

# MUSIC LYRICS ANALYSIS: LYRICS-BASED SONG RECOMMENDER

ALEX BELL  
METIS DATA SCIENCE  
NOVEMBER 16, 2018



# MOTIVATION

- Recommend songs based on lyrical content
- Many listeners like music across genres
- Useful for Movie Industry to find topical songs
- Alternative to ratings-based collaborative filtering

# DATA

- Lyrics to 516,167 songs from [lyrics.com](https://lyrics.com)
- Multitude of languages and genres
- Songs spanning many generations





# METHODOLOGY: PREPROCESSING

- Lyrics data formatted into tables
- Lyrics tokenized into words
- Remove punctuation
- Change to all lower case
- Remove stop words from 50 languages

# METHODOLOGY: VECTOR SPACE MODELS

- Lyrics transformed into vector space
- Document word matrix
- Can now proceed to calculating relevance between words and songs

# METHODOLOGY: BUILDING A MODEL

- LDA: Latent Dirichlet Allocation
- Calculate probabilities of words in topics
- Evaluate model performance
- Tune parameters
- Lyrics grouped into 40 topics

# INFERRING THE TOPIC FROM KEYWORDS

woman perfect  
read working  
sugar warm  
wash america

AMERICANA

red river wine  
glass roses  
chain diamonds  
water paint

LIFESTYLE

fear free  
darkness alive  
burn survive  
human riding

GOTHIC

girl wanna  
move lets  
groove sexy  
fever body

CLUB  
DANCE

heaven jesus  
children  
grow moon

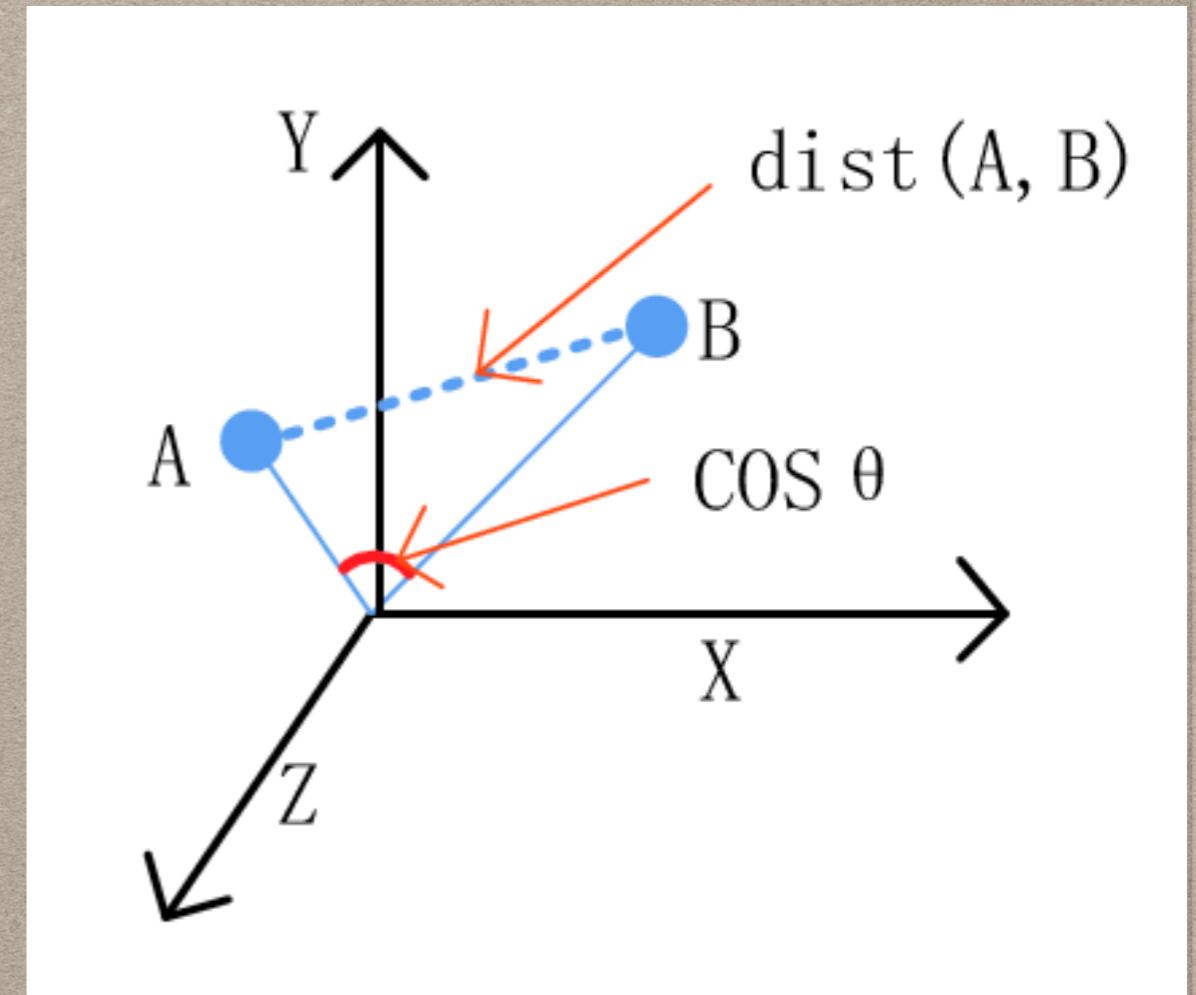
SPIRITUAL

lady falling  
wide answer  
road haven  
romance bridge

LOVE  
BREAKUPS

# METHODOLOGY: PAIRWISE SONG SIMILARITY

- Measure cosine similarity between songs
- Calculates similarity of songs by relative distance of words



**INPUT WORDS FOR RESULTS:**

# TRUCKS BEER AMERICA

BAND	SONG	LYRICS
George Strait	That's My Kind of Woman	<p>She can drive a truck and rope. She feels right at home by my side. That's my kind of woman. That's my cup of tea ...</p>
