

Voice of Market (VoM): Unlocking Market Insights from Social Media

Sponsor Name: Costco Wholesale

Faculty Advisor: Dr. Elio Zhang

Team Number: MSDS 25.4



Meet Our Team



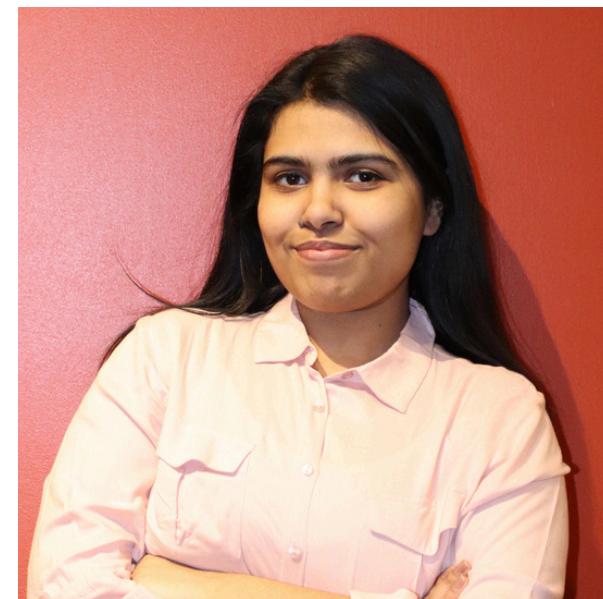
**Mustafa
Bhavanagarwala**



**Hrishikesh Deepak
Dhole**



Lakshit Gupta



Priyal Sunil Joshi



Tejaswi Neelapu

Background

About Costco:

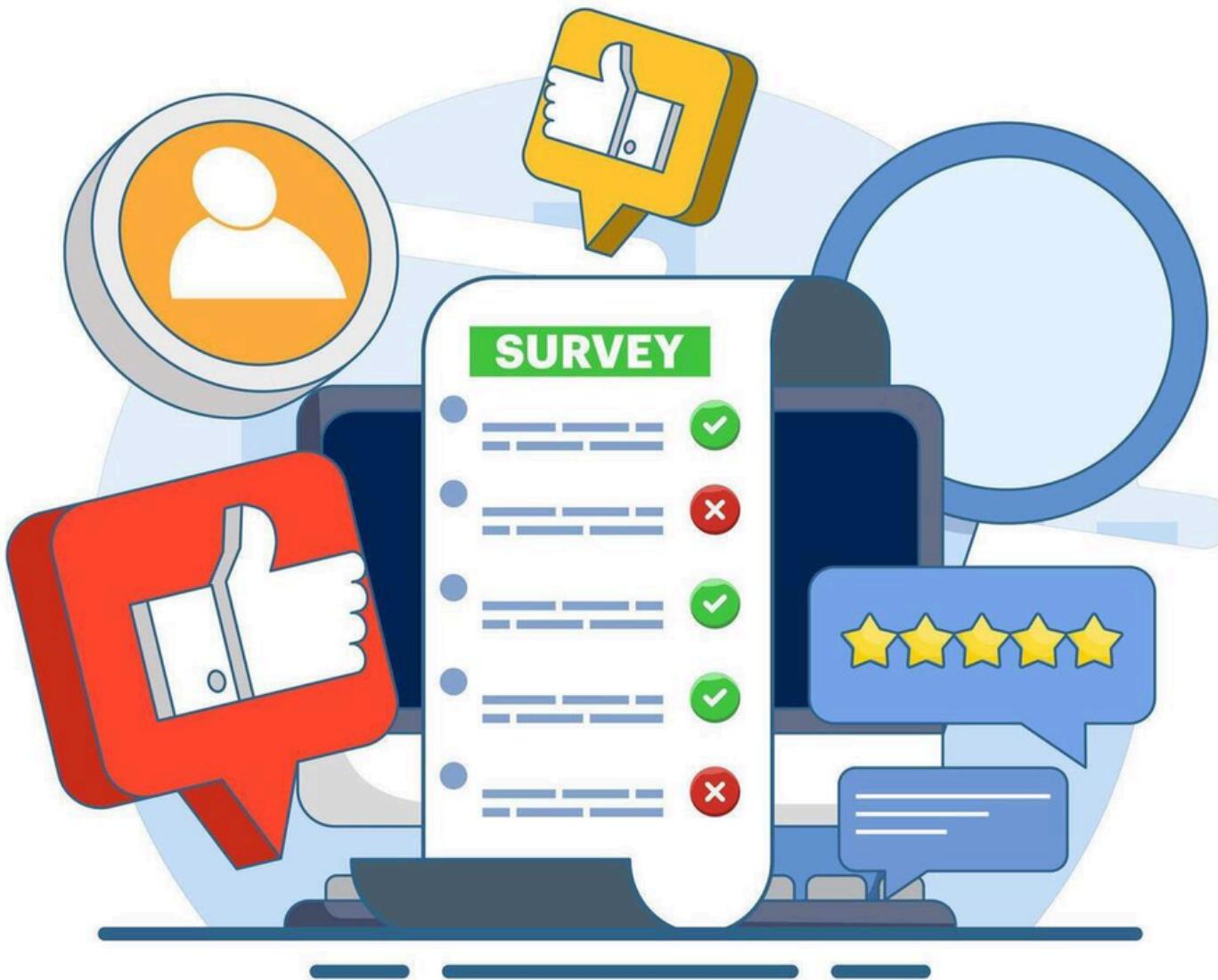
- **Global Retail Leader:** 3rd largest retailer with 137M+ loyal customers (2024).
- **Value & Quality:** Focuses on high-quality products at competitive prices, driving strong customer trust.
- **Efficient Business Model:** Membership-based structure ensures customer retention and operational efficiency.



Problem Statement

Existing Review Methods Have Limitations

- Surveys and focus groups are slow and biased
- Insights are incomplete and outdated
- Lack of real-time data analysis



Problem Statement

Challenges in Analyzing Social Media Review

- Social media data is unstructured and vast
- Insights from Reddit, Instagram, and X remain hidden
- Difficult to make real-time decisions



So how can we unlock insights from social media review??

