Can we choose recipes that draw more traffic to the Tasty Bytes homepage ?

Testing a homepage selection rule

HELD-OUT TEST RESULTS

86.7%

held-out precision

78 / 90

selected recipes were High

80% target

95% interval: 78.1–92.2%

The decision is which recipe to feature each day

Today

The manager selects a favorite from available recipes.

Popular selections have coincided with up to 40% more traffic elsewhere on the site.

Proposed

Use model scores to rank available recipes.

The primary metric is precision: the share of selected recipes that lead to High traffic.

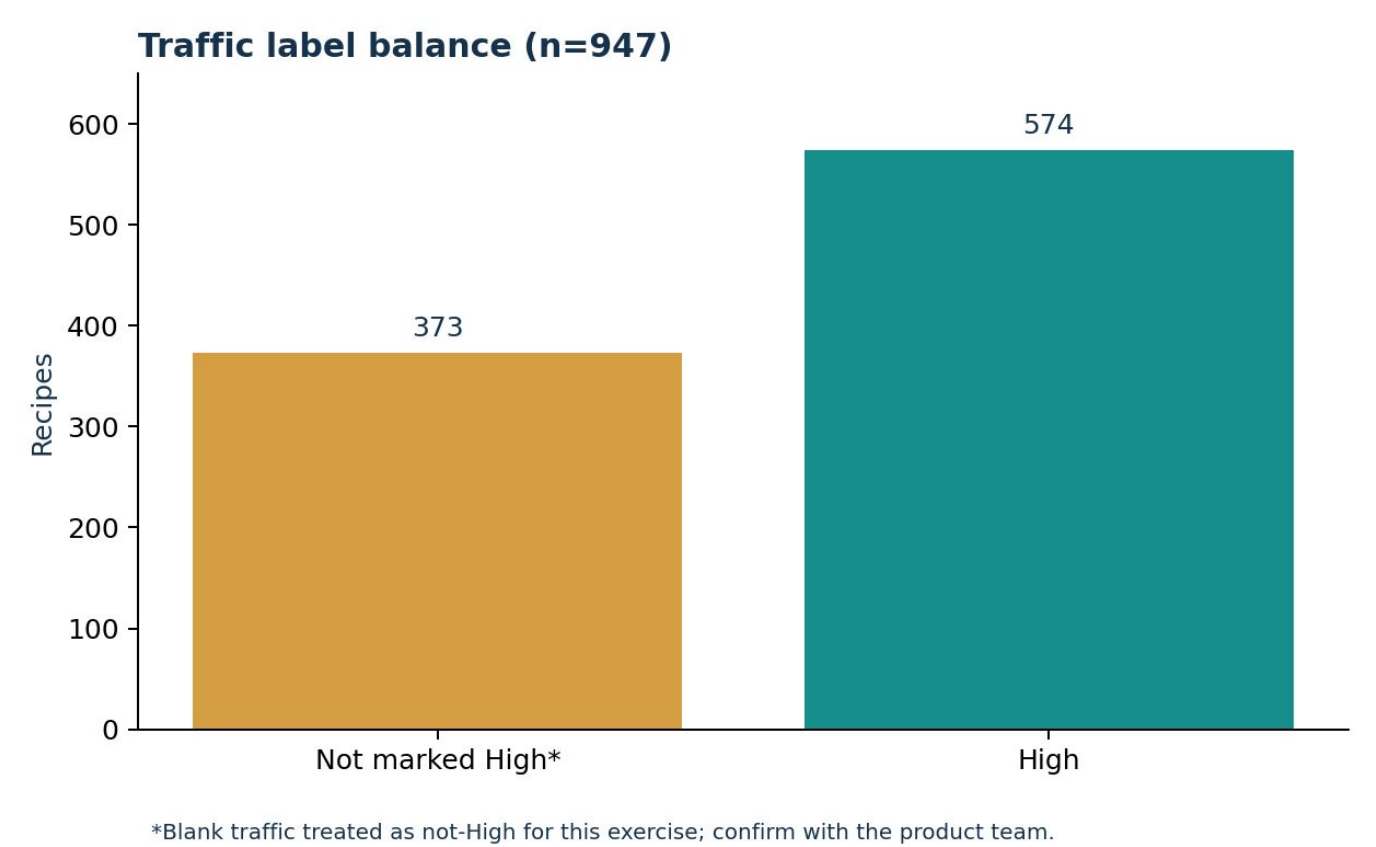
Decision standard

At least 80% of selected recipes should draw High traffic.

