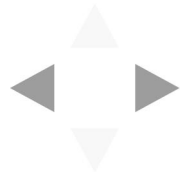
Michael Albers

Udacity Data Analytics Nanodegree course

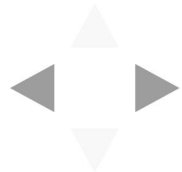
PISA Data

The PISA data set for the data visualizations comprise of mostly student responses to a questionnaires taken by 15-old year old students in 65 countries. This project focuses on PISA data for China, Singapore, and the United Statues. This data set also has the math literacy score for each student along with other quantitative variables such as number of hours studying away from school, number of math classes taken in a week, maximum years of parents education, derived scores for math self concept, math work ethic, and math teacher support.

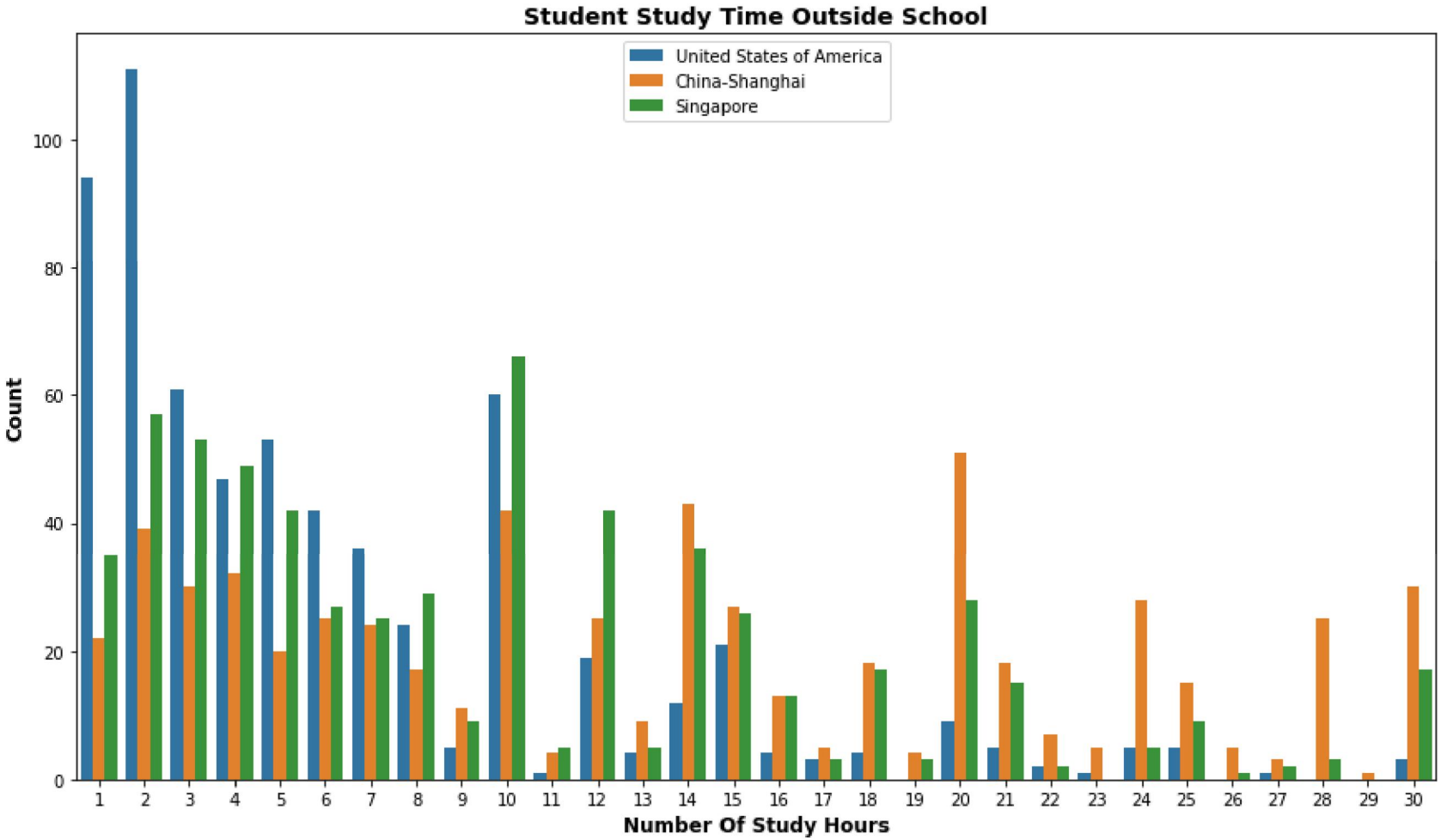
Becuase there are missing values on every record in almost all columns in the data set, a separate dataframe is created for each group of charts and cleaned by removing records with missing data. The features of this data set will be the differences in student attitudes, opinions, and scores among the United States versus both China and Singapore.



