

CLUES-WEASEL clustering accuracy
(mean: 0.6427)

0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

CLUES-WEASEL wins here
[37W, 1T, 23L]

0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9

Shape-DBA clustering accuracy
(mean: 0.5991)

Shape-DBA wins here
[23W, 1T, 37L]

* Dashed lines represent the median

Wilcoxon test for equality of medians: p-value = 0.002
Paired t-test for equality of means: p-value $\leq 1e-3$

