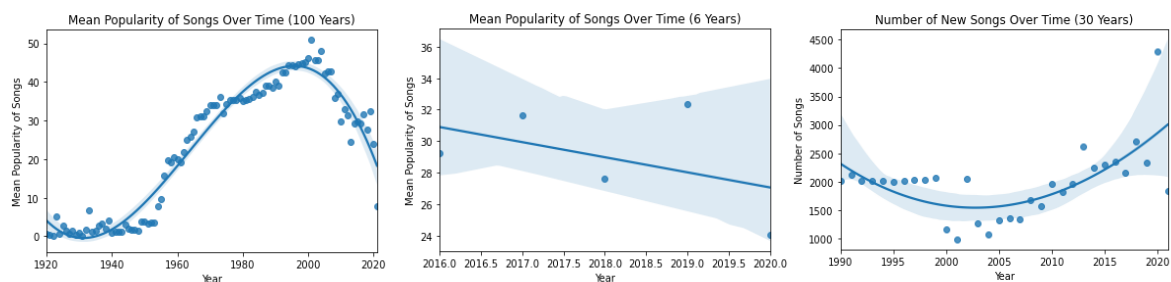


TRENDS IN THE MUSIC INDUSTRY

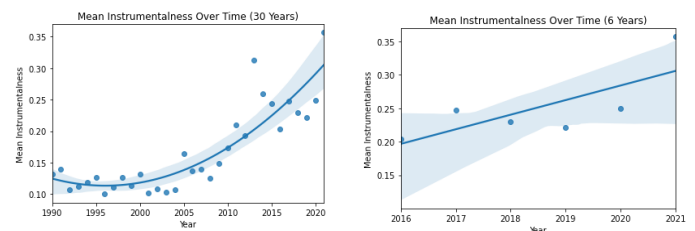
POPULARITY OVER TIME.

One of the strongest indicators in predicting the popularity of songs was found to be the year in which it was published. This rising trend up until 2000 coincides greatly with the rise of popular music in the 60s and 70s followed by the creation of streaming services in the 2000s allowing more people to start publishing their music leading to more songs and a lower mean popularity. The trend of mean popularity dropping continues over the last five years* along with the increase in songs published, with 2020 being the highest number in the last 30 years.



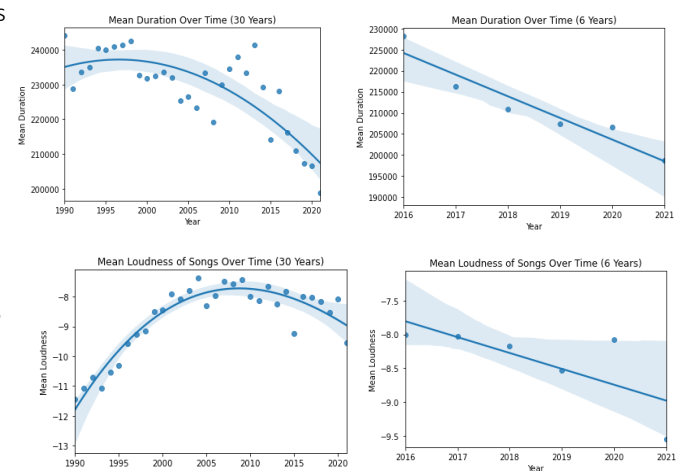
INSTRUMENTALNESS

Music with no / fewer vocals appears to be increasing in popularity with more being produced now than ever before and this trend following over the last five years. Even though the model predicts instrumentalness as a bad trait it may be worth looking at an instrumental act to capitalise on this trend.



DURATION

Songs are getting shorter and shorter over time and this trend appears to be accelerating in the recent years, dropping by approximately 30 seconds over 30 years. Bands that produce shorter songs may potentially do better over the coming years by following this trend and the model also predicts this.



LOUDNESS

Loud songs seem to have recently peaked in numbers but remain high only slowing a gradual dip, in future lower loudness may be preferable but for now louder acts appear to have higher popularity and are certainly more common so looking for this may be beneficial.

BEST ARTIST TO LAUNCH ON SPOTIFY

How?

