



# Seaweed Snack ‘Bugak’ overseas market trend research & market positioning

# Table of content

---

## 1. Topic

- Goal
- Background

## 2. Trend research

- Time series analysis
- Instagram tag trend

## 3. Data analysis

- Consumer reaction
  - Sentiment analysis
  - Keyword analysis

## 4. Business strategy

# Topic



과학기술정보통신부  
Ministry of Science and ICT



경기대학교



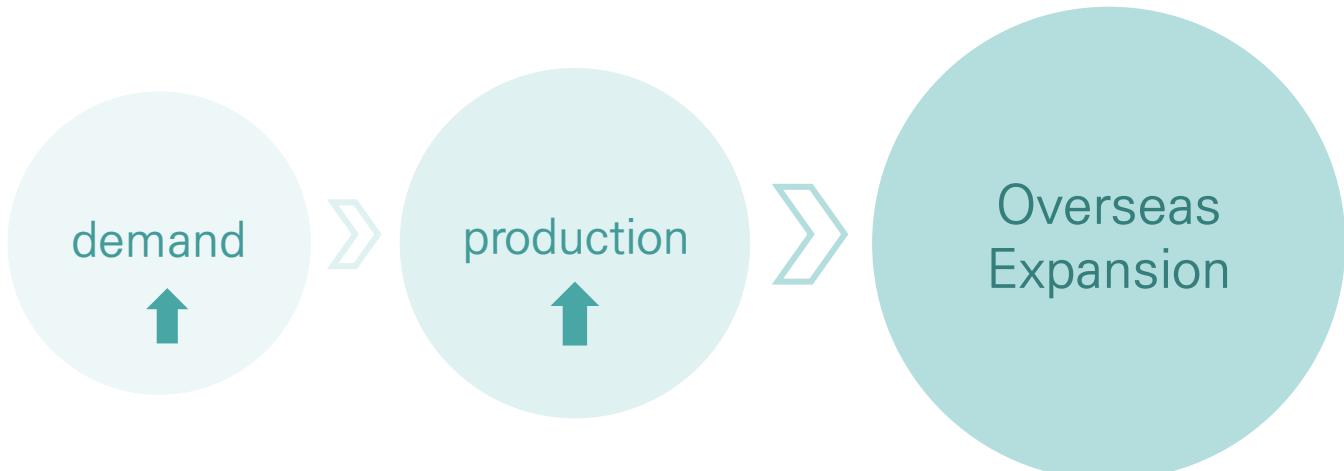
# Goal



- Seaworld corp.
- Food processing corporation
- Mixed crisps
- Wish to expand crisp business overseas

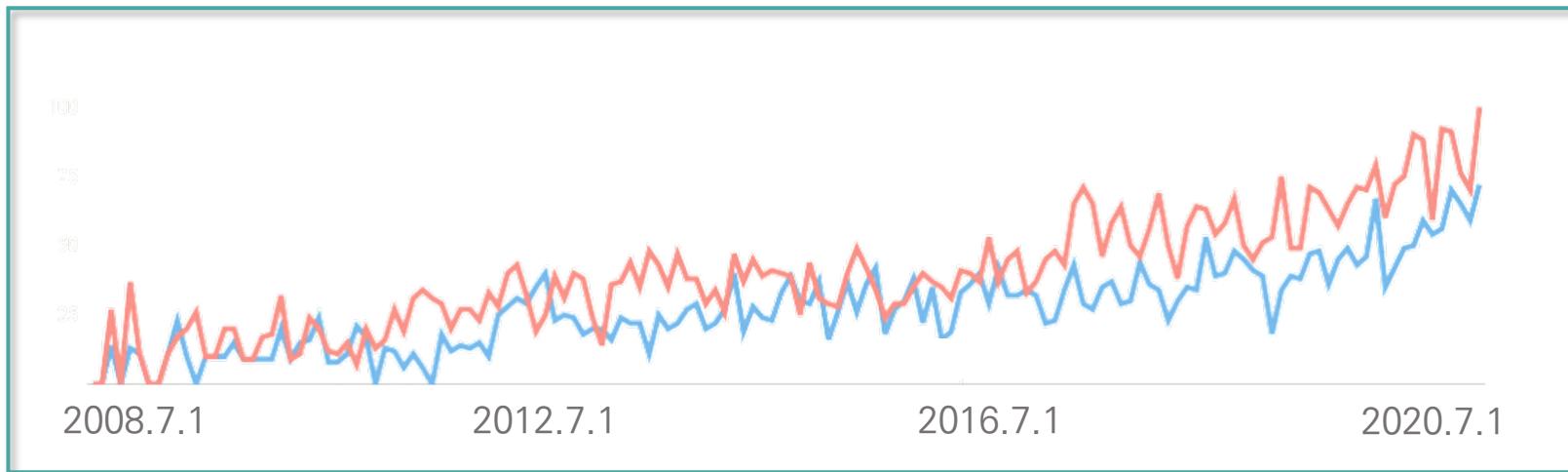


- What is 'Bugak'?
  - Fried seaweed coated with glutinous rice paste
  - Korean dish: vegetarian crisps



## Background

Google trend: ‘Bugak’ related terms search volume



● seaweed snack

● seaweed chips



Growing interest in seaweeds

# Data collection

## Instagram

period 2012.3.23 ~ 2020.8.14

Meta data Date, account, content, comment, # of likes

Raw data 21,313

## Amazon

Meta Data date, reviewer, rate, title, review, product

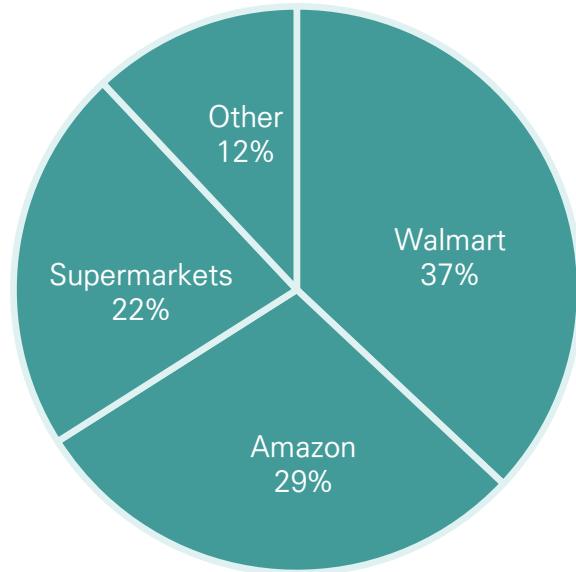
Raw data 4,410

## Walmart

Meta data date, reviewer, rate, title, review, product

raw data 7,810

## Most Recent Digital Grocery Store or Service Used by US Digital Shopper, Sep 2019



- Amazon and Walmart take up to 66% shares of online grocery store

→ Instagram data for Trend research,  
Amazon and Walmart product reviews for actual  
consumer reaction analysis

# Trend research



과학기술정보통신부  
Ministry of Science and ICT



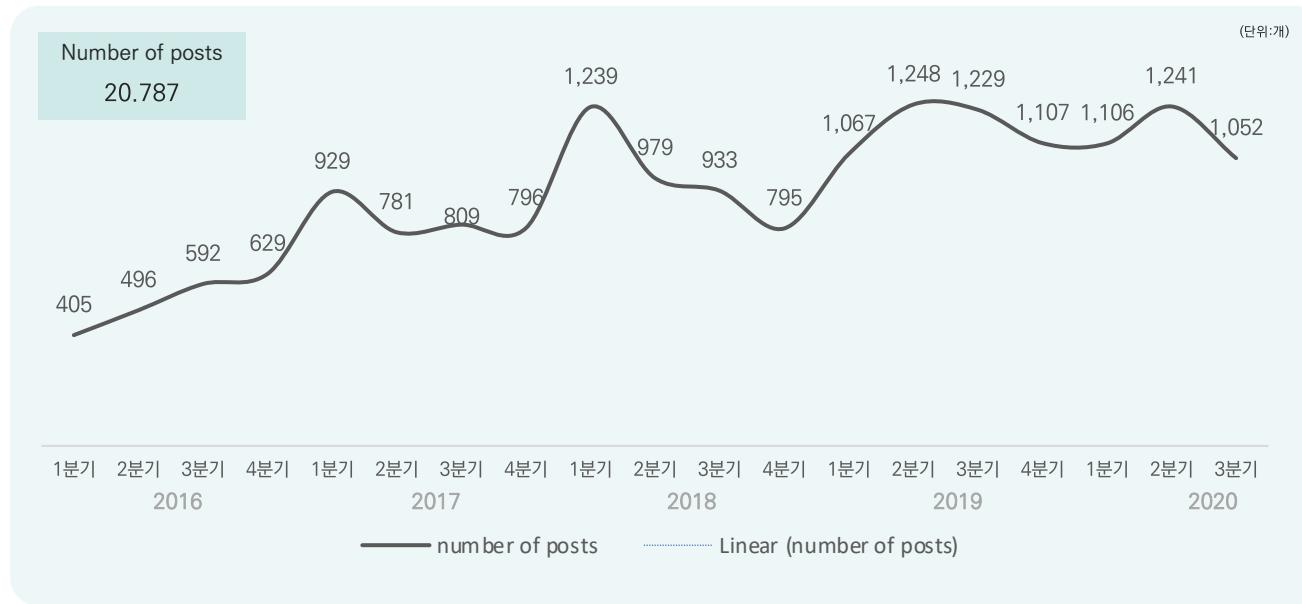
경기대학교



K-data  
한국데이터산업진흥원

# Trend research

## Instagram posting trend



### Collected keywords

#seaweedsnack  
#seaweedchip  
#seaweedcrisps  
#bugak

‘Bugak’ related terms  
search volume



Foreigners’ interest



# Trend research

## Past three years Instagram tag frequency

2018

2018.01~2018.12

no.	Tag	Freq
1	seaweedsnack	824
2	seaweed	654
3	seaweedchips	346
4	foodie	189
5	snack	170
6	foodporn	161
7	<b>vegan</b>	<b>149</b>
8	seaweedcrisps	131
9	food	127
10	viral	119
11	moodrepublik	118
12	bazaar	118
13	bazaaronlinemalaysia	118
14	malaysiaig	118
15	malaysiainstastop	118
16	seaweedpopia	116
17	<b>healthysnack</b>	<b>113</b>
18	<b>healthy</b>	<b>110</b>
19	foodlover	110
20	foodblogger	109

