



Spotify Music Data Exploration

Andrés De La Rosa, Anand Patel, Lina Yang

04/16/2021 | W200 | Sec 12: Fri 4:00-5:30 pm PT

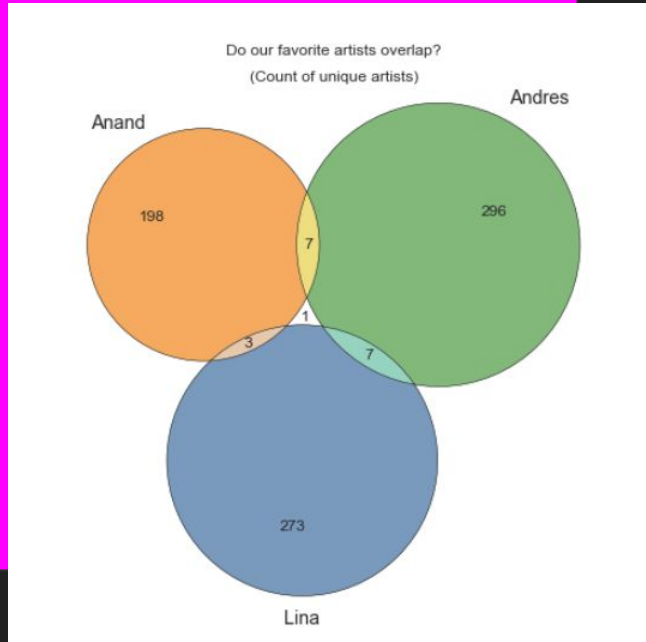
how we got our data

3 music libraries

1,200 songs

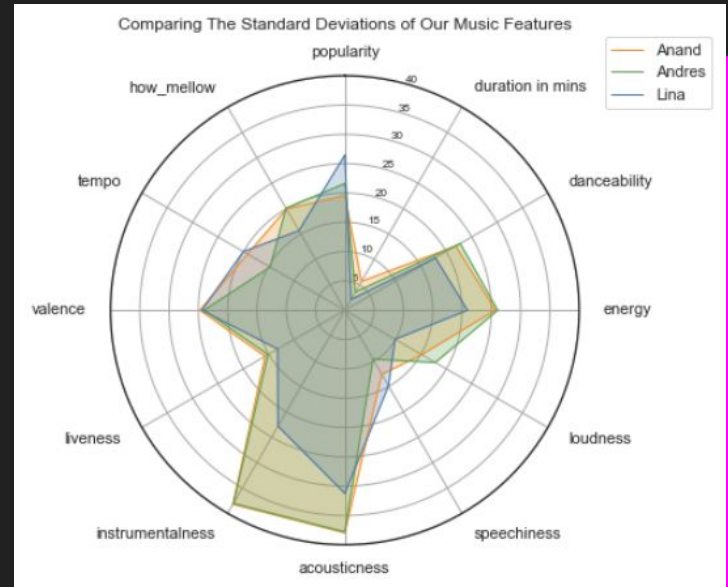
1,009 Pitchfork albums

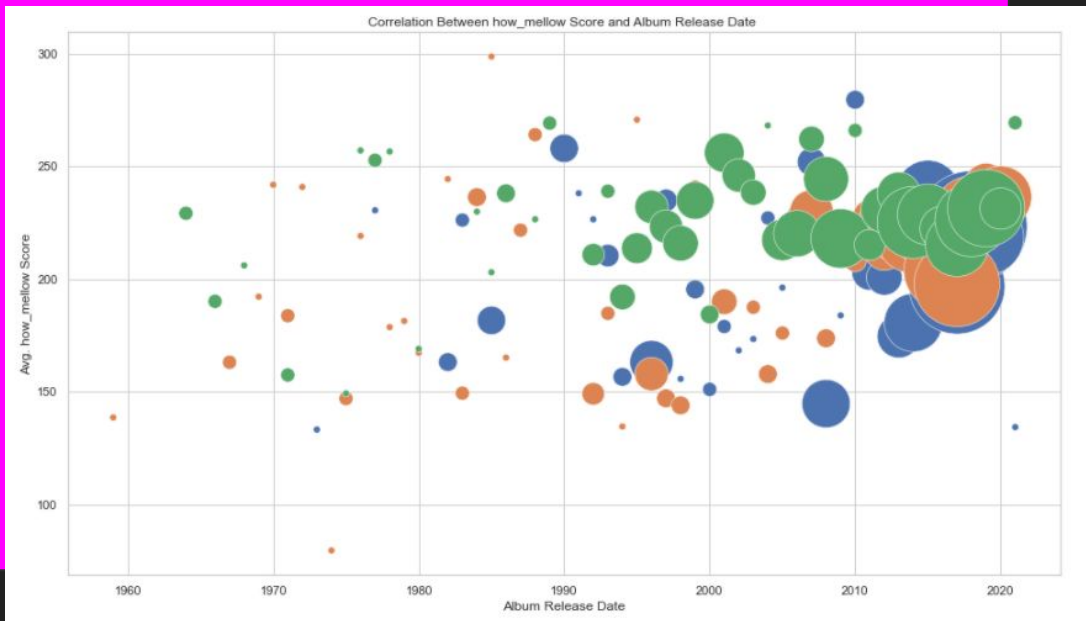
hypothesis



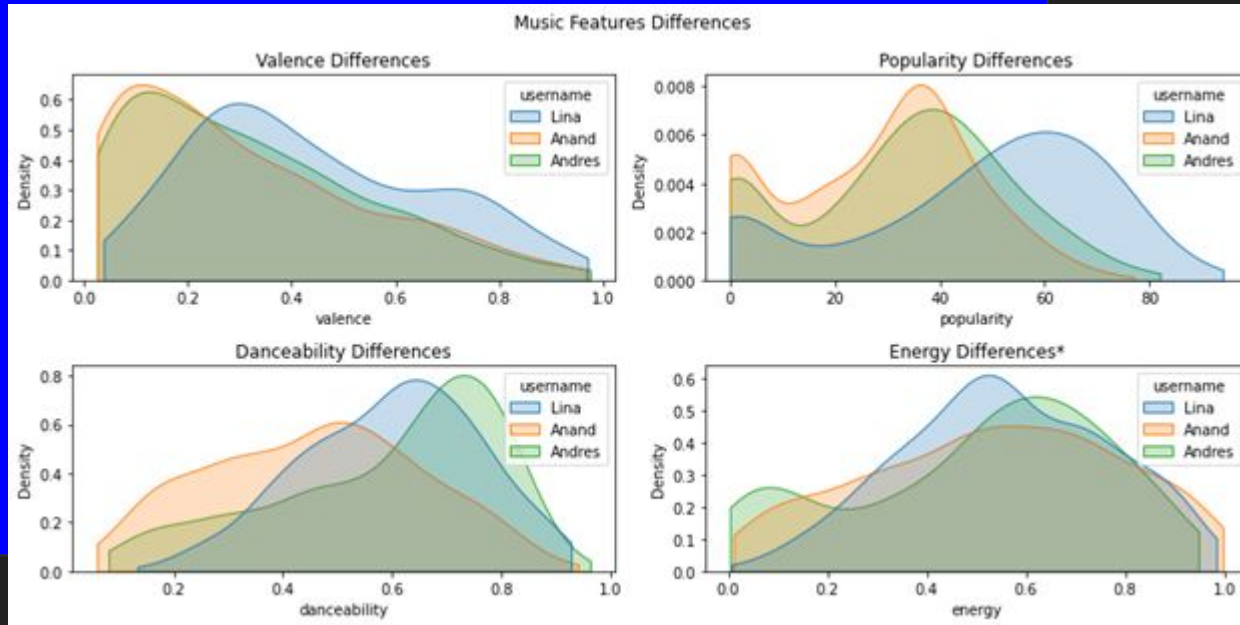
Do our favorite music and artists intersect within each of our top 400 songs?

What is the relative weight of our music preferences across quantitative variables?

