in4mation insights

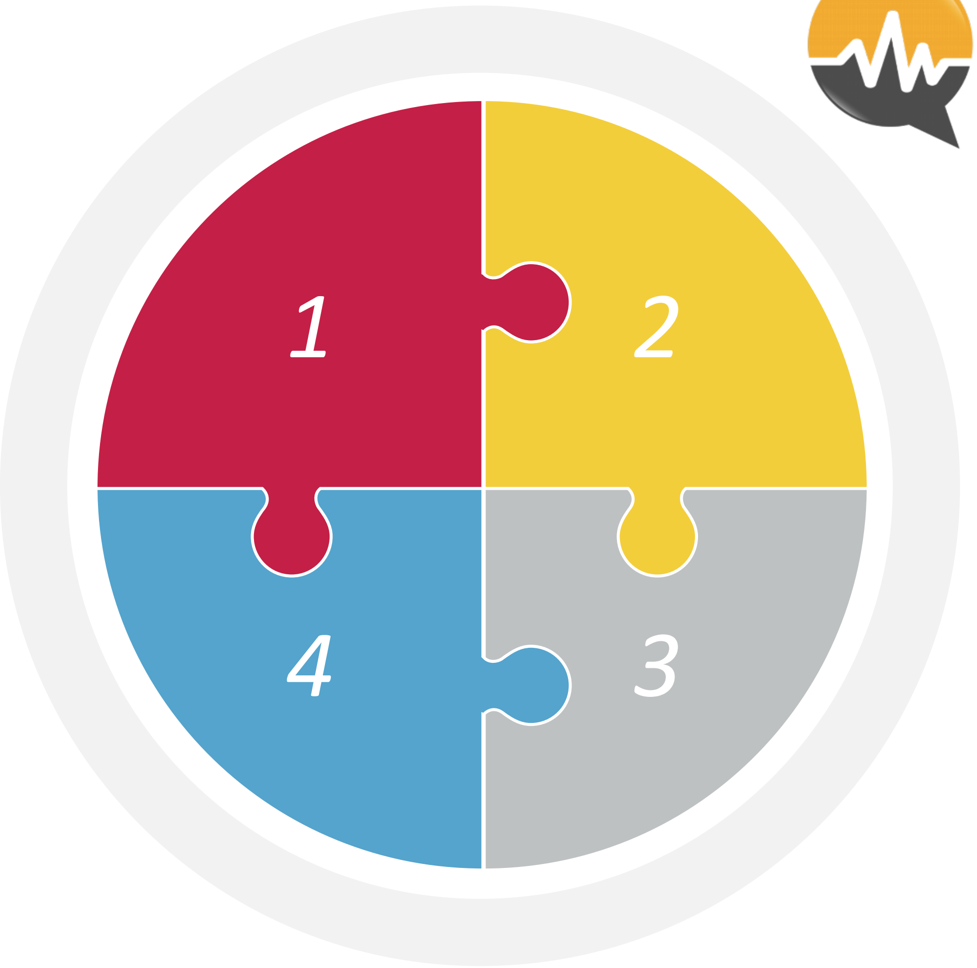
Using Social Media Data to Predict New Product Success

3 April 2019

Overview of Presentation

- Getting Social Media (SM) to *Prediction Grade*
- Machine Learning Using Motives
- The Motive Model
- Preparing Sales Data to Play Nicely With SM
- Proof that SM Can Predict New Product Sales
- Key Takeaways & New Directions

in4mation insights and Converseon partnership: We take Converseon "research grade" social data and build it into our predictive models



Category Expertise

New Product Forecasting
Food & Beverages

Collaborative Methodology

Converseon Convey API allows all of us to participate in refining the rules, so results are relevant to our issues

Multiple Metrics

Beyond sentiment—emotions and intensity using NLP for better predictors of sales and more precise social context

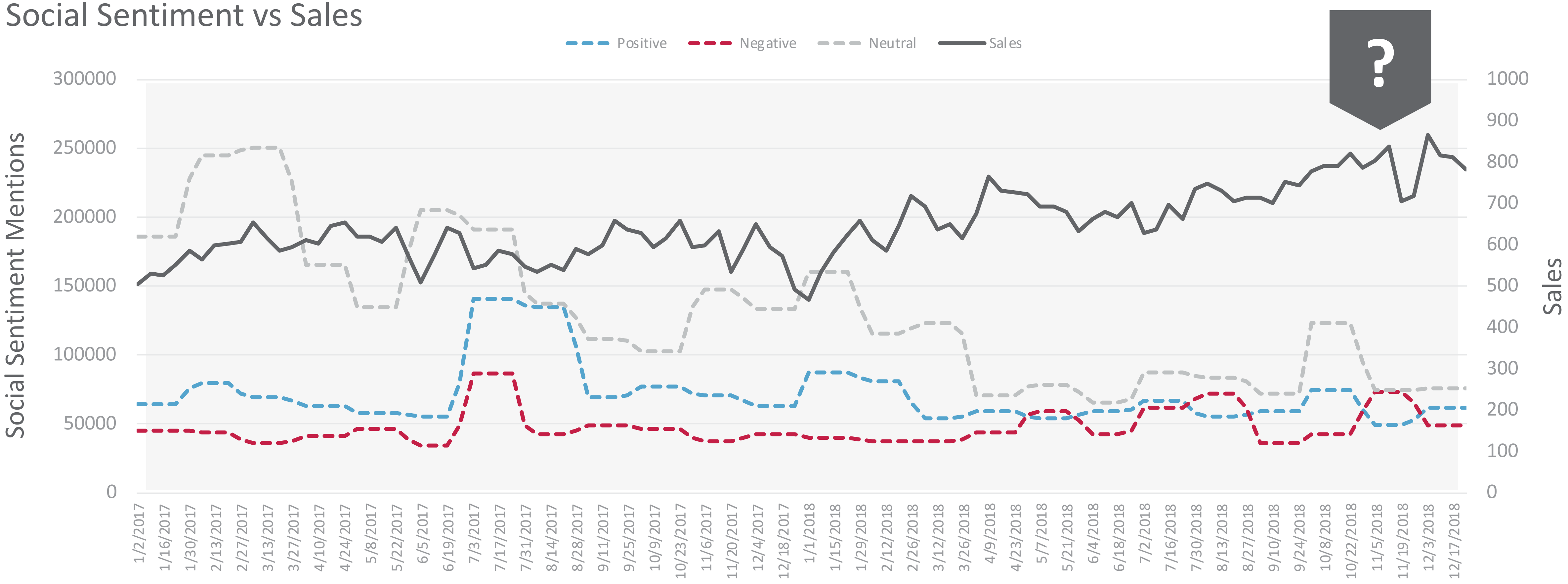
Modeling Integrators

Bayesian Experts
Innovative Approaches
Link Attributes to Sales
Segmenting Markets

Using Natural Language Processing (NLP) for sales prediction

How do we develop social media time series that are prediction grade?

