

Analyzing Media Coverage and Public Reaction to the Epstein File: Insights from X Data

TxT Group

Rivaldi

Akbar Kanugraha

Indira Faisa Afgani

Mutiara Ayu Alzahra



List of Contents

This section outlines the structure of the report, guiding readers from the data overview and sentiment analysis to the key findings and recommendations based on X data.



Overview



Data Understanding



Sentiment Analysis



Conclusions & Recommendations



Overview

Project Context

This analysis was conducted to support the **Editorial Research and Strategy team** of *The American Times*. The team aims to understand public reactions toward the issue in order to develop reporting that better reflects public concerns and ongoing discussions.



Why?

Global Reaction

The Epstein files sparked global media coverage and public debate

Accountability

Questions around accountability and institutional transparency made it highly controversial

X = Public Square

X (Twitter) became a key platform for real-time public reaction

Data = Public Perspective

Analyzing this data reveals how the public truly perceives the case



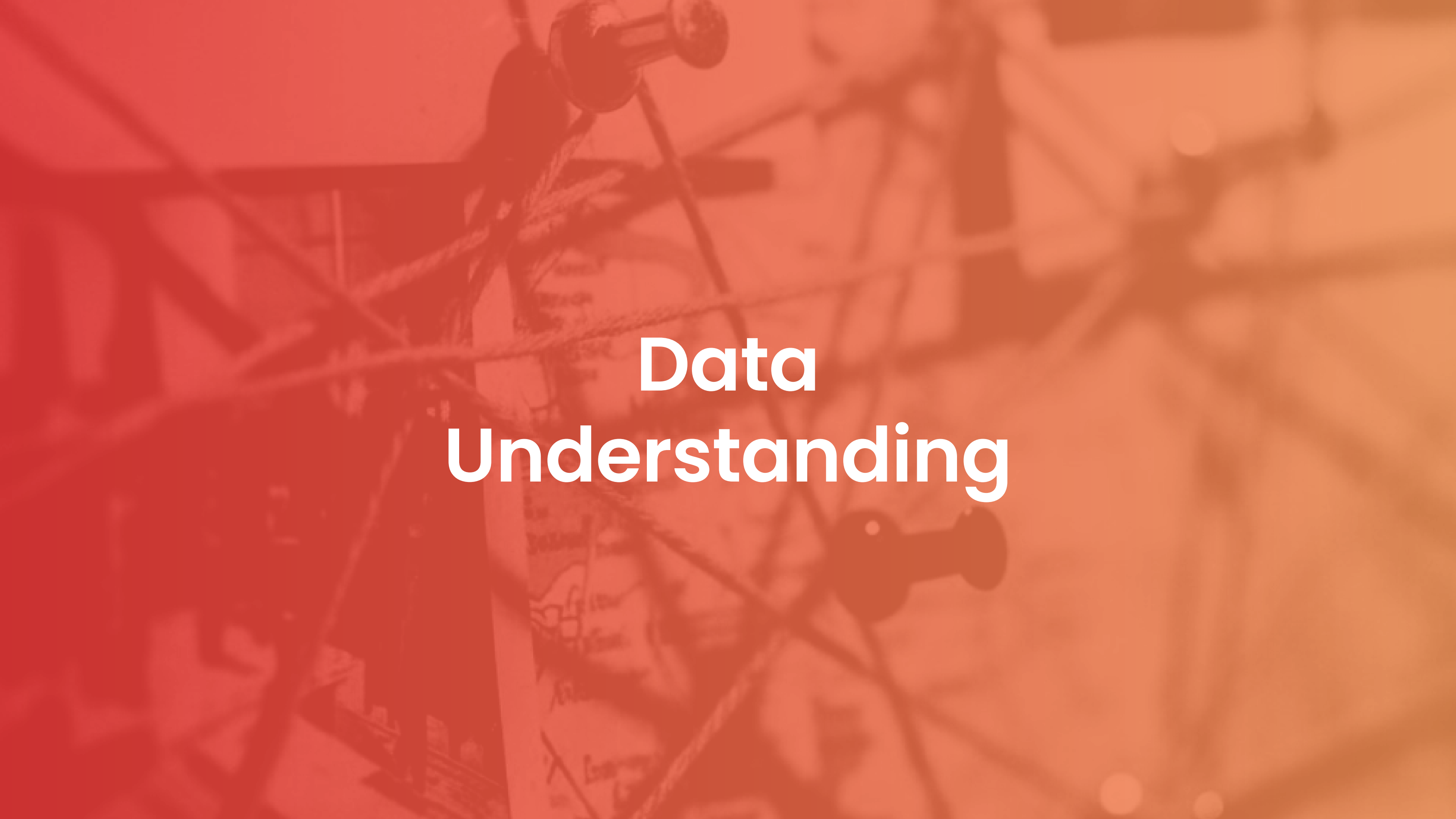
What?

01

How does the public respond to media coverage of the Epstein case on X?

02

What key concerns and discussion themes should the Editorial Research team highlight to ensure media coverage better represents public perspectives?

