Integrating Text Analytics and Statistical Analyses for Data-Driven Insights

Victoria Dickman-Burnett, PhD
Natalie Wozniak

Insights Association: Converge 2020
December 2, 2020

Agenda

01/
About us

02/
Why Text Analytics

03/
Our Approach

04/
Integrating TA Into Analysis

84.51°'s DATA OFFERS UNMATCHED CUSTOMER KNOWLEDGE BASED ON ACTUAL PURCHASE BEHAVIOR

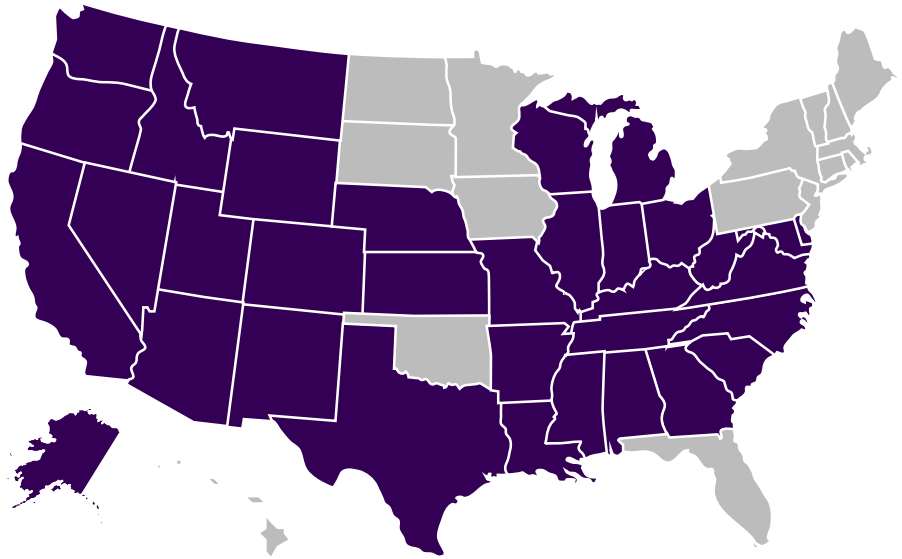
DATA FROM NEARLY

60MM
Households

1 out of 2 households in the U.S.
over 95% of transactions captured

ACROSS 2,800 STORES IN 35 STATES

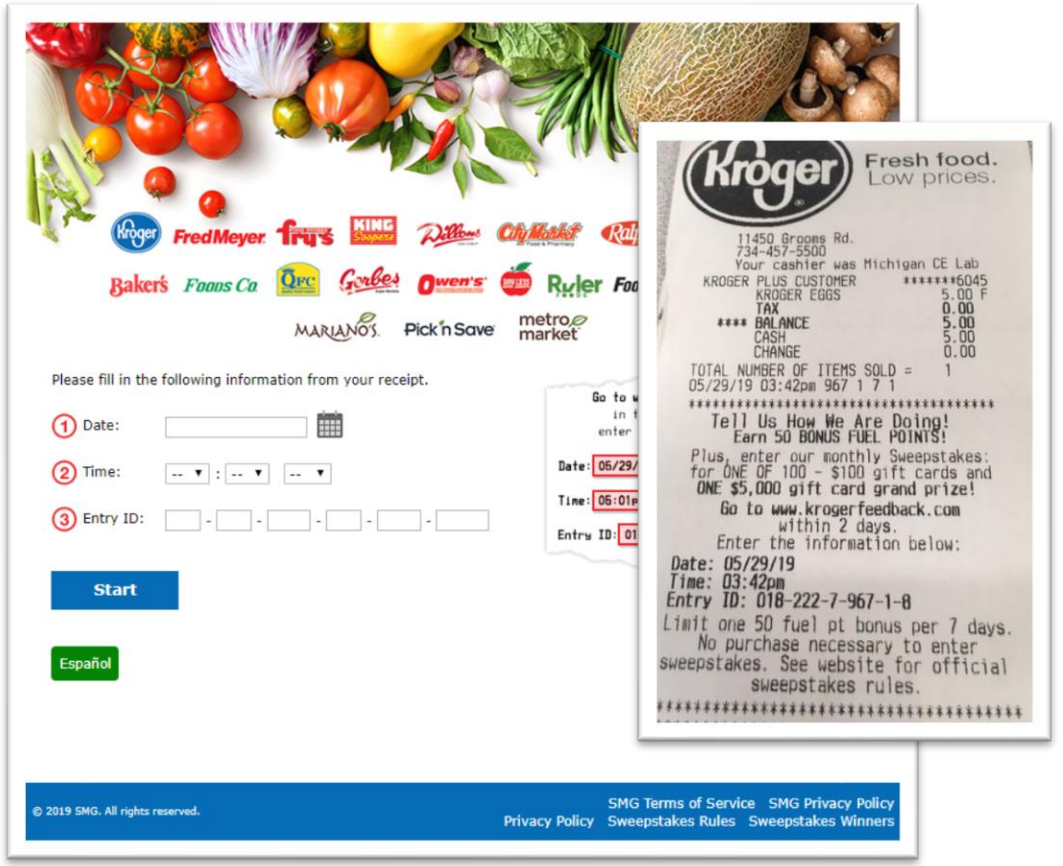
Kroger
FRESH FOR EVERYONE™
Baker's
City Market
Copps
Dillons
Food 4 Less
Foods Co
Fred Meyer
Fry's Food Stores
Gerbes
Harris Teeter
Jay C Food Stores
King Soopers
Mariano's
metro market
Pay Less
Owen's
Pick 'n Save
QFC
Roundy's
Ralphs
Ruler Foods
Smith's



