

WHAT MAKES A SONG POPULAR?



Kate Chen, Christine Go, Carley Lerner, Bailey Heit

PURPOSE

determine the qualities of a song that influence the popularity on spotify



ACOUSTICNESS

ENERGY

SPEECHINESS

GENRE

kaggle

DANCEABILITY

LIVENESS



LOUDNESS

ENERGY

KEY

TEMPO

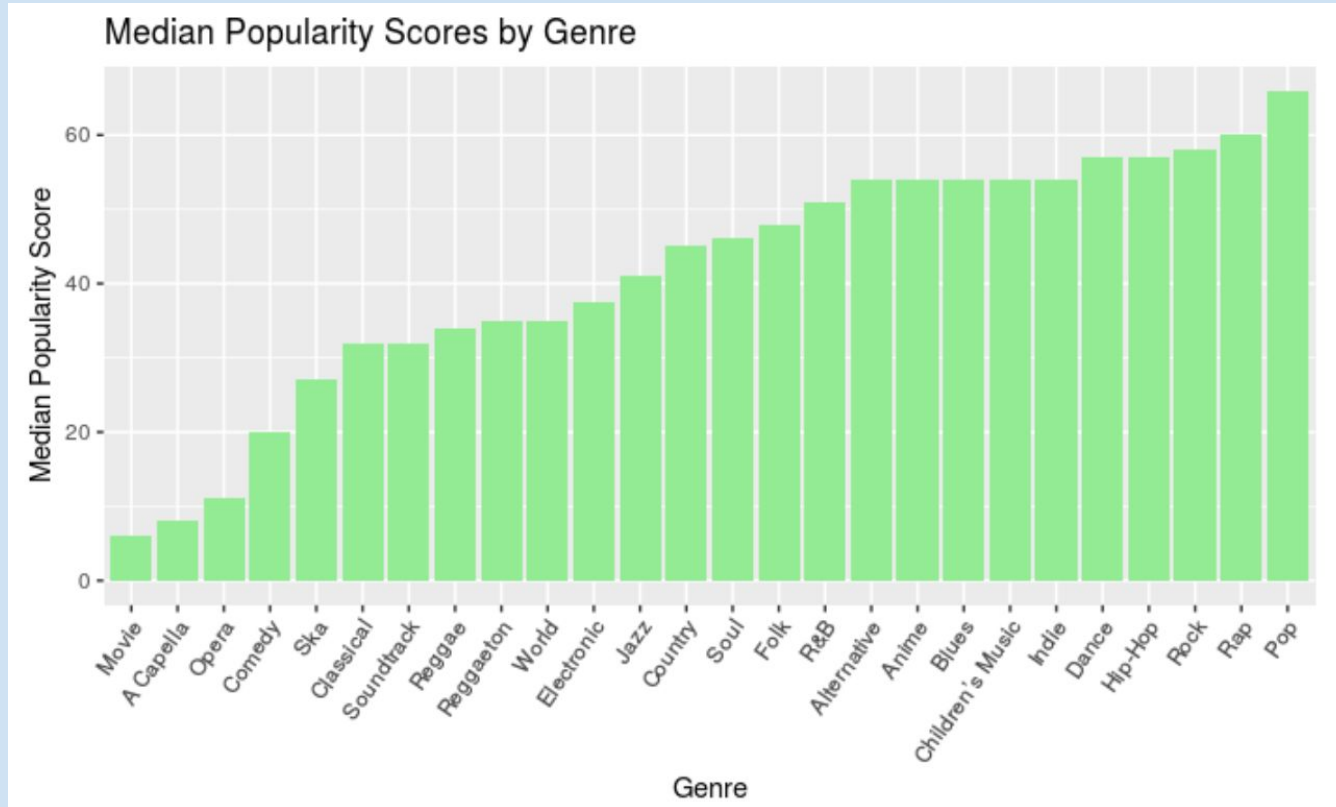
A decorative graphic of a rainbow with musical notes and a treble clef. The rainbow is composed of several horizontal bands of color: pink, purple, blue, green, yellow, and orange. A black treble clef is positioned at the right end of the rainbow. Several musical notes are placed along the rainbow's length. These notes are stylized with faces: some are red with a smile, some are yellow with a smile, some are blue with a sad face, and some are green with a sad face. The notes are connected by a black line, suggesting a melody.

1.

