

Settings

Transformation: Lowercase, Parse html, Remove urls
Tokenization: Tweet
Filtering: Stopwords (Language: English, File: None), Numbers (Includes numbers: No)), Regexp (\\.|,:;|!|\\?|\\(|\\)|\\|\\+|\\'|\"|'|\"|'|\\|\\|...|-|-|—|\\\$|&|*|>|<|\\V|\\V|\\V))

Term Frequency: Count
Document Frequency: (None)
Regularization: (None)

Constructed features

```
Python Script. Método del codo

import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

# Convertimos la tabla de Orange a matriz numpy (bolsa de palabras)
X = in_data.X

# Lista para guardar inercia para cada número de clusters
inertia = []

# Calcular inercia para k de 1 a 15
for k in range(1, 16):
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(X)
    inertia.append(kmeans.inertia_)

# Graficar método del codo
plt.figure()
plt.plot(range(1, 16), inertia, marker='o')
plt.xlabel('Número de clusters')
plt.ylabel('Inercia')
plt.title('Método del codo')
plt.show()
```

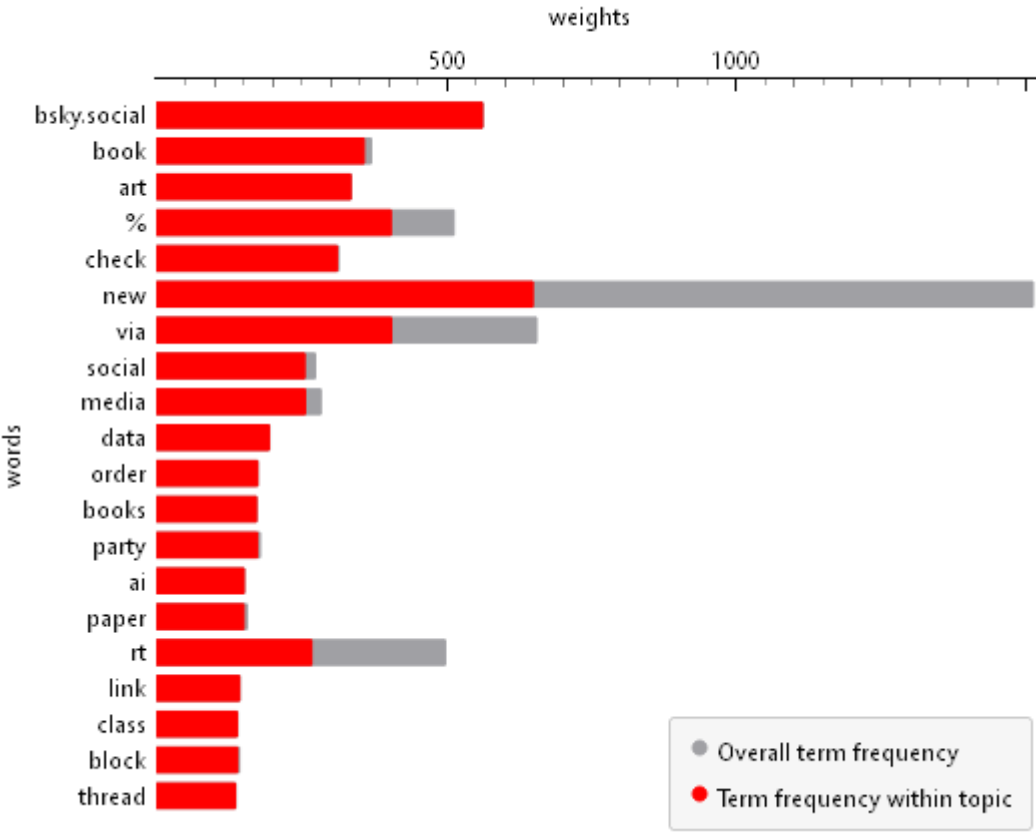
Latent Dirichlet Allocation

Number of topics: 3

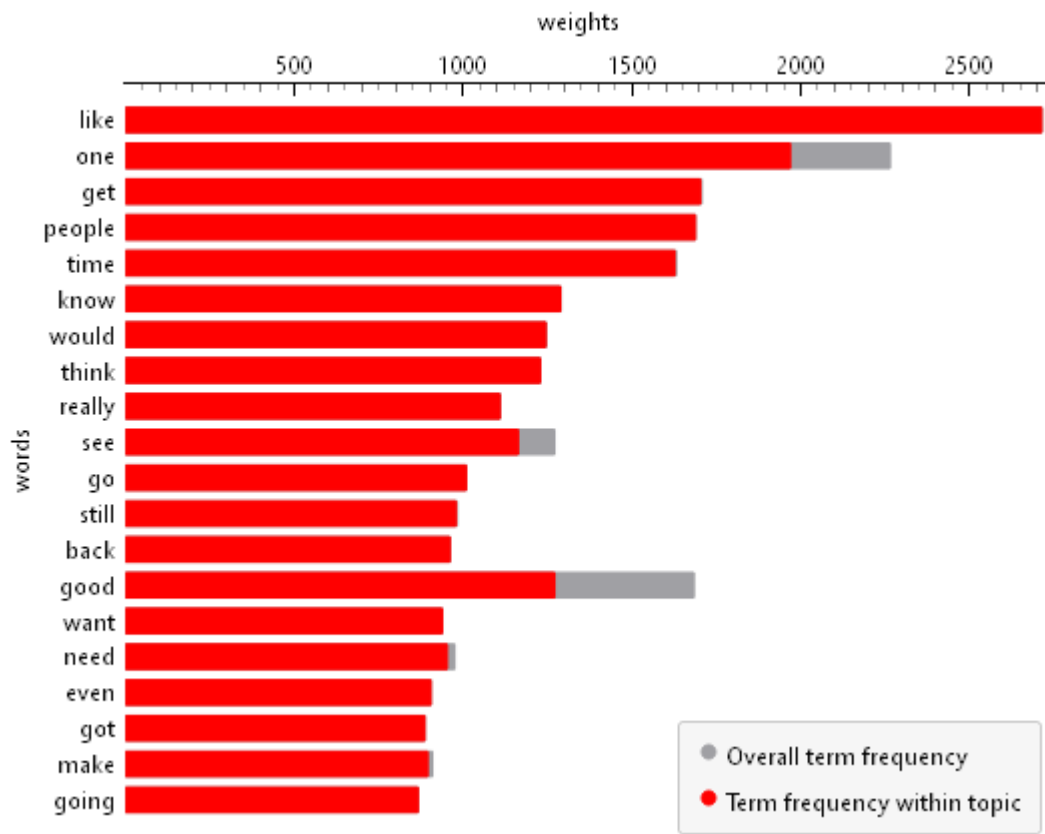
Topics

- 1: new, bsky.social, via, %, book, art, check, one, rt, day
- 2: like, one, get, people, time, know, good, would, think, see
- 3: thanksgiving, happy, day, , thank, today, 🥰, family, morning, good

