WHEN MBTI MEETS SPOTIFY

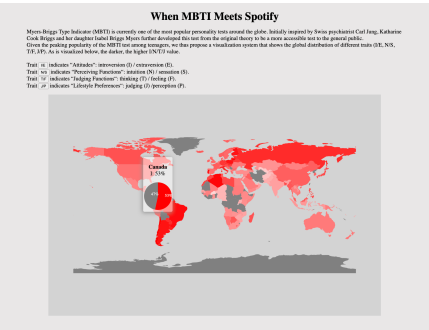
Tian Jin (tj1059), Shan Lu (sl7766), Jade Zhou (yz6213)

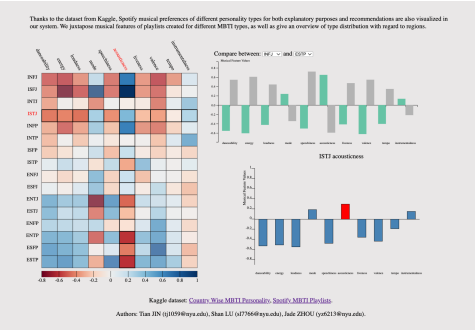
OVERVIEW

Myers-Briggs Type Indicator (MBTI) is currently one of the most popular personality tests in China. Initially inspired by Swiss psychiatrist Carl Jung, Katharine Cook Briggs and her daughter Isabel Briggs Myers developed this test that is more accessible to the general public than the original theory.

As the popularity of MBTI grows among teenagers in recent years, the music industry in China is also entering a new stage. The numbers and prices of musical festivals and livehouses surged following the media exposure of local rock and indie bands. Combining those two heated topics, we aim to promote a better understanding of our musical preferences through the lens of our MBTI personality type.

We propose a visualization system that organically shows the musical preferences of different personality types for both explanatory purposes and recommendations. Our system juxtaposes musical features of playlists created for different MBTI types, as well as gives an overview of type distribution with regard to regions.





DATA

Datasets (from Kaggle):

- The first dataset (Country Wise MBTI Personality) contains country-wise MBTI type distributions for 158 countries. It separates each MBTI type into two according to the Identity trait “Assertive” (A) and “Turbulent” (T), in other words, it only contains “INTJ-A” and “INTJ-T” rather than “INTJ”. The distribution value is collected accordingly.
- The second (Spotify MBTI Playlists) stores more than 4000 records of Spotify playlists with one of the MBTI types specified in the title. Each record has a categorical variable specifying the MBTI type for which this playlist is created along with 10 attributes describing the musical features of the playlist, including danceability, loudness, valence, etc.

Pre-processing (mainly using Pandas):

- Country Wise MBTI Personality: In order to make it consistent with the Spotify dataset, which contains pure 16 types rather than 32 categories (2 for one type as explained above), we first conducted a round of aggregation and regenerated a dataset regarding the

distribution of pure 16 types. Besides, we manually calculated the distribution of each trait (I/E, N/S, T/F, J/P) in order to visualize their distribution on geomap.
- Spotify MBTI Playlists: We grouped the playlists based on their MBTI types, and averaged the musical preferences to get aggregated data for each MBTI type. We also omitted uninformative variables like the keys in which the music is written. Finally, we standardized the values of the musical features to make them comparable on the same scale.

GOALS and TASKS

Usage Scenarios:

- For the general public: Jyoti is a senior student at NYU Shanghai who recently learned about the MBTI personality test and found that her personality type is INTJ. This week, she realized that she hasn't been updating her playlist for a while so she asked her friends, Shan and Jade, for recommendations. Not knowing where to start from, Shan and Jade decided to look for some clues from her MBTI personality type. They checked this visualization system and saw a matrix with MBTI types representing the rows and music features forming the columns. They clicked on the row for INTJ in the matrix and a comprehensive set of music features that INTJs prefer popped up. They then picked several songs that fit the description, with low danceability, high speechiness, and medium tempo. It turned out that Jyoti was happy with most of the selections. They also compared Shan's MBTI type, INFP, with Jyoti's type out of curiosity by checking the boxes for INFP and INTJ as well as one of the music features.
- For the music companies: Use MBTI as a statistical feature to help them understand the market better. For example, they can tailor marketing strategies to different audience groups upon knowing the distributions of their MBTI.

