



Indie Spotify

The ADA-Unibo Lab project presentation about Spotify and music trends

A.Y 2023-2024

The Research

In our research we're discovering trends and features of modern popular music via the data of Spotify.

Our primary focus is exploring the difference in indie and popular music.

The Research Questions

Taking a look at the dataset some questions sparked:

- What is popularity, how does it relate to followers?
- What is the definition of indie?
- How genres distribute?
- How do labels work?

The Dataset

The [dataset](#) we decided to use lives on Kaggle

- Created by Tony Gordon Jr. using the [Spotify API](#)
- 49 columns in the general dataset
- 5 files in total (~450MB)
- CSV format

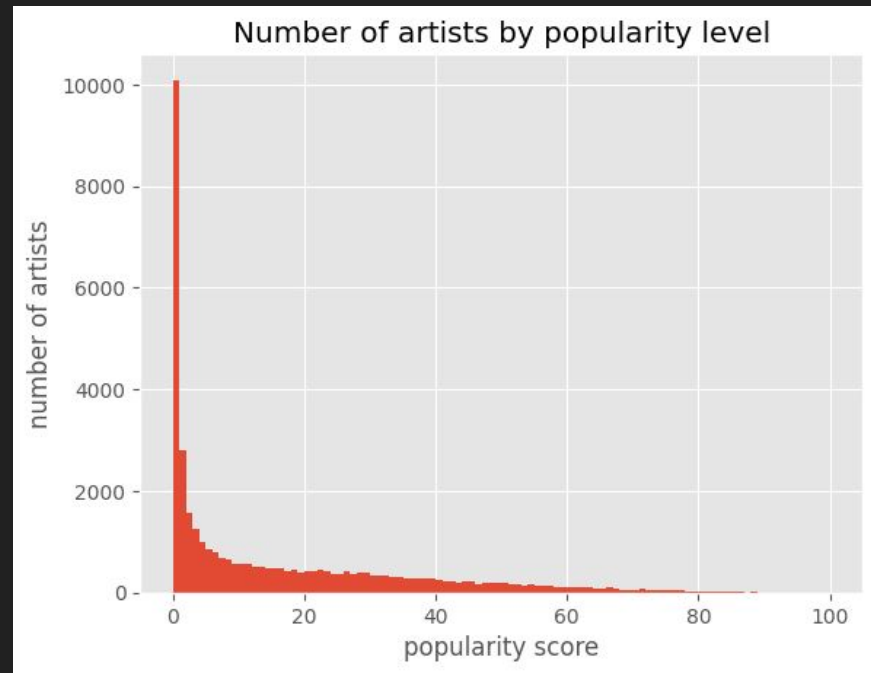
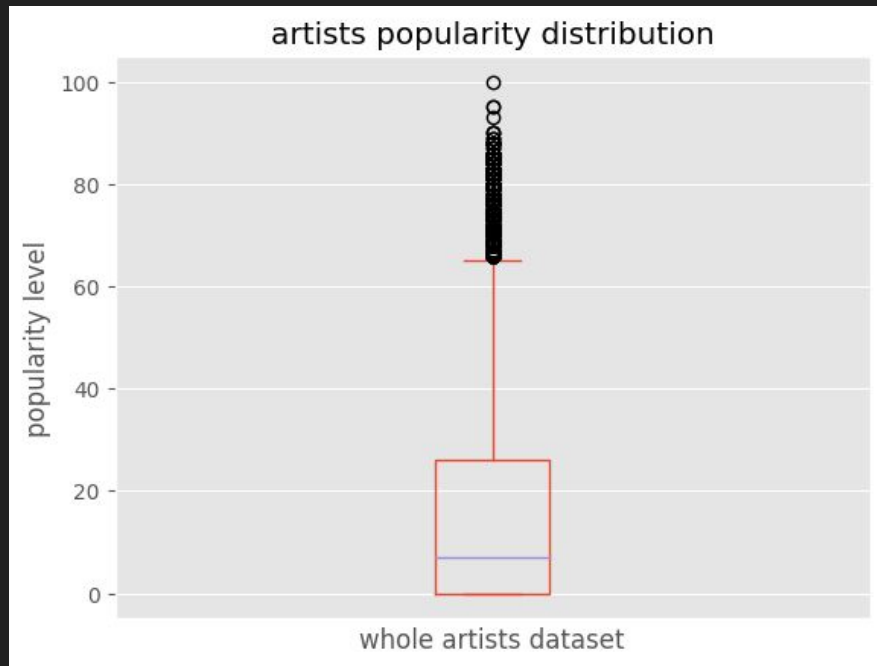
The approach

The approach we used involved simple but powerful data analysis tool.

