



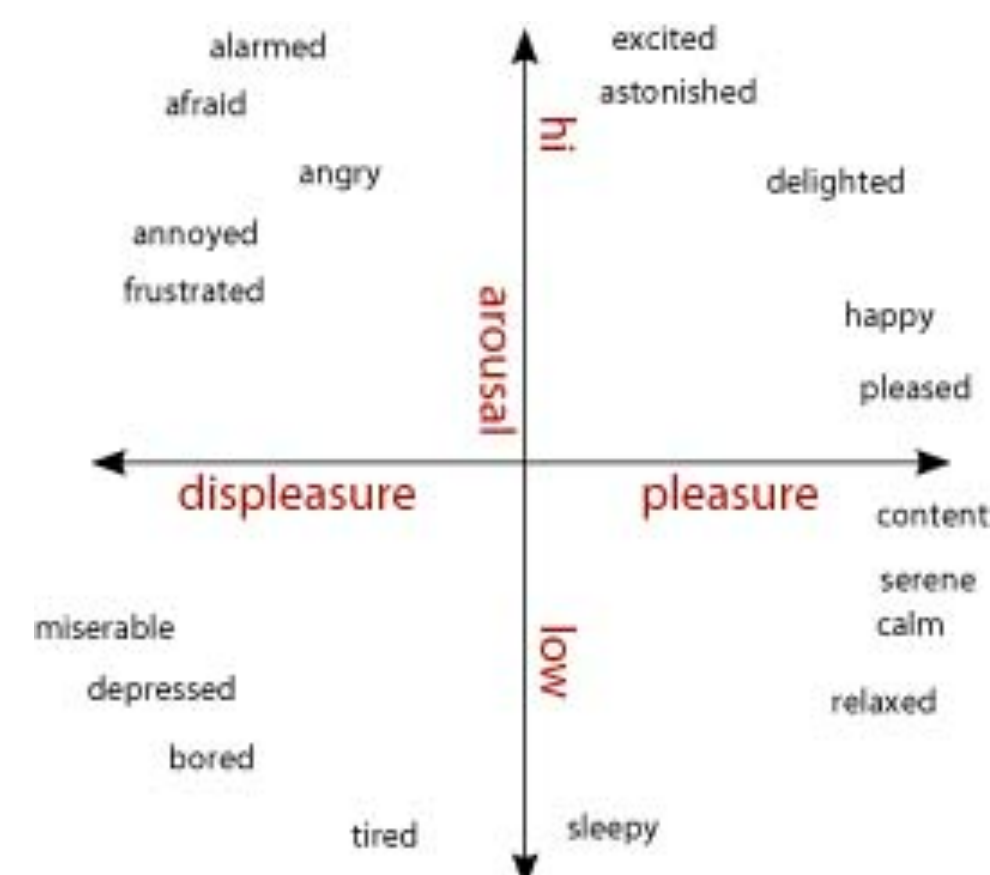
Emotion Classification of Song Lyrics

Dani Mednikoff & Jen Zhu
CS 585: Natural Language Processing, Fall 2017



Introduction

- Our goal was to accurately classify songs according to the emotions expressed through the lyrics
- We used a system of **valence** and **arousal**:



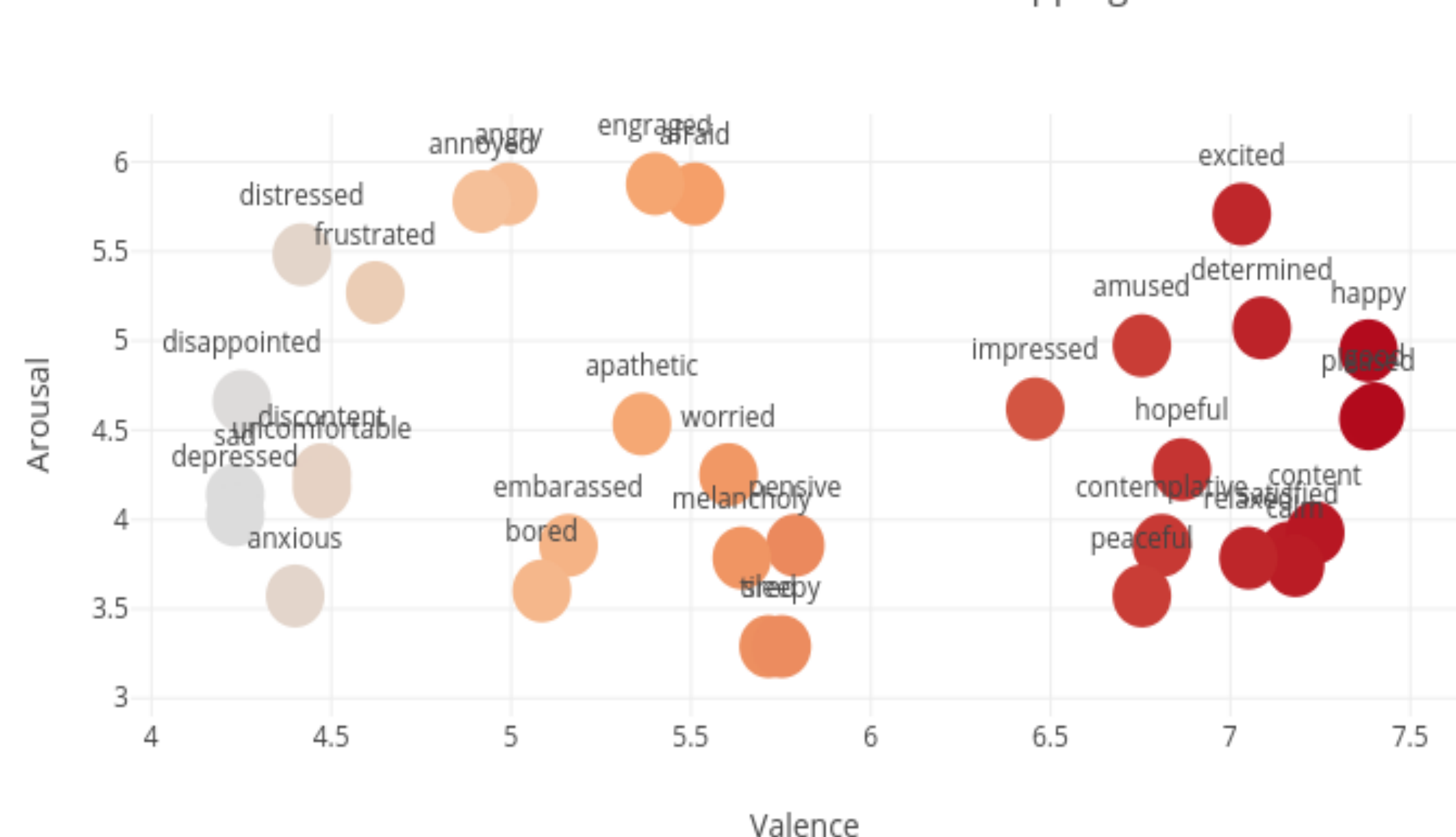
Tools & Data

- Song Lyrics:**
 - Kaggle Billboard Top 100 from 1964 – 2015 dataset, removing instrumental and non-English songs
 - 5,057 songs
 - Includes song title, rank, year, artist and lyrics
- Word Dictionary:**
 - 14,000 words
 - Valence and arousal scores
- Emotion Dictionary:**
 - 34 Emotions
 - Valence and arousal scores

Approach

- Created a **bag of words** for each song from data set
- Calculated **average valence** and **arousal values** for the song:
 - Locate word's valence and arousal scores, multiplied by the number of times word was repeated
 - Divided avg by # of words located
- Scaled** each **emotion's valence** and **arousal range**