The background of the slide features a blurred image of a map with several pushpins. A semi-transparent red overlay covers the entire image, creating a warm, monochromatic aesthetic. The text is centered in a bold, white, sans-serif font.

Data Understanding

Twitter Post Scraping

Auth Token Setup

A Twitter account was set up
to access the API

Scraping With TweetHarvest

Used Python tool to extract
relevant posts from X

Created by Helmi Satria

Storing The Data(Raw)

Collected posts saved as .csv
and uploaded to Google Cloud

Data Understanding

	id_str	created_at	full_text	retweet_count	favorite_count	tweet_url
0	2031520251653860000	2026-03-10 23:59:06+00:00	@sploder6 @DavidNachman3 @59SouthLee @BackupJe...	0	0	https://x.com/undefined/status/203152025165386...
1	2031520170703790000	2026-03-10 23:58:47+00:00	@Israel2252 Not all Jews but all Zionists are ...	0	0	https://x.com/undefined/status/203152017070379...
2	2031519658818540000	2026-03-10 23:56:45+00:00	@Parodyjeffx # I @realDonaldTrump wants the Am...	0	0	https://x.com/undefined/status/203151965881854...
...	...	...	...	...	...	...

The data covers the period from November 1, 2025, to March 12, 2026.

Dataset Breakdown

id_str	Unique identifier for each tweet (Twitter's internal ID)	retweet_count	How many times the tweet was retweeted
created_at	Timestamp of when the tweet was posted	favorite_count	How many likes/hearts the tweet received
full_text	The actual content of the tweet (truncated here with "...")	tweet_url	Direct link to the tweet on X/Twitter

Cleaning & Reprocessing

1

Text Cleaning

Remove mentions, Duplicates, hashtags, URLs, numbers, and punctuation.

2

Case Folding

Convert all text into lowercase.

3

Tokenization

Split text into individual words.

4

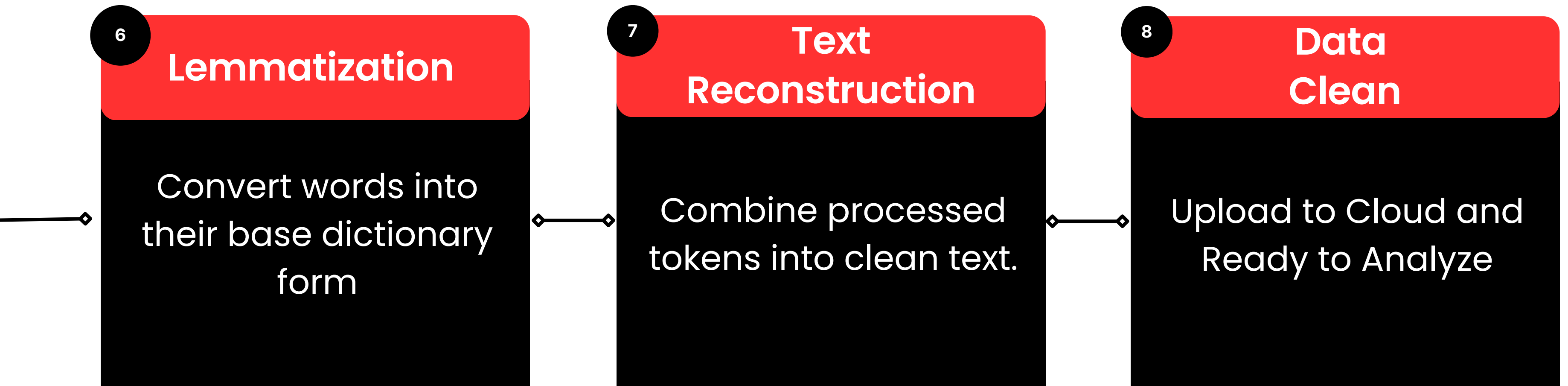
Slang Normalization

Replace informal words with their standard forms.

5

Words Filtering

Remove stop words and dominant primary keywords.