2019

2019.01~2019.12

no.	Tag	Freq
1	seaweedsnack	987
2	seaweed	706
3	<b>vegan</b>	<b>448</b>
4	plantbased	364
5	seaweedchips	361
6	<b>healthysnack</b>	<b>359</b>
7	<b>healthysnacks</b>	<b>283</b>
8	snack	276
9	foodie	267
10	<b>healthylifestyle</b>	<b>242</b>
11	<b>veganlife</b>	<b>239</b>
12	<b>glutenfree</b>	<b>232</b>
13	nongmo	228
14	seaweedcrisps	220
15	<b>healthy</b>	<b>218</b>
16	<b>healthyfood</b>	<b>209</b>
17	foodstagram	182
18	seaweedsnacks	167
19	amazon	165
20	seaveggies	162

2020

2020.01~2020.07

no.	Tag	Freq
1	seaweedsnack	597
2	seaweed	567
3	seaweedchips	445
4	seaweedsnacks	304
5	<b>vegan</b>	<b>141</b>
6	nori	105
7	<b>rumputlaut</b>	<b>103</b>
8	seaweedcrisps	99
9	foodie	86
10	koreanfood	83
11	plantbased	82
12	snack	80
13	seaweedsalad	70
14	snackrumputlaut	65
15	snacks	62
16	seaweedlover	60
17	chips	59
18	<b>glutenfree</b>	<b>58</b>
19	ohmygim	57
20	<b>healthyfood</b>	<b>56</b>

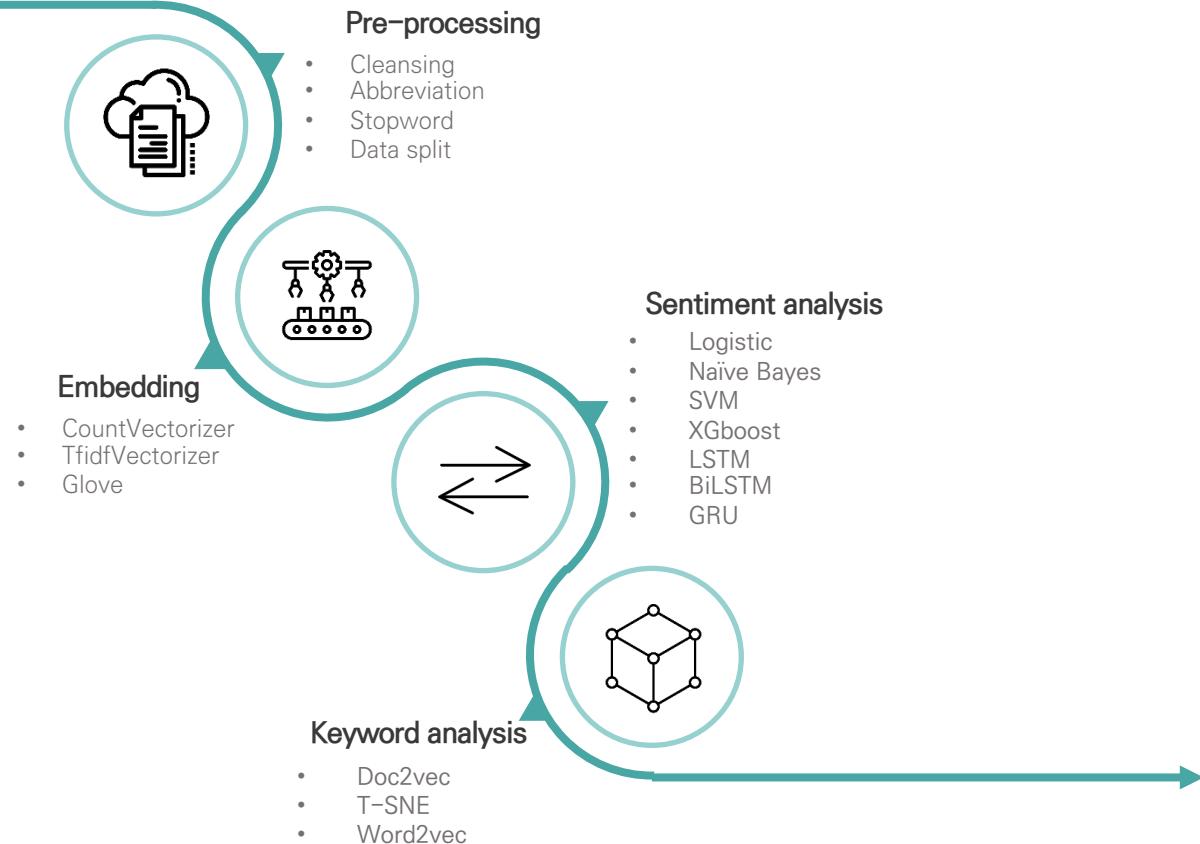
Vegan 

Healthy eating 

# Data analysis

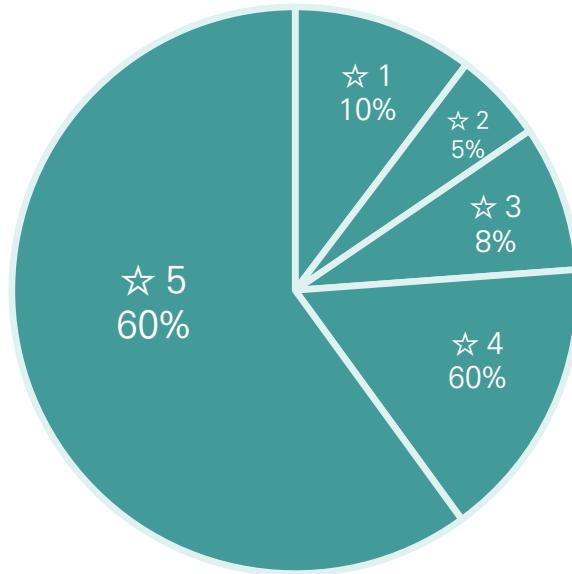


# Data analysis processing steps



# Data distribution

Distribution of 'Seaweed snack' keyword ratings



Rating	count
1.0	596
2.0	299
3.0	482
4.0	934
5.0	3467

Number of 5 stars is significantly higher



Data imbalance handling needed

## Potential causes of imbalanced data:

Seaweed snack is not well known to the international food market.

Therefore, most of the consumers might have already known about the product before purchase or could be re-purchasers.

# Data Search

## Pos/Neg Word Cloud

- 1~2 stars : Neg & 3~5 stars : Pos



### Overview of Positive reviews



Yummy! my favourite snack. Salty and crunchy!



My vegetarian daughter loved them!



### Overview of Negative reviews



Didn't like the fishy taste and texture..



Extremely fishy tasting and could not taste the sesame at all

# Data preprocessing

## Data Cleaning

### 1. Text trimming

1.0 out of 5 stars → 1.0

Dmitry  
★★★★★ This is very dangerous product ! Do not buy !  
Reviewed in the United States on May 4, 2019  
Flavor: Nutty flavor | Size: 0.17 Ounce (Pack of 24) | Verified Purchase  
I've been buying this product for a year, I bought it many times and I would definitely give it 5 stars, but... today when I was enjoying it something very sharp stuck deep inside of my tongue.  
  
It was piece of very thin still wire of about 0.4 inch of length. It was definitely inside of the seaweed sheet. At first I thought that it is a fish bone in it it was wire. I'm scared to death now. I was not expecting this to happen. This thing is very very dangerous. I hardly can imagine what could have happened if I swallowed it. Also I was used to give it to my kids...

### 2. Emoji, number, special character removal

Great flavor  Love it   → Great flavor Love it

### 3. Uppercase to lowercase

 Delicious and crispy/crunchy. → delicious and crispy/crunchy.

" This shouldn't even get one star! This the worst thing i've fishy!\n\nAnnie Chun's Seaweed Snacks, Roasted Wasabi, 0.35-Ounce! I tried to return these and they won't let you!!!!!! be fed to a seal!"

" This shouldn't even get one star! This the worst thing i've fishy!\n\nAnnie Chun's Seaweed Snacks, Roasted Wasabi, 0.35-Ounce! I tried to return these and they won't let you be fed to a seal!"

# Data preprocessing

## Data Normalization

### 1. abbreviation

in contraction format

Original	Contracted
will not	won't
is not	ain't
are not	aren't
did not	didn't
does not	doesn't
had not	hadn't
I would	I'd
should have	should've
what is	what's

### 2. Stopword removal

Remove words that may lead to wrong decision, wrong conclusion and poor analysis

예시
better
amazing
disappointed
yummy
loves
great
wonderful
terrible
gross

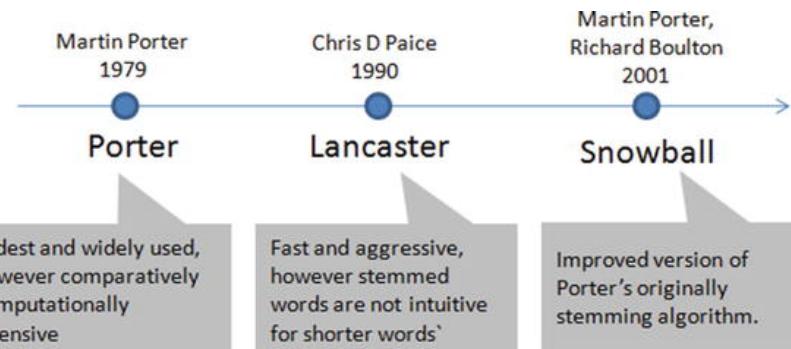
## Tokenization

### 1. stemming

- Snowball stemmer used
- Stemmed text added to data

### 2. Tokenization

Data tokenized



# Data preprocessing

## Data splitting

- X : Cleaned review / Y: 0 (neg), 1 (pos)
- After reviewing 3-star data → classified as positive
- Star 1, 2 Negative / Star 3, 4, 5 positive

### 3-star review example

-  Good taste
-  Wow Considering the popularity of this brand…
-  LO /E how delicious these are but honestly there are so little …
-  A little salty but I enjoyed it.

