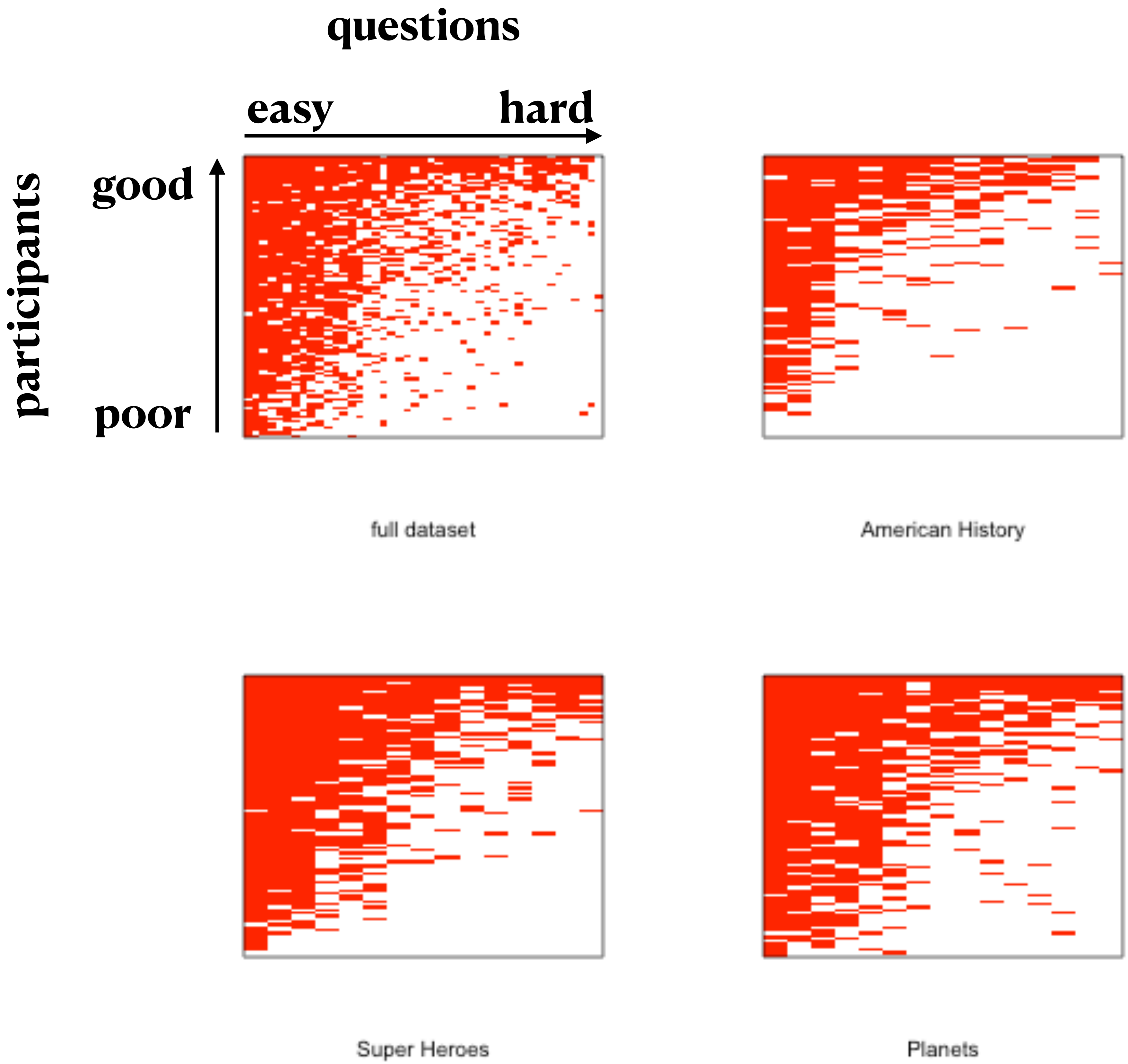
project 'trivia': results

the intuition we started from

Hard questions are almost only answered right by high-performing subjects, so a right answer to a hard question is a good predictor of right answers to easier ones.

