CS 171 Final Project Process Book

Team name: D#ata

Group member 1 (leader): **Max Benegas** (mbenegas@college.harvard.edu)
Group member 2: **Caitlyn Dang** (dang@college.harvard.edu)
Group member 3: **Justin Gonzalez** (justingonzalez@college.harvard.edu)
Group member 4: **Kate Schember** (kschember@college.harvard.edu)

Table of Contents

Title: Music Around the World

Abstract

What are we trying to do?
We plan to learn how the most popular music around the world differs today. We want to use Spotify metrics (listed below) to determine how different music characteristics vary by region.

Motivation:
We are all Spotify users and we're interested in seeing how our music listening trends may be similar or different from trends around the world. Listening to music is a universally loved activity, but the types of music people enjoy varies around the world. This makes for interesting, relevant data to explore and share.

Characteristics
- Danceability
- Duration (if we find something interesting)
- Energy
- Instrumentalness
- Key (if we find something interesting)
- Tempo (if we find something interesting)
- Valence (=happiness)

What goal do we have?
Our intended audience is the general population and casual music listeners. We want our audience to gain a broader understanding of contemporary world music culture. On the tech side of things, we want our visualizations to include: at least one interactive map; filtering option for region and be able to include extra detailed information for the specified regions chosen; filtering option for particular metrics (ex. energy) so users can compare regions by that metric specifically.

Planned datasets and where to acquire them?
- Spotify API
- Spotifyr (R Library) : We will use this library to access Spotify's "[Country] Top 50" playlists for all countries that have available data (~50-60 countries) and use the playlist's song characteristics for analysis and visualization.

