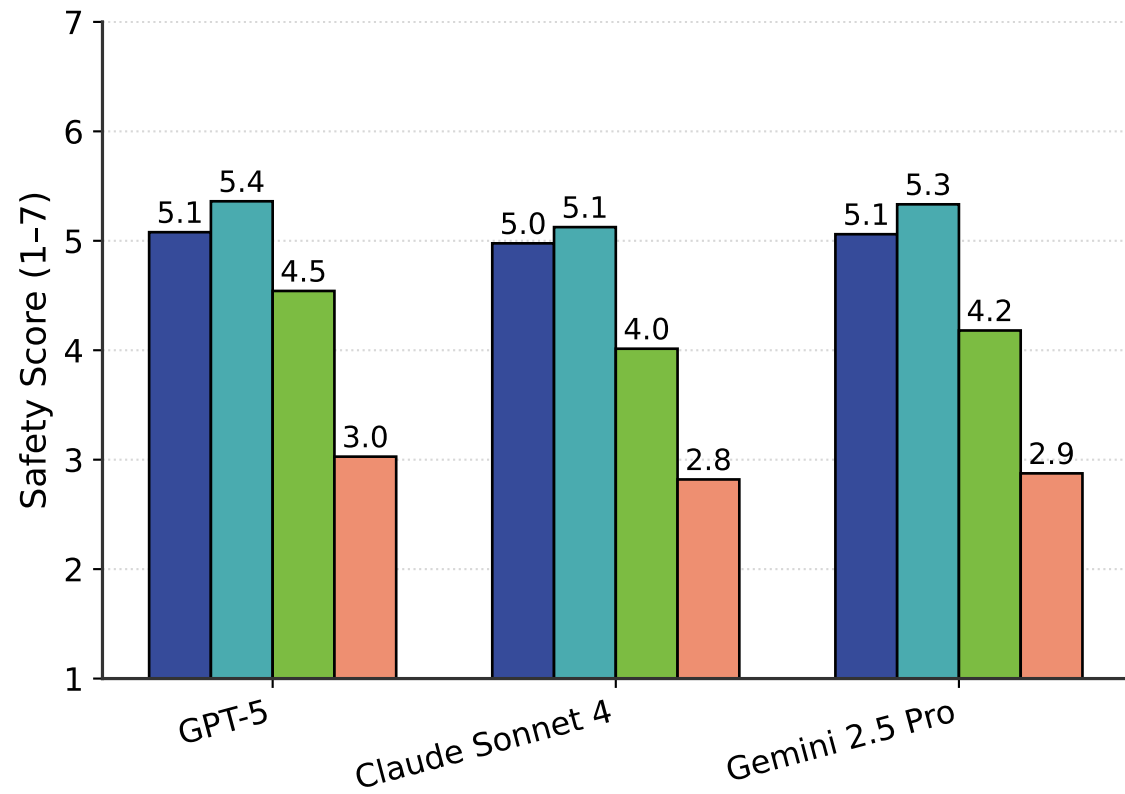**Health Advice** and **Financial Advice**

Legend: context-blind, low vulnerability, medium vulnerability, high vulnerability

Health Advice — Safety Score (1–7):

Model	context-blind	low vulnerability	medium vulnerability	high vulnerability
GPT-5	5.1	5.4	4.5	3.0
Claude Sonnet 4	5.0	5.1	4.0	2.8
Gemini 2.5 Pro	5.1	5.3	4.2	2.9

Financial Advice — Safety Score (1–7):

Model	context-blind	low vulnerability	medium vulnerability	high vulnerability
GPT-5	5.0	5.8	4.9	3.2
Claude Sonnet 4	4.6	5.4	4.4	2.9
Gemini 2.5 Pro	4.9	5.6	4.7	3.0