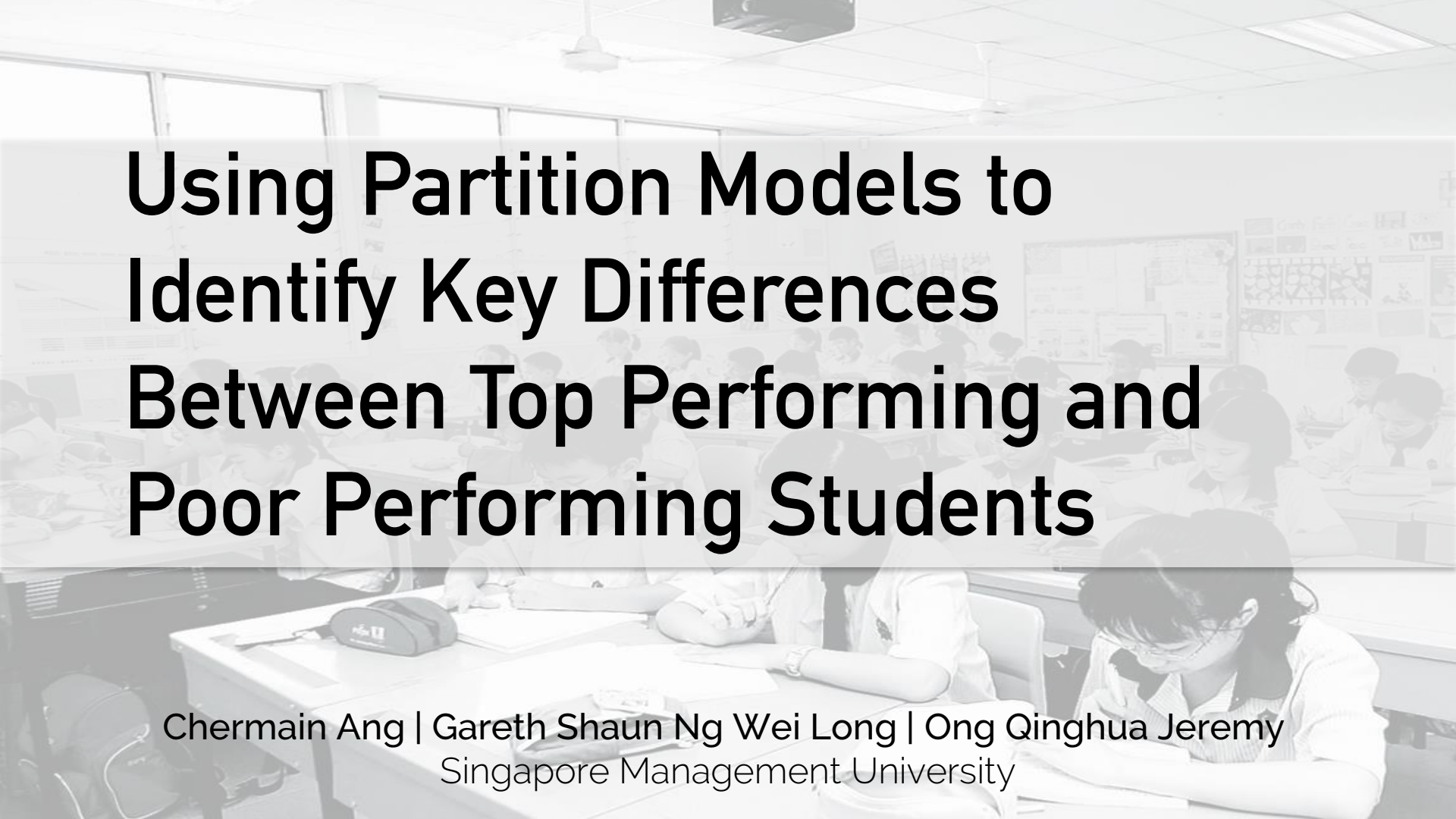


Using Partition Models to Identify Key Differences Between Top Performing and Poor Performing Students

Chermain Ang | Gareth Shaun Ng Wei Long | Ong Qinghua Jeremy
Singapore Management University

Agenda

Project Overview

Data Preparation

Data Analysis

Insights

Conclusion

Background

PISA

Triennial Global Education Survey

Evaluate education systems by testing skills & knowledge of 15 year old students

Math, Reading Science

#1 Reading
Math
Science

Motivation

> Singapore managed to achieve excellence **without wide differences** between children from wealthy and disadvantaged families.
>
> *– Andreas Schleicher*
>
> *(OECD's Education Director)*

IS IT TRUE???