Pete Seeger	East Virginia	<p>I was born and raised in east Virginia, North Carolina I did go. There I met some pretty little woman. But her name and age I ...</p>
Graham Parker	Heat in Harlem	<p>I ain't stopping fingers keep bopping. Swing band rocking from the street to the bar. West coast's dated music's overrated.</p>

# ROCK DEVIL LOUD

BAND	SONG	LYRICS
F-Minus	Vultures	Corporate market rule! Monetary steeple. Crucified for power. Profit over people. Capitalist vultures circle a dying nation.
Rob Zombie	Dragula	Superstition... now add jealousy. Dead I am the one, exterminating son. Slipping through the trees, strangling the breeze...
Asche & Spencer	Night into day	These are the words that cry out of me in the night. Like a buffalo herd from a shotgun blast scattered in flight. I got one

# SANTA JINGLE SNOW

BAND	SONG	LYRICS
Mike Bogle	Banana Tree	Banana tree I don't know. Banana tree maybe. Banana tree I don't know. The best way to see it. Banana tree I don't know.
Agent Sparks	Polly Anne	Superstition... now add jealousy. Dead I am the one, exterminating son. Slipping through the trees, strangling the breeze...
John Michael Montgomery	Winter Wonderland	Sleigh bells ring, are you listenin'? In the lane, snow is glistenin'. It's a beautiful sight, we're happy tonight. Walkin' in a ...

# LOVE HEARTACHE LEAVE

BAND	SONG	LYRICS
The Outfield	Your Love	Josie's on a vacation far away. Come around and talk it over. So many things that I want to say. You know I like my girls a li...
The Magic Numbers	Let Somebody In	Said you want to love but you don't know how. And you want to feel but you're not allowed. And you want to cry But you don't
Squeeze	If I Didn't Love You	If I didn't love you, I'd hate you. Watching you play in the bath A soap suds stickle back navy. A scrubbing brush landing craft

