The Price of Zero: Public Sentiment on Amazon's 2018 Federal Tax Avoidance

Ali Al-Rikabi   10166070
Daniela Mañozca Cruz   30262558
David Fakolujo   30273636
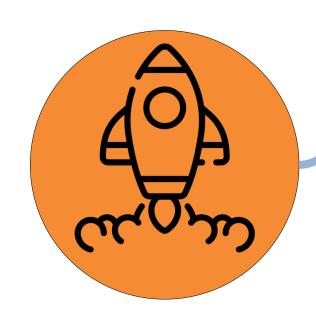Luisa Alejandra Sierra Guerra   30261956

Table of Contents

✓History, Scandal & Relevance

✓Objectives

✓Guiding Questions

✓Analysis Roadmap

✓Visualizations and Explanations

✓Key Findings

✓Ethics and Recommendations

amazon History

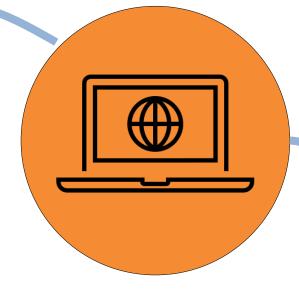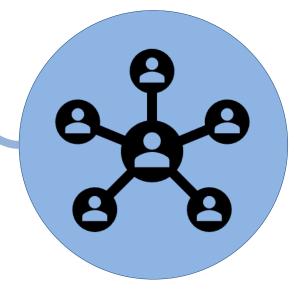Founded in a garage in 1994 as "Cadabra".

Amazon began as an online bookstore delivering across the U.S. and 45 countries.

Amazon emerged during the dot-com boom and quickly evolved beyond books.

By 1999, it introduced third-party sellers

As e-commerce took off globally, Amazon helped define it. With over 200 million Prime members and 2.14 billion online shoppers worldwide, Amazon now leads the sector, setting the standard for digital retail, innovation, and customer experience

UNIVERSITY OF CALGARY

1

The 2018 Amazon Tax Scandal

+$100 Billion

Since 2014, Amazon annual revenue has exceeded $100 billion, reaching $637 billion in 2024. Naturally, many assume that such large corporations pay substantial amounts in taxes.

Accounting strategy

Tax liability isn't based on revenue alone. Depending on a company's financial history and accounting strategy, businesses can carry forward past losses and apply them to reduce taxable income in future years.

$232.9 Billion

In 2018, Amazon used this mechanism, along with federal tax policy changes, to reduce its federal income tax liability to zero, despite generating over $232.9 billion in revenue that year.

$11.2 Billion

In 2018, Amazon reported $11.2 billion in U.S. profits—yet paid zero dollars in federal income tax. This sparked widespread public outrage and media scrutiny.

$129 Million

Despite being one of the most profitable companies in the world, Amazon not only avoided paying federal income tax but received a $129 million tax rebate.

Objectives

Identify the primary topics associated with Amazon's tax scandal.

Evaluate the overall sentiment polarity and subjectivity of the dataset to assess the general tone of public opinion.

Explore which categories of viewer comments generated the highest engagement, based on the volume of responses per topic.

Determine how specific categories of comments contribute to shaping consumer attitudes toward Amazon

Map the structure and interconnections among key discussion topics to uncover the dynamics behind audience sentiment.

Guiding Questions

What are the primary topics discussed in the YouTube comments related to Amazon's $0 tax payment in 2018?

What is the overall public sentiment toward Amazon's $0 federal tax payment in 2018? Positive, negative, neutral?.

