Lanalytics package

MELANIE STEFAN

SALVADOR GARCIA

Structure of the Focus Group (4 sections):

(5 minutes) Introduction of R, Rstudio and Shiny

(25 minutes) Introduction of the Learning Analytics Package and its Shiny Interface

- There are some questions for each tab of the Interface

(20 minutes) User experience evaluation

- Likert evaluation of the user experience in the dashboard.

(10 minutes) Open questions and suggestions

Introduction to R and Rstudio

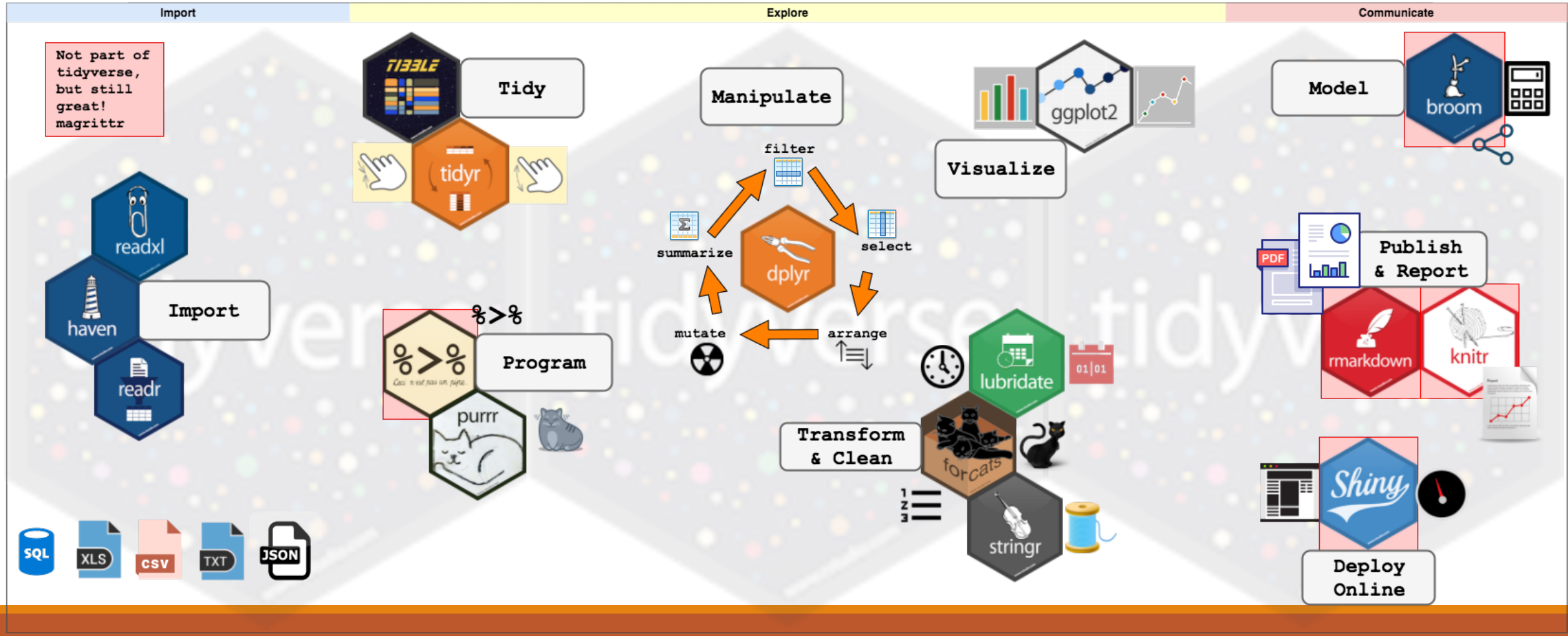
https://cran.r-project.org/

- Have an amazing number of packages related to Psychometrics
- https://cran.r-project.org/web/views/Psychometrics.html
- Well documented, flexible and easy to use
- Free and open source!

https://www.rstudio.com/

-The most common (and really useful) IDE for R

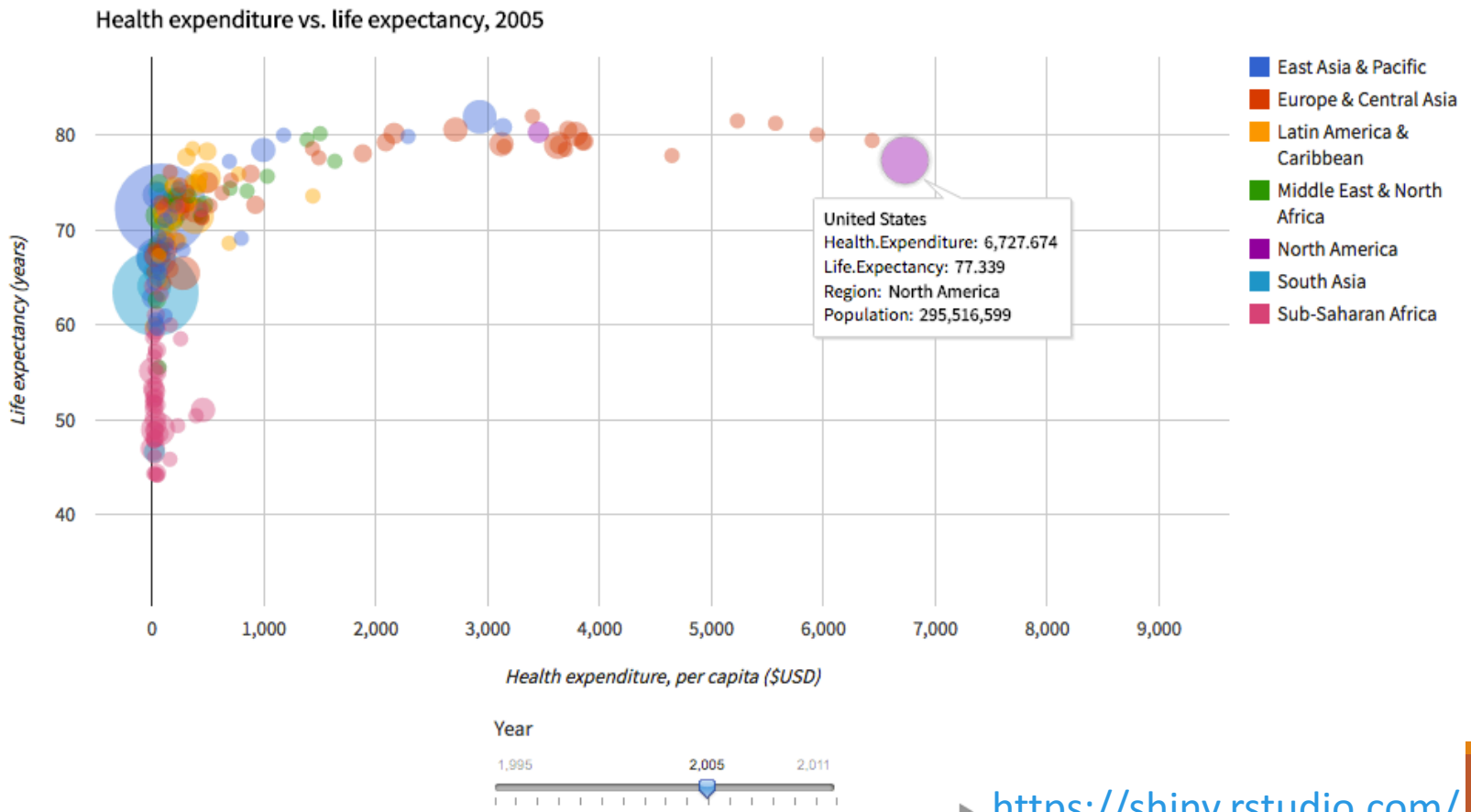
Data analysis in R (tidyverse)



https://www.tidyverse.org/

R Shiny

Google Charts demo



Basically 2 reasons to use R over other interactive visualization software:

- Can work with any package of R!
- Not directly code in HTML

- https://shiny.rstudio.com/

Lanalytics package

Objectives of Learning Analytics Package and Dashboard

1) Understand the items in low-stake online quizzes to improve the design of future quiz items. (Number of questions, time of quiz, easiness).
2) Monitor the performance of the students in different levels: individual, per quiz and per group

Basic Structure of package:

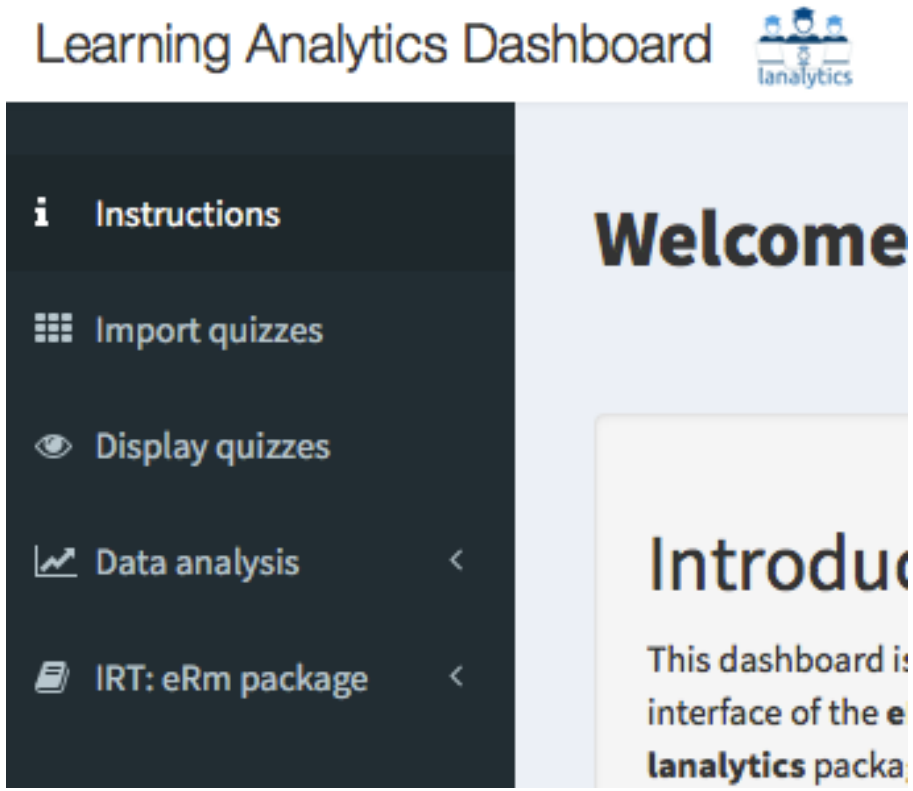
- **Quiz object** (long format table with an entry
- per student-question)

- **Course object** (Grouped quiz objects in one object)

	email.address	question	responded.at	score
1	ADRXM8548M@gmail.com	1	2017-02-11 22:11:46	1
2	ADRXM8548M@gmail.com	2	2017-02-11 22:12:00	1
3	ADRXM8548M@gmail.com	3	2017-02-11 22:12:16	0
4	ADRXM8548M@gmail.com	4	2017-02-11 22:12:41	1
5	ADRXM8548M@gmail.com	5	2017-02-11 22:13:01	1
6	ADRXM8548M@gmail.com	6	2017-02-11 22:13:49	1

	quiz	order.answer	time.per.question
1	datasets/sample_dataset/Q01.csv	1	NA secs
2	datasets/sample_dataset/Q01.csv	2	14 secs
3	datasets/sample_dataset/Q01.csv	3	16 secs
4	datasets/sample_dataset/Q01.csv	4	25 secs
5	datasets/sample_dataset/Q01.csv	5	20 secs
6	datasets/sample_dataset/Q01.csv	6	48 secs

