Which Brands Will Make It?

A Framework for Scoring Emerging Brands

Kristin Henderson

Advisor: Bivin Sadler

hendersonk@smu.edu

Master of Science in Data Science
Southern Methodist University

DataScience@SMU

The Problem

Distributors, brokers, and retailers must decide which emerging brands to invest in.

The Problem

Distributors, brokers, and retailers must decide which emerging brands to invest in.

Some will **grow** rapidly. Some will **fail**. Most sit in between.

How do you tell them apart early?

The Problem

Distributors, brokers, and retailers must decide which emerging brands to invest in.

Some will grow rapidly. Some will fail. Most sit in between.

How do you tell them apart early?

Emerging Brand
Limited history
Uncertain trajectory

→

Decision Framework
Multi-criteria scoring
+ statistical benchmarks

→

Opportunity Tiers
Growing | Stable
Decline | Discontinued

The Problem

Distributors, brokers, and retailers must decide which emerging brands to invest in.

Some will grow rapidly. Some will fail. Most sit in between.

How do you tell them apart early?

Emerging Brand
Limited history
Uncertain trajectory

→

Decision Framework
Multi-criteria scoring
+ statistical benchmarks

→

Opportunity Tiers
Growing | Stable
Decline | Discontinued

Prior work: Focused on demand forecasting

Our approach: Evaluate brands on multiple performance signals, not just sales predictions

The Data

Weekly retail sales data from a brokerage agency

5 years of data (2021-2026), ~260 weeks

22,500 brands

Primary Variables:
Sales (Case and Dollar)
%ACV (All Commodity Volume)

The Data

Weekly retail sales data from a brokerage agency

5 years of data (2021-2026), ~260 weeks

22,500 brands

Primary Variables:
Sales (Case and Dollar)
%ACV (All Commodity Volume)

What is %ACV?

The Data

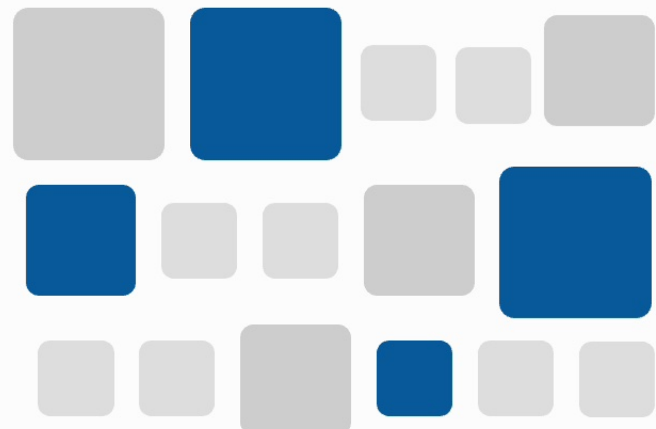
Weekly retail sales data from a brokerage agency

5 years of data (2021-2026), ~260 weeks

22,500 brands

Primary Variables:
Sales (Case and Dollar)
%ACV (All Commodity Volume)

What is %ACV?



Product in **4** of 16 stores, but those 4 represent

35% of total market volume = 35% ACV

Building the Truth Set

The dataset has no outcome labels. We had to create them.

Building the Truth Set

The dataset has no outcome labels. We had to create them.

BUILD THE POOL

Started with a large single category

Building the Truth Set

The dataset has no outcome labels. We had to create them.

BUILD THE POOL

Started with a large single category

→

PLOT AND LABEL

Plotted trajectories and classified each brand

Building the Truth Set

The dataset has no outcome labels. We had to create them.

BUILD THE POOL

Started with a large single category

PLOT AND LABEL

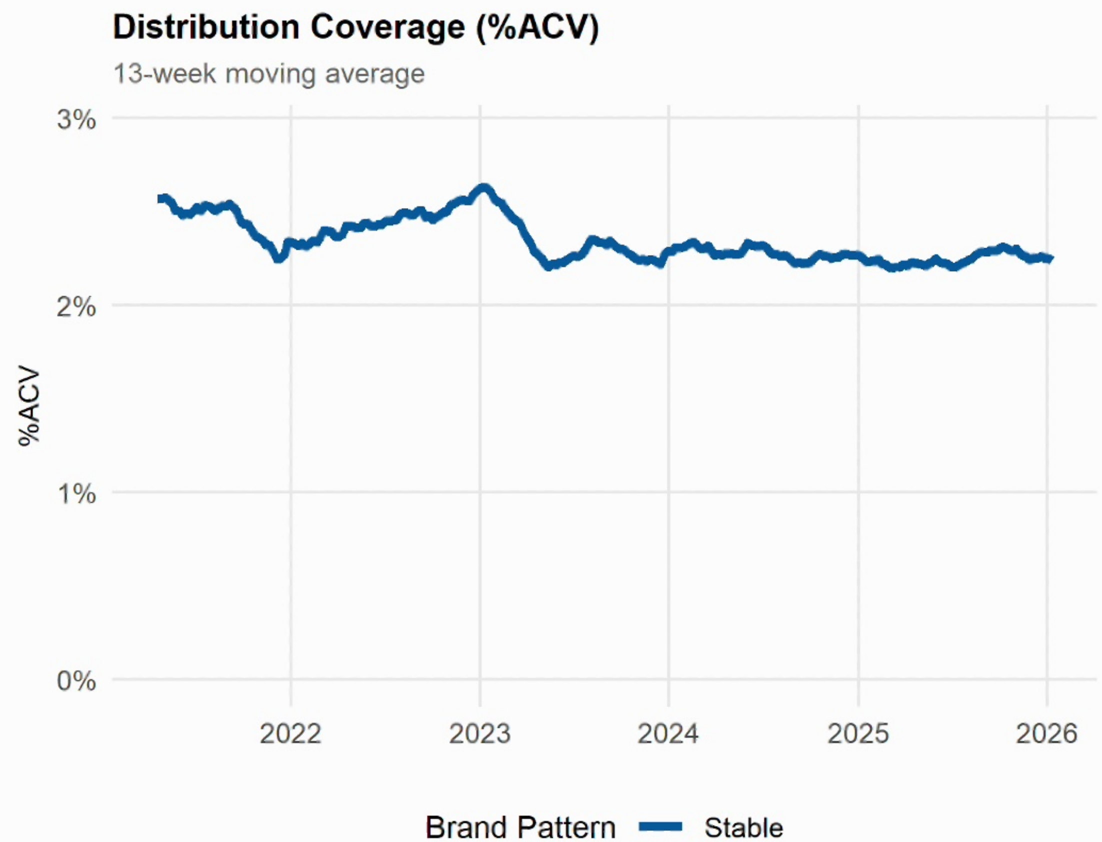
Plotted trajectories and classified each brand

RESULT

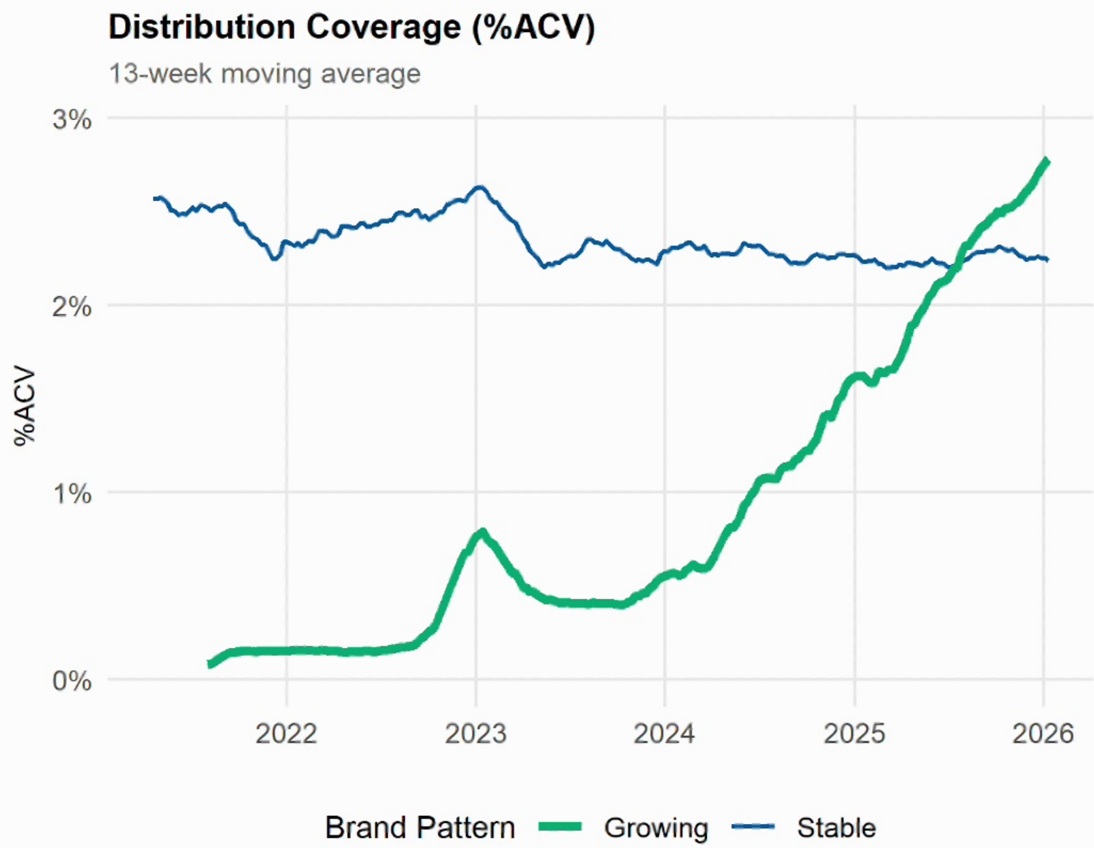
112 labeled brands, roughly balanced across all four tiers

29 Discontinued	30 Decline	28 Stable	25 Growing

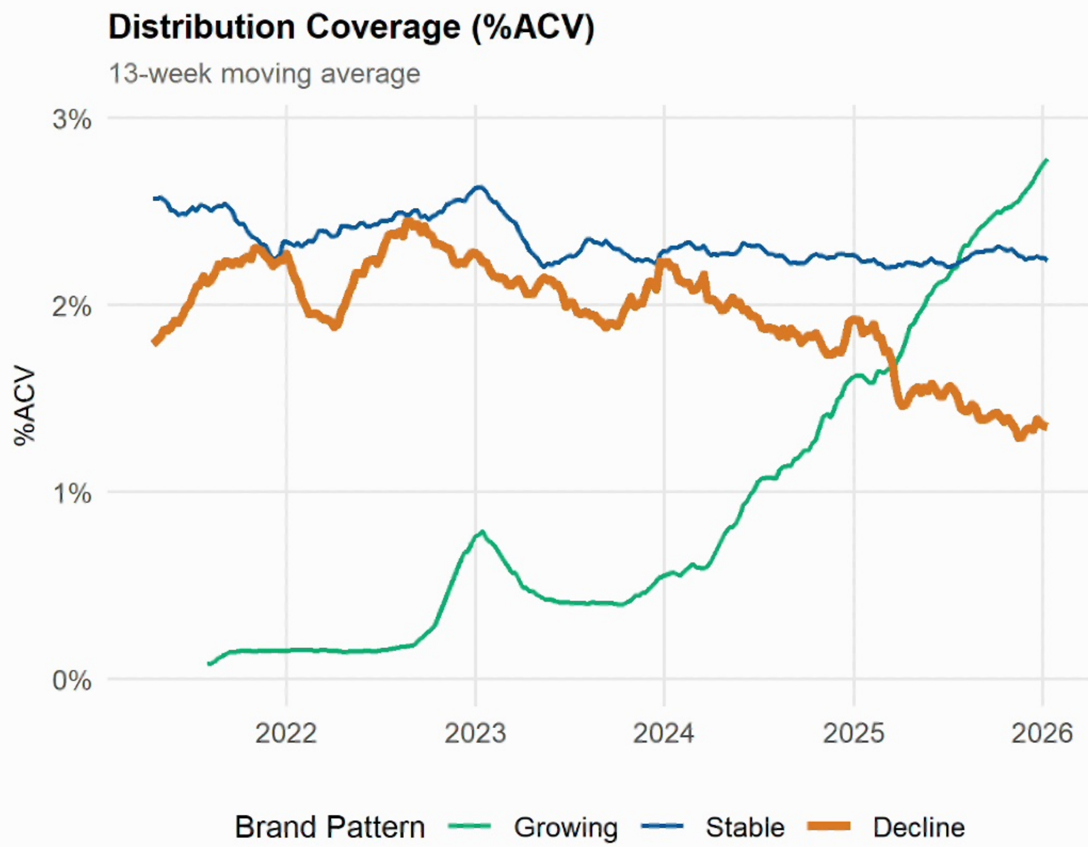
What Does Brand Performance Look Like?



