Power

PHASE 2 OF 6:
BUILDING A MODERN TAM

Now Find Out What the Data Is Actually Telling You.

How to run pattern analysis on your enriched dataset, define guardrails that separate a qualified opportunity from a waste of everyone's time, and use AI to find the thresholds your gut feel never would have landed on.

Power

The Analysis Most Teams Never Run

Phase 1 left you with a fully enriched dataset:

Baseline firmographics validated against external sources, creative vertical-specific signals built through Claygents, and a clean separation between your best clients and your worst ones. You know what fields are reliable. You know which segments have enough records to work with.

Now comes the part most teams never get to. Not because it's hard, but because they never finished Phase 1 properly. Pattern analysis is where the data stops being a spreadsheet and starts being a system.

The question you're trying to answer is different from anything your CRM has ever asked: which of these signals actually predict whether a company will succeed as a client? Not which signals seem reasonable. Not which ones your best rep swears by. Which ones, when you run the distribution across every closed-won and every churned account, actually separate the two groups with statistical confidence. That answer is almost never what teams assume going in.

Power

Power

Your Failures Are Half Your Dataset

Most teams enrich and analyze only their closed-won clients

That gives you a description of what a good client looks like. It doesn't give you the edges: **where good clients end and where poor-fit ones begin. The edges are the whole point.**

Pull your churned accounts too. Specifically: **clients who left within the first 90 days, clients who stayed through one contract and disappeared, prospects who cleared every firmographic filter and got disqualified early in discovery.** Enrich them with the same fields you used for your wins.

What you're looking for is contrast. The enriched signals that cluster in your success group but are absent in your failure group are your most defensible guardrails. The signals that appear at roughly equal frequency in both groups are noise, regardless of how intuitive they sound when someone's explaining them in a pipeline review. If you skip this step, your floors will be set too low and your ceilings too high. You'll know what you're looking for. You won't know when to stop including companies. That is how a TAM quietly becomes an ad spend problem.

One question to settle before you run any analysis:

- What does "success" mean in your business?
- Retention through the first renewal?
- ARR expansion past a threshold?
- A specific NPS range?

The answer determines which accounts belong in your success group and which belong in your failure group. Teams that skip this definition spend days analyzing a comparison that was never properly constructed.

Power

Your Ranges Are Right. Where They Come From Is the Problem.

Once your enrichment is complete, with baseline fields plus creative signals and wins plus failures, you're ready to find the guardrails. **This is where AI does something human analysis genuinely can't: it looks at every signal, across every record in every segment, and identifies where the meaningful thresholds actually are.**

Here's where most teams go wrong: they set ranges by feel. We want clients between $10M and $100M. We want 50 to 500 employees. Those numbers sound reasonable. But if you ask where they came from, the honest answer is usually someone's memory of a few good deals, a round number that got written on a whiteboard, or a range that's been copy-pasted from one planning doc to the next for three years. Nobody ran the distribution. Nobody checked whether those boundaries actually reflect the clients who succeeded.

What you need per signal, per segment, is five numbers, not one:

Point	What it means
Floor	The hard minimum. Below this, the conversation ends regardless of anything else.
Sweet Low	Bottom of where your best clients actually cluster. Below here = manual review.
Median	The middle of your success group. Your clearest "yes" zone.
Sweet High	Top of the strongest fit range. Above here, signals start to diverge.
Ceiling	Upper bound. Above this, flag for a different conversation. Do not auto-exclude.

Power

In practice, this might surface a revenue floor of **$10M**, a sweet low of **$18M**, a median of **$45M**, a sweet high of **$95M**, and a ceiling of **$180M**. A company at **$12M** clears the floor but falls below the sweet low and goes to manual review. A company at **$60M** scores well. A company at **$210M** gets flagged for a different conversation, not removed. Three different outcomes from three positions in the distribution, handled by the same rule set automatically. No pipeline debate.

The moment your team stops asking "is this company big enough" and starts asking "where does this company sit in our actual distribution" is the moment qualification becomes a system instead of a negotiation.

THE PROMPT: Feed this into Claude or GPT after enrichment is complete

Act as a senior strategic data analyst.
I will provide a dataset of companies from a specific industry segment. Your job is to:

1. DATA PREP
- Clean and normalize all numeric fields (remove symbols, commas, text).
- Identify all fields with at least 4+ valid numeric values.