Helpful vs. Not Helpful Customer Review – Why It Matters?

Not Helpful Review

"Love Kirkland coffee!"

- Too vague
- Not actionable
- Limited analytical value

Helpful Review

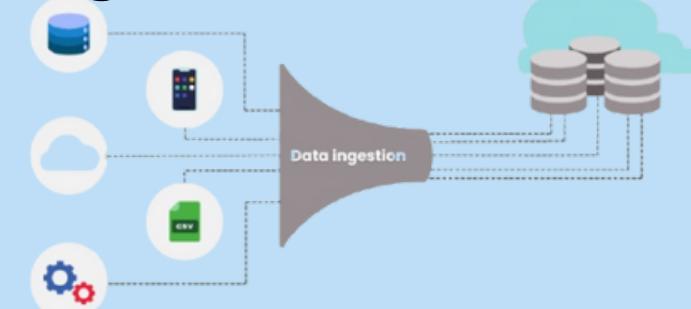
"I love Costco's Kirkland Signature coffee it's affordable and tastes great! However, I wish there were more dark roast options."

- Constructive review
- Specific & actionable
- Useful for analysis

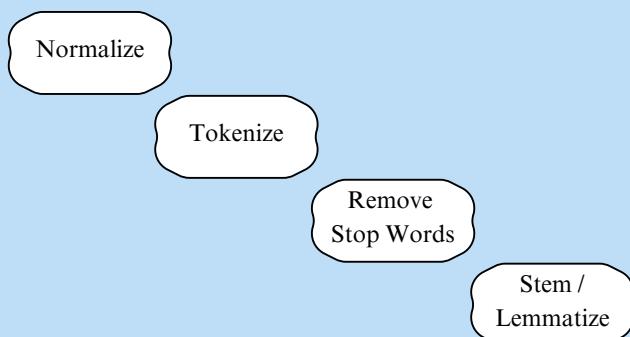
Data Science Lifecycle

Data Acquisition and Preparation

Data Cleaning



Preprocessing

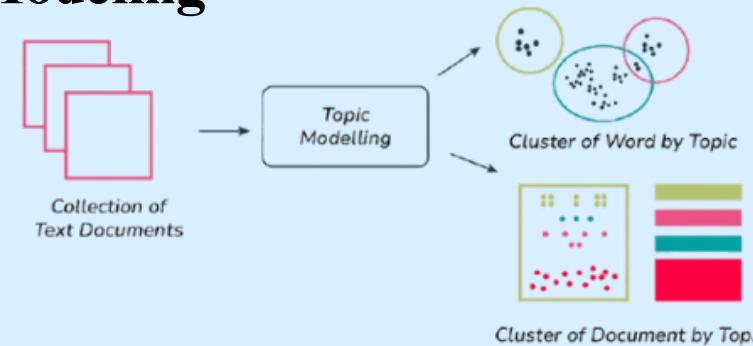


Exploratory Data Analysis

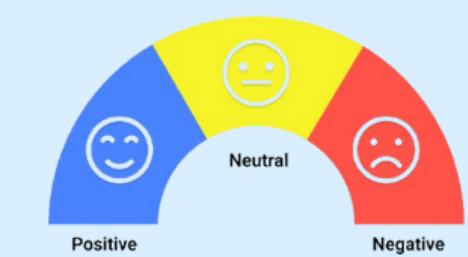


NLP Modeling & Insight Extraction

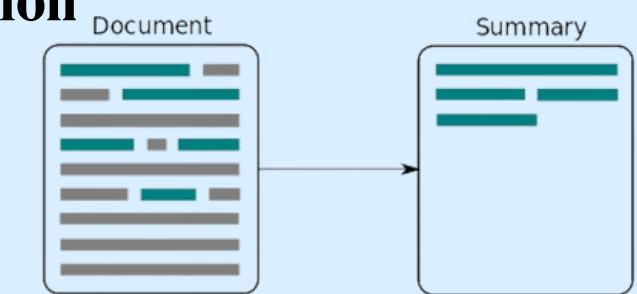
Topic Modeling



Sentiment Analysis



Summarization and Trend Identification



Insight Delivery & Deployment

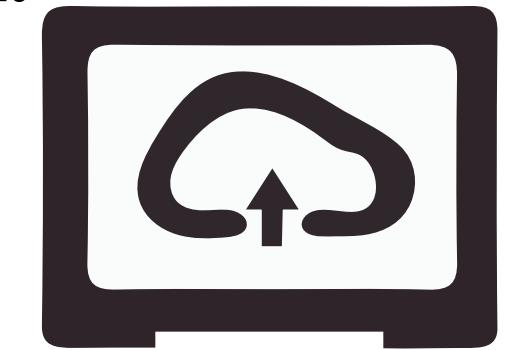
Visualization & Reporting



Chatbot



Model Deployment



Data Collection

Data Sources:

- Social media platforms
- Review sites Sitejabber
- Costco's product reviews

Challenges:

- Platform restrictions
- Data fetch rate limitations
- Inconsistent Data Formats

Web Scrapped Data:

- 200K+ rows



1 – 8 of 252 Reviews



Kirkland salted mixed nuts.

VERIFIED PURCHASER

4 months ago

Nice variety of nuts. Macadamias were larger and more plentiful than in earlier packages. Preferred the plastic jar to the resealable bag.

Yes, I recommend this product.

Helpful? (2) (0) Report



 **Costco Insider** 
@Costco_Insider

Follow

The February 2024 Hot Buys Coupons start tomorrow! costcoinsider.com/costco-februar...
What are you going to buy that's on sale?

