Performance (Plausible@1) vs Cost

- o4-mini-2025-04-16-high (2025-04-16)
- o3-mini-2025-01-31-high (2025-01-31)
- deepseek-r1 (2025-01-20)
- claude-3-7-sonnet-20250219 (2025-02-19)
- claude-3-5-sonnet-20241022 (2024-10-22)
- gpt-4.1-2025-04-14 (2025-04-14)
- claude-3-5-sonnet-20240620 (2024-06-20)
- gemini-2.5-pro-preview-03-25 (2025-04-04)
- deepseek-v3-0324 (2025-03-24)
- deepseek-v3 (2024-12-26)
- gemini-1.5-pro-002
- gpt-4o-2024-11-20 (2024-11-20)
- gpt-4o-2024-08-06 (2024-08-06)
- gemini-2.0-flash-001 (2025-02-05)
- llama-4-maverick (2025-04-05)
- gemini-1.5-pro-001 (2024-05-24)
- llama-3.1-405b-instruct (2024-07-23)
- grok-2-1212 (2024-12-12)
- qwen-2.5-coder-32b-instruct (2024-11-12)
- qwen-2.5-72b-instruct (2024-09-19)
- deepseek-v2.5 (2024-09-05)
- command-a (2025-03-13)
- mistral-large-2411 (2024-11-18)
- mistral-large-2407 (2024-07-24)
- llama-3.3-70b-instruct (2024-12-06)
- llama-3.1-nemotron-70b-instruct (2024-10-15)
- deepseek-r1-distill-llama-70b (2025-01-20)
- mistral-small-2503 (2025-03-17)
- deepseek-r1-distill-qwen-32b (2025-01-20)
- codestral-2501 (2025-01-13)
- gemma-3-27b-it (2025-03-12)
- codestral-2405 (2024-05-23)

X-axis: Total Cost ($)
Y-axis: Total Plausible@1