Sentiment Polarity

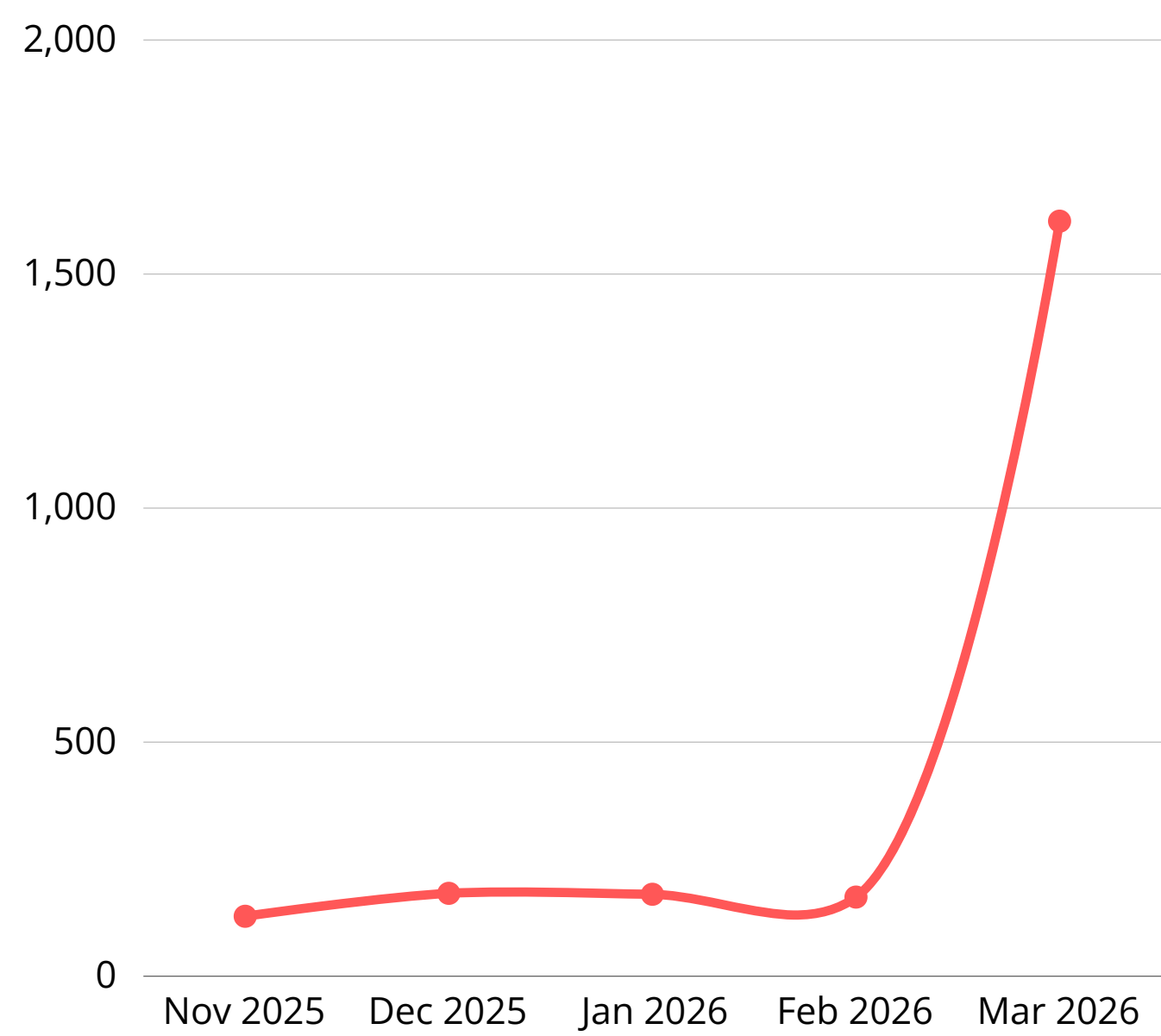


Sentiment Scoring Vader

*"The scoring can detect
up to 7,500 words"*

Scores	Category
Score > 0.1	Positive
$-0.1 \leq \text{Score} \leq 0.1$	Neutral
Score ≤ -0.1	Negative