Which three comment categories or themes generated the highest levels of engagement (replies frequency)?
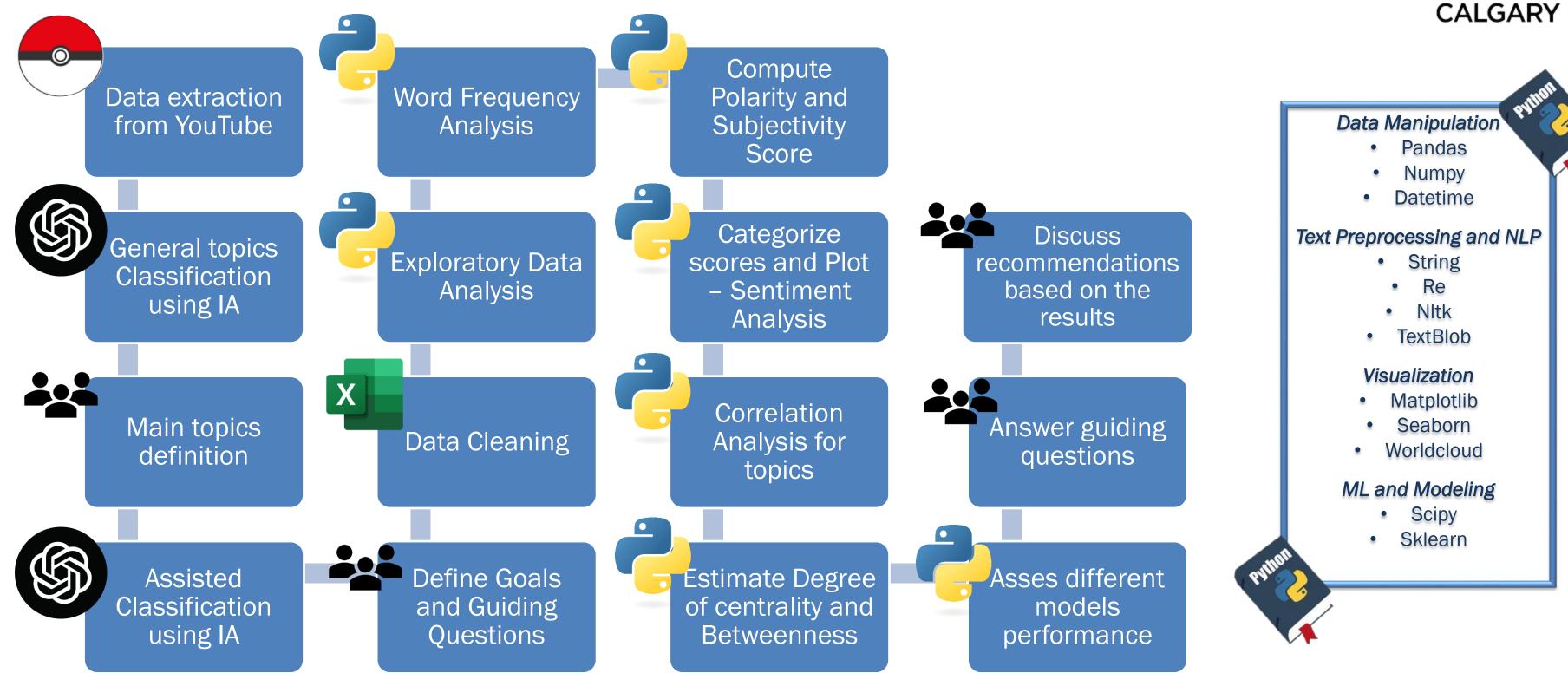
Do certain types of comments appear to influence consumer perception or attitudes toward Amazon? If so, how?

Are there observable connections or patterns among the most frequently discussed topics?
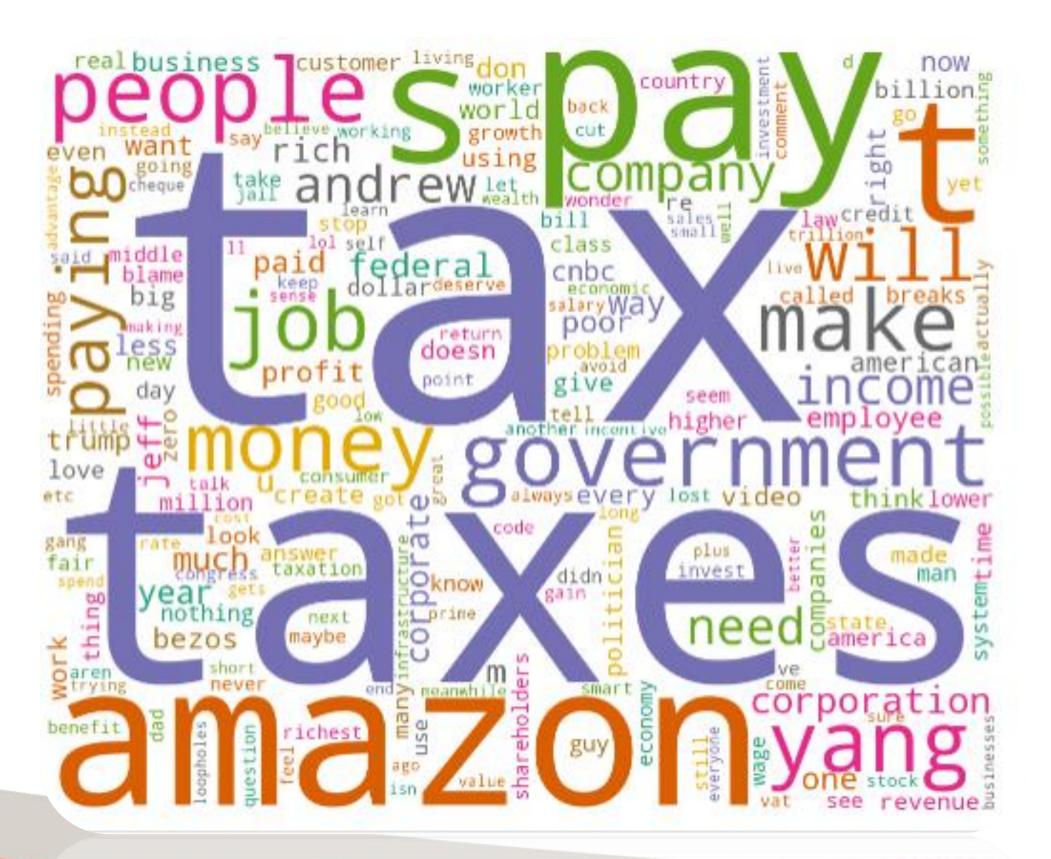
Analysis Roadmap

