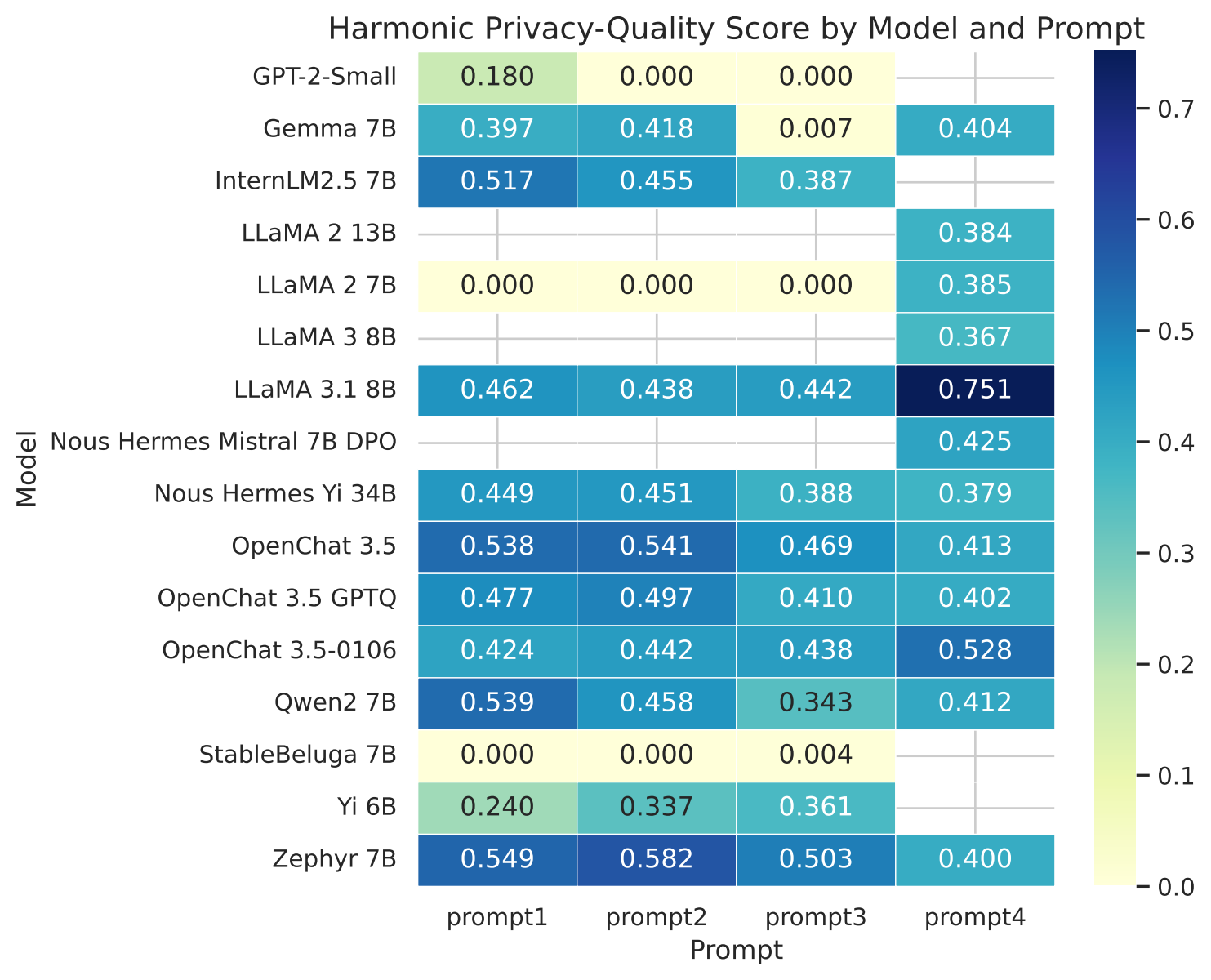
Harmonic Privacy-Quality Score by Model and Prompt
GPT-2-Small
Gemma 7B
InternLM2.5 7B
LLaMA 2 13B
LLaMA 2 7B
LLaMA 3 8B
LLaMA 3.1 8B
Nous Hermes Mistral 7B DPO
Nous Hermes Yi 34B
OpenChat 3.5
OpenChat 3.5 GPTQ
OpenChat 3.5-0106
Qwen2 7B
StableBeluga 7B
Yi 6B
Zephyr 7B
0.180 0.000 0.000
0.397 0.418 0.007 0.404
0.517 0.455 0.387
0.384
0.000 0.000 0.000 0.385
0.367
0.462 0.438 0.442 0.751
0.425
0.449 0.451 0.388 0.379
0.538 0.541 0.469 0.413
0.477 0.497 0.410 0.402
0.424 0.442 0.438 0.528
0.539 0.458 0.343 0.412
0.000 0.000 0.004
0.240 0.337 0.361
0.549 0.582 0.503 0.400
Model
prompt1
prompt2
prompt3
prompt4
Prompt
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0