DECLARATION: I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at http://www.tcd.ie/calendar. I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write.

Report: High-Dimensional Exploration of Music Genres

1 Introduction and Task Abstraction

Music streaming services typically categorise tracks using broad genre labels such as "Pop" or "Rock", which fail to capture finer emotional nuances. A listener seeking high energy music might find suitable candidates across Electronic, Metal, or even Classical repertoires; yet current interfaces silo these tracks into disconnected categories. This project addresses that limitation by organising the Spotify library according to continuous *audio features* (Valence, Energy and Danceability) rather than relying solely on genre metadata.

The resulting visualization is designed to support two primary analytical tasks:

- **Exploration:** Identifying clusters of songs that share sonic characteristics despite different genre labels (for example, tracks characterized by sadness in both Pop and Folk).
- **Comparison:** Comparing the audio signatures of different genres to understand their structural differences.

2 Dataset and Complexity

The dataset contains **114,000 tracks**, each with 11 continuous audio features. To maximize the utility of the visualization, I performed the following preprocessing steps (`clean_data.py`):

- **Meta-Genre Grouping:** The raw data contained over 100 specific subgenres. I aggregated these into 10 broad categories (for instance, mapping 'techno', 'house', and 'dubstep' into 'Electronic/Dance') to ensure color encodings were distinguishable across the visualization.
- **Feature Selection:** I focused on five key psychological features: Valence, Energy, Danceability, Acousticness, and Instrumentalness.

Dataset Sample (Head):

Track Name	Artist	Val.	Egy.	Dnc.	Aco.	Ins.	Pop.	Meta-Genre
Comedy	Gen Hoshino	0.71	0.46	0.67	0.03	0.00	73	Country/Folk
Ghost - Acoustic	Ben Woodward	0.26	0.16	0.42	0.92	0.00	55	Country/Folk
To Begin Again	...Michaelson	0.12	0.35	0.43	0.21	0.00	57	Country/Folk
Can't Help...	Kina Grannis	0.14	0.05	0.26	0.90	0.00	71	Country/Folk
Hold On	Chord Over...	0.16	0.44	0.61	0.46	0.00	82	Country/Folk

Feature Definitions: The visualization relies on the following audio features extracted by Spotify's API:

- **Valence (0.0 to 1.0):** A measure of musical positiveness. High valence indicates happy or cheerful music, while low valence indicates sad or depressed tones. This feature serves as the horizontal axis in the Mood Map.
- **Energy (0.0 to 1.0):** Represents a perceptual measure of intensity and activity. High energy tracks typically feel fast, loud, and noisy. This feature serves as the vertical axis in the Mood Map.
- **Danceability (0.0 to 1.0):** Describes how suitable a track is for dancing based on a combination of tempo, rhythm stability, beat strength, and overall regularity.
- **Acousticness (0.0 to 1.0):** A confidence measure of whether the track is acoustic, with higher values indicating greater acoustic character.
- **Instrumentalness (0.0 to 1.0):** Predicts whether a track contains vocals. Values near 1.0 suggest purely instrumental tracks (such as Classical), while values near 0.0 suggest vocal-heavy tracks (such as Rap).
- **Popularity (0 to 100):** A measure of track popularity on the Spotify platform, with 100 being the most popular.

- **Meta-Genre:** A categorical feature engineered by grouping the original 114 specific genres (such as 'salsa' and 'techno') into 10 broader categories to facilitate visual comparison.

This scale introduces two specific challenges:

- **Performance versus Scale:** Rendering all 114,000 points in a browser causes significant latency (over one second) and severe overplotting, which degrades the user experience.
- **Class Imbalance:** The raw dataset is heavily skewed, with approximately 25,000 Electronic tracks compared to only 3,000 Hip-Hop tracks.
- **Solution (Balanced Sampling):** I implemented a stratified sampling strategy that selects exactly 1,500 tracks per genre. This produces a final dataset of 15,000 points, ensuring minority genres receive equal visual representation while maintaining high performance (latency below 0.1 seconds).

Significant preprocessing was necessary to prepare the data. I used Python to clean the dataset and scikit-learn to normalise numeric features. The complexity lies not just in volume but in the semantic ambiguity of music genres. This visualization aims to resolve that ambiguity by focusing on continuous audio features rather than categorical labels alone.

