Spotify's Official Playlists and Questions of Fairness

Music consumption has drastically changed over the years. With digitisation completely changing the way we consume music, we have become heavily dependent on streaming services. Currently, the most popular streaming service is Spotify, with it taking 31% of the total market (LeClair, 2022). A stand-out feature of Spotify (and its most advertised one) is its playlist. This paper will look into Spotify's playlists and the factors that influence the selection of songs contained in them. A case study of Spotify's editorial playlists created for the Indian audience will also be done to support this paper.

Music wasn't always a commodity to own. In most societies, for thousands of years, music was considered a 'public good' – a universally accessible resource that all members of a society may draw upon to fulfil their individual and collective needs. But with advancements in technology, we were able to reproduce music. Reproducibility brought along with it the ability for music to be bought, transacted and sold. It became a commodity that could be owned rather than performed and enjoyed. But music is still a commodity that has no value unless it is shared. Streaming services do this act of sharing an exponentially easier one.

Streaming services like Spotify have made playlists its defining feature. CEO Daniel Ek said in February 2016 that he wants "to soundtrack every moment of your life," and he was serious about it. You could find a playlist for any impossible situation you fall into. These playlists are algorithmically generated or are curated by Spotify's editorial staff, consisting of 150 people. Engagements with these playlists increased considerably, with these official playlists making up one-third of all listening activity on Spotify (Powell, 2018). This goes on to show how much of an influence these playlists have.

Out of the two types of playlists that Spotify officially makes, the algorithmically created playlists, "tailor-made" for your tastes, raise many questions. The most prominent one is the "fairness" of these algorithms. The short answer to this question is that these algorithms are not fair to every group of users and artists. But this specific problem is not because of the big bad corporation wanting to get more money. It is because of inherent flaws in the recommender systems in place for streaming services. One main flaw is the feedback loop. A feedback loop occurs when users are recommended popular songs, and users engage with these recommendations, thereby increasing the popularity of these songs. These algorithms then take this increased popularity into account to form recommendations, forming a loop (Knibbe, 2021).

The other type of playlists rolled out with the label made by Spotify is the editorial playlists. These are curated playlists for specific genres, situations, "vibes", etc. Most of them are updated on a regular basis. Many of these playlists have a huge following, and a song ending up on one of these playlists would themselves be a substantial promotional vehicle. Playlists like Today's

Top Hits, for example, have around 31 million likes. The position in which your song is placed in the playlist also determines your exposure and streams. With people deciding which goes where one's apprehension of these playlists increases. It also does not help that in the "about recommendation" section of such playlists, Spotify says, "In some cases, commercial considerations may influence our recommendations, but listener satisfaction is our priority, and we only ever recommend content we think you'll want to hear." A lot of analysis has been done on these playlists (Knibbe, 2022), and a disproportionate number of artists from particular labels are seen to get frequent spots in them. Does this mean that any random "artist" can gain enough exposure and influence if enough money is thrown?

Case Study:

While studying Spotify's editorial playlist, I wanted to observe the attributes of the songs included in them. Specifically, the Spotify editorial playlists curated for the Indian audience. I chose six different playlists, all curated by Spotify - Top 50 India, Top Hits Hindi, Bollywood Butter, Punjabi 101, Hot Hits India, and New Music India. While it could be argued that the selection of these playlists showcases a selection bias, these six playlists were specifically chosen as they were the playlists that had the most likes amongst Spotify curated playlists for India. Top Hits Hindi led the number of likes with around 763k likes, followed by Bollywood Butter with approximately 600k likes. Surprisingly (to me), the playlist that came in third was Punjabi 101, with 539k likes. Only one is updated daily (Top 50 India) from these playlists, while the rest are updated weekly. The difference in the playlists due to the frequency of updates is quite negligible as Top 50 India has more or less contained the same fifty songs, with a maximum of one track removed and one new track added to it, once every couple of days.

I used a freely available application called Exportify (https://github.com/watsonbox/exportify), which provided a simple interface to export your playlists into a CSV file. It uses the Spotify Web API to export playlists[1].

