SOCIAL IMPACT

In times of increasing isolation, more and more people have been reporting feelings of loneliness and deterioration in mental health. Music platforms can act as an escape out of the daily routine and a site for solo listeners to collectivize. Our personal experiences engaging with music throughout the pandemic showed us that our listening habits and taste in music are shaped by factors beyond our personal preferences. This is particularly relevant for young people.

In response to this query, our analysis delved into the effects of COVID-19 incidence (troughs vs peaks) on a number of song audio attributes generated by Spotify. Our analysis outlines some anticipated external stakeholder groups:

- Spotify and other existing music streaming platforms: Our analysis may generate a negative reaction across these different platforms. The monopolization of the music streaming industry implies that any explicitly 'negative' outcome will be received with backlash from the corporate world.
- Music experts, avid lovers, and creators: Our analysis may come across as reductionist towards the process of music creation and production. Many factors go into producing a track that may have nothing to do with the physical alterations in our built environments. The strong potential for confounding in our analysis may repel those strongly passionate about the meaning and influence of music in our lives.

In addition, our analysis is rife with assumptions. These include assumptions of no copyright or legal conflicts. We also did not question normative standards, including the role of major streaming platforms in society, human connections to music and continued availability of adequately powerful data.

We expect unfair outcomes in some scenarios. The features we used to assess variation in tracks may not be representative of many important aspects beyond audio features. We expect no significant negative impact on community wellbeing. We did not address environmental harm resulting from data collection and storage.

ETHICS

Our data certainly contains historical biases. COVID-19 data suffers from a number of biases. To begin, there is selection bias whereby some demographics are more likely to report their COVID-19 infection than others. For example, young people in good health might test positive with an at-home test and never report their infection to a central health authority, whereas an older or immunocompromised person might seek out medical attention, resulting in a recorded infection. These discrepancies mean that the peaks and troughs we identify in national infection trends might not correspond to the true peaks and troughs. These data also suffer from reporting bias: some municipalities are far more equipped to collect and report data than others. Some

cities—especially poor ones that were already beleaguered by other social problems—likely faced challenges in accurate reporting when they were overwhelmed by COVID. Moreover, the bureaucrats who run these programs may face perverse incentives to exaggerate numbers to invite federal aid or suppress numbers to prevent federal intervention.

We also expect our Billboard data to exhibit biases. To assemble its weekly list of top songs, Billboard draws on record sales, radio plays, and streaming data. Physical record sales would seem to provide a very weak indication of music trends when music stores were shut down for COVID. This might be negligible if the other two categories effectively capture listening trends, especially if Billboard adjusted its weights following the closure of music stores. Unfortunately, radio plays and streaming data do not account for the remainder of the musical universe. For example, many people still own records, CDs, and downloaded music files, none of which are captured by radio plays or streaming data. In fact, we might expect that in times of great stress, people take comfort in these familiar media. This alone raises suspicion that our data systematically excludes important listening behavior, but the concern is compounded by the fact that it is older generations—precisely those most at-risk for COVID—who are most likely to listen to music in ways not recorded by Billboard. The exclusion of these demographics introduces bias into our assessment of what songs are popular in a given week. Tastemakers and influencers, on the other hand, are disproportionately young, healthy people, who are among the least at-risk for Covid.

Finally, the metrics we derived from Spotify, such as valence, loudness, and energy, are also subject to biases. To preface, public understanding of these metrics is limited due to their proprietary nature. We understand in a general sense what is measured by "energy" in that we can often deem one song "higher energy" than another, but this assessment is intuitive and ordinal. It is unclear precisely how Spotify constructs a cardinal measure for these attributes. Further, we have reason to suspect that these measures reflect the privileged identities of the engineers who work at Spotify. People might felicitously disagree over whether a song is positive or high-energy because these assessments are subjective and culturally-dependent. When Spotify intervenes, we expect these disagreements are resolved in favor of hegemonic White assessments of what qualifies a song as, for example, high-energy.

The above biases exist upstream of our work with the data, but we are also aware of biases that arise in our interpretation of the data. For the COVID data, we chose to assess peaks and troughs rather than absolute case numbers That is, we found local maxima and minima and compared music trends at those points, but this was a deliberate choice on our part, and we chose this method at the exclusion of alternatives like "high and low infection rates," for example. In the Spotify data, we focused on valence and energy over other measures like "acousticness" because these were already represented as real numbers in [0, 1], which lent itself naturally to the analyses we conducted. We also examined loudness for its objectiveness. When collecting lyrics

for our top songs, we scraped lyrics from Genius.com. By the nature of how Genius formats their URLs, our scraping process was unable to collect lyrics for some songs, especially songs by multiple artists. This omission likely biases our lyric results against collaborative genres, such as hip-hop, and inhibits our ability to examine whether collaborative music increased in popularity during peak-Covid periods.

For our Covid, Billboard, and Spotify data, we are certainly using the data in a manner in which the subjects have consented. Our World in Data, the source of our Covid information, explicitly states, "a key part of our mission is therefore to build an infrastructure that makes research and data openly available and useful for all." Our World in Data, partnering with Johns Hopkins University, has been the most-trusted global source of Covid data during the pandemic. Our work is consistent with their goals. Billboard publishes its results to the world every week, and Spotify's API is built for this kind of project. The only source from which we do not have clear consent for data collection is Genius.com, whose API does not allow for lyrics extraction due to liability concerns. That said, as a lyrics website built around crowd-sourced public commentary, we believe that our use of their site is consistent with Genius' mission, and our work does not present any legal challenge to their organization.

Our results are vulnerable to misinterpretation. For example, we find that there is a statistically-significant increase in loudness and energy of top songs during peak COVID periods compared to COVID troughs. We are wary of reactionary narratives that mis-appropriate our results in service of claims like "COVID must not be very severe in these communities because they listened to joyful music during peak COVID periods. These communities don't deserve government support because they were probably playing this music at big parties!" COVID-deniers, anti-vaxxers, Tea Party supporters, and policymakers with competing priorities all might have reason to co-opt our results in this way. The reality, of course, is much more complex: the caged bird sings not out of joy but rather to cope with its miserable conditions, and people can listen to dance music without attending super-spreader parties, but this nuance might be obscured if our results are weaponized against marginalized communities struggling with COVID.