
name: Opportunity Score Improvements overview: Improve opportunity scoring to align with recognized sales/CRM practices (composite scoring, value weighting, consistent semantics) and add clear documentation so you can explain the changes to your team.

todos: [] isProject: false

Opportunity Score Improvements and Team Explanation

Current state (recap)

- **Four different formulas** in the app; no single definition.
- **Mixed semantics:** analytics uses "higher = more opportunity"; outreach/brand use "lower = more opportunity."
- **No revenue or recency** in any score, despite data being available.
- **Brand matching bug:** non-buyers get +10 (higher score) but list is sorted ascending, so existing buyers rank first instead of conquest opportunities.

Recognized practices to align with

Practice	What it is	How we can use it
Lead/Opportunity scoring	Single composite (often 0–100) from fit + value + behavior.	Combine category gap, revenue tier, and (optional) recency into one score.
Value weighting	Prioritize high-revenue or high-potential accounts.	Use revenue (and optionally deal size) so big accounts rank higher when opportunity is equal.
RFM (Recency, Frequency, Monetary)	Segment by last purchase, order count, spend.	Use recency (e.g. days since last order) so active customers rank higher for expansion.
White space / wallet share	"Share of customer wallet we don't have yet."	Keep category-gap idea but normalize and optionally weight by revenue.

Suggested improvements by feature

1. Analytics: Business-category opportunity (keep concept, add clarity and optional revenue weight)

Current: `opportunity_score = (total_categories - categories_bought) /`

`total_categories` (0–1, higher = more opportunity).

Suggested:

- **Keep** this as the core “category gap” (white space) metric.
- **Rename in UI** to “**Category expansion opportunity**” or “**Unrealized category share**” and add a short tooltip: “*Share of product categories this business type does not yet buy; higher = more room to expand.*”
- **Optional enhancement:** Add a **revenue-weighted** variant so business categories that already generate more revenue get a higher “priority” score (e.g. $\text{gap} \times \log(1 + \text{total_revenue_for_segment})$), and expose it as a separate column or toggle. Document that this follows “white space × value” thinking.

Files: [src/analytics.py](#) (`identify_opportunities`), [app.py](#) (chart title and axis label).

2. Customer-level opportunity (outreach, location, brand_product_matcher): composite score 0–100

Current: `opportunity_score = number of product categories bought` (lower = more opportunity). Revenue is available but not in the score.

Suggested formula (recognized practice: composite lead/opportunity score):

- **Category gap (expansion potential):**

```
category_gap = (total_product_categories - categories_bought) /  
total_product_categories
```

(0–1; higher = more categories they don’t buy yet).

- **Revenue tier (value):**

Normalize revenue per customer (e.g. min-max or percentile) into a 0–1 **revenue_score** so larger accounts score higher.

- **Optional recency (engagement):**

If `transaction_date` is available: `recency_score = 1 - (days_since_last_order / max_days)` so more recent buyers score higher.

- **Composite:**

```
opportunity_score = w1 * category_gap + w2 * revenue_score + w3 *  
recency_score
```

with default weights e.g. `w1=0.5, w2=0.4, w3=0.1` (or `w3=0` if no dates). Scale to **0–100** and **round** for display.

- **Semantics: Higher = more opportunity** (prioritize top of list). Sort **descending** by `opportunity_score`.

This gives a single, explainable number that combines “room to grow” (gap), “value” (revenue), and optionally “engagement” (recency).

Files: [src/outreach_automation.py](#) (`find_target_businesses_for_outreach`), [src/location_analytics.py](#) (`find_similar_businesses_by_location`), [src/brand_product_matcher.py](#) (consumes outreach score; switch to descending sort if/when outreach returns new score).

3. Brand matching: fix logic and use same composite idea

Current:

`opportunity_score = len(customer_categories)` if buys brand category, else `len(customer_categories) + 10`. Sorted ascending, so **existing buyers** rank first (opposite of "opportunity = don't buy yet").

Suggested:

- **Conquest vs upsell:**

- o **Conquest score:** Prioritize customers who **do not** buy the brand's category (opportunity to win new category).
- o **Upsell score:** Prioritize customers who **do** buy the brand's category (opportunity to sell more / switch to brand).

- **Single "opportunity" list:**

Use the **same 0–100 composite** as in (2) (category_gap + revenue + optional recency), but:

- o **Conquest:** Filter to "does not buy brand category" and sort by composite **descending**.
- o **Upsell:** Filter to "buys brand category" and sort by composite **descending**.
- **Remove** the ad hoc ⁺¹⁰ and the "lower = more opportunity" convention here so behavior matches outreach and is easy to explain.

Files: [src/brand_matching.py](#) (`find_businesses_for_brand`). Optionally add a parameter `priority="conquest" | "upsell"` and sort/filter accordingly.