Once these playlists were exported, we see that the following attributes of each song can be viewed.

```
Index(['Track URI', 'Track Name', 'Artist URI(s)', 'Artist', 'Album URI',
       'Album Name', 'Album Artist URI(s)', 'Album Artist Name(s)', 'Year',
       'Album Release Date', 'Disc Number', 'Track Number',
       'Track Duration (ms)', 'Explicit', 'Popularity', 'Genres',
       'Danceability', 'Energy', 'Key', 'Loudness', 'Mode', 'Speechiness',
       'Acousticness', 'Instrumentalness', 'Liveness', 'Valence', 'Tempo',
       'Time Signature', 'Label', 'Copyrights'],
      dtype='object')
```

[1] https://developer.spotify.com/documentation/web-api/

While most of these indexes are self-explanatory[2], the attribute that holds the most importance to us is the attribute labelled 'Popularity'. Spotify describes the calculation of this variable as follows - "The popularity of a track is a value between 0 and 100, with 100 being the most popular. The popularity is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are. Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past. Duplicate tracks (e.g. the same track from a single and an album) are rated independently. Artist and album popularity is derived mathematically from track popularity. Note that the popularity value may lag actual popularity by a few days: the value is not updated in real time"[3]. This popularity index is considerably more useful when looking at algorithmic playlists created by Spotify. But as we move ahead, we will see that even in these "hand-picked" curated playlists, the popularity index of the songs in that playlist generally tend to cluster around the same numbers.

The CSVs of all the exported playlists were combined into one large CSV file. This was a file with 382 rows (that is 382 songs) and 30 columns. All the entries in this file were not unique and I did not remove duplicates to see how many times one song would appear throughout these six different playlists.

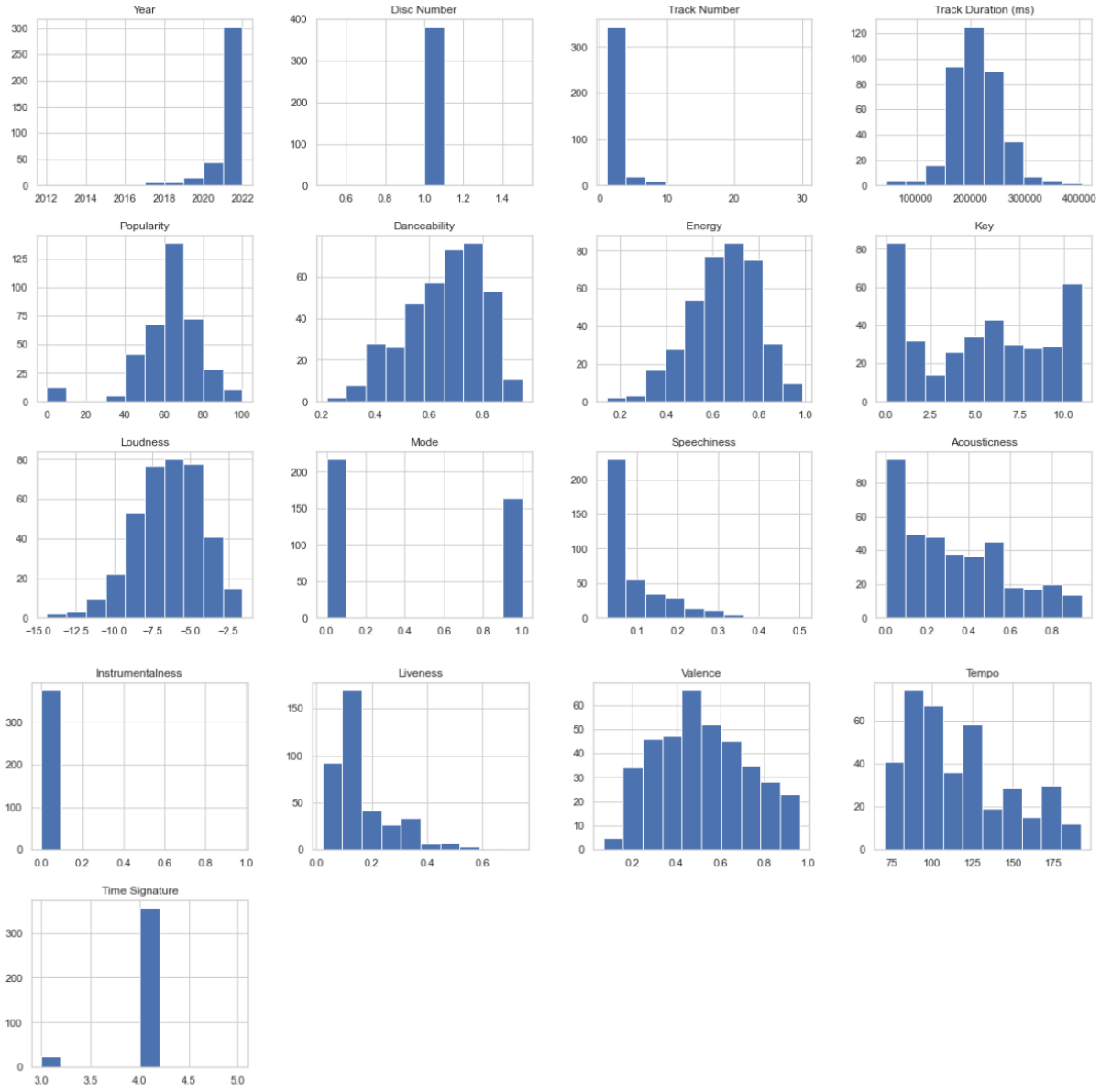
```
Mehabooba                                                    4
Meri Jaan                                                    4
Kesariya Audio Teaser (From "Brahmastra")                    4
Atak Gaya - Arijit Singh                                     4
Kahani (From "Laal Singh Chaddha")                           3
Ranjha (From "Shershaah")                                    3
Chauffeur                                                    3
Mere Yaaraa (From "Sooryavanshi")                            3
Voodoo (with J Balvin & Tainy)                               3
Meri Jaan Meri Jaan (From "Bachchhan Paandey")                3
Bhool Bhulaiyaa 2 Title Track (From "Bhool Bhulaiyaa 2")      3
Srivalli (From "Pushpa The Rise Part - 01")                   3
Heer Raanjhana (From "Bachchhan Paandey")                     3
```

