

PROJECT 4

SPOTIFY CUSTOMER BEHAVIOR

May 19, 2024

By



SADE BECKLES

PATRICIA DAHER

ARISLEYDA REYES





Spotify was established on April 23, 2006, by Swedish entrepreneurs Daniel Ek and Martin Lorentzon



Has emerged as a leading global platforms for audio streaming and media services available in mobile, tablet and laptop devices.



Spotify offers users access to a variety library of audio content, including music, podcast and the most recently audiobooks.



December 2024, Spotify reported over 678 millions montly active users, which 268 millions where paying subscribers.

Exploring Customer Behavior

TO

Determine Indicators of paid subscribers

FOR

FOCUSED MARKETING

THROUGH

Analyzing a sample of Spotify consumer data

VIA

Machine Learning

PROBLEM

WHAT MAKES A FREE-SUBSCRIBER
BECOME A PAID SUBSCRIBER ?

GOAL

Understand consumer behavior
to optimize marketing
to increase subscribers

APPROACH

Supervised
and
Unsupervised Machine Learning

DATASET

DATA SOURCES

KOGGLE – Spotify Data

Many Columns &
Qualitative Data !

	Age	Gender	spotify_usage_period	spotify_listening_device	spotify_subscription_plan	premium_sub_willingness	preffered_premium_plan	preferred_listening_content	fav_music_genre	music_time_slot	music_Influencial_mood	music_lis_frequency	music_expl_method	music_recc_rating	pod_lis_frequency	fav_pod_genre	preffered_pod_format	pod_host_preference	preffered_pod_duration	pod_variety_satisfaction
0	20-35	Female	More than 2 years	Smart speakers or voice assistants	Free (ad-supported)	Yes	Family Plan-Rs 179/month	Podcast	Melody	Night	Sadness or melancholy	leisure time	Playlists		Daily	Comedy	Interview	Both	Both	Ok
1	12-20	Male	More than 2 years	Computer or laptop	Free (ad-supported)	Yes	Individual Plan-Rs 119/month	Podcast	Rap	Afternoon	Social gatherings or parties	Workout session	Playlists	2	Several times a week	Comedy	Interview	Both	NaN	Satisfied
2	35-60	Others	6 months to 1 year	Smart speakers or voice assistants	Free (ad-supported)	Yes	Student Plan-Rs 59/month	Podcast	Pop	Night	Relaxation and stress relief	Study Hours, While Traveling	Playlists	4	Once a week	Sports	Interview	NaN	Both	Satisfied
3	20-35	Female	1 year to 2 years	Smartphone, Smart speakers or voice assistants	Free (ad-supported)	No	NaN	Music	Melody	Night	Relaxation and stress relief, Social gathering...	Office hours, Workout session, leisure time	recommendations, Playlists	4	Never	NaN	NaN	NaN	NaN	Ok
4	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	NaN	Music	Melody	Night	Relaxation and stress relief	leisure time	recommendations, Playlists	4	Rarely	Lifestyle and Health	Story telling	Well known individuals	Both	Ok
...
515	20-35	Female	More than 2 years	Smartphone, Computer or laptop, Wearable devices	Free (ad-supported)	Yes	Student Plan-Rs 59/month	Music	Pop	Morning	Uplifting and motivational, Sadness or melancholy	While Traveling	recommendations	5	Rarely	Lifestyle and Health	Conversational	Both	Shorter	
516	20-35	Female	More than 2 years	Smartphone, Computer or laptop	Free (ad-supported)	Yes	Student Plan-Rs 59/month	Music	Pop	Night	Relaxation and stress relief, Uplifting and mo...	While Traveling, Workout session, leisure time	Others	2	Rarely	Comedy	Story telling	Both	Shorter	
517	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	Student Plan-Rs 59/month	Music	Pop	Night	Relaxation and stress relief	leisure time	recommendations	3	Never	NaN	Story telling	Both	Shorter	
518	20-35	Female	6 months to 1 year	Smartphone	Premium (paid subscription)	Yes	Individual Plan- Rs 119/month	Music	Melody	Night	Relaxation and stress relief, Uplifting and mo...	Workout session, leisure time	recommendations, Others	2	Several times a week	Comedy	Conversational	Both		
519	20-35	Female	More than 2 years	Smartphone, Computer or laptop	Premium (paid subscription)	Yes	Student Plan-Rs 59/month	Music	Pop	Night	Relaxation and stress relief, Sadness or melan...	Office hours, Study Hours, While Traveling, lei...	Playlists, Radio	3	Never	NaN	NaN	NaN		

520 rows x 20 columns

```
column_names = ['Age', 'Gender', 'spotify_usage_period',
'spotify_listening_device', 'spotify_subscription_plan',
'premium_sub_willingness', 'preffered_premium_plan',
'preferred_listening_content', 'fav_music_genre', 'music_time_slot',
'music_Influencial_mood', 'music_lis_frequency', 'music_expl_method',
'music_recc_rating', 'pod_lis_frequency', 'fav_pod_genre',
'preffered_pod_format', 'pod_host_preference',
'preffered_pod_duration', 'pod_variety_satisfaction']
```





```
column_names = [ 'Age', 'Gender', 'spotify_usage_period',  
'spotify_listening_device', 'spotify_subscription_plan',  
'premium_sub_willingness', 'preferred_premium_plan',  
'preferred_listening_content', 'fav_music_genre', 'music_time_slot',  
'music_Influencial_mood', 'music_lis_frequency', 'music_expl_method',  
'music_recc_rating', 'pod_lis_frequency', 'fav_pod_genre',  
'preferred_pod_format', 'pod_host_preference',  
'preferred_pod_duration', 'pod_variety_satisfaction' ]
```

method	music_recc_rating	pod_lis_frequency	fav_pod_genre	preferred_pod_format	pod_host_preference	preferred_pod_duration	pod_variety_satisfaction
Playlists	3	Daily	Comedy	Interview	Both	Both	Ok
Playlists	2	Several times a week	Comedy	Interview	Both	NaN	Satisfied
Playlists	4	Once a week	Sports	Interview	NaN	Both	Satisfied
dations, Playlists	4	Never	NaN	NaN	NaN	NaN	Ok
dations, Playlists	4	Rarely	Lifestyle and Health	Story telling	Well known individuals	Both	Ok
...
dations	5	Rarely	Lifestyle and Health	Conversational	Both	Shorter	
Others	2	Rarely	Comedy	Story telling	Both	Shorter	
dations	3	Never	NaN	Story telling	Both	Shorter	
dations, Others	2	Several times a week	Comedy	Conversational	Both		
s, Radio	3	Never	NaN	NaN	NaN		



