Austin Coffee Shop Customer Segmentation

Janie Chen

Agenda

The Business Overview

Business Problem

Coffee shops have difficulty positioning to attract customers.

Importance

- ❖ Improve Customer Satisfaction
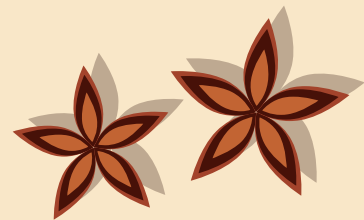- ❖ Gain Competitive Advantage
- ❖ Tailor Marketing Strategies

Business Goal

Identify and analyze customer preferences and behaviors across local boutique and chain coffee shops in Austin, TX to uncover key themes and improve targeted marketing strategies.
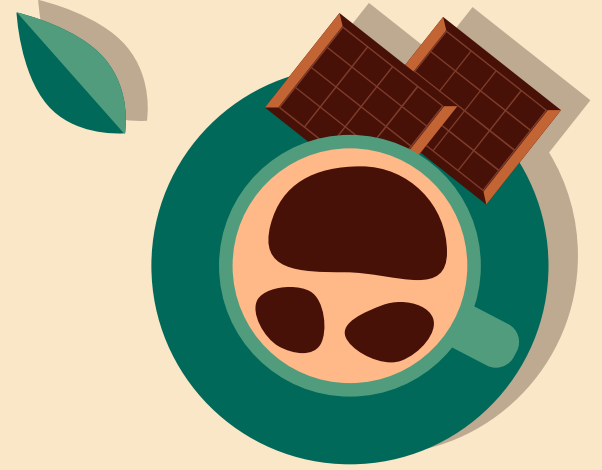
Our Approach

Step 01

Scraping & Splitting Data

Collect 5,880 Google reviews and ratings from 16 coffee shops

Analyze chain (6 stores) & boutique (10 stores) stores separately

Step 02

Topic Modeling

Use LDA to group review content into 5 key topics for boutiques and 5 for chain

Find each review's topic percentage

Step 03

Sentiment Analysis

Look at overall sentiment for each review

Assign a positive or negative label based on sentiment score

Step 04

Clustering

Use topic percentages, sentiment scores, and ratings to group reviews into 5 customer segments
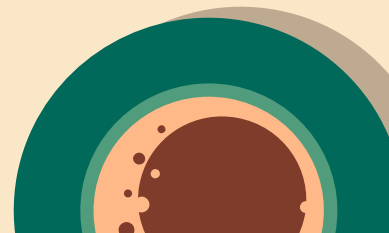
K-Means Clustering Steps

1. Scale the Data
2. Find the optimal k-clusters based on the sum of distances
3. Cluster with the original 7 dimensions
4. Print Summary Dataframes
5. PCA to visualize in 2 dimensions

