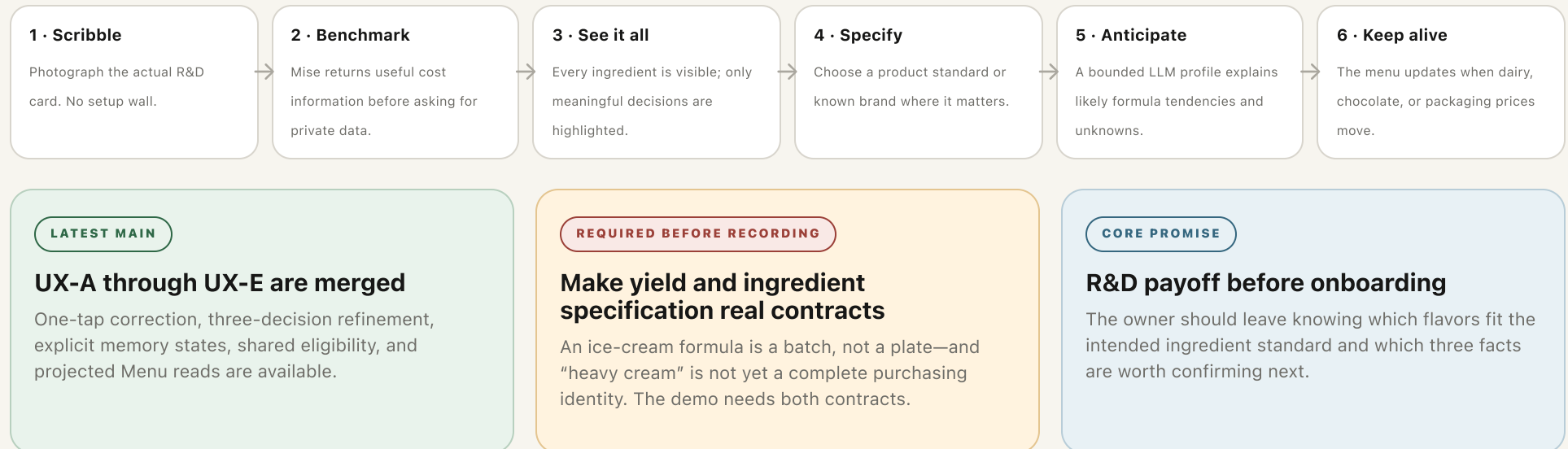


From a scribbled ice-cream formula to a launch-ready menu.

A recording plan for a new shop owner doing R&D before opening: capture one real kitchen note, get an honest benchmark, choose the ingredient specifications that define the product, and watch the answer become a living recipe book.



Do not demo “recipe software.” Demo a pre-opening decision getting easier.

PERSONA

Maya is opening Juniper & Cream in Austin.

She has test batches, proposed menu prices, and no invoice history. She needs to decide which flavors can support the shop before she commits to vendors and printed menus.

Her question: “Can I sell this Brown Butter Honeycomb scoop for \$6—and which cream should I build it around?”

- ✓ Cost the test recipe using an honest market benchmark.
- ✓ Show uncertainty instead of treating missing lines as \$0.
- ✓ Tell her which answers materially improve the decision.
- ✓ Keep the recipe useful when supplier quotes change.

RECORDING OUTCOME

The viewer should understand seven things without narration gymnastics.

1. Mise reads the kitchen artifact the owner already has.
2. It produces an immediate, caveated result—not a setup form.
3. It turns “heavy cream” into a chef-meaningful specification choice—not a catalog dump.
4. It shows the entire formula, while highlighting only the few decisions worth attention.
5. It describes likely product tendencies from verified inputs and names what it cannot know.
6. Every cost can be opened and traced to evidence.
7. The first recipe becomes a living menu, not a disposable report.

NARRATIVE BOUNDARY

Zero-state means benchmark, not “true cost.”

Maya has no purchase history yet. Say “market benchmark,” “known cost,” “lower bound,” and “coverage.” Only her own prices can earn a kitchen-backed result.