PROCESS

Each ClusterModel View we Created is Based on a
Different Combination of Features

ETL



WHO

DEMOGRAPHIC OF USER

Age

Gender

spotify_subscription_plan

premium_sub_willingness



DEMOGRAPHICS

HOW

INTERACTION WITH APP

spotify_usage_period

music_time_slot

music_lis_frequency

music_recc_rating

spotify_listening_devices

potify_subscription_plan

premium_sub_willingness

1. preffered_premium_plan



INTERACTION

WHAT

CONTENT THEY LISTEN TO

preferred_listening_content

fav_music_genre

music_Influential_mood

spotify_subscription_plan

premium_sub_willingness

1.pod_lis_frequency

2.fav_pod_genre

3.preffered_pod_format

4.pod_host_preference

5.preffered_pod_duration

6.pod_variety_satisfaction



CONTENT

PROCESS

JUPYTER NOTEBOOK

PYTHON

SCIKIT LEARN

STANDARD SCALAR

ETL TOOLS

CHAT
GBT



Aris

GET DUMMIES



Patricia

Sade



tell me about
yourself



PROCESS

JUPYTER NOTEBOOK

PYTHON

SCIKIT LEARN

STANDARD SCALAR

ETLTOOLS

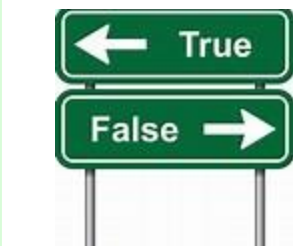
PCA

CLUSTERING

K MEANS

RANDOM FORREST

NEURAL NETWORKS



**SUPERVISED
MACHINE
LEARNING**

We used **Neural Networks** to create learning predictive models that study the behavior of Users willing to subscribe, and **Random Forrest** to identify feature relevance.

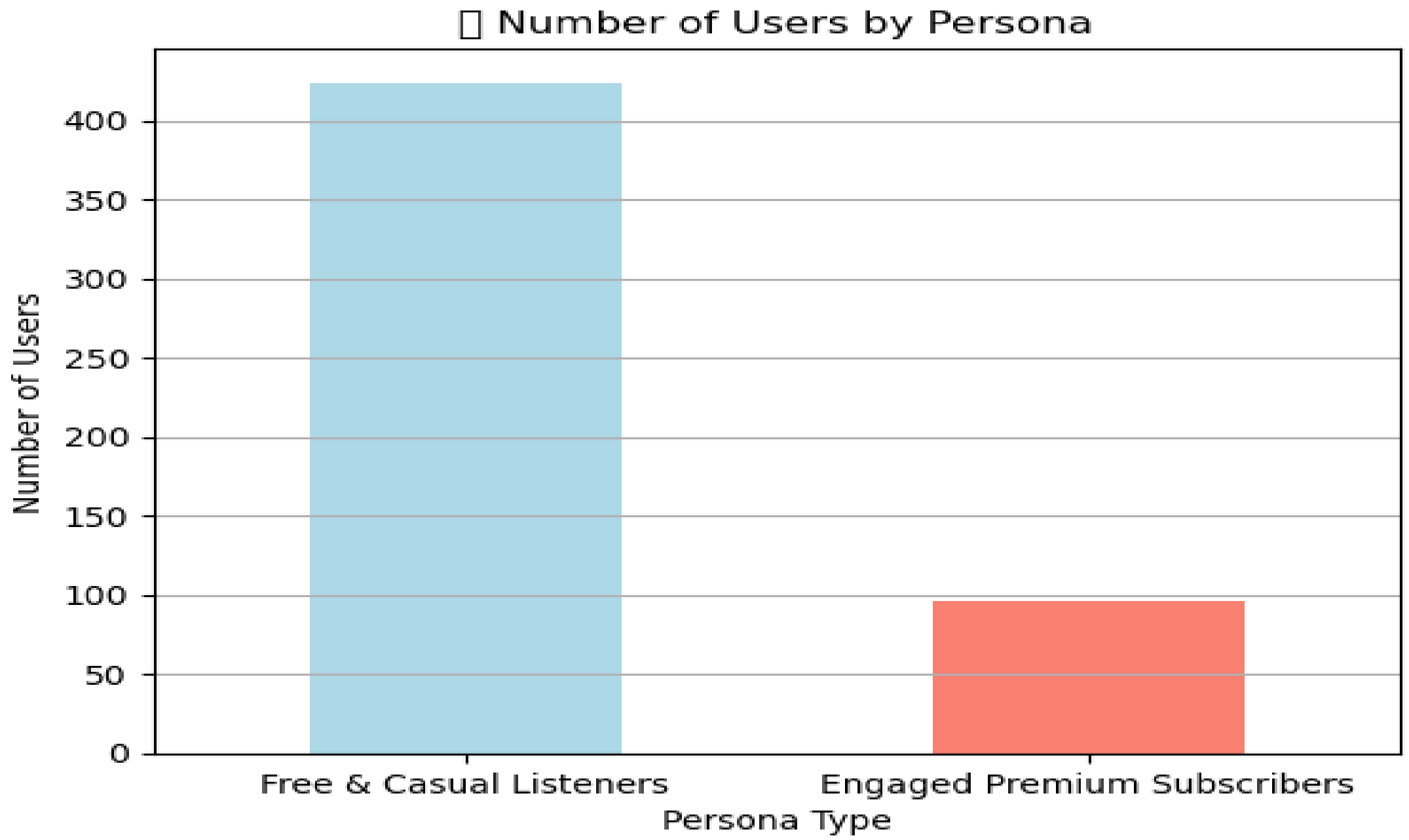
**UNSUPERVISED
MACHINE
LEARNING**

We used **K-Means Clustering** as our main machine learning model to identify **listener personas** based on behaviors, mood, demographics, and subscription preferences. We also **applied PCA to visualize cluster separation** and help interpret **complex patterns**.