VISUALIZATIONS

exploratory data analysis to gauge correlations
between different variables and song popularity

Popularity by Genre



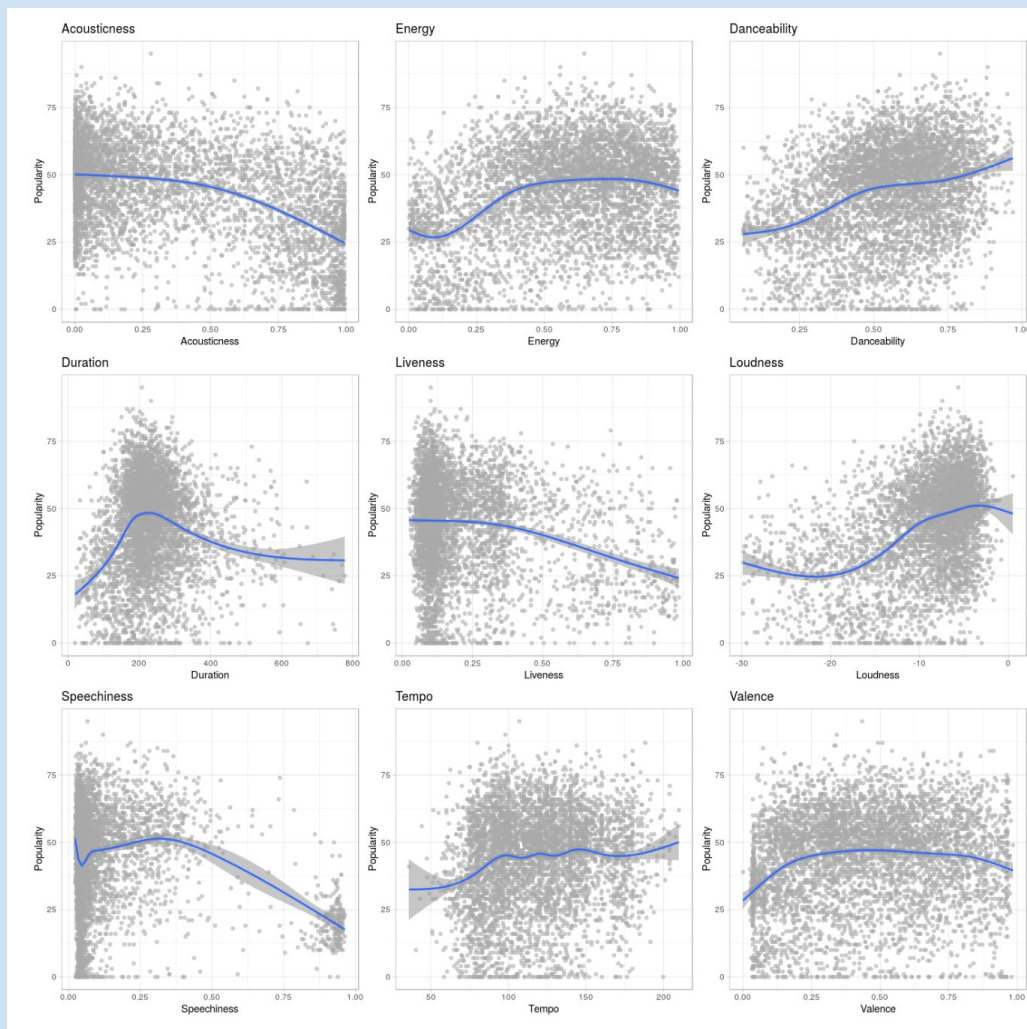
Trendlines

Positive relationship:

- Energy
- Danceability
- Loudness
- Tempo

Negative relationship:

- Acousticness
- Speechiness
- Liveness



A decorative graphic of a rainbow with musical notes and a treble clef. The rainbow is composed of several horizontal bands of color: pink, purple, blue, green, yellow, and orange. A black treble clef is positioned at the right end of the rainbow. Several musical notes are placed along the rainbow's path. From left to right, there are two green notes with pink faces, a black note with a black face, two red notes with red faces, a yellow note with a yellow face, a blue note with a blue face, and two orange notes with orange faces. The background is a solid light blue.