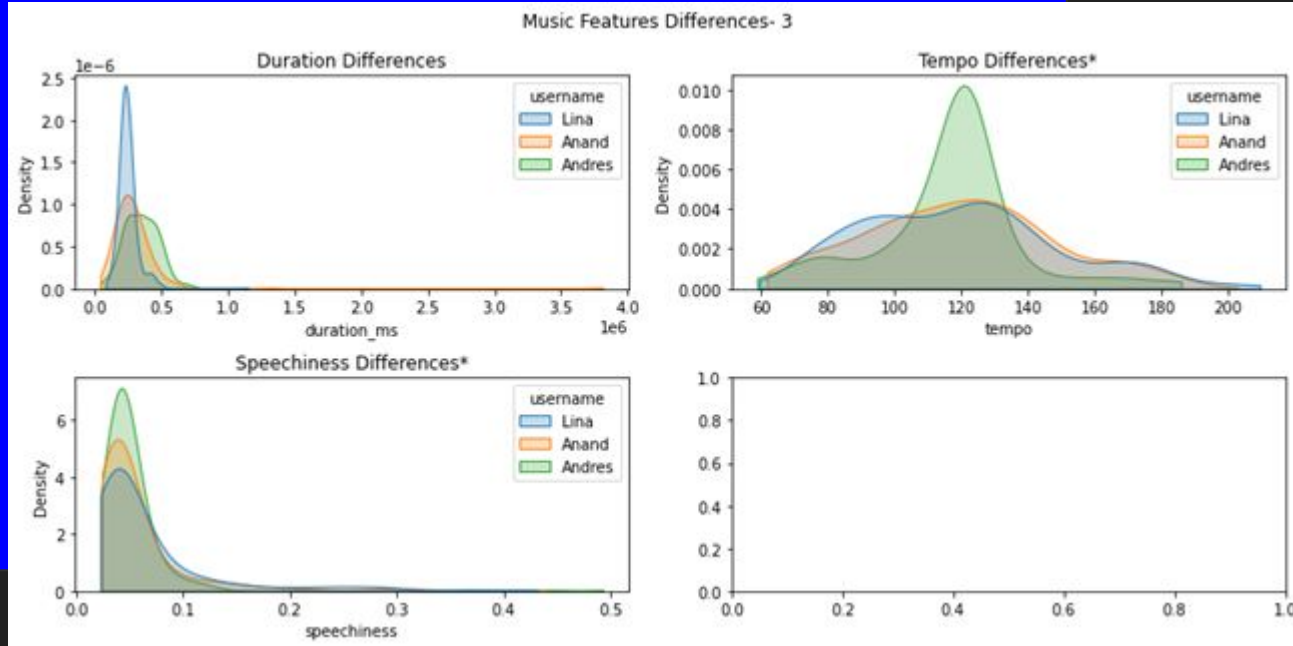




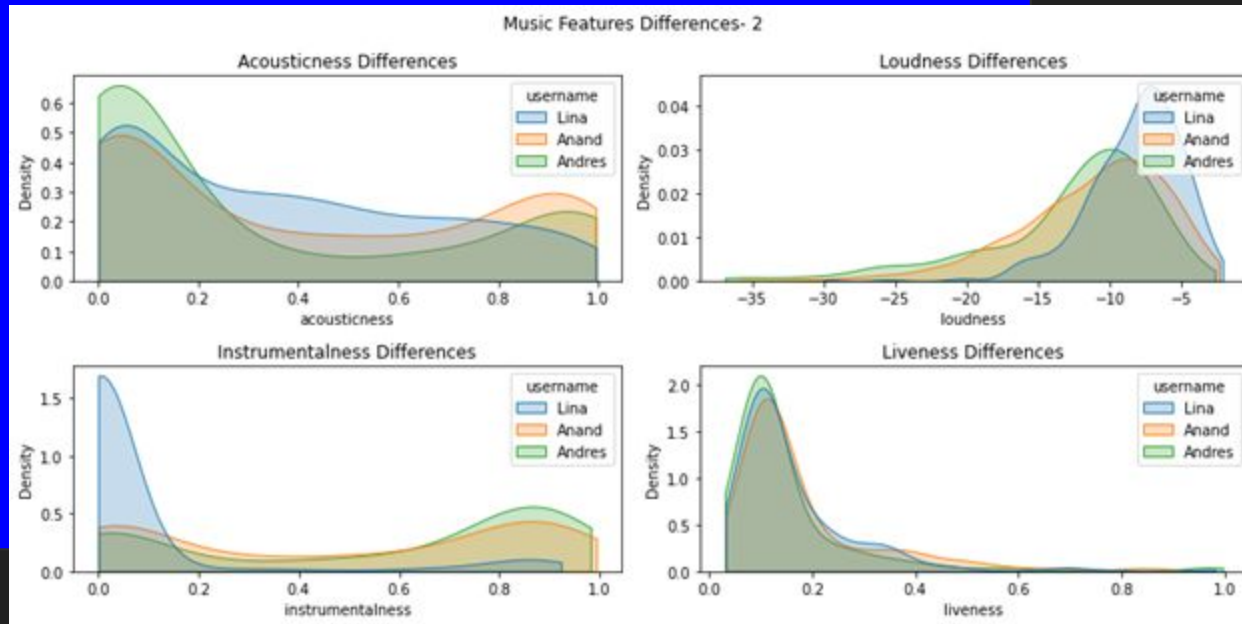
How does the mellowness of our music correlate