Investigating The Foremost Factors Determining Artists' Success

Om Fale
ofale1@binghamton.edu
Binghamton University
Binghamton, NY, USA

Disha Shetty
dshetty3@binghamton.edu
Binghamton University
Binghamton, NY, USA

Yashaswi Hasarali
yhasara1@binghamton.edu
Binghamton University
Binghamton, NY, USA

Harshitha Guru Raj
hgurura1@binghamton.edu
Binghamton University
Binghamton, NY, USA

Shubham Patil
spatil36@binghamton.edu
Binghamton University
Binghamton, NY, USA

ABSTRACT

In today's dynamic and ever-evolving music industry, understanding the factors that contribute to an artist's success is of paramount importance. This interdisciplinary project combines the power of social media and data science through the utilisation of two prominent APIs, Reddit and Spotify, to explore the key determinants behind artists' achievements. This project aims to analyse vast datasets from these platforms to uncover patterns, trends, and insights that shed light on the factors influencing an artist's trajectory in the music industry. Within this virtual sea of text data, including posts, comments, and discussions in subreddits like 'r/spotify,' 'r/MusicRecommendations,' 'r/musician,' 'r/singer,' and more, there is a valuable subset of content that revolves around Spotify music, tracks, albums, and artists. We will leverage Reddit data to gain a deep understanding of fan engagement, sentiment analysis, and emerging trends within the online music community. Additionally, we will tap into Spotify's extensive music streaming data to assess metrics such as play counts, follower growth, and playlist placements.

KEYWORDS

Reddit API, Spotify API, MongoDB, Data Collection, Python, REST

ACM Reference Format:
Om Fale, Disha Shetty, Yashaswi Hasarali, Harshitha Guru Raj, and Shubham Patil. 2023. Investigating The Foremost Factors Determining Artists' Success. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The contemporary music industry stands at a crossroads of unprecedented change, driven by technological advancements, digital platforms, and evolving consumer preferences. In this era of rapid transformation, understanding the intricate factors that determine an artist's success has never been more crucial. Artists, both emerging and established, navigate a complex web of variables that shape their careers. From crafting captivating melodies to cultivating a devoted fan base and securing lucrative partnerships, the road to success is as diverse as it is challenging. To gain insights into this intricate web, we employ a fusion of social media analysis and data science methodologies, harnessing the capabilities of two prominent APIs: Reddit and Spotify. Reddit, as one of the largest and most diverse online communities, provides a treasure trove of fan discussions, reviews, and opinions. We delve into this wealth of data to discern patterns in fan engagement, sentiment analysis, and emerging trends within the music community. In parallel, we leverage Spotify's vast music streaming data, which encompasses play counts, follower growth, and playlist placements, offering a window into an artist's performance in the digital realm. In this project, we'll be gathering real-time music-related data to determine what musical style or performer is popular and trending when - we'll combine all this information from these sources within a certain time frame.

2 DATA SOURCE

We will be using Reddit data from Reddit Stream API wherein we'll be collecting information about top 10 streaming songs and the artists and for this we will be using Spotify API.

2.1 Spotify API

We will be using the Spotify API, to check for fetch the data from a Spotify URL. We will extract a unique track ID and send that as a request parameter to query the Spotify URL and in response we will get the comprehensive data about the artist's name, album, popularity, ranking, followers and such. Later on, with the help of this data we can check the posts related to the artists that are currently trending on Reddit.

2.2 Reddit API

The Reddit REST API provides access to user-generated content on reddit.com, including discussions and stories that have been rated by the community. We plan to use this API to collect user reviews and opinions about songs trending on Reddit, gaining insights into public sentiment and engagement. The API also offers advanced features such as access to user account information and sub-reddit moderation, enabling us to understand user demographics and target specific music-related communities for analysis. This in general,

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

gives data and insights about the trends and discussions on the platform.

3 API METHODS

We will access data resources through standard HTTPS requests to API endpoints. These APIs primarily use HTTP methods like GET and POST. Reddit APIs allow us to find, retrieve, and engage with various resources, primarily utilising the GET method for data retrieval. Spotify follows REST principles, where data is retrieved using GET requests, and the response data is typically in JSON format.

4 MEASUREMENTS AND ANALYSIS

We will be analysing data to identify trending artists or genres based on their popularity on Reddit social media platforms. We'll gather data from both Reddit and Spotify APIs and continuously update our database in real-time as new information becomes available. To analyse the collected data, we'll periodically retrieve it from the database and create graphs or charts using Python's plotting and visualisation libraries. Additionally, we plan to conduct sentiment analysis by assessing positive and negative reactions in tweets related to trending artists or albums.

5 DATA COLLECTION ESTIMATES

Right now the average number of comments in a thread in reddit is a rough estimation. Every 30 minutes we see around 5-10 posts related to music recommendations, but there there is a high chance that there might be some fluctuations from cross-subreddit posts. The data will be updated roughly 10 minutes by our system and we are using MongoDB for collection and storage of this data which supports JSON data storage.

6 DATA FLOW AND IMPLEMENTATION

We first retrieve the data from the Spotify and Reddit APIs using those respective APIs. We save the relevant information in mongoDB database. We will utilize the data such as artist's information, genre, trending songs and albums, music information for analysis and visualization. We will use python "requests" library to fetch the information from the API and we will use the same for authorization of the respective APIs required. Using this track ID, we can then send a GET request to the Spotify API to obtain all the track information, including the artist, album, genre, and other details that we will need to evaluate our data. Using certain keywords we will get the informations from the Reddit.

7 REFERENCES

Martin Pichl, Eva Zangerle, Günther Specht.Combining Spotify and Twitter Data for Generating a Recent and Public Dataset for Music Recommendation,2014.

Data Study on Popular Artists, Weijie Deng, Mar 27, 2018.

Greasley AE, Lamont AM. 2006. Music preference in adulthood: why do we like the music we do? In Proc. of the 9th Int. Conf. on Music Perception and Cognition (ICMPC9), Bologna, Italy, 22–26 August 2006 (eds M Baroni, AR Addessi, R Caterina, M Costa), pp. 960–966. Society for Music Perception Cognition/European Society for the Cognitive Sciences of Music.

Bhattacharjee S, Gopal RD, Lertwachara K, Marsden JR, Telang R. 2005. The effect of P2P, file sharing on music markets: a survival analysis of albums on ranking charts. NET Institute, Working Paper No. 05-26.