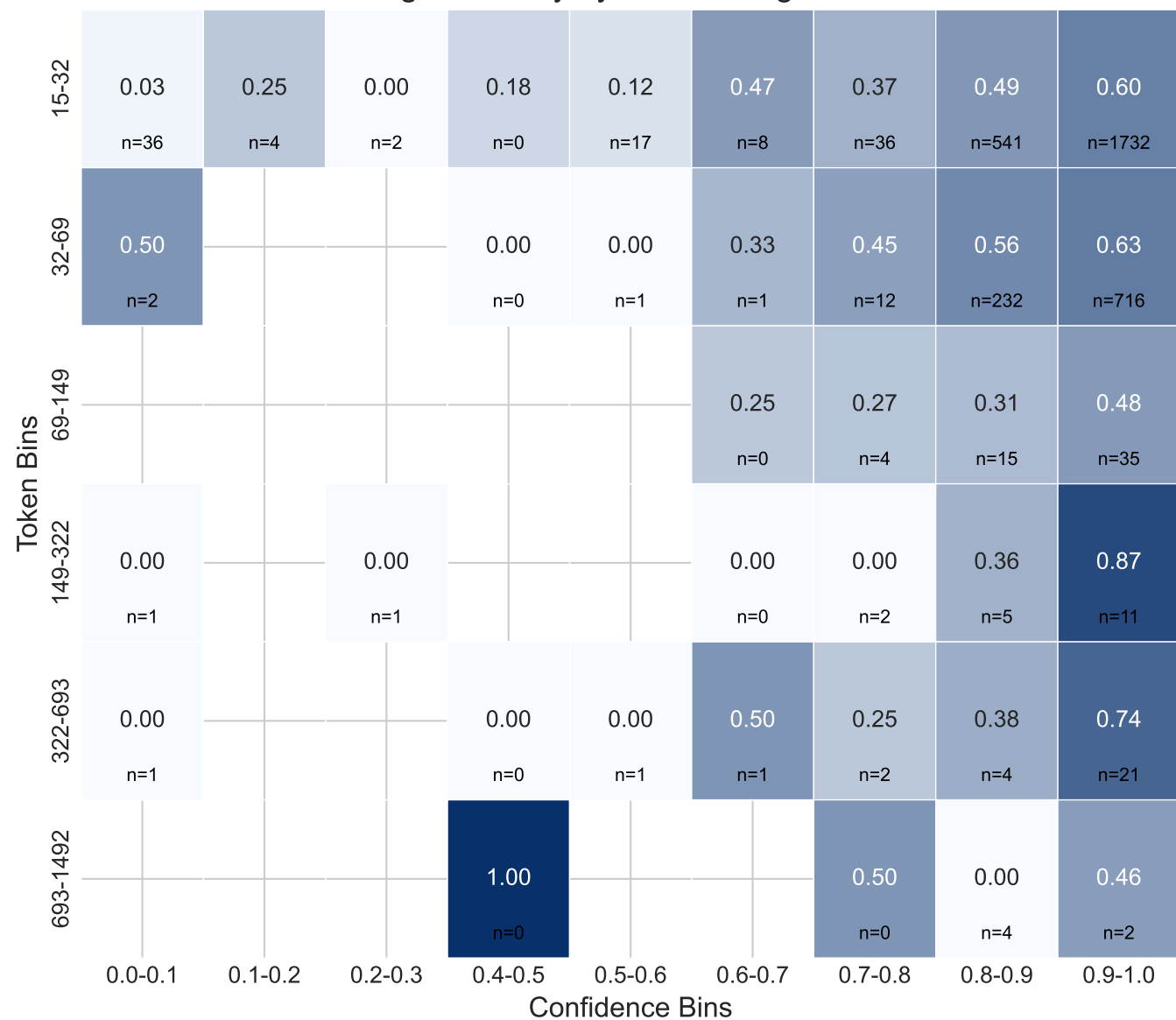
