Performance Distribution by Model Size Tier
Box Plot with Individual Data Points

Benchmark Score (%)

Model Tier

Large/Flagship (n=8) | Medium (n=9) | Small (n=2) | Unknown (n=3)

- - - High Performance
- - - 50% Threshold