Team Datadventure Process Book

COM-480 Data Visualization

Introduction

It seems almost cliché that music has always been deeply connected to human emotion. Whether it's a comforting melody during hard times or an energizing beat that lifts our mood, music often feels like more than just background noise. This data story explores this connection. Can the genres we gravitate toward offer clues about how we're feeling or even predict underlying struggles like depression or anxiety? Our data story uses data on listening habits and mental health indicators, diving into patterns that suggest certain genres are more commonly associated with particular emotional states. But we also go a step further questioning whether music is the cause or simply a reflection of something deeper. For this, our focus expands beyond music itself to consider the lifestyle factors that may sit behind both musical taste and mental health. The goal is not to make sweeping conclusions, but to make the user think, unpack meaningful correlations and reflect on how they might help us better understand the emotional lives of ourselves and others.

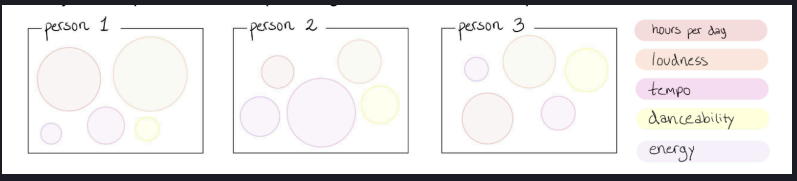
Description

This project is set up into three main sections: 1. hypothesizing and reflecting, 2. exploring and 3. learning. The story starts after the user clicks on "Click here to start", which is necessary to avoid problems with the background sounds and browsers' autoplay restrictions for playing sound without user interaction with the website.

Below, we describe each section, as well as a picture of our mockup and the final website design. As the sketches were on a white background, and we ended up having a dark background on the website, we needed to change the colors.

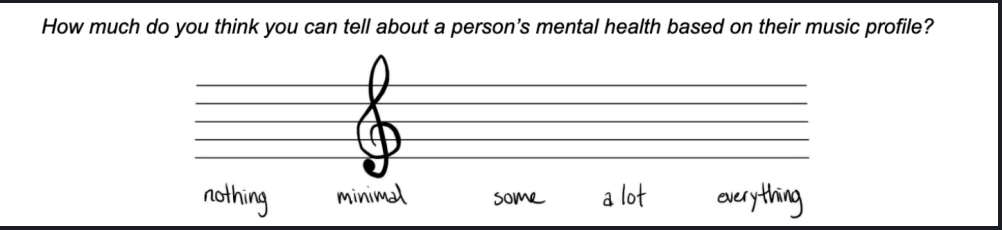
1. Hypothesizing and Reflecting

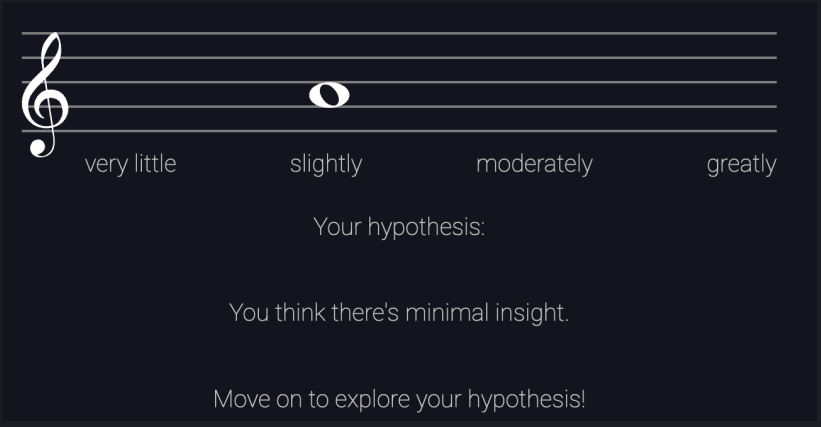
First, we invite the user to get into the right headspace, and hypothesize on the problematic of "What effect do music habits have on mental health?", by asking them engaging questions. They are presented with three personas, of which only one self-diagnoses as depressed. There is no real pattern to be found that the user can use to distinguish the mentally-ill individual. Nevertheless, the user is prompted to try and guess the depressed one, to get the user more engaged. We also added green/red edges and sounds to give feedback to the user depending on the user guessing correct or wrong.