What Does Brand Performance Look Like?

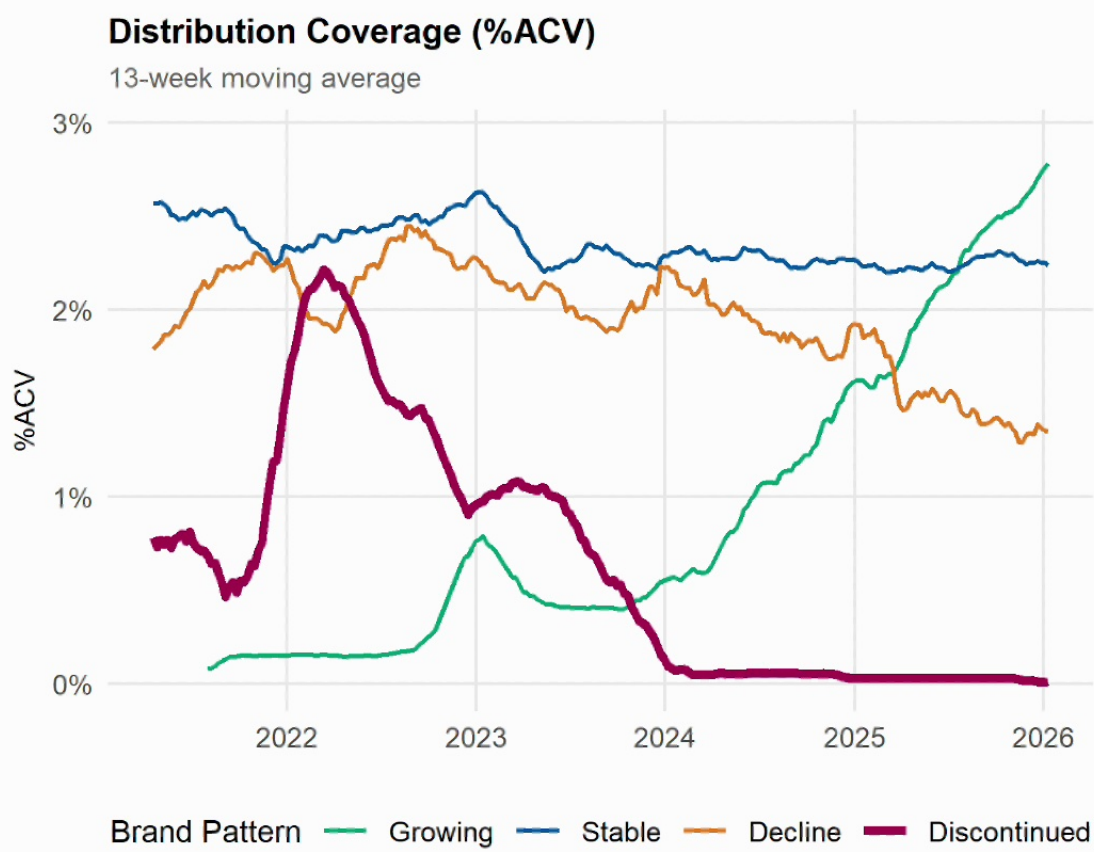


What Does Brand Performance Look Like?



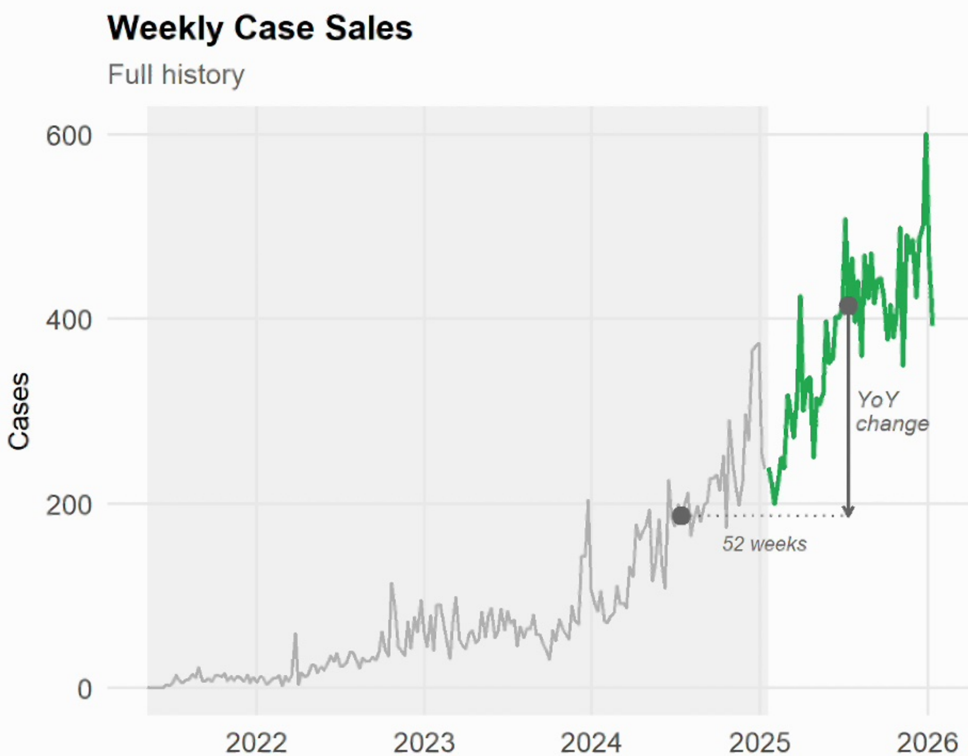
DataScience@SMU

What Does Brand Performance Look Like?