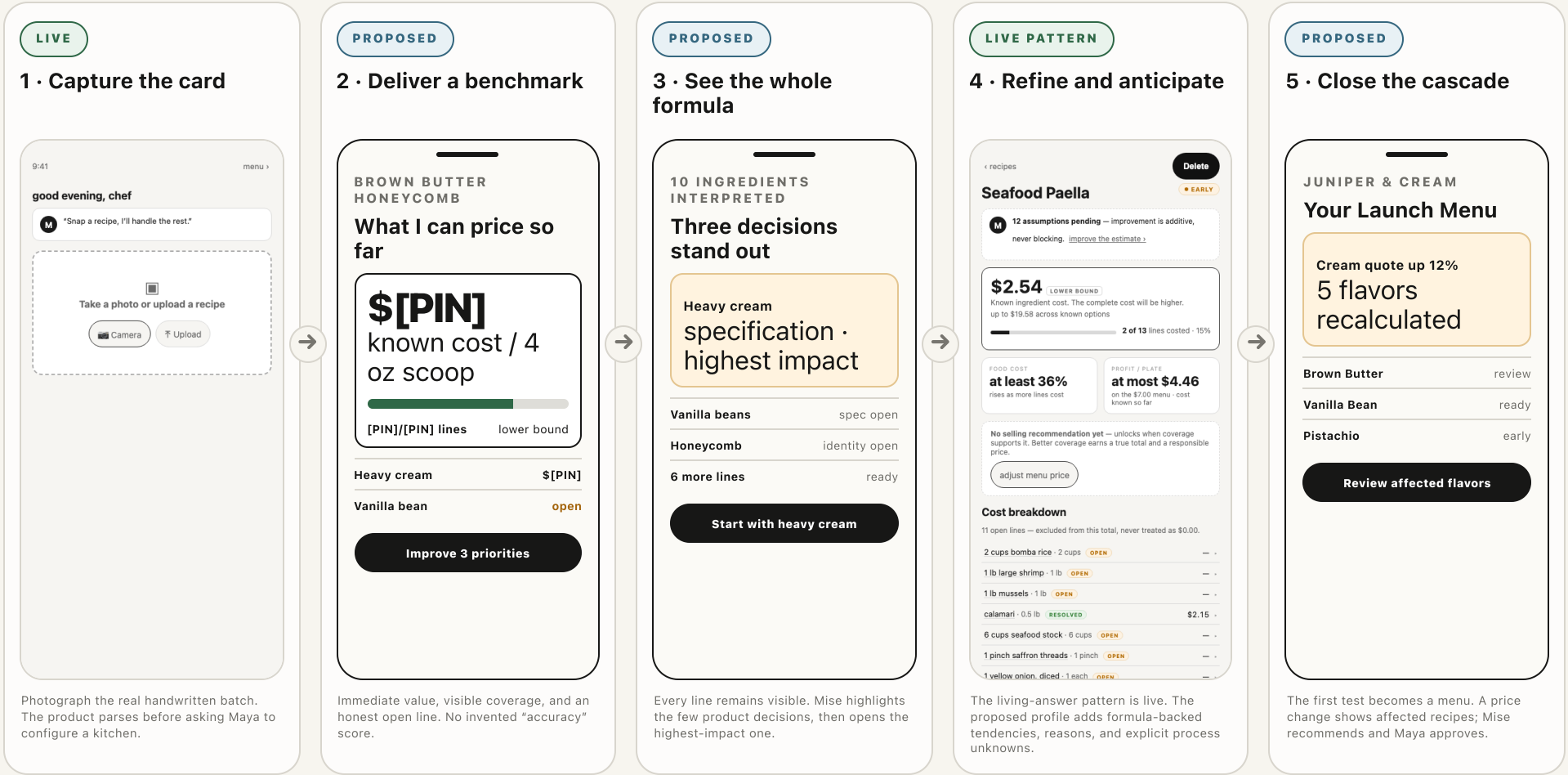
Ideal length

100–120s

Long enough to show the loop; short enough to preserve the magic of “scribble in, decision out.”

Five screens carry the owner from “I have a recipe” to “I have a launch book.”

REAL CURRENT SCREEN **PROPOSED THEMED SCREEN** The scene pages that follow enlarge each screen and provide Edith's exact action and voiceover.



Visual rule for the recording: the real handwritten card creates the real filmed recipe. Any screen not yet live is labeled proposed in this playbook and must be implemented and run before Edith records it as product.

The latest product is closer—but the demo still has seven hard prerequisites.