#costco #costcoinsider #costcobeys
#costcohottbuys #costcocoupons
#costcocouponbook #costcosavingsbook
#costcoshopping


ORROW, 2/24/24! NEW
IN-WAREHOUSE HOT BU
Valid 2/24/22 - 3/31/24
Costco February / March 2024 Hot Buys Coupons

From costcoinsider.com

17:24 · 2/23/24 · 3.7K Views

7 replies · 2 retweets · 1 like

Most liked replies

[REDACTED] · 3/13/24
Hello! I just wanted to say that Costco is making a huge mistake getting rid of the Verizon Kiosks. As

Data Collection



costco

Comments

[REDACTED] 12w
How much do gingerbread houses cost? 

costco 12w · Author
[REDACTED] We advise you to contact your local Costco as pricing can vary by location. Another great option is to view this information on the Costco Mobile App. Simply click on the 'Warehouse' icon and enter the product description to view pricing and availability.

X

Instagram

Before I complain about premier chocolate protein shake with medicine taste. This caffeine Latte has no medicine taste (had it at room temp, not cold)! Love the taste. It's like killing 2 birds with one stone where i save on another cup of joe! Just one thing, it's more sweeter despite the whole bottle is only 1g of sugar.



34 44 Share

Join the conversation

Sort by: Best

[REDACTED] 12h ago · ⚡ Sticked comment

[REDACTED] 11h ago
What's also really good is mixing one of these with a can of the Kirkland Cold Brew

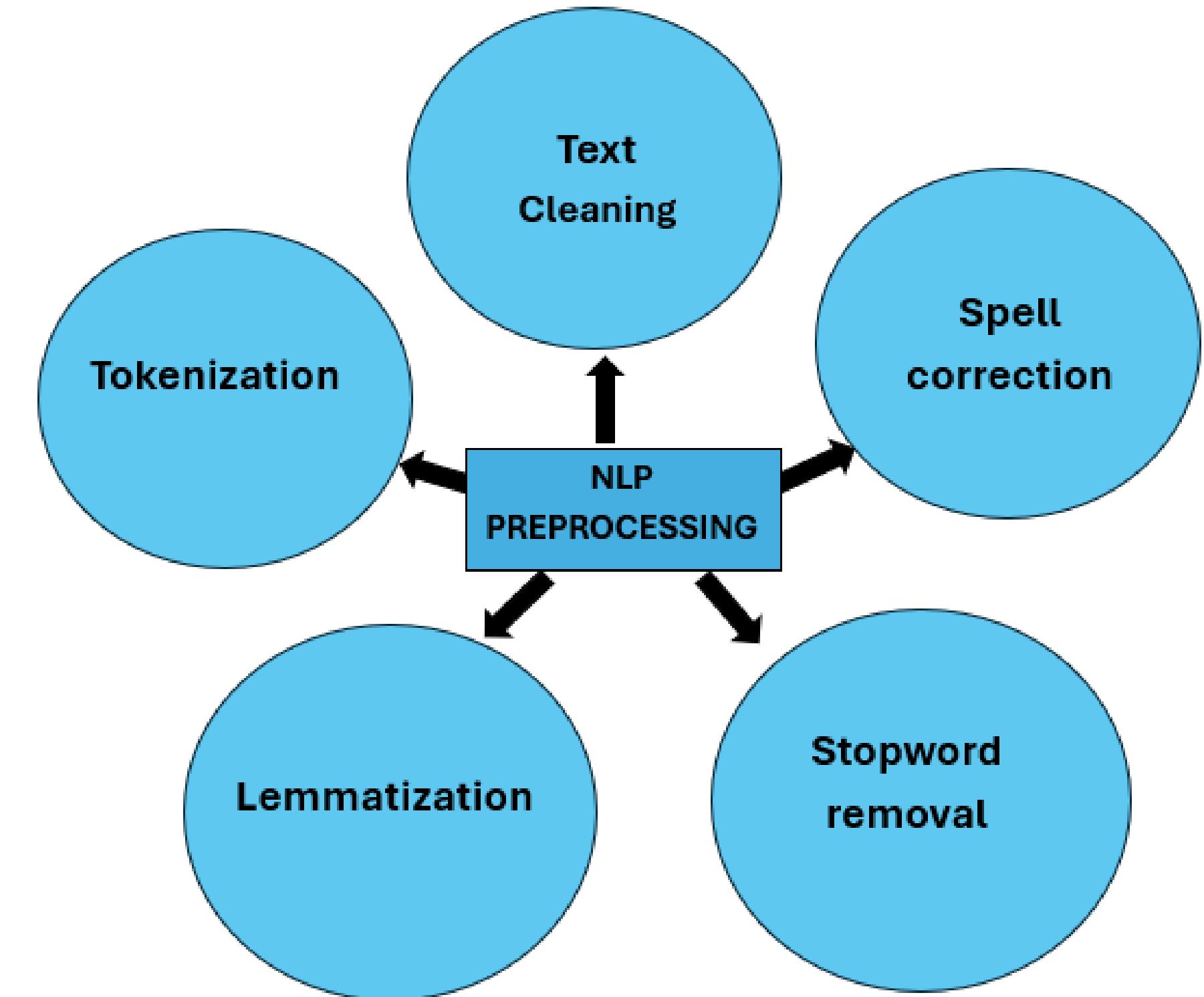
[REDACTED] 11h ago
Yessss!!

Reddit

Preprocessing

Text Preprocessing:

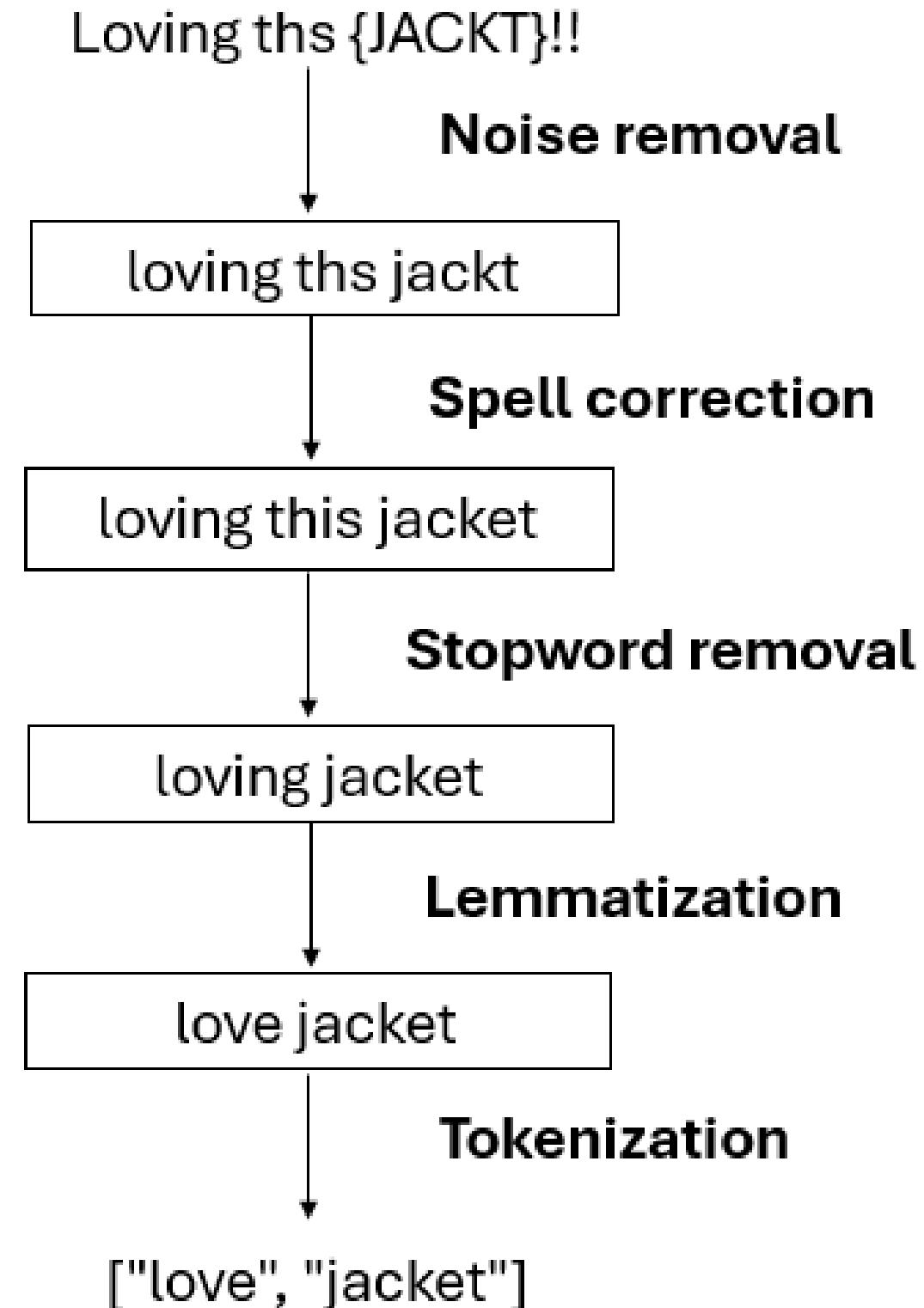
- Removing noise in data
- Spell correction
- Stopword removal
- Lemmatization
- Tokenization



Preprocessing Example

Example Review:

"Loving ths {JACKT}!!"



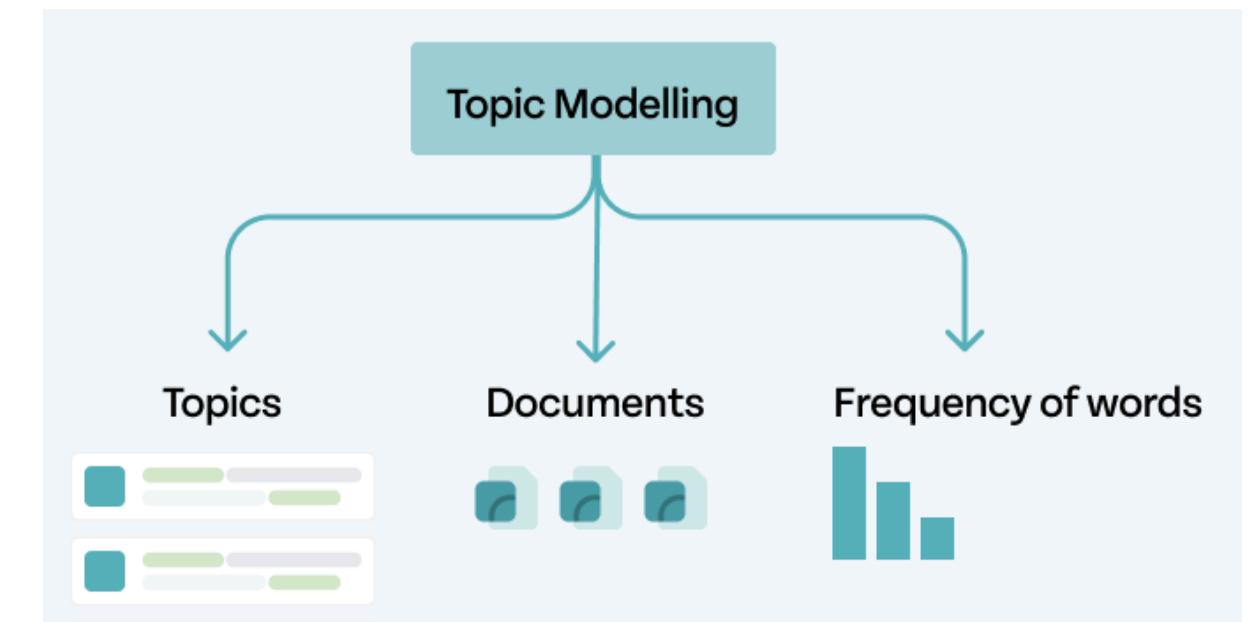
After cleaning:

- 175K+ rows remaining

Specifications

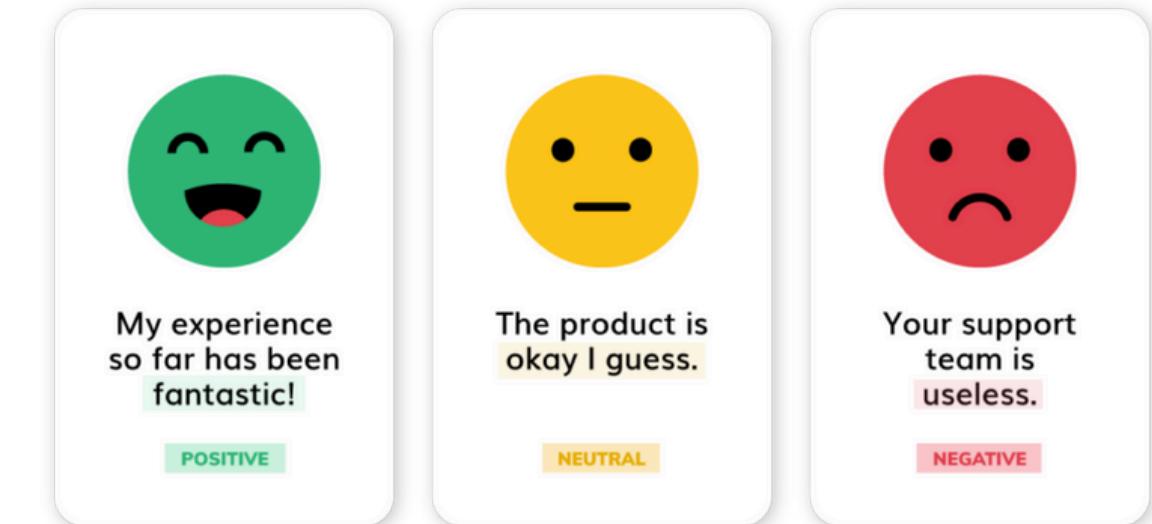
Topic Modeling

- Identification of discussion themes
- Latent Dirichlet Allocation (LDA)
- Categorizing review



Sentiment Analysis

- Reviews as positive, negative, or neutral.
- VADER for short-form sentiment analysis
- BERT for deep contextual sentiment understanding.



Specifications

AI-Powered Summarization

- Extractive Summarization
- Abstractive Summarization

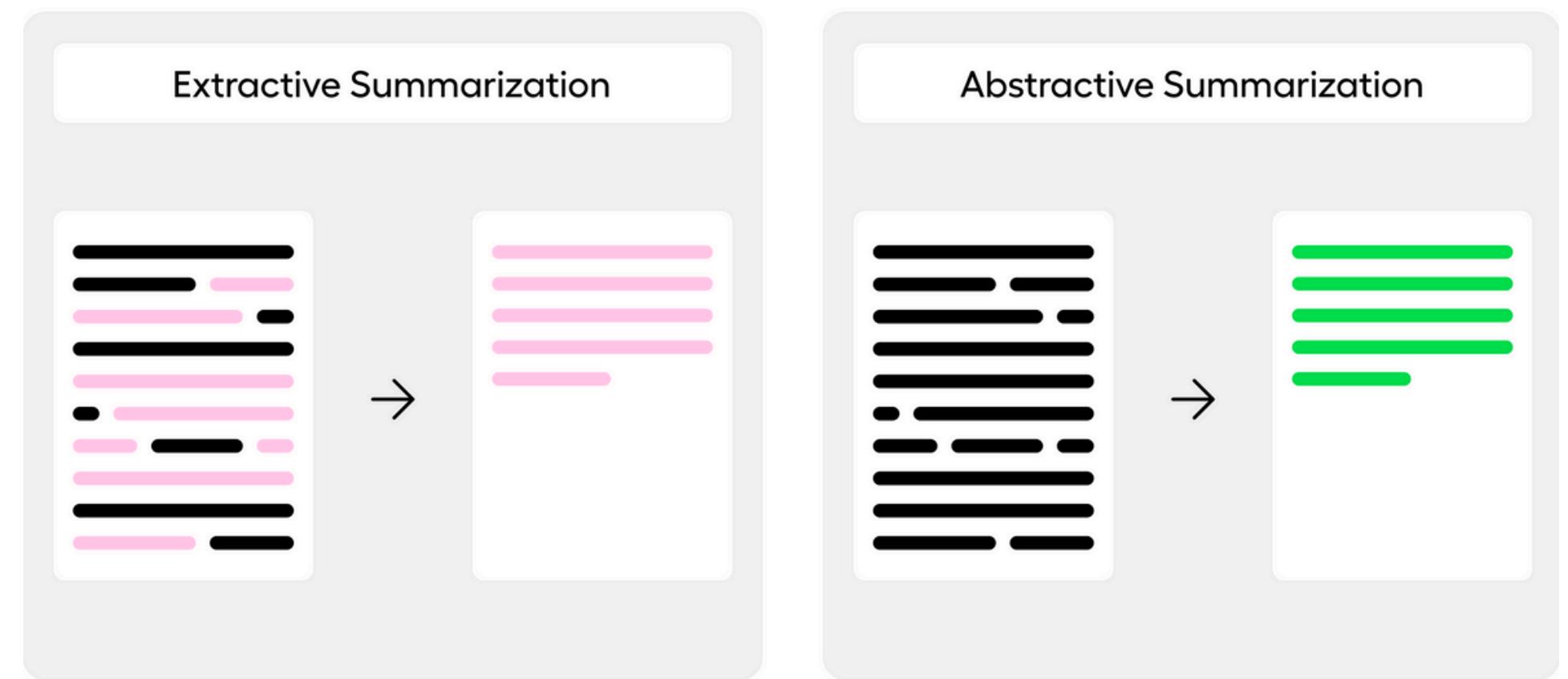
Visualization & Reporting

- Power BI dashboard
- Interactive knowledge graph

Compliance & Ethics

- General Data Protection Regulations & ethical AI principles.
- Anonymizing data.