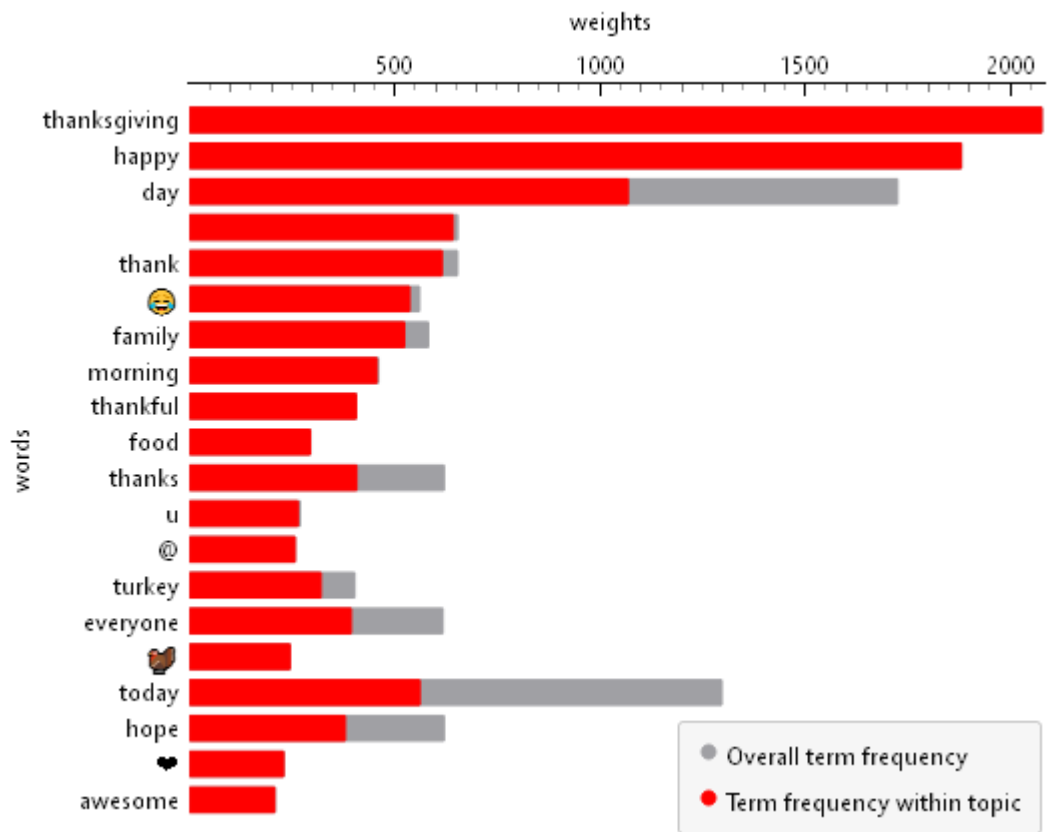
Relevance: 0.5
Shown topic: Topic 1



Relevance: 0.5
Shown topic: Topic 2



Relevance: 0.5
Shown topic: Topic 3



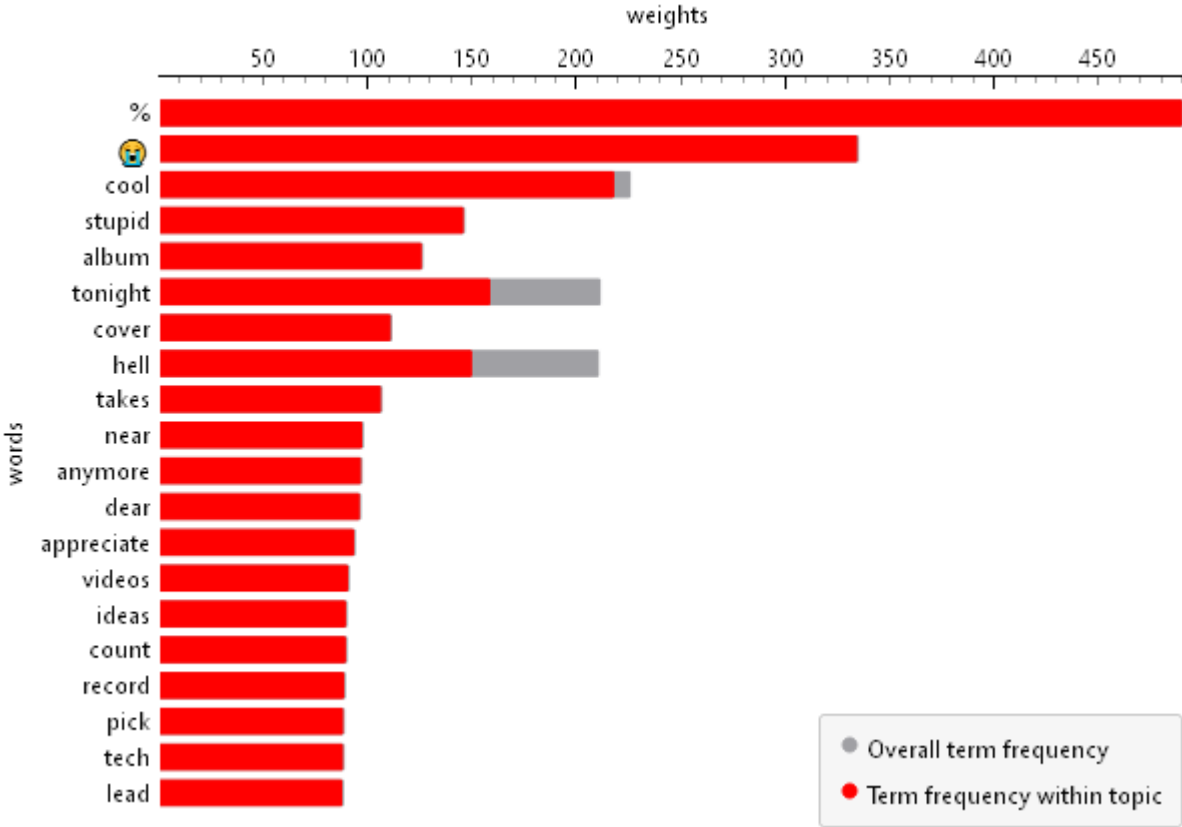
Latent Dirichlet Allocation

Number of topics: 9

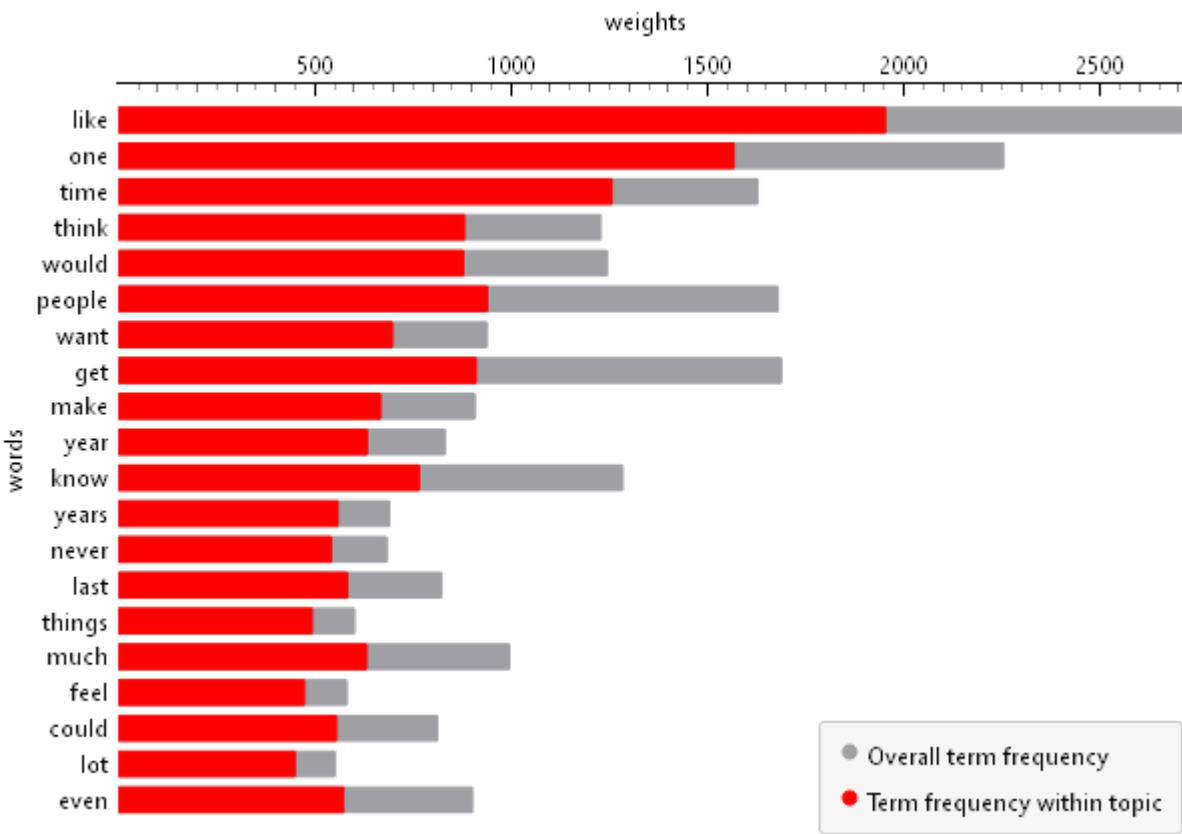
Topics

- 1: %, 🤖, cool, tonight, hell, get, stupid, new, nice, album
- 2: like, one, time, people, get, think, would, know, want, good
- 3: thanksgiving, happy, day, today, good, family, thank, hope, morning, thankful
- 4: new, via, 📷, photo, car, check, sounds, hard, age, okay
- 5: people, like, get, know, see, something, really, right, way, yes
- 6: bsky.social, via, x, half, love, came, wow, oh, lmao, glad
- 7: 😊, , thanks, women, ❤️, via, song, @, latest, video
- 8: love, rt, media, social, new, list, celebrate, following, back, data
- 9: one, day, new, power, book, days, books, level, interesting, order