**FIND SIMILAR SONGS BY  
CHOICE / CREATE A PLAYLIST:**

# ... SKY ROCKETS IN FLIGHT, AFTERNOON DELIGHT! ...

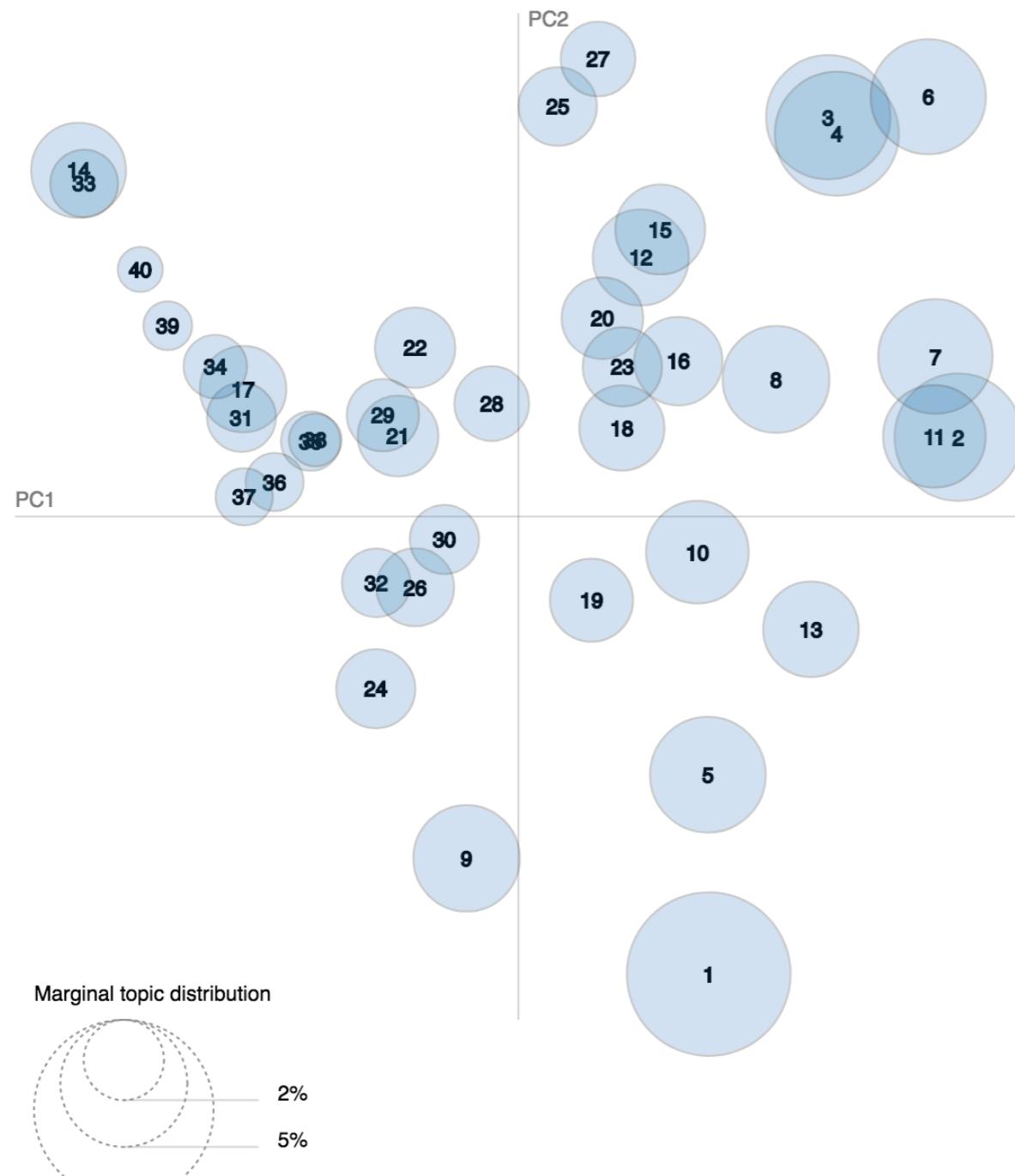
BAND	SONG	LYRICS
Barry Manilow	Saturday Night	<p>Saturday night is the loneliest night in the week. 'Cause that's the night that my sweetie and I Used to dance cheek to cheek</p>
Frank Sinatra	A Lovely Way To Spend An Evening	<p>Some like a night at the movies, some like a dance or a show. Some are content with an evening spent home by the radio. Some</p>
The Wright Touch Big Band	Summer Wind	<p>The summer wind came blowin' in from across the sea. It lingered there to touch your hair and walk with me. All summer long</p>

