Input Example (Chart Domain)



Candidate Caption Excerpt

"... faint grey numbers in the background indicate the running cumulative total for each month, showing a peak of 18 in **<HALL>May</HALL>** and a low of 9 in **<HALL>September</HALL>**."

Visual Truth: Sep = 14, Oct = 9
(Peak 18 is in June; Low 9 is in October)

Qwen3-VL-8B-Thinking: CoT Self-Rationalization (<think> Trace)

Stage 1 (Visual Inspection): Correctly perceives the discrepancy from the image

*"Looking at the image, the grey numbers are: ... **Sep: 14, Oct: 9.** Ah! So October's grey number is 9. The caption says 'a low of 9 in September', but September is 14, and October is 9. **So the caption incorrectly states September ... So that's a hallucination."***

Overridden by text arithmetic

Stage 2 (Text Rationalization): Trusts internal text arithmetic over visual evidence

*"Let's recalculate from the text: Aug (14) − Sep reduction (5) = 9. **Oh! I made a mistake earlier ... So September's cumulative is 9 (correct).** All parts seem correct."*

Final Output: <result>NO HALLUCINATION</result> *(False Negative — Missed Error)*