Lanalytics dashboard

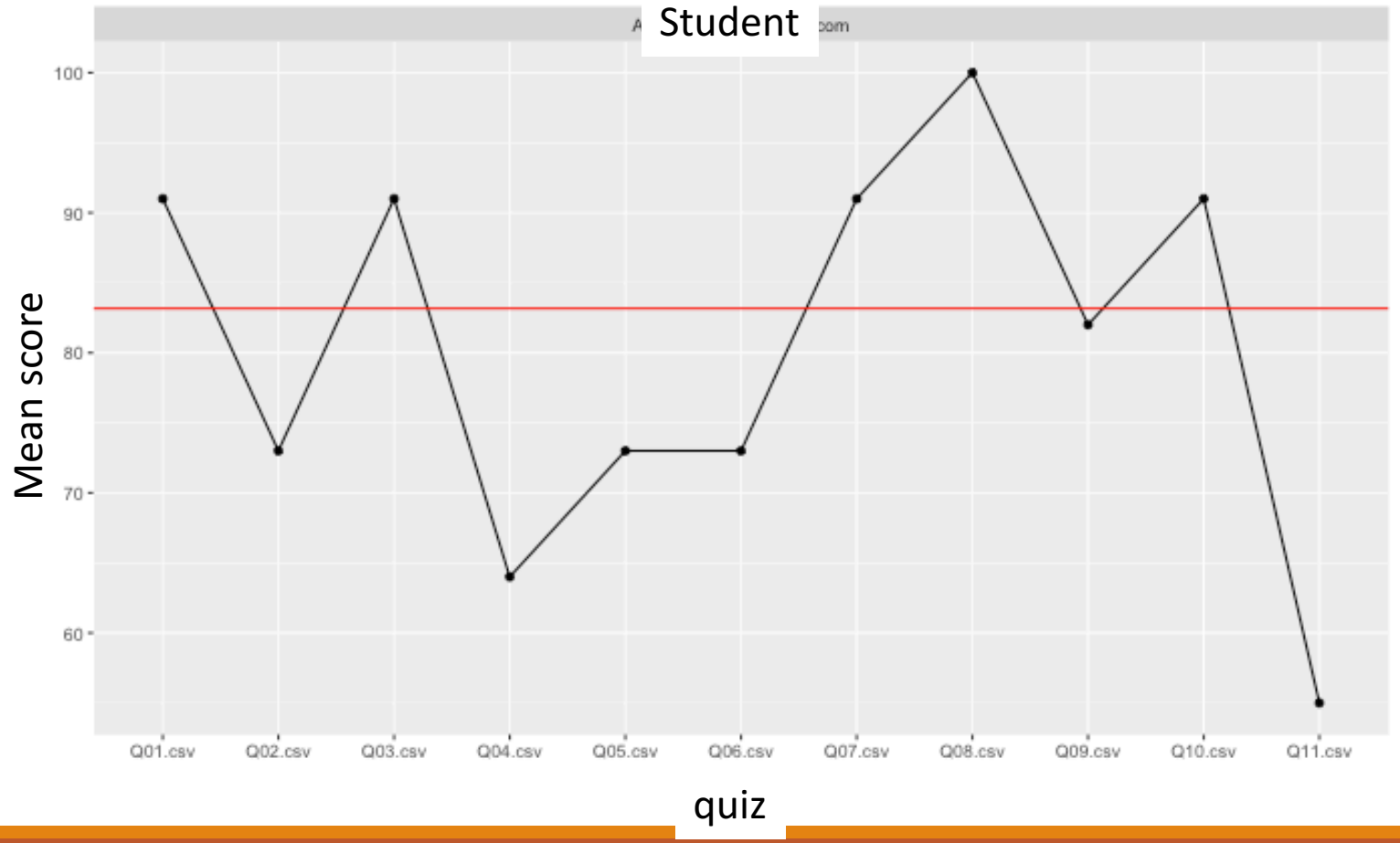


- **5 Tabs to make analysis and display plots**

- Performs analysis at distinct levels: individual, group and quiz.

- Uses **rasch models** (eRm package) to infer a person parameter and an item parameter.
  - The person parameter measure the ability of the person
  - The item parameter measure the difficulty of the item
  - Both are measured in the same scale (a latent variable). That means that we compare them!