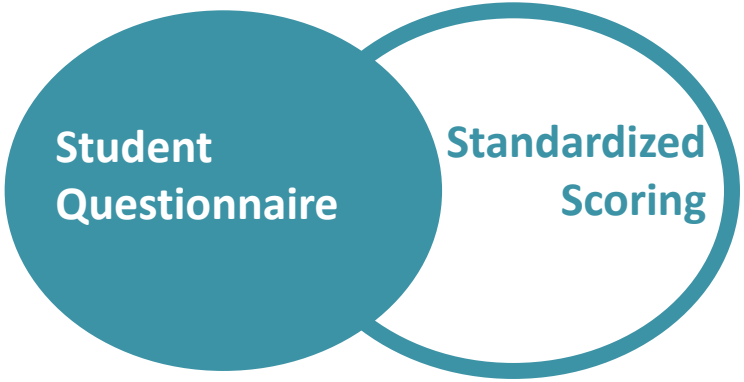
Objective

Math **Reading** **Science** **Overall**

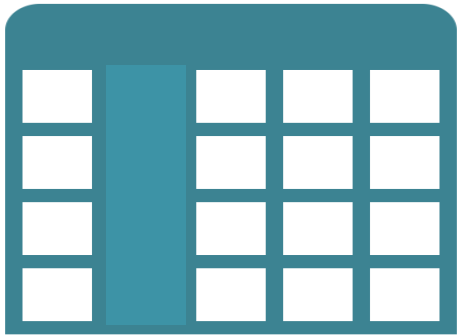
Top Performers

Poorer Performers

Data Preparation



1 Combining Tables



2 Column Selection

- Ordinal
- Nominal
- Continuous

3 Classification of Variables

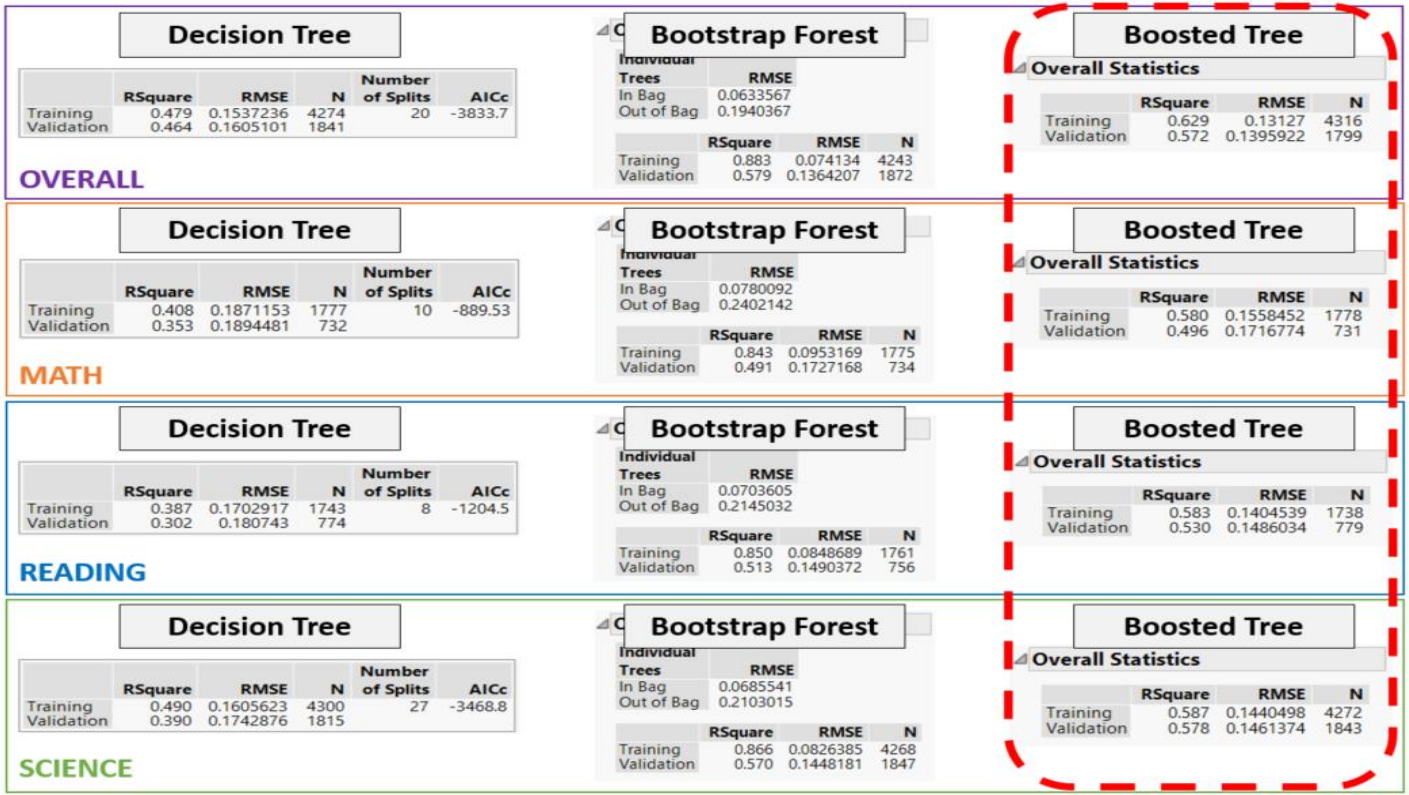
Model Selection

	Multiple Linear Regression	Partition Models
Benefits	• High level of familiarity • Flexibility of model	• Easy to interpret results • Able to handle large amounts of data
Requirements	Relationship between Independent and Dependent Variables has to be linear	Does not require a prior good model
Model Choice		✓

Model Selection

Decision Tree	Bootstrap Forest	Boosted Tree
▪ Data is recursively partitioned ***according to the relationship*** between the dependent and independent variables	▪ Many decision trees are built, and ***subsequently combined*** to form a more powerful model	▪ ***Fitting*** many small decision trees ***sequentially*** (layer by layer), to build a large decision tree