# CONCLUSIONS & FUTURE WORK

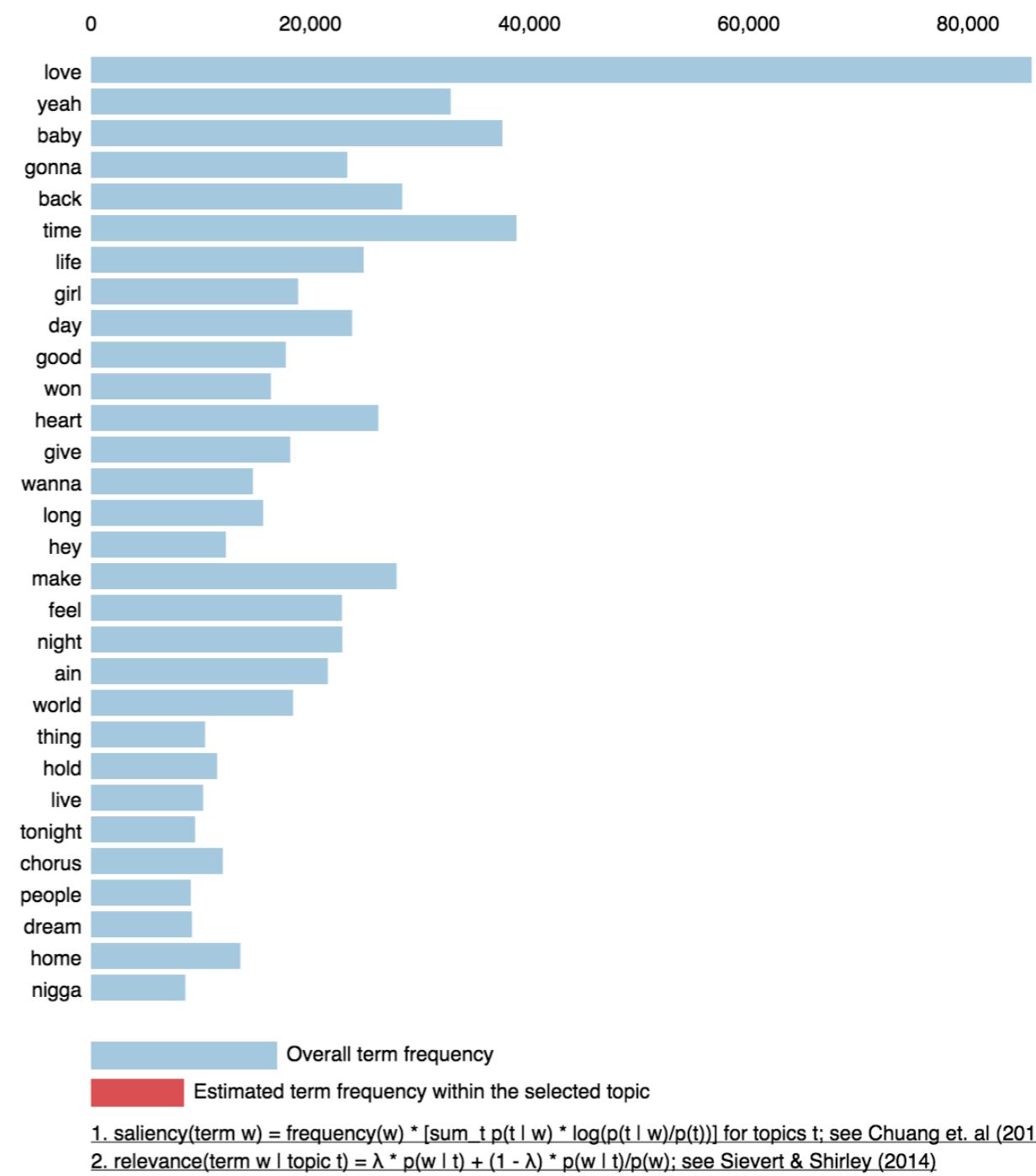
- Model performs well with cosine similarity
- Exposes listeners to a wide variety of songs
- Repetitive songs skew results
- Process songs into unique word sets
- Create flask app for user input in browser

