UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA
DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE

Make Sense of Your Data

Rafael Peñaloza

Bologna, October 25th, 2025

Università di Milano-Bicocca

System Comparisons

if you torture the data long enough, it will confess to anything

– Ronald Coase

numbers don't tell a story; scientists do

Measurements

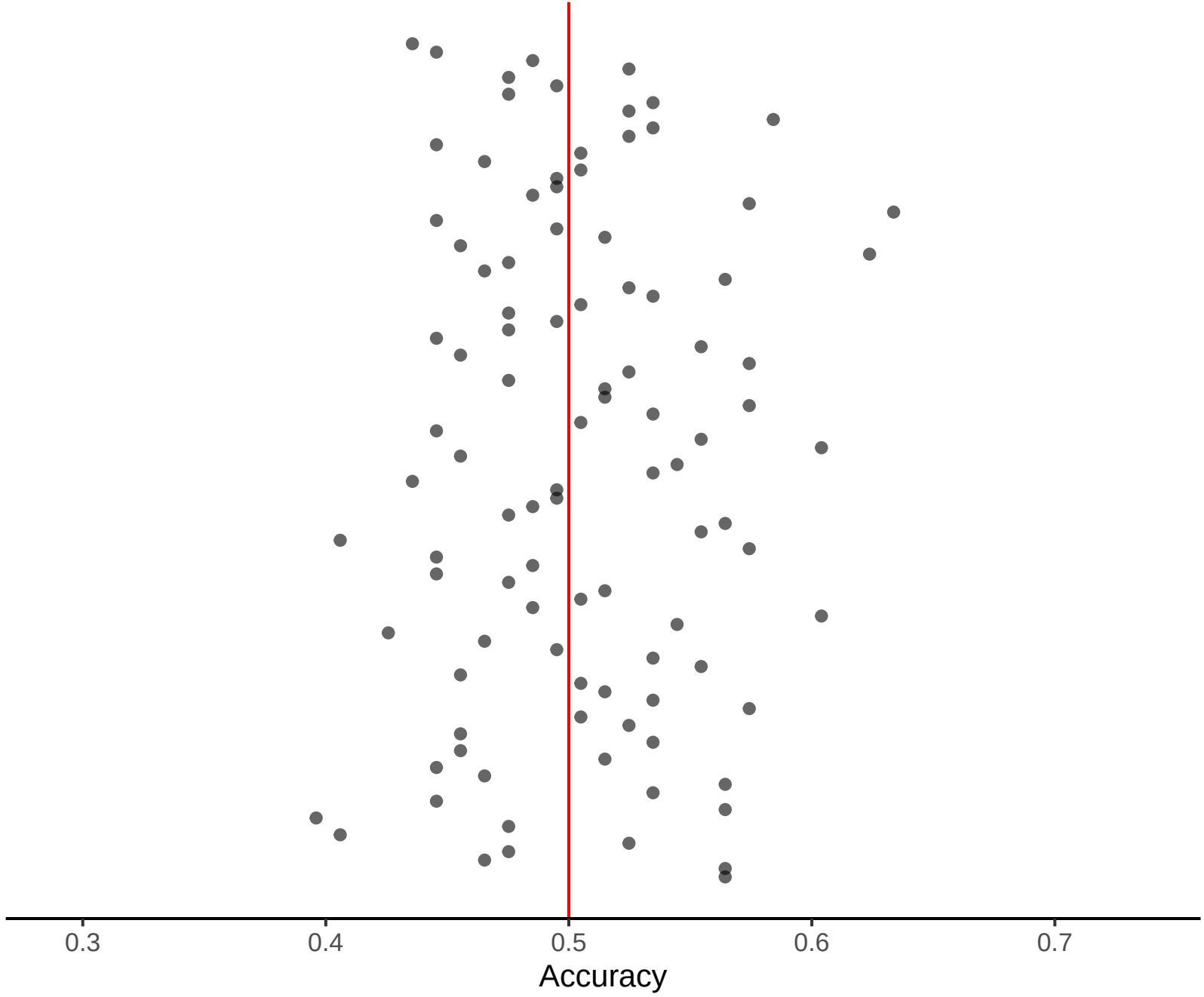
To measure a property of a system:

