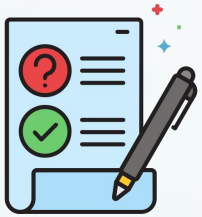
Visualizing Cronbach's Alpha for a Large Number of Assessments

Doing data-driven assessment improvement

Aditya Sharma, Playpower Labs
Nirmal Patel, Playpower Labs

Data-Driven Assessment Improvement



1
Use assessments to measure student knowledge

3
Improve assessments using data and **<u>human intervention</u>**

2
Collect and analyze assessment data to find possible improvements in assessment quality

Assessments

Assessments or tests give us a measure of student knowledge

Student scores on assessments are high stakes and are used for important decisions, so it is paramount to have high quality assessments

How do we know whether our assessments are high quality? We ensure that they are reliable and valid. What does that mean?

- **Reliability**: The assessment will give us a consistent measures of student knowledge (i.e. will not contain random errors over repeated administrations)
- **Validity**: The assessment will measure the construct it is supposed to measure

Unreliable & Unvalid

Unreliable, But Valid

Reliable, Not Valid

Both Reliable & Valid

Cronbach’s Alpha

Widely used metric to assess the reliability of a test

Tells us how well the items within a test covary

A measure between 0 and 1, 0 meaning that items of the test are entirely independent and 1 meaning that items of the test have high covariance

High alpha is generally good, low alpha is bad

Common to consider alpha of 0.7 and above as acceptable

Cronbach's alpha	Internal consistency
$0.9 \leq \alpha$	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	Questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	Unacceptable

https://en.wikipedia.org/wiki/Cronbach%27s_alpha

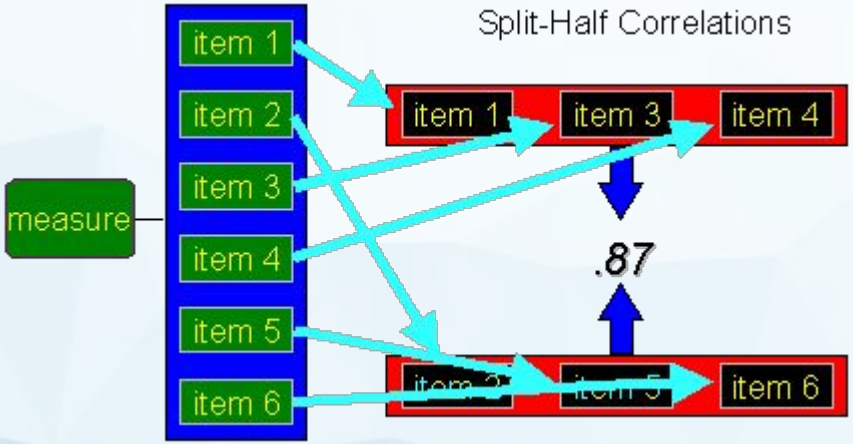
Example

Let us say we have a test with 6 questions, and we administer it to a number of students. At the end of the test, we get a dataset like the one below:

Student	item1	item2	item3	item4	item5	item6
S1	1	1	0	1	0	1
S2	1	1	1	0	1	1
S3	0	0	1	0	0	1
S4	0	1	1	0	0	0
...	...	...	...	...	...	...

Split Half Reliability

Cronbach's Alpha is average of all possible split half reliabilities of a test. What is a split half reliability of a test?

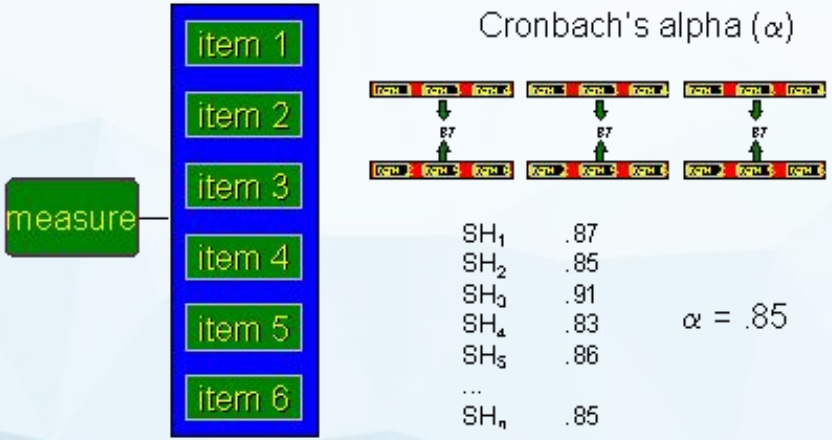


Split half reliability of a test is calculated by randomly dividing items of a test in two halves, calculating student scores for each halves, and finding the correlation of those scores.

https://socialresearchmethods.net/kb/reltypes.php

Cronbach’s Alpha

Theoretically, Cronbach’s Alpha is average of all possible split half reliability correlations. Although, the mathematical formulas don’t reflect this directly.



https://socialresearchmethods.net/kb/reltypes.php

Suppose there are K items in the test (Y_1 to Y_k) and the their scores add up to a total score of X. Then, Cronbach’s Alpha for that test is calculated using the following formula:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} \sigma_{Y_i}^2}{\sigma_X^2}\right)$$

https://en.wikipedia.org/wiki/Cronbach%27s_alpha

Using Alpha to Improve Assessments

There are multiple ways to use Cronbach's Alpha to help improve assessment quality:

1. **Find assessments with low Alpha**: We simply calculate Alpha for a set of assessments and investigate into assessments that have questionable measure of Alpha.
2. **Find assessment items that are 'causing' the Alpha to decrease**: For assessments with questionable Alpha, we can calculate Alpha of the assessments after removing each item and see whether removal of an item leads to increase in the Alpha of the assessment.