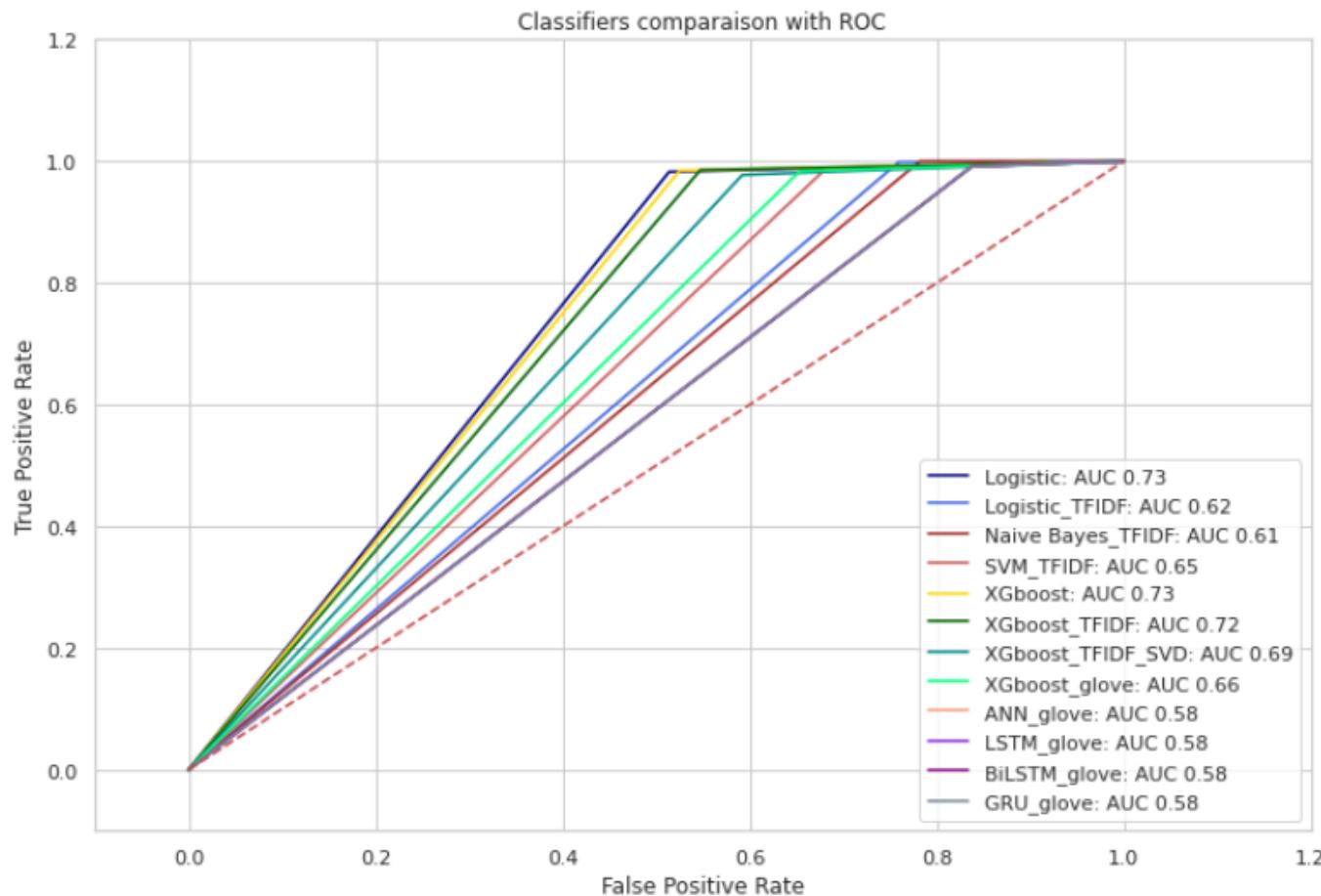
- Train, Test set (8:2)
  - stratify = Y used when splitting train and test sets to preserve the proportion of target as in original dataset
- Cross-validation
  - Stratified k-fold cross-validation used
  - split by 5, then shuffled

## Embedding

- CounterVectorizer
- Tf-idf Vectorizer

# Sentiment analysis: model performance comparison

## performance comparison



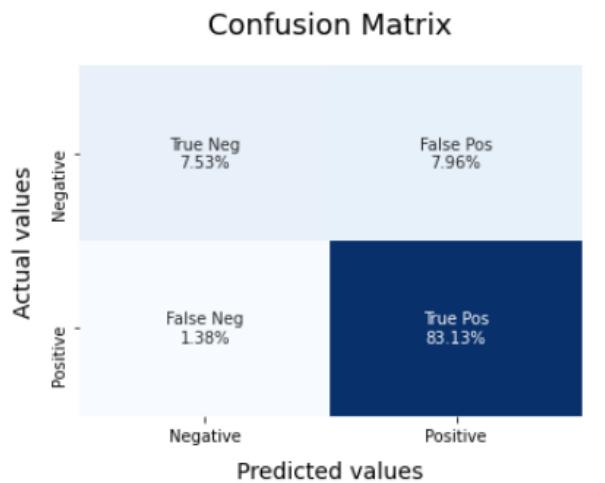
No significant difference between models

Logistics and XGBoost are top two models in ROC curve

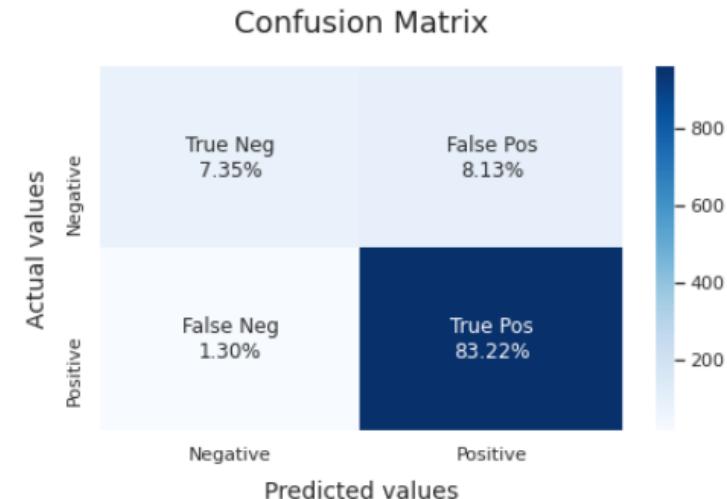
# Sentiment analysis: model performance comparison

## Logistics and XGBoost

### Logistic Regression



### XGBoost



	Precision	Recall	f1-score
Neg	0.84	0.49	0.62
Pos	0.91	0.98	0.95

Accuracy 90.66%

	Precision	Recall	f1-score
Neg	0.85	0.47	0.61
Pos	0.91	0.98	0.95

Accuracy 90.57%

# Sentiment analysis: model performance comparison

## Logistics and XGBoost

\*value: regression coefficient

Logistic Regression

delicious	1.583751
snack	1.207807
tasty	1.164487
favorite	1.131115
calories	1.027322
kids	1.005802
son	0.997377
addictive	0.971398
daughter	0.962564
flavors	0.961170
fast	0.952714
easy	0.941521



XGBoost

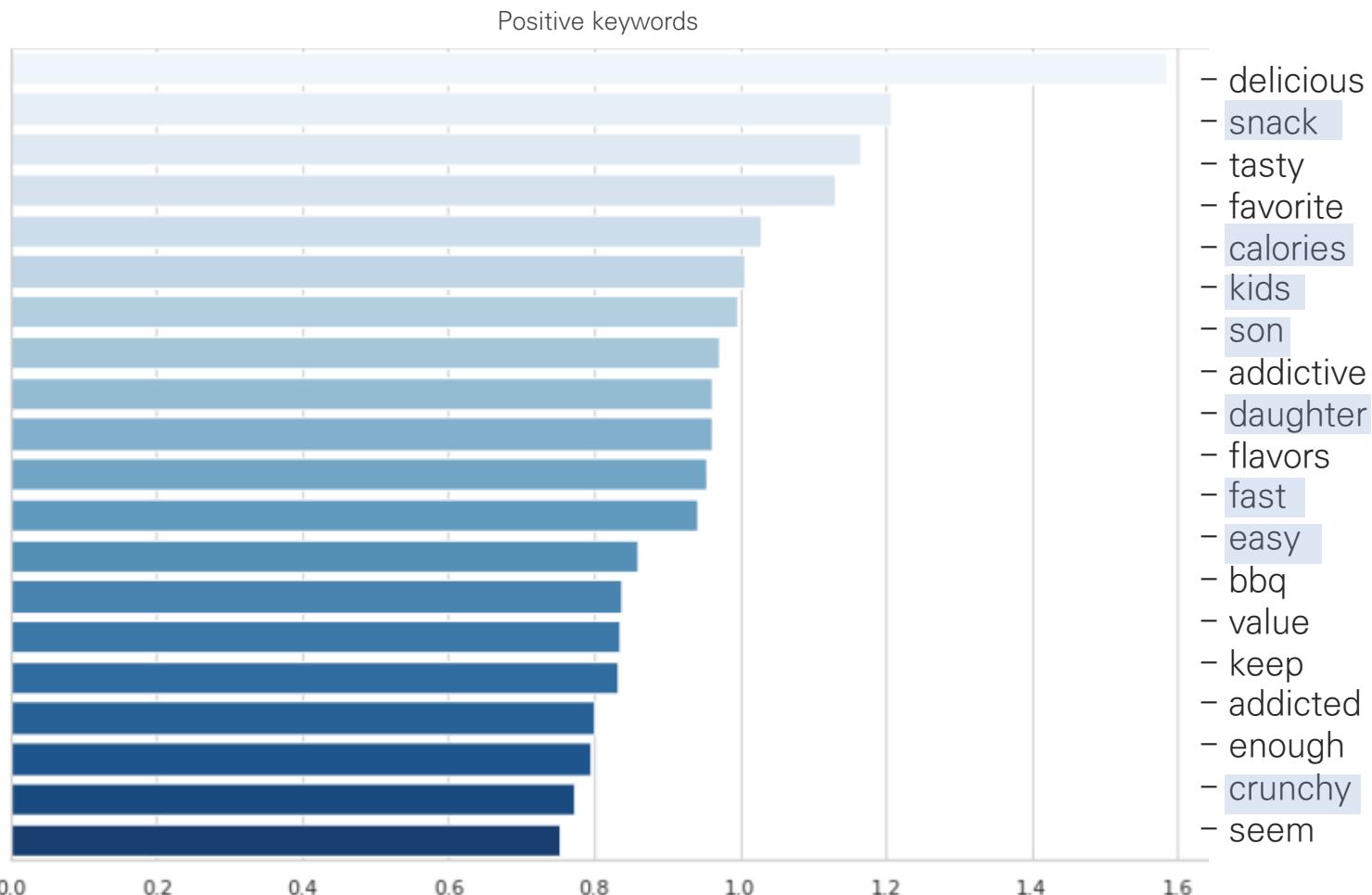
taste enjoyed	14.6049
salty wanted	14.0864
never tried	14.0551
people salt	13.5574
taste expensive	13.1973
flavor order	13.0977
flavor expensive	13.0537
addicting	12.4555
company much	12.4246
repurchase	12.1066
delicious	11.6648
daughter	11.5219