Social Sentiment vs Sales



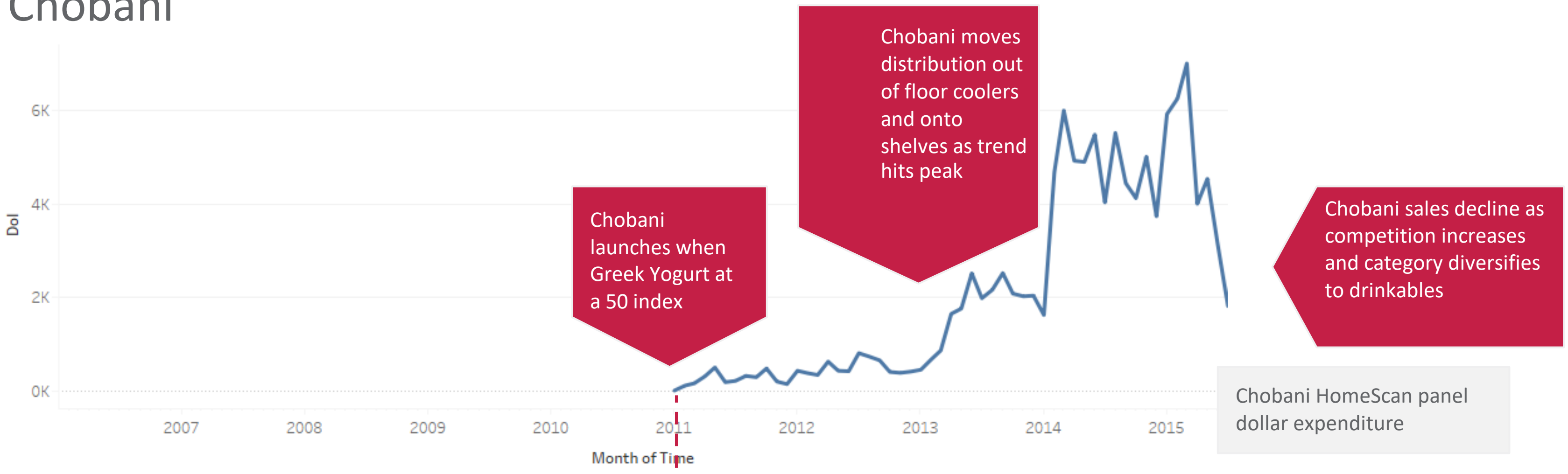
Can social media be used to predict new product success?

YOGURT EXAMPLE

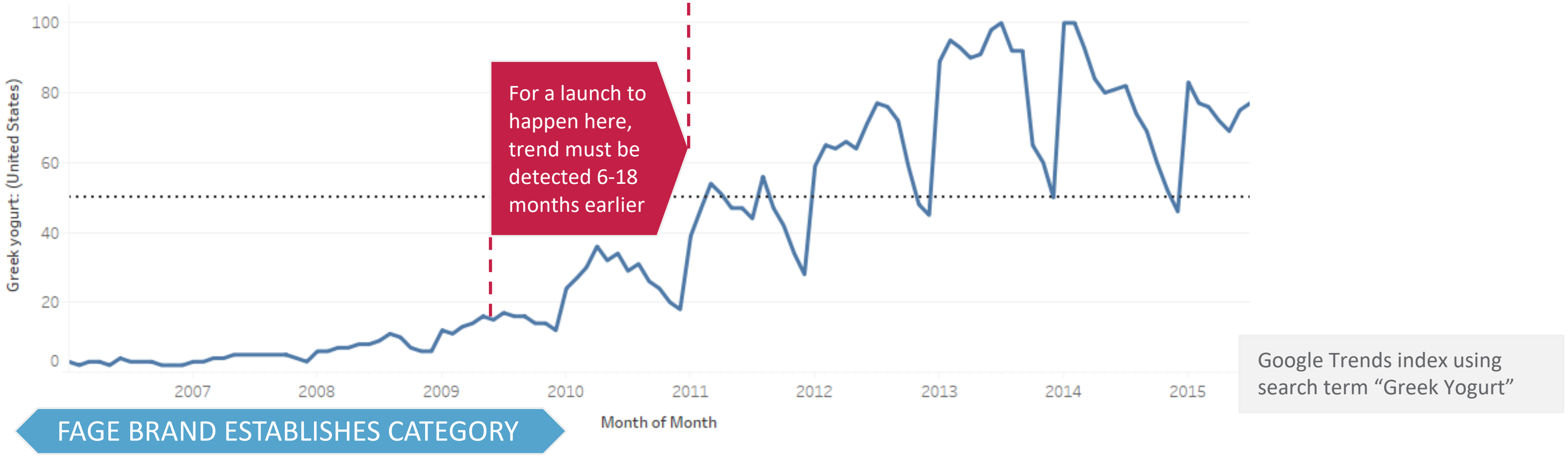
Depending on how fast you can bring a product to market, a new trend must be detected 6-18 months before launch.

Even identifying potential acquisition targets requires significant lead time.

Chobani



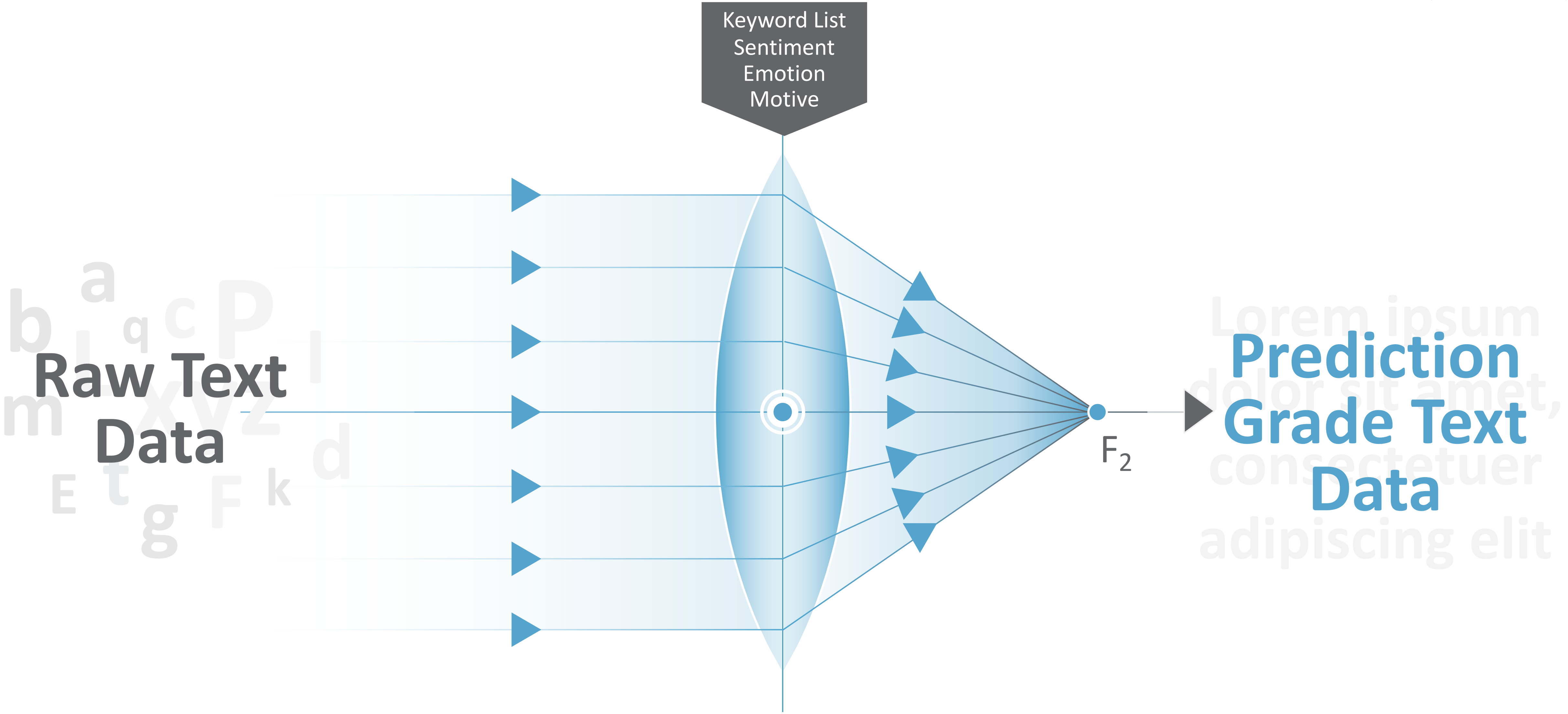
Greek Yogurt



What does text data look like?

