ASSIGNMENT 2 FRONT SHEET

Qualification	**BTEC Level 5 HND Diploma in Computing**		
Unit number and title	Unit 14: Business Intelligence		
Submission date		**Date Received 1st submission**	
Re-submission Date		**Date Received 2nd submission**	
Student Name	Nguyen Duy Lam Tung	**Student ID**	GCH210001
Class	GCH1005	**Assessor name**	Nguyen Duc Giang

Student declaration

I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.

Student's signature	***Tung***

Grading grid

P3	P4	P5	P6	M3	M4	D3	D4

☼ **Summative Feedback:**		☼ **Resubmission Feedback:**
Grade:	**Assessor Signature:**	**Date:**
IV Signature:		

BUSINESS INTELLIGENCE

TEAM 1

UNIVERSITY OF GREENWICH

Alliance with FPT Education

OUR TEAM

Nguyễn Việt Tấn

Nguyễn Duy Lâm Tùng

Nguyễn Thái Bình

WHAT IS BUSINESS INTELLIGENCE?

Definition: BI is an intelligent management reporting system capable of combining technology to assist business executives in managing the enormous volume of data arriving from numerous sources and successfully utilizing that data source. BI offers multidimensional, anticipatory perspectives of specific patterns from historical data. Optimizing each company decision for maximum efficiency is the ultimate goal of business intelligence.

BENEFIT OF BI

- Enhanced customer insights - By analyzing user data, Spotify can better understand customer preferences and behaviors to create more personalized experiences.

- Operational efficiency - Spotify can use data to identify inefficiencies in their processes and operations and find areas for optimization.

- Increased revenues - Analytics on customer lifetime value, churn, and acquisition costs can help Spotify maximize revenues.

- Informed strategy - Market intelligence, competitive analysis, and financial modeling enable data-driven strategic planning.

- Faster innovation - Spotify can test and analyze new features and offerings to rapidly iterate.

REAL EXAMPLE:

How Spotify could use BI to analyze listening patterns and improve recommendations

UNIVERSITY OF GREENWICH
Alliance with FPT Education

Challenge:
- Spotify had generic music and podcast recommendations that didn't reflect users' personal tastes.
- This resulted in low engagement with recommended content and missed opportunities to cater to users.

Solution:
- Implemented a BI system to collect and analyze data on users' listening histories, likes/dislikes, playlists, etc.
- Used machine learning algorithms to identify patterns and clusters among users based on their activity.
- Developed models to categorize users into different taste profiles and preference groups.
- Built a recommendation engine that matches users to relevant content based on their learned profiles.
- Recommendations are now highly personalized based on each user's unique listening behavior.

Results:
- Recommendation click-through rates increased by 25% after implementing personalized models.
- Average time spent listening per user grew by 20% as recommendations better matched individual interests.
- Positive sentiment about recommendations on social media improved by 15% over generic options.
- Conversion rates from free to paid subscriptions increased 10%, attributable to better customization.

COLLECTION TECHNIQUES

UNIVERSITY OF GREENWICH

Alliance with FPT Education

Using analytics, a method of gleaning useful information from large amounts of data, businesses may spot patterns and opportunities.

Cleansing

The method of identifying and removing incorrect information from a data set is known as cleaning (or data cleansing).Dirty data is information that has a negative effect on your market.

COLLECTION TECHNIQUES

UNIVERSITY OF GREENWICH

Alliance with FPT Education

spotify_id	name	followers	popularity	genres	chart_hits
48WvrUGo	Byklubben	1738	24	['nordic hc	['no (3)']
4lDiJcOJ2G	Kontra K	1999676	72	['christlich	['at (44)', 'de (111)', 'lu (22)', 'ch (31)', 'vn (1)']
652XIvIBN	Maxim	34596	36	[]	['de (1)']
3dXC1YPbi	Christophe	249233	52	['dancehal	['at (1)', 'de (1)']
74terC9ol9	Jakob Hell	21193	39	['classic sv	['se (6)']
0FQMb3m	Madh	26677	19	[]	['it (2)']
71BhXa24	Juice	11312	37	['swedish c	['se (4)']
3TG1RXLa	Nehuda	36252	31	['francoto	['fr (1)']
7MFvm8p	VovaZiLvo	14619	43	['ukrainian	['ua (1)']
5I82NM6jI	Nata Reco	188	12	[]	['do (1)']
5OkJ74jazI	Yomi	7246	30	['norwegia	['no (2)']
6swIRY9N	Kauniit & U	798	9	['suomi ro	['fi (2)']
5qBZETtyz	Danny Elfn	327256	63	['orchestra	['pl (2)', 'pt (2)', 'tw (1)']
0AjnB78Sil	Attraction	11174	21	['japanese	['jp (1)']
5RYLofQH	MaRina	6408	29	[]	['pl (1)']
3xs0LEzcP	Rockwell	40344	58	[]	['us (1)', 'gb (1)', 'at (1)', 'be (1)', 'ca (1)', 'ee (1)', 'de (1)', 'is (1)', 'ie (1)', 'lv (1)', 'lt (1)', 'sk (1)']
4kYtSSCtiz	Cosmos In	17	0	[]	['hn (3)']
2tGETO39	PÃ¡lmi Gur	3207	27	['classic ic	['is (4)']

In the first image you can see that the data is a mess, messy and full of counter information and after cleaning up in the 2nd image, everything looked very easy to distinguish data collection

	A	B	C	D	E
1	spotify_id	name	followers	popularity	genres
2	6eUKZXaKl	Ed Sheeran	1.02E+08	90	['pop', 'uk pop']
3	66CXWjxzI	Ariana Gra	83045090	88	['dance pop', 'pop']
4	6qqNVTkY	Billie Eilish	68407227	88	['art pop', 'electropop', 'pop']
5	3TVXtAsR1	Drake	66852536	95	['canadian hip hop', 'canadian pop', 'hip hop', 'rap', 'toronto rap']
6	1uNFoZAH	Justin Bieb	65590075	90	['canadian pop', 'pop']
7	7dGJo4pcI	Eminem	59184634	90	['detroit hip hop', 'hip hop', 'rap']
8	06HL4z0C\	Taylor Swi	58554324	94	['pop']
9	4YRxDV8w	Arijit Singh	58523986	86	['desi pop', 'filmi', 'modern bollywood']
10	4q3ewBCX	Bad Bunny	55669387	100	['reggaeton', 'trap latino']
11	3Nrfpe0tU	BTS	54532917	91	['k-pop', 'k-pop boy group']
12	5pKCCKE2	Rihanna	50002769	87	['barbadian pop', 'dance pop', 'pop', 'urban contemporary']
13	1Xyo4u8u	The Week	49387909	93	['canadian contemporary r&b', 'canadian pop', 'pop']
14	0du5cEVh	Bruno Mar	43103624	87	['dance pop', 'pop']
15	1dfeR4Ha	Queen	42644068	84	['classic rock', 'glam rock', 'rock']

ANALYSIS TECHNIQUES

Dashboards

A BI dashboard is a data visualization and analysis tool that displays on one screen the status of key performance indicators (KPIs) and other important business metrics and data points for an organization,department, team orprocess.

Dashboards are an integral component of most BI software platforms and are widely used to deliver analytics information to business executives and workers.

Descriptive Analytics

Descriptive analytics refers to techniques used to aggregate, summarize, and present historical data in a meaningful way to describe what has occurred in the past. The goal is to provide insights into the current state of affairs.

