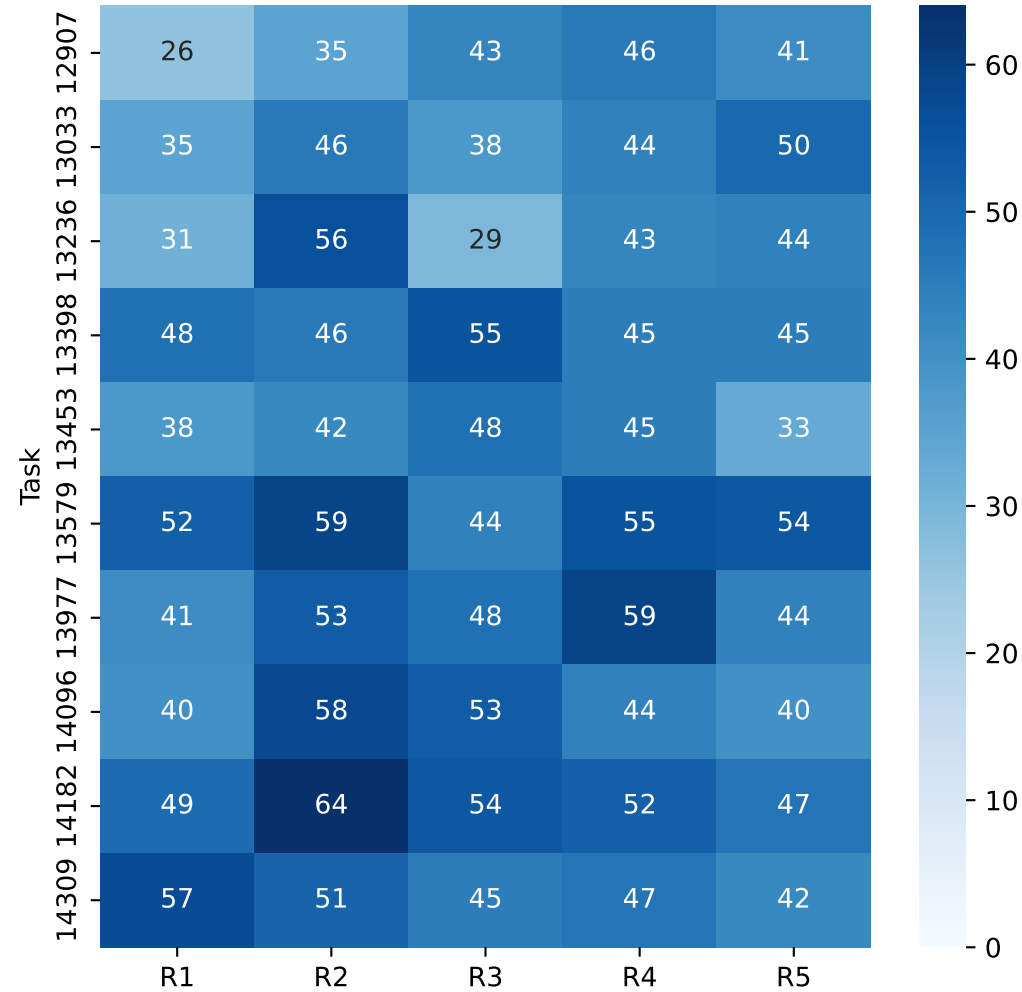
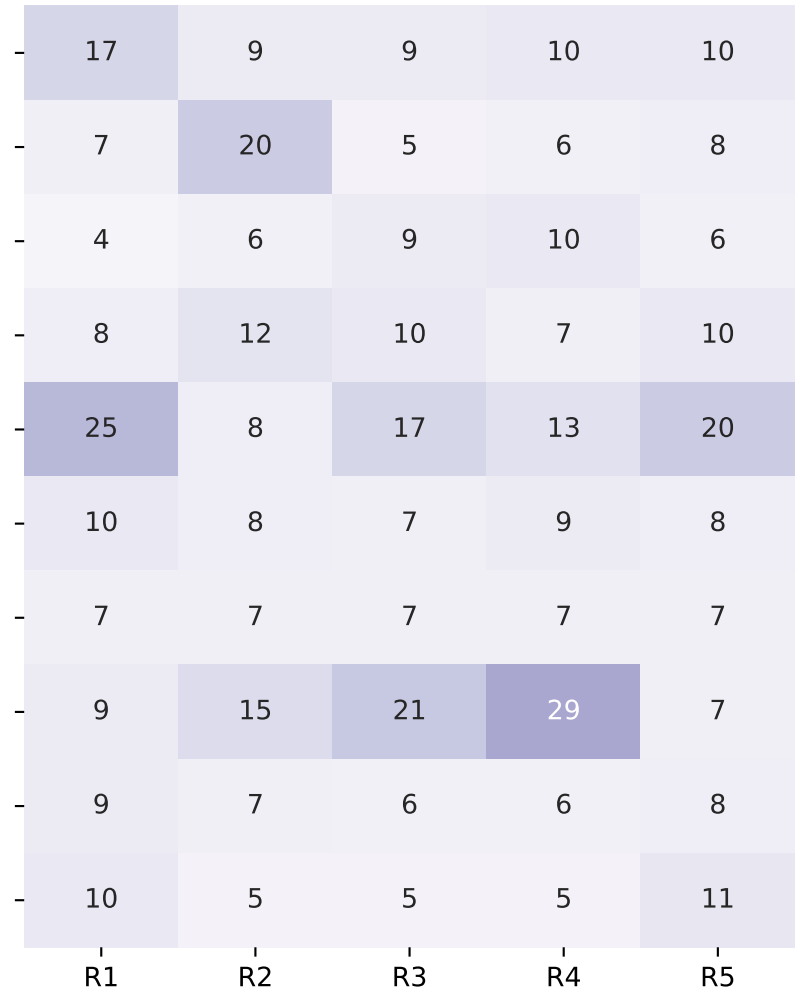


Step Counts: 10 Tasks × 5 Runs

Claude



GPT-5



Llama

