FindMeFM Data Storyboard

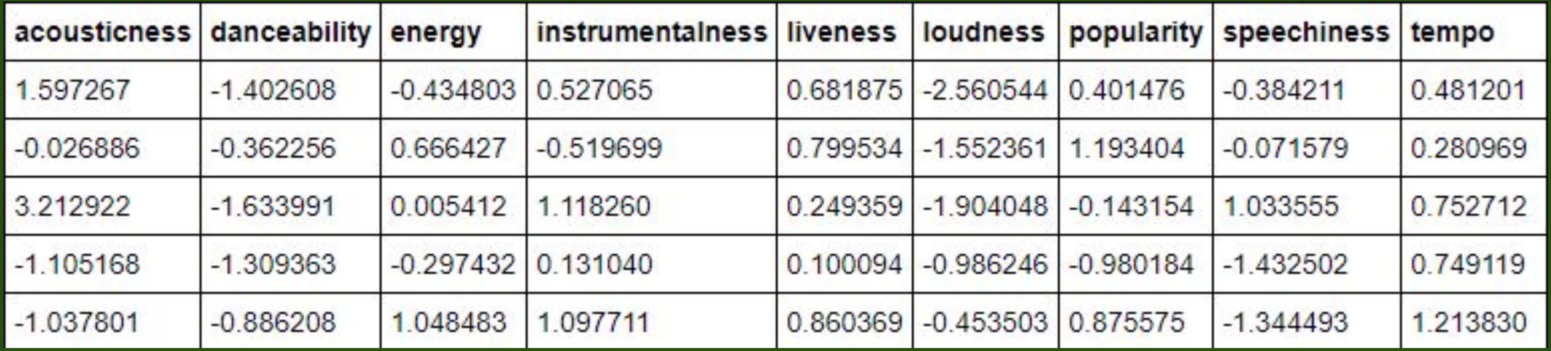
(Draft)

Question

Can machine learning predict what songs a user will enjoy based on the audio features of songs they like?

Data Sources

1. “Audio features of ~600k songs released in between 1922 and 2021” (created using Spotify API)
2. Feedback from user (“Did you like this recommendation?”)