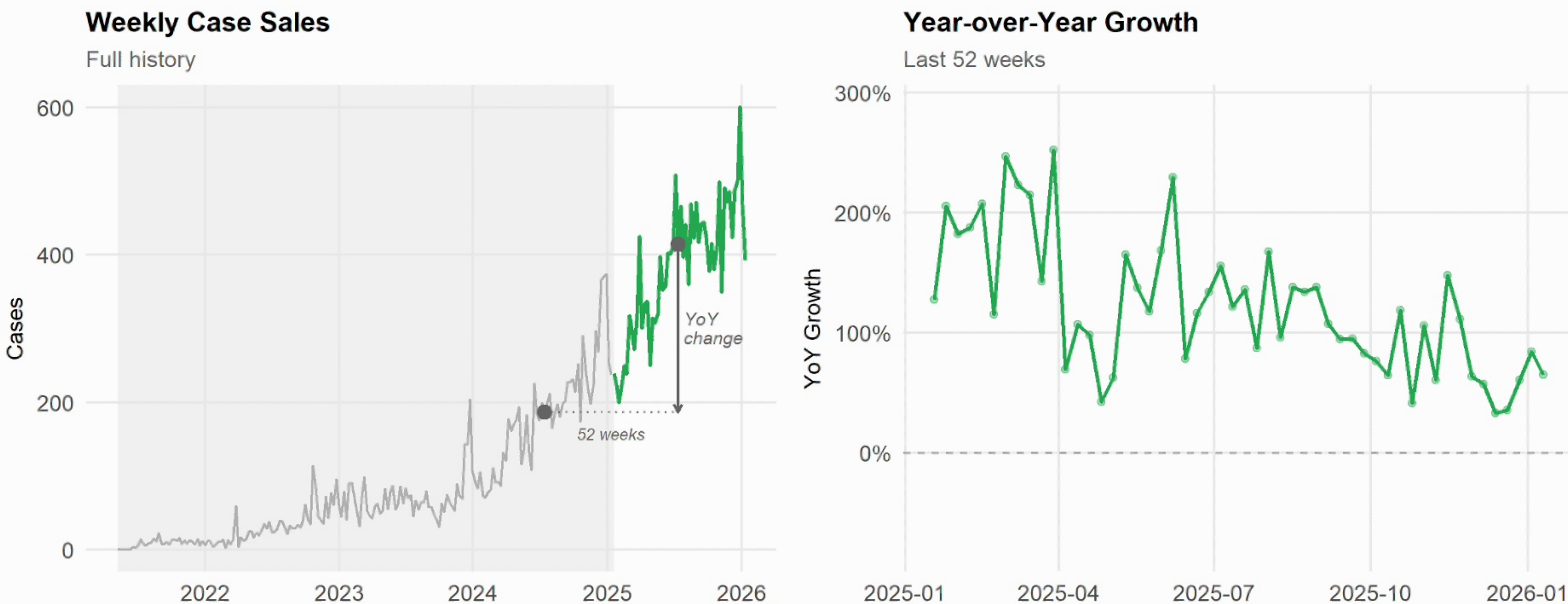
DataScience@SMU

From Raw Sales to Demand Features



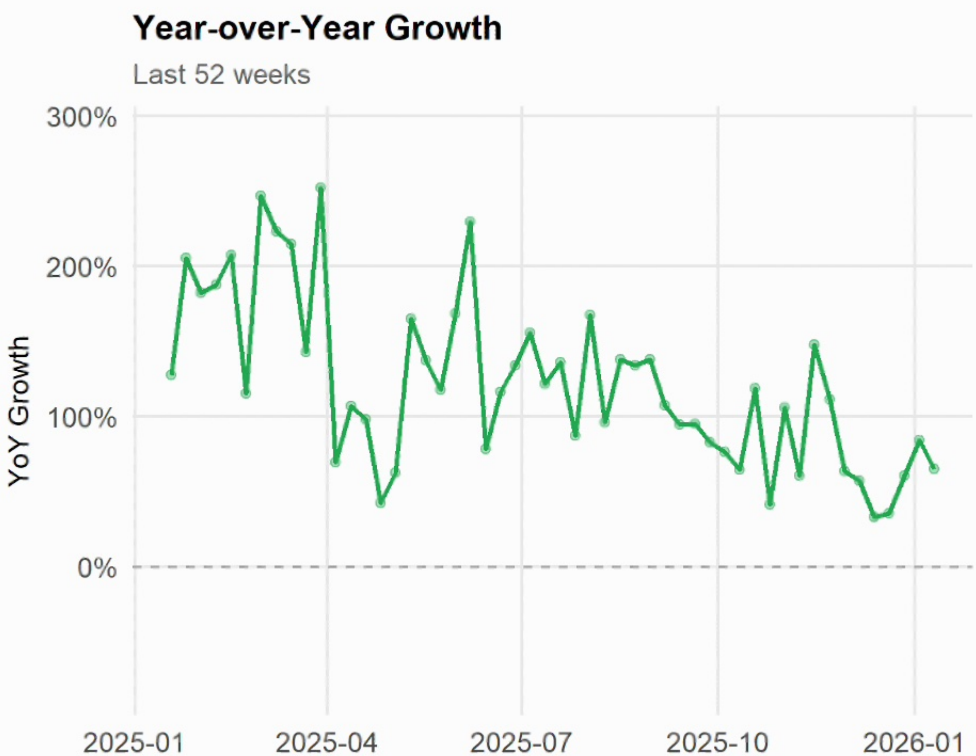
DataScience@SMU

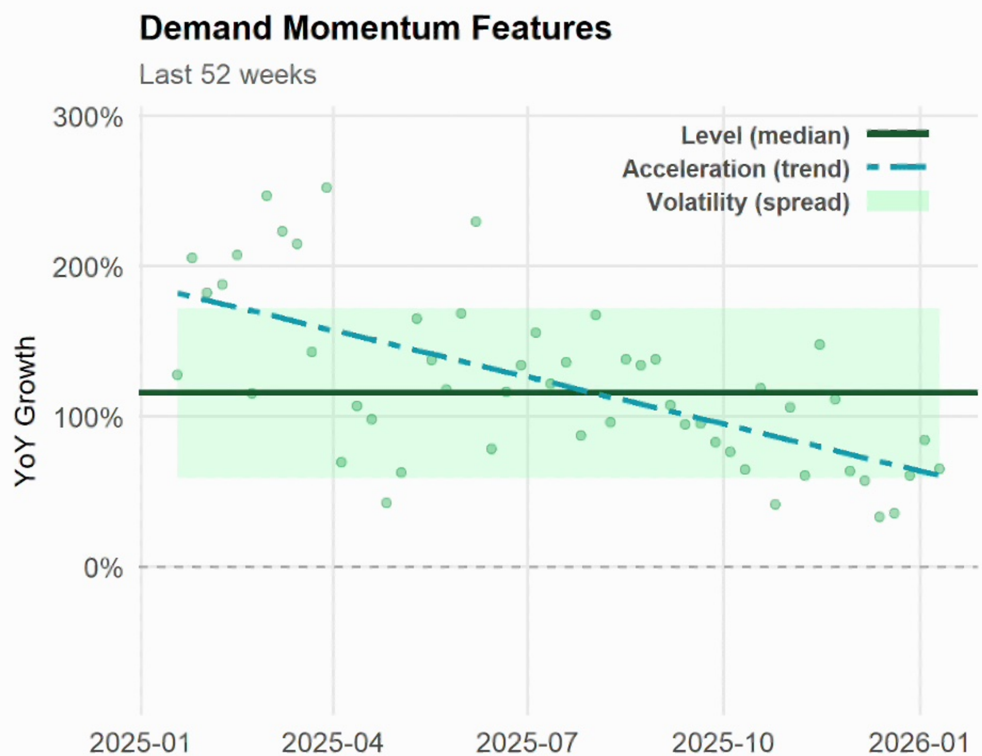
From Raw Sales to Demand Features



DataScience@SMU

From Raw Sales to Demand Features





DataScience@SMU

A Scoring Framework Built for This Decision

Multi-Criteria Decision Analysis (MCDA) evaluates brands across multiple criteria using a weighted sum:

$$\text{Opportunity Score} = w_1(\text{Demand Momentum}) + w_2(\text{Distribution}) + w_3(\text{Promotion}) + w_4(\text{Category})$$

A Scoring Framework Built for This Decision

Multi-Criteria Decision Analysis (MCDA) evaluates brands across multiple criteria using a weighted sum:

$$\text{Opportunity Score} = \boldsymbol{w_1}(\textbf{Demand Momentum}) + w_2(\text{Distribution}) + w_3(\text{Promotion}) + w_4(\text{Category})$$

Level

Median YoY growth

How fast is the brand growing?

Acceleration

Trend of YoY growth

Is growth speeding up or slowing down?

Volatility

Spread of YoY growth

How consistent is the growth?

A Scoring Framework Built for This Decision

Multi-Criteria Decision Analysis (MCDA) evaluates brands across multiple criteria using a weighted sum:

$$\text{Opportunity Score} = \boldsymbol{w_1}(\textbf{Demand Momentum}) + w_2(\text{Distribution}) + w_3(\text{Promotion}) + w_4(\text{Category})$$

Level
Median YoY growth
How fast is the brand growing?

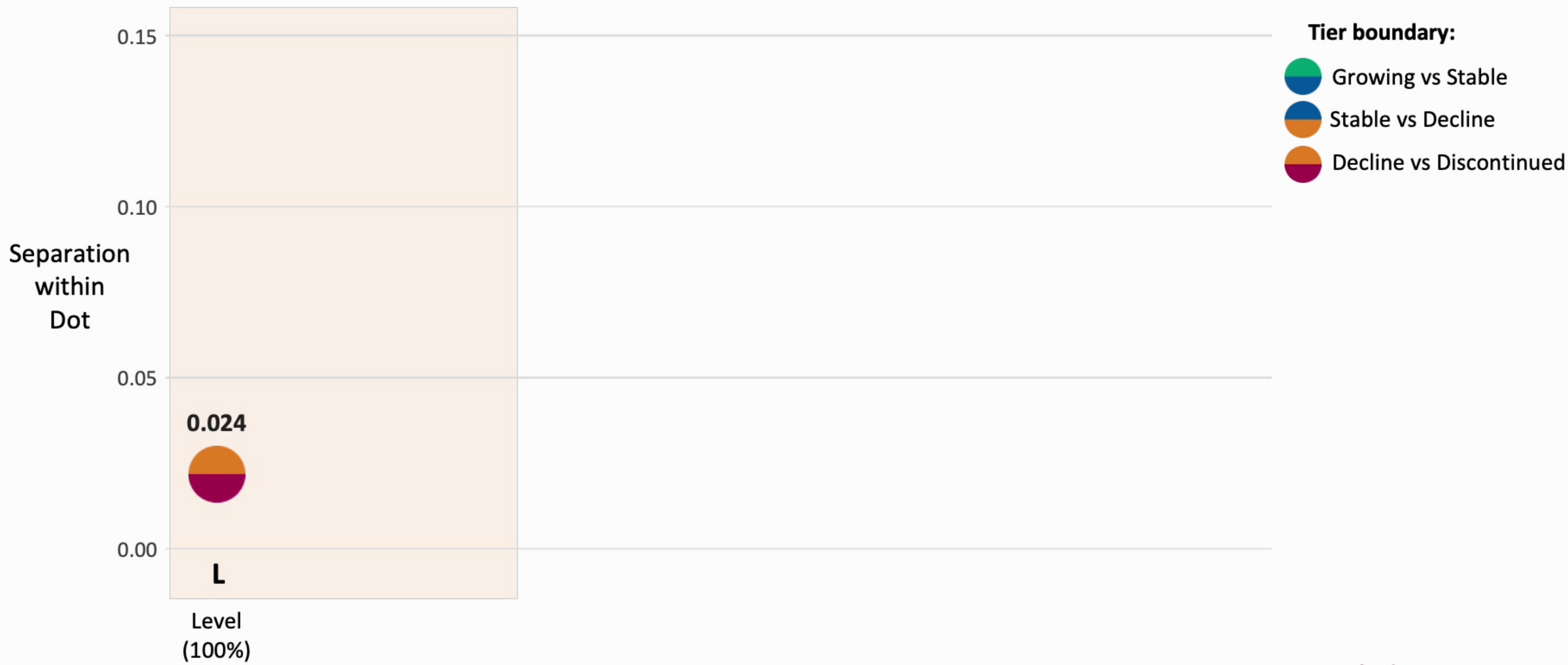
Acceleration
Trend of YoY growth
Is growth speeding up or slowing down?

Volatility
Spread of YoY growth
How consistent is the growth?

Demand Momentum
Weighted combination of level, acceleration, and volatility

Which Feature Configuration Best Separates Tiers?

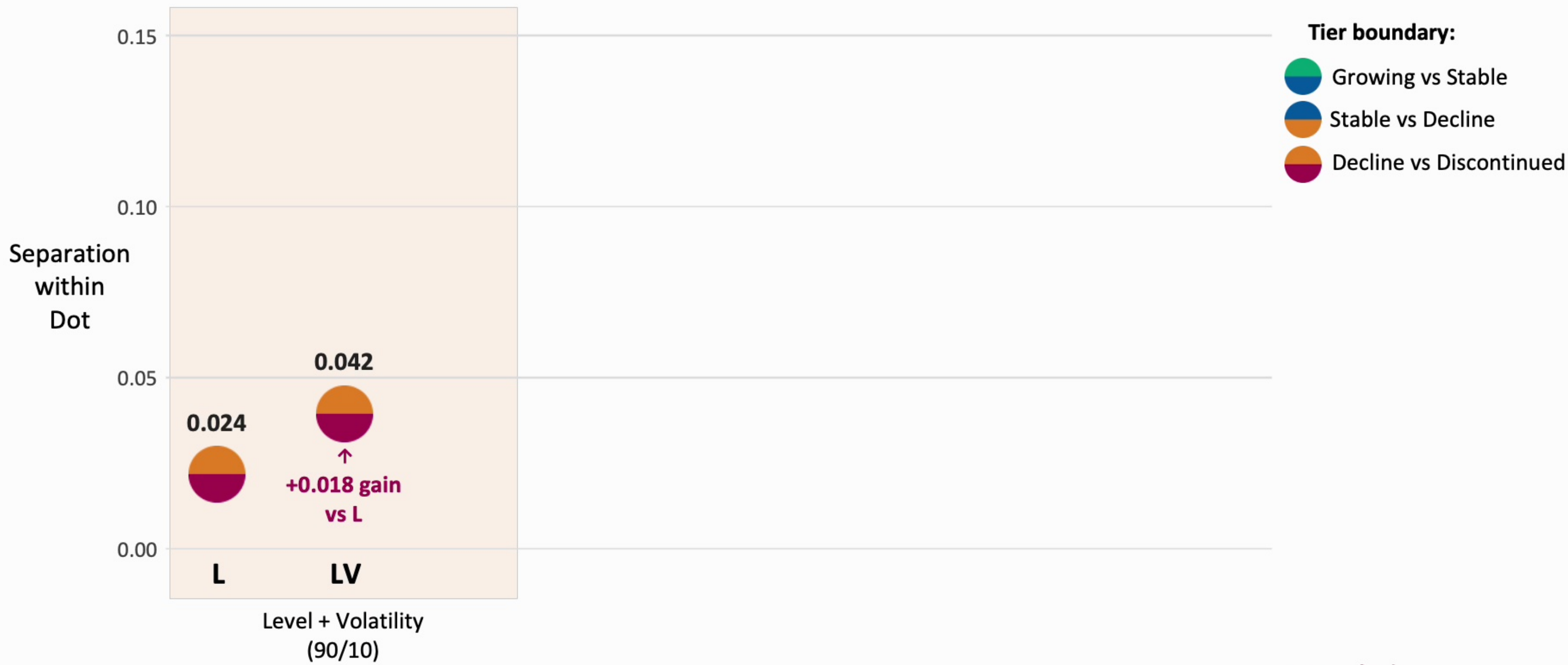
Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



DataScience@SMU

Which Feature Configuration Best Separates Tiers?