Model Evaluation



DISCUSSION OF RESULTS

M R S O

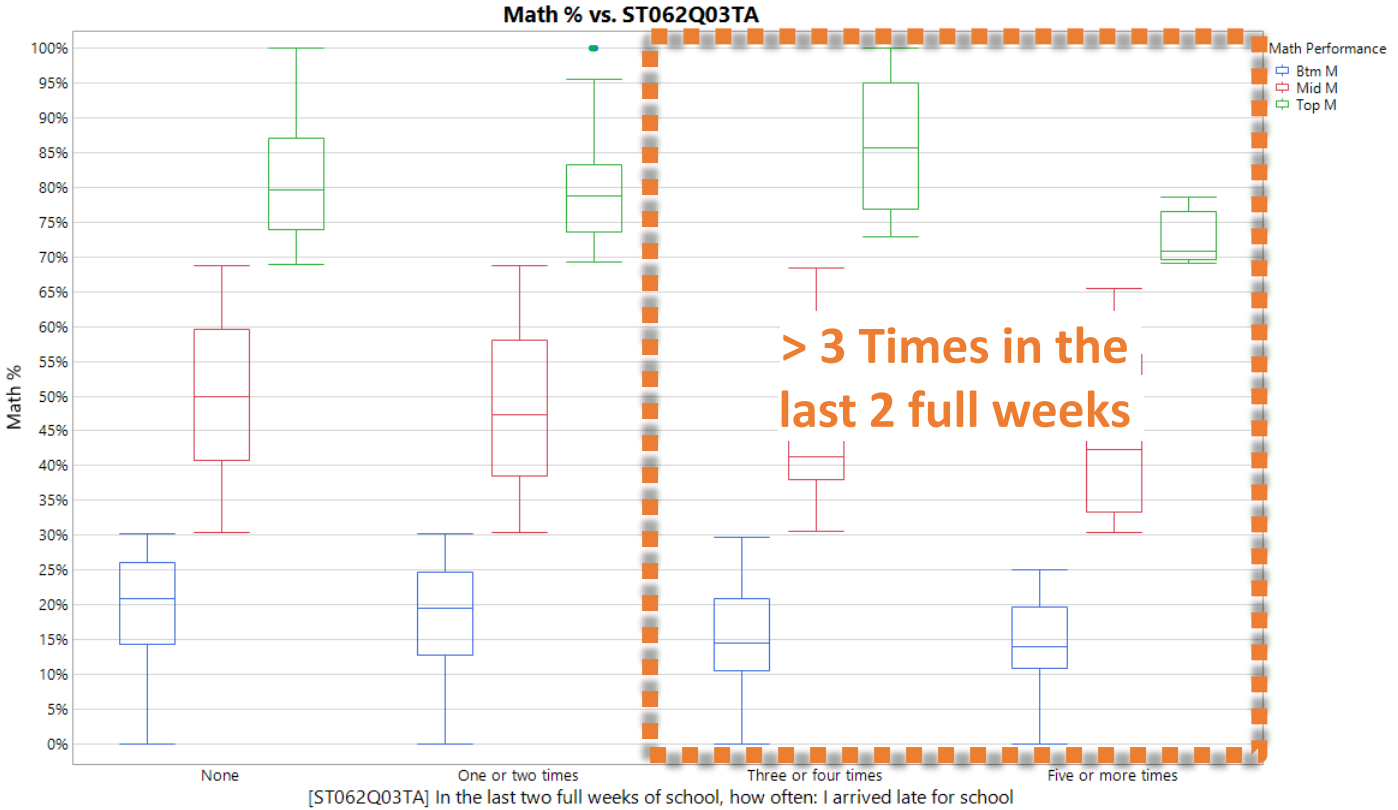
Column Contributions for Math Component

Column Contributions

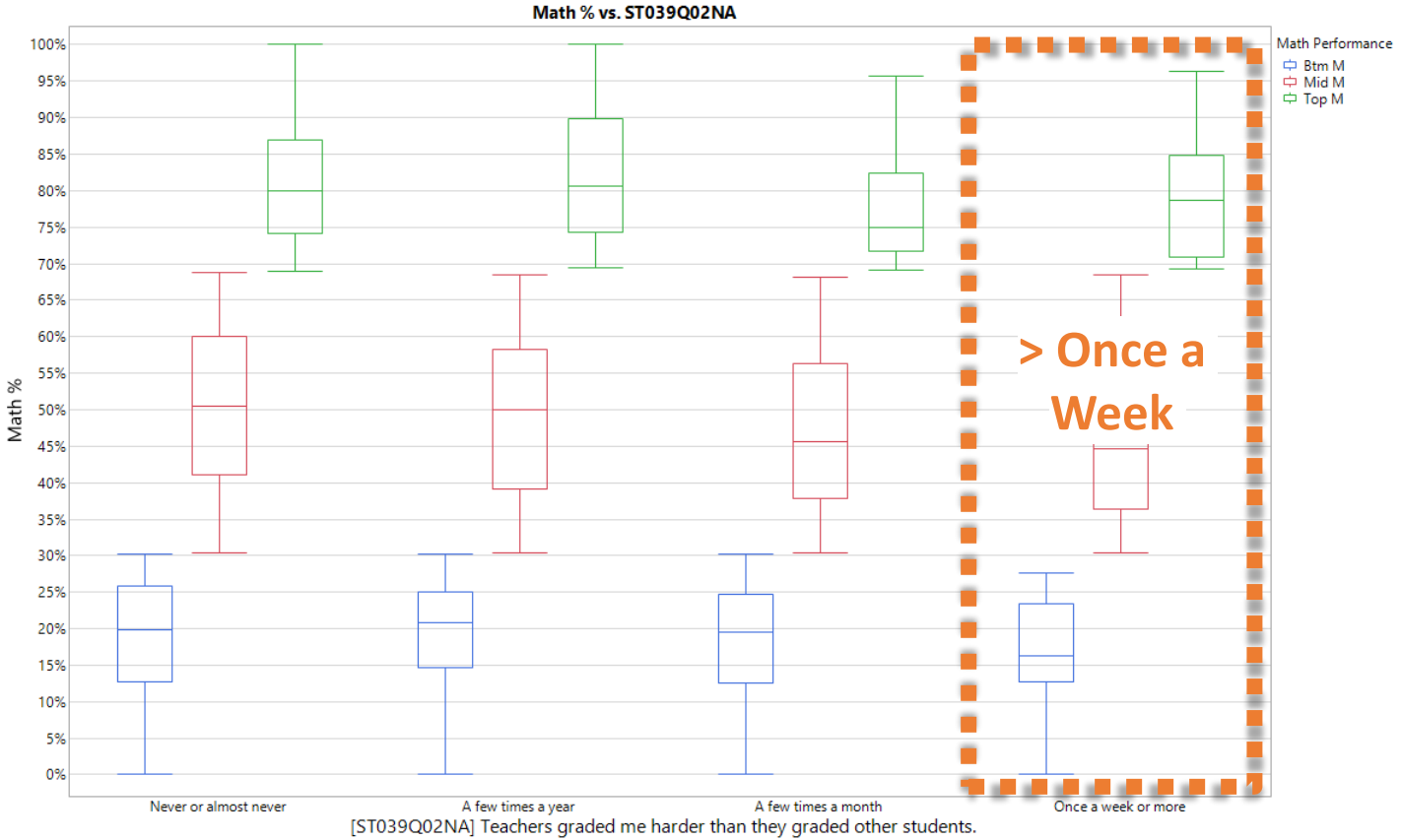
Term	Number of Splits	SS	Portion
ST092Q01TA	8	64.1883924	0.2043
SMINS	10	30.6143619	0.0975
ST064Q01NA	5	30.3087272	0.0965
LMINS	10	28.9515978	0.0922
TMINS	5	20.1208559	0.0641
ST059Q03TA	8	19.7610814	0.0629
ST111Q01TA	3	15.2403646	0.0485
LANGN	3	8.68736467	0.0277
ST062Q03TA	5	5.96854288	0.0190
hisei	2	5.88055787	0.0187
ST039Q02NA	4	5.07109326	0.0161
ST065Class	2	4.93497631	0.0157
IC009Q08TA	4	4.67098352	0.0149
IC008Q03TA	4	4.64548769	0.0148
ST013Q01TA	2	4.08940806	0.0130
IC007Q01TA	4	4.05073582	0.0129
ST097Q04TA	2	3.62523336	0.0115
ST092Q02TA	3	3.49601075	0.0111
ST118Q04NA	4	3.25360927	0.0104
FISCED	2	3.2530317	0.0104
IC002Q01NA	4	2.97832429	0.0095

Question ID	Question Description
ST092Q01TA	How informed are you about this environmental issue? The increase of greenhouse gases in the atmosphere
SMINS	Learning time (minutes per week) - <science>
ST064Q01NA	<school science> courses? I can choose the <school science> course(s) I study.
LMINS	Learning time (minutes per week) - <test language>
TMINS	Learning time (minutes per week) - in total
ST059Q03TA	Number of <class periods> required per week in <science>
ST111Q01TA	Which of the following do you expect to complete?
LANGN	Language at home (3-digit code)
ST062Q03TA	In the last two full weeks of school, how often: I arrived late for school
hisei	Index highest parental occupational status
ST039Q02NA	Teachers graded me harder than they graded other students.
ST065Class	Student coded science class (from ST065Q01NA)
IC009Q08TA	Digitial devices available at school: USB (memory) stick
IC008Q03TA	Use digital devices outside school for using email.
ST013Q01TA	How many books are there in your home?
IC007Q01TA	On a typical weekend day, for how long do you use the Internet outside of school?
ST097Q04TA	How often does this happen in your <school science> lessons? Students cannot work well.
ST092Q02TA	How informed are you about this environmental issue? The use of genetically modified organisms (<GMO>)
ST118Q04NA	I get very tense when I study for a test.
FISCED	Father's Education (ISCED)