About Our Team & Use Case

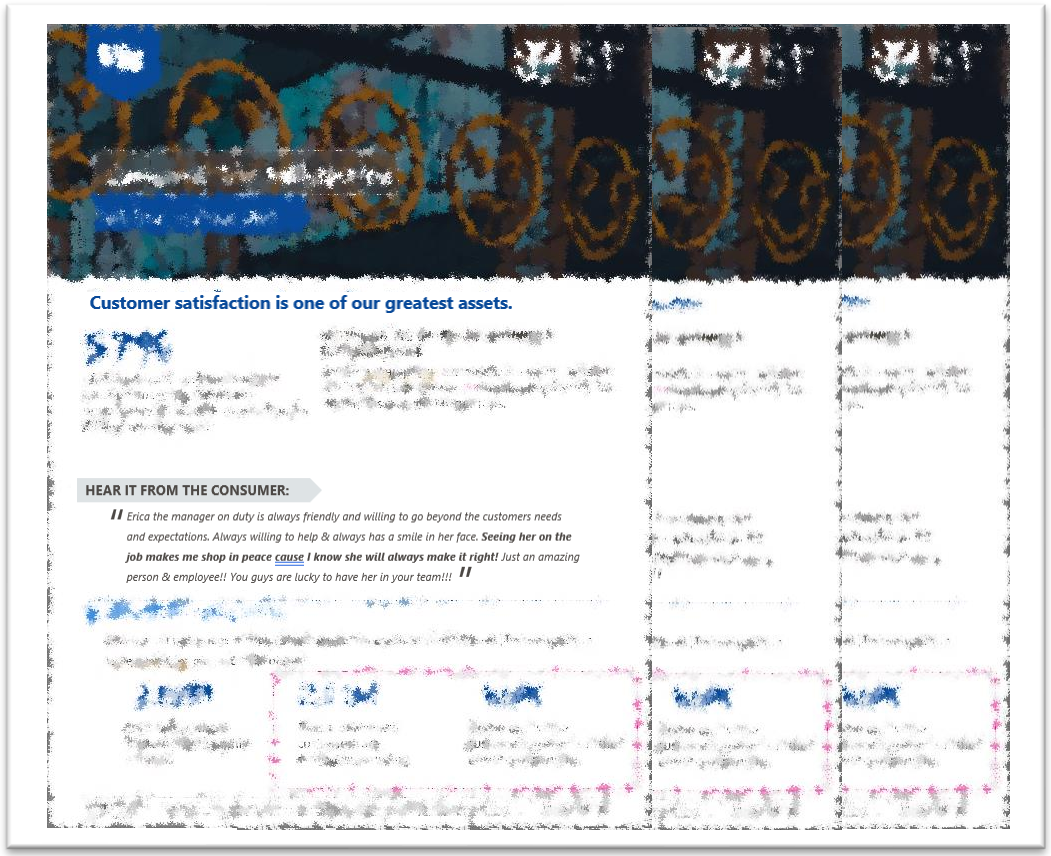
- The Foundational Research Pillar of Consumer Research at 84.51° is responsible for large, typically ongoing, research projects to support the business