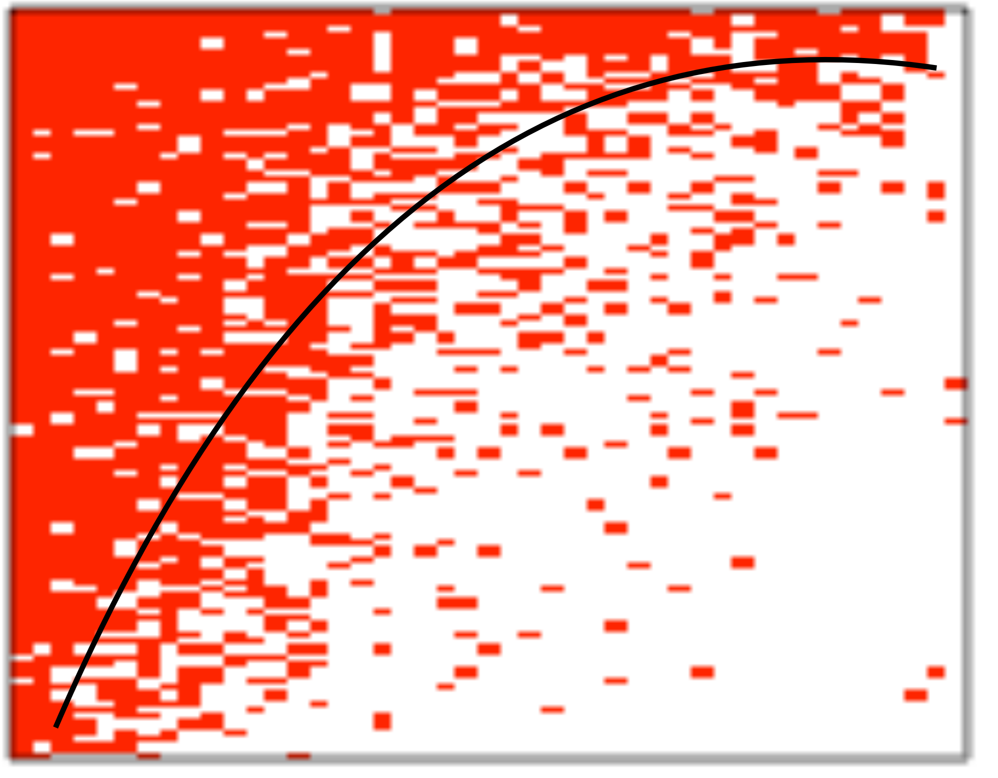
raw data


questions
easy
hard
participants
good
poor
full dataset
American History
Super Heroes
Planets

nestedness measures

NODF: finds how many rows/columns pairs where the below/right row/column is less full, counts that, and inside those pairs, counts the full cells that overlap.

Temperature: simulates a perfectly nested matrix, with a sharp dividing line, then counts the cells that are unexpectedly full or empty relative to that. A surprising cell counts more the further it is from the divide.
low temperature = high nestedness



permutation methods

Starting from the original matrix, they generate permuted binary matrices, keeping equal...

	r1	**curveball**
matrix size (n of questions & participants)	✓	✓
matrix fill (overall n of good answers)	✓	✓
column fill (each individual question's difficulty)	✓	✓
row fill (each individual participant's success)	✗	✓

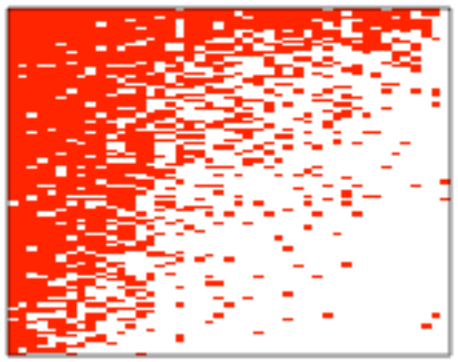
in other words

	r1	**curveball**
Keeping the individual player performances identical to the real data	✗ no constraint at all on variance	✓ exact same level of perf for every individual
Keeping everything else equal	✓	✓

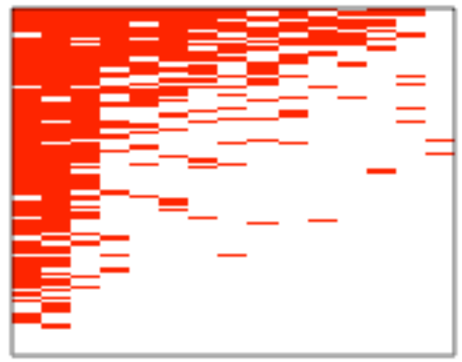
the intuition we started from

Hard questions are almost only answered right by high-performing subjects, so a right answer to a hard question is a good predictor of right answers to easier ones.