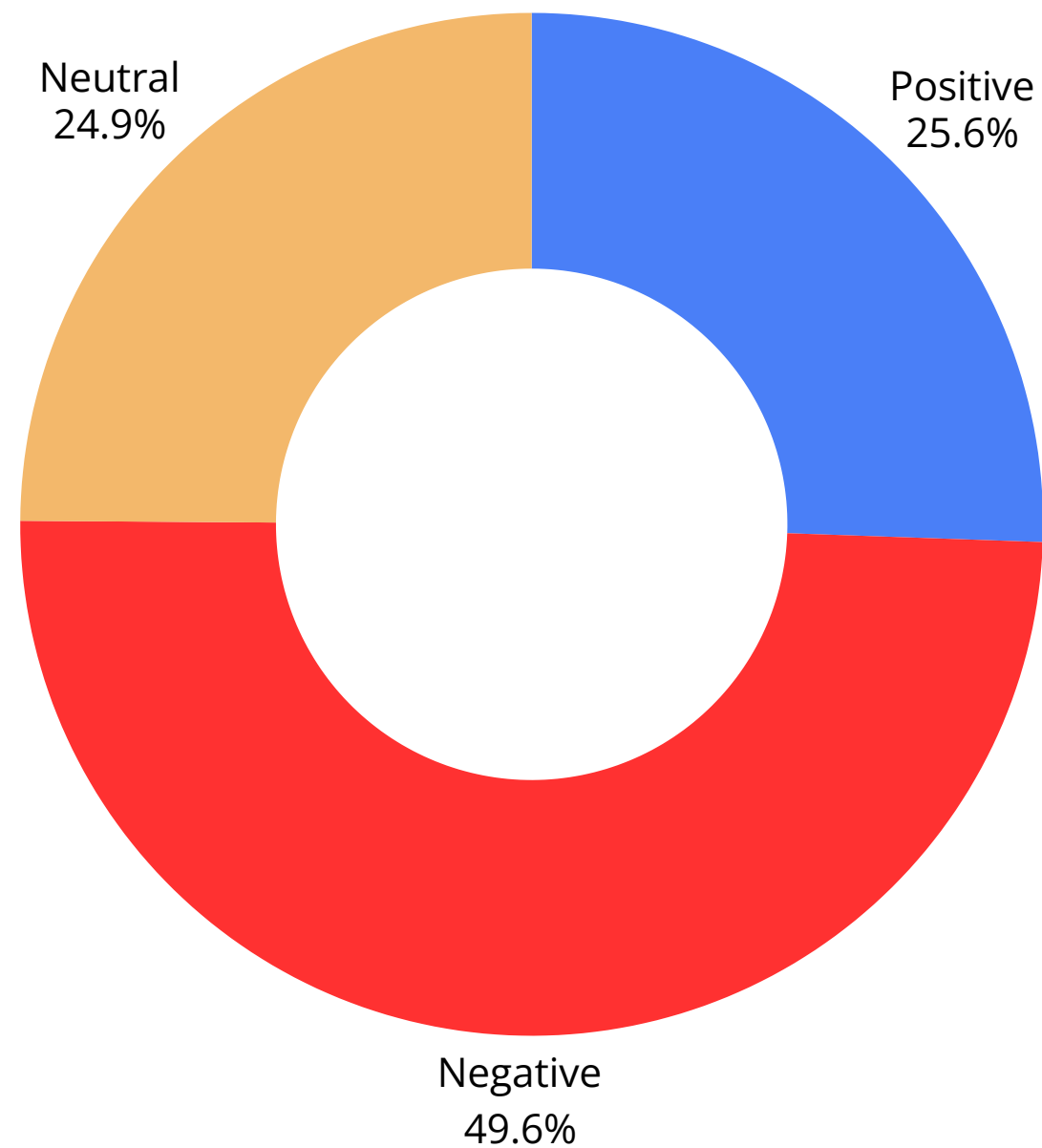
Distribution by Time (EDA)



Summary:

The line chart shows the number of posts over time related to the topic on X. From November 2025 to February 2026, the discussion level remains relatively stable. However, in March 2026, there is a sharp increase in post volume, indicating a surge in public attention likely triggered by new developments or increased media coverage.