Extractive vs Abstractive Summarization



Word Cloud Analysis of Instagram

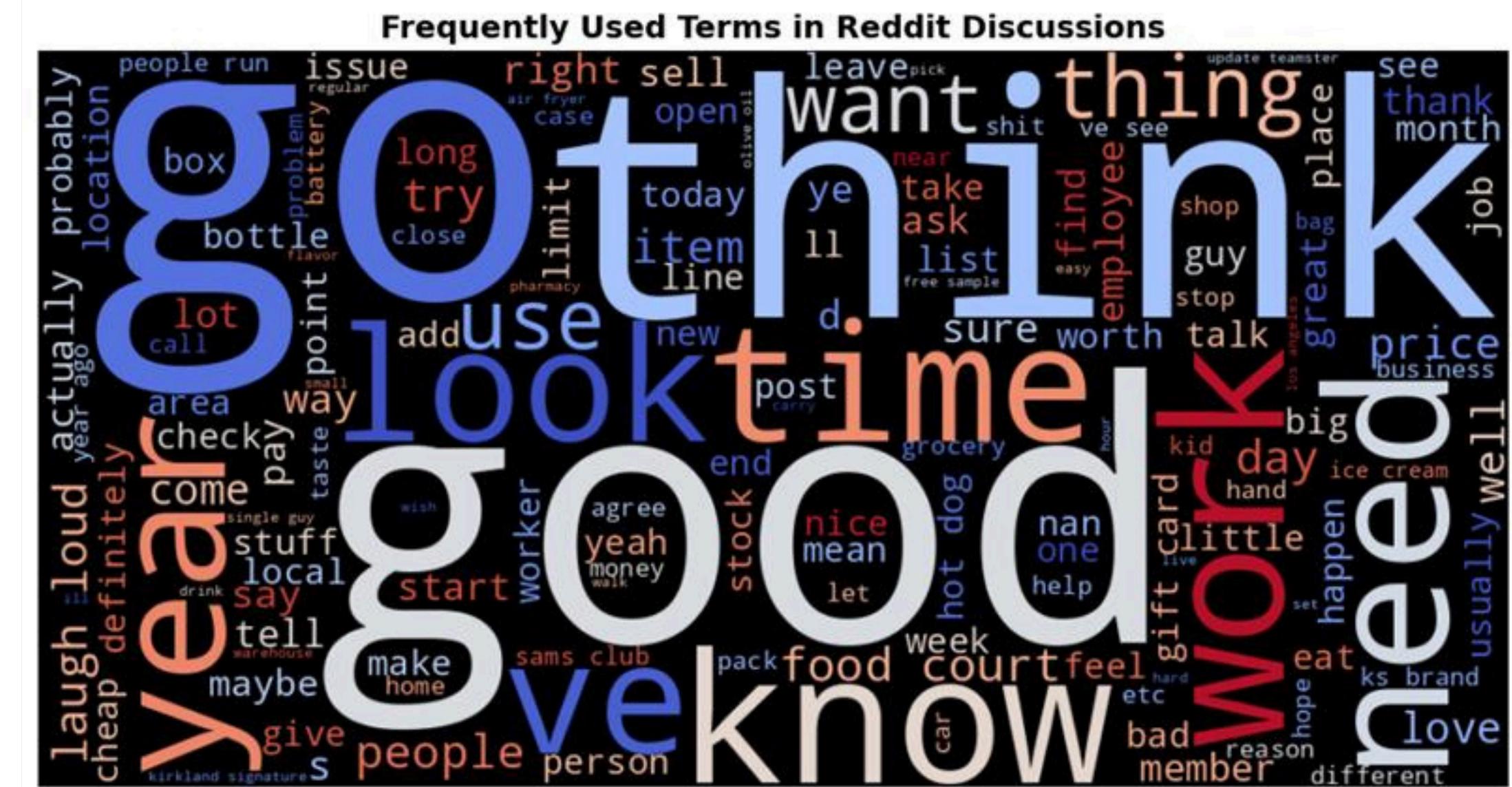
- **Key Themes:** Focus on "Costco," "love," "please," and "bring back."
 - **Customer Engagement:** High frequency of "please" and "love" suggests brand loyalty.
 - **Common Topics:**
 - Product requests
 - Positive sentiment