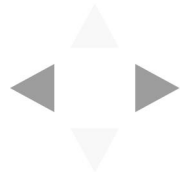
Students Study Time Outside Of School



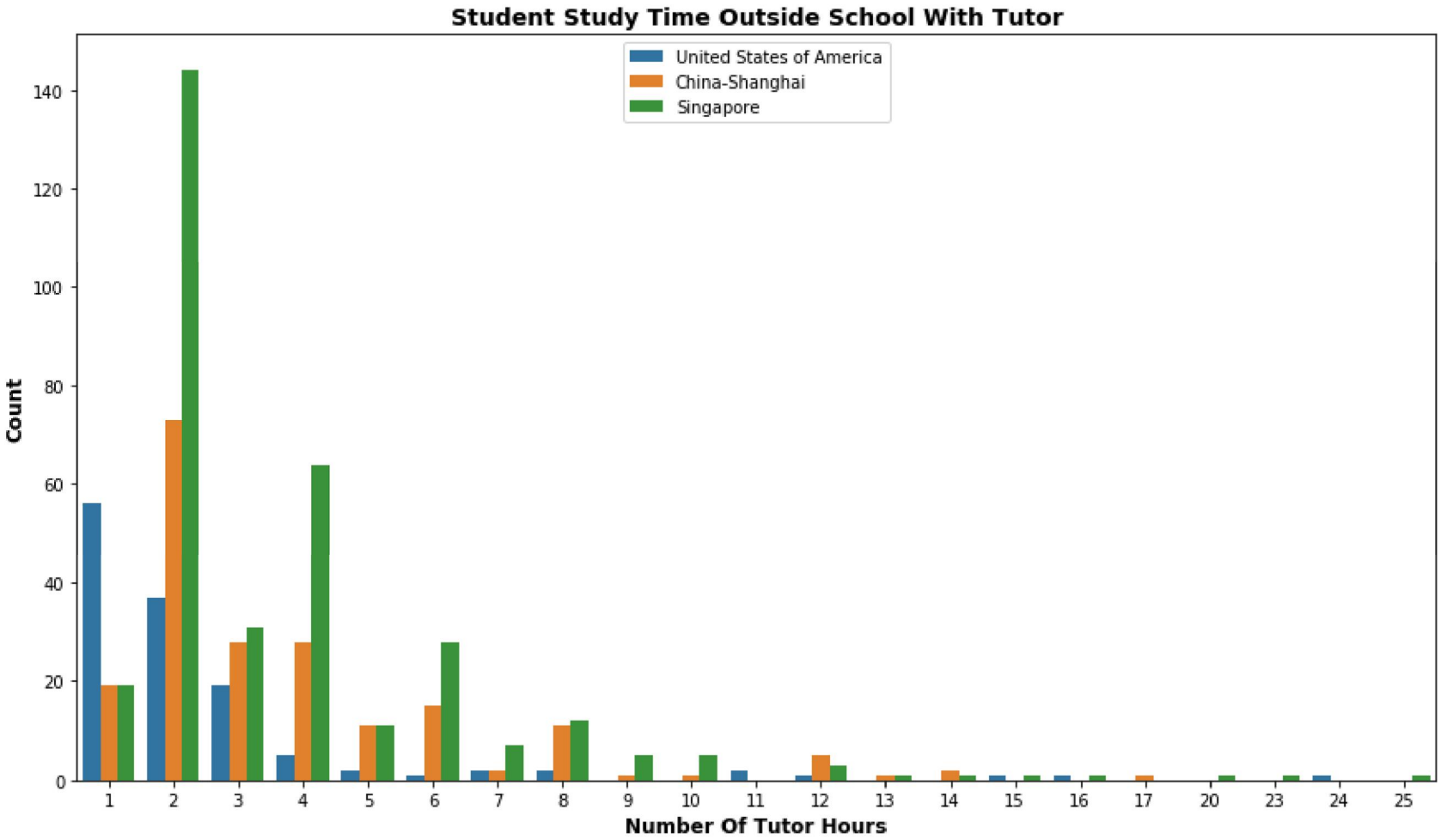
In the United States of America, the number of 15-year old students outnumber China and Singapore for study time outside of school between one and five hours. But as the number of out of school hours studying increases above ten, China and Singapore each have more students studying outside of school than U.S. students. This shows that more China and Singapore students study outside of school for longer than five hours per week than students in the United States. Does this correlate to higher math literacy scores for students from China and Singapore?


