

Queries and Errors per Model

Model	Name	Queries	Both Responses Errors	Neither Response Errors
			Errors	Errors
	Claude 3 Haiku	29830	0	0
	Claude 3 Sonnet	25770	0	0
	Claude 3.5 Sonnet	12960	0	0
	Claude 3 Opus	14250	0	0
	Gemini 1.5 Flash	24140	0	0
	Gemini 1.5 Pro	13620	0	0
	Gemma 2 27B	24450	0	4
	Code Llama 7B	42900	3	13
	Code Llama 13B	40770	3	30
	Code Llama 34B	40220	0	0
	Llama 2 Chat 7B	40450	0	149
	Llama 2 Chat 13B	37030	0	548
	Llama 3 Instruct 8B	30900	0	1
	Llama 3 Instruct 70B	24290	0	0
	Llama 3.1 Instruct 8B	40870	0	2
	Llama 3.1 Instruct 70B	29380	0	6
	Llama 3.1 Instruct 405B	14595	0	0
	Mistral 7B	25460	2	6
	Mixtral 8x7B	23220	0	0
	Mixtral 8x22B	24240	0	0
	Mistral Large 2	18285	0	0
	GPT-3.5 Turbo (1106)	36070	0	1
	GPT-3.5 Turbo (0125)	37840	0	0
	GPT-4 (0314)	13965	0	0
	GPT-4 (0613)	13575	0	0
	GPT-4 Turbo (1106)	14670	0	0
	GPT-4 Turbo (2024-04-09)	13785	0	0
	GPT-4o mini	24990	0	0
	GPT-4o (2024-05-13)	14310	0	0
	o1 Mini (2024-09-12)	25300	0	0
	o1 Preview (2024-09-12)	10510	0	0

Queries

Both Responses Errors Neither Response Errors