we were at parrot cay the first night and met our servers richard was the head server santana the server and tita the assistant server all were excellent the entire cruise really really attentive and friendly we learned an important lesson the first night don t order fish ever more later in this report dw ordered the grouper and it was so fishy she couldn t eat it we do enjoy good food and dining out at home and she s not a picky eater however she earned that reputation the first night after she couldn t eat the fish she just wanted a chicken breast she really didn t eat it either so the staff insisted on bringing another dish it was a vegetarian mushroom dish that she said was very good a general note on the food for the entire trip it was good but unremarkable we all found something we could enjoy every night with the exception of the fish more later nothing was really bad

Motive as a lens

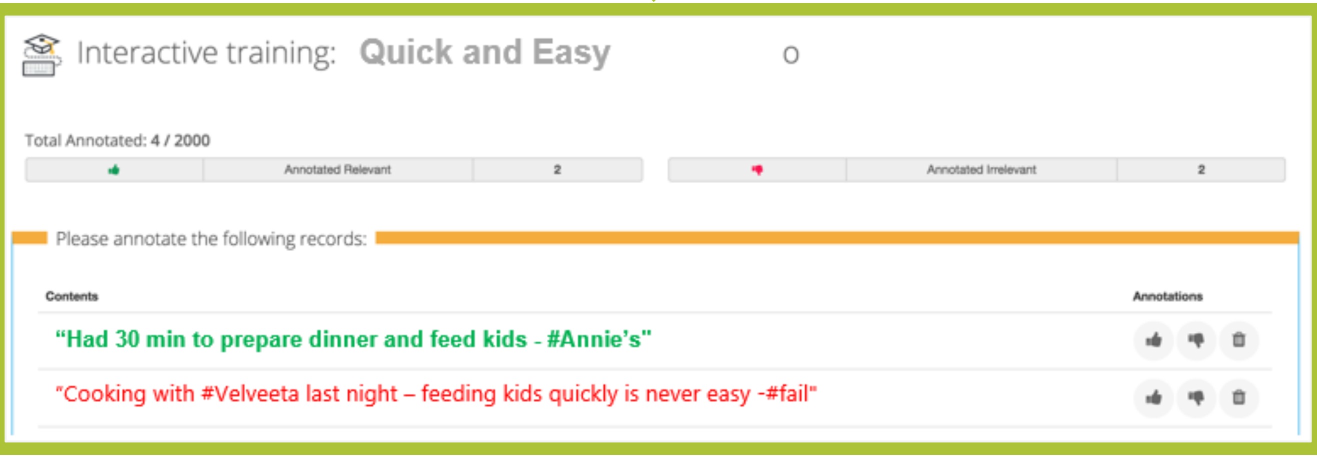


Why Converseon's Convey API is better than other SL platforms

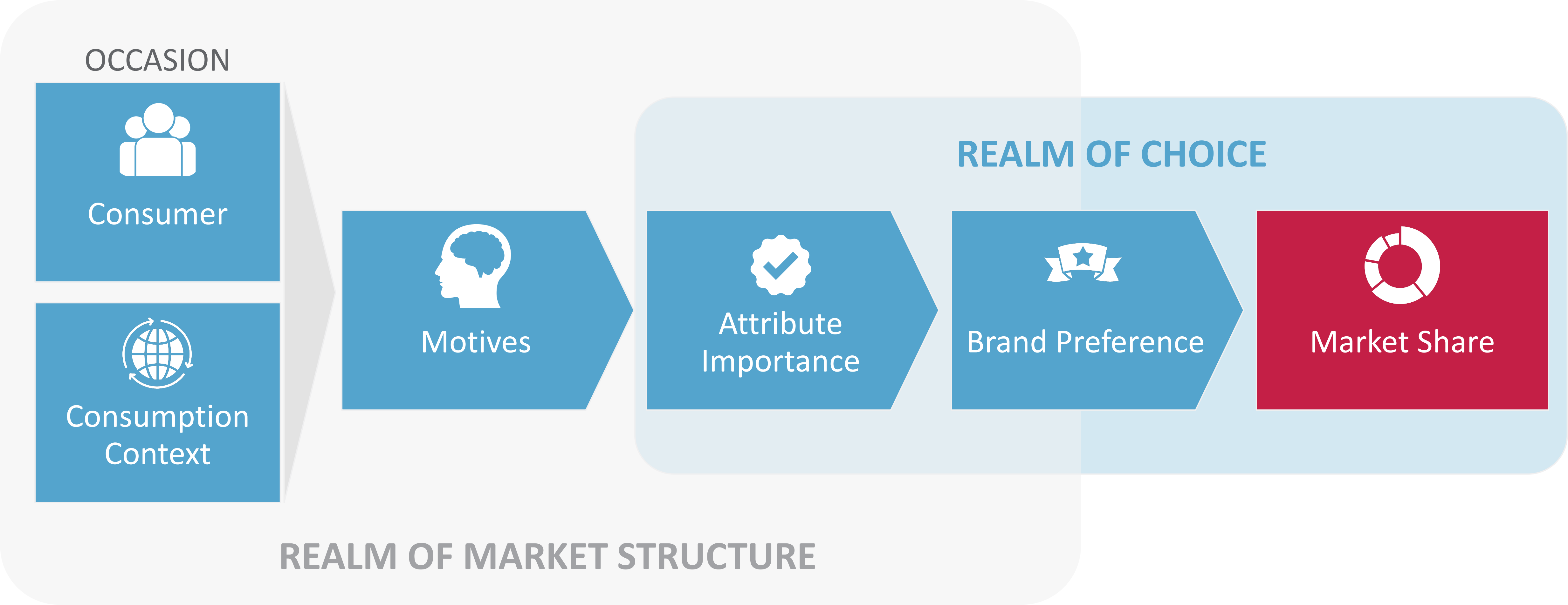
Develop a variety of strategies to capture motives

Motive Wording	Short Handle	Examples of Machine Learning Analysis
"It's a family favorite recipe"	Favorite Recipe	= Keyword List + Positive Sentiment / Joy
" I wanted to satisfy a craving"	Craving	= Anticipation + Positive Sentiment (to be validated)
" I needed something that was quick and easy to prepare"	Quick & Easy	Convey Custom Classifiers

Jointly tune the semi-supervised model



The model predicts **market share** by telling us which motives drive preference for which feature levels