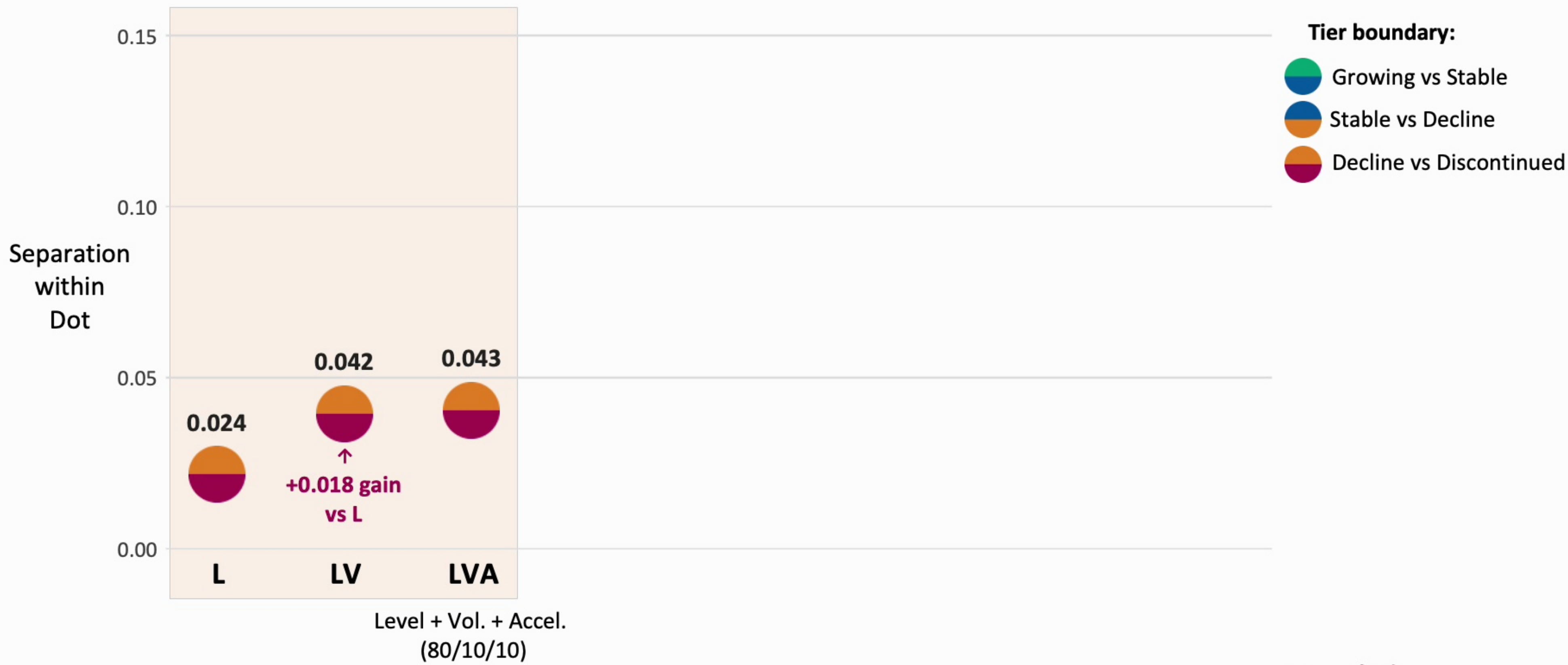
Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



DataScience@SMU

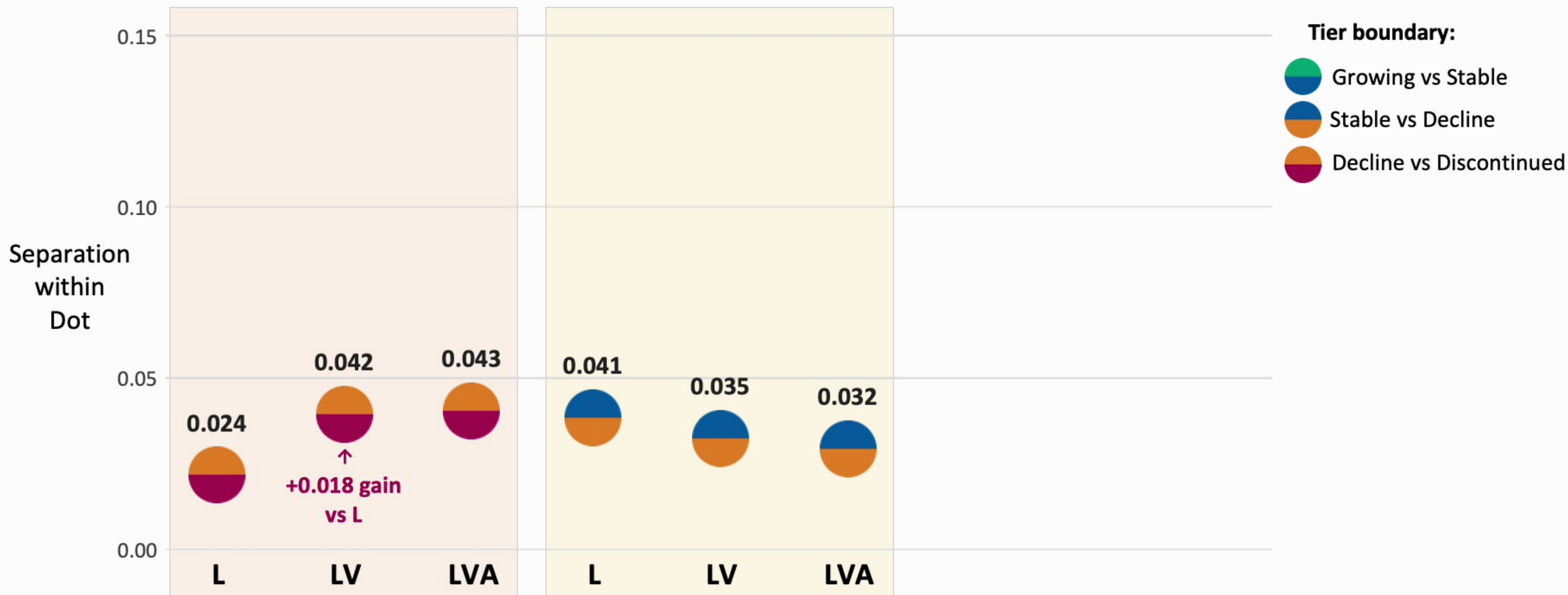
Which Feature Configuration Best Separates Tiers?

Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



Which Feature Configuration Best Separates Tiers?

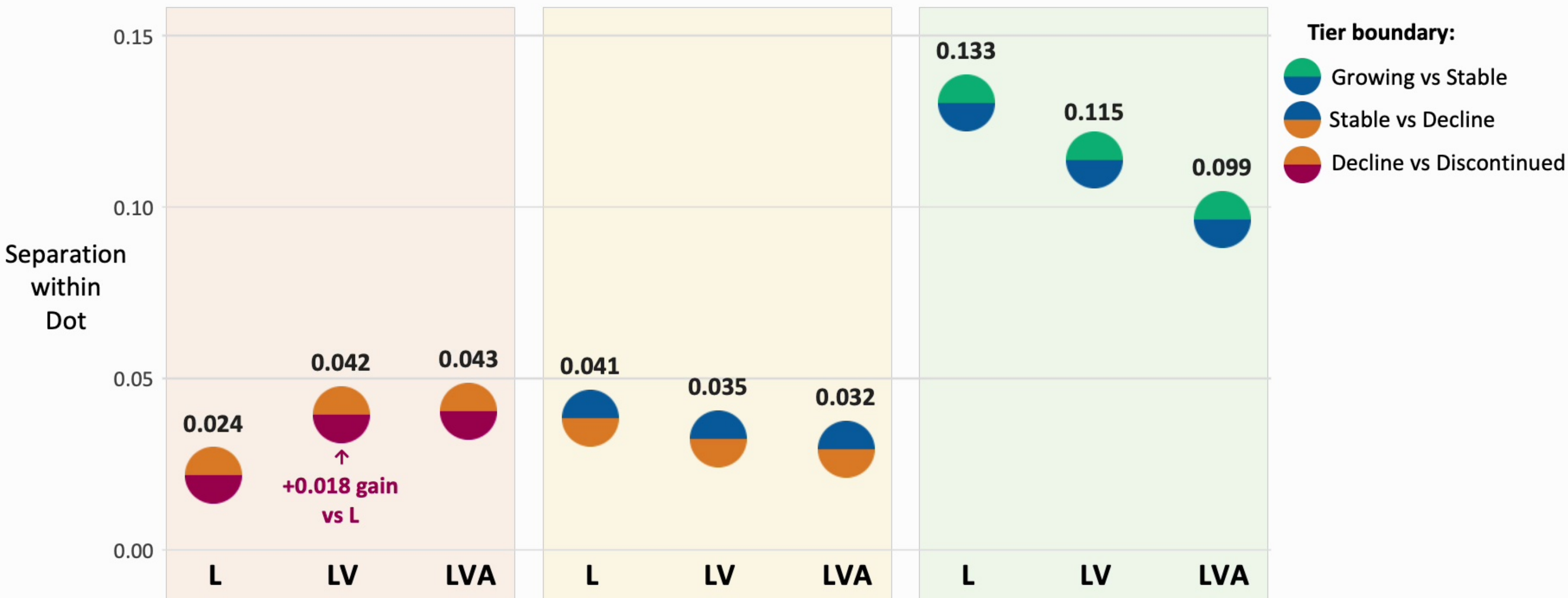
Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



DataScience@SMU

Which Feature Configuration Best Separates Tiers?