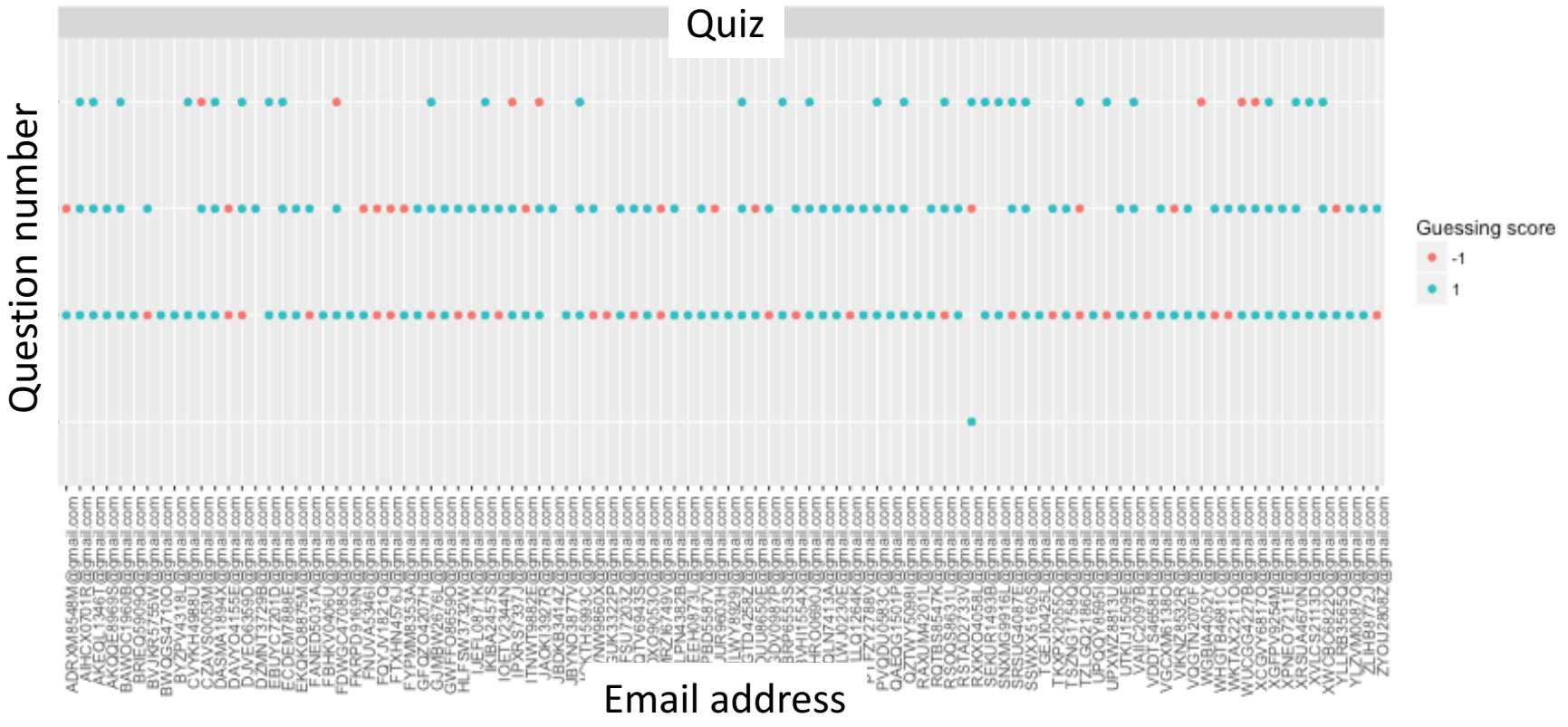
Individual analysis



- You can filter one student email and see its performance in all quizzes
- You can monitor which students are not making progress or which are improving their scores.
- You can upload a file with the final grade of the exam and it can be shown like the red line in this example

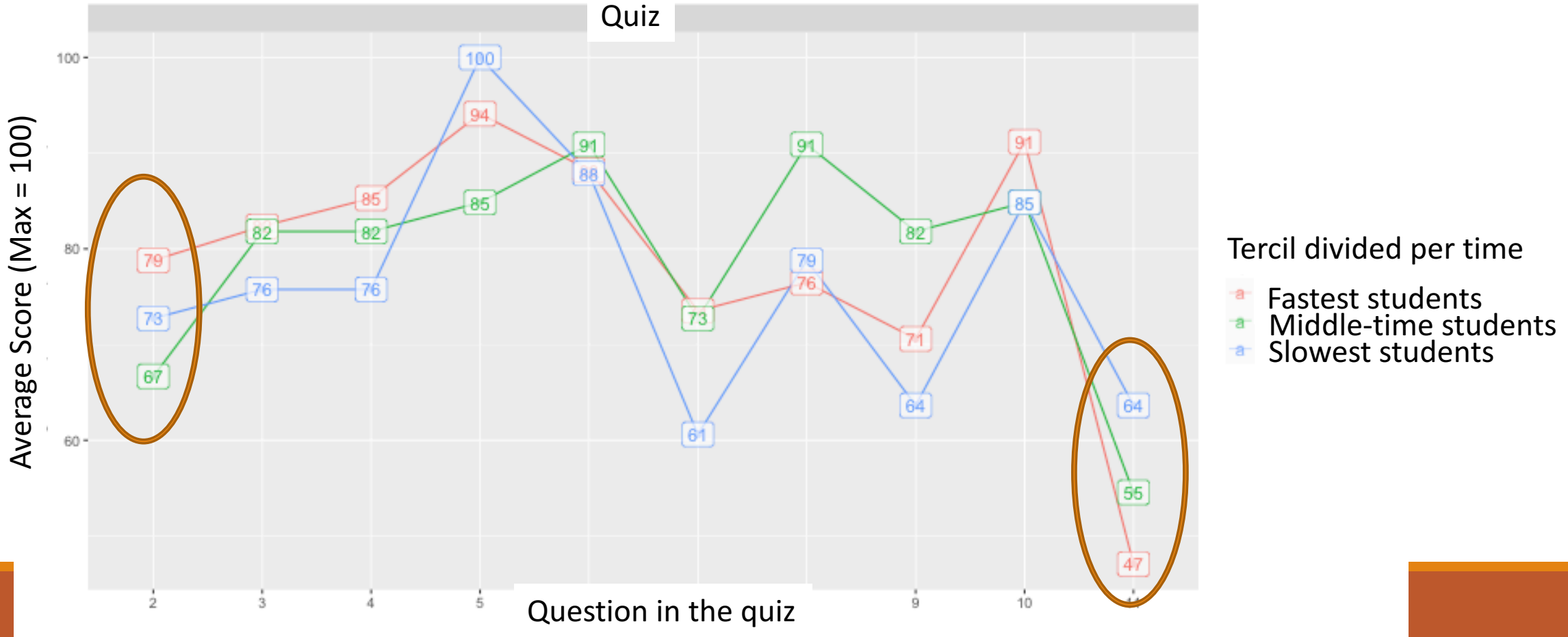
Group Analysis

- Sometimes a student answers a question very fast is suspicious!
- The probability that he gets this question wrong should be is high. But what happen in he gets many correct answers answering very fast the questions?
- In this plot the questions that are answered in less than 20 seconds (or below an individual threshold set which is set by answering really questions) is shown.
- 1 means that he answer the question correctly and -1 means that he answer the question incorrectly.



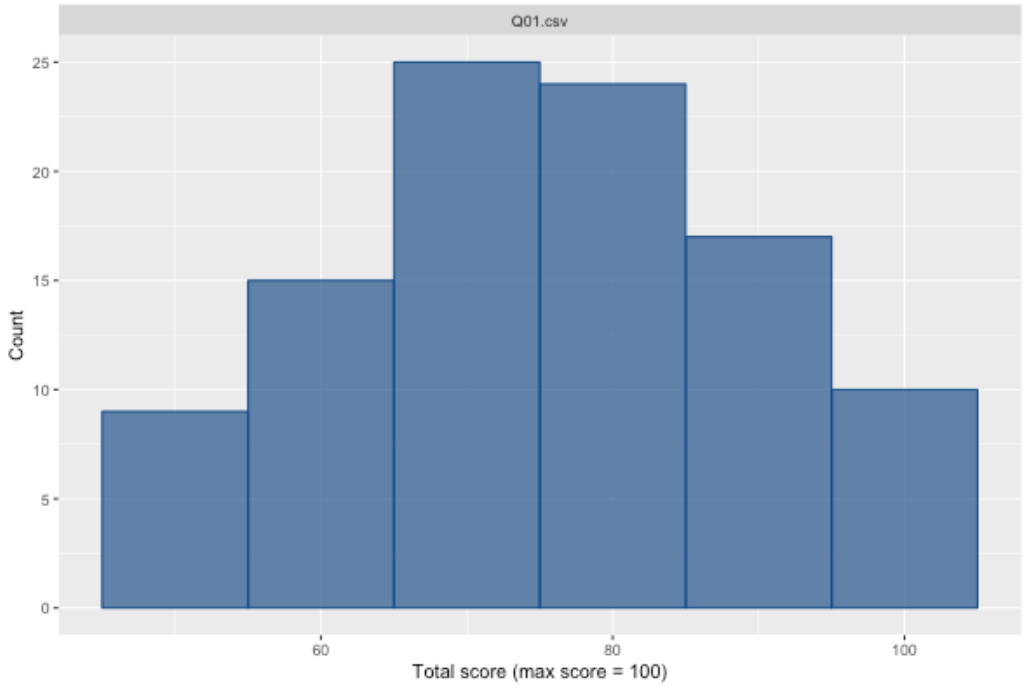
Group Analysis

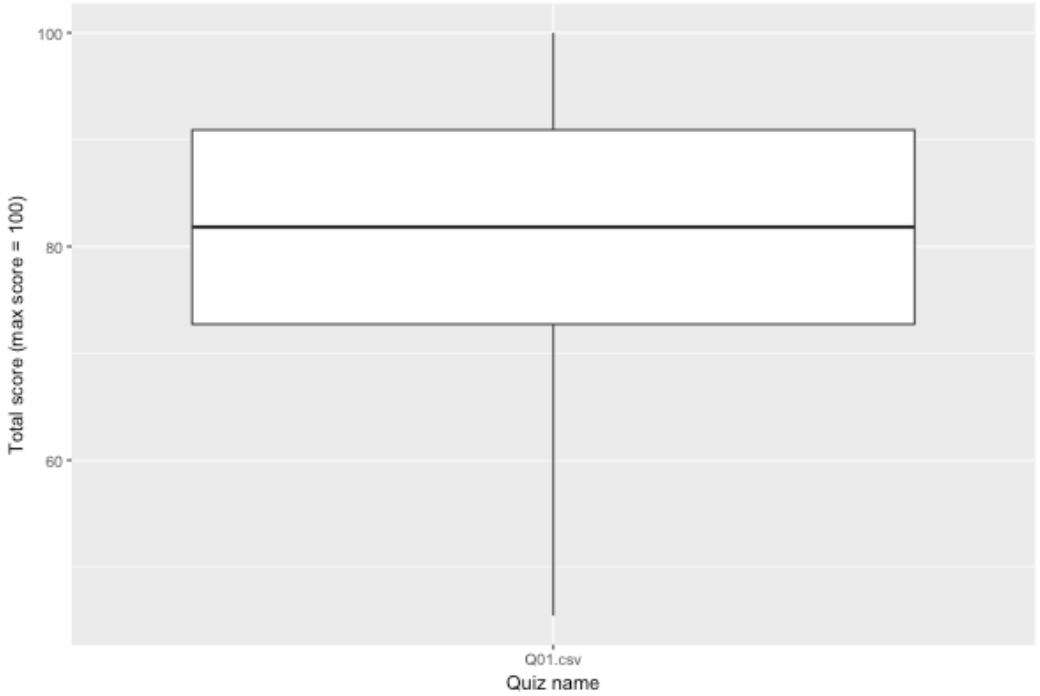
- The quizzes are designed for an specific time
- Students are taking too long to answer each question?
- This extra time really help them to answer correct the question?
- Or they take more time because they don't know the answer?
- Analyzing the spent time vs the average score should give us an idea.