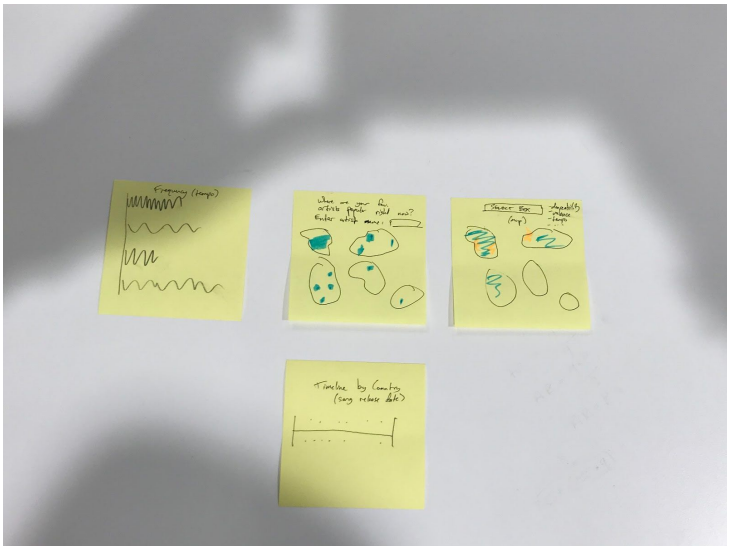
Week 2

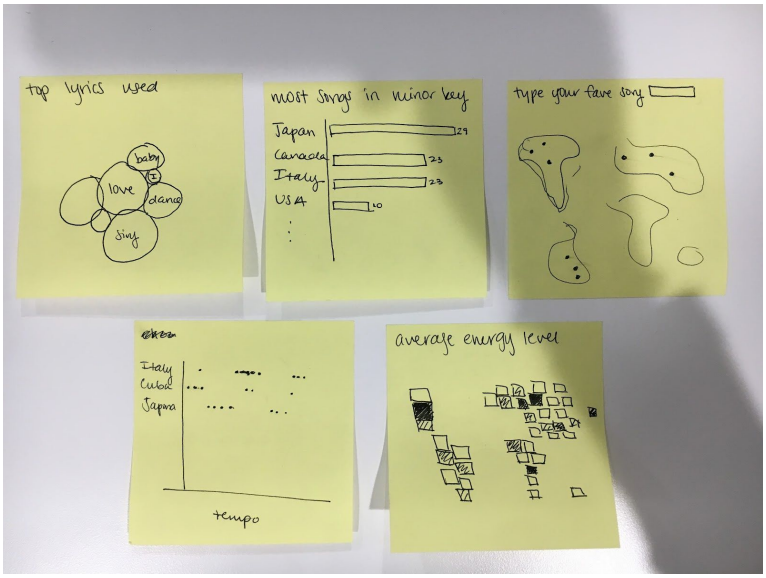
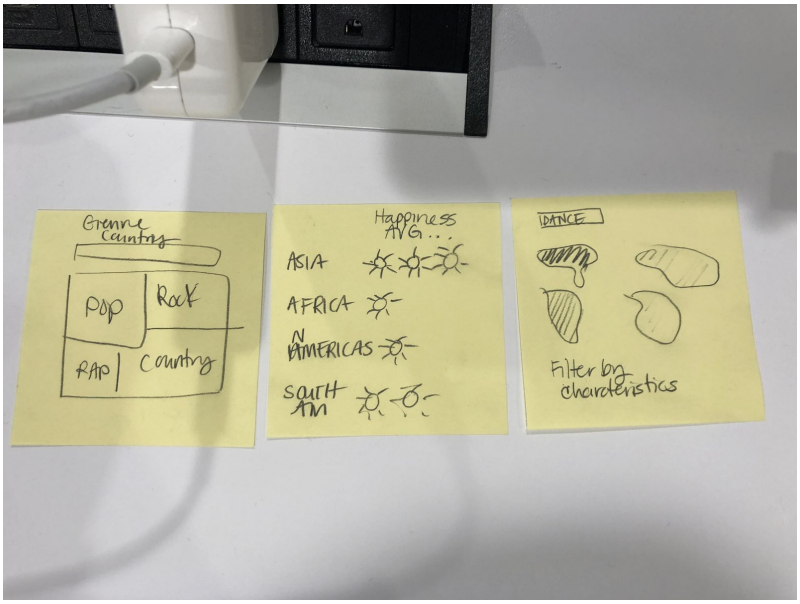
Outline:

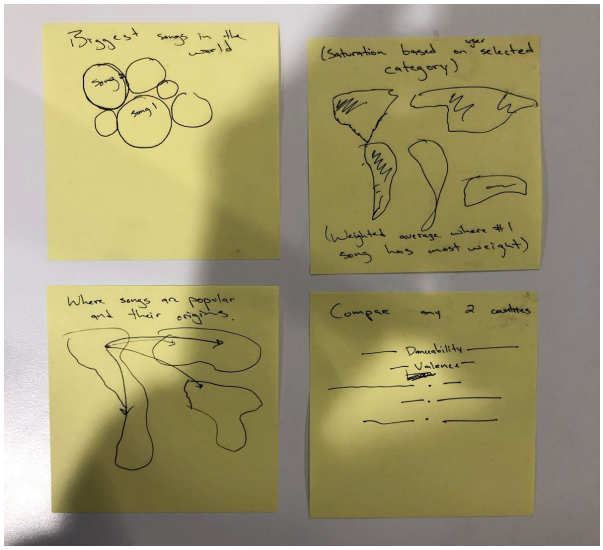
1. Analyze the client, audience, and specific goals.
    a. Client: music publication (e.g. Spotify blog, *Rolling Stone*, Pitchfork)
    b. Audience: general audience, geared towards music listeners
    c. Goals:
        i. Clearly express the difference in music listening habits and preferences between different countries across the world, including differences in:
            1. Liveliness of songs (tempo, energy, "danceability"),
            2. Tone of songs (major/minor, cheerful vs. sad),
            3. Instrumental composition (wordiness, instrumentality), and
            4. Lyrical composition (lyrical word usage and meanings).
        ii. Include facts and anecdotes relating to the data and larger trends (e.g. there's an article about how minor chords are used more in certain countries—can we see this in our data? If so, we can cite that to add interesting, smaller stories to our broader story)
        iii. Create a central, interactive map visualization that can be filtered by different data types
        iv. Make sure the audience has something specific to relate to in the visualizations/story
            1. One way we will implement this (we should think of more too) is to have a search box for songs and artists that will show up on a map or other visualization so people can see where they favorite songs are being listened
2. List at least 10 interesting questions your visualizations should be able to answer.
    a. Which countries have the most top songs in common?
    b. Which countries have the least top songs in common?
    c. How many of the top songs are lively and how does this trend change between countries?
    d. How many of the top songs are in a major key and how does this trend change between countries?
    e. Are top songs more wordy or instrumental and how does this trend change between countries?
    f. Are there noticeable trends in lyrical word usage or meaning in each country or overall?
    g. Are there any characteristics that most countries' top songs have in common?
    h. What is the characteristic that varies most between countries' top songs?
    i. Which countries have the most top songs within one genre? (e.g. which country has the most pop songs in their top playlist, most country songs, etc.)
    j. Do intra-continental countries share more trends than countries across continents?