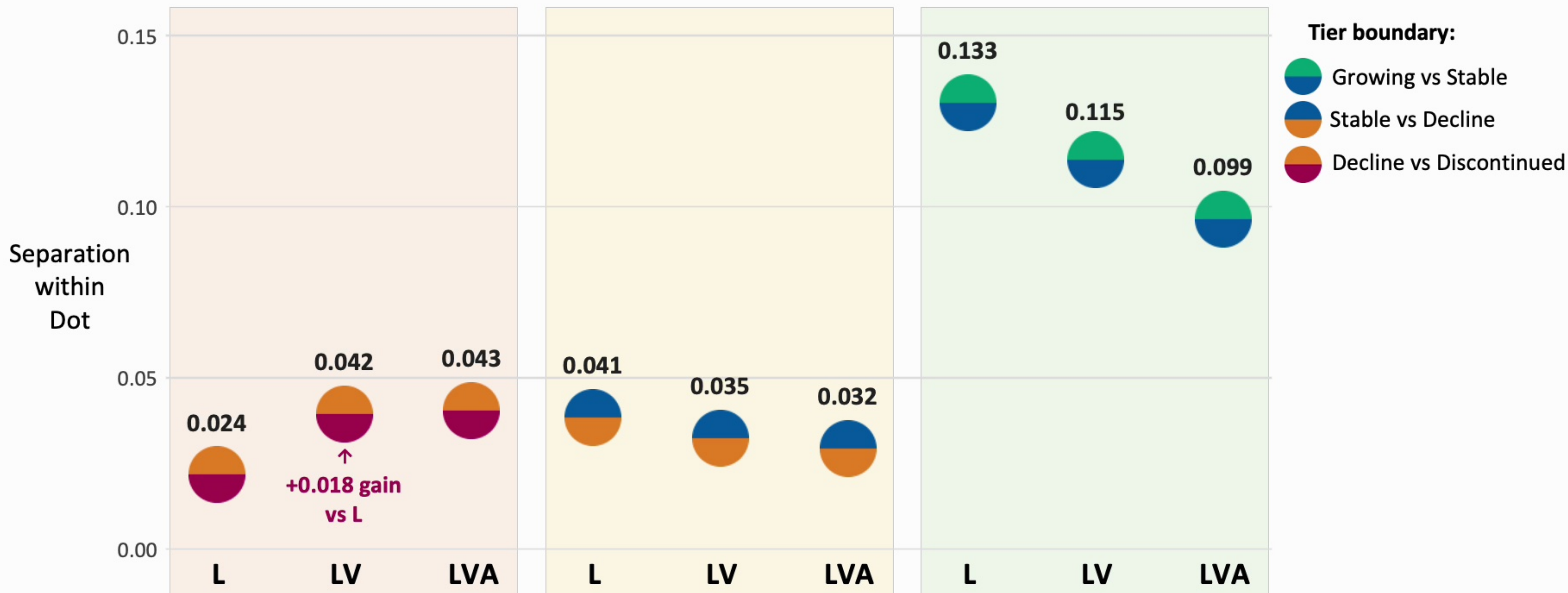
Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



DataScience@SMU

Which Feature Configuration Best Separates Tiers?

Separation = difference in mean MCDA score between adjacent tiers. Larger gap = better separation.



LV is the most balanced config: gains at the bottom tier without much loss at the top.

DataScience@SMU

Two Models, Two Levels of Detail

Each model addresses a different decision a broker faces

Is this brand worth investing in?

Logistic Regression

Low-performing (disc + decline) vs. Viable (stable + growing)

Best config: Level + Volatility

Level $p < 0.001$ *** | Volatility $p = 0.038$ *

83% Balanced Accuracy
83% Macro F1

Metrics from leave-one-out cross-validation (n = 112).

Two Models, Two Levels of Detail

Each model addresses a different decision a broker faces

Is this brand worth investing in?

Logistic Regression

Low-performing (disc + decline) vs. **Viable** (stable + growing)

Best config: Level + Volatility

Level $p < 0.001$ *** | Volatility $p = 0.038$ *

83% Balanced Accuracy

83% Macro F1

Which tier is it in?

Ordinal Logistic Regression

Discontinued < **Decline** < **Stable** < **Growing**

Best config: Level only

Level $p < 0.001$ ***

58% Balanced Accuracy

92% Adjacent Accuracy | 59% Macro F1

Metrics from leave-one-out cross-validation (n = 112).

Models Get It Close

Low-Performing vs. Viable

	Truth: Low-Perf	Truth: Viable
Pred: Low-Perf	47	7
Pred: Viable	12	46

83% balanced accuracy

7 false positives | 12 false negatives

Rows = predicted, Columns = truth. LOOCV (n = 112).

Models Get It Close

Low-Performing vs. Viable

	Truth: Low-Perf	Truth: Viable
Pred: Low-Perf	47	7
Pred: Viable	12	46

83% balanced accuracy

7 false positives | 12 false negatives

Ordinal: 4-Tier

	Discontinued	Decline	Stable	Growing
Discontinued	18	9	0	0
Decline	2	14	5	0
Stable	8	7	18	10
Growing	1	0	5	15

92% adjacent accuracy

Discontinued is hardest tier (47% recall)

Rows = predicted, Columns = truth. LOOCV (n = 112). Green = correct Yellow = off by one **Red** = off by 2+

Can We Predict Outcomes in Advance?

Balanced accuracy with data truncated 6, 12, and 18 months before the end of the observation window.

Approach	Task	Full Data	6 mo	12 mo	18 mo
MCDA Score LV	Low-Perf. vs. Viable	83%	82%	79%	72%

Can We Predict Outcomes in Advance?

Balanced accuracy with data truncated 6, 12, and 18 months before the end of the observation window.

Approach	Task	Full Data	6 mo	12 mo	18 mo
MCDA Score LV	Low-Perf. vs. Viable	83%	82%	79%	72%
Logistic	**Low-Perf. vs. Viable**	**83%**	**83%**	**83%**	**82%**

Can We Predict Outcomes in Advance?

Balanced accuracy with data truncated 6, 12, and 18 months before the end of the observation window.

Approach	Task	Full Data	6 mo	12 mo	18 mo
MCDA Score LV	Low-Perf. vs. Viable	83%	82%	79%	72%
Logistic	**Low-Perf. vs. Viable**	**83%**	**83%**	**83%**	**82%**
Ordinal	4-Tier (balanced)	58%	57%	56%	55%
Ordinal	4-Tier (adjacent)	92%	93%	91%	91%

Binary logistic model is the most stable. It barely degrades even 18 months out.

Validation: Hand-Labeled Brands

Sample of 10 hand-labeled brands with 6-month forecast predictions, sorted by MCDA score

Extension	Truth
Brand A	Growing
Brand B	Growing
Brand C	Growing
Brand D	Stable
Brand E	Stable
Brand F	Decline
Brand G	Stable
Brand H	Decline
Brand I	Discontinued
Brand J	Discontinued

Validation: Hand-Labeled Brands

Sample of 10 hand-labeled brands with 6-month forecast predictions, sorted by MCDA score

Extension	Truth	MCDA Score (6mo)
Brand A	Growing	0.411
Brand B	Growing	0.369
Brand C	Growing	0.268
Brand D	Stable	0.198
Brand E	Stable	0.195
Brand F	Decline	0.159
Brand G	Stable	0.144
Brand H	Decline	0.101
Brand I	Discontinued	0.088
Brand J	Discontinued	0.045

Validation: Hand-Labeled Brands

Sample of 10 hand-labeled brands with 6-month forecast predictions, sorted by MCDA score

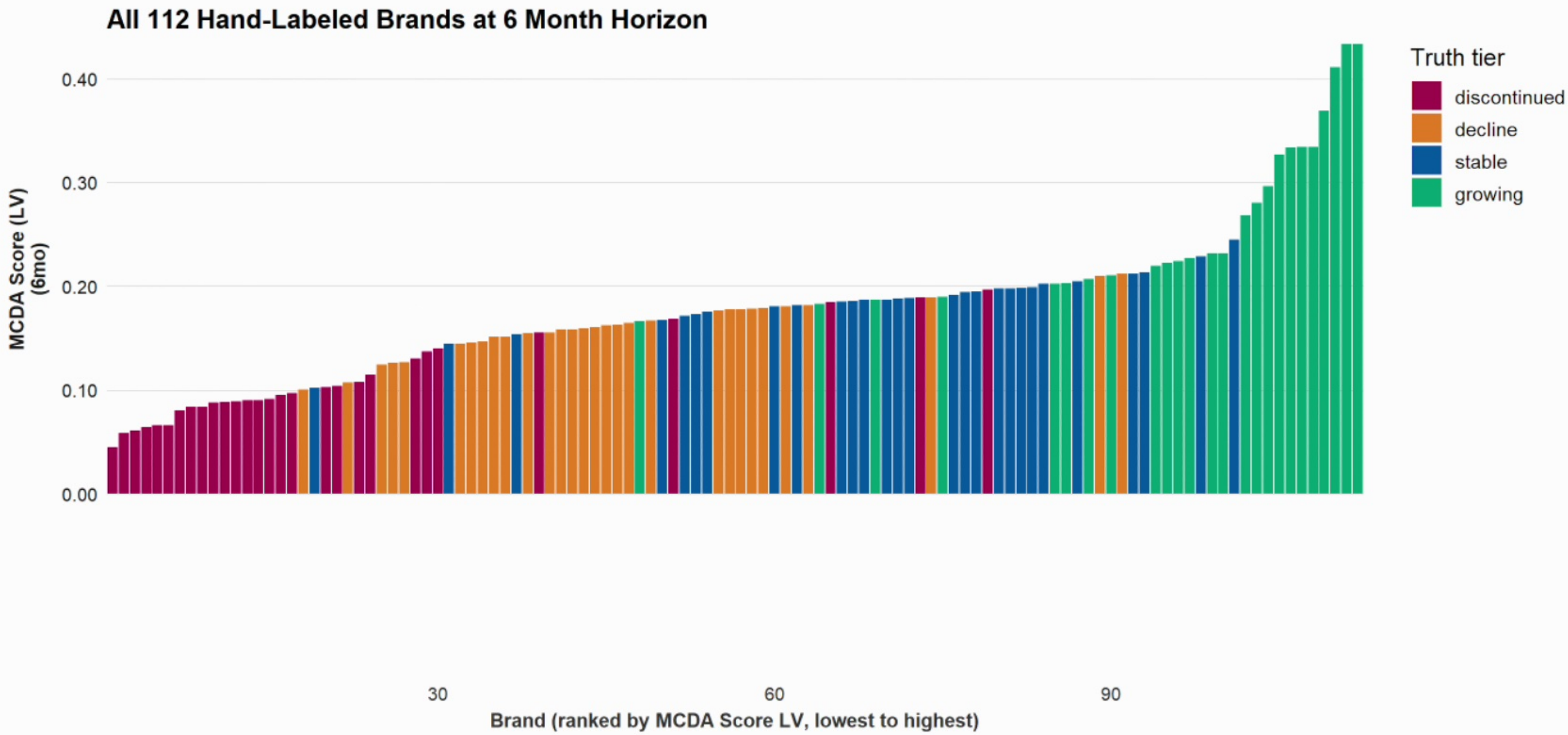
Extension	Truth	MCDA Score (6mo)	Logistic (6mo)
Brand A	Growing	0.411	Viable
Brand B	Growing	0.369	Viable
Brand C	Growing	0.268	Viable
Brand D	Stable	0.198	Viable
Brand E	Stable	0.195	Viable
Brand F	Decline	0.159	Low Performance
Brand G	Stable	0.144	Low Performance
Brand H	Decline	0.101	Low Performance
Brand I	Discontinued	0.088	Low Performance
Brand J	Discontinued	0.045	Low Performance

Validation: Hand-Labeled Brands

Sample of 10 hand-labeled brands with 6-month forecast predictions, sorted by MCDA score