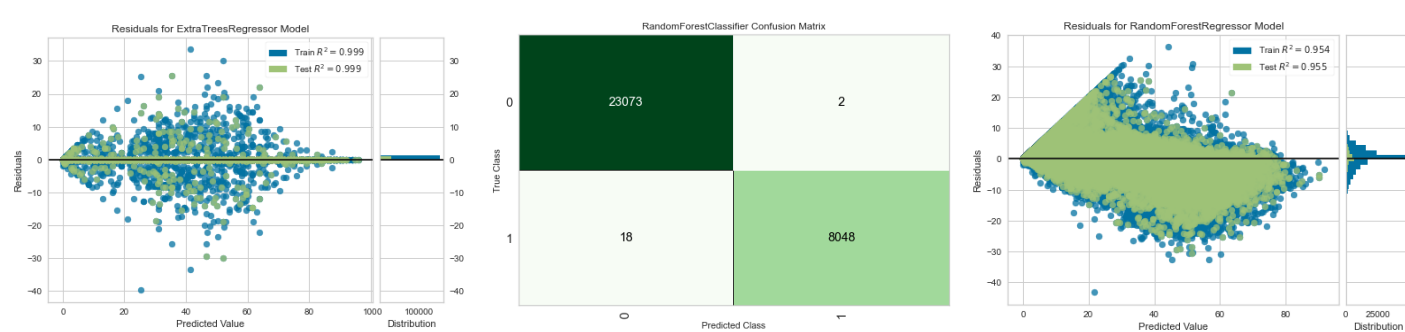
After analysing the data, we have produced several models aiming to predict the popularity of a song based on its characteristics. We found the most successful prediction of pure popularity to come from an Extra Trees regressor and the best categorizer of popular or not to be a Random Forest classifier. Both these methods work by creating many decision trees to try and determine the popularity of the input data. For the regressor we valued low mean absolute error to try and get a precise prediction of

*Large and sudden drop in number of new songs in 2021 purely because we do not have complete data for the year.

popularity, while for the classifier we valued precision in order to be sure our answer was correct. To aid in reliability we also used a second regression algorithm (Random Forest) and a neural network to reinforce our confidence in our popularity projections.

For the regression model we obtained a mean absolute error of 8.3 popularity points when training, and for the classification model we achieved an accuracy of 83.5% and a precision of 74%. The neural network had slightly higher mean absolute error of 9.0 popularity points but would still be useful in confirming our decision.

Generating a plot of residuals for the regressors and a confusion matrix for the classifier and its test data we see a very high precision for the classifier and good fit for the regressors.



WHO?

Unanimously it was predicted the most successful artist would be Elliot Tempest. This gives us a high degree of confidence that you should put that artist forward and invest in them for your entry into streaming services.

Artist	Extra Trees Regressor	Random Forest Regressor	Random Forest Classifier	Neural Network
Elliot Tempest	45	51	Popular	30
Amy Apollo	15	21	Not Popular	26
Pocket Rockets	11	17	Not Popular	3

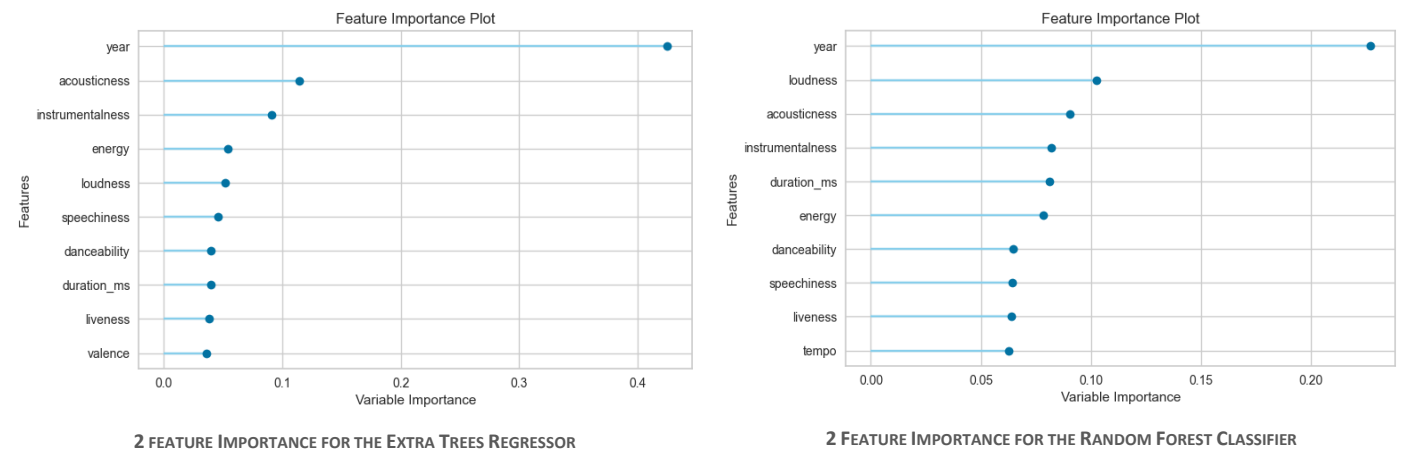
We defined popular for the classification algorithm as in the top 25% of all songs for popularity so all models indicate that Elliot Tempest should be the most popular at approximately 40 popularity, Amy Apollo next at approximately 20 points, and Pocket Rockets last at around 10 points.