Visualization Language:

- Geomap with tooltip: Using pie charts (symbols) and polygons (area/countries) marks with color saturation channels to display the distribution density of the four MBTI traits in each country and hue channels for the pie charts.
- Heatmap with tooltip: Using area (square) marks with color hue and saturation channel to display the value range of each cell. For the bar chart tooltip, it uses vertical bars as marks with length (for magnitudes of value), position (indicating if values are negative/positive) and hue (highlight the selected bar) channels.
- Barchart allowing comparison: Using vertical bars as marks with length (for magnitudes of value), position (indicating if values are negative/positive) and hue (each MBTI type is assigned with a color) channels.

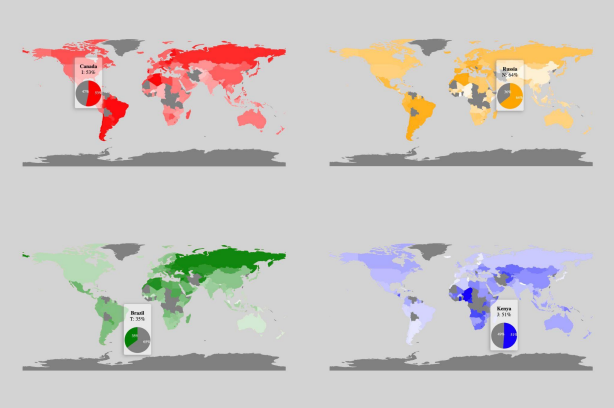
VISUALIZATION

In this project, we constructed the following three main visualizations with additional charts and selections as tooltips:

I. GEOMAP w. Tooltip (PIECHART):

As our original dataset contains regional distribution of all MBTI types, it is quite straightforward to apply a geomap for visualization. Considering that 16 types are a bit too many

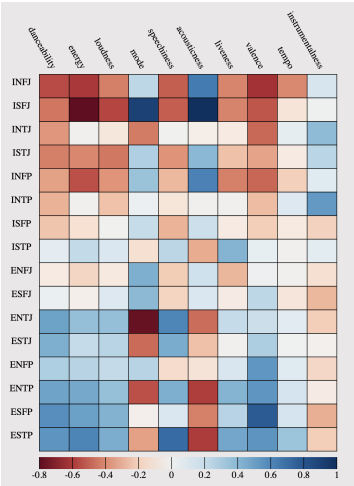
to be directly visualized and demonstrated differences in frequency, we determined to visualize the distribution of the four traits instead. As a result, it allows the audience to get a clear overview of how the four traits of MBTI distribute across the globe, meanwhile observing the differences between each country/region by clicking the four buttons to switch traits respectively. Additionally, a tooltip of piechart is also applied whenever the mouse hovers over a country/region, which better indicates the percentage of distribution at this place.

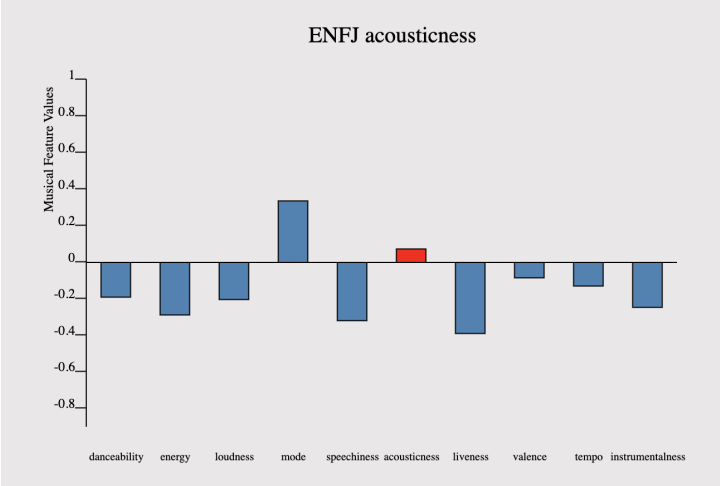


II. HEATMAP w. Tooltip:

This is the main visualization that displays the relationship between MBTI types and their musical preferences. The x-axis enlists the musical features provided by the Spotify dataset. The y-axis sits all 16 MBTI types. Each cell in the matrix represents the average, standardized musical taste of an MBTI type in terms of a specific musical feature, with a diverging color scale serving the channel.