SAFE TO DEMONSTRATE NOW

Capture	Camera/upload, parsed-line preview, silent costing.
Result first	Known cost, lower-bound language, coverage, open impact.
Refinement	Top three material decisions, recost and re-rank after each.
One tap	Recipe-only correction applies immediately with Undo; kitchen rule is deliberate.
Memory	Recipe answer, kitchen rule, and paused-rule states are visible.
X-ray	Chef-language identity, price source, calculation, assumptions, relevant repair.
Menu	Coverage/readiness, attention dollars, price-move cascade, cached projections.

RECORDING BLOCKERS

LLM disabled locally	No LLM_API_KEY or Anthropic key is configured here. An uploaded photo falls back to Mushroom Risotto.
Batch yield absent	Capture does not extract servings/yield. The current engine path sums captured ingredient lines as the displayed serving cost.
Specification/brand is not first-class	The chooser can resolve identities, but R&D needs verified butterfat/formulation/package attributes and an optional known brand/SKU presented together.
No bounded product-profile contract	The LLM parses and flags today; there is no endpoint that turns verified formula facts into a versioned, caveated sensory tendency profile.
Wrong vocabulary	Current screens say dish, plate, and menu price. Ice cream needs batch, scoop, cup, and pint language.
Dirty rich DB	The current Menu contains duplicate/non-themed recipes and implausible legacy outcomes. Record only against a clean synthetic tenant.
Cold refinement	Rich-DB Paella questions took about 19 seconds locally. The mock corpus must stay small and the route must pass a recording-time threshold.

Observed Menu

316 ms in the local mobile audit after projections were warmed. Environment-specific, not an SLA.

Do not disguise waits

Pre-warm seeded views. A newly captured recipe must still complete honestly; edit only dead time, never a failed state.

Money boundary

The script uses placeholders until the ice-cream fixture checker pins every displayed number.

Seed a small menu that makes the cascade obvious.

ORGANIZATION

Juniper & Cream

SYNTHETIC

Austin, TX · one planned location · founder/admin · pre-opening R&D · no invoice history.

Menu seeded for the recording

RECIPE	DEMO ROLE	STATE
Brown Butter Honeycomb	Hero capture	New benchmark
Vanilla Bean	Shared dairy sensitivity	Ready
Dark Chocolate Sea Salt	Commodity volatility	Estimate
Strawberry Buttermilk	Seasonal fruit exposure	Estimate
Pistachio Cardamom	High-cost specialty line	Early
Lemon Basil Sorbet	Dairy-free comparison	Ready

EVIDENCE DESIGN

Use the smallest dataset that exercises the real ladder.

- ✓ Clean synthetic market aggregates for dairy, eggs, sugar, flavorings, mix-ins, cups, and spoons.
- ✓ Verified cream attributes—such as butterfat percentage, treatment, formulation, and package—plus fictional brand/SKU examples.
- ✓ One honest unpriceable or low-coverage line so zero-state honesty remains visible.
- ✓ Three carefully seeded material ambiguities—never twelve contaminated catalog rows.
- ✓ A small set of owner-stated supplier quotes for the recipes marked Ready.
- ✓ One later heavy-cream quote change that touches five flavors and fires the forever loop.

PRIVACY

No real vendor prices in the recording fixture.

Create synthetic price points or vendor-free market aggregates. If the demo shows a vendor, it must be Maya's fictional own supplier—not a name/price pairing from another kitchen.

Never upload the answer key

The photographed recipe contains quantities and yield only. Expected prices, conversions, and pinned results stay in the fixture checker—not in the capture asset.

Brown Butter Honeycomb should exercise costing, yield, ambiguity, and packaging at once.

Brown Butter Honeycomb — test #3

2 qt heavy cream — 36%? local dairy?

2 qt whole milk

2 lb sugar

18 egg yolks

4 oz skim milk powder

2 vanilla beans

1 tsp kosher salt

8 oz unsalted butter — brown it

1 lb honeycomb candy, crushed

48 × 6 oz cups + 48 spoons

after churn: ~48 × 4 oz scoops?

menu test = \$6 / scoop

WHAT THE MODEL MAY EXTRACT

- ✓ Recipe name and raw ingredient lines exactly as written.
- ✓ Quantities, units, and a proposed serving/yield statement.
- ✓ Culinary language and tentative specifications such as “36%?” for interpretation—not silent acceptance.

