Flavor Analysis Brief – Sparkling Ice

Omnium

Adam Mahmoud

August 14, 2025

Background

Talking Rain's Sparkling Ice brand has already achieved high distribution in the bottled water category, with strong presence across the food channel. To maintain growth, innovation in flavors is key. One question that emerged: *Why does Sparkling Ice not currently offer a cola flavor?* Given consumer familiarity with cola in adjacent beverage categories, evaluating its potential performance was a priority.

Problem / Goal

The goal was to develop a quantitative method to estimate how specific flavors—particularly cola—would perform for Sparkling Ice. We focused on creating **flavor indices** that measure a flavor's velocity relative to category or brand benchmarks.

Data Filters

- **Geography**: Total US – Food
- **Timeframe**: L52 Aggregated WE 07/13/2025
- **Distribution filter**: ACV $> 10\%$
- **Categories**: Bottled Water, Carbonated Beverages
- **Subcategories**: Flavored seltzer/sparkling/mineral water, low calorie soft drinks, regular soft drinks
- **Pack sizes**: Compared separately:

  - Singles: 12–20 oz
  - 12-pack: 144 oz and 204 oz (204 oz includes Sparkling Ice due to their singles bottles being 17 oz)

- **Flavor binning**: The raw IRI data contained ∼1,500 unique flavors. To make the analysis meaningful, our team manually grouped these into **49 standardized flavor bins** (e.g., Cola, Vanilla Cola, Strawberry Kiwi, All Other).

Approaches Overview

We tested four different approaches, ultimately focusing on the **Flavor–Brand model** (approach 4) for deeper insight:

1. **Indexing by average dollar velocity** (dollars / stores selling / weeks in distribution) **across all flavors**
   For each flavor, calculate an index as:

$$\text{Index} = \frac{\text{Flavor Dollar Velocity}}{\text{Average Dollar Velocity Across All Flavors}}$$

   This measures how a flavor performs relative to the category average.

2. **Percentile index**
   Rank each flavor by its dollar velocity among all flavors, then express the ranking as a percentile:

$$\text{Percentile Index} = \frac{\text{Number of Flavors with Lower Velocity}}{\text{Total Number of Flavors}} \times 100$$

   This shows the relative standing of a flavor in the distribution of velocities.

3. **Indexing off a specific flavor (e.g., Cola)**
   Choose a reference flavor (Cola) and calculate:

$$\text{Index} = \frac{\text{Flavor Dollar Velocity}}{\text{Cola Dollar Velocity}}$$

   This directly measures how each flavor compares to the benchmark flavor.

4. **Flavor–Brand model** (primary focus)
   Detailed in the following section.

Flavor–Brand Model

Part 1: Brand–Specific Cola Index

1. Filtered to brands with both a cola flavor and at least one other flavor.

2. For each brand:
   (a) Calculated average velocity **excluding cola**.
   (b) Calculated the **Cola Index** = Cola velocity ÷ brand average velocity.
3. Averaged these indices across brands (excluding Coca-Cola and Pepsi, as their portfolios skew results heavily).
4. Result: **Average Cola Index** = 1.02 for singles and 2.14 for 12-pack.
5. Applied these multipliers to brands without cola:

$$\text{Experimental Cola Velocity} = \text{Brand Avg. Velocity} \times \begin{cases} 1.02, & \text{singles} \\ 2.14, & \text{12-pack} \end{cases}$$

Part 2: Flavor–Brand Regression

Goal: Estimate the relative performance of all flavors while controlling for brand effects, ensuring that each flavor receives its own index value.
Data filter: Brands with at least two flavors were included to ensure the model could distinguish flavor effects within each brand.
Model specification: Several regression configurations were tested in Python (integrated with Excel for flexible filtering and output formatting). We compared models with and without intercepts, and with raw velocity versus log-transformed velocity.

Models with an intercept implicitly set one flavor as the baseline and expressed all others relative to it — meaning that the baseline flavor itself would have no coefficient or index value. Removing the intercept allowed every flavor, including what would have been the baseline, to have its own coefficient and index.

Running regressions on raw velocities caused high-velocity outliers to dominate the results, making indices less stable and less comparable across flavors. Using the natural log of velocity reduced the influence of these outliers, normalized residuals, and produced proportional, scale-free coefficients that better reflect relative differences in performance.

Based on these considerations, we selected the following specification:

```
log(Velocity) ~ 0 + C(Brand) + C(Flavor)
```

This setup provides interpretable indices for all flavors, consistent proportional scaling, and more stable results across different brand and pack size filters.
Calculation steps:

1. Extracted the `C(Flavor)` coefficients from the regression output.
2. Converted log coefficients into **Flavor Indices**:

$$\text{Flavor Index} = e^{\text{log coeff} - \text{mean log coeff}}$$

3. Sorted results to identify the highest- and lowest-performing flavors.

Brand Filter	Cola Index (12-pack)	Cola Index (Singles)
All brands	8.99	2.30
Excluding Coca-Cola & Pepsi	3.31	1.05
Alternative[1] + low-calorie soft drinks	3.78	1.90
Similar brands[2]	1.65	1.02
Alternative brands	1.50	1.02

For 12-pack, the Cola Index shifts substantially depending on which brands are included, with values ranging from 1.50 to 8.99. This indicates a strong influence from large soda manufacturers and certain brand groupings. In contrast, the singles results are far more stable, remaining close to 1.02 across most brand filters, suggesting that the relative performance of cola flavors in singles is less dependent on the specific brands chosen for comparison.

Key Limitation

Flavor indices are **highly sensitive** to brand set selection. Ensuring the comparison set is truly representative of Sparkling Ice's positioning is critical to making reliable predictions.

Result

Both the brand-specific and regression approaches supported the hypothesis that cola-style flavors could perform competitively for Sparkling Ice. Findings suggested potential for a **soda flavor line** including cola, Dr. Pepper-style, root beer, and cream soda.

Because the flavor indices represent expected performance relative to a brand's existing portfolio, they can be used to project sales opportunity for hypothetical new flavors. For example, for Sparkling Ice, multiplying the brand's average velocity by the estimated flavor index for each of the four soda-style flavors provided an initial projection of potential velocity, and in turn, gross sales opportunity.

This methodology is also broadly applicable beyond Sparkling Ice. The underlying pipeline can be adapted to other brands and categories, with the most manual steps being the creation of fair brand comparison sets and consistent flavor binning. Once these inputs are established, the approach provides a repeatable framework for estimating how any potential flavor might perform in the market.

[1] Alternative brands include Sparkling Ice, Poppi, Olipop, Zevia

[2] Similar brands include Sparkling Ice, La Croix, Poppi, Polar, Olipop, Bubly, Spindrift, Waterloo, Zevia, Topo Chico, Bubbl'r, Clearly Canadian