Project Progress Report 2

(due May 24th 11:59p.m)

You can start working on the project once your report is accepted and graded by your TA. The entire final project is worth **35%** of your final grade and this report accounts for **10%**. This project is done individually.

Submission Guideline

Download this google doc, fill the table. **Type** your answers, no handwritten answers will be accepted (except for the very last question). Submit it in **PDF** format on Gradescope.

If you need some inspirations please feel free to take a look at:
Showcase of Information is Beautiful Awards
Bloomberg Year In Graphics Review
The Pudding
The New York Times

Project Guidelines

Note: The guideline has been further clarified from Progress Report 1, so double-check whether your dataset choice still satisfies the updated guideline below.

1. You may use more than one dataset, however, regardless if you use one or multiple datasets, your visualizations must make use of at least three following data types - **link, position, and attribute.**
2. You cannot use any dataset from the class (Labs, Assignments, Lecture Exercises)
3. You can make your own dataset (Web scrape etc.) provided point 1. is satisfied.

Part 1 - Story and Narrative

Link to the dataset	https://www.kaggle.com/datasets/andrewmvd/spotify-playlists
Example item from the dataset	"user_id", "artistname", "trackname", "playlistname" "9cc0cfd4d7d7885102480dd99e7a90d6","Elvis Costello","(The Angels Wanna Wear My) Red Shoes","HARD ROCK 2010"
Story you want to deliver	(a story should be in a form of a list of facts, insights, and messages - refer to the lecture slide) Since Spotify offers a vast collection of music, it contains diverse genres and playlists created by users worldwide. In my project I delve into the intriguing world of Spotify's organization of music, focusing on how artists are grouped in playlists based on genres by asking several questions: What are the most common genres found in user-generated playlists on Spotify, and how do they vary across different user locations and languages? Facts: • Analyze the genre distribution in user-generated playlists. • Identify the top genres and their frequencies. • Compare genre distributions across different user locations and languages. Insights: • Use a choropleth map to show the variation in genre preferences across different user locations. Messages: • The diversity of music preferences among Spotify users reflects the cultural and linguistic differences across various regions. Can we see how several different artists are classified and how their genre changes in different cultures/locations, Facts: • The genre distribution of my selected artists across different user cultures/locations.

	Insights: • Use a world map to show the most popular playlist genre the artist appeared in in each country Messages: • Music genres are not fixed and can be perceived differently in various cultures, resulting in diverse genre classifications for the same artist. Can we identify the rise of certain genres/artists or the influence of specific artists? Facts: • Identify changes in the popularity of genres by looking at how popular a few genres are within the entire dataset of playlists, highlighting the most popular right now (top 3) and those that are rising (top 7). Aggregate the top 7's from each country and present the most popular (top 3) and the rising (remaining 4). • Track the presence and influence of specific artists in playlists. Choosing artists that are categorized as in the remaining 4 genres, then tracking how many countries in which they appear in over x% of the playlists. Insights: • Use a bar chart or area chart to visualize the most popular genres and the rising genres by frequency of playlist dominated (> 50%) by that genre. • Use a force graph to display the percentage of playlists that include a certain artist from genres of interest. Message: • Playlists reflect the evolving tastes and preferences of users, showcasing the impact of artists and the emergence of new musical trends.
Describe your target audience.	(using the questions the lecture slide listed) The audience is anyone interested in music trends and emerging genres/artists. Familiarity: Spotify is a popular service and music/genres are common knowledge Do they care: Very likely, it is a very popular and widely used service that offers a lot of insight into cultural trends.

	Take aways: How different cultures, spread across multiple countries/regions can view the same art differently and how their demands can push forward new global rising trends. Knowledge on visualization: They are expected to habe seen most of the structures I will be using, most will be standard and not require extensive prior knowledge to get information/meaning from. They can encounter the visualizations through the website/page it will be displayed on. Mathematical background: There is very little mathematics needed to understand the facts and figures I aim to produce, with things like percentages and ratios being the likely highest level of numbers present. Device: This would best work on a computer screen like a webpage so as to gain the full accessibility, it will likely offer a similar but slightly inferior experience on a mobile device.
The goal of your project outcome. And why?	My project is exploratory due to the aim to discover patterns, relationships, and insights within the Spotify dataset. I am trying to explore and investigate the organization of music on Spotify by analyzing the categorization of artists in playlists based on genres. I'm interested in understanding the differences across playlists created by users from different locations and languages. This involves examining patterns, trends, and potential variations in genre categorization, playlist composition, and user preferences.
Narrative structure you plan to use	The structure will be more interactive than linear Akin to Martini-glass.
Elaborate your choice of narrative structure.	I plan to lead in with my initial questions about the dataset, choose the artists to follow and genres to track then display the data and present some loose conclusions about the topics based on the results, but the main goal is for the reader to see the displays. It will be informative enough that the reader to see the trends for other genres and one or two other artists than the ones I choose to follow.

Narrative genre you plan to use	Likely a partitioned poster where the poster is a large map of countries.
Elaborate your choice of narrative genre.	The map will be a choropleth with information in countries about their top genres based on playlists and display their frequencies. There will also be an option to look at a few artists and once you select them, you should be able to look at the same countries but this time you can see what genres the artist has most frequently been categorized as within the country. Below the choropleth should be a bar chart or area chart to visualize the most popular genres and the rising genres by frequency of playlist dominated (> 40% of songs inside the playlist are of that genre) by that genre.