3 Encodings and Design Rationale

The dashboard follows Shneiderman's mantra of "Overview first, zoom and filter, details on demand" [2] using three linked views arranged in a vertical flow: one primary chart spanning the top, and two supporting charts positioned side by side below.

A. Mood Map (Primary View): Valence versus Energy

- **Rationale:** I mapped Valence to the horizontal axis and Energy to the vertical axis because these features correspond to the *Circumplex Model of Affect*, a well-established psychological model for emotion [1]. This creates an intuitive semantic map where users can locate music by emotional character: happy music occupies the upper right (high valence, high energy), while sad music occupies the lower left (low valence, low energy).
- **Encoding:** A scatter plot spanning 1400 pixels displays all sampled tracks. I applied the Tableau20 color scheme to ensure the 10 genre categories remain visually distinct.

B. and C. Interactive Detail Views The views are linked through brushing interaction. Selecting a region in the Mood Map simultaneously updates the two charts below:

- **Genre Composition (Bar Chart):** Shows the distribution of genres within the current selection. This often reveals unexpected diversity; for instance, the "Happy" quadrant typically contains substantial amounts of Latin and Electronic music alongside the expected Pop tracks.
- **Feature Profile (Line Chart):** Displays the average values of five audio features for each genre in the selection. This provides a multidimensional "fingerprint" that allows users to compare genres quantitatively. For example, Metal tracks typically exhibit high energy but low danceability relative to Pop tracks in the same emotional quadrant.

Interactive Mechanisms: The visualization implements two complementary interaction techniques. First, **rectangular brushing** allows users to select arbitrary regions in the Mood Map by clicking and dragging; the selected area is highlighted with a semitransparent grey overlay, while unselected points fade to 15% opacity. This focused attention mechanism makes patterns within the selection immediately apparent. Second, **legend filtering** enables users to toggle individual genres on or off by clicking legend items. Multiple genres can be selected simultaneously, allowing for direct comparison of specific categories (for example, comparing only Rock and Classical). These interactions update all three views in real time (latency below 100 milliseconds), supporting fluid hypothesis testing without interrupting the user's analytical flow. The combination of spatial selection (brushing) and categorical filtering (legend) provides flexible exploration across both continuous and discrete dimensions of the data.

This composition of multiple chart types constitutes a novel visualization approach. Rather than employing a traditional grid layout, the design emphasizes the Mood Map as the primary navigation tool, with supporting charts providing quantitative detail on demand. The brushing interaction tightly couples these views, enabling seamless exploration. By grounding the primary axes in psychological theory, the visualization provides a semantically meaningful map of the musical space.

