

Claude 4 Just Beat Every Human Tester on Real Programming Jobs

Summary

Anthropic released Claude 4 today, December 12, 2025, and independent testers gave it real coding tasks from companies. It solved 89% of them correctly on the first try – better than the average senior developer (81%). Companies can now plug it straight into GitHub to write and fix code automatically. Price stays the same as Claude 3.5, so millions more people will switch.

What It Means

Until yesterday, AI could help with simple code or fix small bugs. Claude 4 changes everything. Testers gave it messy, real-world projects – think banking apps, game backends, startup code full of bugs – and it understood the goal, wrote clean code, and passed all tests faster than most humans. For small companies

this is like hiring a top senior developer for \$20/month. Big companies will use it to cut teams or speed up launches. Normal programmers won't disappear, but anyone who doesn't use AI tools will feel very slow in 2026. Even weekend side-project builders can now ship apps that used to take months. It's the moment AI went from "toy" to "actual coworker".

Key Takeaways

- Beats senior developers on real company tasks
- Works inside GitHub – fixes and writes code automatically
- Same price as old version – instant upgrade for everyone
- Small startups now code 3–5× faster
- Big companies already testing it on live projects

Our Take (2026 Outlook) * Speculative

By summer 2026 half of all new code will be written with Claude 4 or something similar. Junior coding jobs will shrink fast, senior jobs will pay even more for people who direct AI.

References

Anthropic Blog • LiveCodeBench • TechCrunch (Dec 12, 2025)