



What Makes a Playlist Successful?

Tatev Aslanyan

Agenda

- Project Goal
- Strategy
- Defining Success Metrics
- Feature Candidates for Successful Playlists
- Data Exploration and Visualization
- Regression Analysis
- Conclusions
- Next Steps



Project Goal (Listen to Playlist More and Longer)

Business Goal: Provide the most successful playlists to Spotify customers to

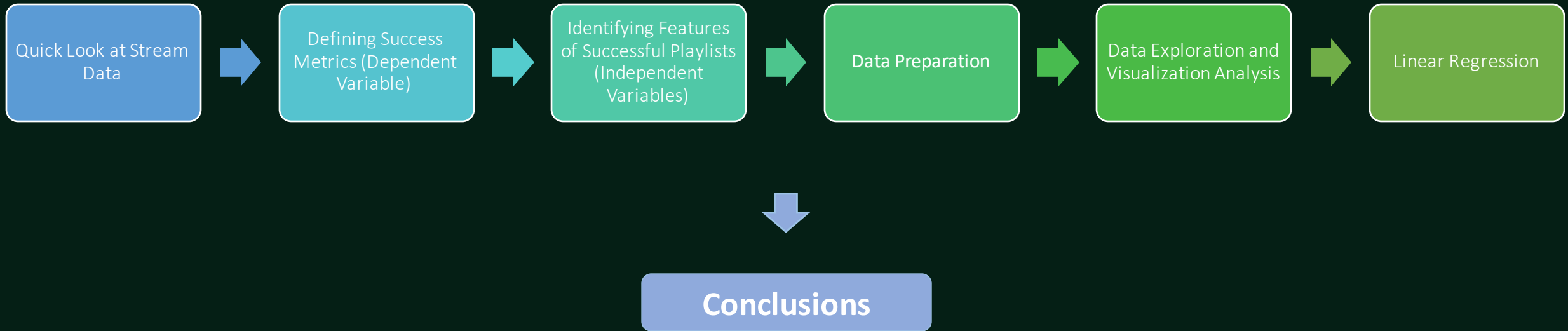
- ↑ usage main goal
- ↑ customer satisfaction indirect result
- ↑ revenue indirect result

Technical Goals: Quantify what makes playlist successful

- How to measure the successfulness of the playlist?
- What are the characteristics of successful playlists?



Strategy



Looking at Steam Data

	playlist_uri	owner	streams	stream30s	...	genre_3	mood_1	mood_2	mood_3
0	spotify:user:36069af6af076ccd9e597184a67b68c9:...	36069af6af076ccd9e597184a67b68c9	27	27	...	Country & Folk	Peaceful	Romantic	Somber
1	spotify:user:d1144a65b1c31c5f9f56b94f831124d5:...	d1144a65b1c31c5f9f56b94f831124d5	0	0	...	Alternative	Excited	Yearning	Defiant
2	spotify:user:6b7fbed9edd6418ddd3b555bba441536:...	6b7fbed9edd6418ddd3b555bba441536	4	2	...	-	Lively	Upbeat	Romantic
3	spotify:user:580b98725077a94c3c8d01d07390426b:...	580b98725077a94c3c8d01d07390426b	12	12	...	Pop	Excited	Aggressive	Defiant
4	spotify:user:1305d39070c95d161cc502e15014897d:...	1305d39070c95d161cc502e15014897d	20	4	...	Electronica	Excited	Defiant	Yearning
...
403361	spotify:user:4672952d42bdd93b9215ce9a40394ea6:...	4672952d42bdd93b9215ce9a40394ea6	18	6	...	Dance & House	Excited	Yearning	Energizing
403362	spotify:user:28c4378e099b4843f5dd42bb848c78ea:...	28c4378e099b4843f5dd42bb848c78ea	0	0	...	Rap	Sensual	Excited	Brooding
403363	spotify:user:1c54302dc7e610a10c51eed81e26a168:...	1c54302dc7e610a10c51eed81e26a168	0	0	...	Alternative	Brooding	Defiant	Sophisticated
403364	spotify:user:adc973443cdf1abecdfb4244e530d451:...	adc973443cdf1abecdfb4244e530d451	0	0	...	Latin	Defiant	Energizing	Aggressive
403365	spotify:user:b3752c94e387192b7950b687453bcf45:...	b3752c94e387192b7950b687453bcf45	74	16	...	Alternative	Defiant	Energizing	Cool

[403366 rows x 25 columns]



Looking at Steam Data

Descriptive Statistics for Some Features

Number of Playlists*	402967
Number of Owners	314898
AVG Number of Tracks of Playlists	202
AVG Number of Albums of Playlists	88
AVG Number of Artists of Playlists	84
AVG Number of Tokens of Playlists	2
AVG Number Streams of Playlists (Today)	21
AVG Number Streams30s per Playlists (Today)	12 (57%)
AVG Number Streams30s per Playlists (This Month)	354

AVG DAU per Playlist	1
MIN DAU per Playlist	0
MAX DAU per Playlist	18290
AVG WAU per Playlist	5
MIN WAU per Playlist	0
MAX WAU per Playlist	71891
AVG MAU per Playlist	15
MIN MAU per Playlist	2
MAX MAU per Playlist	206756

* Excluded 399 Spotify Playlists



Defining Success Metrics

When can we say that the playlist is successful?