- [Colab](#) for code sharing and live cooperation
- Pandas for exploration and analysis
- Numpy for linear algebra tasks
- Matplotlib to visualize

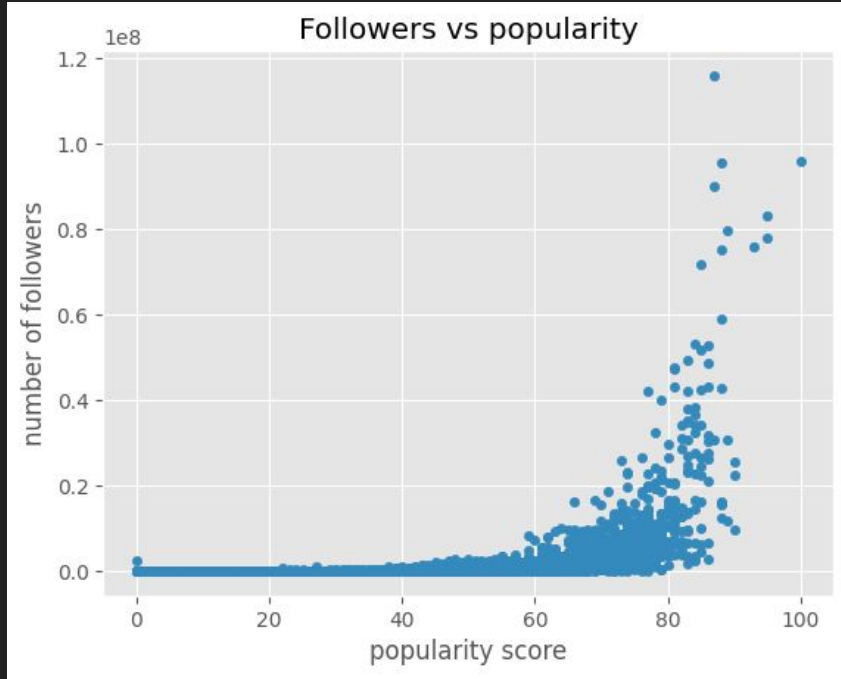
Last but not least the KISS philosophy (Keep It Simple Stupid)

The skewed artist popularity



These two graphs show how skewed and submerged the lowest levels of popularity are in the dataset

Popularity is current trend, Followers are something else



Hypothesis: Popularity and followers are not the same. Popularity measures current trends relating to time.

From the graph we can infer that:

- Popularity and Followers are somehow correlated but different
- The highest number of followers does not correspond to the highest popularity
- Followers are an absolute measure, therefore Popularity must be somehow time related (current trending artists)

Popularity is current trend, Followers are something else

	name	artist_popularity	followers
37011	Taylor Swift	100	95859165
37010	Drake	95	83298497
37009	Bad Bunny	95	77931484
37008	The Weeknd	93	75945958
37007	Peso Pluma	90	9775920

	name	artist_popularity	followers
36994	Ed Sheeran	87	115998928
37011	Taylor Swift	100	95859165
36995	Ariana Grande	88	95710972
36993	Billie Eilish	87	89996504
37010	Drake	95	83298497

Being **Taylor Swift** the most popular at the time and date of this dataset (2023) we can say with a certain degree of certainty that our hypothesis is correct.

What do we mean by “indie” anyway?

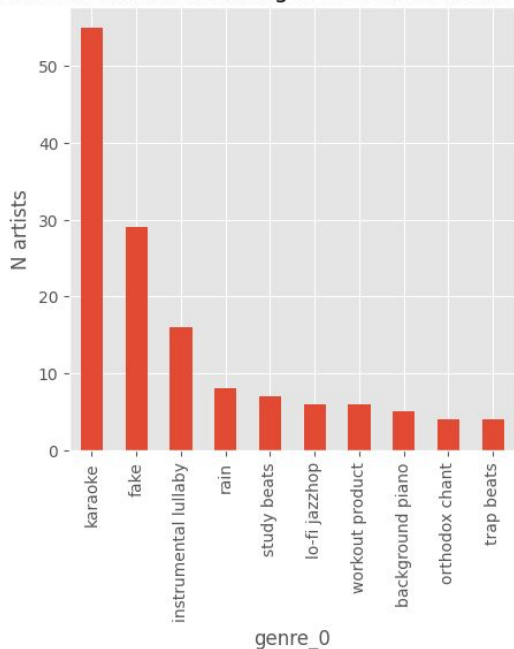
In the scope of this project we investigate and treat as indie the submerged and ignored part of the Spotify data.

We decided that the right approach was not selecting those artists that have the “indie” word stand-alone or associated to another genre in the genre columns. (In order to avoid indie artists of the past that are popular and well known today).