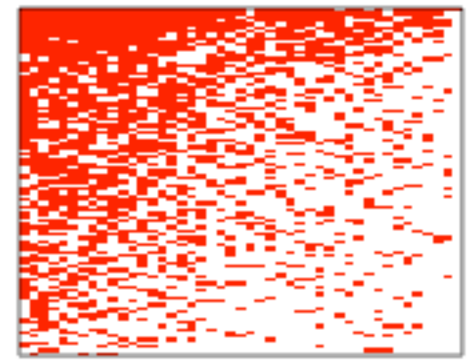
randomised matrices — method r1



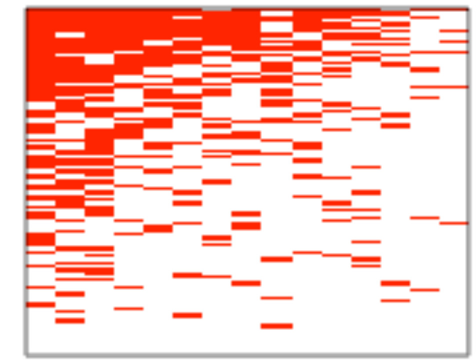
full dataset



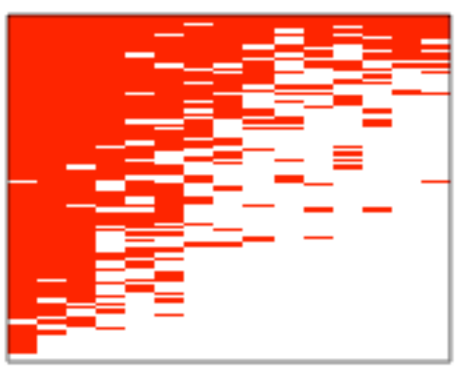
American History



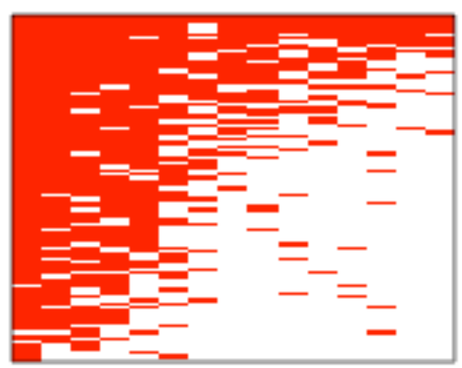
full dataset, r1



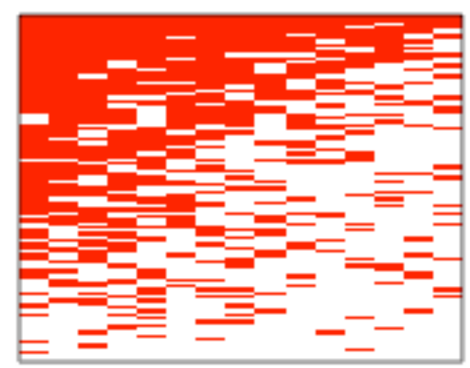
American History, r1