Student Study Time Outside School
United States of America
China-Shanghai
Singapore
Count
0
20
40
60
80
100
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
Number Of Study Hours

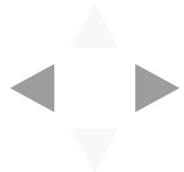
Students Study Time Outside Of School With Tutor



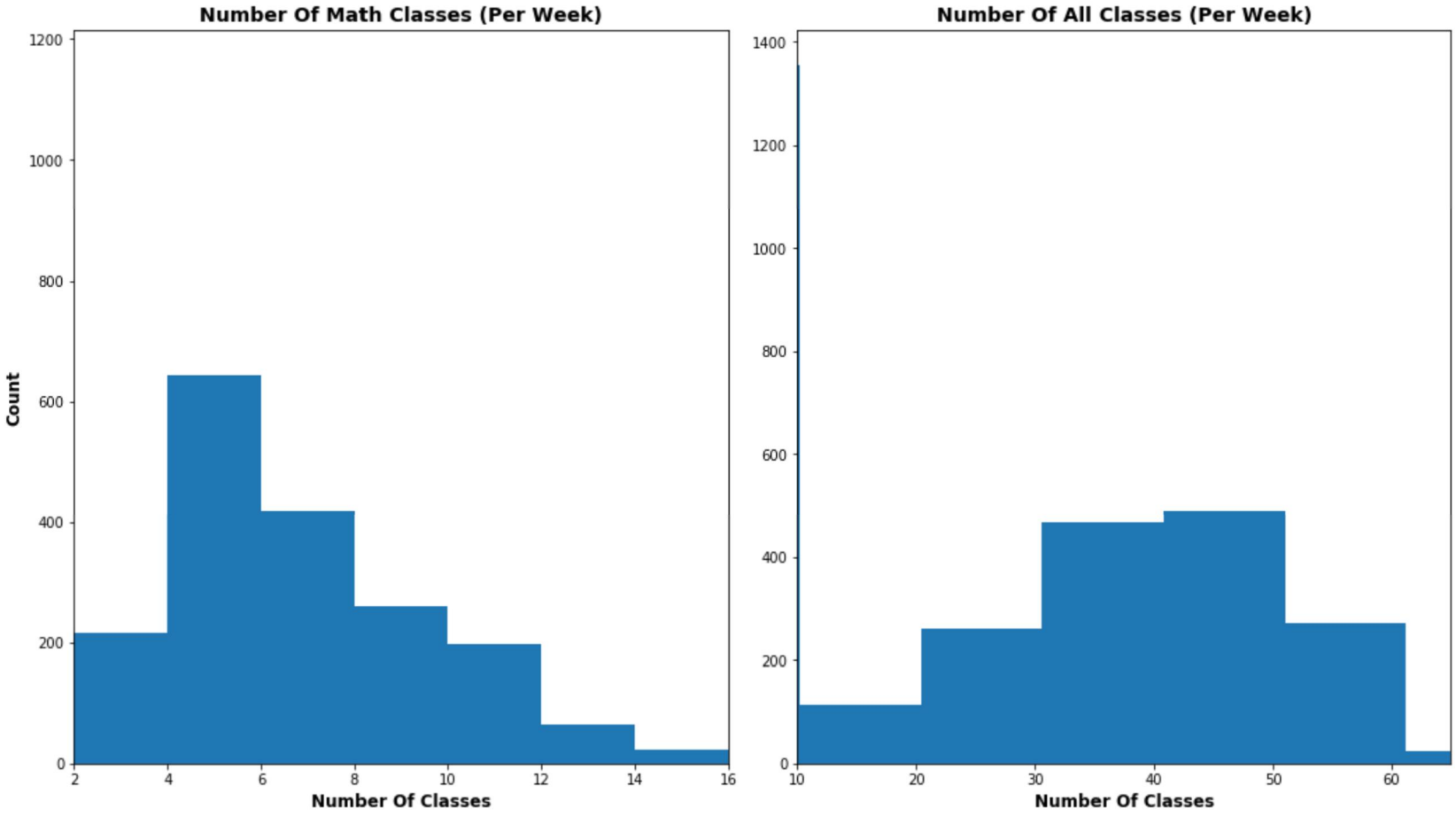
In China and Singapore, considerably more 15-year old students spend time studying outside of school with a tutor. Does this correlate to higher math literacy scores for China and Singapore students?