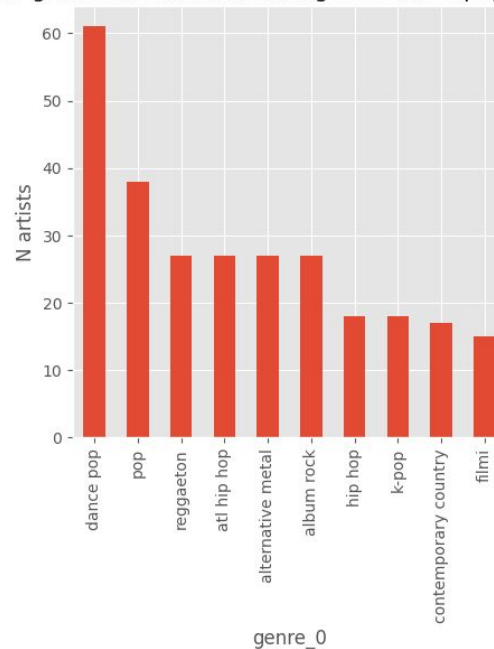
Our approach was to select the **1000 least followed ones**, avoiding the artists with 0 followers that are not very representative in our opinion.

Indie and Popular artists have different taste

top 10 genres distribution among 1000 least followed (indie) artists

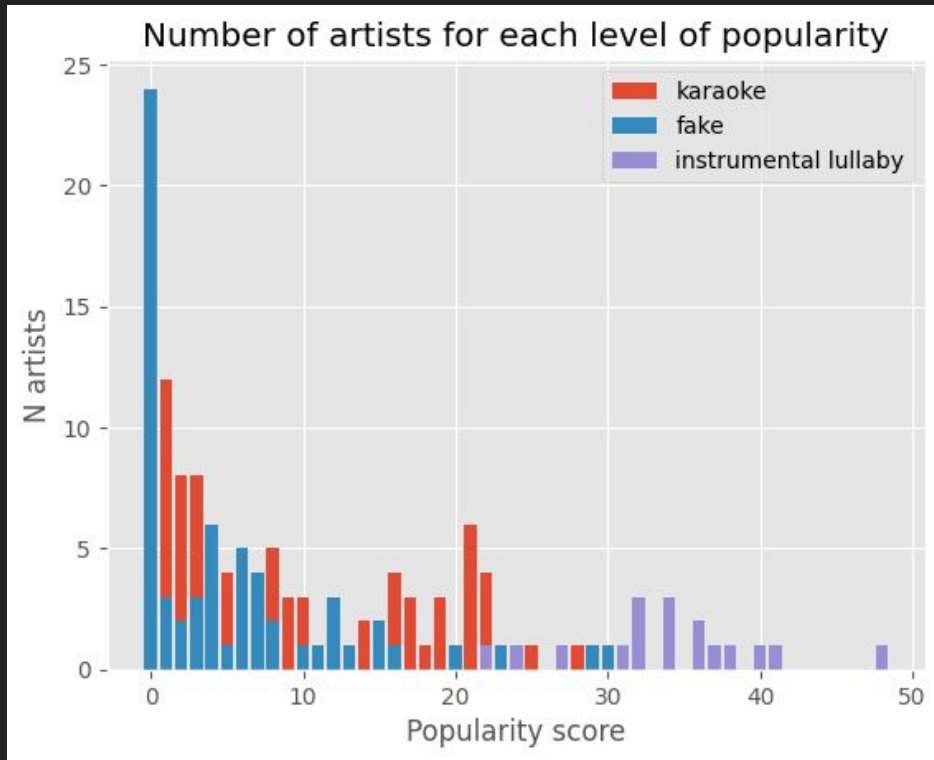


top 10 genres distribution among 1000 most popular artists



Our hypothesis on genre distribution was correct, different popularity leads to different leading genres. Popular ones have few surprises in store, while indie ones were quite unexpected.