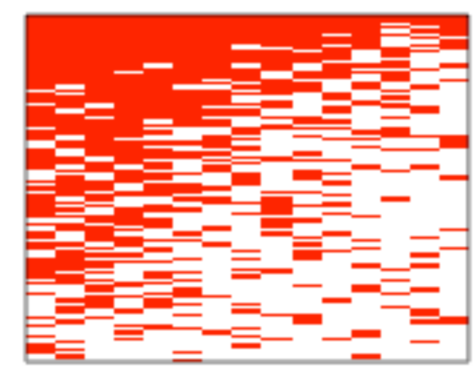
Super Heroes



Planets



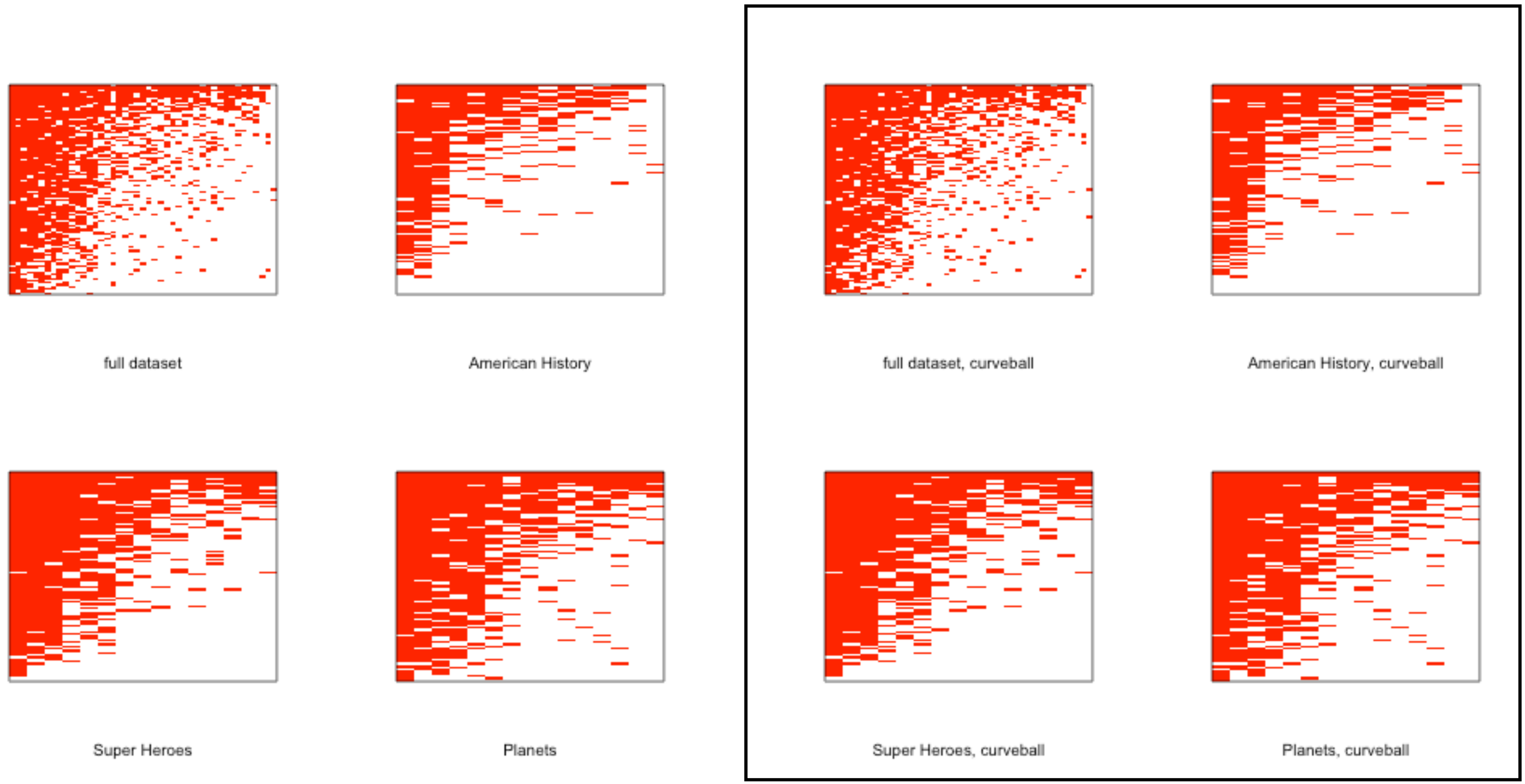
Super Heroes, r1



Planets, r1

(each matrix is sorted with best participant / easiest question on top/left)

randomised matrices — method curveball



(each matrix is sorted with best participant / easiest question on top/left)