WHAT MUST REMAIN DETERMINISTIC

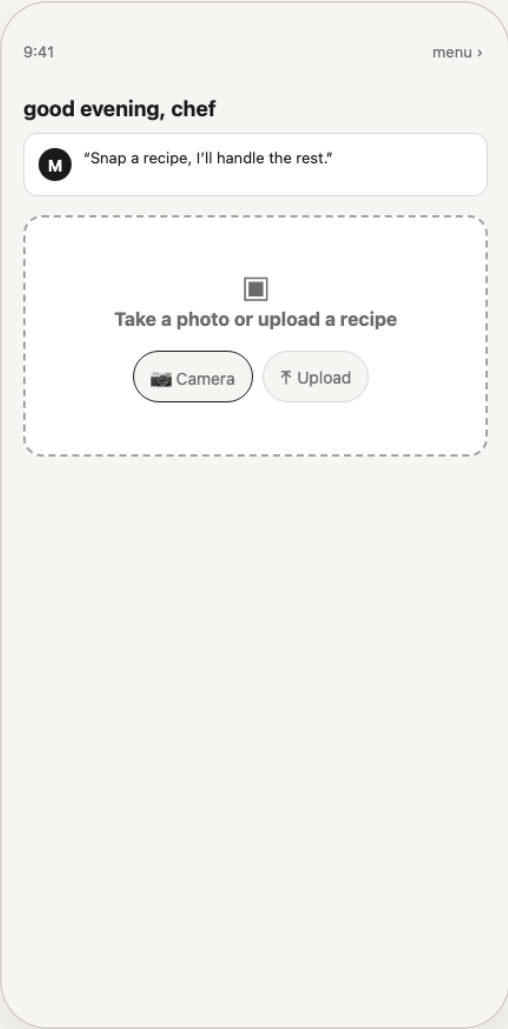
- ! Whether “48 × 4 oz scoops” is accepted as the batch yield.
- ! Batch cost ÷ confirmed yield + packaging cost per serving.
- ! Every ingredient price, unit conversion, cost, margin, and dollar swing.

Capture acceptance

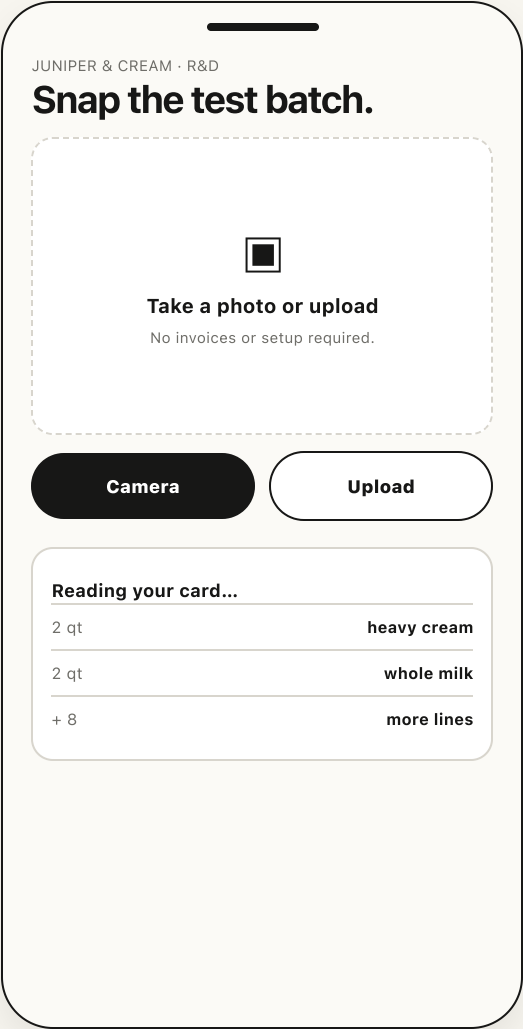
Run this image through the configured vision model three times. It passes only if the ingredient lines and yield are stable enough to produce the pinned fixture result—or if the UI clearly asks Maya to confirm the material difference.

The question mark after “48 scoops” is intentional. Mise should not silently bless a tentative R&D yield; it should convert it into one chef-readable confirmation before any per-scoop economics are shown.