- execute over a dataset
- measure individual results
- compute the average

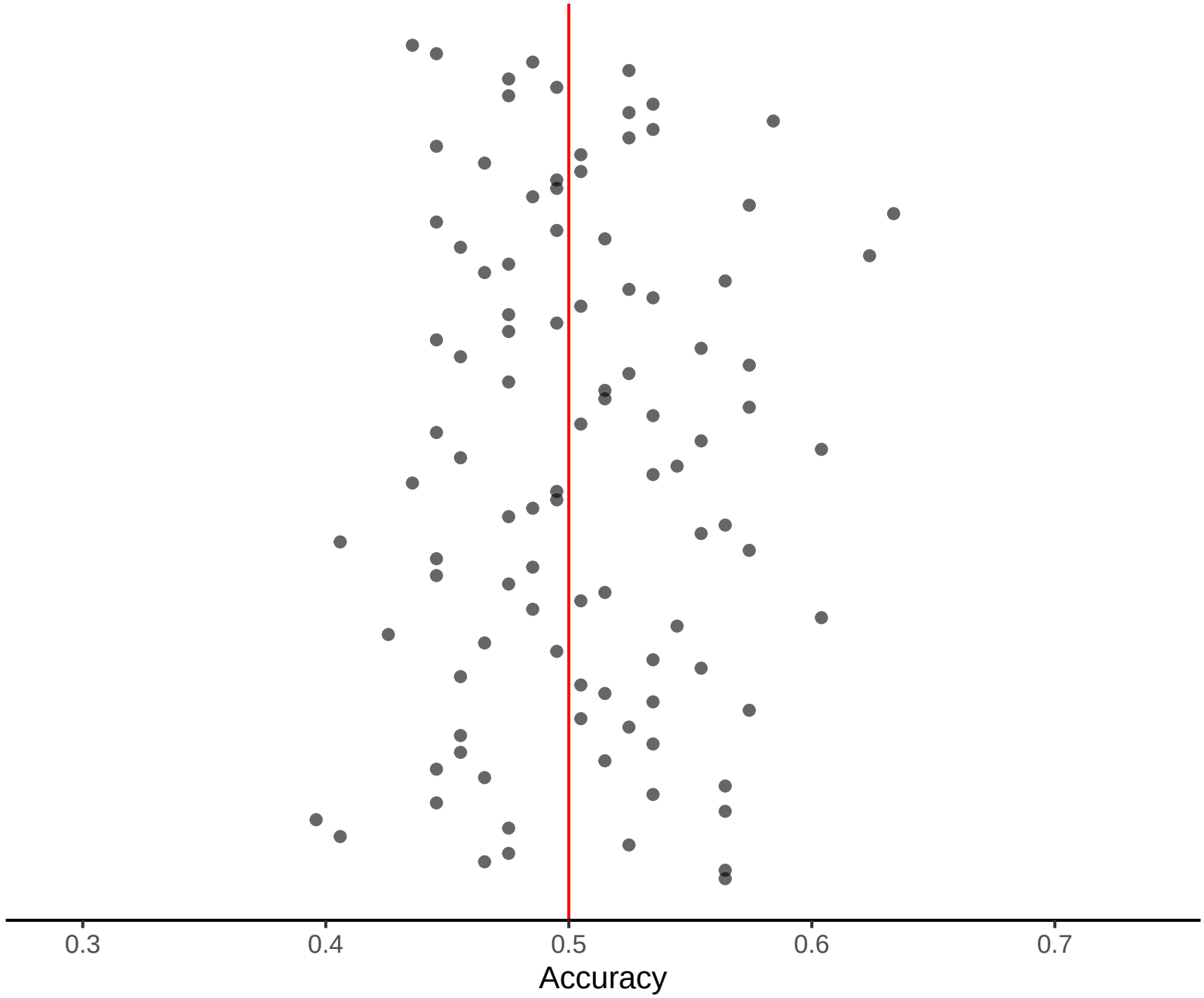
In statistical terms:

- measure over a sample
- estimate the population mean

Point estimators are not good for comparing


0.3
0.4
0.5
0.6
0.7
Accuracy

Use interval estimators


0.3
0.4
0.5
0.6
0.7
Accuracy

Confidence intervals

What is a confidence interval?

The interval estimator (constructed through the experiment)
has a X% probability of containing the true parameter

Around 95% of the intervals actually contain 0.5
the true accuracy of the method

What about comparisons?