Sentiment Distribution

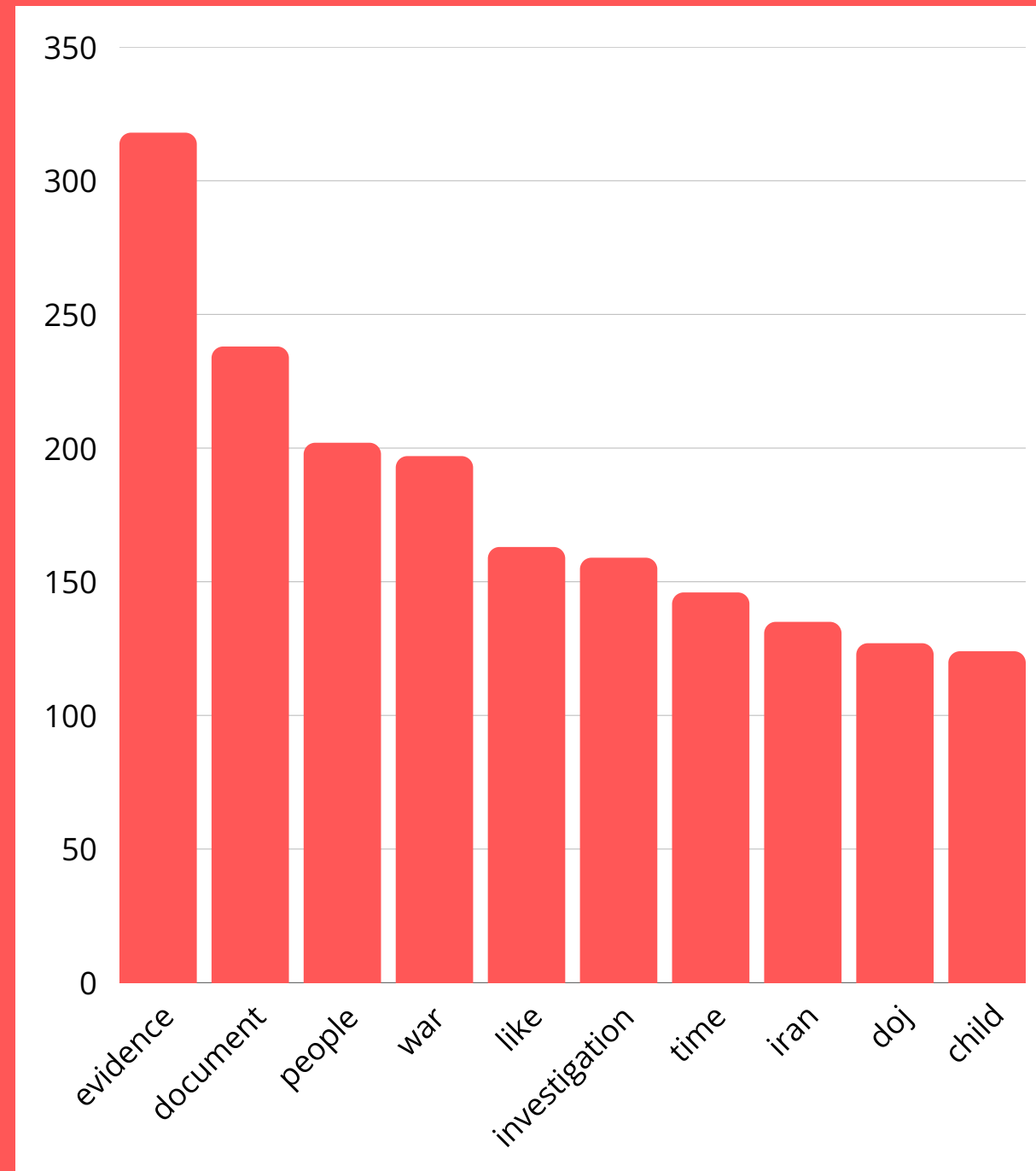


Summary:

Sentiment analysis reveals that nearly half of the posts (49.6%) express **negative sentiment**, reflecting widespread criticism and concern among users. **Positive sentiment** accounts for 25.6%, while 24.9% of posts are neutral, primarily sharing information or news.

Overall, the conversation on X leans notably negative, suggesting that public discourse around the issue is largely driven by criticism and ongoing debates rather than support or neutral exchange.