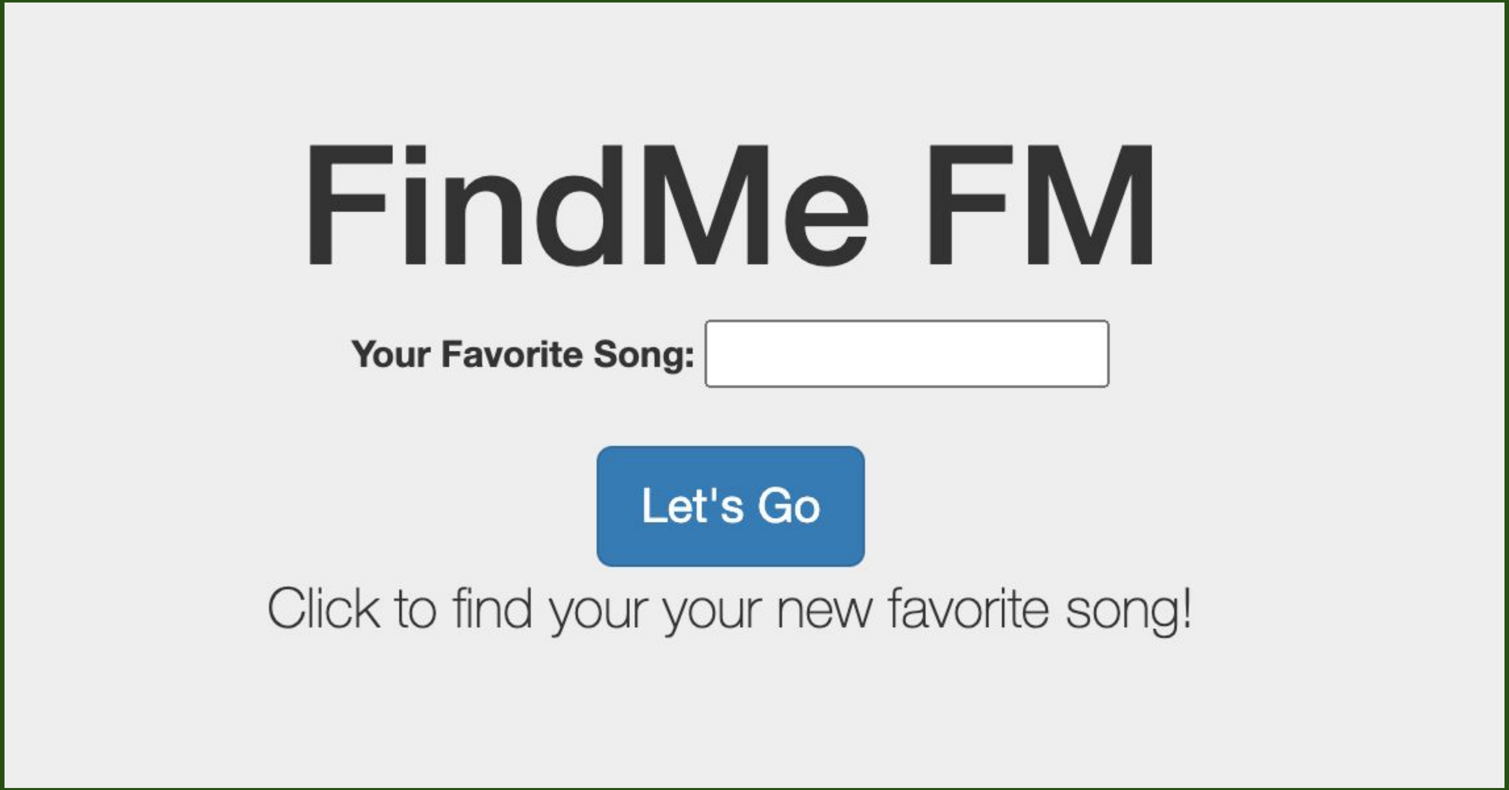


acousticness	danceability	energy	instrumentalness	liveness	loudness	popularity	speechiness	tempo
1.597267	-1.402608	-0.434803	0.527065	0.681875	-2.560544	0.401476	-0.384211	0.481201
-0.026886	-0.362256	0.666427	-0.519699	0.799534	-1.552361	1.193404	-0.071579	0.280969
3.212922	-1.633991	0.005412	1.118260	0.249359	-1.904048	-0.143154	1.033555	0.752712
-1.105168	-1.309363	-0.297432	0.131040	0.100094	-0.986246	-0.980184	-1.432502	0.749119
-1.037801	-0.886208	1.048483	1.097711	0.860369	-0.453503	0.875575	-1.344493	1.213830

Before User Input

1. Use Principal Component Analysis (PCA) to determine audio features with the largest variability
2. Build an equation that minimizes the distance between features of the input song and the recommended song
3. Build a new dataset with the recommended songs linked to the input songs