Student Study Time Outside School With Tutor
United States of America
China-Shanghai
Singapore
Count
0
20
40
60
80
100
120
140
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
20
23
24
25
Number Of Tutor Hours

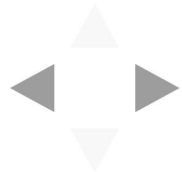
Aggregate Number Of Math Classes Taken Per Week



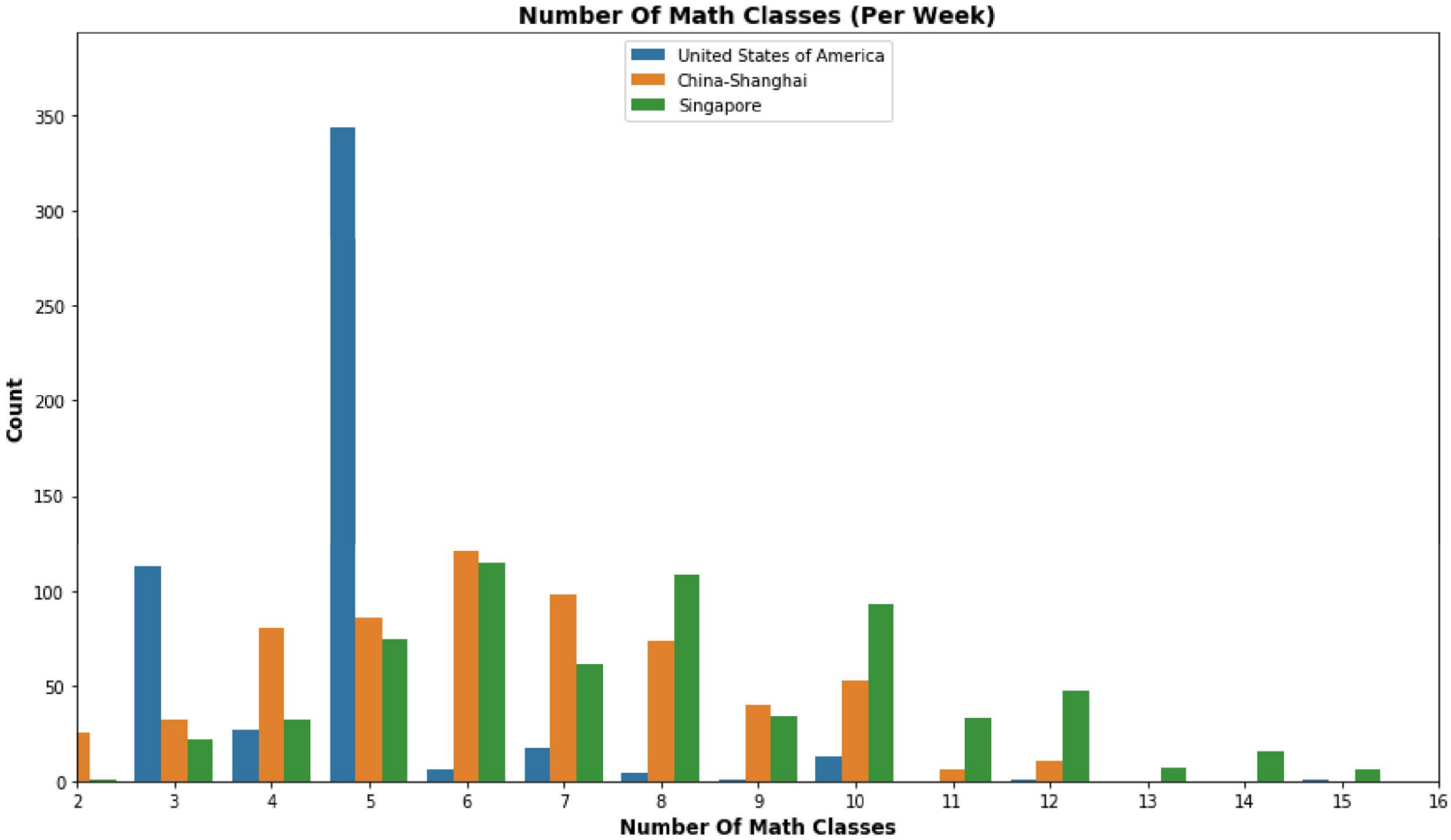
Overall, most 15-year students take about 5 math classes per week. There are many 15-year olds in this dataset who take more than 5 math classes per week. The total number of classes taken per week among all students ranges between 15 to 60. The next plot will show the number of math classes taken per week by 15-year old students for each country.