Next, we ask the user to form their own hypothesis on the issue. We hope that this makes them think about the issue, and makes the next sections more interesting, since they are curious whether their hypothesis is correct.

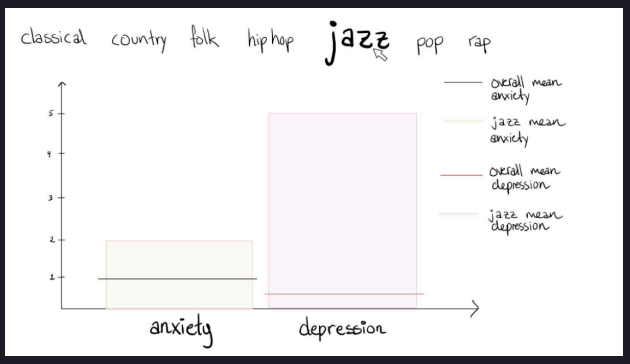
In the final implementation, we found that the slider begs to be played and fiddled with, but does not really encourage reflection, which is why we added a timer, a shrinking ball, that takes 10 seconds to disappear. During this time, the user has to stare at the website, and think about the question. Then, the user can play with the scale, which makes a sound depending on where the note lands. This increases the interactivity of the element, and makes it more engaging. The location of the slider also triggers a hypothesis to display accordingly. This message subtly makes the user take this story more seriously. By putting words in their mouth, they might feel obliged to represent this hypothesis for the coming minutes.