Comparing systems

System	Measure	95% CI (100)
S1	0.45	[0.402, 0.498]
S2	0.53	[0.482, 0.578]

Insufficient evidence to claim differences in behaviour

System	Measure	95% CI (1000)
S1	0.45	[0.435, 0.465]
S2	0.53	[0.515, 0.545]

Significant difference (*)

Verifying differences

How to check that two systems' expected (i.e., average) behaviour is different?

If tested on same benchmark:

- estimate mean difference (CI estimation)
- verify that CI does not contain 0

If tested on different benchmarks:

- estimate each mean behaviour (two CIs)
- verify that the CIs do not intersect

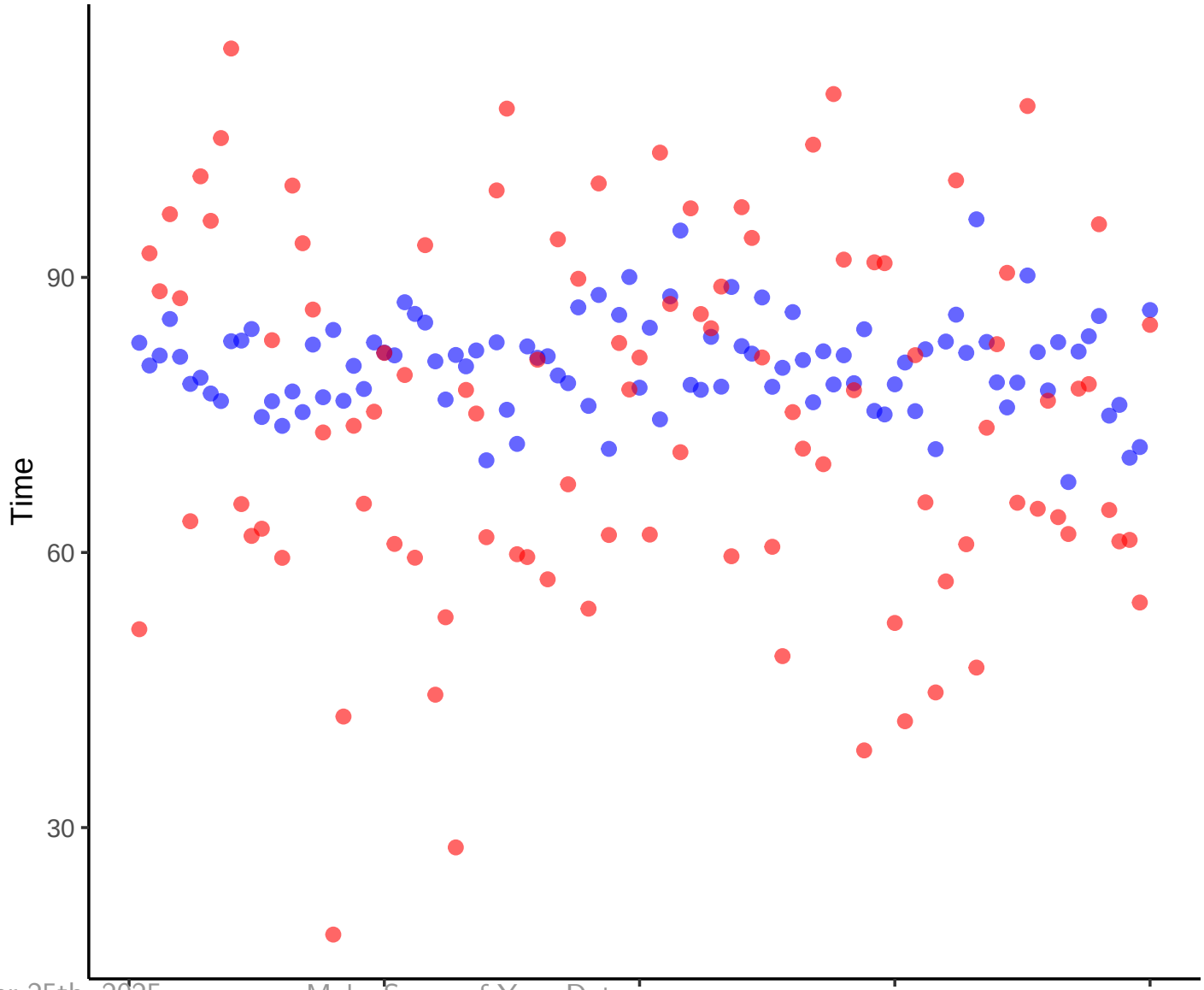
the least minimum

but remember
numbers don't tell stories

Execution time I

System	95% CI
S1	[71.25503,78.96345]
S2	[79,52834,81.56037]

