

Pre & Post COVID Black Friday Consumer Experience

Text Analytics Final Project



Introduction



GOAL

**Identify the
impacts of
COVID-19 on
customers
perception
of Black
Friday**

KEY QUESTIONS

**What is the general
sentiment towards
Black Friday?**

**How did COVID-19
impact that
sentiment?**

**How sentiment
insights drives value?**

RECOMMENDATION

**Leverage
Sentiment
Analysis as a
strategic aid**

IMPACTS

- 1. Understanding
the trends and
positioning
strategically**
- 2. Identification
of value shift
for customers**
- 3. Keep an eye
on the
competitive
landscape**





Preprocessing and data gathering



**Step
1**

**Gather data
from reddit
dump**

Data from the amazonprime,
blackfriday, bestbuy,
anticonsumption and
walmart subreddits

**Step
2**

**Convert JSON
extract to CSV**

**Step
3**

**Filter only
relevant
comments**

- Filter blackfriday period or posts with blackfriday keyword
- Remove stop words and non-alpha words

**Step
4**

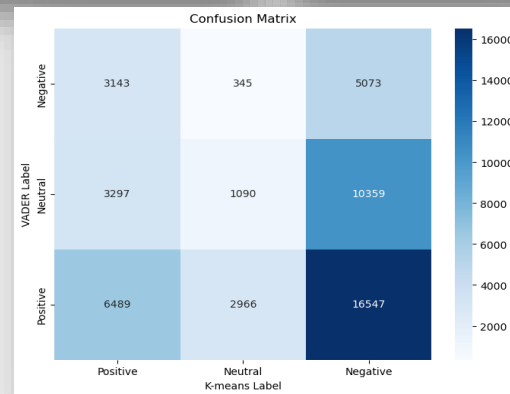
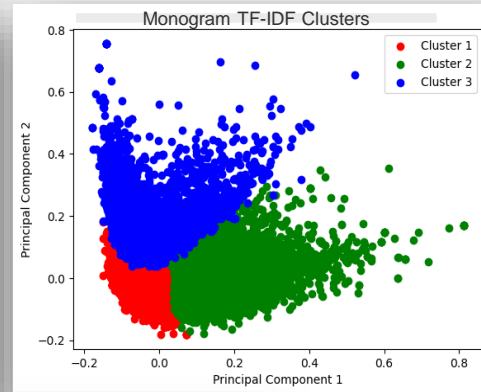
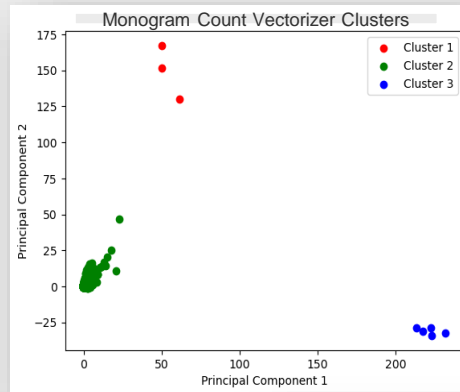
**Preprocessing for
NLP to create 16
vectorized
datasets**

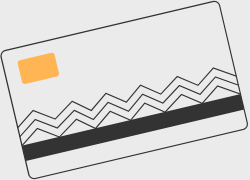
- Lemmatized and Original text
- Monograms, Bigrams, Trigrams and (1,3)grams
- For each CountVectorizer and TFIDF

= 16 datasets

Sentiment Analysis

- Conducted VADER sentiment analysis on the Reddit posts.
- Experimented with K-means clustering on the different combinations of datasets we produced.
- The best-performing K-means clustering used the (1:3) lemmatized TF-IDF dataset, achieving a 42% similarity with VADER sentiment labels.
- 54% of the dissimilarity is attributed to K-Means suggesting the post is negative, while VADER suggests otherwise.
- After a manual validation check, VADER sentiments are the chosen result.
- Sentiment Distribution: 53% Positive, 30% Neutral, and 17% Negative.





Online vs In-Store?

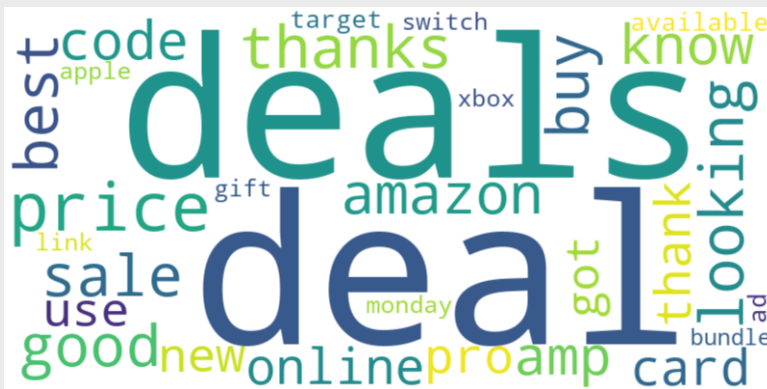


Optimal Preprocessing
steps and number of topics

In-Store
Shopping
47%



Online
Shopping
53%



Data
preprocessing

Number of
topics

Evaluation

TF-IDF(40%
Max-df, 2 min-
df,
unlemmatized
with Two topics

- Vectorization:
Count vectorizer
and TF-IDF
- Tokenization:
Unigrams,
bigrams, and 1,3
- Lemmatization
- Stop words
removal

- Model Perplexity
- Visual inspection
- Topic distribution

- Visual inspection
- Topic distribution

Blackfriday, www,
http, https, black,
Friday, com

Is Black Friday Checking Out?