Open on the artifact. Let the first interaction feel effortless.



Live screen. Replace the generic environment with the Juniper & Cream demo tenant; the interaction itself is already correct.



Small copy addition. Use the tenant name and "test batch"; preserve the zero-setup capture.

VOICEOVER

"Maya is opening an ice-cream shop. Her recipes are still handwritten because she's still testing. She takes one photo—Mise starts costing before she has to set anything up."

What Edith does on screen

- 1. Begin with the physical recipe card in frame.
- 2. Open Mise directly to Capture.
- 3. Photograph the card; hold on the parsed-line preview.
- 4. Let the product transition into silent costing.

Recording rule

Do not type the recipe into the DB and pretend the photo created it. The filmed artifact must be the input that created the filmed recipe.

Pay off the capture before asking Maya to improve anything.

BROWN BUTTER HONEYCOMB

Here's what I can price so far ESTIMATE

\$[PIN]

Known ingredient cost per 4 oz scoop. The complete cost may be higher.

[PIN] of [PIN] lines

fixture coverage

LOWER BOUND

Heavy cream

category benchmark · specification open

\$[PIN]

Whole milk

market benchmark

\$[PIN]

Vanilla beans

largest open impact

open

Improve the estimate — 3 priorities >

or use this benchmark now

VOICEOVER

"Before asking Maya twenty questions, Mise gives her a useful benchmark. It shows what is known, what is still open, and never pretends a missing ingredient costs zero."

Show

Cost per **4 oz scoop**, coverage, evidence state, top open line, and Maya's \$6 test price.

Do not show

"Accuracy 82%," "true cost," or a finished margin unless the fixture has ground truth and kitchen-backed coverage.

NUMBER DISCIPLINE

Replace every [PIN] from the checker output.

No illustrative number survives into the recording. If the result changes after a dataset update, repackage the fixture and regenerate the shot list.

Product principle: payoff creates the willingness to answer. A blank "we need more information" state would lose the pre-opening owner here.

Show the whole formula. Highlight the few decisions hiding inside it.

10 INGREDIENTS INTERPRETED

Your formula is mapped

6 ready 3 product decisions · 1 assumption

Heavy cream
HIGH IMPACT

specification open · choose 36%, 40%, or known brand

Vanilla beans
DECIDE

grade/size remains open

Honeycomb candy
DECIDE

product identity remains open

Whole milk	benchmark ready
Sugar	benchmark ready
Egg yolks	benchmark ready
Milk powder	benchmark ready
Butter	benchmark ready
Salt	assumption carried
Cups + spoons	package estimate

Start with heavy cream

LLM SYNTHESIS

PRELIMINARY

Expected product profile

Formula-backed tendencies before a test note or process record exists.

Body & richness · likely fuller

Egg yolks and milk powder push the formula toward more body than the named reference base; the open cream specification could move this profile.

Sweetness · likely pronounced

Sugar plus honeycomb candy contribute two visible sweetness sources.

Texture · dense/smooth tendency

Formula suggests this direction; overrun, churn temperature, aging, and freezing process remain unknown.

Updates live: choosing 40% cream refreshes the profile and names exactly which input changed.

VOICEOVER

"Mise does not hide the recipe behind three questions. Maya can see every interpreted ingredient, which lines are already usable, and the few choices that still define the product. It also gives her a preliminary formula profile—clearly separated from an actual tasting result."

The profile's evidence contract

- ✓ Deterministic input packet: quantities, verified identities/specifications, formula ratios, and a named comparison base when one exists.
- ✓ LLM job: synthesize qualitative tendencies and plain-language reasons.
- ✓ Output carries the inputs used, model/version, process unknowns, and "directional" state.

What the model may not do

It may not invent fat percentage, call a brand "premium," promise a sensory outcome, compute money, or compare against "typical ice cream" without a named benchmark formula.

Attention rule: the list is complete; the action queue is selective. Visibility never becomes a ten-question verification wall.

Three decisions—not the entire ingredient database.

< back

M 10 things would move your plate cost — answer them right here.

DISH NAME · TAP TO EDIT

Seafood Paella

First, the biggest movers top 3 of 10
 Answer one and the next-biggest takes its place — you never face more than 3 at once.

