

Gemini 3.7 Flash

Google

AT A GLANCE

PRICE PER 1M

\$0.75 / \$3.75

input / output · cached
\$0.075

CONTEXT

1.05M

tokens in

MAX OUTPUT

66K

tokens per reply

TAKES

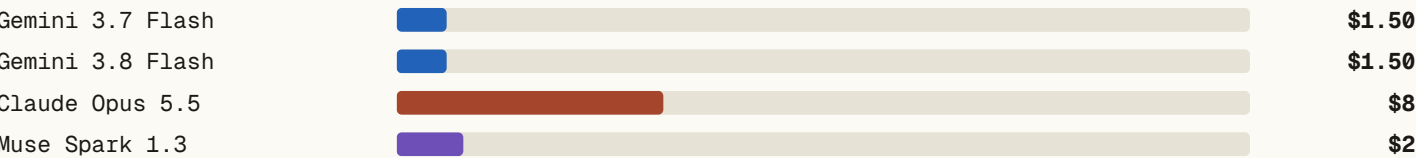
**text, image, video,
file, audio**

returns text

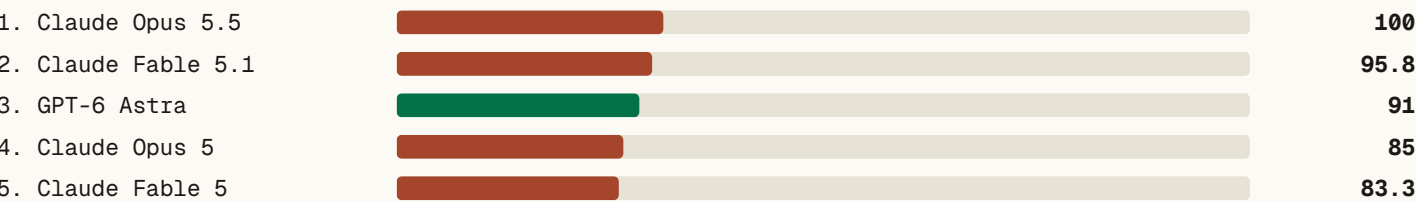
HOW IT COMPARES

MODEL	ROLE	IN / 1M	OUT / 1M	CONTEXT	BENCH INDEX
Gemini 3.7 Flash	new	\$0.75	\$3.75	1.05M	not rated yet
Gemini 3.8 Flash	it follows	\$0.75	\$3.75	1.05M	50
Claude Opus 5.5	today's #1	\$4	\$20	1M	100
Muse Spark 1.3	best value	\$1.25	\$4.25	1.05M	79.1
vs Gemini 3.8 Flash		same price	the model it follows: \$0.75 / \$3.75, bench index 50		
vs Claude Opus 5.5		81% cheaper	today's #1: \$4 / \$20, bench index 100; context 1M		
vs Muse Spark 1.3		input 40% cheaper, output 12% cheaper		best value: \$1.25 / \$4.25, bench index 79.1	

PRICE PER 1M TOKENS, BLENDED 3 IN : 1 OUT



WHERE THE PUBLIC LEADERBOARDS STAND TODAY



Gemini 3.7 Flash is not rated yet: the public boards usually score a new model within days, and its page on artificialwatch.com gains the scores then. The bench index is the median of each model's percentile across the public leaderboards that rate it (LMarena, LiveBench, SimpleBench, LLM Stats, Vals AI, Artificial Analysis).

Sources: price, limits and modalities for Gemini 3.7 Flash as the OpenRouter catalog lists google/gemini-3.7-flash; the other models as their makers list them (models.dev); rankings from the ArtificialWatch Bench Index. Facts only: no forecast of ours, no claims from the lab's announcement.