After reviewing results, logistics is the most ideal model for

- 
- classifying positive and negative keywords
  - \*most importantly, analyzing consumers' evaluation (compliment, complain)

# Sentiment analysis

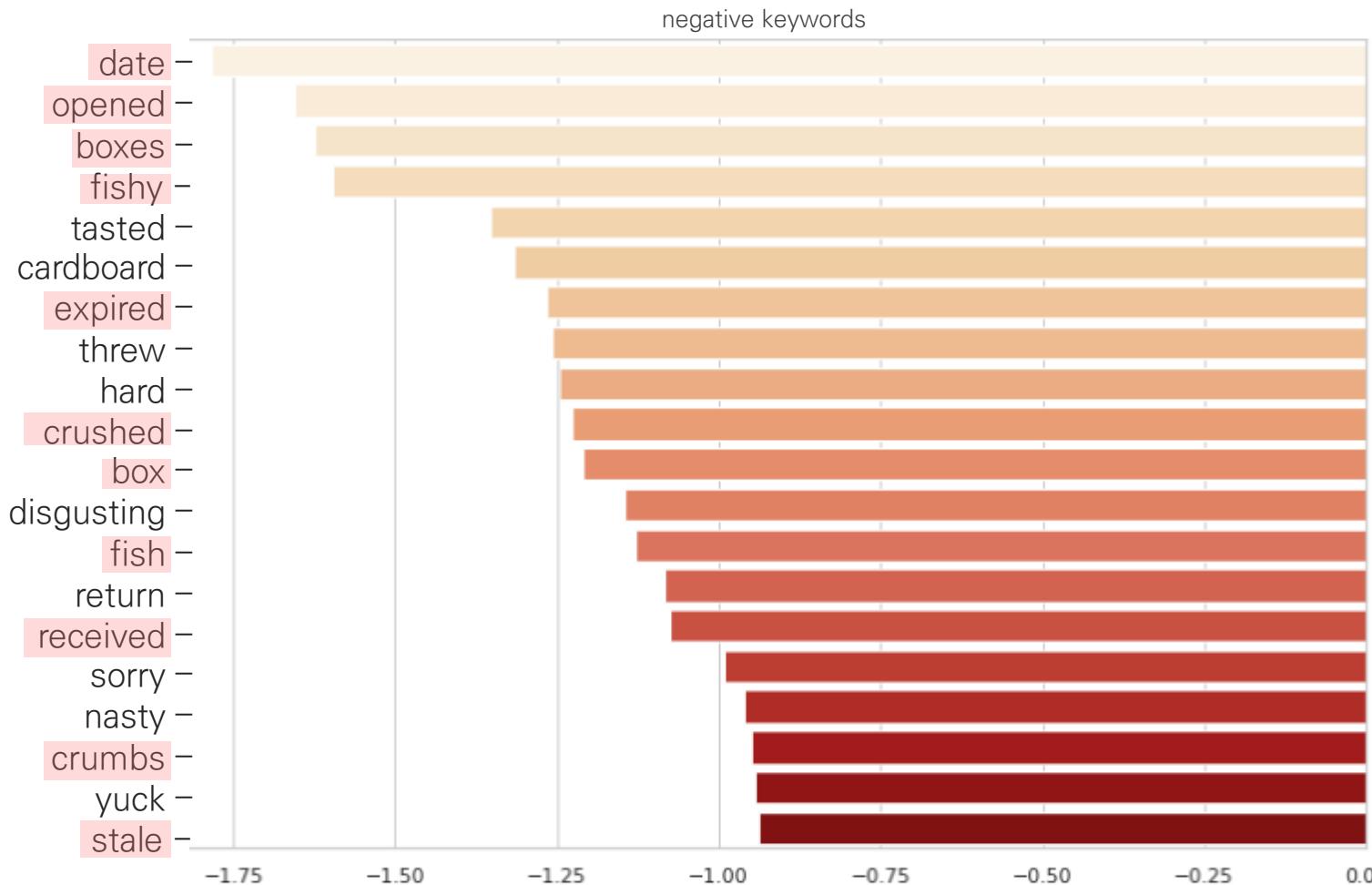
## sentiment analysis results: positive keywords



**Family, healthy, convenience, flavor**  
related keywords found in the positive set

# Sentiment analysis

## sentiment analysis results: negative keywords



**Freshness, smell, packaging**  
related keywords found in the negative set

# Keyword analysis method

---

## Keyword similarity

- Keyword overall distributed representation

### Doc2Vec

1. Map coordinates to a document in Doc2Vec
2. Treat each document id as one paragraph token
3. keywords appeared in the same document are located closer to each other – vector formation

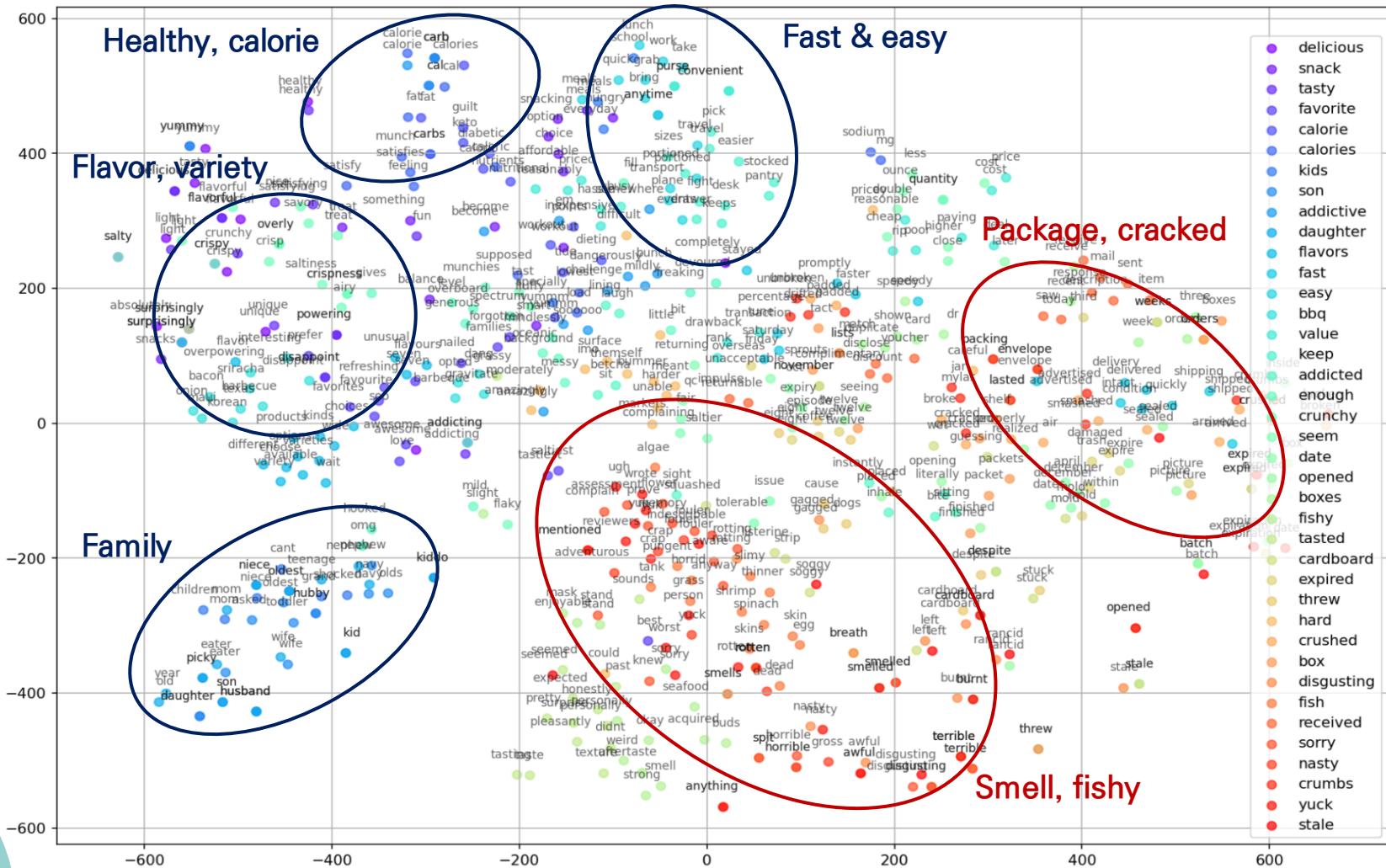
### T-SNE plot

1. For visualization purposes, converted high dimensional data into two dimensions
  2. explore actual distribution of classified keywords
- Individual keyword analysis

