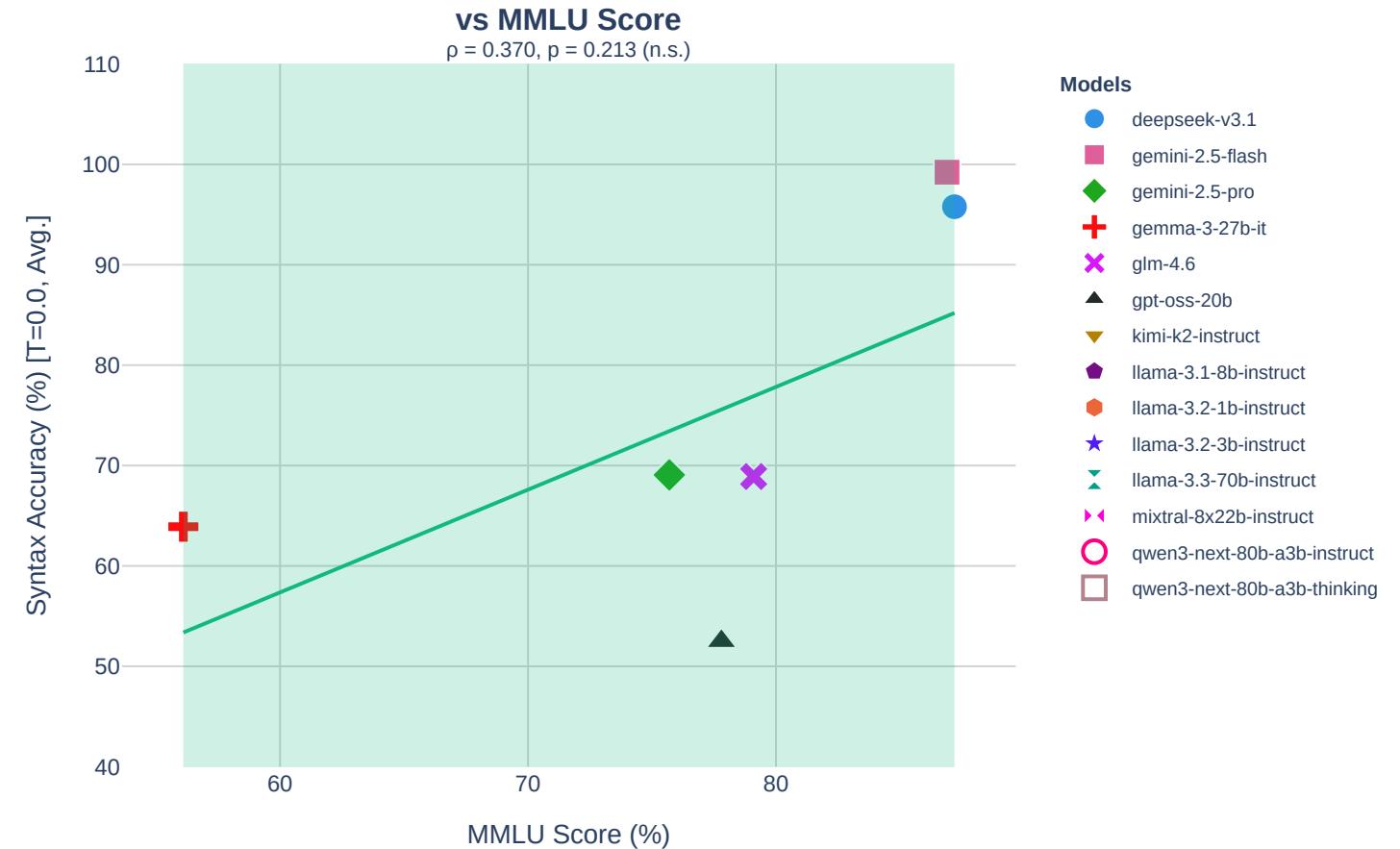
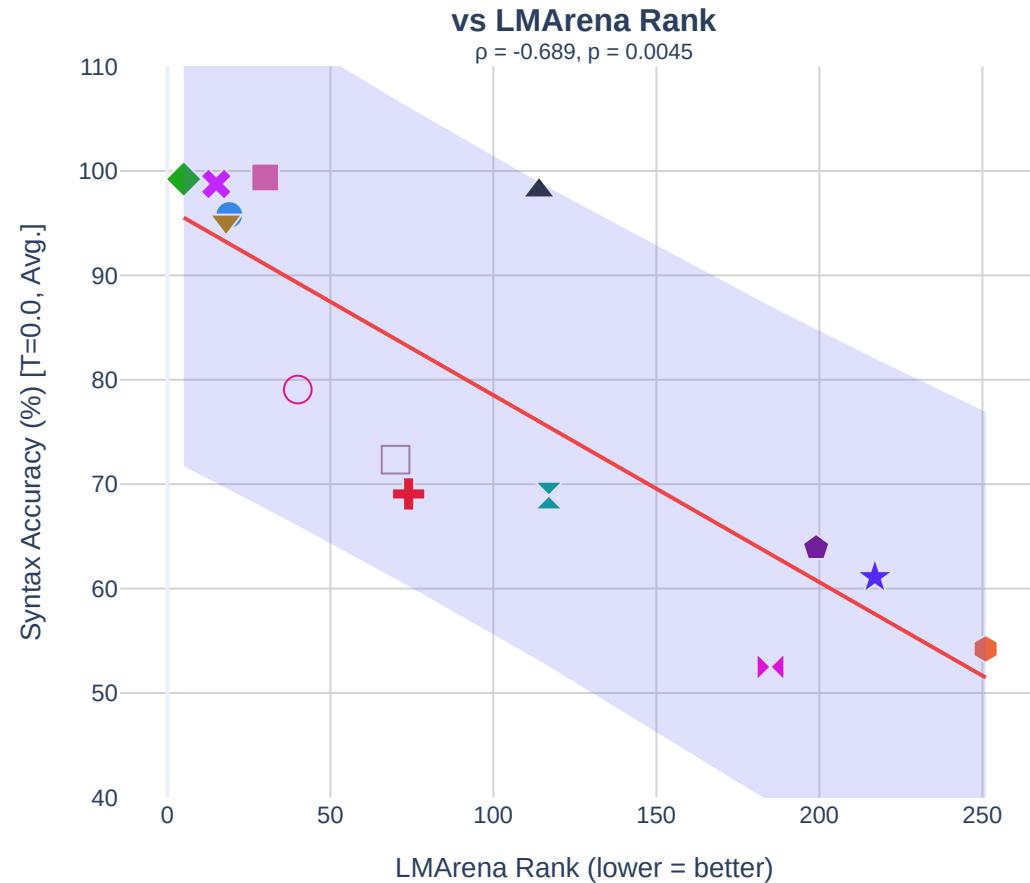


Benchmark Correlations: Syllogistic Reasoning vs External Benchmarks



- Models**
- deepseek-v3.1
 - gemini-2.5-flash
 - ◆ gemini-2.5-pro
 - ✚ gemma-3-27b-it
 - ✖ glm-4.6
 - ▲ gpt-oss-20b
 - ▼ kimi-k2-instruct
 - llama-3.1-8b-instruct
 - llama-3.2-1b-instruct
 - ★ llama-3.2-3b-instruct
 - ▲ llama-3.3-70b-instruct
 - mixtral-8x22b-instruct
 - qwen3-next-80b-a3b-instruct
 - qwen3-next-80b-a3b-thinking