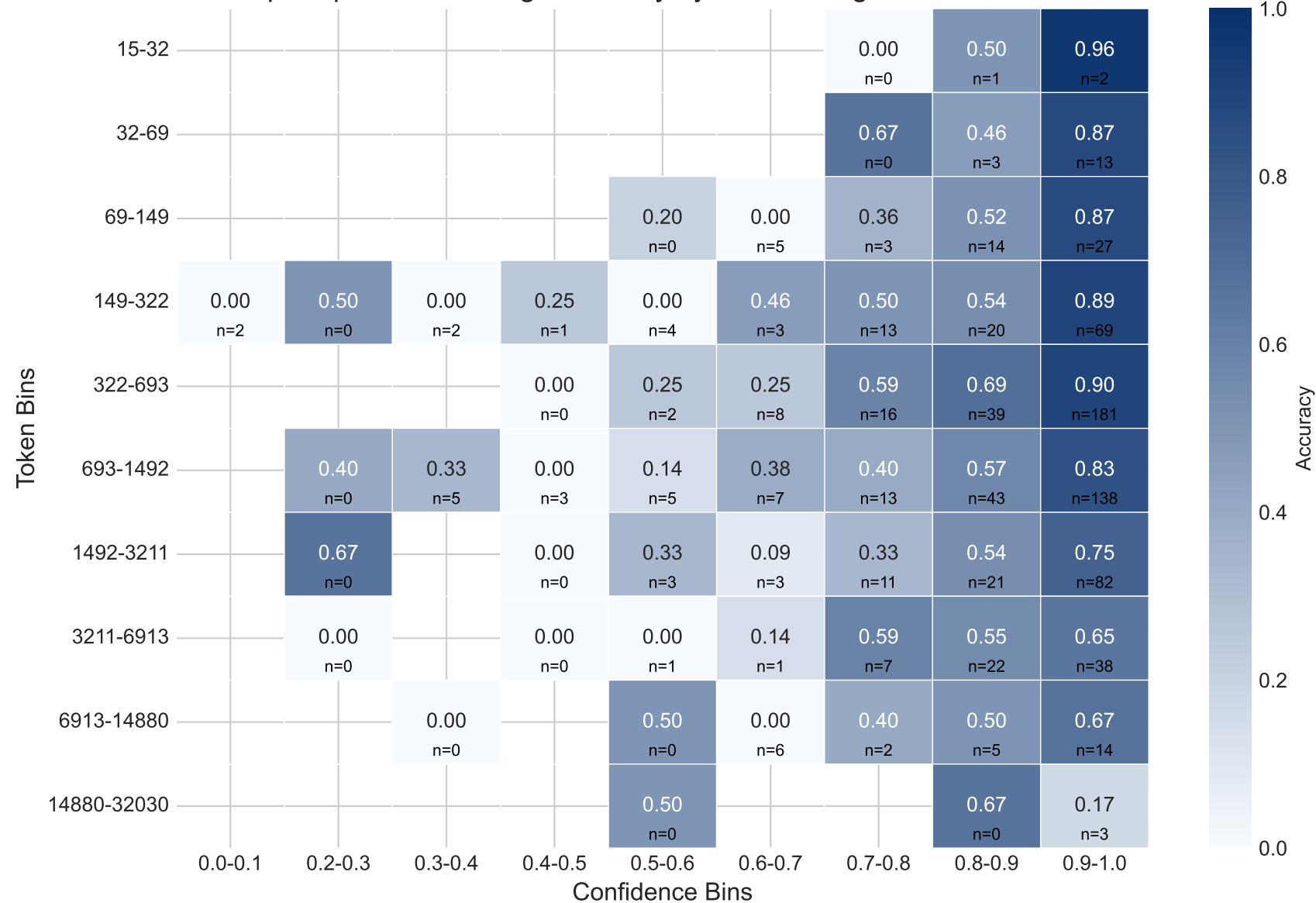