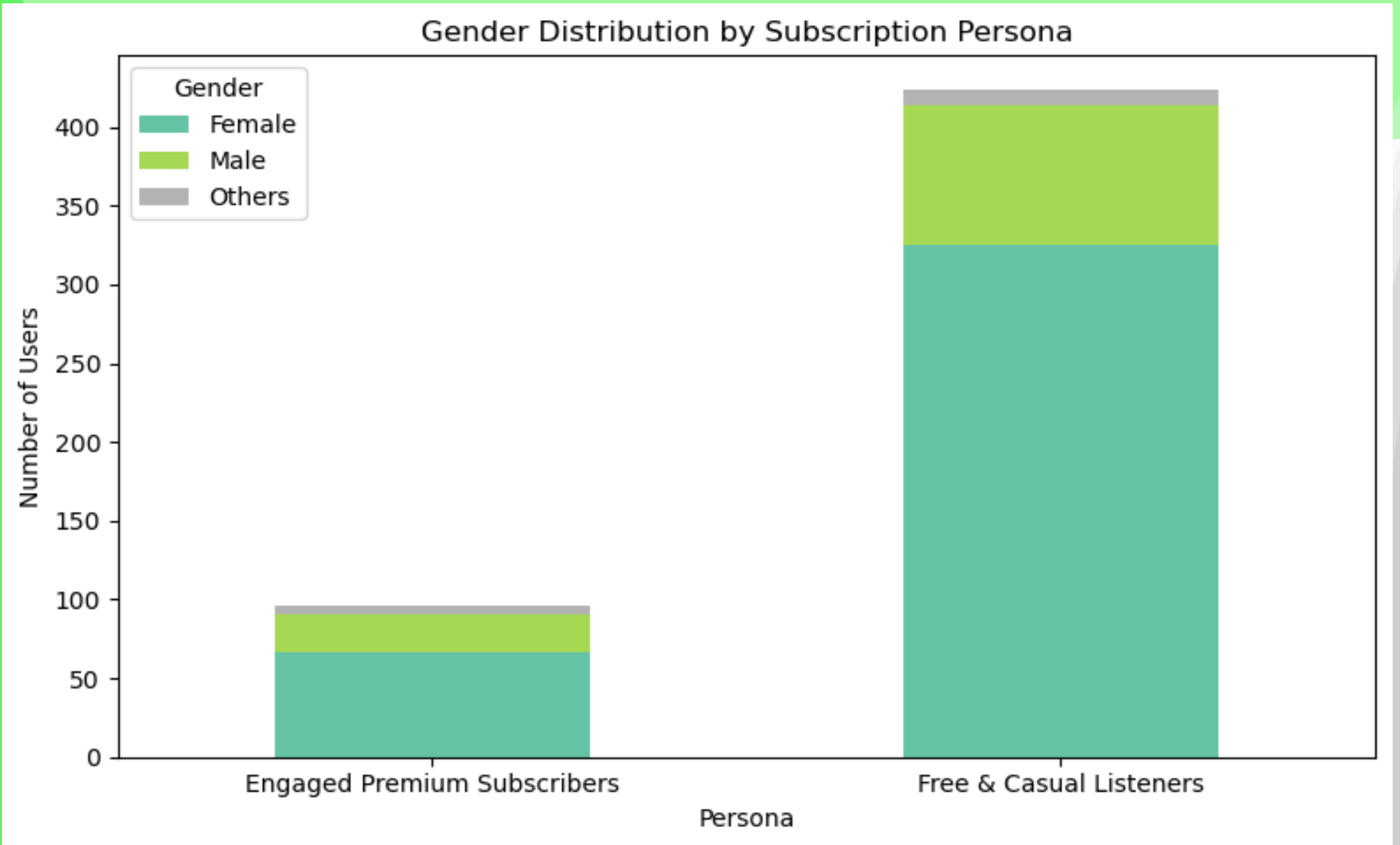
DEMOGRAPHICS OF USERS

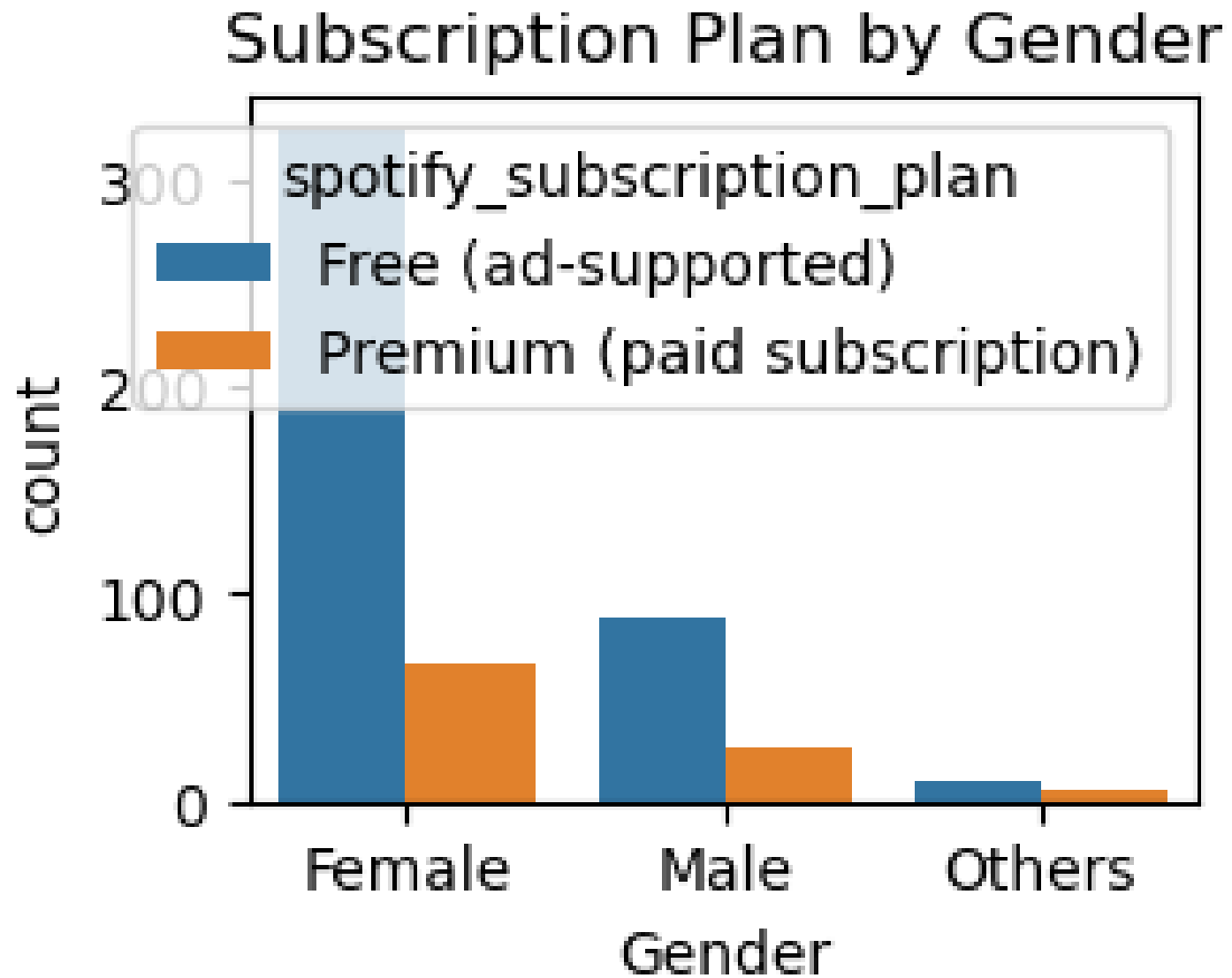
Age
Gender

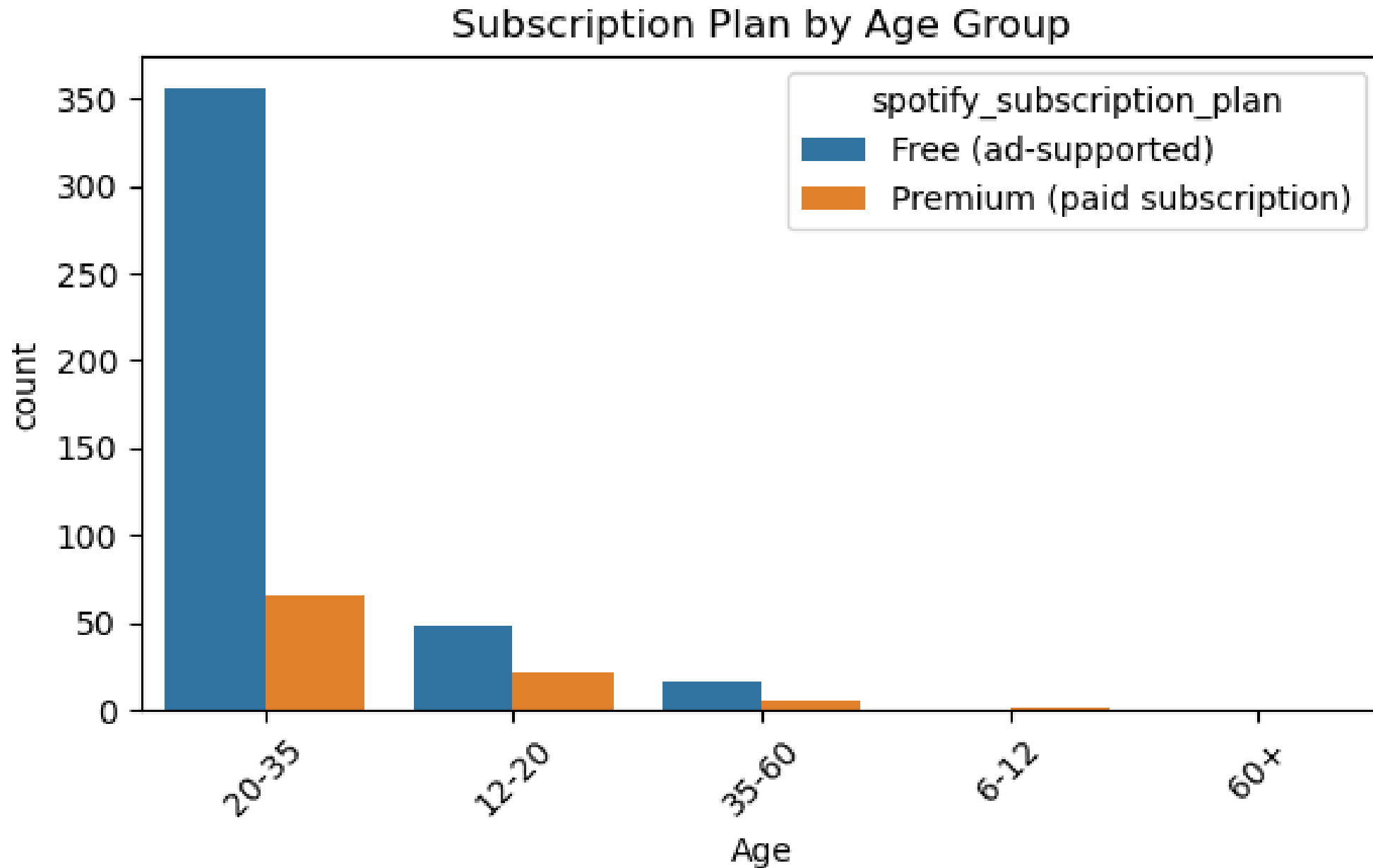