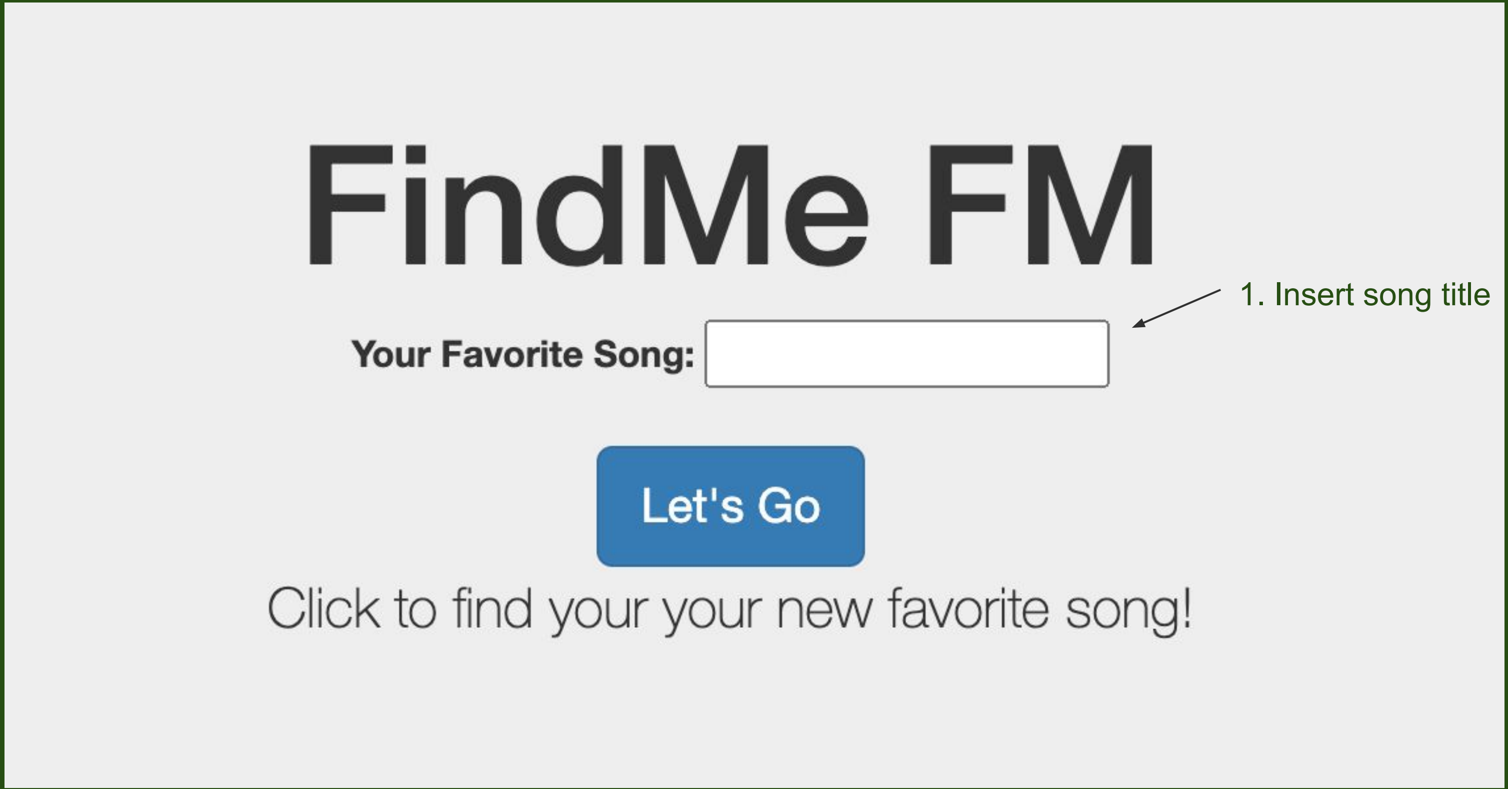
Flask App Visualization (Interactive Feature)



FindMe FM

Click to find your your new favorite song!

Flask App Visualization (Interactive Feature)