Token Usage vs Confidence and Accuracy Analysis by Reasoning Strategy

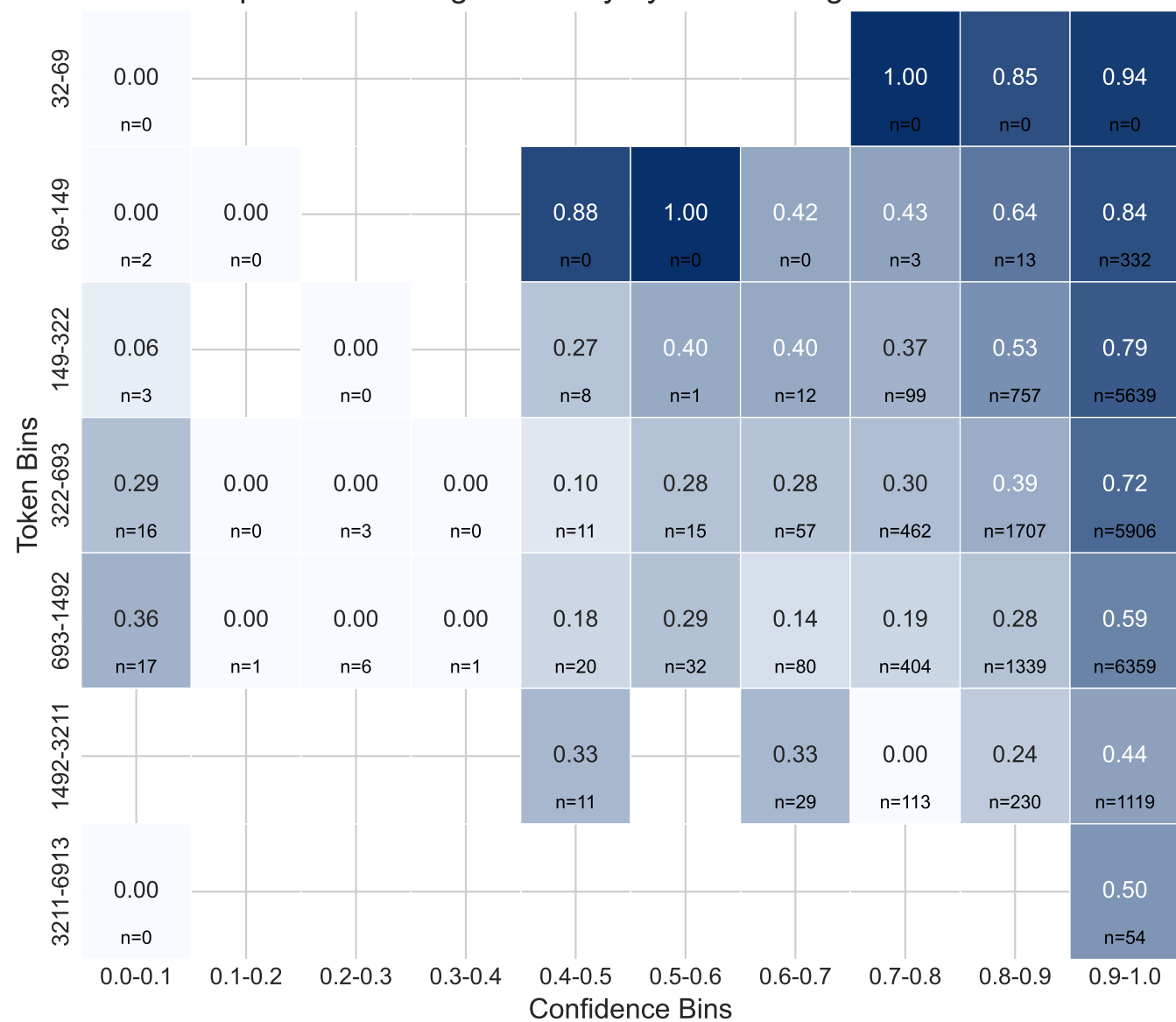
No Reasoning: Accuracy by Token Usage and Confidence



