

Comparison of non\_llm\_string\_similarity across models  
Friedman  $\chi^2 = 51.9200$ ,  $p = 0.0000$  (significant)