4. Consistent semantics and UI

- **Standard: Higher score = more opportunity** everywhere. Use 0–100 for customer-level scores so "85" is clearly "high opportunity."
- **UI:**
 - o Use the same label everywhere for customer-level score, e.g. **"Opportunity score (0–100)"** with tooltip: *"Combines category expansion potential, account value, and recent activity; higher = higher priority."*
 - o For analytics, use **"Category expansion opportunity"** (or "Unrealized category share") and keep 0–1 or show as %.
- **Sorting:**
 - o Analytics: keep **descending** (higher first).
 - o Outreach, location, brand: **descending** by opportunity score (top of list = best opportunities).
- **Avg Opportunity Score metric:** Once on 0–100, the "Avg Opportunity Score" in the app is interpretable (e.g. "average 62" means "moderate opportunity set").

Files: [app.py](#) (labels, tooltips, any sort order that still assumes "lower = better"), [src/outreach_automation.py](#), [src/location_analytics.py](#), [src/brand_matching.py](#), [src/brand_product_matcher.py](#).

How to explain the desired changes to your team

One-line pitch

- **Before:** "Opportunity score means different things in different places; sometimes low is good, sometimes high is good, and we don't use revenue or recency."
- **After:** "One clear definition for customer lists: a 0–100 score where higher always means 'better'"

opportunity,’ combining how much room they have to buy more categories, how valuable they are, and how recently they bought.”

Short “why” for stakeholders

1. **Industry alignment:** Sales and marketing usually use composite scores (fit + value + behavior) and “higher = better”; we’re aligning with that so the number is intuitive and defensible.
2. **Revenue in the score:** Right now we could send outreach to tiny accounts first. Adding revenue ensures we prioritize valuable accounts when opportunity is similar.
3. **Consistent meaning:** Same scale (0–100) and same “higher = more opportunity” everywhere, so product, sales, and support all mean the same thing when they say “high opportunity.”
4. **Brand fix:** Today, “opportunity” in brand matching actually ranks current buyers first; we’ll fix that so conquest (don’t buy yet) and upsell (buy already) are explicit and correct.

Optional one-pager for the team (you can paste into Confluence/Notion/README)

Title: Opportunity score – definition and how we use it

- **Business category view (analytics):**

“Category expansion opportunity” = share of product categories that business type doesn’t buy yet (0–100% or 0–1). **Higher = more room to expand** into new categories for that segment.

- **Customer/outreach view (target lists, brand matching, location):**

“Opportunity score (0–100)” = composite of:

- o **Expansion potential** (how many categories they don’t buy yet),
- o **Account value** (revenue size),
- o **Recent activity** (optional; how recently they bought).

Higher = higher priority for outreach. We sort lists so the best opportunities are at the top.

- **Brand conquest vs upsell:**

We separate “don’t buy this category yet” (conquest) from “already buy this category” (upsell). Both use the same 0–100 score so we prioritize the best accounts in each list.

If someone asks “is this an industry standard?”

- “There isn’t one universal formula, but the **ideas** are standard: composite scoring (fit + value + behavior), revenue weighting, and white-space or category-gap. We’re implementing those in a simple, transparent way we can explain and tune (e.g. weights).”

Implementation order (for when you leave Plan mode)

1. **Define shared scoring helper** (e.g. in `src/analytics.py` or new `src/opportunity_scoring.py`): input = per-customer dataframe with columns for categories bought, total categories, revenue, optional `last_order_date`; output = 0–100 score and optional components (`category_gap`, `revenue_score`, `recency_score`). Use min-max or percentile normalization for revenue; handle missing `transaction_date` by setting recency weight to 0.
2. **Outreach:** In `find_target_businesses_for_outreach`, compute total revenue and (if available) days since last order per customer; call shared helper; replace current `opportunity_score` with 0–100; sort descending.
3. **Location:** Same in `find_similar_businesses_by_location`; sort descending.
4. **Brand matching:** Remove +10 logic; use shared helper; add conquest vs upsell filter and sort descending

for the chosen mode.

5. **Brand product matcher:** Use new score from outreach; sort descending.
 6. **Analytics:** Keep formula; rename in UI to “Category expansion opportunity” and add tooltip; optionally add revenue-weighted column.
 7. **App UI:** Update all labels and tooltips to “Opportunity score (0–100)” or “Category expansion opportunity,” fix any “lower = better” copy, and ensure tables/charts sort descending where appropriate.
 8. **Documentation:** Add the one-pager (or link to it) in the repo (e.g. `docs/OPPORTUNITY_SCORE.md` or a section in existing guide) so the team has a single reference.
-

Summary

- **Analytics:** Keep category-gap formula; rename and add tooltip; optionally add revenue-weighted variant.
- **Customer-level:** Introduce a single **0–100 composite** (category gap + revenue + optional recency); **higher = more opportunity**; use in outreach, location, and brand matching; fix brand conquest/upsell logic.
- **Team explanation:** Use the one-line pitch, the “why” bullets, and the one-pager so everyone understands what the score means and why the changes make it more meaningful and aligned with common practice.