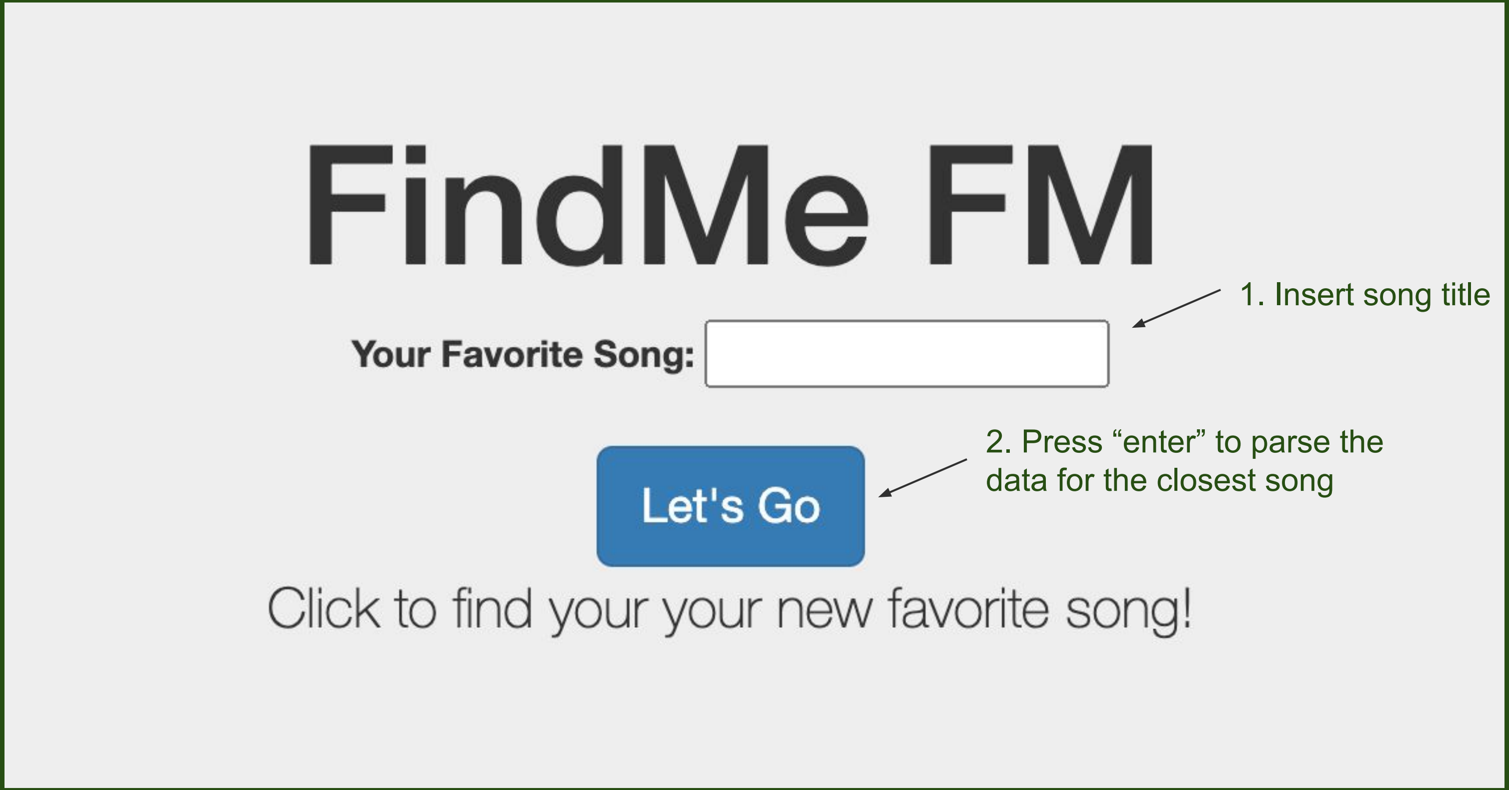
FindMe FM

Your Favorite Song:

Click to find your your new favorite song!

