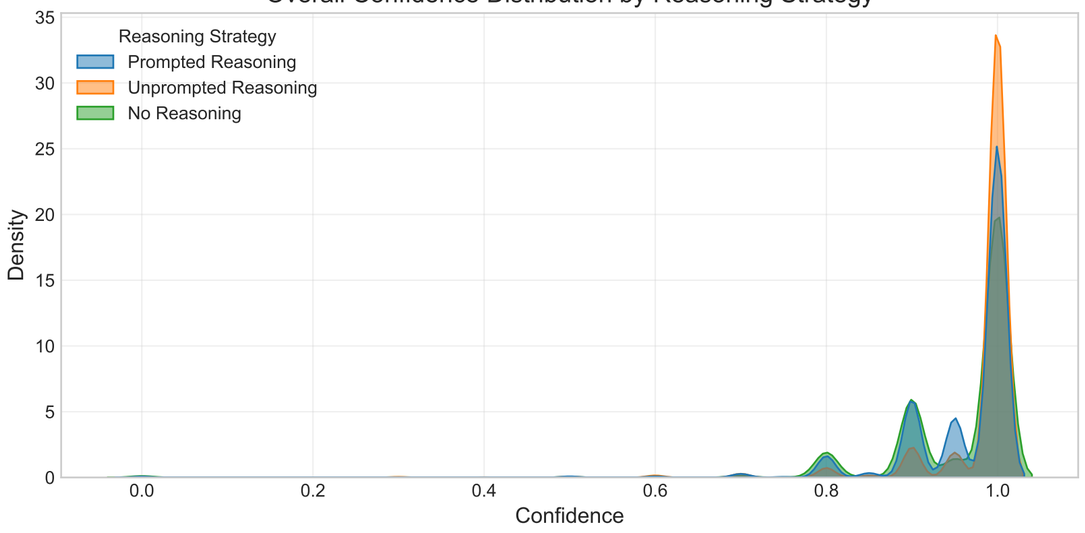
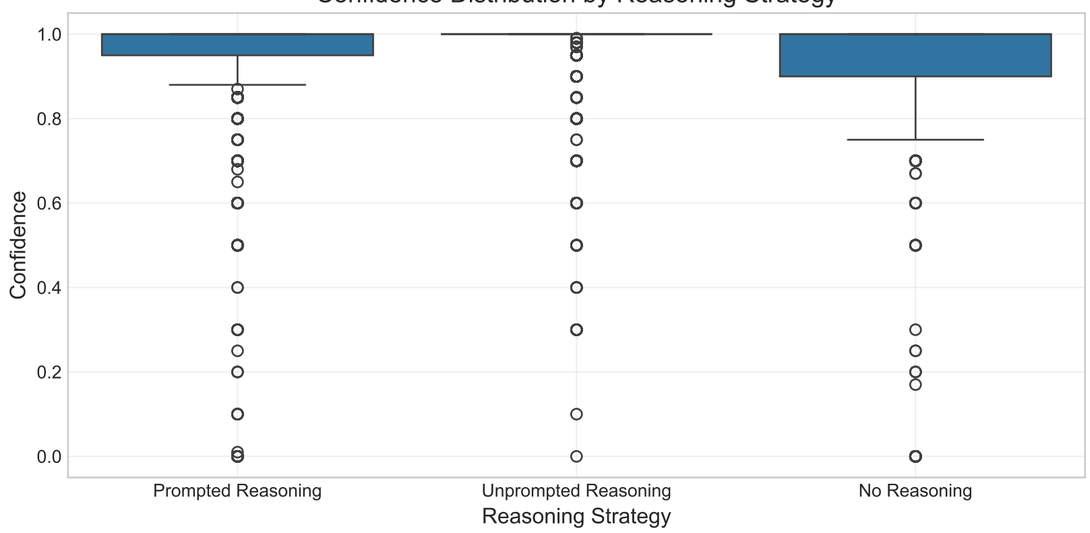
Confidence and Token Usage Analysis by Reasoning Strategy

