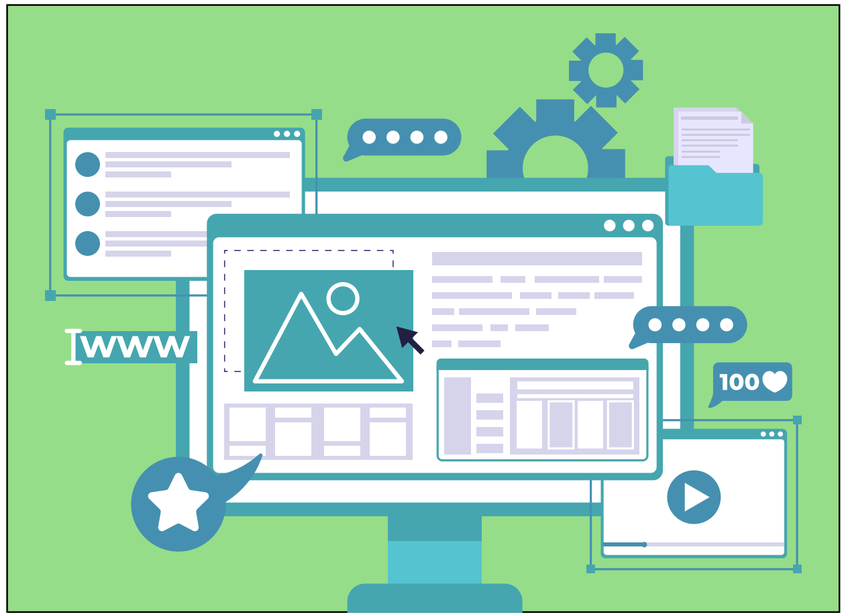
APRIL THEORY

Social Media Sentiment Analysis:

DONE BY:NOURHAN HASSAN

Intriduction

- Analyzed social media data around the April Theory
- Used NLP to assess sentiment and key discussion topics
- Focused on posts throughout the month of April



DATA OVERVIEW

Dataset & Tools

- Dataset: april_theory_final_dataset.csv
- Columns: text, created_at, top_words
- Tools used: Python, Pandas, NLTK, VADER, Matplotlib, Seaborn

```python
import pandas as pd
import random
from datetime import datetime, timedelta

# Sample tweet-like content
sample_texts = [
    "I really believe in the April Theory. There's something magical about spring resets!",
    "April is the new January. Fresh air, longer days, fresh mindset. #AprilTheory",
    "Started my new habits in April and it's already going better than New Year's.",
    "April Theory just makes sense. Nature resets, so do we.",
    "Not sure about April Theory. It feels like just another month to me.",
    "Who else is starting fresh this April? Loving the energy! #FreshStart",
    "Forget January, April is when I actually feel alive again.",
    "April Theory hits different when the sun is out and everything is blooming.",
    "Can confirm: my April routine feels way more sustainable than my January resolutions.",
    "April Theory is overhyped. Just my two cents.",
    "April Theory is working wonders for my mental health!",
    "Trying April Theory this year—so far, I'm meditating daily and journaling. 🌱",
    "Spring cleaning inside and out. April Theory activated.",
    "Let's normalize starting over in April, not just January.",
    "April Theory: sunlight, flowers, dopamine. It tracks.",
    "Day 5 of April Theory and I'm already sleeping better.",
    "Not buying into April Theory, but I do love the weather.",
    "Just a vibe check: anyone else feeling way more motivated this month?",
    "This April reset feels more natural than the whole New Year pressure.",
    "April > January. That's the tweet."
]

# Generate timestamps
```