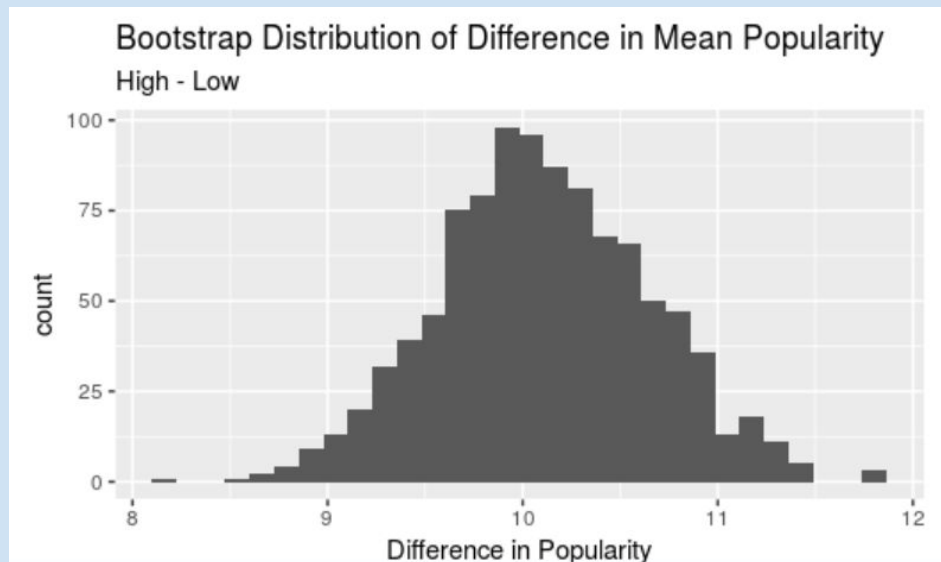
2.

BOOTSTRAPPING

determine with confidence the extent to
which variables impact popularity

High V. Low Danceability

Comparing **the highest 25%** and **lowest 25%** values to generate a bootstrap distribution of the difference in mean popularity scores



lower_bound
<dbl>

9.38982

upper_bound
<dbl>

11.57811

Confidence Intervals

95% Confidence intervals were calculated for the difference in mean popularity between the **highest 25%** and **lowest 25%** values of each of these variables.

Variables	Lower Bound	Upper Bound
Danceability	10.0	12.3
Energy	11.1	13.8
Loudness	18.4	20.9
Acousticness	-20.8	-28.4

Variables	Lower Bound	Upper Bound
4/4 v. Not 4/4	10.4	13.3
Rap v. Pop	3.82	7.04

A decorative graphic of a rainbow with musical notes and a treble clef. The rainbow is composed of several horizontal bands of color: pink, purple, blue, green, yellow, and orange. A black treble clef is positioned at the right end of the rainbow. Several musical notes are placed along the rainbow's length. From left to right, there are two pink notes, a black note, two red notes, a yellow note, a blue note, and two orange notes. Each note has a small, colorful smiley face on it. The background is a solid light blue.

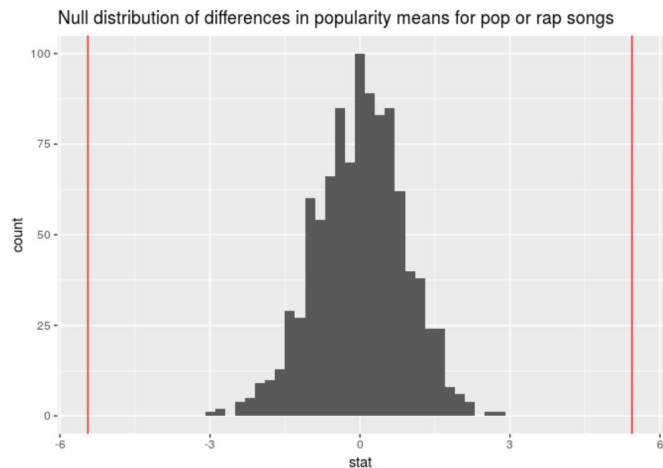
3.

PERMUTATIONS

exploring the difference between relationships of different factors to song popularity

Pop V. Rap Genres

```
## # A tibble: 26 x 2
##   genre      medianpop
##   <chr>         <dbl>
## 1 Pop           66
## 2 Rap           60
## 3 Rock          58
## 4 Dance         57
## 5 Hip-Hop       57
## 6 Alternative   54
## 7 Anime         54
## 8 Blues         54
## 9 Children's Music 54
## 10 Indie        54
## # ... with 16 more rows
```



```
## # A tibble: 1 x 1
##   pvalue
##   <dbl>
## 1     0
```

The p-value is 0, so there is evidence that **there is a difference** in mean popularities between pop and rap songs

Continuous Numerical Variables

Variable	Observed Difference	P-Value
Energy	10.35	0
Danceability	9.2	0
Speechiness	21.43	0

Within each of these variables, there is a **significant difference** between high and low values of each, indicating that the trends are impactful on popularity.

A decorative graphic of a rainbow with musical notes and a treble clef. The rainbow is curved and has seven distinct color bands: red, orange, yellow, green, blue, purple, and pink. Various musical notes are placed along the rainbow. Some notes are colored (red, yellow, blue) and have smiley faces. Others are black. A large black treble clef is positioned at the end of the rainbow on the right side.

4.