    i. Are there other groupings?
3. Describe what kind of story you would like to tell. Come up with some candidates for your main messages.
    a. We want to convey the universality of music. Across tens of thousands of languages spanning every country in the world, music is a universal language, a human quality that transcends barriers. Every culture and subculture has its own relationship with music—different artists, trends, and messages. The four of us are deeply entrenched in American music culture, and it's easy to forget the idiosyncrasies that exist between the music cultures of different nations. With this project, we will highlight this essential principle;

Sketch:
Individually, create at least 5 different sketches for possible visualization to answer your questions.

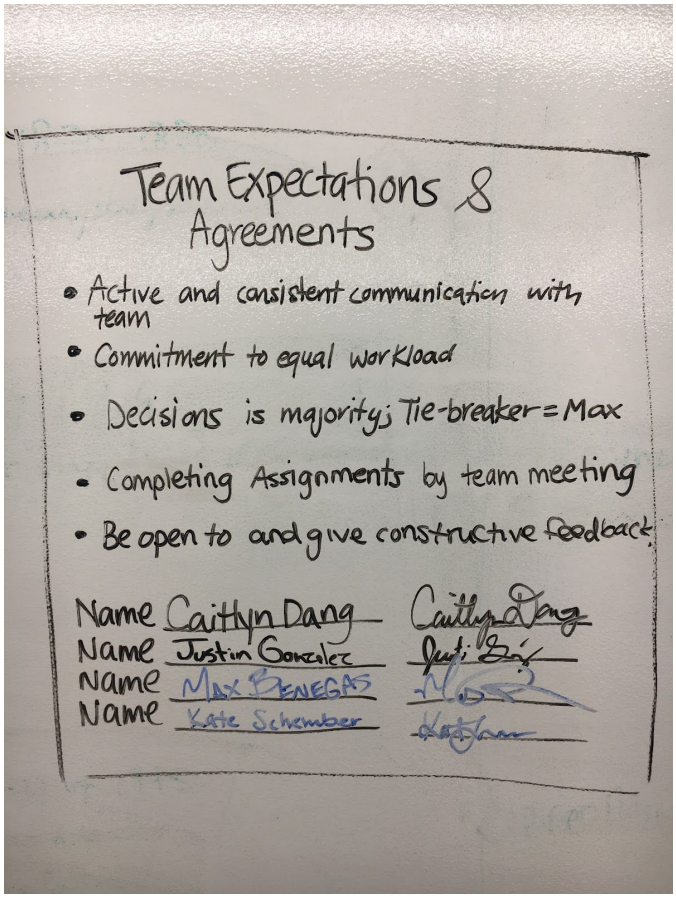
Max	

Kate	
Caitlyn	

Justin	

Project Plan

Group members who contributed: Caitlyn Dang, Justin Gonzalez, Max Benegas, Kate Schember



Definition of goals and tasks of the final project (1-2 pages)

Goals

1. Clearly express the difference in music listening habits and preferences between different countries across the world
    a. Liveliness of songs (tempo, energy, “danceability”),
    b. Tone of songs (major/minor, cheerful vs. sad),
    c. Instrumental composition (wordiness, instrumentality), and
    d. Lyrical composition (lyrical word usage and meanings).
2. Include facts and anecdotes relating to the data and larger trends
    a. For example, research/articles about how music varies (ex. How chords that are used changes from country to country)
3. Create a central, interactive map visualization that can be filtered by different data types
    a. Data types are the listed characteristics in goal #1
4. Make sure the audience has something specific to relate to in the visualizations/story
    a. For example, take the audience through a journey around the world with music