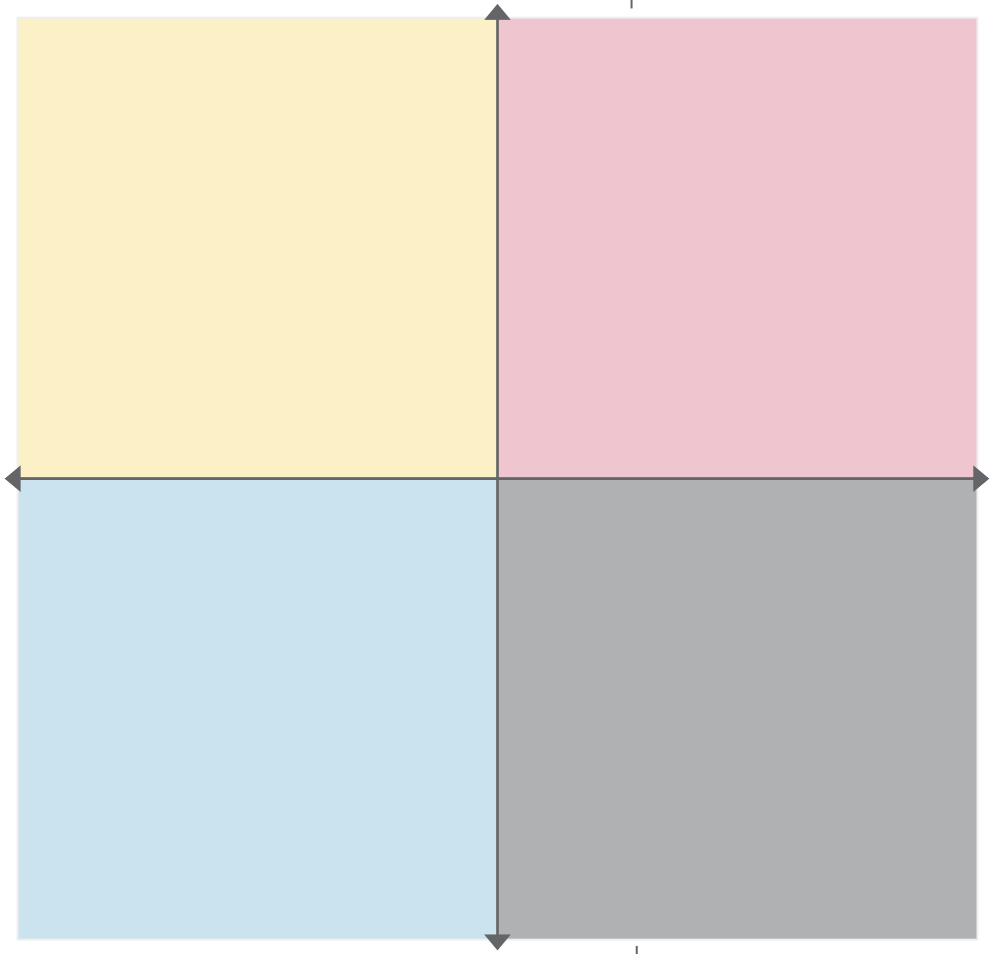


Mental model for developing motives

EMOTIONAL | Feel a certain way

TO **MOVE AWAY** FROM AN **UNDESIRABLE** STATE

- Solve a problem
- Prevent a problem
- Maintain a current situation, routine



TO **MOVE TOWARDS** A **DESIRABLE** STATE

- Explore interests
- Sensory pleasure

FUNCTIONAL | Achieve a specific objective

Motives break down into different types….

EMOTIONAL

Emotional / Avoid Problems

I wanted a brand I can trust
I wanted to calm myself
I wanted to feel safe
I did not want to be surprised

Emotional / Desirable

I wanted to be in a good mood
I wanted to take a break
I wanted to take care of family
I wanted some variety

TO **MOVE AWAY** FROM AN **UNDESIRABLE** STATE

TO **MOVE TOWARDS** A **DESIRABLE** STATE

Functional / Avoid Problems

I wanted something with no artificial ingredients
I wanted something convenient
I wanted to satisfy a craving
I wanted something less processed

Functional / Desirable

I wanted to eat organic food
I wanted something with recognizable ingredients
I wanted something filling
I want something that can be eaten everyday

FUNCTIONAL

Defining the motive filters: *Recognize Ingredients*

Relevance—Topic Modeling

Assumed Relevant

High-Confidence Results	Ingredients Driving Excitement/Enjoyment about Food Greens & Ingredients That Are Specifically Perceived as Healthy
Interesting High-Confidence Sub-Themes	Breakfast Motivated to Shop at Farmer's Markets & Eat Locally/Sustainably Sourced Food
Lower-Confidence Results	Motivated to Eat Organic Food (Very Common: "I Love Organic Food.") Health-Conscious Eating & Recipe Sharing without Detailed Ingredient Profiles Veganism: Posts about Seeking Vegan Food Options & Eating Vegan Healthy/ Clean Eating and Fitness/ Weight Loss
Results Assumed Irrelevant	Posts Generally about Veganism ("Ethical/Environmental" Veganism, No Mention of Specific Food / Not about Current Eating Activities) Lists of Ingredients in Recipes (Lacking Consumer Commentary, often Recipe Sharing on Facebook)

