SPOTIFY MUSIC: TRACKLIST STREAMING NETWORK

GROUP 27

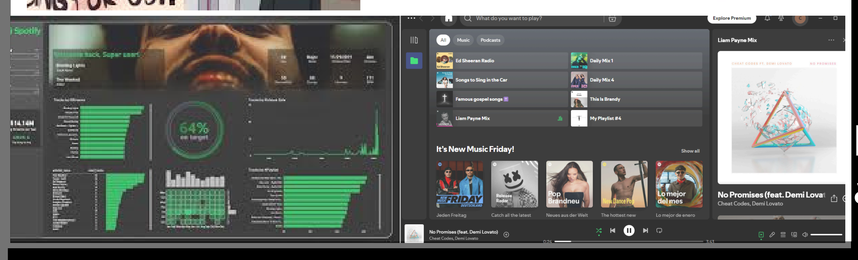
- GEORGE OBAZI ISEK
- KAINAT TANVEER BUKHARI
- STEPHANIE CHIOMA IBEANU
- FADASH MOHAAMED KUTTIKUTHIA

Research Question:

Is there a correlation between track popularity and the discovery of new artists in a tracklist database?

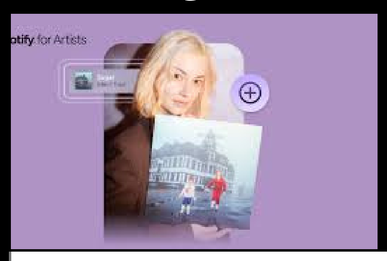
Motivation:

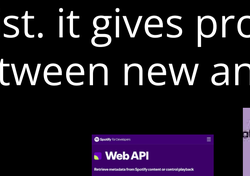
Streaming platforms are essential for promoting new artists. This research question analyzes whether high track popularity are influential tools for introducing new artists to a larger fan base.



Data:

- Data was collected from Spotify Web API. The spotify library was used to retrieve data using the client ID and Secret we generated.
- Data cleaning and processing were made.
- From our previous data, we were able to create a new variable "Is New Artist", which is part of our research question.
- Tools: Pandas, Spotify library for API , Matplotib, Python, Seaborn .



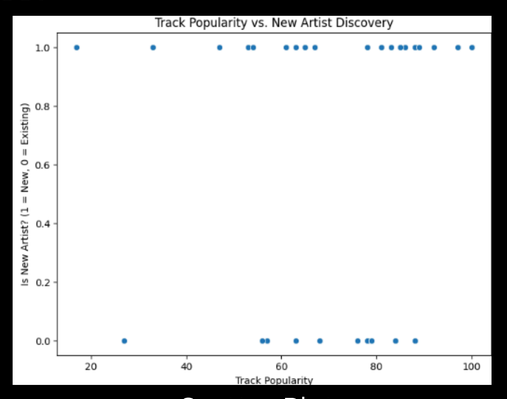


Methods:

- Point-Biserial correlation was used to perform statistical test, whether there a is relation between track popularity and discovery of new artists. It was used since a variable is binary.
- We also created scatter plot, bar plot and historam to explore data patterns in our visualisation.

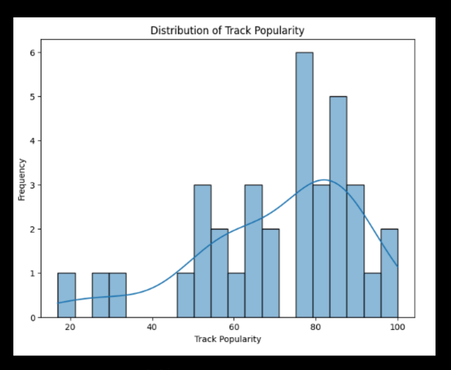
Results:

- From our analysis using Point-Biserial correlation, we found the correlation of 0.0508 with P-Value 0.771, which shows a very weak correlation between track popularity and new artists and there is statiscallly insignificant among variables.
- Scatter plot: It shows no group clustering. it was scattered randomly. Therefore no correlation among the two variables.



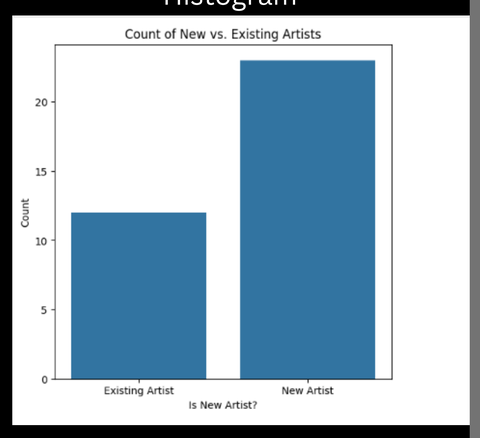
Scatter Plot

- Histogram: It shows that the track popularity of 70-75 has the highest frequency in the distribution.



Histogram

- Bar plot: This shows the comparison between new artist and existing artist. it gives proper proportion between new and existing aritsts.



Bar Plot

Conclusion:

- In our analysis we were able to find out from our correlation that track popularity on its own cannot leads to the discovery of new artists.
- This analysis was limited to only spotify platform dataset, therefore making the size of our dataset small.
- For future analysis or research, other data from Apple Music, YouTube etc. should be included.