BI TOOLS

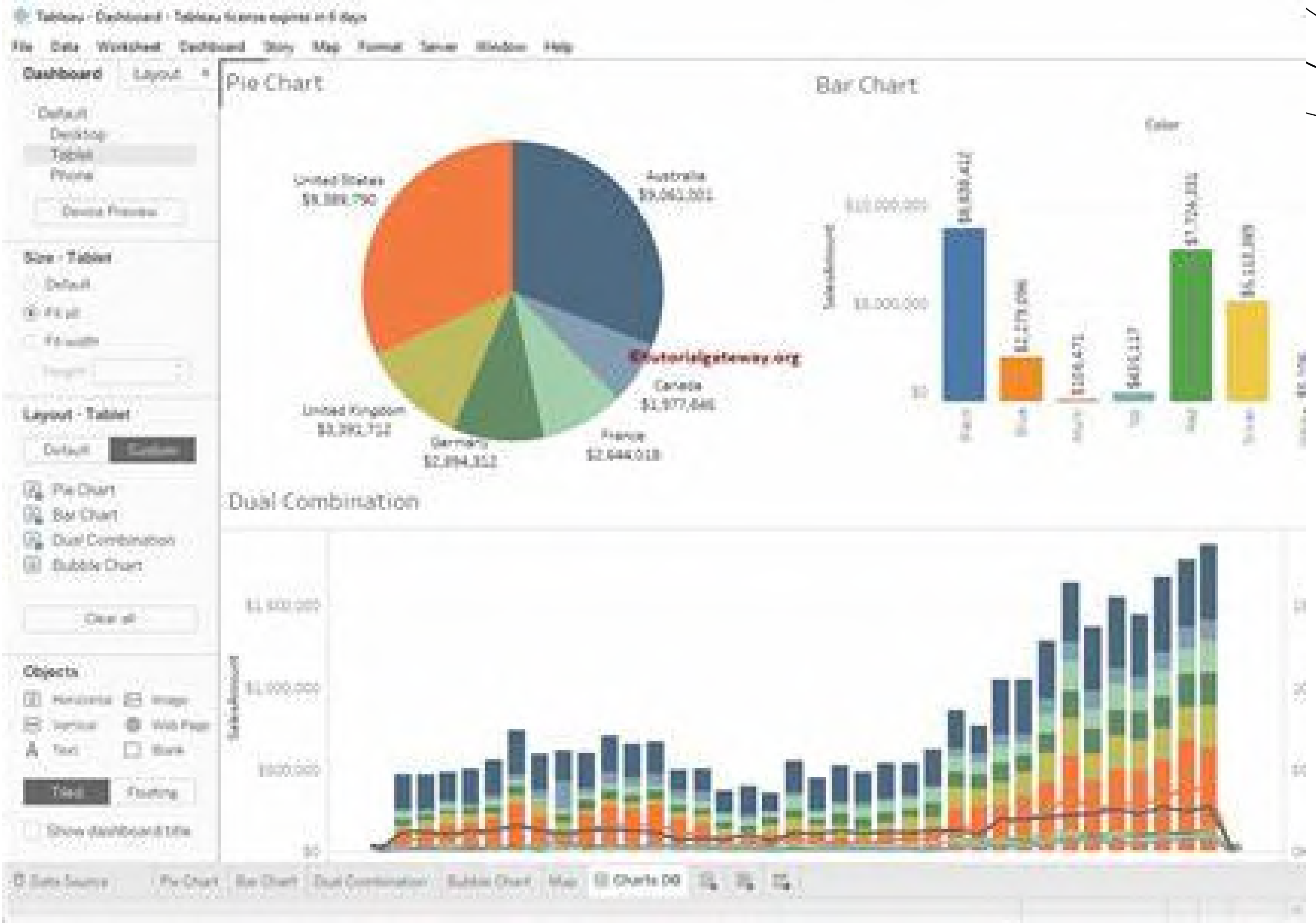
Development Tools

- Semantics in Python are interpretable and dynamic. It is perfect for RAD since it has high-level, built-in data structures and dynamic import and binding.
- Python is simple to understand and use, which lowers the cost of program maintenance. Reusing code and applications is made easier by Python's modules and packages. For all popular operating systems, the Python parser source code and binaries are available for free.

BI TOOLS

Data Visualization Tools

- Tableau excels at data analysis and visualization, making it a valuable Business Intelligence tool. Users can use the tool to analyze data, visualize it, and share it without much help from IT.
- Among the various data sources that Tableau may connect to include Microsoft Excel, Oracle, Microsoft SQL, Google Analytics, and SalesForce. Each user will have access to an intuitive dashboard that has been carefully designed.



TASK 2

DATASET OVERVIEW

Spotify User Behavior Dataset:

- Source: Kaggle public dataset
- Scope: User streaming history for 1,056 users
- Size: 232,531 rows
- Information: User ID, song details, genres, timestamps

Spotify Artist Network Dataset:

- Source: Kaggle public dataset
- Scope: Collaboration network of Spotify artists
- Size: 8,024 rows
- Information: Artist IDs, links between collaborating artists

DATA PRE-PROCESSING

Data Cleaning - Fixing or removing incorrect, incomplete, inaccurate, irrelevant, duplicated, or improperly formatted data.

Data Reduction - Reducing the data volume by aggregating, eliminating redundancies, etc.

LESS POPULAR SINGERS AND NO MUSIC GENRE

	A	B	C	D	E	F	G	H
1	spotify_id	name	followers	popularity	genres	chart_hits	recent_plays	liked_songs
2	48WvrUGoijadXXCsGo	Byklubben	1738	24	['nordic house', 'russel	['no (3)']	321981	236154
3	4lDiJcOJ2GLCK6p9q5B	Kontra K	1999676	72	['christlicher rap', 'gerr	['at (44)', 'de (111)', 'lu	405774	86348
4	652XlvIBNGg3C0KIGEJ	Maxim	34596	36	[]	['de (1)']	70341	109744
5	3dXC1YPbnQPsfHPVkn	Christopher Martin	249233	52	['dancehall', 'lovers ro	['at (1)', 'de (1)']	277698	269853
6	74terC9ol9zMo8rfzhS(	Jakob Hellman	21193	39	['classic swedish pop',	['se (6)']	159430	68485
7	0FQMb3mVrAKlyU4H5	Madh	26677	19	[]	['it (2)']	185322	251529
8	71BhXa24Zf5zcikUb00	Juice	11312	37	['swedish drill', 'swedis	['se (4)']	307622	112614
9	3TG1RXLaEhHz5SIPMV	Nehuda	36252	31	['francoton']	['fr (1)']	273893	85044
10	7MFvm8pwjLdmVBZdc	VovaZiLvova	14619	43	['ukrainian hip hop', 'ul	['ua (1)']	450286	221010
11	5I82NM6jN4Y267iHw\	Nata Record	188	12	[]	['do (1)']	129962	299847
12	5OkJ74jazRprW3kCTjf	Yomi	7246	30	['norwegian pop rap']	['no (2)']	114894	168721
13	6swIRY9NcGlW0PwO5	Kauniit & Uhkarohkeat	798	9	['suomi rock']	['fi (2)']	433428	52764
14	5qBZETtyzfYnXOobDXl	Danny Elfman	327256	63	['orchestral soundtrac	['pl (2)', 'pt (2)', 'tw (1)'	171603	106939
15	0AjnB78SikKKhqYoqFI(	Attractions	11174	21	['japanese dance pop']	['jp (1)']	221110	174459
16	5RYLofQHxZcrEl9Bl23I	MaRina	6408	29	[]	['pl (1)']	34987	79721
17	3xs0LEzcPXtgNfMNcH	Rockwell	40344	58	[]	['us (1)', 'gb (1)', 'at (1)'	409397	51993
18	4kYtSSCtizbVfZkgAj4Xb	Cosmos Insania	17	0	[]	['hn (3)']	110356	281971
19	2tGETO39QXKpmeQM	PÃ¡lmi Gunnarsson	3207	27	['classic icelandic pop'	['is (4)']	459039	87194
20	5mjhFqCEMhIwO4ZOY	GeeSixFive	517	0	[]	['za (1)']	343636	80799
21	6cd9X2t4HN8zjoxObla	Suur Papa	6752	29	['estonian hip hop', 'es	['ee (3)']	27159	107434
22	2NUz5P42WqkxilbI8oc	Vybz Kartel	1026598	63	['dancehall', 'jamaican	['cr (3)', 'pa (1)']	389385	161443
23	3PGwQavDRylE3cEDRI	CAMO	394	10	[]	['dk (2)']	201771	137865
24	4Lm0pUvmisUHMdoky	Apocalyptica	864846	60	['alternative metal', 'b	['fi (2)']	458467	176270
25	6M2wZ9GZgrQXHCFfjv	Dua Lipa	36163788	88	['dance pop', 'pop', 'uk	['us (17)', 'gb (25)', 'ar (	122427	239989
26	3W78vKR2pmGmYbtN	Akky	35	1	[]	['ca (1)', 'in (1)']	460853	221979
27	0SIXZXJCAhNU8sxK0qr	Stefanie Sun	561300	57	['c-pop', 'mandopop', '	['hk (2)', 'sg (3)', 'tw (36	154335	207040
28	7jXoGtR69J2iYCefc58N	Twins	92440	45	['c-pop', 'cantopop']	['hk (3)']	156591	257384
29	2sjnyyL9NXijL3Fr2eLik	Aspova	339292	57	['turkish hip hop', 'turk	['tr (12)']	483740	96239