Assumed Irrelevant

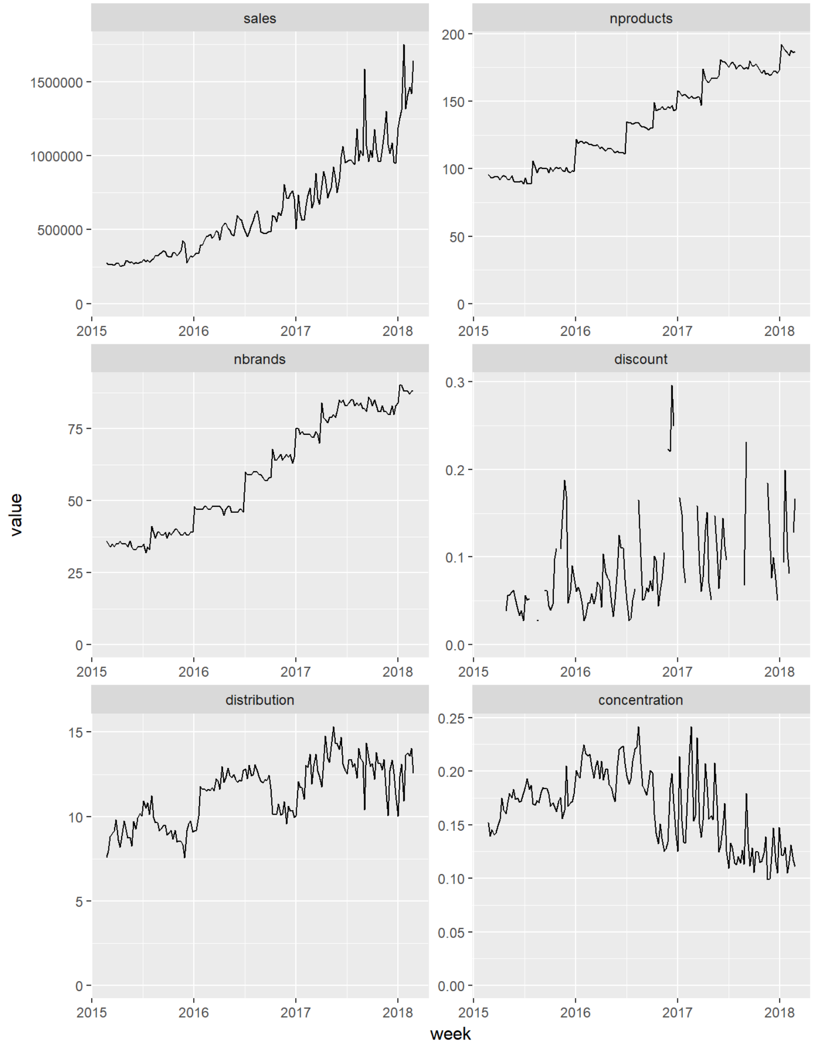
Motive Model—High Correlation

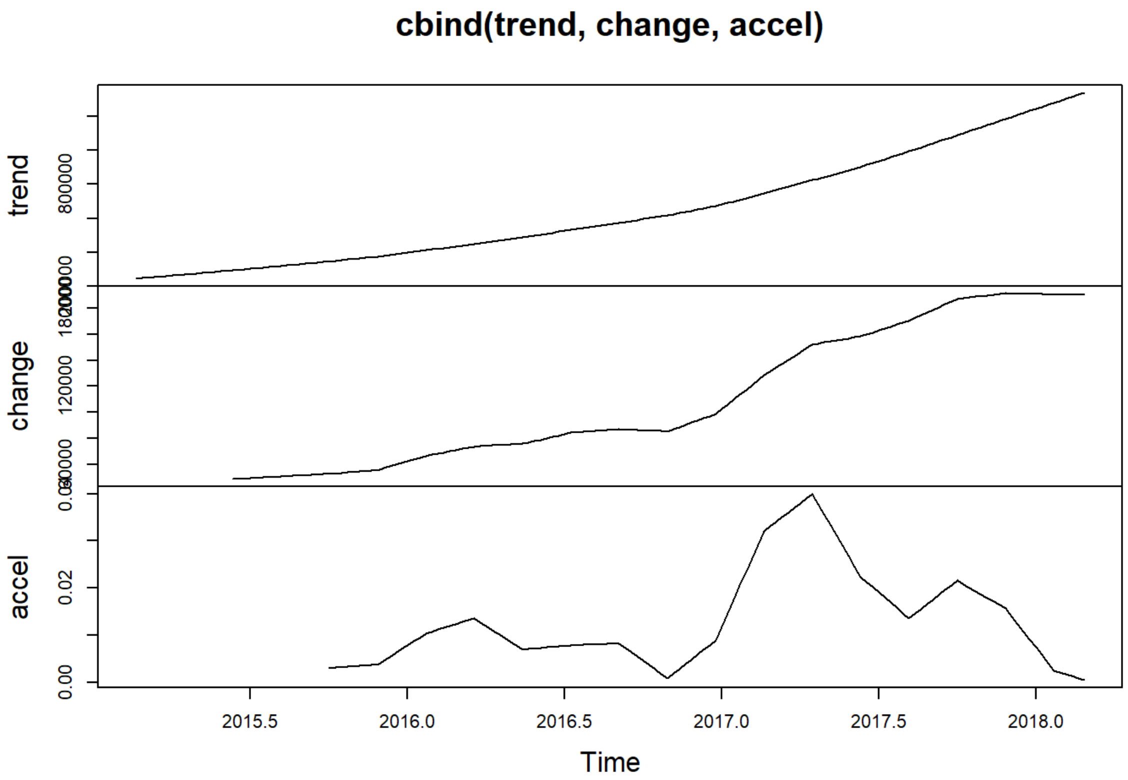
Clean Eating

- Recognize ingredients
- Not overly processed
- Fresh ingredients
- Feel smart about food choices
  - Use of color green
  - LUVO PLANTED
  - EVOL
  - ANNIE'S