Tasks

Deadlines

Sunday, November 18th → Prototype 1 Due
Sunday, November 25th → Prototype 2 Due (95% of the project done)
Sunday, December 2nd → Project Due

1. Prototype 1 Tasks
    a. Meeting with TF Monday
    b. Aggregate data (Top 50 songs in each country, playlists created by spotify)
        i. For each country (61 countries), there is a list of 50 songs
        ii. Each song has the characteristics listed above in Goal #1
        iii. Prefer JSON, but CSV works
    c. Map out a story we want to tell
    d. Pick visualizations / Assign to each person
    e. Each person implement visualizations & creates their part of the story
    f. Put it all together in group meeting
2. Prototype 2 Tasks
    a. Meeting with TF Monday
    b. Edit story/visualization transitions
    c. Add extra interactivity
    d. Streamline cohesive design
    e. Troubleshoot any remaining issues
    f. User testing!
3. Before Project is due
    a. Make sure code is commented well
    b. More user testing

c. Tweaking small things/last minute edits
d. Should be ready to submit!

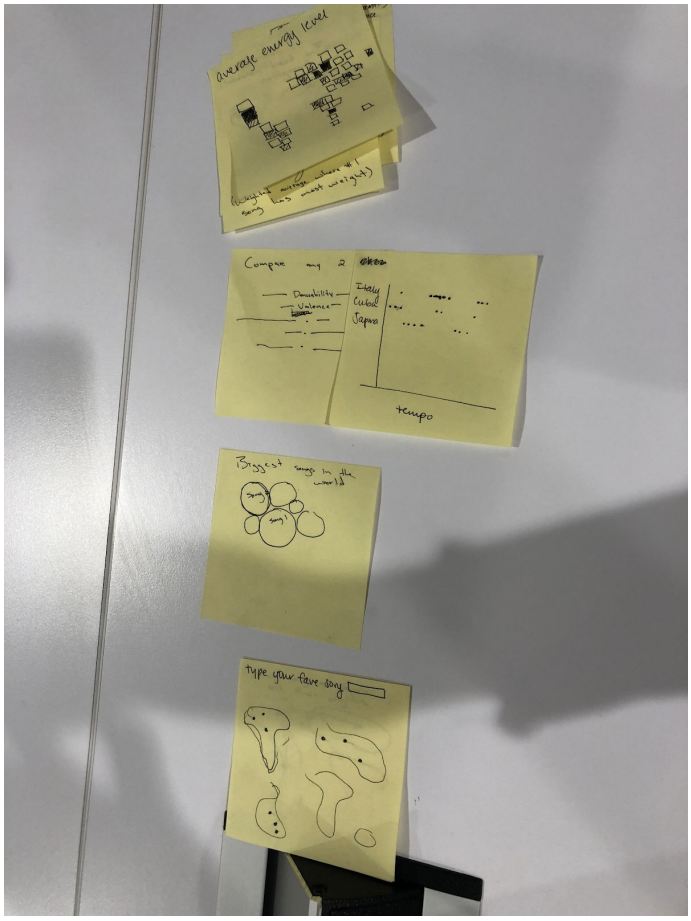
A description of your data and where you will get the data from (at least concrete ideas on where to acquire the data)
Spotify API

Spotifyr (R Library) : We will use this library to access Spotify's "[Country] Top 50" playlists for all countries that have available data (~50-60 countries) and use the playlist's song characteristics for analysis and visualization.