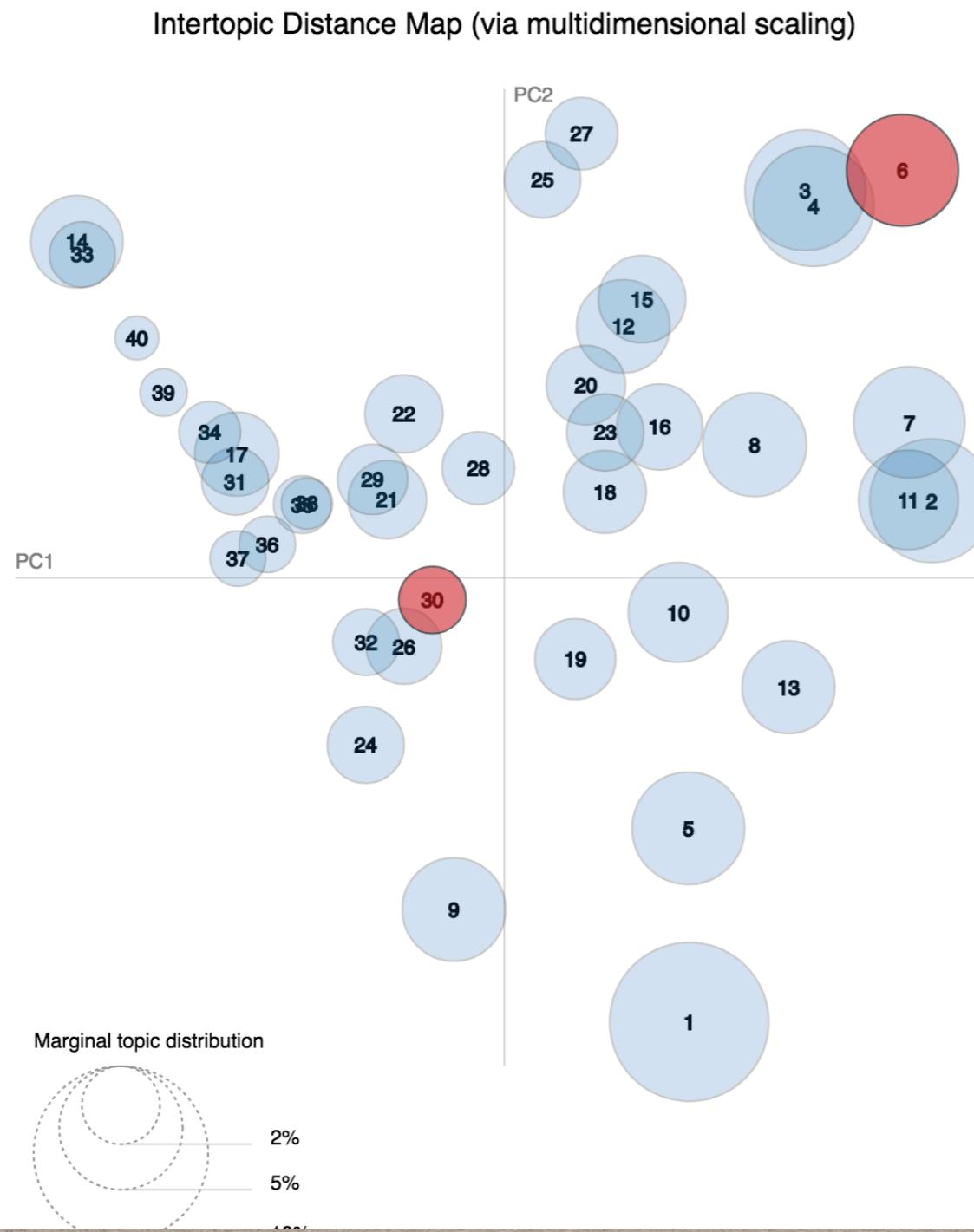
**THANK YOU**

Intertopic Distance Map (via multidimensional scaling)

