



LOGAN

An R Package for Log-file Analysis in International Large-Scale Assessments



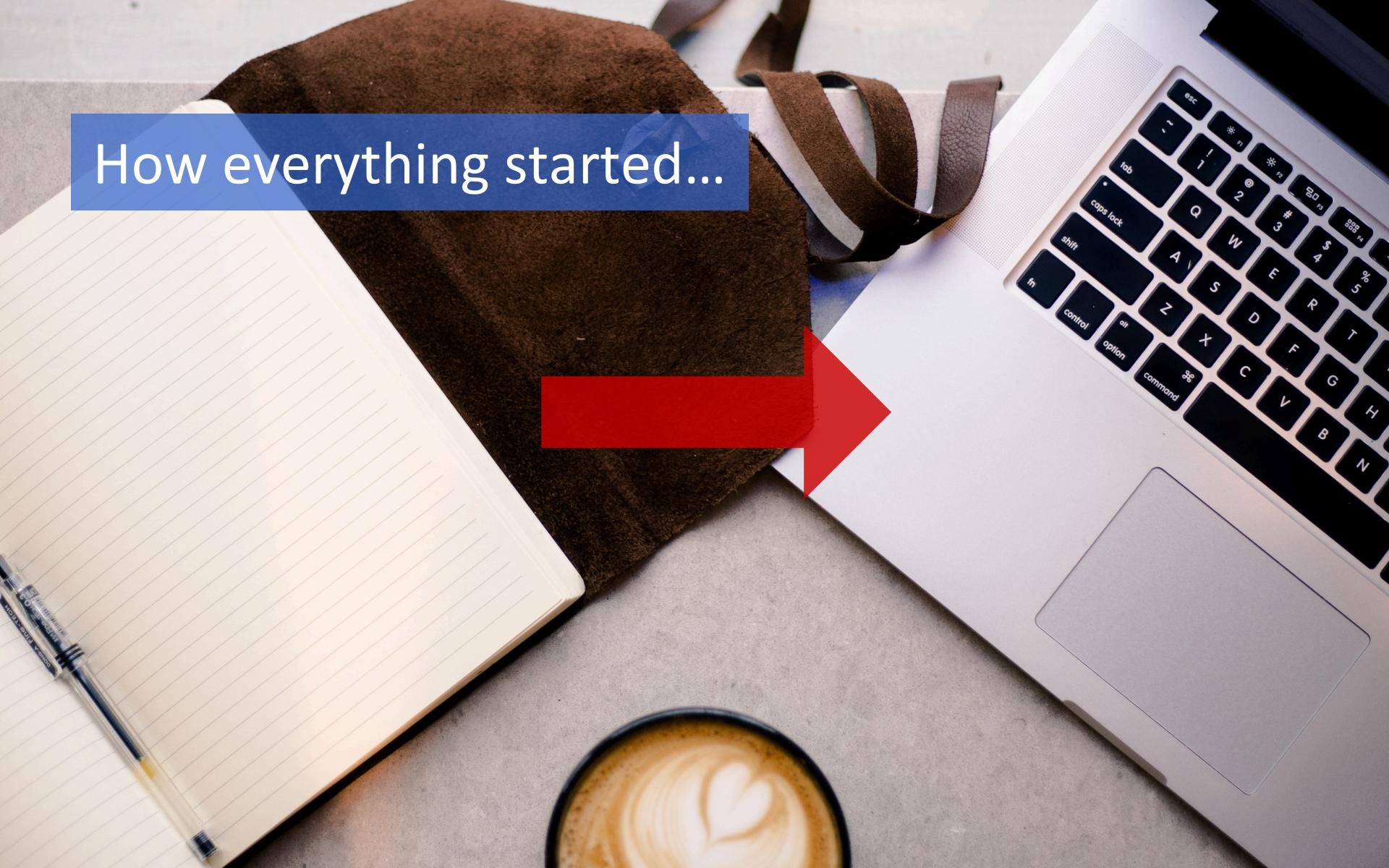
Denise Reis Costa (d.r.costa@cemo.uio.no)
Ronny Scherer (ronny.scherer@cemo.uio.no)



June 2020

CEMO





How everything started...



LOGAN R package thanks...