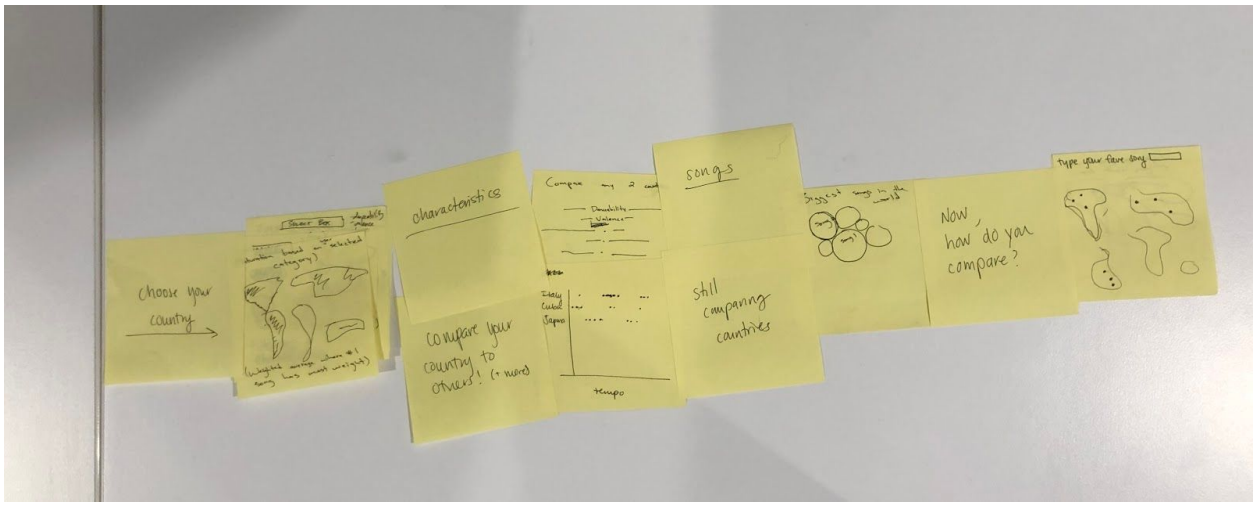
We have already aggregated our data.

At least 3 sketches of visualization ideas for your project

1. Saturation map with a dropdown of all the characteristics
2. Direct comparison between 2 countries
3. Biggest/most common songs
4. Favorite song/artist map

