INTRODUCTION

- What about a song determines its popularity? What makes a song ‘good’ is subjective but there are certainly trends in how music is perceived and these trends are complex and interesting.
- We used machine learning to determine the features of songs that are most important for predicting song popularity.

DATASET (DS)

- The Million Song Dataset (MSD) contains almost 500 GB of song data and metadata from which we extract features for our learning models.
- We used a subset of the MSD containing 10,000 songs to train and develop our learning models.
- To measure popularity, we used “hotttness”, which is a metric developed by Echonest and used in the MSD.

FEATURE EXTRACTION (FE)

- **Macro-level features**: included song key, tempo, duration, length of song name, time signature, fadeout duration, loudness.
- **Micro-level features**: extracted features from the MSD “segment” feature arrays with various statistics.
- **Bag-of-words features**: used term frequencies of 100 most common words as features for our models.
- **Location features**: Artist origin is converted to indicator features for the top 10 most common locations.

IMPLEMENTATION & EXPERIMENTS (I&E)

- In the MSD, songs are split into hundreds of ‘segments’ that are roughly uniform in timbre and pitch. The timbre is analyzed for each of these segments, yielding a value corresponding to each of the 12 spectral basis functions above. Pitch information is also included for each segment. We manipulated this information to extract information about the song’s melodic and harmonic content. This yields incredibly descriptive information about what a song sounds like.

RESULTS & ANALYSIS (R&A)

- Most common artist locations sorted by increasing hotttness:
  - United States
  - Texas
  - California
  - Washington
  - Florida
  - Georgia
  - New York
  - Illinois
  - Missouri
  - Ohio
- Highest weighted micro-level features:
- Most negative weighted features:

FUTURE WORK (FW)

- We are considering implementing classifier models to make direct predictions about song popularity.
- Time permitting, we would like to implement a neural network for predicting song hotttness.

REFERENCES