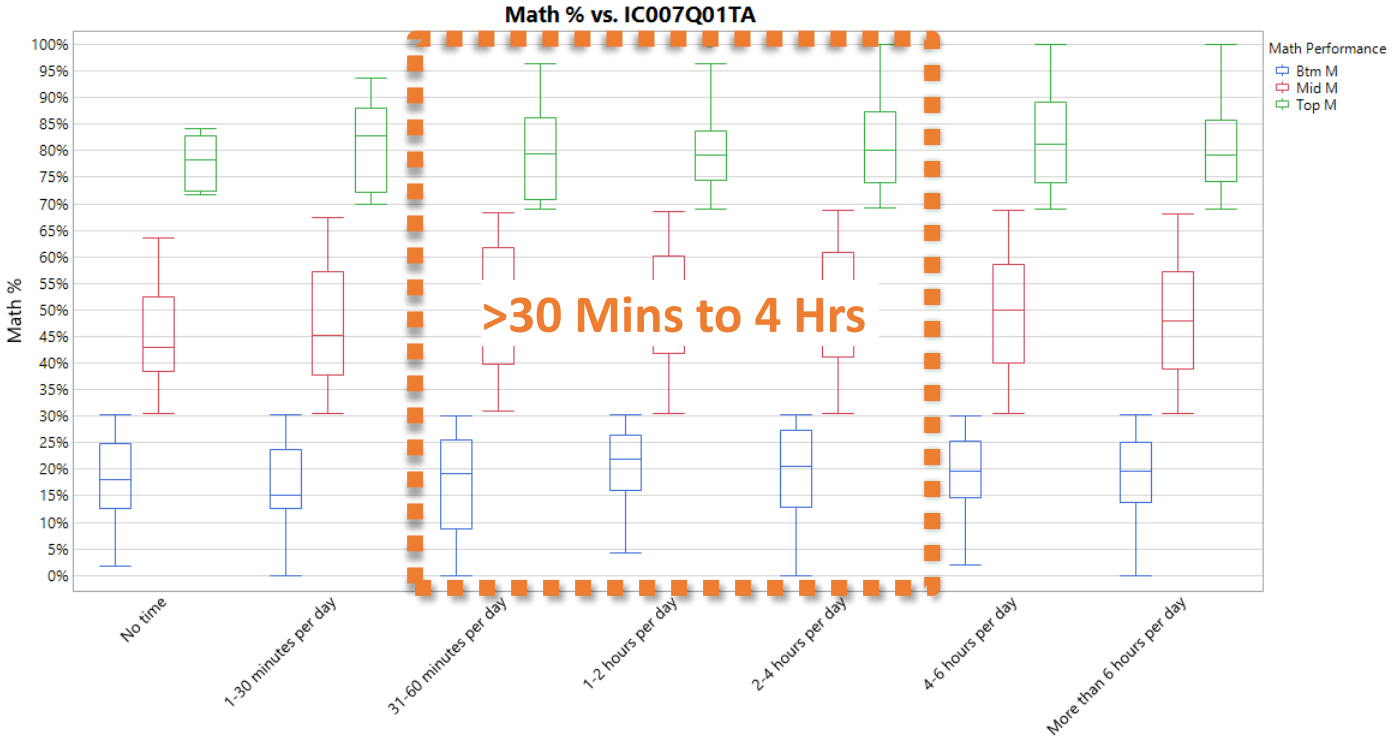
Punctuality



Opinion on Teacher's grading

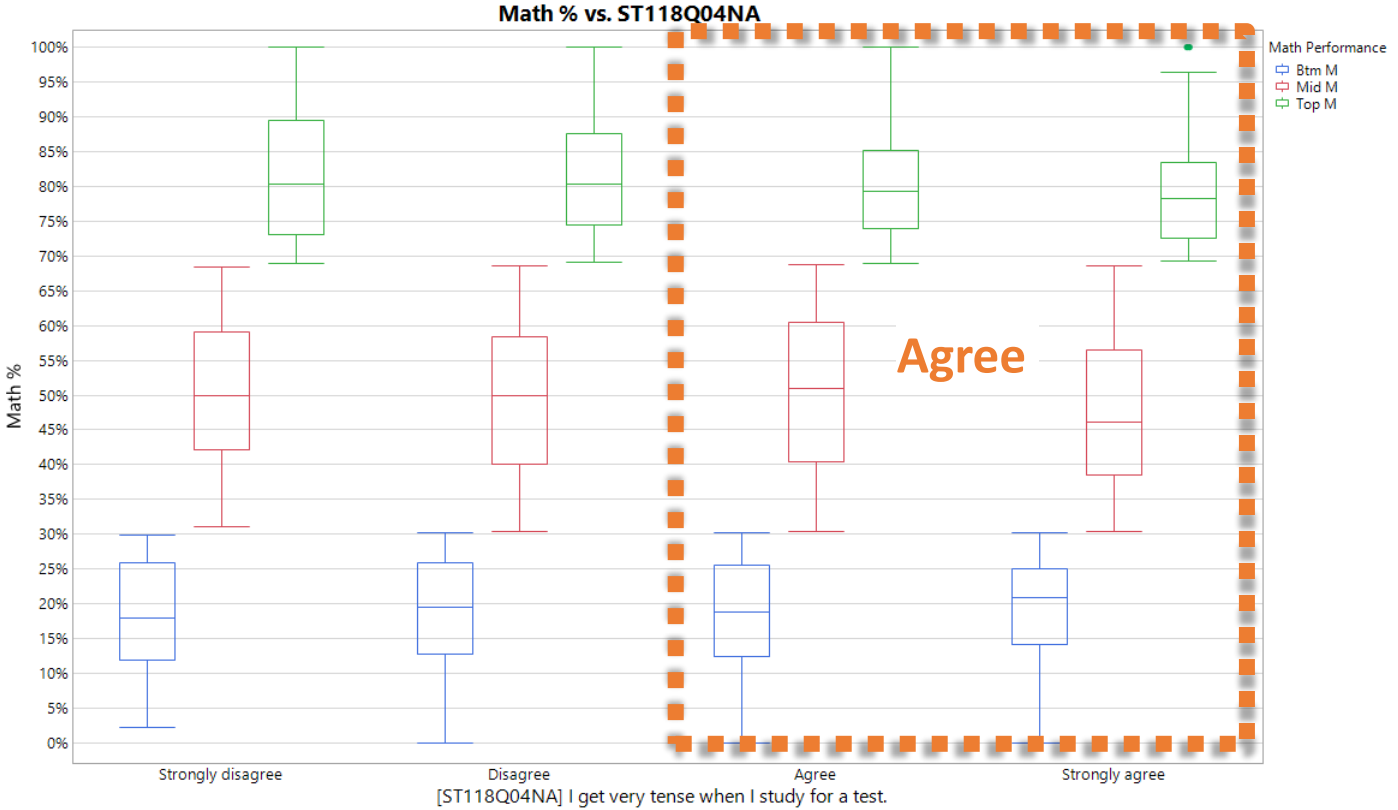


Weekend Internet Usage

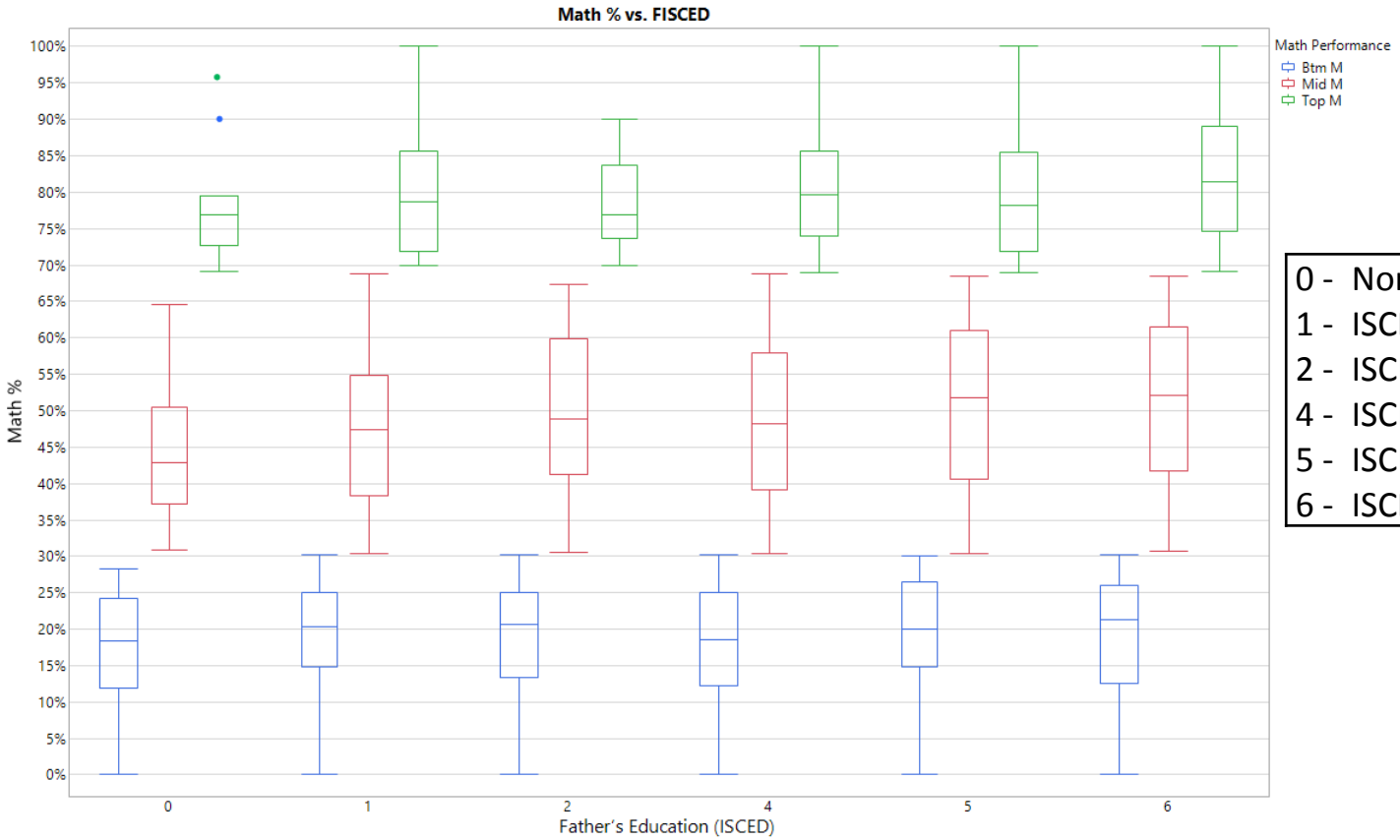


[IC007Q01TA] On a typical weekend day, for how long do you use the Internet outside of school?

Tense?


Math % vs. ST118Q04NA
Math Performance
Btm M
Mid M
Top M
Math %
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Agree
Strongly disagree
Disagree
Agree
Strongly agree
[ST118Q04NA] I get very tense when I study for a test.