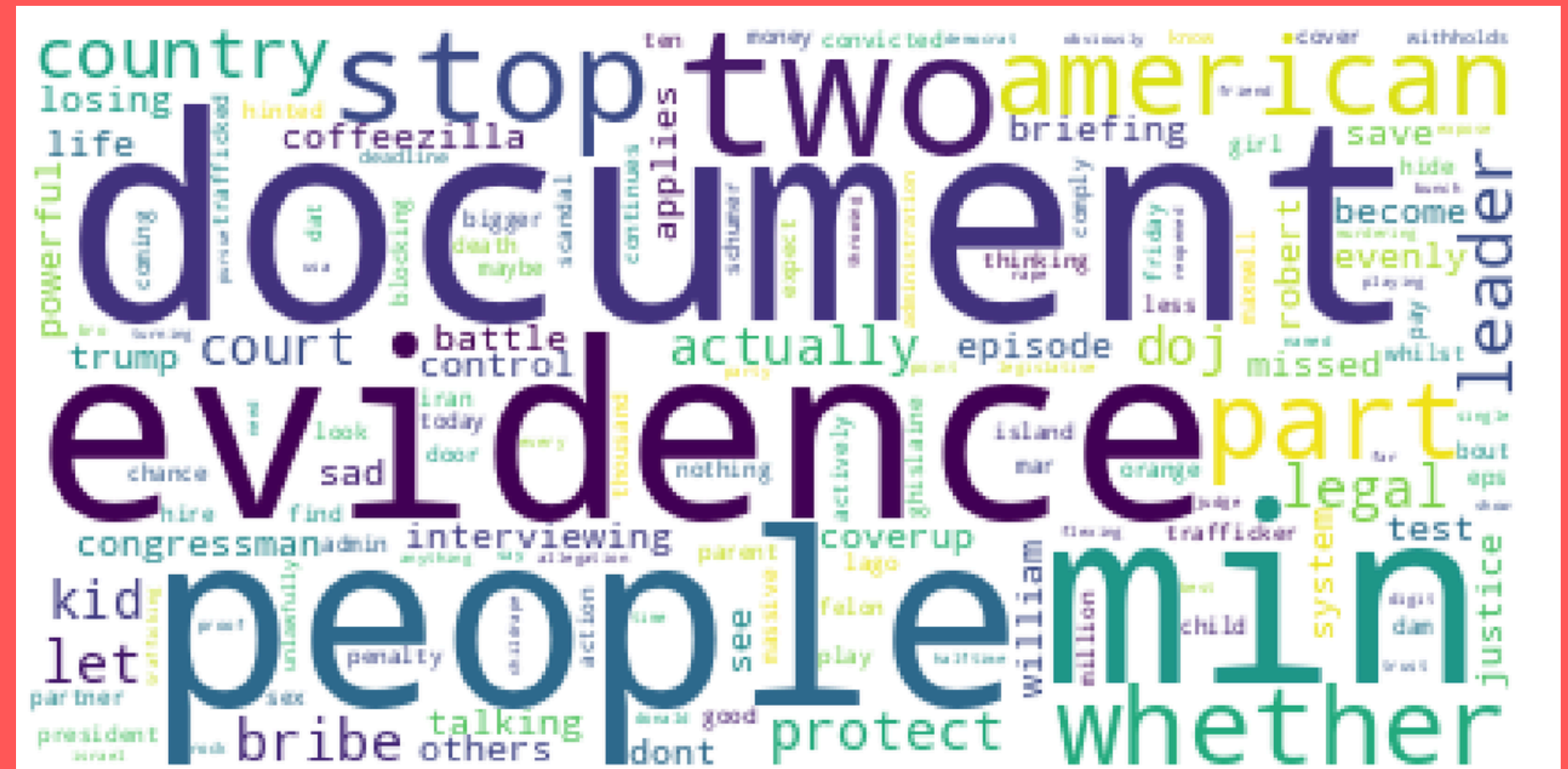
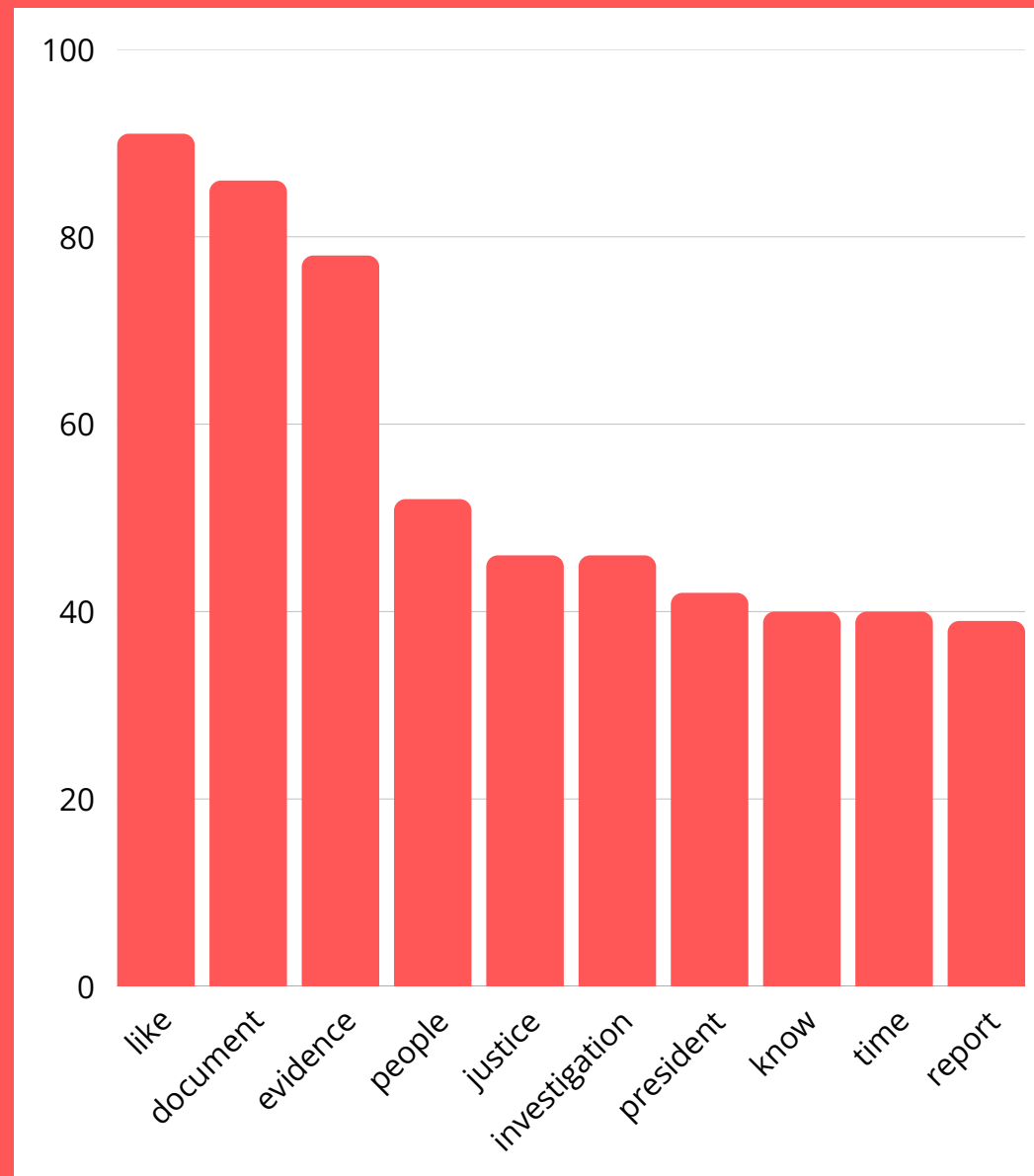
Top 10 Most Common Words