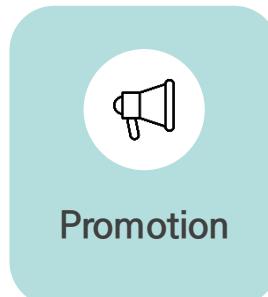
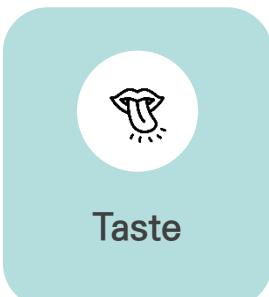
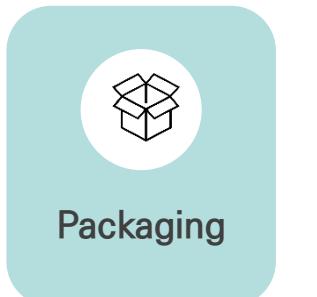
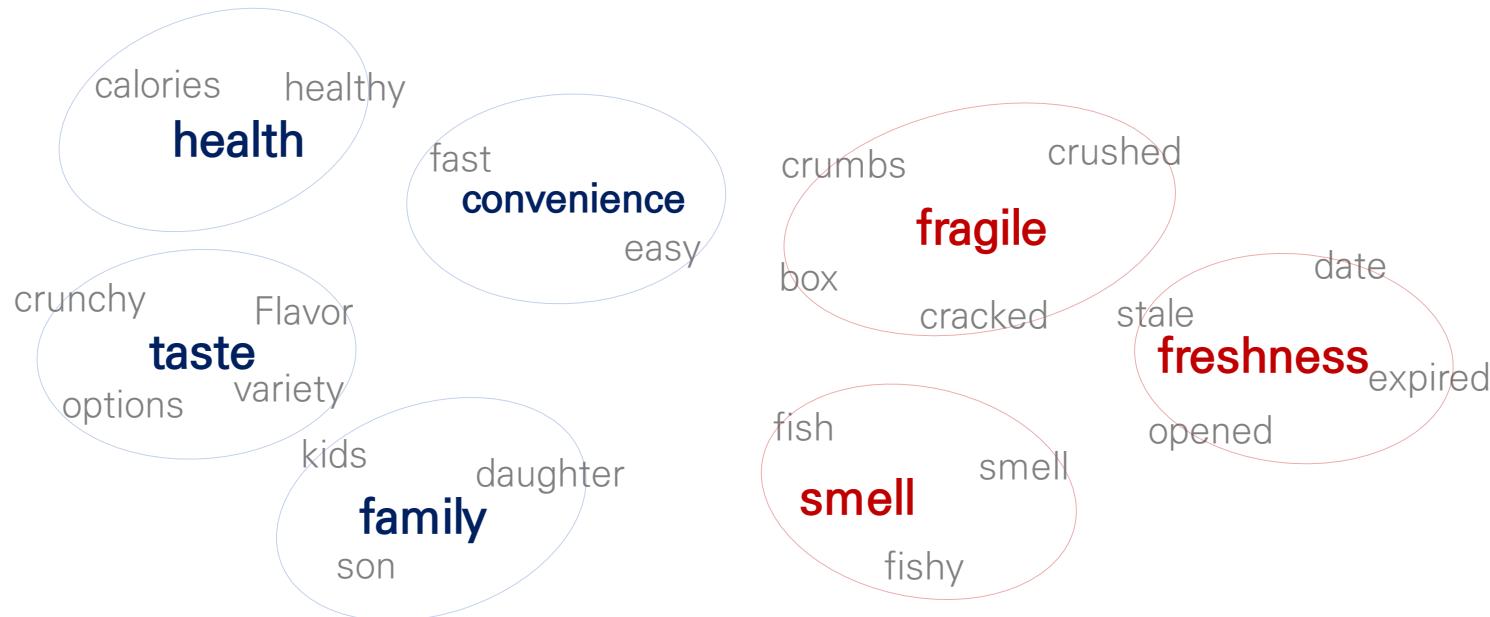
### Word2Vec

1. distributed representations of context words (target word surrounding words) are combined
2. CBOW model used

# Keyword analysis

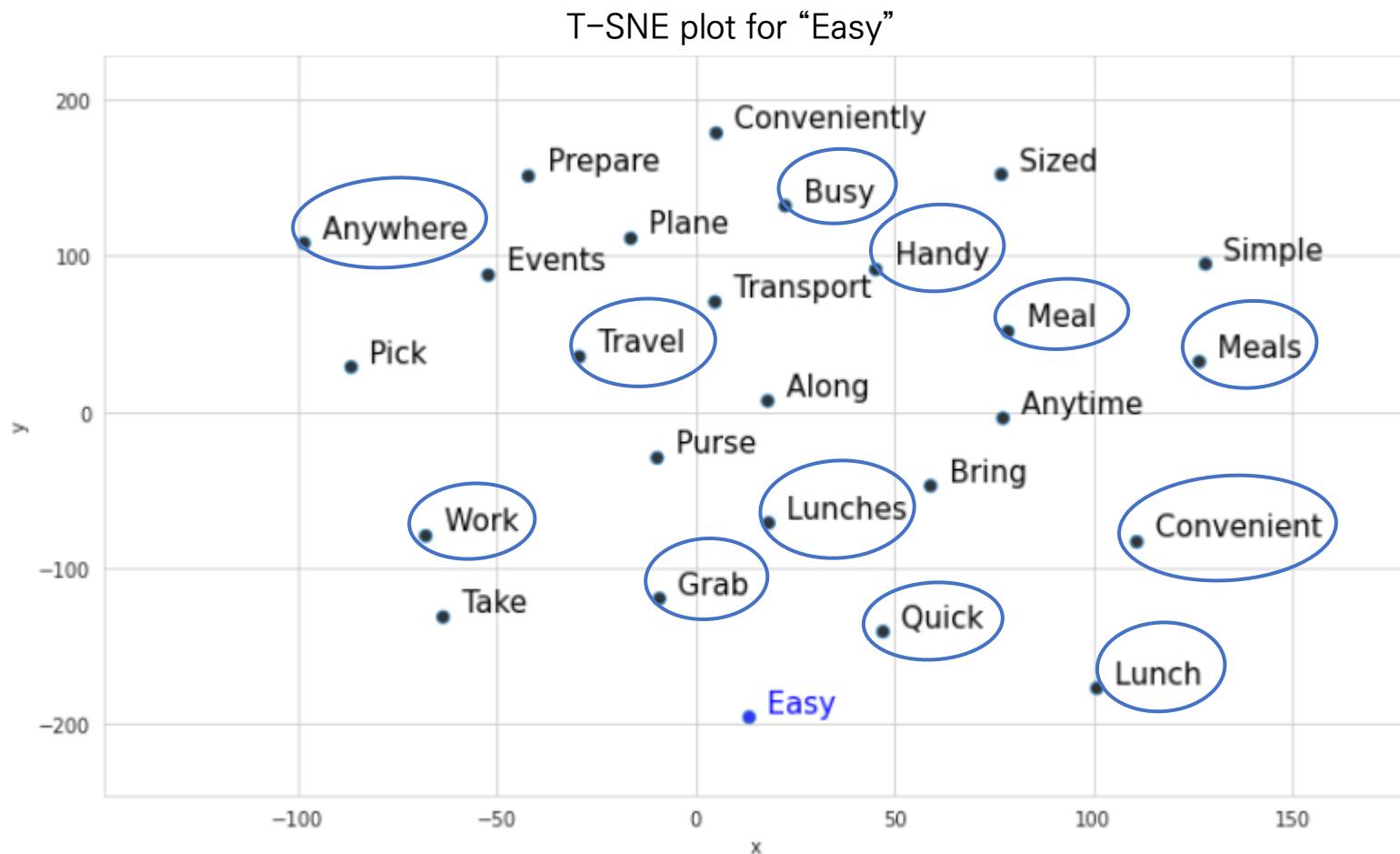


# Keyword analysis



# Keyword analysis – Packaging

## Convenience

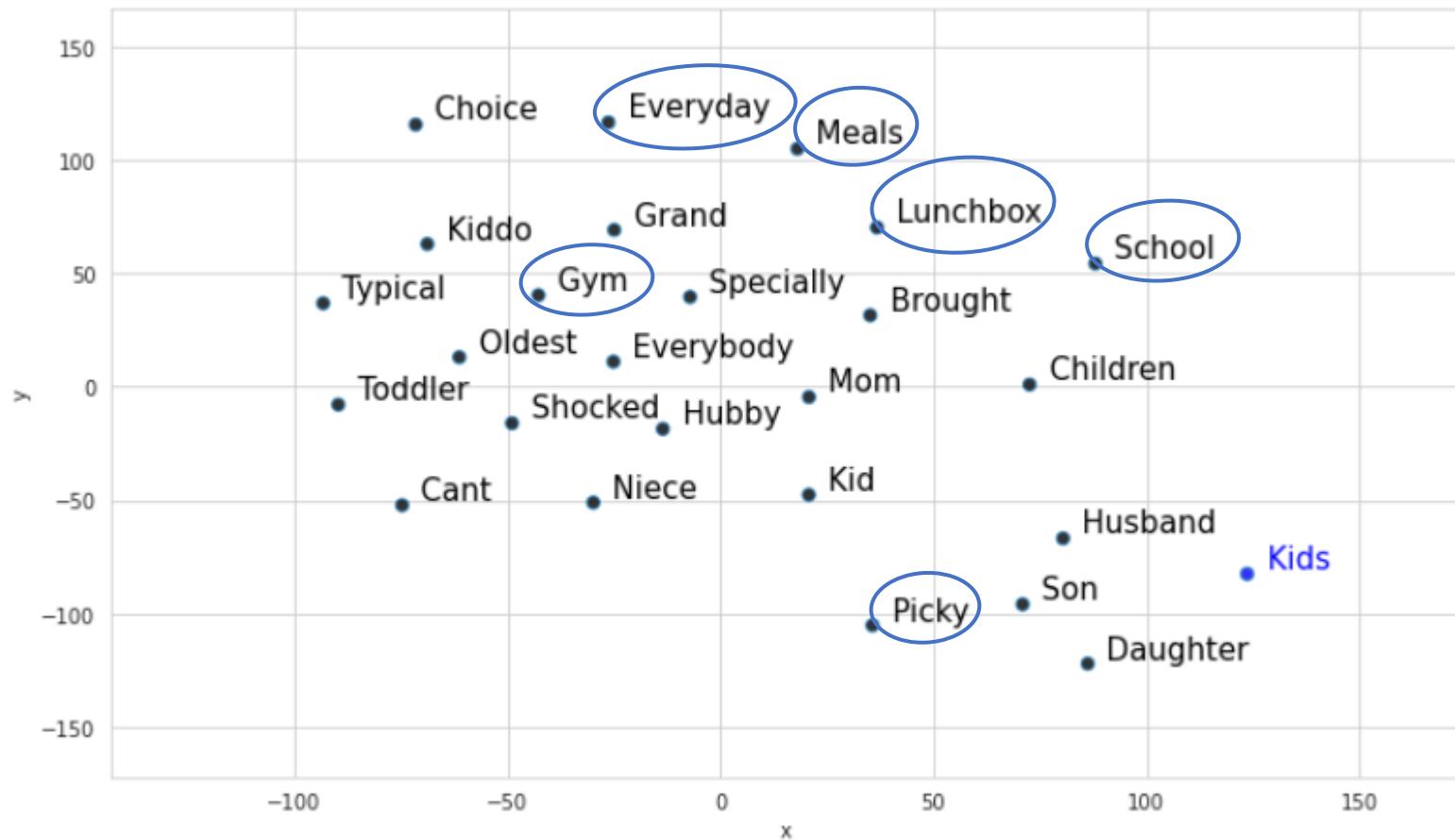


Convenient meal and quick Ready to Eat Foods for outdoor activities,  
workplace lunches during a busy day