Data extraction from YouTube

Word Frequency Analysis

Compute Polarity and Subjectivity Score

General topics Classification using IA

Exploratory Data Analysis

Categorize scores and Plot – Sentiment Analysis

Discuss recommendations based on the results

Main topics definition

Data Cleaning

Correlation Analysis for topics

Answer guiding questions

Assisted Classification using IA

Define Goals and Guiding Questions

Estimate Degree of centrality and Betweenness

Asses different models performance

Data Manipulation
- Pandas
- Numpy
- Datetime

Text Preprocessing and NLP
- String
- Re
- Nltk
- TextBlob

Visualization
- Matplotlib
- Seaborn
- Worldcloud

ML and Modeling
- Scipy
- Sklearn

UNIVERSITY OF CALGARY

What are the primary topics discussed in the YouTube comments related to Amazon's $0 tax payment in 2018?
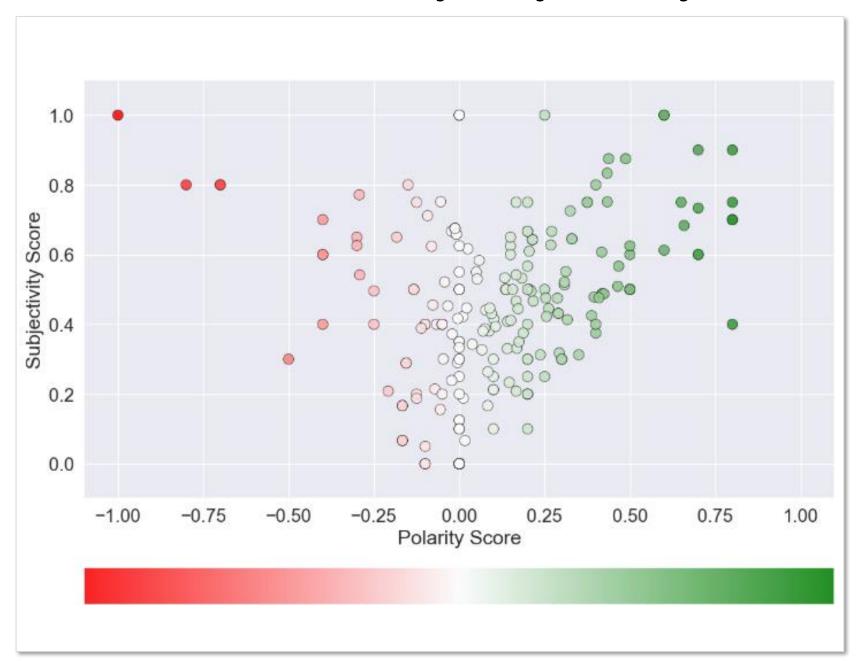