We notice that out of the 382 songs, 49 of them have been duplicated at least once. Out of these, four songs appeared four times and nine songs appeared thrice throughout these playlists. With the exception of Chauffeur by Diljit Dosanjh (the significance of which we will come to in a bit) the other twelve songs were songs from Bollywood movie soundtracks. We could very likely put

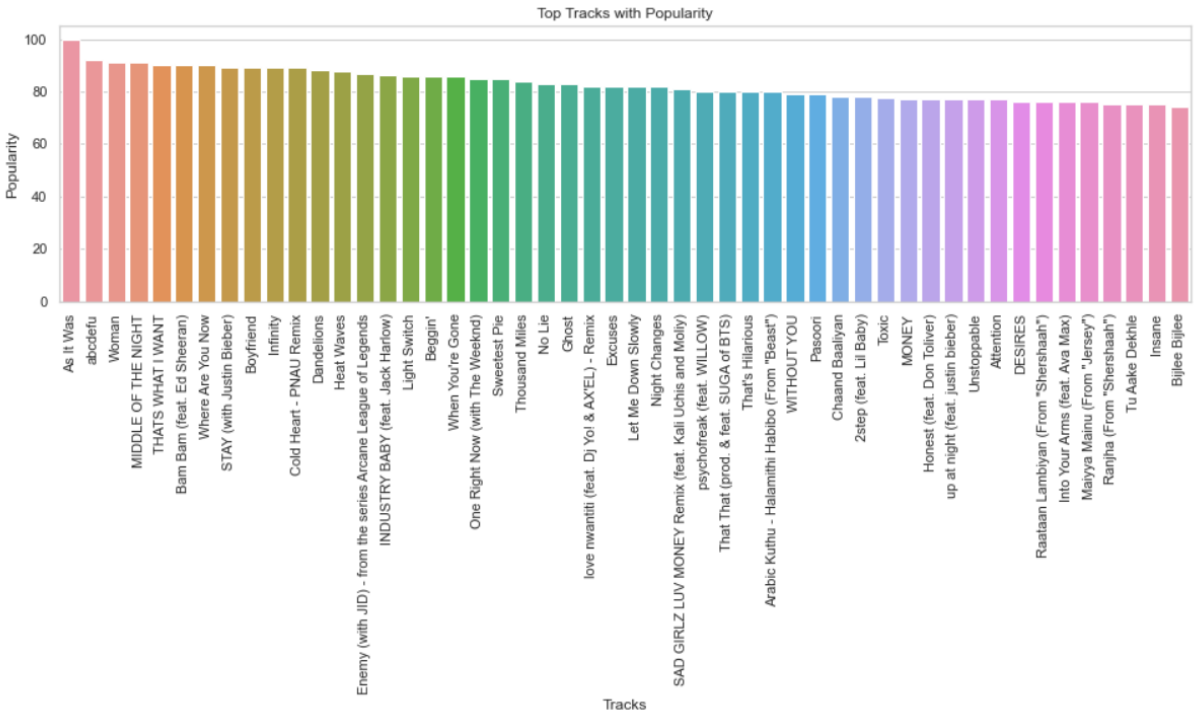
[2] A detailed explanation of all these attributes can be found here : https://developer.spotify.com/documentation/web-api/reference/#/operations/get-several-audio-features