? Which seafood stock?
 swings this cost ~\$58.90
 now reads as **6 cups seafood stock**

choose the right one >

? Which large shrimp?
 swings this cost ~\$10.33
 now reads as **1 lb large shrimp**

choose the right one >

? Which bomba rice?
 swings this cost ~\$1.41
 now reads as **2 cups bomba rice**

choose the right one >

Auto-resolved by dollars show 2
 2 questions auto-resolve — their deltas were pennies. Tap "show" to override.
 ✓ 1 line matched from your kitchen's memory and catalog.
 Costed so far: **\$2.54** — a lower bound. Answering the movers above sh

Use this estimate >

Live behavior. The queue is capped at three and re-ranks after each answer. The mock fixture must produce ice-cream-specific questions in this same pattern.

FIRST, THE BIGGEST MOVERS · 1 OF 3

Which cream are you building around?

Butterfat, formulation, package, and a known brand can materially change the scoop benchmark.

36% foodservice cream
 benchmark specification · quart cartons · \$[PIN] / scoop

40% cream
 Maya's R&D choice · verified higher-butterfat specification · \$[PIN] / scoop

Choose a brand or supplier item
 curated, priceable matches only—never the raw catalog

I'm not sure yet
 carry the market range and keep testing

One tap fixes this recipe and recosts it. Undo remains available. "Make it a kitchen rule" is a separate choice.

VOICEOVER

"Mise ranks ambiguity by dollars, then translates it into a product decision. Maya chooses 40% cream for this R&D batch—not because Mise calls it better, but because that is the product standard she wants to test. One tap recosts the scoop and refreshes the expected profile."

Seed the three expected decisions

1. **Heavy cream specification** — butterfat/formulation first; package and brand inside the choice.
2. **Vanilla bean specification** — grade/size plus count-to-weight uncertainty.
3. **Honeycomb identity** — purchased candy or intended brand, not cereal or a contaminated catalog row.

Memory moment

After the first pick, briefly show "Fixed for this recipe," Undo, and the optional "Make it a kitchen rule." Do not infer that the most expensive brand is "best," or automatically apply an R&D answer to every future flavor.

Fixture acceptance

The dollar gate—not the script—must put these in the top three. If a different decision ranks higher, fix the fixture evidence or record the honest result.

VOICEOVER

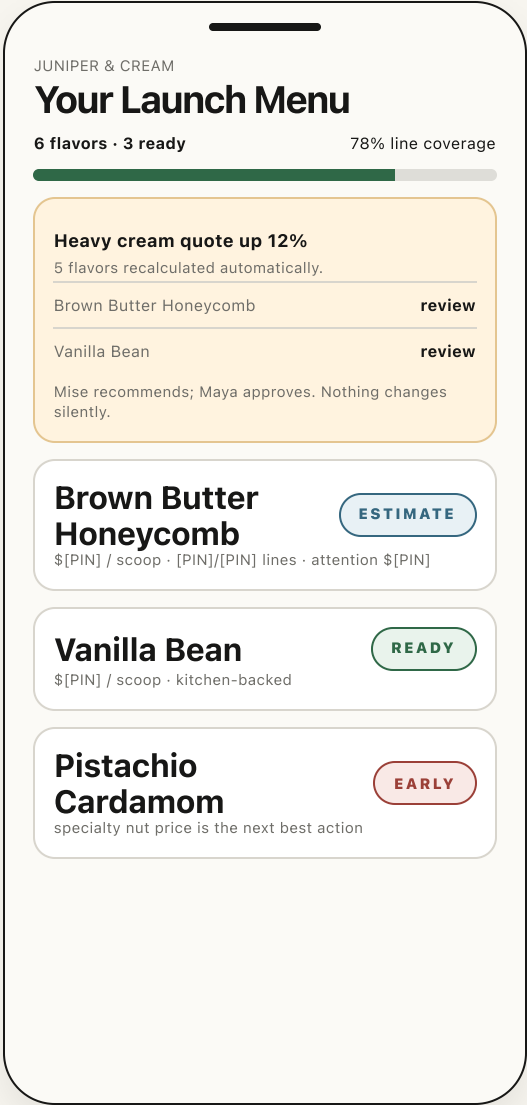
"Now Maya has a cost per scoop she can use for R&D, and the expected profile has refreshed around the confirmed 40% cream. She can open the ingredient and see the specification, whether a brand is known, the price source, and the package assumption carrying the number."