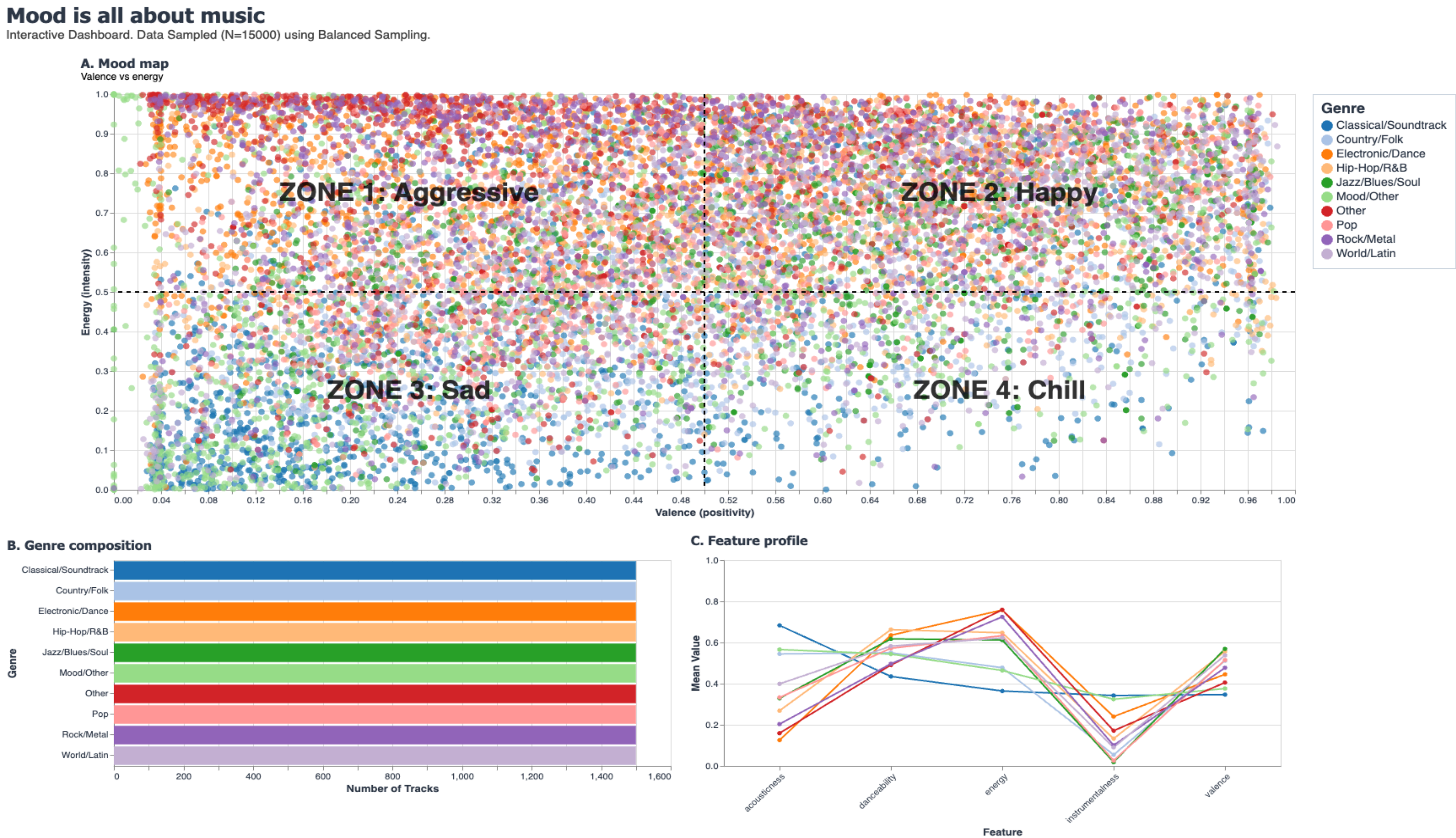

Fig-01: Mood is all about music

(A) Mood Map: A semantic scatter plot mapping Energy vs. Valence to reveal emotional zones. (B) Genre Composition: A bar chart showing the genre breakdown of the selected tracks. (C) Feature Profile: A parallel coordinate plot displaying the audio signature (e.g., high energy, low acousticness) of the selection. Note: The Brushing & Linking interaction is demonstrated by the selection in the "Happy" quadrant (Zone 2), filtering views B and C.