Frequently Used Terms in Instagram Discussions



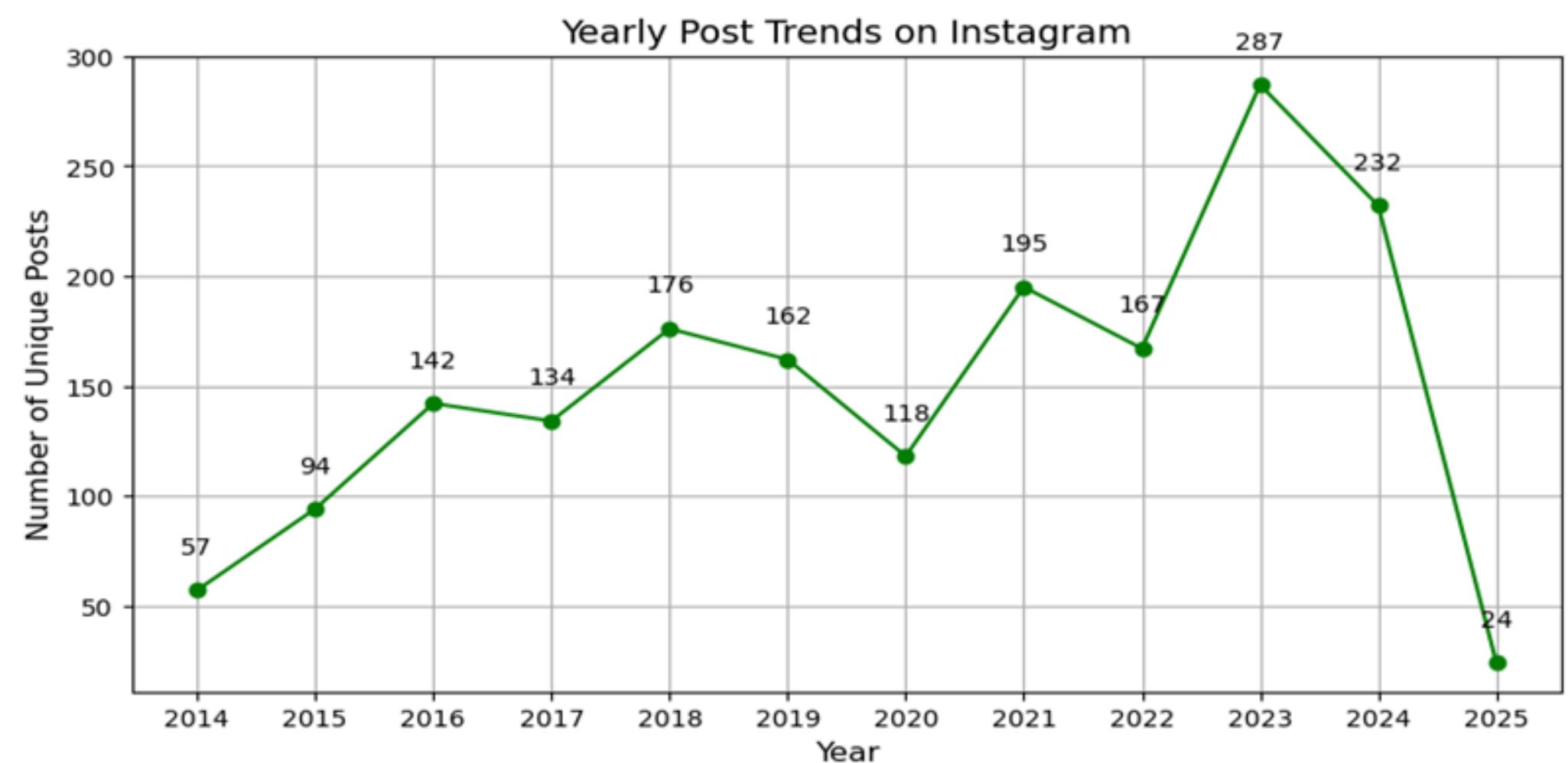
Word Cloud Analysis of Reddit

- **Key Themes:** More discussion-driven compared to Instagram
 - **Customer Behavior:**
 - Opinion-heavy, indicating debates and feedback
 - Less about brand loyalty, more about experiences and product/service analysis.
 - **Common Topics:**
 - Store experiences
 - Product discussions



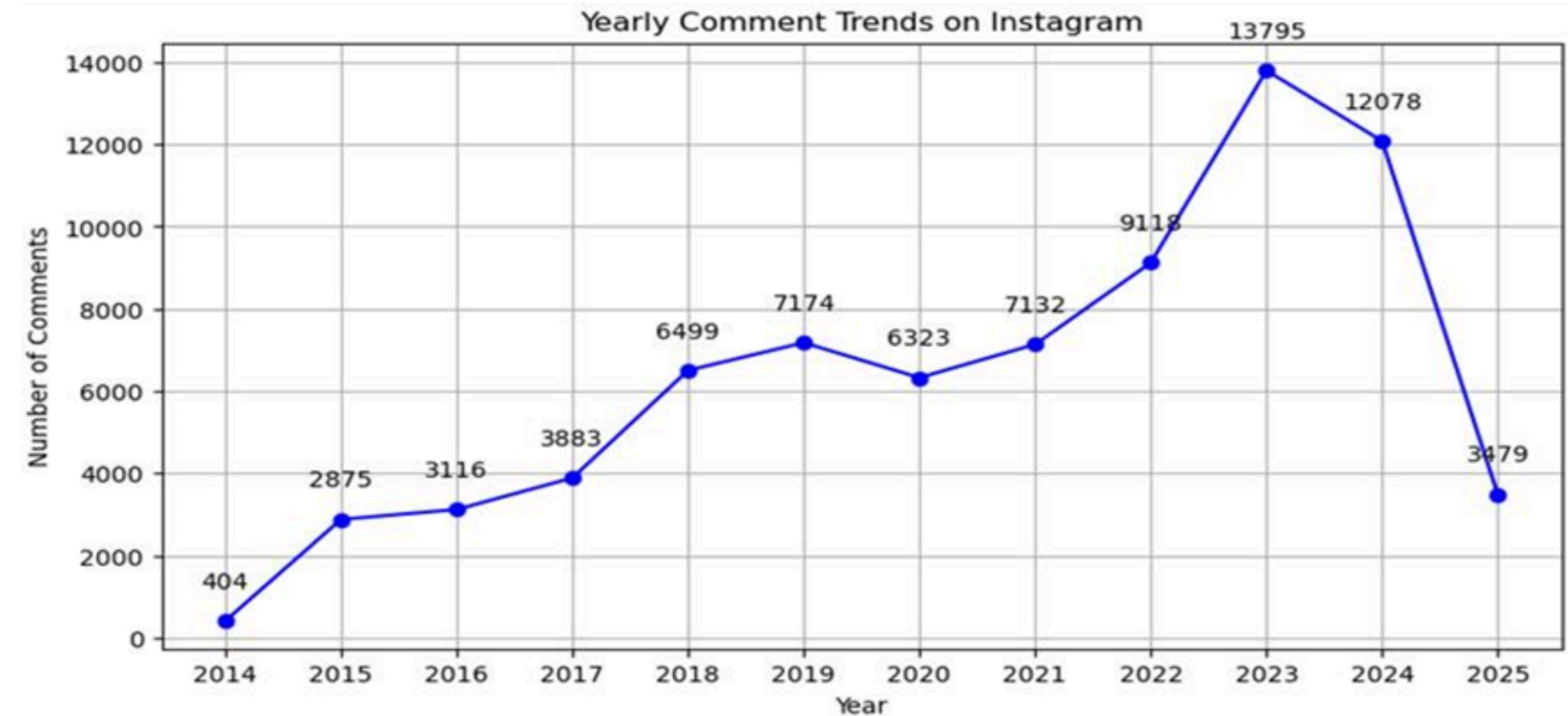
Yearly Trends in Instagram Posts

- Increase in posts from 2014 to 2023.
- Fluctuations may correlate with marketing campaigns or trends.
- Steady growth in posts reflects Costco's increasing online engagement.