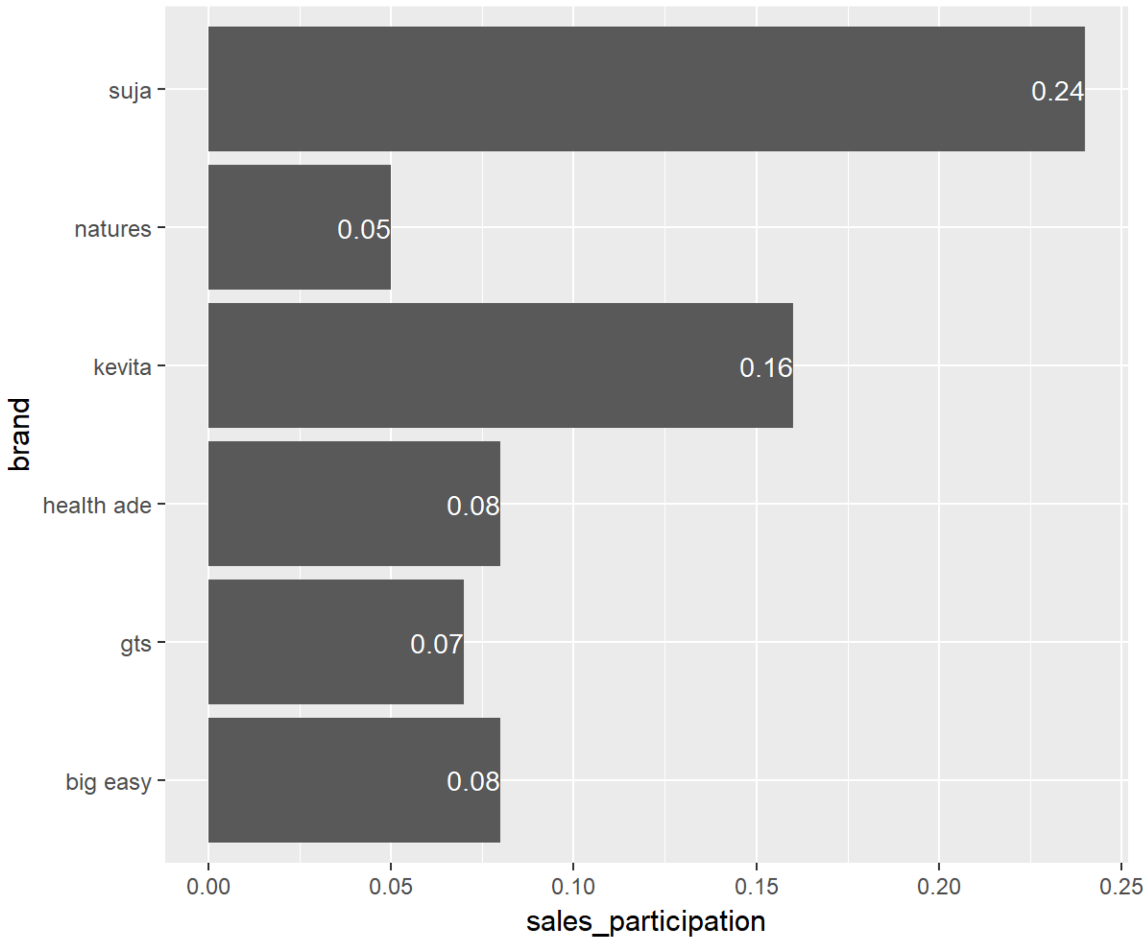

BUTTER
PURE CANE
Sugar
LIGHT BROWN
GREEK
Bell
Tea
Wine

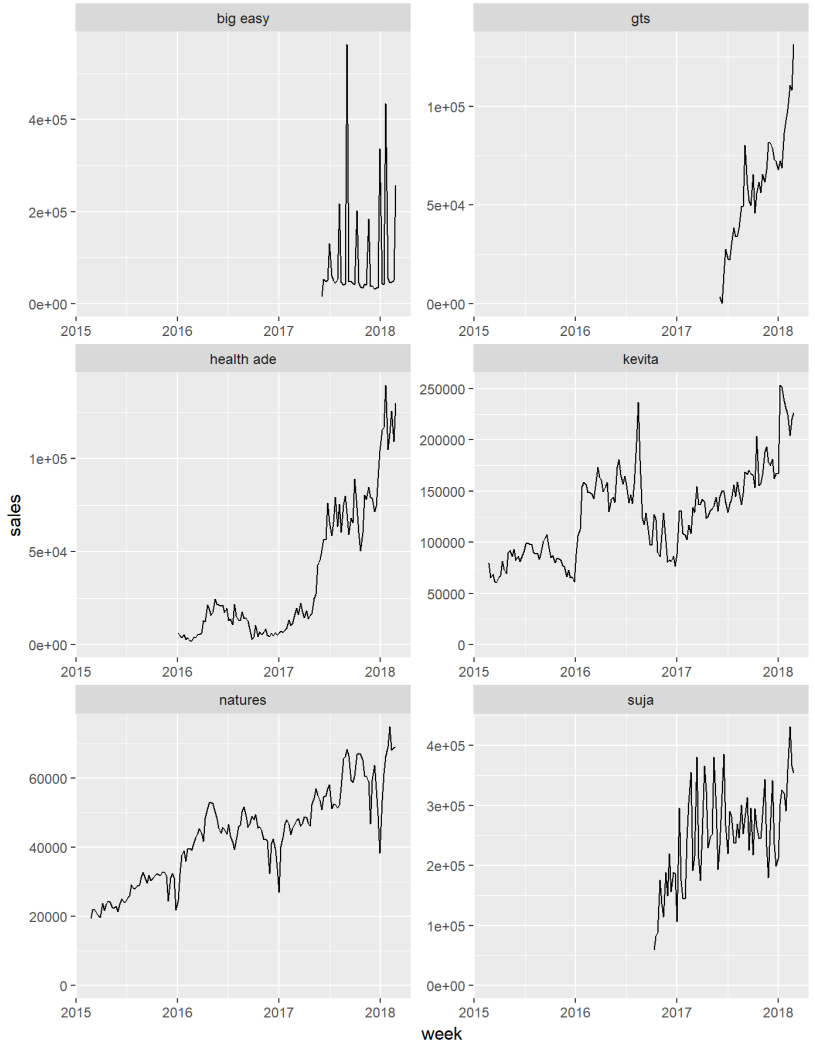
What does the sales data look like?





Is this a one-off or an incipient trend?





Using NLP on sales data – descriptive labels

Identify

Attributes were identified by parsing the UPC text description of each product, *e.g.*

180 SNACKS ALMOND POP BLUEBERRS NUTRITIONAL SNACK BR RICE RP 1 OZ - 0180332000041

Identifiable Attributes

- brand is 180 Snacks
- product is a snack in the shape of bar
- contains rice, almond & blueberries
- size is 1 oz

Text Normalization

In order to identify attributes, text was normalized to account for different spelling, stemming, etc. *e.g.* BLUEBERRS, BLUEBERRY, BLUEBERRIES.

Bigram Attributes

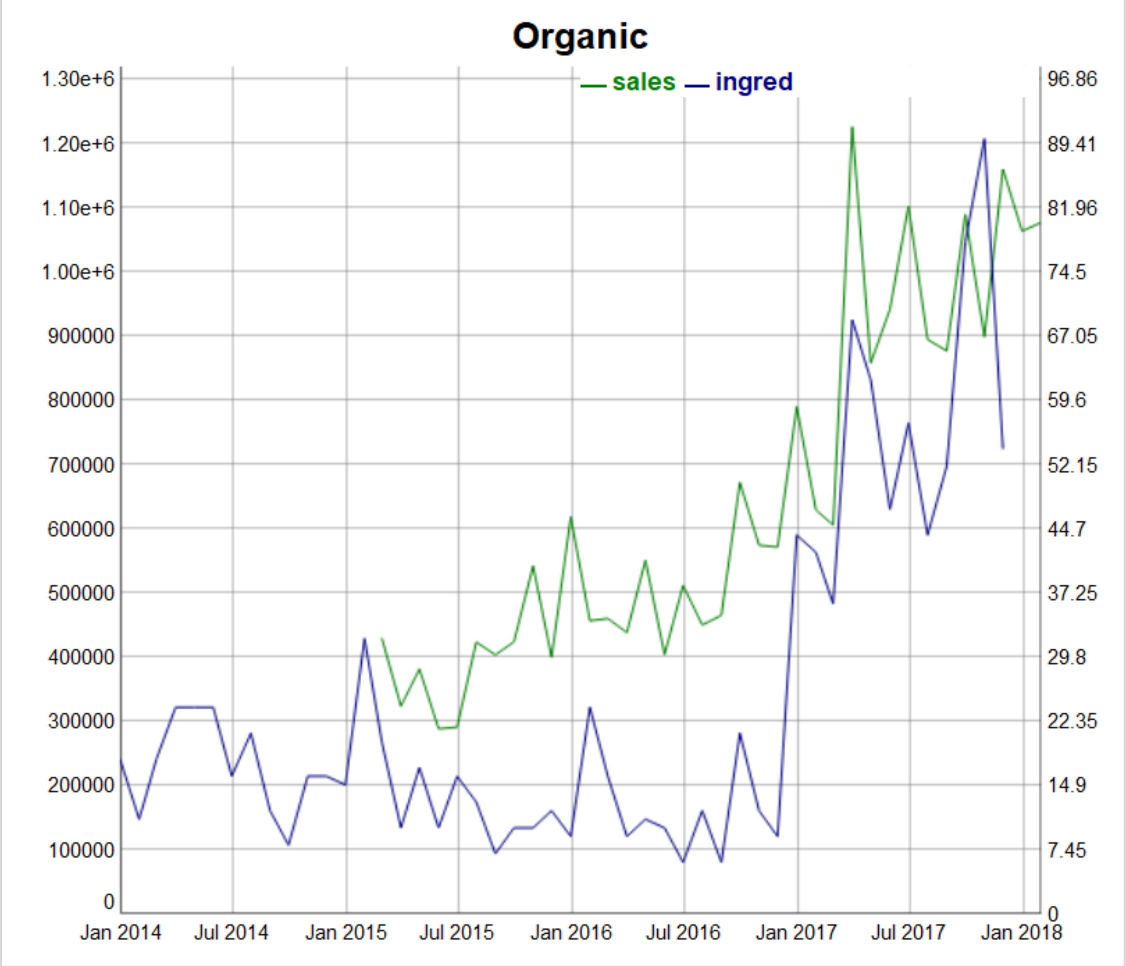
In order to identify bigram attributes/concepts, word correlations were computed, *e.g.*, almond milk, non-dairy, meat substitute, gluten free.