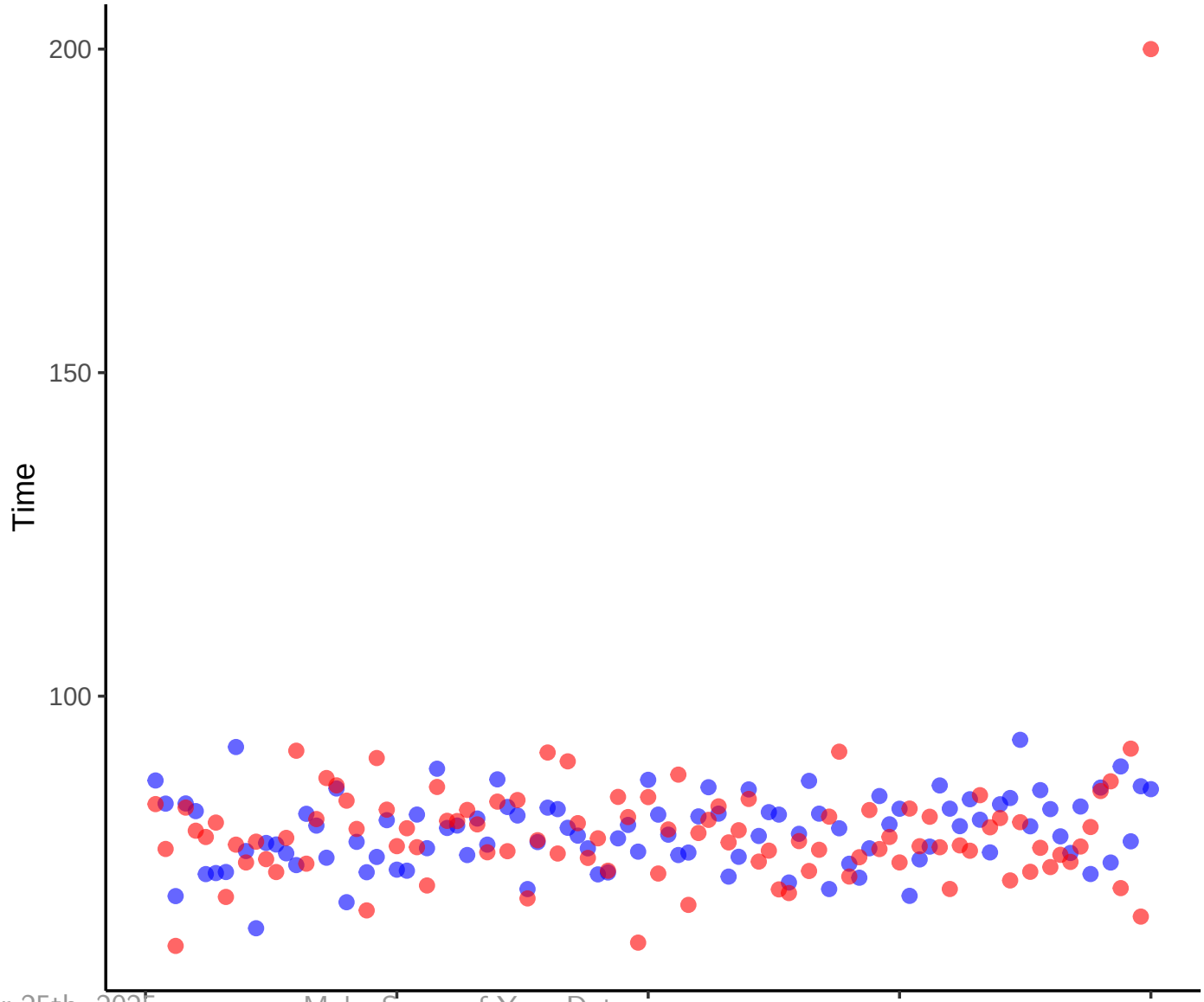
Execution time I



Execution time II

System	95% CI
S1	[78.00547,80.18993]
S2	[76.81135,82.22503]

