Music-Evoked Autobiographical Memories

Julia Stelman

Professor: Petr Janata
Background

- Inspiration: Alzheimer’s study
- Data source: Prof. Janata’s psychology students and other subject volunteers
- Data Collection: MEAMCentral “Music-Evoked Autobiographical Memories”
- Data format: Linked Data; SPARQL triples
Goals

- Efficiently integrate Python NLTK tools to perfect sentiment analysis.

- Determine significant trends between emotions in memories and sentimental characteristics of associated songs.
What is MEAMCentral?
Investigation

- What information can be extracted from user-inputted memory?

- What can be discovered with that data and what other data is available?
  
  - Focus: Do memories with similar emotional trends correspond to songs with similar features?

  - Tools: Spotify Web API via “Spotipy” python library

  - Method: Track metadata, Mood and Genre playlists, Recommendation feature.

*sophomore year of college, i think. i seem to remember the guys in my singing group going through a phase in...*
"sophomore year of college, i think. i seem to remember the guys in my singing group going through a phase in which they were obsessed with barenaked ladies, the band. i seem to remember bnl being played a lot during tours or retreats. i can't pinpoint a specific memory, but the memory of the era feels satisfyingly authentic."
The memory narratives are each tokenized. The tokens are then lemmatized and stemmed. Stop words are removed.

Songs that didn’t come from Spotify were removed. (Some came from itunes). This is something I’m working on fixing because about 40% of the data is lost due to this.
Most popular words overall:

(stop words removed)

Word size is based on frequency of the root word in the whole dataset.
I have reviewed the data for content of ten different emotions.

The ten rows at the bottom each pertain to a specific emotion.

Their elements are frequencies of related words.
Breakdown of Emotion Words by Emotion Group
Spotify

Genres & Moods

Amplify: Women
Hip-Hop
Mood
Pop
Workout
Country
Latin
Focus
What does it all mean?

• I have over 4,000 unique memories to work with – tied to 1 song & 1 user.
• Just over 2,000 are identifiably associated with a track on Spotify
• There are roughly 100,000 songs on Spotify’s collection of Mood/Genre playlists
• Less than 700 of them are identifiably associated with a song on a Mood/Genre playlist
• There are fewer than 100 unique songs in this pool
• Not much to work with 😞
New strategy

• Plan to use the recommendation seed feature on Spotipy

• Can get a set of 10 songs that Spotify considers “similar” to a given track

• Look for direct connections to other songs in the network

• Indirect connections – two songs in the network connect to same external song

• Can possibly check for presence of recommended songs in Mood/Genre playlists. Can be (an)other variable(s) for analyzing the “seed” track.
New

Using recommendation seed feature.
Mood playlists: Happy vs Sad more like Good vs Bad mood
Mood playlists: Happy vs Sad more like Good vs Bad mood

As you can see, each playlist is has a title and a description.

The following wordcloud consists of words from every mood playlist’s description.

Take a look:
Most popular words used by Spotify in the descriptions of their “Mood” playlists:

- Not very many emotion-related words.
Adding Recommendations

- For each memory triggered by a Spotify-compatible song, ask SpotiPy to recommend 10 songs similar to the trigger song
- Save these recommendations and their seeds (corresponding trigger songs), in a DataFrame join with the DataFrame you saw last.
- Here it is again:
Data Preview

This is the data without any recommendations added. Each memory column is unique.

This has been transposed to preserve the visibility of variable names.
Adding Recommendations

• For each memory triggered by a Spotify-compatible song, ask SpotiPy to recommend 10 songs similar to the trigger song.

• Save these recommendations and their seeds (corresponding trigger songs), in a DataFrame join with the DataFrame you saw last.

• Here it is again:

• And here is what it looks like after adding recommendations:
**Data Preview**

This is the data with recommendations added*.

*I selected these columns intentionally to show a broader variety of entries in emotion frequency and in the bottom four rows. They are not in sequential order.

In the full table columns with even one occurrence of any emotion are few and far between.
Adding Recommendations

- I used these details as new variables on which I could analyze the seed tracks (“songid” row), and everything attached to them.

- That includes memory content data
  - emotional frequencies

- The following wordclouds pertain to 9 out of the 10 emotions I examined:
Word Clouds (Unweighted)  Each memory narrative that contains a word related to
the emotion of interest is pasted in the document once.

anger

fear

passion

compassion

happy

solitude

confusion

humiliation

sorrow
Word Clouds (Weighted)

Each memory narrative is pasted in the document as many times as they contain a word related to the emotion of interest.
Adding Recommendation Details

- Check for presence of recommendation songs in Mood/Genre/Action playlists
- Save the details in a DataFrame and join with the DataFrame you saw last.
- Here it is again:
Data Preview

This is the data with only the Spotify IDs of recommendations added. Each memory column is unique.

This has been transposed to preserve the visibility of variable names.
Adding Recommendations Details

- Check for presence of recommendation songs in Mood/Genre/Action playlists
- Save the details in a DataFrame and join with the DataFrame you saw last.
- Here it is again:

  - And here is what it looks like after adding recommendation details:
This is the data with recommendation details added*.

*I selected these columns intentionally to show a broader variety of entries in emotion frequency and in the bottom four rows. They are not in sequential order.

In the full table columns with even one occurrence of any emotion are few and far between.
Adding Recommendation Details

- I then experimented with recommendation-containing playlists, and their details, as variables on which to analyze seed tracks.
  - Includes Playlist’s **folder** (i.e. “mood”, “chill”, “travel”, “classical”)
  - Playlist’s **name** and Playlist’s **description**

- The following wordclouds pertain to all ten of the emotions I examined:
  - Playlist **folders** are in blue, playlist **descriptions** are in orange
Word Clouds (Unweighted)