CODE EXPLAIN

UNIVERSITY OF GREENWICH

Alliance with FPT Education

Filter for less popular singers and no music genre

```python
import pandas as pd

file_path = 'nodes.csv'

df = pd.read_csv(file_path)

df = df[df['followers'] >= 10000000]
df = df[df['popularity'] >= 60]
filtered_df = df[df['genres'] == '[]'].copy()

df = df.sort_values(by='followers', ascending=False)
df['music_genres'] = df['genres'].apply(lambda x: x.replace('[','').replace(']','').split(','))
df = df.explode('music_genres')
df['music_genres'] = df['music_genres'].str.strip()

genre_counts = df.groupby('music_genres')['spotify_id'].count().reset_index(name='count')
genre_counts_sorted = genre_counts.sort_values('count', ascending=False)

columns_to_drop = ['chart_hits']
df_filtered = df.drop(columns=columns_to_drop).copy()

df_filtered = df_filtered.drop(columns=['music_genres'])

df_filtered = df_filtered.drop_duplicates()

output_file_path = 'filtered_spotify_pj.csv'
df_filtered.to_csv(output_file_path, index=False)

genre_counts_output_file_path = 'genre_counts_sorted.csv'
genre_counts_sorted[['music_genres', 'count']].to_csv(genre_counts_output_file_path, index=False)
```

Using Python's pandas library, I process 'nodes.csv' data. The code below outlines the steps: reading data into 'df' via 'pd.read_csv', filtering by follower and popularity, and isolating artists with music genres. Sorting 'df' by followers, genres are transformed and counted. The count results are grouped by 'music_genres' and named 'count'. The data is sorted by count. Columns are pruned, creating 'df_filtered'. After removing 'music_genres', duplicates are removed. Data is saved in 'filtered_spotify_pj.csv' and 'genre_counts_sorted.csv'.

CLEAN LESS POPULAR SINGERS AND NO MUSIC GENRE

UNIVERSITY OF GREENWICH
Alliance with FPT Education

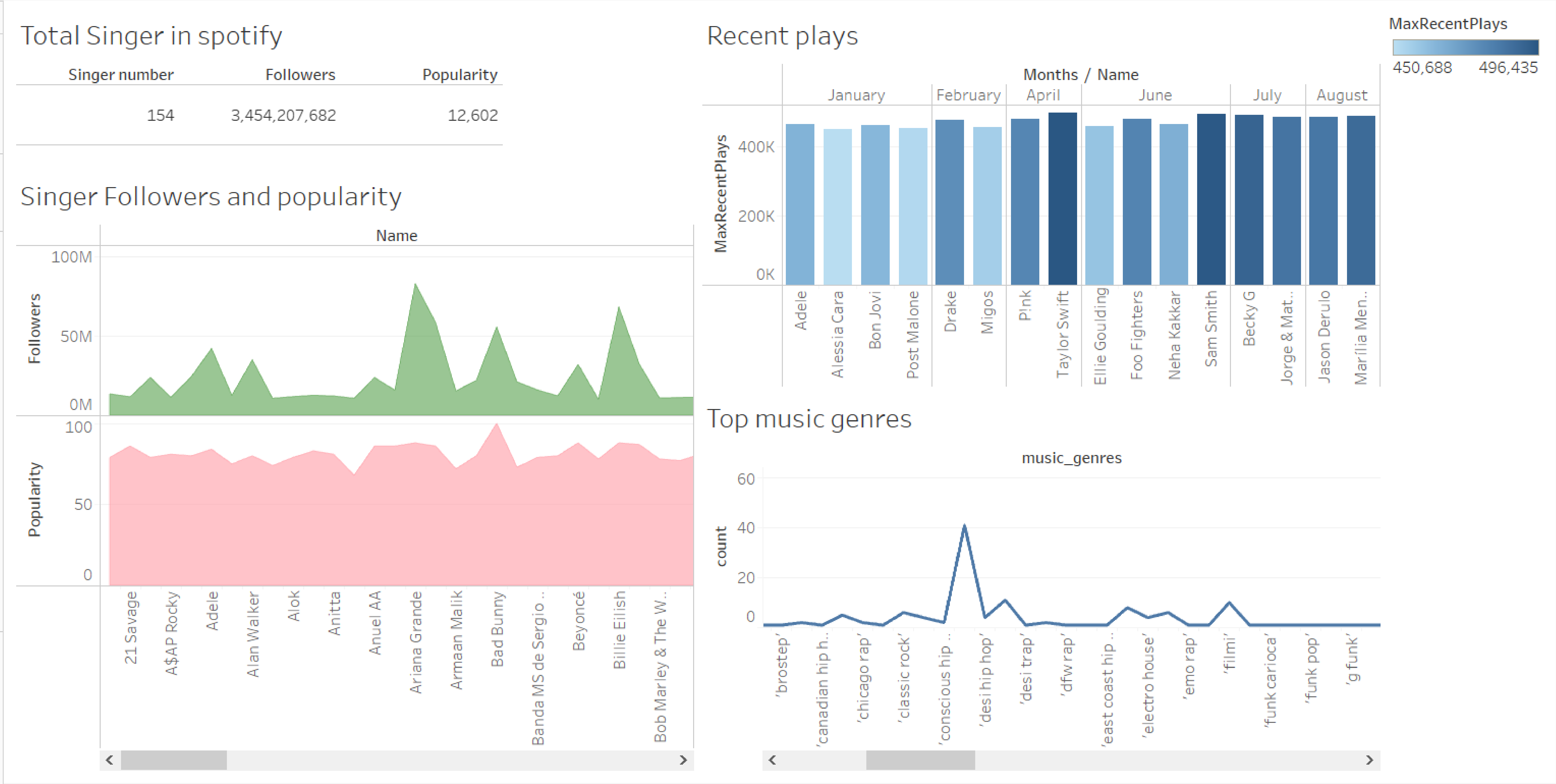
spotify_id	name	followers	popularity	genres	recent_plays	liked_songs
6eUKZXaKkcviH0Ku9w2n3V	Ed Sheeran	102156853	90	['pop', 'uk pop']	10096	192939
66CXWjxzNUsdJxJ2JdwvnR	Ariana Grande	83045090	88	['dance pop', 'pop']	414709	90895
6qqNVTkY8uBg9cP3Jd7DAH	Billie Eilish	68407227	88	['art pop', 'electropop', 'pop']	21150	241230
3TVXtAsR1Inumwj472S9r4	Drake	66852536	95	['canadian hip hop', 'canadian po	476964	195742
1uNFoZAHBGtllmzznpCI3s	Justin Bieber	65590075	90	['canadian pop', 'pop']	245850	105454
7dGJo4pcD2V6oG8kP0tJRR	Eminem	59184634	90	['detroit hip hop', 'hip hop', 'rap']	193172	218620
06HL4z0CvFAxyc27GXpf02	Taylor Swift	58554324	94	['pop']	496435	295715
4YRxDV8wJFPHPTeXepOstw	Arijit Singh	58523986	86	['desi pop', 'filmi', 'modern bollyw	191310	53779
4q3ewBCX7sLwd24euuV69X	Bad Bunny	55669387	100	['reggaeton', 'trap latino']	221376	68294
3Nrfpe0tUJi4K4DXYWgMUX	BTS	54532917	91	['k-pop', 'k-pop boy group']	100404	240809
5pKCCKE2ajJHZ9KAiaK11H	Rihanna	50002769	87	['barbadian pop', 'dance pop', 'po	194826	239688
1Xyo4u8uXC1ZmMpatF05PJ	The Weeknd	49387909	93	['canadian contemporary r&b', 'c	316691	165242
0du5cEVh5yTK9QJze8zA0C	Bruno Mars	43103624	87	['dance pop', 'pop']	155907	86000
1dfeR4HaWDbWqFHLkxsg1d	Queen	42644068	84	['classic rock', 'glam rock', 'rock']	438118	166667
4dpARuHxo51G3z768sgnrY	Adele	42174621	84	['british soul', 'pop', 'pop soul', 'uk	464753	134090
53XhwfbYqKCa1cC15pYq2q	Imagine Dragons	41751320	88	['modern rock', 'rock']	186589	119172
7n2wHs1TKAczGzO7Dd2rGr	Shawn Mendes	39186962	84	['canadian pop', 'dance pop', 'pop	434498	238866
4gzpq5DPGxSnKTe4SA8HAU	Coldplay	38371125	88	['permanent wave', 'pop']	42958	217073
246dkjvS1zLTtiykXe5h60	Post Malone	38085069	89	['dfw rap', 'melodic rap', 'rap']	452504	243593
15UsOTVnJzReFVN1VCnxy4	XXXTENTACION	37935983	87	['emo rap', 'miami hip hop']	81930	153426
04gDigrS5kc9YWfZHwBETP	Maroon 5	37808723	85	['pop']	216683	289779
5f4QpKfy7ptCHwTqspnSJI	Neha Kakkar	37695348	74	['desi pop', 'filmi', 'modern bollyw	466019	90671
0C8ZW7ezQVs4URX5aX7Kqx	Selena Gomez	37269207	82	['dance pop', 'pop', 'post-teen po	405436	125038
6M2wZ9GZgrQXHCFfjv46we	Dua Lipa	36163788	88	['dance pop', 'pop', 'uk pop']	122427	239989
1vyhD5VmyZ7KMfW5gqLgo5	J Balvin	35224872	88	['reggaeton', 'reggaeton colombia	214799	285057
7vk5e3vY1uw9plTHJAMwjN	Alan Walker	35053203	80	['electro house']	94247	272813
64KEffDW9EtZ1y2vBYgq8T	Marshmello	34229273	83	['brostep', 'dance pop', 'edm', 'po	242179	287646
1i8SpTcr7yvPOmcqrbnVXY	Ozuna	33511653	86	['puerto rican pop', 'reggaeton', 't	154049	116642