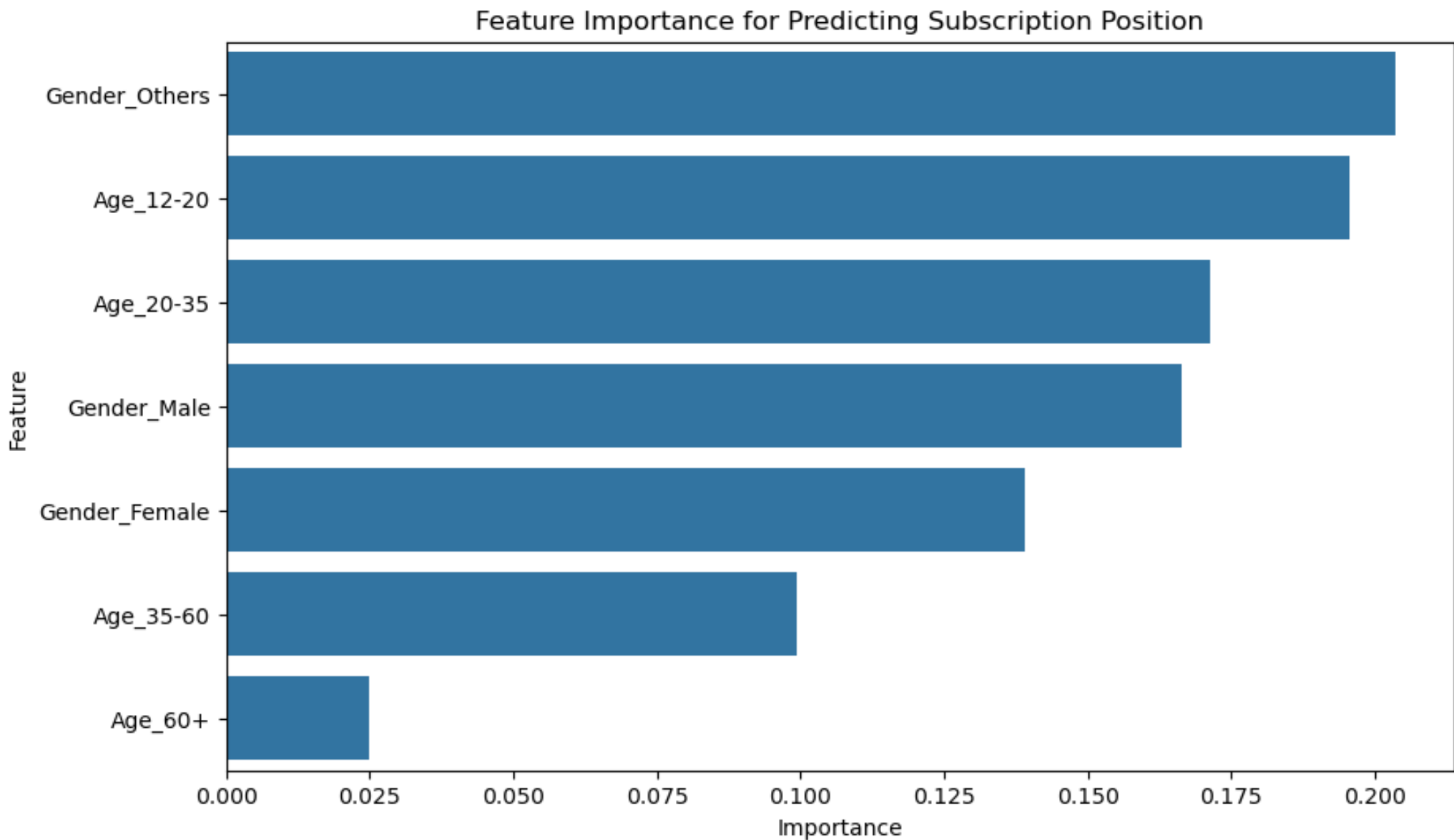
DEMOGRAPHICS







FEATURE IMPORTANCE

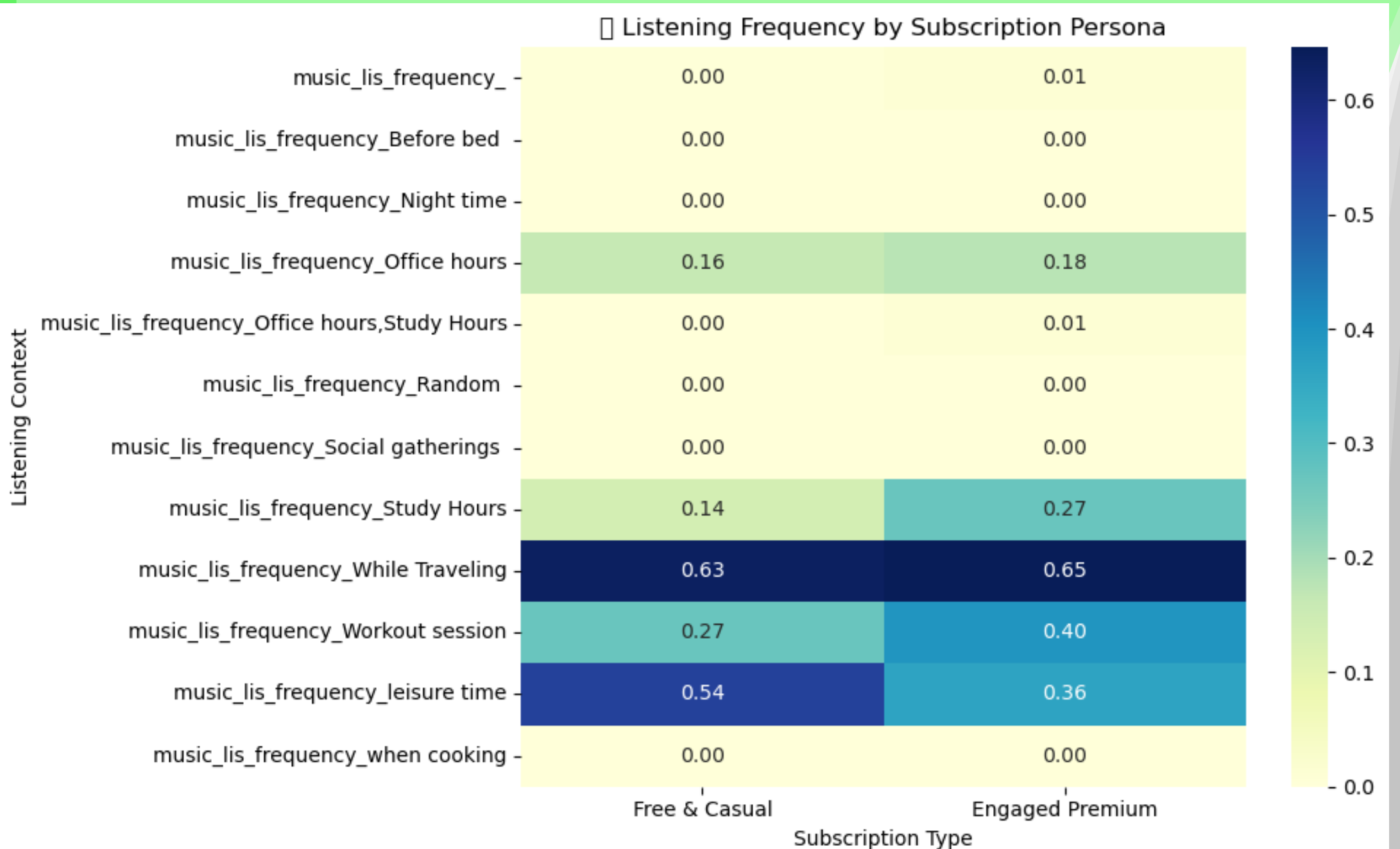


INTERACTION WITH APP

spotify_usage_period
music_time_slot
music_lis_frequency
music_recc_rating
spotify_listening_devices
potify_subscription_plan
premium_sub_willingness
preffered_premium_plan