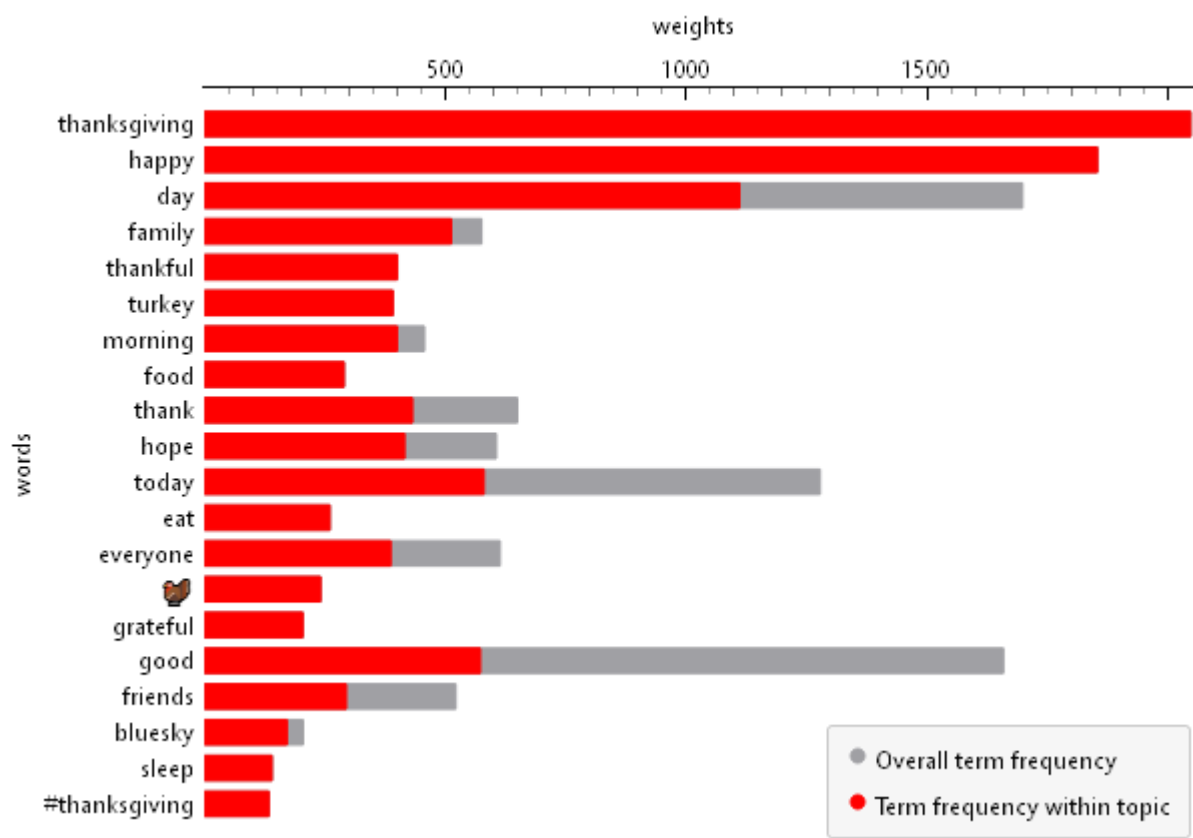
Relevance: 0.5
Shown topic: Topic 1



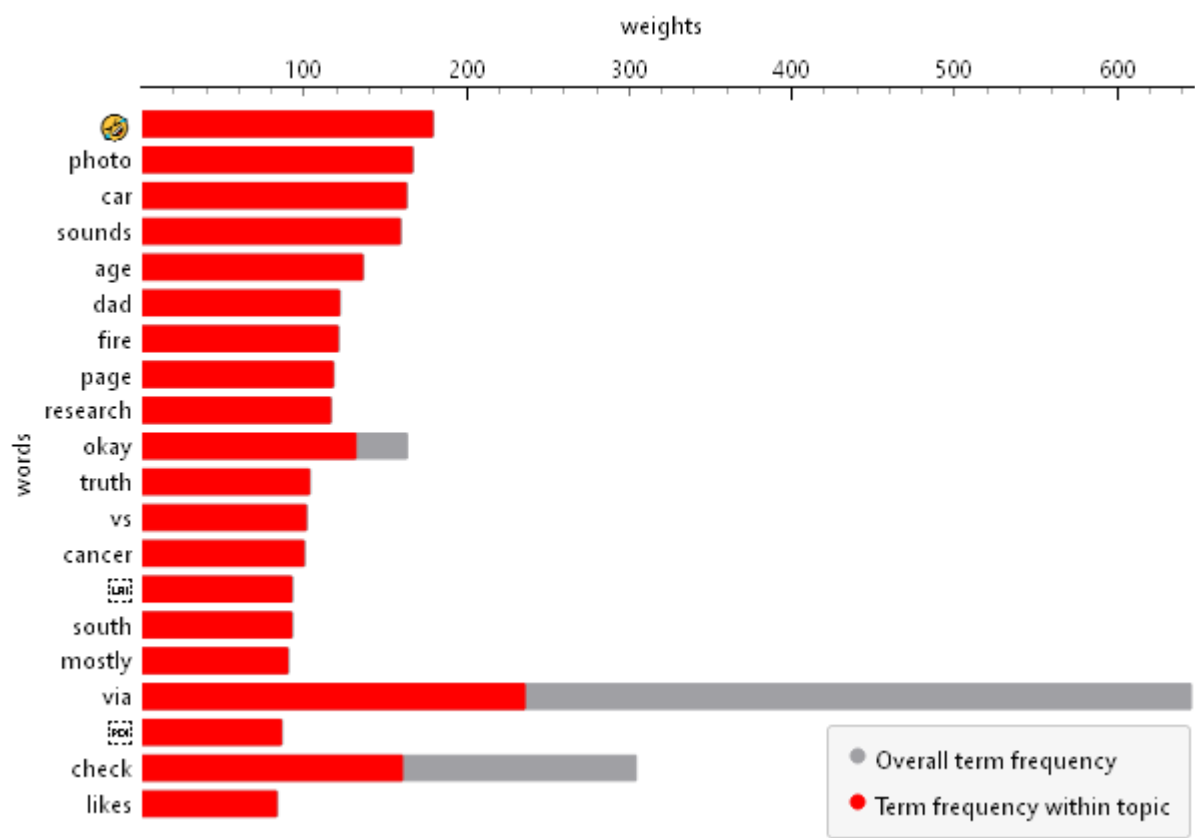
Relevance: 0.5
Shown topic: Topic 2