TEXT PREPOCESSING

1. Removed URLs, special characters, and numbers

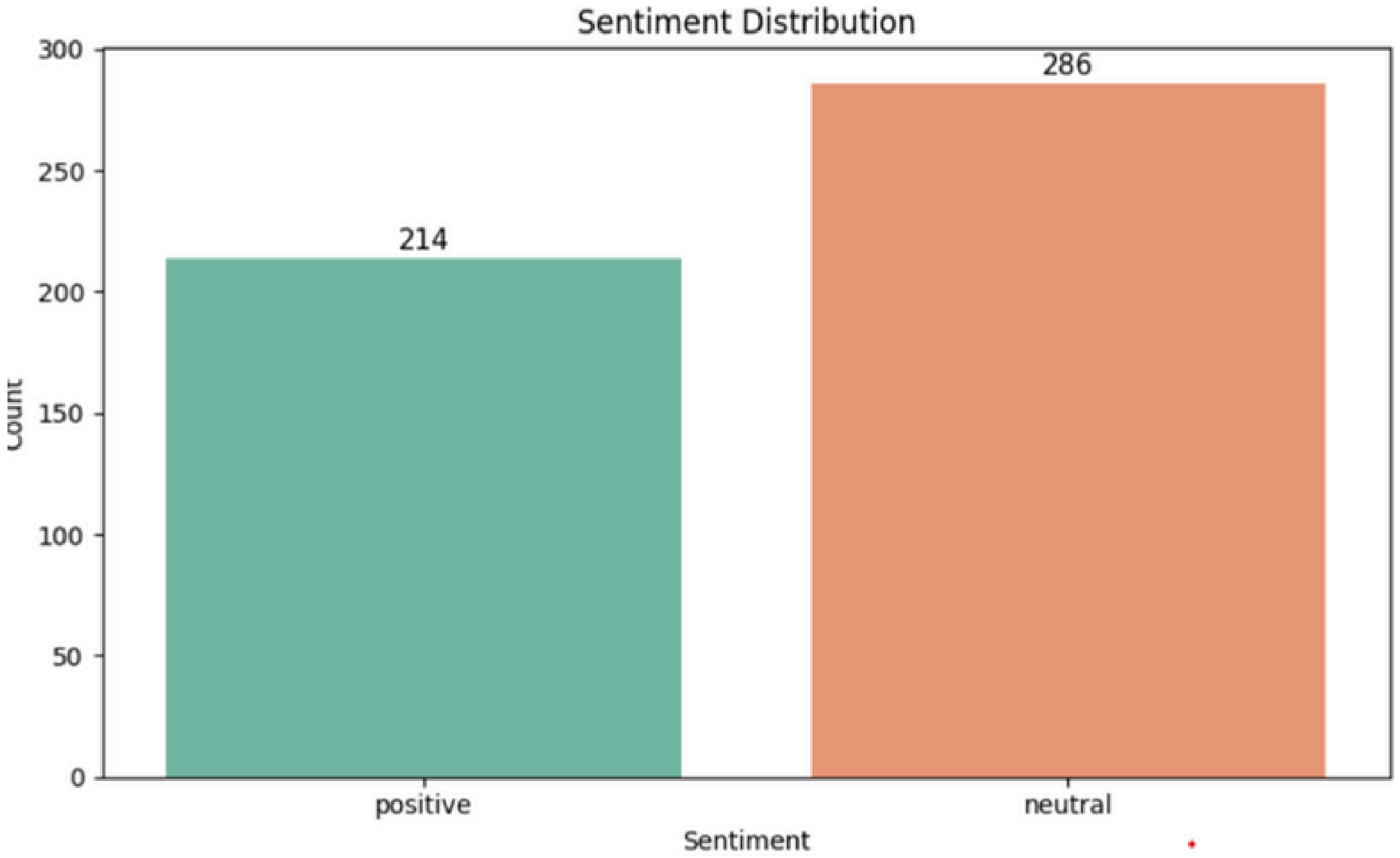
2. Lowercased text and removed stopwords

3. Used lemmatization for cleaner word forms

SENTIMENT ANALYSIS METHOD

- Used VADER for sentiment scoring
- Calculated compound score for each post
- Categorized posts as:
  -  Positive (≥ 0.05)
  -  Neutral (-0.05 to 0.05)
  -  Negative (≤ -0.05)