The heatmap allows for interaction. When one hovers the mouse over any cell, the row and column to which the cell belongs will be highlighted with a thickened stroke and a larger opacity compared with other cells. The names (i.e. MBTI and musical features) of the row and column will also turn red to make clear what the cell refers to. A tooltip pops up on the right showing a barchart of all the musical features for the selected MBTI type, with the selected musical feature highlighted in red. We did not choose the opposite because we believe that the general public will be more interested in their own MBTI types, instead of the comparison of a specific musical feature for all MBTI types.

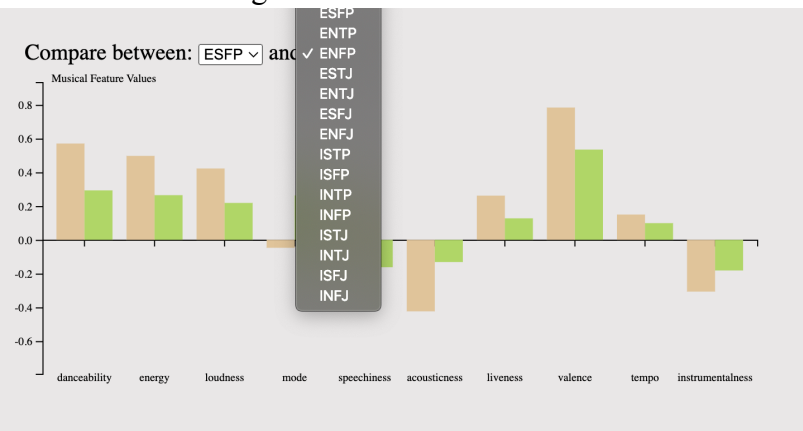




III. BARCHART allowing Comparison:

This visualization is a barchart with two dropdown menus where the users can interact by choosing two MBTI types to compare the musical feature values. This would update the barchart to show two bars for each selected option. The group bar chart makes it easy for comparison that the users can easily see which MBTI has higher values in some musical features.

The x-axis represents the musical features,with negative values extending downwards and positive values upward. The y-axis represents the values of these musical features. The height of each bar shows the magnitude of the values, and each MBTI type is assigned a color using a color scheme. The color encoding helps users to differentiate between different personality types and notice the changes.



REFLECTION

Throughout the development process, we remained committed to our original proposal, and we successfully implemented all three complex views as initially proposed. This allowed us to effectively cooperate while creating the system.

Yet, there is still room for improvement. One part we identified is that we should include detailed explanations for each musical feature. Although most of the features are easy to understand by their names, it would be clearer if a detailed description is provided next to the visualization.

Besides, issues also exist when we attempt to combine each other's work into one project. As our geomap and the interactive barchart were written in pure HTML while the heatmap used React, we realized that we needed to modify the code by wrapping all the functions into a single function and then exporting it, which was quite demanding and time-consuming. Eventually, our solution was to directly put the .js files into the dist directory so that they could directly work in index.html. Yet it also posed new problems, since we found it difficult to organize the position of all divs at the same time — as the screen size changes, the position would shift a bit. As a reflection, we could have agreed upon a thorough framework for the project at the very beginning using a consistent package, rather than part in HTML while part using Reach.

One pity that we have after finishing this project lies in the fact that the Spotify dataset lacks information on regional distribution regarding each MBTI type with certain musical features. In other words, if we have that information, we could link the geomap to the heatmap and thus further visualize the relationship between each type, each musical feature and each country.