- ✓ **High frequency of Tax-Related Terms**: The most prominent words are "tax", "taxes" ,"pay".
- ✓ **Strong Focus on Amazon**: The word "Amazon" appears centrally, reinforcing that public discourse was intensely centered on the company's actions.
- ✓ **Public Sentiment around Fairness**: Words like "people", "job", "money", and "corporation" suggest that commenters are drawing contrasts between everyday taxpayers and large corporations.
- ✓ **Political and System criticism**: Mentions of "government", "federal", and "yang" indicate that the conversation extended into political territory.
- ✓ **Calls for action**: Terms such as "need", "make", "change", and "breaks" imply not just criticism, but a desire for change in tax policy or corporate responsibility.
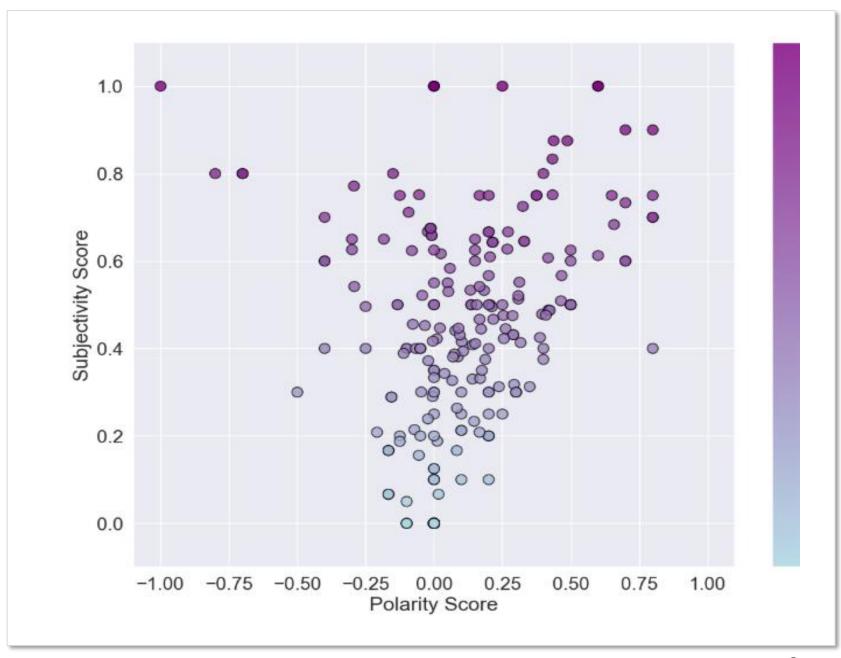
What is the overall public sentiment toward Amazon's $0 federal tax payment in 2018? positive, negative, neutral?
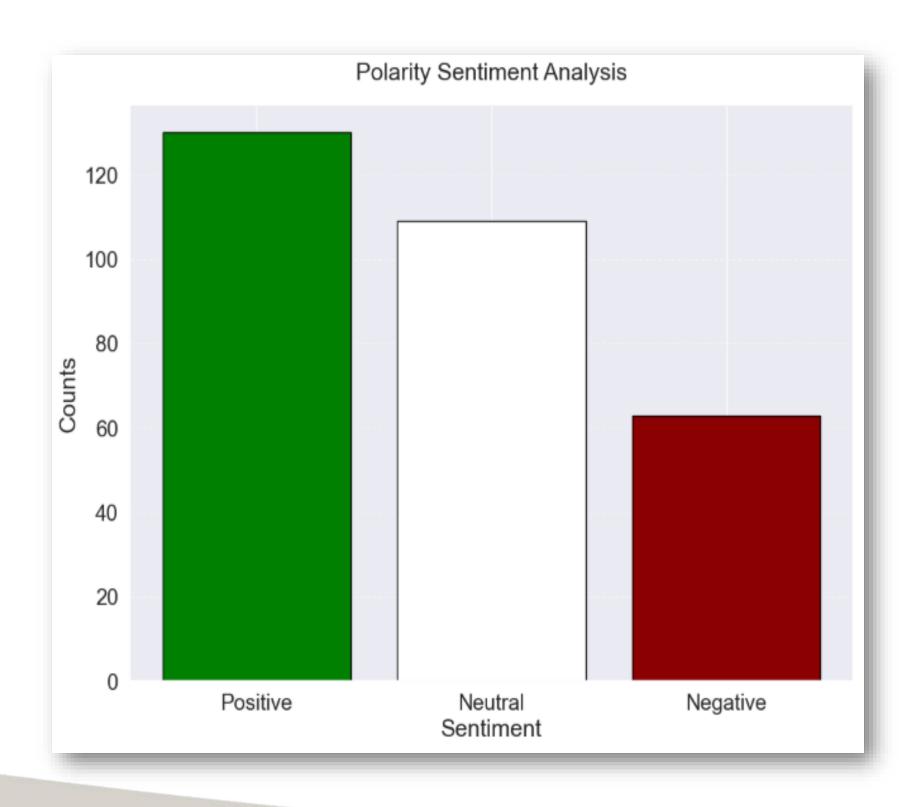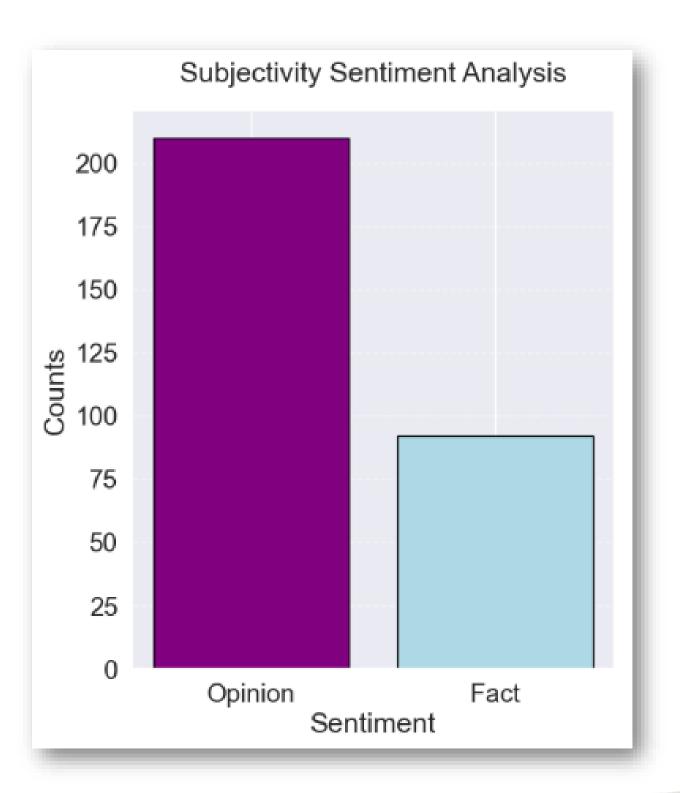

