Family Bias Analysis
(Δ = FaviScore_same_family - FaviScore_cross_family)
Aggregated Ratings

Family Bias Δ (FaviScore units)

0.539*
0.312
0.260
0.198
-0.300
-0.492**

Judge Model

llama2-13b   mistral-7b   llama2-70b   mixtral-8x7b   gpt-4o-mini   gpt-3.5-turbo