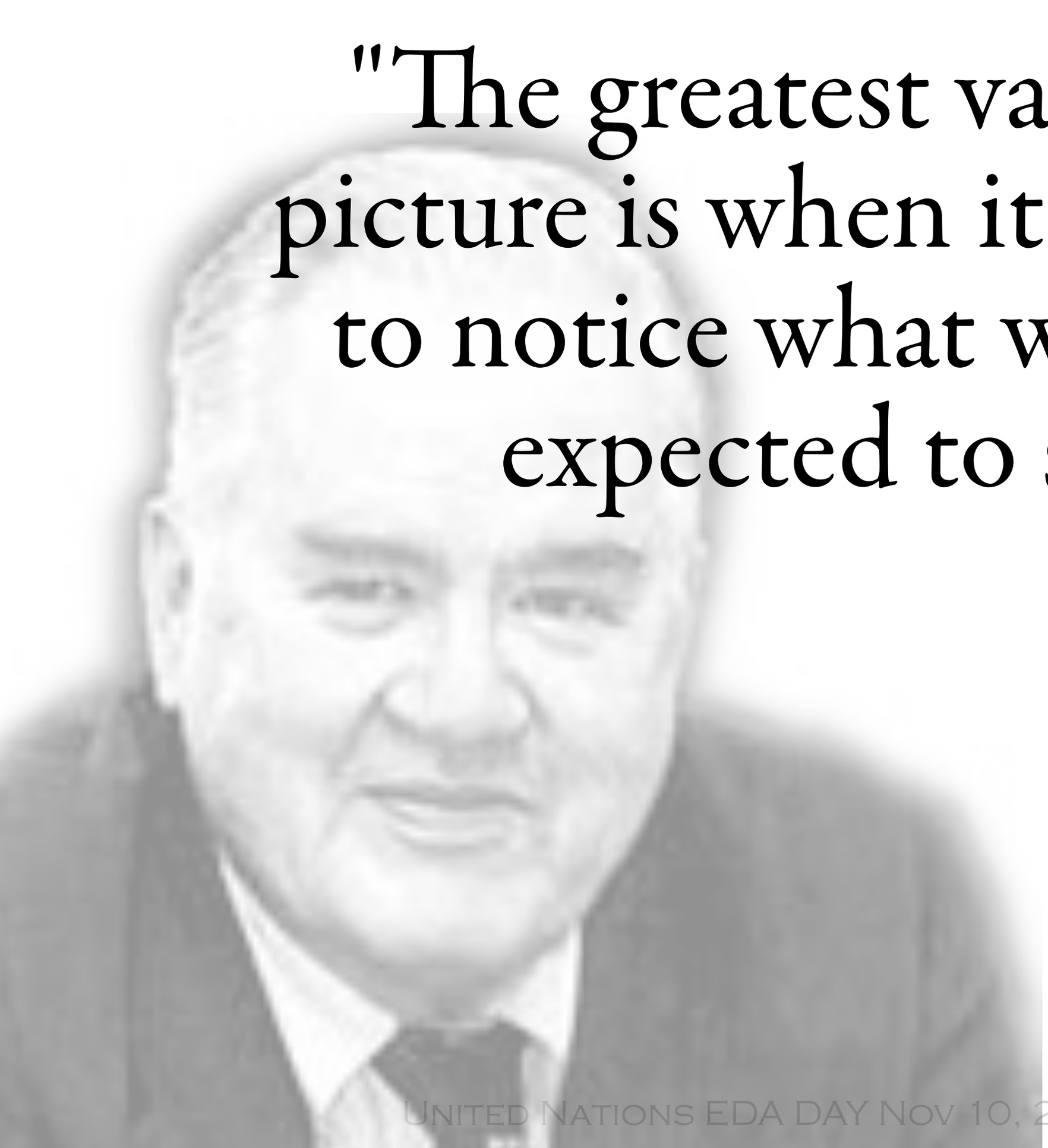
An Exploratory Data Analysis of OECD's 2012 PISA Surveys

Di Cook
Iowa State University

Joint exploration with
Luke Fostveldt, Alex Shum,
Ian Lyttle



United Nations EDA DAY Nov 10 2014



"The greatest value of a picture is when it forces us to notice what we never expected to see."

Education - my expectations

- Gender gap: “boys are better at math”
- USA doesn’t perform well compared to other countries
- Would expect parents, time studying and socioeconomic status to be important.

DATA

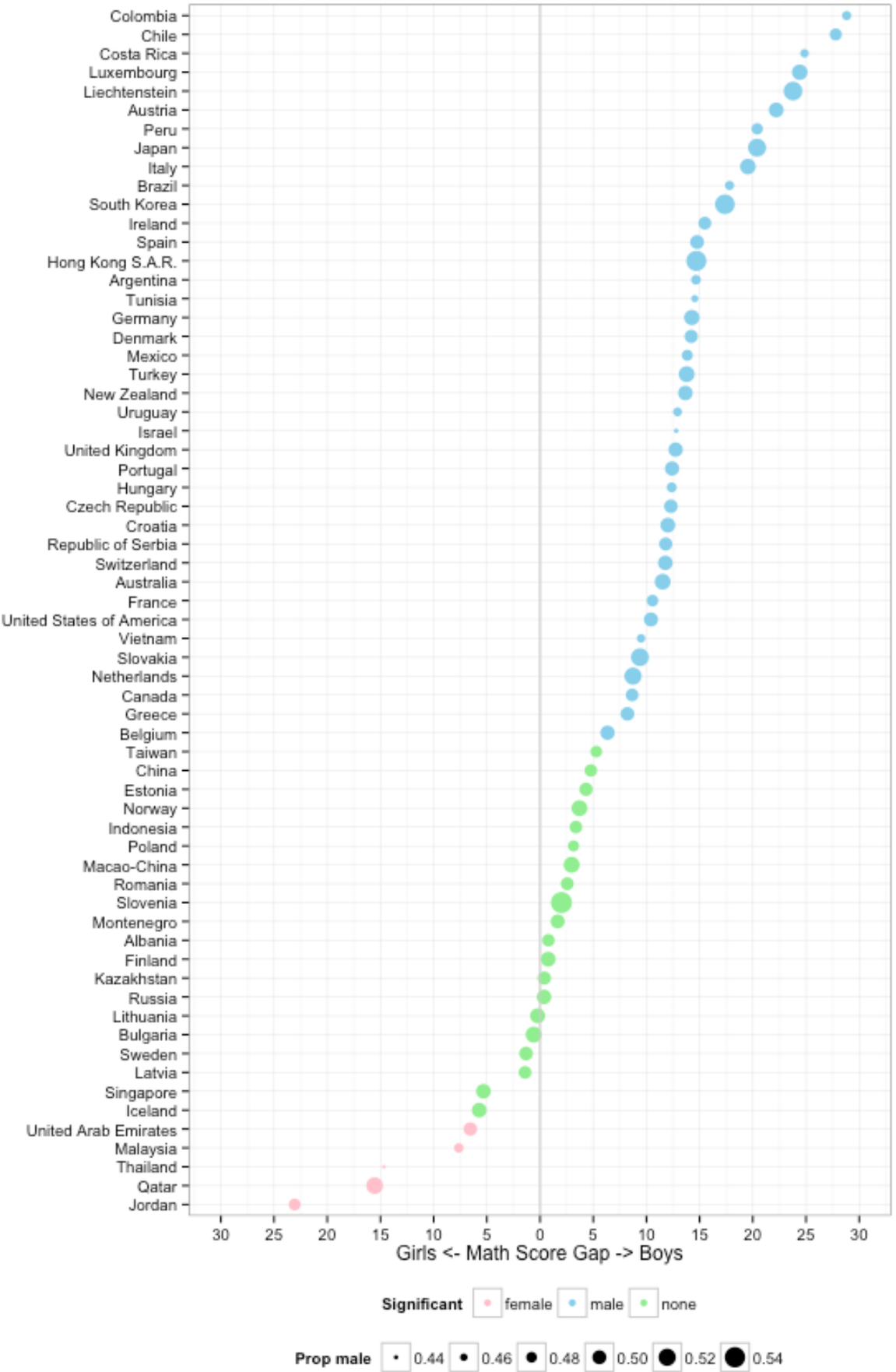
Results from 2012 OECD PISA tests: "the world's global metric for quality, equity and efficiency in school education".

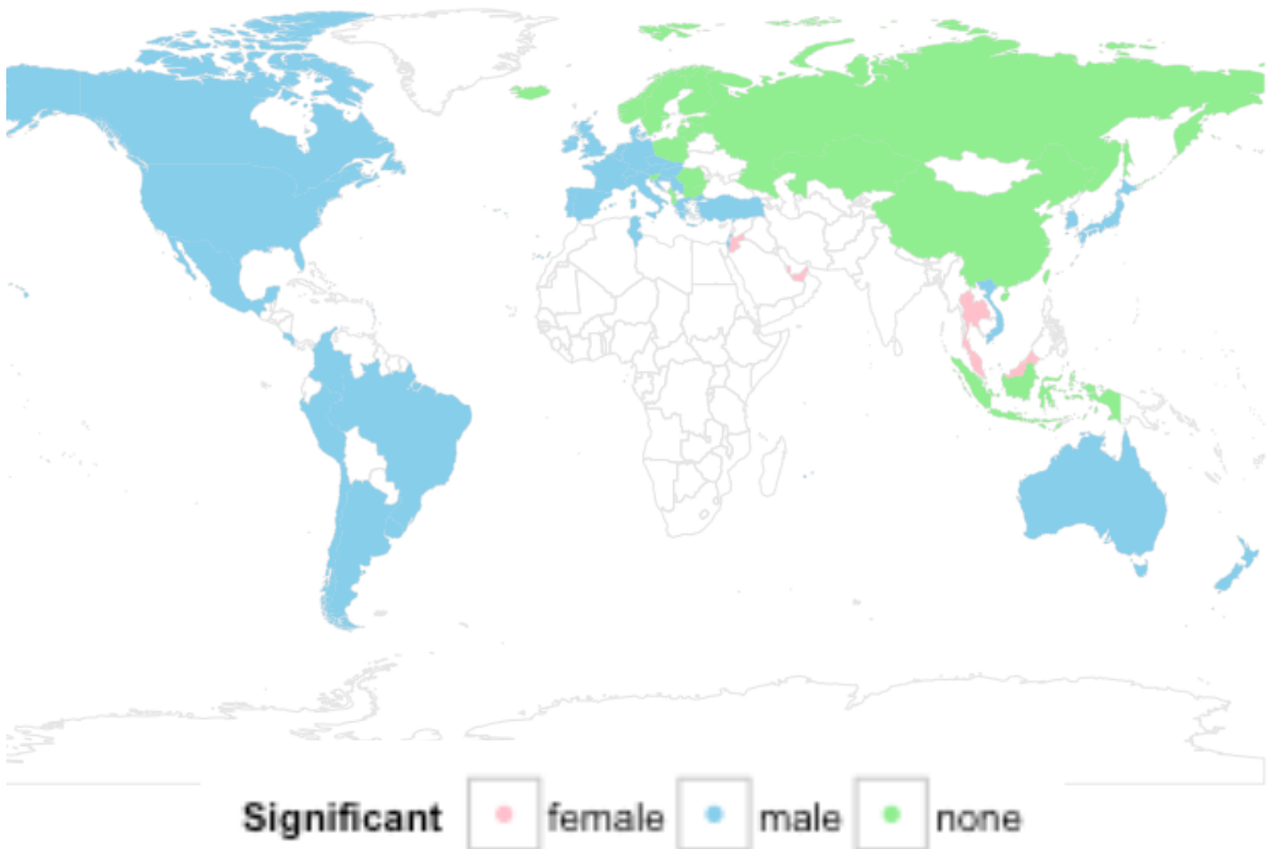
- 485,490 students math, science and reading test scores
- 65 countries, between 100-1500 schools in each
- Student questionnaires about their environment (635 vars)
- Parents surveyed on work, life, income (143 vars)
- Principals provide information about their schools (291 vars)

“Show the data."