	A	B
1	music_genres	count
2	'pop'	60
3	'dance pop'	41
4	'rap'	20
5	'uk pop'	16
6	'rock'	16
7	'post-teen pop'	12
8	'trap latino'	12
9	'reggaeton'	12
10	'modern bollywood'	11
11	'desi pop'	11
12	'hip hop'	11
13	'filmi'	10
14	'trap'	9
15	'pop rap'	8
16	'edm'	8
17	'arrocha'	7
18	'permanent wave'	7
19	'sertanejo universitario'	6
20	'modern rock'	6
21	'classic rock'	6
22	'pop dance'	6
23	'electropop'	6
24	'reggaeton colombiano'	6
25	'canadian pop'	5
26	'hard rock'	5
27	'sertanejo'	5
28	'latin pop'	5
29	'r&b'	4

DASHBOARD PURPOSE

A dashboard that helps Spotify and its customers see how many singers work with Spotify, number of followers and popularity, as well as the most used music genres and the most listened artists right now

DASHBOARD


Total Singer in spotify
Singer number
Followers
Popularity
154
3,454,207,682
12,602
Singer Followers and popularity
Name
Followers
100M
50M
0M
Popularity
100
50
0
21 Savage
A$AP Rocky
Adele
Alan Walker
Alok
Anitta
Anuel AA
Ariana Grande
Armaan Malik
Bad Bunny
Banda MS de Sergio ..
Beyoncé
Billie Eilish
Bob Marley & The W..
Recent plays
Months / Name
January
February
April
June
July
August
MaxRecentPlays
400K
200K
0K
Adele
Alessia Cara
Bon Jovi
Post Malone
Drake
Migos
P!nk
Taylor Swift
Ellie Goulding
Foo Fighters
Neha Kakkar
Sam Smith
Becky G
Jorge & Mat..
Jason Derulo
Marília Men..
MaxRecentPlays
450,688
496,435
Top music genres
music_genres
count
60
40
20
0
'brostep'
'canadian hip h..
'chicago rap'
'classic rock'
'conscious hip ..
'desi hip hop'
'desi trap'
'dfw rap'
'east coast hip ..
'electro house'
'emo rap'
'filmi'
'funk carioca'
'funk pop'
'g funk'