Father's Education Level



0 - None
1 - ISCED 1
2 - ISCED 2
4 - ISCED 3A, ISCED 4
5 - ISCED 5B
6 - ISCED 5A, 6

Column Contributions for Reading Component

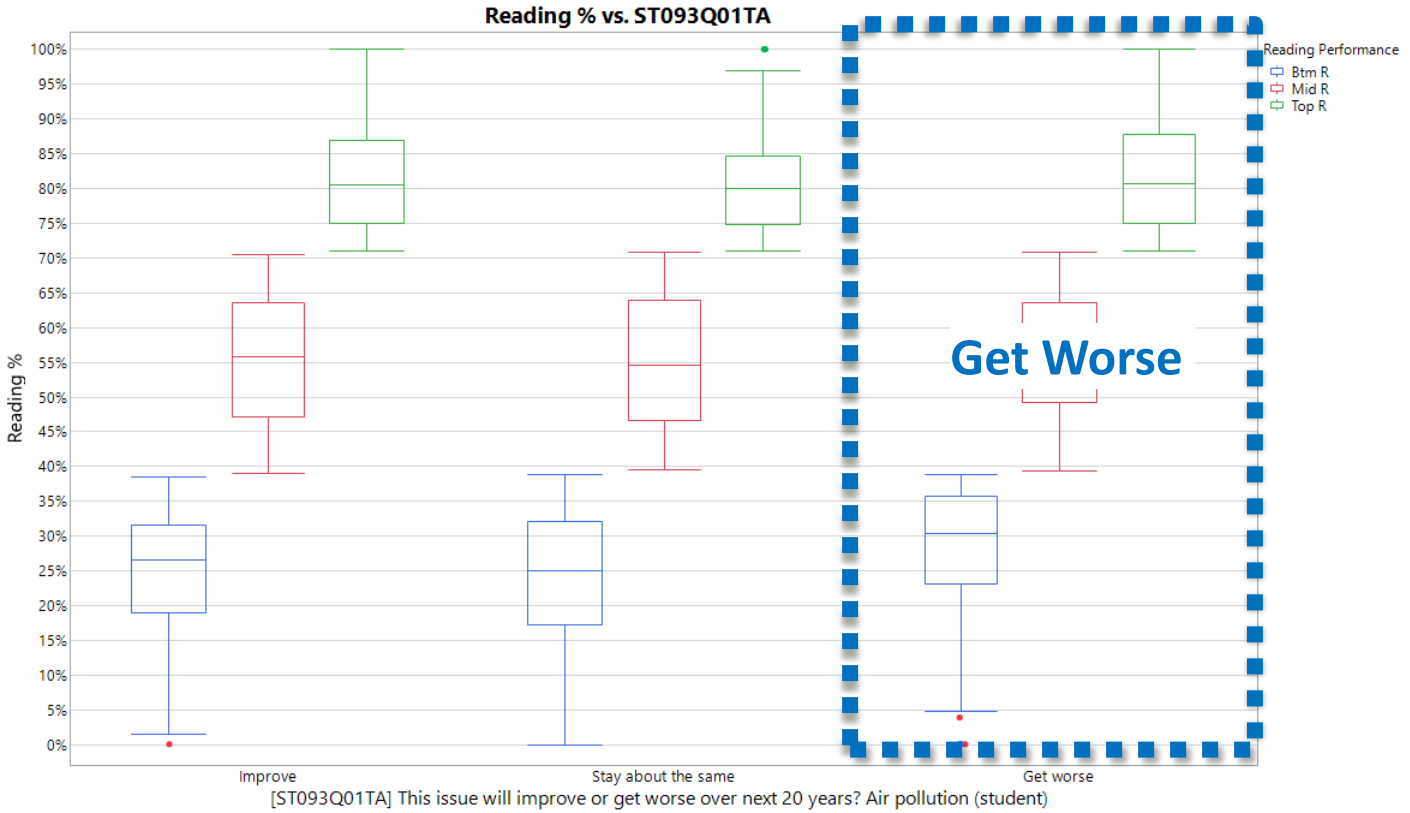
Column Contributions

Term	Number of Splits	SS	Portion
ST092Q01TA	4	42.2843034	0.1677
ST064Q01NA	5	23.5122222	0.0932
SMINS	8	16.4610877	0.0653
ST013Q01TA	5	12.7887107	0.0507
ST111Q01TA	3	12.2670334	0.0487
ST065Class	3	12.2149655	0.0484
LMINS	7	12.1850499	0.0483
ST093Q01TA	5	10.0383823	0.0398
ST131Q06NA	5	9.45569548	0.0375
ST011Q07TA	4	7.60689145	0.0302
IC008Q03TA	6	6.99096798	0.0277
ST076Q11NA	5	6.44476684	0.0256
ST092Q05TA	2	6.20299338	0.0246
ST121Q01NA	6	5.63320985	0.0223
TMINS	2	4.84828843	0.0192
ST059Q03TA	3	4.60604058	0.0183
ST092Q02TA	3	4.10116776	0.0163
BFMJ2	3	3.46898286	0.0138
IC002Q01NA	5	3.23465807	0.0128
IC008Q13NA	4	3.07852248	0.0122
IC009Q08TA	3	2.73806516	0.0109
ST082Q13NA	4	2.7245493	0.0108
IC013Q04NA	3	2.38756848	0.0095

Question ID	Question Description
ST092Q01TA	How informed are you about this environmental issue? The increase of greenhouse gases in the atmosphere
ST064Q01NA	<school science> courses? I can choose the <school science> course(s) I study.
SMINS	Learning time (minutes per week) - <science>
ST013Q01TA	How many books are there in your home?
ST111Q01TA	Which of the following do you expect to complete?
ST065Class	Student coded science class (from ST065Q01NA)
LMINS	Learning time (minutes per week) - <test language>
ST093Q01TA	This issue will improve or get worse over next 20 years? Air pollution (student)
ST131Q06NA	It is good to try experiments more than once to make sure of your findings.
ST011Q07TA	In your home: Classic literature (e.g. <Shakespeare>)
IC008Q03TA	Use digital devices outside school for using email.
ST076Q11NA	Before going to school did you: Exercise or practice a sport
ST092Q05TA	How informed are you about this environmental issue? The consequences of clearing forests\other land use
ST121Q01NA	<NAME 1> is motivated? Gives up easily when confronted with a problem and is often not prepared
TMINS	Learning time (minutes per week) - in total
ST059Q03TA	Number of <class periods> required per week in <science>
ST092Q02TA	How informed are you about this environmental issue? The use of genetically modified organisms
BFMJ2	ISEI of father
IC002Q01NA	How old were you when you first used a digital device?
IC008Q13NA	Use digital devices outside school for downloading new apps on a mobile device.
IC009Q08TA	Digitial devices available at school: USB (memory) stick
ST082Q13NA	To what extent do you disagree or agree about yourself? I find that teamwork raises my own efficiency.