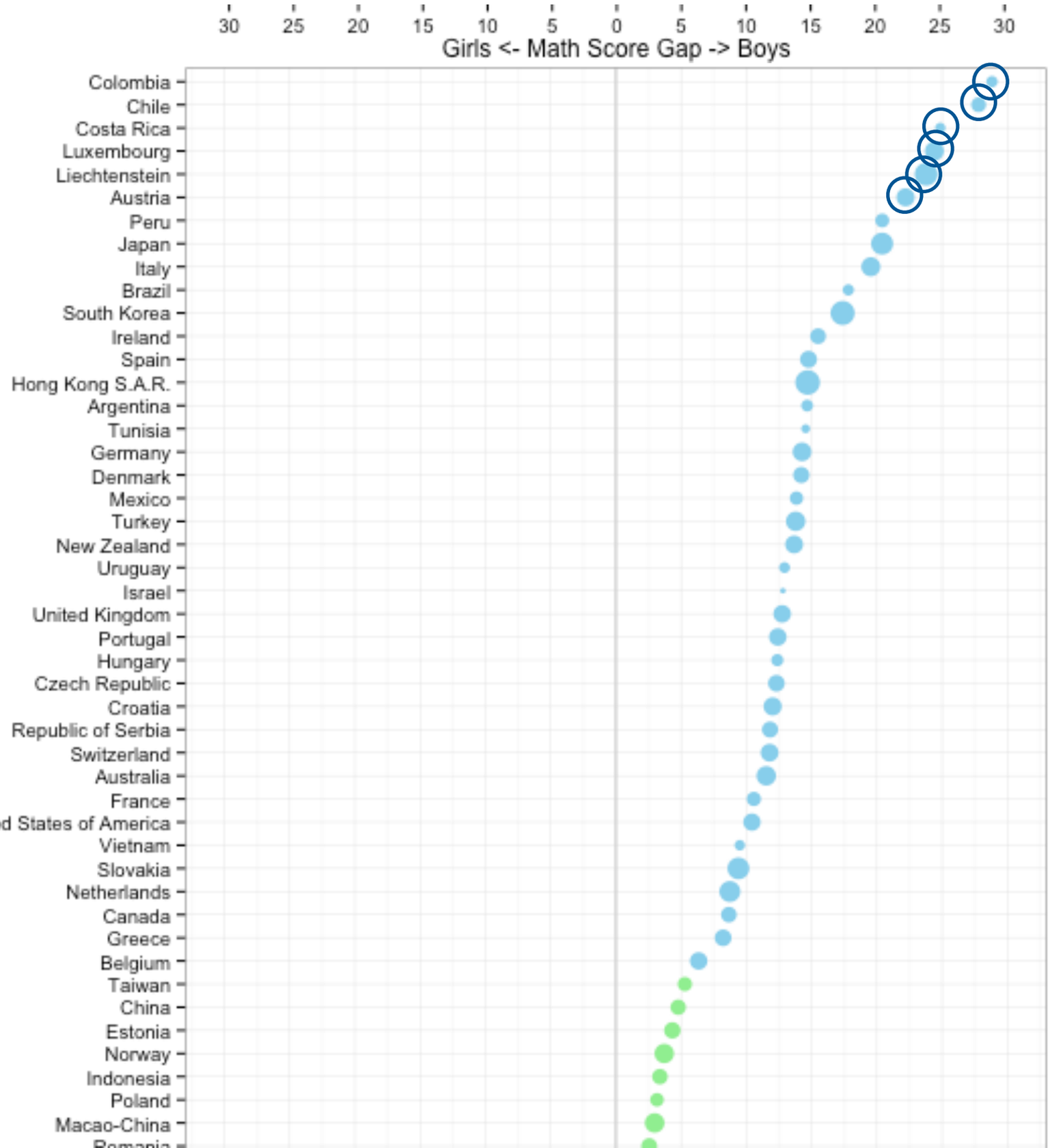
Gender & Math

1. Compute means by country and by gender.
2. Show mean difference by country





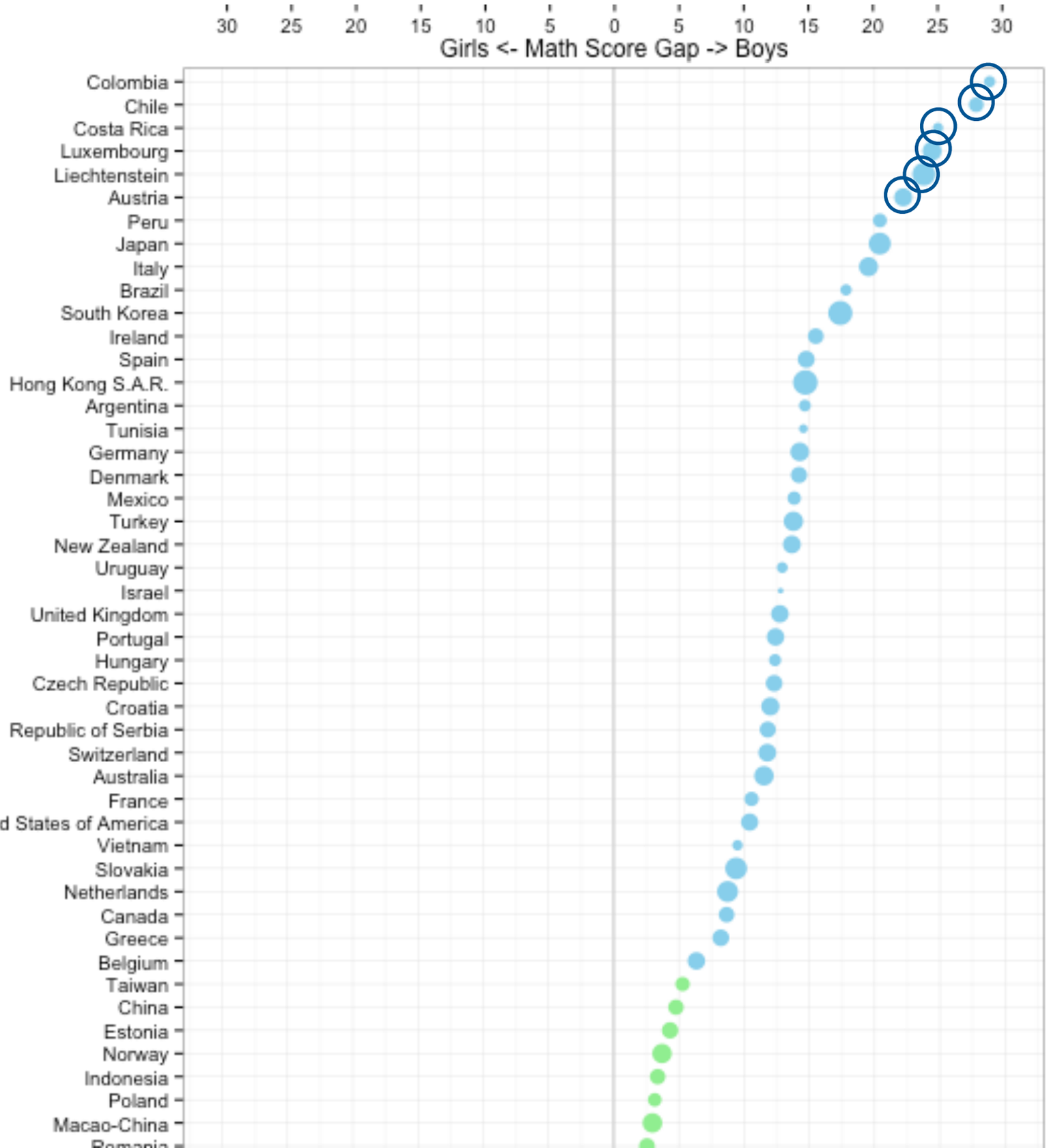
Gender & Math



Colombia
Chile
Costa Rica
Luxembourg
Liechtenstein
Austria

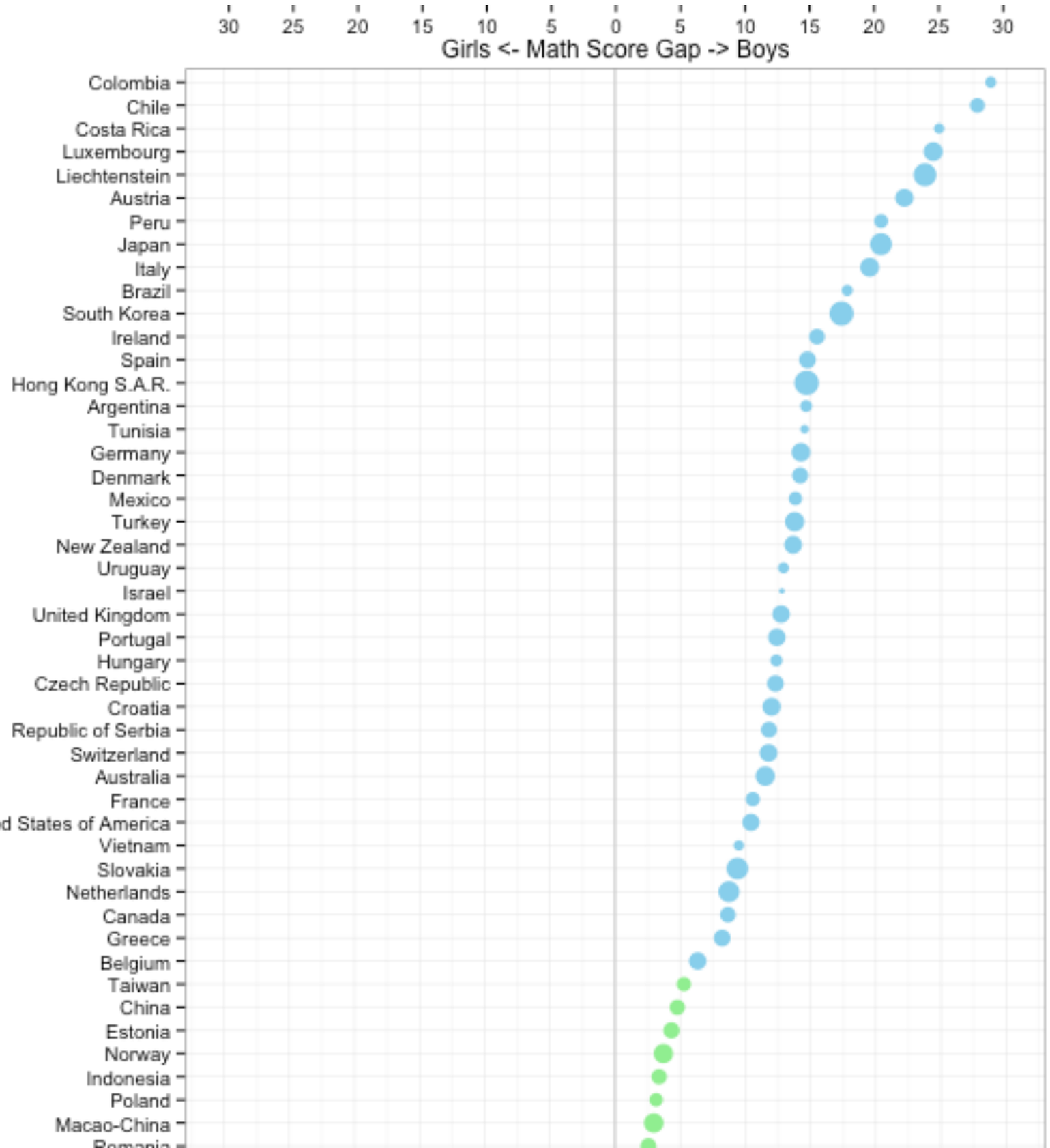
2014

Gender & Math



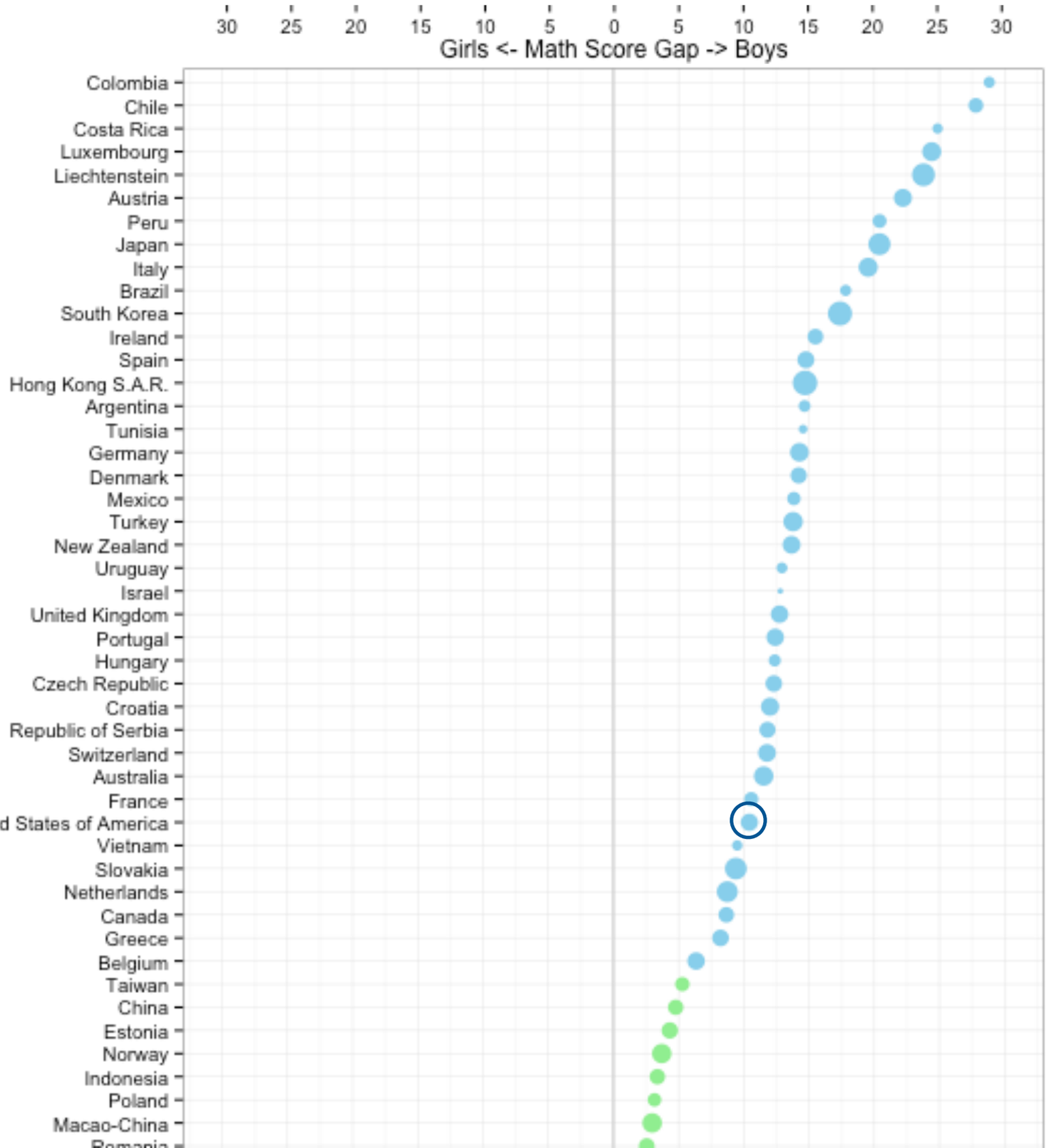
2014

Gender & Math



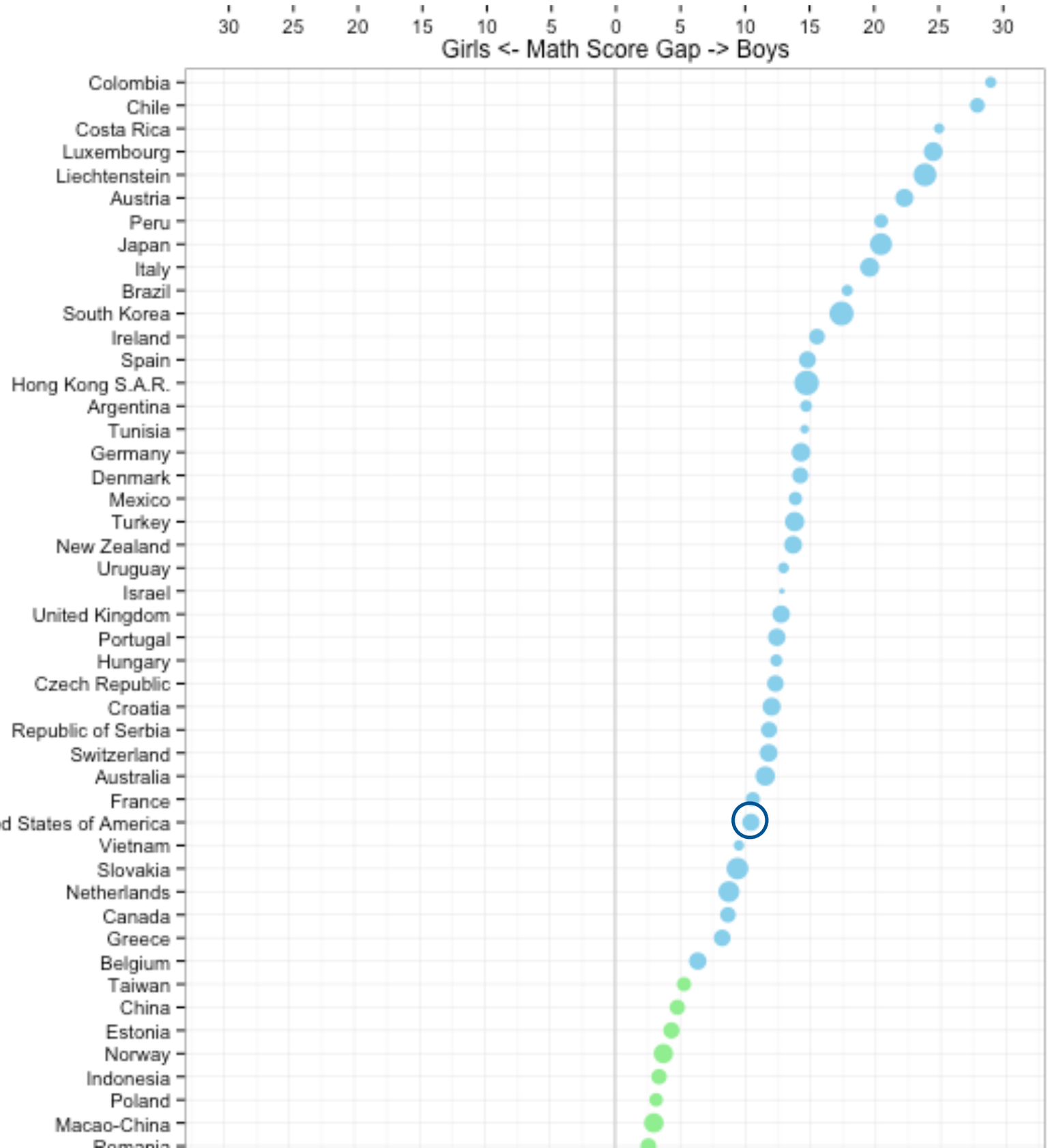
2014

Gender & Math



2014

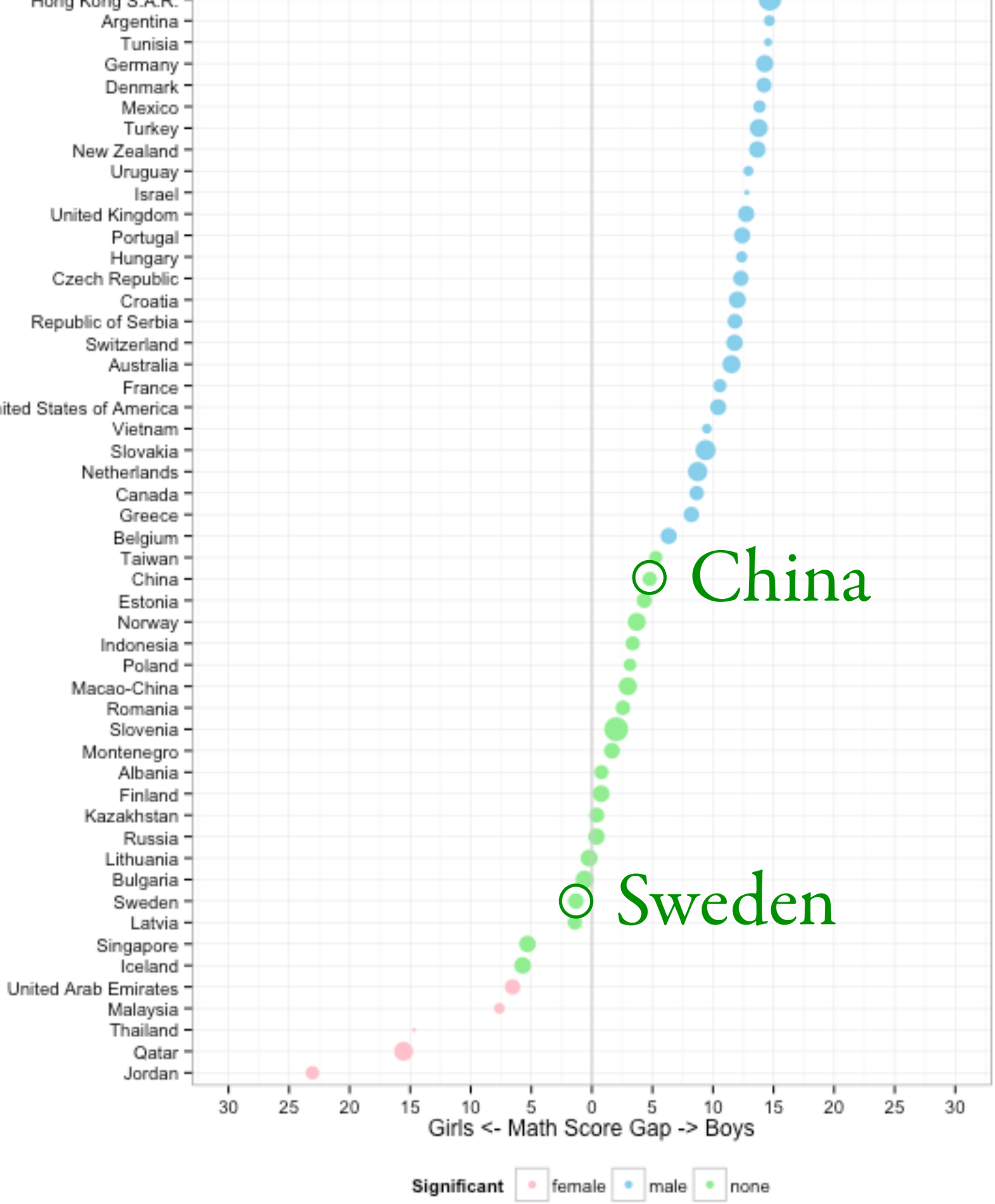
Gender & Math



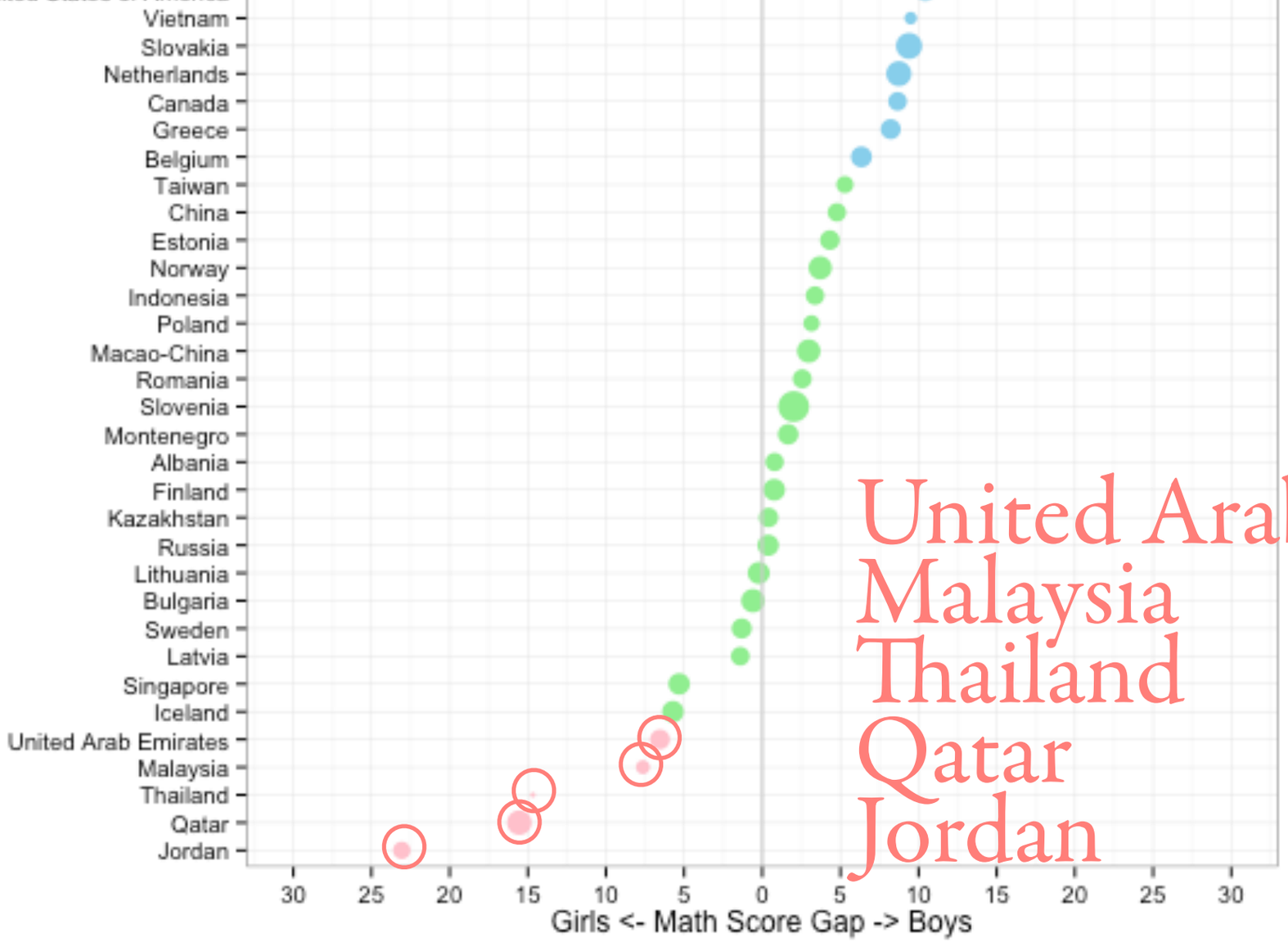
USA

2014

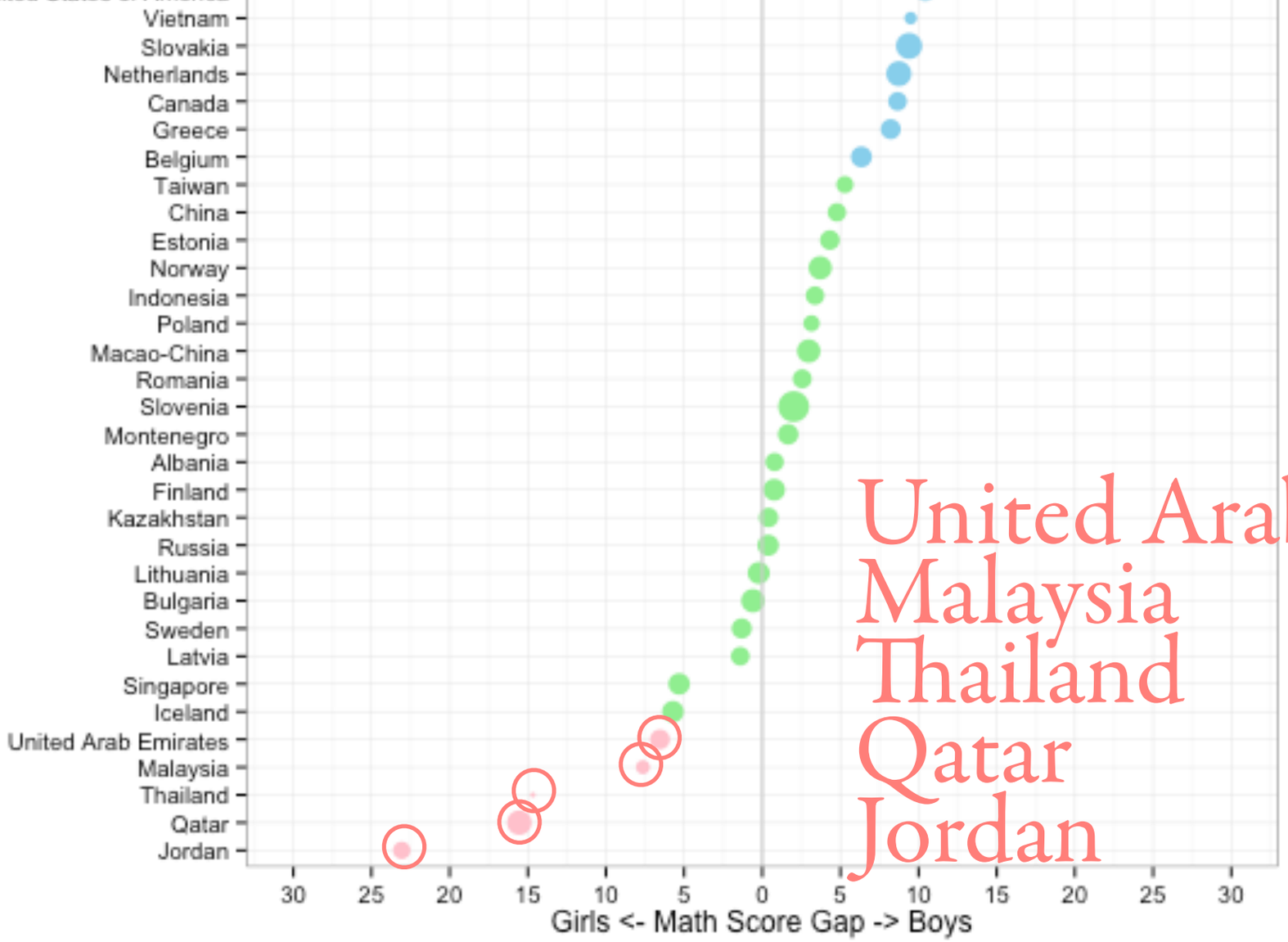
Gender & Math


Hong Kong S.A.R.
Argentina
Tunisia
Germany
Denmark
Mexico
Turkey
New Zealand
Uruguay
Israel
United Kingdom
Portugal
Hungary
Czech Republic
Croatia
Republic of Serbia
Switzerland
Australia
France
United States of America
Vietnam
Slovakia
Netherlands
Canada
Greece
Belgium
Taiwan
China
Estonia
Norway
Indonesia
Poland
Macao-China
Romania
Slovenia
Montenegro
Albania
Finland
Kazakhstan
Russia
Lithuania
Bulgaria
Sweden
Latvia
Singapore
Iceland
United Arab Emirates
Malaysia
Thailand
Qatar
Jordan
China
Sweden
30 25 20 15 10 5 0 5 10 15 20 25 30
Girls <- Math Score Gap -> Boys
Significant
female
male
none

Gender & Math

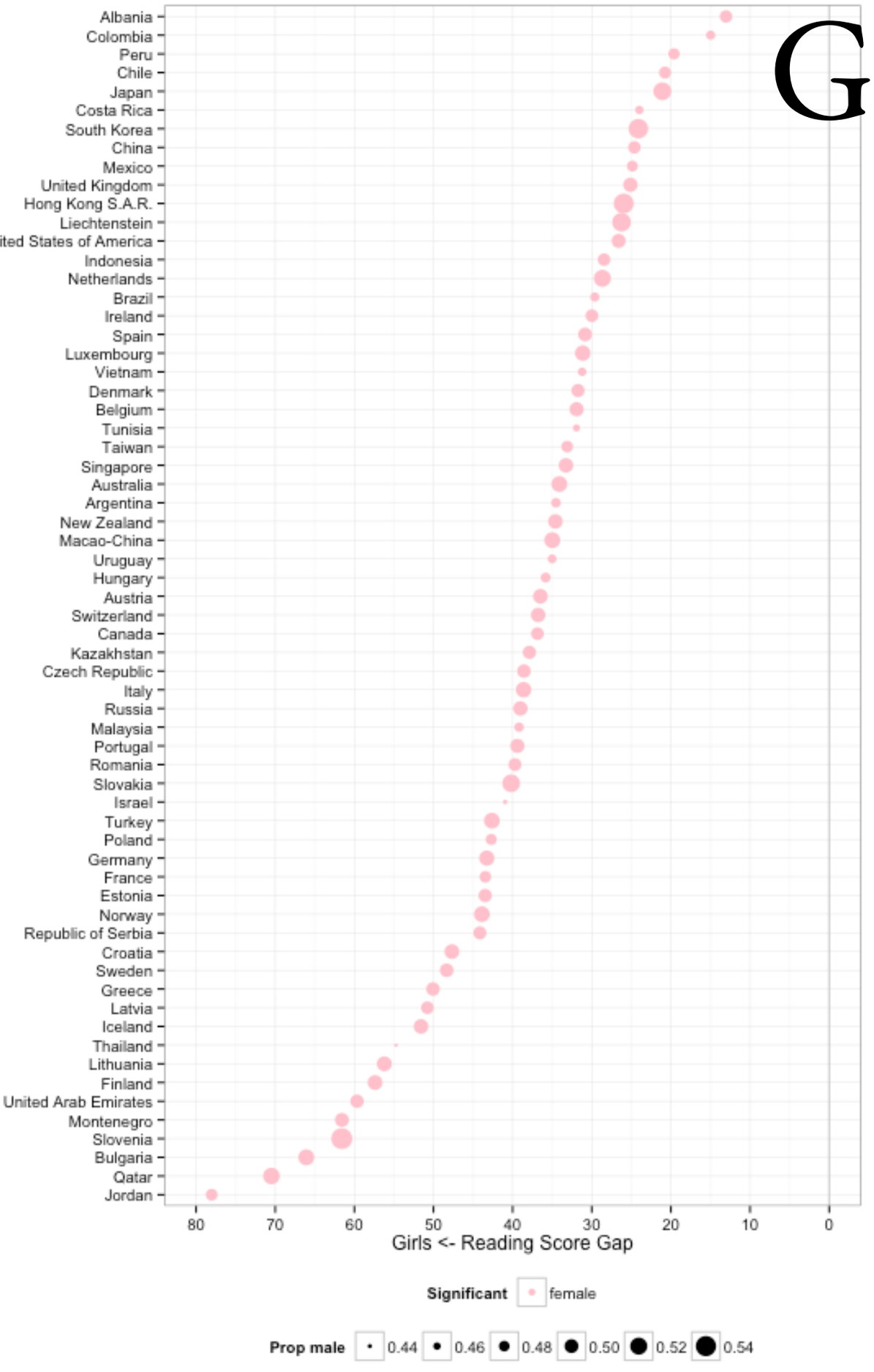

Vietnam
Slovakia
Netherlands
Canada
Greece
Belgium
Taiwan
China
Estonia
Norway
Indonesia
Poland
Macao-China
Romania
Slovenia
Montenegro
Albania
Finland
Kazakhstan
Russia
Lithuania
Bulgaria
Sweden
Latvia
Singapore
Iceland
United Arab Emirates
Malaysia
Thailand
Qatar
Jordan
United Arab Emirates
Malaysia
Thailand
Qatar
Jordan
30
25
20
15
10
5
0
5
10
15
20
25
30
Girls <- Math Score Gap -> Boys

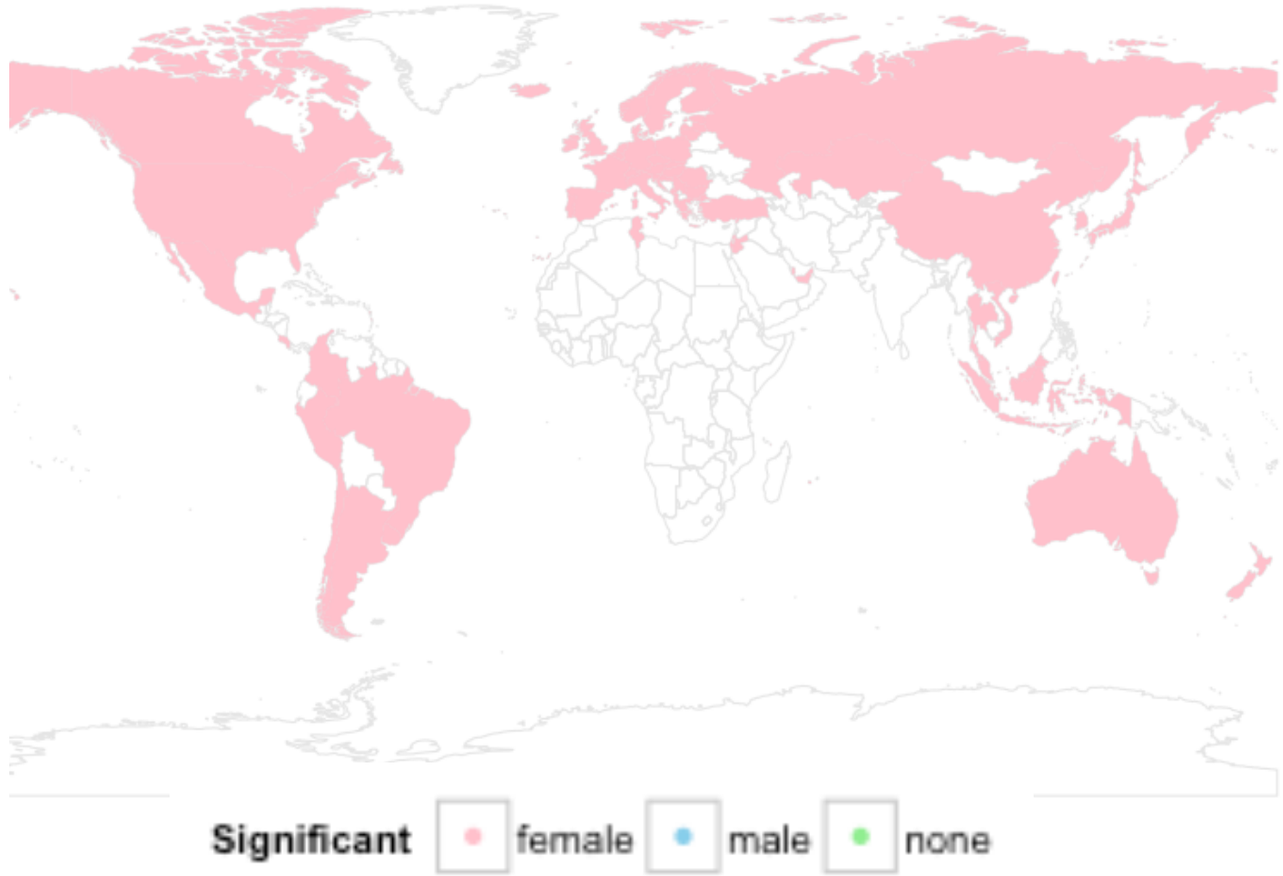
Gender & Math



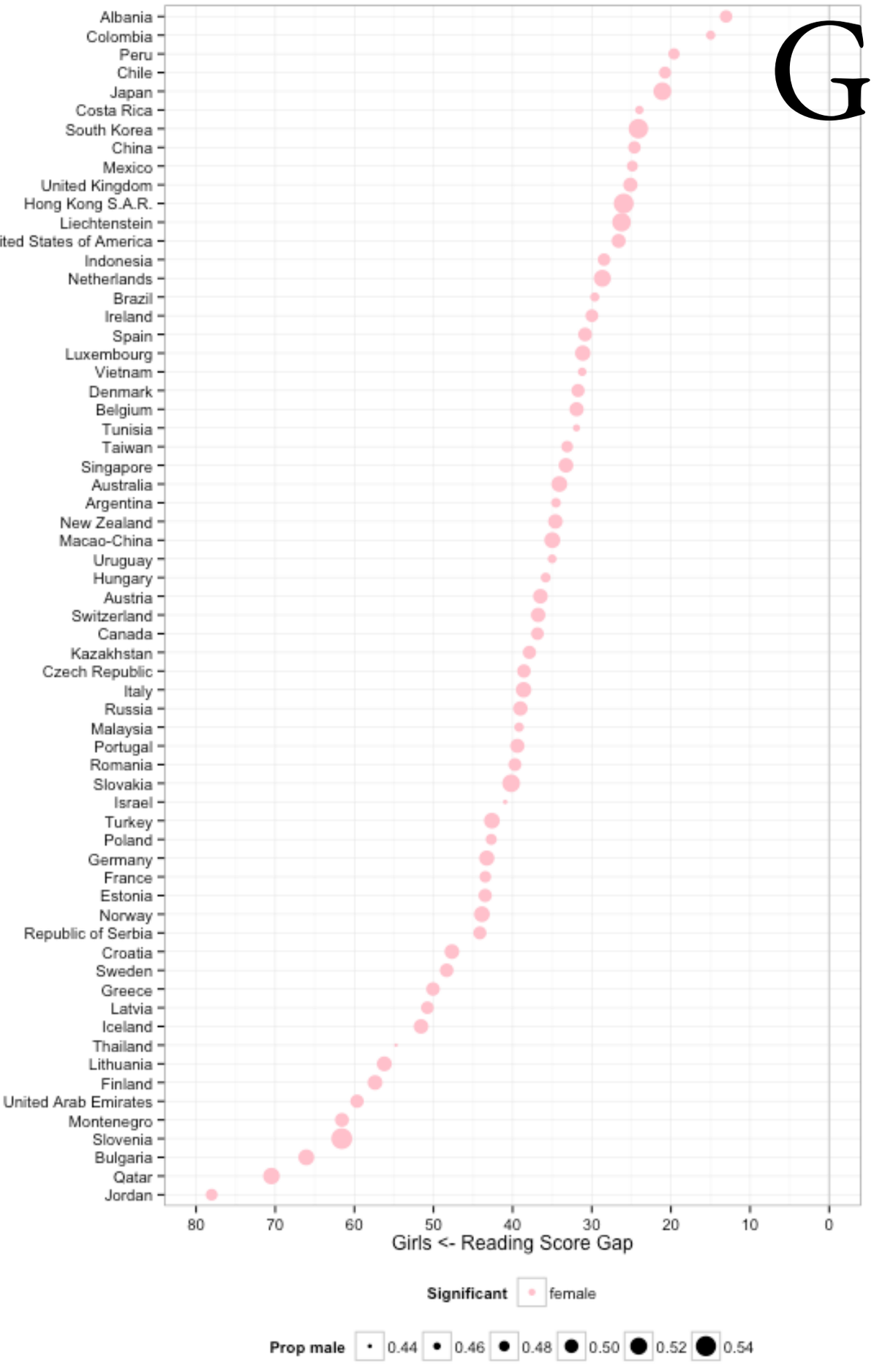
Surprised?

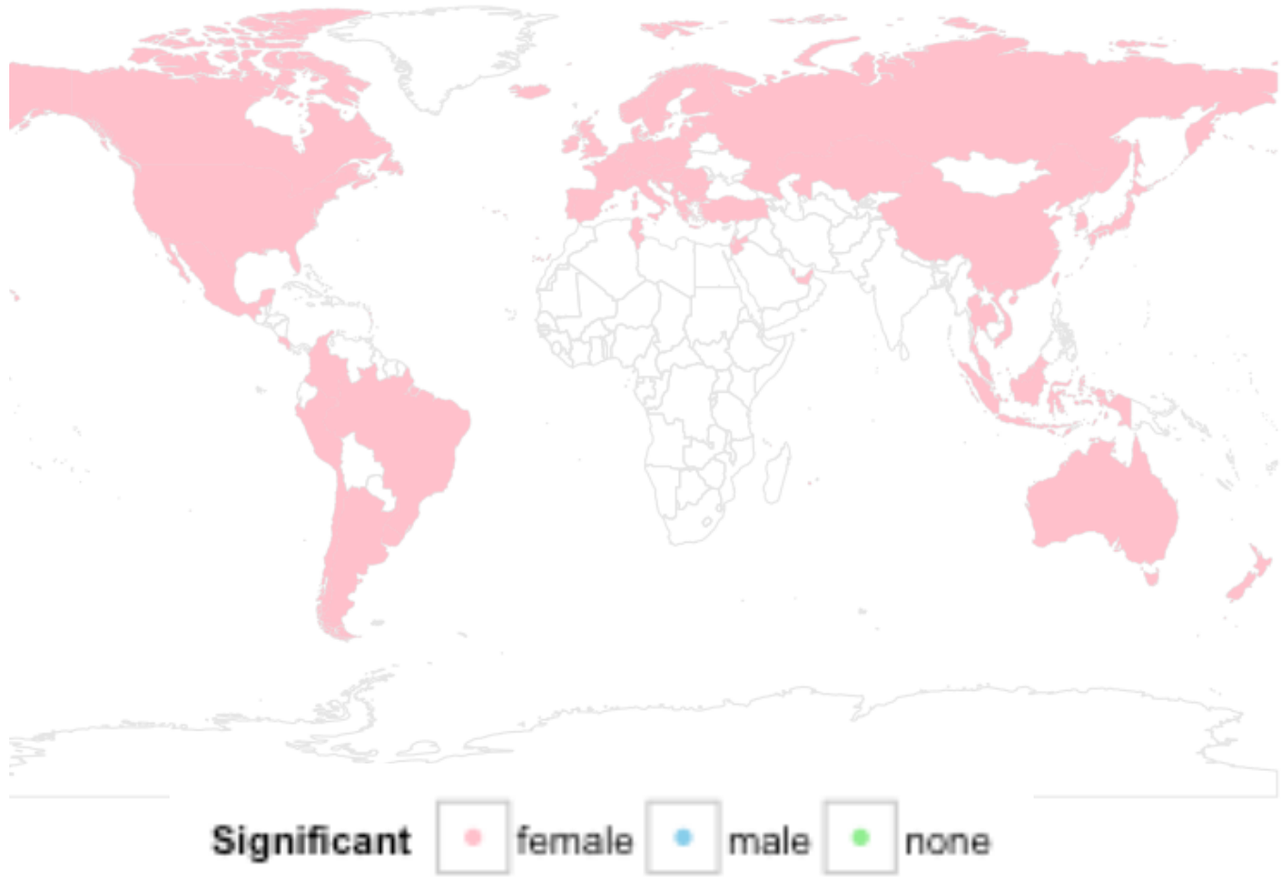
Gender & Reading





Gender & Reading





Surprised again?

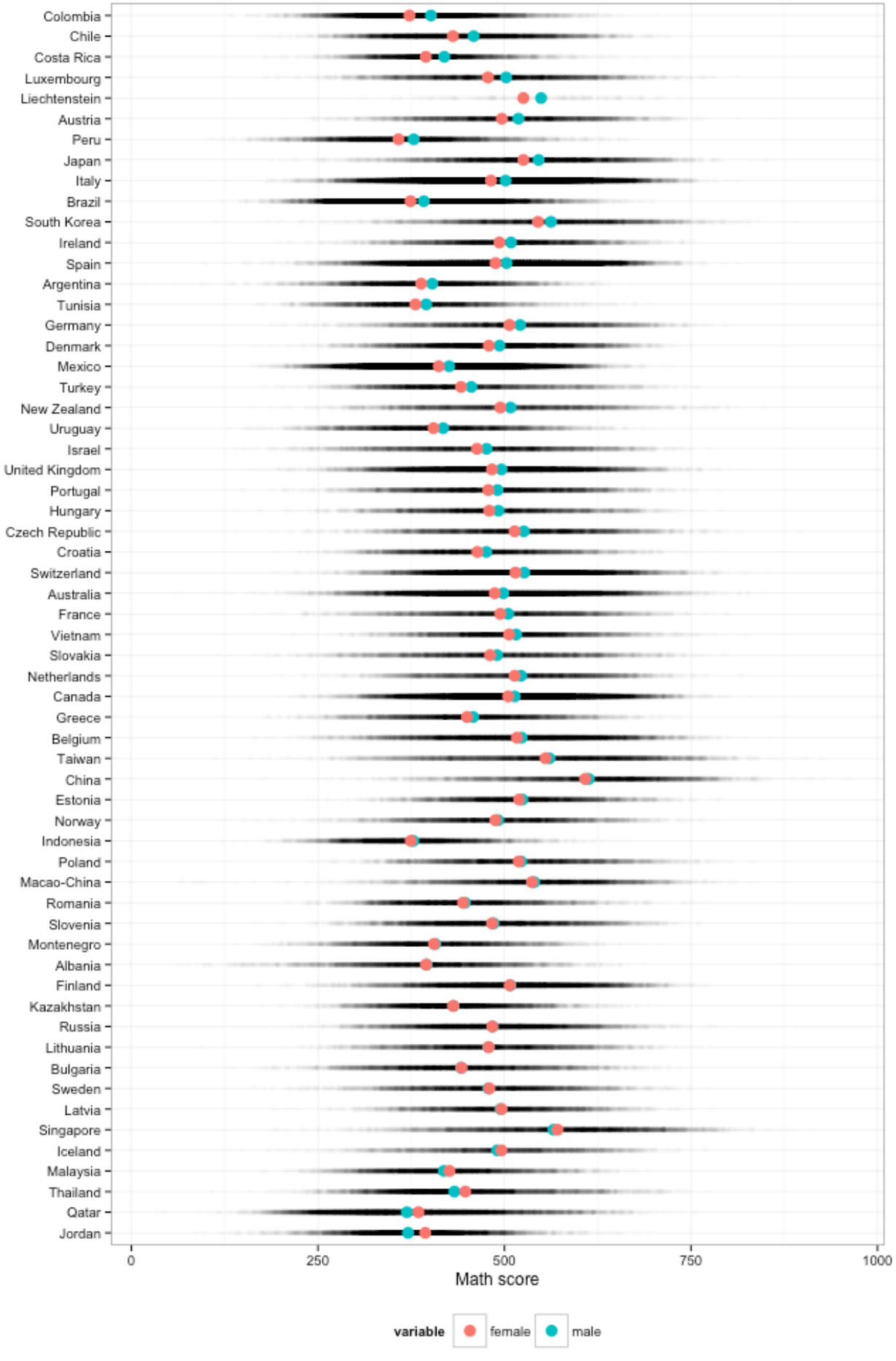
In reading scores, the world is PINK!
In EVERY COUNTRY tested GIRLS score significantly BETTER THAN BOYS. Why don't we talk about the gender gap in reading?

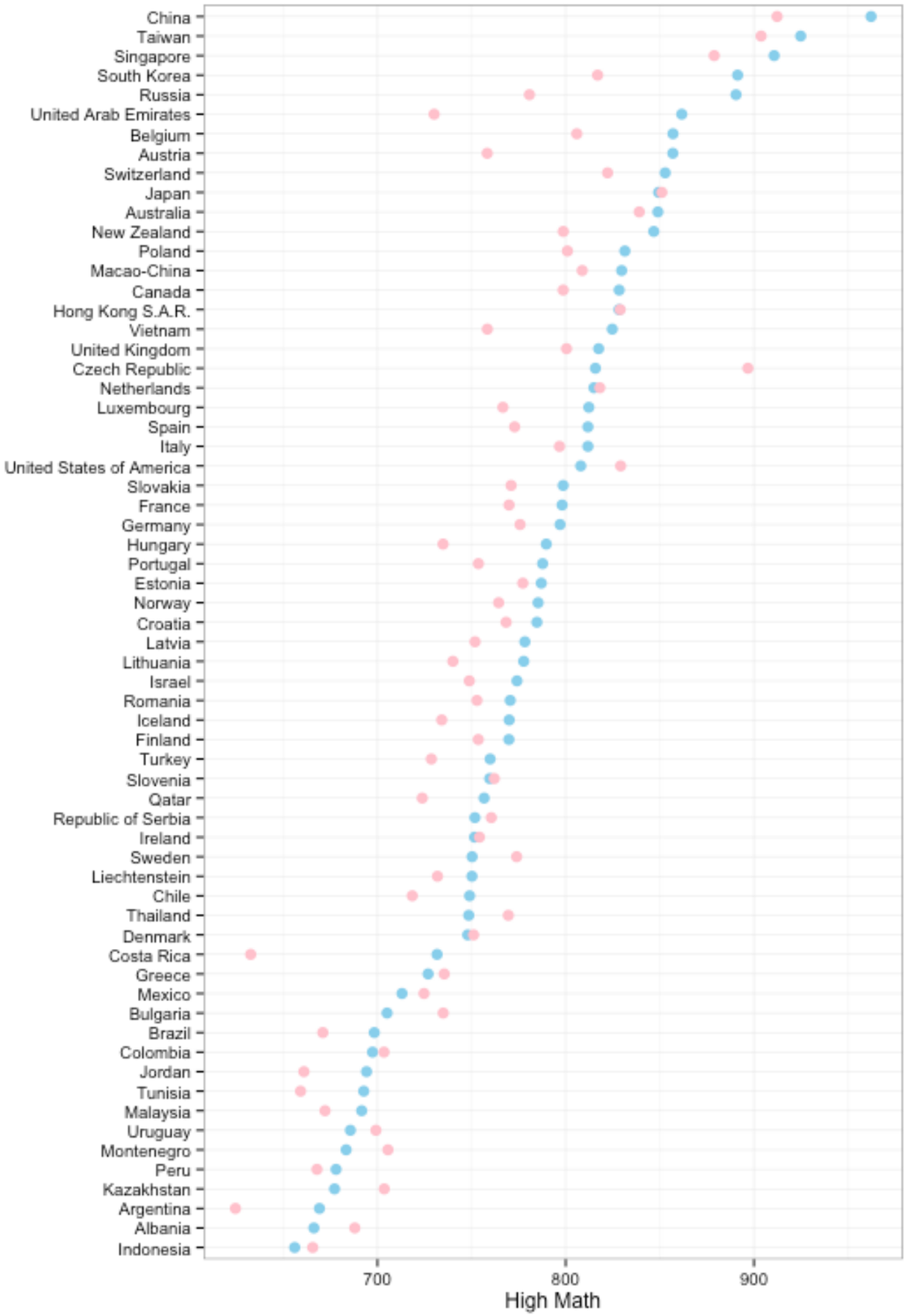
Individuals

- Test scores range from 0-1000.
- Gaps for math were at most 30 points.
- For reading at most 80 points.
- Differences are tiny.

Individuals

- Test scores range from 0-1000.
- Gaps for math were at most 30 points.
- For reading at most 80 points.
- Differences are tiny.





Individuals

- Even countries with a math gender gap, sometimes have a girl who scored the top mark.

Individuals

- Even countries with a math gender gap, sometimes have a girl who scored the top mark.

Individuals

- Even countries with a math gender gap, sometimes have a girl who scored the top mark.

Czech Republic

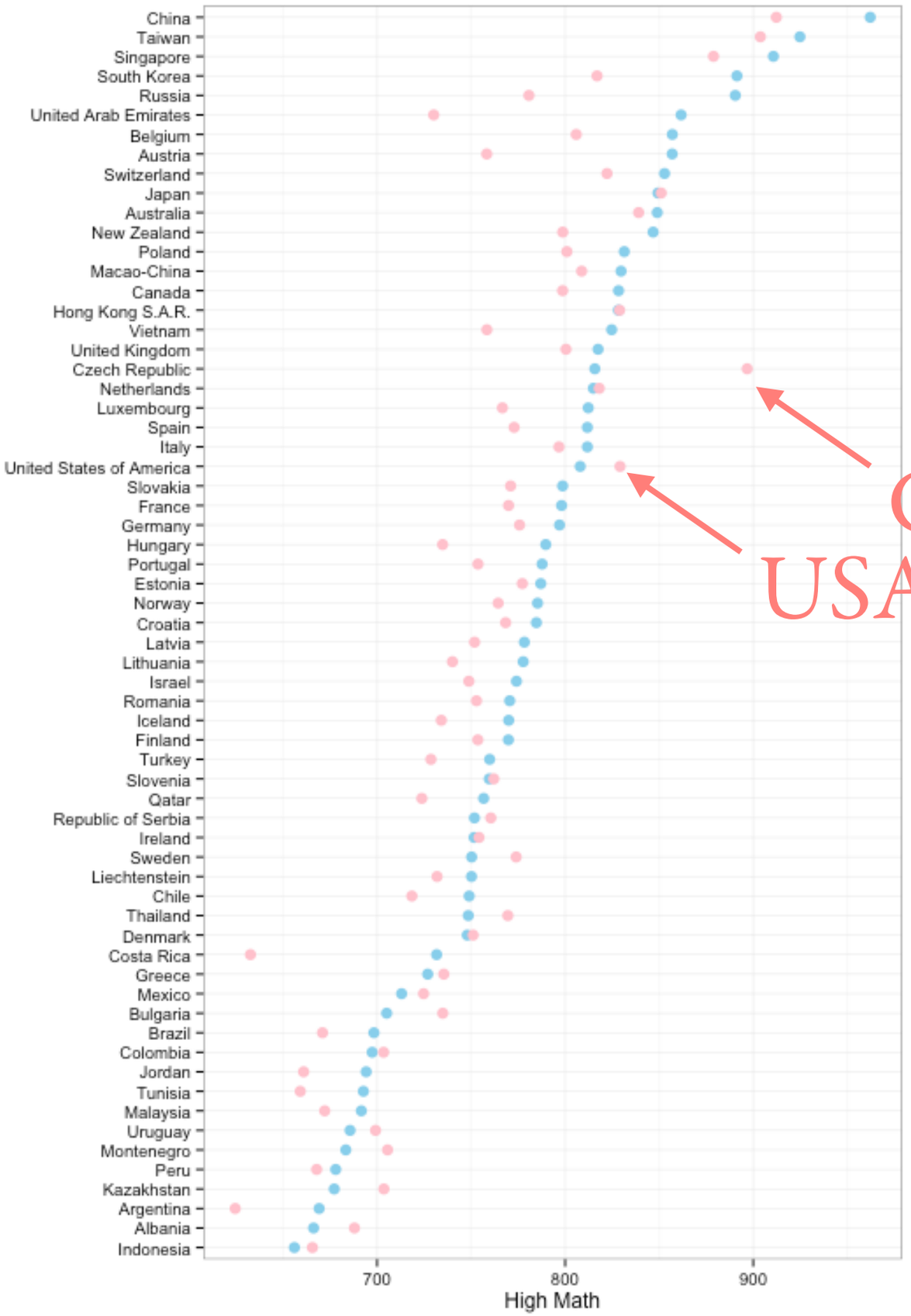
Individuals

- Even countries with a math gender gap, sometimes have a girl who scored the top mark.

Czech Republic

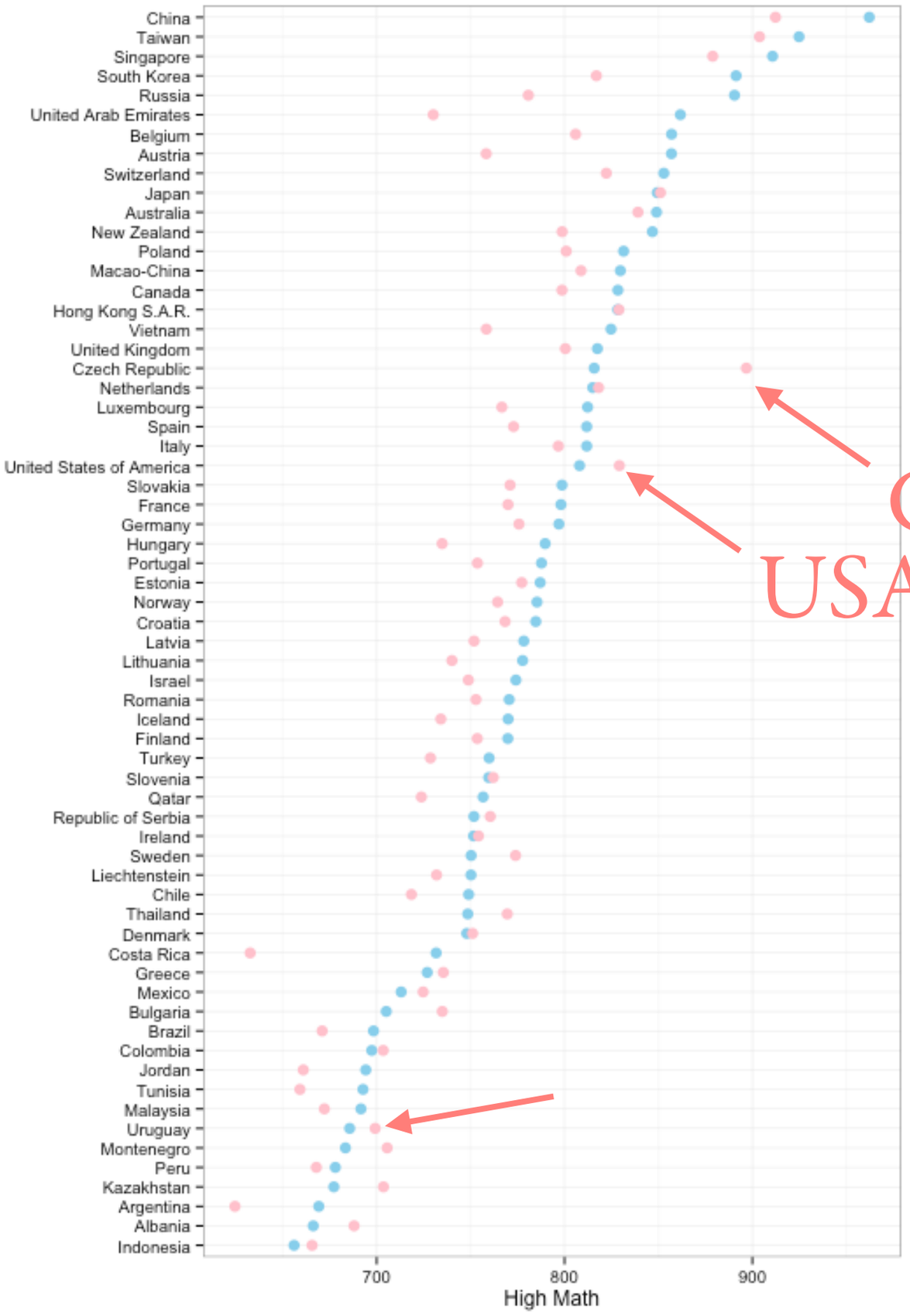
Individuals

- Even countries with a math gender gap, sometimes have a girl who scored the top mark.