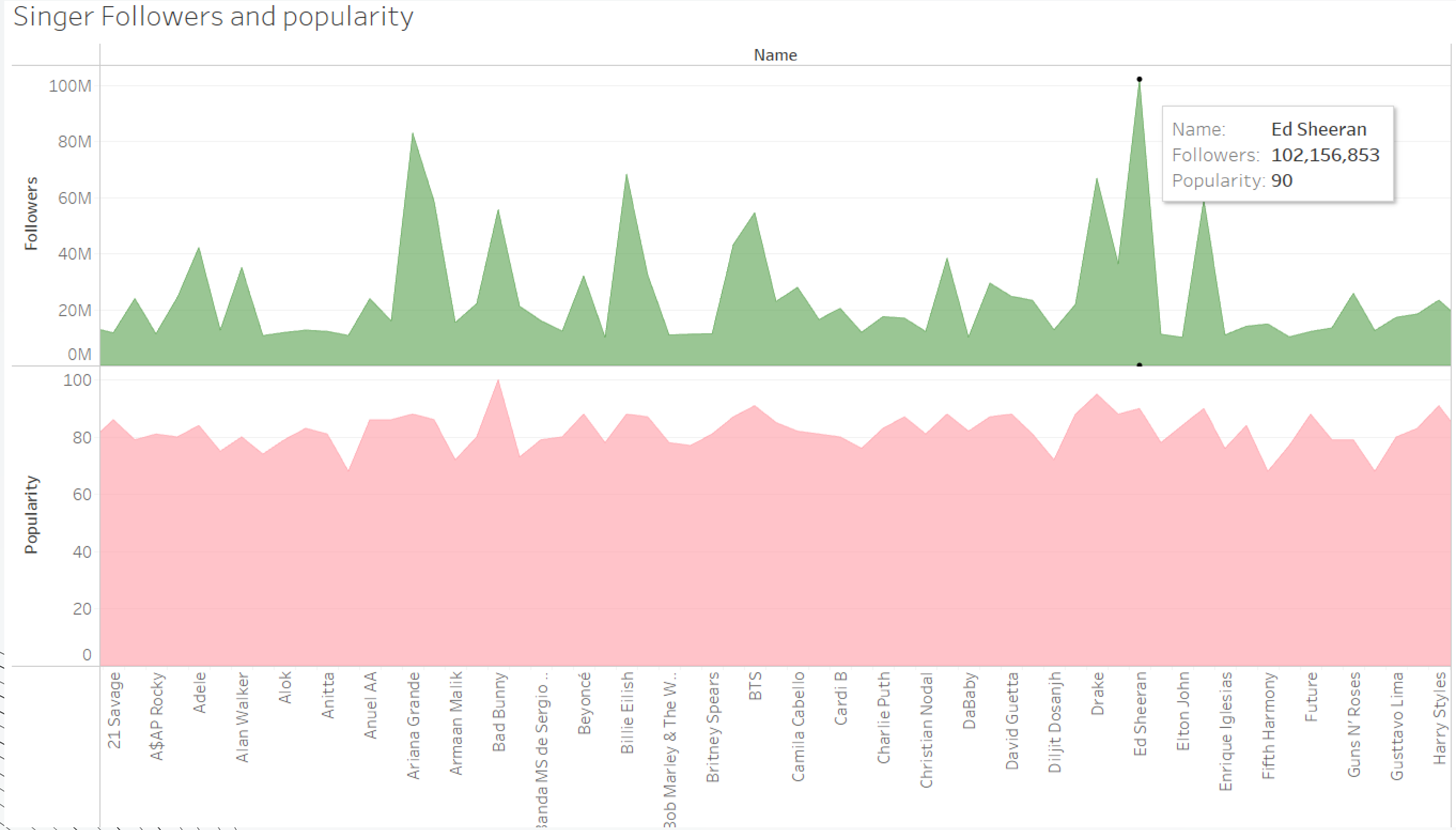
CHART

Total Singer in spotify

Singer number	Followers	Popularity
154	3,454,207,682	12,602

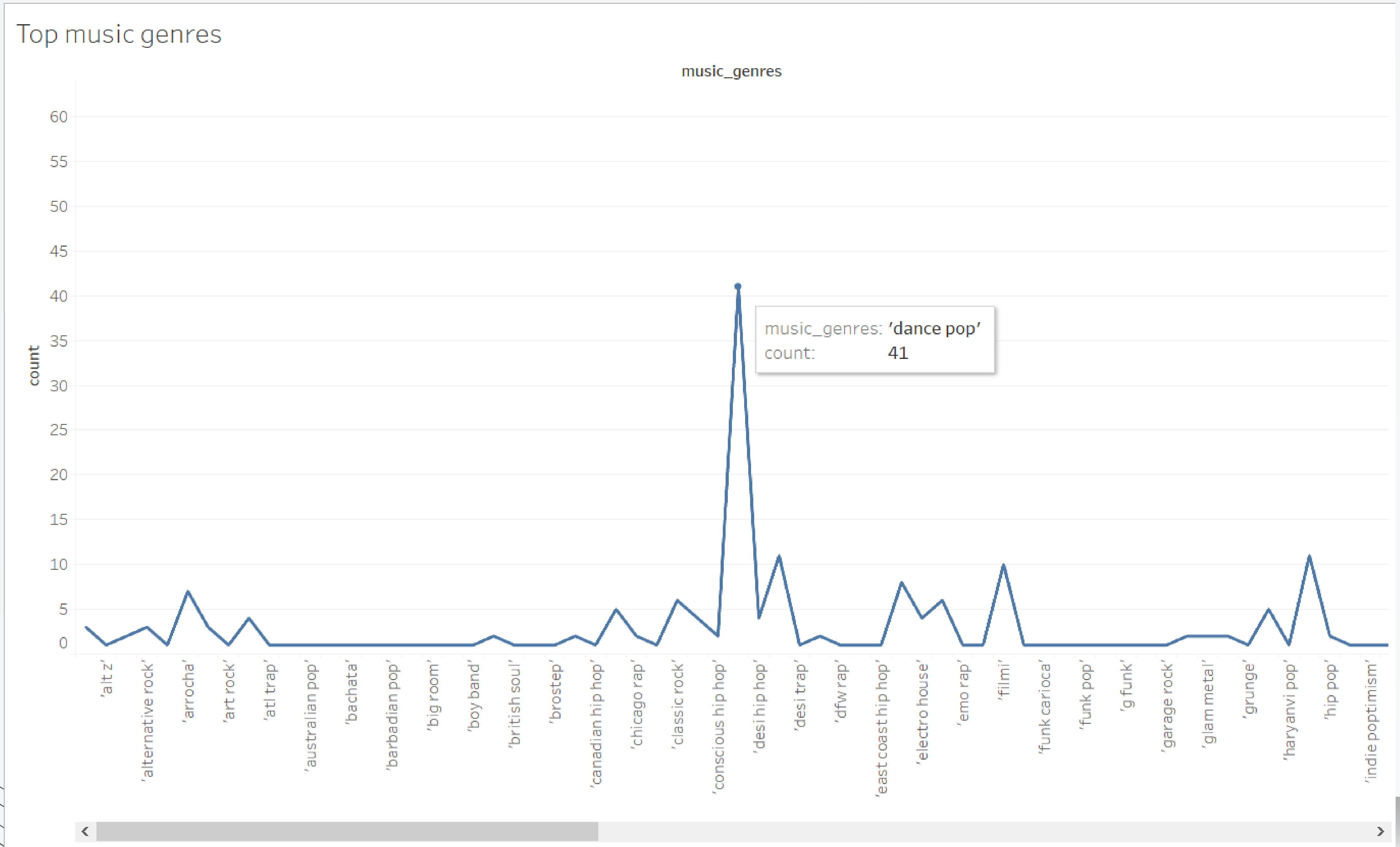
Here is a table where I have statistics about the total number of current singers who have an account with spotify as well as the total number of followers and popularity

CHART



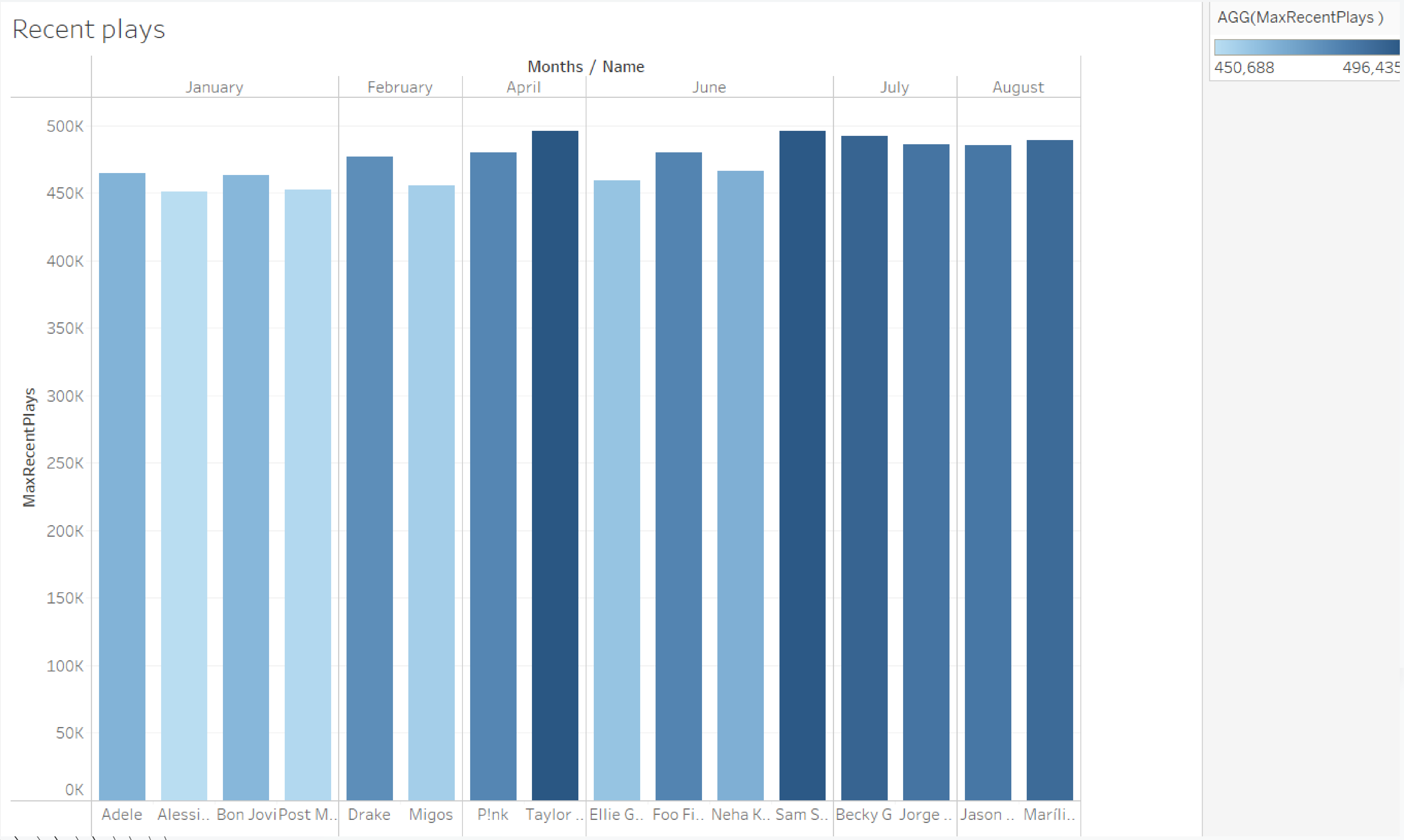
This is a table showing but people with their number of followers and popularity like Ed sheeran have nearly 103 million followers and up to 90% of users know him

CHART



In this chart, I have calculated and filtered out the genres of music used by singers to know which music will be popular example dance pop have 41 singer use this music genre.

CHART



In this chart I will show the number of people listening to songs from January to August 2023 and when I spotify analyze this data and show it to the client the client can see that in the last months through singers with high listenership means that they are popular that month and have a high reputation

PRE-PROCESS DATA

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	Age	Gender	y_usage_p	_listening_	subscriptio	n_sub_wil	ed_premiu	_listening	_music_ge	sic_time_s	Influencial	c_lis_frequ	c_expl_me	ic_recc_ra	_lis_freque	_pod_ger	red_pod_f	host_prefe	ed_pod_d	riety_satis	t_creation	_activity_
2	20-35	Female	More than	Smart spea	Free (ad-su	Yes	Family Plan	Podcast	Melody	Night	Sadness or	leisure tim	Playlists	3	Daily	Comedy	Interview	Both	Both	Ok	########	########
3	12-20	Male	More than	Computer	Free (ad-su	Yes	Individual	Podcast	Rap	Afternoon	Social gath	Workout s	Playlists	2	Several tin	Comedy	Interview	Both	None	Satisfied	########	########
4	35-60	Others	6 months t	Smart spea	Free (ad-su	Yes	Student Pla	Podcast	Pop	Night	Relaxation	Study Hou	Playlists	4	Once a we	Sports	Interview	None	Both	Satisfied	########	########
5	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	Office hou	recommer	4	Never	None	None	None	None	Ok	########	########
6	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	leisure tim	recommer	4	Rarely	Lifestyle a	Story tellin	Well know	Both	Ok	########	########
7	20-35	Male	More than	Smartphor	Free (ad-su	No	None	Music	Pop	Night	Uplifting a	Workout s	Others	3	Never	None	None	None	None	Ok	########	########
8	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Morning	Relaxation	Office hou	recommer	3	Never	None	None	None	None	Ok	########	########
9	20-35	Female	Less than 6	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Social gath	leisure tim	recommer	5	Several tin	Lifestyle a	Conversati	Both	Longer	Satisfied	########	########
10	20-35	Female	Less than 6	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Afternoon	Relaxation	While Trav	Playlists, R	4	Rarely	Comedy	Story tellin	Well know	Shorter	Satisfied	########	########
11	20-35	Female	More than	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	Office hou	recommer	4	Never	None	None	None	None	Ok	########	########
12	20-35	Female	More than	Smartphor	Free (ad-su	No	None	Music	Melody	Afternoon	Relaxation	While Trav	Playlists	3	Several tin	Lifestyle a	Story tellin	Both	Both	Satisfied	########	########
13	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	While Trav	recommer	3	Several tin	Comedy	Interview	Both	Shorter	Ok	########	########
14	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Morning	Relaxation	Office hou	Playlists	3	Rarely	Lifestyle a	Story tellin	Well know	Shorter	Ok	########	########
15	20-35	Female	More than	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	Office hou	recommer	1	Rarely	Comedy	Story tellin	Well know	Shorter	Ok	########	########
16	20-35	Female	More than	Smartphor	Free (ad-su	Yes	None	Music	Pop	Morning	Relaxation	Office hou	recommer	5	Rarely	Lifestyle a	Educationa	Well know	Shorter	Very Satisf	########	########
17	20-35	Female	More than	Smartphor	Free (ad-su	No	None	Music	Pop	Afternoon	Relaxation	Office hou	Others	2	Never	None	None	None	None	Very Dissa	########	########
18	20-35	Female	Less than 6	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Afternoon	Relaxation	While Trav	Playlists, R	1	Daily	Comedy	Educationa	Well know	Longer	Satisfied	########	########
19	20-35	Female	Less than 6	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Social gath	Social gath	recommer	4	Never	None	None	None	None	Ok	########	########
20	20-35	Female	More than	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	leisure tim	recommer	3	Never	None	None	None	None	Ok	########	########
21	20-35	Female	1 year to 2	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Night	Sadness or	While Trav	recommer	4	Never	None	None	None	Shorter	Ok	########	########
22	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	While Trav	recommer	4	Never	None	None	None	None	Ok	########	########
23	20-35	Male	More than	Smartphor	Free (ad-su	Yes	Individual	Music	Classical &	Morning	Relaxation	Office hou	recommer	5	Never	None	None	None	None	Ok	########	########
24	20-35	Female	Less than 6	Smartphor	Free (ad-su	No	None	Music	Melody	Night	Relaxation	While Trav	recommer	4	Once a we	Comedy	Story tellin	Well know	Shorter	Satisfied	########	########
25	20-35	Female	More than	Smartphor	Premium (	Yes	Individual	Music	classical	Night	Relaxation	While Trav	recommer	1	Rarely	Comedy	Interview	Both	Shorter	Very Satisf	########	########
26	20-35	Male	More than	Smartphor	Premium (	Yes	Family Plan	Music	Rock	Night	Relaxation	Study Hou	recommer	3	Rarely	Lifestyle a	Conversati	Both	Both	Ok	########	########
27	20-35	Female	1 year to 2	Smartphor	Free (ad-su	No	None	Music	Pop	Night	Relaxation	While Trav	Others	4	Never	None	None	None	None	Ok	########	########
28	12-20	Male	6 months t	Smartphor	Premium (	Yes	Student Pla	Podcast	Rap	Morning	Relaxation	Office hou	recommer	4	Several tin	Lifestyle a	Story tellin	Both	Longer	Satisfied	########	########
29	35-60	Male	Less than 6	Smartphor	Free (ad-su	No	None	Music	Old songs	Night	Relaxation	leisure tim	recommer	1	Never	None	None	None	None	Very Dissa	########	########
30	20-35	Female	Less than 6	Smartphor	Free (ad-su	No	None	Music	Pop	Night	Relaxation	Office hou	recommer	4	Never	None	None	None	None	Ok	########	########
31	20-35	Female	6 months t	Smartphor	Free (ad-su	No	None	Podcast	Melody	Night	Relaxation	Study Hou	recommer	3	Daily	Comedy	Conversati	Both	Longer	Satisfied	########	########

CODE EXPLAIN

clean_data 1 | Spotify_data.xlsx

clean_data > ...

```python
import pandas as pd

df = pd.read_excel("Spotify_data.xlsx")

df = df[~(df == "None").any(axis=1)]

df = df[df['spotify_subscription_plan'] != "Premium (paid subscription)"]

df = df[df['account_status'] == "Active"]

df['spotify_listening_device_filter'] = df['spotify_listening_device'].apply(lambda x: x.split(', '))
df = df.explode('spotify_listening_device_filter')
df['spotify_listening_device_filter'] = df['spotify_listening_device_filter'].str.strip()

genre_counts = df.groupby('spotify_listening_device_filter')['Age'].count().reset_index(name='count')

columns_to_drop = ['preffered_premium_plan','music_lis_frequency','music_expl_method','music_recc_rating','churn_reason']
df = df.drop(columns=columns_to_drop)

df.to_excel("Clean_Updated_Spotify_data.xlsx", index=False)

genre_counts_output_file_path = 'device_count.csv'
genre_counts[['spotify_listening_device_filter', 'count']].to_csv(genre_counts_output_file_path, index=False)
```

CODE EXPLAIN

Filter for less popular singers and no music genre

1. Imports pandas library and reads in an Excel file into a DataFrame (df)Removes any rows where all values are "None".
2. Filters the df to only include rows where spotify_subscription_plan is not Premium (paid subscription).
3. Filters df to only include rows where account_status is Active.
4. Splits the spotify_listening_device column on ',' to expand each device into separate rows.
5. Strips whitespace from the spotify_listening_device values.
6. Groups by device and counts the number of occurrences, storing the result in genre_counts.
7. Drops some unnecessary columns from df.
8. Writes the cleaned df Excel file as Clean_Updated_Spotify_data.xlsx.
9. Writes the device counts to a CSV file called device_count.csv.

CLEAN PREMIUM USERS AND SOME UNUSED COLUMNS

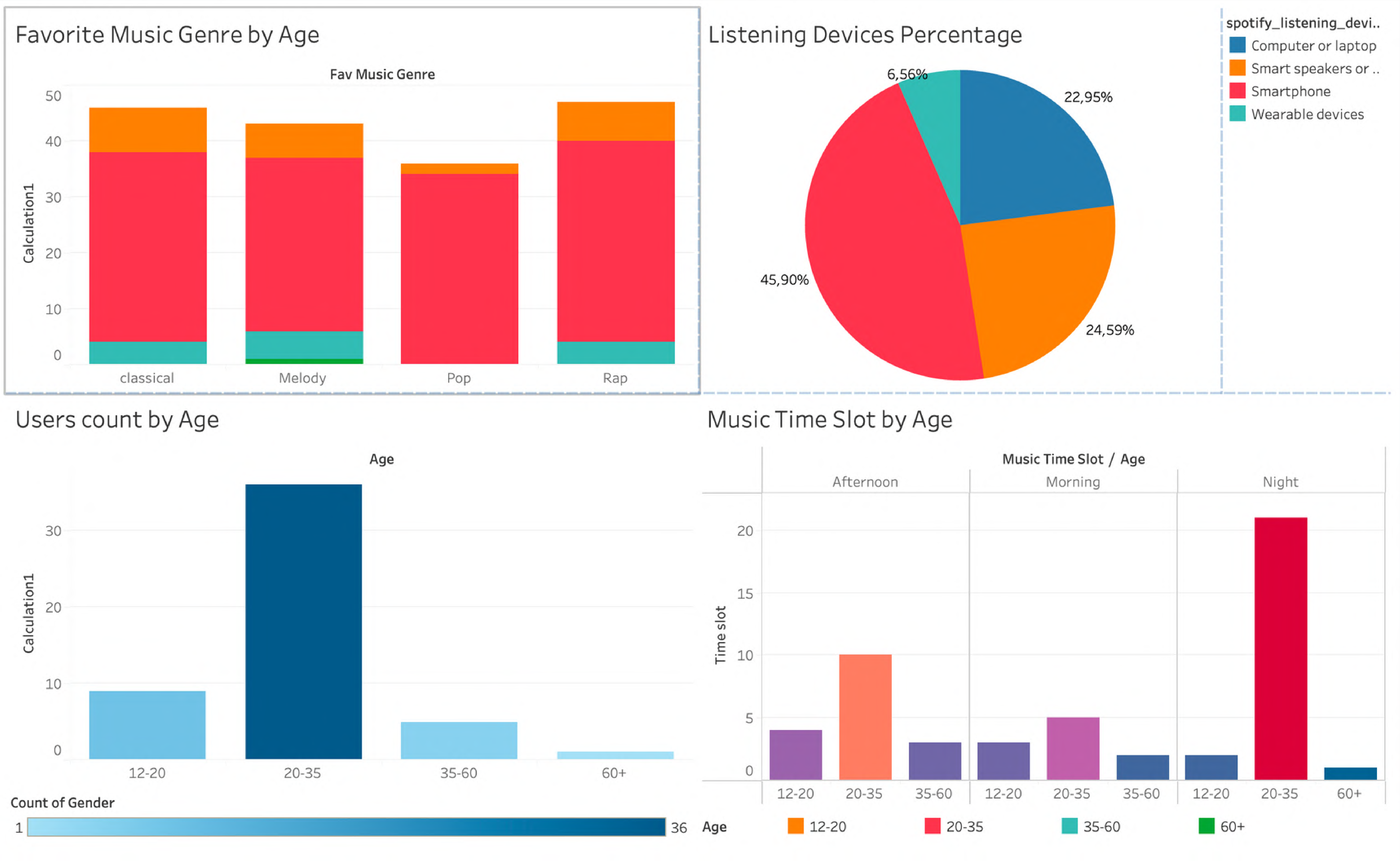
UNIVERSITY OF GREENWICH
Alliance with FPT Education