Genre popularity evolves in strange ways

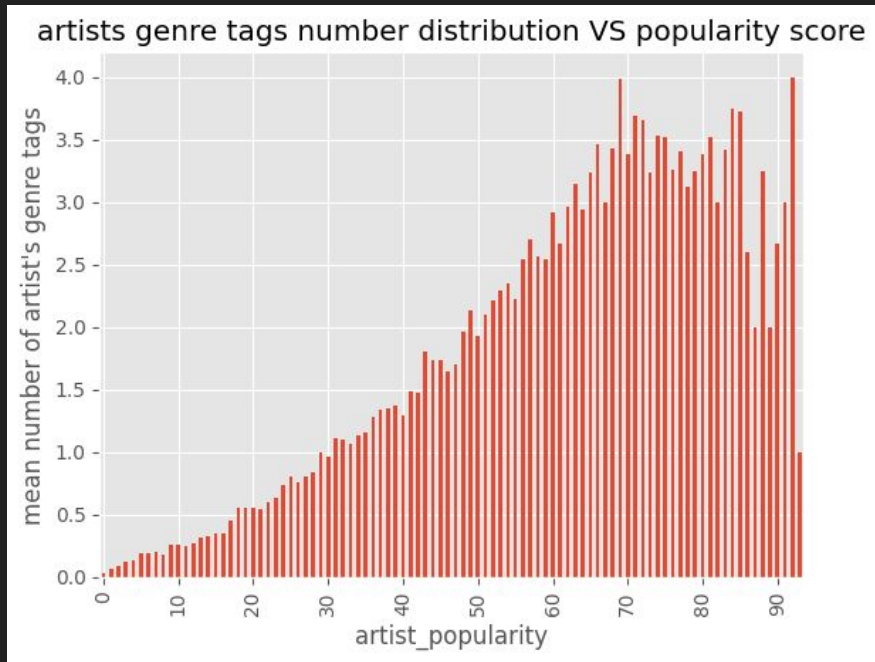


In this graph we sampled different peculiar genres from the top 10 indie genres seen in the previous chart.

Popularity for these genres is varied and present different curves:

- **Fake** is a placeholder genre for covers, unidentified reproduction of existing music and it makes sense to have high number of “artists” in the 0 tier.
- **Karaoke** is distributed in a more uniform way to then disappear at almost 30 popularity
- **Instrumental lullaby** is a very niche but consistently popular genre that starts at 20 and finishes with an artist at almost 50 popularity

Indie music has many genres, indie artists don't.

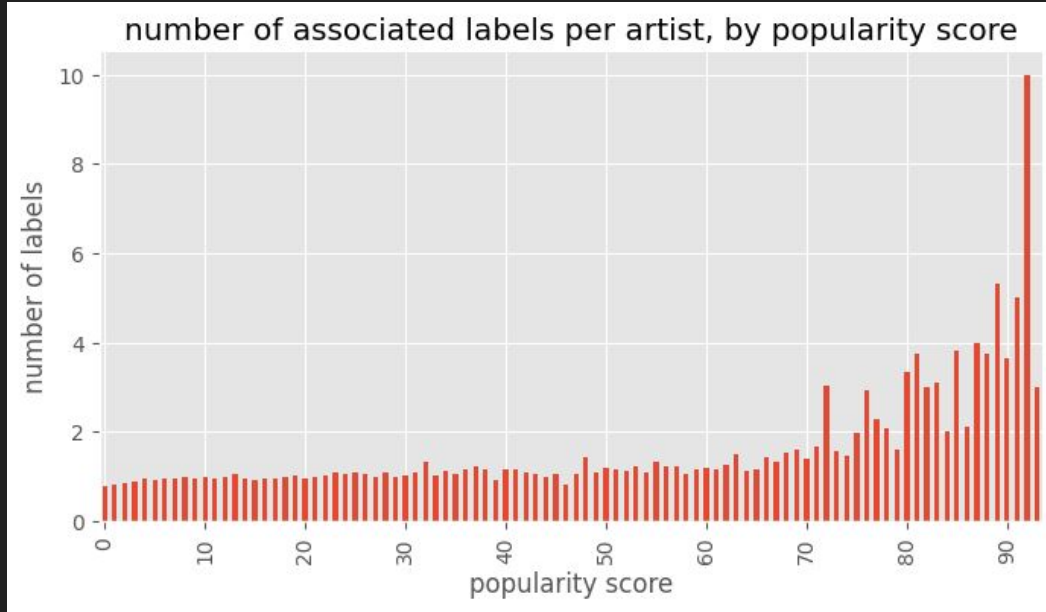


11.06% of total genres are also present among most popular artists (top 1000)

22.51% of total genres are also present among indie artists (least 1000)

This graph seems to collide with the percentages showed above, but they're not actually. It just explains that indie and emerging artists have more variety when they pick a genre, but having a smaller production they tend to have fewer genres for each one of them.

Popular artists have few labels to choose from.



For the most part of the popularity range, the distribution grows apparently linearly

For very popular artists the distribution starts growing explosively

This goes to show that major artists tend to get much more attention from the labels

Conclusions and take away of the Project

This study allows us to understand some key features of music trends and how they are stored and treated by the main music provider today.

- Submerged parts of the dataset represent and allow us to peak into different genres and interests from higher popularity tiers.
- Spotify is easy-access for new artists but it's not that easy to become popular.
- Labels follow a different pattern in Spotify from the one we can expect.
- Indie and popular music are different in their genre distribution
- Indie music is more creative with genres but one indie artist is usually less so than one popular artist
- Labels treat indie and average segment equally, and treat the most popular artists differently

Thank you for your attention!

Francesco Alaimo - Nikolai Gorbachev