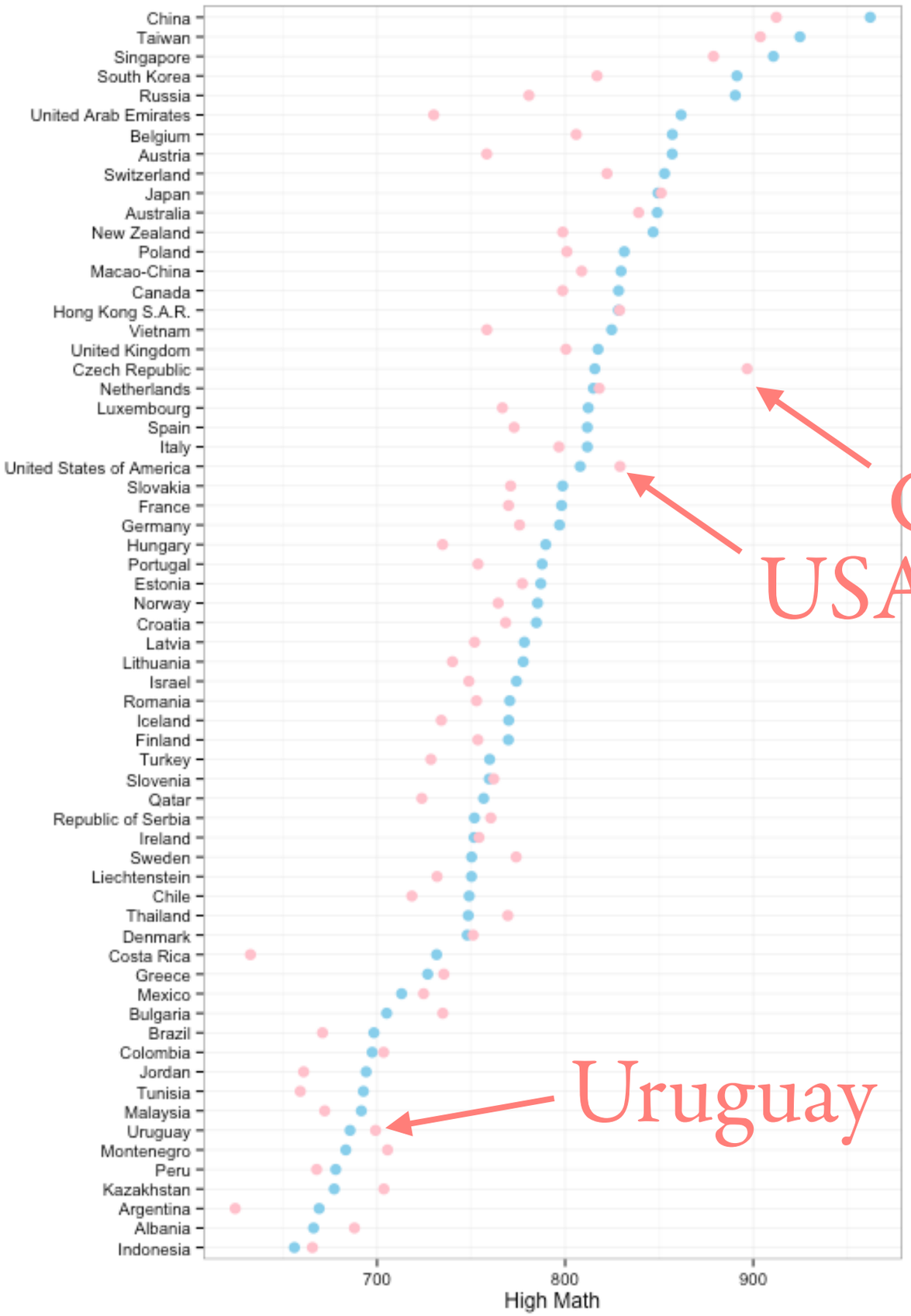
Individuals

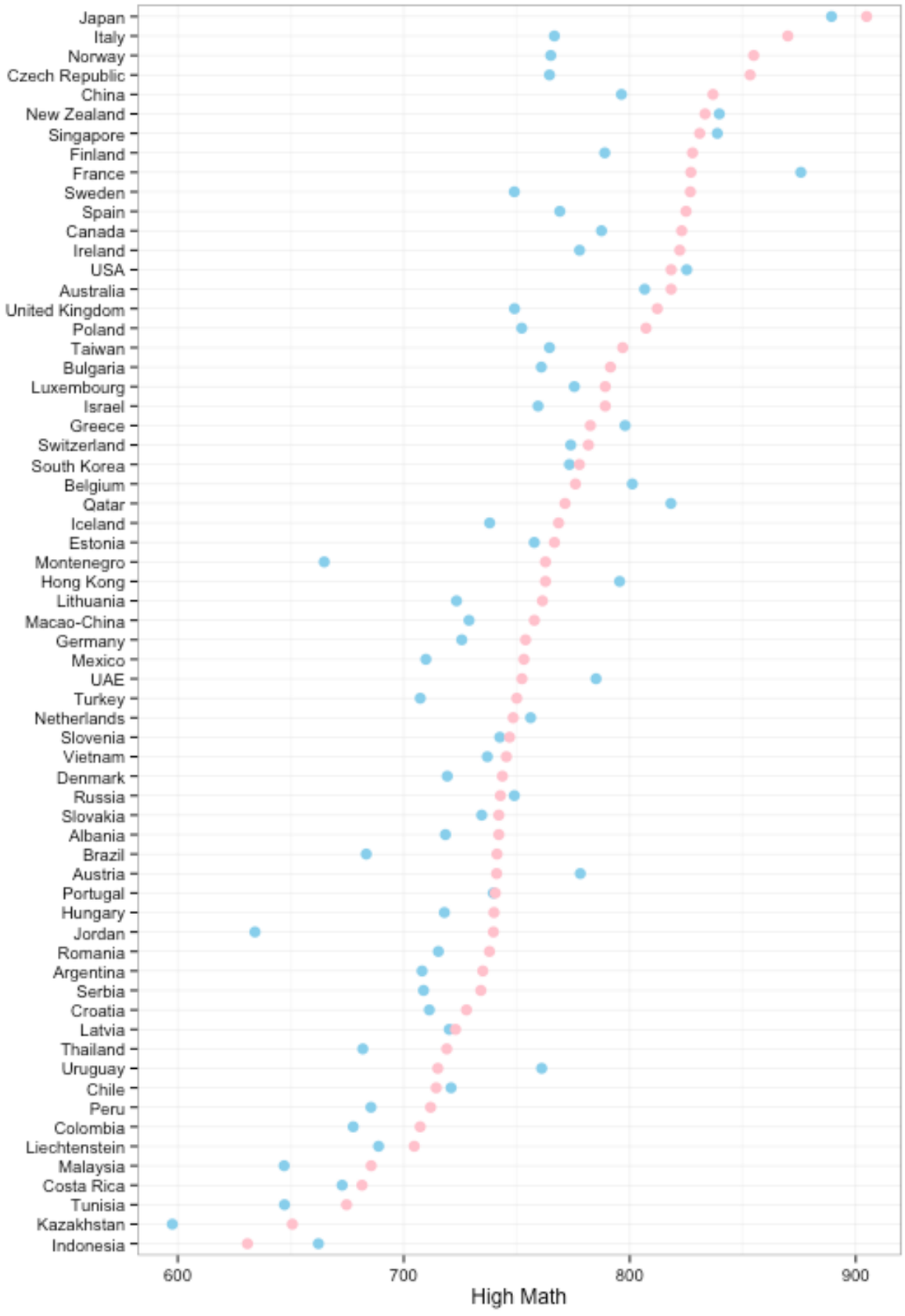
- Even countries with a math gender gap, sometimes have a girl who scored the top mark.



Individuals

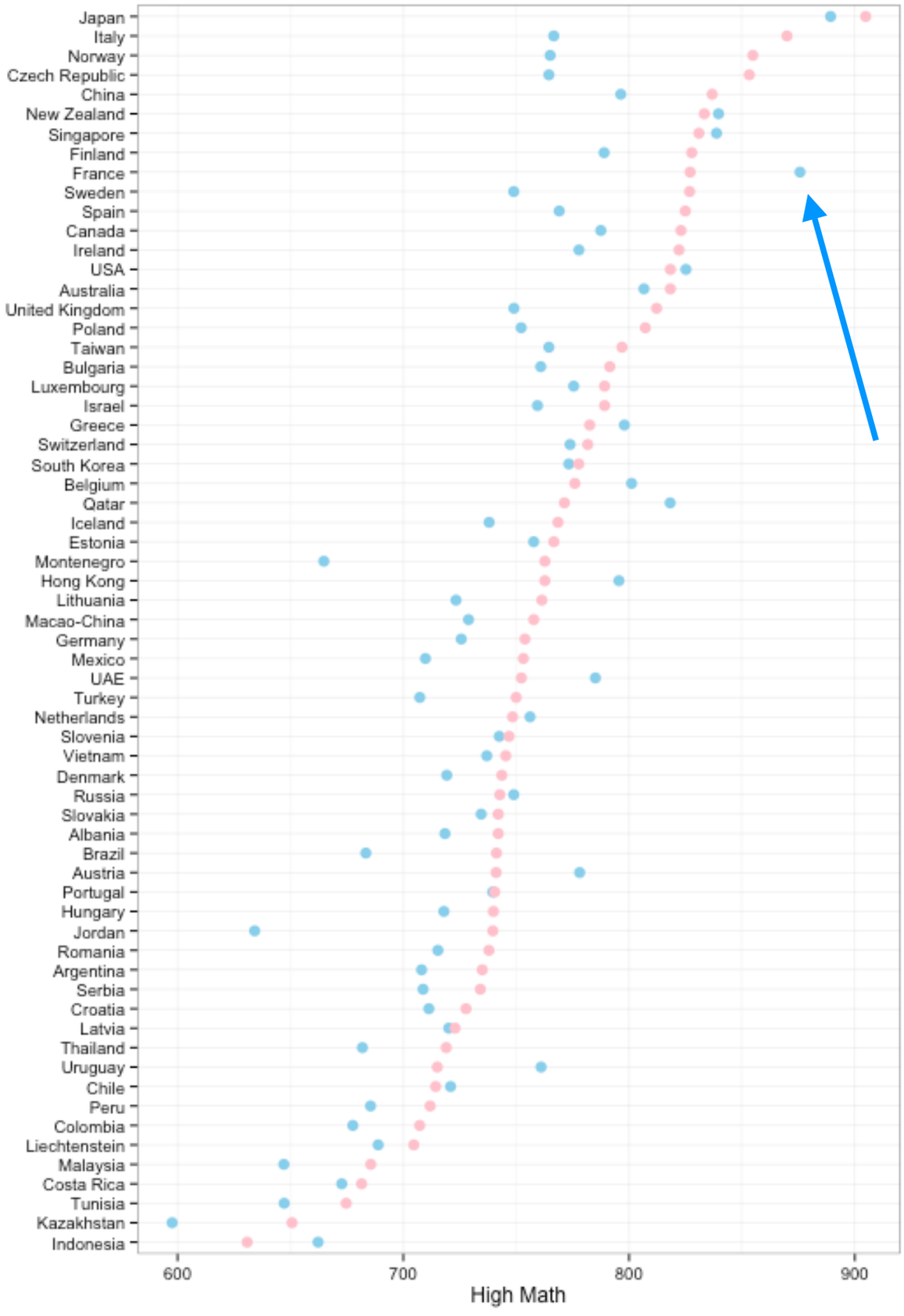
- Even countries with a math gender gap, sometimes have a girl who scored the top mark.





Individuals

- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

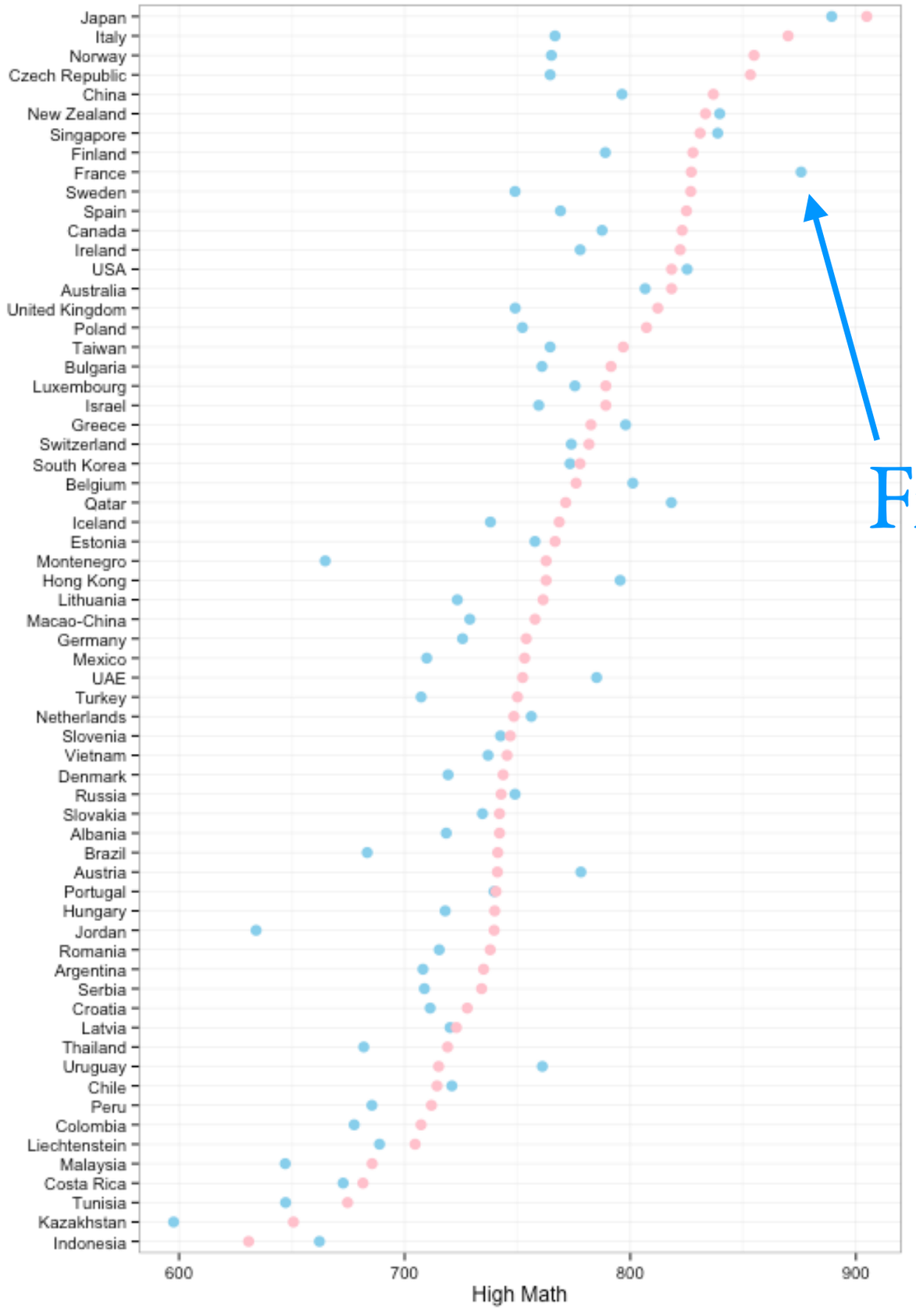


Individuals

- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

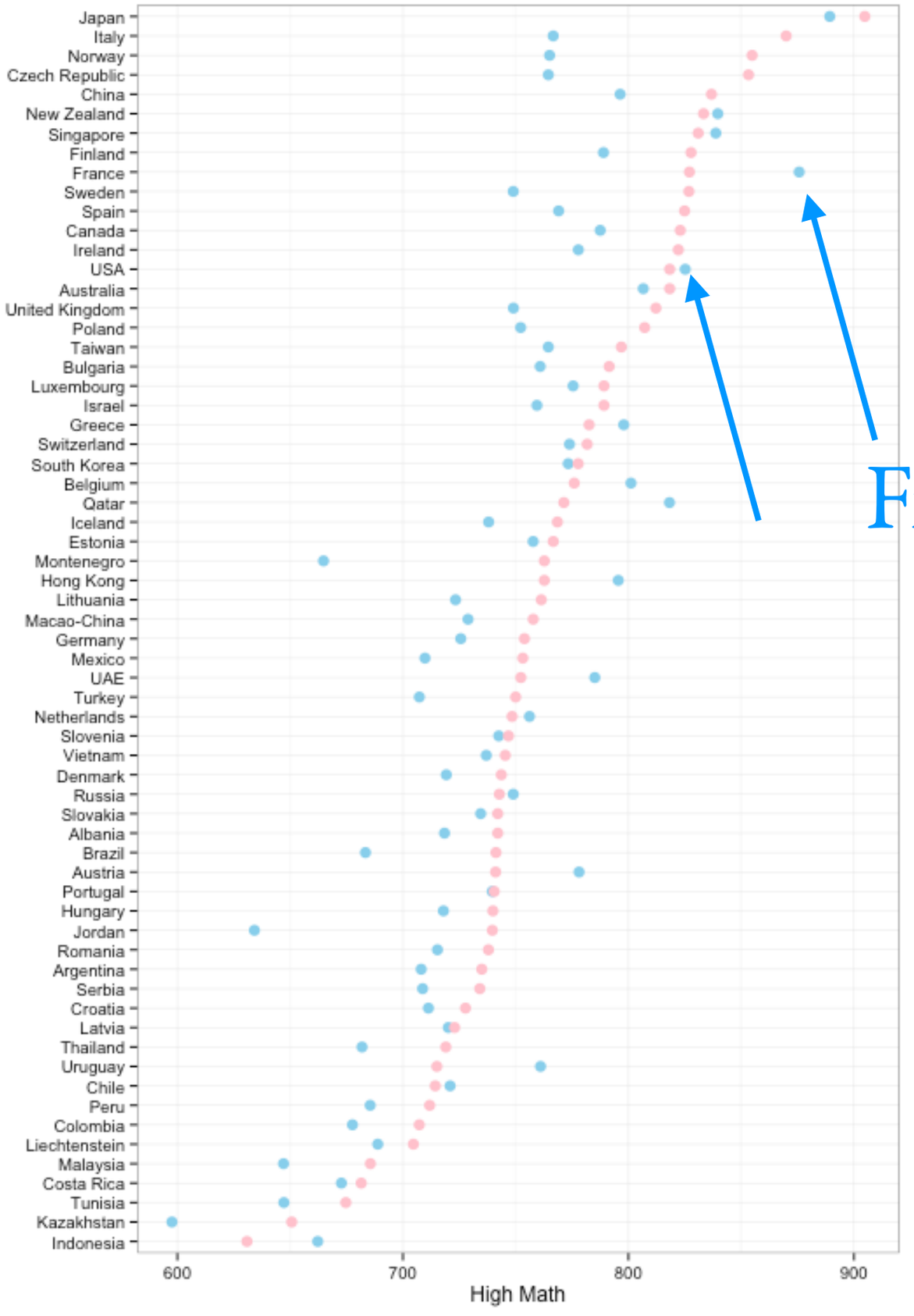
Individuals

- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

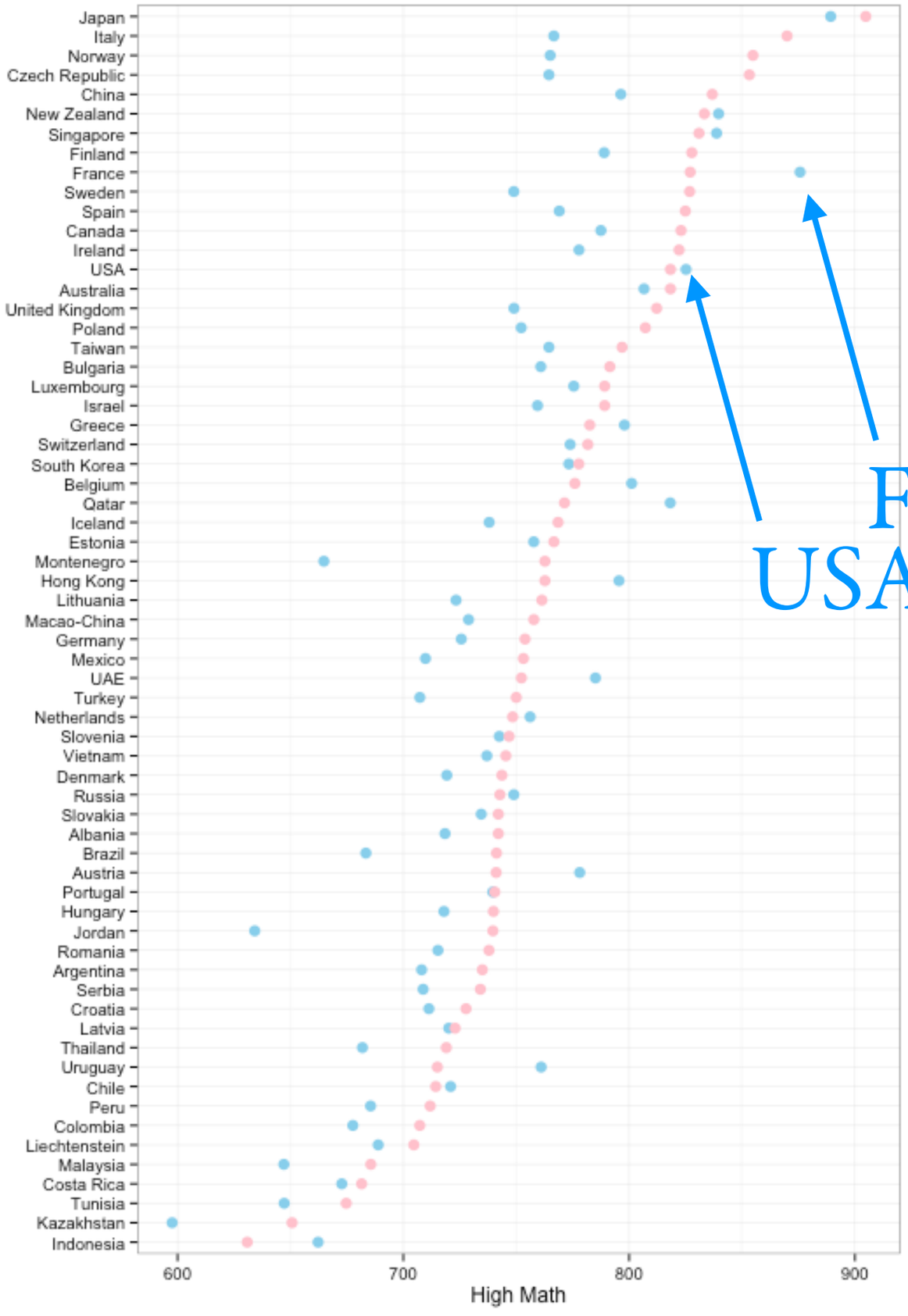


Individuals

- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.