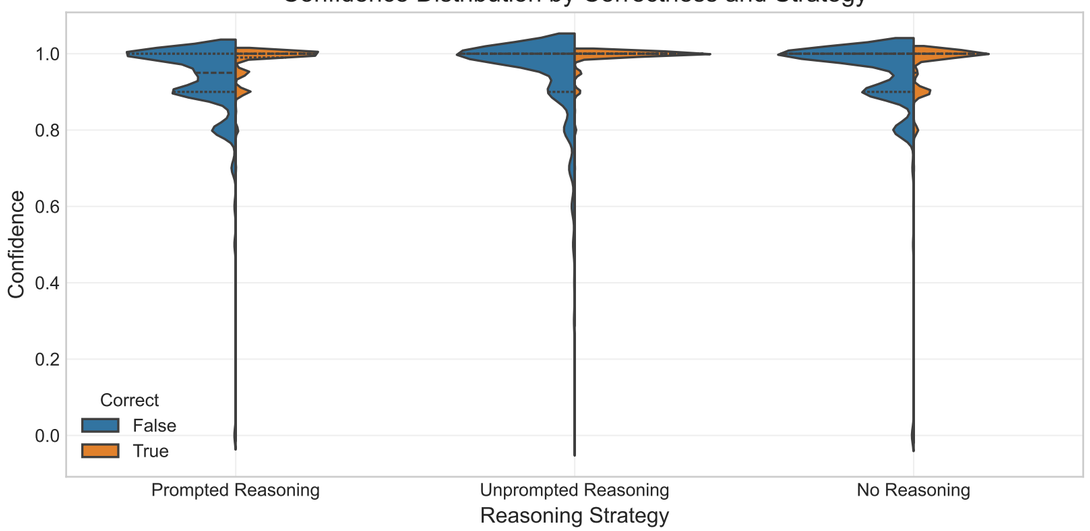
Overall Confidence Distribution by Reasoning Strategy



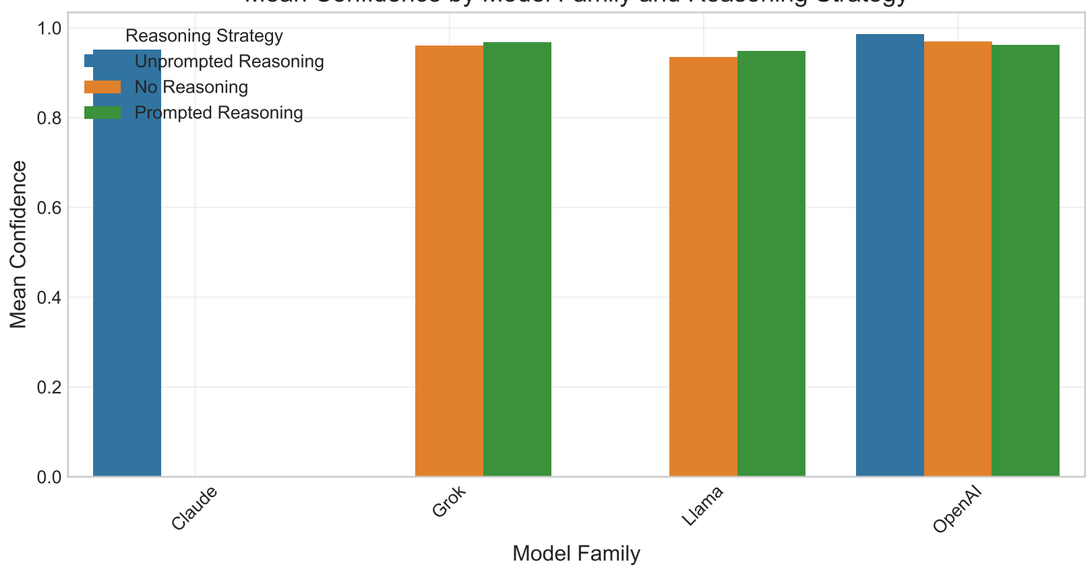
Confidence Distribution by Reasoning Strategy