SENTIMENT ANALYSIS METHOD CHART

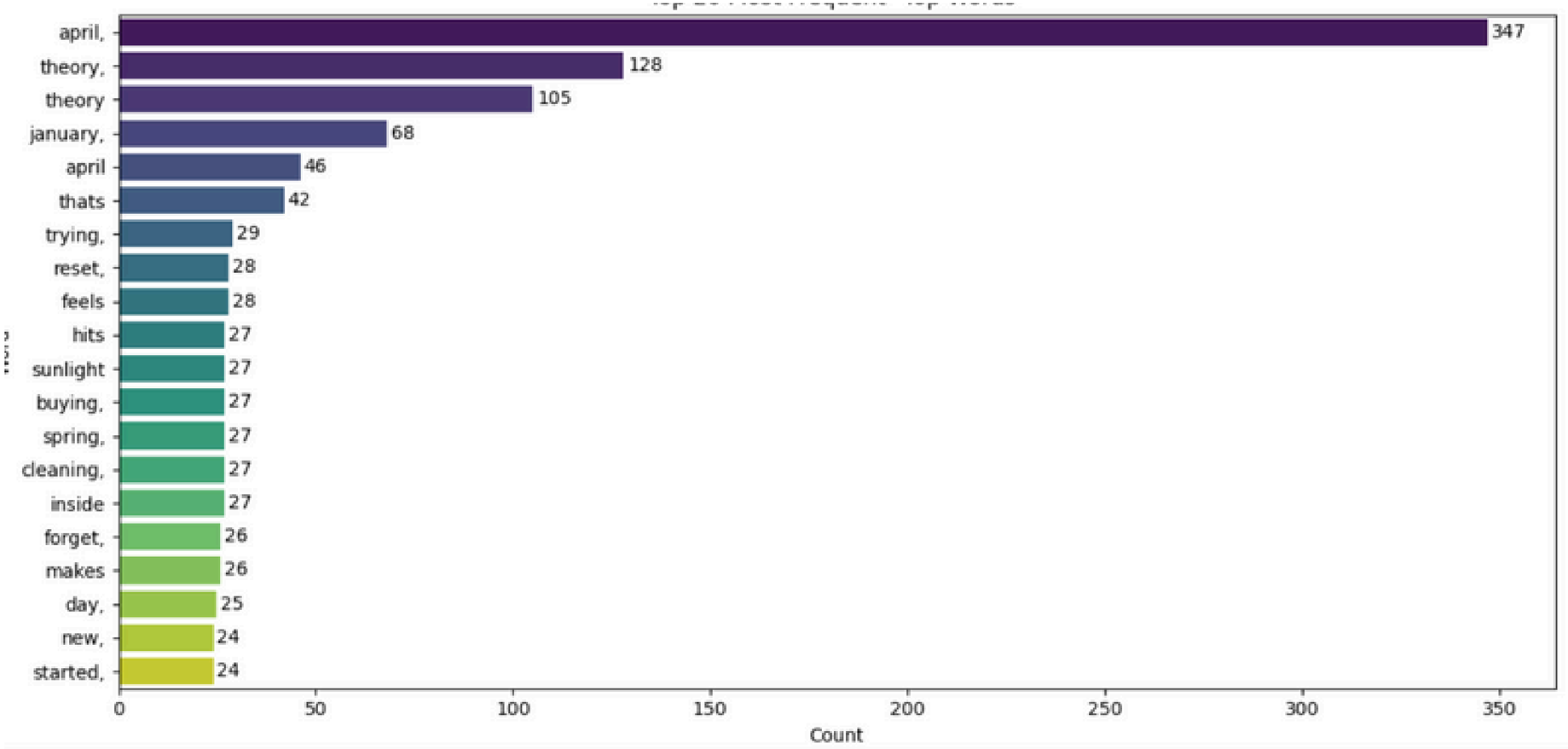


SENTIMENT ANALYSIS METHOD CODE

```
# STEP 3: Sentiment Analysis with VADER
nltk.download('vader_lexicon')
from nltk.sentiment import SentimentIntensityAnalyzer

sia = SentimentIntensityAnalyzer()
df["sentiment_score"] = df["cleaned_text"].apply(lambda x: sia.polarity_scores(x)["compound"])

def classify_sentiment(score):
    if score >= 0.05:
        return "Positive"
    elif score <= -0.05:
        return "Negative"
    else:
        return "Neutral"

df["sentiment"] = df["sentiment_score"].apply(classify_sentiment)
```