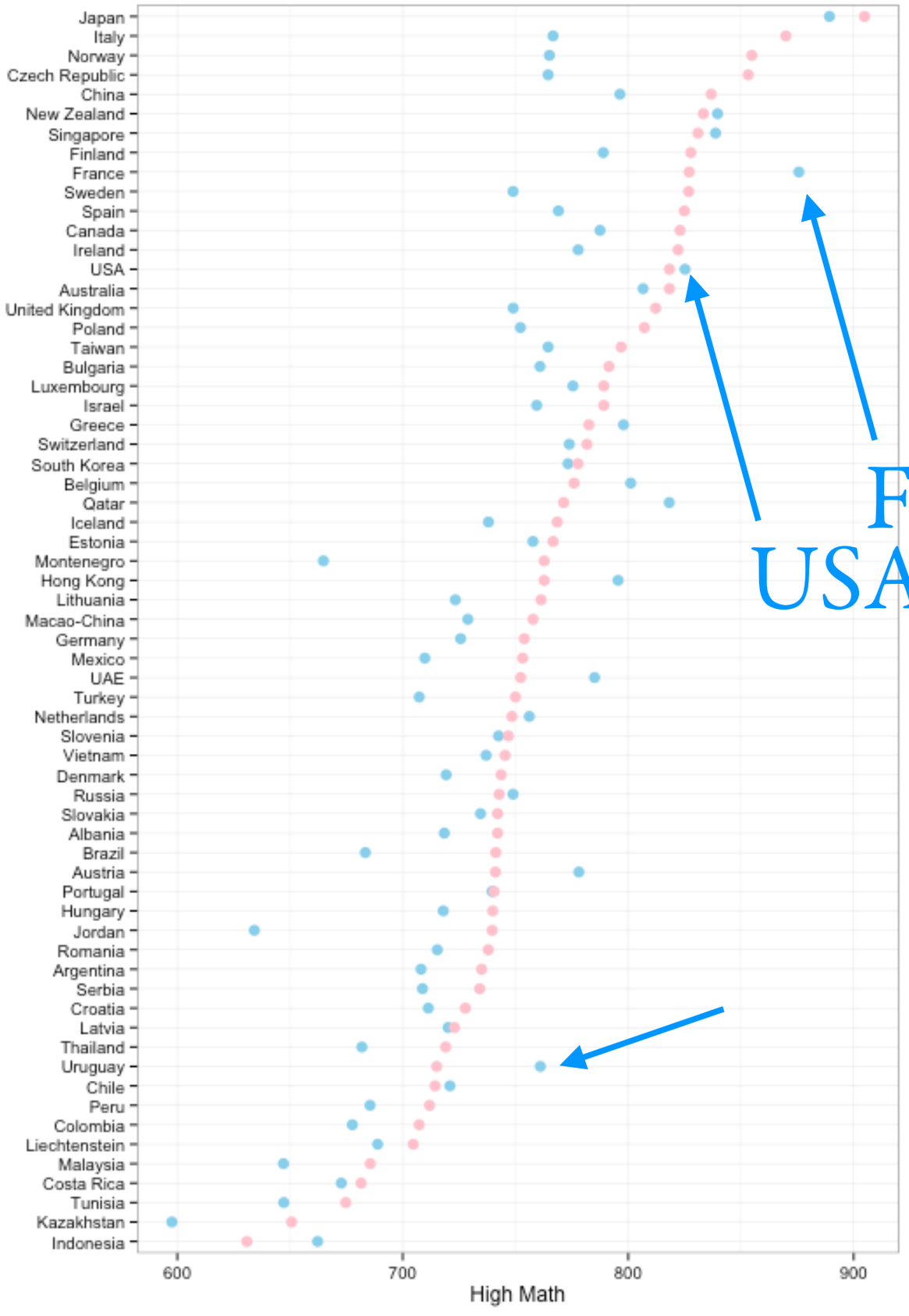
Individuals



- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

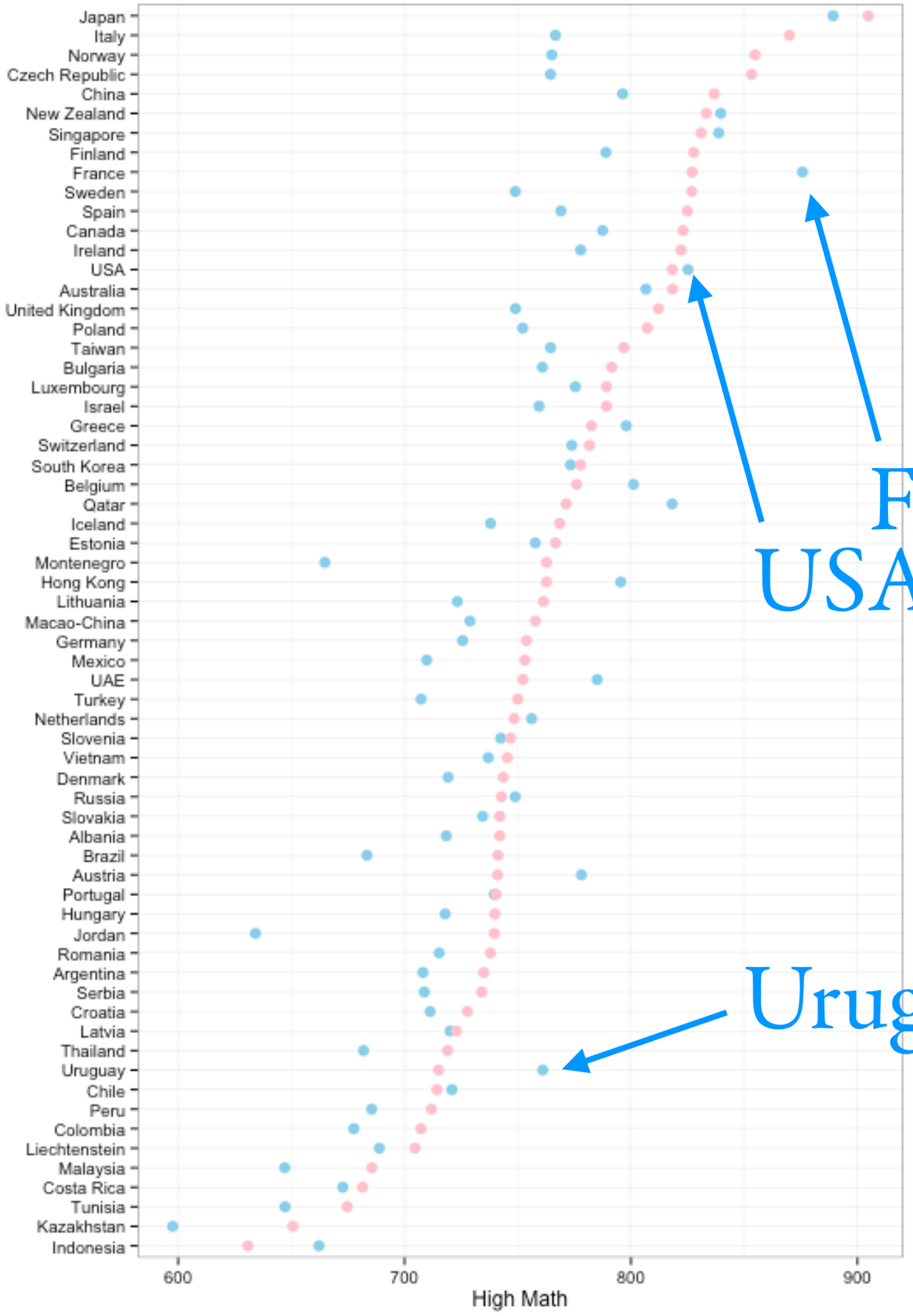
Individuals

- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

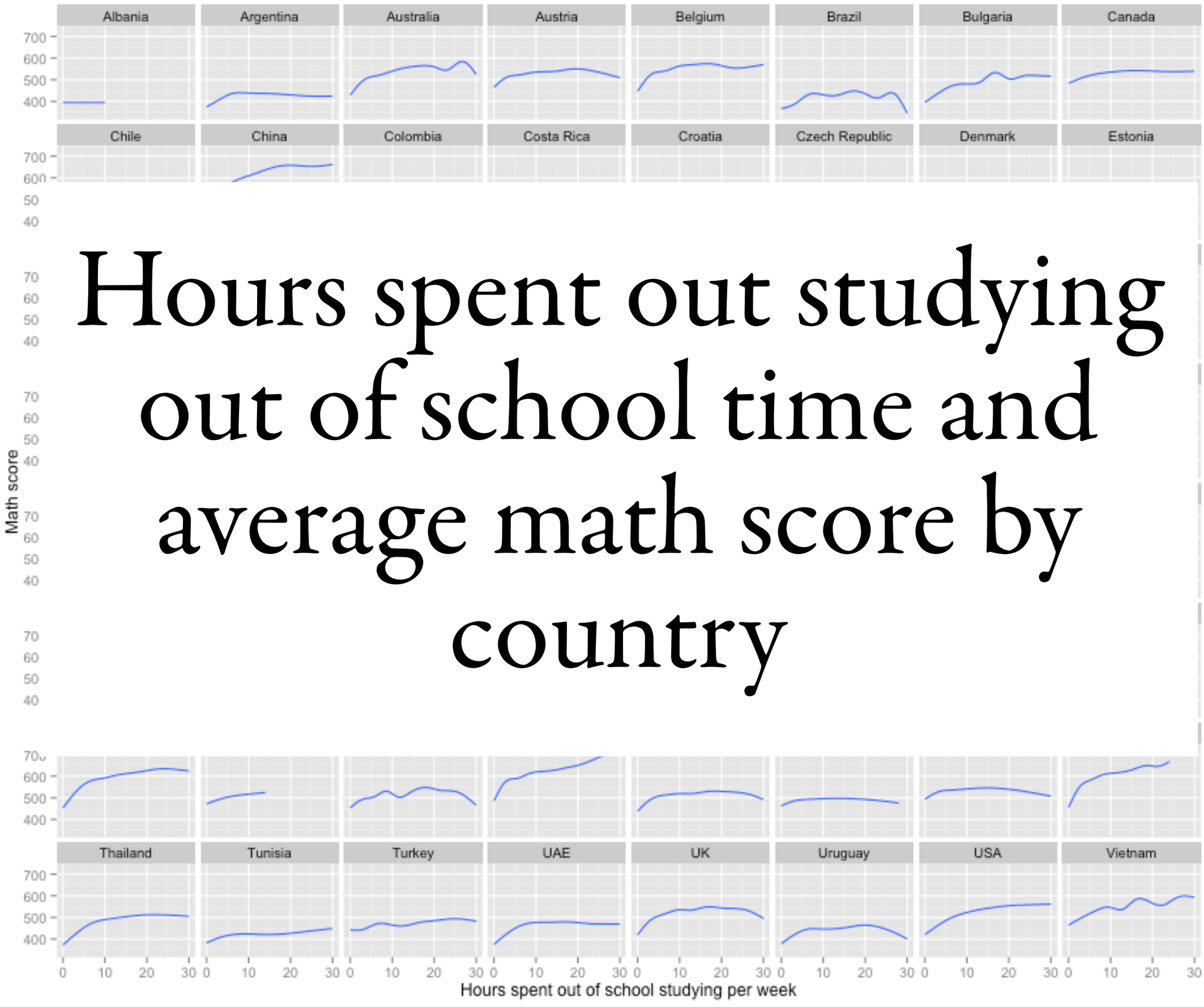


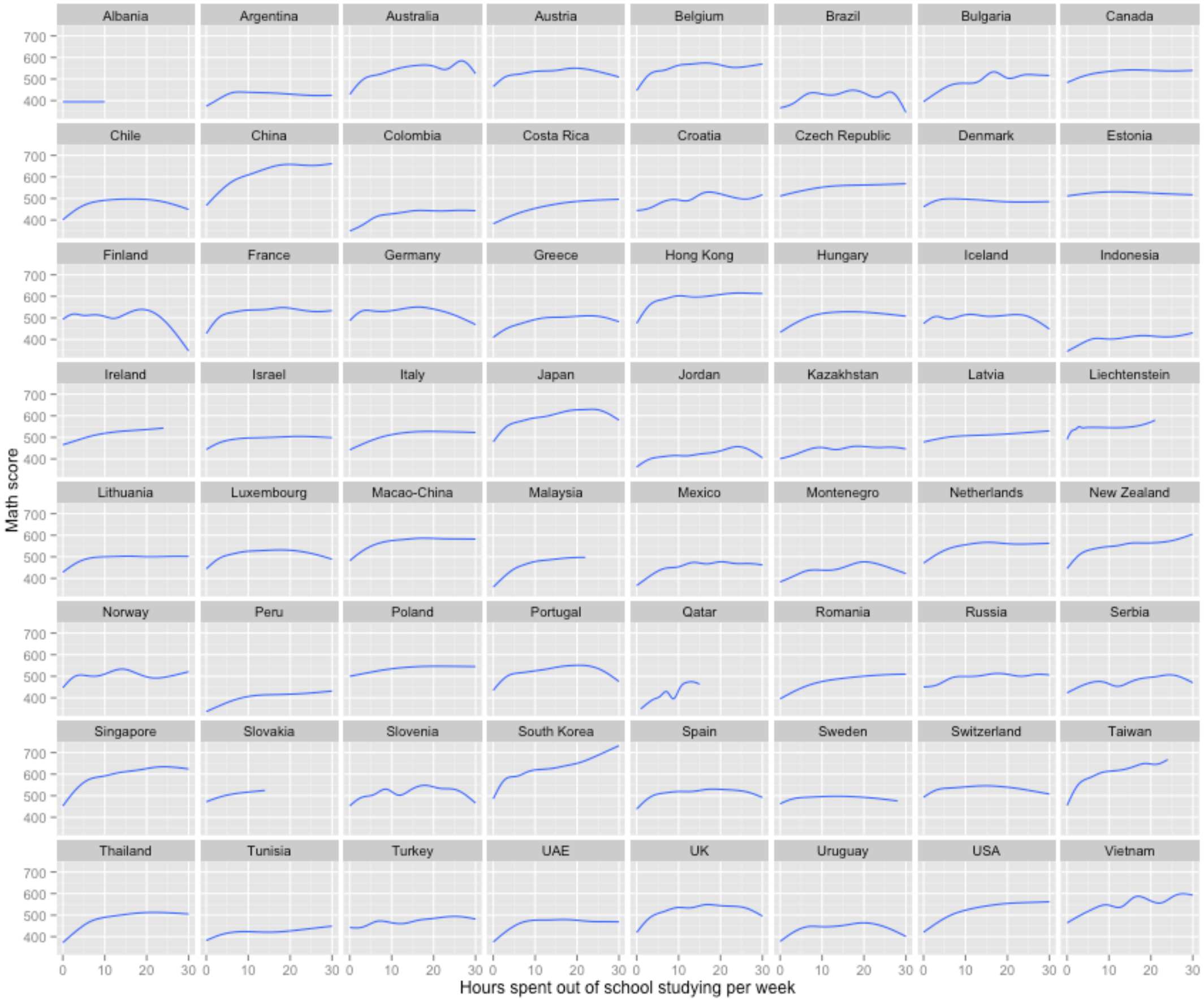
Individuals

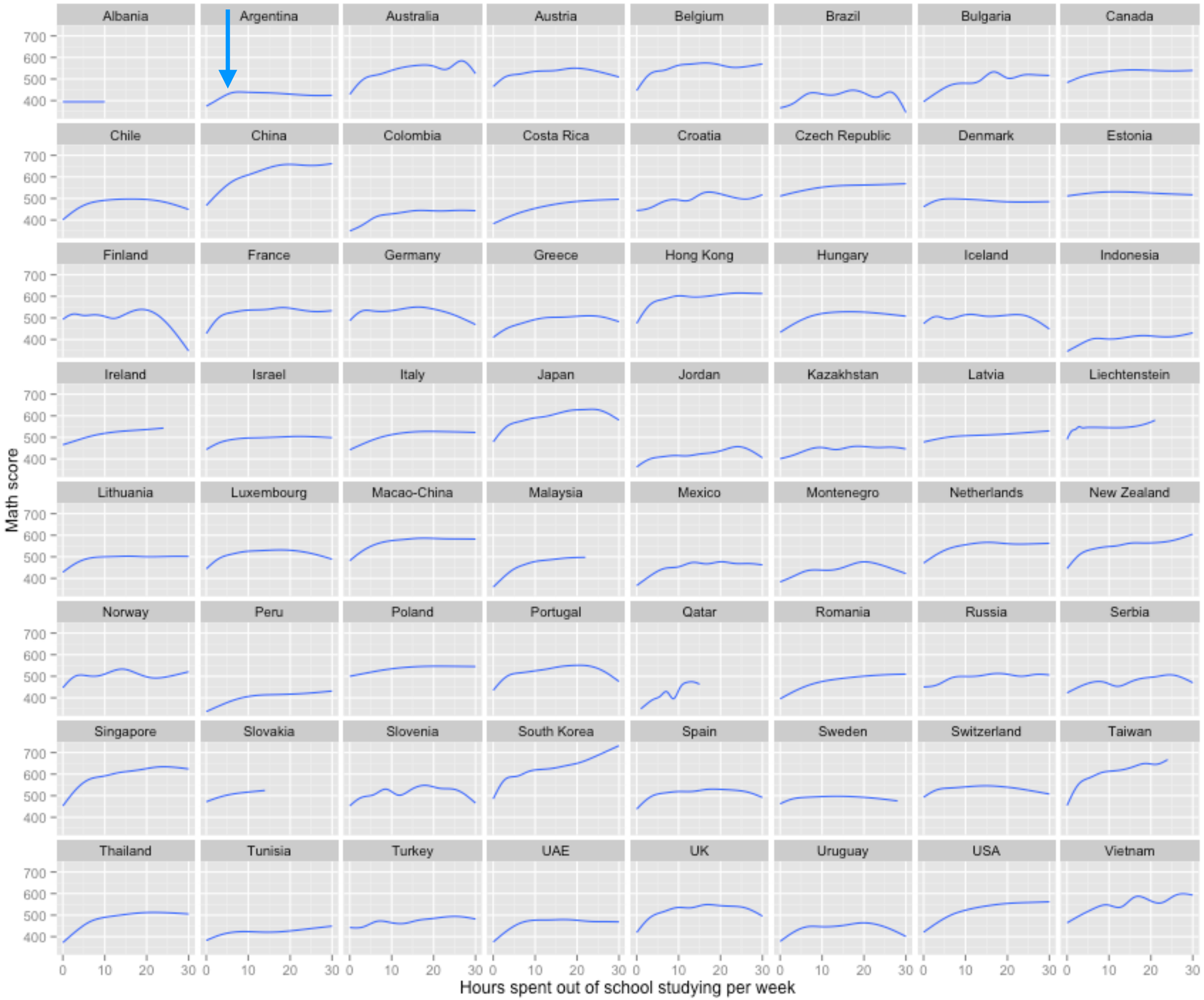
- And although there is a gender gap in reading in all countries, individually some top reading scores are obtained by boys.

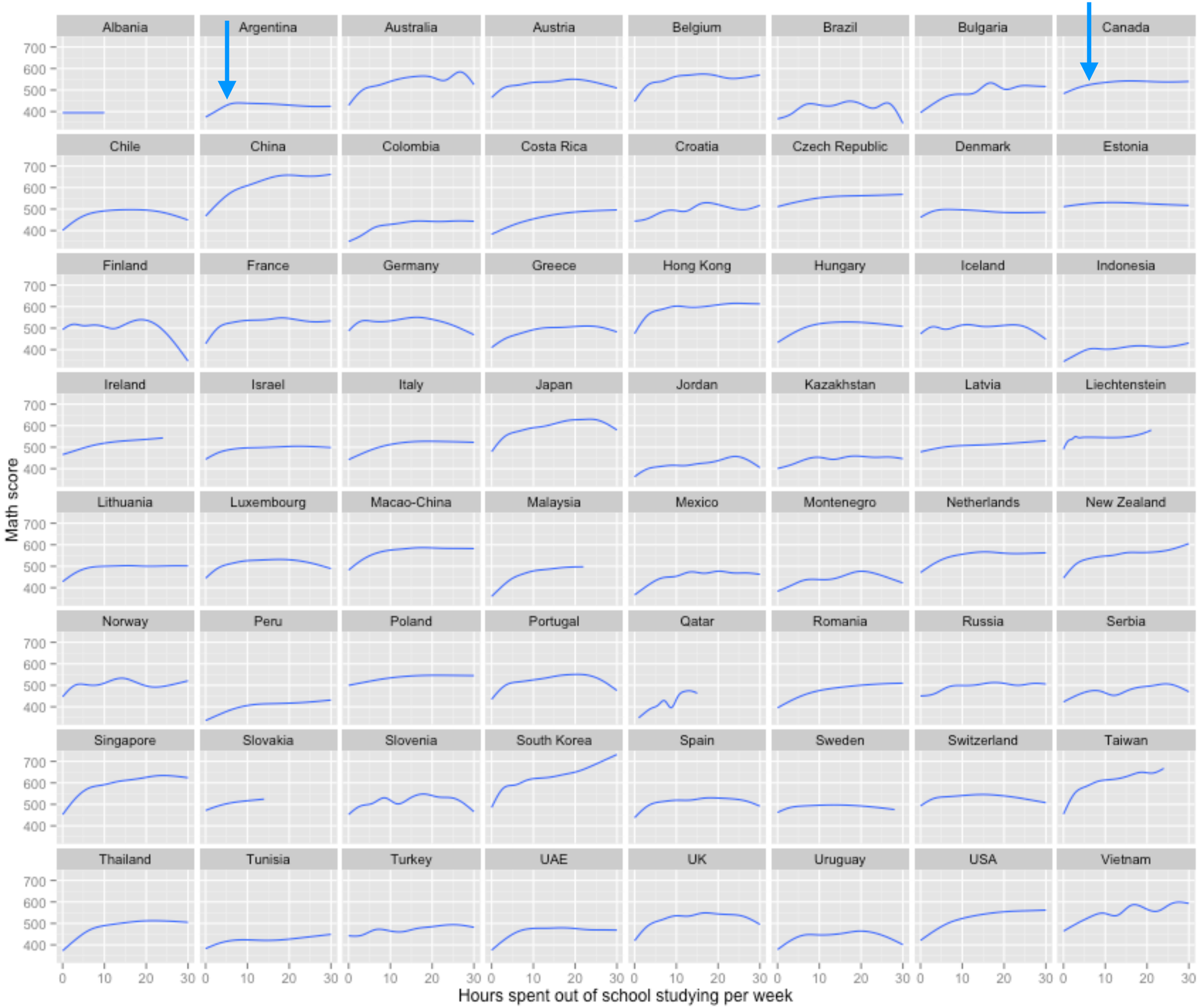


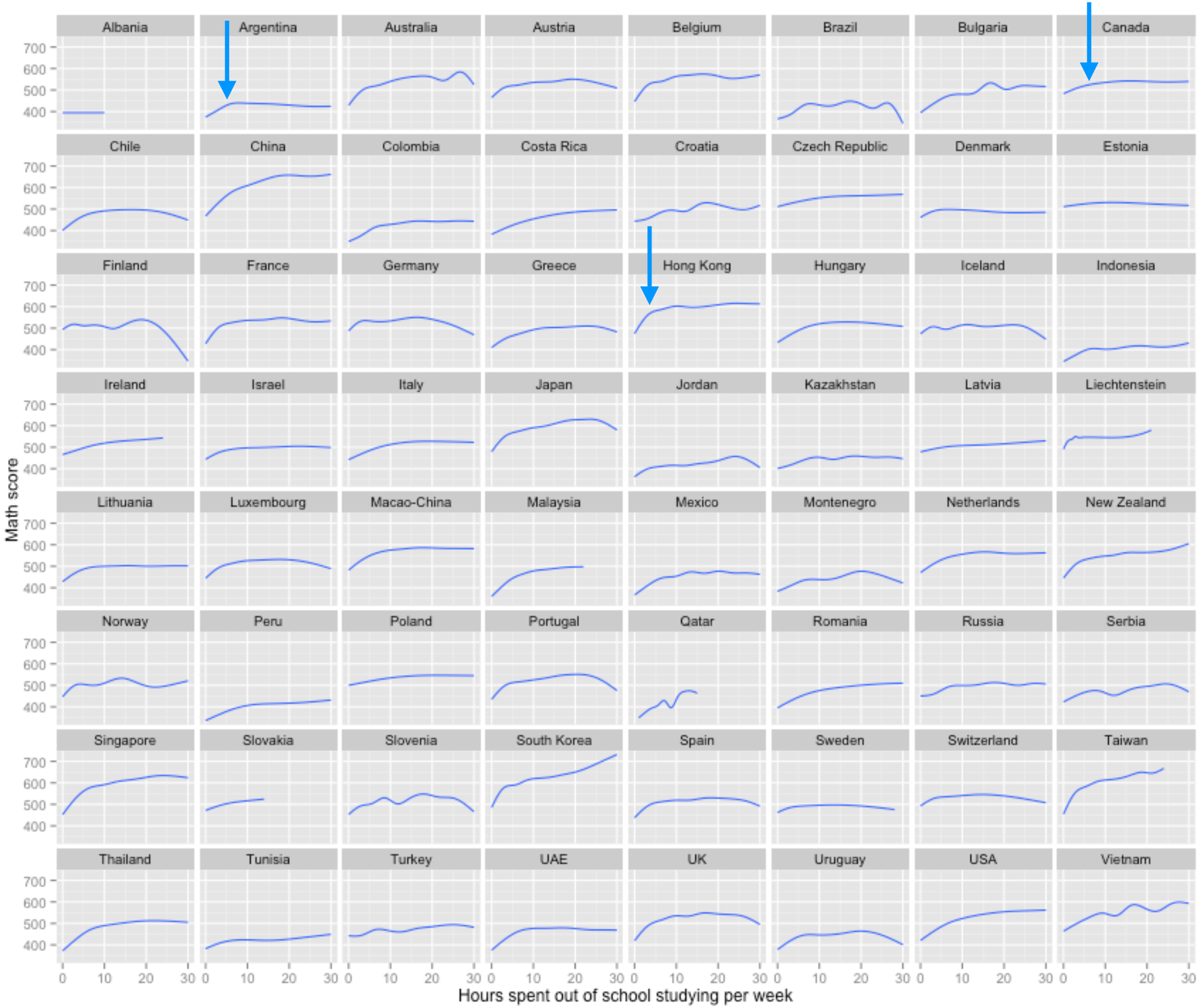
Hours spent out studying out of school time and average math score by country

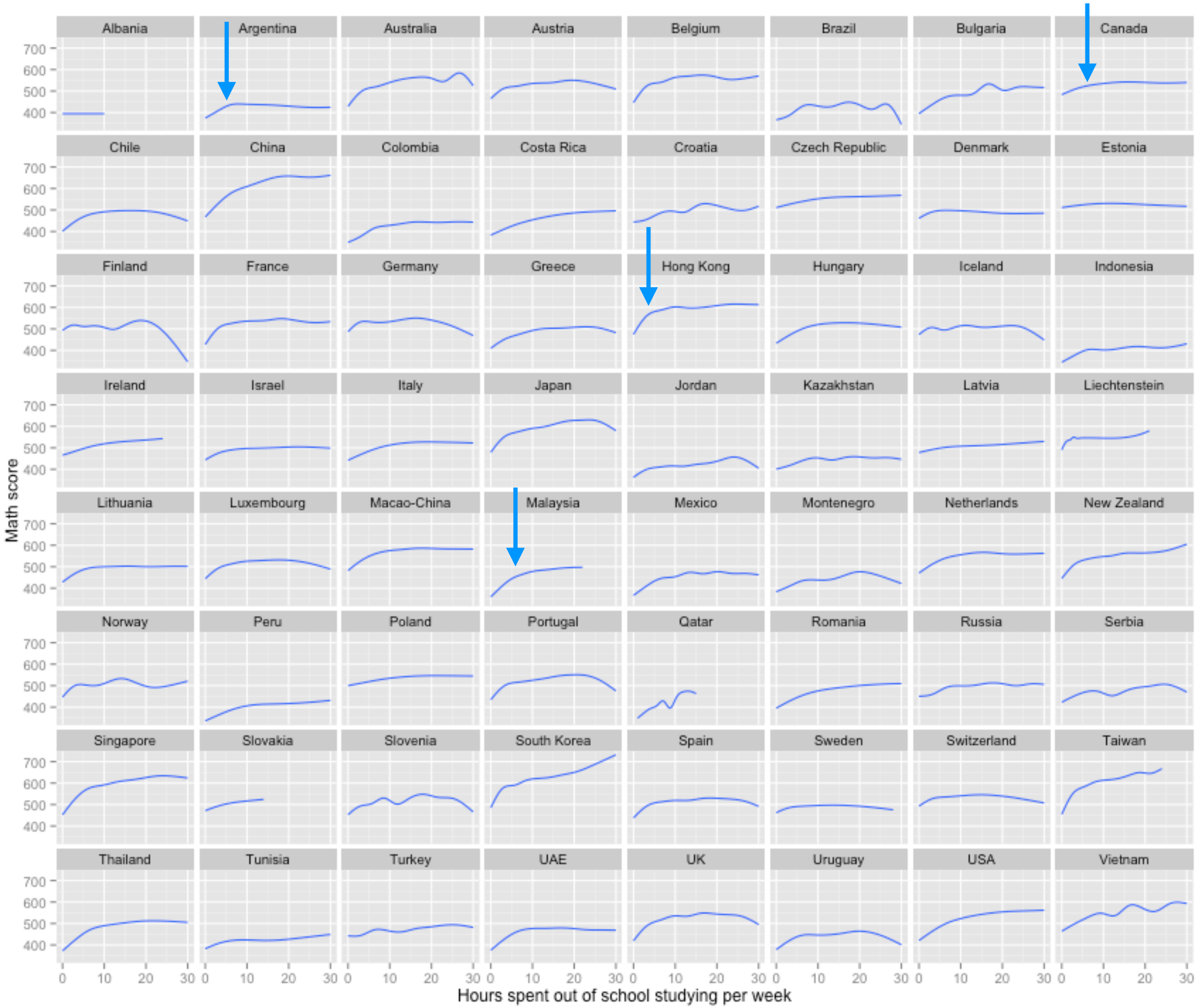


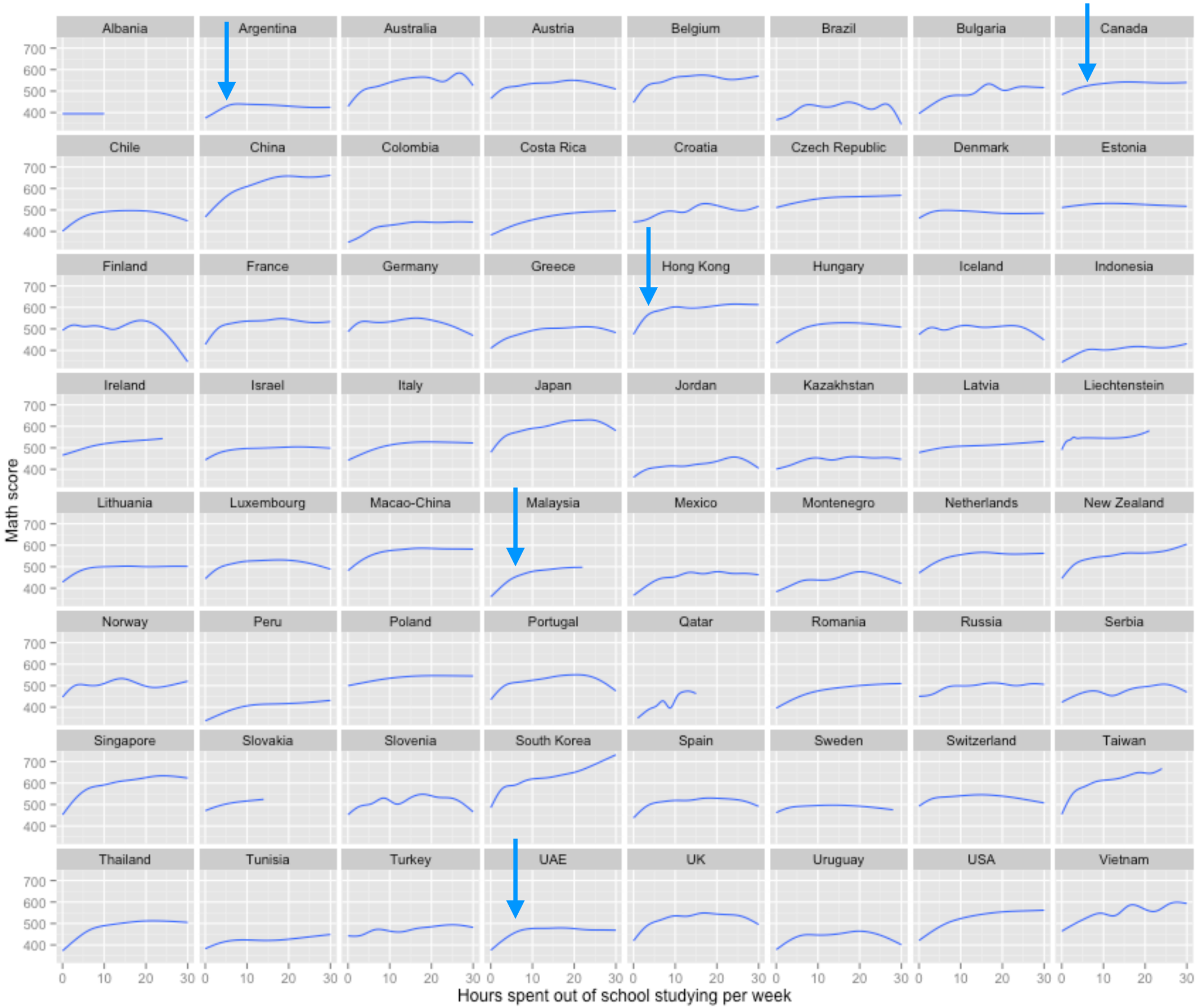

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

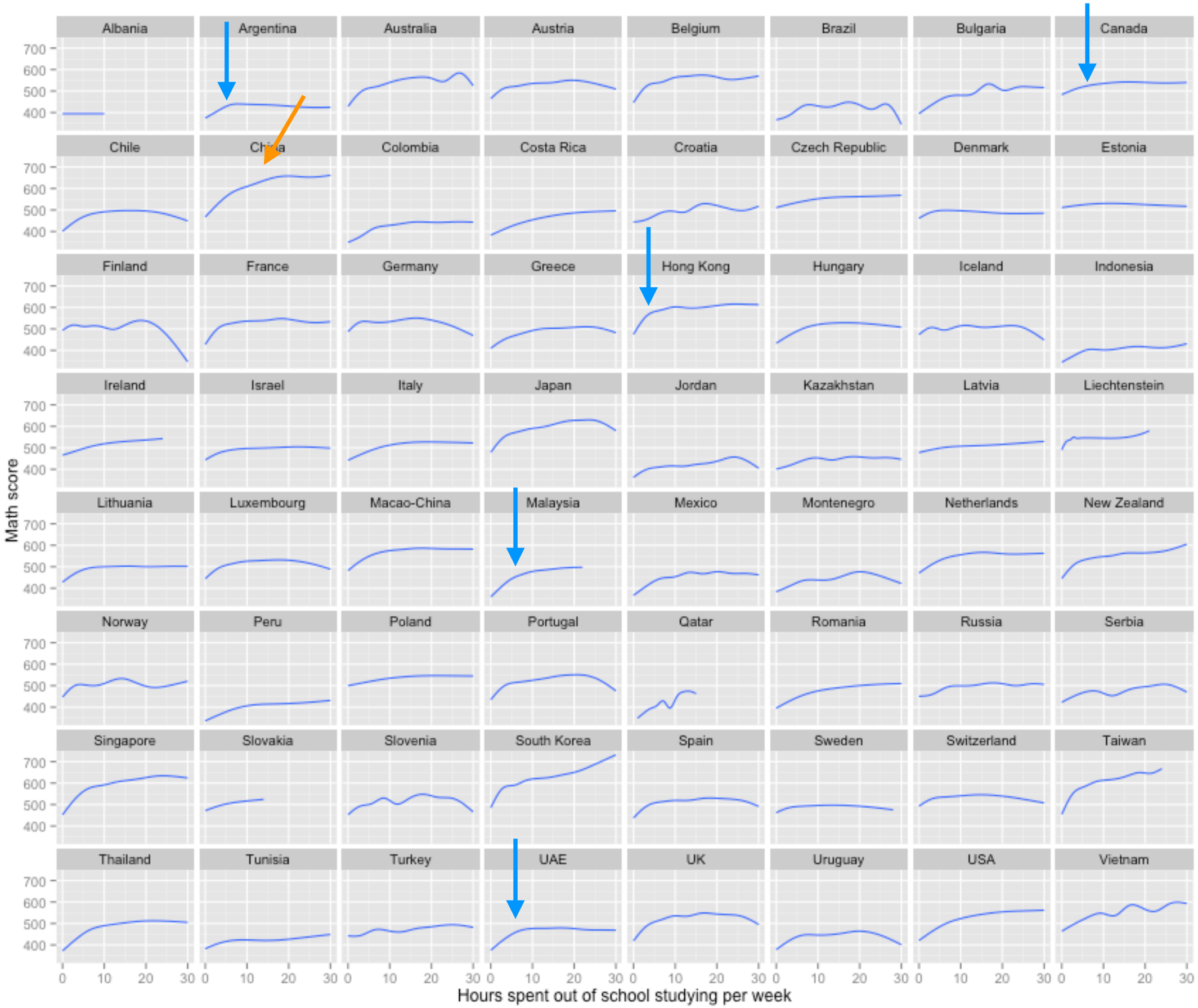

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

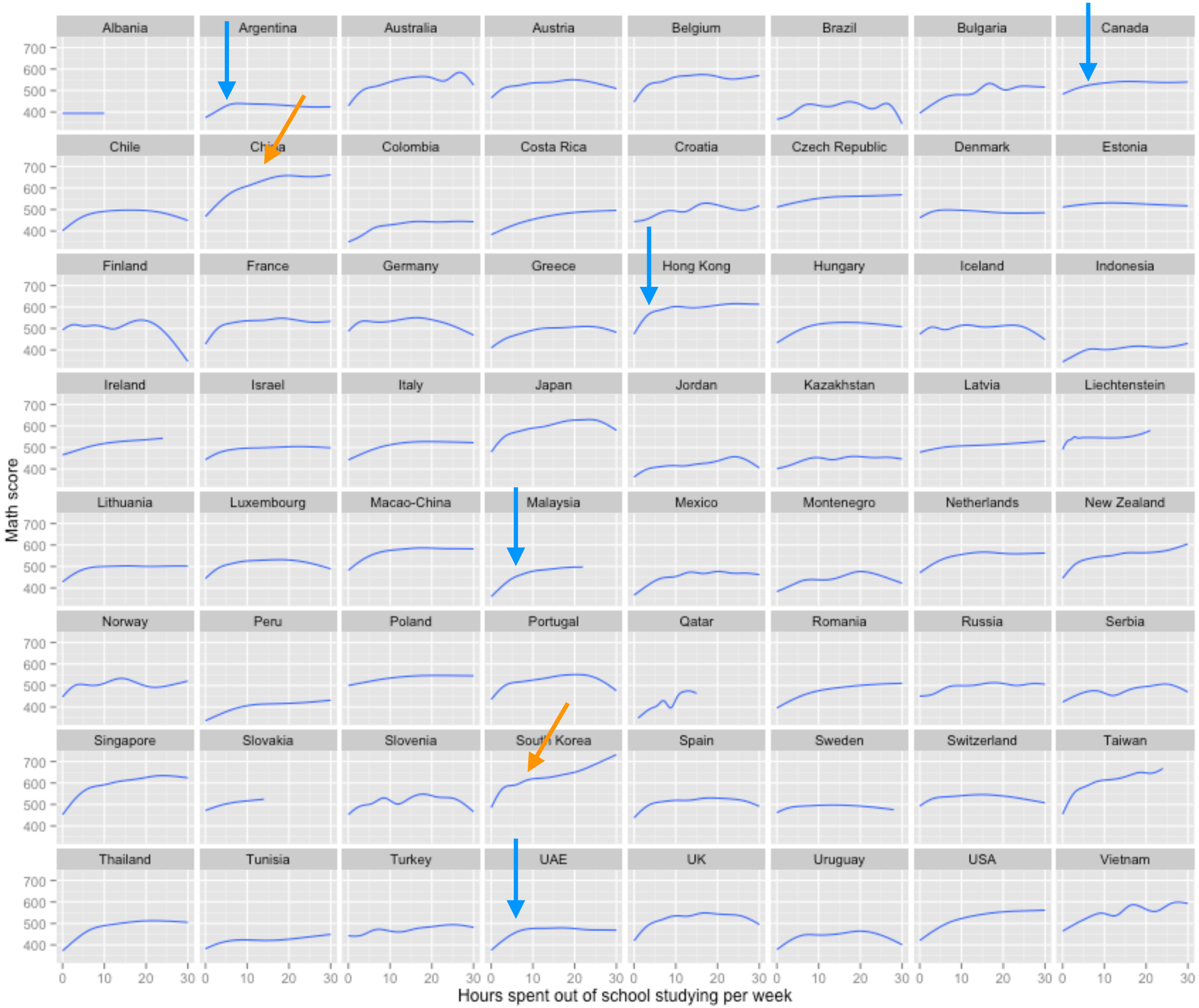

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

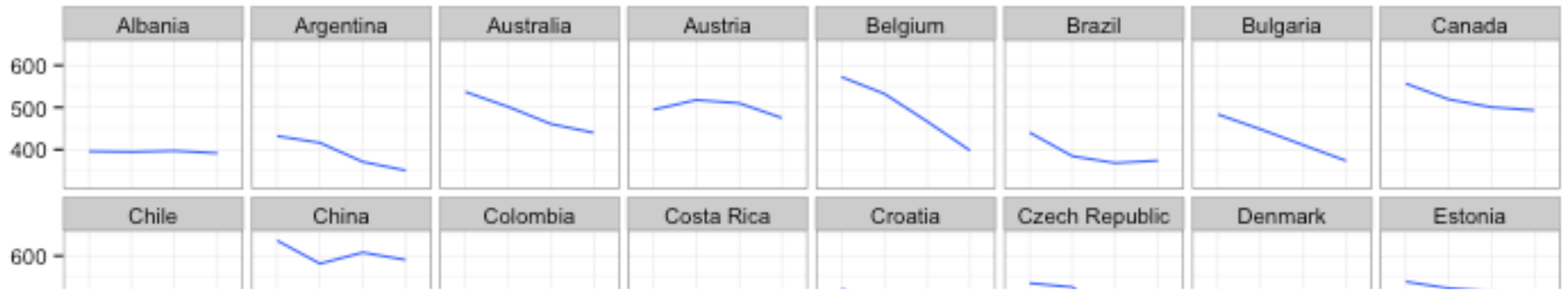

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
400
500
600
700
0
10
20
30
Hours spent out of school studying per week


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

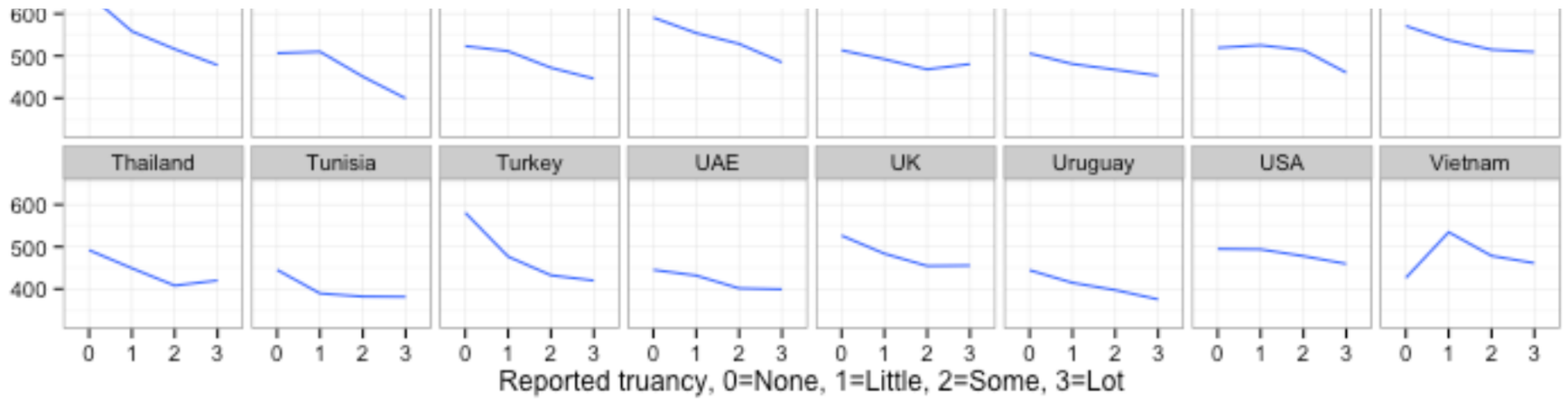

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

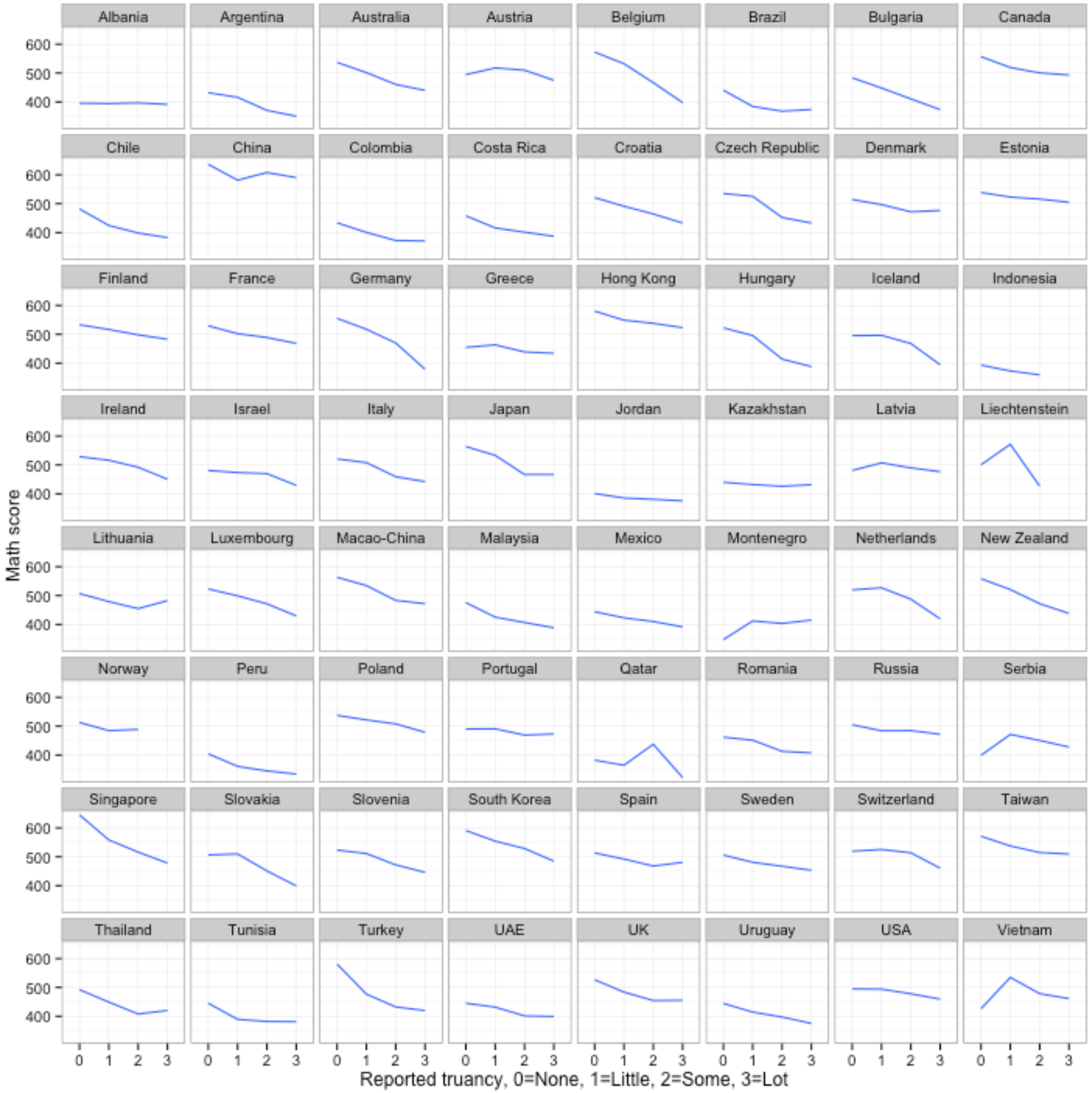

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
400
500
600
700
0
10
20
30
Hours spent out of school studying per week