Part 2 - Outline

Story you want to deliver	Since Spotify offers a vast collection of music, it contains diverse genres and playlists created by users worldwide. In my project I delve into the intriguing world of Spotify's organization of music, focusing on how artists are grouped in playlists based on genres by asking several questions: What are the most common genres found in user-generated playlists on Spotify, and how do they vary across different user locations and languages? Facts: • Analyze the genre distribution in user-generated playlists. • Identify the top genres and their frequencies. • Compare genre distributions across different user locations and languages. Insights: • Use a choropleth map to show the variation in genre preferences across different user locations. Messages:

	• The diversity of music preferences among Spotify users reflects the cultural and linguistic differences across various regions. Can we see how a few selected artists are classified and how their genre changes in different cultures/locations, Facts: • The most popular genre of my selected artists across different user cultures/locations. Insights: • Use a world map to show the most popular playlist genre the artist appeared in in each country Messages: • Music genres are not fixed and can be perceived differently in various cultures, resulting in diverse genre classifications for the same artist. Can we identify the distributions of a few artists within the dataset? Facts: The genre distribution of my selected artists across different user cultures/locations. Insights: • Use a bar chart or area chart to visualize the distribution of my selected artists as a percentage. • Use a stacked bar chart to display the percentage of genre classification for each artist with a tooltip to display the different genres present in each bar. Message: • Playlists reflect the evolving tastes and preferences of users, showcasing the different amount of ways playlist makers viewed an artist.
Specifications on each plot **in the order of how you lay out on your project**	(for each plot, include 1) clear task abstraction, 2) attributes used, 3) marks, 4) channels, and 5) how this plot adds to the story) Plot 1: **Clear Task Abstraction**: Visualize the genre distribution in user-generated playlists across different user locations using a choropleth map.

	Attributes Used: User locations (geographic data) and genre frequencies (quantitative data). **Marks**: Choropleth map, where each location is represented by a polygonal shape (e.g., country) and text label based on the frequency or proportion of the genres within each country. **Channels**: Geographic Location: Encoded using the position and shape of the polygons on the map. Text labels: Encoded to display the top 7 most common genres for each country. **How This Plot Adds to the Story**: By using a choropleth map to visualize the genre preferences across different user locations, we can highlight the cultural and linguistic diversity reflected in the music choices of Spotify users. Plot 2 (on the same choropleth): **Clear Task Abstraction:** Visualize how several different artists are classified and how their genre changes in different cultures/locations using a choropleth map. The goal is to understand the variation in genre classification for different artists across various regions and cultural backgrounds. **Attributes**: Artist Name- The name of the artist whose genre classification is being analyzed. Country/Location- The geographic regions or countries where the artist's music is popular or streamed. **Marks**: Text Labels will be used as marks on the choropleth map. Each text label represents the name of an artist and shows the most popular playlist genre the artist appeared in. **Channels**: Position: The position of the text label on the choropleth map represents the location or country where the artist's genre classification is being analyzed. Each text label is positioned near the corresponding country or region. Text: Same as the Text labels in Marks **How This Plot Adds to the Story**: This allows viewers to identify specific artists and understand how their music is categorized in different regions. It adds depth to the story by highlighting the diversity in genre classification for different artists, offering insights into cultural influences and preferences. Plot 3: **Clear Task Abstraction:** Identify the distribution of genres and the amount of differing ways the artists were perceived. Additionally, it aims to analyze the presence and influence of

	specific artists across different countries by looking at the amount of different classifications. **Attributes**: Genres- The different musical genres present in the playlists. Artists- The specific artists whose presence and influence are being tracked. Countries- The countries or regions where the playlists are sourced from. **Marks**: Bar Chart will be used to visualize the popularity of genres and the rising genres. The height or area of the bars represents the frequency or prevalence of playlists with a specific genre. A Stacked bar chart will show how prevalent our specific artists are within each country. **Channels**: Position: The position of the bars on the bar chart represents the genres. The higher the position, the more popular or prevalent the genre. Color: The different colors/sections of the bars allow the user to see how many different genres users thought the artist fell into. **How This Plot Adds to the Story**: By using a bar chart or area chart, this plot visualizes the popularity of artists and their spread of genres. It allows viewers to identify the most popular genres for each country and observe the rise of certain genres. The stacked bar chart shows the reader how multi-faceted an artist may be.
Elaborate the choice of their marks and channels for each vis	A map is a better vessel for visualizing international data as it provides familiar shapes for contextualizing the data. Bar chart/ stack bar chart are explained above.

Following sample answer about a single plot shows how detailed your answers to part 2 should be.

1. Plot 1
    1) Task: This chart a) analyzes trend between Height and Weight of patients with heart diseases and b) locates outliers within the patients

2) Attributes: Height, Weight
3) Marks: point mark
4) Channels:
    - aligned vertical position channel for Height
    - aligned horizontal channel for Weight
5) How this plot adds to the story:
   My visualizations aim to deliver health characteristics of patients with heart disease. This plot will provide more specific insights on Height and Weight.

Part 3 - Prototype

Provide a photo or screenshot of your prototype. A prototype should depict how you place different components of your visualization. You may use pen-paper, or using tools like excalidraw, figma etc.

A basic, barebones sample prototype for this project