# Keyword analysis – Packaging

## Family

T-SNE plot for “Kids”

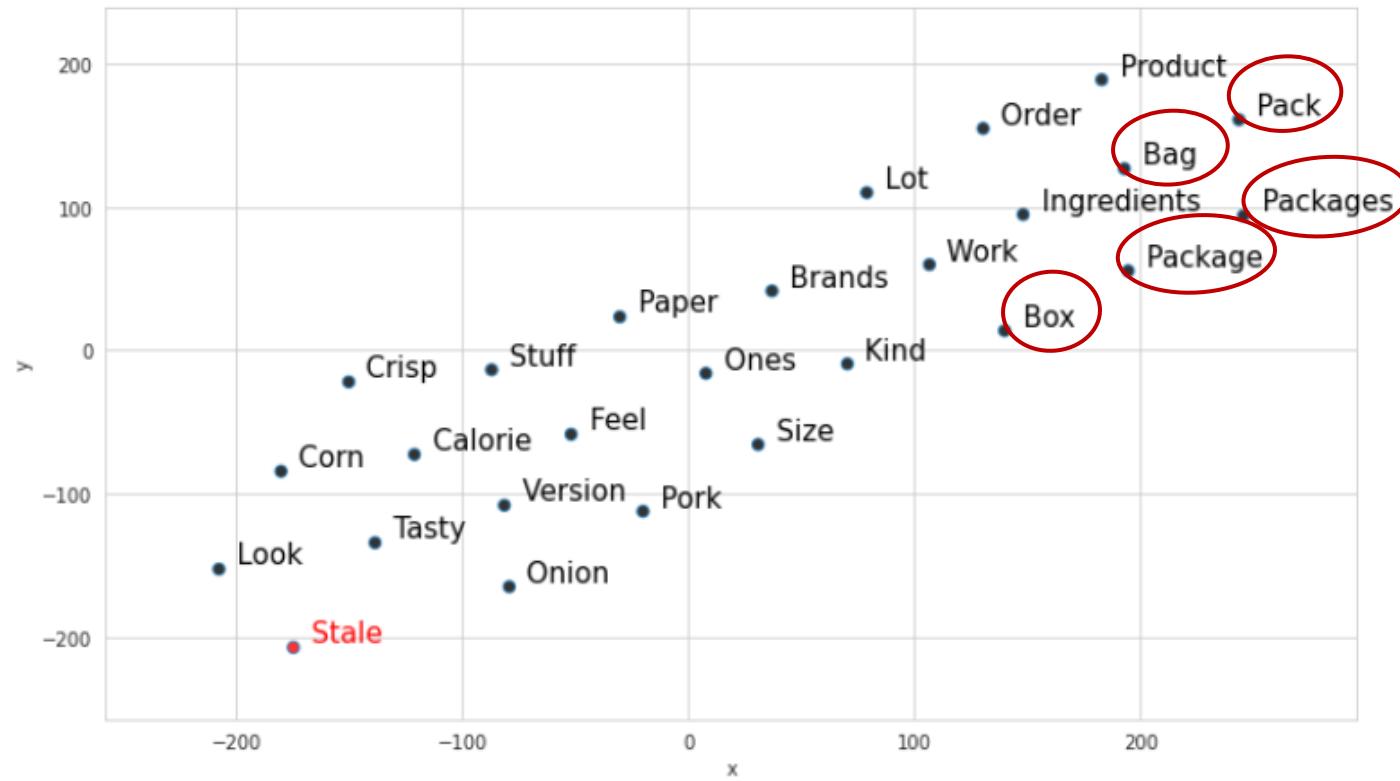


Family snack especially great for [lunch box snack](#) for kid

# Keyword analysis – Packaging

## Freshness

T-SNE plot for “Stale”



‘package’ found in keyword set of ”Stale”

(Package, Box, Pack, Bag)

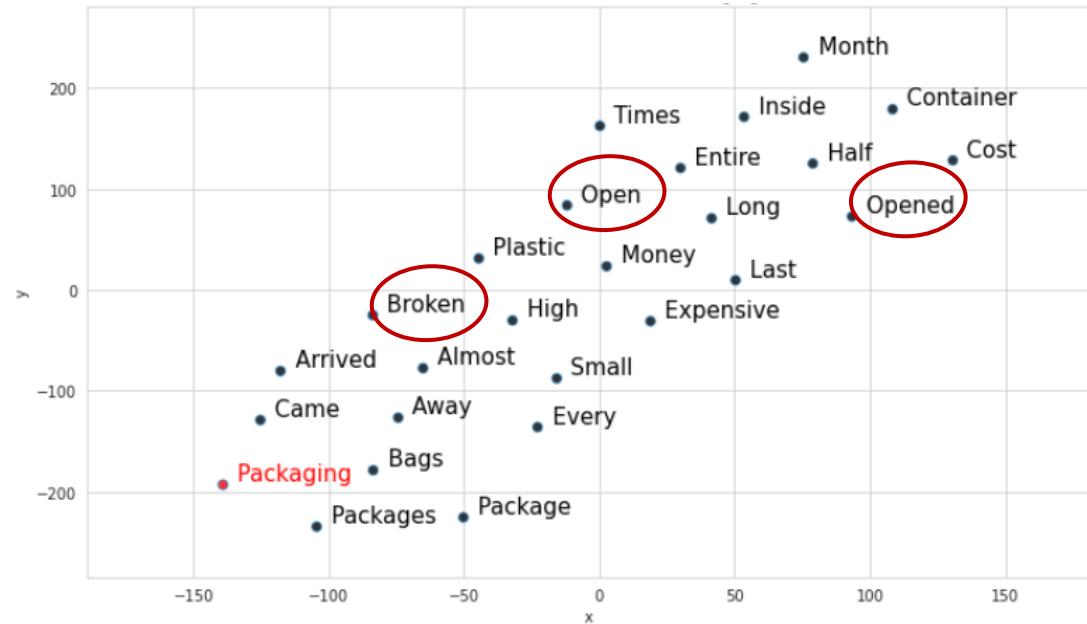
# Keyword analysis – Packaging

## Fragile

Trigram for “Opened”

- already opened package
- packaged opened stale
- bags opened broken
- packing opened inner
- mistake opened box
- additionally opened weird

T-SNE plot for “Packaging”



Common complains: Product found broken and opened

# Keyword analysis – Taste

## Taste

word2vec of “Smell”

<b>rotten</b>	<b>0.82246</b>
<b>fishy</b>	<b>0.81818</b>
<b>aftertaste</b>	<b>0.81075</b>
<b>skins</b>	<b>0.80918</b>
<b>smells</b>	<b>0.80685</b>
<b>fish</b>	<b>0.79640</b>
<b>slimy</b>	<b>0.79327</b>
<b>smelled</b>	<b>0.77978</b>
<b>dead</b>	<b>0.77903</b>
<b>strong</b>	<b>0.77052</b>



### Overview of Negative reviews

Terrible flavor and aftertaste, **very fishy!** Tried a couple and that was it. Now I have 19 packages of disgusting product. Don't believe I can return these.



So sorry I ordered this product. It is absolutely nasty with a **bad fishy aftertaste** that will stay with you for hours. Would not recommend this product to anyone especially for some child to take to school as a snack – OMG!



### Overview of Positive reviews

Tasty teriyaki flavor, and I **hate fishy smells and tastes**, so this is wonderful!