Sentiment Analysis by Polarity

Sentiment Analysis by Subjectivity

What is the overall public sentiment toward Amazon's $0 federal tax payment in 2018? positive, negative, neutral? (Cont.)

Do certain types of comments appear to influence consumer perception or attitudes toward Amazon?

Sentiment Analysis - Sarcasm or Critical Humor

Sentiment Analysis - General Institutional Distrust

Sentiment Analysis - Criticism of the Tax System
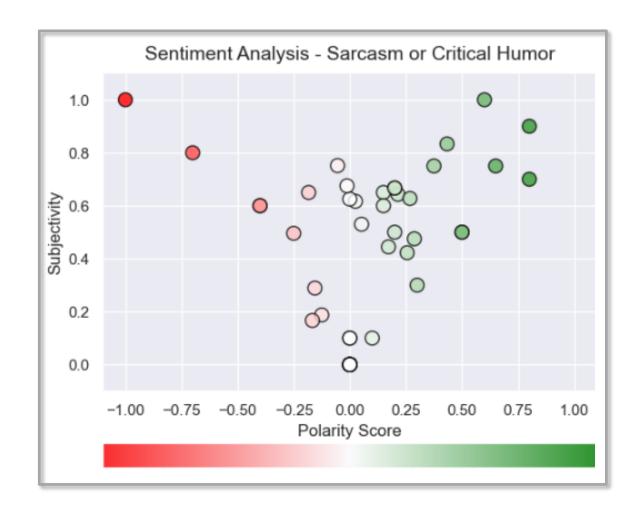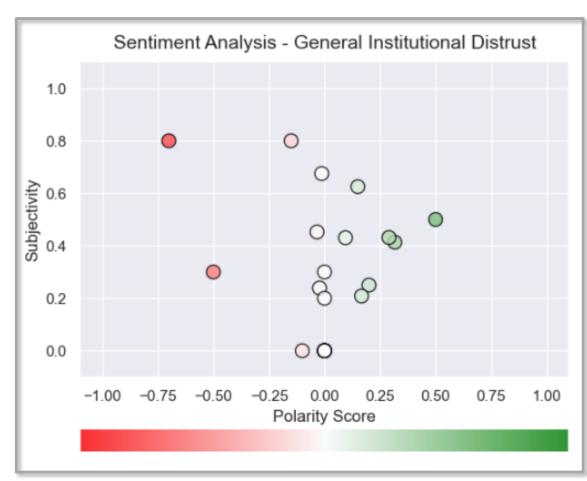
*Although the three plots presented correspond to topic categories with highest frequency of negative polarity scores, most comments shows neutral or slightly positive polarity, **even in sarcastic or critical contexts**. This likely **reflects the algorithm's inability to detect implied negativity when emotional tone is subtle or ironic.***

Do certain types of comments appear to influence consumer perception or attitudes toward Amazon? (Cont)
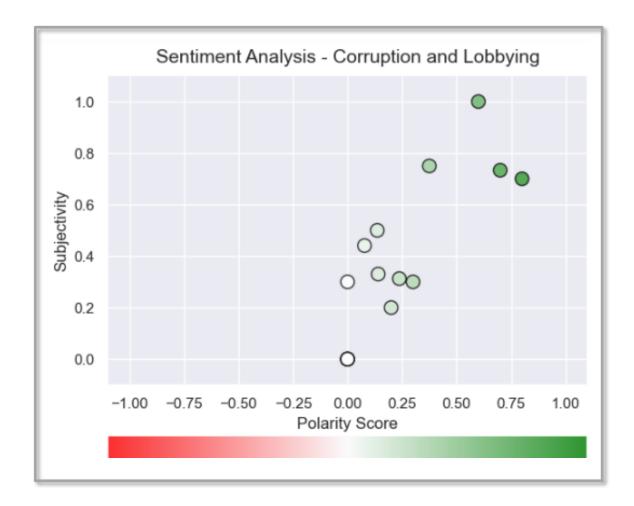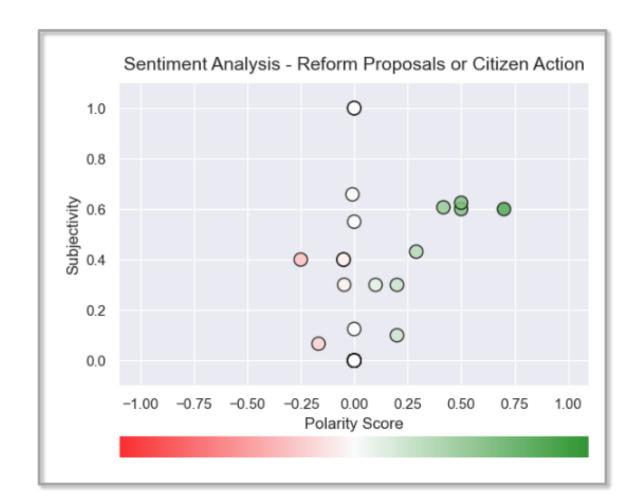
Comments labeled as "Moral or Ethical Outrage" appear overwhelmingly positive in polarity, contradicting the tone of critique seen in raw comments. This reinforces how **emotionally neutral wording masks discontent, leading to misclassification.**