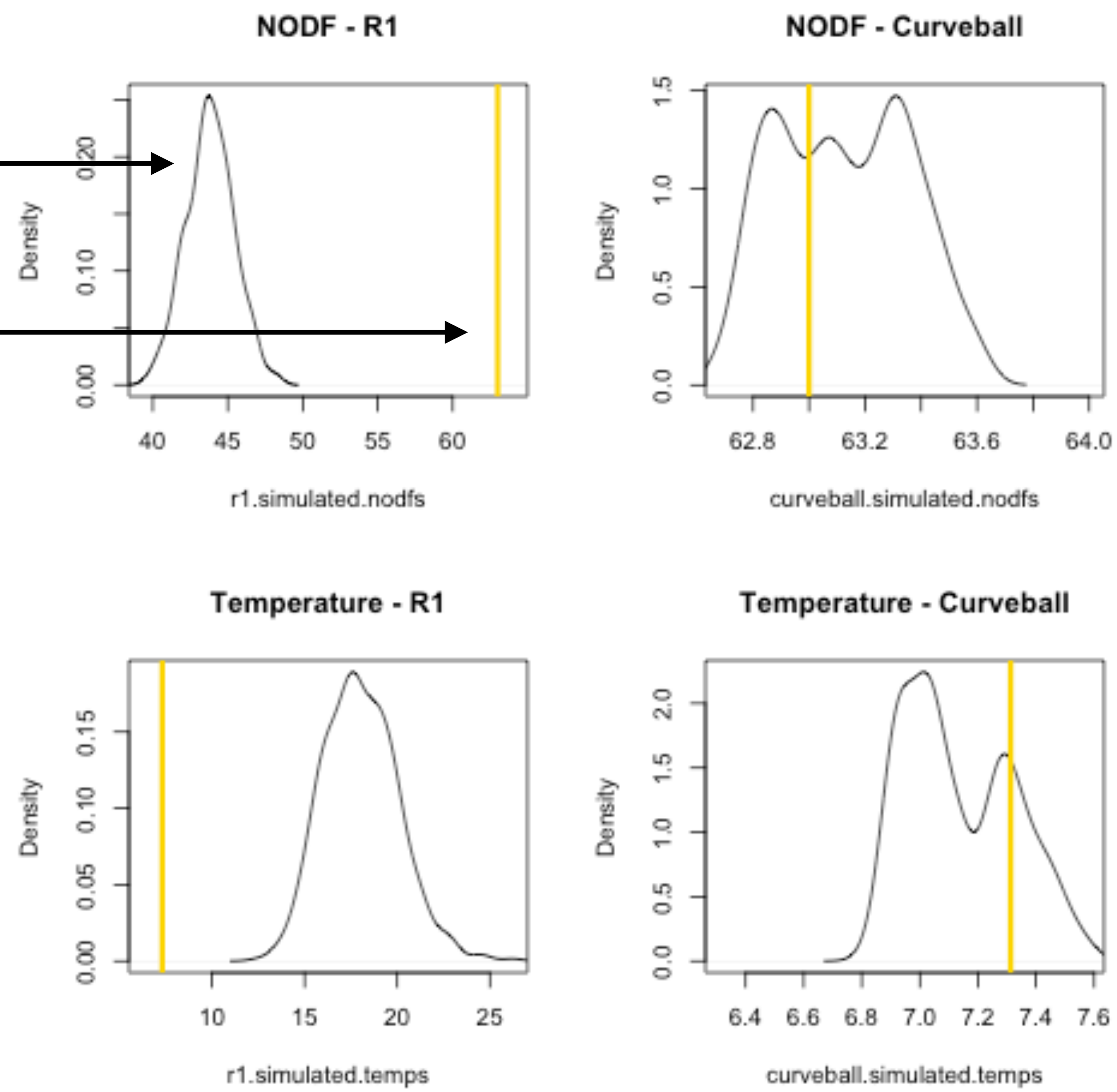
results — summary

	r1	curveball
NODF	✓	✗
Temperature	✓	✗

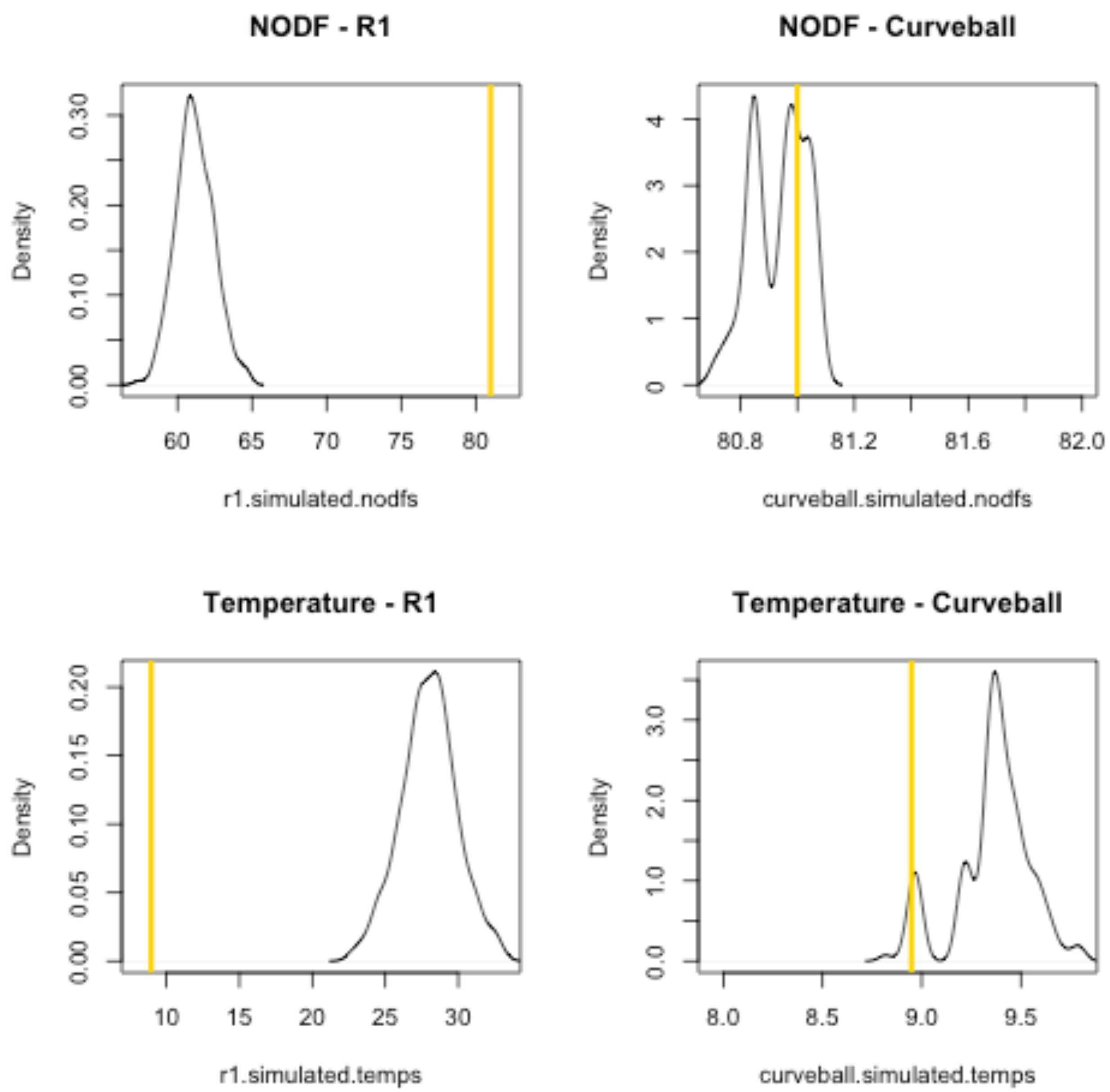
History

density curve for the simulated matrices' values

actual data value

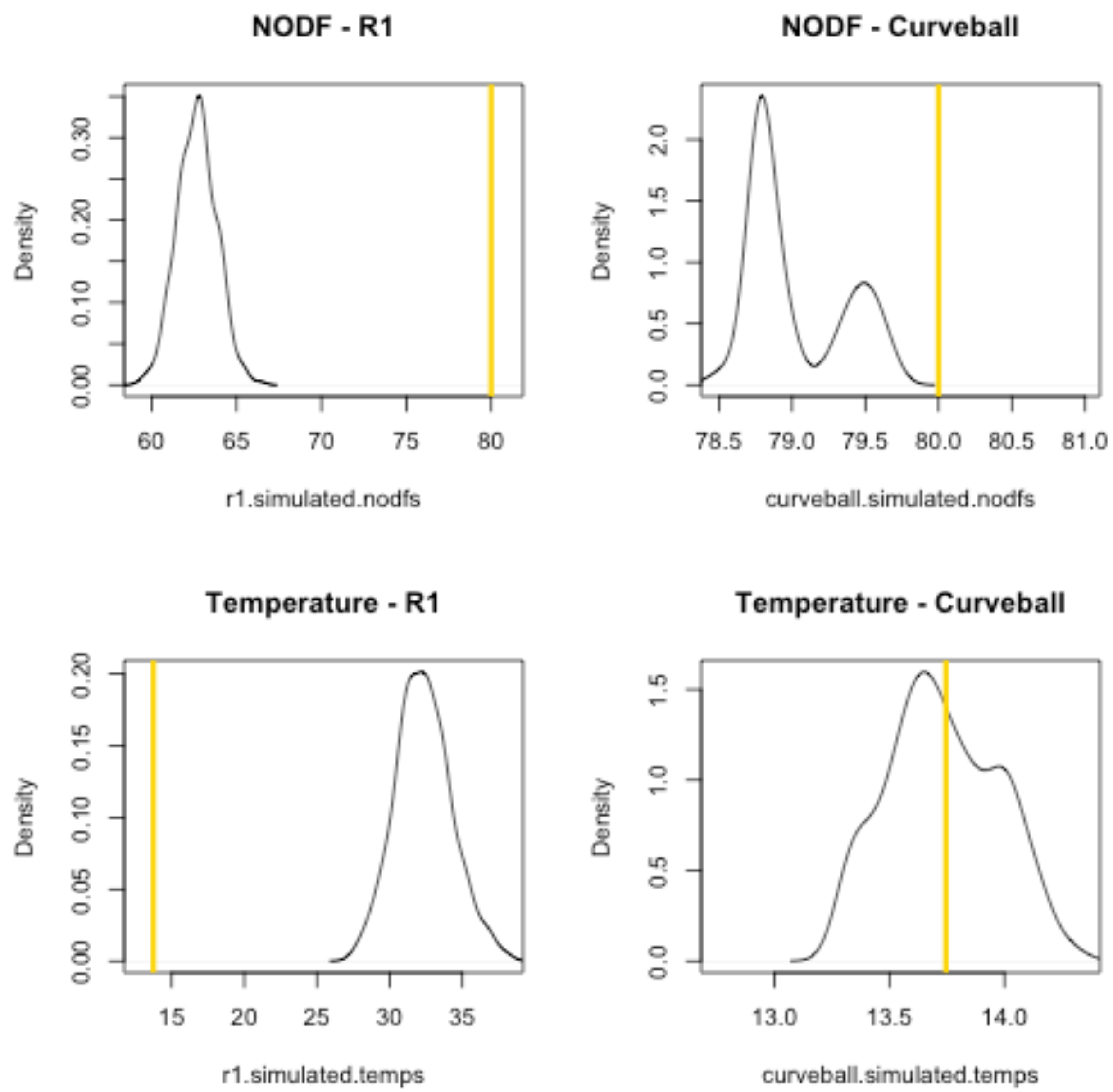


Super Heroes