with time intervals?



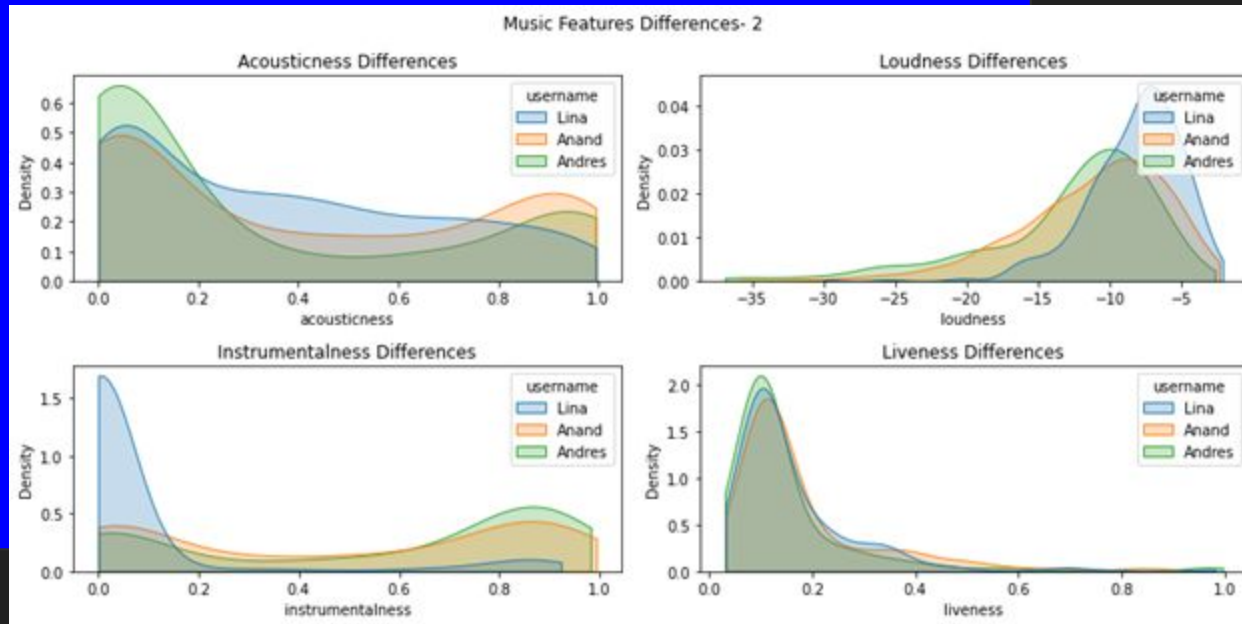
How different are our team members' musical preferences?