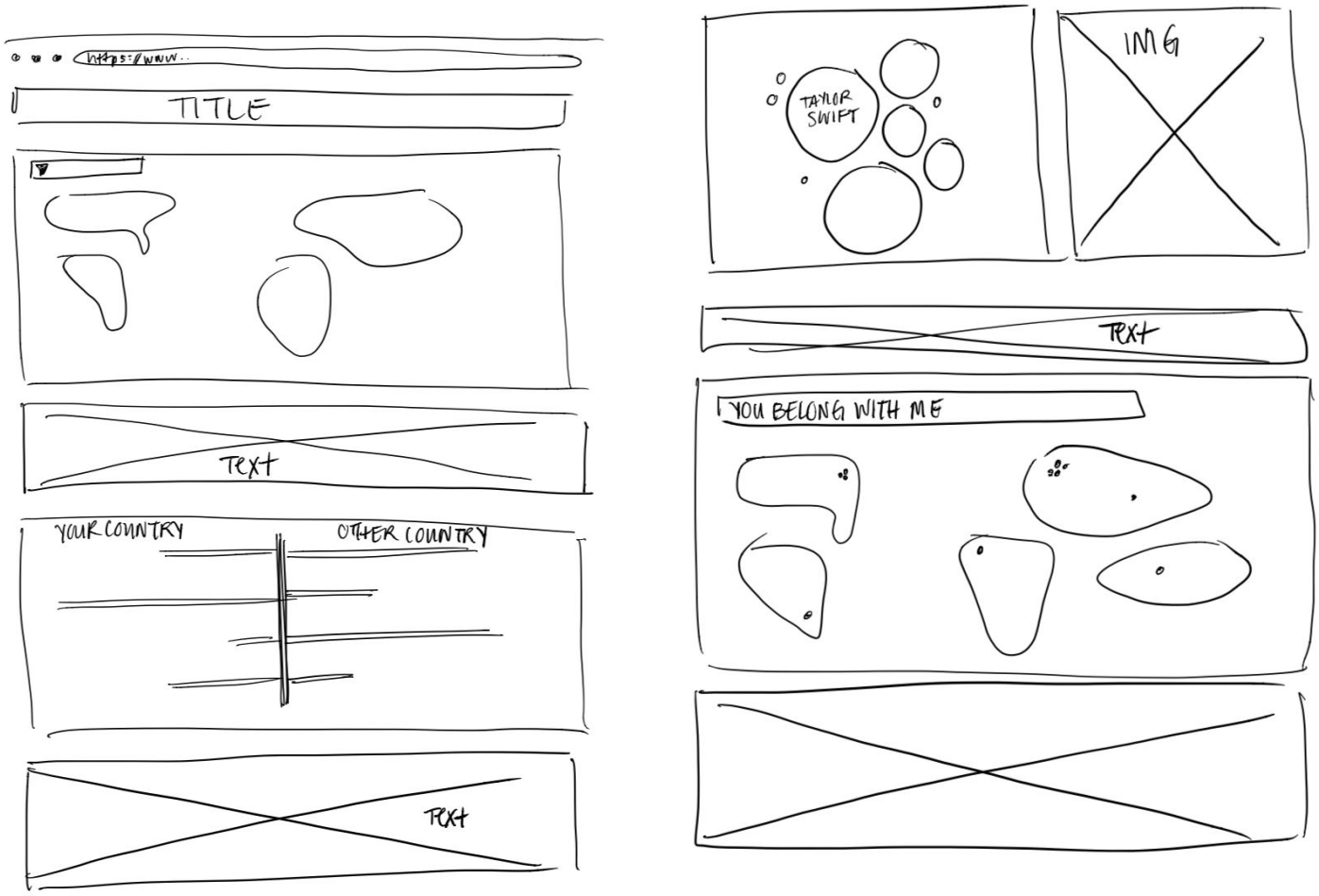


A sketch of an interaction storyboard



A sketch of your webpage layout/storytelling

Disclaimer about which countries have spotify



A project timeline (with milestones when you are planning to finish which feature)

Monday, November 12 - Wednesday, November 14:

- Map out clear story to tell
- Meet with TF (Michael) to discuss general plan and solidify story plan
- Make final decisions about which visualizations to include and more specific sketches of those
- Fix any issues with aggregated data
- Get data in a format we can all work from
- Break down visualization code work among group members

Wednesday, November 14 - Sunday, November 18:

- Implement visualizations
    - Each person must complete their assigned visualization to their best ability
- Rough drafts of story text
    - Each person must complete the text section and find images if necessary that relate to their assigned visualization

- Group meeting over the weekend
    - Put all code into one document
    - Style page
    - Add story transitions where necessary
- **Sunday, November 18: PROTOTYPE 1 DUE**

Monday, November 19 - Tuesday, November 20:
- Meet with TF to do user testing and general feedback on project progress
- Discuss what remaining issues there are and what edits should be made based on the user testing and TF feedback
- Assign specific jobs for each group member to complete for prototype 2

Tuesday, November 20 - Saturday, November 24:
- Fix remaining issues in visualizations
- Add extra interactivity elements
- Edit story and add transitions as necessary
- Make sure there is a cohesive design, make edits as necessary

Saturday, November 24 - Sunday, November 25:
- Send individual code to group and put together in one document
- **Sunday, November 25: PROTOTYPE 2 DUE** (95% of the project should be done for this submission)

Monday, November 26 - Sunday, December 2:
- Make sure all code is commented well
- More user testing
- Tweaking small elements, adding details, last minute edits
- Meet with group and ensure everything looks good
- **Sunday, December 2: FINAL PROJECT DUE**

A feature list (with must-have, good-to-have, and optional items)

Our visualization **must** allow the user to:
- View detailed data for the top 50 songs in each provided country
- Filter by various song metrics
- Easily compare data between countries

It would be **good** for our visualization to allow the user to:
- View detailed data for each song in our database
- Initially select their country or a country of interest and cater their experience/story to that selection

Some **optional** features that would be nice to have would allow the user to:

- Select a favorite song and see which country their music taste most aligns with
- View lyric data for the top songs in different countries
- See a song's country of origin and where it is currently most popular in the world

A description of team roles

Max

- Implement visualizations: main map
- Gather, clean, and update data
- Edit story
- Make final design decisions
- Style website

Justin

- Implement visualizations: country comparison chart
- Assist in writing story text
- Add website transitions
- Style website

Caitlyn

- Implement visualizations: most common songs
- Assist in writing story text
- Lead user feedback process
- Style website

Kate

- Implement visualizations: song search map
- Compile code for all visualizations
- Edit story
- Style website