AIME 2024

AIME-2024 · math benchmark · 2024

Source: https://policywindow.org/wiki/aime-2024

Generated 2026-11-19T01:06:08 UTC

Summary

30 problems from the 2024 American Invitational Mathematics Examination — high-school competition math.

At a glance

Score range

0–100 % accuracy

Contamination risk

low

Methodology

https://www.maa.org/math-competitions/american-invitational-mathematics-examination-aime

Saturation

saturating

Details

Released after most current models' training cutoffs. Top reasoning models 75-90%; non-reasoning 10-30%. Currency (2026-06-21): Frontier has climbed past the article top o3 96.7% figure - GPT-5 (~95.7%), Grok 4 (~94.3%), and Gemini 3 Deep Think (98-99%) now top AIME leaderboards, reinforcing the near-saturation thesis; the table could add a post-2024-model row, but caution per the existing iter-449f audit note that many headline figures (e.g. OpenAI 94.6%, Grok 4 100%) are AIME 2025, not AIME 2024.

How to cite this article

APA

Policy Window. (2024). AIME 2024 [Wiki article — Benchmark]. https://policywindow.org/wiki/aime-2024

CHICAGO

Policy Window. 2024. "AIME 2024." Wiki article (Benchmark). https://policywindow.org/wiki/aime-2024.

HARVARD

Policy Window (2024) 'AIME 2024', Wiki article — Benchmark, available at: https://policywindow.org/wiki/aime-2024.

OSCOLA

Policy Window, 'AIME 2024' (Wiki article — Benchmark, 2024) <https://policywindow.org/wiki/aime-2024> accessed [date].

BIBTEX

```
@misc{policywindow-aime-2024,
title  = {AIME 2024},
author = {Policy Window},
year   = {2024},
howpublished = {AIME-2024 (2024)},
url    = {https://policywindow.org/wiki/aime-2024},
```

```
note   = {Primary source: https://www.maa.org/math-competitions/american-invitational-mathematics-examination-aime}
}
```