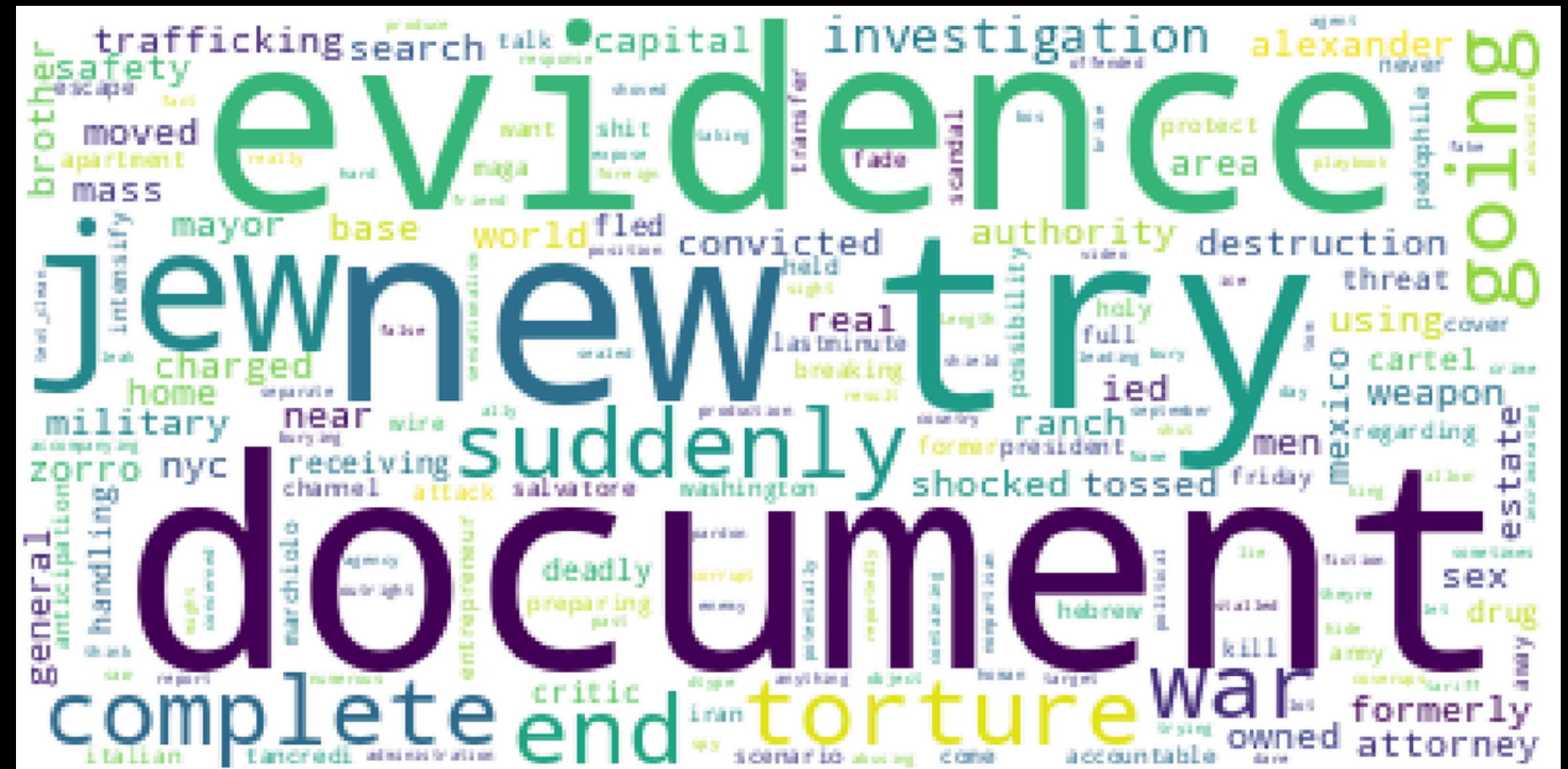
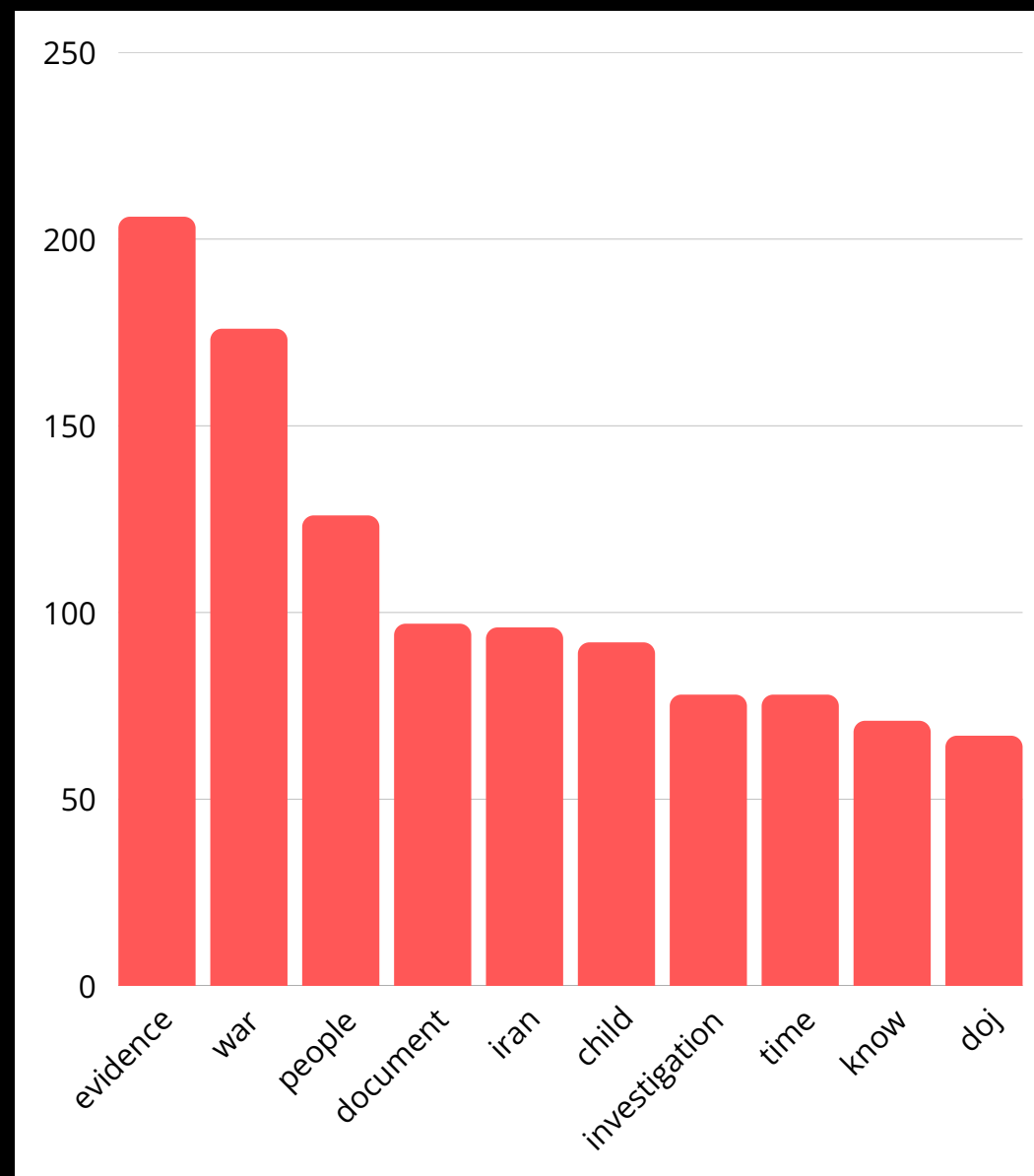
Summary:

The bar chart shows the top ten most frequent words found in X. The word **evidence** appears most frequently, followed by **document**, **people**, and **war**. Other commonly used words, such as **investigation**, **time**, and **DOJ** suggest that discussions often focus on documentation, investigative processes, and institutional involvement. Overall, these keywords highlight the main themes and concerns discussed by users online.

Positive Distribution



Negative Distribution



Comment Sample

My heart is broken. America killed hundreds of children. ~~Trump~~ started this war to cover up the perverted pedophile case. That devil will be gone soon and I believe it. He's the killer of the children who died in Iran just like on Epstein Island.



STOP THE DISTRACTIONS ~~TRUMP~~ WE ARE WAITING FOR THE F**KING EPSTEIN FILES THAT YOU LIED AND SAID YOU ARE GOING TO PROVIDE UNREDACTED NOW! STOP LYING ABOUT SOMOLIANS AND STOP ATTACKING ~~WALZ~~ WITHOUT ANY EVIDENCE! TRAITOR ~~TRUMP~~ NEEDS TO RESIGN NOW HE IS ATTACKING OPPONENTS!!

Negative Comment

Oh THANK YOU Donald I truly mean it! You ll make the Democrats job SO MUCH EASIER. ~~Trump~~ will accidentally gift-wrap a political win for we Dems should he refuse to release the Epstein files because it s an active investigation. If he only wants Democrats investigated



Im happy to report that unlike the president of the United States of America im not a pedophile nor am I in the epstein files like ~~Donald J Trump~~ is



Positive Comment

Conclusion & Recommendations

Sentiment Analyst Conclusions

1

Negative sentiment dominates at 49.6%, indicating emotionally charged, adversarial public engagement

2

Positive (25.6%) and Neutral (24.9%) sentiments are roughly equal, suggesting a minority engages informatively or affirmatively

3

Overall pattern reflects moral outrage and systemic critique rather than neutral deliberation

Improve Transparency in Institutional Reporting

Public discussions frequently mention terms such as **DOJ**, **evidence**, and **documents**.

Recommendations:

- Provide clear reporting on investigations, legal processes, and official developments.
- Explain complex legal documents and evidence in a way that is easy for the public to understand.
- Distinguish between verified information and unconfirmed claims circulating on social media.

Maintain a Victim-Centered Perspective

The appearance of terms such as **child** indicates that public discussions also reflect concerns about victims and their rights.

Recommendations:

- Ensure that media coverage remains victim-centered and respectful.
- Provide updates on legal outcomes and victim advocacy efforts.

Sentiment Analyst Recommendations

Thank You!



[illegible]