2. GUARDRAILS ANALYSIS
For each numeric field, calculate:
- Guardrail Low (10th percentile)
- Sweet Spot Low (25th percentile)
- Median (50th percentile)
- Sweet Spot High (75th percentile)
- Guardrail High (90th percentile)
- Number of observations (n)
Output in a structured table: Feature | p10 | p25 | p50 | p75 | p90 | n_obs

3. STATISTICAL SIGNIFICANCE
- Identify a primary outcome variable (prefer Company Revenue if available).
- For numeric fields: Run Spearman correlation vs. outcome.
- For categorical fields: Run Kruskal-Wallis test if applicable.
- Rank the top statistically significant features based on p-values and effect size.

4. FEATURE PRIORITIZATION
- Identify the top 5-10 most important fields based on statistical significance, data completeness, and business relevance.

5. INSIGHTS
- Summarize the ideal sweet spot company profile, red flag ranges (outside guardrails), and key differentiators of high-value companies.

6. OUTPUT FORMAT
- Table 1: Guardrails for all numeric fields
- Table 2: Top statistically significant fields
- Bullet summary of insights (concise, strategic)

Important: If dataset is too small (<5 rows), explicitly state that guardrails are unreliable.
Do not hallucinate thresholds. Only compute from data.
Favor non-parametric methods (Spearman, Kruskal-Wallis) due to small sample sizes.

Message ChatGPT

Power

Run this per segment, not across your full client base combined. A Series B company with 60 people and a VP of Marketing hired six months ago has a completely different fit profile than a 400-person company with a mature sales team and no growth marketing function. Both are B2B. Both might be in your TAM. The signals that predict success for one will not predict success for the other, and combining them into a single scoring model produces thresholds that are miscalibrated for both.

One more thing the AI will surface that human analysis misses: signals with insufficient coverage. If you enriched 300 clients for LinkedIn follower count but only 90 records returned a usable value, the model will flag that signal as unreliable. You want to know this before it ends up with weight in your scoring rubric.

Power

What You Actually Have at the End of This

At this point you have something most teams have never built:

an evidence-based picture of who you serve well,
who you don't, and where the lines between them are.

Two types of guardrails will come out of this analysis.

- **The first type is the hard cut:**

any signal so clear that anything below a certain threshold is out, regardless of how it performs on everything else. Revenue floors are the most common. These come directly from the p10 of your success group for that segment. Not from a gut call. Not from the most recent painful churn that showed up in Slack.

That last point matters more than it sounds. The instinct to add new hard cuts spikes after a bad client experience. Someone churned in 60 days. Someone adds a rule: no more companies under $15M, no more early-stage anything. The rule feels logical because it's attached to something real. But one bad account is an anecdote. A pattern across 40 accounts is a rule. Every hard cut in your guardrails should be traceable to the analysis, not to post-mortem decisions made in the heat of a bad quarter.

- **The second type of guardrail is everything else:**

the signals that matter but don't individually disqualify. These get weighted, scored, and built into a rubric, which is what Phase 3 covers. For now, what you need is a ranked list of signals with their five-point ranges, documented by segment, with a clear note on which ones will function as hard cuts and which will feed the scoring system.

Power

Before You Move to Phase 3

Check that you have four things in place: five-point guardrail ranges for every statistically significant signal, documented per segment; a ranked list of top signals per segment with the AI's reliability notes; a defined set of hard-cut signals per segment grounded in your p10 distribution; and a clear record of which enrichments didn't survive the pressure test and why.

If any of those are missing, what you build next will inherit the gaps. The guardrails aren't just a deliverable. **They're the filter you're about to apply to tens of thousands of companies you've never met.**

Next: Building the Universe You'll Filter Against

Most teams build their TAM by starting narrow. A database, a few filters, a list that feels right. The problem isn't the tool. It's the sequence. You filtered before you knew what you were filtering for.

Now you do. Phase 3 is about going intentionally broad, pulling the full market universe before applying a single guardrail, so that when you do filter, you're not guessing at coverage. You're carving a defensible TAM out of a complete picture, not hoping the right companies made it into a pre-filtered list.

The guardrails you just built tell you who belongs. Phase 3 makes sure you're looking at everyone who could.

This is Part 2 of a 6-part series on building a modern, data-driven Total Addressable Market.

Next: Phase 3 - Market Universe Creation