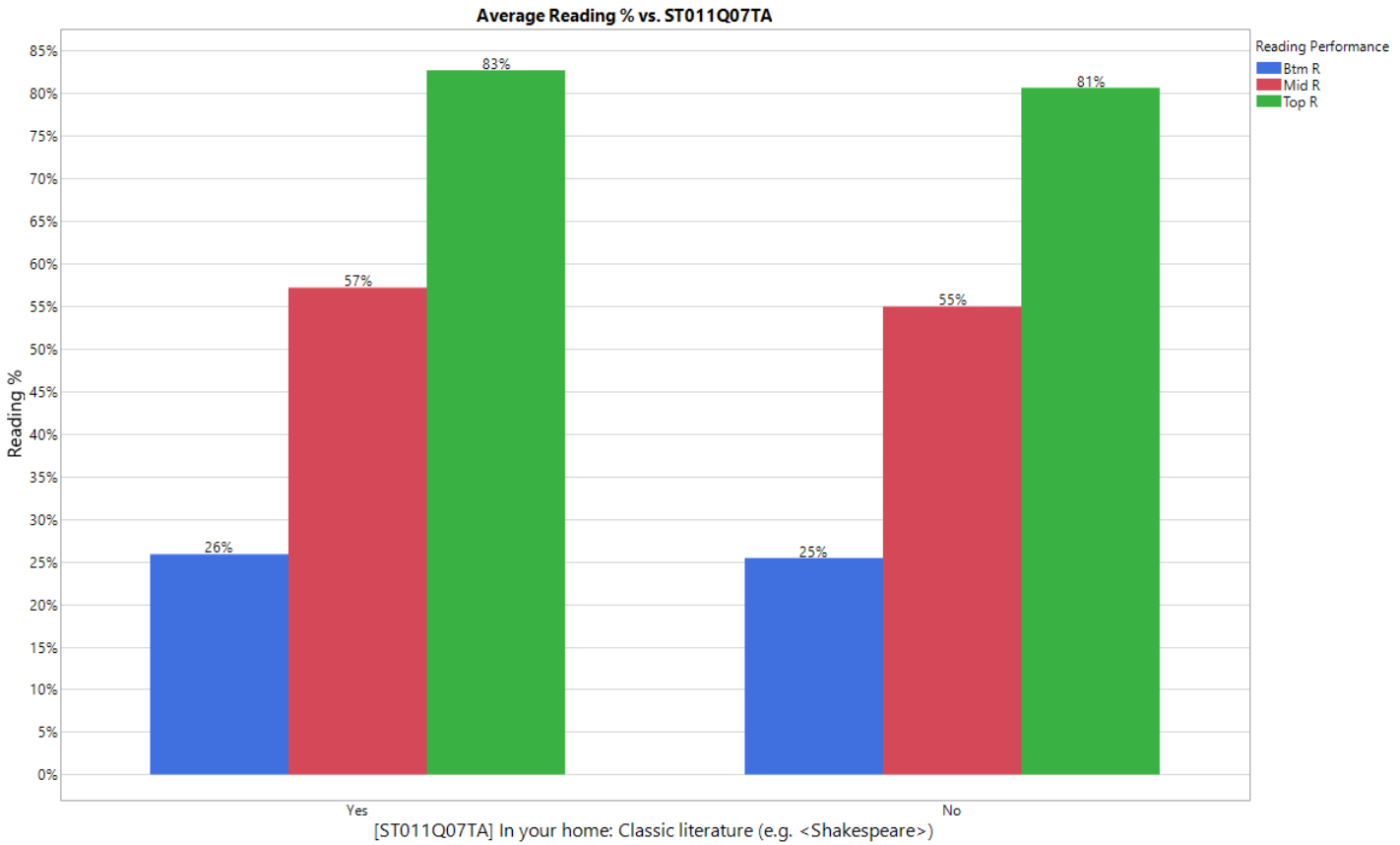
Future Optimism

~5%

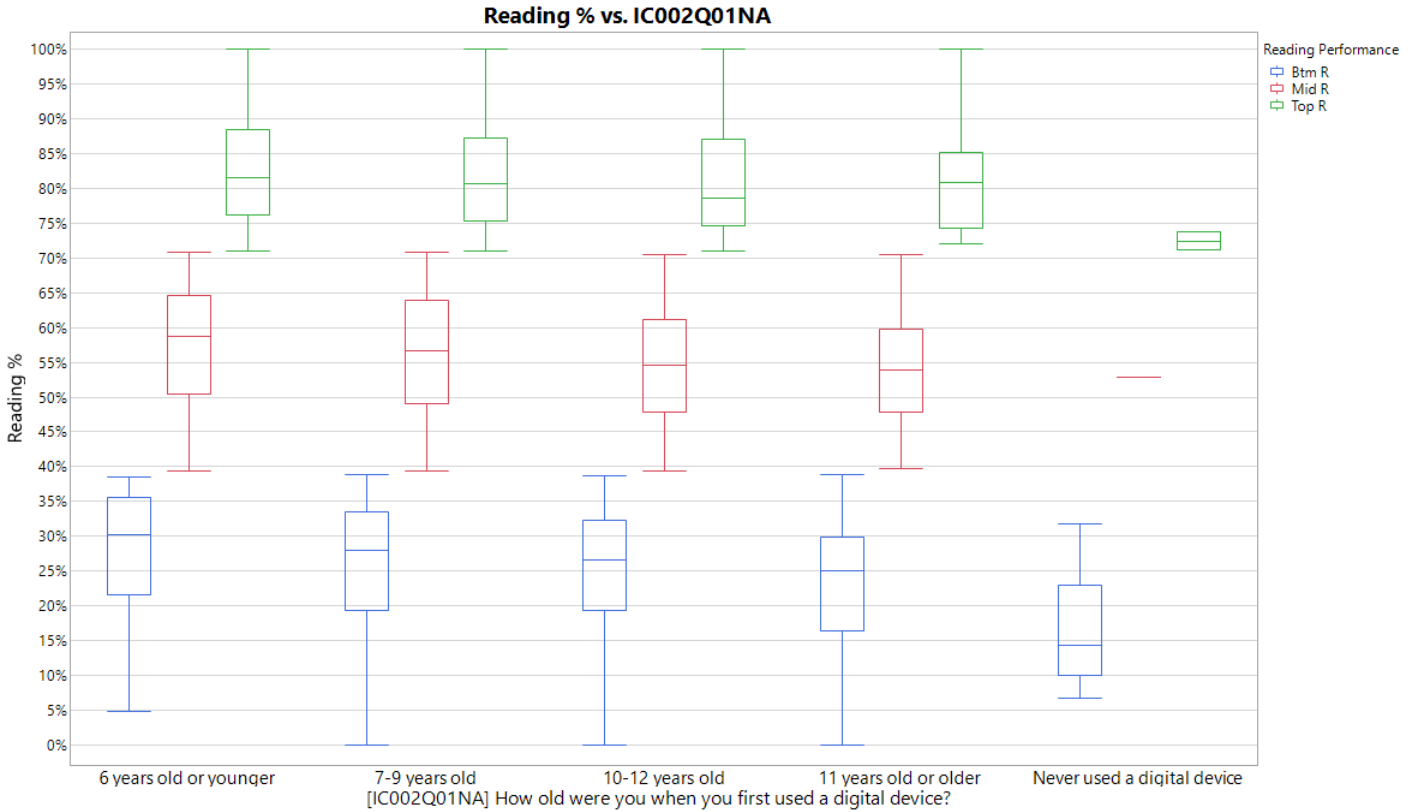


Classic Literature

~2%



Exposure to Digital Devices



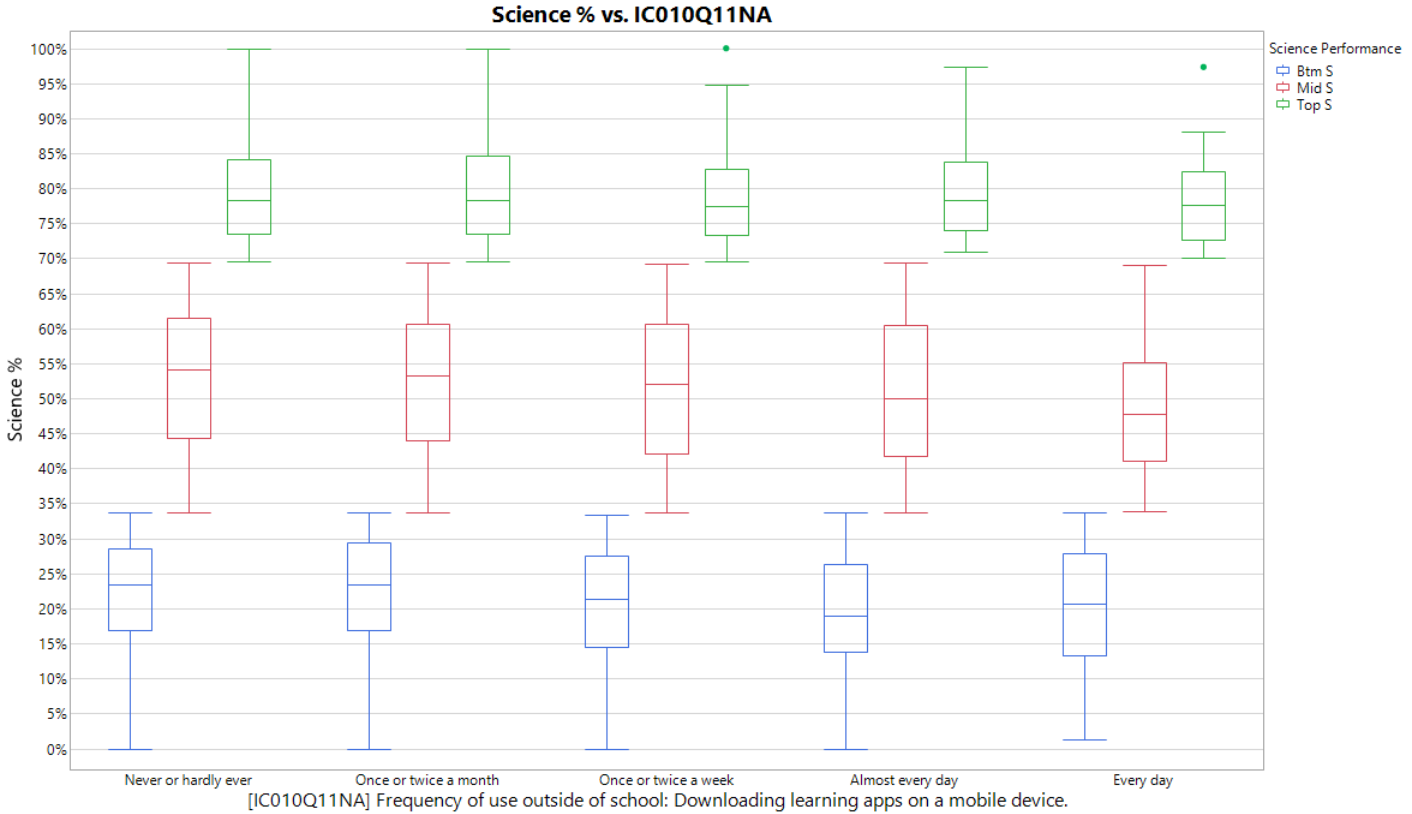
Column Contributions for Science Component