INTERACTION

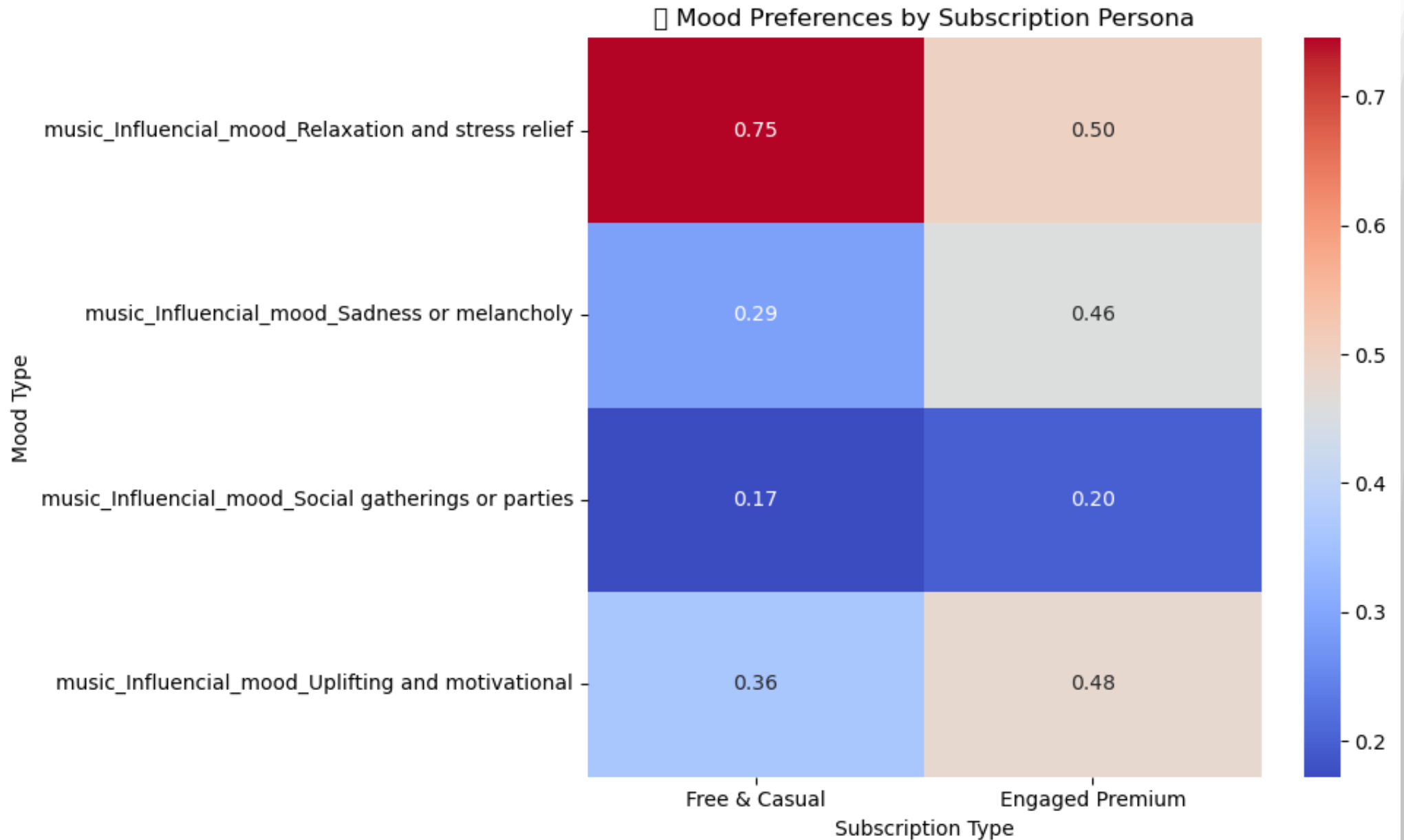


CONTENT LISTENED TO

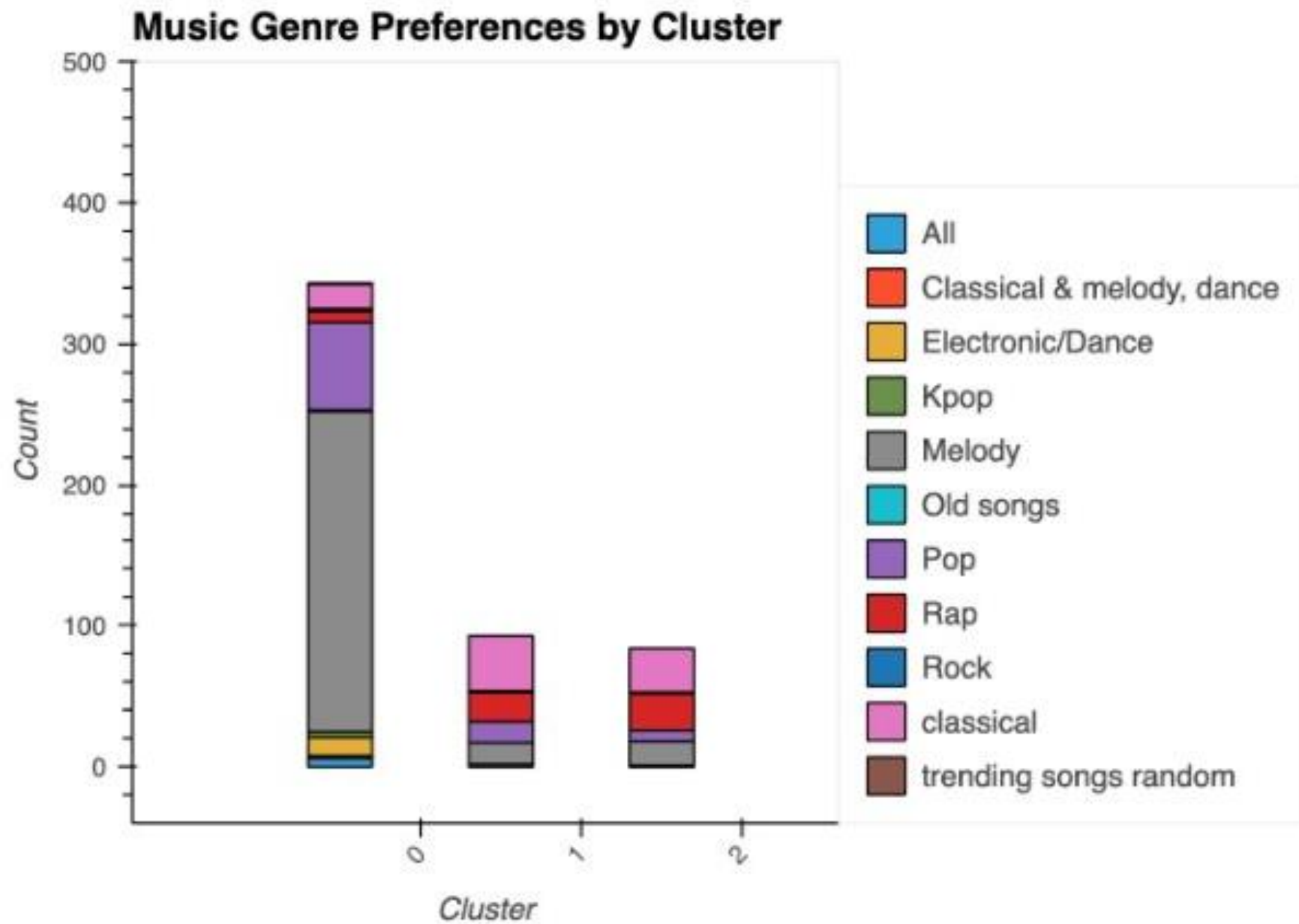
preferred_listening_content
fav_music_genre
music_Influencial_mood
spotify_subscription_plan
premium_sub_willingness