1. Insert song title

Flask App Visualization (Interactive Feature)



FindMe FM

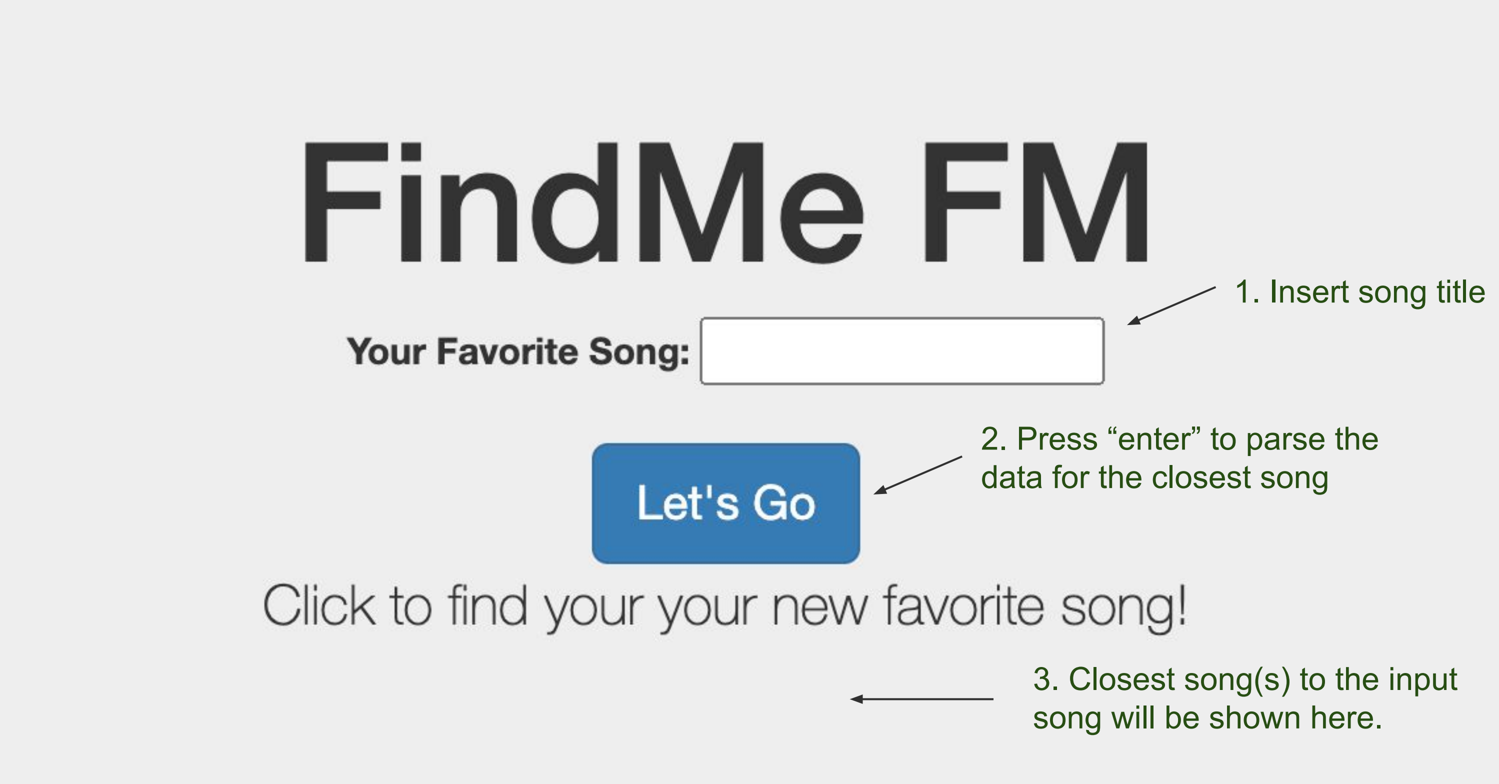
Your Favorite Song:

Click to find your your new favorite song!