Number Of Math Classes (Per Week)
Count
Number Of Classes
0
200
400
600
800
1000
1200
2
4
6
8
10
12
14
16
Number Of All Classes (Per Week)
Number Of Classes
0
200
400
600
800
1000
1200
1400
10
20
30
40
50
60

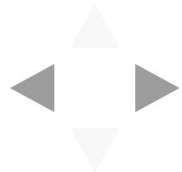
Number Of Math Classes Taken Per Week By Country



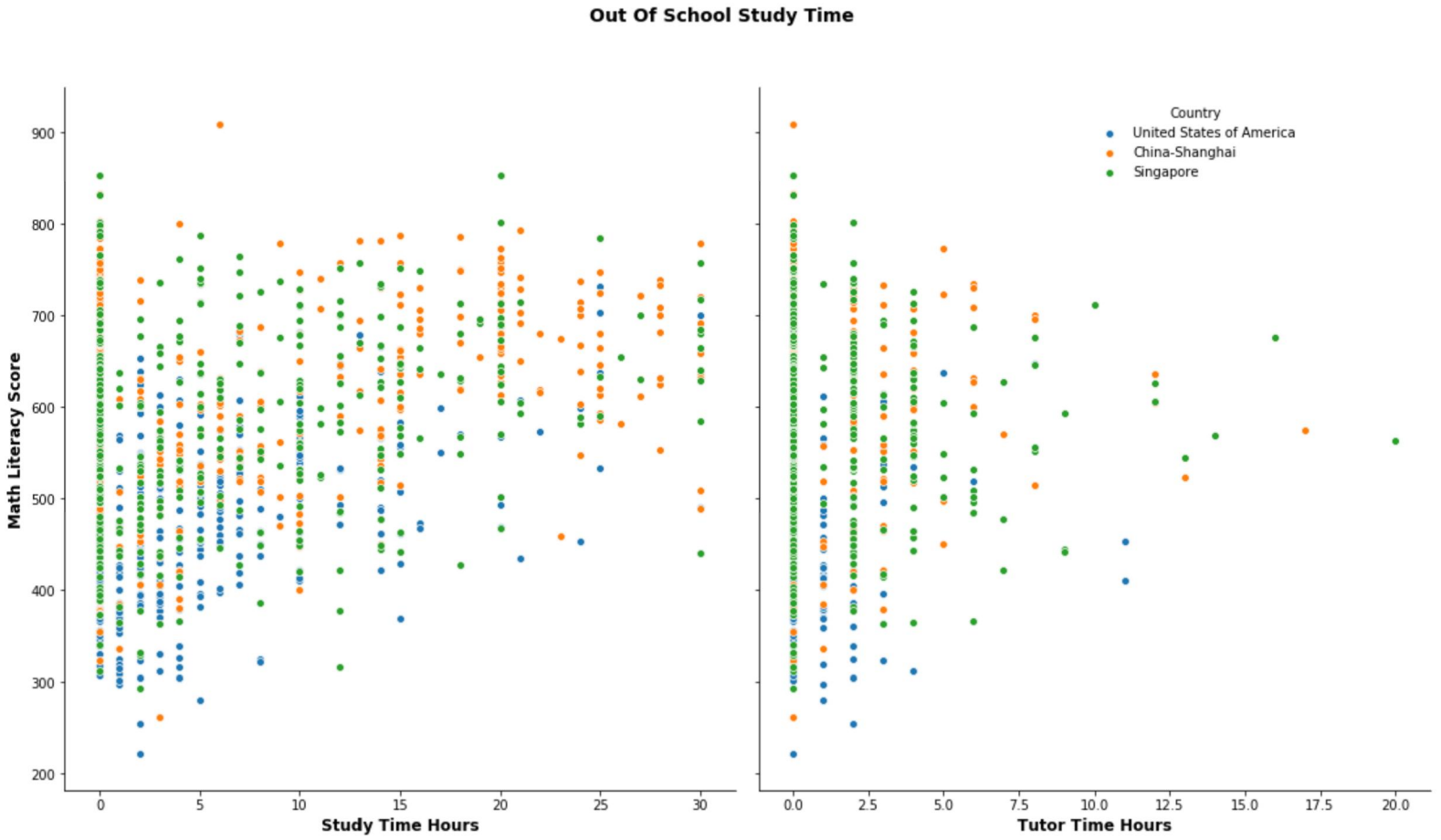
Here we see that most 15-year old students in the United States take between 1 and 5 math classes per week. Most 15-year old students from China and Singapore take between 5 and 10 math classes per week, far more than students in the United States. Does this correlate to higher math literacy scores in China and Singapore?