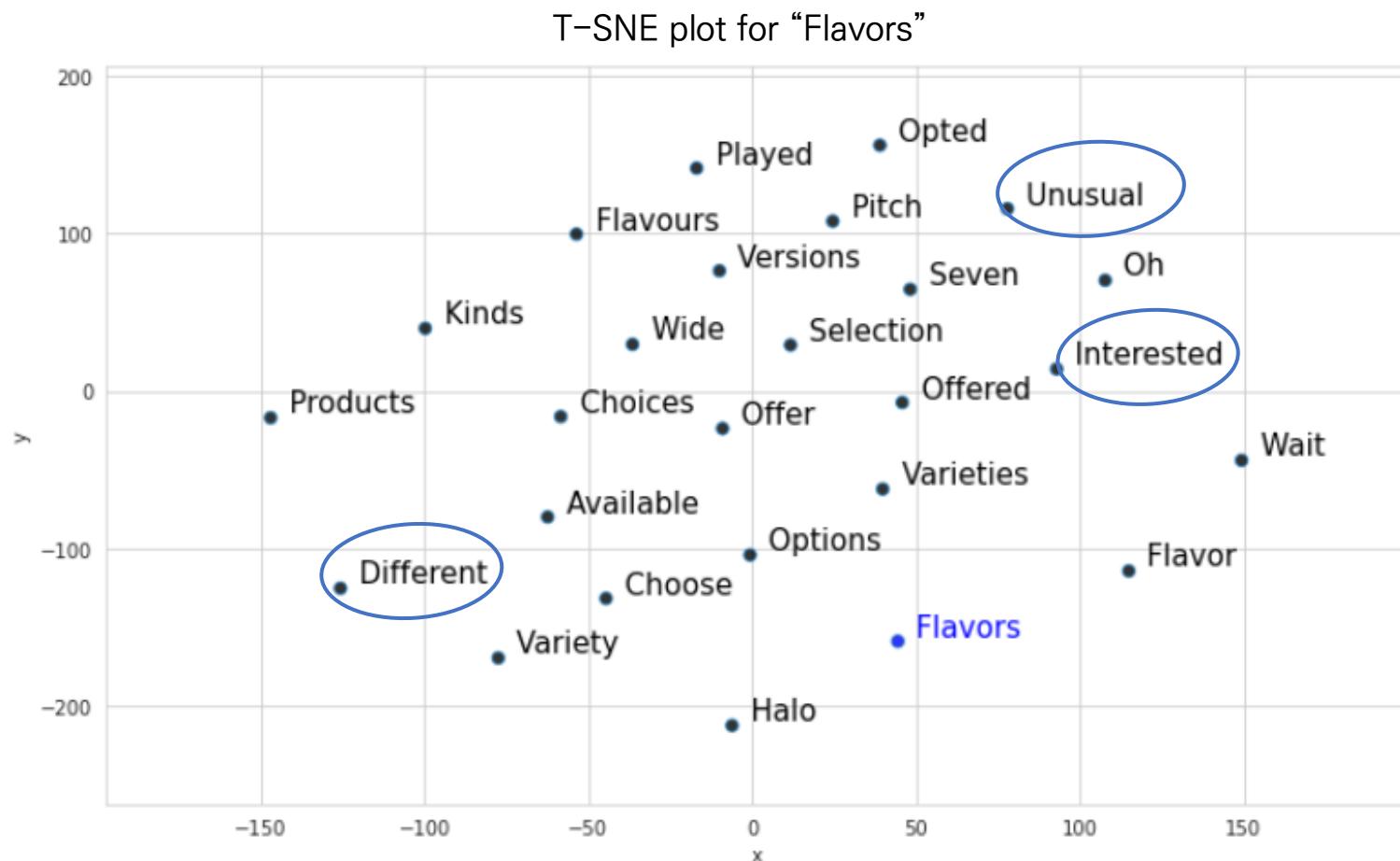


Nice smooth tasting seaweed...no fishy taste...!

Aversion to foods with **seafood smells**

# Keyword analysis – Taste

## Taste



unusual, different, interested etc. keywords found  
Many consumers **trying** unfamiliar seaweed snack

## Keyword analysis – Taste

### Taste

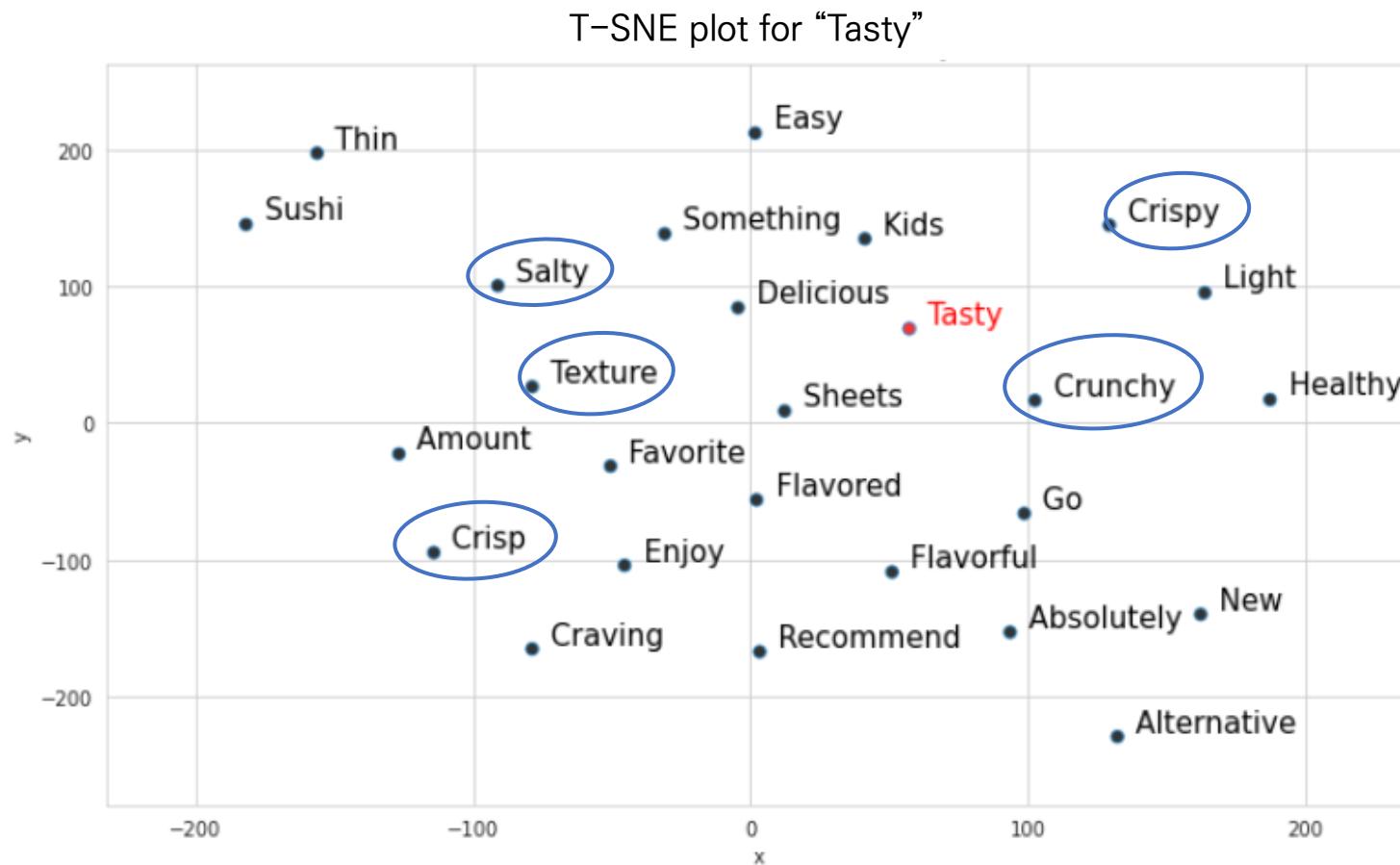
Positive                              Negative  
Flavors, delicious, tasty, tasted + Fish, smell, fishy  
= word2vec result

yummy	0.6494
snacking	0.5350
choices	0.5311
flavorful	0.5309
choose	0.5210
sooooooo	0.5198
option	0.5168
convenient	0.5011
options	0.5001
pick	0.4998

Consumers wants a variety of flavors that can cover fishy  
(choices, choose, option, options)