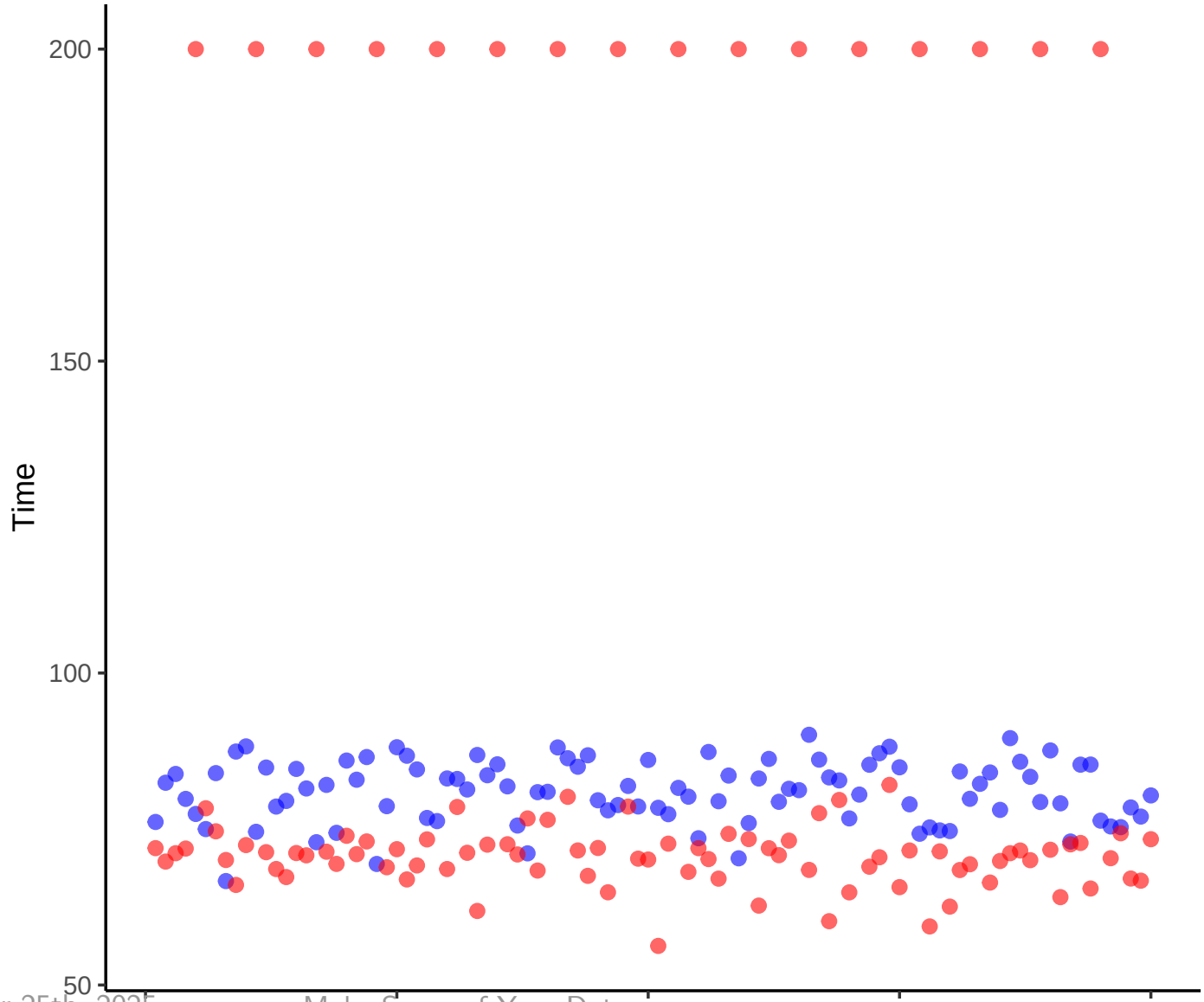
Execution time II



Execution time III

System	95% CI
S1	[79.96821,81.92348]
S2	[81.66366,100.6783]

Execution time III



Now you have no excuse

give voice to your work

rafael.penaloza@unimib.it