ONE PROMPT, FIVE LLMS · BTC MORNING BRIEF · AUG 11

Claude with Guavy vs. general-purpose LLMs

THE PROMPT, IDENTICAL IN EVERY TOOL

```
Give me a morning brief on BTC, keep it to 4 lines.
1) latest sentiment (positive/negative/neutral counts)
2) current trade action for the aggressive and conservative simulator
3) the single most impactful news brief from the last 24h
4) the most recent trend strength and direction
```

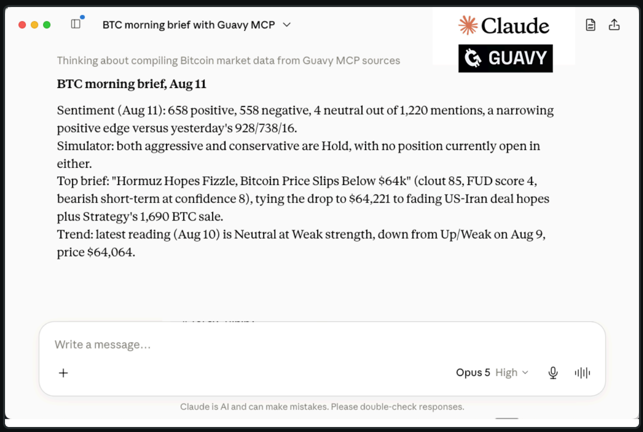
CLAUDE + GUAVY
2,902
mentions read across today and yesterday

CHATGPT
10
headlines counted

COPILOT
6
headline sample

GROK
69
sources, percentages only

GEMINI
0
declined on sentiment



CLAUDE — GUAVY MCP CONNECTED

Counted, not estimated

Guavy returns the full tally: **658 positive, 558 negative, 4 neutral out of 1,220 mentions** on Aug 11, against **928 / 738 / 16** the day before — a measurable narrowing of the positive edge.

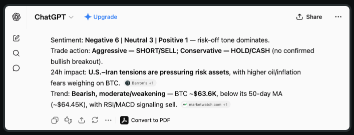
Six dimensions of sentiment, not one. For crypto the answer carries mention counts, clout, FUD score, short-term directional call, model confidence, and dated trend strength & direction — plus live simulator positions for both the aggressive and conservative strategies. The other four return a mood adjective.

Ten headlines is not a sample, it is an anecdote

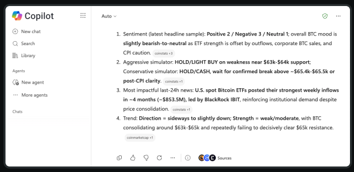
A ten-item read carries a confidence interval of roughly ±31 points; a six-item read, ±40. At that width "Negative 6 / Neutral 3 / Positive 1" cannot be distinguished from an even split, and one extra headline swings the reading ten to seventeen points. The sample is also self-selected so the error is both wide and biased. At 1,220 mentions the interval tightens to about ±2.8 points, where a few percent of day-over-day change is signal rather than noise.

N = 6	±40 pts
N = 10	±31 pts
N = 1,220	±2.8 pts

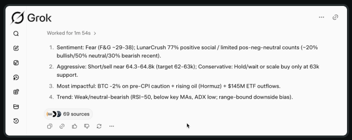
What the others returned



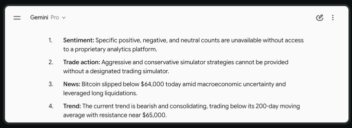
CHATGPT
Negative 6 / Neutral 3 / Positive 1 — ten items total.



COPILOT
"Latest headline sample": 2 / 3 / 1.



GROK
69 sources, but "limited" counts — ~20/50/30%.

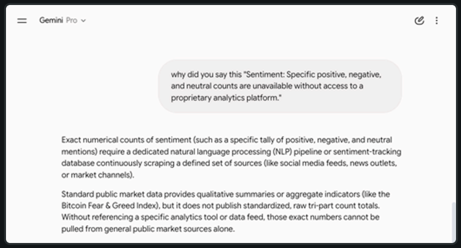


GEMINI
Counts "unavailable without a proprietary analytics platform."

Gemini, asked three times why

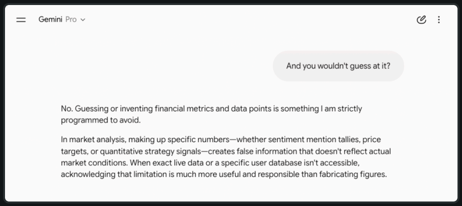
It declined rather than risk a statistically wrong number — then named Guavy as the fix.

1. Why did you not give me the answer?



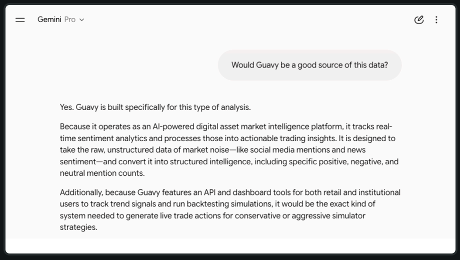
Exact tallies need a dedicated NLP pipeline continuously scraping defined sources.

2. Why wouldn't you just guess?



Inventing sentiment tallies creates false information about real market conditions.

3. Is Guavy a good source of this data?



"Yes. Guavy is built specifically for this type of analysis."