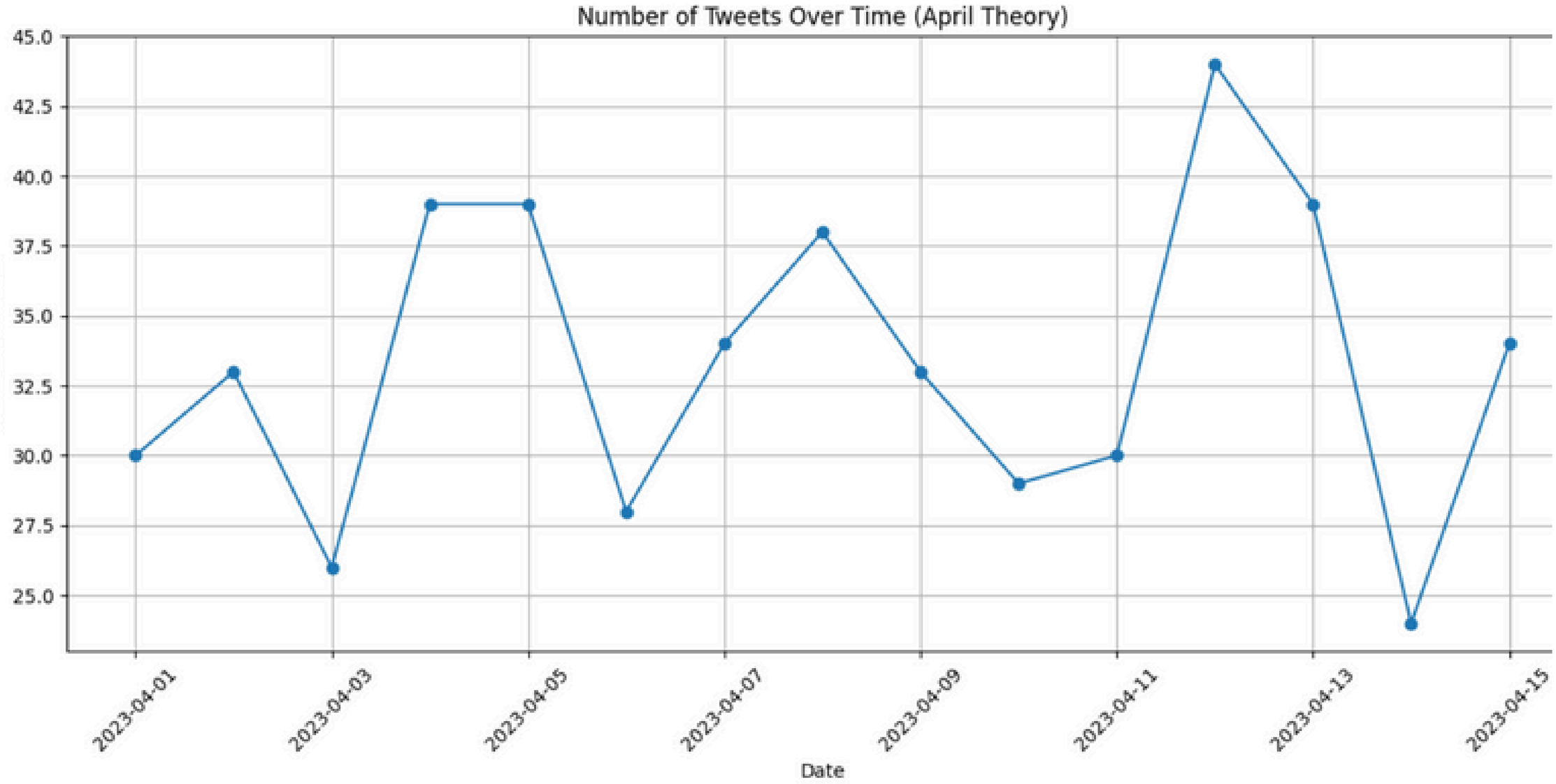
WORD FREQUENCY (CHART)

MOST USED WORDS



- “APRIL”,
- “CHANGE”,
- “NEW”,
- “START”,
- “SPRING”

TWEET ACTIVITY OVER TIME



TWEET ACTIVITY OVER TIME

```python
import pandas as pd

# Load your tweet data
df = pd.read_csv('april_theory_final_dataset.csv')

# Convert to datetime
df['created_at'] = pd.to_datetime(df['created_at'])

# Group by date
df_grouped = df.groupby(df['created_at'].dt.date).size()

# Plot
import matplotlib.pyplot as plt

plt.figure(figsize=(12, 6))
df_grouped.plot(kind='line', marker='o')
plt.title('Number of Tweets Over Time (April Theory)')
plt.xlabel('Date')
plt.ylabel('Number of Tweets')
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

KEY INSIGHTS

1. Sentiment was overwhelmingly positive throughout April
2. Common themes: renewal, fresh starts, and emotional reset
3. Aligns with seasonal optimism and mental health reflections

CONCLUSION

- April triggers a visible emotional shift in social conversations
- The April Theory reflects deeper patterns of collective mindset
- Social media sentiment analysis reveals real-time emotional trends