By: Sarah McCoy, 11671063

Final Project:
Spotify Tracks

The Objective:

This project's purpose is to apply what we learned from the course and to practice our data communication skills. My goal is to tell a story through data and statistics.

Dataset Description:

The dataset I will use to conduct my analysis was taken from kaggle.com.The original dataset has a range of 125 different genres and their audio features that were popular between the years of 2022-1970. For this project, I randomly selected 149 tracks to test the relationship between popularity and music features.

Chosen Variables:

- **Popularity:** The popularity of a track is the dependent variable being measured and tested in this study. The popularity of a track is a value between 0 and 100, with 100 being the most popular.
- **Duration**: is the track length in milliseconds
- **Danceability**: describes the track's appropriateness for dancing based on rhythm and beat. A value of 0 is the least danceable and 1.0 is the most danceable.
- **Loudness**: is the overall intensity of sound in a track measured in decibels (dB)
- **Energy**: is a measure from 0.0 to 1.0 and represents a measure of activity. An energetic track would give the listener a feeling of the track moving fast and producing an energetic feeling like a rock song for example.
- **Tempo:** The estimated tempo is the track beats per minute (BPM).

Introduction:

Spotify has mentioned the popularity of songs on Spotify is based on the frequent number of recent listens (*How we generate popular tracks*). However, the ranking of songs doesn't provide insight to the artist or the listener as to why people enjoy listening to the song. What if there is a formula that causes a song to increase in popularity? What if there are specific music attributes that affect a song's popularity? As a data analyst at City Records music label, your job is to discover and propose questions about what makes listeners hum along to a song, play it repeatedly, and whether there is a formula to produce a popular song.

Module 1:

The sample data of 149 Spotify songs are normally distributed. Find the probability of selecting a song that scored above 80 in popularity. Round to the nearest three decimal places.

x	80
SD	21.756
$\bar{x}$	48.987
z	1.425
Probability Above	.077
Probability Below	.923

n = 149

SD = 21.756

$z = (x - \bar{x}) / SD$

$z = 1.425$

The probability of selecting a song that scored above 80 is .077.

n	149

Module 2:

SD	21.756
Za/2	1.96
E	3.493

City Records music label requires you to report what score a Spotify track would need to make it into the top 40 percentile. Additionally, conduct a 95% confidence interval for popularity and calculate if we have the necessary sample size to ensure we have an accurate representation of the population.

E = margin of error = $Za/2\ \sigma/\sqrt{n}$ = 3.493

n = ($Za/2$ * σ / E)^2 = 149

Score needed to make it into the Top 40	95% Confidence Interval for popularity	Necessary Sample Size
54.499	51.51 to 57.99	n= 149

Module 3:

At City Records, we believe popular songs that rank above 54.499 are short, catchy, and high-energy songs that would be played at a party. Is there sufficient evidence to support our claim at a significance level of .05. Calculate the test statistic and p-values, round the answer to three decimal places, and clarify the decision ($\mu1$ represents songs above 54.499, $\mu2$ represents all songs).

Ho: $\mu1 = \mu2$

1. Popular songs have a shorter duration than the average songs (Ha: $\mu1 < \mu2$).
2. Popular songs have more danceability than average songs (Ha: $\mu1 > \mu2$).
3. Popular songs are louder than average songs (Ha: $\mu1 > \mu2$).
4. Popular songs have more energy than average songs (Ha: $\mu1 > \mu2$).
5. Popular songs have more tempo than average songs (Ha: $\mu1 > \mu2$).

	1. Duration	2. Danceability	3. Loudness	4. Energy	5. Tempo
Test statistic t	.313	.392	.629	-1.155	-.853
P-Value	.378	.348	.266	.126	.198

I performed a one-sample t-test to determine if there was a statistically significant difference in means. I found the means to not have a significant difference. All p-values were above .05. We failed to reject the null hypotheses. There is insignificant evidence to support the idea that all popular songs are short, high-energy party songs.