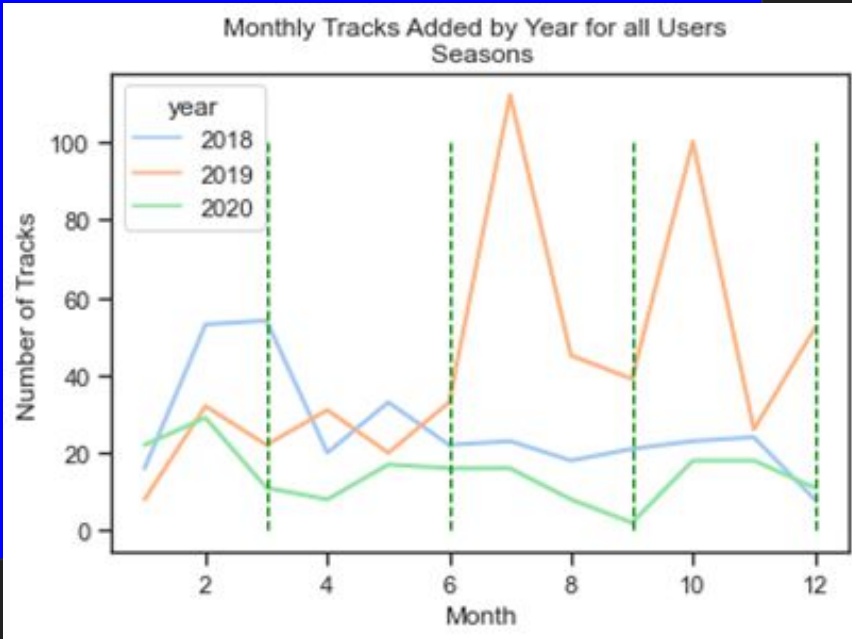
How different are our team members' musical preferences?



How different are our team members' musical preferences?

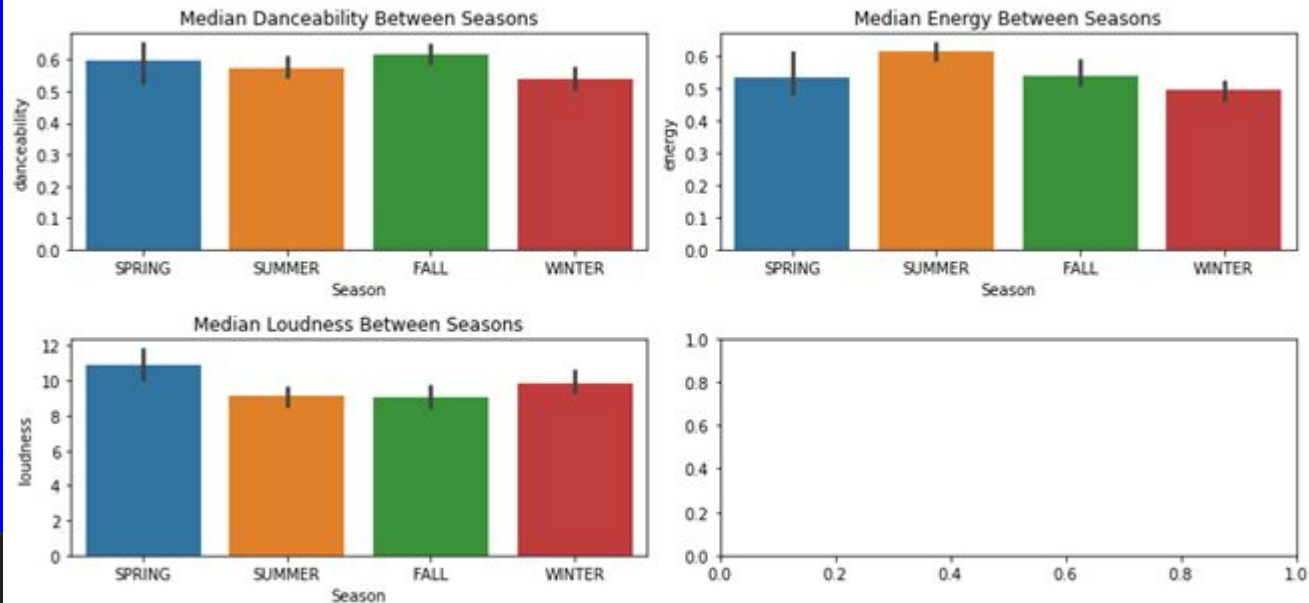


How different are our team members' musical preferences?



In which period of the year, we tend to save music into our Spotify library?

Median Music Features Between Seasons



Do the characteristics of the songs saved correspond to the season?

Pitchfork

our music
recommender



Building our Pitchfork Album Recommender

- Normalize audio features to be from 0 to 1.
- Make audio features vector (avg values, weights) from the subset of a user's library
- Get the distance from a P4K album's audio features to the user's library vector.
 - Sort by distance to get the best recommendations first.
- **Get the top 100 album recommendations for each user**
- Functionally programmed - adjustable to:
 - Turn off weights
 - Denormalize or exclude audio features
 - Return songs instead of albums