- Customer Trip Satisfaction survey: 14M responses/year



How We Started

Enterprise Leadership Use:

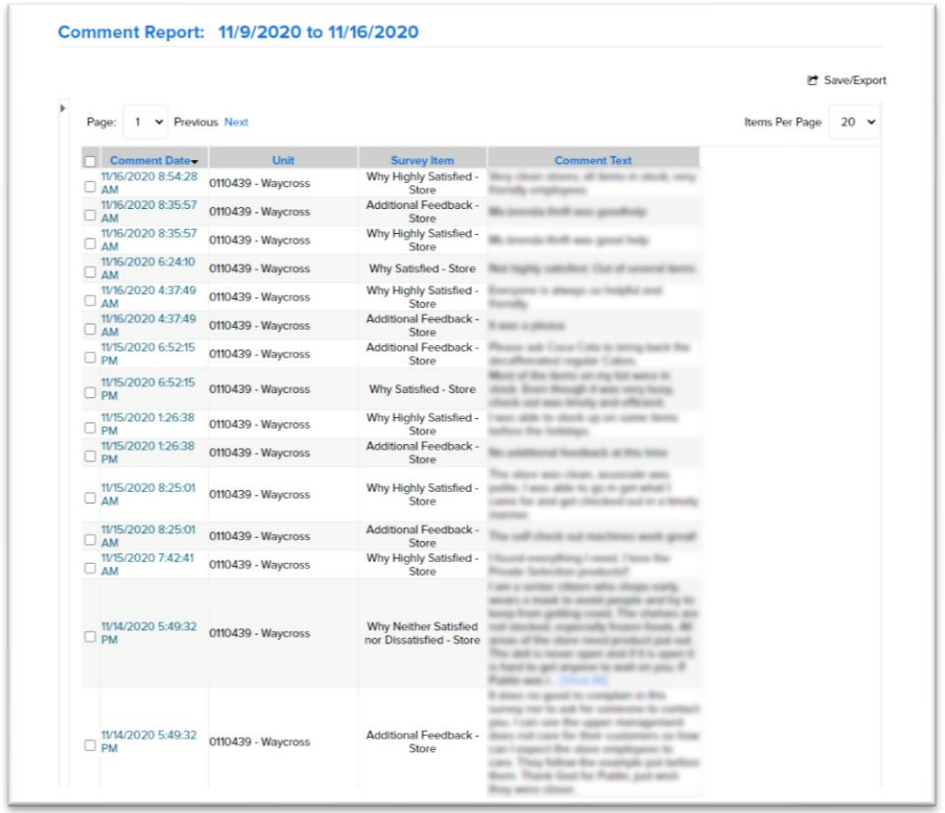


Customer satisfaction is one of our greatest assets.

HEAR IT FROM THE CONSUMER:

Erica the manager on duty is always friendly and willing to go beyond the customers needs and expectations. Always willing to help & always has a smile in her face. ***Seeing her on the job makes me shop in peace cause I know she will always make it right!*** *Just an amazing person & employee!! You guys are lucky to have her in your team!!!*

Store Management Use:



Comment Report: 11/9/2020 to 11/16/2020

Save/Export

Page: 1 Previous Next Items Per Page 20

Comment Date	Unit	Survey Item	Comment Text
11/16/2020 8:54:28 AM	0110439 - Waycross	Why Highly Satisfied - Store	
11/16/2020 8:35:57 AM	0110439 - Waycross	Additional Feedback - Store	
11/16/2020 8:35:57 AM	0110439 - Waycross	Why Highly Satisfied - Store	
11/16/2020 6:24:10 AM	0110439 - Waycross	Why Satisfied - Store	
11/16/2020 4:37:49 AM	0110439 - Waycross	Why Highly Satisfied - Store	
11/16/2020 4:37:49 AM	0110439 - Waycross	Additional Feedback - Store	
11/15/2020 6:52:15 PM	0110439 - Waycross	Additional Feedback - Store	
11/15/2020 6:52:15 PM	0110439 - Waycross	Why Satisfied - Store	
11/15/2020 1:26:38 PM	0110439 - Waycross	Why Highly Satisfied - Store	
11/15/2020 1:26:38 PM	0110439 - Waycross	Additional Feedback - Store	
11/15/2020 8:25:01 AM	0110439 - Waycross	Why Highly Satisfied - Store	
11/15/2020 8:25:01 AM	0110439 - Waycross	Additional Feedback - Store	
11/15/2020 7:42:41 AM	0110439 - Waycross	Why Highly Satisfied - Store	
11/14/2020 5:49:32 PM	0110439 - Waycross	Why Neither Satisfied nor Dissatisfied - Store	
11/14/2020 5:49:32 PM	0110439 - Waycross	Additional Feedback - Store	

Why Text Analytics?

1. High volume of unstructured data
2. High degree of subjectivity
3. Challenges from data to insights

Why Text Analytics?

1. High volume of unstructured data
2. High degree of subjectivity
3. Challenges from data to insights

Why Text Analytics?

1. High volume of unstructured data
2. High degree of subjectivity
3. Challenges from data to insights

Why We Chose Topic Modeling

There are many different types of text analytics:

- Natural Language Processing (NLP)
- Manual coding
- Sentiment analysis
- Bigrams
- Etc.

Benefits of Topic Modeling:

- Comparatively easy to understand
- No black boxes
- Enough researcher control

Drawbacks:

- Sentiment not included
- Manual coding still required

© 84.51° 2020 | Confidential

Our Approach

Our Topic Model Process

1. Comments are imported and broken down to their root words in Python
2. The model analyzes word relationships to determine the optimal number of topics
3. Each comment is assigned to a topic
4. Output for interpreting the model is generated
5. Researcher reads a sample of about 20 representative comments for each topic to name the topic

OSAT_Comment_raw
AWFUL EXPERIENCE WHILE HEADING DOWN THE CRACKER ISLE I WAS CUT OFF BY A STAFF MEMBER THAT PROCEEDED TO YELL IN MY OPINION TO A CO WORKER THAT STUFF NEEDED TO GET DONE THEN ON THE SHAMPOO ISLE TO EMPLOYEES WE RE TALKING ABOUT HOW HORRIBLE A PARTICULAR MEMBER OF MANAGEMENT WAS THEN IN PHARMACY AGAIN COWORKERS WERE TALKING ABOUT MANAGEMENT PLAYS FAVORITISM AWFUL SHOPPING EXPERIENCE GLAD I DON T LIVE IN THE AREA AND I WAS JUST PASSING THROUGH
IT IS BAD ENOUGH THAT RETAIL STORES ARE REPLACING HUMANS WITH SELF SERVICE COUNTERS BUT TO HAVE NO CHECK OUT LANES OPEN BEFORE IS RIDICULOUS THIS IS ONE OF THE CRAZIEST TIMES OF THE DAY WHEN PEOPLE ARE HURRYING INTO THE STORE BEFORE WORK THE SELF SERVICE REGISTERS ALWAYS HAVE A GLITCH AND REQUIRE SOMEONE TO COME OVER AND ASSIST IN SOME FUNCTION OR ANOTHER EVERYONE IS ANGRY AND RUNNING LATE AND THEN THEY GET IN THEIR CAI STATE IN TRAFFIC AND ON TO THEIR JOBS JUST STU

X 100,000

Comments_norm
[aw, experi, head, cracker, isl, cut, staff, member, proceed, yell, opinion, co, worker, stuff, need, get, done, shampoo, isl, employe, talk, horribl, particular, member, manag, pharmaci, cowork, talk, manag, play, favorit, aw, shop, experi, glad, live, area, pass]
[bad, enough, retail, store, replac, human, self, servic, counter, check, lane, open, ridicul, one, craziest, time, day, peopl, hurri, store, work, self, servic, regist, alway, glitch, requir, someon, come, assist, function, anoth, everyon, angri, run, late, get, car, continu, state, traffic, job, stupid, someon, part]