Credits

Ronny Scherer - CEMO/University of Oslo
Sigrid Blömeke - CEMO/University of Oslo
Rolf Vegar Olsen - CEMO/University of Oslo
Bjørn Andersson - CEMO/University of Oslo
Fredrik Jensen - ILS/University of Oslo
Nani Teig - ILS/University of Oslo
Jelena Radisic - ILS/University of Oslo
Tove Stjern Frønes - ILS/University of Oslo
Waldir Leoncio Netto – OCBE/University of Oslo
Maria Bolsinova - ACTNext
Jesper Tijmstra - Tilburg University
Qiwei He - ETS
Samuel Greiff - University of Luxembourg
Frank Goldhammer - DIPF



Package ‘LOGAN’

May 4, 2019

Title Log File Analysis in International Large-Scale Assessments

Version 1.0.0

Date 2019-05-02

Description Enables users to handle the dataset cleaning for conducting specific analyses with the log files from two international educational assessments: the Programme for International Student Assessment (PISA, <<http://www.oecd.org/pisa/>>) and the Programme for the International Assessment of Adult Competencies (PIAAC, <<http://www.oecd.org/skills/piaac/>>). An illustration of the analyses can be found on the LOGAN Shiny app (<<https://loganpackage.shinyapps.io/shiny/>>) on your browser.

Author Denise Reis Costa [aut, cre],
Waldir Leoncio [aut]



LOGAN Architecture

Data entry: raw data (e.g., xml) or semi-processed log file (e.g., PISA 2012 SPSS format)

Example of functions: import data, count of log events per variable, clean events

Example of functions: summary of total time by variable, summary of the association between test performance and specific respondent's strategy

Log file data

Module 0:
Data
Preparation

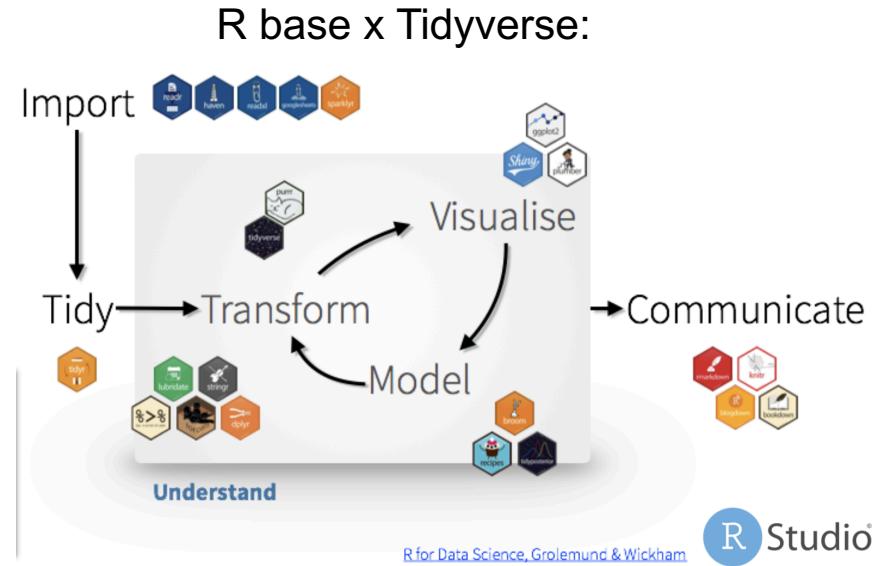
Module 1:
Time

Module 2:
Actions



LOGAN Features

- Item-level analysis
- Modules
- Tidyverse package
- Generate automatic reports with log-file analyses





LOGAN v1.0.0

Module	Functions	Description
0: Data preparation	CleanActions	Clean events
	ConcatActions	Concatenate events
	DataActionsbyID	Transform dataset to wide format
	ImportSPSS	Read SPSS process data
	RangeNumberActionsbyVar	Check number of students and actions by a specific variable
1: Time	TrimVar	Trim variables
	NumericTimeVar	Convert time to numeric
	PlotTimeonTaskbyVar	Plot the density of Time on task by specific variable
	SummaryTOTbyVar	Summarize time on task by a specific variable
	TOTVar	Calculate time on task
2: Cognitive-related actions	VarTimebyID	Extract the start or end time
	DescriptiveStrategy	Report descriptive statistics by strategy
	PlotStrategybyCatPerformance	Plot the percentage of students by strategy and performance
	VarActionSearch	Retrieve the frequency of specifics events in a variable of Actions

Embedded Example:

- PISA 2012 problem-solving task «Climate Control»
- Item CP025Q01

en-GB Programme for International Student Assessment 2012

CLIMATE CONTROL

You have no instructions for your new air conditioner. You need to work out how to use it.

You can change the top, central and bottom controls on the left by using the sliders (↔). The initial setting for each control is indicated by ▲.

By clicking APPLY, you will see any changes in the temperature and humidity of the room in the temperature and humidity graphs. The box to the left of each graph shows the current level of temperature or humidity.