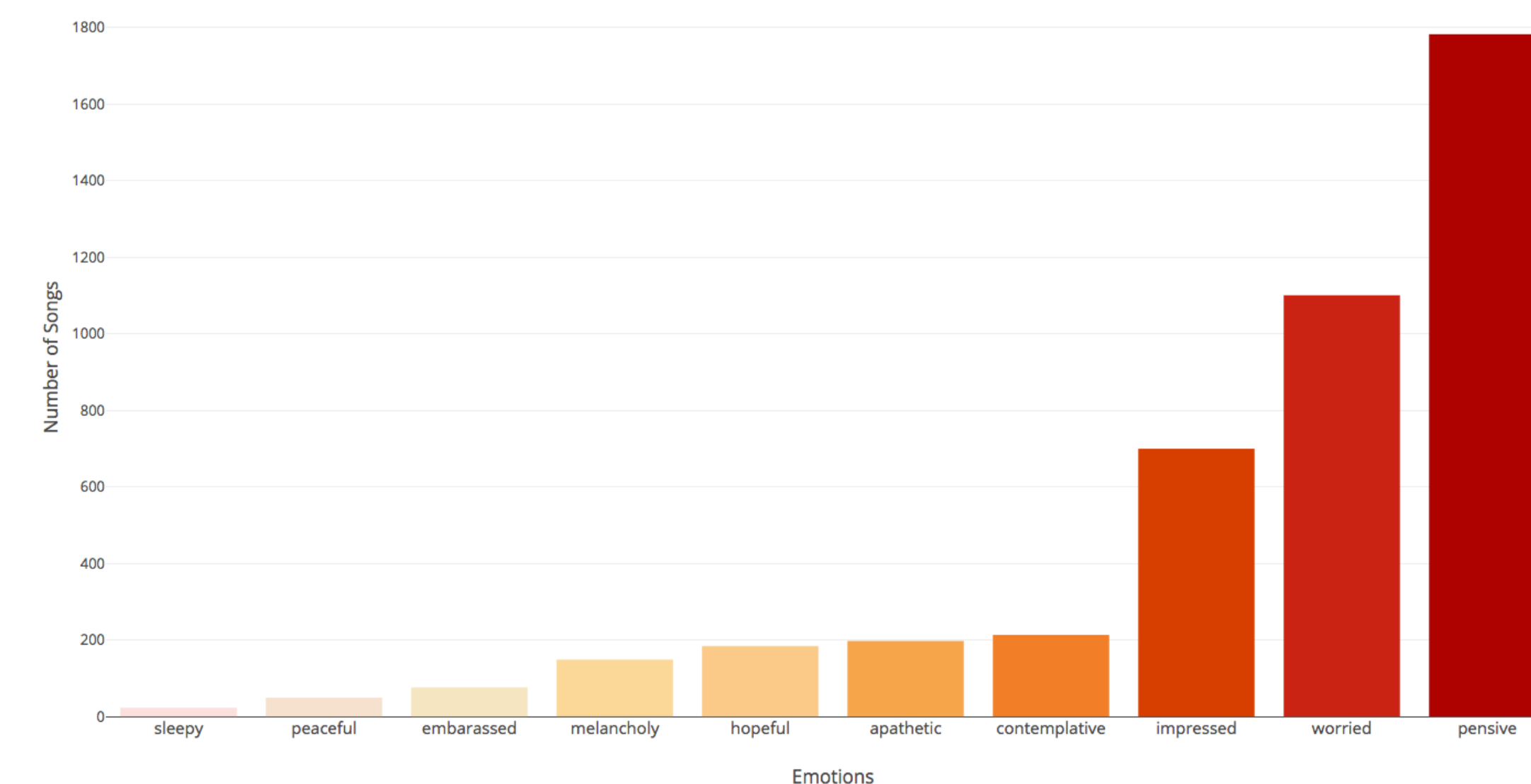
Emotion Valence and Arousal Mappings



- Used **Euclidean distance** to map the nearest emotion to a song's average valence and arousal scores

Results

Top 10 Emotions vs Number of Songs



- Top 10 Emotions and Number of Songs:** *sleepy (23), peaceful (50), embarrassed (76), melancholy (149), hopeful (184), apathetic (198), contemplative (214), impressed (700), worried (1101), pensive (1783)*
- Examples:** *“Happy”: amused, “Halo”: peaceful, “Teardrops on My Guitar”: contemplative*

Future Implementations

- Address words not located via synonym lexicons
- Using additional information such as artist, year, rank to discover patterns