AI Cost Analysis

1. Development & Testing Costs

1.1. LLM API costs

Source	Cost	Notes
OpenAI API	$0	Low usage and new account
OpenAI API	$21.28	Preloaded for use, not used

1.2. Total tokens consumed

Source	Tokens
OpenAI API	~78k

1.3. Number of API calls made

Source	Calls
OpenAI API	~90

1.4. Other AI-related costs

Type	Cost
Claude Code Max x20	$213.20
Web Domain Hosting	Free - Pre-existing self-paid domain (aaroncarney.me)

1.5. Total Costs

$234.48

2. Production Cost Projections

This projection estimates the operational costs of integrating AI into the collaborative board demo using the GPT-4o-mini model. This model is chosen for its high efficiency-to-performance ratio, making it ideal for the high-volume, iterative prompts common in digital whiteboarding.

2.1 Core Assumptions

Our cost analysis is built on three primary behavioral and technical pillars:

- **User Engagement Pattern:** Based on **B2B Productivity Benchmarks (IDC/Gartner)**, we assume a typical user engages in a "burst" pattern. They don't use AI constantly, but rather in clusters during active brainstorming sessions. We've modeled for a typical active user performing roughly a dozen sessions per month.
- **Context Density:** Following **SaaS FinOps Frameworks (Bessemer)**, we account for "Context Inflation." On a collaborative board, each new request must process more data as the board fills up. Our math uses a weighted average of the total data exchanged per user action to ensure the projection holds as boards get complex.
- **Model Efficiency:** We are utilizing **GPT-4o-mini-2024-07-18**, priced at $0.15 per million input and $0.60 per million output tokens. This allows us to handle structured data (like board coordinates and JSON) at a fraction of the cost of larger models.

2.2 Projected Monthly Costs

The following estimates reflect costs to provide AI features to a single user over a 30-day period based on common user profiles

User Profile	Usage Level	Monthly Cost Est.
Casual	Occasional sticky note generation	<$0.01
Typical	Regular brainstorming & clustering	$0.02 - $0.04
Power User	Heavy board summarization & diagramming	$0.10 - $0.15

Key Takeaway: Because the **GPT-4o-mini** model is so lean, the primary "cost" of the demo isn't the AI generation itself, but rather the cumulative context of the board. Even with high-density sessions, the cost per user is negligible—well below the typical margins for a standard SaaS subscription.

Following our unit cost analysis, we have modeled how these expenses scale as the user base grows. To provide a realistic forecast, we've applied a distribution model based on standard **SaaS Power User Curves (Reforge)** and **B2B Engagement Benchmarks (IDC)**.

2.3 User Distribution Assumptions

Users are not a monolith; their AI consumption varies significantly. Based on the **Pareto Principle (80/20 rule)** common in collaborative SaaS, we have categorized the projected 2026 user base into three profiles:

- **Casual (60%):** Viewers or light collaborators who rarely trigger AI.
- **Typical (30%):** Active team members using AI for daily brainstorming and cleanup.
- **Power Users (10%):** Facilitators and "Divas" who use AI for complex board summarization and diagramming.

2.4 Scaling Cost Forecast

Using the high efficiency of the **GPT-4o-mini** model ($0.15/1M input, $0.60/1M output), we have integrated the "Context Inflation" factor—the tendency for costs to rise as boards become more data-dense over a session (**Bessemer Venture Partners 2026 FinOps**).

Monthly Active Users	Monthly AI Operational Cost
100	**~$2.50 - 43.50**
1,000	**~$25 - $35**
10,000	**~$250 - $350**
100,000	**~$2500 - $3500**

2.5 Strategic Analysis

- **Extreme Margin Efficiency:** Because GPT-4o-mini is priced so aggressively, the total cost for 100,000 users is equivalent to a single entry-level salary. This suggests that AI costs are a **negligible percentage of COGS** (Cost of Goods Sold) for this demo.
- **Budget Ceiling:** Even if our "Power User" segment grew to 20% of the population, the total cost at 100k users would likely stay under $\$5,000/\text{month}$.
- **Free-to-Paid Viability:** The blended average cost per user is so low ($<\$0.04$) that a freemium model is highly sustainable without needing aggressive usage caps.

3. Sources

- **Reforge/CleverTap:** Power User Curve distribution models.
- **Bessemer Venture Partners:** 2026 SaaS Benchmarks on AI Gross Margins.
- **IDC/Gartner:** 2026 Productivity software engagement cycles.
- **OpenAI:** GPT-4o-mini technical pricing specifications.