1. Insert song title

2. Press “enter” to parse the data for the closest song

Flask App Visualization (Interactive Feature)



FindMe FM

Your Favorite Song:

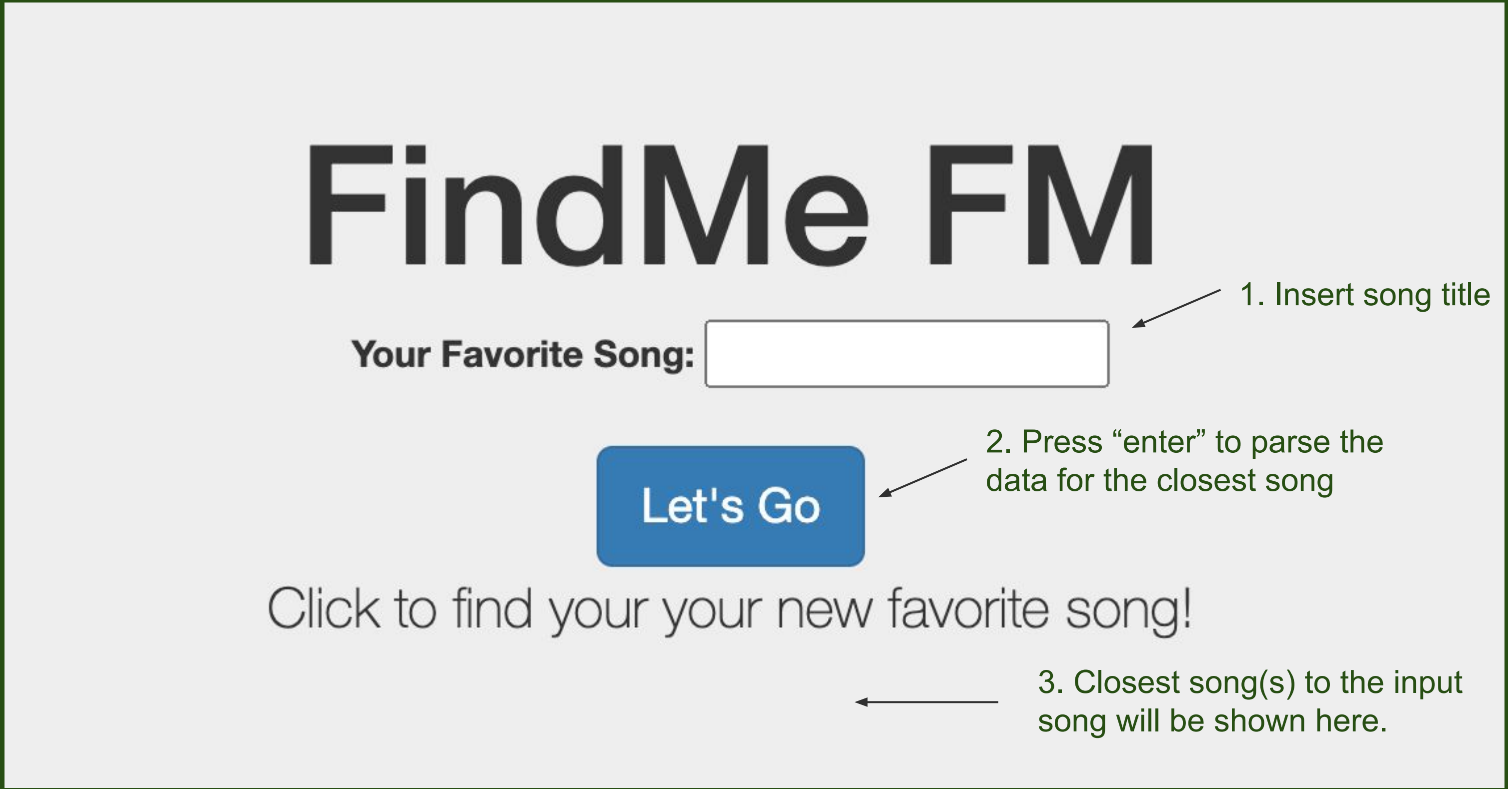
Click to find your your new favorite song!

1. Insert song title

2. Press “enter” to parse the data for the closest song

3. Closest song(s) to the input song will be shown here.

Flask App Visualization (Interactive Feature)



Click to find your your new favorite song!

1. Insert song title

2. Press “enter” to parse the data for the closest song

3. Closest song(s) to the input song will be shown here.

4. Future addition: User feedback. “Did you like this recommendation?”. We will use this to refine the recommendation model.