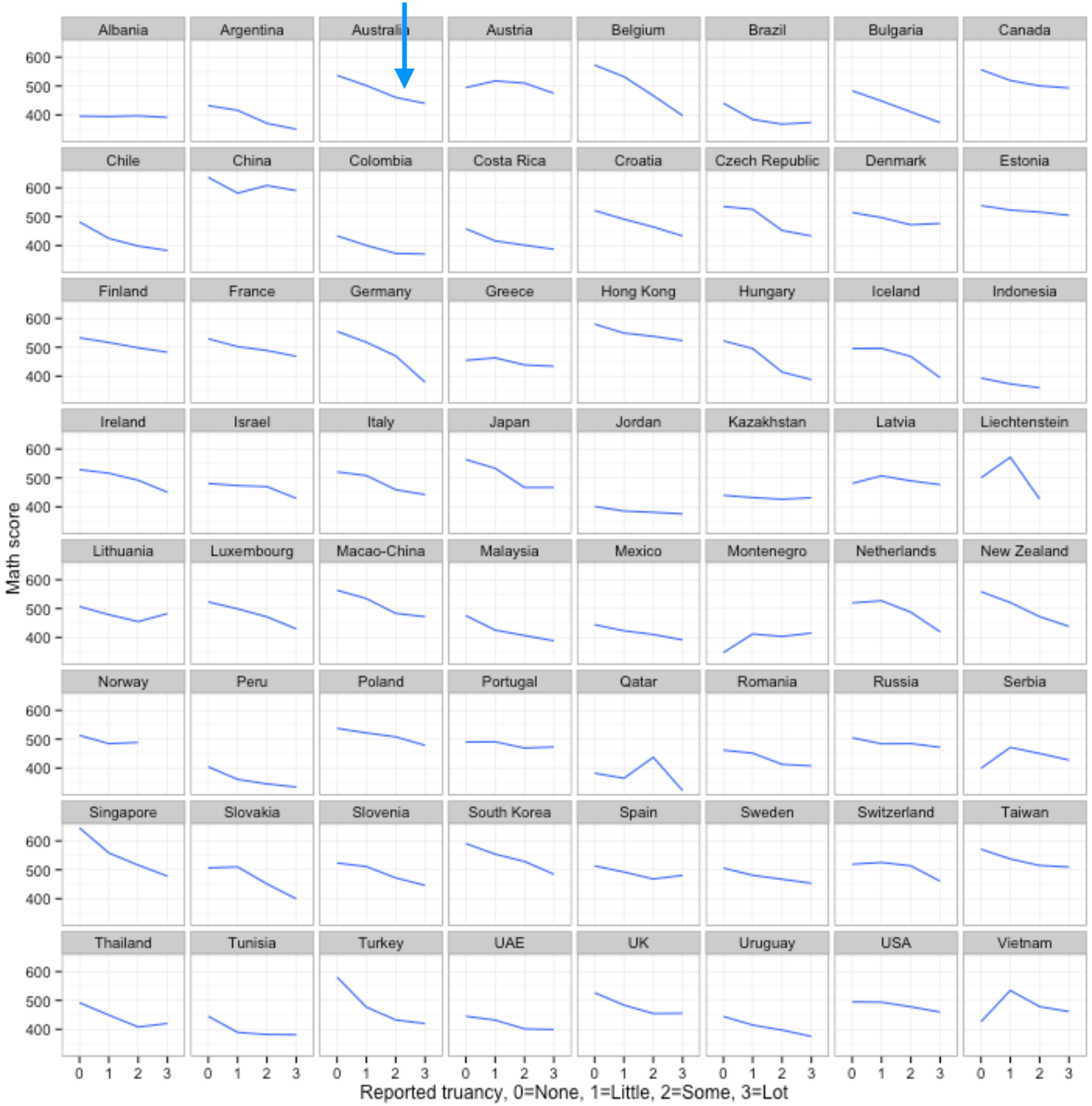

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
700
600
500
400
0
10
20
30
Math score
Hours spent out of school studying per week

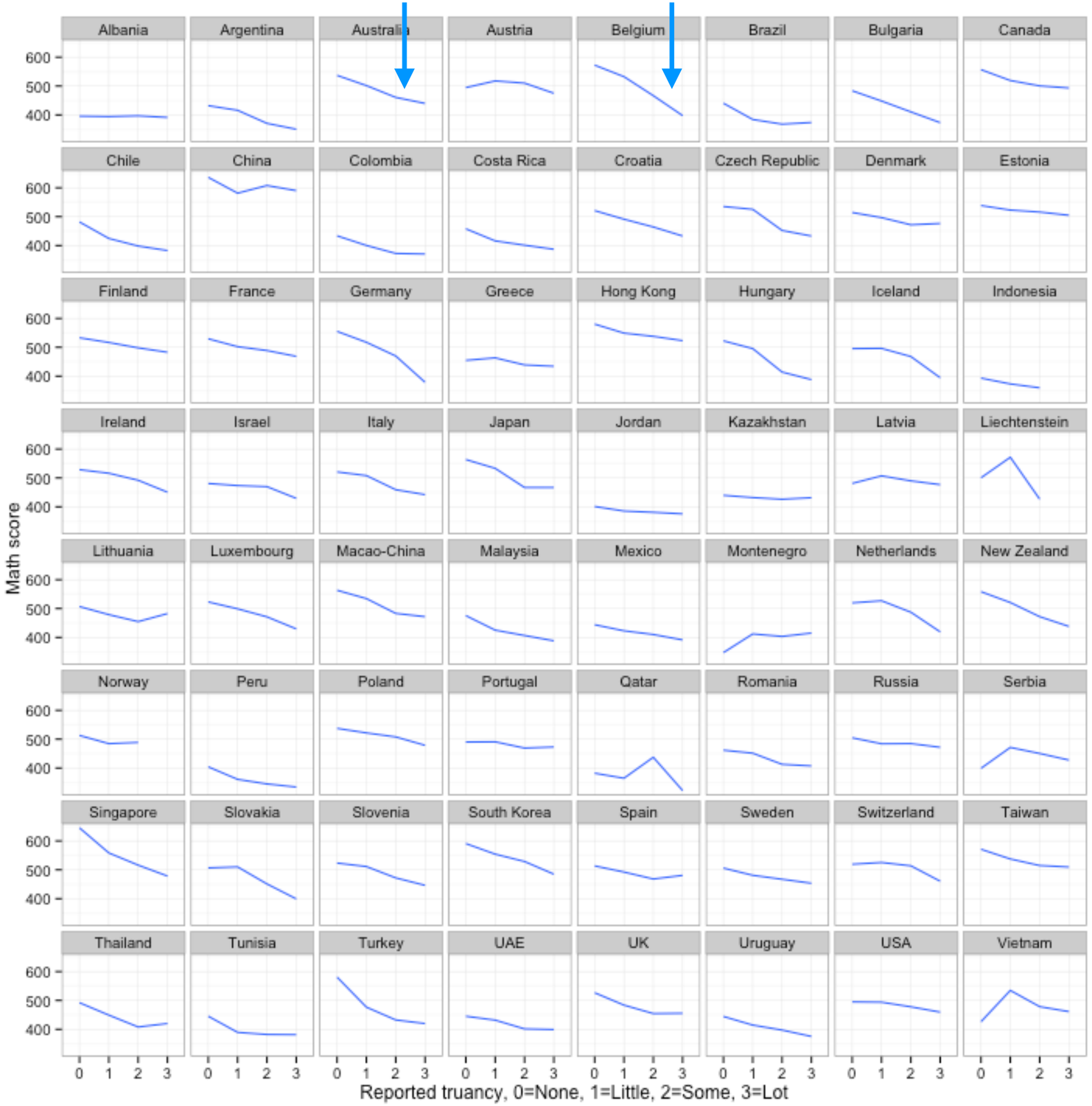


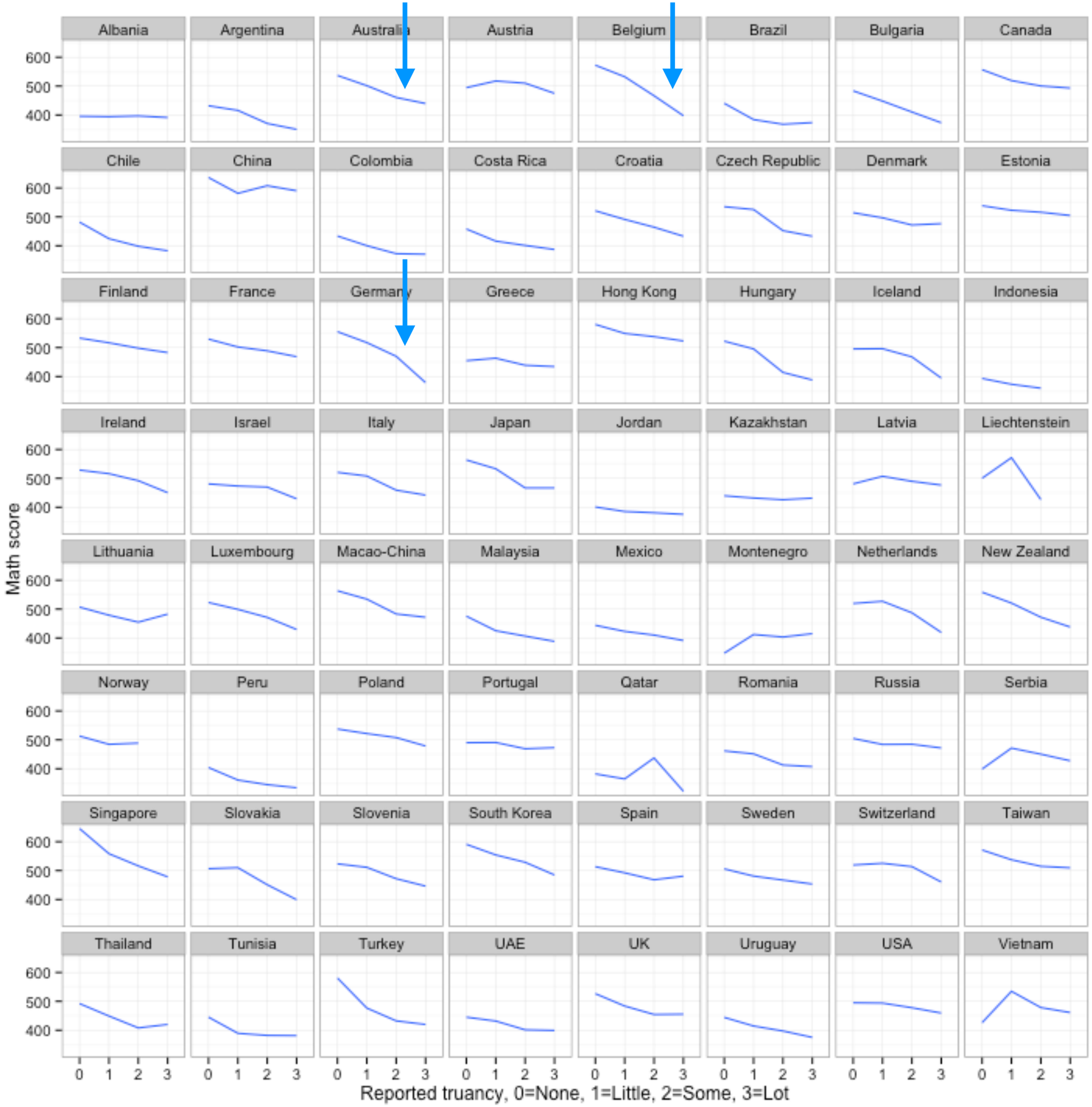
Reported truancy and average math score by country

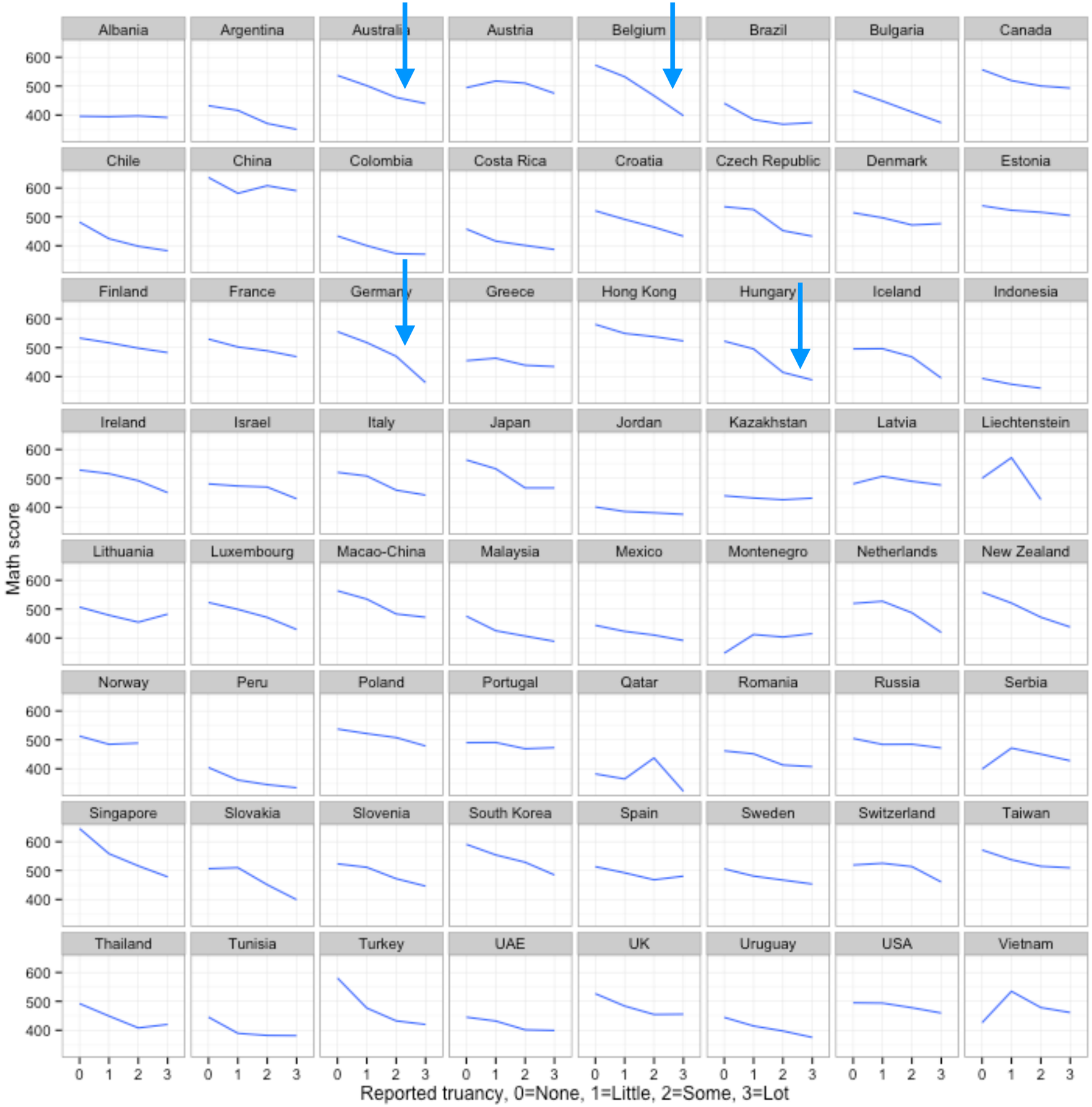


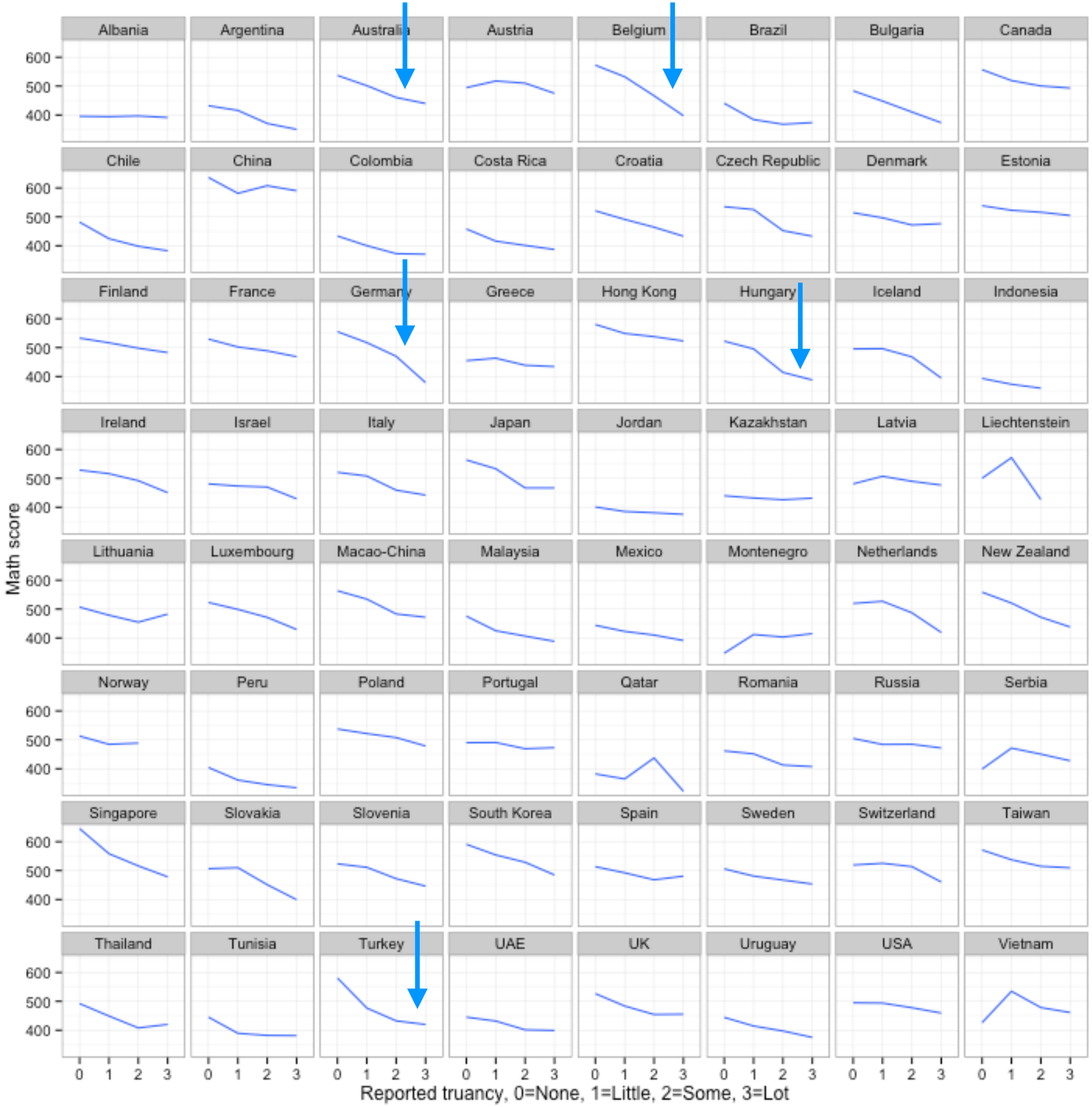

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0
1
2
3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

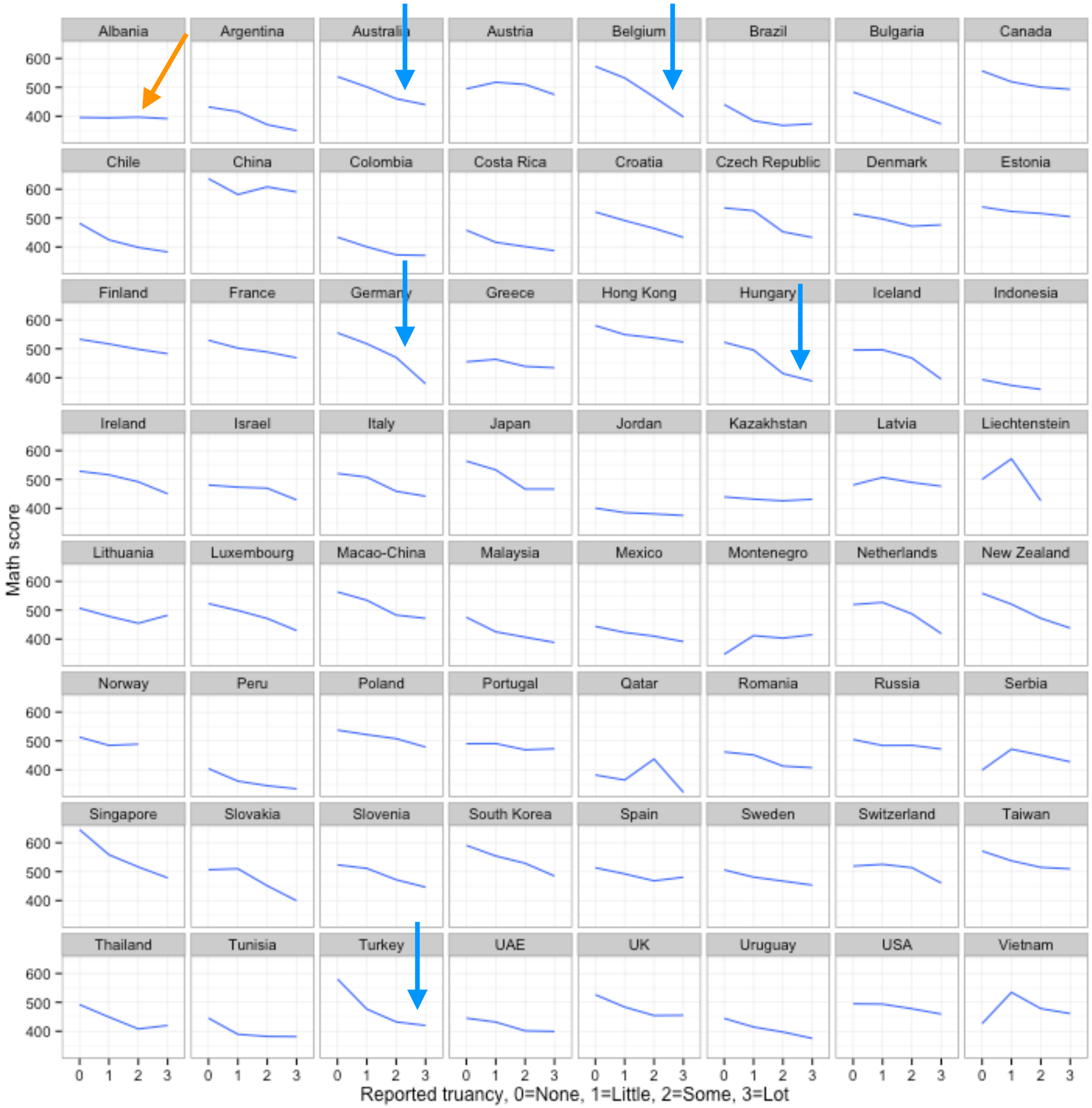

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

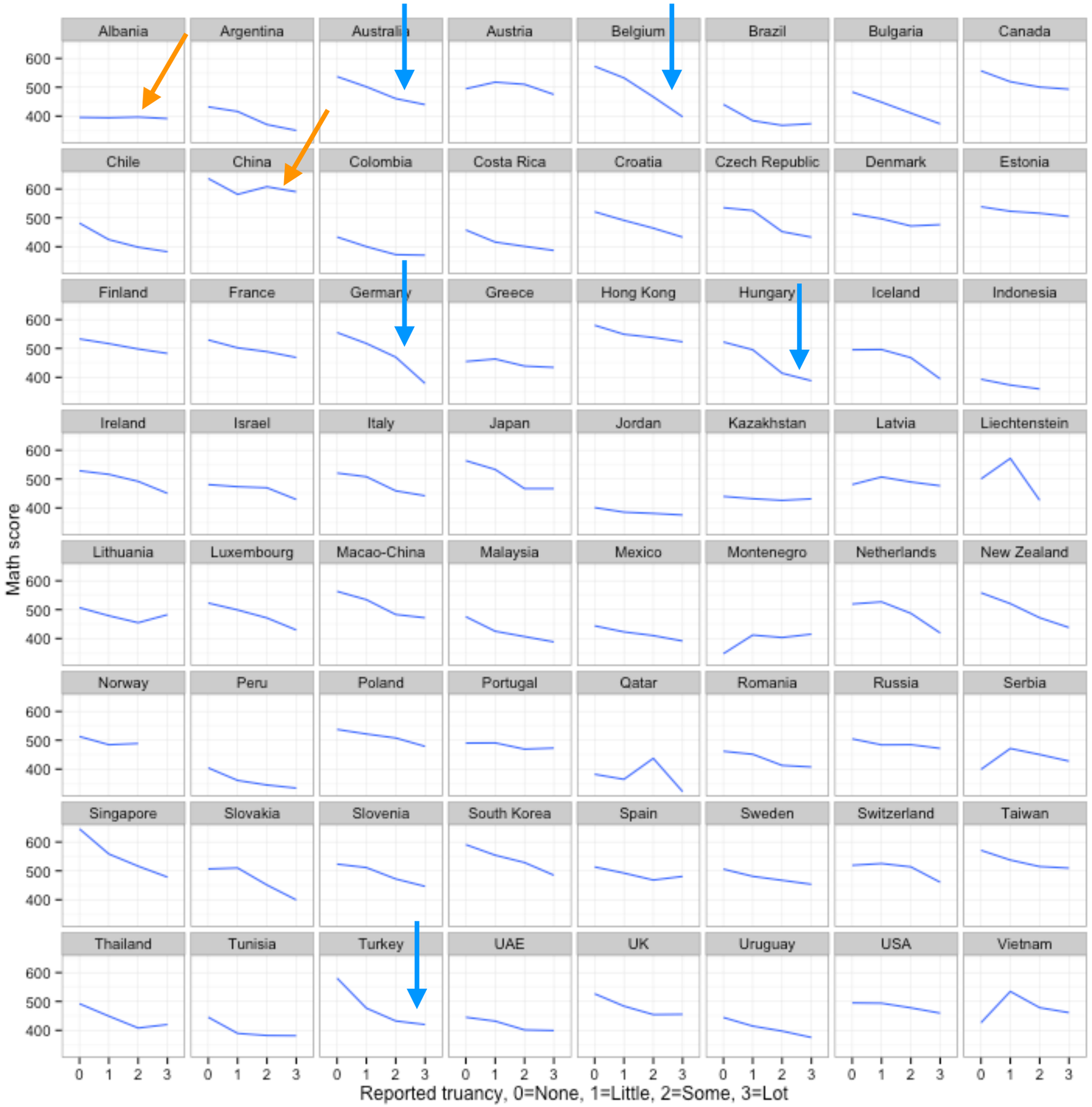

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

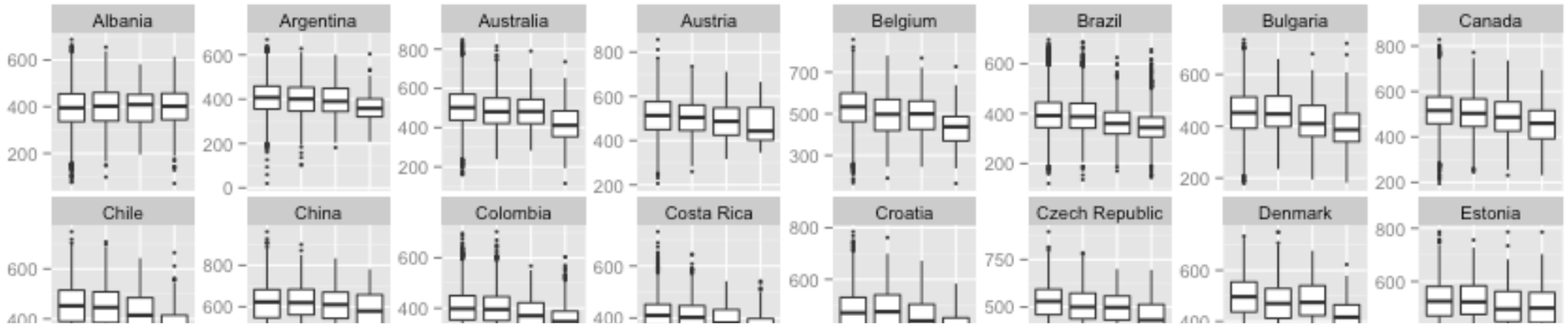

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

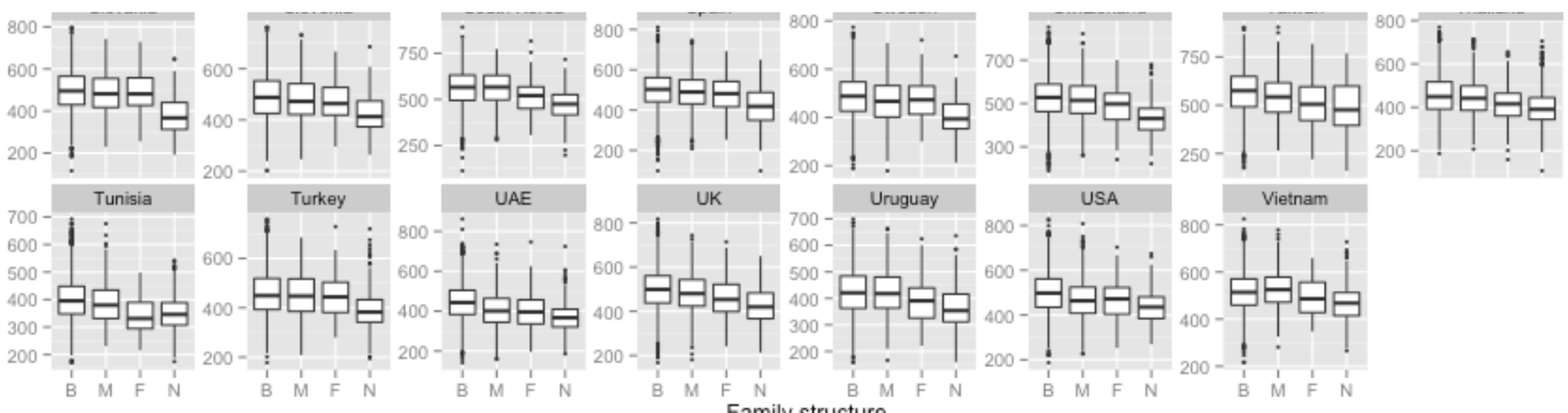

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

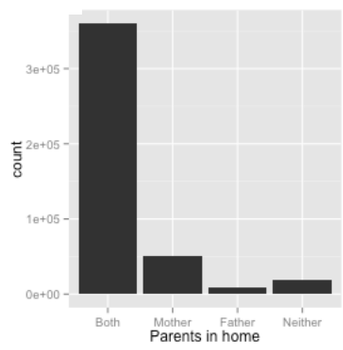

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0
1
2
3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

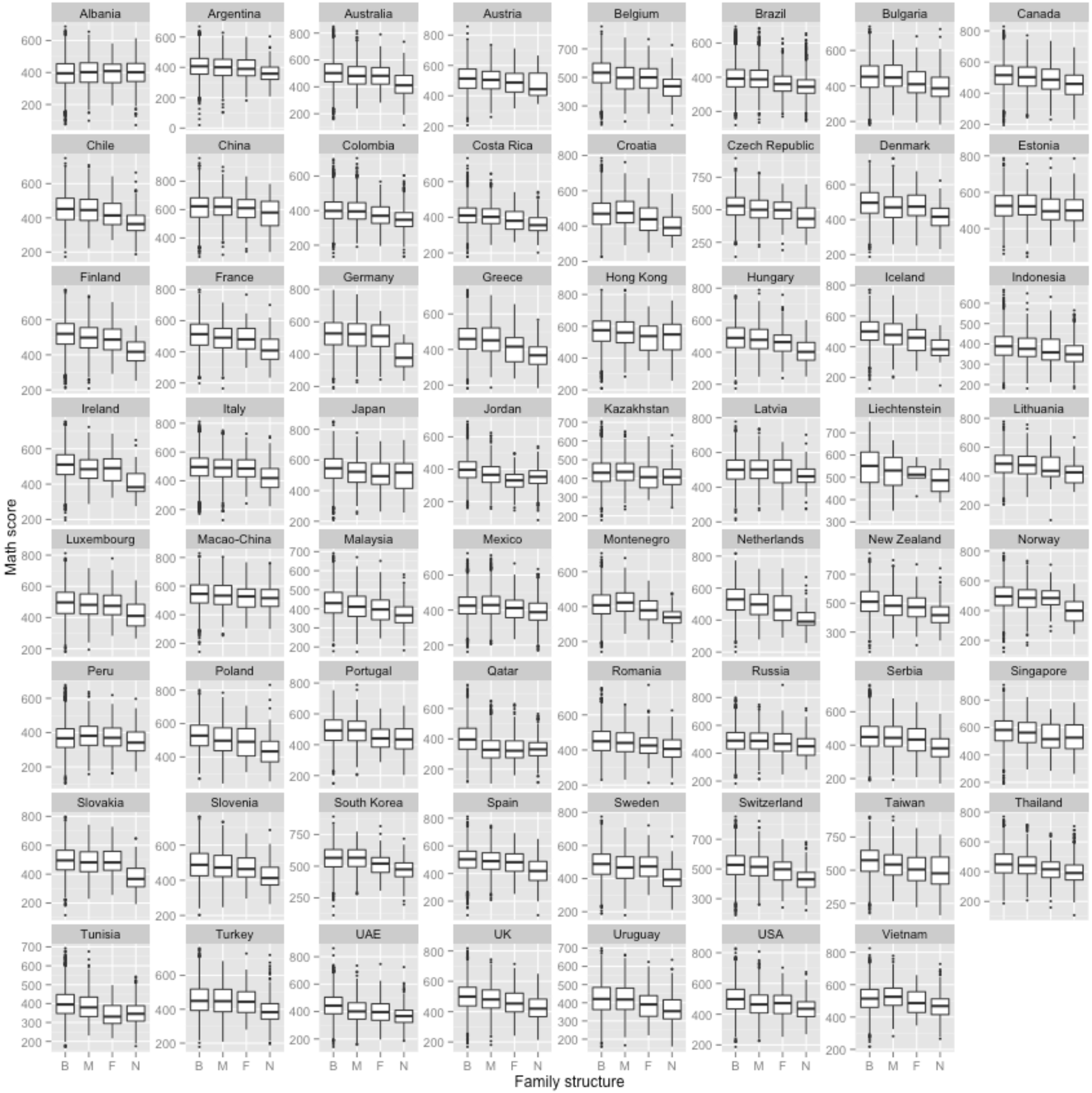

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
0 1 2 3
Reported truancy, 0=None, 1=Little, 2=Some, 3=Lot

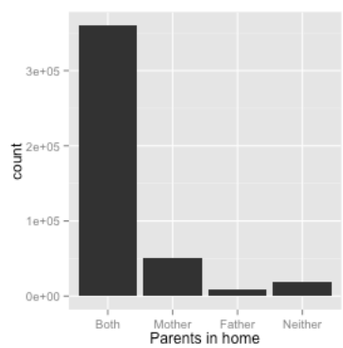


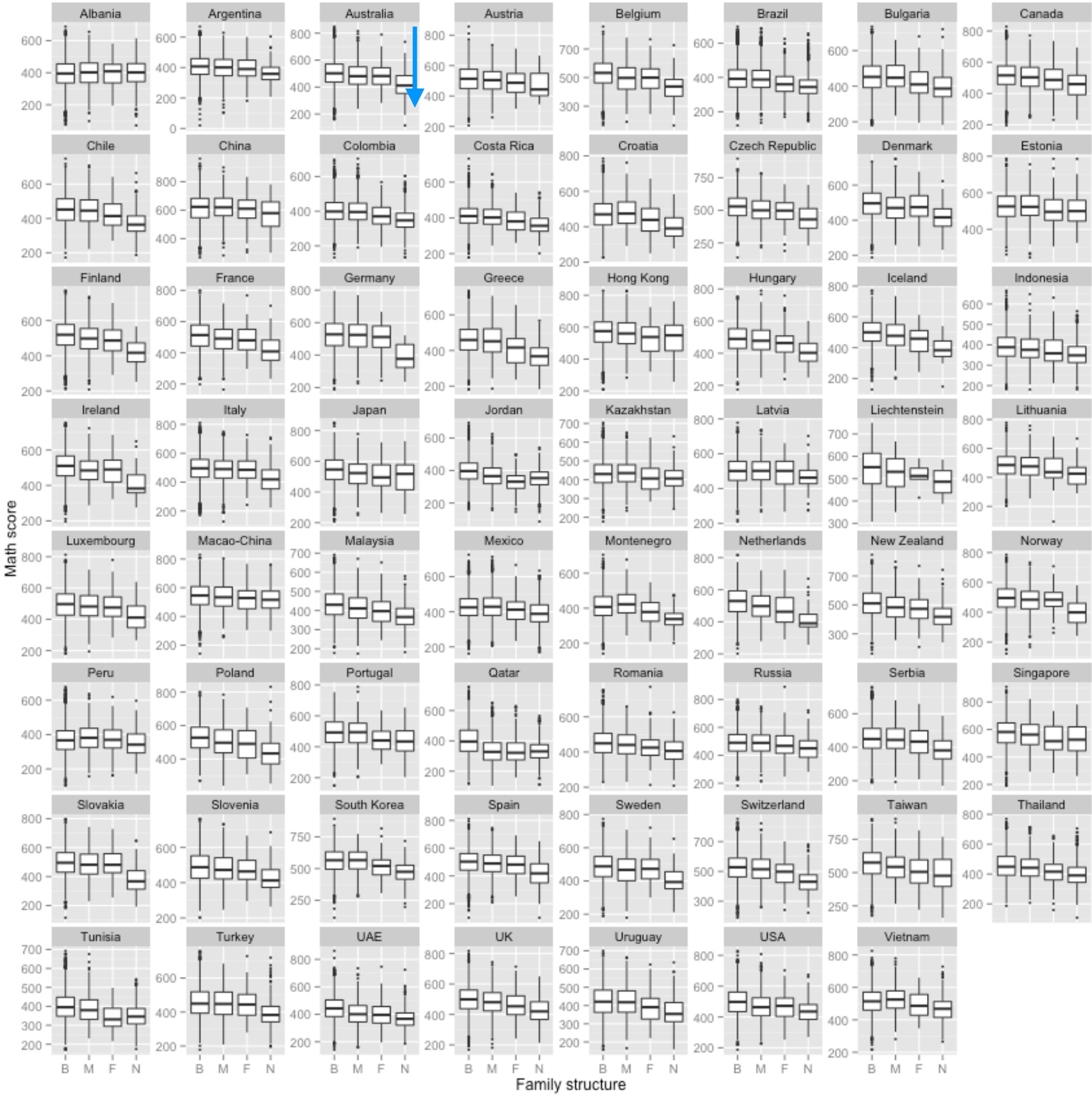
Parents in the home and average math score by country