K-Means Clustering

Elbow Method for Optimal K

Chose 5 clusters as the optimal

K-Means Cluster Summary

Boutique Stores

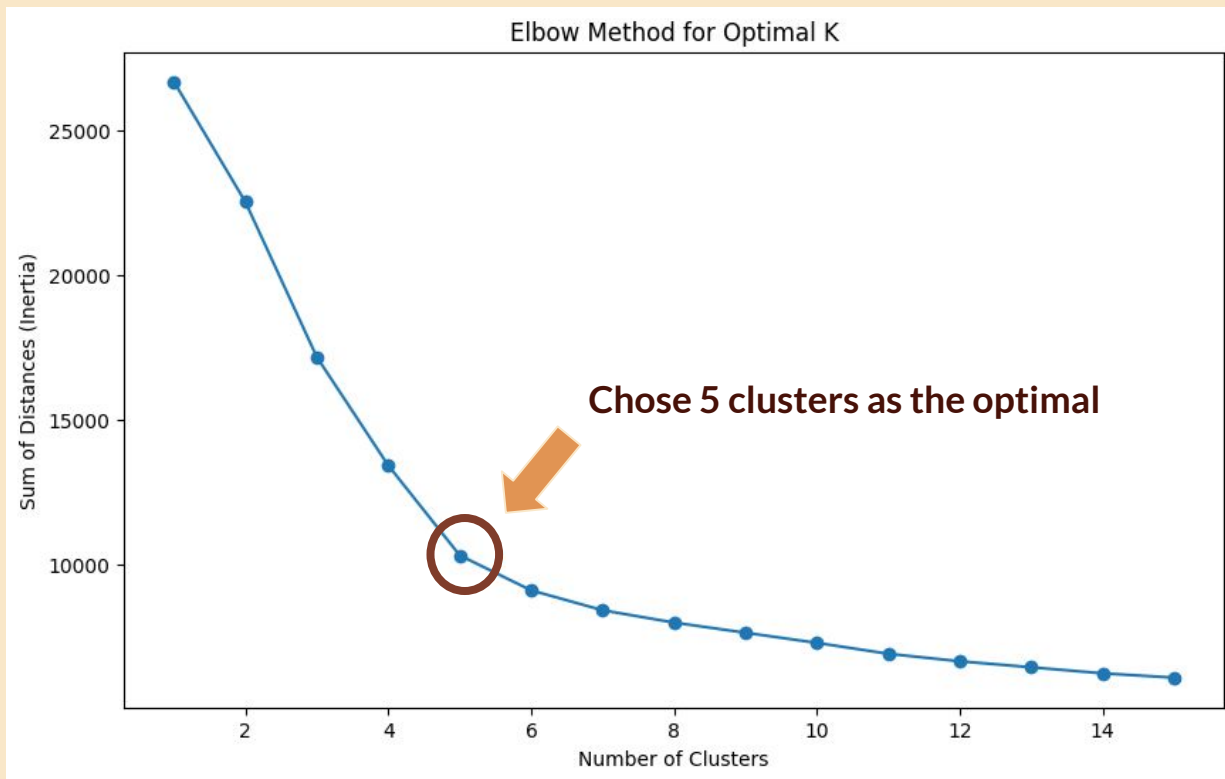| Cluster	Atmosphere	Beverages and Food Variety	Service and Quality	Entertainment	Complaints	sentiment_score	Rating
1	0.133604	0.087971	0.077298	0.599298	0.101829	0.810809	4.791908
2	0.093652	0.068774	0.679707	0.050480	0.107386	0.760923	4.879620
3	0.151840	0.158445	0.105651	0.064525	0.519540	0.039011	2.193211
4	0.108762	0.624743	0.101651	0.060000	0.104844	0.740829	4.559066
5	0.710272	0.067041	0.081689	0.051220	0.089778	0.787264	4.679352

Chain Stores

Cluster	Customer Service and Ambiance	Food and Donuts	Order Process and Wait Times	Positive Experience and Loyalty	Specialty Coffee	sentiment_score	Rating
1	0.074353	0.644796	0.124243	0.074477	0.082132	0.237972	3.123636
2	0.108013	0.069822	0.097854	0.073935	0.650375	0.731881	4.641129
3	0.695431	0.064818	0.096168	0.068018	0.075564	0.655122	4.588353
4	0.090687	0.078980	0.087726	0.668199	0.074407	0.635155	4.505376
5	0.074663	0.139460	0.606487	0.118069	0.061321	-0.227137	1.583138