Open heavy cream because it connects ingredient identity, functional specification, optional brand/SKU, package conversion, market evidence, and the option to replace the benchmark with Maya's own supplier quote.

- ✓ **Benchmark:** known cost/lower bound and coverage.
- ✓ **Owner price entered:** chef-stated evidence, still scoped and dated.
- ✓ **Kitchen-backed:** true cost only after every material line rests on Maya's evidence.

Body/richness tendency strengthened. Changed input: cream specification confirmed at 40% butterfat. Still unknown: overrun, aging, churn temperature, and freezing process.

Zoom out: one test batch becomes a menu that keeps recalculating.



VOICEOVER

“That recipe now lives in Maya’s launch menu. Mise shows what is ready, what is still an estimate, and where the next answer matters most. When her cream quote moves, every affected flavor is recalculated—and Maya approves the response instead of rebuilding the book.”

Current benefits shown

Portfolio coverage, readiness, attention dollars, fast projected Menu, price-move cascade, recommended next action.

Natural next action

After value, invite Maya to add supplier quotes or invoices so benchmark evidence is progressively replaced by her own.

Better than “popular nearby flavors”

Show **economic sensitivity** the product actually knows: which flavors are most exposed to dairy, cocoa, pistachio, and packaging. This is actionable and ledger-backed.

Do not claim local demand intelligence

Mise currently has no location-aware price cohort or consumer-flavor dataset. “Austin loves honeycomb” would be invented. Popularity requires a named external demand source and a separate provenance contract.

Build serving and specification contracts first. Treat location intelligence as a later data product.

PRIORITY	CHANGE	WHY THE OWNER NEEDS IT	ACCEPTANCE	OWNER
P0	Batch → serving contract	Ice cream is produced by batch but sold by scoop.	Capture stores confirmed yield; deterministic engine returns batch cost and cost per serving. No model arithmetic.	Engine + Product
P0	Specification-aware identity	“Heavy cream” is too broad for serious R&D.	Chooser exposes verified butterfat/formulation/package attributes, optional brand/SKU, and resulting engine cost—never “premium” by price alone.	Product + Catalog + FE
P0	Bounded formula profile	The owner wants to anticipate the product before committing to a test.	Deterministic facts feed the LLM; output names tendencies, reasons, unknown process variables, model/version, and directional state. No model dollars or invented specs.	Product + Agent
P0	Serving vocabulary	“Plate” makes the demo feel repurposed.	Recipe carries a display label such as scoop/cup/pint; all cost and margin labels use it consistently.	Product + FE
P0	Packaging inside yield	Cups and spoons materially change a low-ticket item.	Batch ingredients plus 48 cups/spoons divide into the confirmed 48 servings and ledger separately.	Engine
P0	Idempotent demo seed + checker	The recording must be reproducible after data changes.	One command seeds the synthetic tenant; one checker pins every displayed number, top-three question, and cascade.	Engineering
P0	Vision-key preflight	Without a key, the photo silently becomes the fallback Risotto demo.	Startup/readiness check fails loudly when the ice-cream recording mode lacks vision.	Engineering
P1	R&D context card	Target portion, menu test price, yield confidence, and opening stage make the output useful.	Skippable after capture; every field states what it improves.	Product + FE
P1	Question/gate latency	A long cold refinement load breaks the recording and the product.	Mock tenant result/refinement loads within the agreed threshold; gate computed during background costing where possible.	Backend
LATER	Location-aware benchmark	Regional costs may improve zero-state usefulness.	Only display a city/region when the evidence pool is actually segmented and carries cohort provenance.	Data + Engine
LATER	Flavor popularity	Demand insight could help pre-opening menu design.	Requires an external demand/POS/search source, freshness, geographic coverage, and explicit provenance. Not part of this demo.	New data product

Recommended cut: P0 serving contract + specification-aware chooser + bounded profile + clean fixture are mandatory. P1 R&D context remains useful. Do not bolt on speculative flavor recommendations.

Make the recording state rebuildable, not a precious local database.

RECOMMENDED PACKAGE

An idempotent management command plus sanitized fixture assets.