Do certain types of comments appear to influence consumer perception or attitudes toward Amazon? (Cont)
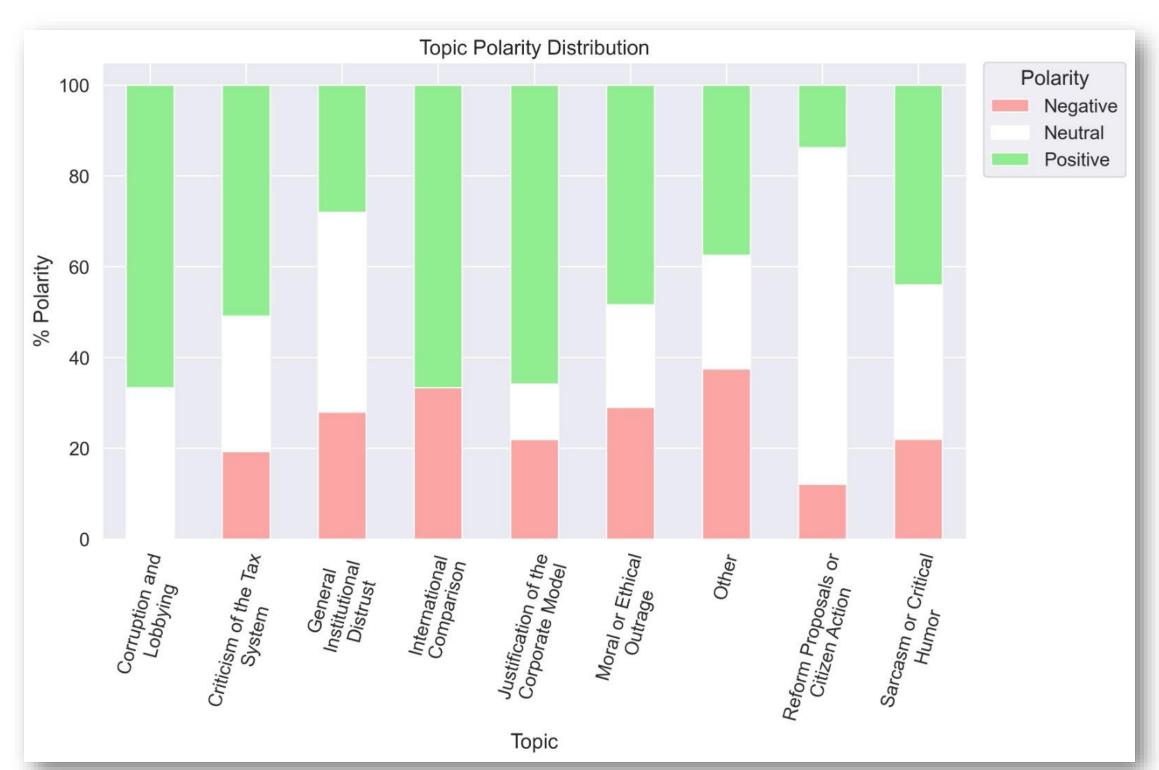
Topics such as "Corruption and Lobbying", which naturally suggest disapproval, are misclassified as positive or neutral, as **the word "corruption" is not treated as emotionally charged unless paired with explicit sentiment terms.**

# Do certain types of comments appear to influence consumer perception or attitudes toward Amazon? (Cont)

Topic Polarity Distribution

- ✓ **Most topics show high levels of "positive" polarity**, even when their actual tone ("Corruption" or "Moral Outrage" comments) should lean negative.
- ✓ **"General Institutional Distrust" and "Other" have the highest share of negative polarity**, suggesting stronger emotional reactions.
- ✓ **"Reform Proposals" and "Justification of the Corporate Model" skew strongly positive**, likely because their wording leans more constructive or neutral.
- ✓ Calls for action: Terms such as "need", "make", "change", and "breaks" imply not just criticism, but a **desire for change in tax policy or corporate responsibility**.