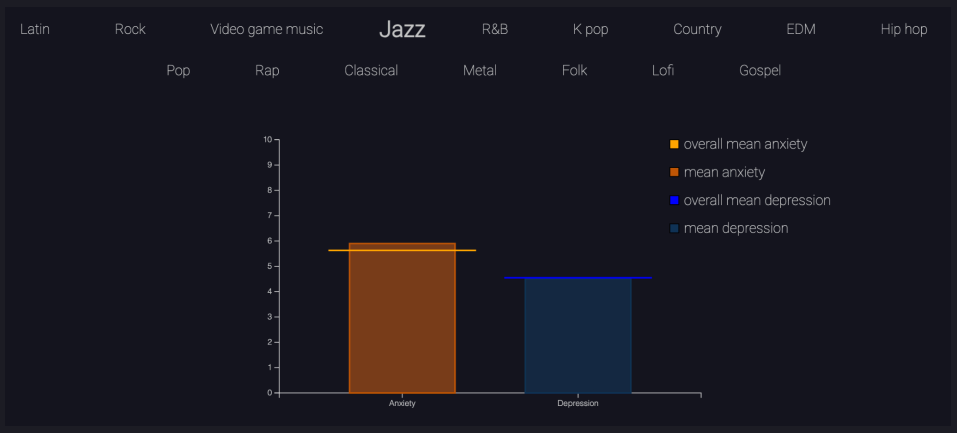




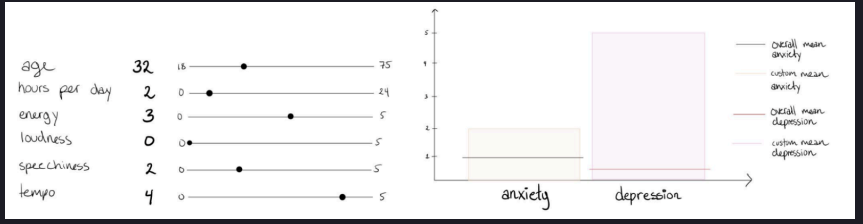
2. Exploring

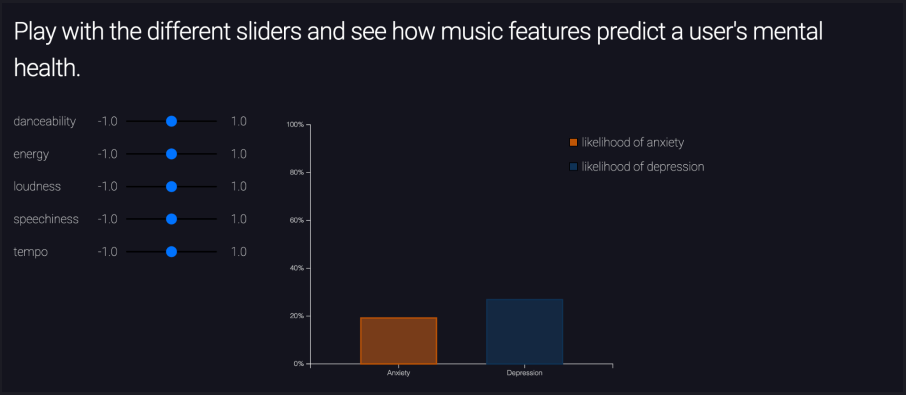
As the user scrolls down, they reach the exploratory part of the data story. For each genre, we calculated the average anxiety and depression based on scores ranging from 0 to 10. Hovering over a genre updates the graph with the values for this genre, and clicking on it selects it permanently. This way the user can compare genres by moving their cursor around the list of genres.





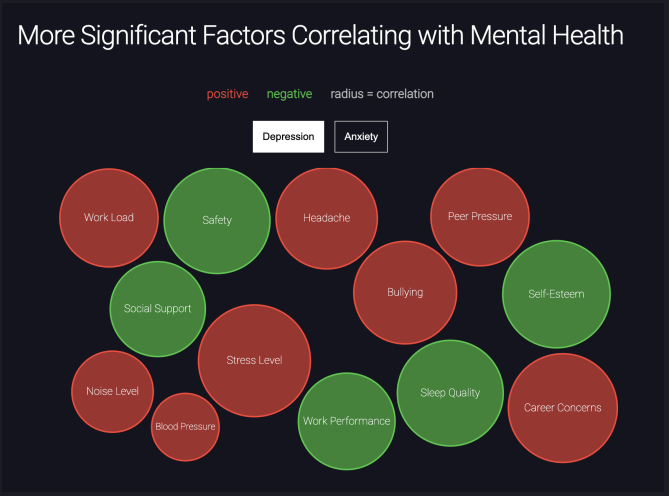
While the graph above is interesting, and cool to play around with, it is hard to generalize the findings. Listeners of lofi for example display worse mental health than listeners of gospel, but what does this actually mean? In the next section, we take up this observation, and let the user fiddle with features of music, like tempo and danceability. The sliders represent features for a linear regression trained on data from the "Mental disorders and music features" dataset, and the output the percentage of users in the data who displayed the disorder. Here, the user might finally learn something, such as greater danceability lowering the chance of being mentally ill. However, the results are sometimes counterintuitive (more energy == more depression?), and the effects of the sliders are quite weak.





3. Learning

After having explored how music preferences relate to mental health, we acknowledge that the previous findings are probably not very causal, and go on to give the user the chance to explore other potential factors correlated with depression and anxiety in a more hands-on way. Through the interactive bubble visualization, they can engage with non-music-related factors, like sleep quality, social support, workload, or bullying , and see how strongly each one correlates with mental health outcomes. We specifically used area of circles to show the correlation coefficients, since it is more difficult to perceive differences in area, and we don't want to overstate the differences in impact between the factors. A healthy lifestyle is not one with a lot of sleep but no social support, but rather all factors deserve attention.



The data used

As the focus for this project was on the visualization and not necessarily on the quality of the data, we decided to use datasets from Kaggle. They have a lot of interesting datasets within various fields that are easy to implement. We ended up using these datasets:

- Mental disorders and music features: This dataset is primarily used in the section where you can use sliders to change danceability, energy etc. to see how our regression model predicts the users mental health.