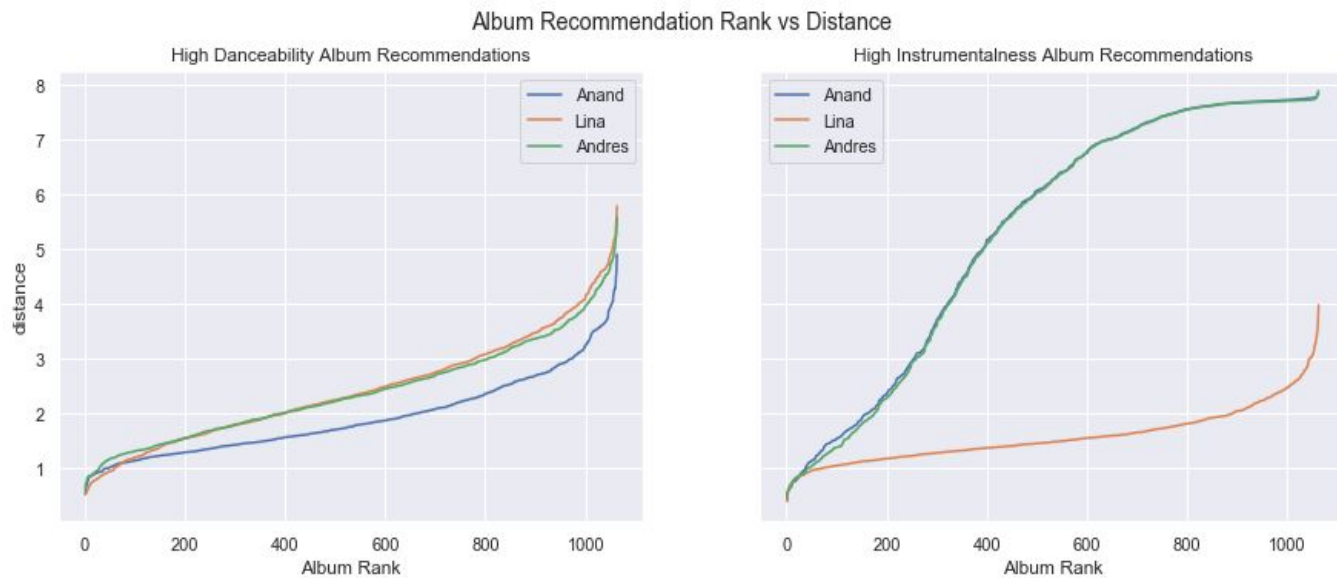
$$w_n = 100 * \frac{\text{abs}(\frac{n.\text{mean}()}{n.\text{std}()})}{\sum_{i=1}^k \text{abs}(\frac{i.\text{mean}()}{i.\text{std}()})}$$

$$\text{album's distance} = \sqrt{\sum_{n=1}^k w_n (v_{n, \text{alb}} - v_{n, 0})^2}$$

3 music libraries

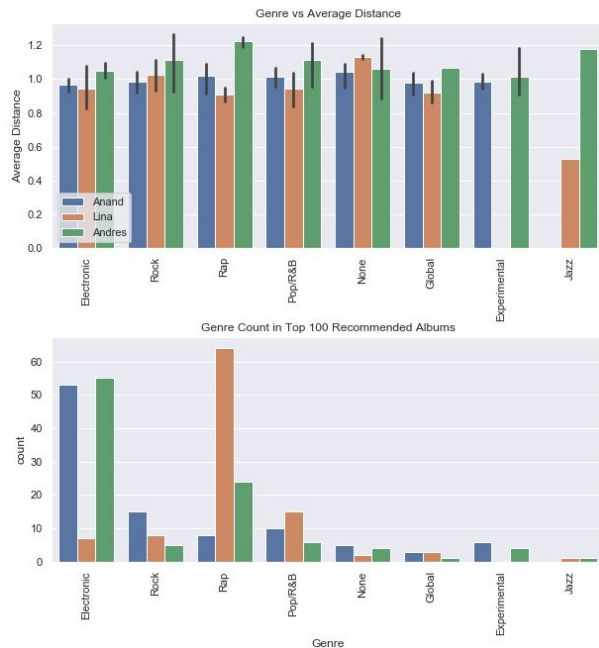
100 Pitchfork albums recs each

**High Danceability &
High Instrumentalness**

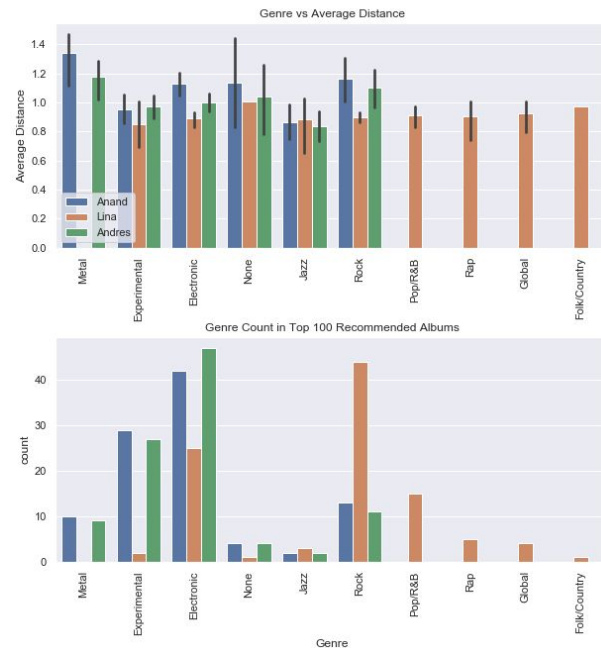


How well do we predict album recommendations for users?