Traffic labels support this test. They do not establish subscription lift or a causal homepage effect.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

The dataset is usable for exploration, with one critical label check



947

historical featured recipes

52

rows missing all four nutrients

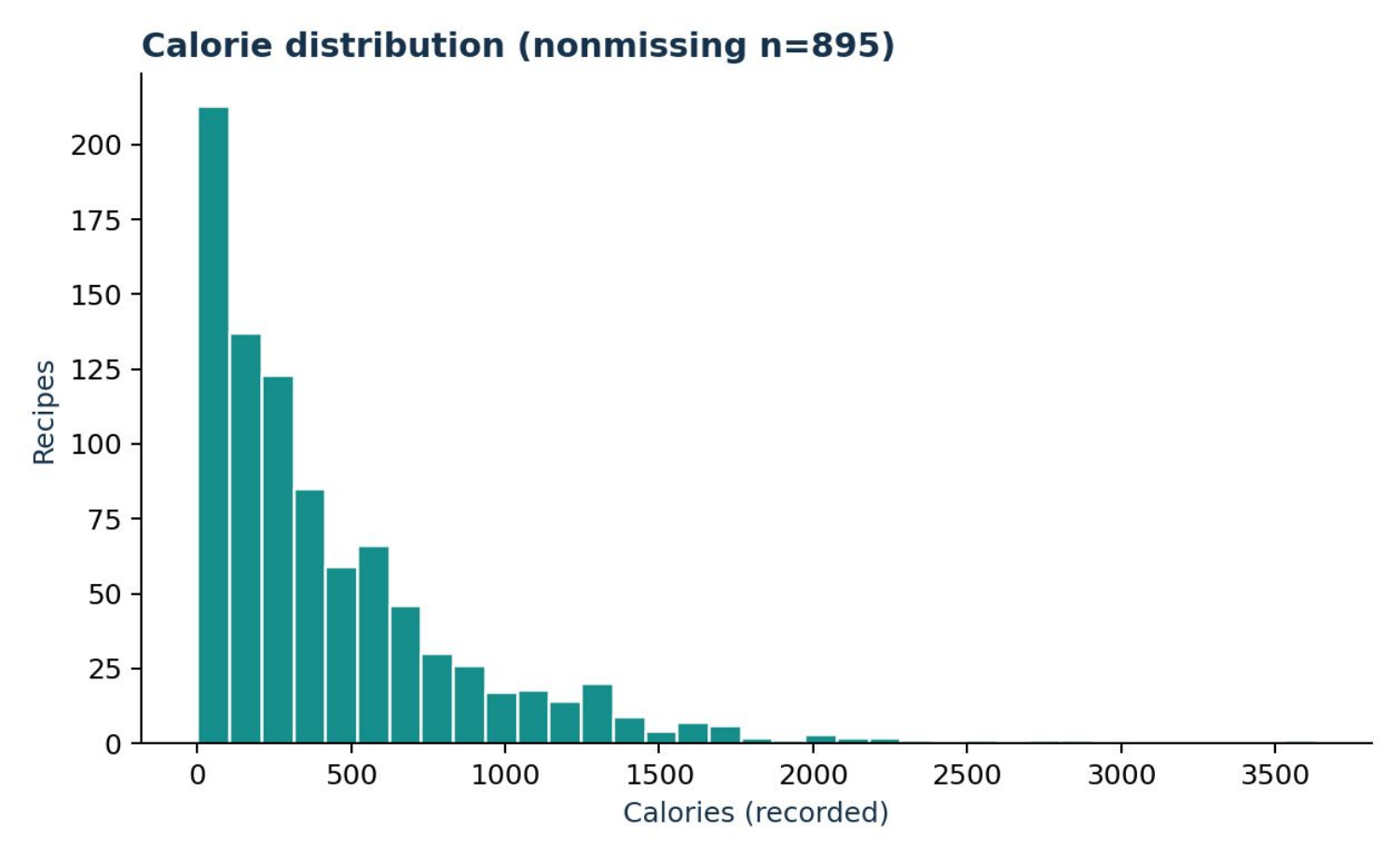
11

observed categories; brief lists 10

Decision: interpret 373 blank traffic values as not-High for this exercise. Confirm with the product team before any use.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