Column Contributions

Term	Number of Splits	SS	Portion
ST092Q01TA	9	207.676908	0.3134
LMINS	9	48.5579711	0.0733
SMINS	8	47.6805886	0.0719
ST064Q01NA	4	34.3913875	0.0519
ST065Class	4	31.3041596	0.0472
ST131Q06NA	3	22.9003653	0.0346
ST013Q01TA	4	17.7507752	0.0268
ST059Q03TA	4	17.7191312	0.0267
ST092Q02TA	5	15.0421448	0.0227
ST103Q01NA	4	14.5357796	0.0219
BFMJ2	2	13.8968569	0.0210
TMINS	2	13.6794789	0.0206
ST121Q01NA	4	9.88575963	0.0149
IC008Q03TA	4	9.4970133	0.0143
ST111Q01TA	2	9.47717741	0.0143
ST129Q02TA	3	9.44575021	0.0143
IC009Q11NA	4	8.07498565	0.0122
ST093Q07NA	2	7.71019708	0.0116
IC010Q11NA	3	6.73946602	0.0102
IC002Q01NA	3	5.84466891	0.0088

Question ID	Question Description
ST092Q01TA	How informed are you about this environmental issue? The increase of greenhouse gases in the atmosphere
LMINS	Learning time (minutes per week) - <test language>
SMINS	Learning time (minutes per week) - <science>
ST064Q01NA	<school science> courses? I can choose the <school science> course(s) I study.
ST065Class	Student coded science class (from ST065Q01NA)
ST131Q06NA	It is good to try experiments more than once to make sure of your findings.
ST013Q01TA	How many books are there in your home?
ST059Q03TA	Number of <class periods> required per week in <science>
ST092Q02TA	How informed are you about this environmental issue? The use of genetically modified organisms
ST103Q01NA	How often does this happen in <school science>? The teacher explains scientific ideas.
BFMJ2	ISEI of father
TMINS	Learning time (minutes per week) - in total
ST121Q01NA	<NAME 1> is motivated? Gives up easily when confronted with a problem and is often not prepared
IC008Q03TA	Use digital devices outside school for using email.
ST111Q01TA	Which of the following do you expect to complete?
ST129Q02TA	Explain why earthquakes occur more frequently in some areas than in others.
IC009Q11NA	Digitial devices available at school: Interactive Whiteboard, e.g. <Smartboard®>
ST093Q07NA	This issue will improve or get worse over next 20 years? The increase of greenhouse gases in the atmosphere
IC010Q11NA	Frequency of use outside of school: Downloading learning apps on a mobile device.

Learning Apps



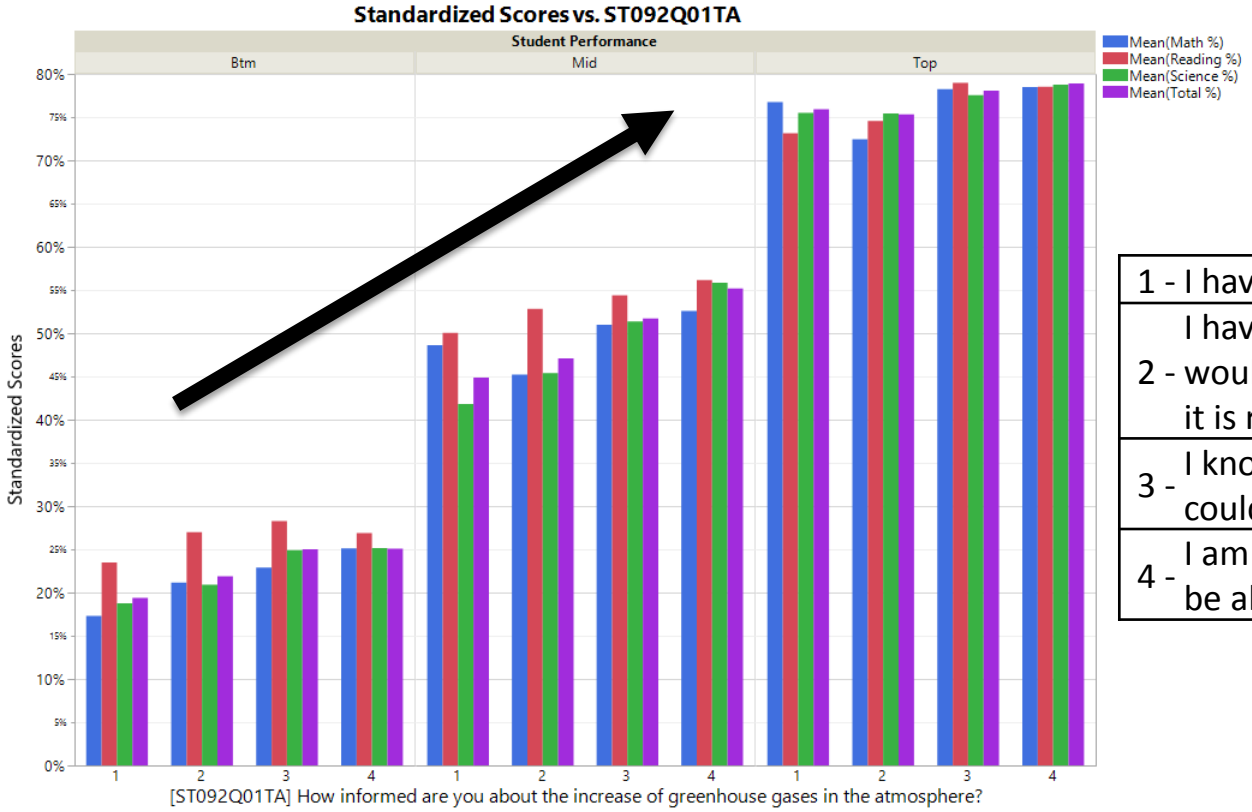
Column Contributions for Overall Results

Column Contributions

Term	Number of Splits	SS	Portion
ST092Q01TA	7	189.599059	0.2855
SMINS	10	62.6240549	0.0943
LMINS	9	54.5334635	0.0821
ST064Q01NA	4	39.6419292	0.0597
ST013Q01TA	6	26.6159236	0.0401
ST065Class	5	25.9147706	0.0390
ST131Q06NA	5	22.4298004	0.0338
TMINS	3	18.7177171	0.0282
ST111Q01TA	2	14.9362663	0.0225
IC008Q03TA	7	14.112212	0.0213
ST093Q07NA	3	13.7311301	0.0207
ST103Q01NA	4	13.4349232	0.0202
ST121Q01NA	5	13.4314221	0.0202
ST059Q03TA	4	11.9464936	0.0180
ST092Q02TA	6	11.7075194	0.0176
hisei	1	10.8498325	0.0163
ST092Q05TA	1	8.78449599	0.0132
IC008Q12TA	4	8.55011456	0.0129
ST076Q11NA	3	8.30944783	0.0125
ST129Q02TA	4	7.42028727	0.0112
LANGN	3	7.39281618	0.0111
IC009Q11NA	4	6.86691362	0.0103
IC009Q08TA	3	5.80597372	0.0087

