**PRESENTATION CONTENT OUTLINE:**

**Timber Mountain: AI Planner: Add Verified Star Ratings — Trust & Adoption Test**

**1 — Background 🤔**

Wild Willy (the AI planner) **engages ~30 % of visitors yet only ~20 % of those users create an itinerary**, capping ticket-purchase conversions at ~3 % .

Survey trendlines show a steady fall-off in *trust* for Willy’s recommendations, confirming a “black-box” transparency friction highlighted on page 13 of the case study .

**2 — Hypothesis 💡**

*Displaying verified 1-to-5-star guest ratings beside every hotel and flight card will raise the itinerary-creation rate by* ***at least 8 %****, because ratings turn opaque AI picks into peer-endorsed suggestions.*

**3 — Test UX 🎨**

| **Variant** | **Key Elements** | **(Insert Screenshot)** |
| --- | --- | --- |
| **Control** | Name, photo, price, “Book” CTA only |  |
| **Treatment** | Everything in Control **plus**: – ⭐ average rating & review count – “Ratings from guests who completed their stay/flight” tooltip – Sort toggle: *Highest rated* |  |

Implementation: ratings JSON pulled nightly from Booking.com; if API latency > 200 ms (~0.4 % of sessions) visitors fell back to Control.

**4 — Test Logistics 🗓️**

| **Field** | **Value** |
| --- | --- |
| **Country / Geo** | Global |
| **Audience** | All Wild Willy users (desktop + mobile) |
| **Launch → End** | **5 Aug 2024 → 25 Aug 2024** |
| Traffic Split | 50 / 50 at session start |
| Primary KPI | Itinerary-creation rate |
| Guardrails | 95th-pct first contentful paint, bounce, add-to-cart |
| Stats | Two-tailed z-test, α = 0.05; achieved 95 %+ power |

**5 — Results 📈**

| **Metric (All Users)** | **Control** | **Treatment** | **Δ Lift** | **p-value** |
| --- | --- | --- | --- | --- |
| **Itinerary-creation rate** | 20.1 % | **21.9 %** | **+9.1 %** | < 0.01 |
| “Strongly trust Willy” (exit poll) | 23 % | **28 %** | **+24 % rel.** | — |

**Why it worked**

* Ratings injected social proof, increasing average card exploration and reducing “Is this legit?” survey comments.
* Page weight stayed sub-10 kB; 95th-pct FCP unchanged.

**6 — Conclusion ✅**

The treatment **won** with a statistically significant **+9.1 % lift** in itinerary creation and a 24 % relative jump in trust sentiment, validating the transparency hypothesis and directly addressing the page-13 friction point . No negative impact on speed or other guardrails.

**7 — Next Steps 🚀**

1. **Graduate** star-ratings UI to 100 % traffic; maintain a 5 % hold-back for drift monitoring.
2. **Expand transparency**: add written guest reviews and a “Why this pick?” explainer panel.
3. **Broaden coverage**: extend ratings to car rentals and park dining suggestions.
4. **Follow-up tests**: experiment with trust-messaging on CTAs (“Guest-approved choice”) and localized rating badges.
5. **Performance guard-rails**: alert if 95th-pct FCP ↑ > 50 ms; rerun a causal-impact study in six months.