Recorded calories are strongly right-skewed



Recorded

895 / 947

recipes have calories

Median

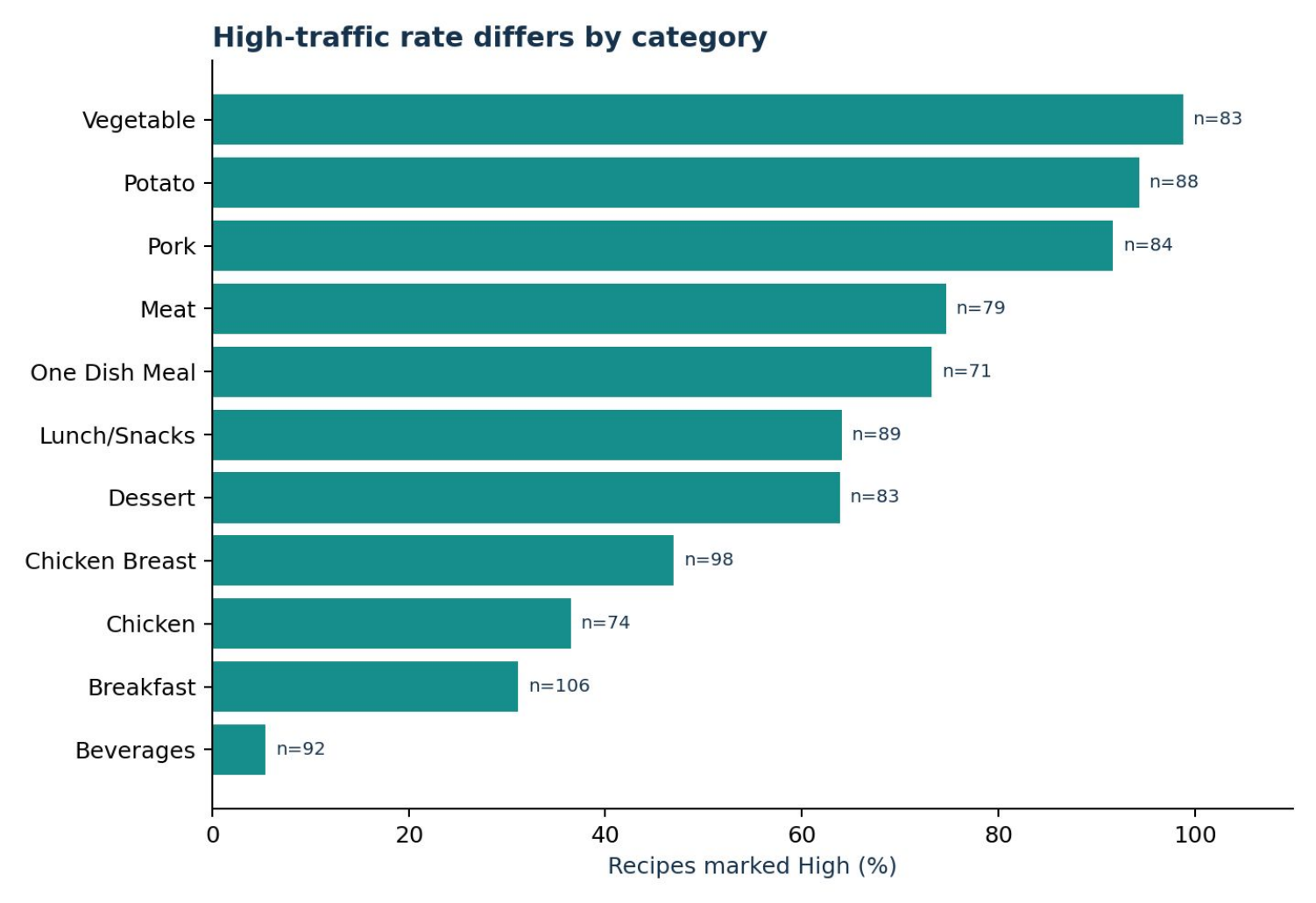
288.6

calories per recipe

The long upper tail makes the mean less representative of a typical recipe.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

