Proposal: An analysis to understand changes in female artist streaming over time

This document proposes a series of analyses that may help us to understand how Spotify influences the streaming of female artists and if Spotify users' organic consumption of female artists changes with increased exposure to female artists in editorial programming.

Recently, the algo-bias team has identified several biases in both Spotify's data sets as well as their recommender models. Particularly, gender streamshare – or the proportion of artists streamed that are female – has emerged as an avenue of bias that is ripe for disruption. We presently know that listeners seek out fewer female artists on their own in comparison to when they are listening to other people's playlists or editorial programming. Additionally, Discover Weekly has notably lower female-artist streamshare than overall or direct listening. This finding in particular illustrates an "issue of algorithmic confirmation bias due to inherent asymmetries in the data inputs (McDonald, 2016)." However, Spotify editorial programming has significantly higher percentage of female-artist streams in general and in almost every country. In response to this, Female Representation Bet efforts have been made to increase the share of female artists and their streams. Still, we do not yet understand the impacts overtime on direct female-artist streamshare from exposure to higher female-artist streamshare on editorial/algotorial programming.

By understanding this, we think we can:

1. Understand if Spotify is helping or hurting female creators in a broad sense.
2. Better understand which A/B tests will provide us with causal evidence of Spotify's potential to modify user taste and behavior.
3. Develop a method for assessing user behavior modification across other categories (e.g., artist tier).
4. Not only understand algorithmic gender-biases at play, but discover new paths to correct those biases.
5. Get a clearer understanding of how editors should program Editorial playlists and Algotorial pools.

Key Questions

RQ1: How does a user's affinity for (or consumption of) female artists change from preferences expressed during onboarding? 1 week after being on the platform? After activation? 1 month later? 1 year later?

RQ2: Can user behavior be modified to increase organic female artist streamshare through exposure to female-artist heavy programming?

RQ3.1: Which listening conditions and user characteristics are related to short term and long term behavior modification?

RQ3.2: Which user demographic groups (e.g., adolescent, US-based, female) exhibit greater modifications in behavior, both long and short term, when exposed to more female artists?

RQ3.3: What is the appropriate dosage (how much listening and how often) for short and long term behavioral change?

Suggested Analyses

1. Compare proportion of female artists (*fem_art_onbrd*) selected during onboarding with downstream female artist streamshare. See if changes in proportion vary by percentage of streams by play context (e.g., does the proportion of time someone spends in a user playlist vs editorial playlist matter?)
    a. Cons: Tend to offer few choices in on-boarding, which may make this a noisy value
    b. Pros: This may help us to establish a baseline "female affinity" score for users before they are exposed to programming.
    c. Unanswered Questions: We're still unsure how onboarding works from a UX perspective. Need clarity.
    d. Pay attention to whether or not the onboarding is standardized across all users, we need to understand what we're choosing from.
        i. Does everyone get the same 10 artists, or do we try to make it personalized? What is the gender share?
2. Create female streamshare 7-day. 14-day, 30-day, and 90-day aggregates to compare variability/female artist listening across time period.
    a. It is likely that 28 day periods will be the most useful increment for individual level data, according to previous research done by [@glennm].
    b. Soul did 28 days too because it's divisible by 7

3. Compare organic female artist stream (*organic_fem_art*) share directly before and after exposure to Women's March promotional shelf (or similar female-artist heavy playlist).
    a. Might be people who didn't log in on those days (match sampling)... compare with how much they interacted before and after
4. Regress *organic_fem_art* (per users, over time) on exposure to *_playlist_female_streamshare*, controlling for user *age*, *gender*, *city*, *genre affinity,* and *fem_art_onbrd*
5. A question of dosage: multivariate time series analysis?
    a. Cons: not sure if this is an appropriate analysis for the questions we're interested in answering. Also ambitious undertaking.

Sampling Strategy

- US Premium users
- Stratified sampling across gender, age, meta genre affinity, and region of the country
- New Users within the last 90 days
    - Activation:
        - Does it matter? How might we think about this? What kinds of economic goals can we attach to this research to give it a broader audience?

Characteristics to Compare By

- Genre Affinity
- Curiosity Score
- User Gender
- Genre Gender score?
    - Also some kind of way to show the method works for multiple genres?
- User Age Group (0-17, 18-24, 25-35, 35+)
    - There may be a more thoughtful way to group folks by age, however we hypothesize that younger users will have larger effect sizes, as individual's musical taste tends to crystalize in their 30s (see Ajay's work here, and others' related work here, here, here, and here)
- User city/region
    - This will be particularly useful if we're using the "Women's March" shelf, since not every region of the country responded to this political moment in similar ways.

Features of Interest

- User Demographics
    - Age

    - Gender
    - Zip Code for weekend streams
        - US Census Data for this zip code to understand home/neighborhood environment
    - Account sign up date
    - Daily Active User?
    - Weekly Active User?
    - Monthly Active User?
- General Listening Traits
    - Curiosity Score
    - Top Meta Genre
        - Meta Genre Affinity Score
        - Meta Genre Female Artist Share
    - Female Artist Share at Onboarding
    - Baseline Female Artist Affinity (during month 1 on platform?)
- Female Artist Affinity Over Time
    - Fem_Art_Affinity_W1 (28 day aggregate)
    - Fem_Art_Affinity_W2 (28 day aggregate)
    - Fem_Art_Affinity_W3 (28 day aggregate)
    - Fem_Art_Affinity_W4 (28 day aggregate)
    - Fem_Art_Affinity_W5 (28 day aggregate)
    - Fem_Art_Affinity_W6 (28 day aggregate)
- Exposure to Female artist heavy promoted playlists (how to structure this?
    - Feminist Friday: spotify:user:spotify:playlist:37i9dQZF1DWU8quswnFt3c
    - Badass Women: spotify:user:spotify:playlist:37i9dQZF1DX2I7Ykltk83m
    - Fierce Femme's: spotify:user:spotify:playlist:37i9dQZF1DX5kjCvsC5isB
    - Femme Fatale: spotify:user:spotify:playlist:37i9dQZF1DX0IyMQV27EGn
    - Latin Divas: spotify:user:spotify:playlist:37i9dQZF1DX5AVYhCeISA6
    - Women of Pop: spotify:user:spotify:playlist:37i9dQZF1DX3WvGXE8FqYX
    - Women of Folk: spotify:user:spotify:playlist:37i9dQZF1DWSIcimvN18p3
    - Women of Rock: spotify:user:spotify:playlist:37i9dQZF1DXd0ZFXhY0CRF
    - Women of Hip Hop: spotify:user:spotify:playlist:37i9dQZF1DX5l9rcXWdrth
    - Women of RnB: spotify:user:spotify:playlist:37i9dQZF1DX1wNY9tfWQsS

Deliverables

Insights deck for the above 4 research questions, including recommendations for future research, other potential algorithmic-bias interventions, and how these insights might impact future editorial and algotorial programming.

Q’s/Thoughts/Next Steps:

1. Super lean back listening? → hard to control for, and you may end up capturing our internal push toward
    a. Look for artists that were searched for or added to a playlist (companion analysis)
2. Percentage of female consumption vs percentage of what people think is female?
3. Onboarding: what percentage of people go through onboarding?

Procedural Notes: