

# Análise de Dashboard: LLM-Stats

## 1. Introdução e Escolha do Caso

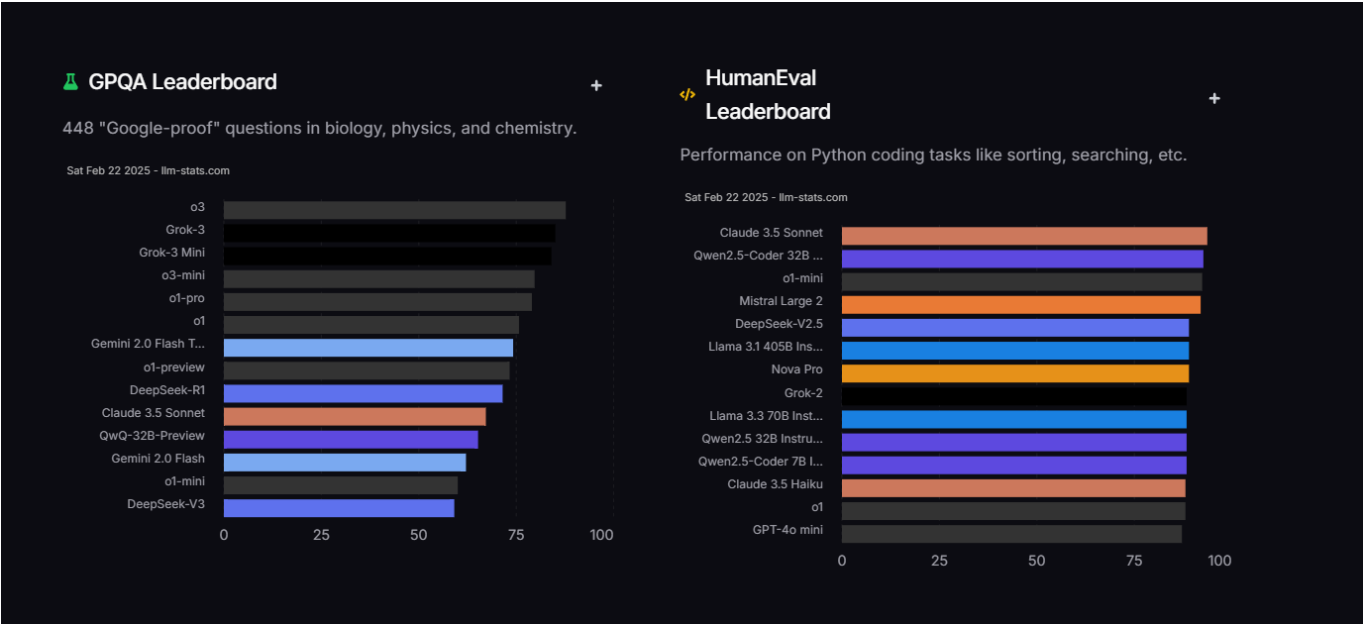
O avanço dos modelos de inteligência artificial tornou essencial a comparação entre diferentes Large Language Models (LLMs). O site LLM-Stats apresenta uma solução inovadora ao consolidar benchmarks de IA em uma única dashboard. Esta análise explora como esse dashboard organiza e apresenta os dados e como ele se compara a outras abordagens, como a da xAI.

## 2. Problema e Contexto

A crescente diversidade de modelos de IA criou um desafio para pesquisadores e empresas: comparar diferentes modelos de forma justa. Benchmarks padronizados são essenciais para avaliar o desempenho real dos LLMs. Dashboards que apresentam dados de forma transparente e intuitiva são ferramentas fundamentais para facilitar essa análise.

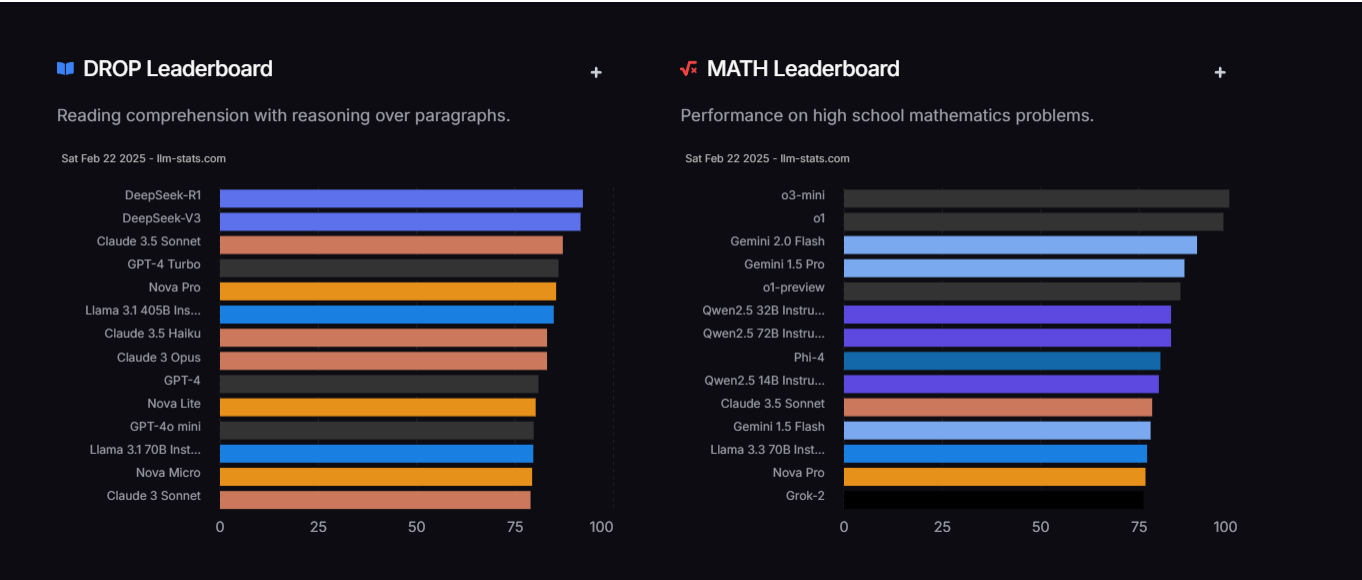
## 3. Descrição da Solução: LLM-Stats

O LLM-Stats se destaca por sua abordagem transparente e detalhada. Ele apresenta rankings de modelos com métricas diversas, utilizando gráficos absolutos que evitam manipulação de dados.



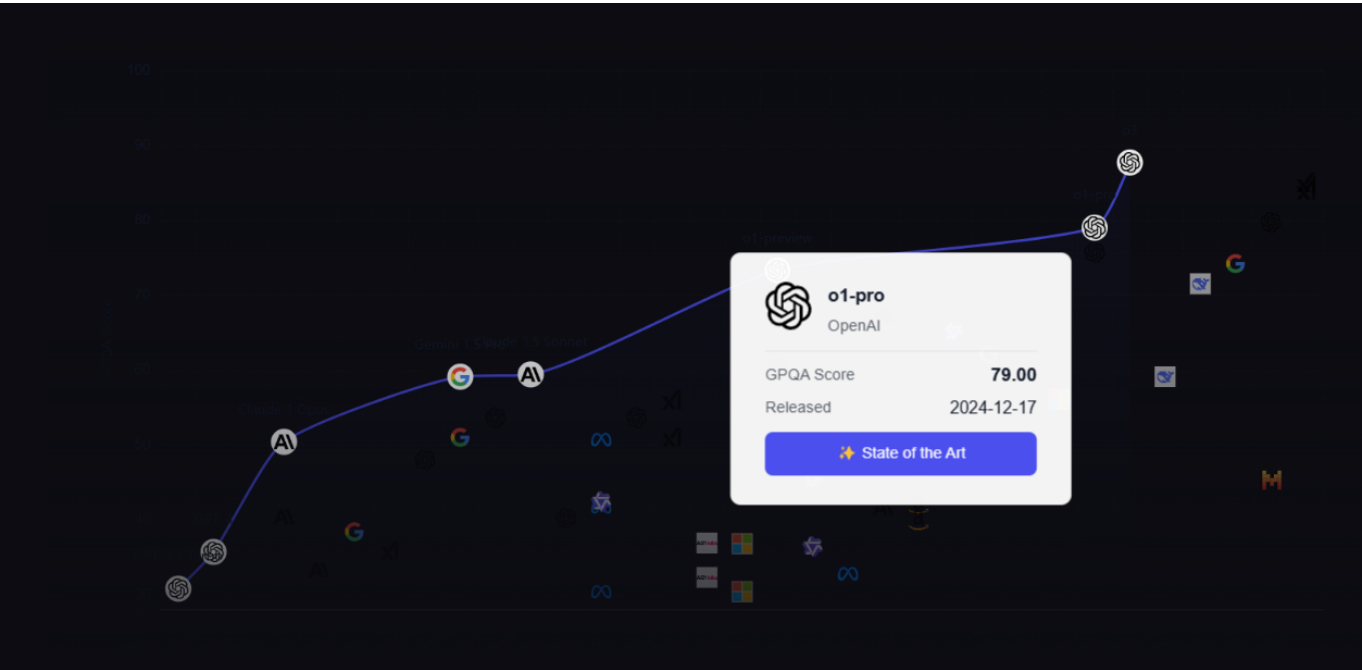
### GPQA Leaderboard

O gráfico GPQA Leaderboard apresenta a performance dos modelos em questões científicas complexas. A escolha de um gráfico de barras horizontais facilita a comparação direta entre os modelos. A escala absoluta utilizada mantém a transparência dos dados, evitando distorções.



MATH Leaderboard

Este gráfico avalia a habilidade dos LLMs em resolver problemas matemáticos. A escolha de barras coloridas para cada modelo auxilia na distinção rápida das performances. A estrutura bem definida permite uma interpretação rápida dos resultados.



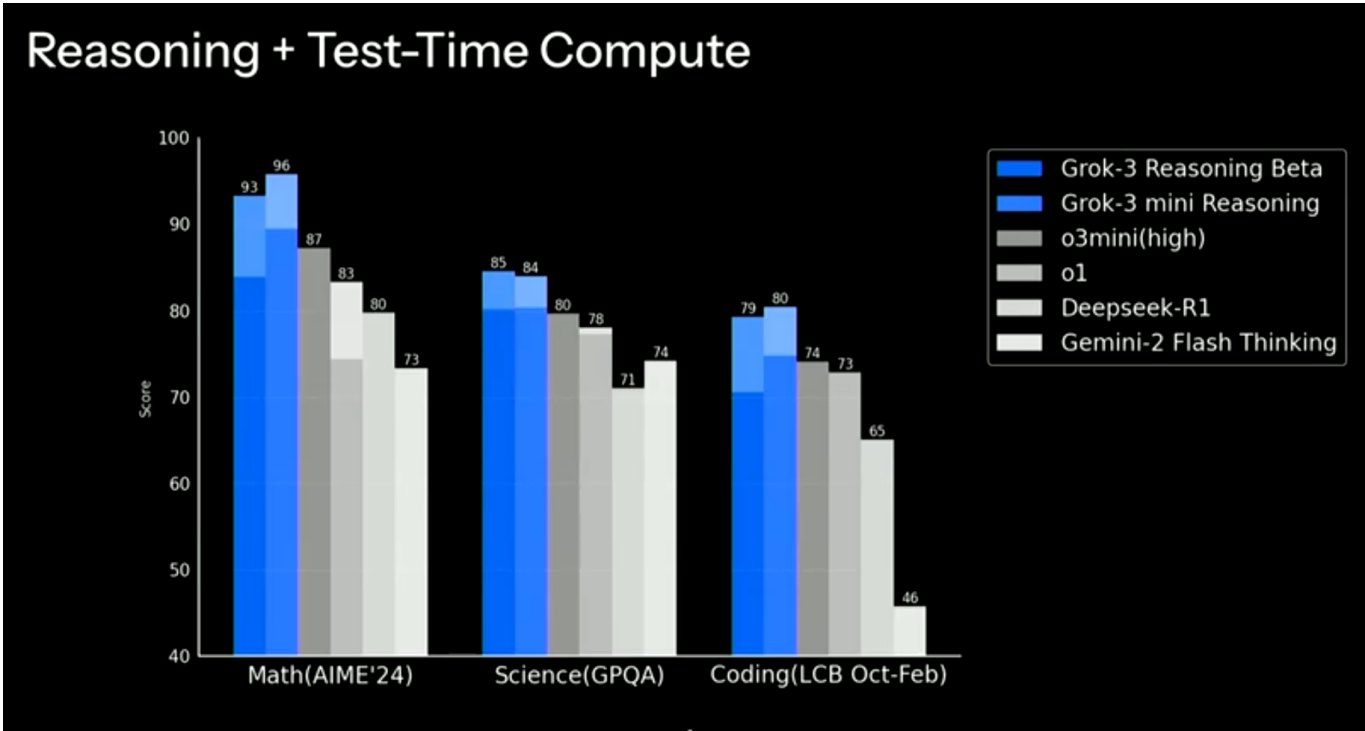
Evolução Temporal dos Modelos

O gráfico de linha ilustra a evolução da performance dos modelos ao longo do tempo. Ele permite uma visão clara do progresso e das tendências do setor, sendo um diferencial do LLM-Stats.

4. Comparação com xAI (Grok-3)

Diferente do LLM-Stats, a xAI utilizou um gráfico recortado para enfatizar a superioridade do Grok-3. Ao omitir a parte inferior do gráfico, a empresa gerou um efeito visual exagerado,

distorcendo a percepção dos dados.



5. Análise Crítica: Transparência e Eficiência

O LLM-Stats acerta ao apresentar dados com clareza e sem manipulações. Seu uso de gráficos bem organizados e cores distintas facilita a leitura e a tomada de decisões. A comparação com a xAI reforça a importância da escala absoluta para evitar distorções visuais.

6. Conclusão

O LLM-Stats se destaca como um dashboard confiável e transparente para avaliação de LLMs. Sua abordagem gráfica bem fundamentada facilita a análise comparativa e se alinha às boas práticas descritas na Aela.io.

Referências

- <https://llm-stats.com/>
- <https://www.aela.io/pt-br/blog/conteudos/dashboard-como-criar-o-design-de-dados-e-informacoes>
- Análise comparativa de dashboards e visualização de dados em inteligência artificial.