Module 4:

City Records is interested in more inferences from the Spotify dataset. Using the ANOVA procedure, find if there is any difference between the means of the song attributes. Explain your analysis.

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
popularity	149	7299	48.9865772	473.337656		
duration_ms	149	33774848	226676.832	4710769748		
danceability	149	88.471	0.5937651	0.02429001		
energy	149	91.2267	0.61225973	0.05328626		
loudness	149	-1207.512	-8.1041074	16.9535481		
tempo	149	17626.839	118.300933	813.847587		
ANOVA						
Source of Variati	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Grc	6.3782E+12	5	1.2756E+12	1624.74695	0	2.22418428
Within Grour	6.9719E+11	888	785128509			
Total	7.0754E+12	893				

After using the ANOVA procedure, I received a p-value of 0. This indicates there is at least one group that differs significantly. These results might seem contradictory to the t-test performed earlier upon request. However, ANOVA tests are typically better at detecting differences than performing individual t-tests. Additionally, there is a chance we encountered an error because we did not adjust the significance level for each test.

<u>Module 5:</u>

City Records desires to predict the popularity of a Spotify song based on song attributes. Provide a regression equation for predicting song popularity and interpret your findings.

EXCEL MODEL 1:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.29895673
R Square	0.08937513
Adjusted R S	0.0575351
Standard Err	21.1211773
Observations	149

ANOVA

	df	SS	MS	F	Significance F
Regression	5	6261.08276	1252.21655	2.80700507	0.01890994
Residual	143	63792.8904	446.104129		
Total	148	70053.9732			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	74.1774449	17.9147133	4.14058788	5.8948E-05	38.7655709	109.589319	38.7655709	109.589319
duration_ms	2.669E-05	2.8056E-05	0.95131911	0.34304755	-2.877E-05	8.2148E-05	-2.877E-05	8.2148E-05
danceability	13.5373983	13.2793436	1.01943279	0.30971914	-12.711776	39.7865732	-12.711776	39.7865732
energy	-28.057655	10.0580061	-2.7895842	0.00599736	-47.939237	-8.1760728	-47.939237	-8.1760728
loudness	1.43891115	0.60062106	2.39570544	0.01788231	0.25166818	2.62615412	0.25166818	2.62615412
tempo	-0.0882431	0.0635449	-1.3886733	0.16709127	-0.2138518	0.0373656	-0.2138518	0.0373656

T-stat = -1.977

Rejection Region = $t < -1.977$ or $t > 1.977$

We can conclude from the first regression model that energy and loudness have statistically significant effects on the dependent variable popularity. However, outliers have a significant impact on the R-squared and adjusted R-square which could indicate why the R-square is low and the model does not explain the relationships between variables as effectively.

In addition, it's essential to remove insignificant variables to improve the model's performance. I used the SAS program to select the best regression model for predicting song popularity and to provide City Records with a more in-depth analysis (Appendix A).

More specifically, I used stepwise and backward stepwise selection which are variable selection procedures (Appendix B, figures 1-4). The procedure consists of testing different combinations of variables that improve the adjusted R-square while limiting RMSE. Using these tools I was able to simplify the regression model and verify accuracy.

To conclude, Model 3 performed the best with an adjusted R-square of .0613, R-square of .0803, and RMSE of 21.0795 (Appendix B, figure 5). Based on statistical evidence, there is a relationship between popularity and the independent variables energy, loudness, and tempo.

Regression equation =

$$\hat{y} = 89.7795999 - 28.142122(\text{energy}) + 1.66243768(\text{loudness}) - .0852923(\text{tempo})$$

APPENDIX A:

SAS Code

```
data music;
input y x1 x2 x3 x4 x5;
datalines;
y       x1      x2      x3      x4      x5
61      230213  0.485 0.723 -6.565      167.845
54      206546  0.397 0.538 -6.503      154.342
0       181872  0.627 0.562 -6.268      95.054
80      242973  0.603 0.67  -3.882      107.993
85      234093  0.913 0.603 -4.892      123.061
61      245978  0.408 0.314 -7.729      105.096
64      223794  0.79  0.728 -6.523      110.059
54      176616  0.675 0.931 -3.432      124.008
72      216480  0.484 0.368 -7.784      92.923
88      242013  0.561 0.431 -8.81 143.875
82      220779  0.559 0.559 -6.425      72.498
84      233720  0.609 0.378 -9.828      100.418
0       269186  0.629 0.664 -10.517     139.975
0       153190  0.604 0.741 -5.571      153.947
82      179172  0.918 0.585 -7.781      131.966
84      204939  0.724 0.818 -3.747      77.004
85      231826  0.76  0.703 -5.412      95.997
3       201573  0.512 0.796 -4.075      171.014
89      174000  0.548 0.816 -4.209      95.39
88      238805  0.761 0.525 -6.9  80.87
84      203807  0.695 0.884 -2.278      103.014
58      225310  0.638 0.731 -4.537      89.928
59      351425  0.365 0.449 -13.516     103.643
58      328854  0.314 0.478 -7.201      199.631
58      241304  0.66  0.926 -3.452      109.997
59      377000  0.798 0.636 -10.52      139.98
59      184370  0.461 0.5   -9.828      76.176
61      236057  0.728 0.861 -5.852      87.467
62      278546  0.522 0.74  -7.018      92.618
69      118250  0.484 0.059 -25.084     89.329
78      377093  0.414 0.428 -11.097     145.075
70      123500  0.307 0.0381      -28.197  72.453
51      305530  0.347 0.799 -6.371      162.575
52      159066  0.632 0.754 -8.207      89.964
36      120413  0.628 0.52  -8.331      89.97
50      244120  0.51  0.42  -8.014      125.898
30      401773  0.423 0.617 -6.458      142.042
44      258482  0.513 0.81  -6.826      121.015
24      350000  0.205 0.181 -18.639     72.184
48      261229  0.523 0.535 -11.394     159.903
55      336932  0.254 0.527 -9.225      147.561
42      242840  0.451 0.952 -3.186      95.985
24      118000  0.254 0.996 -21.444     107.921
42      209920  0.571 0.853 -6.589      126.016
84      205946  0.649 0.716 -5.371      99.988
79      192024  0.886 0.672 -4.394      91.976
30      268736  0.628 0.027 -13.64      122.007
```

46	205658	0.49	0.674	-7.276	122.807
60	256680	0.316	0.413	-12.216	155.69
51	265386	0.47	0.271	-7.636	145.902
73	230666	0.676	0.461	-6.746	87.917
55	149610	0.42	0.166	-17.235	77.489
57	210826	0.438	0.359	-9.734	76.332
71	201933	0.266	0.0596	-18.515	181.74
82	198853	0.618	0.443	-9.681	119.949
58	214240	0.688	0.481	-8.807	98.017
74	229400	0.407	0.147	-8.822	141.284
80	242946	0.703	0.444	-9.331	150.96
56	205594	0.442	0.632	-6.77	78.899
74	244800	0.627	0.363	-8.127	99.905
69	240165	0.483	0.303	-10.058	133.406
52	198712	0.489	0.314	-9.245	124.234
62	248448	0.691	0.234	-6.441	87.103
56	188133	0.755	0.78	-6.084	120.004
58	244986	0.489	0.561	-7.933	83.457
56	129750	0.706	0.112	-18.098	110.154
77	180750	0.658	0.9	-3.479	156.096
74	198173	0.749	0.862	-3.494	97.982
53	272639	0.462	0.217	-11.436	120.971
79	173947	0.902	0.582	-5.902	107.005
68	192280	0.692	0.881	-2.563	103.053
64	189546	0.683	0.947	-1.717	100.024
66	537653	0.372	0.426	-8.421	123.993
76	227478	0.904	0.723	-5.224	145.013
64	203056	0.765	0.77	-3.16	103.005
52	150139	0.495	0.484	-10.923	126.632
50	213571	0.659	0.633	-5.343	125.939
54	183100	0.852	0.967	-1.661	124.024
59	149213	0.85	0.734	-6.415	103.027
69	225012	0.72	0.824	-5.086	165.996
51	266959	0.667	0.545	-8.088	147.988
53	341177	0.403	0.56	-7.46	155.701
74	221573	0.82	0.852	-2.567	109.036
52	255186	0.664	0.637	-4.338	93.137
52	261000	0.591	0.828	-10.309	89.981
67	178529	0.878	0.777	-3.702	139.989
52	250124	0.59	0.908	-4.137	87.989
67	139213	0.852	0.762	-4.448	131.958
63	185508	0.519	0.352	-8.332	78.943
40	164571	0.809	0.464	-8.716	140.006
38	238064	0.667	0.645	-8.727	123.994
39	200145	0.541	0.84	-10.416	137.04
39	162384	0.734	0.551	-10.879	144.058
38	259339	0.388	0.498	-9.165	84.567
56	220506	0.857	0.628	-5.58	101.987
39	214256	0.537	0.417	-11.937	73.916
39	203600	0.404	0.448	-10.731	98.503
40	281566	0.389	0.516	-9.313	165.68
39	214200	0.716	0.671	-9.517	100.021
39	247720	0.495	0.969	-3.376	179.729
39	120122	0.811	0.731	-4.085	100.017
40	140700	0.727	0.423	-12.411	100.038

```
40          192000  0.78  0.721 -7.042              79.996
40          194567  0.772 0.652 -8.125              127.027
39          232041  0.5   0.912 -5.725              139.865
39          229200  0.707 0.885 -5.133              120.032
40          480000  0.717 0.898 -7.52 137.024
49          193596  0.541 0.839 -7.141              123.984
29          358669  0.632 0.983 -6.276              129.995
21          204687  0.433 0.816 -5.507              131.287
30          279000  0.62  0.585 -7.233              124.137
53          162580  0.709 0.74  -5.783              124.144
21          354666  0.563 0.879 -7.16 180.009
47          147230  0.393 0.94  -3.894              129.958
36          182553  0.658 0.857 -5.625              140.985
49          236712  0.519 0.874 -4.219              145.944
3           165249  0.684 0.429 -10.107             96.016
27          141665  0.697 0.535 -6.314              91.905
26          186635  0.641 0.4   -6.74 84.981
10          143004  0.774 0.657 -6.587              96.978
8           113454  0.74  0.503 -9.669              82.473
25          144494  0.756 0.4   -8.383              88.983
31          134579  0.742 0.254 -15.696             150.029
8           155581  0.648 0.439 -10.652             172.073
10          144381  0.703 0.56  -6.529              95.092
51          191293  0.462 0.722 -5.134              173.941
34          295333  0.428 0.312 -14.24              126.75
32          230840  0.558 0.492 -10.173             111.488
55          284493  0.283 0.657 -8.557              97.991
23          208828  0.54  0.77  -5.597              133.942
50          231853  0.641 0.757 -4.867              87.995
63          332506  0.46  0.162 -17.142             122.767
23          341200  0.565 0.772 -8.436              89.984
25          270706  0.573 0.908 -4.369              117.107
61          205306  0.568 0.849 -10.472             130.746
24          195626  0.7   0.496 -8.64 100.058
23          222738  0.583 0.643 -6.531              120.116
23          428266  0.524 0.951 -6.851              108.1
23          321200  0.478 0.471 -7.923              204.643
23          186826  0.675 0.452 -11.492             127.901
23          212573  0.786 0.517 -14.507             136.923
25          219866  0.622 0.714 -11.152             159.948
25          245706  0.373 0.849 -6.352              114.792
24          184466  0.496 0.507 -8.384              101.117
23          230272  0.576 0.672 -9.332              112.799
25          264773  0.515 0.926 -8.01 136.62
55          244306  0.53  0.873 -10.572             126.16
0           266383  0.505 0.87  -9.392              112.102
52          184866  0.678 0.375 -6.992              89.994
run;
proc reg data = work.music;
    model y = x1 x2 x3 x4 x5 /cli clm ;
run;
proc reg data = work.music;
    model y = x1 x2 x3 x4 x5 /vif r ;
run;
```

```
proc reg data = work.music;
     model y = x1 x2 x3 x4 x5 / selection=stepwise
slentry=0.25 slstay=0.15;
run;

proc reg data = work.music;
     model y = x1 x2 x3 x4 x5 / SELECTION=BACKWARD
slstay=0.15;
run;
 proc reg data = work.music;
     model y = x1 x2 x3 x4 x5 /    selection = adjrsq Best = 13
CP RMSE;
run;
proc reg data = work.music;
     model y = x1 x2 x3 x4 x5 /    selection = rsquare Best = 2
adjrsq CP RMSE;
run;
```

APPENDIX B:

The REG Procedure
Model: MODEL1
Dependent Variable: y

Number of Observations Read	150
Number of Observations Used	149
Number of Observations with Missing Values	1

Stepwise Selection: Step 1

Variable x5 Entered: R-Square = 0.0164 and C(p) = 9.4654

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1146.33396	1146.33396	2.45	0.1200
Error	147	68908	468.75945		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.52752	7.59023	29809	63.59	<.0001
x5	-0.09756	0.06238	1146.33396	2.45	0.1200

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable x4 Entered: R-Square = 0.0304 and C(p) = 9.2586

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2130.76983	1065.38491	2.29	0.1049
Error	146	67923	465.22742		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	66.13714	8.48815	28244	60.71	<.0001
x4	0.62713	0.43112	984.43586	2.12	0.1479
x5	-0.10201	0.06222	1250.43151	2.69	0.1033

Bounds on condition number: 1.0024, 4.0097

Figure 1: Stepwise Selection 1 and 2

Stepwise Selection: Step 3

Variable x3 Entered: R-Square = 0.0803 and C(p) = 3.4284

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5623.87491	1874.62497	4.22	0.0068
Error	145	64430	444.34551		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	89.77960	11.82873	25598	57.61	<.0001
x3	-28.14212	10.03717	3493.10508	7.86	0.0057
x4	1.66244	0.56024	3912.57660	8.81	0.0035
x5	-0.08529	0.06110	865.79231	1.95	0.1649

Bounds on condition number: 1.788, 13.717

Stepwise Selection: Step 4

Variable x5 Removed: R-Square = 0.0679 and C(p) = 3.3692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	4758.08260	2379.04130	5.32	0.0059
Error	146	65296	447.23213		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	80.70075	9.91212	29645	66.29	<.0001
x3	-29.50953	10.02165	3877.74429	8.67	0.0038
x4	1.68392	0.56184	4017.38407	8.98	0.0032

Bounds on condition number: 1.771, 7.0841

All variables left in the model are significant at the 0.1500 level.

The stepwise method terminated because the next variable to be entered was just removed.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x5		1	0.0164	0.0164	9.4654	2.45	0.1200
2	x4		2	0.0141	0.0304	9.2586	2.12	0.1479
3	x3		3	0.0499	0.0803	3.4284	7.86	0.0057
4		x5	2	0.0124	0.0679	3.3692	1.95	0.1649

Figure 2: Stepwise Selection 3 and 4

The REG Procedure
Model: MODEL1
Dependent Variable: y

Number of Observations Read	150
Number of Observations Used	149
Number of Observations with Missing Values	1

Backward Elimination: Step 0

All Variables Entered: R-Square = 0.0894 and C(p) = 6.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6261.08276	1252.21655	2.81	0.0189
Error	143	63793	446.10413		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	74.17744	17.91471	7648.21797	17.14	<.0001
x1	0.00002669	0.00002806	403.72783	0.91	0.3430
x2	13.53740	13.27934	463.61069	1.04	0.3097
x3	-28.05765	10.05801	3471.48421	7.78	0.0060
x4	1.43891	0.60062	2560.37207	5.74	0.0179
x5	-0.08824	0.06354	860.27324	1.93	0.1671

Bounds on condition number: 2.029, 37.794

Backward Elimination: Step 1

Variable x1 Removed: R-Square = 0.0836 and C(p) = 4.9050

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	5857.35493	1464.33873	3.28	0.0131
Error	144	64197	445.80985		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	82.23789	15.77918	12109	27.16	<.0001
x2	8.95158	12.36944	233.48002	0.52	0.4704
x3	-28.04084	10.05467	3467.33446	7.78	0.0060
x4	1.52109	0.59418	2921.63805	6.55	0.0115
x5	-0.07668	0.06235	674.23065	1.51	0.2208

Bounds on condition number: 1.987, 24.238

Figure 3: Backward Elimination 0 and 1

Backward Elimination: Step 2

Variable x2 Removed: R-Square = 0.0803 and C(p) = 3.4284

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5623.87491	1874.62497	4.22	0.0068
Error	145	64430	444.34551		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	89.77960	11.82873	25598	57.61	<.0001
x3	-28.14212	10.03717	3493.10508	7.86	0.0057
x4	1.66244	0.56024	3912.57660	8.81	0.0035
x5	-0.08529	0.06110	865.79231	1.95	0.1649

Bounds on condition number: 1.788, 13.717

Backward Elimination: Step 3

Variable x5 Removed: R-Square = 0.0679 and C(p) = 3.3692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	4758.08260	2379.04130	5.32	0.0059
Error	146	65296	447.23213		
Corrected Total	148	70054			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	80.70075	9.91212	29645	66.29	<.0001
x3	-29.50953	10.02165	3877.74429	8.67	0.0038
x4	1.68392	0.56184	4017.38407	8.98	0.0032

Bounds on condition number: 1.771, 7.0841

All variables left in the model are significant at the 0.1500 level.

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x1	4	0.0058	0.0836	4.9050	0.91	0.3430
2	x2	3	0.0033	0.0803	3.4284	0.52	0.4704
3	x5	2	0.0124	0.0679	3.3692	1.95	0.1649

Figure 4: Backward Elimination 2 and 3

The REG Procedure
Model: MODEL1
Dependent Variable: y

Adjusted R-Square Selection Method

Number of Observations Read	150
Number of Observations Used	149
Number of Observations with Missing Values	1

Number in Model	Adjusted R-Square	R-Square	C(p)	Root MSE	Variables in Model
3	0.0613	0.0803	3.4284	21.07950	x3 x4 x5
4	0.0582	0.0836	4.9050	21.11421	x2 x3 x4 x5
5	0.0575	0.0894	6.0000	21.12118	x1 x2 x3 x4 x5
4	0.0573	0.0828	5.0392	21.12406	x1 x3 x4 x5
2	0.0552	0.0679	3.3692	21.14786	x3 x4
3	0.0548	0.0740	4.4164	21.15148	x2 x3 x4
4	0.0515	0.0771	5.9284	21.18916	x1 x2 x3 x4
3	0.0491	0.0683	5.3019	21.21579	x1 x3 x4
4	0.0265	0.0528	9.7394	21.46594	x1 x2 x3 x5
3	0.0221	0.0419	9.4542	21.51475	x2 x3 x5
3	0.0211	0.0409	9.6105	21.52592	x1 x2 x3
2	0.0205	0.0337	8.7419	21.53250	x2 x3
2	0.0171	0.0304	9.2586	21.56913	x4 x5

Figure 5: Summary of all Regression Models

Reference

How we generate popular tracks. Spotify. (n.d.).
https://support.spotify.com/us/artists/article/how-we-generate-popular-tracks/