[3] https://community.spotify.com/t5/Content-Questions/What-is-the-popularity-range-for-a-song-to-be-considered/td-p/5221230

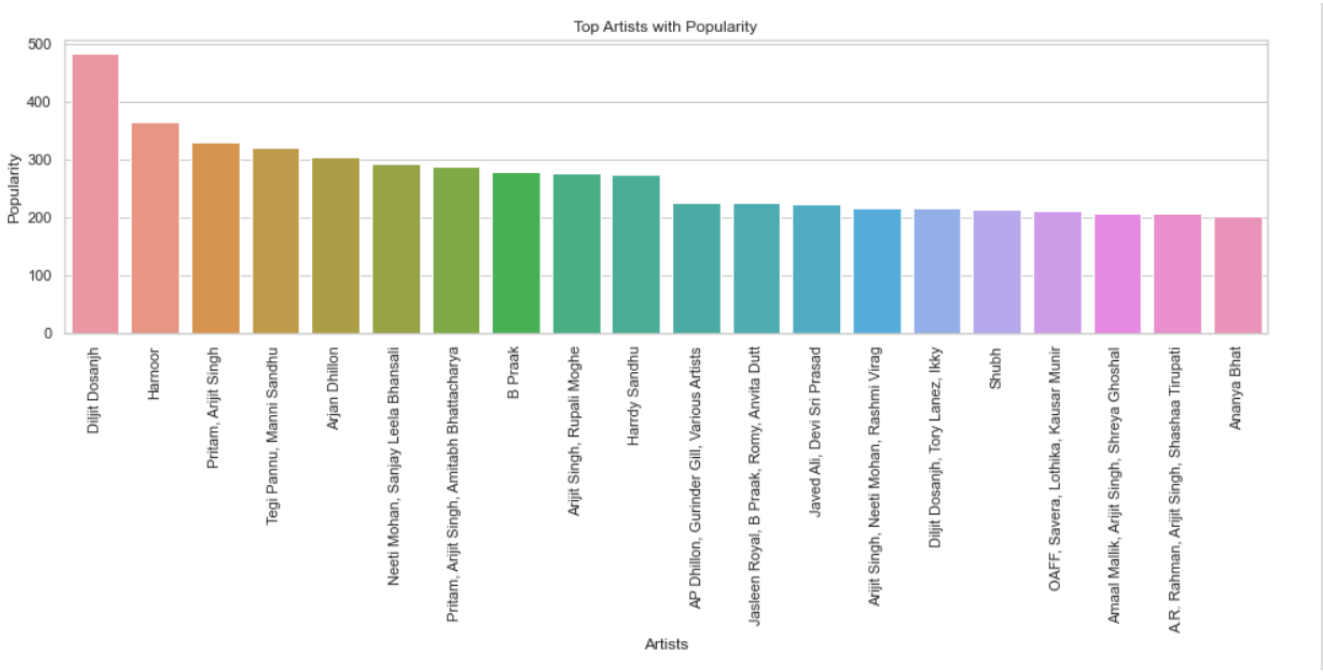
the blame for this on the users and the music culture in India that is so inherently tied with its movie industry.