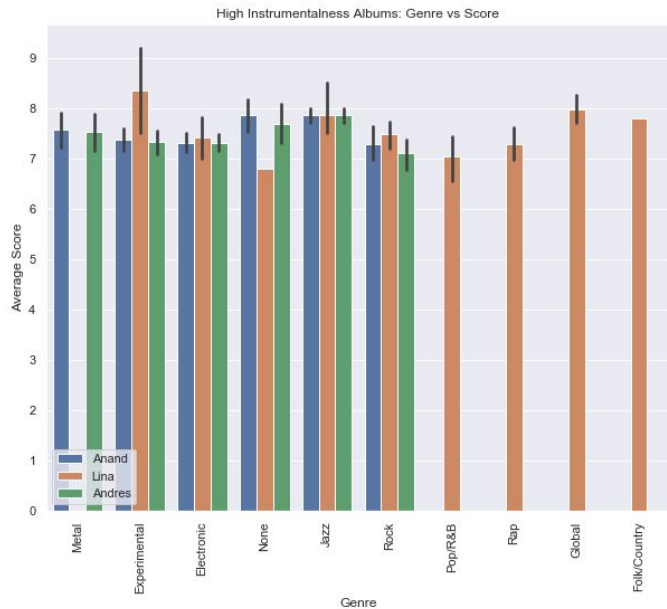
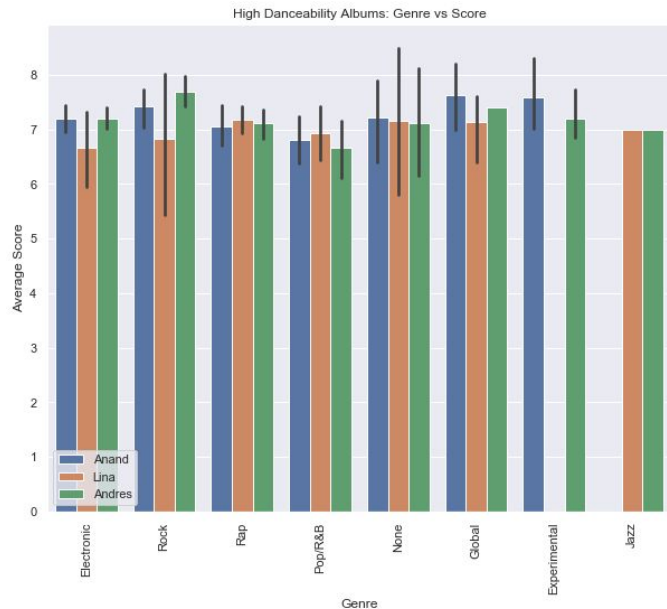
100 High Danceability Album Recommendations: Genre Study



100 High Instrumentalness Album Recommendations: Genre Study

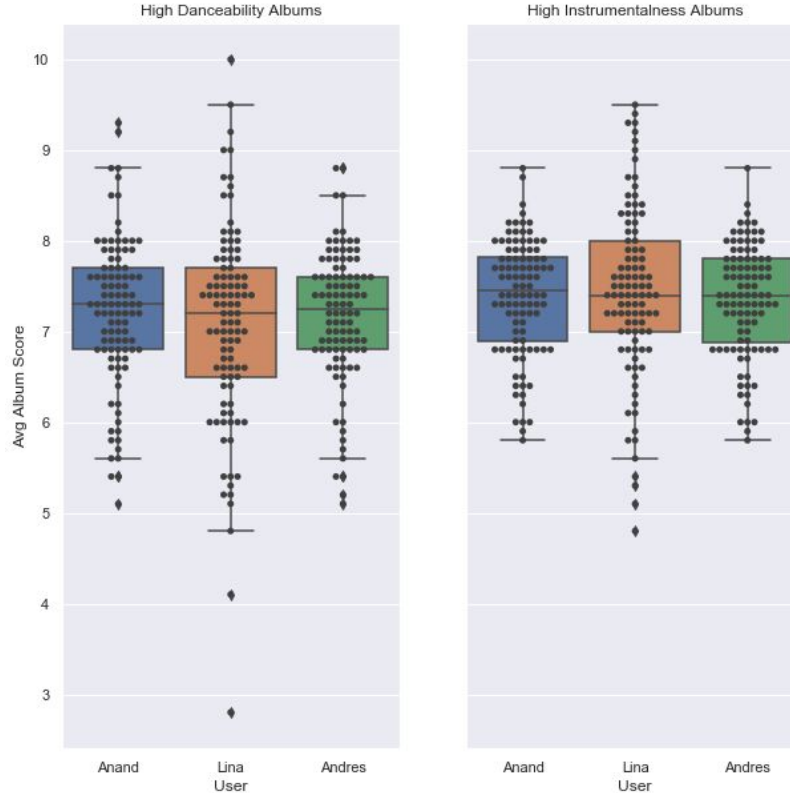


How many different genres do we get recommended and does that have to do with rank?

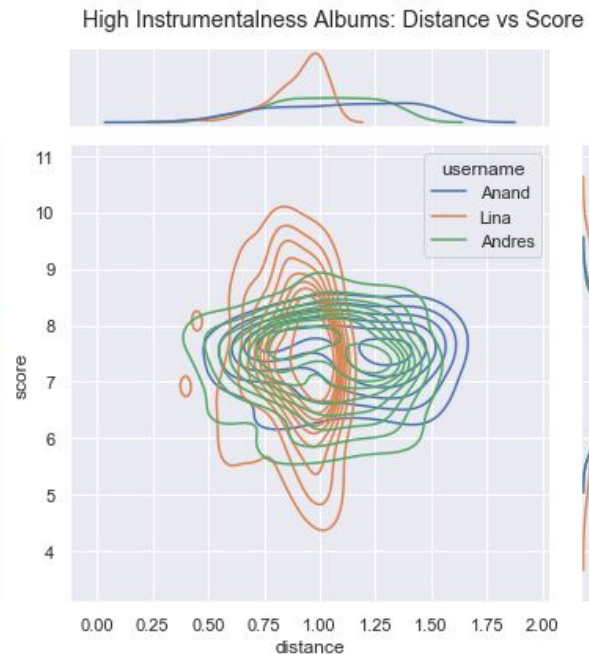
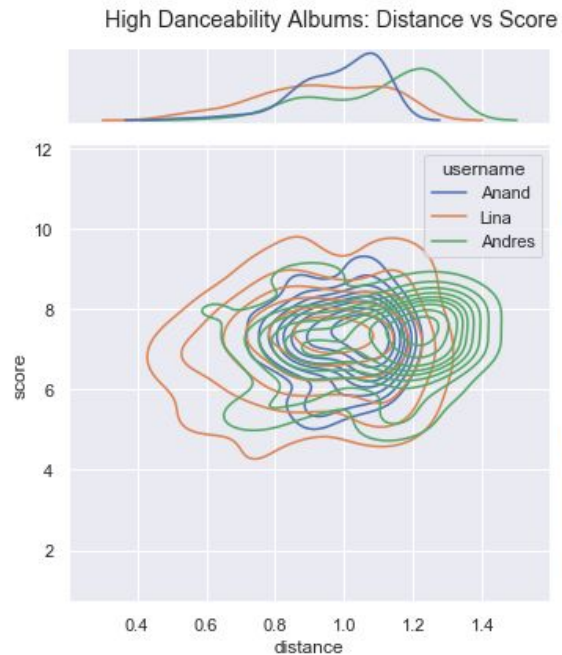


Does our music taste better align with P4K critics for certain genres?