Category has a strong descriptive association with traffic



Vegetable

82 / 83

marked High

Beverages

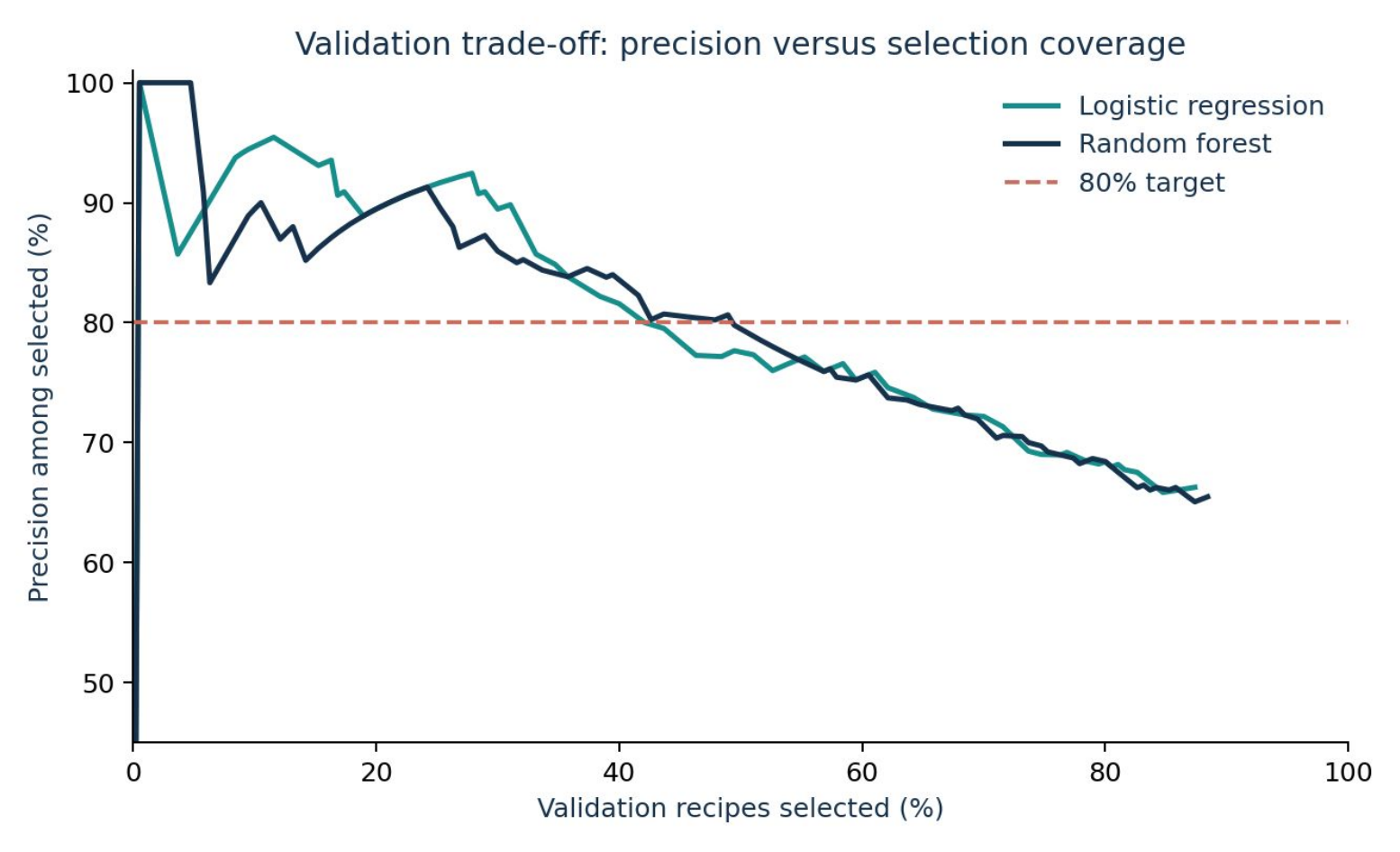
5 / 92

marked High

Category is predictive in this sample; it is not evidence that changing category causes more traffic.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

A conservative threshold meets 80% in validation



Validation set · n=190

Random forest

75 / 93 High · 80.6% precision

48.9% selection coverage

Logistic regression

64 / 80 High · 80.0% precision

42.1% selection coverage

Choice: random forest at score threshold 0.6425; final test set was untouched until selection.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