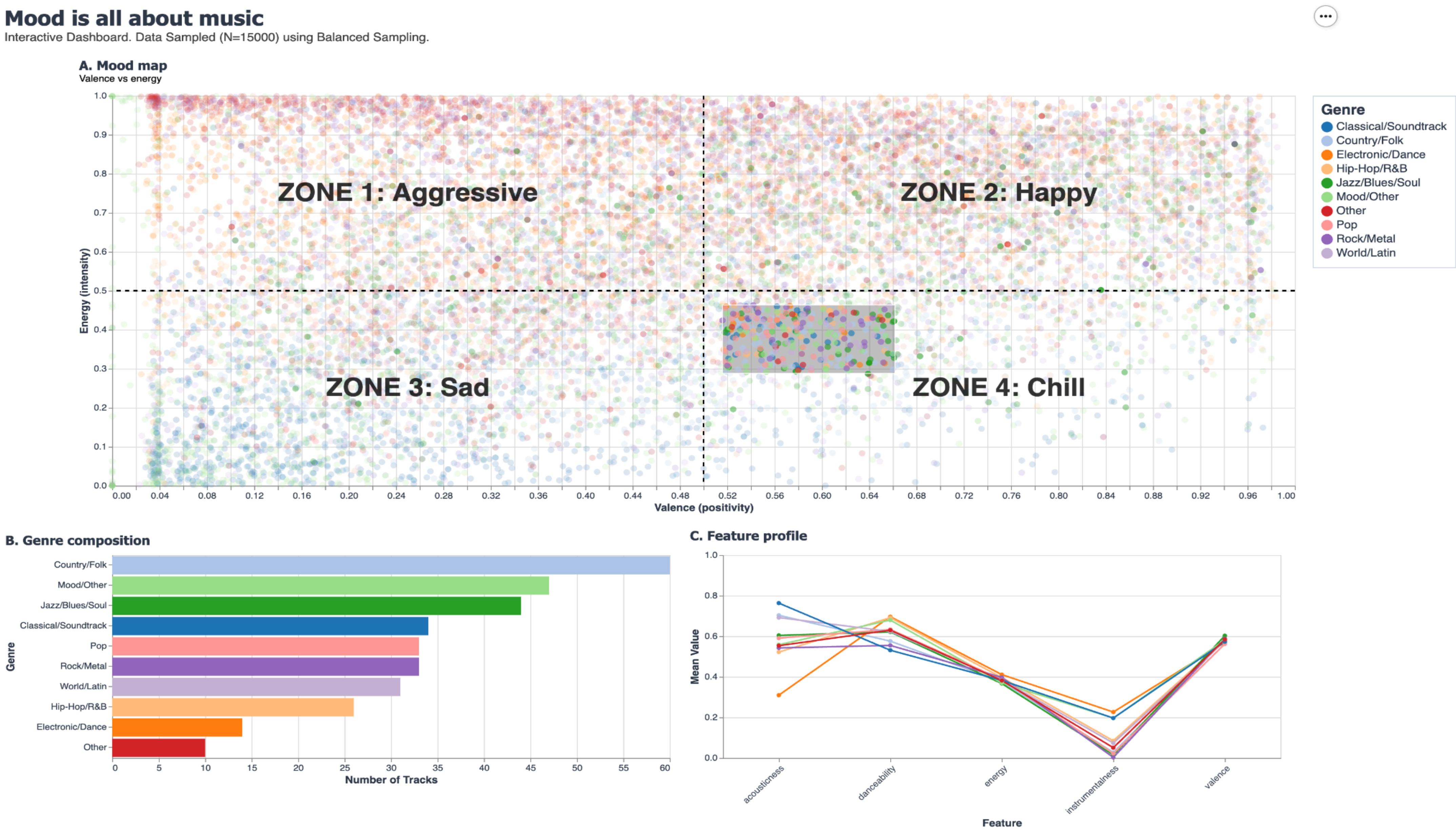

Fig-02: Mood is all about music

(A) Mood Map: A semantic scatter plot mapping Energy vs. Valence, where any selection made on the plot now drives updates across all views.(B) Genre Composition: A bar chart that reflects only the genres of the currently selected tracks. (C) Feature Profile: A parallel-coordinates plot showing the audio signature (e.g., high energy, low acousticness) of the active selection.

4 Critical Analysis

Strengths: The dashboard effectively demonstrates that **musical mood transcends genre boundaries**. For example, the "Sad" quadrant clearly contains overlapping clusters from Rock, Country, and Pop, supporting the hypothesis that continuous audio features convey more emotional information than categorical genre labels. The coordinated design facilitates both overview and detailed inspection: users can quickly identify regions of interest in the mood map and immediately examine the genre composition and audio characteristics of their selection. Balanced stratified sampling was essential to this discovery process; without it, minority genres would have been overwhelmed by the Electronic majority. The interactive brushing mechanism enables fluid exploration, encouraging users to test hypotheses without waiting for page reloads or navigating between separate views.

Weaknesses: The primary limitation of the sampling strategy is that it fundamentally alters the original distribution. Although balanced sampling improves visibility of minority genres, it obscures the dominance of Electronic music in the original dataset and may mislead users about genre prevalence. The reduction from 114,000 to 15,000 points necessarily discards information; local outliers and subtle patterns present in the complete dataset are lost. Furthermore, the line chart showing average feature values can be challenging for non-expert users to interpret, particularly when many genres overlap. Future iterations could address these limitations by allowing users to toggle between proportional and balanced sampling modes, or by providing contextual statistics about the original distribution.

References

[1] James A. Russell. "A circumplex model of affect." Journal of personality and social psychology 39.6 (1980): 1161.

[2] Ben Shneiderman. "The eyes have it: A task by data type taxonomy for information visualizations." Proceedings of the IEEE Symposium on Visual Languages. 1996.