	A	B	C	D	E	F	G	H	I
1	Age	Gender	spotify_usage_period	spotify_listening_device	spotify_subscription_plan	premium_sub_willingness	preferred_listening_content	fav_music_genre	music_time_
2	20-35	Female	More than 2 years	Smart speakers or voice ass	Free (ad-supported)	Yes	Podcast	Melody	Night
3	20-35	Female	More than 2 years	Smartphone, Computer or la	Free (ad-supported)	Yes	Music	Melody	Night
4	35-60	Male	1 year to 2 years	Smartphone, Smart speaker	Free (ad-supported)	Yes	Podcast	Melody	Morning
5	12-20	Female	Less than 6 months	Smartphone	Free (ad-supported)	Yes	Podcast	Rap	Morning
6	20-35	Male	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Podcast	Rap	Morning
7	20-35	Female	1 year to 2 years	Computer or laptop	Free (ad-supported)	Yes	Music	Melody	Night
8	12-20	Female	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Podcast	classical	Afternoon
9	20-35	Female	More than 2 years	Computer or laptop, Smart s	Free (ad-supported)	Yes	Podcast	classical	Morning
10	35-60	Female	1 year to 2 years	Smart speakers or voice ass	Free (ad-supported)	Yes	Music	classical	Afternoon
11	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	Yes	Music	Melody	Night
12	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	No	Podcast	Melody	Afternoon
13	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	Music	classical	Night
14	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	Yes	Music	Melody	Morning
15	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	Yes	Music	Pop	Night
16	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	Music	Pop	Afternoon
17	20-35	Female	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Podcast	classical	Afternoon
18	20-35	Female	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Podcast	classical	Afternoon
19	20-35	Female	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Music	Rap	Morning
20	12-20	Female	6 months to 1 year	Computer or laptop	Free (ad-supported)	Yes	Podcast	Rap	Morning
21	12-20	Female	6 months to 1 year	Smart speakers or voice ass	Free (ad-supported)	No	Podcast	classical	Afternoon
22	35-60	Female	1 year to 2 years	Computer or laptop, Smart s	Free (ad-supported)	No	Music	classical	Morning
23	35-60	Female	6 months to 1 year	Smart speakers or voice ass	Free (ad-supported)	No	Podcast	Melody	Afternoon
24	20-35	Female	1 year to 2 years	Smart speakers or voice ass	Free (ad-supported)	Yes	Podcast	classical	Afternoon
25	12-20	Female	6 months to 1 year	Smart speakers or voice ass	Free (ad-supported)	No	Podcast	classical	Afternoon
26	20-35	Female	6 months to 1 year	Smartphone	Free (ad-supported)	Yes	Music	Melody	Night
27	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	Music	Melody	Night
28	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	Music	classical	Night
29	20-35	Female	More than 2 years	Smartphone, Computer or la	Free (ad-supported)	No	Music	Pop	Night
30	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	No	Podcast	Melody	Night
31	60+	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	Music	Melody	Night

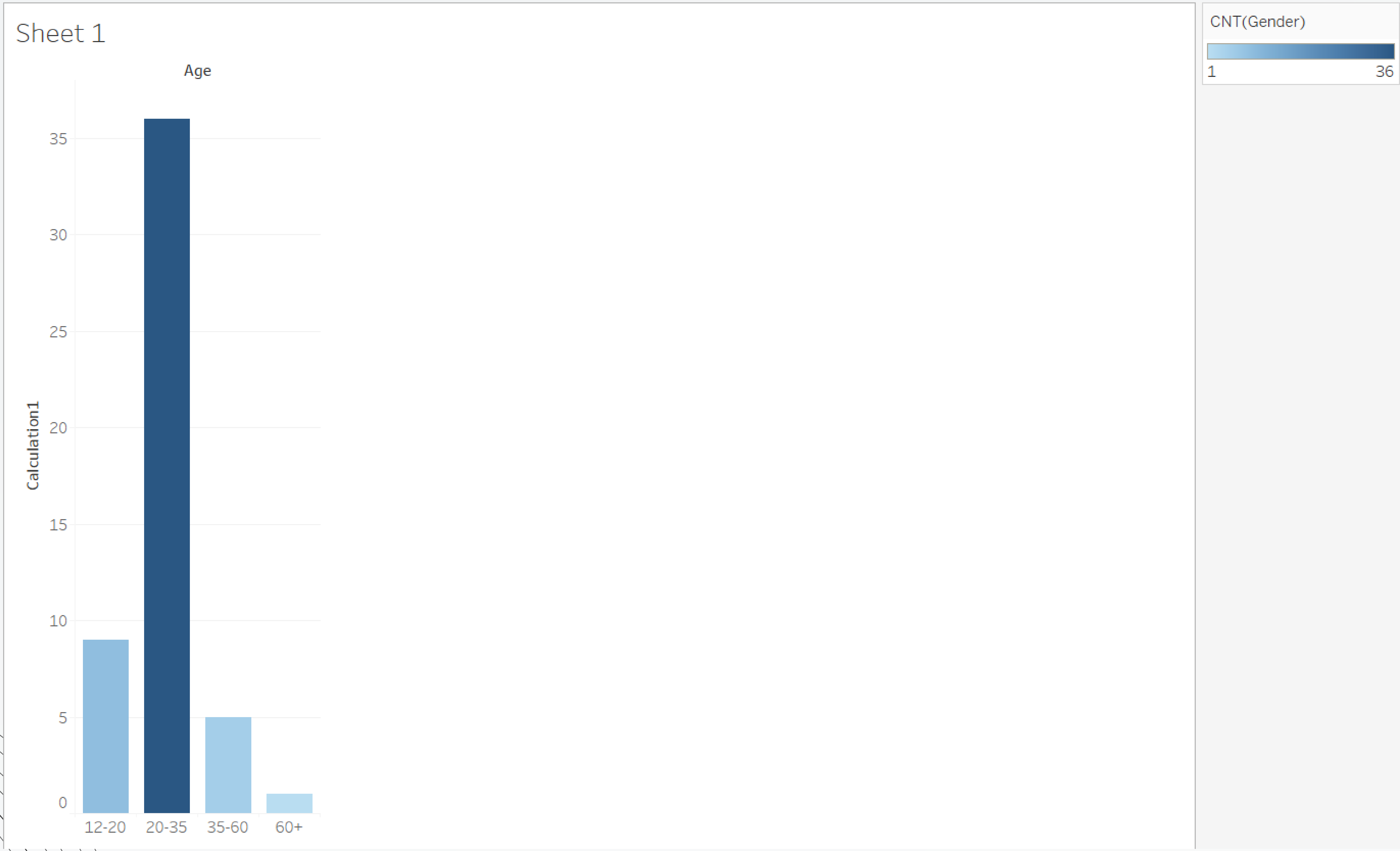
DASHBOARD PURPOSE

Dashboard provides data on user insights such as the number of users by age, active time of each age, favorite music genre by age and device used to help advertisers deliver content to the proper target

