Coding-agent usage report

Last 7 days · Claude Code · example data

Example data

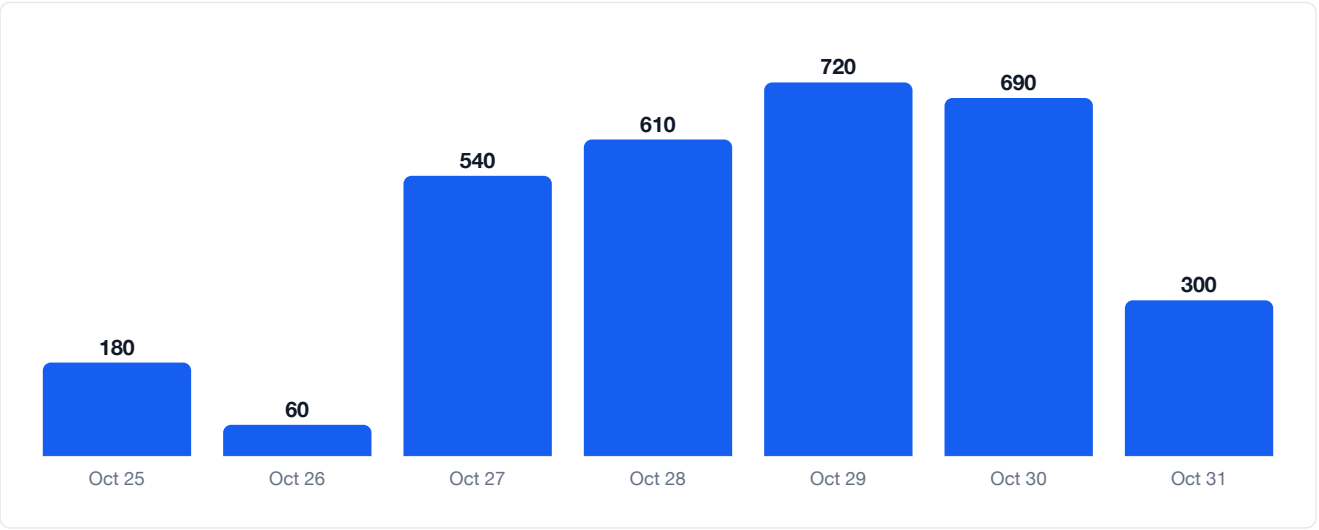
Invented figures in the layout of agentmeter --json. Assembled by hand.

Summary

Requests	Sessions	Input	Cache read
12,480	58	3.1 G tokens	97 %

Observed	2026-08-02 to 2026-10-31, 41,260 requests in all
Output	9.4 M tokens, 11 compactions (auto 11, manual 0)
Data quality	every line readable, 5 of 5 rules measurable

Input per day (M tokens)



Projects by input

Project	Sessions	Requests	Input share	Cache read
parser-lib	17	3,920	31 %	98 %
web-frontend	14	3,105	24 %	97 %
infra-scripts	9	2,210	18 %	95 %
docs-site	11	1,845	15 %	97 %
other (3)	7	1,400	12 %	96 %

Findings

MEDIUM

baseline-load: the median session start is 62k tokens

Today's configuration is 27k of the median start (44 %). Another 24k is in every start, 11k varies per session. Project web-frontend adds 6k through project instructions or MCP servers.

Check which skills and MCP servers are loaded at start but not called in the period.

MEDIUM

cache-rebuild: 96 rebuilds after pauses over 1 h

96 of 171 unexpected rebuilds (18.7 M tokens) follow pauses longer than one hour, which no cache lifetime covers. The remaining 75 are in sub-agent files that wrote their caches with the 5-minute lifetime.

Look at whether a longer cache lifetime reaches sub-agents. The logs do not show it.

LOW

long-session: 4 sessions span more than 24 hours

The longest session ran 53 hours with 4 compactions. Long sessions are not waste by themselves; they carry a large context into every request.

Check whether the long sessions were one task or a window left open.

Rules without a finding

- big-context: nothing found (2 % of input from contexts over 200k, threshold 10 %)
- fine: cache read share is 97 %, no day below 90 %