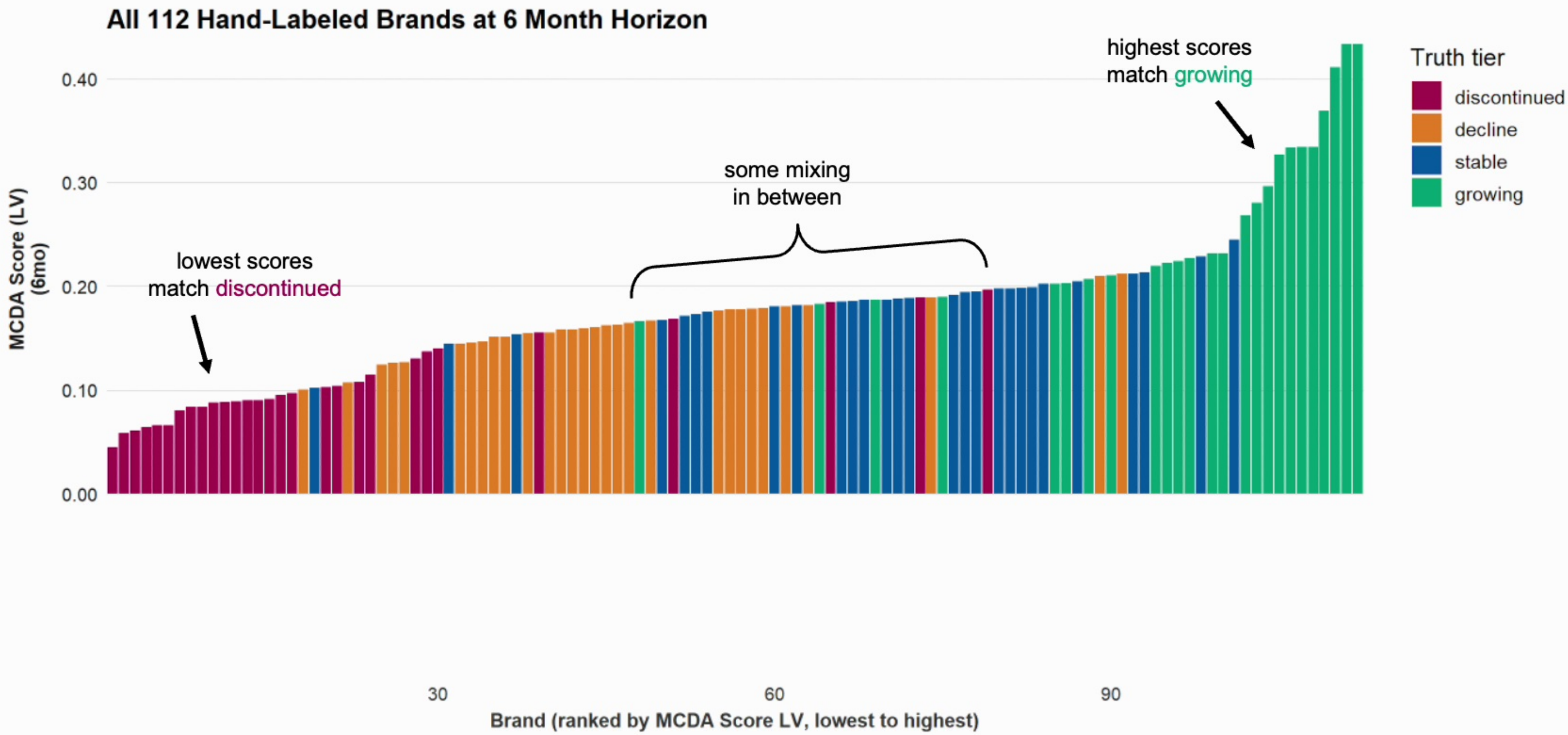
Extension	Truth	MCDA Score (6mo)	Logistic (6mo)	Ordinal (6mo)
Brand A	Growing	0.411	Viable	Growing
Brand B	Growing	0.369	Viable	Growing
Brand C	Growing	0.268	Viable	Growing
Brand D	Stable	0.198	Viable	Stable
Brand E	Stable	0.195	Viable	Stable
Brand F	Decline	0.159	Low Performance	Decline
Brand G	Stable	0.144	Low Performance	Decline
Brand H	Decline	0.101	Low Performance	Discontinued
Brand I	Discontinued	0.088	Low Performance	Discontinued
Brand J	Discontinued	0.045	Low Performance	Discontinued

Validation: Full Hand-Labeled Set

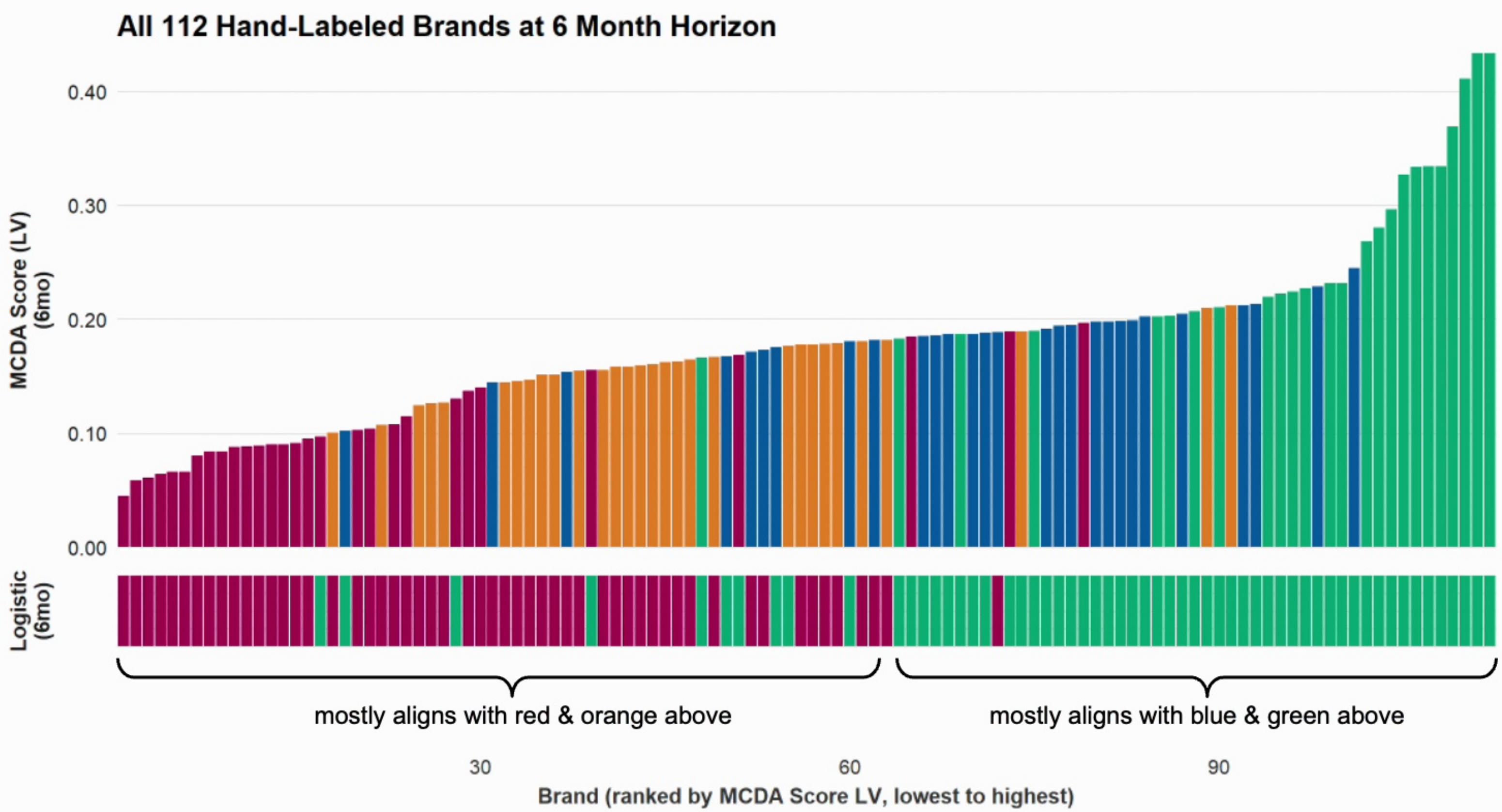


DataScience@SMU

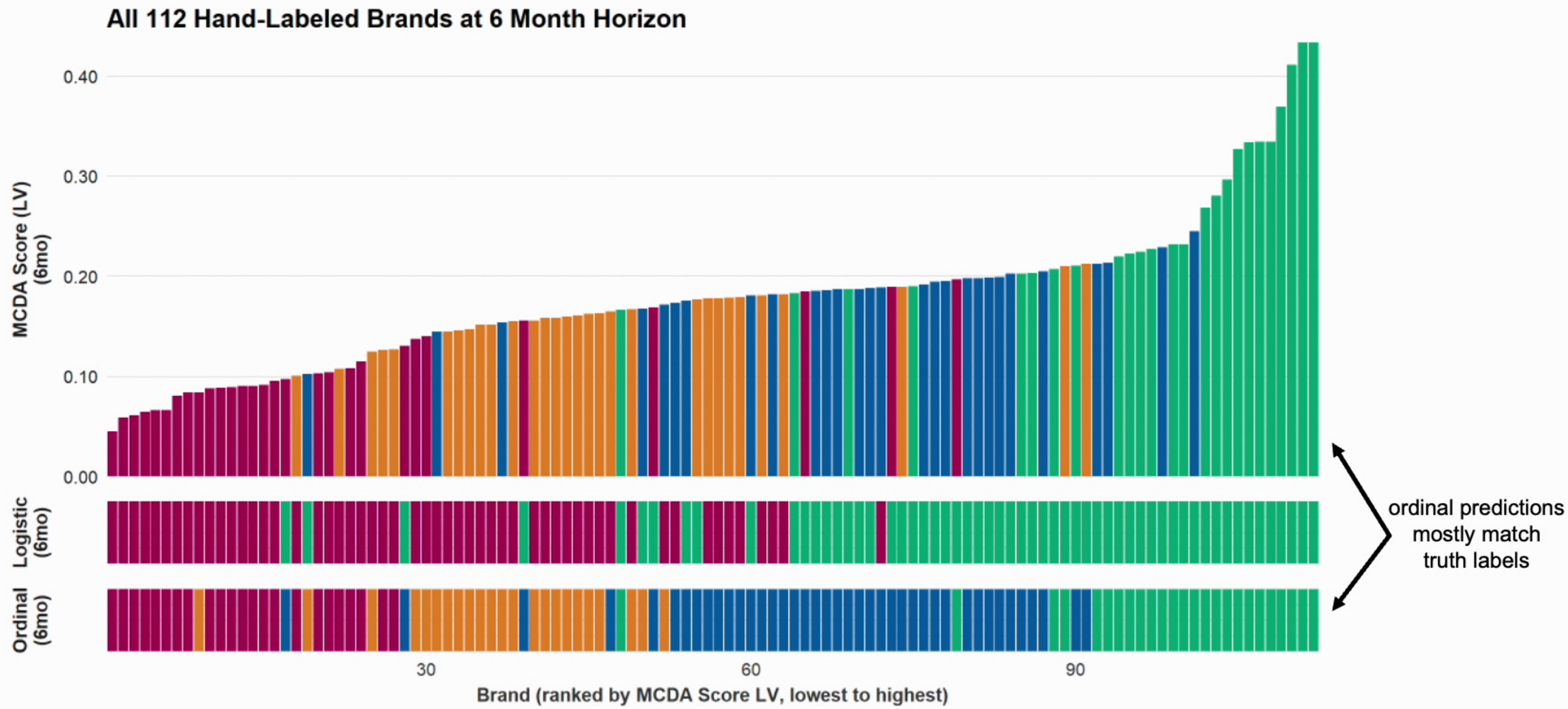
Validation: Full Hand-Labeled Set



Validation: Full Hand-Labeled Set



Validation: Full Hand-Labeled Set



DataScience@SMU

Takeaways

WHAT: A simple demand feature separates brand tiers.

Median year-over-year case growth distinguishes low-performing from viable brands.

Takeaways

WHAT: A simple demand feature separates brand tiers.

Median year-over-year case growth distinguishes low-performing from viable brands.

WHEN: The framework works as a forward-looking signal.