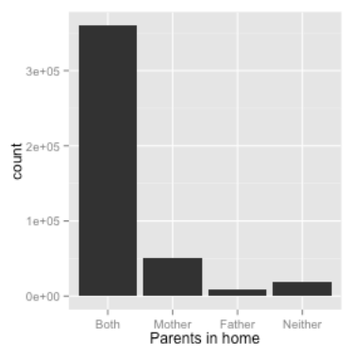


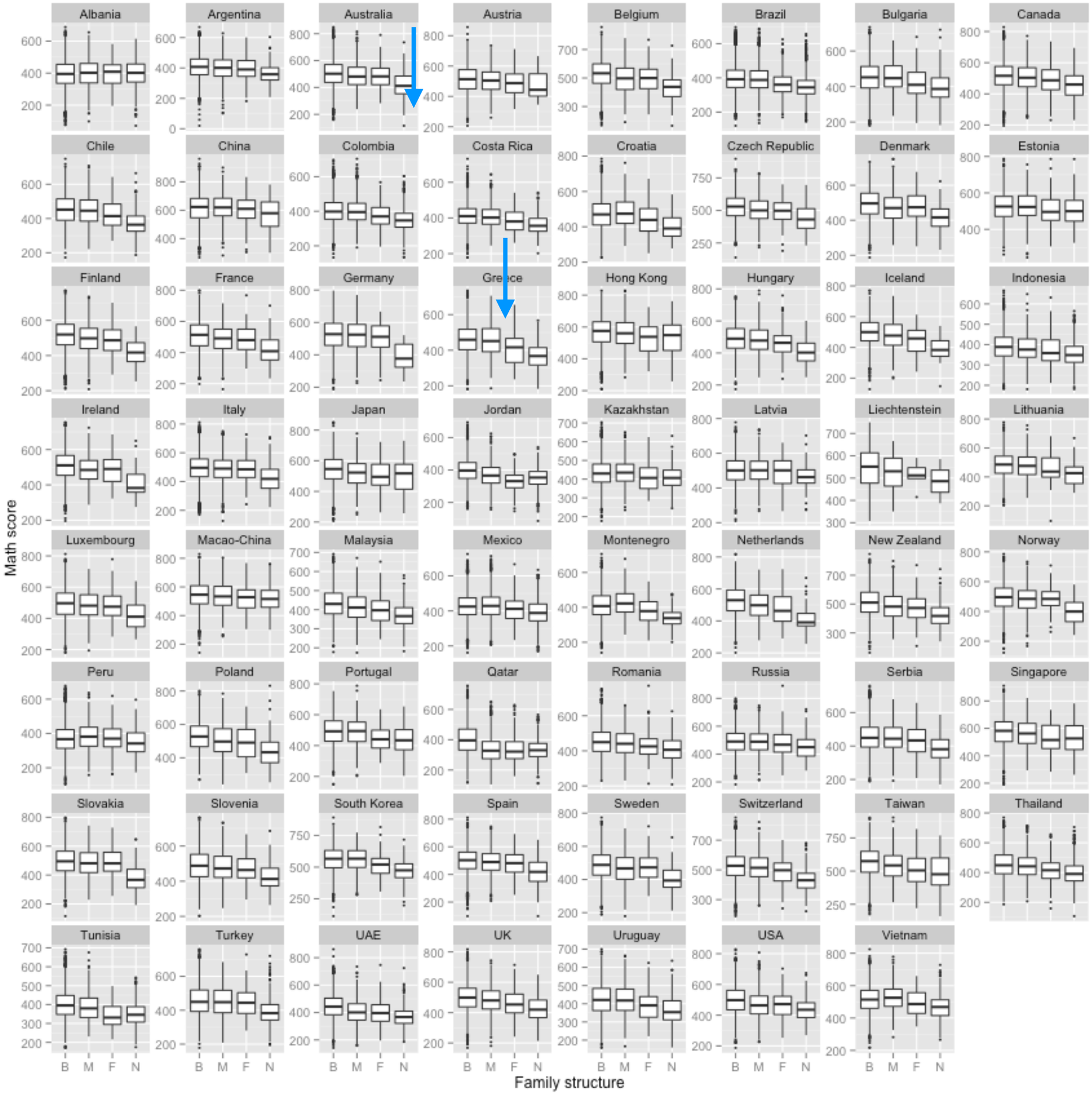


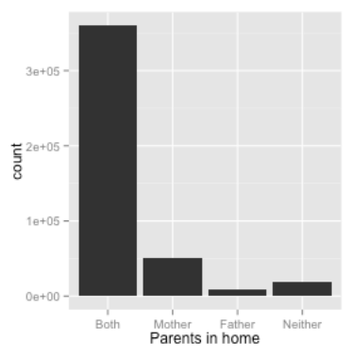

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure


count
Both
Mother
Father
Neither
Parents in home

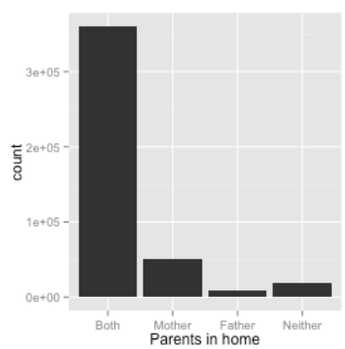

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure


count
3e+05
2e+05
1e+05
0e+00
Both
Mother
Father
Neither
Parents in home

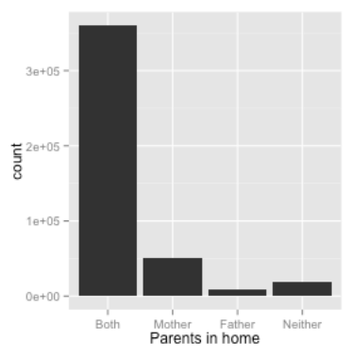

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
B M F N
Family structure
Math score


count
Both
Mother
Father
Neither
Parents in home

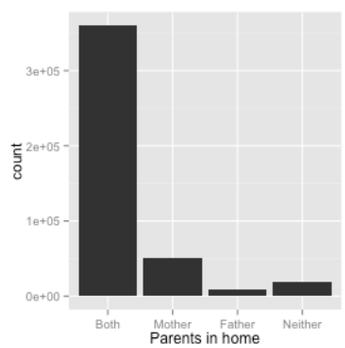
Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
B M F N
Family structure
Math score


count
Both
Mother
Father
Neither
Parents in home

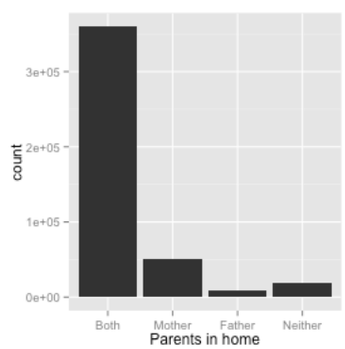
Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
B M F N
Family structure
Math score

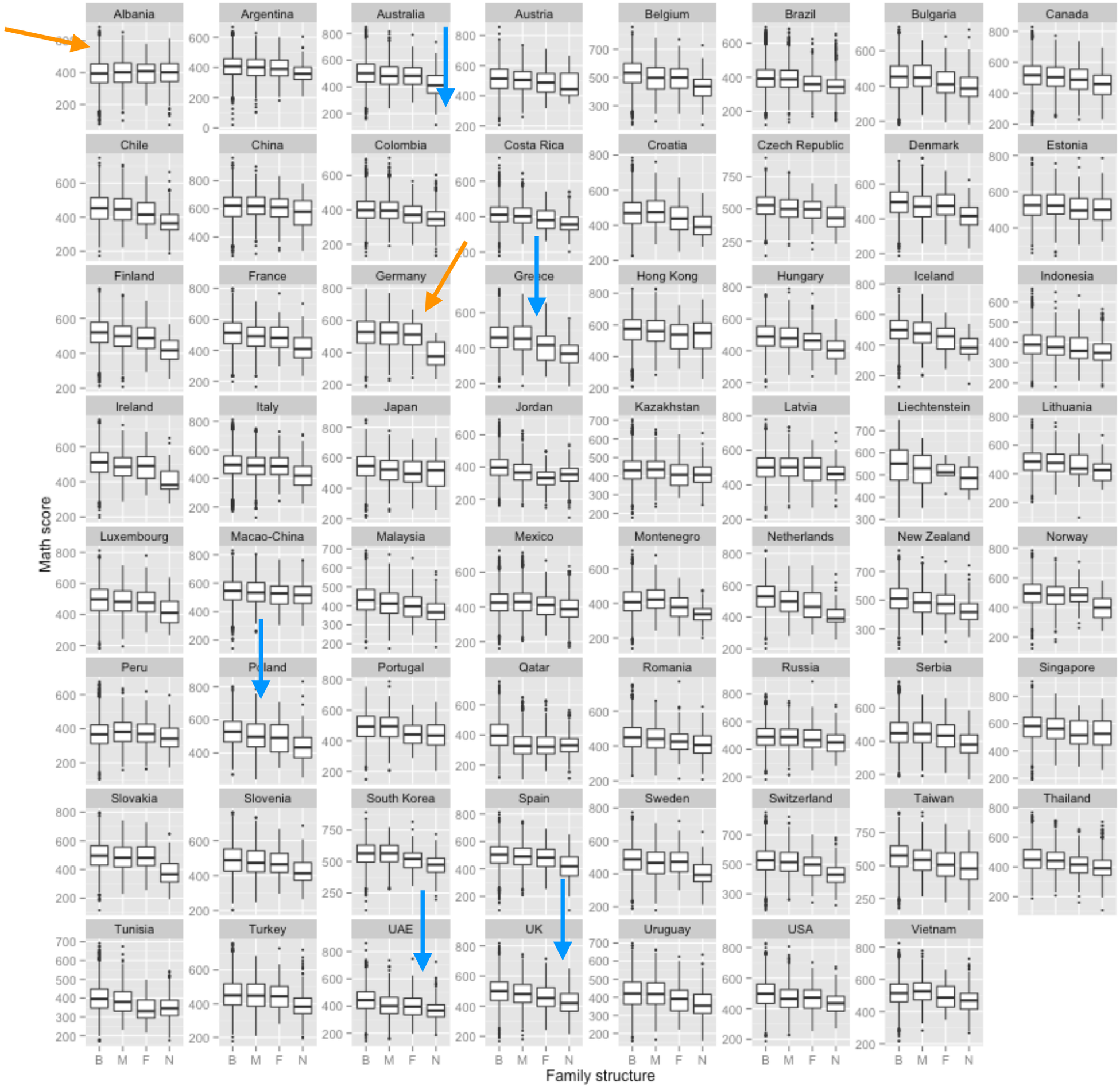

count
3e+05
2e+05
1e+05
0e+00
Both
Mother
Father
Neither
Parents in home

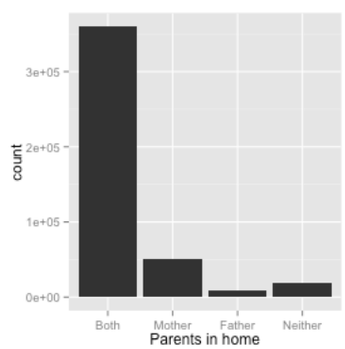
Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure

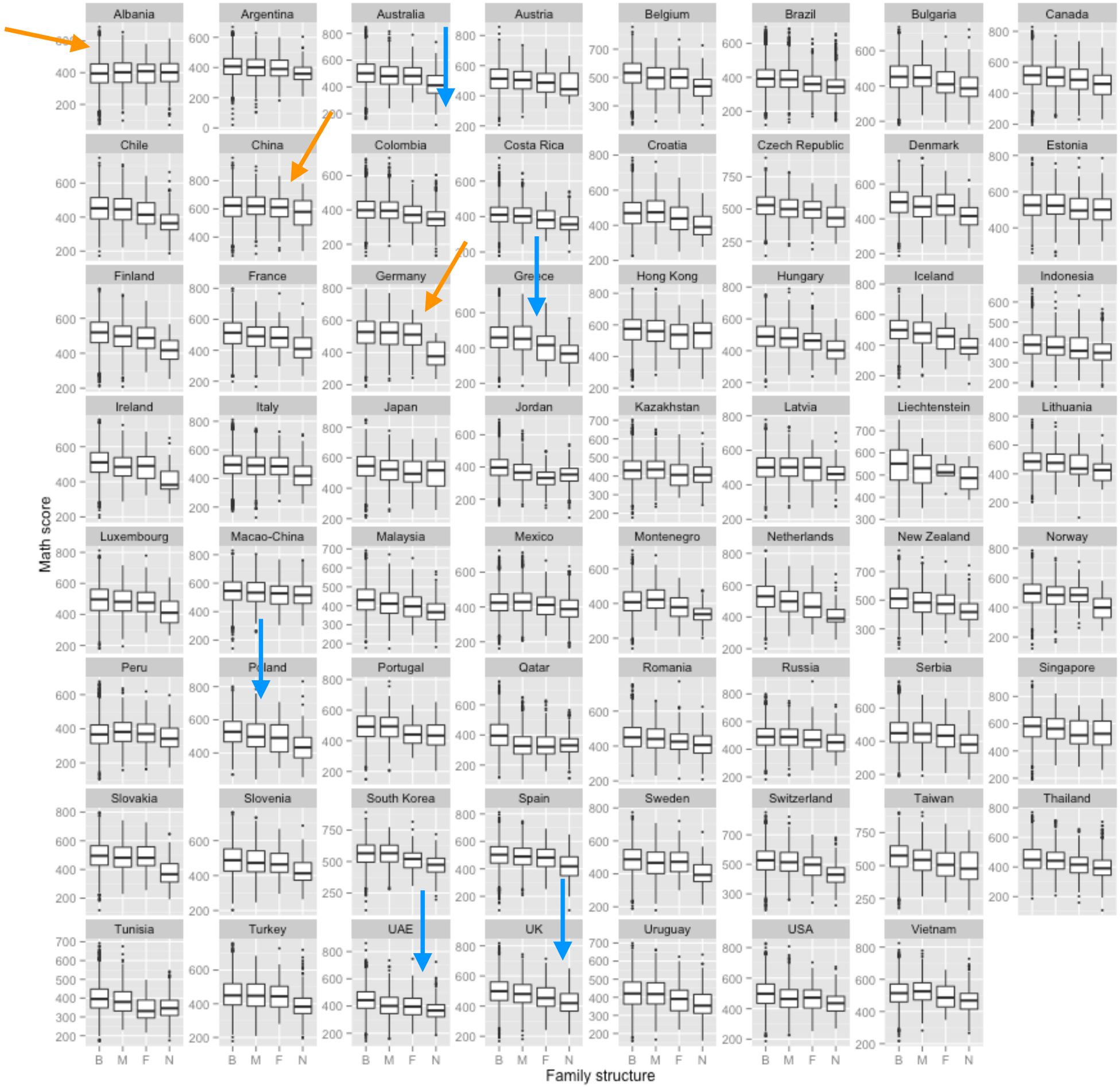

count
Both
Mother
Father
Neither
Parents in home

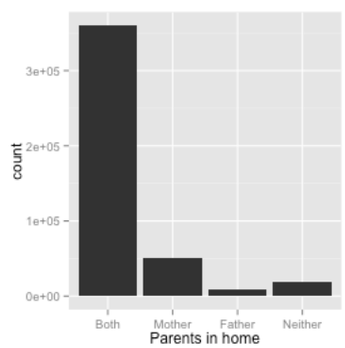
Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure

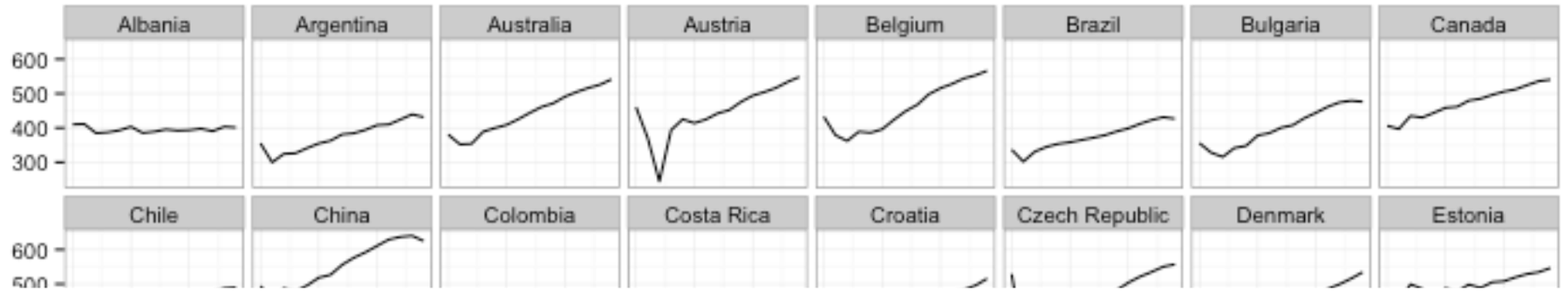

count
Both
Mother
Father
Neither
Parents in home


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure

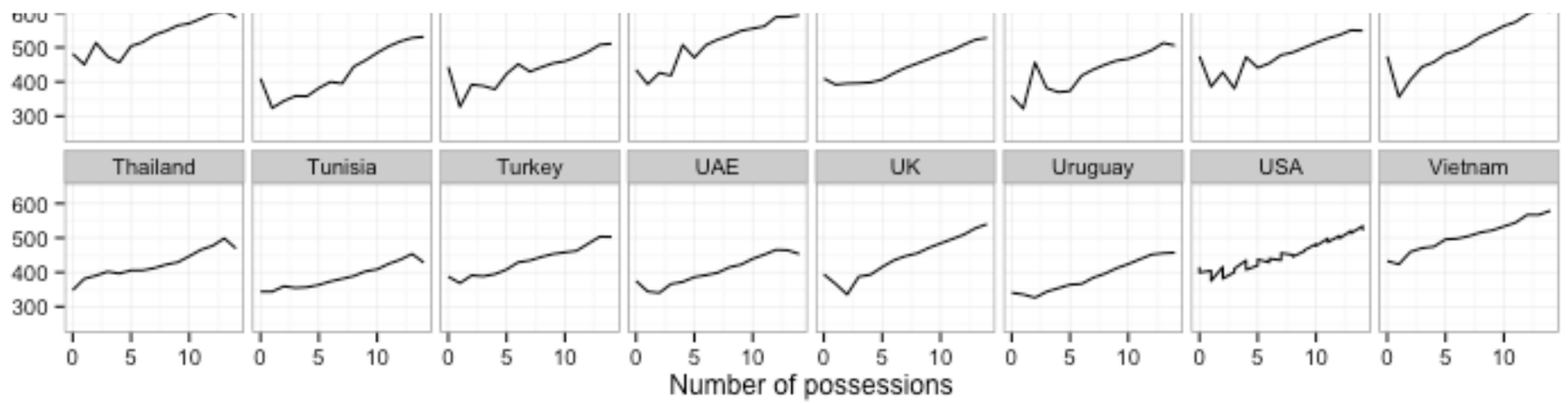

count
3e+05
2e+05
1e+05
0e+00
Both
Mother
Father
Neither
Parents in home

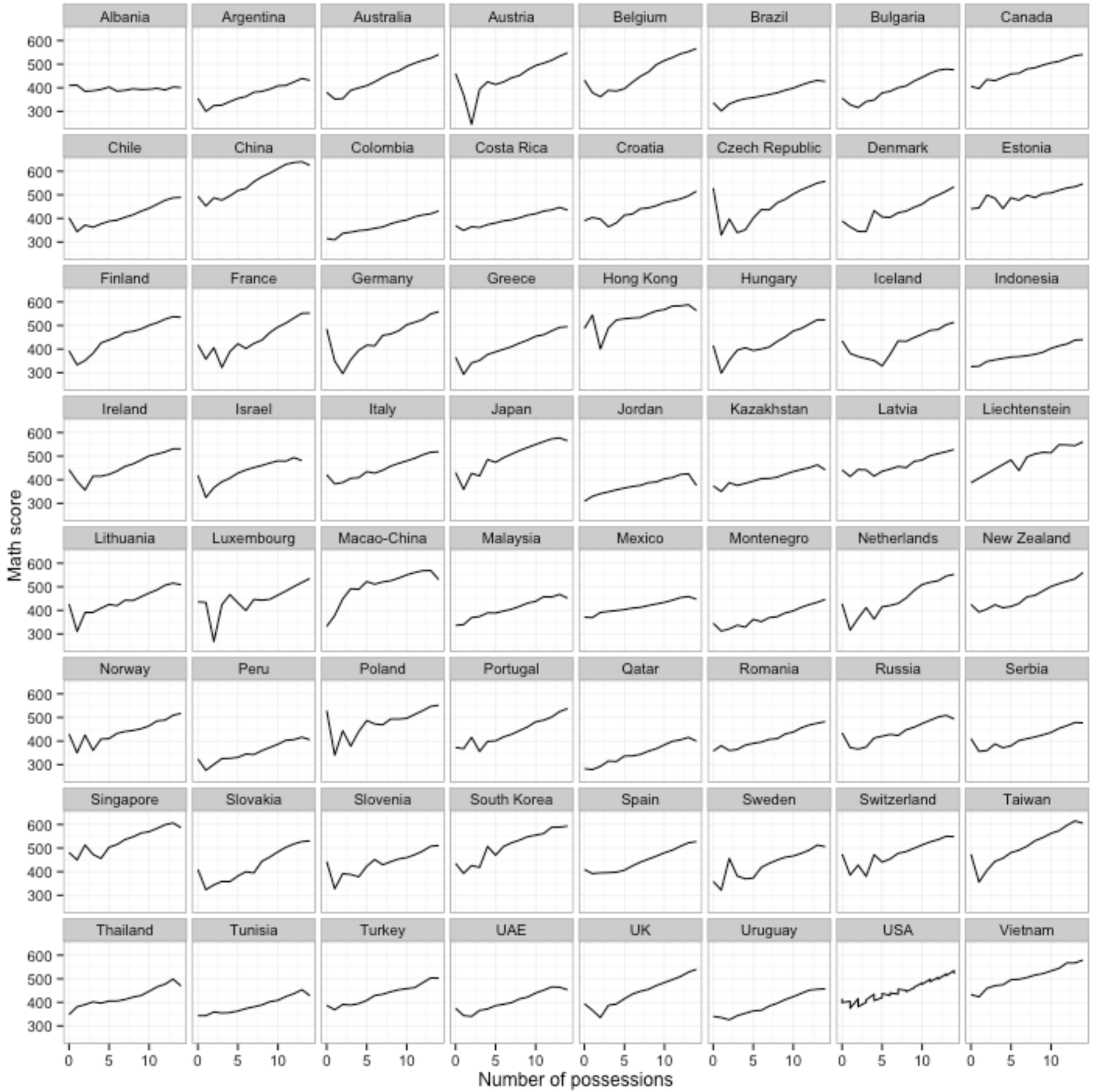

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
Math score
B M F N
Family structure

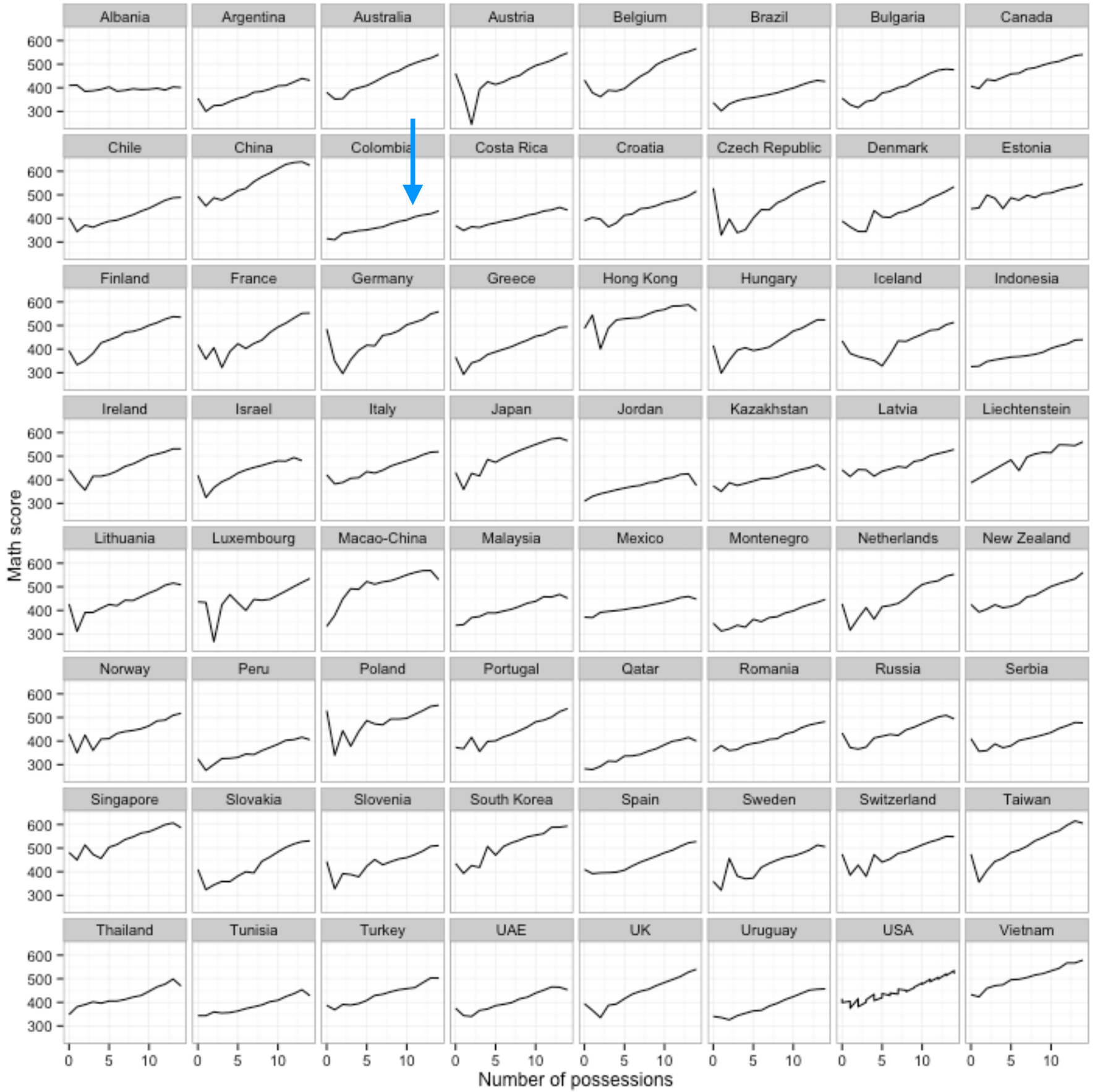

count
3e+05
2e+05
1e+05
0e+00
Both
Mother
Father
Neither
Parents in home

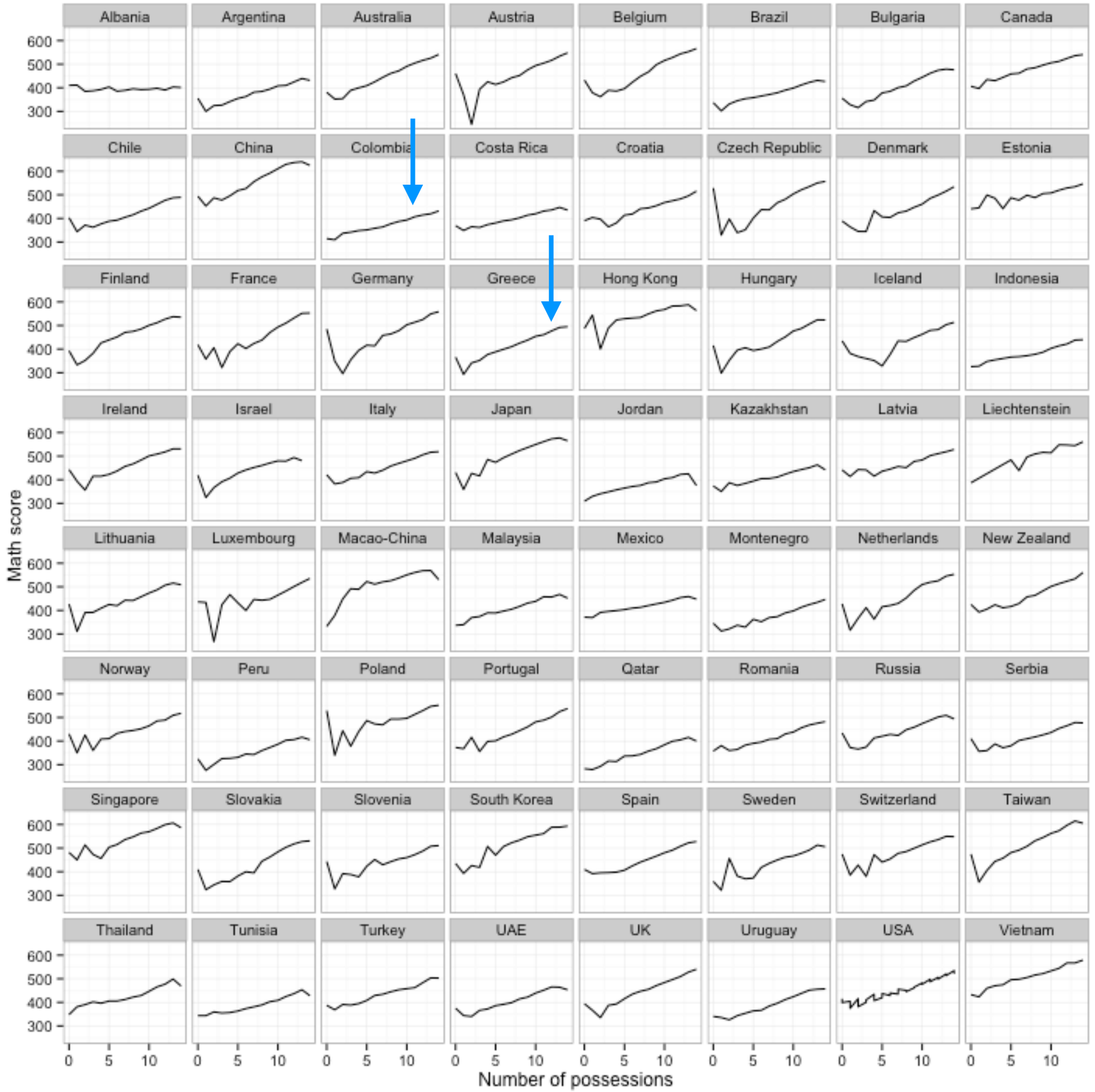


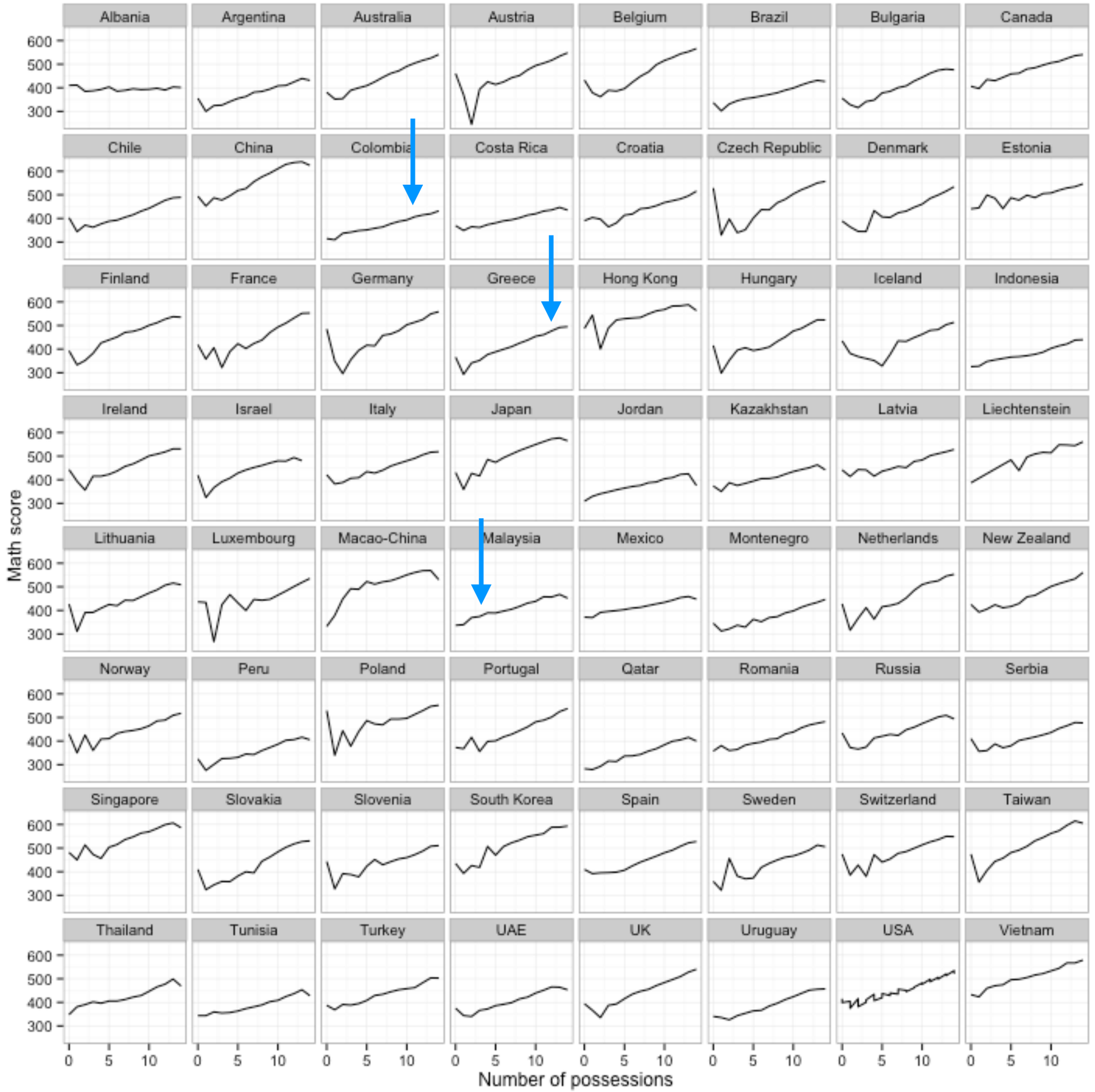
Possessions and average math score by country