- Users listen to the successful playlists a lot
- Users listen to the successful playlist for a longer time (longer than 30 seconds)

What are possible levels for the success?

- Success at playlist level
- Success at the playlist owner level



Defining Success Metrics

Short Term

- **DAUs**

Number of daily active users with streams more than 30 seconds from playlist

- **WAUs**

Number of weekly active users with streams more than 30 seconds from playlist

- **Streams30s**

Number of streams from playlist today with more than 30s

Long Term

- **MAUs**

Number of monthly active users with a stream over 30 seconds from playlist this month. Or total number of active users from the playlist in the last 2 months (mau_both_months)

- **Monthly_streams30s**

Number of streams over 30 seconds by playlist

- **Monthly_owner_streams30s**

Number of streams over 30 seconds by playlist owner this month

Today: WAU

In Future: Combination of DAU, WAU, MAU, Streams30s, Monthly_streams30s, Monthly_owner_streams30s



Candidate Features

Playlist Variety

- Number of Tracks
- Number of Artists
- Number of Albums

Playlist Owner Descriptors

- Number of playlists per owner (proxy for owner effort)
- Owner has more than 1 playlist (proxy for owners activity)

Playlist Categorical Features

- Moods: Leading Mood
- Genres: Leading Genres
- Tokens Sentiment (Sentiment Analysis for tokens)

Playlist Ratios

- Streams30s to Streams Ratio
- Streams30 to Monthly Streams Ratio
- Monthly_owner_stream30s to monthly_stream30s Ratio



Hypothesis

Varsity Matters

Genre, moods, number of tracks, albums, artists

Effort Matters

Length of title

Sentiment Matters

Sentiment Analysis moods, genre, tokens

Engagement Matters

Number of Playlists Owned by User



Data Preparation

Checking for Outliers

- Removing 399 Playlists of Spotify with above than average stream
- Removing the rows from 99th percentile

Checking for Missing Values

- There are no null values in the data but there are missing values, especially in Mood, Genre variables

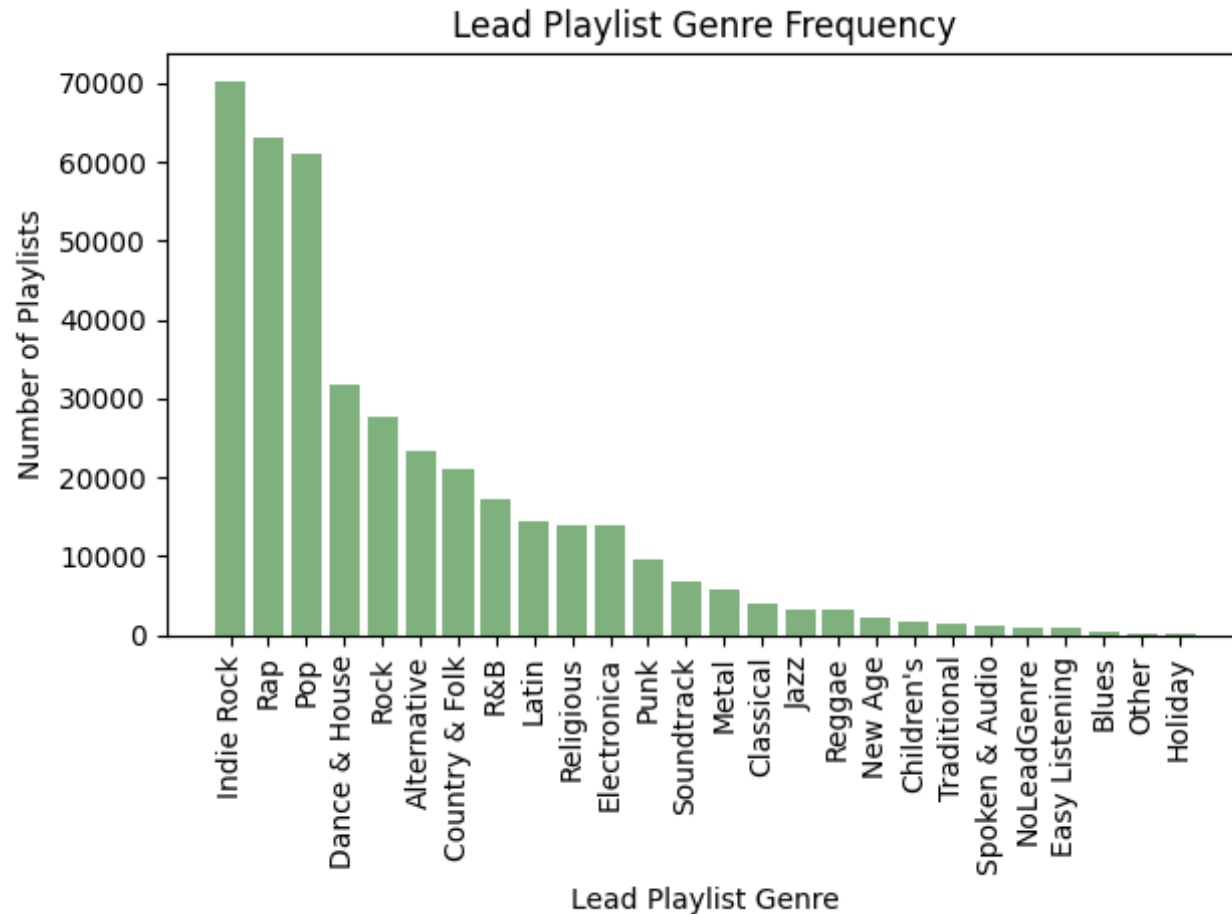
Information Retrieval and Sentiment Analysis for Tokens