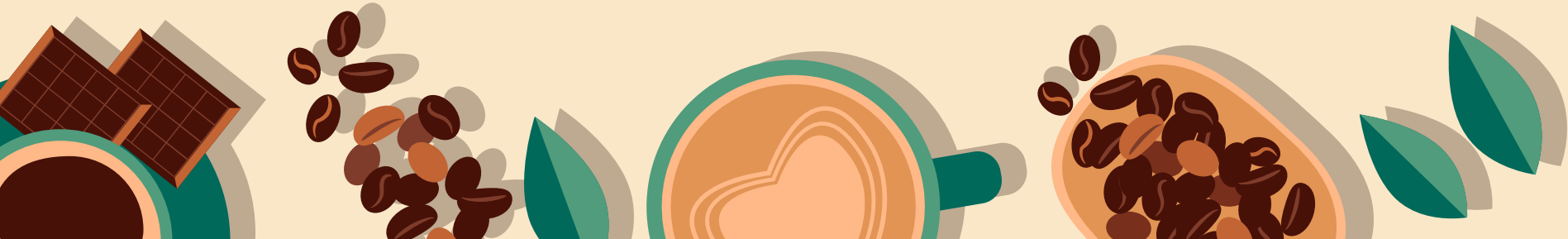# K-Means Cluster Visualization
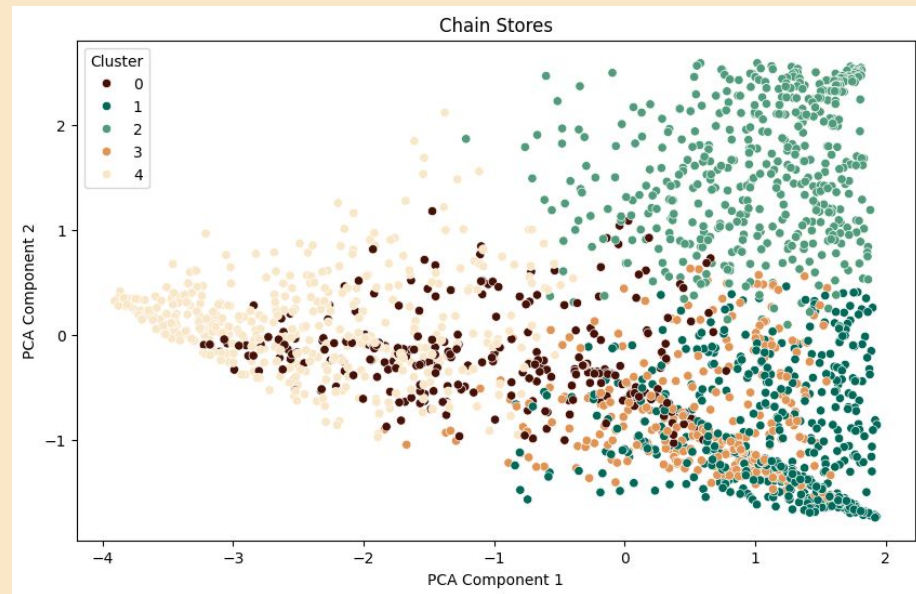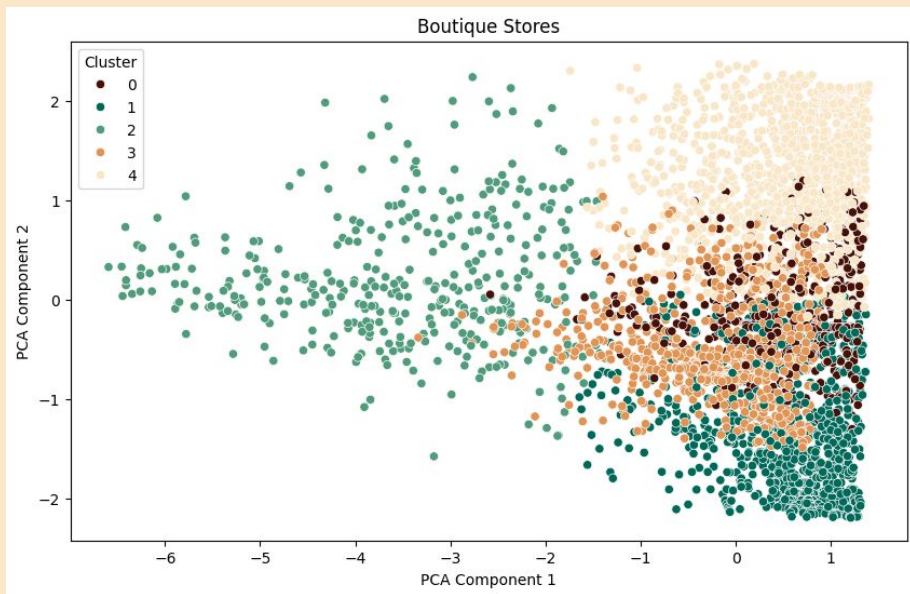
Interpreting the Results