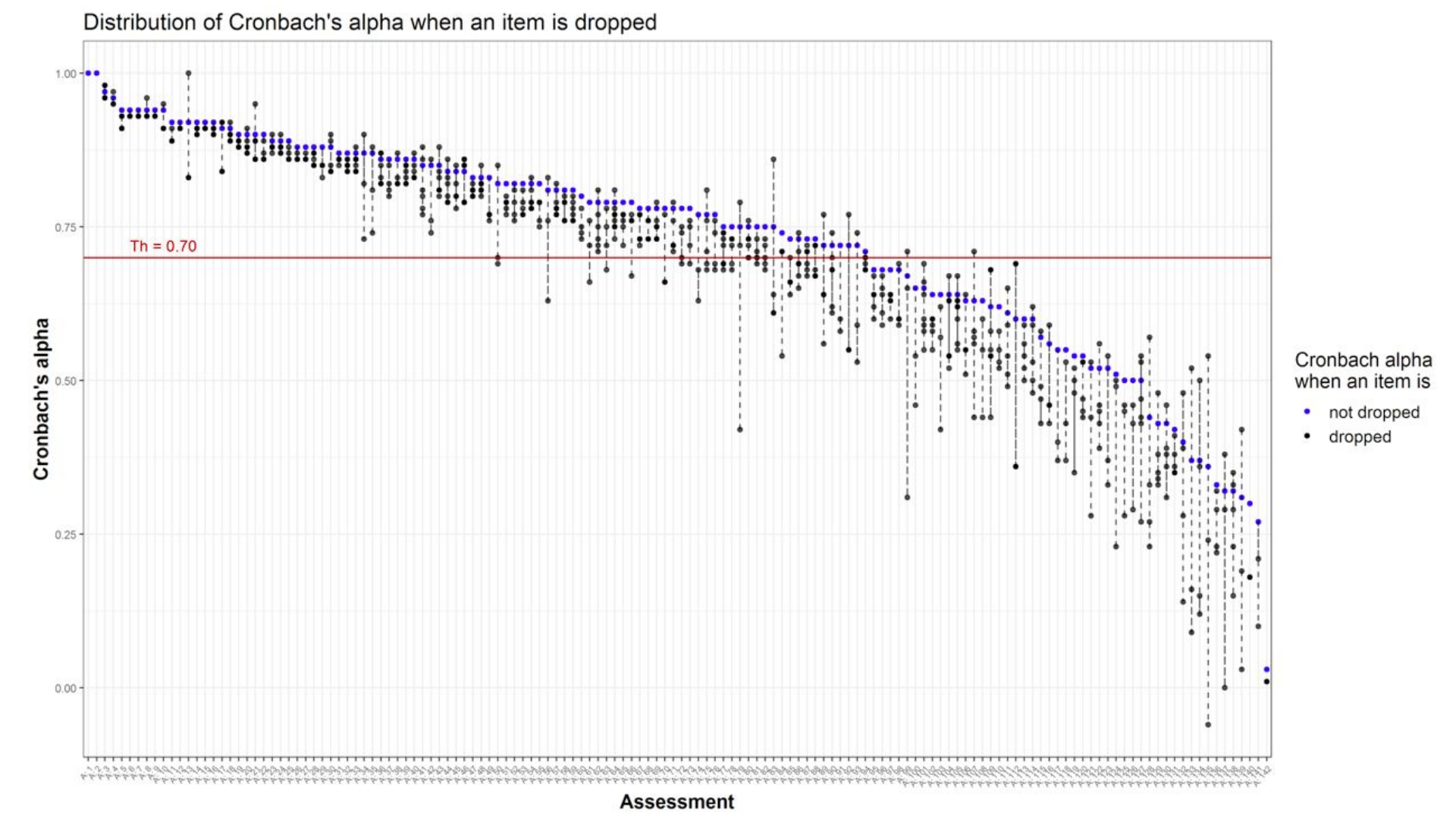
Big Data

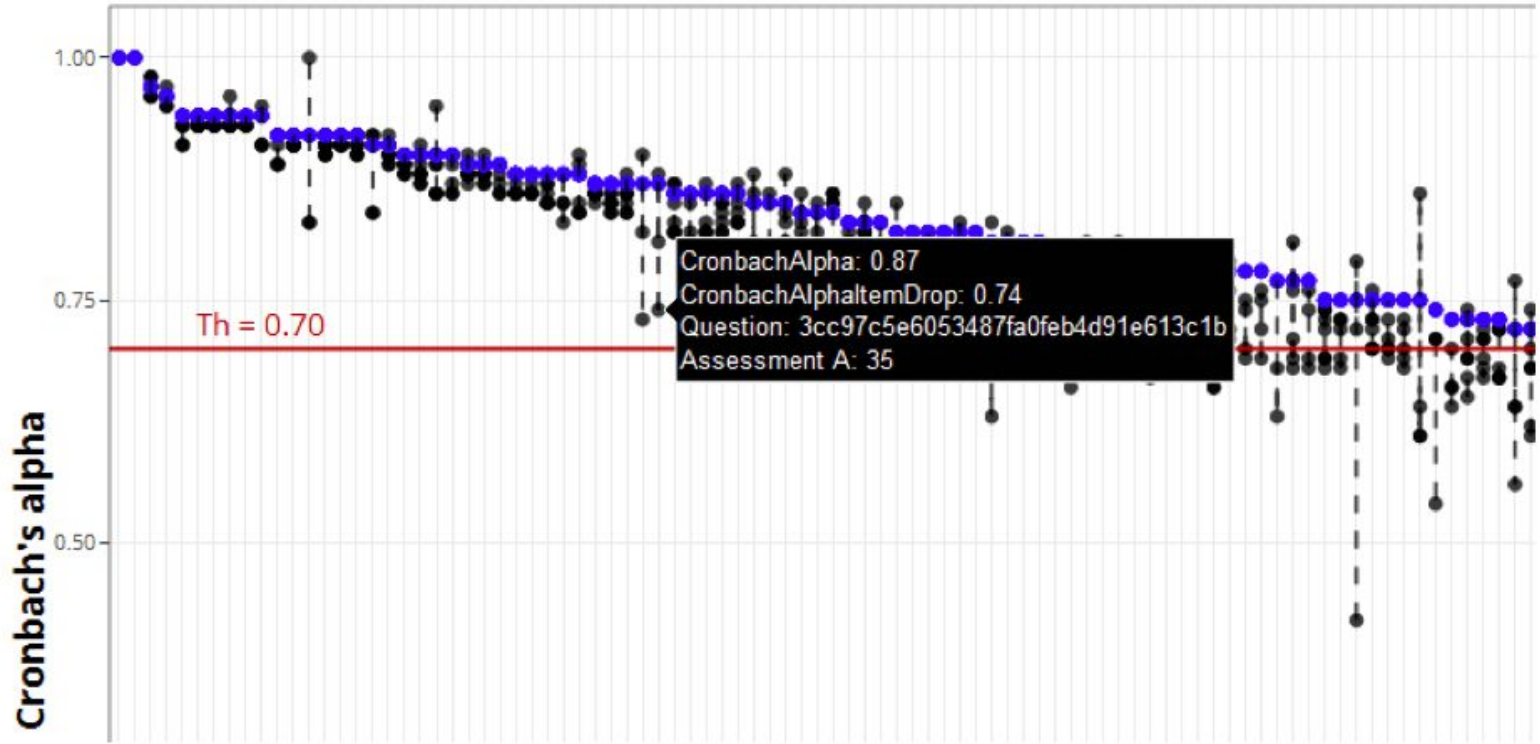
If you have data from a large number of assessments, you will find that digging through data using tables is difficult.

Getting an overall picture of assessment quality using numeric tables is hard, and doesn't allow you to **quickly identify trends and communicate information to others**.

That is why we created a data visualization that plots assessment quality metrics of many assessments together.

We plot overall alpha and alpha after item removals of many assessments, all at once!


Distribution of Cronbach's alpha when an item is dropped
Th = 0.70
Cronbach's alpha
1.00
0.75
0.50
0.25
0.00
Assessment
Cronbach alpha when an item is
not dropped
dropped

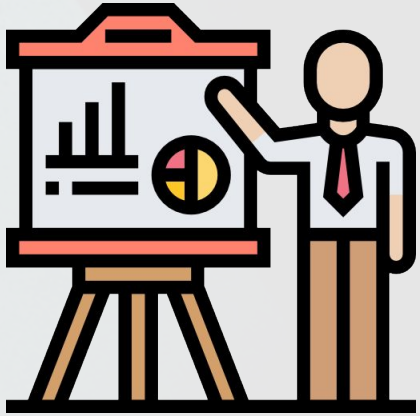

1.00
0.75
0.50
Cronbach's alpha
Th = 0.70
CronbachAlpha: 0.87
CronbachAlphaItemDrop: 0.74
Question: 3cc97c5e6053487fa0feb4d91e613c1b
Assessment A: 35

Purpose of the Visualization

Data visualizations are an important part of the overall data-storytelling process. The presented visualization makes it easy to identify trends and patterns in the assessment data.

To send a list of items that need correction or removal, data in a tabular format is ideal. But to make people ‘see’ what is going on with the assessments in a program, visualization can become handy.

The visualization is also a useful tool for exploratory data analysis.



Code

Please find the code of the visualization here:

https://github.com/nirmalpatel/edviz-2019