The results for **??** show that both evolutionary algorithms outperform the competition across all datasets and against all baselines. The fact that sparsity and similarity are the main drivers for this consistent improvement indicates a higher structual alignment between counterfactual and factual.

This result is remarkable as it shows that the algorithm can outperform baselines regardles of the process log, and its length.

The underperformance of the random model was expected. In **??** we indicated that viable algorithms must at least reach a viability of 2. Furthermore, we expected that the search space for the Random-Search Generator is too vast to possibly find viable results. The fact that 8 of 9 datasets showed that the random model cannot exceed the threshold of two supports this claim. A different support is the observation that every Casebased-Search Generator reaches at least 2.

The results give us hints about the empirical upper and lower bounds of the viability measure. None of the results median viabilities exceeds 3. Possibly, because of reasons we discussed in **??**. Mainly, the u