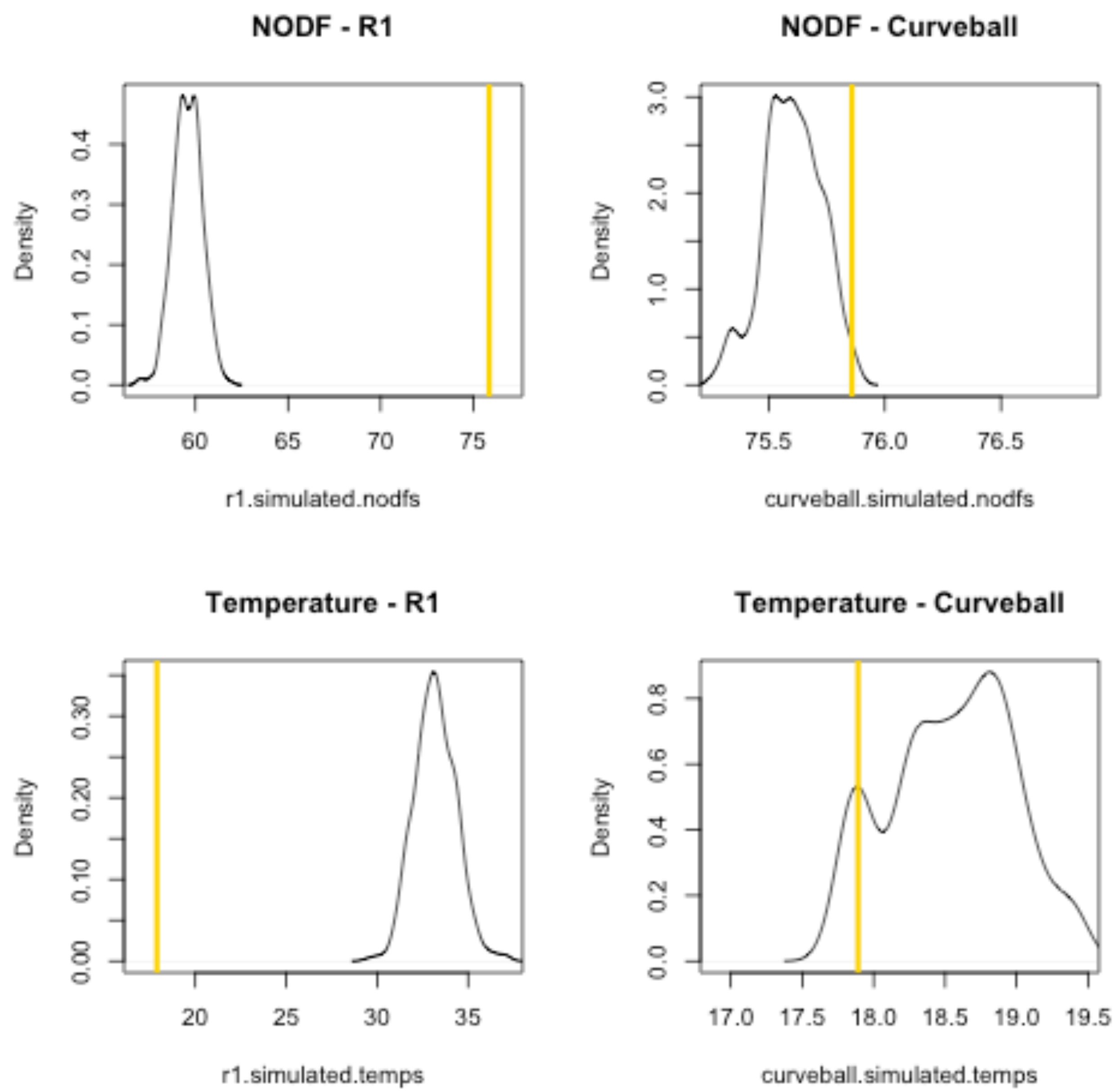
this one: $p = 0.02$,
not meeting our preregistered alpha threshold

Planets



<= Only significant nestedness result for curveball

Full dataset



this one: $p = 0.01$, **not** meeting our preregistered alpha threshold

$p = 0.13$

conclusion:
nestedness driven by between-subjects variance

Hard questions are almost only answered right by high-performing subjects, so a right answer to a hard question is a good predictor of right answers to easier ones.

=> Yes, we have nestedness in that sense.

Our data **is** more nested than randomly shuffled matrices that keep most other things equal *but* randomise variance between subjects.

Our data is **not** more nested than randomly shuffled matrices that keep most other things equal but keep each subject's performance level exactly identical to its real level.