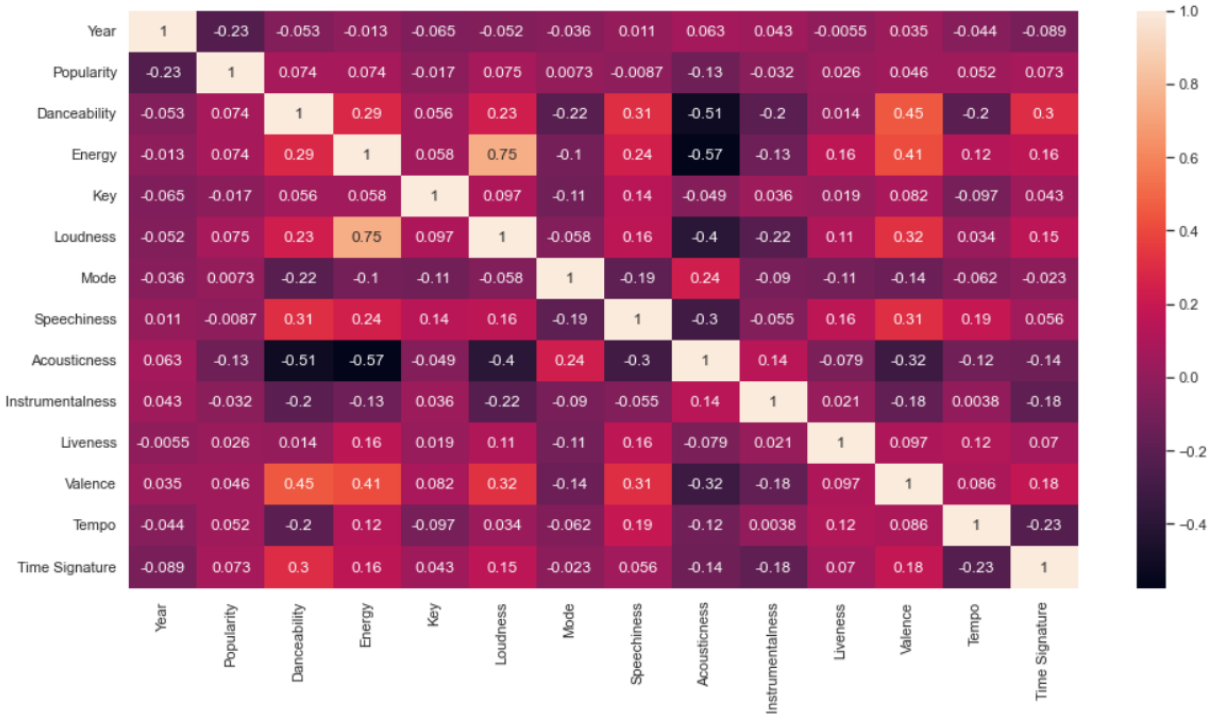
These are some of the other graphs that we obtain while trying to analyse the attributes of the songs. There are multiple inferences that we are able to make from this. Most songs included in these playlists were released in 2022. Most songs are considerably less acoustic, but rather, they are more danceable and loud. While a majority of the songs have a neutral valence, a larger number of songs are on the positive side than negative. And most importantly, a large number of songs included in these playlists have a popularity index between 60 and 70, with most other songs either a little above or below this measure.



Only one song, among all the songs present in the derived dataset, is given a popularity index of 100 - the recently released Harry Styles single As It Was (this is a little suspicious as there are multiple other tracks in these lists that are played a lot recently, but this would just be speculation). The above graph does not do much but shows us that the English songs present in the playlists have a higher popularity index. They are a considerably larger audience, not just limited to India. The Indian song with the highest popularity index is number 32 on this list. But what is interesting here is that almost every one of these songs, with one or two exceptions in between, are songs that are trending on social media platforms like Instagram and TikTok. From here, we start to get an idea of the checks a song must pass to be included in these playlists.