- Converting text to list of strings
- Sentiment Score low: Negative vs high: Positive (NLP on tokens)



Data Exploration & Visualization

Playlists Leading Genre



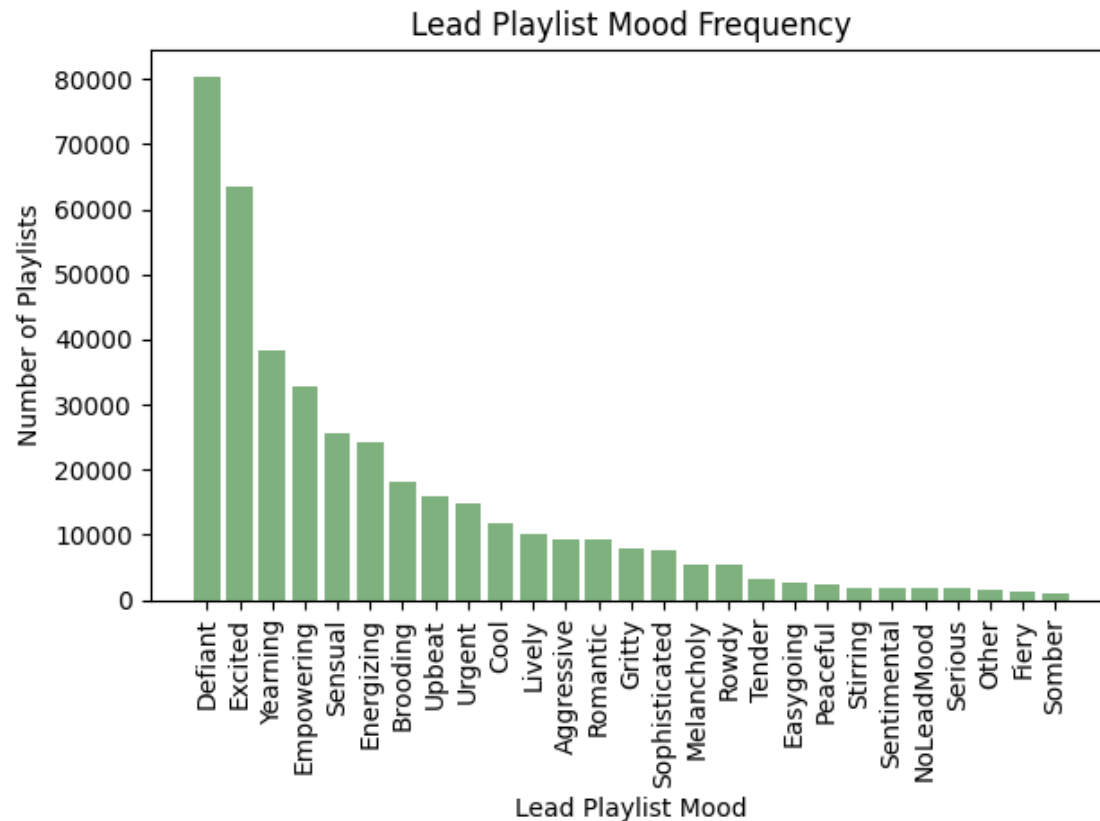
Popular Genres Leading Playlists

- 3 most popular leading genres across all playlists: "Indie Rock", "Rap", "Pop"
- ↓
- Playlists from these 3 genres are more successful?
- ↓
- Binary features from these 3 for genres for future predictive model



Data Exploration & Visualization

Playlists Leading Mood

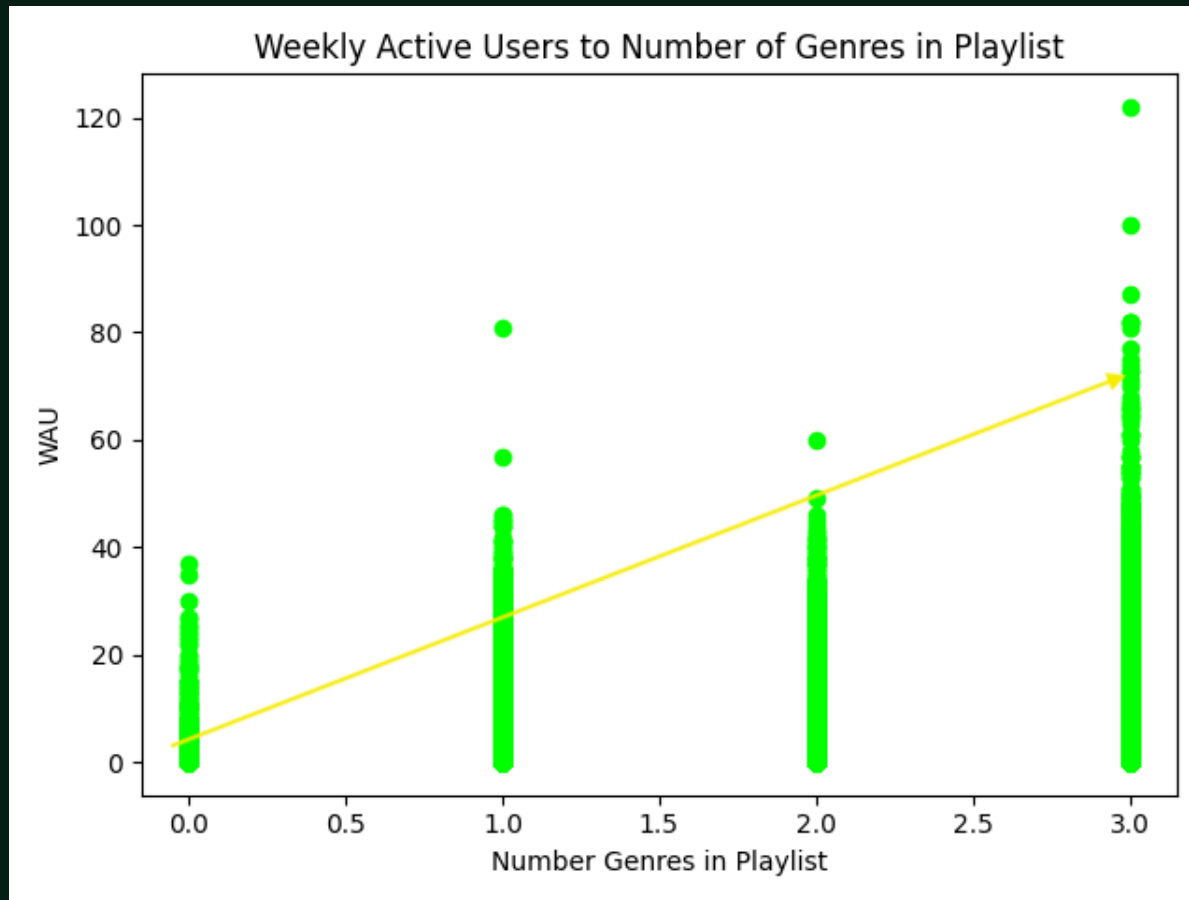


Popular Moods Leading Playlists

- 2 most popular leading moods across all playlists: "Defiant", "Excited"
- ↓
- Playlists with these 2 moods are more successful
- ↓
- Binary features from these 2 moods for future predictive model



Data Exploration & Visualization

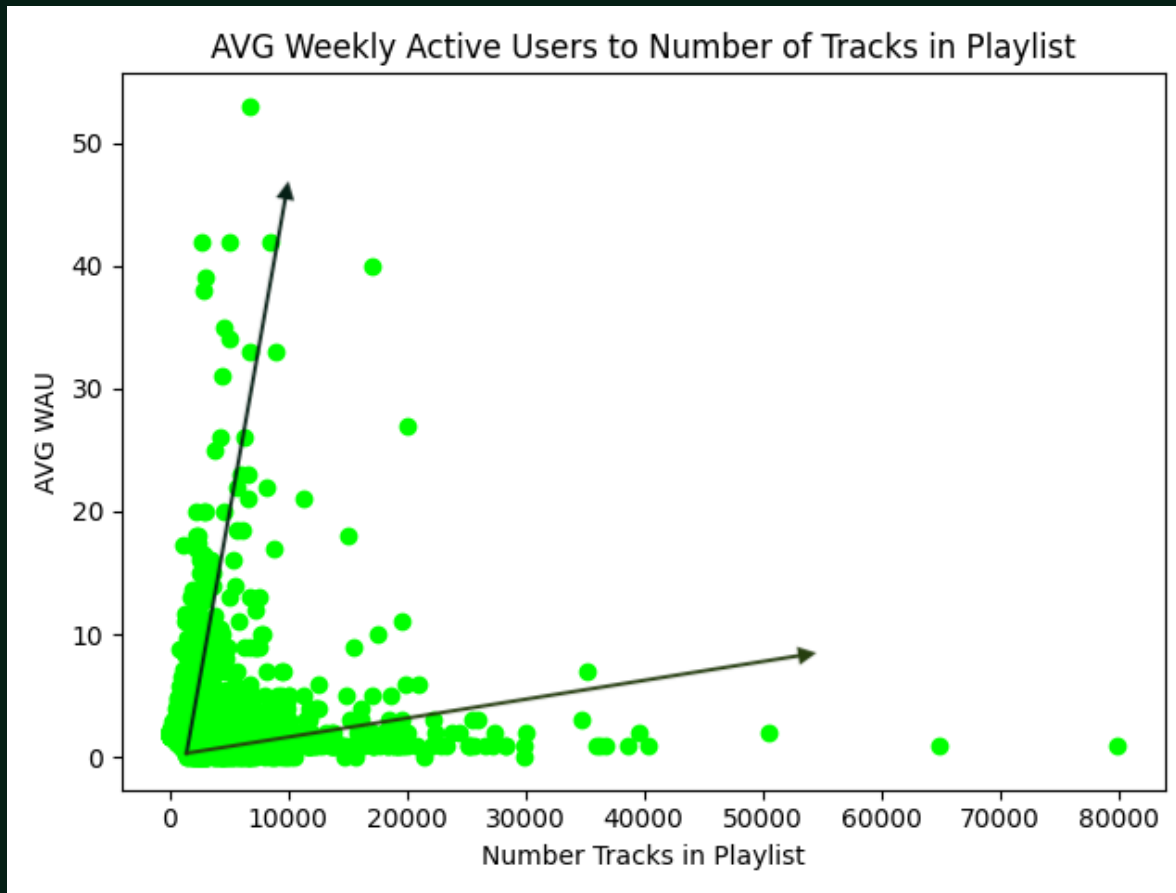


Variety in genre matters

- Positive relationship between number of genres and WAU
- Playlists with 3 genres have higher WAU than playlists with less than 3 genres



Data Exploration & Visualization

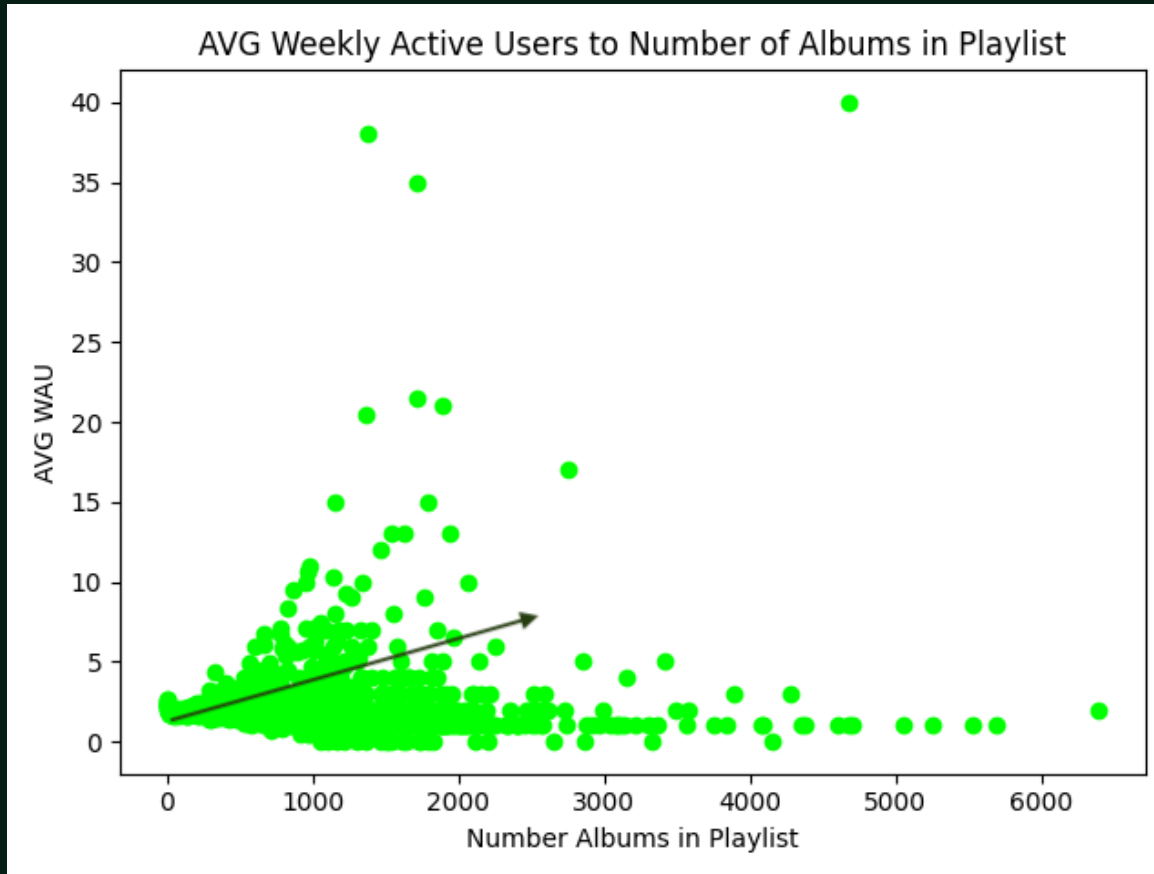


Variety in tracks doesn't matter

- No clear relationship between number of tracks in playlists and its WAU



Data Exploration & Visualization

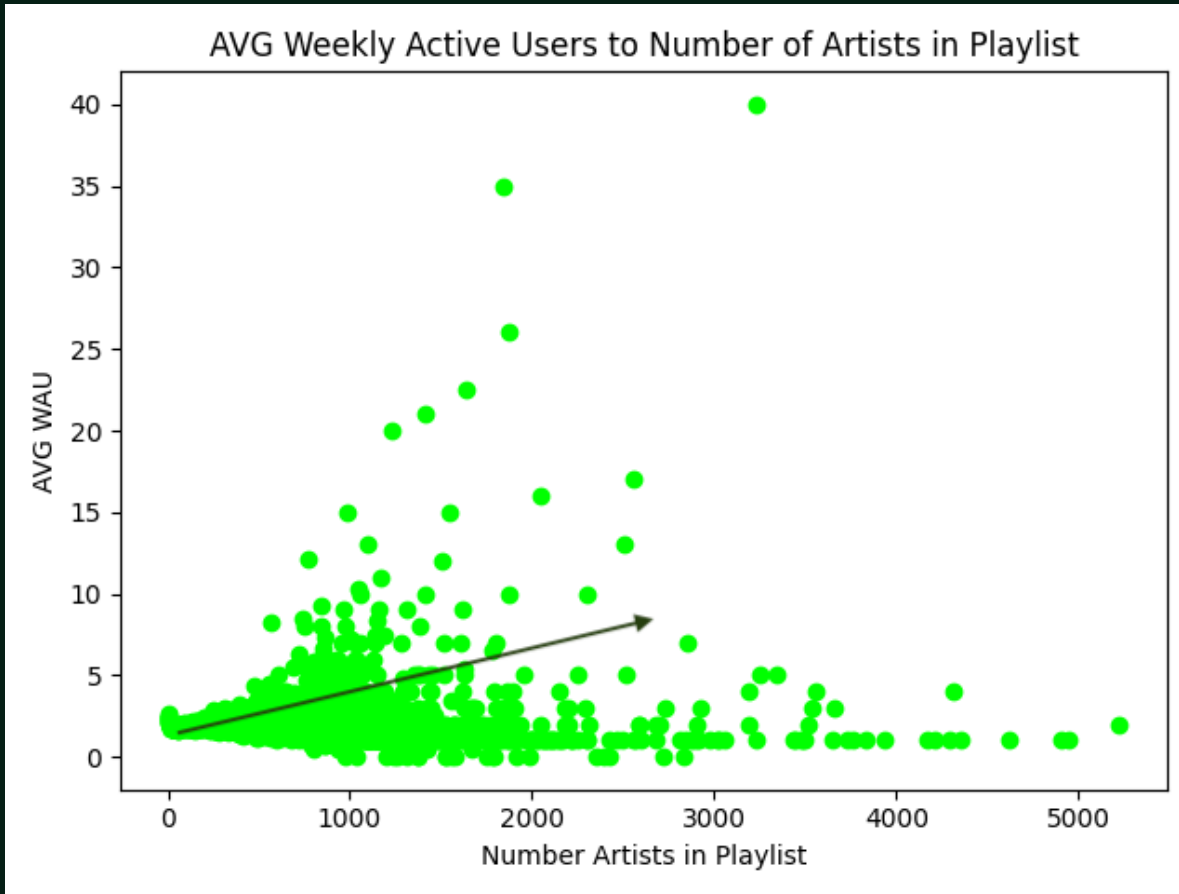


Variety in Albums matters

- Positive relationship between number of albums in playlist and WAU
- Playlists with more albums show higher WAU than playlists with very small amount of albums
- But this effect wears off
but correlation is not causation



Data Exploration & Visualization



Variety in Artists matters

- Positive relationship between number of artists in playlist and WAU
- Playlists with more artists show higher WAU than playlists with very small amount of artists
- But this effect wears off
but correlation is not causation



Data Exploration & Visualization

Sentiment Analysis using NLP

Most Positive Tokens with Highest Sentiment Score (1)

	tokens	sentiment_score_tokens
26832	["best"]	1.0
332300	["best", "rock", "roll"]	1.0
41139	["awesome"]	1.0
121742	["awesome"]	1.0
21059	["greatest", "hits"]	1.0
...
146580	["best"]	1.0
258191	["awesome", "sauce"]	1.0
104147	["kalapana", "best"]	1.0
398253	["anita", "baker", "best", "anita", "baker"]	1.0
5463	["awesome", "mix"]	1.0

Most Negative Tokens with Lowest Sentiment Score (-1)

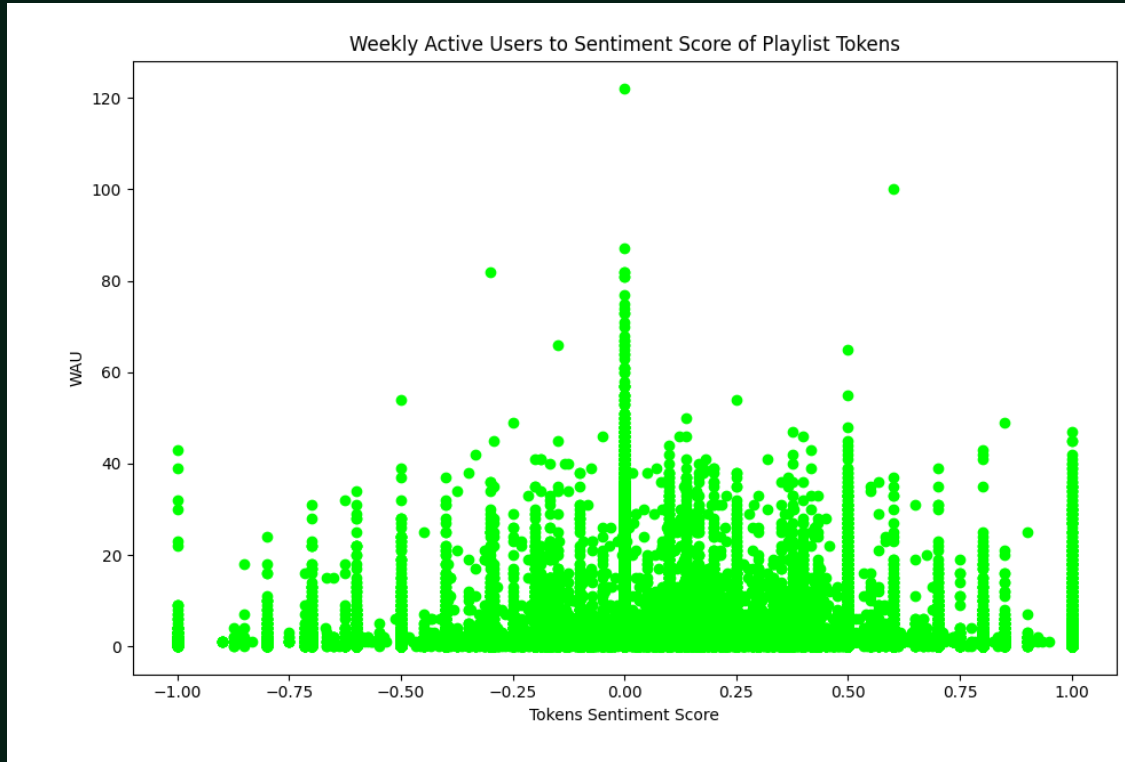
	tokens	sentiment_score_tokens
170780	["awful", "playlist"]	-1.0
49591	["awful", "feelings"]	-1.0
112677	["spank", "rock", "everything", "boring", "eve..."]	-1.0
65024	["make", "nasty"]	-1.0
216320	["make", "work", "terrible", "mix"]	-1.0
...
73398	["bleak"]	-1.0
212903	["worst", "enemy"]	-1.0
323691	["songs", "terrible"]	-1.0
357532	["insane", "clown", "posse"]	-1.0
126980	["terrible", "pop"]	-1.0



Data Exploration & Visualization

Sentiment Analysis using NLP

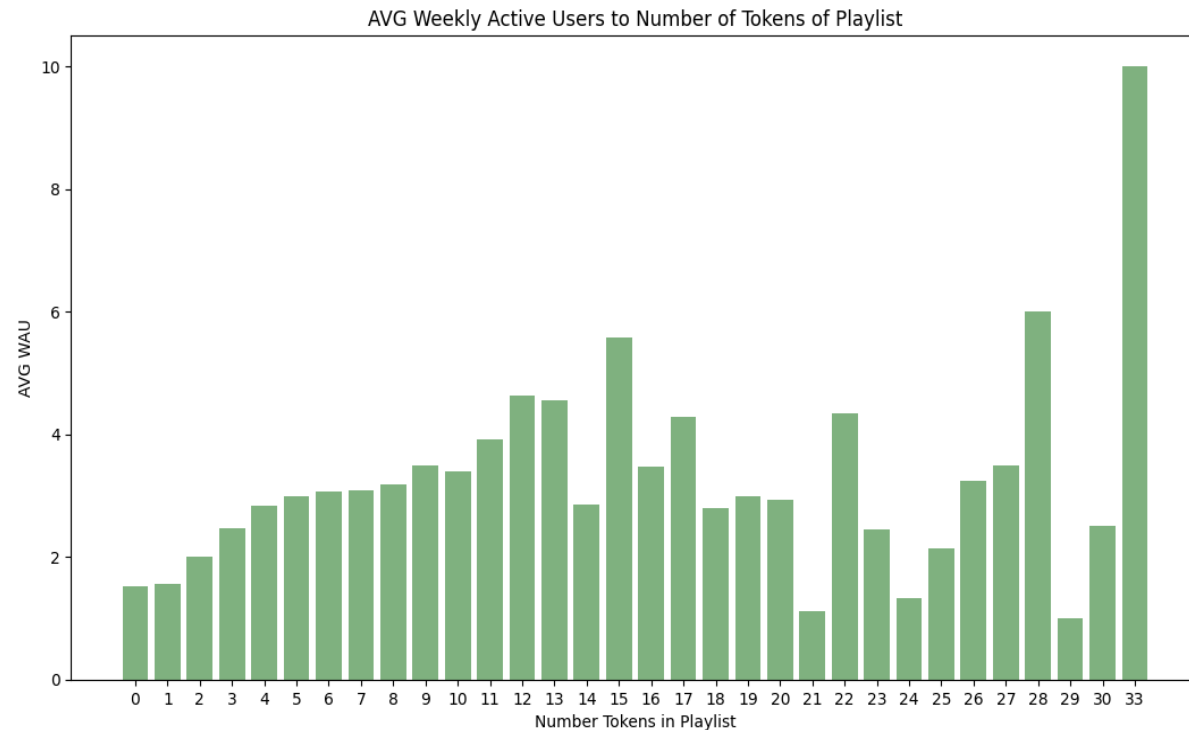
Sentiment of Playlist Tokens doesn't matters



- No clear relationship between sentiment of the playlists tokens and its WUA



Data Exploration & Visualization



Effort Doesn't really matters

- No clear pattern convincing that the length of the tokens (effort proxy) has an impact on the usage of the playlist
- There is some positive pattern from 0-11 as num tokens but after this effect wears off



Linear Regression (preliminary)

Sample size: **15975**

This means 15975 or more measurements/surveys are needed to have a confidence level of 99% that the real value is within $\pm 1\%$ of the measured/surveyed value.

Confidence Level: ?	99% ▾
Margin of Error: ?	1%
Population Proportion: ?	50%
Population Size: ?	398928
<div>Calculate ▶</div> <div>Clear</div>	



OLS Regression Results

Dep. Variable:	wau	R-squared (uncentered):	0.852
Model:	OLS	Adj. R-squared (uncentered):	0.852
Method:	Least Squares	F-statistic:	6546.
Date:	Tue, 22 Mar 2022	Prob (F-statistic):	0.00
Time:	19:45:40	Log-Likelihood:	-28607.
No. Observations:	15975	AIC:	5.724e+04
Df Residuals:	15961	BIC:	5.735e+04
Df Model:	14		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
genre_1_Indie Rock	0.0212	0.033	0.648	0.517	-0.043	0.085
genre_1_Rap	-0.0089	0.041	-0.217	0.828	-0.089	0.071
genre_1_Pop	-0.1739	0.034	-5.119	0.000	-0.240	-0.107
mood_1_Defiant	-0.0483	0.037	-1.319	0.187	-0.120	0.023
mood_1_Excited	0.0081	0.033	0.248	0.804	-0.056	0.072
stream30s	0.0172	0.001	27.587	0.000	0.016	0.018
n_tracks	-2.015e-05	2.5e-05	-0.805	0.421	-6.92e-05	2.89e-05
n_local_tracks	-4.616e-05	0.000	-0.294	0.768	-0.000	0.000
n_artists	-0.0002	0.000	-0.744	0.457	-0.001	0.000
n_albums	0.0007	0.000	2.789	0.005	0.000	0.001
num_genres	0.2295	0.009	26.395	0.000	0.212	0.247
num_tokens	0.0041	0.009	0.470	0.638	-0.013	0.021
sentiment_score_tokens	-0.0127	0.059	-0.215	0.830	-0.128	0.103
users	0.1099	0.000	228.984	0.000	0.109	0.111

Omnibus:	8386.991	Durbin-Watson:	1.990
Prob(Omnibus):	0.000	Jarque-Bera (JB):	14094583.336
Skew:	-0.927	Prob(JB):	0.00
Kurtosis:	148.504	Cond. No.	3.30e+03

Linear Regression

- **Pop Genre:** significant and **negative** effect
- **Streams30s:** significant and **positive** effect
- **Num_albums:** significant and **positive** effect
- **Num_genres:** significant and **positive** effect
- **Users:** significant and **positive** effect



Conclusions

What makes playlist successful?

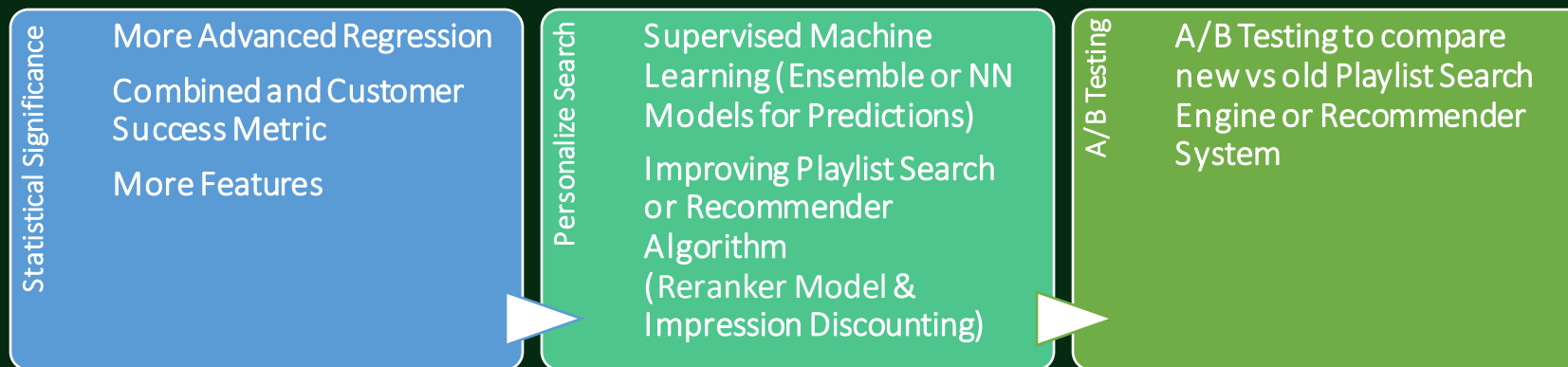
**Variety in
albums**

**Variety
in genres**

But we need more features and more advanced success metric



Next steps





Thank you!