Top 100 Recommended Albums: Average Score Study

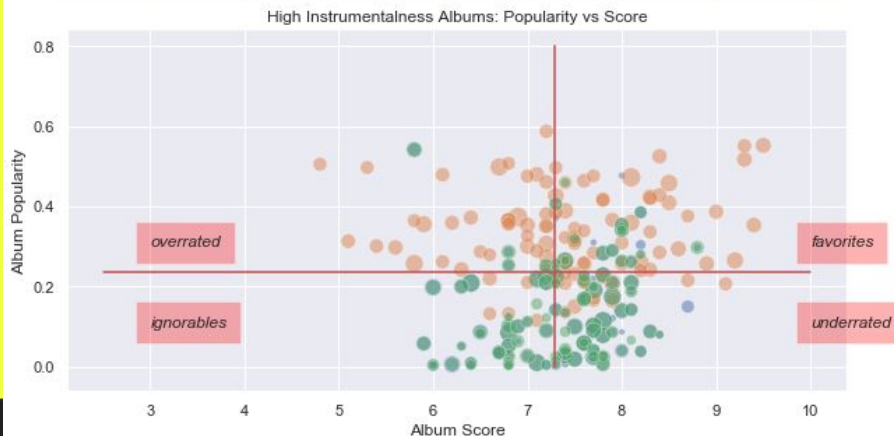
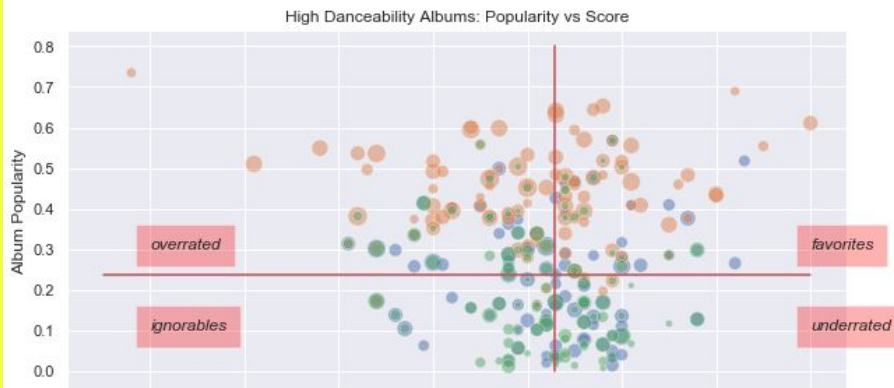


Whose album recommendations have the highest scores?



Is album recommendation rank influenced by the album score?