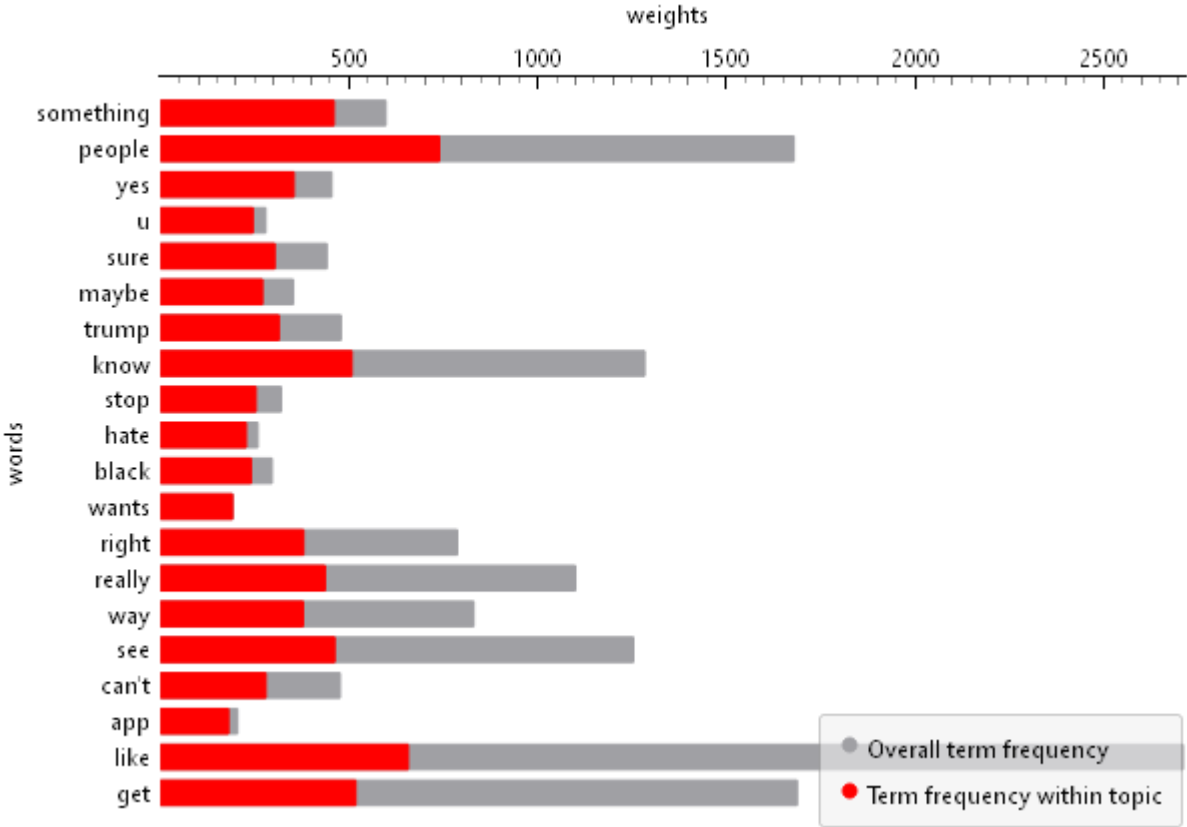
Relevance: 0.5
Shown topic: Topic 3



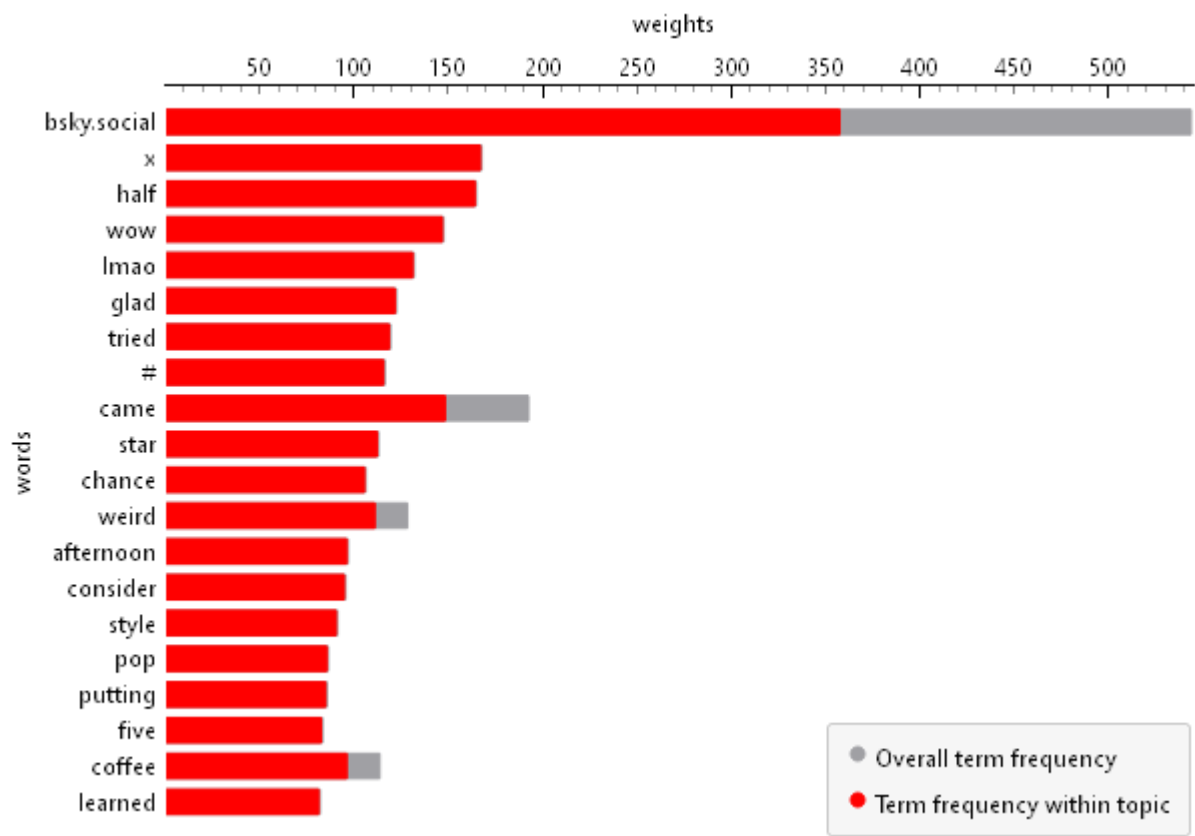
Relevance: 0.5