Tenant	Juniper & Cream organization, user, menu, R&D context.
Recipes	Six themed recipes with stable semantic identifiers—not raw drifting IDs.
Catalog	Only identities, verified specifications, synthetic brands/SKUs, conversions, packaging, and activity needed by the menu.
Evidence	Synthetic market aggregates and a few fictional owner quotes.
Memory	Only deliberate demo rules; no residue from trial runs.
Cascade	One reproducible heavy-cream price movement with known affected recipes.
Assets	The exact no-cost recipe card image used in the recording.

REPEATABLE WORKFLOW

1. Set a versioned demo-data constant.
2. Seed/update by stable tenant and recipe keys.
3. Run the deterministic fixture checker.
4. Build recipe projections.
5. Run the capture asset three times in a disposable transaction or explicitly delete created IDs.
6. Capture screenshots and record only after all pins pass.

When the dataset changes

Bump the demo-data version, re-run the sanitizer/seed, regenerate the expected-output manifest from the deterministic engine, and re-run the recording checklist. Never hand-edit a dump until it "looks right."

Avoid a shared rich DB dump

It is harder to audit, easy to contaminate with test memory, and risks carrying private vendor data. Keep the rich corpus separate; package the narrow synthetic demo graph.

Suggested commands can be named `seed_ice_cream_demo` and `check_ice_cream_demo`, but the exact implementation belongs in Edith's build plan—not as shell folklore in the recording notes.

A two-minute recording with no unsupported claims.

TIME	SCREEN/ACTION	VOICEOVER OBJECTIVE	PROOF ON SCREEN
0–8s	Show scribbled card → open Capture.	<i>"Maya is still testing recipes before opening."</i>	Physical artifact and zero-setup capture.
8–18s	Photograph/upload → parsed-line preview.	<i>"Mise reads what she already uses."</i>	Correct name, ingredients, yield; no prices in input.
18–35s	Result-first benchmark.	<i>"Useful now; honest about what remains open."</i>	Pinned per-scoop lower bound, coverage, top open impact.
35–50s	Show the complete formula + preliminary profile.	<i>"Everything is visible; only three decisions need attention."</i>	All ten lines, state per line, formula-backed tendencies, process unknowns.
50–70s	Choose cream specification or known brand.	<i>"Only the decisions that move money or define the product."</i>	Verified attributes, optional brand/SKU, top 3 of N, one-tap recost, Undo.
70–75s	Show recipe-only memory → optional kitchen rule.	<i>"Answers become memory only at the right scope."</i>	Recipe state and deliberate kitchen-rule action.
75–95s	Living answer → refreshed profile → cream X-ray.	<i>"The expected profile and every number explain themselves."</i>	Changed input, profile reason, identity, specification, brand state, source, package.
95–105s	Return to themed Menu.	<i>"One formula becomes a launch book."</i>	Six flavors, coverage/readiness, next best action.
105–114s	Open heavy-cream price movement.	<i>"Mise closes the cascade; Maya approves."</i>	Five affected flavors, reliable vs directional changes.
114–120s	End on "Improve with your prices."	<i>"Supplier quotes and invoices make the benchmark hers."</i>	Progressive data invitation after demonstrated value.

Technical preflight

- ✓ HEAD equals approved recording SHA.
- ✓ Vision key configured; fallback disabled.
- ✓ Seed/check commands green.
- ✓ Menu projections built.
- ✓ Capture and refinement meet time threshold.

Product preflight

- ✓ Every dollar equals checker output.
- ✓ Batch yield visibly confirmed.
- ✓ Cream choice shows verified specification and optional brand—not a price-defined quality label.
- ✓ Every profile claim names its formula inputs or an unknown; profile is directional, never an "accuracy" score.
- ✓ No more than three decisions shown.
- ✓ Benchmark never called true cost.
- ✓ One X-ray and one memory echo work live.

Privacy/claim preflight

- ! No real vendor data or cross-kitchen attribution.
- ! No answer key inside the uploaded card.
- ! No "popular in Austin" claim.
- ! No hidden manual DB correction between shots.
- ! No screenshot of a flow that was not run.

Definition of done: a new viewer can repeat the owner’s story in one sentence: “Mise turns my test recipe and ingredient standards into an honest cost per scoop, asks only what matters, and keeps the whole launch menu current.”