Top Control
Central Control
Bottom Control

Temperature: 33
Humidity: 26

APPLY RESET

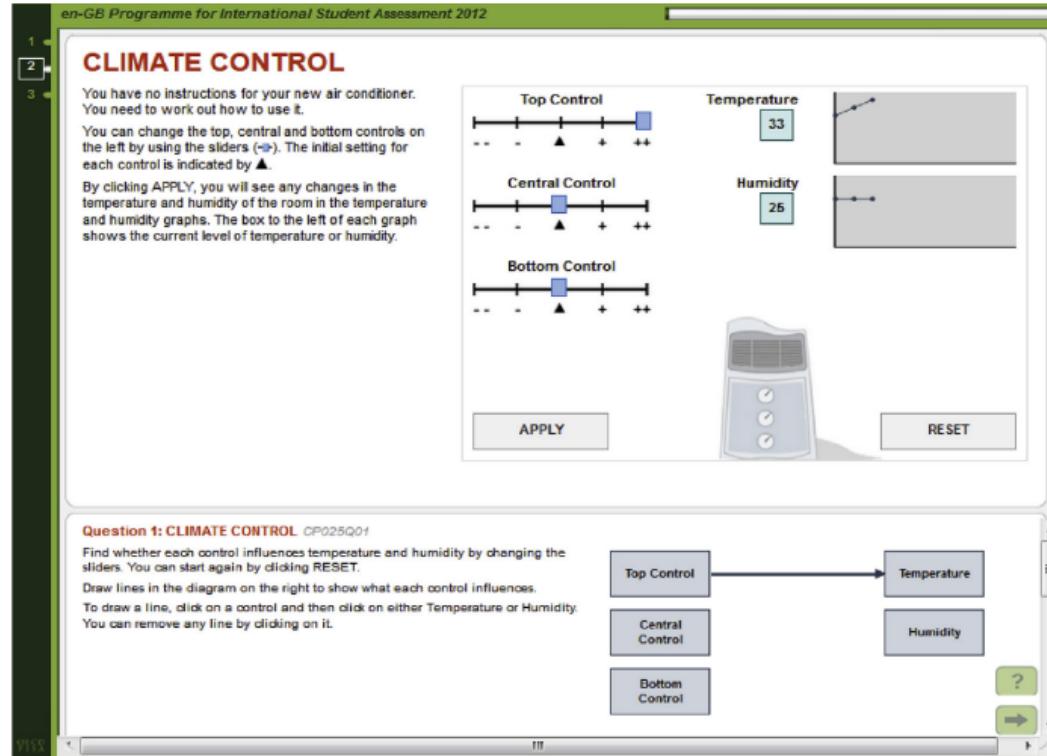
Question 1: CLIMATE CONTROL CP025Q01

Find whether each control influences temperature and humidity by changing the sliders. You can start again by clicking RESET.

Draw lines in the diagram on the right to show what each control influences. To draw a line, click on a control and then click on either Temperature or Humidity. You can remove any line by clicking on it.

Top Control → Temperature
Central Control → Humidity
Bottom Control → Temperature

?





Module 0: Data preparation

Log-file data:

The interactive table below shows the public log-file data from this item. It was downloaded from the

Show entries

cnt	schoolid	StdIDStd	event	tim
ARE	0000189	04852	START_ITEM	1288
ARE	0000189	04852	ACER_EVENT	1291
ARE	0000189	04852	ACER_EVENT	1338
ARE	0000189	04852	ACER_EVENT	1346

	cnt	n.STID	max.event	min.event
25	MAC	519	224	3
26	MNE	927	188	2
27	MYS	949	326	1
28	NLD	895	644	2
29	NOR	406	1077	1
30	POL	384	154	2
31	PRT	489	243	2
32	QCN	408	151	3
33	QCY	1238	254	2
34	QUK	1	25528	25528
35	RUS	505	254	2
36	SGP	469	165	2
37	SRB	873	147	1
38	SVK	486	182	2
39	SVN	671	449	1
40	SWE	422	3511	2

