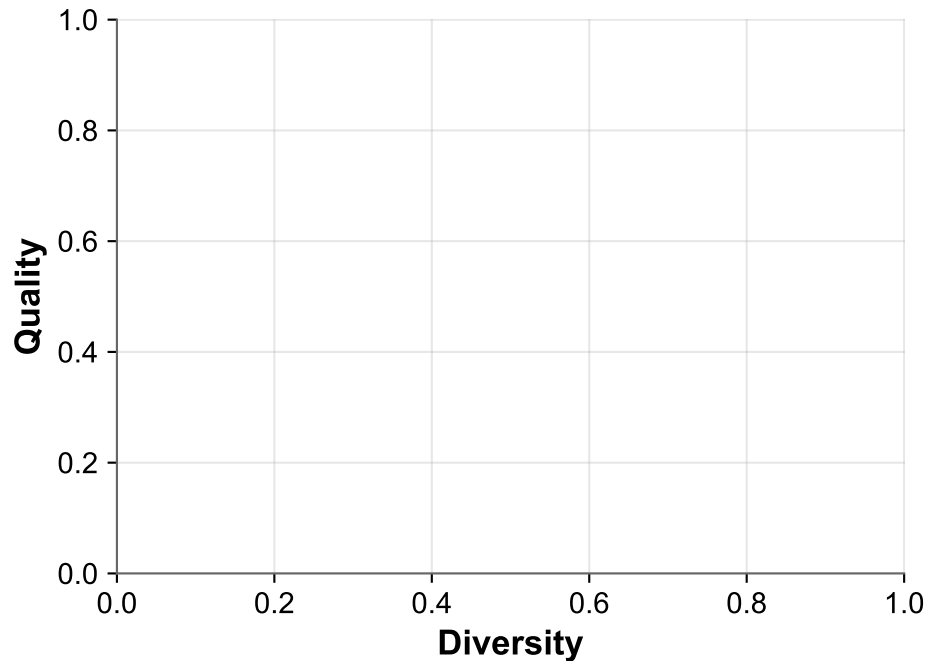


Number of Candidates (k) Ablation Study: Diversity vs Quality Analysis

Model: GPT-4.1



Model: Gemini-2.5-Flash

