

Project Earworm: Analyzing Similarity and Shared Fanbases Among Universal Music Group Artists

PROBLEM & CLIENT

Universal Music Group (UMG) is the world's leading music company, capturing 34.1% of the digital music market industry. Given the importance of audience engagement and participation, UMG aims to understand which artists share fanbases to inform potential collaborations. We also aimed to identify factors that influence artist similarity.

DATA



. UMG Artist Sample

Contains a sample of UMG artists and artist information

2. Custom Twitter Dataset

Contains a subset of 167 artists with <100,000 Twitter followers & Twitter follower data (full list of follower IDs)



Team 73 Will Connors, Jose Diaz, Carlotta Hicks, and Madison Ramos

ADDITIONAL DATA PROCESSING

We created a similarity matrix with pairwise similarity defined by percent follower overlap, taken with regards to the smaller of two artists. We then created an artist similarity dataset storing the similarity between every pair of artists with artist genres. This information informed our results.





With our similarity dataset, we were able to identify significant connections between artists that share fanbases. This chord diagram (left) displays the strength of connections between every pair of artists identified as similar (sharing >5% of their followers). We also charted average similarity between genres (above). Most artists are most similar to other artists in their genre.

HIGHLIGHTS

- Our approach visually identifies connections between artists based on shared fanbases, to better inform potential collaborations.
- Our approach demonstrates how we can use Twitter followers to identify factors (e.g., genre) that influence artist similarity and shared fanbases.

CHALLENGES & LIMITATIONS

- Due to delayed data acquisition, we started data cleaning and EDA later than expected.
- The artist sample we received had incorrect usernames for many artists, so we had to validate information prior to any analysis.
- Due to non-academic Twitter API rate limits, we were unable to expand our sample to include artists with >100,000 followers.

CONCLUSIONS

We identified follower overlap between 167 signed UMG artists to inform potential collaborations. Through our dashboard, UMG can identify artists that share overlap and the extent of their overlap. We also showed how genre plays a role in shared fanbases, and our dashboard allows users to view the top genres that a given genre has follower overlap with.

FUTURE WORK

- Expand UMG artist sample & incorporate unsigned artists. We could strengthen our analysis with larger and non-UMG artists.
- Identify additional similarity factors. We would look into additional factors like artist audio features on Spotify (e.g., danceability).
- Create a predictive model. We would create a model to predict the likelihood that two given artists would share fanbases.