Are there observable connections or patterns among the most frequently discussed topics?

✓ Strong positive correlations exist between Criticism of the Tax System, Sarcasm or Critical Humor, and Corruption and Lobbying.

✓ General Institutional Distrust shows strong negative correlations with most topics (especially International Comparison and Other at -0.9 to -1.0),

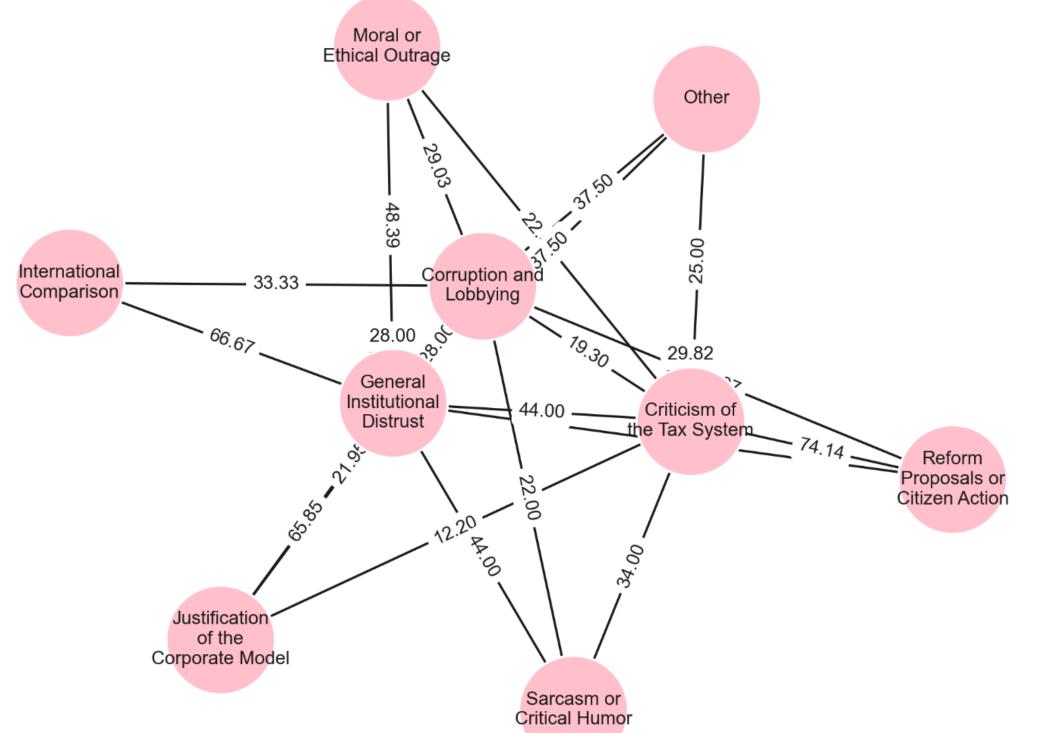✓ Justification of the Corporate Model correlates well with many topics, suggesting it often co-occurs in conversations spanning different sentiment types.

This matrix reveals that while many topics share similar sentiment distributions, some stand apart, likely due to phrasing that conceals tone, confusing automated sentiment tools.

Correlation Matrix of Topic Polarity Distributions

	Corruption and Lobbying	Criticism of the Tax System	General Institutional Distrust	International Comparison	Justification of the Corporate Model	Moral or Ethical Outrage	Other	Reform Proposals or Citizen Action	Sarcasm or Critical Humor
Corruption and Lobbying	1.0	1.0	0.0	0.5	0.8	0.7	-0.0	0.0	1.0
Criticism of the Tax System	1.0	1.0	-0.2	0.7	0.9	0.8	0.2	-0.2	1.0
General Institutional Distrust	0.0	-0.2	1.0	-0.9	-0.6	-0.7	-1.0	1.0	0.1
International Comparison	0.5	0.7	-0.9	1.0	0.9	1.0	0.9	-0.9	0.5
Justification of the Corporate Model	0.8	0.9	-0.6	0.9	1.0	1.0	0.6	-0.6	0.7
Moral or Ethical Outrage	0.7	0.8	-0.7	1.0	1.0	1.0	0.7	-0.7	0.7
Other	-0.0	0.2	-1.0	0.9	0.6	0.7	1.0	-1.0	-0.1
Reform Proposals or Citizen Action	0.0	-0.2	1.0	-0.9	-0.6	-0.7	-1.0	1.0	0.1
Sarcasm or Critical Humor	1.0	1.0	0.1	0.5	0.7	0.7	-0.1	0.1	1.0

✓ **Criticism of the Tax System is the most central node in the network**, connected to the largest number of other topics.