5 years tendency	Features	Coefficients	Interpretation
	In Store	-0.138068	Decrease
	Online	0.141088	Increase
	Anticonsumption	-0.105235	Decrease
	Amazon Prime	0.017945	Increase
	Best Buy	0.040561	Increase
	Walmart	-0.112774	Decrease

Is Black Friday Checking Out?

5 years tendency	Features	Coefficients	Interpretation
	In Store	-0.138068	Decrease
	Online	0.141088	Increase
	Anticonsumption	-0.105235	Decrease
	Amazon Prime	0.017945	Increase
	Best Buy	0.040561	Increase
	Walmart	-0.112774	Decrease

Cyber Monday is Taking over Black Friday

Getting the Cat Out of the Bag

COVID EFFECT STUDY

Hypotheses

Feature	Coefficient BEFORE COVID-19	Hypothesis Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24963	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.26945	Decrease	-0.05509	Decrease
subreddit_amazonprime	0.24879	Increase	-0.21033	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

Getting the Cat Out of the Bag

Hypotheses

Black Friday is becoming an online event

Feature	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24963	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.26945	Decrease	-0.05509	Decrease
subreddit_amazonprime	0.24879	Increase	-0.21033	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

Getting the Cat Out of the Bag

COVID EFFECT STUDY

Hypotheses

Black Friday is becoming an online event

Buy less, buy better

Feature	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24963	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.05509	Decrease	-0.05509	Decrease
subreddit_amazonprime	0.24879	Increase	-0.21033	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

Getting the Cat Out of the Bag

Hypotheses

Black Friday is becoming an online event

Buy less, buy better

We can suppose competition with Prime Day

Feature	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24963	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.05509	Decrease	-0.05509	Decrease
subreddit_ama	-0.05509	Decrease	-0.05509	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

Getting the Cat Out of the Bag

COVID EFFECT STUDY

Hypotheses

Feature	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24963	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.05509	Decrease	-0.05509	Decrease
subreddit_ama	-0.05509	Decrease	-0.05509	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

Black Friday is becoming an online event

Buy less, buy better

We can suppose competition with Prime Day

Inventory problems often reported

Cat's out the bag

COVID EFFECT STUDY

Feature	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
in_store_shopping_prob	-0.25425	Decrease	-0.23051	Decrease
online_shopping_prob	0.24062	Increase	0.24718	Increase
subreddit_Anticonsumption	-0.26945	Decrease	-0.05509	Decrease
subreddit_amazonprime	0.24879	Increase	-0.21033	Decrease
subreddit_bestbuy	0.06539	Increase	0.09920	Increase
subreddit_walmart	-0.20013	Decrease	-0.13808	Decrease

How can we benefit from these hypotheses?

Implications

Tailoring Business Strategies

- Utilize sentiment analysis to refine marketing, inventory, and customer service

Economic and Policy Insights

- Analyze changes in consumer behavior to guide economic policies and stimulate shopping period activities

Societal and Cultural Trends

- Examine shifts towards sustainability and online shopping, impacting public policy and consumer attitudes

Understanding Market Dynamics

- Explore the impact of global events like COVID-19 on Black Friday sentiments, aiding in future planning and sustainability discussions.
-

Delivering a Better Black Friday Experience

Cyber Monday

Difference between the two event is fading

Consumer Frugality

Consumer shop and compare online, then buy at Black Friday

Consumer's influence

Rise in engagement with others when browsing... opinion matters!



Online over in-person...

Increasing quantity of good are bought online

... but not in all scenarios

Consumer's preference for online vs in-person purchase varies based on the goods

GOAL

**Identify the
impacts of
COVID-19 on
customers
perception
of Black
Friday**

KEY QUESTIONS

**What is the general
sentiment towards
Black Friday?**

**How did COVID-19
impact that
sentiment?**

**How sentiment
insights drives value?**

RECOMMENDATION

**Leverage
Sentiment
Analysis as a
strategic aid**

IMPACTS

- 1. Understanding
the trends and
positioning
strategically**
- 2. Identification
of value shift
for customers**
- 3. Keep an eye
on the
competitive
landscape**



APPENDIX



Logistic regression coefficient

Feature	Coefficient Overall	Coefficient BEFORE COVID-19	Interpretation Set 1	Coefficient AFTER COVID-19	Interpretation Set 2
score	-0.324609	0.01974	Increase	-0.75727	Decrease
<u>in_store_shopping_prob</u>	-0.138068	-0.25425	Decrease	-0.23051	Decrease
<u>online_shopping_prob</u>	0.141088	0.24963	Increase	0.24718	Increase
<u>Year_var</u>	0.021753	0.00000	No effect	0.00000	No effect
<u>subreddit_Anticonsumption</u>	-0.105235	-0.26945	Decrease	-0.05509	Decrease
<u>subreddit_amazonprime</u>	0.017945	0.24879	Increase	-0.21033	Decrease
<u>subreddit_bestbuy</u>	0.040561	0.06539	Increase	0.09920	Increase
<u>subreddit_blackfriday</u>	0.162523	0.15258	Increase	0.32097	Increase
<u>subreddit_walmart</u>	-0.112774	-0.20013	Decrease	-0.13808	Decrease

Model Performance

Accuracy: 0.8609556607834696

Precision: 0.868834287825605

Recall: 0.9478033051189036 F1

Score: 0.9066024096385542