Topic Model takes raw comments to root words and groups like roots into topics

```
0, '0.073*"coupon" + 0.065*"produc" + 0.061*"fresh" + 0.048*"save" + 0.029*"digit" + 0.027*"meat" + 0.025*"deal"')
1, '0.139*"everyth" + 0.136*"need" + 0.118*"item" + 0.096*"found" + 0.081*"stock" + 0.054*"want" + 0.049*"got"')
2, '0.235*"good" + 0.140*"price" + 0.084*"product" + 0.038*"select" + 0.034*"qualiti" + 0.033*"food" + 0.033*"experi
3, '0.295*"great" + 0.104*"servic" + 0.084*"price" + 0.054*"custom" + 0.043*"store" + 0.039*"fuel" + 0.039*"point"')
4, '0.136*"store" + 0.130*"friendli" + 0.103*"clean" + 0.074*"employe" + 0.071*"help" + 0.061*"staff" + 0.055*"alway
5, '0.065*"fast" + 0.051*"time" + 0.049*"checkout" + 0.048*"check" + 0.041*"line" + 0.036*"wait" + 0.027*"busi"')
6, '0.070*"help" + 0.049*"cashier" + 0.031*"friendli" + 0.024*"self" + 0.020*"custom" + 0.020*"check" + 0.018*"nice"
7, '0.100*"alway" + 0.086*"kroger" + 0.084*"shop" + 0.065*"store" + 0.060*"love" + 0.040*"like" + 0.019*"best"')
8, '0.103*"easi" + 0.096*"find" + 0.063*"abl" + 0.062*"get" + 0.057*"quick" + 0.056*"quickli" + 0.056*"need"')
9, '0.039*"went" + 0.032*"one" + 0.028*"thing" + 0.026*"get" + 0.025*"problem" + 0.022*"shop" + 0.020*"lot"')
```


© 84.51° 2020 | Confidential

Example Output:

1 _Documents	2 Perc_Documents	3 Topic	4 Topic_Keywords
6907	0.2158	4.0	store, friendli, clean, employe, help, staff, alway, well, stock, nice
5249	0.1640	1.0	everyth, need, item, found, stock, want, got, sale, look, list
4253	0.1329	8.0	easi, find, abl, get, quick, quickli, need, item, store, shop
3703	0.1157	3.0	great, servic, price, custom, store, fuel, point, locat, conveni, select
3621	0.1132	7.0	alway, kroger, shop, store, love, like, best, enjoy, satisfi, go
3604	0.1126	2.0	good, price, product, select, qualiti, food, experi, sale, excel, varieti
3464	0.1082	6.0	help, cashier, friendli, self, custom, check, nice, checkout, servic, groceri
2072	0.0648	9.0	went, one, thing, get, problem, shop, lot, time, everi, buy
1943	0.0607	5.0	fast, time, checkout, check, line, wait, busi, even, reason, open
1907	0.0596	0.0	coupon, produc, fresh, save, digit, meat, deal, special, sale, love

© 84.51° 2020 | Confidential

Example Topic Naming

OSAT_Sco	OSAT_Comment_raw	Topic	Perc_Cont	Topic_Keywords
5	READY ON TIME AND ITEMS WERE CORRECT. RECEIVED TEXT MESSAGE WHEN ITEMS WERE NOT AVAILABLE AND A CHOICE IF A SUBSTITUTE IS AVAILABLE.	0	0.746	order, time, item, pick, substitut, everyth, got, readi, avail, stock
5	I PLACED MY ORDER THIS MORNING AND WAS ABLE TO PICK UP SAME DAY. THERE WAS ONLY ONE ITEM I DIDN'T ACCEPT THE SUB FOR AND ONE ITEM COMPLETELY OUT OF STOCK BUT THATS OK.	0	0.7439	order, time, item, pick, substitut, everyth, got, readi, avail, stock
5	EVERYTHING WAS TIMELY. ORDER WAS CORRECT. HAD ALMOST EVERYTHING I ORDERED IN STOCK. SUBSTITUTE ITEMS WERE ACCEPTABLE.	0	0.7359	order, time, item, pick, substitut, everyth, got, readi, avail, stock
5	ORDER WAS COMPLETED AS ORDERED. I CHOSE NOT TO SUBSTITUTE SO THERE WERE A COUPLE OF ITEMS THAT I DIDN'T GET BUT WHAT I GOT WAS EXACTLY AS I ORDERED.	0	0.7344	order, time, item, pick, substitut, everyth, got, readi, avail, stock
5	GOT EVERYTHING ORDERED EXCEPT ONE ITEM. ON A SUBSTITUTE ITEM I RECEIVED A LARGER SIZE FOR THE PRICE OF THE SMALLER ONE I HAD ORDERED.	0	0.7336	order, time, item, pick, substitut, everyth, got, readi, avail, stock
5	IT WAS VERY WELL ORGANIZED. I WAS INFORMED IN ADVANCE THAT ONE ITEM WAS UNAVAILABLE AND I WAS ASKED TO APPROVE ONE SUBSTITUTION.	0	0.7284	order, time, item, pick, substitut, everyth, got, readi, avail, stock