# Keyword analysis – Taste

맛

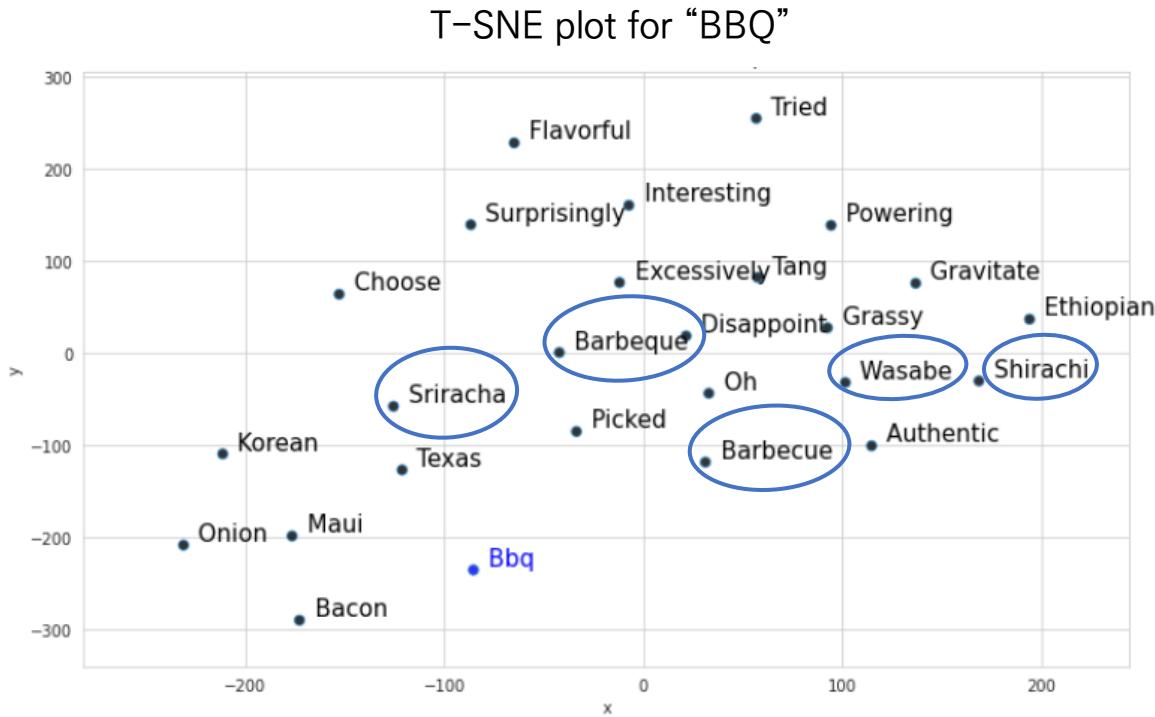


Crispy and Crunchy textures are important to consumers

Most preferred flavor: Salty

# Keyword analysis – Taste

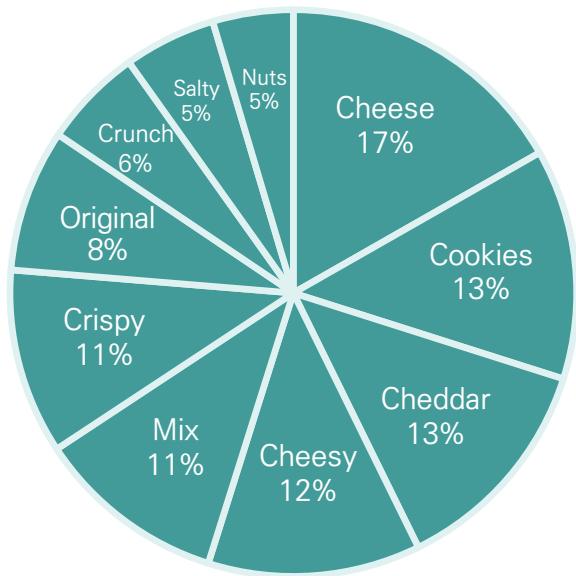
## Taste



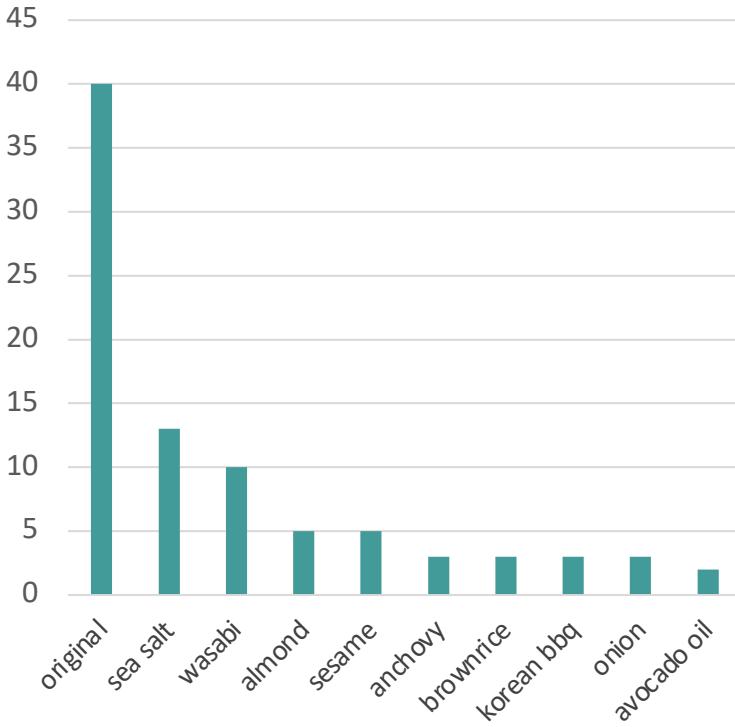
Spicy flavors (Sriracha and wasabi)  
and Korean bbq flavor are popular

## Keyword analysis – Taste

Top 40 Best selling snack flavors



Seaweed snack flavor available in the market

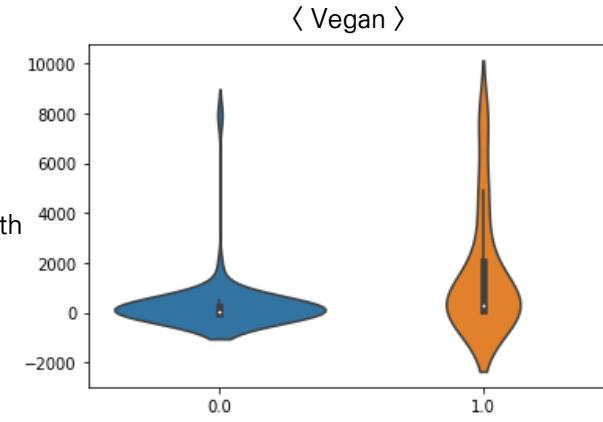
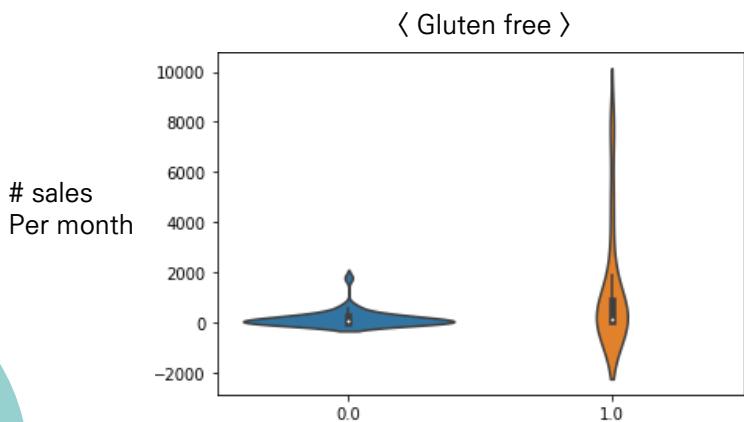
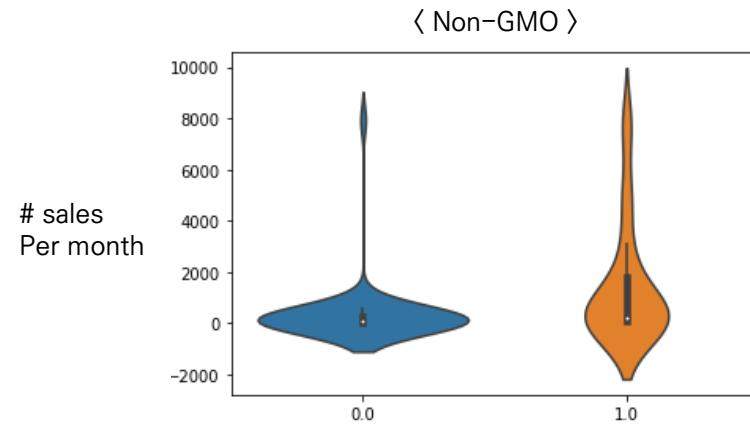
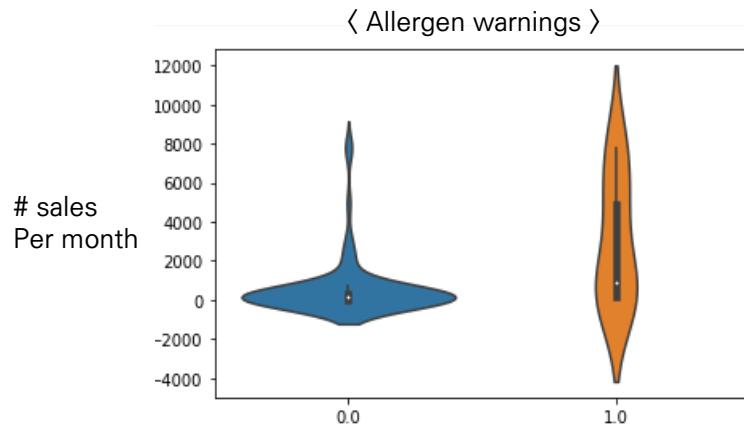


Cheese flavored snacks dominate best selling snacks but,  
No cheese flavor found in seaweed snacks

# Keyword analysis – Promotion

## Effect of Labeling on sales

- Violinplot
- 0 : not indicated / 1 : indicated



Allergen warnings, Non-GMO, Gluten free, vegan indicated product = Sales ↑

# Business strategy



과학기술정보통신부  
Ministry of Science and ICT

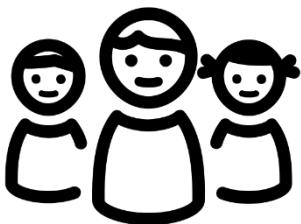
KGU 경기대학교

Kdata  
한국데이터산업진흥원

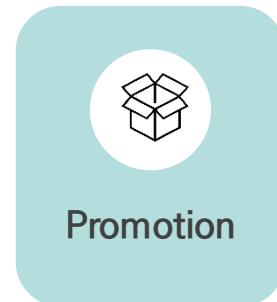


# Business strategy

## Target setting

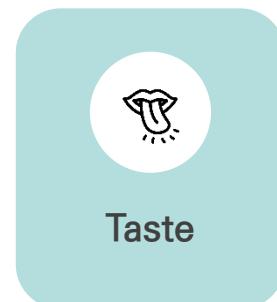


Kids/family



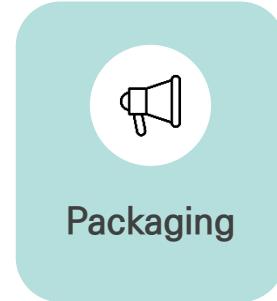
Promotion

- New term needed to describe seaweed snack: not to get confused with seaweed sheets ex) vege chips
- Healthy snack, made of Korea originated seaweed
- Target RTE (Ready to Eat) market



Taste

- Spicy and BBQ flavors to remove fishy smell of seaweed
- Add cheese flavor
- Add nuts to maximum crunchy texture

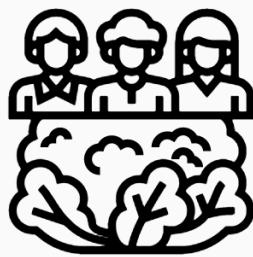


Packaging

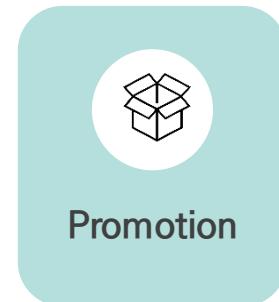
- Kid-friendly Individual packs: keep product fresh and easy to open

# Business strategy

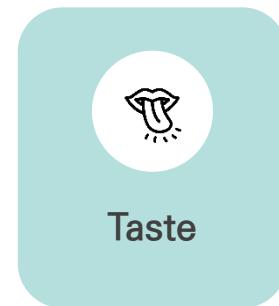
## Target setting



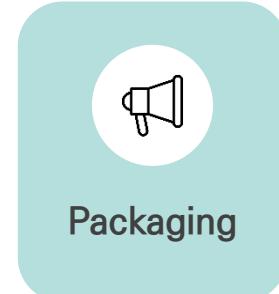
Vegan



- Indicate Non-MSG, Non-GMO, No artificial spices
- Emphasize 'Organic' plant-based protein source



- Enhance unseasoned original flavor



- Key global packaging trend: Resealable
- Eco-friendly packaging: No More Plastic



**THANK YOU**