01

Customer Segmentation

02

Chain vs. Boutique

03

Topic & Store Ratings

Boutique Store Cluster Customer Profiles

	Cluster 1 (14%)	Cluster 2 (25%)	Cluster 3 (10%)	Cluster 4 (19%)	Cluster 5 (32%)
Atmosphere	13.4%	9.4%	15.2%	10.9%	71.0%
Food & Beverage Variety	8.8%	6.9%	15.8%	62.5%	6.7%
Service & Quality	7.7%	68.0%	10.6%	10.2%	8.2%
Entertainment	59.9%	5.0%	6.5%	6.0%	5.1%
Complaints	10.2%	10.7%	52.0%	10.5%	9.0%
Sentiment Score (-1 to 1)	0.811 - Very Pos	0.760 - Very Pos	0.039 - Neutral	0.741 - Very Pos	0.787 - Very Pos
Rating (out of 5)	4.79	4.88	2.19	4.56	4.68
Customer Profile	Here for the Show Not the Drink	Wants Good Coffee and Fast	Disappointed Dudes	Adventurous Consumers	Vibes are Everything

Chain Store Cluster Customer Profiles

	Cluster 1 (13%)	Cluster 2 (24%)	Cluster 3 (24%)	Cluster 4 (18%)	Cluster 5 (21%)
Customer Service & Ambiance	7.4%	10.8%	69.5%	9.1%	7.5%
Food & Donuts	64.5%	7.0%	6.5%	7.9%	13.9%
Order Process & Wait Times	12.4%	9.8%	9.6%	8.8%	60.1%
Positive Experience & Loyalty	7.4%	7.4%	6.8%	66.8%	11.8%
Specialty Coffee	8.2%	65.0%	7.6%	7.4%	6.1%
Sentiment Score (-1 to 1)	0.238 - Pos	0.732 - Very Pos	0.655 - Very Pos	0.635 - Very Pos	-0.227 - Neg
Rating (out of 5)	3.12	4.64	4.59	4.51	1.58
Customer Profile	**Quick Bites**	**Novelty Item Seekers**	**Happy Dine-Ins**	**The Regulars**	**Impatient Karens**

Comparing Chain v. Boutique
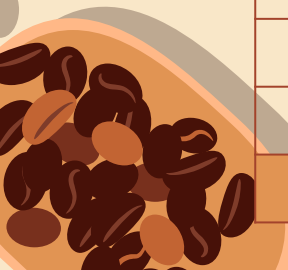
Boutique Cafe Topics	Chain Cafe Topics
Atmosphere	Customer Service and Ambiance
Food and Beverages Variety	Food and Donuts
Service and Quality	Order Process and Wait Times
Entertainment	Positive Experience and Loyalty
Complaints	Specialty Coffee

Different → (Entertainment / Complaints) and (Positive Experience and Loyalty / Specialty Coffee)

Boutique Cafe Customer Profiles	Chain Cafe Customer Profiles
Vibes are Everything (32% of reviews)	Happy Dine-Ins (24%)
Disappointed Dudes (10%)	Impatient Karens (21%)
Adventurous Consumers (19%)	Novelty Item Seekers (24%)
Wants Good Coffee and Fast (25%)	Quick Bites (13%)
Here for the Show Not the Drink (14%)	The Regulars (18%)

Per Store Average Topic Frequencies

Chain Stores

Coffee Shop Name - Chain	Customer Service and Ambiance	Food and Donuts	Order Process and Wait Times	Positive Experience and Loyalty	Specialty Coffee
Dunkin'	0.133768	0.263681	0.315425	0.203832	0.083294
Starbucks	0.358329	0.124144	0.152374	0.257367	0.107785
Summer Moon Coffee	0.254203	0.083099	0.127324	0.127624	0.407750

Boutique Stores

Coffee Shop Name	Atmosphere	Beverages and Food Variety	Complaints	Entertainment	Service and Quality
Caffé Medici	0.372679	0.176854	0.159496	0.105054	0.185917
Epoch Coffee	0.404271	0.121117	0.239182	0.083474	0.151955
Figure 8 Coffee Purveyors	0.234658	0.176147	0.121312	0.091570	0.376314
Flat Track Coffee	0.209182	0.220870	0.126228	0.070594	0.373125
Fleet Coffee	0.188842	0.182979	0.109474	0.085296	0.433409
Houndstooth Coffee	0.211799	0.230142	0.125334	0.073828	0.358897
Jo's Coffee – South Congress	0.202663	0.368297	0.178356	0.118828	0.131855
Mozart's Coffee Roasters	0.209019	0.124565	0.137537	0.454878	0.074001
Radio Coffee & Beer	0.583592	0.093064	0.124891	0.095273	0.103180
Stouthaus Coffee Pub - Sunset Valley	0.445882	0.165215	0.093028	0.101128	0.194748