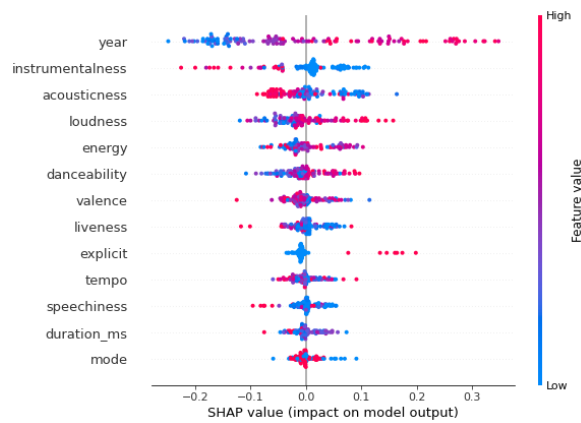
WHY?

Looking at feature importance for the classification and regression models used, we see the same characteristics as important for popularity: year, instrumentalness, acousticness, loudness, danceability, and duration.

Since we cannot control the year, we look at the other categories to see what is preferable. Commonly agreed is that a higher loudness is good, a higher danceability is good, lower acousticness, duration, and instrumentalness are also all good.

Elliot Tempest has the lowest instrumentalness out of all the acts, he is the loudest and highest energy, and has the lowest acousticness. I believe these to be the main factors as to why Elliot was predicted to be the most successful despite having lower danceability.



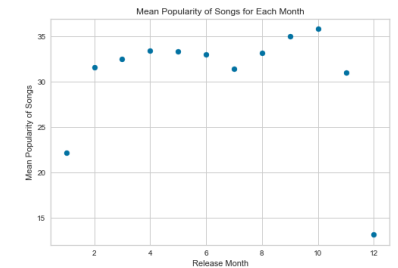
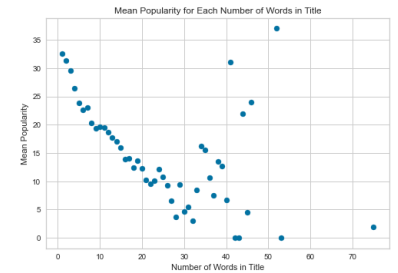


WHAT?

Since a song title is yet undecided, I decided to investigate the affect of title on popularity. I observed clear trends in popularity for shorter titles with fewer words and characters. This could be because these titles are catchier and easier to remember. This leads me to recommend shorter, perhaps single word, titles that would be easy to remember and catchy for an audience e.g “Love”. If a longer title is desired, I would only advise to keep it as short as possible.

WHEN?

When looking at mean popularity against release month October gives the highest mean popularity, probably due to the run up to Christmas. Therefore, I would recommend releasing in that month to have the best chance of a popular song.



COLLABORATION

To decide a suitable artist for collaboration, I first clustered the songs into 6 broad genres with similar characteristics. From there I found all songs in the same genre as Elliot Tempest, selected only the songs released in the last two years, and found all the unique artists from that subsection of songs. I calculated the mean popularity of that artist’s songs in the subsection and ensured they had at least 2 songs in the section before simulating a collaboration by combining average song characteristics. Then after producing the collaboration score, calculating the cost I was left with approximately 15 artists.

Out of those artists I recommend Lewis Capaldi if money is no object, the projected collaboration score had a popularity of 71, he has a popularity of 82 over the last two years and has 5 songs in the same genre in the last two years. However, his cost is the highest of all the artists with 63.2 (cost ranges from 19.6 to 63.2). If that is too high a lower budget option would be Shawn Mendes at a cost of 47.5 (collaboration score of 55, popularity of 82, 3 songs in the genre) or the lowest budget option of Sia for only 24 (collaboration score of 55, popularity of 78, 2 songs in the genre).

WHICH ARTISTS PERFORM BEST AND ARTISTS TO LOOK OUT FOR

As shown in the “Why?” section, we see that an ideal artist would be loud and danceable with short songs that do not have many acoustic or instrumental elements. This seems to broadly line up with the pop genre or perhaps hip hop. The models also seem to indicate explicitness is not a bad thing for artists, so you need not be worried about that decreasing popularity. It would appear liveness and speechiness are also not massive indicators of popularity so you could experiment by trying new artists with live recordings or more hip hop / rap style artists.

Regarding trends, perhaps quieter artists will have higher popularity in the future despite the model. In a similar regard, the rise of instrumental acts can only lead to some being popular and in future a model trained only on recent songs may predict instrumental songs to lead to a higher popularity.

For now though, the safe bet is to look for a loud, energetic act with short, danceable songs with vocal content and non-acoustic instruments.