Prediction accuracy is stable even when computed up to 18 months before the present.

Takeaways

WHAT: A simple demand feature separates brand tiers.

Median year-over-year case growth distinguishes low-performing from viable brands.

WHEN: The framework works as a forward-looking signal.

Prediction accuracy is stable even when computed up to 18 months before the present.

WHY: This is a proof of concept designed to grow.

The framework is validated by statistical models and is designed to expand as new criteria are tested.

Thank You

Questions?

Connect with me on LinkedIn

Kristin Henderson
hendersonk@smu.edu
kdhenderson.github.io

Appendix

Brand Hierarchy

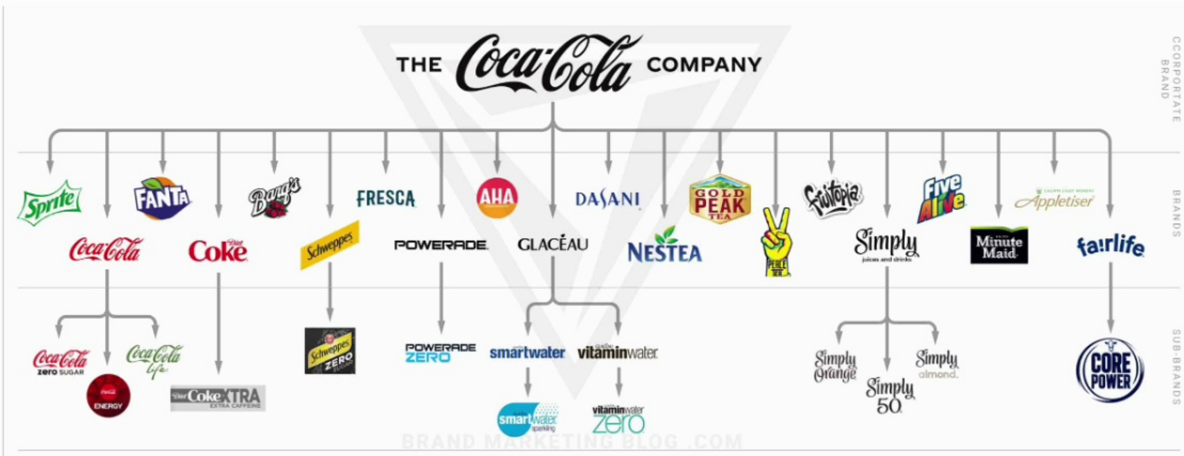
Each brand is identified at four levels. Analysis is conducted at the extension level.

Brand Hierarchy

Each brand is identified at four levels. Analysis is conducted at the extension level.

Brand Owner

Coca-Cola Corporation

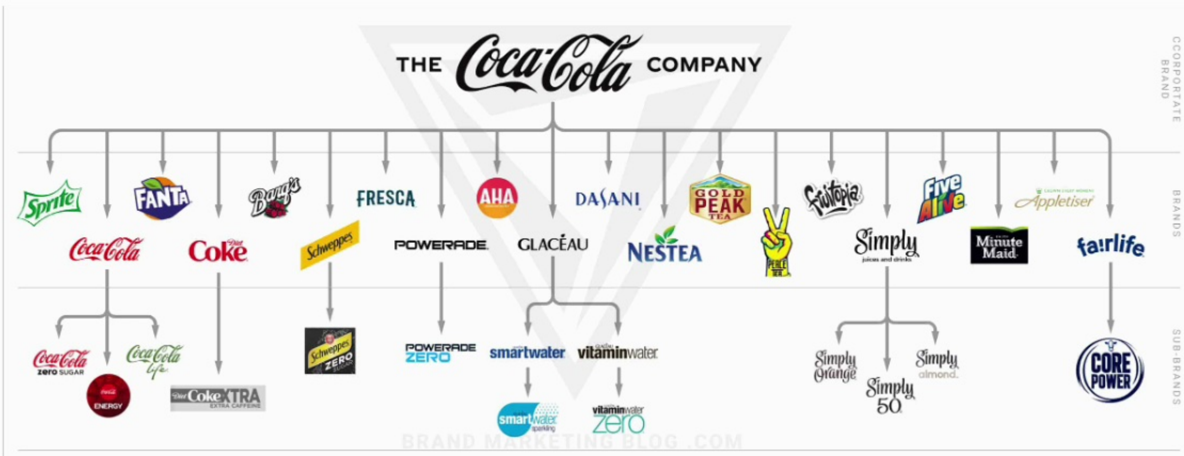


Brand Hierarchy

Each brand is identified at four levels. Analysis is conducted at the extension level.

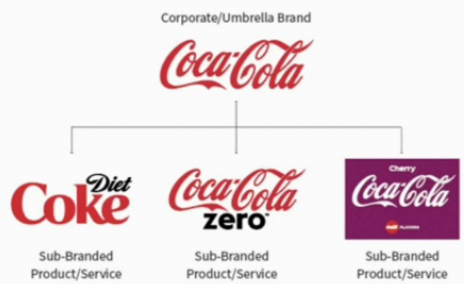
Brand Owner

Coca-Cola Corporation



Brand Family

Coca-Cola



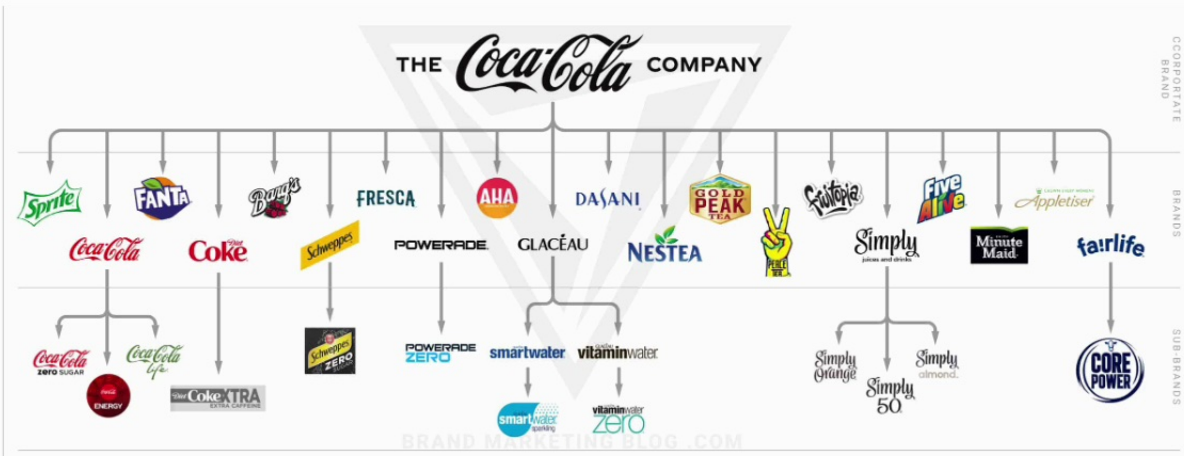
DataScience@SMU

Brand Hierarchy

Each brand is identified at four levels. Analysis is conducted at the extension level.

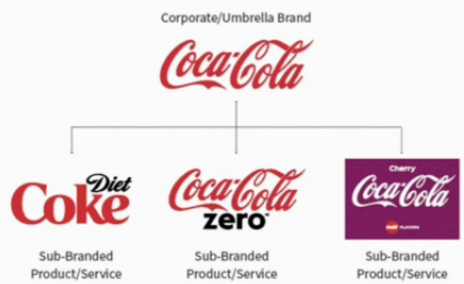
Brand Owner

Coca-Cola Corporation



Brand Family

Coca-Cola



Brand Extension

Cherry Coke

Analysis level

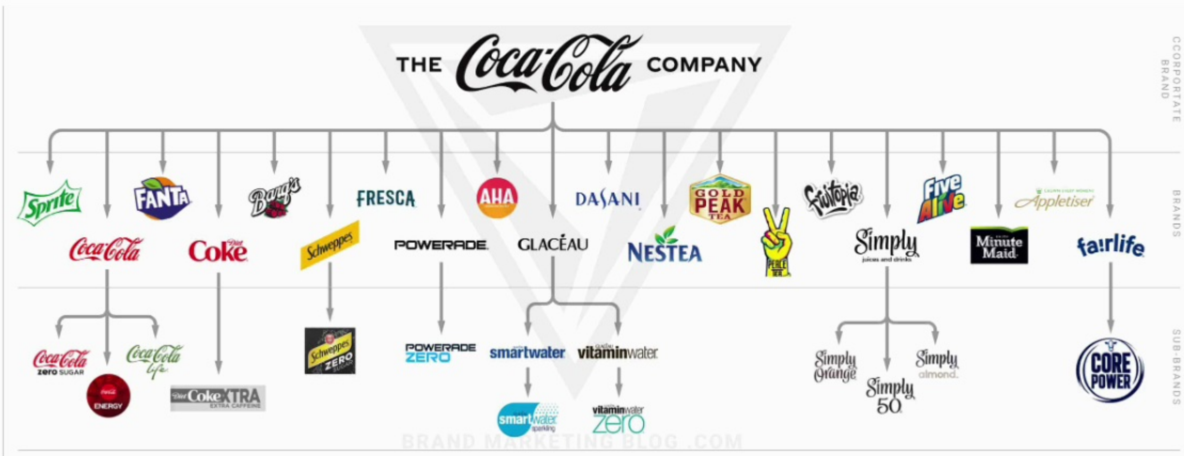
DataScience@SMU

Brand Hierarchy

Each brand is identified at four levels. Analysis is conducted at the extension level.

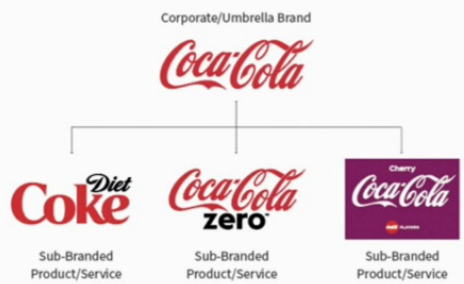
Brand Owner

Coca-Cola Corporation



Brand Family

Coca-Cola



Brand Extension

Cherry Coke

Analysis level

UPC

12 oz can

DataScience@SMU

Why Case Sales?