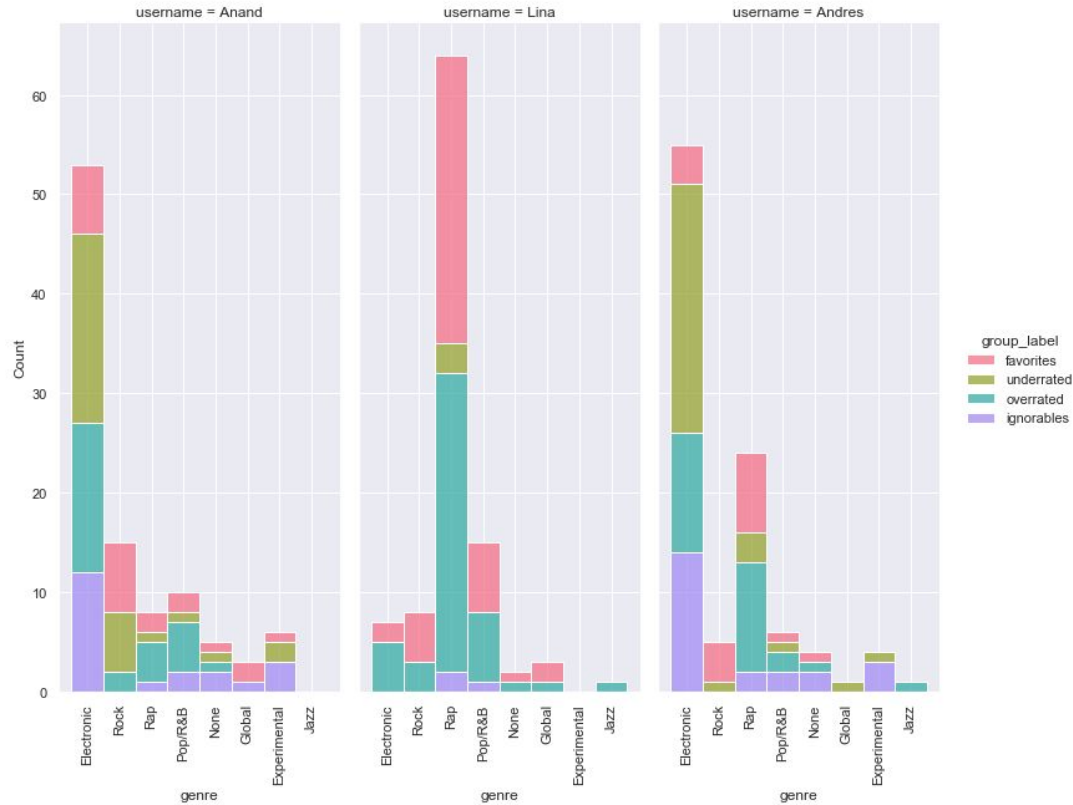
Recommended Albums: Popularity vs Score



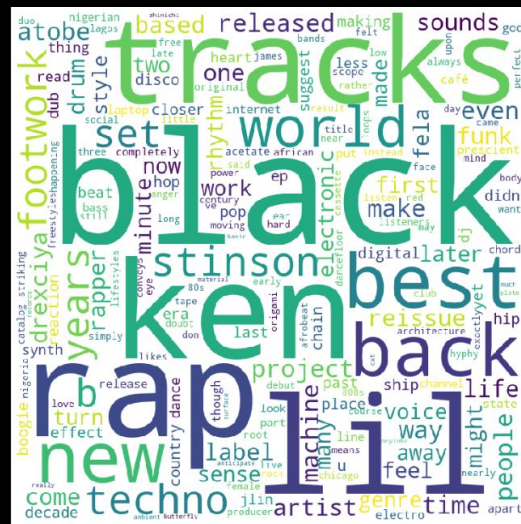
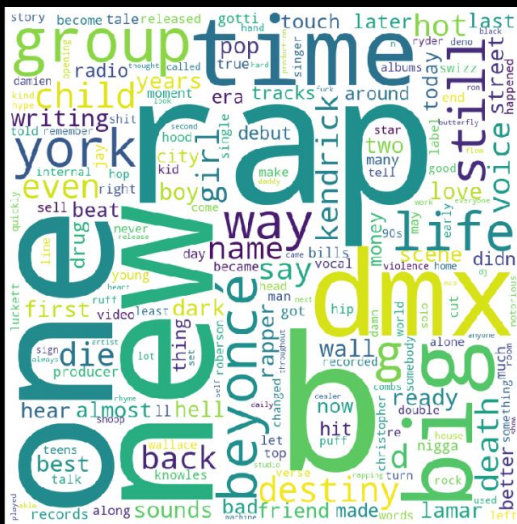
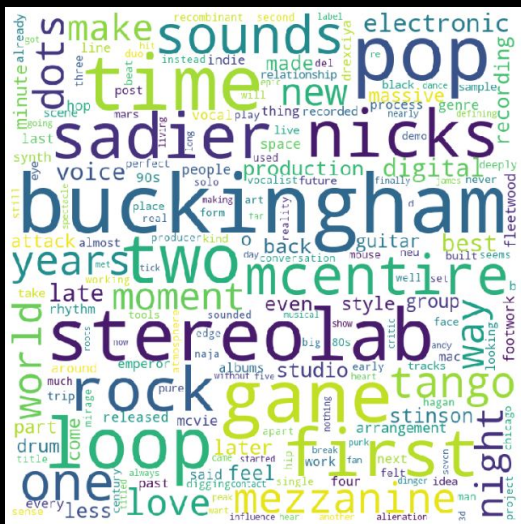
How would Pitchfork
categorize our music
taste?

Anand
Lina
Andres

High Danceability Album Recs: Genre vs Pitchfork Group Label

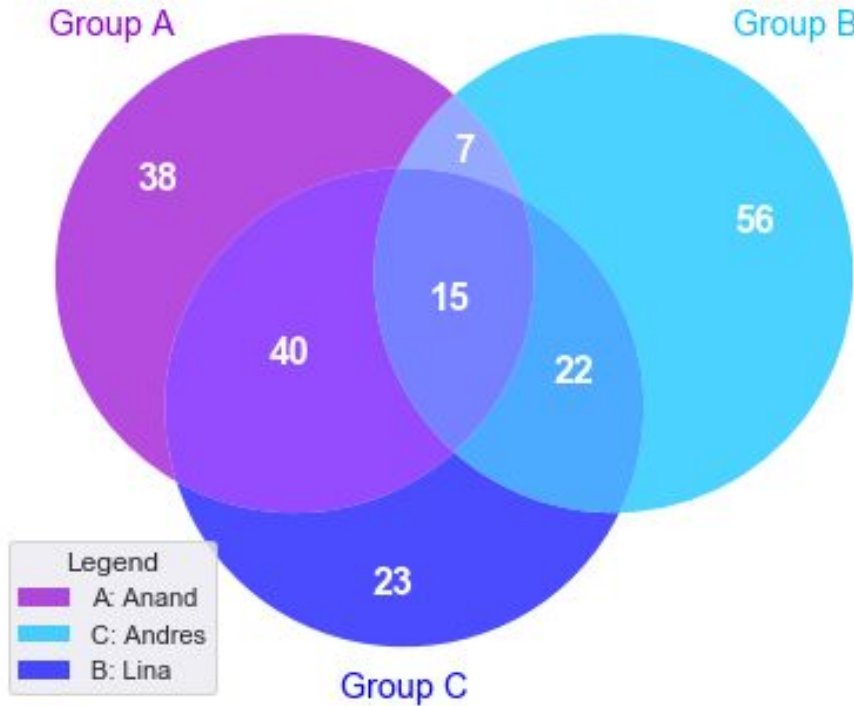


Do these categories vary by genre?



What do the reviews look like for our most acclaimed danceable recommendations?

100 High Danceability Album Recommendations Overlap



Do our specific album recommendations overlap?

thank you.