Are Shorter Songs More Popular!?

DSO 545 Final Project

Michelle (Wen) Lee, Rania Soetirto, Bella Kuo, Denisse Wohlstein, Fan(Frances) Zuo

2020

v

197 seconds or 3.2 minutes

Trends in Music Production

1930 to 1950: steady increase in average song length from 195 seconds to 259 seconds

1990: oversaw longest song averages

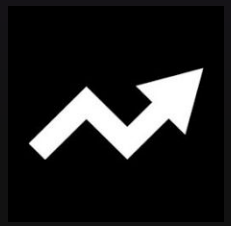
2020: steady decrease in average song length to 197 seconds

How do Music Streaming Services Generate Revenue?

Spotify Premium was worth an average of about **0.68 of a cent** in royalties per streams in 2015 and continues to grow.

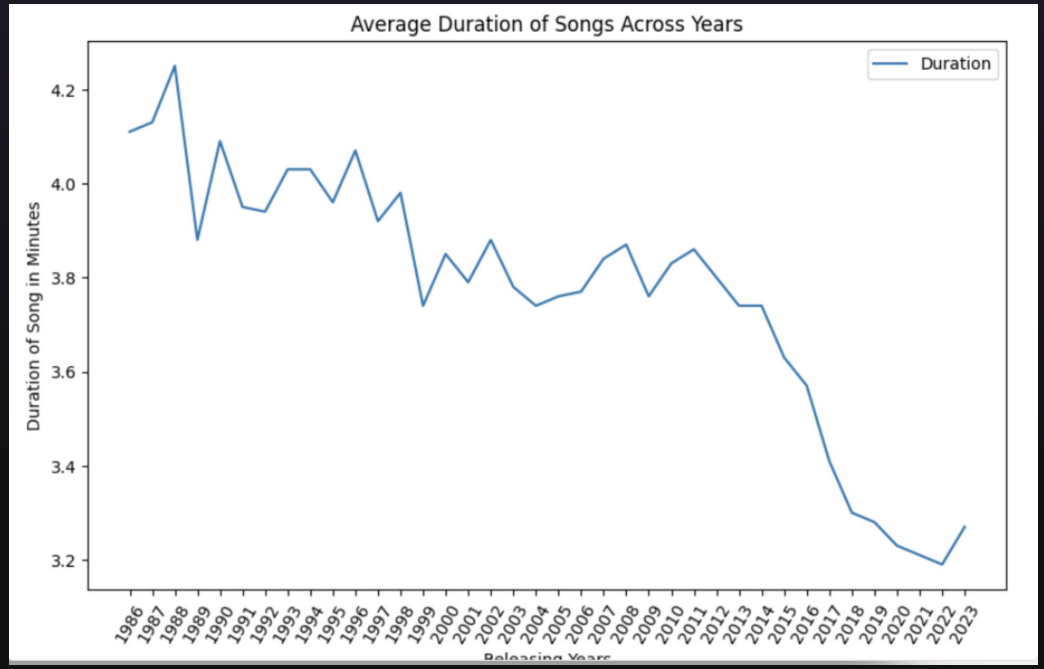
Why are Songs Getting Shorter?

Psychological:
25% percent of listeners will push the skip button in the first five seconds



Market Demand:
To accommodate for shorter-duration content in social media and emerging trends

Historical Perspective



Introduction

Definition

The dataset has 22 numerical variables related to characteristics and categorical variables for artists, genre, and album type.

Popularity (Popularity score)

Considers factors such as total plays, recent frequency, and overall popularity, thus offering a comprehensive measure of a song's popularity across the entire streaming platform.

Duration_min

Duration in this data set are denoted in minutes.

kaggle

Hypothesis

Null Hypothesis (H0):

- The duration of a song does not exert a statistically significant impact on the popularity of a song.

Alternative Hypothesis (Ha):

- Alternative Hypothesis (Ha): There exists a significant statistical correlation between the duration of a song and its popularity.

Agenda

Methods

Methods

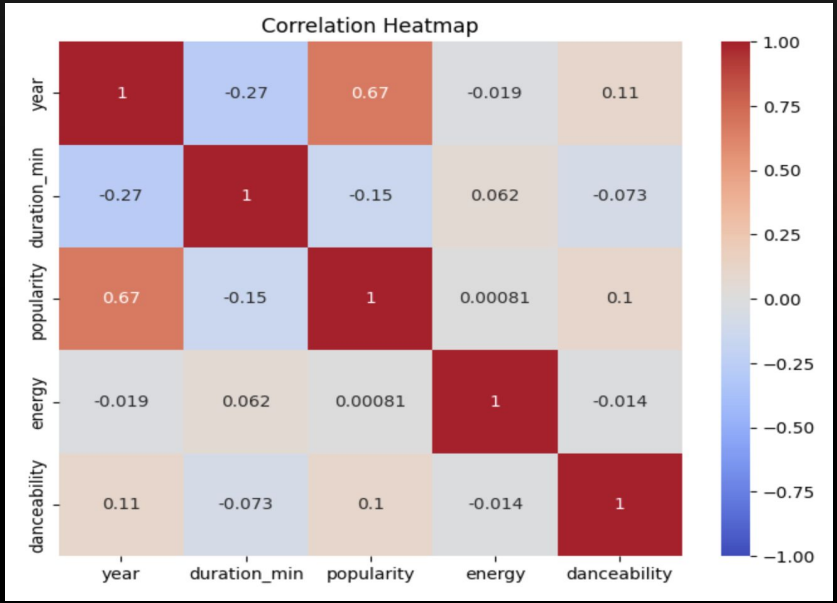
1 Correlation Analysis
2 Simple Regression
3 PCA Analysis
4 Multiple regression
5 Prediction

Descriptive & Data Cleaning

1. This dataset has 22 numerical variables and 2 of them are replicated (duration_ms and duration_min)
2. Remove outliers based on the duration_min
    a. Keeping the dataset in between ±3 standard deviations
3. Drop null / NA after we cap out outliers
4. Remove unnecessary categorical variables, such as hyperlinks

	popularity	year	duration_min	energy	danceability
count	11324.00	11324.00	11324.00	11324.00	11324.00
mean	68.05	2004.31	3.77	0.65	0.61
std	9.32	11.02	0.92	0.21	0.16
min	45.00	1986.00	0.60	0.00	0.00
25%	61.00	1995.00	3.19	0.52	0.51
50%	68.00	2004.00	3.71	0.67	0.62
75%	75.00	2014.00	4.30	0.82	0.73
max	100.00	2023.00	7.02	1.00	0.99

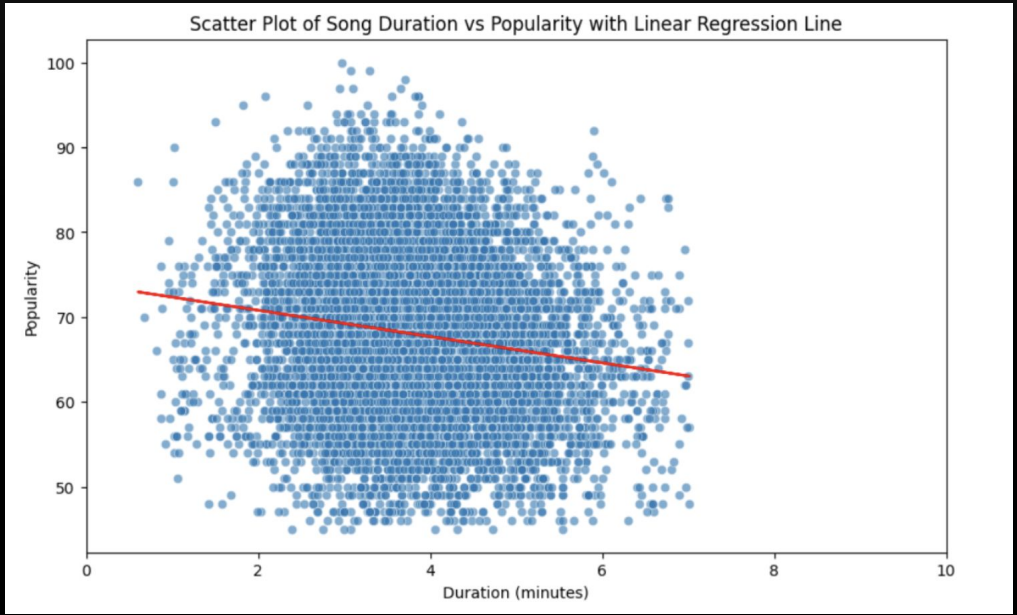
Correlation Analysis



- Negative correlation (-0.15) between popularity and duration
- Duration tends to decrease over time, indicated by a negative correlation (-0.27) by year

Simple Regression

- The trend is not strong, implying other factors are likely influencing popularity,
- There's still a negative correlation between duration and and popularity.



Simple Regression

OLS Regression Results

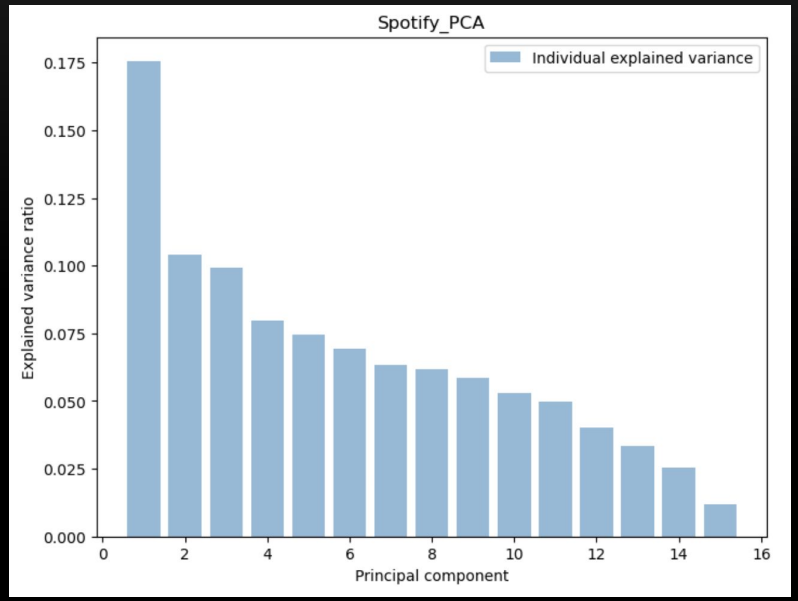
Dep. Variable:	popularity	R-squared:	0.023
Model:	OLS	Adj. R-squared:	0.023
Method:	Least Squares	F-statistic:	270.5
Date:	Mon, 27 Nov 2023	Prob (F-statistic):	4.49e-60
Time:	23:43:55	Log-Likelihood:	-41212.
No. Observations:	11324	AIC:	8.243e+04
Df Residuals:	11322	BIC:	8.244e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	73.8966	0.366	201.965	0.000	73.179	74.614
duration_min	-1.5496	0.094	-16.445	0.000	-1.734	-1.365

Omnibus:	143.800	Durbin-Watson:	0.537
Prob(Omnibus):	0.000	Jarque-Bera (JB):	97.274
Skew:	0.101	Prob(JB):	7.54e-22
Kurtosis:	2.594	Cond. No.	17.4

- There's a statistically significant negative coefficient for song duration.
- Each additional minute of a song's duration is associated with a 1.549 point decrease in popularity score.
- The t-statistic for duration_min is highly significant ($p < 0.001$), reinforcing the robustness of the negative relationship.

PCA Analysis




```
array([0.176, 0.104, 0.099, 0.08 , 0.074, 0.069, 0.063, 0.062, 0.058,
       0.053, 0.05 , 0.04 , 0.034, 0.025, 0.012])
```

PCA Analysis

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
popularity	0.0	-0.0036	0.9950	0.0922	0.0292	0.0220	-0.0042	-0.0028
duration_min	0.0	0.0012	-0.0211	-0.0104	-0.0048	0.9975	-0.0323	0.0117
danceability	0.0	0.0009	0.0018	-0.0049	-0.0008	-0.0167	0.0379	-0.0928
energy	-0.0	-0.0011	0.0006	-0.0388	-0.0112	0.0075	0.0261	-0.0753
instrumentalness	-0.0	0.0002	-0.0012	0.0130	0.0047	0.0022	0.0063	0.0168
key	-0.0	-0.0005	-0.0003	-0.2998	0.9539	0.0010	-0.0172	-0.0004
liveness	-0.0	-0.0000	-0.0003	-0.0043	-0.0017	-0.0032	0.0083	-0.0027
loudness	0.0	-0.0154	0.0971	-0.9476	-0.2980	-0.0103	-0.0072	0.0218
mode	-0.0	-0.0001	-0.0042	0.0083	-0.0153	-0.0315	-0.9954	-0.0611
speechiness	0.0	-0.0001	0.0004	-0.0020	0.0000	-0.0024	0.0272	-0.0182
tempo	-0.0	-0.9999	-0.0051	0.0145	0.0041	0.0013	0.0001	-0.0002
time_signature	0.0	0.0001	-0.0006	-0.0120	-0.0038	0.0157	0.0472	-0.9747
valence	-0.0	-0.0004	-0.0022	-0.0141	-0.0018	-0.0427	0.0190	-0.1177
acousticness	-0.0	0.0009	0.0000	0.0343	0.0091	-0.0320	-0.0476	0.1296
principal_artist_followers	1.0	-0.0000	-0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000

Multiple Regression

OLS Regression Results

Dep. Variable:	popularity	R-squared:	0.181
Model:	OLS	Adj. R-squared:	0.180
Method:	Least Squares	F-statistic:	356.7
Date:	Tue, 28 Nov 2023	Prob (F-statistic):	0.00
Time:	22:59:43	Log-Likelihood:	-40164.
No. Observations:	11311	AIC:	8.034e+04
Df Residuals:	11303	BIC:	8.040e+04
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	81.7882	1.060	77.129	0.000	79.710	83.867
duration_min	-1.6504	0.080	-20.641	0.000	-1.807	-1.494
principal_artist_followers	1.603e-07	4.38e-09	36.625	0.000	1.52e-07	1.69e-07
loudness	0.4581	0.022	21.183	0.000	0.416	0.501
tempo	-0.0057	0.003	-2.137	0.033	-0.011	-0.000
key	-0.0371	0.023	-1.643	0.100	-0.081	0.007
mode	-1.5015	0.172	-8.739	0.000	-1.838	-1.165
time_signature	-0.9402	0.229	-4.099	0.000	-1.390	-0.491

Omnibus:	83.732	Durbin-Watson:	0.803
Prob(Omnibus):	0.000	Jarque-Bera (JB):	67.399
Skew:	0.115	Prob(JB):	2.31e-15
Kurtosis:	2.700	Cond. No.	2.83e+08

Characteristics of a Song with High Popularity:

- Shorter duration still yields a higher popularity score compared to songs with a longer duration
- Louder-sounding music also have a higher popularity

Prediction

P-value of the F-statistic: 0.0000000000000000000

OLS Regression Results

Dep. Variable:	duration_min	R-squared:	0.071
Model:	OLS	Adj. R-squared:	0.071
Method:	Least Squares	F-statistic:	870.4
Date:	Mon, 27 Nov 2023	Prob (F-statistic):	2.25e-184
Time:	23:43:55	Log-Likelihood:	-14688.
No. Observations:	11324	AIC:	2.938e+04
Df Residuals:	11322	BIC:	2.940e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	48.4101	1.513	31.996	0.000	45.444	51.376
year	-0.0223	0.001	-29.503	0.000	-0.024	-0.021

Omnibus:	218.157	Durbin-Watson:	1.957
Prob(Omnibus):	0.000	Jarque-Bera (JB):	345.446
Skew:	0.192	Prob(JB):	9.71e-76
Kurtosis:	3.765	Cond. No.	3.64e+05

Year	Prediction
2030	3.14
2040	2.92
2050	2.70

- Negative relationship between year and duration_min
- As year increase by 1 the duration_min will decrease by 0.02 min
- P-value < 0.05 : reject the null ; R-Squared: Not statistically fit

Conclusion / Discussion

Our predictive analysis shows a weak fit in regression models for popularity and duration:

- Interestingly, duration might not significantly impact a song's overall popularity.

However, our descriptive analysis highlights an ongoing trend:

- the average song duration in 2023 is at a record low for the past two decades.

We encourage industry professionals to explore alternative variables to boost music streams and revenue.

Thank You

Trojan Fight On

USC MSBA Class of 2024

0:23

-3:25

References

Gangiredla, D., Moghaddas, D., Soman, O., Nguyen, T., & Jamal, Z. (2020, December 22). *Spotify Trends Analysis.* DataRes at UCLA. Retrieved November 21, 2023, from https://ucladatares.medium.com/spotify-trends-analysis-129c8a31cf04

Maria Eriksson, Rasmus Fleischer, Anna Johansson, Pelle Snickars, & Patrick Vonderau. (2019). *Spotify Teardown : Inside the Black Box of Streaming Music*. The MIT Press.

Wright, D. (2023, April 13). Pop Songs Really Are Shorter Than Ever Now. *VICE*. https://www.vice.com/en/article/qjv8pq/pop-songs-shorter-than-ever#

Zillman, C., & Eshelman, C. (2015, June 4). Spotify: How much do artists really earn? *Fortune*. https://fortune.com/2015/06/04/heres-how-much-artists-really-make-on-spotify/

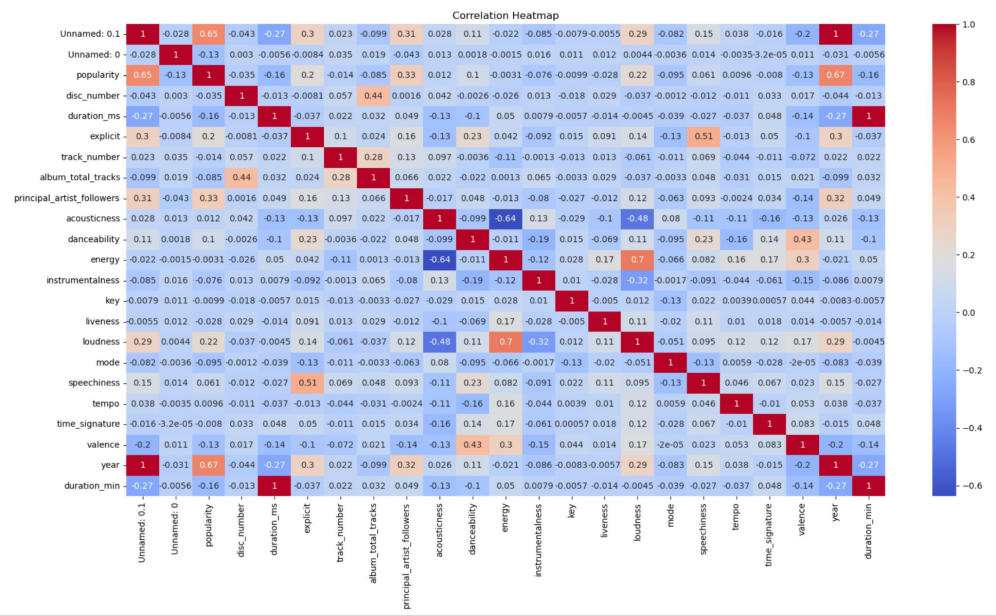
Appendix

Despacito

Luis Fonsi, Daddy Yankee

0:23 -3:25

Correlation Analysis



- Negative correlation (-0.15) between popularity and duration
- Duration tends to decrease over time, indicated by a negative correlation (-0.27) by year.