CONTENT



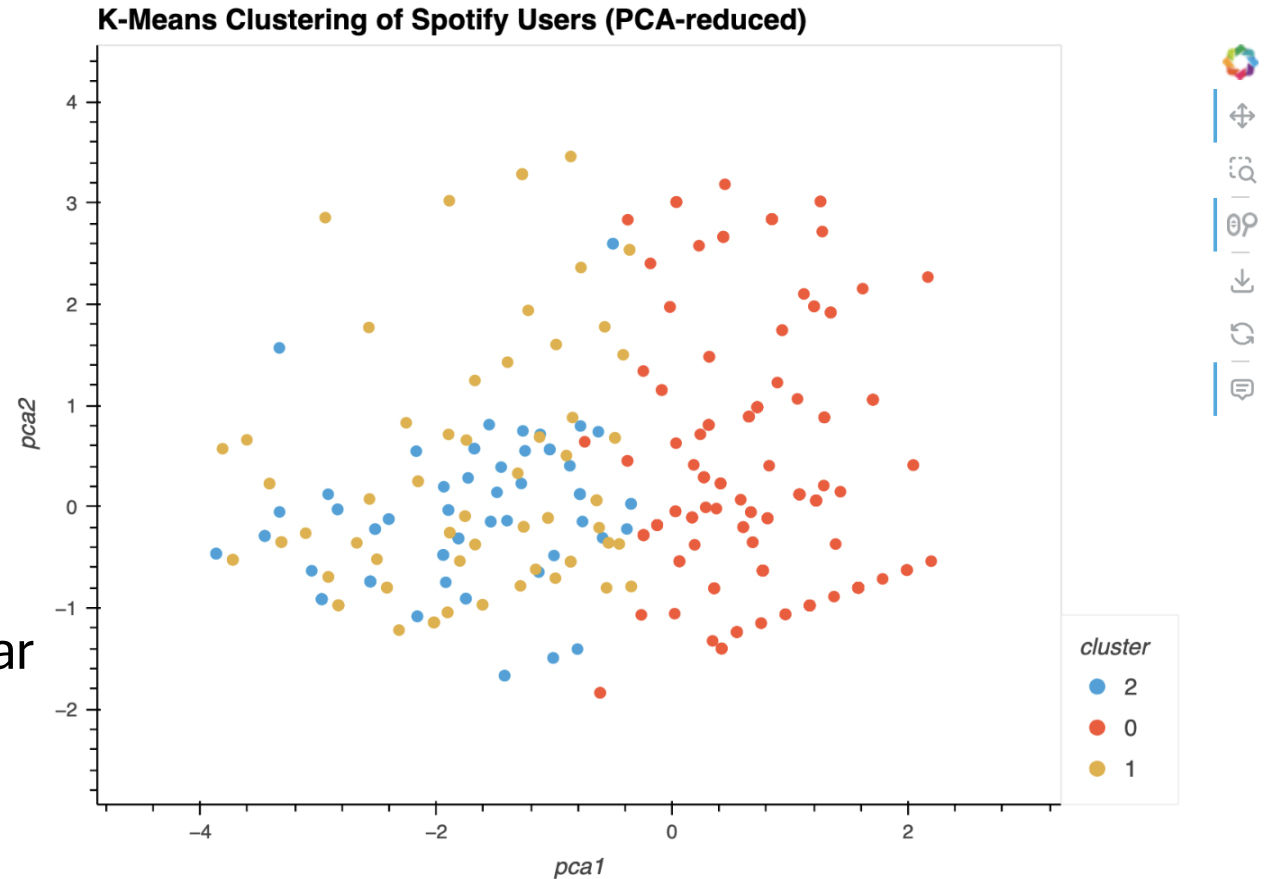
FEATURE IMPORTANCE



FEATURE IMPORTANCE

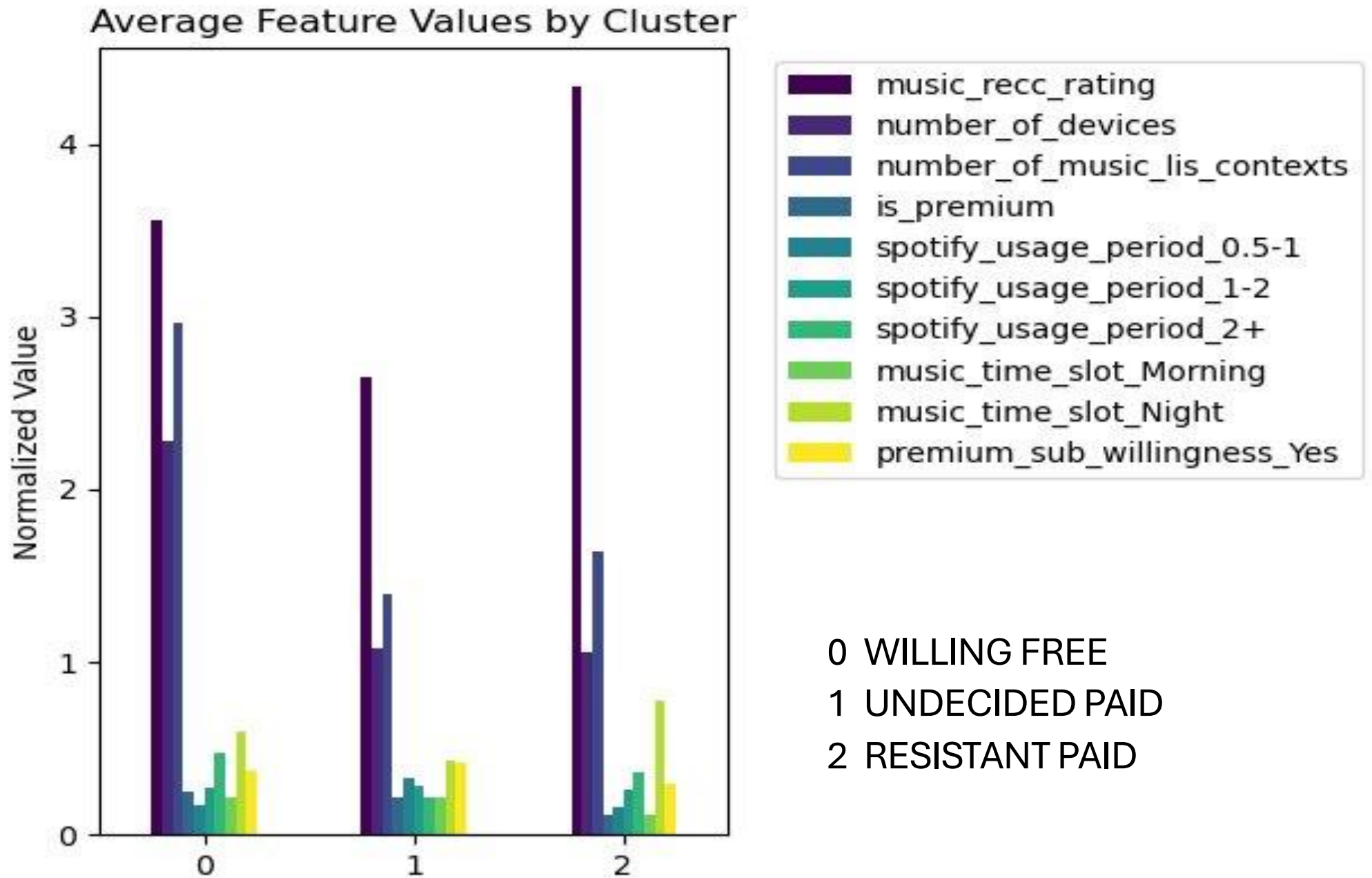
CLUSTERING BASED ON ALL FEATURES:

- . Listening Preference
 - . Music Taste
 - . Subscription Info
 - . Mood
- PCA reduced dimensions to 2
 - Points close together are similar in features
 - Color Distinct segments
 - K means



fav_music_genre	All	Classical & melody, dance	Electronic/Dance	Kpop	Melody	Old songs	Pop	Rap	Rock	classical	trending songs random
cluster											
0	6.0	2.0	13.0	4.0	227.0	1.0	62.0	8.0	2.0	17.0	1.0
1	0.0	0.0	2.0	0.0	15.0	0.0	15.0	21.0	1.0	39.0	0.0
2	0.0	0.0	1.0	0.0	17.0	0.0	8.0	26.0	1.0	31.0	0.0

FEATURE IMPORTANCE



KEY FINDINGS

**TARGET DEMOGRAPHIC**

- . THE MAJORITY OF USERS ARE **WOMEN**
- . GENDER **OTHER** IS AN IMPORTANT TARGET GROUP
- . THE MAJORITY OF USERS ARE OF **AGE 25 –35**

**INTERACTION BEHAVIOR**

- . MOST USERS USE SPOTIFY AT **NIGHT**
- . USERS WHO USE SPOTIFY ON **DEVICES OTHER** THAN SMARTPHONES ARE LIKELY TO BECOME SUBSCRIBERS
- . PEOPLE WHO USE SPOTIFY DURING **TRAVELING, WORKOUT** AND **RELAXATION** ARE MOST LIKELY TO BECOME SUBSCRIBERS

**CONTENT**

- . USERS MOSTLY USE SPOTIFY FOR **MUSIC MORE** THAN **PODCASTS**
 - IT MAY BE BENEFICIAL TO OFFER MORE EXCLUSIVE FEATURES FOR SUBSCRIBERS TO INCREASE THEIR SATISFACTION
- . **MELODY** IS THE MOST POPULAR **GENRE** OF MUSIC

BYE



THANK YOU!