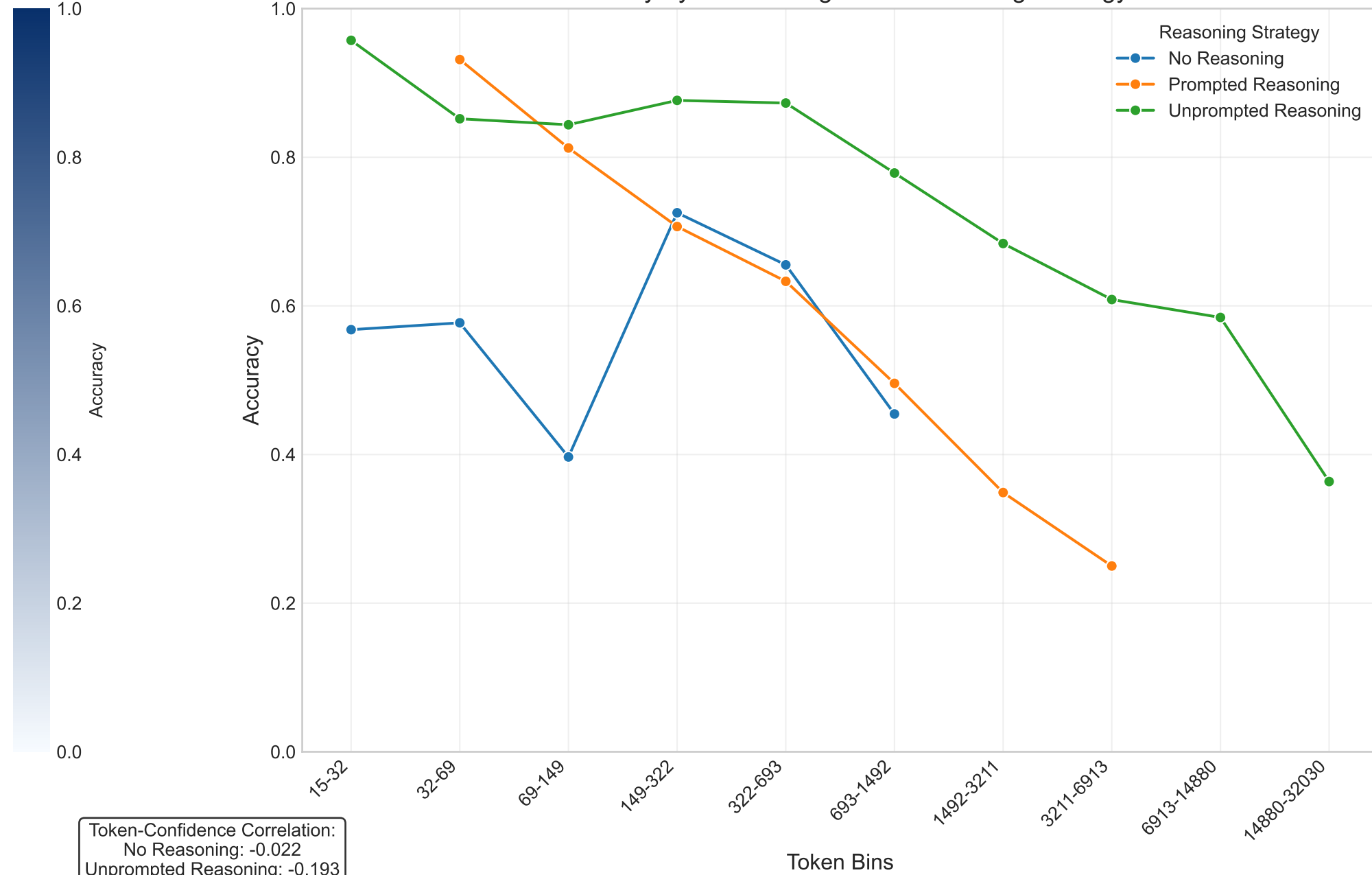
Unprompted Reasoning: Accuracy by Token Usage and Confidence



Prompted Reasoning: Accuracy by Token Usage and Confidence



Accuracy by Token Usage and Reasoning Strategy



Token-Confidence Correlation:
No Reasoning: -0.022
Unprompted Reasoning: -0.193
Prompted Reasoning: -0.174

Token-Accuracy Correlation:
No Reasoning: 0.005
Unprompted Reasoning: -0.161
Prompted Reasoning: -0.205