✓ **Corruption and Lobbying is another high-betweenness topic**, acting as a connector between polarizing or less-connected themes such as Moral Outrage, Institutional Distrust, and Other.

✓ **Topics like International Comparison and Justification of the Corporate Model appear more peripheral**, suggesting they are part of more specialized or isolated comment threads rather than core discourse.

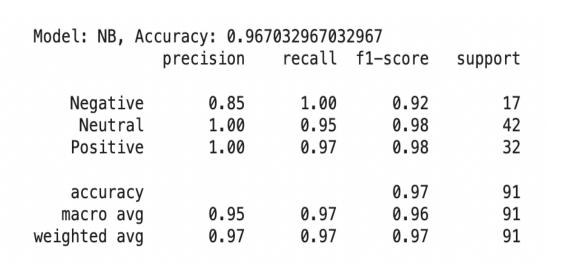# Key Findings – Predicting Polarity Analysis

```
Model: LR, Accuracy: 0.8351648351648352
Model: KNN, Accuracy: 0.8241758241758241
✅ Model: NB, Accuracy: 0.967032967032967
Model: SVC, Accuracy: 0.5604395604395604
Model: RFC, Accuracy: 1.0
Model: DTC, Accuracy: 1.0
Model: GBC, Accuracy: 1.0
```
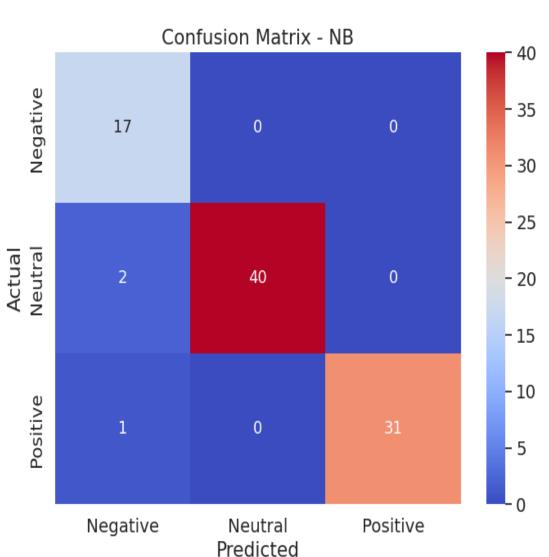
```
Model: NB, Accuracy: 0.967032967032967
              precision    recall  f1-score   support

    Negative       0.85      1.00      0.92        17
     Neutral       1.00      0.95      0.98        42
    Positive       1.00      0.97      0.98        32

    accuracy                           0.97        91
   macro avg       0.95      0.97      0.96        91
weighted avg       0.97      0.97      0.97        91
```


Confusion Matrix - NB

✓ A classification task was performed to predict sentiment polarity (Negative, Neutral, Positive) using **Polarity**, **Subjectivity**, and **Topic** as predictors.

✓ Multiple classification models were tested, and **Naive Bayes** was selected as the preferred model due to its high accuracy and generalization without overfitting.

✓ The Naive Bayes model was further evaluated using a **classification report** and **confusion matrix**, showing strong performance across **precision**, **recall**, and **F1 score**.

Ethics and Recommendations

1. Enhance Transparency in Tax Practices

Action:
- Release detailed reports on tax strategies and market contributions.

Rationale:
- Address criticism on fairness and corporate tax responsibility.
- Build trust by demystifying tax practices.

2. Engage in Corporate Responsibility Initiatives

Action:
- Increase investments in community and social programs.
- Focus on areas benefiting from tax savings.

Rationale:
- Counteract perceptions of inequality.
- Provide tangible societal benefits.

3. Engage in Corporate Responsibility Initiatives

Action:
- Participate in tax reform discussions and policy-making.
- Support equitable tax systems.

Rationale:
- Align with public desires for change.
- Demonstrate commitment to fair play.

4. Improve Communication and Public Relations

Action:
- Launch a campaign highlighting contributions beyond taxes.
- Focus on job creation and economic impact.

Rationale:
- Address need for better understanding of Amazon's role.
- Reshape perceptions to emphasize positive impacts.

References

- Our World in Data. (n.d.). Our World in Data. Retrieved March 2, 2025, from https://ourworldindata.org/

- World Health Organization. (n.d.). World Health Organization. Retrieved March 2, 2025, from https://www.who.int/

- World Bank. (n.d.). World Bank Open Data. Retrieved March 2, 2025, from https://data.worldbank.org/

- United Nations Development Programme. (n.d.). United Nations Development Programme. Retrieved March 2, 2025, from https://www.undp.org/

- Institute for Economics & Peace. Global Peace Index March 2, 2024. https://www.visionofhumanity.org/maps/#/

THANK YOU!