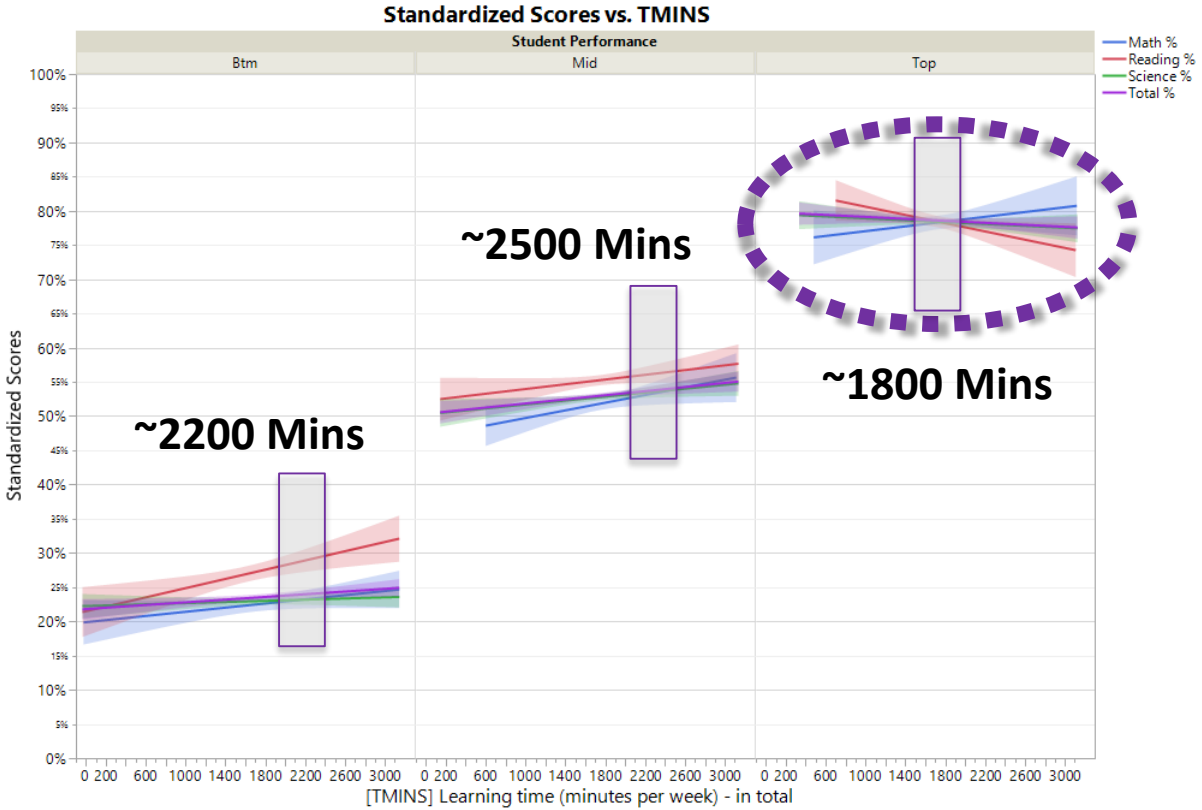
Question ID	Question Description
ST092Q01TA	How informed are you about this environmental issue? The increase of greenhouse gases in the atmosphere
SMINS	Learning time (minutes per week) - <science>
LMINS	Learning time (minutes per week) - <test language>
ST064Q01NA	<school science> courses? I can choose the <school science> course(s) I study.
ST013Q01TA	How many books are there in your home?
ST065Class	Student coded science class (from ST065Q01NA)
TMINS	Learning time (minutes per week) - in total
ST111Q01TA	Which of the following do you expect to complete?
IC008Q03TA	Use digital devices outside school for using email.
ST059Q03TA	Number of <class periods> required per week in <science>
ST092Q02TA	How informed are you about this environmental issue? The use of genetically modified organisms

Environmental Awareness

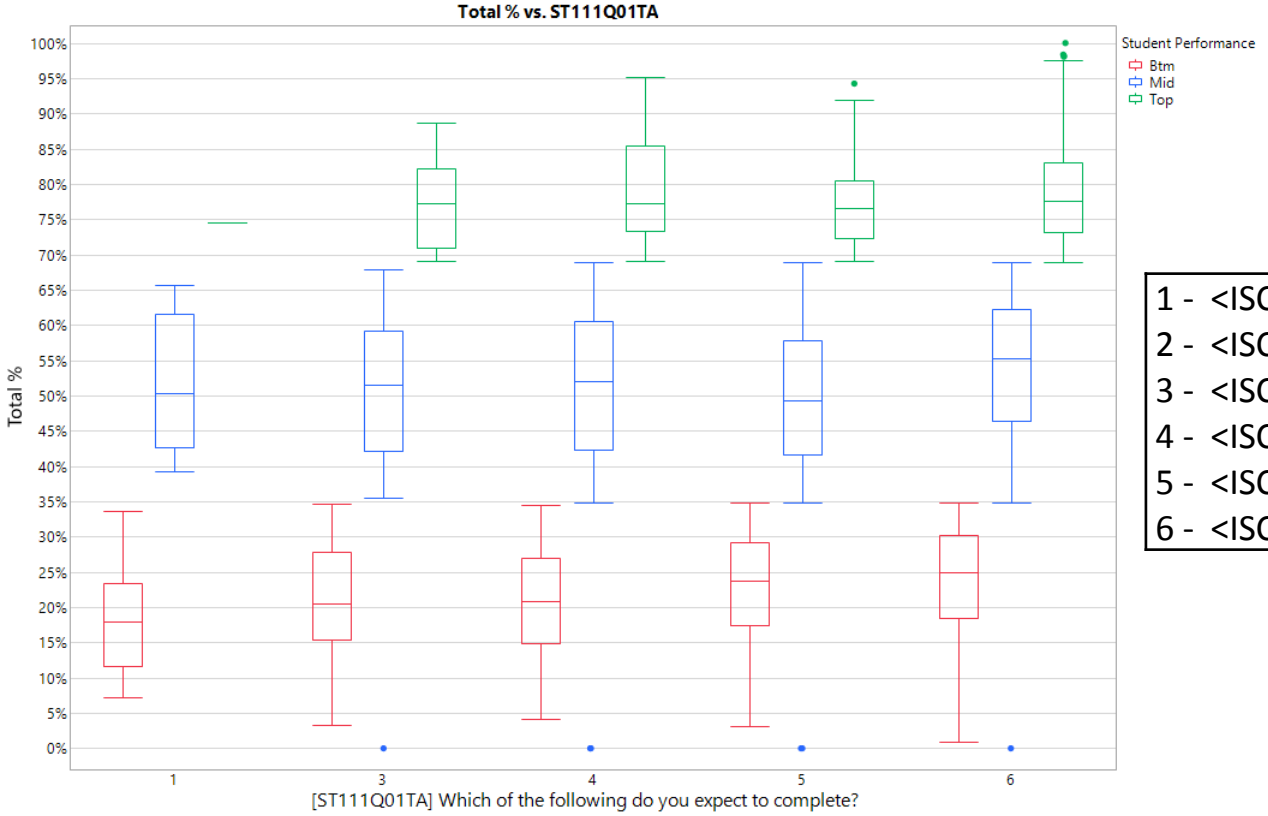


1 -	I have never heard of this
2 -	I have heard about this but I would not be able to explain what it is really about
3 -	I know something about this and could explain the general issue
4 -	I am familiar with this and I would be able to explain this well

Time Spent Learning



Education Expectations



1 - <ISCED level 2>
2 - <ISCED level 3B or C>
3 - <ISCED level 3A>
4 - <ISCED level 4>
5 - <ISCED level 5B>
6 - <ISCED level 5A or 6>

Conclusion

- ❑ Socioeconomic factors
- ❑ Learning Environment
- ❑ Exposure to Digital Devices


Thank You