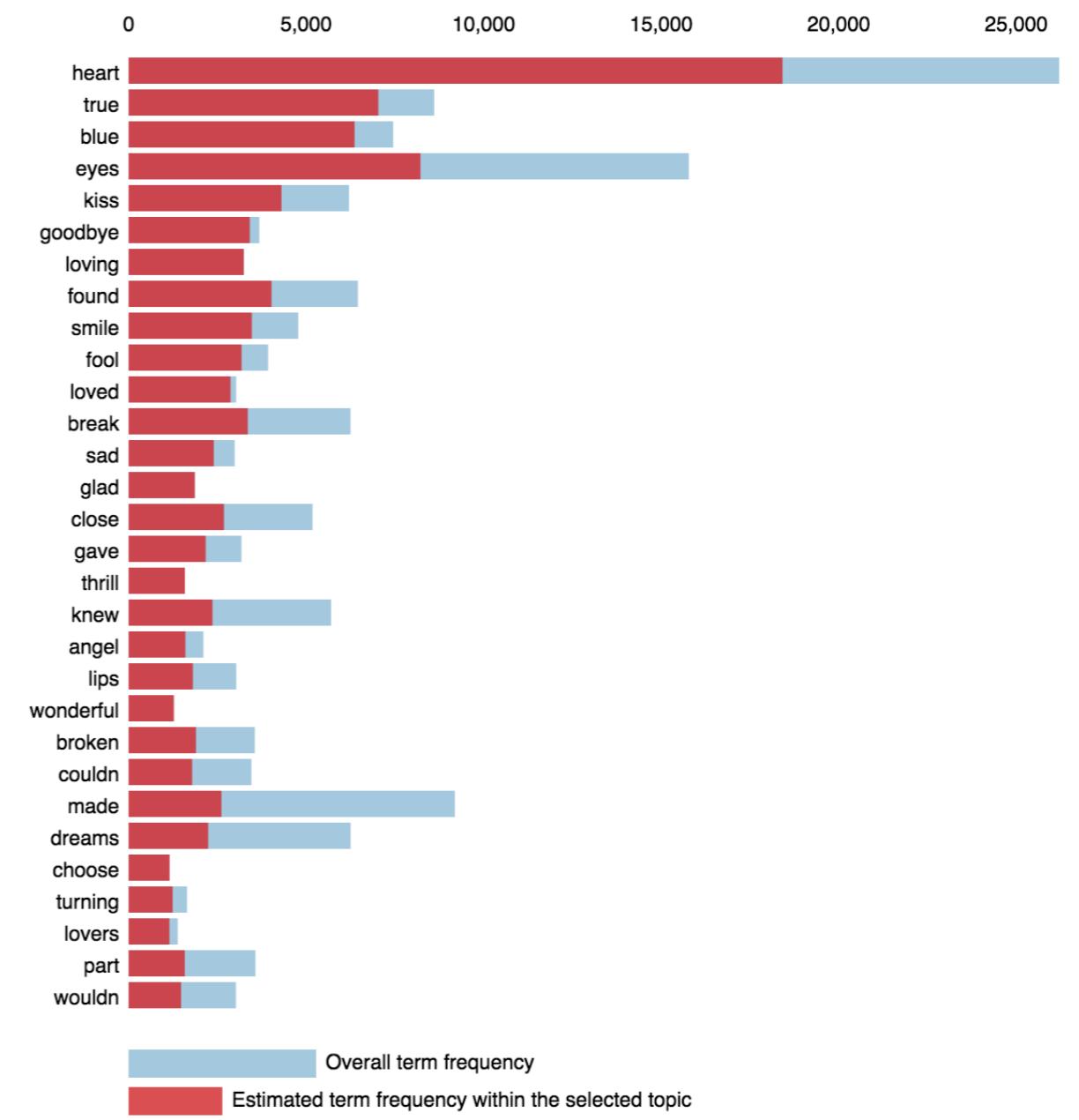


Top-30 Most Salient Terms<sup>1</sup>





Top-30 Most Relevant Terms for Topic 6 (4.1% of tokens)



1. saliency(term w) = frequency(w) \* [sum\_t p(t | w) \* log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) =  $\lambda * p(w | t) + (1 - \lambda) * p(w | t) / p(w)$ ; see Sievert & Shirley (2014)