What we can visualize

- Formula used to determine distance between songs
- The most significant audio features (variability, etc)
- The clusters we build (K-Means/ K-Nearest Neighbor)
- Initial User Feedback Data

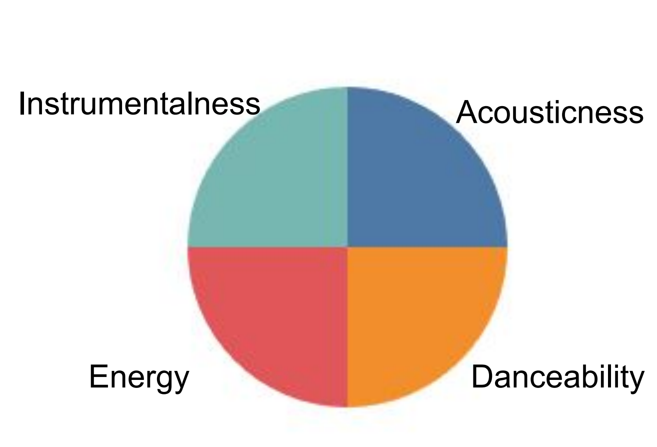
Distance Formula

Formula used to determine/minimize distance between song input and output

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

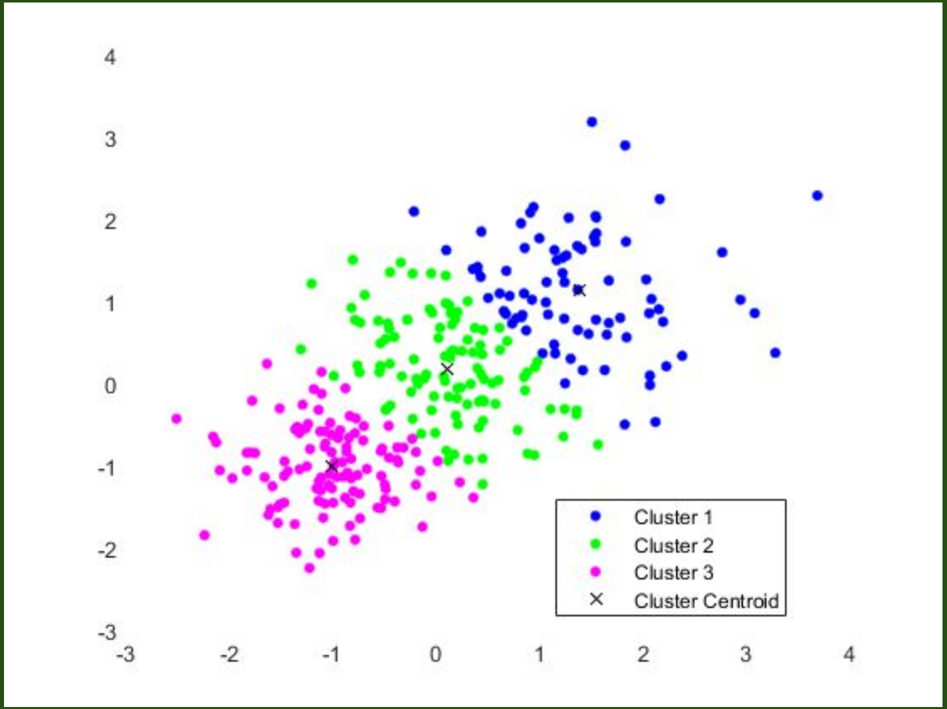
Audio Features

Pie chart to visualize the magnitude of the audio features with most variety

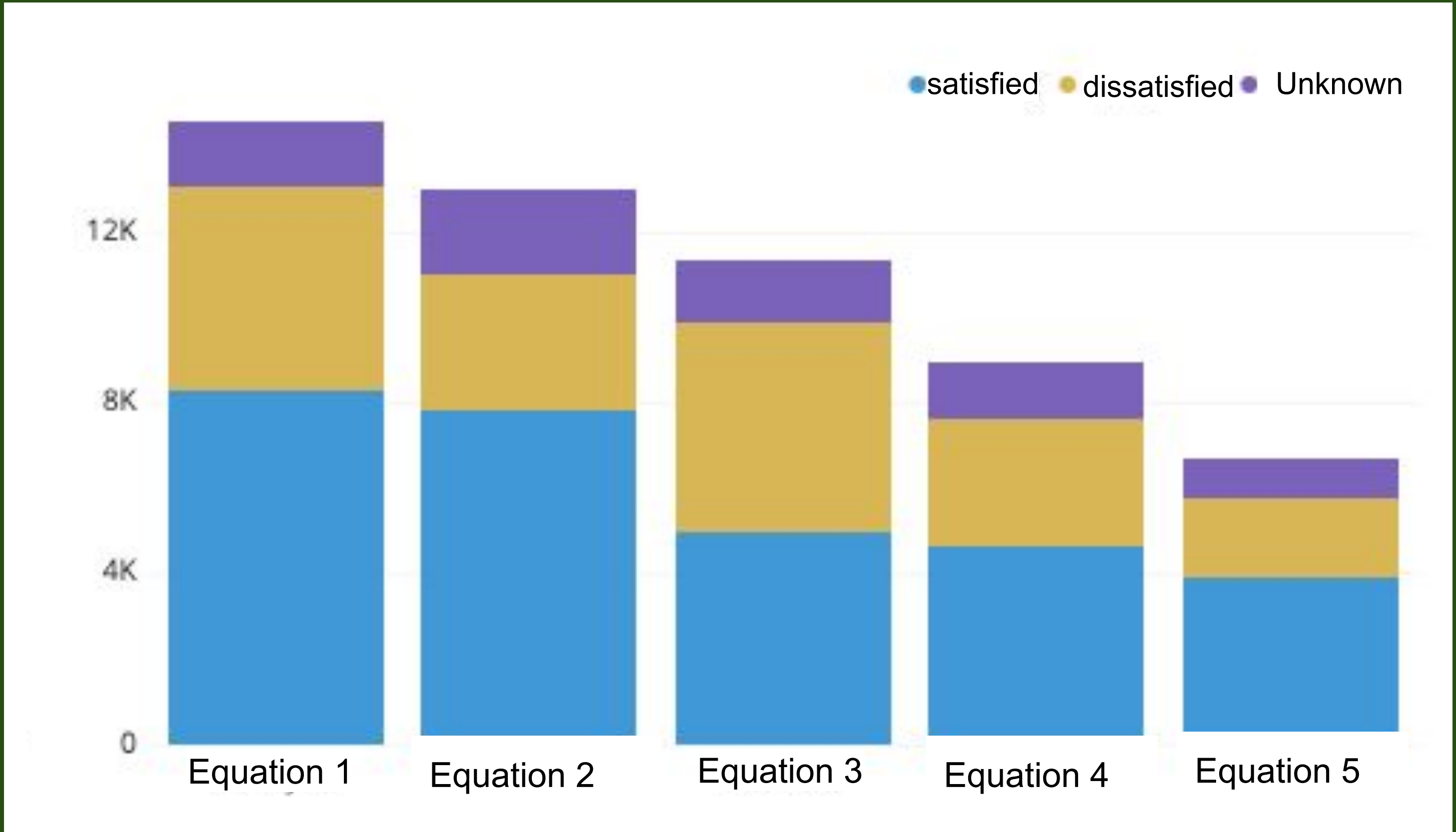


K-Means/K-Nearest Neighbor Cluster Visualization

Visualization of the clusters and centroids for each audio feature.



Initial User Feedback Data


● satisfied
● dissatisfied
● Unknown
12K
8K
4K
Equation 1
Equation 2
Equation 3
Equation 4
Equation 5