The selection rule reached 86.7% held-out precision

86.7%

precision among selected

78 High among 90 selected

47.4%

selection coverage

90 selected from 190 recipes

67.8%

recall of High recipes

78 found from 115 High

60.5%

all-High baseline precision

→

86.7%

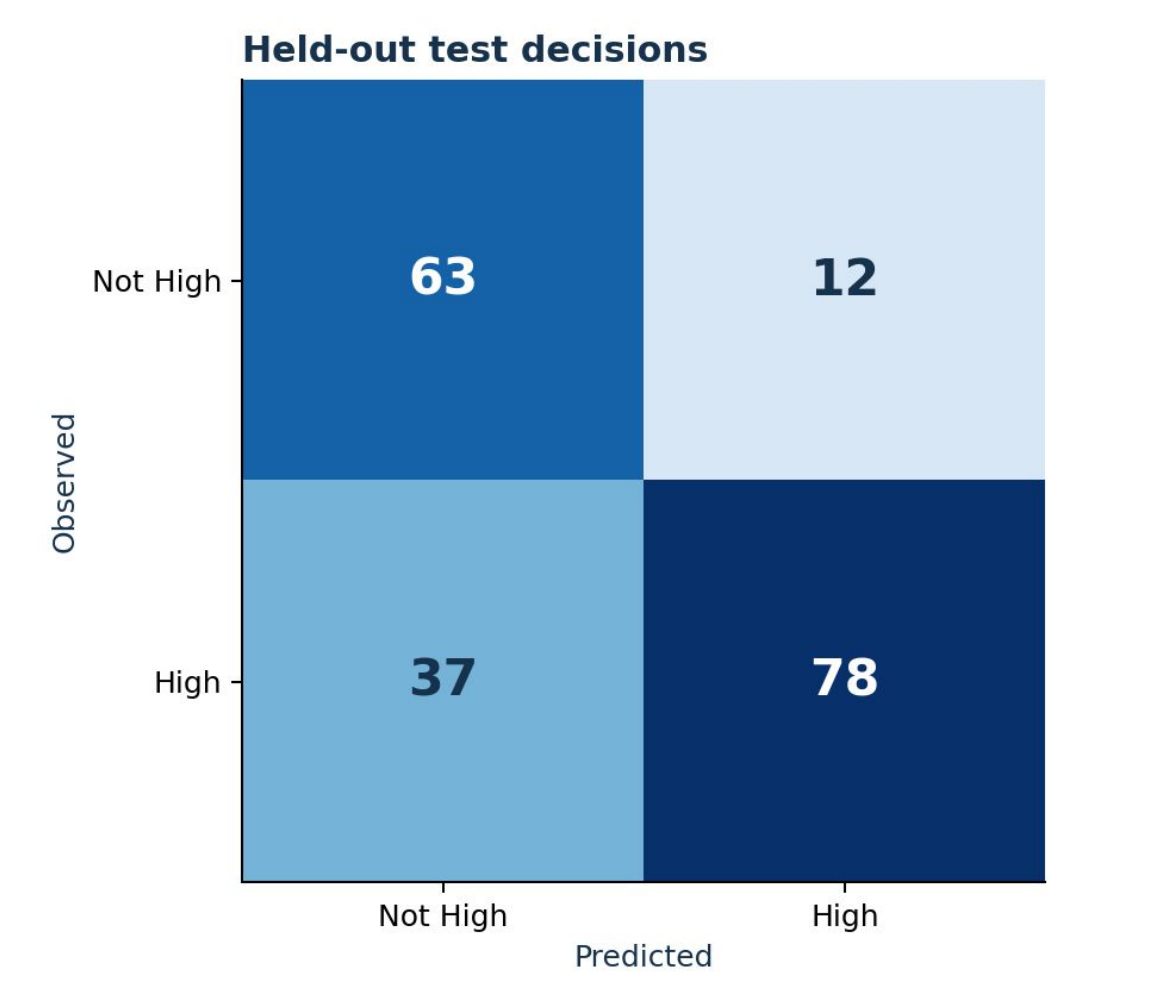
model selection precision

+26.1 pp

Uncertainty matters: Wilson 95% precision interval is 78.1–92.2%, so 80% is not yet statistically secure.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

The model makes errors in both directions



Selected by the model

78 High-traffic recipes

12 not-High recipes

37 High recipes were not selected.

Use the score to narrow choices, with an editorial fallback when no candidate clears the threshold.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

Test the rule prospectively before changing the homepage

01

Confirm the label

Resolve whether a blank value means normal traffic or missing measurement.

02

Log every candidate

Capture features, score, placement, exposure, traffic and subscription outcome.

03

Shadow the editor

Compare ranked suggestions with actual choices without changing the live page.

04

Randomize the pilot

Test model-assisted selection against current practice; report lift and uncertainty.

Pilot dashboard: selected-recipe precision and coverage, total site traffic, subscriptions and guardrails by period.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026

Decision request: approve a monitored, reversible pilot

What I recommend

1 Confirm the blank-label meaning.

2 Run the model in shadow mode.

3 Launch a randomized test only after the data and tracking checks pass.

What this analysis supports

A selection rule to test in a pilot.

Not yet proven: an 80% live hit rate, causal traffic lift, or subscription gain.

Source: recipe_site_traffic_2212.csv (n=947); analysis dated 25 Oct 2026