Quiz analysis

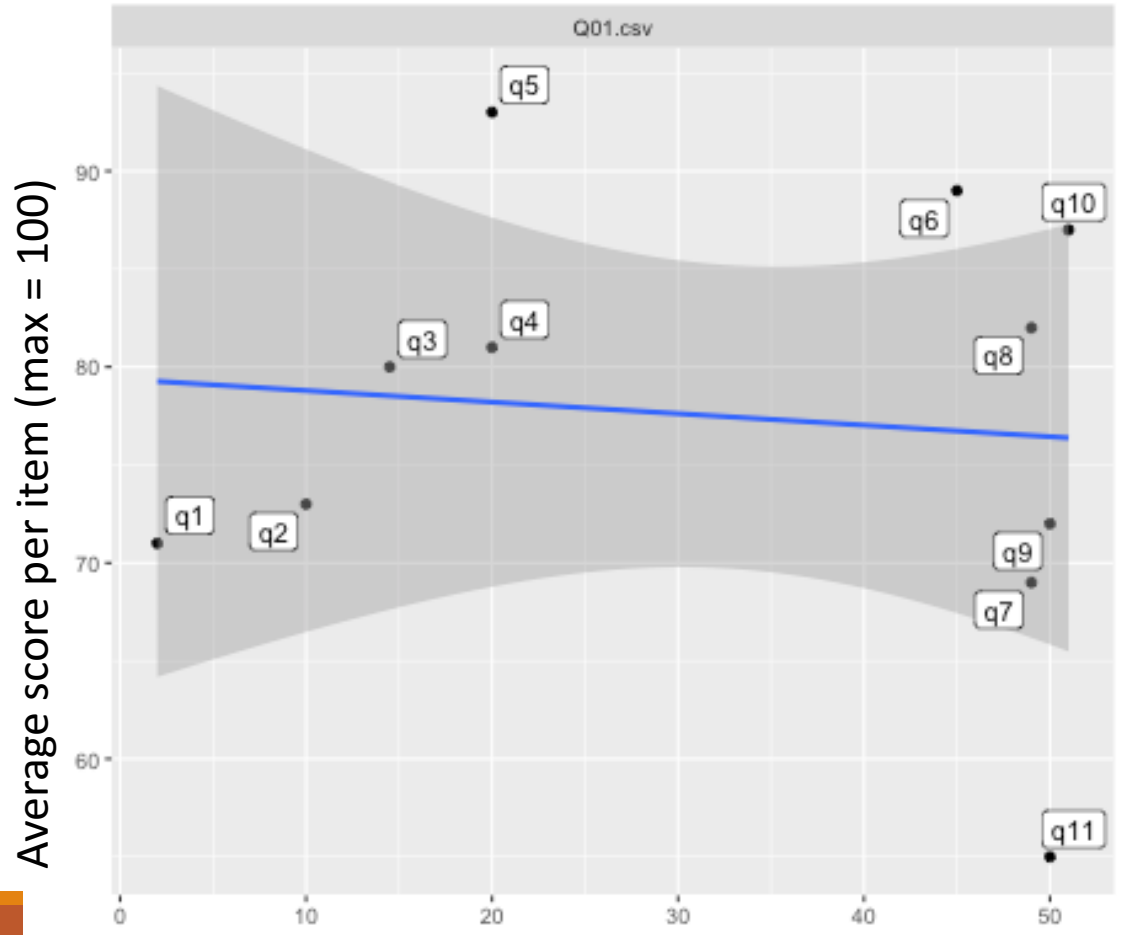
- Show basic descriptive plots like histograms and boxplots.
- Can study the dispersion, skewness and outliers.