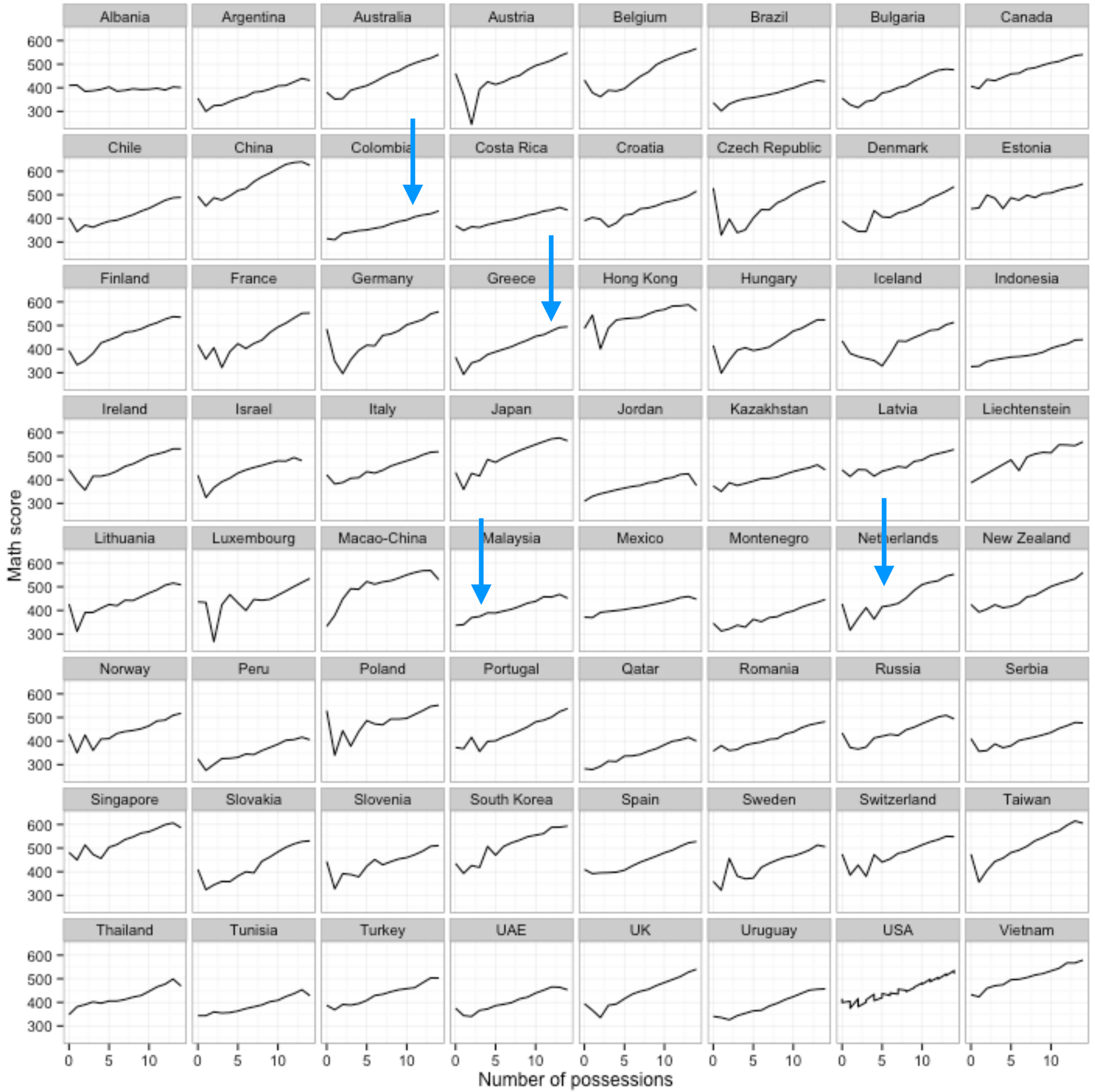


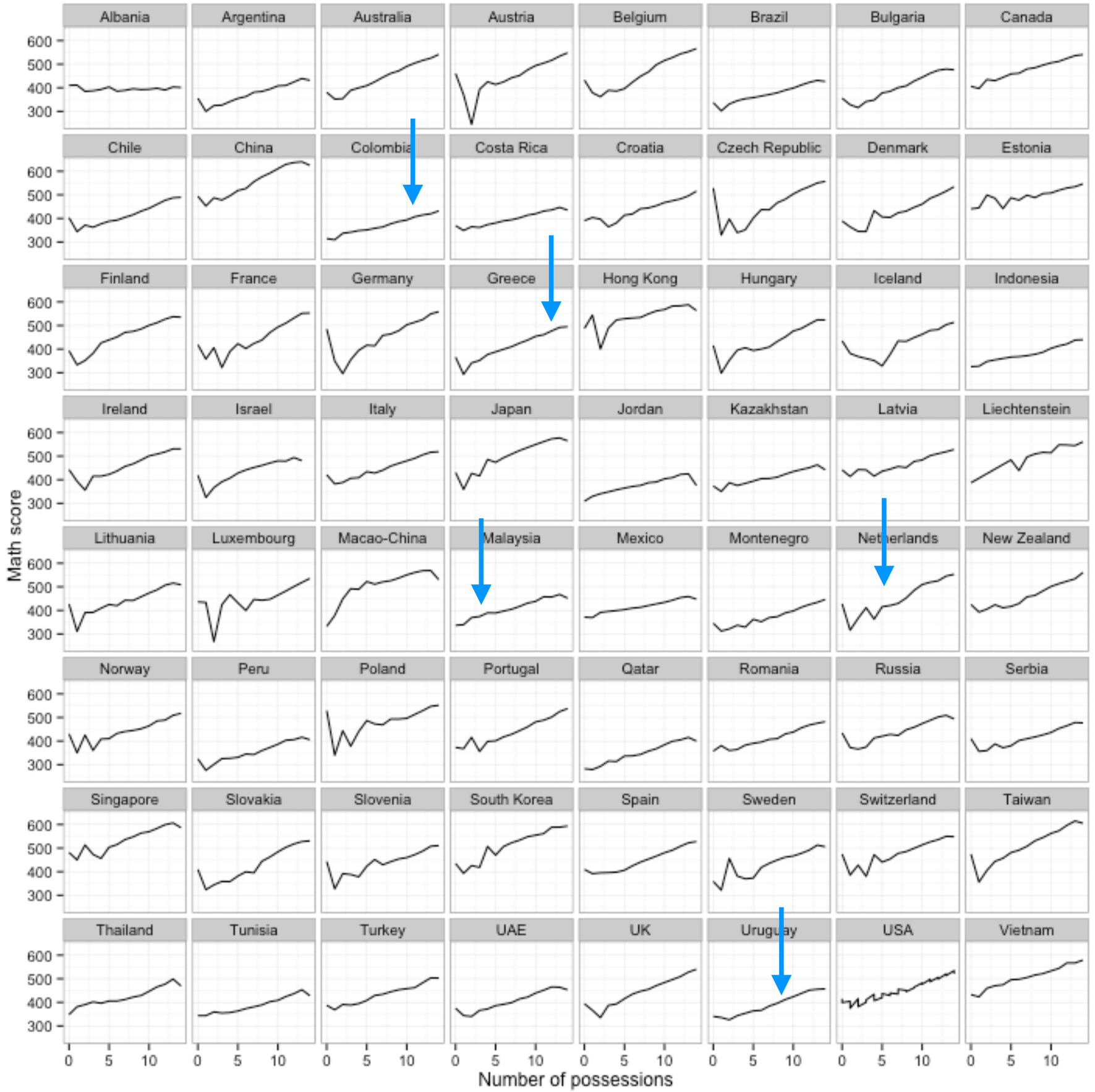

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Number of possessions

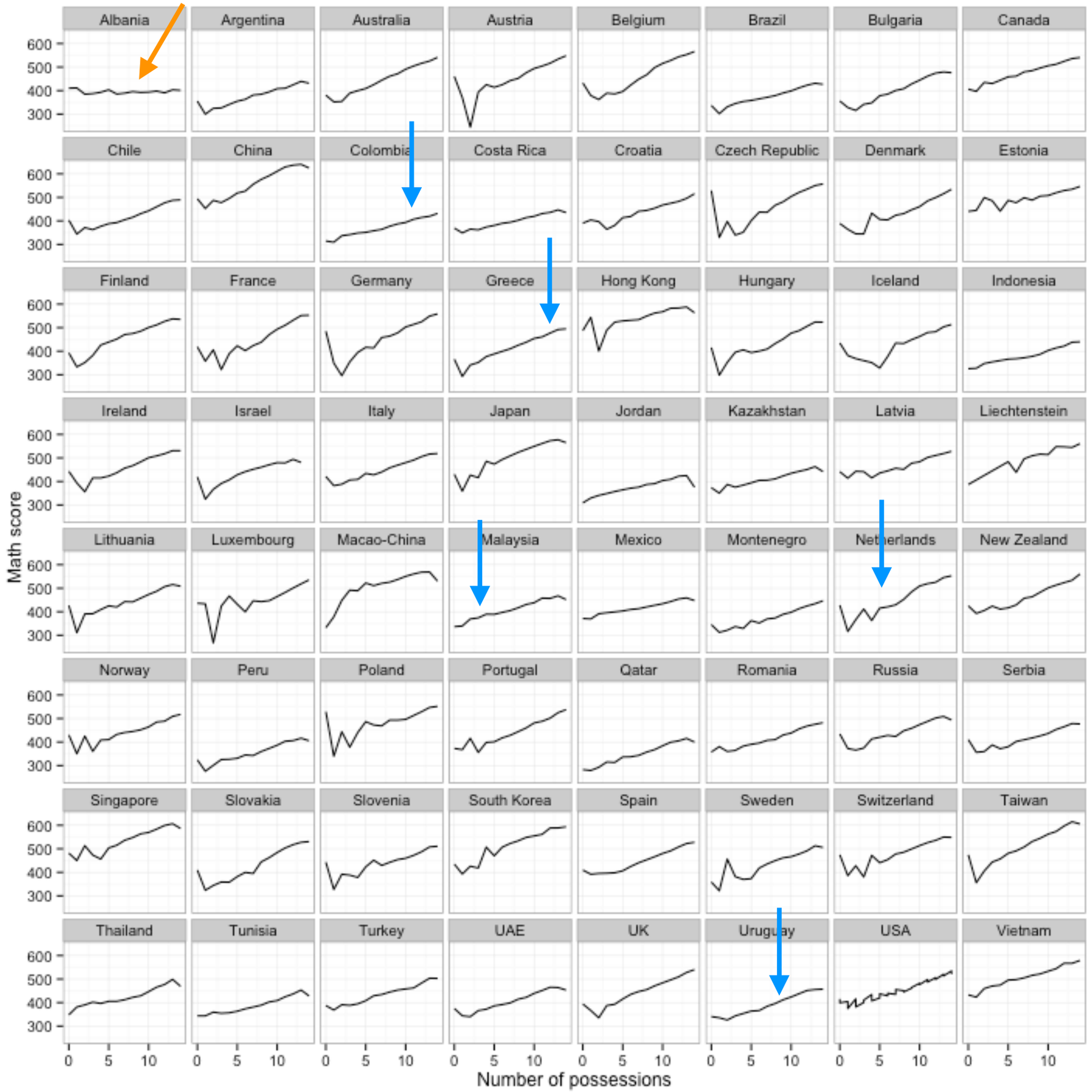

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Number of possessions

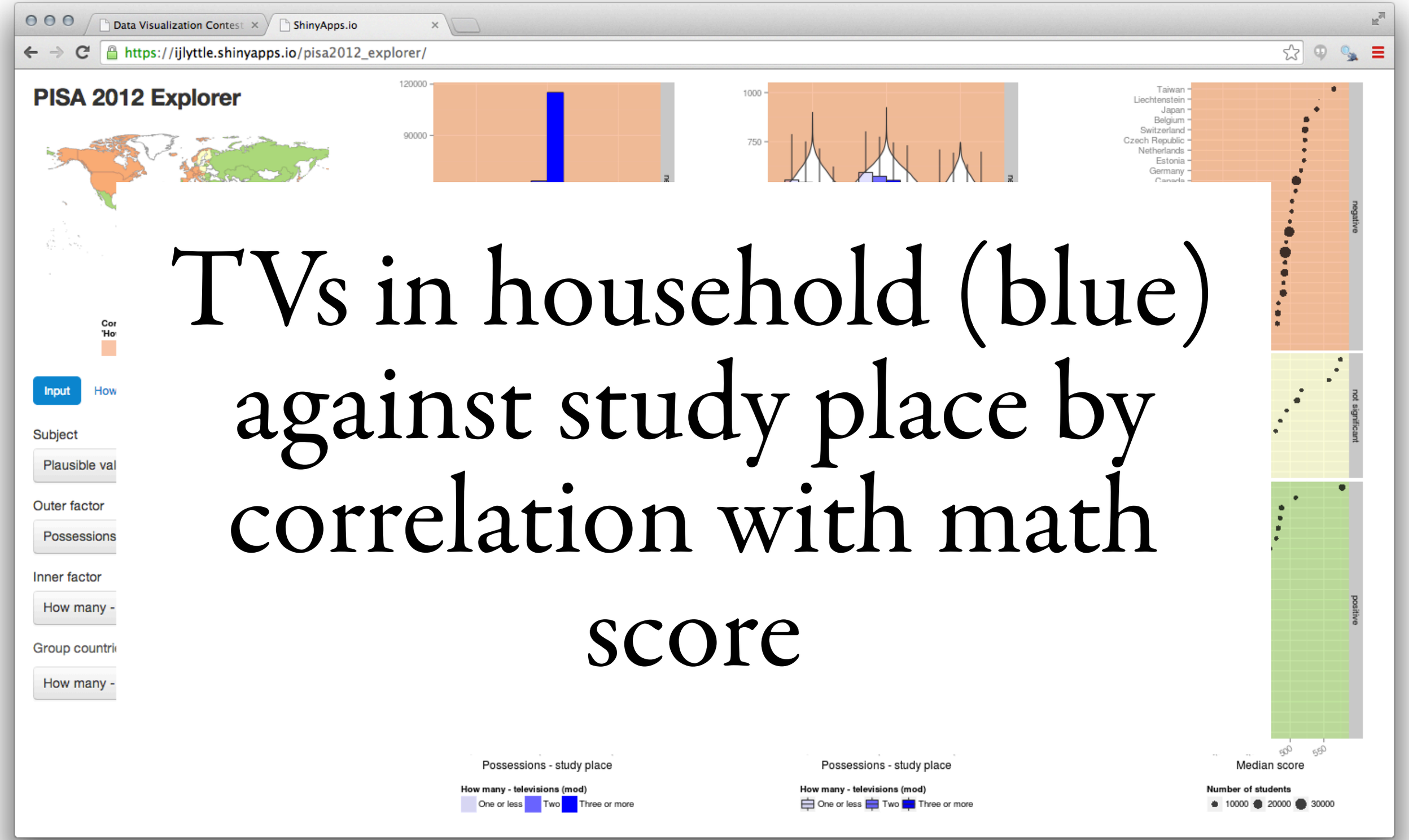

Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Number of possessions


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Math score
Number of possessions


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Math score
Number of possessions


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Math score
Number of possessions


Albania
Argentina
Australia
Austria
Belgium
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hong Kong
Hungary
Iceland
Indonesia
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Latvia
Liechtenstein
Lithuania
Luxembourg
Macao-China
Malaysia
Mexico
Montenegro
Netherlands
New Zealand
Norway
Peru
Poland
Portugal
Qatar
Romania
Russia
Serbia
Singapore
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand
Tunisia
Turkey
UAE
UK
Uruguay
USA
Vietnam
600
500
400
300
0
5
10
Math score
Number of possessions



TVs in household (blue) against study place by correlation with math score

http://bit.ly/pisa2012_explorer

Data Visualization Contest | ShinyApps.io

https://ijlyttle.shinyapps.io/pisa2012_explorer/

PISA 2012 Explorer

Correlation between 'Score' and 'How many - televisions (mod)'

negative | not significant | positive

Input | How-to | Example | Details | About

Subject

Plausible value 1 in mathematics

Outer factor

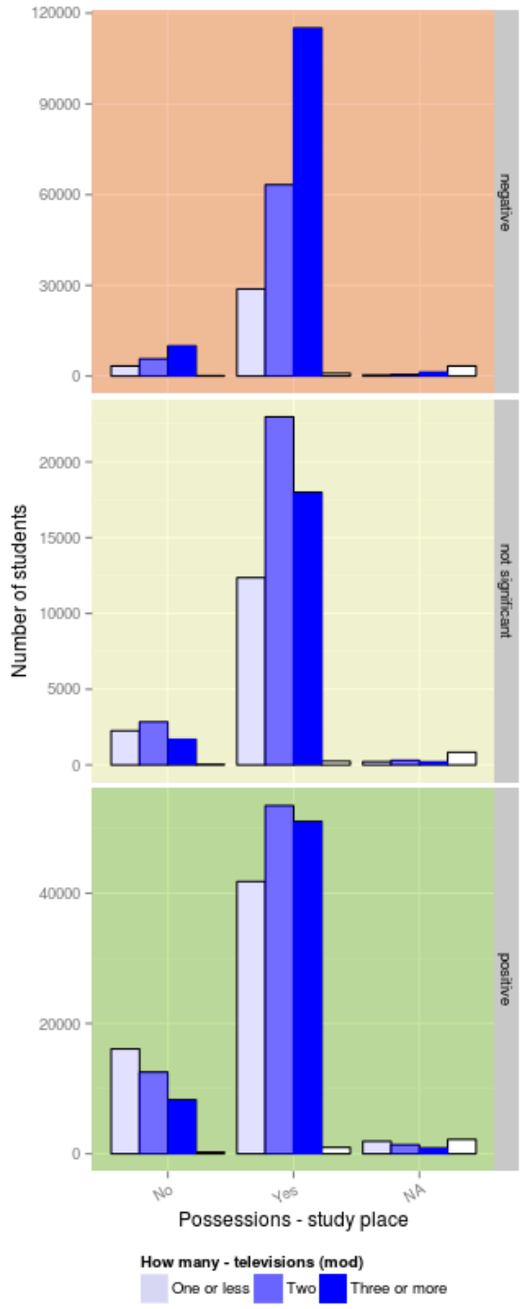
Possessions - study place

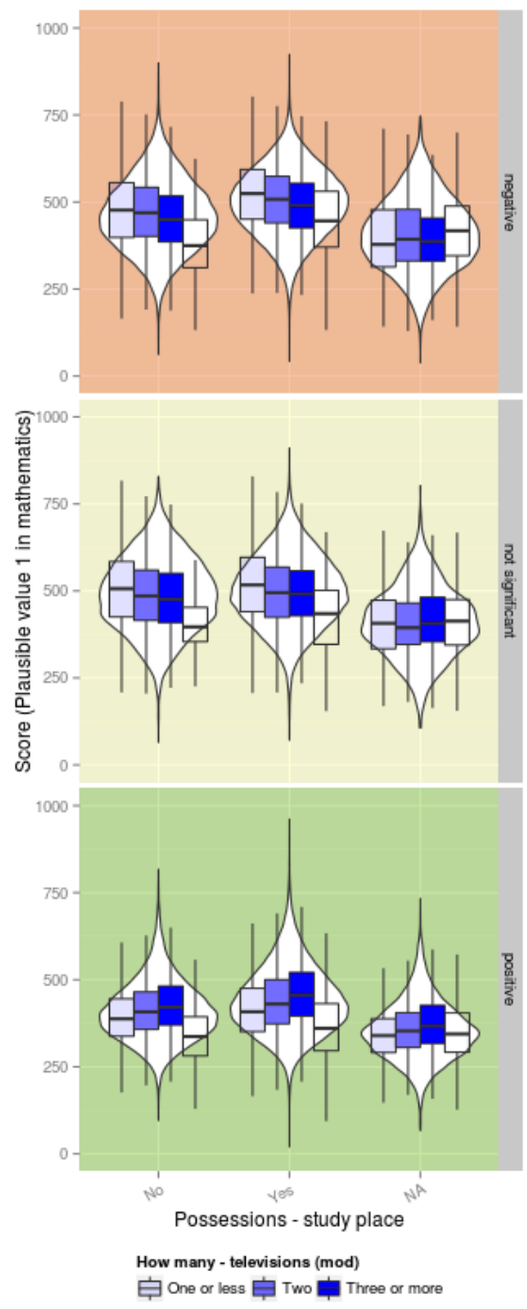
Inner factor

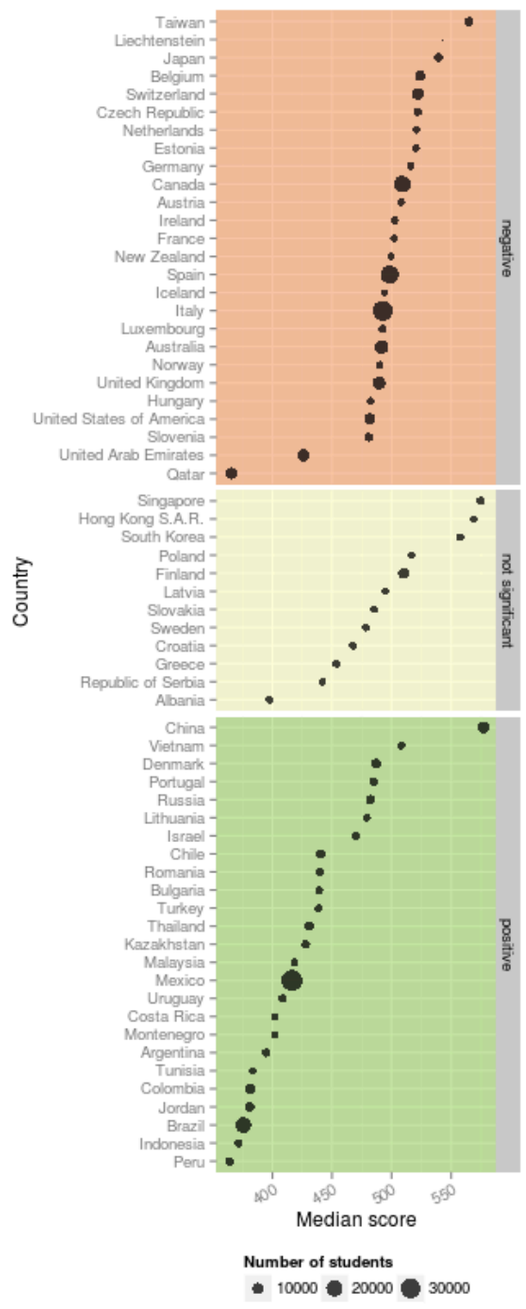
How many - televisions (mod)

Group countries by correlation of **score** and

How many - televisions (mod)







http://bit.ly/pisa2012_explorer

Data Visualization Contest | ShinyApps.io

https://ijlyttle.shinyapps.io/pisa2012_explorer/

PISA 2012 Explorer

Correlation between 'Score' and 'How many - televisions (mod)'

negative | not significant | positive

Input | How-to | Example | Details | About

Subject

Plausible value 1 in mathematics

Outer factor

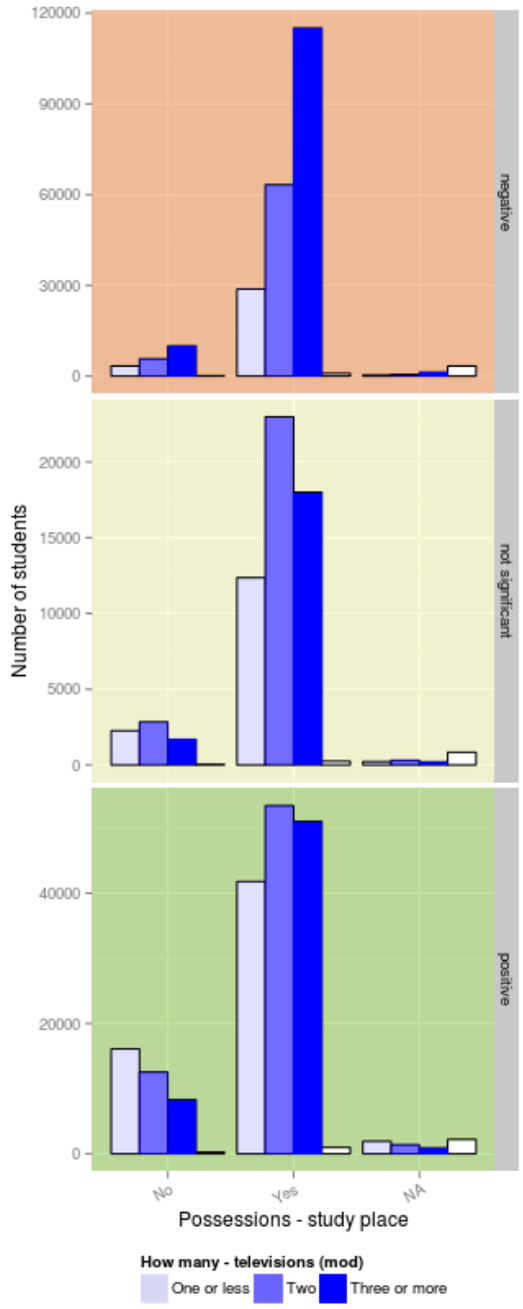
Possessions - study place

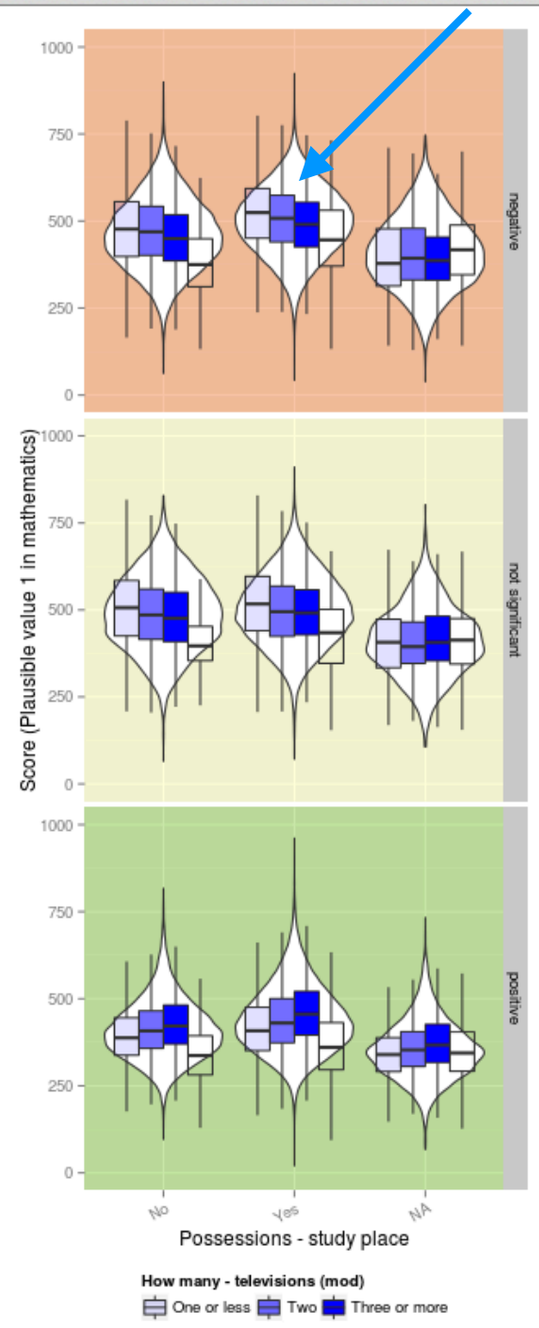
Inner factor

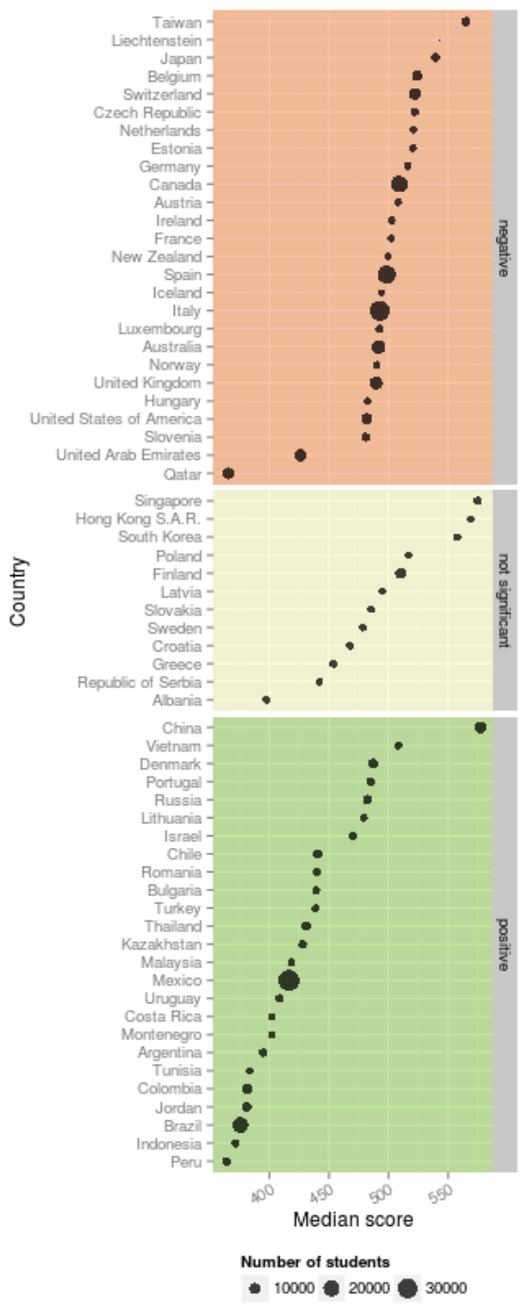
How many - televisions (mod)

Group countries by correlation of **score** and

How many - televisions (mod)







http://bit.ly/pisa2012_explorer

Data Visualization Contest | ShinyApps.io

https://ijlyttle.shinyapps.io/pisa2012_explorer/

PISA 2012 Explorer

Correlation between 'Score' and 'How many - televisions (mod)'

negative | not significant | positive

Input | How-to | Example | Details | About

Subject

Plausible value 1 in mathematics

Outer factor

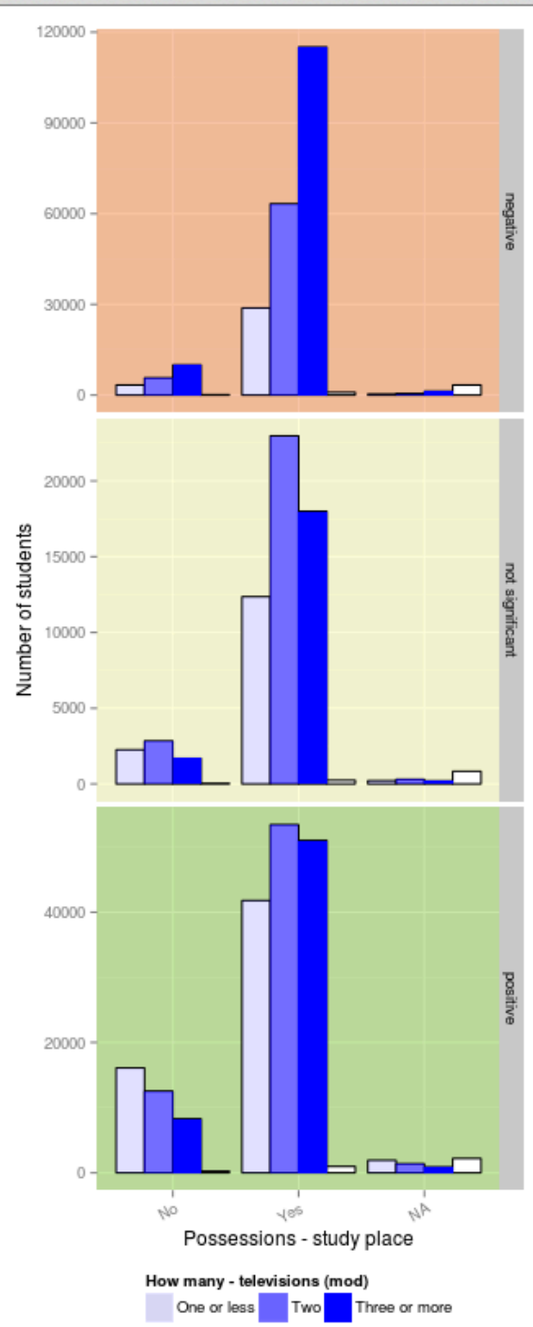
Possessions - study place

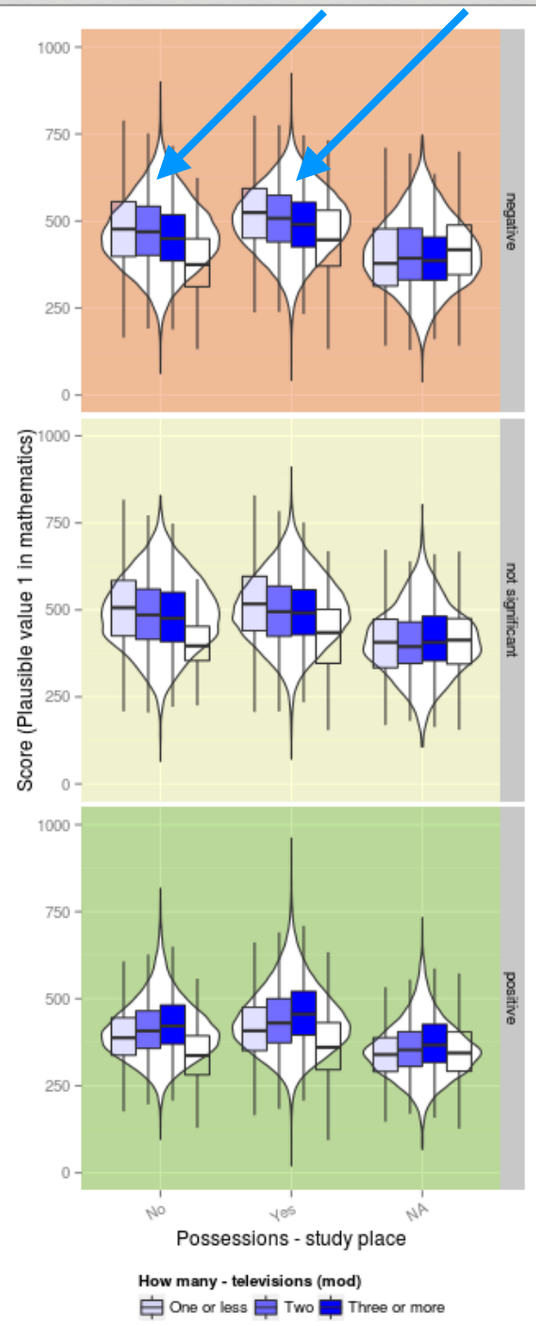
Inner factor

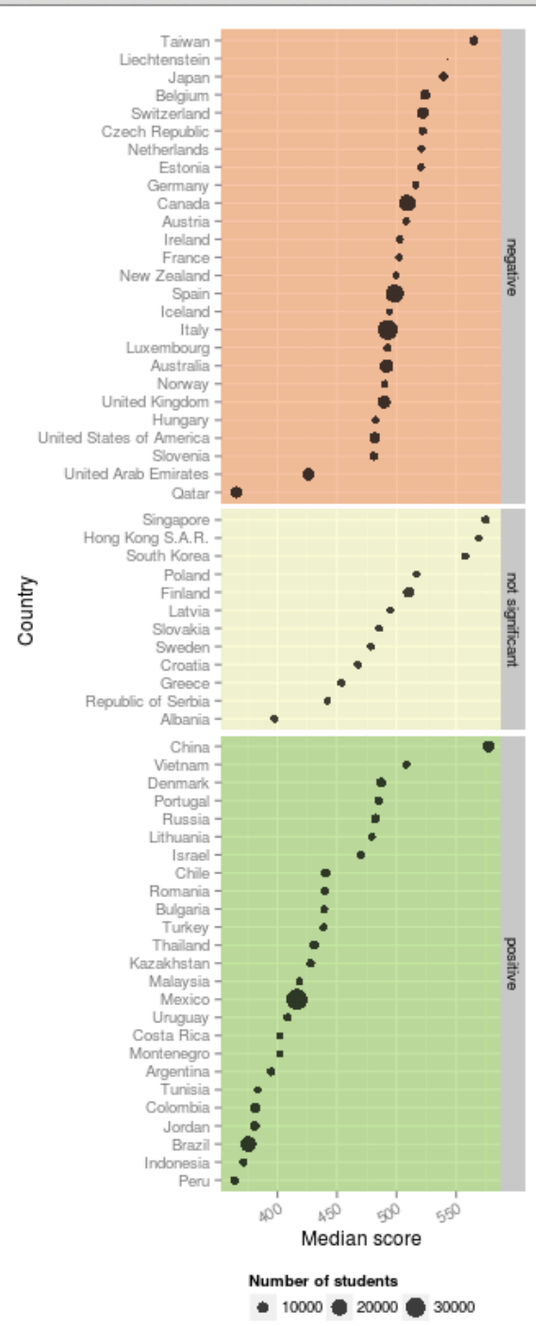
How many - televisions (mod)

