

Text Analytics API: Quick Reference Guide

Follow Along During Demo

Setup Checklist

- Get your API key from [gnews.io](#) (free tier = 100 calls/day)
 - Install libraries: `vaderSentiment`, `wordcloud`, `requests`, `pandas`
 - Pick a company to analyze (Tesla, Apple, your employer?)
-

The 4-Step Pipeline

Step 1: Fetch the Data

```
python
url = f"https://gnews.io/api/v4/search?q={COMPANY}&token={API_KEY}"
response = requests.get(url)
data = response.json()
```

You're getting: Live news articles about your company

Step 2: Structure It

```
python
df = pd.DataFrame(data['articles'])
df = df[['title', 'description', 'url', 'source']]
```

You're creating: A clean table like Excel

Step 3: Analyze Sentiment

```
python
analyzer = SentimentIntensityAnalyzer()
df['sentiment_score'] = df['description'].apply(
    lambda x: analyzer.polarity_scores(x)['compound']
)
```

You're measuring: How positive/negative each article is (-1 to +1)

Step 4: Visualize Themes

```
python
all_text = " ".join(df['description'])
wordcloud = WordCloud().generate(all_text)
plt.imshow(wordcloud)
```

You're showing: What topics dominate the conversation

Sentiment Score Decoder

| Score Range | Meaning | Business Action |
|--------------|---------------|--------------------|
| > 0.5 | Very Positive | Leverage good PR |
| 0.1 to 0.5 | Positive | Monitor trends |
| -0.1 to 0.1 | Neutral | Normal operations |
| -0.5 to -0.1 | Negative | Investigate issues |
| < -0.5 | Very Negative | Crisis management |

What Each Variable Contains

- `API_KEY`: Your personal access code
 - `COMPANY`: The search term (can be any text)
 - `data`: Raw JSON from the API
 - `df`: Structured DataFrame with articles
 - `sentiment_score`: -1 (bad) to +1 (good)
 - `wordcloud`: Visual frequency map
-

Common API Response Fields

```
python
data['articles'][0].keys()
# Returns: title, description, url, source, publishedAt, content
```

- **title**: Article headline
 - **description**: 2-3 sentence summary
 - **url**: Link to full article
 - **source**: Publication name
 - **publishedAt**: Timestamp
-

VADER Sentiment Components

```
python
scores = analyzer.polarity_scores(text)
# Returns: {'neg': 0.0, 'neu': 0.7, 'pos': 0.3, 'compound': 0.6}
```

- **neg**: % negative
 - **neu**: % neutral
 - **pos**: % positive
 - **compound**: Overall score (what we use)
-

Troubleshooting

No data returned?

- Check API key is correct
- Verify company name spelling
- Try without spaces: "BankofAmerica" not "Bank of America"

Weird sentiment scores?

- VADER struggles with: Sarcasm, industry jargon, non-English
- Solution: Check a few descriptions manually

Word cloud shows common words?

- Add stop words: `WordCloud(stopwords={'the','and','for'})`
-

Try These Variations

1. Filter by Sentiment

```
python
```

```
positive_news = df[df['sentiment_score'] > 0.5]
negative_news = df[df['sentiment_score'] < -0.5]
```

2. Find Most Positive/Negative

```
python
best_news = df.loc[df['sentiment_score'].idxmax()]
worst_news = df.loc[df['sentiment_score'].idxmin()]
```

3. Group by Source

```
python
df.groupby('source')['sentiment_score'].mean()
# See which outlets are most positive/negative
```

4. Export Results

```
python
df.to_csv(f'{COMPANY}_sentiment.csv', index=False)
```

Business Questions to Consider

1. **Trend Analysis:** Is sentiment getting better or worse?
2. **Source Bias:** Do certain outlets favor/criticize this company?
3. **Topic Correlation:** Which words appear with negative sentiment?
4. **Competitor Comparison:** How does Apple sentiment compare to Samsung?
5. **Event Impact:** Did the product launch improve sentiment?

Your Turn: Mini Project Ideas

- **Easy:** Compare two competitors' sentiment
- **Medium:** Track sentiment change after major announcement
- **Hard:** Predict stock movement from sentiment scores
- **Expert:** Build dashboard updating every hour

Remember

You're not just analyzing text—you're extracting business intelligence from unstructured data. This skill is worth \$100K+ in the job market.

Key insight: 80% of business data is text (emails, reviews, reports). You just learned to analyze it.