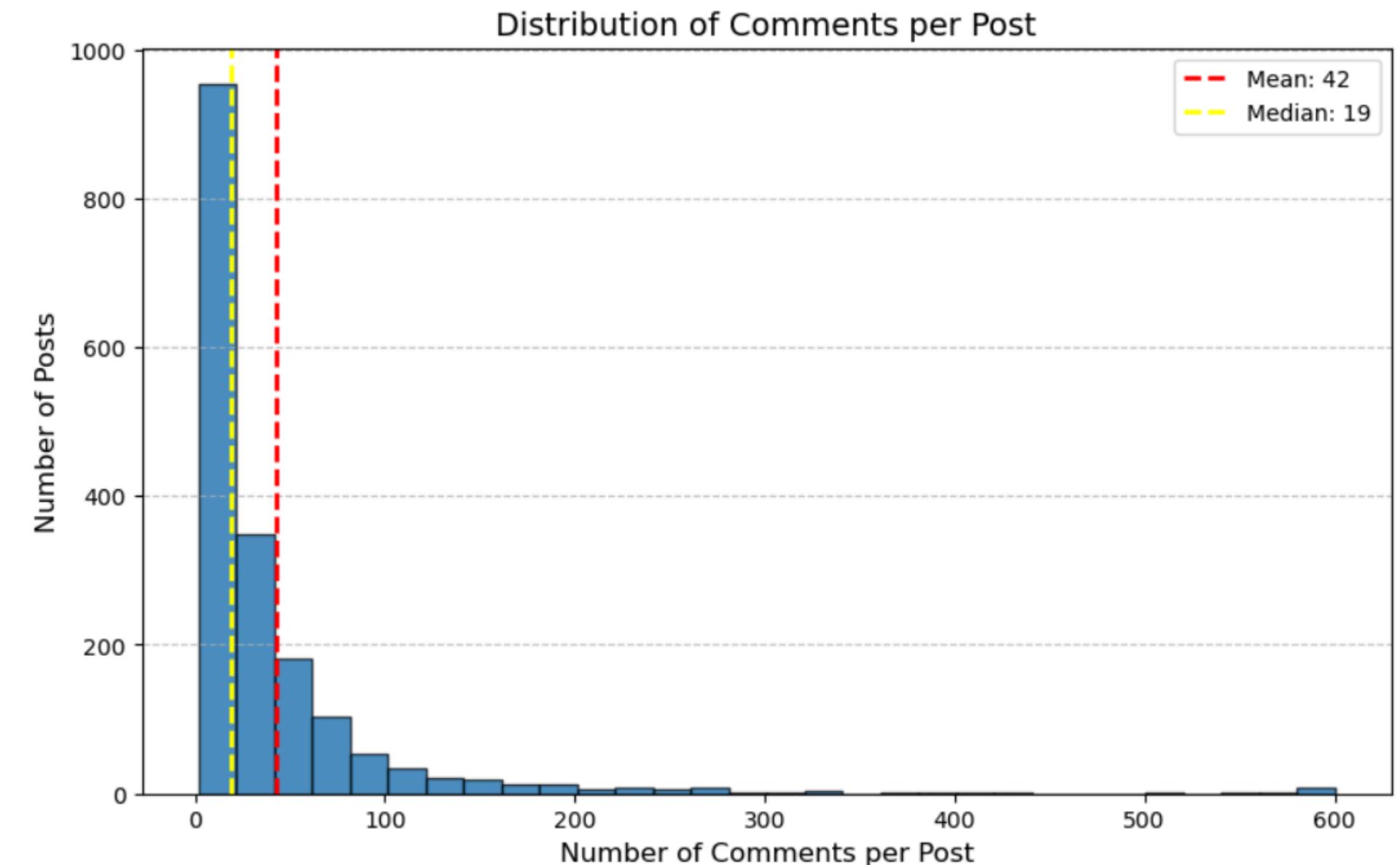
Yearly Trends in Instagram Comments

- Comments show significant growth in 2023.
- Engagement spikes suggest promotional campaigns or viral discussions.
- Customer engagement is rising, highlighting the importance of social media discussions.



Distribution of Comments per Post

- Most posts get few comments; some go viral.
- Engagement is highly skewed.
- A few posts drive most interactions.
- Optimizing viral content can boost engagement.



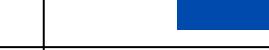
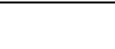
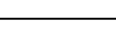
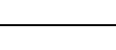
Project Timeline

Progress:

- Data from 4+ Sources
- NLP Preprocessing
- Data Structuring & Organization

Next Steps:

- Advanced Topic Modeling with Sentiment Analysis model
- AI-powered Chatbot
- Interactive Dashboard and a knowledge graph.

PROCESS	QUARTER 1			QUARTER 2		
	Jan	Feb	Mar	Apr	May	Jun
Project kickoff & setup						
Project Definition and Role Assignment						
Technical Setup and Resource Identification						
Data collection & preprocessing complete						
Data Extraction and Storage						
Data Cleaning and Segmentation						
Sentiment analysis & knowledge graph						
Sentiment Analysis Implementation						
Knowledge Graph Development & Ranking						
Dashboard development & testing						
Dashboard Design and Integration						
Stakeholder Testing and Refinements						
Summarization chatbot prototype ready						
Chatbot Development and Training						
Integration and Internal Testing						
Final deliverables & presentation						
Final Deliverable Compilation						
Presentation and Stakeholder Feedback						

Conclusion

Benefits of our proposed solution:

- Extracts Actionable Insights
- AI-Driven Automation
- Knowledge graph enhances marketing decisions
- Scalable & Reusable Framework



Acknowledgements

We would like to express our sincere gratitude to those who have guided and supported us throughout this project.

- **Prakash Bhandi (Liaison)**
- **Xuan Zhang (Liaison)**
- **Costco Wholesale (Sponsor Organization)**
- **Dr. Elio Zhang (Faculty Advisor)**
- **Project Center Team (Seattle University)**
- **Dr. Ariana Mendible (Course Advisor)**

Finally, we extend our appreciation to the audience for your time, engagement, and valuable feedback as we continue refining our work.

Come see us at the Projects Day 2025!



Date: June 6th
Location: Seattle
University Law school

Thank you!!

Any Questions?