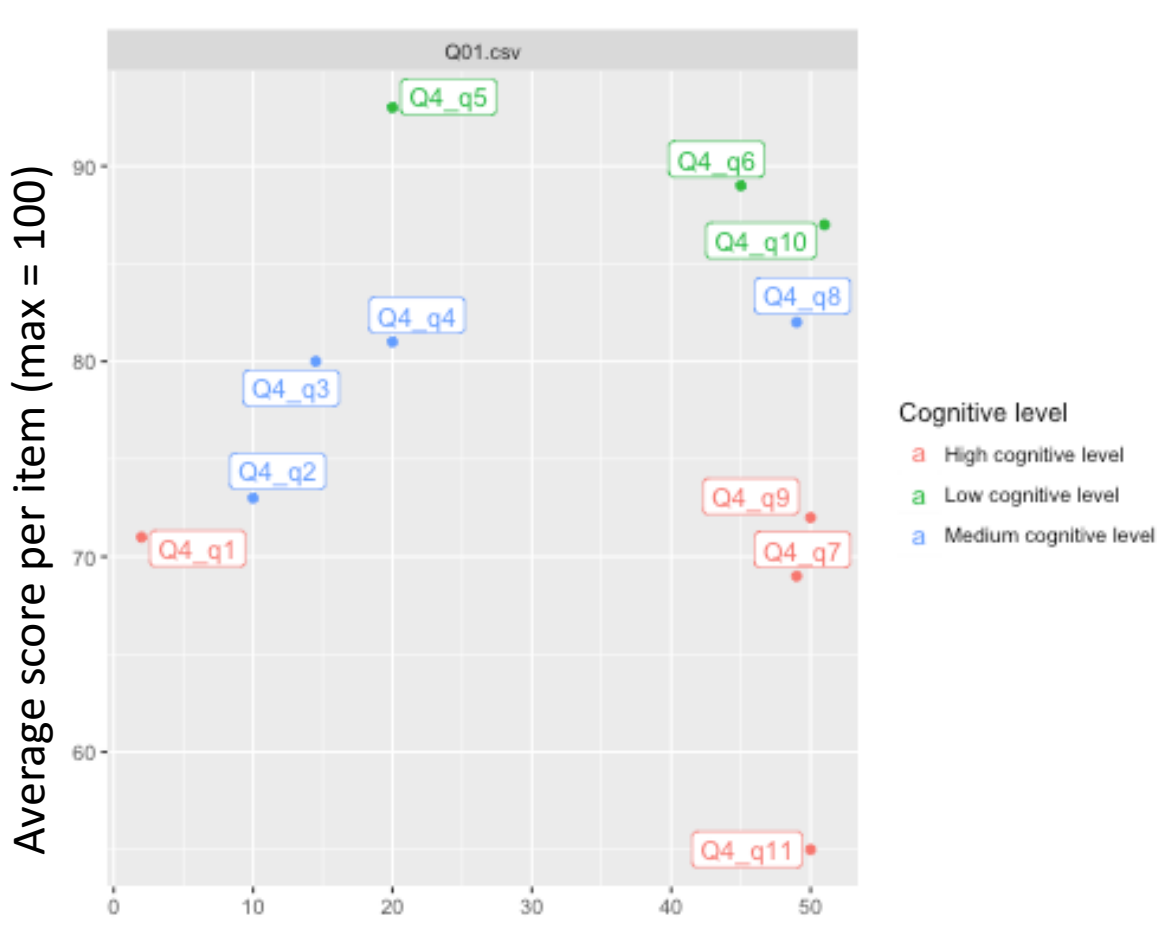




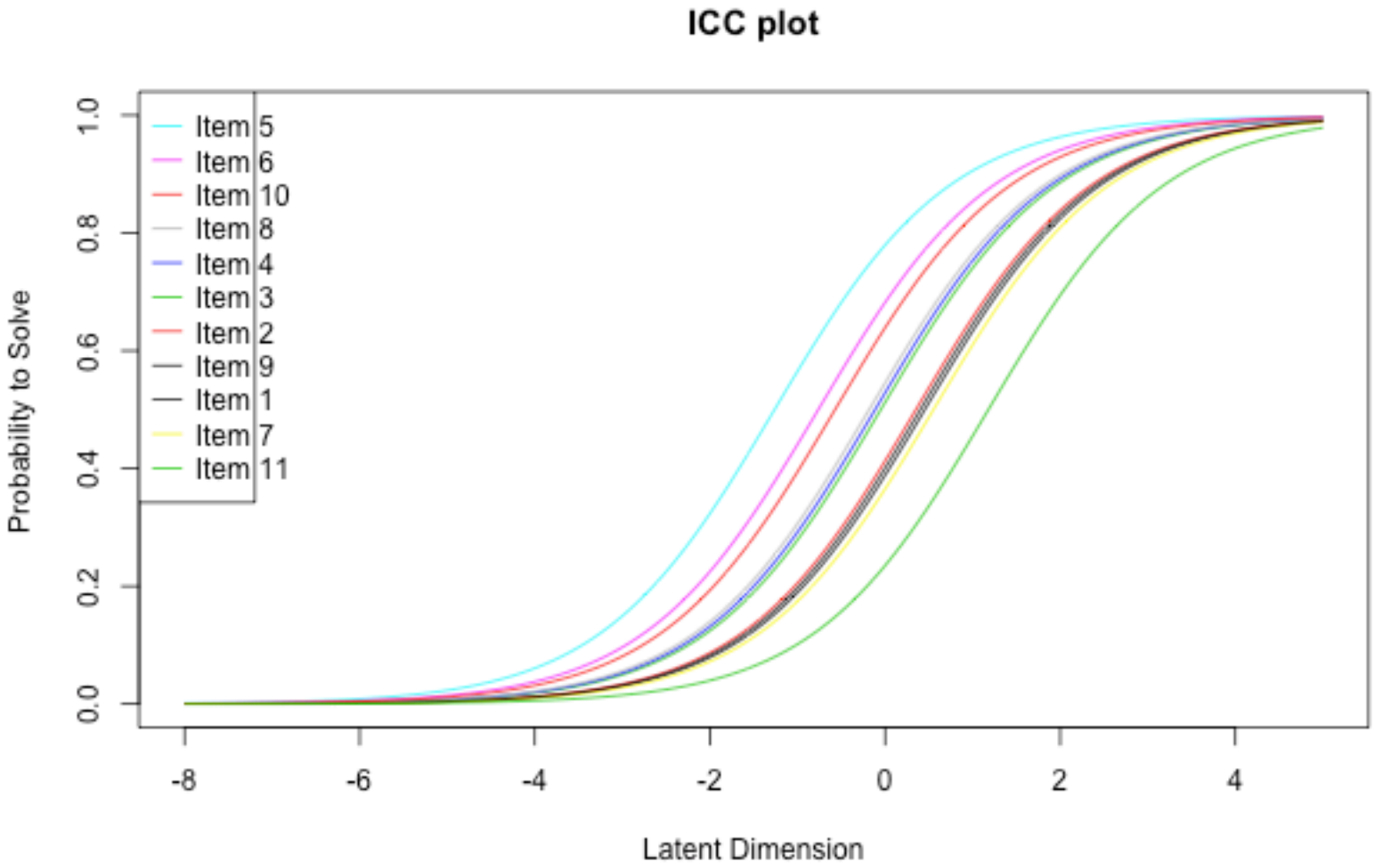
Quiz analysis

- Easiness-Time plot shows relation between the spent time in each question versus the average score. (Easy questions are really answered faster?)
- Easiness-Time-Level plot shows the relation between cognitive level (high cognitive items takes more time!