Group countries by correlation of **score** and

How many - televisions (mod)







http://bit.ly/pisa2012_explorer

Data Visualization Contest | ShinyApps.io

https://ijlyttle.shinyapps.io/pisa2012_explorer/

PISA 2012 Explorer

Correlation between 'Score' and 'How many - televisions (mod)'

negative | not significant | positive

Input | How-to | Example | Details | About

Subject

Plausible value 1 in mathematics

Outer factor

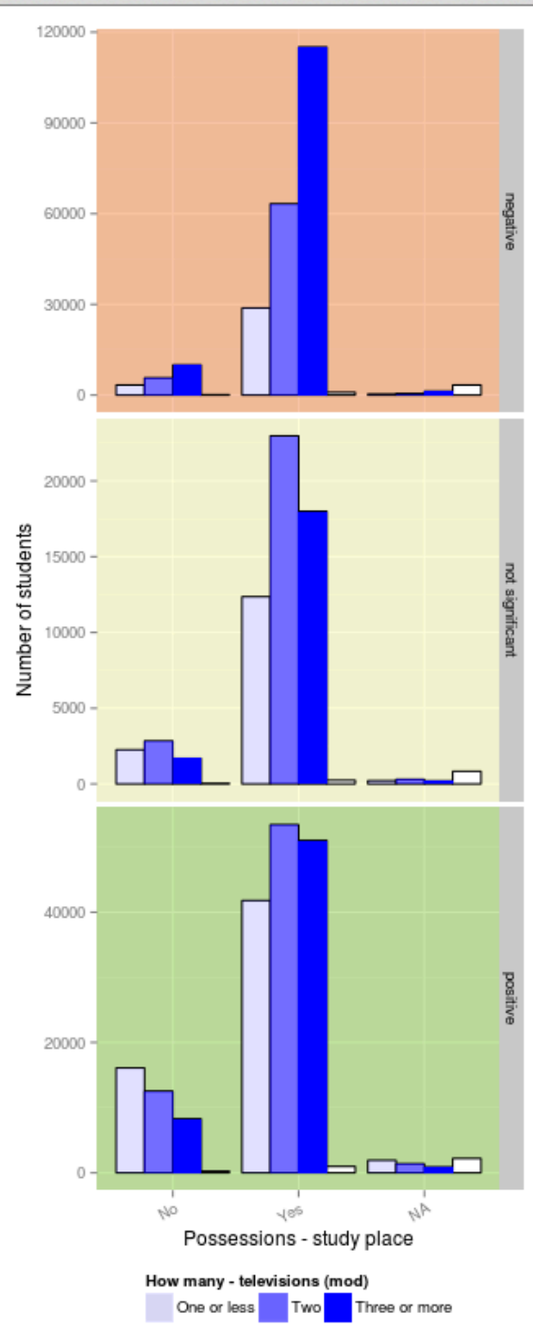
Possessions - study place

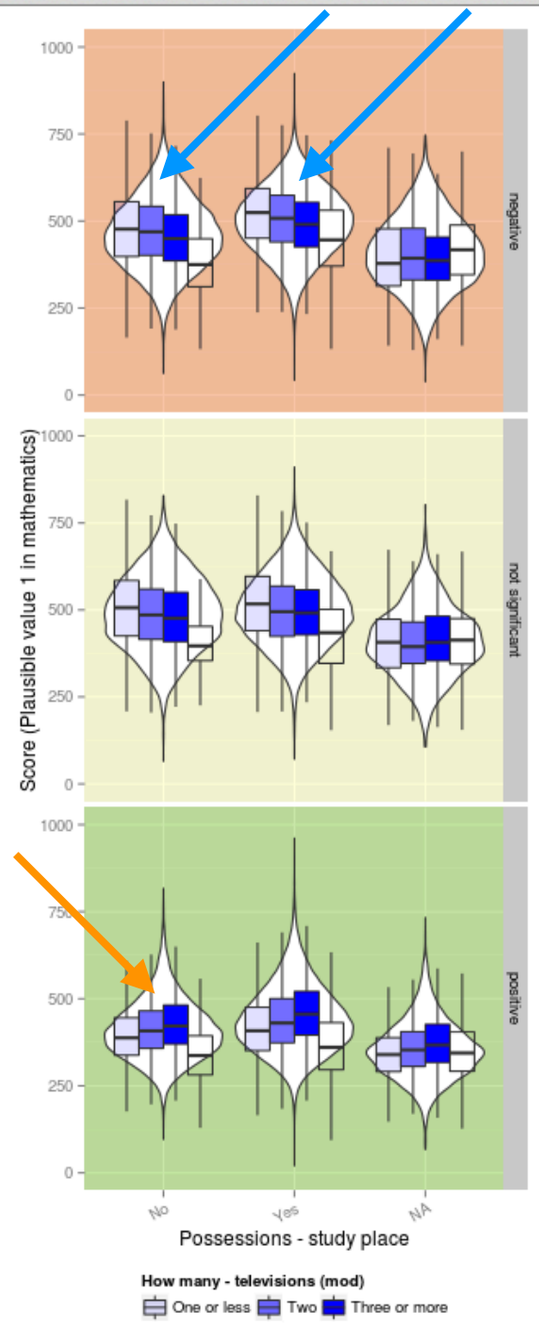
Inner factor

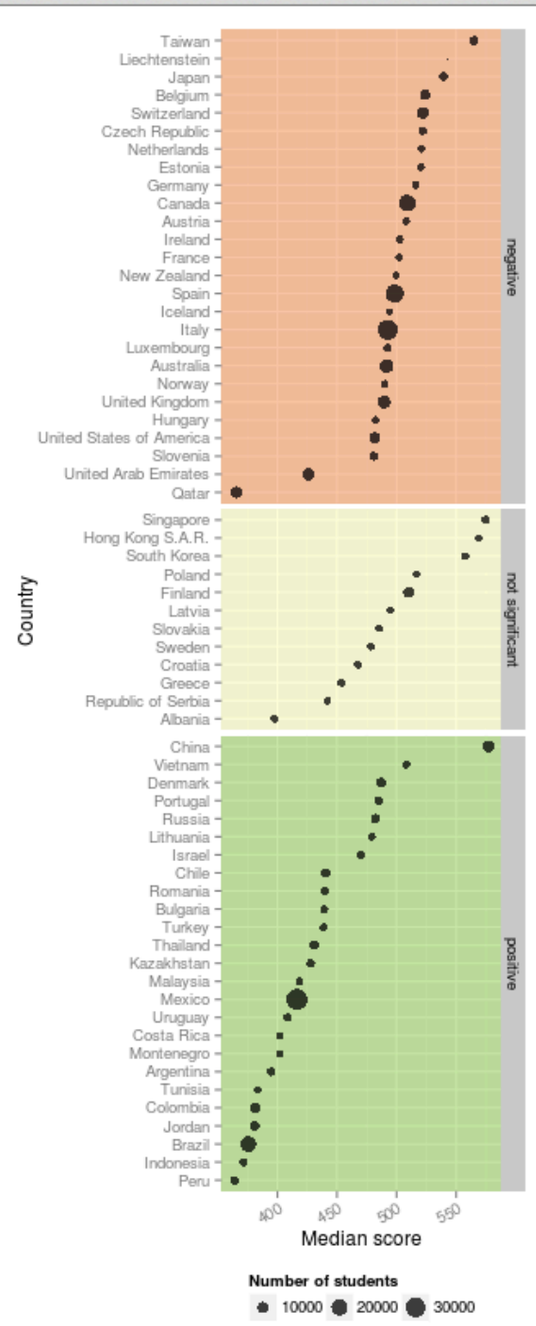
How many - televisions (mod)

Group countries by correlation of **score** and

How many - televisions (mod)







http://bit.ly/pisa2012_explorer

Data Visualization Contest | ShinyApps.io

https://ijlyttle.shinyapps.io/pisa2012_explorer/

PISA 2012 Explorer

Correlation between 'Score' and 'How many - televisions (mod)'

negative | not significant | positive

Input | How-to | Example | Details | About

Subject

Plausible value 1 in mathematics

Outer factor

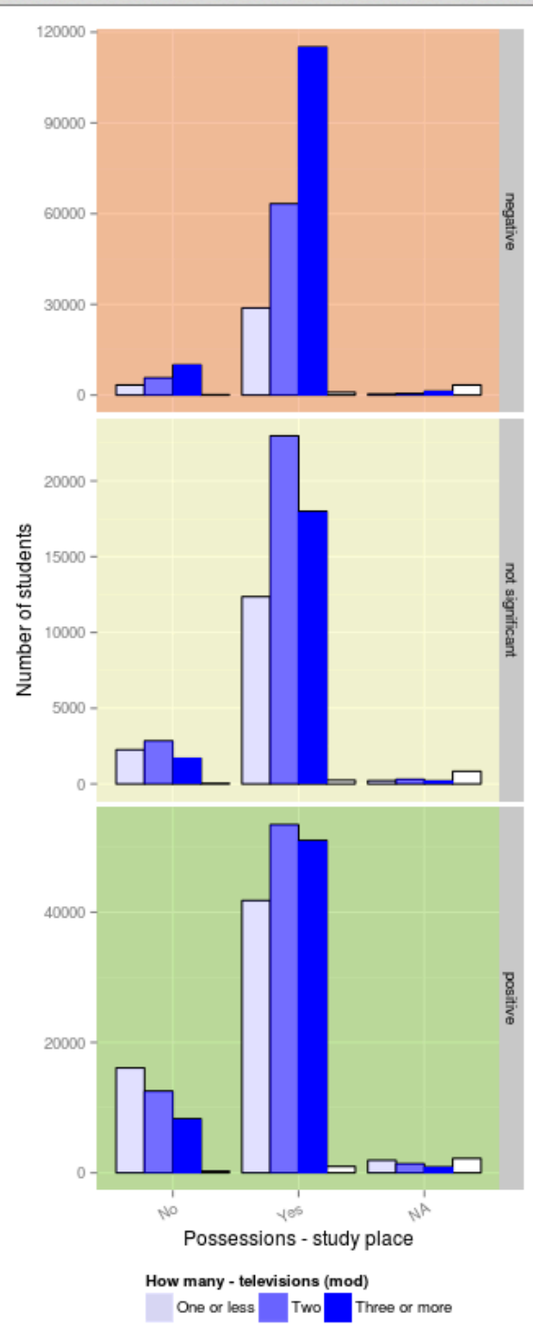
Possessions - study place

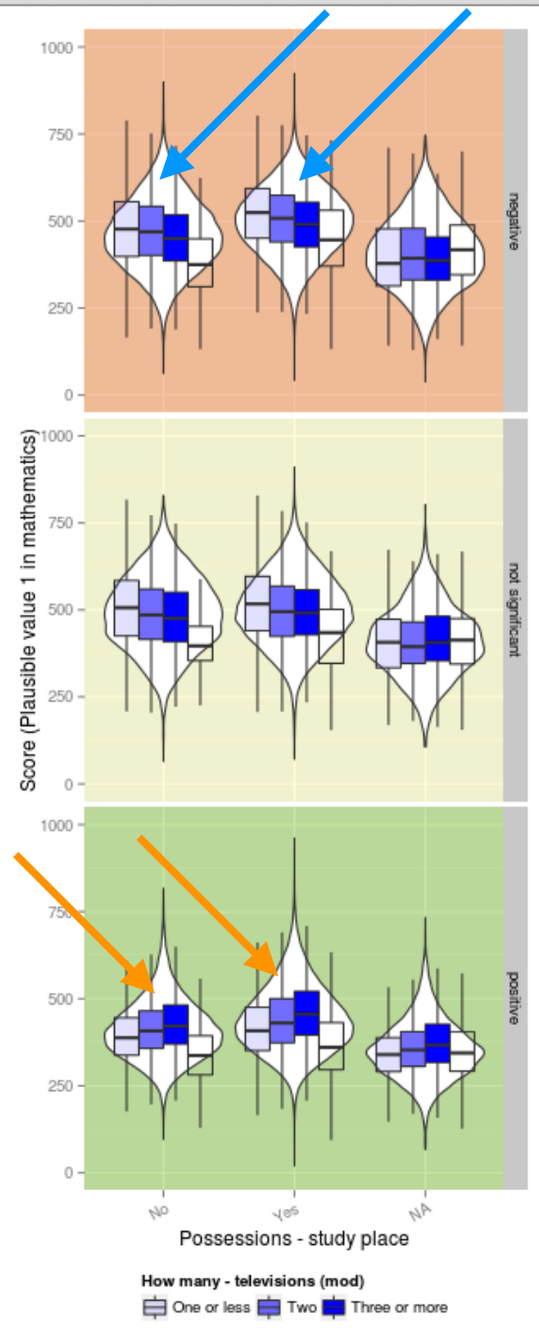
Inner factor

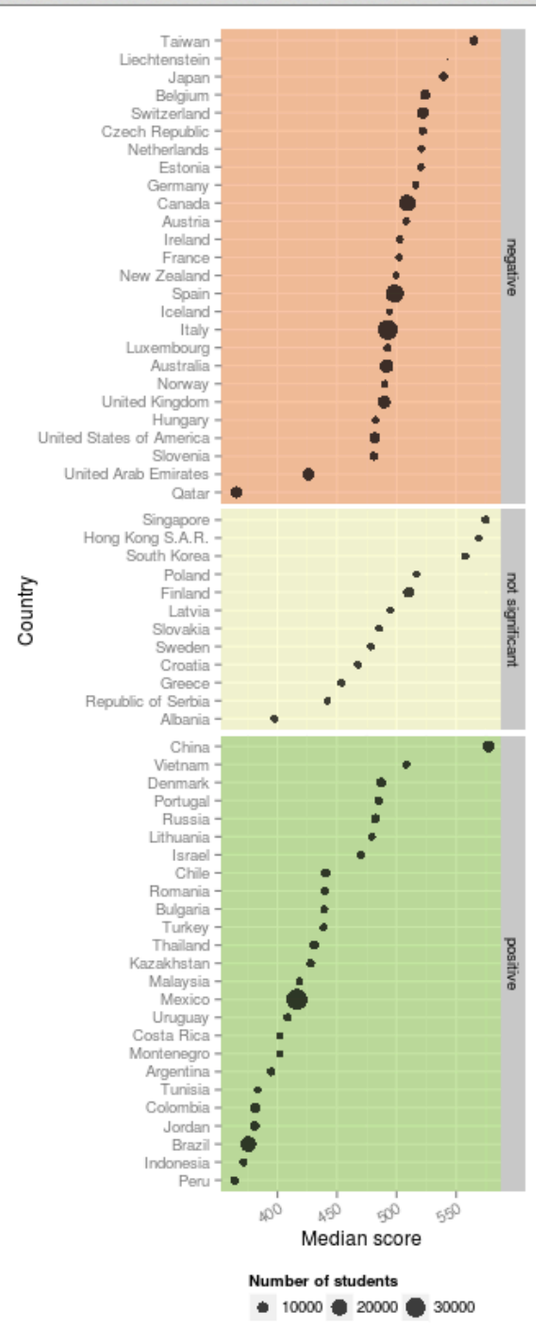
How many - televisions (mod)

Group countries by correlation of **score** and

How many - televisions (mod)





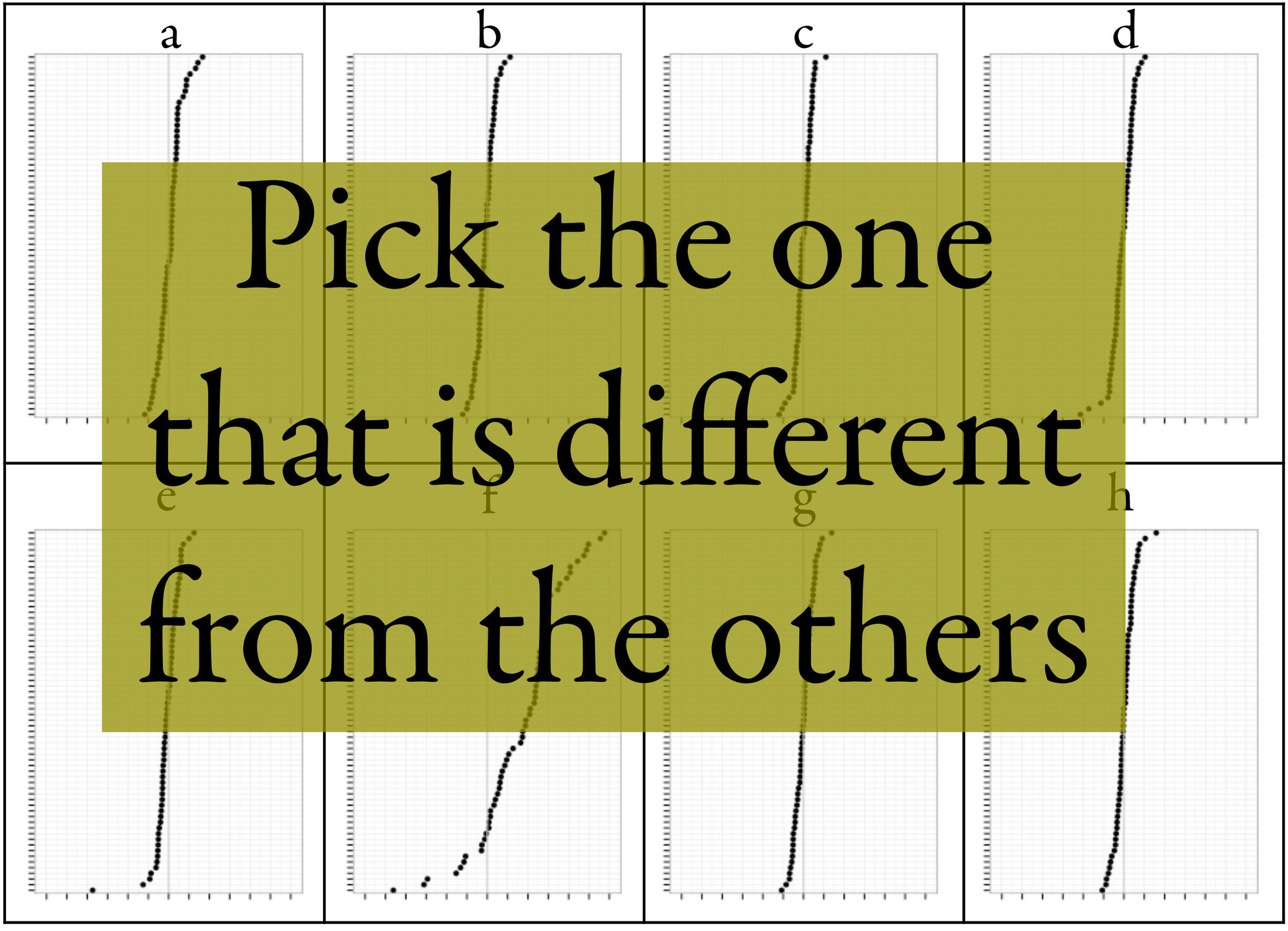


http://bit.ly/pisa2012_explorer

EDA & Inference

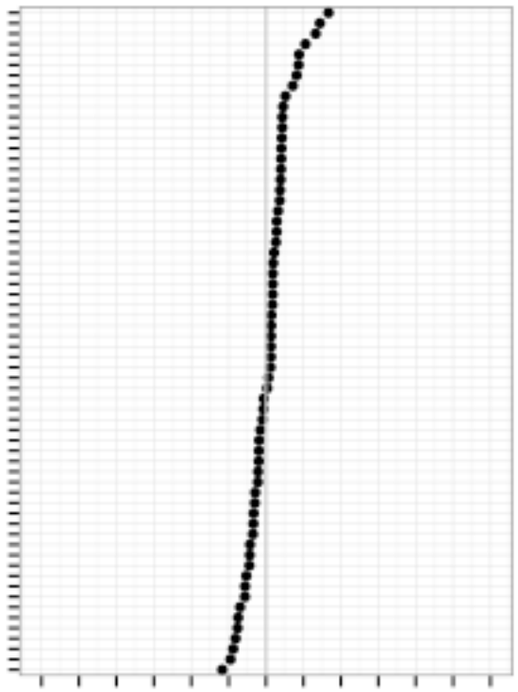
- I have just reported a huge number of conclusions based purely on exploratory methods.
- Do you believe them?
- Exploratory and inferential ARE NOT mutually exclusive.
- I could have showed you the exploratory statistics in the context of NOTHING HAPPENING. From this we could have obtained statistical significance or lack of it.

Here is the math gap exploration placed in the CONTEXT of there being NO MATH GAP ...

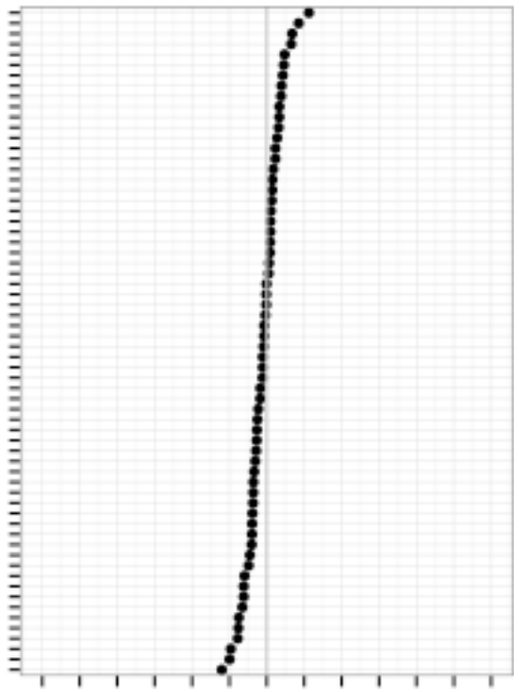


Pick the one that is different from the others

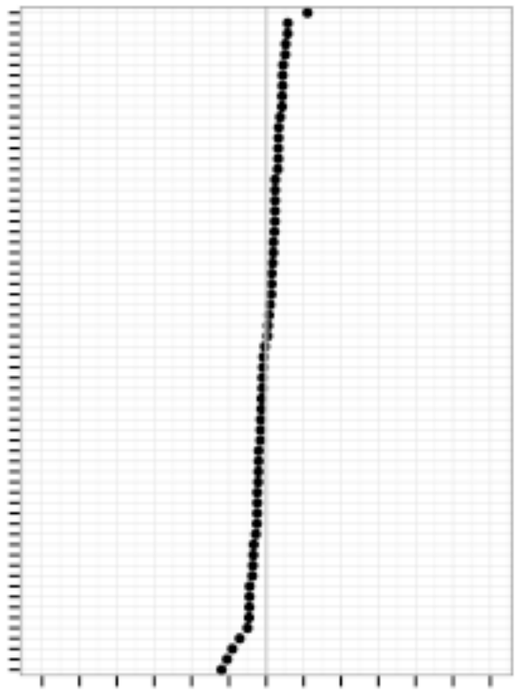
a



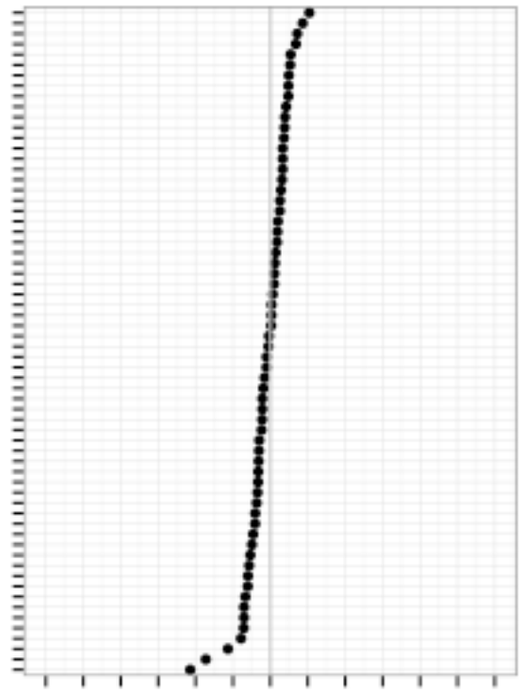
b



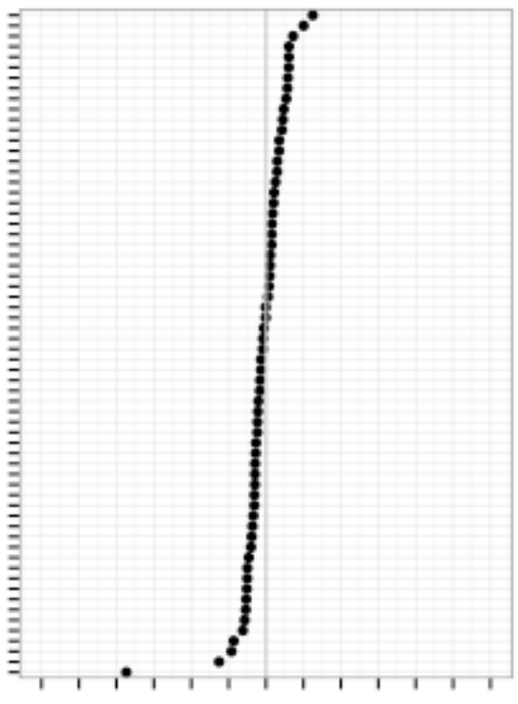
c



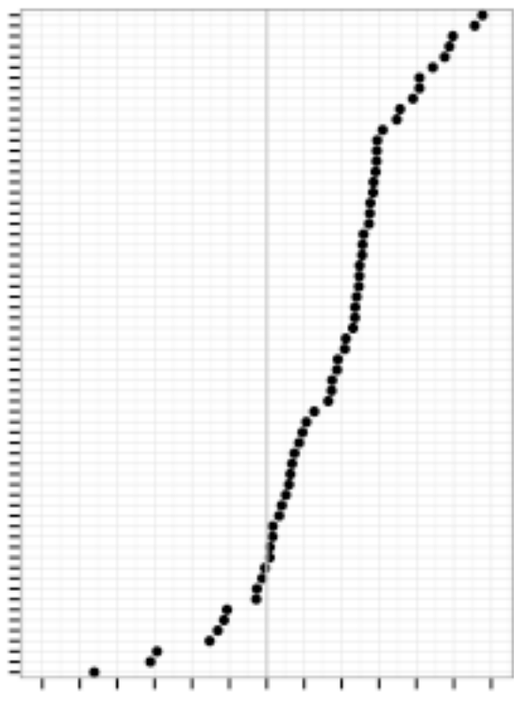
d



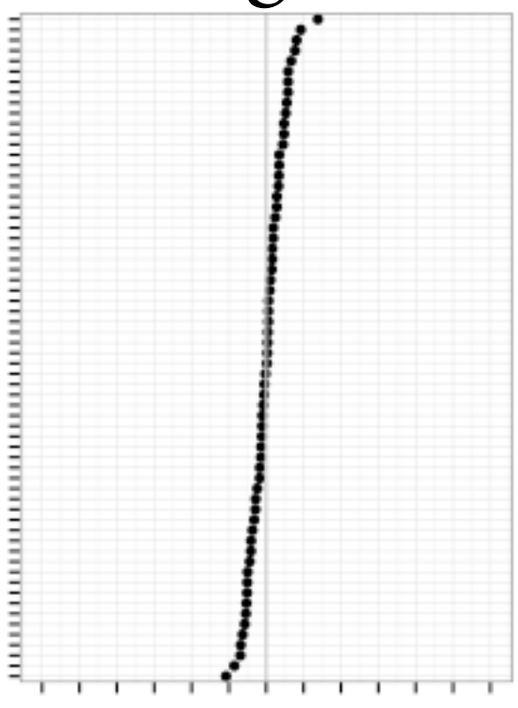
e



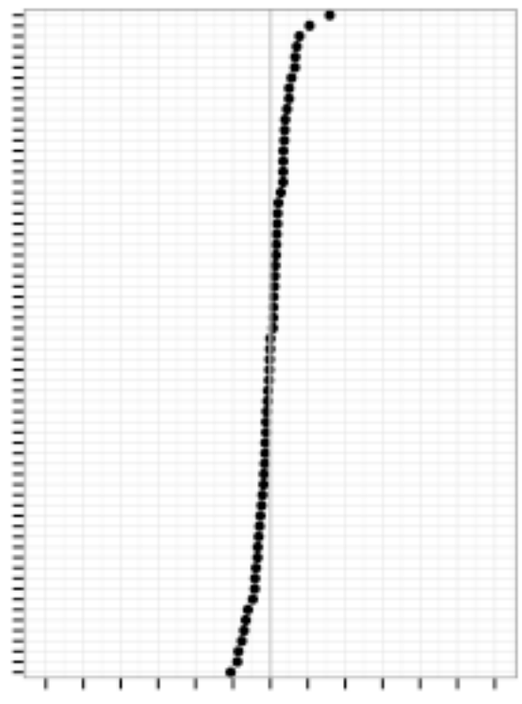
f

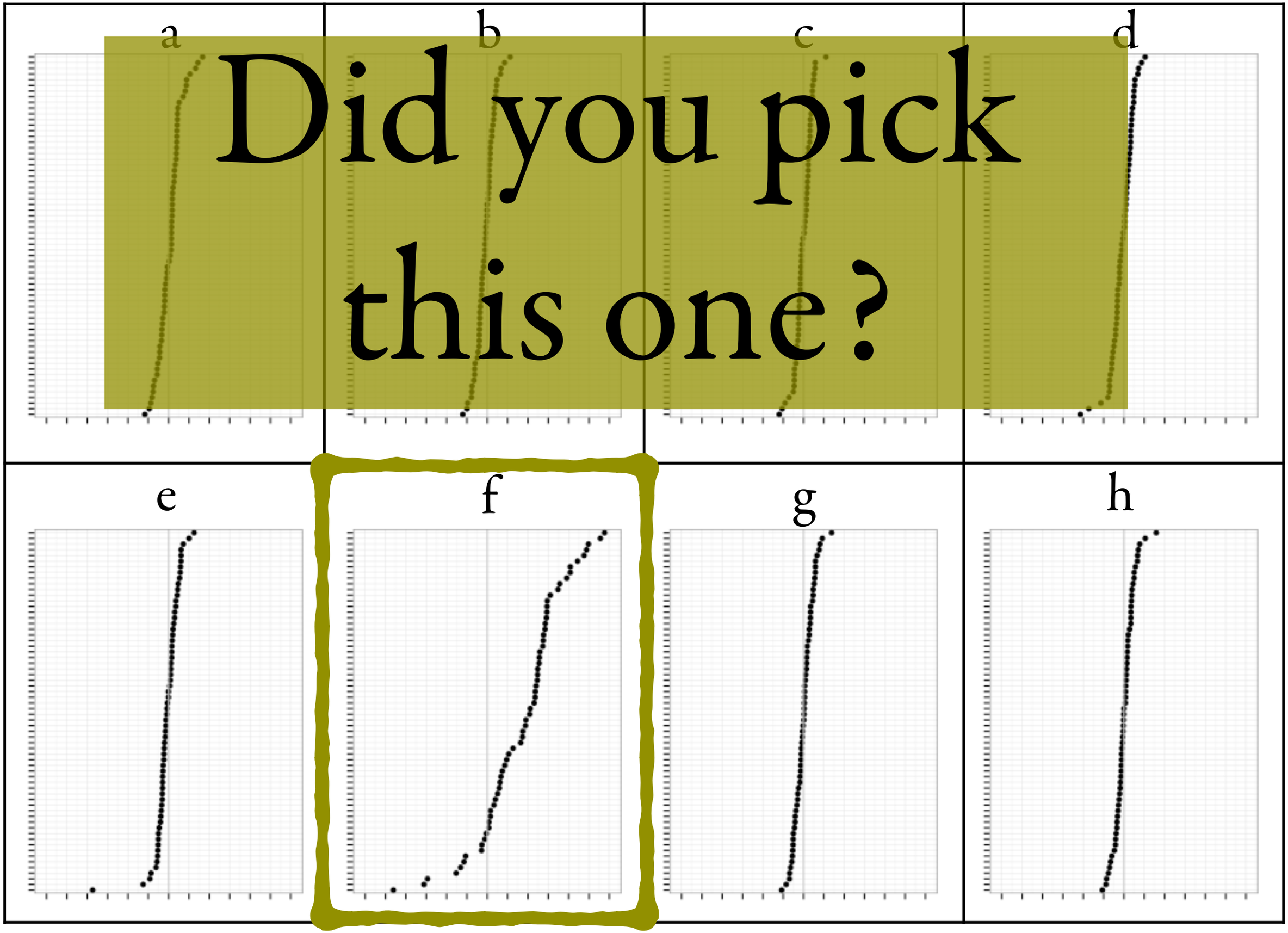


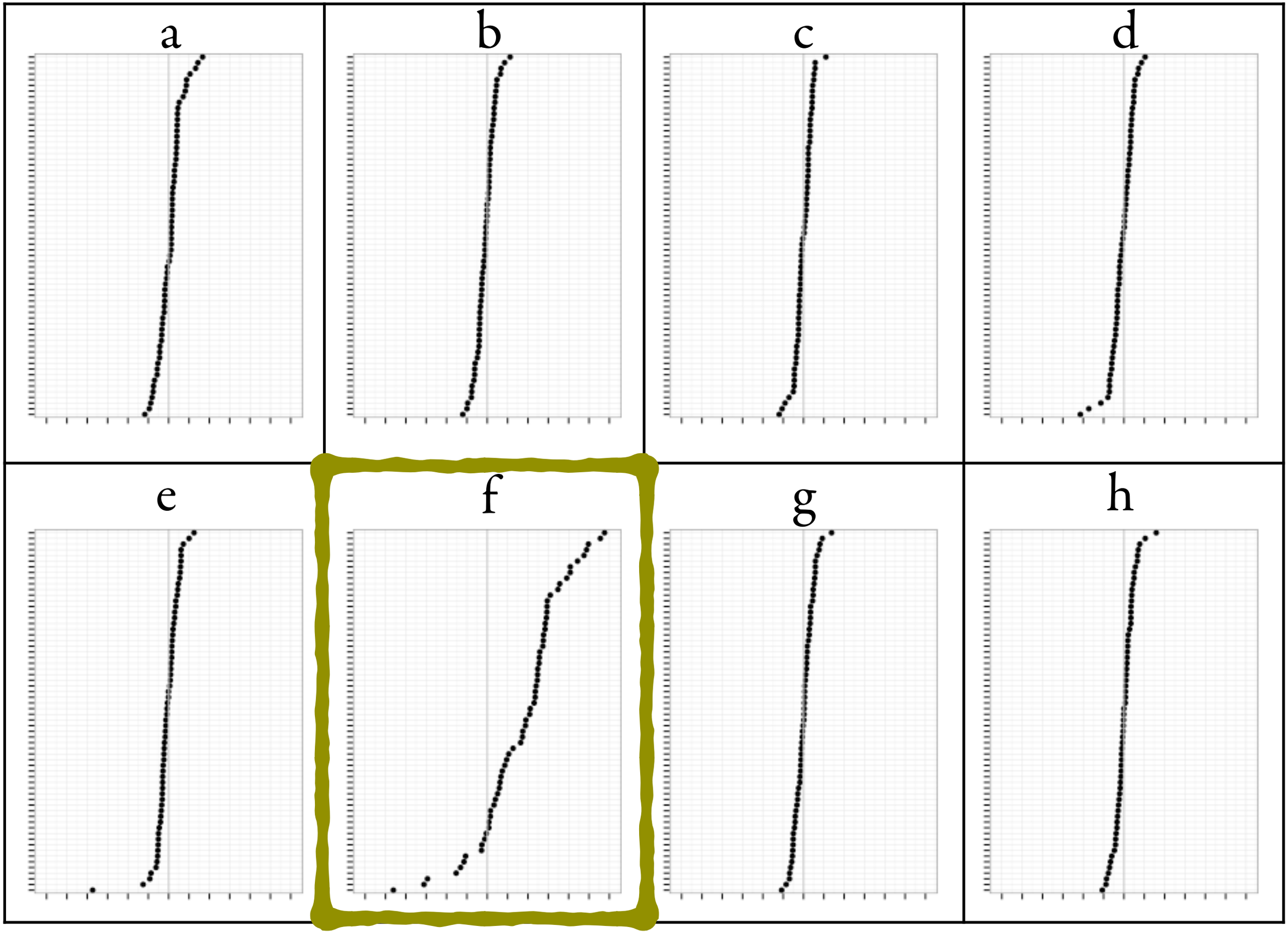
g



h




a
b
c
d
Did you pick this one?
e
f
g
h


a
b
c
d
e
f
g
h

EDA & Inference

If the plot that is picked is the plot of the real data, this is statistical significance, and a *p*-value can be placed on the discovery.

- Buja et al (2009) RSPT A (econ eg)
- Wickham et al (2010) InfoVis/TVCG
- Hofmann et al (2012) InfoVis/TVCG
- Majumder et al (2013) JASA
- Roy Chowdhury et al (2014) Comput. Stat.
- Zhao et al (2014) IJITAS

Nulls by permutation

- Hold country fixed (subset by country)
- Permute the gender labels, so that the math scores are randomly assigned to a boy or girl
- Recalculate the difference between the means
- Plot the mean difference by country again

Mapping data to plots

- Making good plots of data is not straight forward!
- There are more bad data plots than good.
- Learning how to effectively plot data should be in CORE statistics/data science curricula

Reproducibility

- Being able to reproduce the analysis is critically important.
- I used RStudio, with knitr and Rmarkdown to embed analysis with documentation.
- This is what I use in my statistics labs today, so that students learn to document their analysis, and enable others to see what they did to the data, and re-run the analysis themselves.

“Show the data."

What we learned

- Gender gap is not universal in math.
- Gender gap is universal in reading.
- Individual differences trump everything else.
- Parents matter! Mothers more than fathers!
- Socioeconomic status matters.
- Turn the TV off, if you live in a developed country!
- Something's funny about Albania's data.

Tools

- `ggplot2`, by Hadley Wickham and the RStudio team (look for their interactive web graphics ggvis soon!)
- `knitr`, by Yihui Xie and the RStudio team. So good and easy now that first year statistics students can use for their homeworks and labs!
- `shiny`, by RStudio, for the interactive web app.