Topic_Keywords	
custom, ask, said, rude, cashier, servic, manag, employe, help, told	→ Unfriendly (Checkout) Employees
store, kroger, shop, time, go, get, back, custom, cash, year	→ Cash back fee
check, self, line, checkout, open, one, wait, cashier, custom, lane	→ Wait time and # of open lanes at (self-) checkout
store, cart, aisl, shop, employe, stock, item, need, find, get	→ Blocked aisles / shelves not stocked /shopping carts
stock, item, produc, kroger, store, product, brand, shelv, buy, select	→ Poor quality of produce / items out of stock
item, price, sale, charg, get, buy, purchas, check, sign, back	→ Items incorrectly priced / didn't pay sale price
deli, wait, order, chicken, meat, minut, went, one, counter, get	→ Wait time at deli / Poor quality deli chicken & meat
coupon, card, use, point, purchas, digit, gift, receipt, credit, get	→ Digital Coupons & Gift cards
bag, groceri, put, fuel, milk, bagger, item, got, pack, home	→ Bagging of grocery items

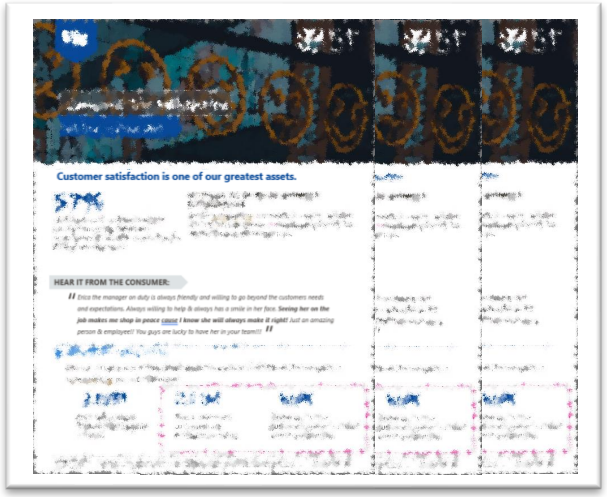
Our Current State

Enterprise Leadership Use:

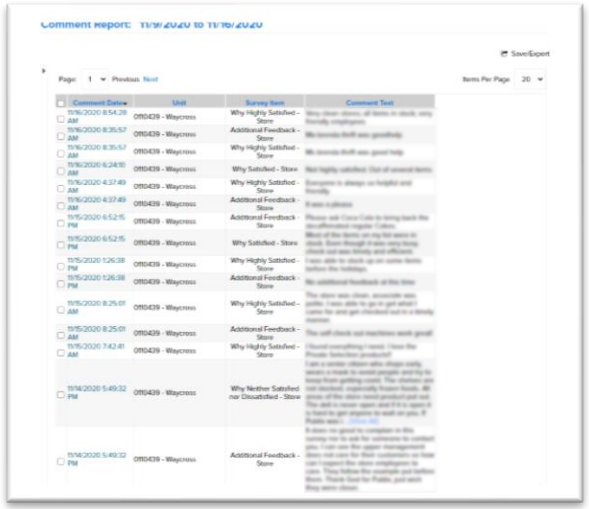
Most Mentioned Topics by:

HIGHLY SATISFIED CUSTOMERS	DISSATISFIED CUSTOMERS
Overall Great Service. Friendly & Helpful Employees	Items Out of Stock/Missing from order
Got everything I needed/In Stock	Consistently disappointed *"Never gets my order right AND Never ready on time"*
Easy, Quick, Convenient Experience	Long Wait Time/Order not ready
Good car-side attendant experience	Store had to reschedule my pickup for later/next day
Appreciate Pickup during pandemic	Unaware I was receiving substitutions

Continued Enterprise Use:



Store Management Use:



Integrating with Analysis

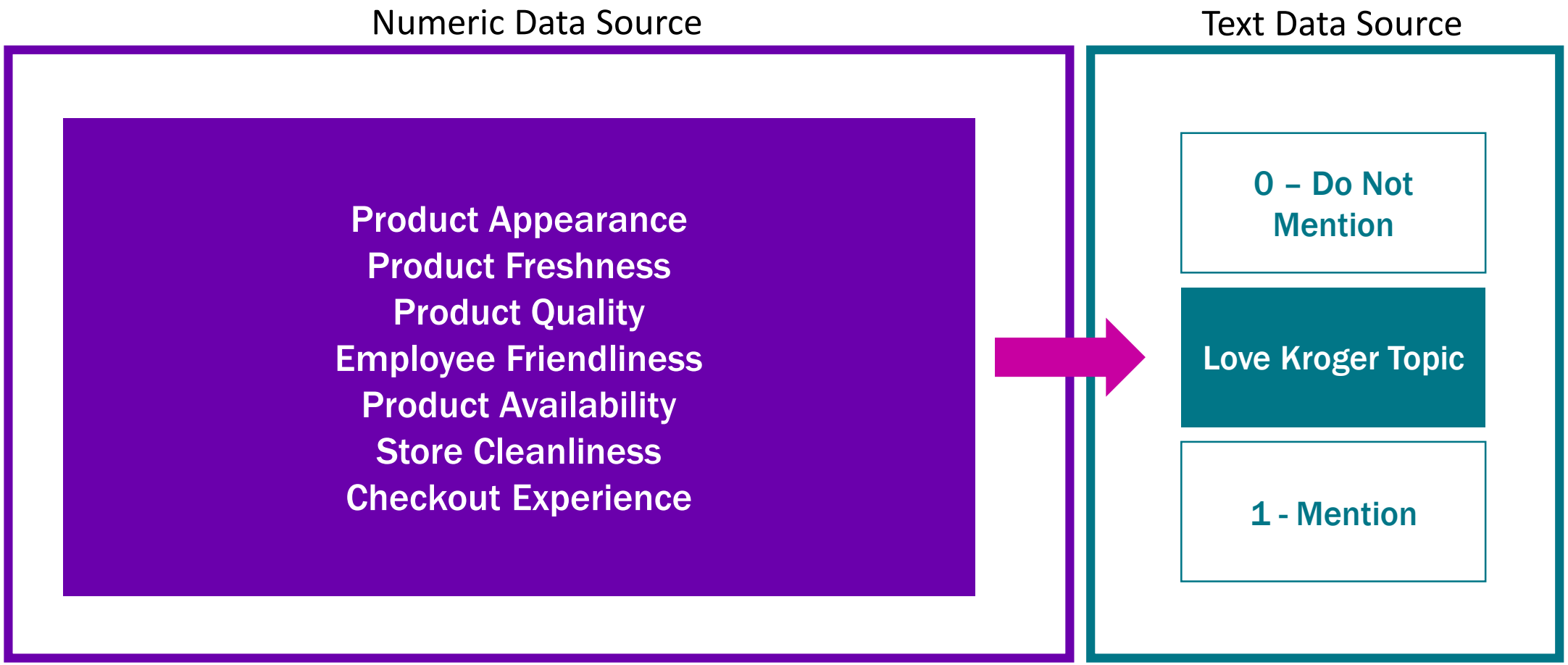
Integrating TA with Quantitative Analysis

Advantages:

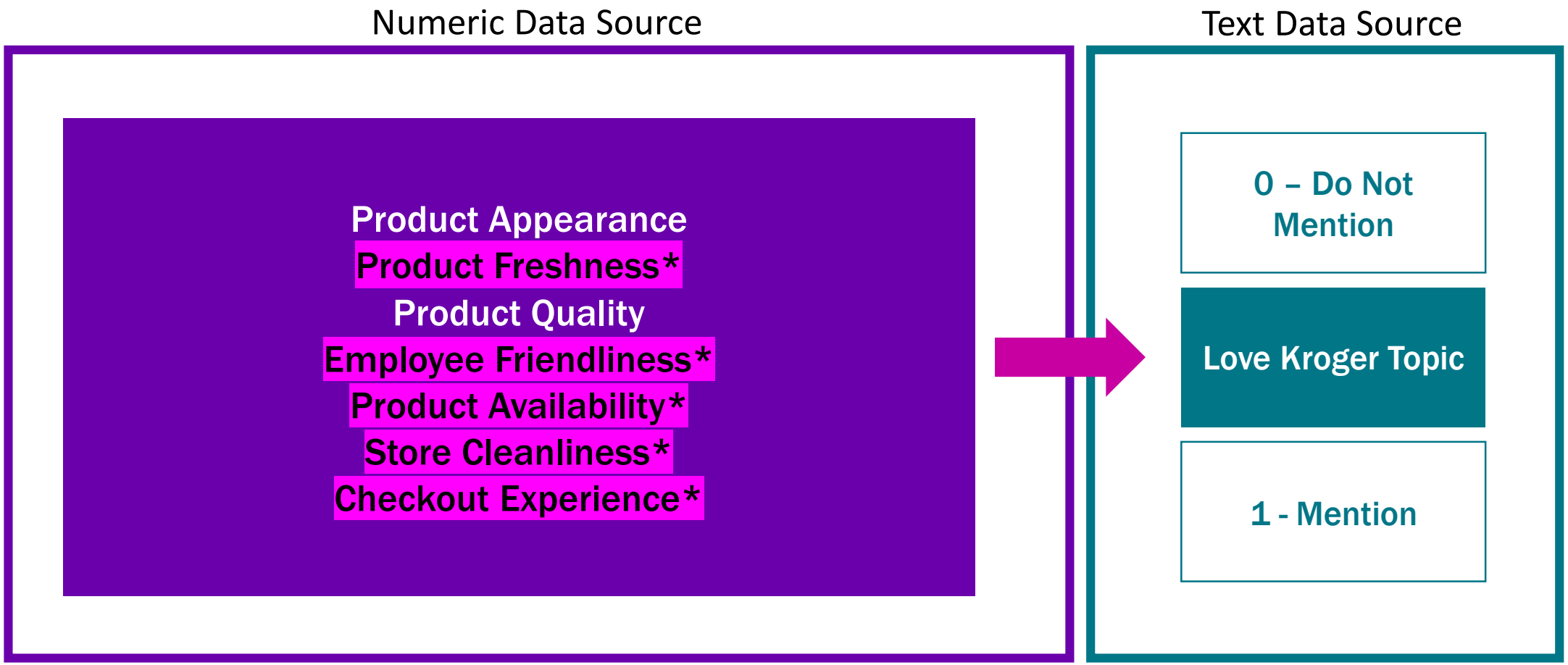
- Ability to examine statistical relationships between topics without additional steps in the survey process
- Flexibility
- Moves TA beyond descriptive categories

© 84.51° 2020 | Confidential

Logistic Regression Example



Logistic Regression Results



Integrating TA into Qualitative analysis

Replacing the first few steps of qualitative analysis

- Supporting and automating theming and clustering in qualitative analysis
- Developing a coding schema

Topics as a starting point for interview guide

First step of instrument creation process

- Text analysis → questionnaire development → cognitive interview → quantitative testing

Search terms to identify needs and trends

© 84.51° 2020 | Confidential

What we want you to take away

1. There are approachable solutions to customer comments/large amounts of text data
2. Topic Modeling allows us to capture specific ideas from a large volume of comments
3. Topic Modeling interpretation is scalable within an organization with limited technical or coding resource
4. Consider new tools that enhance the analysis you already do

© 84.51° 2020 | Confidential

THANK YOU

Contact:
Victoria.Dickman-Burnett@8451.com
Natalie.Wozniak@8451.com