

CLUES-WEASEL wins here
[47W, 0T, 24L]

CLUES-WEASEL clustering accuracy
(mean: 0.6180)

KASBA clustering accuracy
(mean: 0.5752)

KASBA wins here
[24W, 0T, 47L]

* Dashed lines represent the median

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