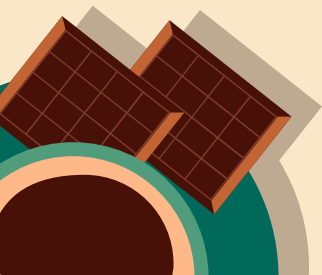# Per Most Frequent Topic Per Store - Complaints

Average Sentiment

Coffee Shop Name	Complaints
Caffé Medici	0.00000
Epoch Coffee	0.75840
Figure 8 Coffee Purveyors	0.70880
Flat Track Coffee	0.69765
Fleet Coffee	0.65880
Houndstooth Coffee	0.30740
Jo's Coffee – South Congress	0.57500
Mozart's Coffee Roasters	0.76485
Radio Coffee & Beer	0.80200
Stouthaus Coffee Pub - Sunset Valley	0.68830

Average Ratings

Coffee Shop Name	Complaints
Caffé Medici	2.573770
Epoch Coffee	2.967742
Figure 8 Coffee Purveyors	3.953488
Flat Track Coffee	4.026316
Fleet Coffee	3.760000
Houndstooth Coffee	3.206897
Jo's Coffee – South Congress	2.958904
Mozart's Coffee Roasters	3.111111
Radio Coffee & Beer	3.600000
Stouthaus Coffee Pub - Sunset Valley	3.400000

The discrepancy between high sentiment scores and low ratings indicates that reviews may be incorrectly captured by VADER.

Example: "Service and being polite don't rank high. Rude guy who served us. Guy who took order was great." has a **sentiment score of 0.76**.

Per Most Frequent Topic Per Store - Ratings

Chain Stores

	Customer Service and Ambiance	Food and Donuts	Order Process and Wait Times	Positive Experience and Loyalty	Specialty Coffee
Dunkin'	4.493976	2.824390	1.771218	3.924419	3.955556
Starbucks	4.457711	3.744681	2.958904	4.472222	3.803922
Summer Moon Coffee	4.607527	3.589744	2.625000	4.342466	4.680905

Boutique Stores

Topic 1	Coffee Shop Name	Atmosphere	Beverages and Food Variety	Complaints	Entertainment	Service and Quality
0	Caffé Medici	4.401361	4.015385	2.573770	4.486486	4.546875
1	Epoch Coffee	4.488764	4.600000	2.967742	4.608696	4.711864
2	Figure 8 Coffee Purveyors	4.780220	4.409836	3.953488	4.807692	4.893750
3	Flat Track Coffee	4.708861	4.560440	4.026316	4.944444	4.802469
4	Fleet Coffee	4.573770	4.584906	3.760000	4.960000	4.887006
5	Houndstooth Coffee	4.500000	4.414894	3.206897	4.666667	4.822368
6	Jo's Coffee – South Congress	4.416667	4.378049	2.958904	4.736842	4.707317
7	Mozart's Coffee Roasters	4.594595	4.157895	3.111111	4.711111	4.600000
8	Radio Coffee & Beer	4.745520	4.500000	3.600000	4.857143	4.935484
9	Stouthaus Coffee Pub - Sunset Valley	4.678947	4.551724	3.400000	4.606061	4.853659

Caption: This PivotTable shows the average ratings per store for each topics reviews. Topics were assigned to each review according to the highest frequency of topic words in each review.

The findings from this table are consistent with our intuition.

Topics & Ratings Per Store - Standouts

Coffee Shop	Most Frequent Topic	Freq Topic Rating	Highest Rated Topic	Highest Rated Topic Rating
Dunkin'	Order and Wait Times	1.77	Customer Service and Ambiance	4.49
Starbucks	Customer Service and Ambiance	4.45	Positive Experience and Loyalty	4.47
Summer Moon Coffee	Specialty Coffee	4.68	Specialty Coffee	4.68
Epoch Coffee	Atmosphere	4.49	Service and Quality	4.71
Mozart's	Entertainment	4.71	Entertainment	4.71
Radio Coffee & Beer	Atmosphere - high frequency (72%)	4.74	Service and Quality	4.85

Caption: This PivotTable shows the most frequent topic by store, the average rating for that topic per store, and the highest rated topic with the rating for that topic for comparison.