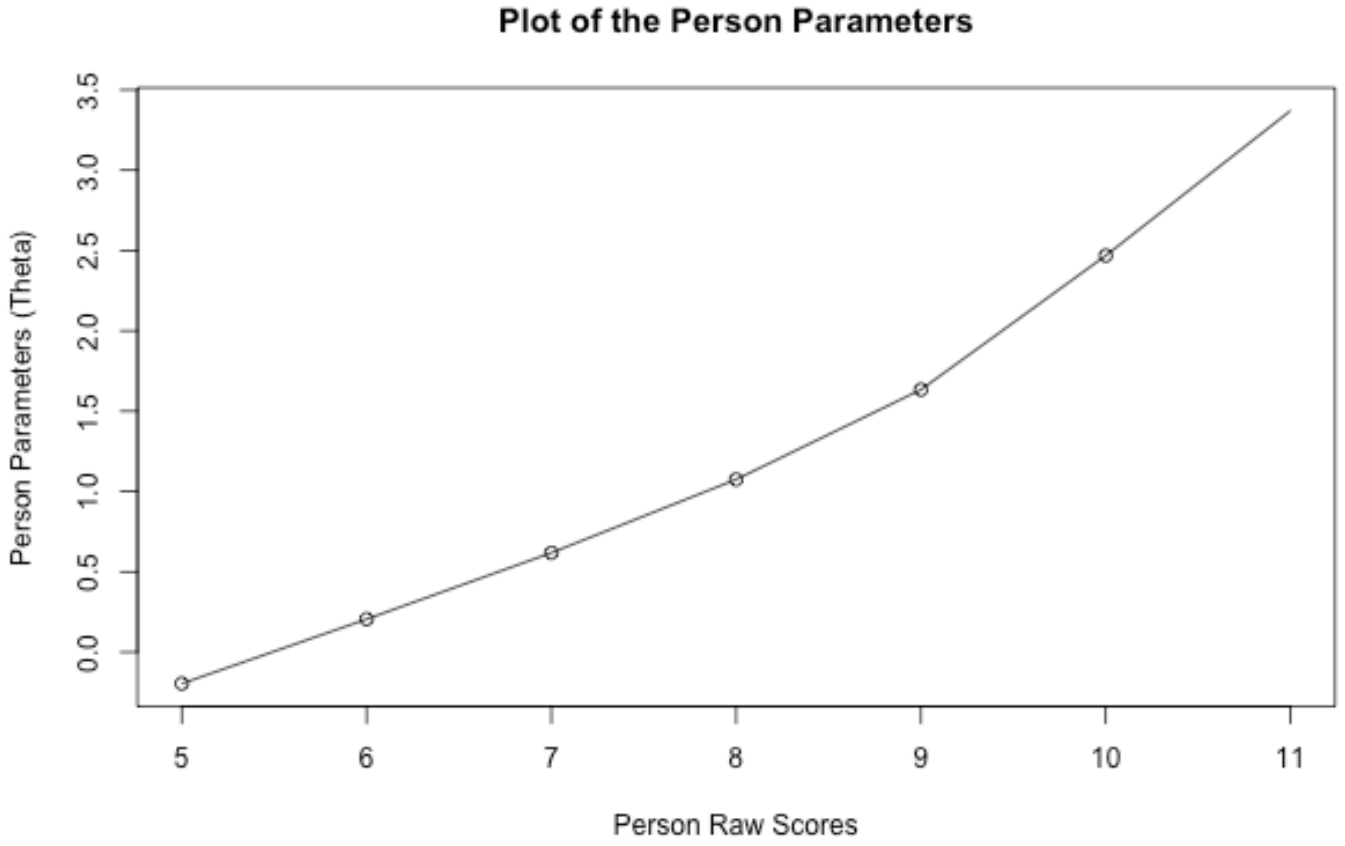


Rasch model - Item Characteristic Curve



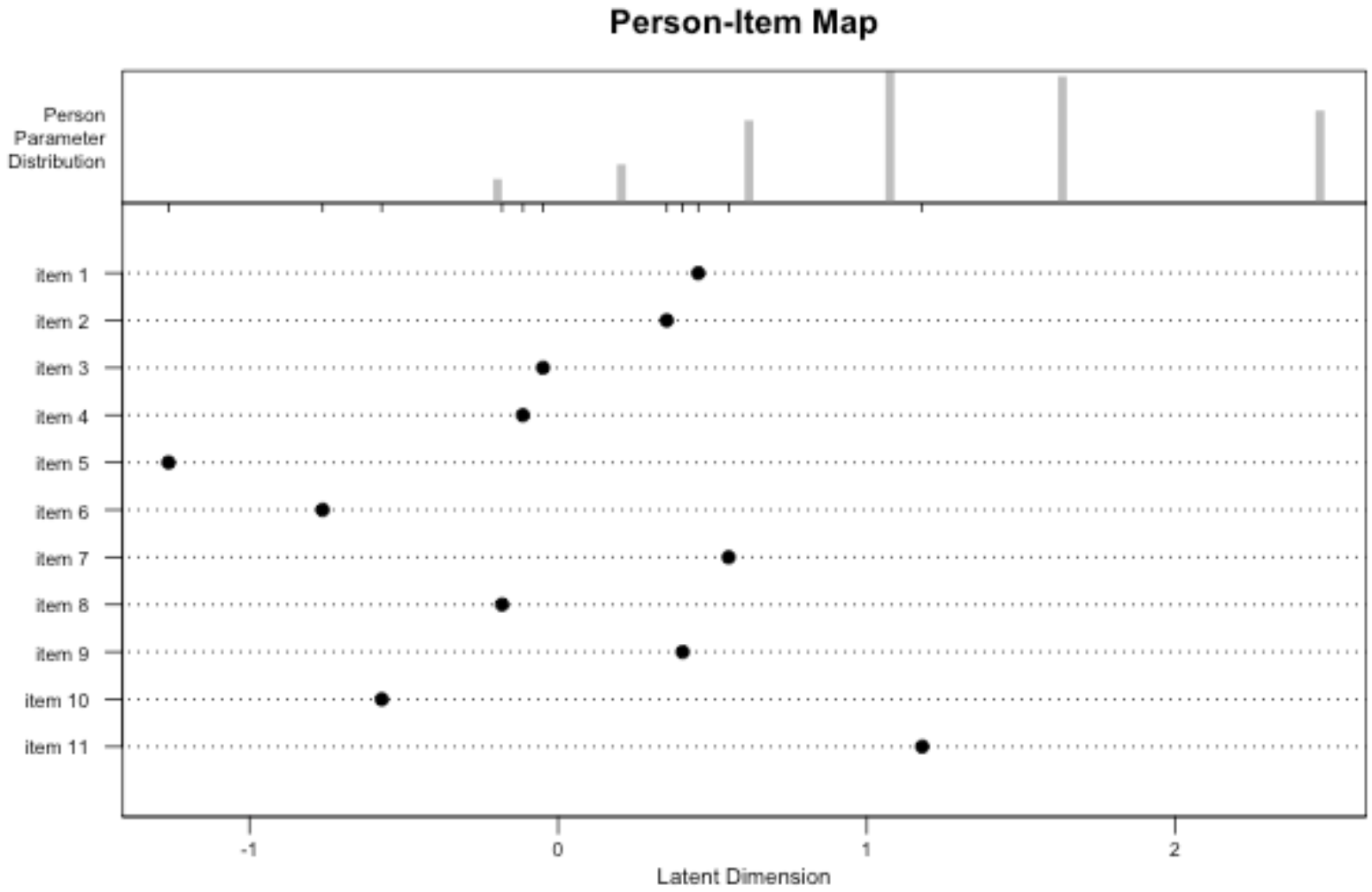
- A way to display the item parameter (difficulty) is to plot it as a sigmoid function. (slope is the parameter)
- The **latent dimension** is sometimes called ability.
- It shows the probability of getting correct the item
- For example, item 5 is the easier in this set (for low levels of ability, it has high probability to be correct)

Rasch model - person parameter



- The raw scores are the total number of correct answers per person.
- The person parameter is in the same scale as the item parameter (**latent dimension)**
- For example, this quiz is skewed to have more correct answers (there are people who have all questions correct, but there are no one who got 0 questions correct)

Rasch model - person item map



Distribution of person parameter

- Person parameters are high compared with the difficulty of the items.
- The dataset was designed so that minimum there is 50% probability to get correct the items.

Value of item parameter

- Easy items compared with the abilities of the students
- This example is of open book questions, so it's normal to get this.
- But, depending on the type of quiz, this might change