- Music & Mental Health Survey Results: We use data from this dataset both in the first section with the “Try to figure out: Which of the listeners self-report as depressed?”-quiz and the third section where you can explore the correlation between genres and mental health. In the quiz we found three example people where one of them had a self reported depression value of 10. The exploratory analysis of this can be found in the notebook finding_personas.ipynb.
- Student Stress Factors: A Comprehensive Analysis: The dataset used to reveal the factors actually correlated to depression and anxiety, in section five “Factors Correlating with Depression and Anxiety”.

Challenges

Coming up with a Data Story and Finding Appropriate Data

Our initial plan in this data story was to explore and visualize the correlation between music habits and mental health. We assumed that there must be some noticeable link. However, looking through multiple different datasets, we found that music listeners actually display pretty similar mental health statistics, despite having different music habits. Most of our findings were not very insightful. The disappointment from our findings sparked a new idea: let our users experience the same journey as we did. We turned the project into a mini-adventure, in which they are not meant to reach any conclusions, except for the fact that there is no obvious one. First, guess which listener is self reporting as depressed. Then watch the myth of music habits and mental health being correlated unravel. The project is built upon exploration of data in which they are not meant to reach any specific conclusions. It was a bold idea, and we had to clarify and plan a lot of things to ensure that the exploration would not be boring. As the idea took shape, we settled on a more artsy, emotional exploration, that is meant to replicate a journey. Through the use of sounds and visuals, the user should reflect on this topic, maybe reaching some conclusions about themselves personally, but no specific conclusions about music and mental health in general. By the end, users are eager for answers so we present factors that have a greater influence on mental health than music habits, providing some answers, and avoiding the user feeling like they wasted their time. Spoiler: the factors are the everyday habits and lifestyle choices you’ve heard a thousand times before. Sleep, exercise, stress and self esteem are some of the factors that matter more than just swapping out your playlist and hoping for a miracle cure.

The Design

When we adapted the visualizations to fit into the website’s format, we ran into some design questions, especially around the color palette. We chose a website template with a dark theme, so we needed colors that would stand out and provide enough contrast to be clearly visible hence the bright color. At the same time, we wanted the colors to feel intuitive (like red = bad mental health, green = good mental health) and also be accessible to people with

colorblindness. That meant avoiding confusing hues using a “colorblind simulator” website, and tweaking brightness and opacity in some places while keeping the overall look consistent. The final product balances clarity, accessibility, and aesthetic cohesion. It’s worth noting that in the third part, the colorblind-friendliness was not our main consideration, as it was not compatible with an intuitive color scheme.

Visualising the Difference between Causation and Correlation

In our story, we faced the challenge of communicating our results without misleading our users in regards to the classic fallacy of correlation vs. causation.

An example is our bubble chart where each bubble represents how strongly each lifestyle factor is correlated with a mental health disorder. The bigger the bubble, the stronger the correlation. The problem is that people naturally look at a big bubble and think, “Ah, this must be causing depression!” when really, it’s just *associated* with it. Similarly, our regression model that predicts depression and anxiety likelihood based on the feature of music a user listens to might also be misleading. It’s interesting, but we had to continuously remind ourselves that just because a genre indicates mental health problems, it doesn’t mean it’s the cause. People who are already feeling down may be drawn to certain genres, or maybe there’s some third factor influencing both. Since our visualizations are intuitive and engaging, they might easily be overinterpreted by excited users. Consequently, we tried to add text and explanations to make it clear that these are just patterns that they are observing.

Dealing with a sensitive topic and not being insensitive accidently

As everyone has mental health, and depression and anxiety being a sensitive topic, it has been a challenge how to decide on the wording to not be insensitive by accident. As correlating problems with mental health to something as “simple” as music and genres, we wanted to make it clear that we understand that the topic is complex and that there is no easy quick fix. Which we hope the last part will help add on.