Shown topic: Topic 4



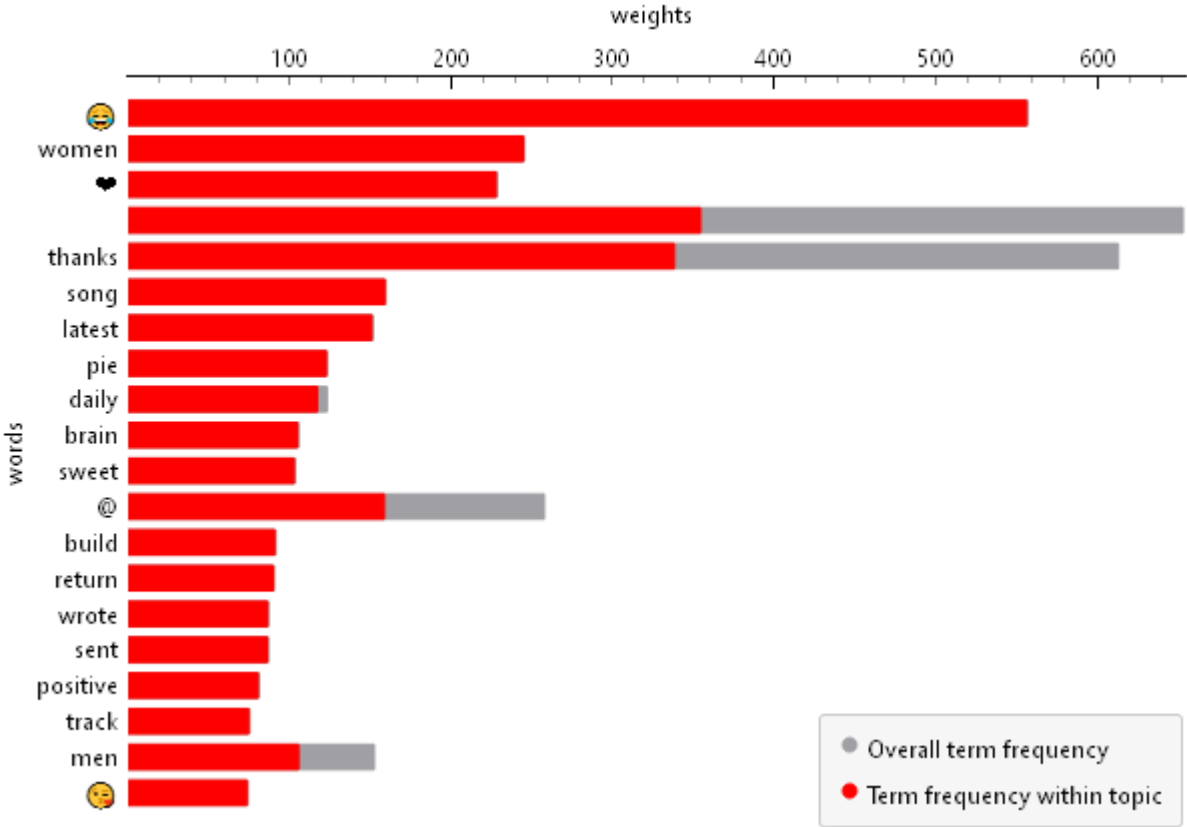
Relevance: 0.5
Shown topic: Topic 5



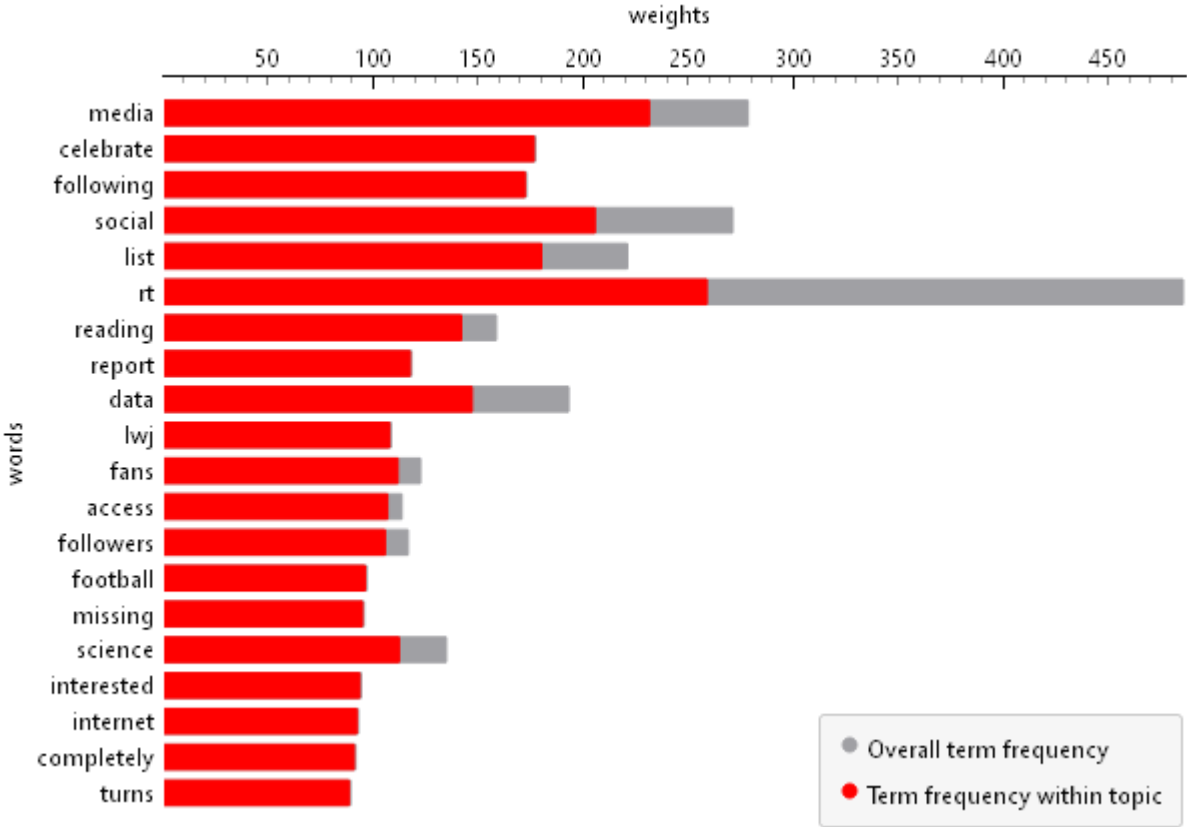
Relevance: 0.5
Shown topic: Topic 6



Relevance: 0.5
Shown topic: Topic 7



Relevance: 0.5
Shown topic: Topic 8



Relevance: 0.5
Shown topic: Topic 9