Co-occurring Attributes

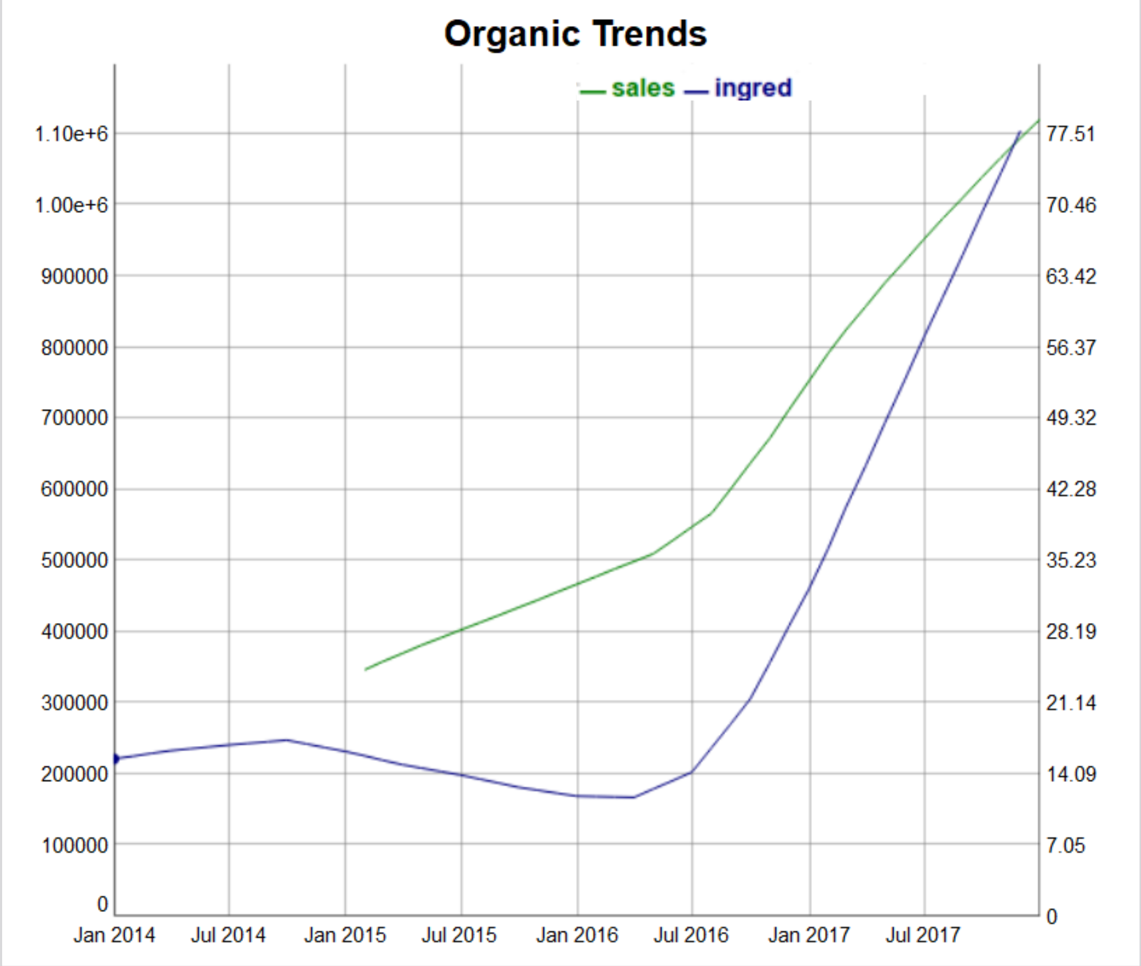
In order to identify co-occurring attributes word/ bi-gram correlations were computed, *e.g.*, (chocolate cookies, ginger tea).

The motive (recognize **ingred**ients) predicts an acceleration in sales of organic foods with recognizable ingredients

Converseon SM trained by motive

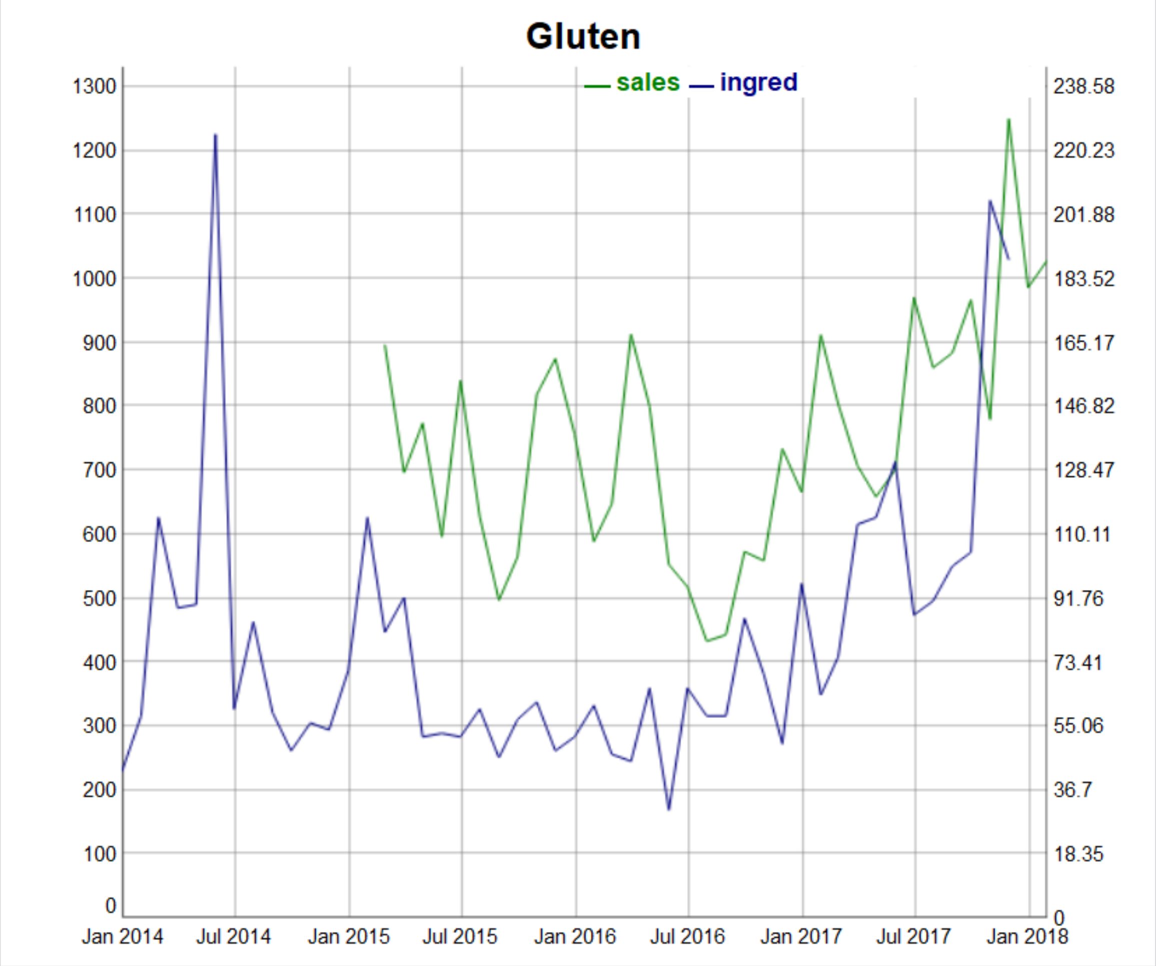


Smoothed time series

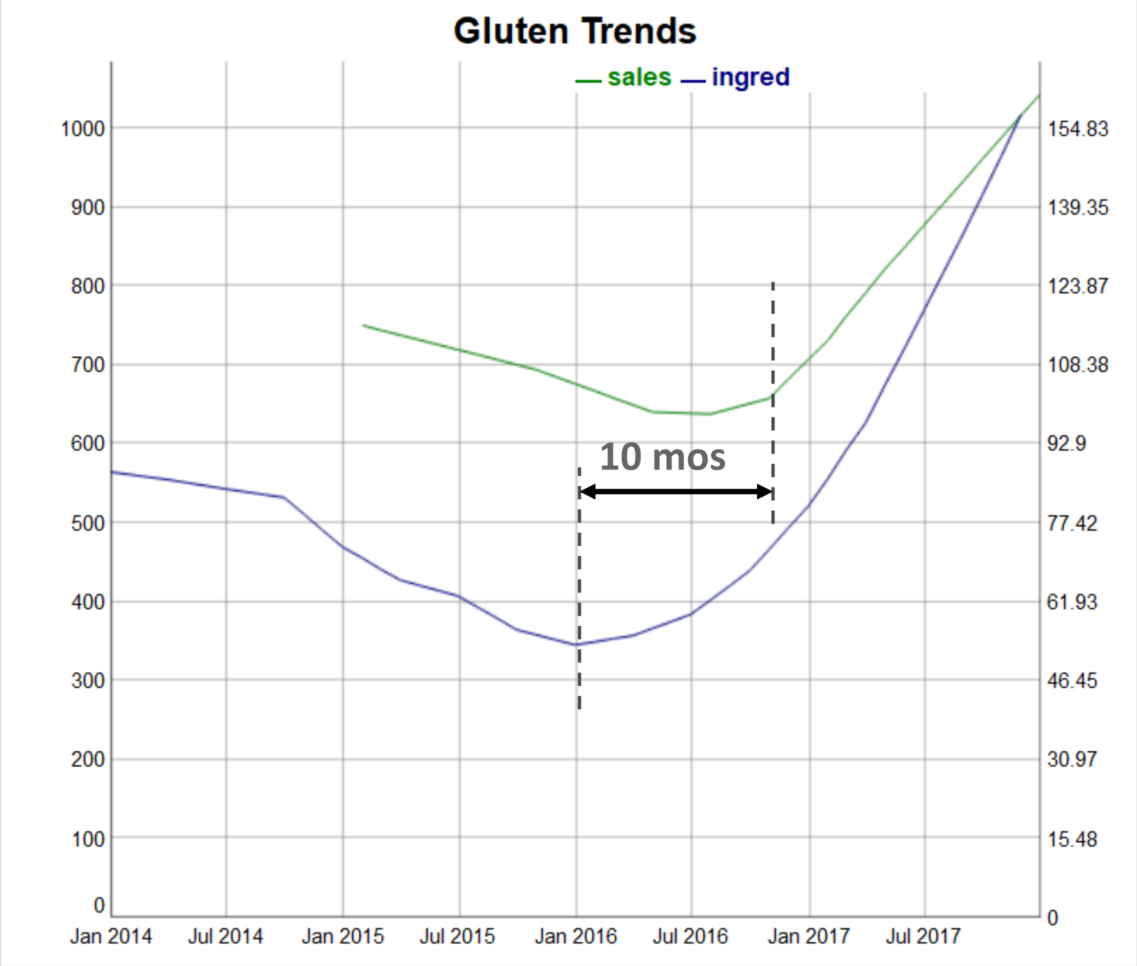


The motive (recognize **ingred**ients) predicts an acceleration in sales of gluten-free foods with recognizable ingredients

Converseon SM trained by motive



Smoothed time series



Designing a bigger experiment

What Did We Learn?

It takes a lot of unique posts to make this work. (>1MM)

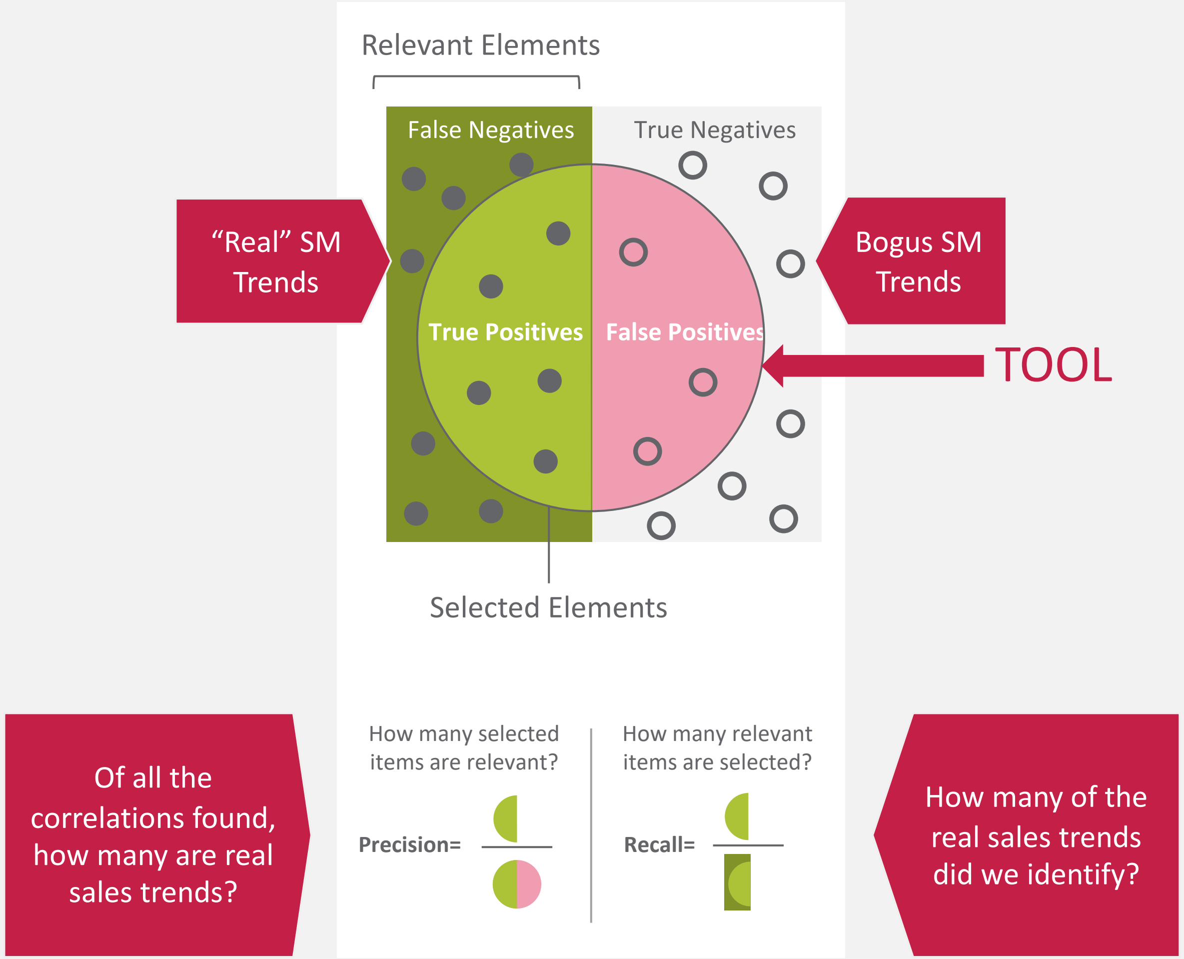
The stronger the filter, the better the result.

Training the machine-learning classifier on motives DOES have strong predictive qualities.

Can predict inflection in sales trends at least 90 days before they happen.

There is a feedback loop: in the beginning SM predicts sales but after a while the sales activity itself may also contribute to the SM trend.

A Better Test—currently Underway



Key Takeaways

There is a **big role** for inclusion of SM time series as predictors in models:
- Marketing Mix/ Business Drivers
- New Products

Modeling leads to a **forward-looking** rather than backward-looking use of SM data.

It takes a relatively **high volume of posts**, **good cleaning procedures,** and **NLP expertise** to make the time series useful.

A combination of **survey data & behavioral data** is better than either one alone.

Questions?

Mark Garratt
mgarratt@in4ins.com

Stuart Schwartz
sschwartz@in4ins.com

www.in4ins.com