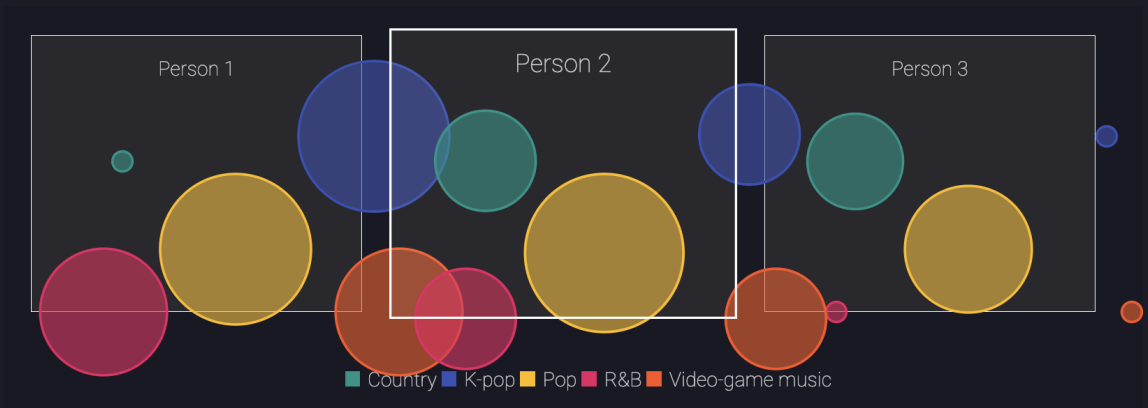
One way we wanted to make the experience a little more approachable and gentle for users was by including common relaxing stimuli, like background mindfulness meditation music, to create a calm atmosphere. We also chose to use non-aggressive, soft shapes, like draggable bubbles, instead of sharp graphs or rigid layouts. The fact that users can move the bubbles around themselves gives a subtle sense of agency and interactivity, which we felt aligned well with the tone we were going for: informative, but also comforting and respectful.

Technical Difficulties

Since we all are not very familiar with html, css, javascript and its d3 library, we faced plenty of small technical challenges that come with getting used to an unfamiliar environment. As we progressed up the learning curve, coding got more fun, but it also revealed bad practices and mistakes from the past. Since it is not very interesting to recount every issue we faced, we will focus on the two biggest, getting to know the template and making the visualizations work with different screen sizes.

We used a html 5 template, specifically html5 UP's "Landed" template. Being unfamiliar with the template as well as the underlying technologies, we were frequently faced with overlapping elements, text being formatted in (initially) unpredictable ways, and the website's scroll dynamics not working as expected. With the help of the internet, chatGPT and the browser's "inspect" tab however, we managed to learn and get used to the template, and use it to its full potential.

A specific coding challenge we faced was ensuring that the bubbles and text boxes in the first quiz remained well-positioned and readable across different screen sizes (see screenshot below). Initially, the positions and sizes were hard coded using absolute pixel values, which caused overlaps or off-screen elements on different sized displays. To fix this, we switched to a responsive layout using relative units and dynamic positioning based on the actual width and height of the container. We also used percentage-based scaling within the SVG, which allowed elements to resize and reposition proportionally, ensuring a consistent user experience across different zoom levels and devices.



Peer Assessment

Yohann	Sina	David
Looking for data and exploring the datasets	Looking for data and exploring the datasets	Looking for data and exploring the datasets
Set up website template	Sketches of the website	Everything related to the hypothesis-slider section
Visualization of Factors correlating with depression and anxiety	Everything related to the first quiz with the personas	Everything related to the genre exploration graph
Text transition from 2nd to 3rd part	Merged depression and anxiety together and adding buttons to change between them	Everything related to the transition text between the two exploratory sections
Screencast	Small fixes in other sections	Everything related to the music feature exploration graph
	Favicon	Start screen & background music
		Website harmonization and small fixes in other sections