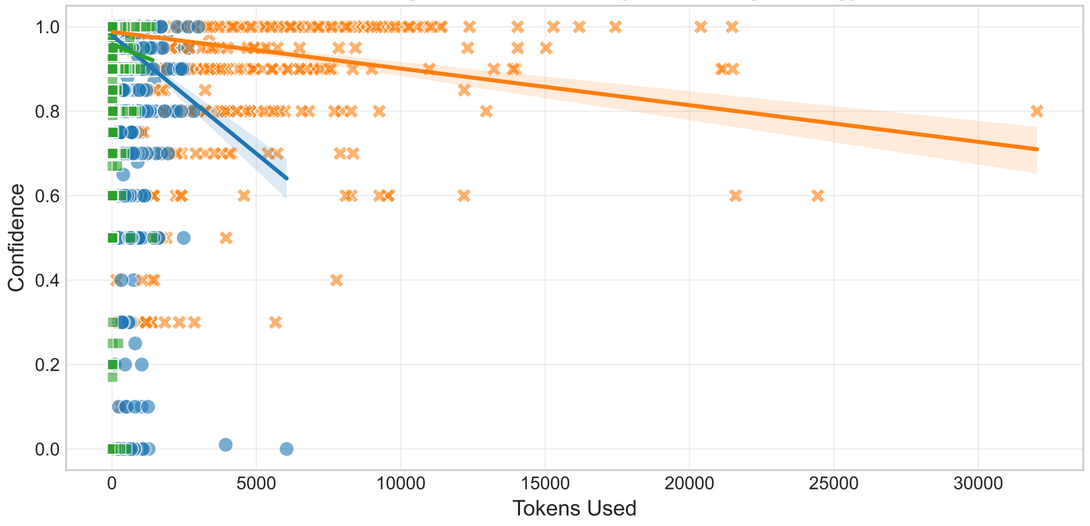
Confidence Distribution by Correctness and Strategy

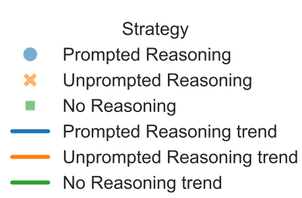


Mean Confidence by Model Family and Reasoning Strategy

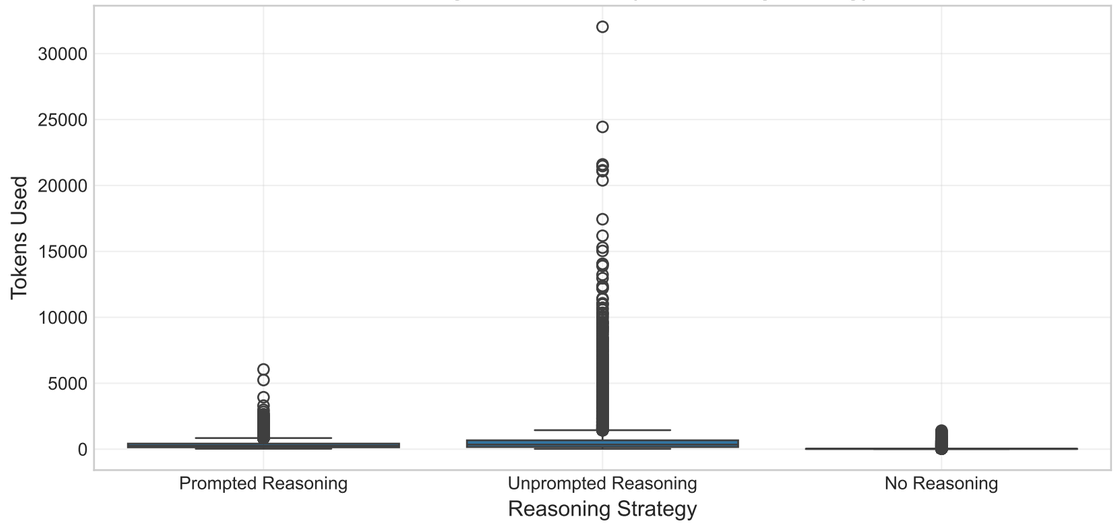


Token Usage vs Confidence by Reasoning Strategy





Token Usage Distribution by Reasoning Strategy



Statistics by Strategy:
No Reasoning: Confidence=0.96 ± 0.09, Tokens=34.4 ± 71.7, Correlation=-0.022
Unprompted Reasoning: Confidence=0.98 ± 0.06, Tokens=702.3 ± 1372.8, Correlation=-0.193
Prompted Reasoning: Confidence=0.96 ± 0.08, Tokens=317.8 ± 244.2, Correlation=-0.174