Business Applications

01 Target Segment Suggestions

02 Case Studies

Boutique Store Suggestions

1

Here for the Show

- Allocate space for stage
- Invest in stage equipment
- Invite local entertainment
- Host open-mics

2

Wants Good Coffee Fast

- Invest in quality coffee beans and brewing equipment
- Hire more staff/better employee training
- Rearrange store to streamline operations & customer flow

3

Disappointed Dudes

- Address complaints
- Diversify the menu
- Decorate to create a better atmosphere

4

Adventurous Consumers

- Diversify the menu
- Add seasonal menu items
- Create a unique, brand-specific drink

5

Vibes are Everything

- Seasonal decor and music
- Curate lighting/art that are on brand
- Employee customer service training

Chain Store Suggestions

1 Quick Bites
- Expand food menu
- Optimize drive-through
- Bundled coffee and sweets deals

2 Novelty Item Seekers
- Seasonal drinks
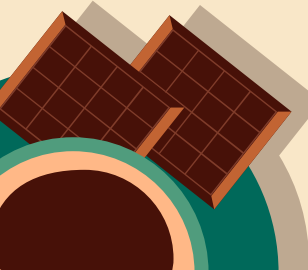- Menu items with uncommon ingredients
- Gimmick items

3 Happy Dine-Ins
- Curate music playlists
- Cozy furniture and design
- Train employee customer service skills

4 The Regulars
- Focus on consistency
- Promotions for repeat visits
- Membership deals

5 Impatient Karens
- Schedule more workers during rush hour
- "Sorry" gifts/deals
- Optimize operation flow

Boutique Store Case Study
Why is Mozart's So Successful?

Boutique Customer Segment 1

Boutique Customer Segment 5

Mozart's

Here for the Show Not the Drink

Vibes are Everything

Most Frequent Topic: Entertainment
Highest Rated Topic: Entertainment

Highest Topic: Entertainment

Highest Topic: Atmosphere

- **Strategy**: Differentiation & Max Customer Targeting (Customer Acquisition)
  - Fulfills the Austin consumer's appreciation for live entertainment
  - Captures ~½ of boutique customer base
    - 14% Segment 1 + 32% Segment 5 (Segment 5 is the largest segment)
- Entertainment & Atmosphere are related
  - Simultaneous targeting (2x the effect!)
  - Two biggest contributors to Mozart's high rating & brand reputation
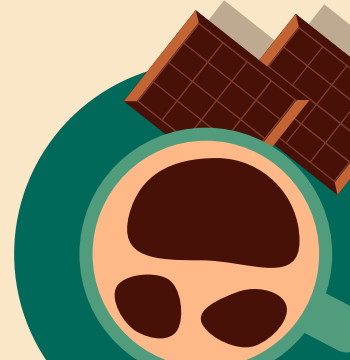
Chain Store Case Study

How to Help Dunkin' Donuts

Dunkin'	→	**Chain Customer Segment 3**	**Chain Customer Segment 5**

Dunkin'

Most Frequent Topic: Order and Wait Times
Highest Rated Topic: Customer Service and Ambiance

Chain Customer Segment 3

Happy Dine-Ins

Highest Topic: Customer Service & Ambiance

Chain Customer Segment 5

Impatient Karens

Highest Topic: Order Process & Wait Times

- **Strategy**: Remedy Customer & Reputation Loss (Customer Retention)
  - Impatient customers leave lowest ratings & have most negative sentiments
  - Impatient customers make up ~¼ of customer base
- Still captures ~½ of chain store customer base
  - 24% Segment 3 + 21% Segment 5 (Segment 3 is the largest segment)
- **Recommendations**:
  - Optimize operation flow and staffing according to demand
  - Consider omnichannel implementations for efficiency

Future Steps

Expand Scope

Analyze more coffee shops in and outside the Austin area.

JST Model

Consider using Joint Topic Sentiment models for more in-depth analysis of topic and sentiment relationships

More Data

Incorporate data from social media, forums, and more!

Thanks!

Do you have any questions?