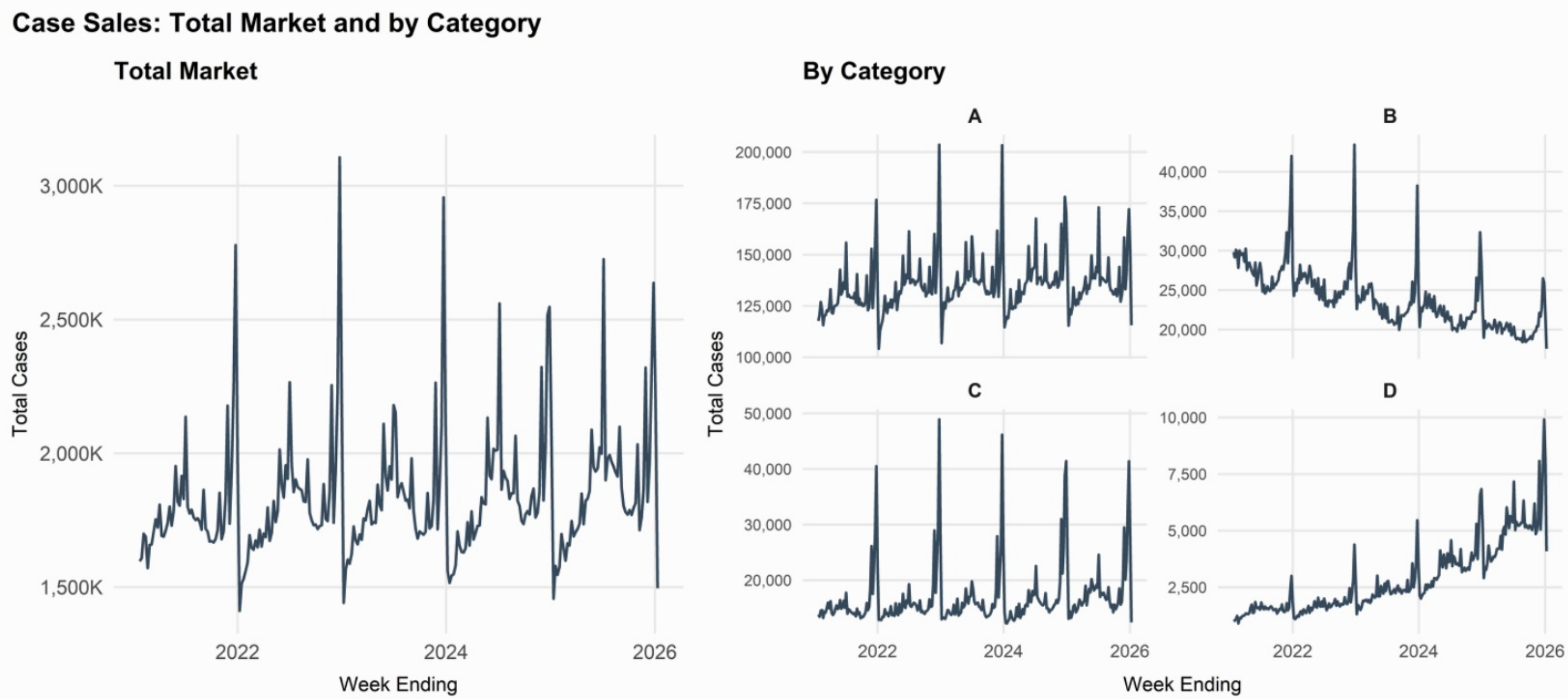
Case sales standardize packaging units, making product movement comparable across formats.

Unit Sales	Case Sales
1 case of 2L bottles = **6 units**	1 case of 2L bottles = **1 case**
1 case of 12oz cans = **24 units**	1 case of 12oz cans = **1 case**
Not comparable across formats: Different container sizes → unit counts are misleading	**Comparable** across formats: Both represent a standard package sold

Case sales are a standardized packaging metric.
Dollar sales are also available but are more sensitive to price promotions. Both are informative.
Divergence suggests pricing/promotional effects.

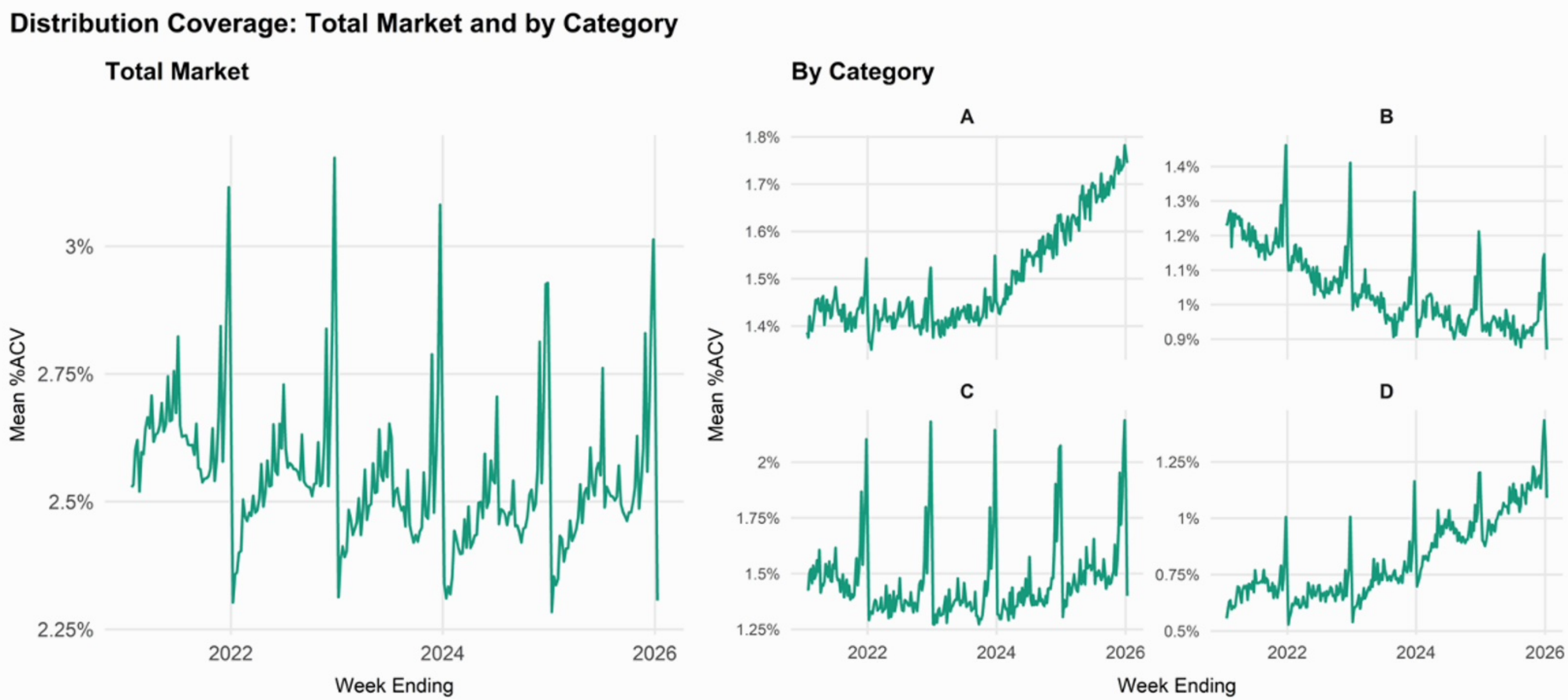
Market and Category Context

Market-level trends and example category trends. Category context is planned as the next framework feature.



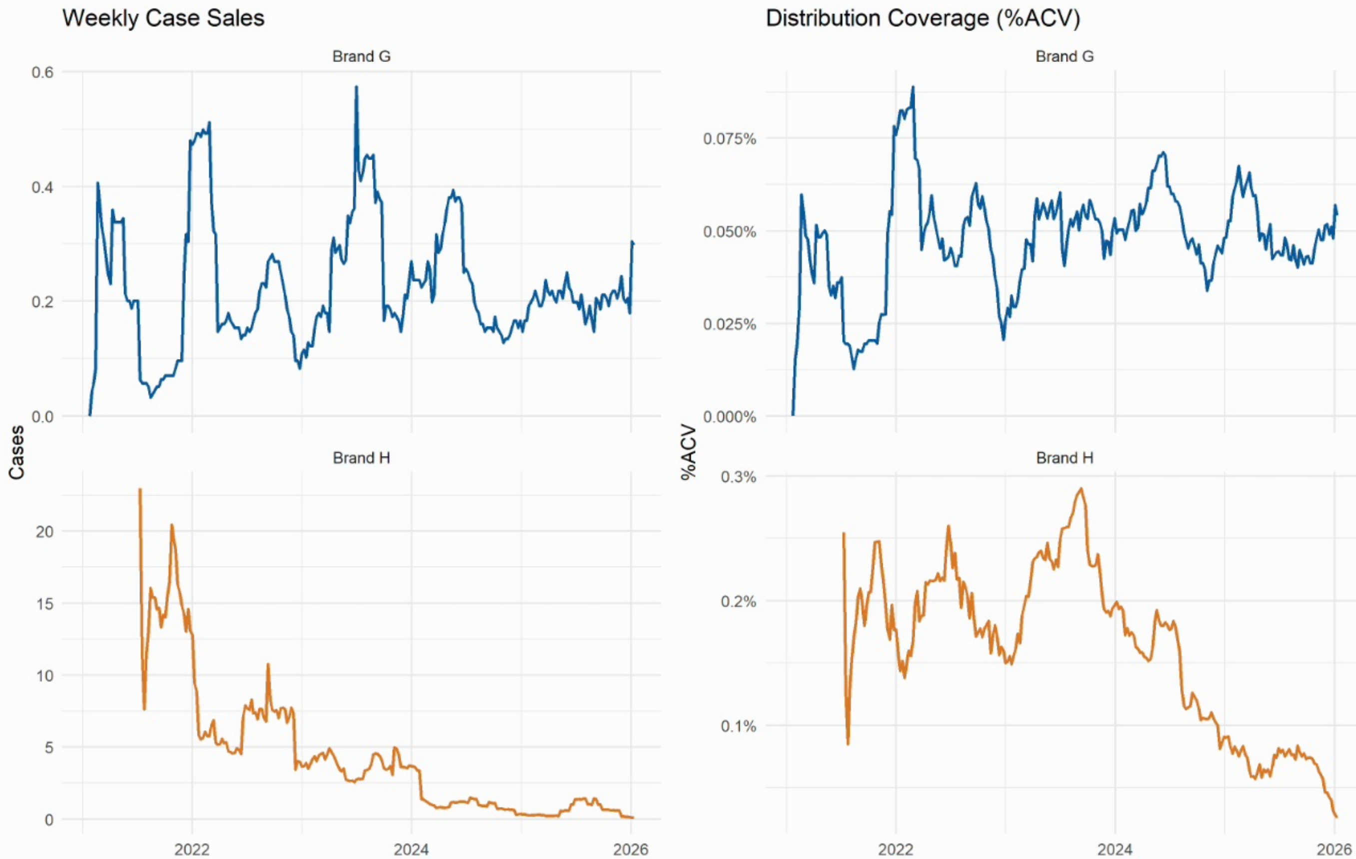
Market and Category Context

Market-level trends and example category trends. Category context is planned as the next framework feature.



DataScience@SMU

What's Going on with Brand G & H?



Validation: Stakeholder Truth Set

18 stakeholder families (10 bankrupt, 8 acquired). **Features computed 6 months prior to event date.**

Family	Outcome	MCDA Score	Logistic	Ordinal
Brand A	Acquired	0.300	Viable	Growing
Brand B	Acquired	0.270	Viable	Growing
Brand C	Acquired	0.203	Viable	Stable
Brand D	Acquired	0.190	Viable	Stable
Brand E	Acquired	0.190	Viable	Stable
Brand F	Bankrupt	0.190	Viable	Stable
Brand G	Bankrupt	0.189	Viable	Stable
Brand H	Bankrupt	0.187	Viable	Stable
Brand I	Bankrupt	0.183	Viable	Stable
Brand J	Acquired	0.172	Low Performance	Decline
Brand K	Bankrupt	0.170	Viable	Stable
Brand L	Acquired	0.158	Viable	Stable
Brand M	Bankrupt	0.141	Low Performance	Decline
Brand N	Bankrupt	0.137	Low Performance	Discontinued
Brand O	Bankrupt	0.135	Low Performance	Decline
Brand P	Acquired	0.133	Low Performance	Decline
Brand Q	Bankrupt	0.130	Low Performance	Decline
Brand R	Bankrupt	0.108	Low Performance	Discontinued