LINEAR MODEL

finding the best fit model to maximize
popularity of a song

Selected Model Excluding Genre

Adjusted R^2 Value: 0.30

Intercept	57.28	Liveness	-11.85
Acousticness	-13.77	Loudness	0.78
Danceability	16.52	Speechiness	-7.72
Energy	-6.24	Tempo	0.012
Instrumentalness	-3.26	Valence	-9.29

Selected Model Including Genre

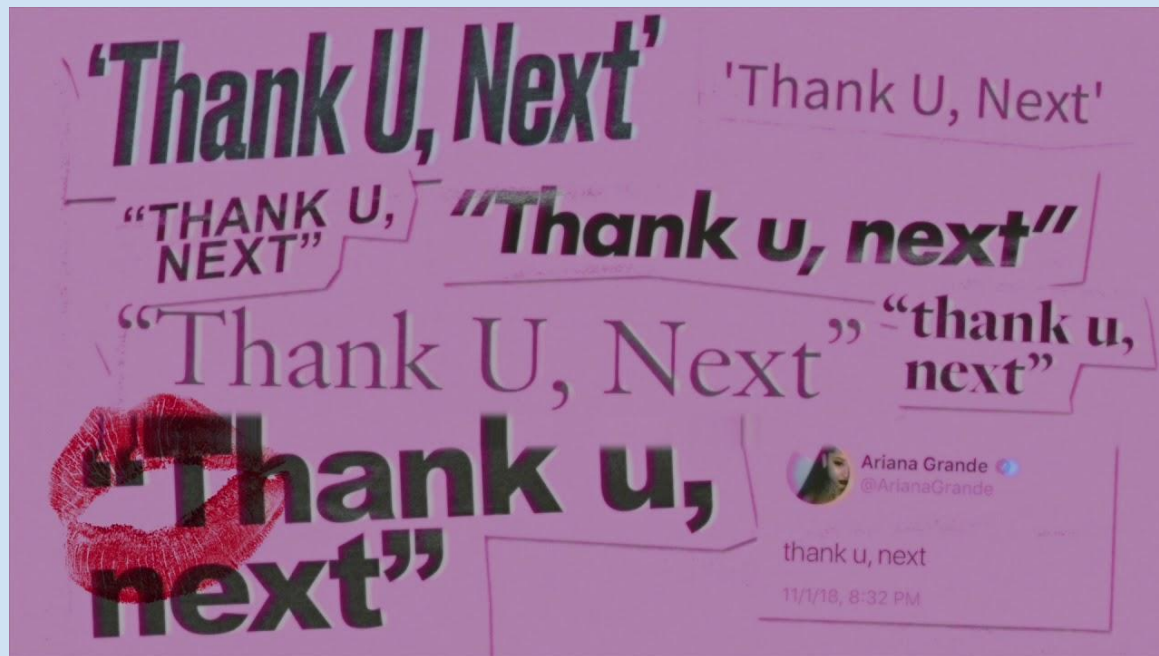
Adjusted R² Value: 0.71

Intercept	9.15	Danceability	3.28	Loudness	0.12
Speechiness	-2.055	Valence	-1.32	Alternative	44.97
Anime	44.67	Blues	45.81	Children's music	45.82
Classical	24.01	Comedy	13.64	Country	37.92
Dance	47.85	Electronic	28.50	Folk	40.34
Hip-hop	49.09	Indie	45.57	Jazz	31.07
Movie	1.94	Opera	4.96	Pop	56.81
R&B	42.54	Rap	51.31	Reggae	26.72
Reggaeton	27.86	Rock	49.95	Ska	19.29
Soul	37.95	Soundtrack	25.83	World	27.13

Discussion

- **71.0% of variability in popularity** can be explained by regression model
 - **Strong positive overall relationship**
- **Genre** is the strongest predictor of song popularity
 - Popularity **increased most** when song is **POP** or **RAP**.
 - **MAKES SENSE!!** 72% of spotify listeners are millennials*

The Sample's MOST Popular Spotify Song?



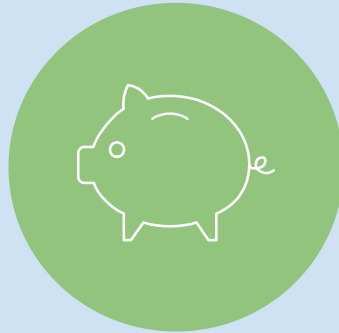
Ariana Grande's Pop
Song:

"Thank u, next"

Implications



Spotify can use data to create playlists based on certain qualities



Important implications in the music industry



Help artists determine how to create popular songs



THANK YOU

go forth and make a popular song!