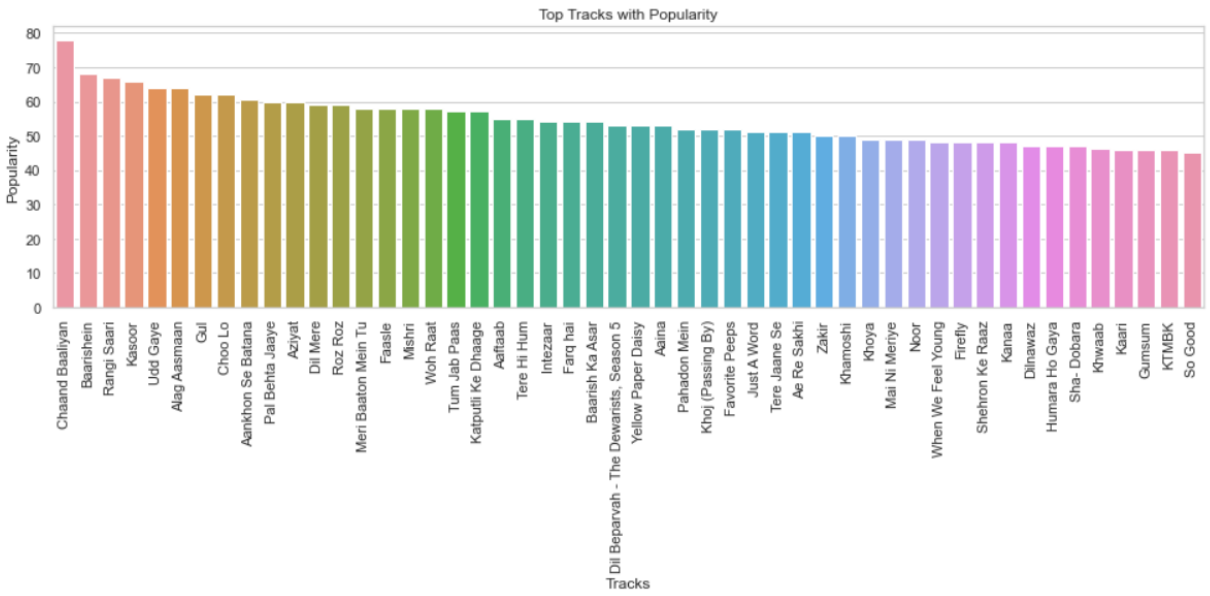


But in contrast to the dominance of English songs in the most popular tracks list, we see a complete list of Indian artists in the graph for artists with the highest popularity index. But we cannot give too much credit to this graph as this dominance can be traced back to our selection bias. Nevertheless, an interesting inference from this graph is the increased number of Punjabi artists. Diljit Dosanjh leads the list, pointing at the fact that his songs have a high popularity index while multiple songs of his (or the same song multiple times) are present in these playlists. Another issue with this graph is that different artists collaborating on one track do not count for each artist's popularity. But rather all the artists together are considered a separate group. We see Arijit Singh appearing five times just in the top 20 itself. But since he was collaborating with other artists for each song, we can’t see the true value of the popularity index for Arijit Singh.



Our final graph is a heatmap that showcases correlations between the attributes of the songs. One surprising inference is that we do not view a very high correlation between the popularity and the year of a song's release. This could be due to the fact that there are many songs present in these playlists, which do have a high popularity index but are not ones that were released in recent years. These are the songs that are currently viral on social media platforms. Examples of this are the 2015 released song Agar Tum Saath Ho and the 2014 released song Night Changes, both present in our dataset. Other inferences from this graph are the high correlation between loudness and energy and the high correlation between valence and danceability and valence and energy. The reasons for these correlations are a lot more straightforward and obvious.

I also exported two lesser liked playlists - Radar India and Indie India, to view in comparison to the above-selected playlists that would more obviously contain well-known songs. But interestingly, many songs in these "indie" playlists as well are songs that are currently viral on social media platforms, with the graph of its most popular songs looking like this:



While the average popularity index has considerably decreased, many of these songs are not ones you can necessarily call indie (certainly not Udd Gaye, with Ashoka itself contributing a considerable amount of streams for it). Many songs in these playlists are still Bollywood songs, just slower and more acoustic than danceable.

The inferences lead to multiple questions about these curated playlists. Are these hand-picked playlists only curated to amplify the reach of already popular songs? If popularity is focused on creating editorial playlists, what is the need for a curator? Can we use an algorithm to do this? Are the users increasingly engaging with songs that are trending on social media?

The more inferences we obtain, the more questions we have on the system's impartiality (or partiality). However, a slight shift in this model is being seen, with more and more independent artists being given spots in these official playlists. But we have a long way to go, if even such a drastic change is possible, for a system that is in parts reminiscent of an ideal one.

References

Knibbe, J. (2021, September 21). *Are music recommendation algorithms fair to emerging artists?* Music Tomorrow. Retrieved May 10, 2022, from https://www.music-tomorrow.com/blog/fairness-and-diversity-in-music-recommendation-algorithms

Knibbe, J. (2022, April 27). *Is It Possible to Get On a Spotify Editorial Playlist as an Independent Artists?* Music Tomorrow. Retrieved May 10, 2022, from https://www.music-tomorrow.com/blog/is-spotify-editorial-playlist-landscape-fair-to-emerging-artists

LeClair, D. (2022, January 20). *New Report Reveals Most Popular Music Streaming Services*. How-To Geek. Retrieved May 10, 2022, from https://www.howtogeek.com/781066/new-report-reveals-most-popular-music-streaming-services/

Powell, A. (2018, March 12). *Inside the Black Market for the Top Spotify Playlists*. The Daily Dot. Retrieved May 10, 2022, from https://www.dailydot.com/upstream/spotify-playlist-black-market/