Number Of Math Classes (Per Week)
United States of America
China-Shanghai
Singapore
Count
0
50
100
150
200
250
300
350
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Number Of Math Classes

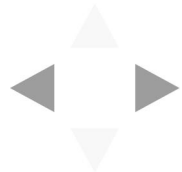
Out Of School Study Time And Math Literacy Score



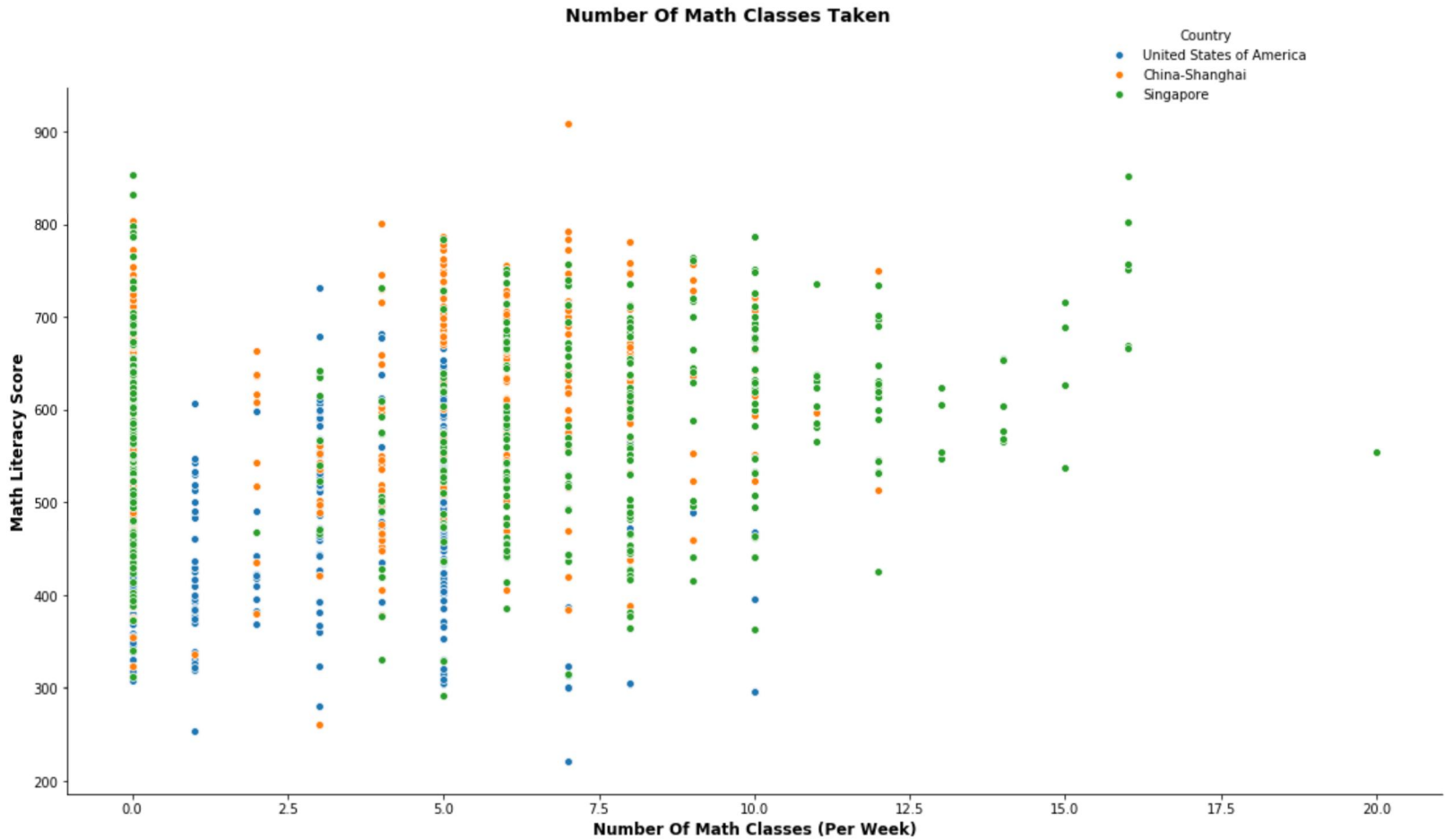
We try to answer the question of whether increased time studying outside of school correlates to higher math literacy scores for 15-year old students. For 15-year old students who spend time outside school studying, between 1 to 30 hours, China and Singapore achieve considerably higher math literacy scores than the United States of America. For out of school time spent with tutors between 2 to 5 hours, students from China and Singapore have the highest math literacy scores.


Out Of School Study Time
Math Literacy Score
Study Time Hours
Tutor Time Hours
Country
United States of America
China-Shanghai
Singapore

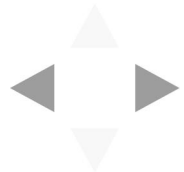
Students Taking Math Classes And Math Literacy Score



Starting at five math classes taken per week, students from China and Singapore have the highest math literacy scores, between 700 and 800. We do not see students from the United States of America taking more than 5 math classes per week. We do not see 15-year old students from the United States reaching math literacy scores between 700 and 800 at any number of math classes taken per week.


Number Of Math Classes Taken
Country
United States of America
China-Shanghai
Singapore
Math Literacy Score
900
800
700
600
500
400
300
200
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Number Of Math Classes (Per Week)

Final outcomes



The features that stand out are:

- Number of hours spent studying outside of school
- Number of hours spent with tutor outside of school
- Number math classes taken per week

In all three of these features, 15-year old students from China and Singapore outperform 15-year old students from the United States in math literacy scores.

In the univariate exploration, it is revealed that the United States 15-year old students spend a lot more time outside of school studying between 1 and 5 hours per week than students from China and Singapore. But what's surprising is that in the multivariate exploration, 15-year old students from China and Singapore score higher on math literacy at all hours spent studying outside of school.

While there are differences in 15-year old students from the United States versus China and Singapore, the data file was comprised of mostly student responses to questionnaires. What is missing in the data file are the parent and school questionnaire responses. Separate questionairres were give to teachers, parents and school employees but these responses were not in the PISA 2012 source data file. This study was therefore skewed on student responses and math scores. Although there are differences drawn out between 15-year old students of the United States versus China and Singapore, it was not convincing enough to explain why a large and wealthy country like the United States ranked 36 out of 65 countries on the math literacy PISA scores.