Showing 1 to 40 of 44 entries



Module 1: Time on task

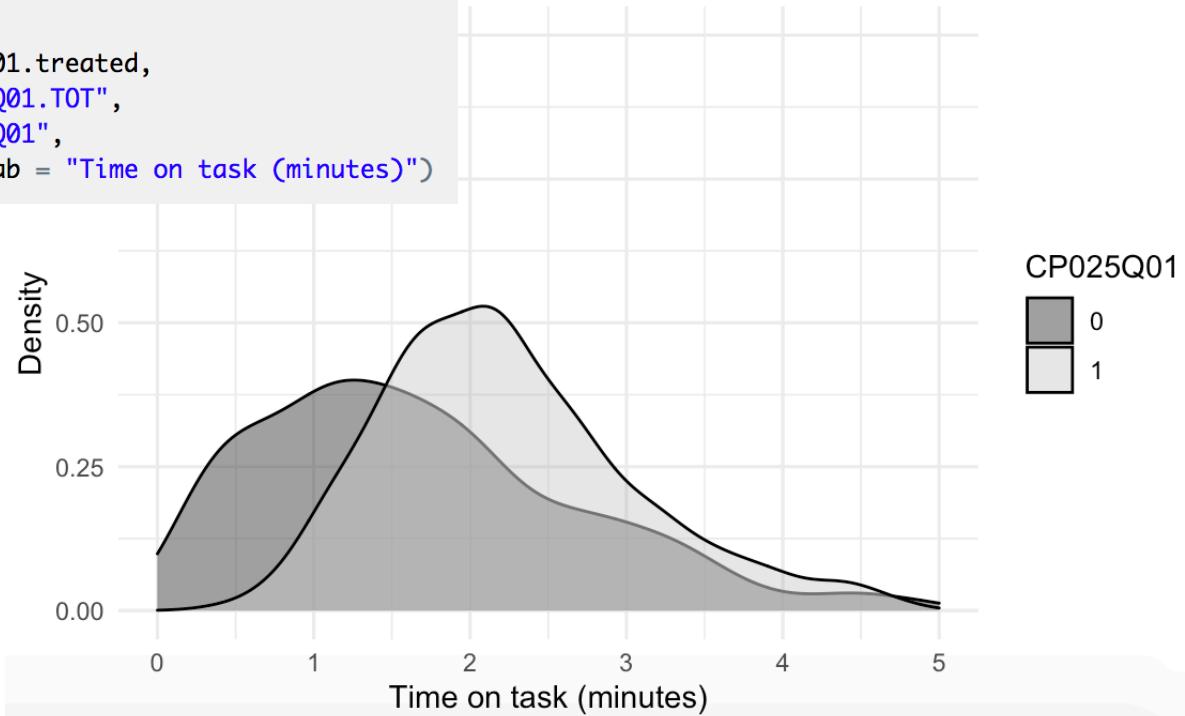
```
# Summary of time on task by a specific variable  
m1$SummaryTOTbyVar(cp025q01.treated,  
                     "CP025Q01.TOT",  
                     "CP025Q01",  
                     TRUE)
```

Summary of Time on Task by CP025Q01 - Individual level

Statistics	ToT Total	CP025Q01=0	CP025Q01=1
Total N	1470	737	733
Min	-3.08	-3.08	-2.2
1st.Qu	1.28	0.92	1.66
Median	1.9	1.51	2.14
Mean	2	1.72	2.28
SD	1.12	1.2	0.95
3st.Qu	2.57	2.32	2.71
Max	9.81	9.81	7.41

Module 1: Time on task

```
# Plot the time data
m1$PlotTimeonTaskbyVar(cp025q01.treated,
                      "CP025Q01.TOT",
                      "CP025Q01",
                      namexlab = "Time on task (minutes)")
```





Module 2: Respondents' actions

```
# Report: Descriptive statistics by strategy
m2$DescriptiveStrategy(cp025q01.treated,
                      "votat",
                      "CP025Q01",
                      "PV1CPRO")
```

Measures of association between votat and CP025Q01 .

votat / CP025Q01	0	1	Total
0	631	128	759
1	106	605	711
Total	737	733	1470

Frequency table - Individual level

Construct	Scale	Relat. Freq.	Frequencies	Total N	Chi-squared = 680.83, df = 1, p-value < 0.01
Strategy: votat	0 1	52% 48%	759 711	1470	Phi coefficient = 0.68
Performance: CP025Q01	0 1	50% 50%	737 733	1470	



Module 2: Respondents' actions

Summary of PV1CPRO by VOTAT1-
Individual level

```
# Report: Descriptive statistics by strategy
m2$DescriptiveStrategy(cp025q01.treated,
  "votat",
  "CP025Q01",
  "PV1CPRO")
```

Biserial/Polyserial correlation = 0.7695

Statistics	Total	0	1
Total N	1470	759	711
Min	177.53	177.53	360.15
1st.Qu	426.7	375.57	505.43
Median	493.26	433.19	552.5
Mean	492.37	433.41	555.31
SD	99.28	82.17	74.18
3st.Qu	562.24	487.57	603.64
Max	836.58	688.05	836.58

Stay tuned!

Project website:

<https://www.uv.uio.no/cemo/english/research/projects/log-file-analysis-in-ilsa-logan/>

LOGAN is open for contributions!!

For employees Norwegian website

UiO • Centre for Educational Measurement
Faculty of Educational Sciences

Home Research Studies About CEMO People

Research

Projects

Log-file Analysis in International Large-Scale Assessments (LOGAN)

Norwegian Contact

Project leader: Denise Reis Costa

Participants

Denise Reis Costa
Waldir Leoncio Netto

Detailed list of participants

Involved research groups

FREMO



How do we extract and analyze log data from computer-based assessments?
The LOGAN project has the answer. Photo: Luke Chesser/Unsplash