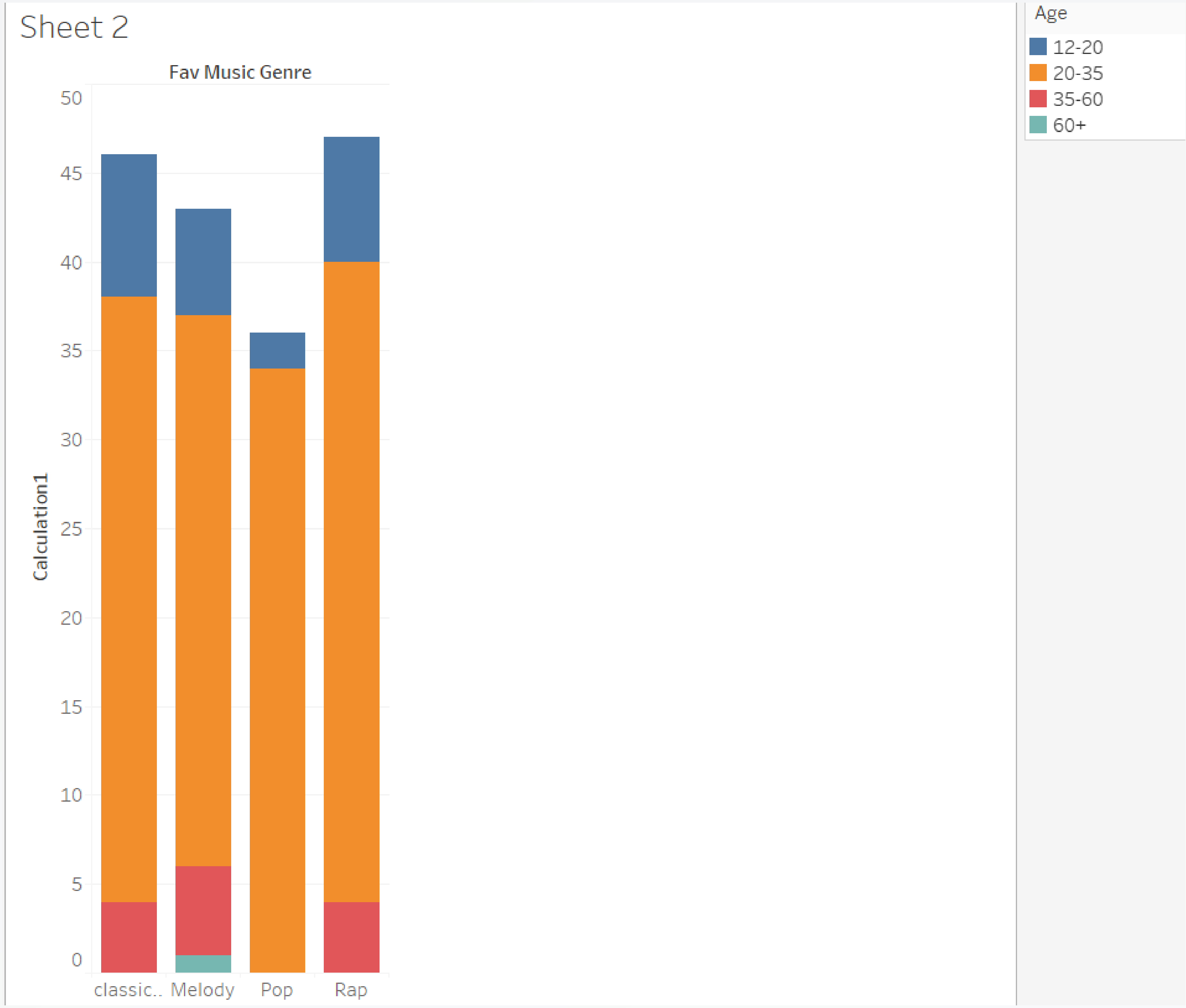
DASHBOARD


Favorite Music Genre by Age
Fav Music Genre
Calculation1
50
40
30
20
10
0
classical
Melody
Pop
Rap
Listening Devices Percentage
6,56%
22,95%
45,90%
24,59%
spotify_listening_devi..
Computer or laptop
Smart speakers or ..
Smartphone
Wearable devices
Users count by Age
Age
Calculation1
30
20
10
0
12-20
20-35
35-60
60+
Count of Gender
1
36
Music Time Slot by Age
Music Time Slot / Age
Afternoon
Morning
Night
Time slot
20
15
10
5
0
12-20
20-35
35-60
12-20
20-35
35-60
12-20
20-35
60+
Age
12-20
20-35
35-60
60+

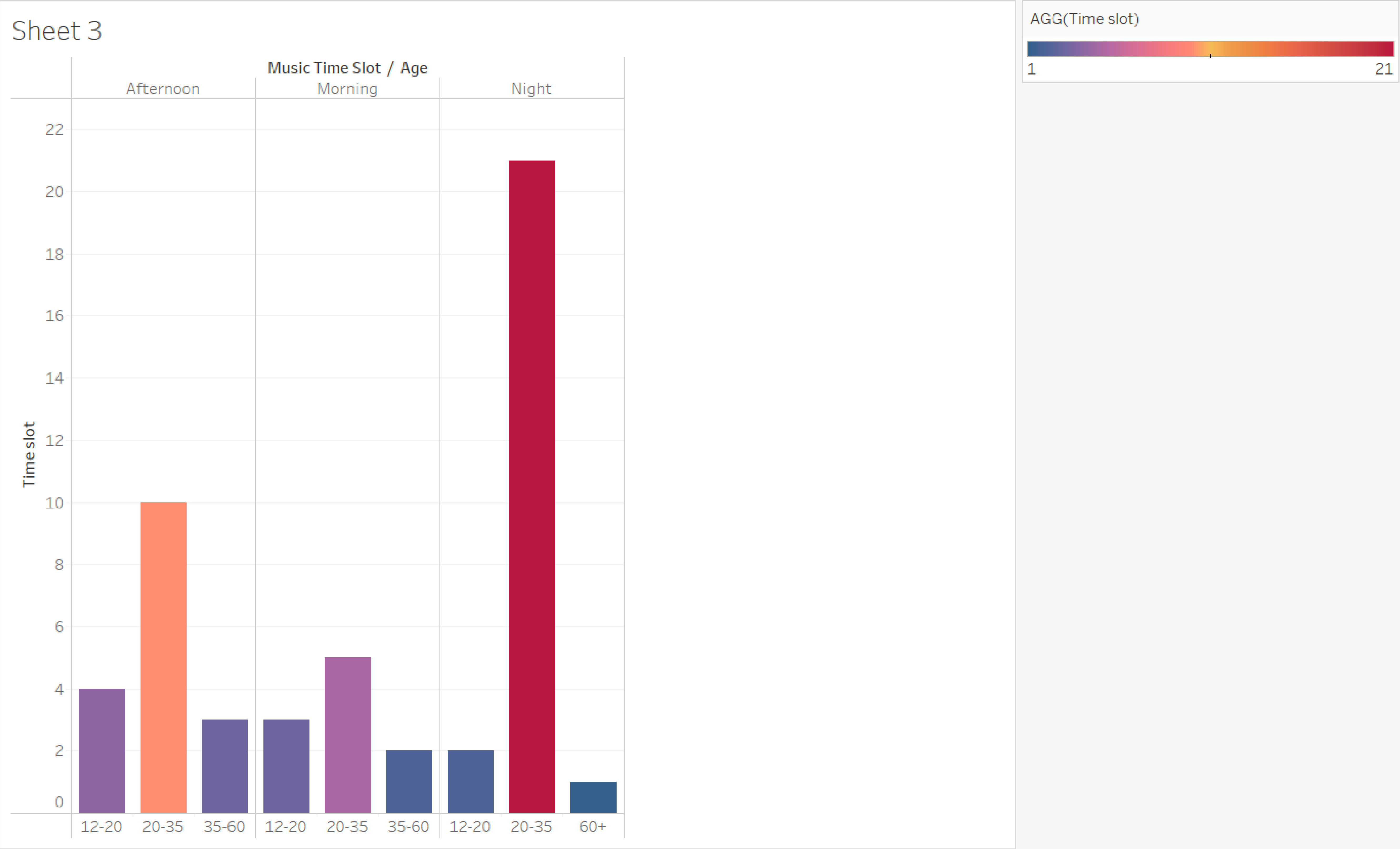
USERS COUNT BY AGE



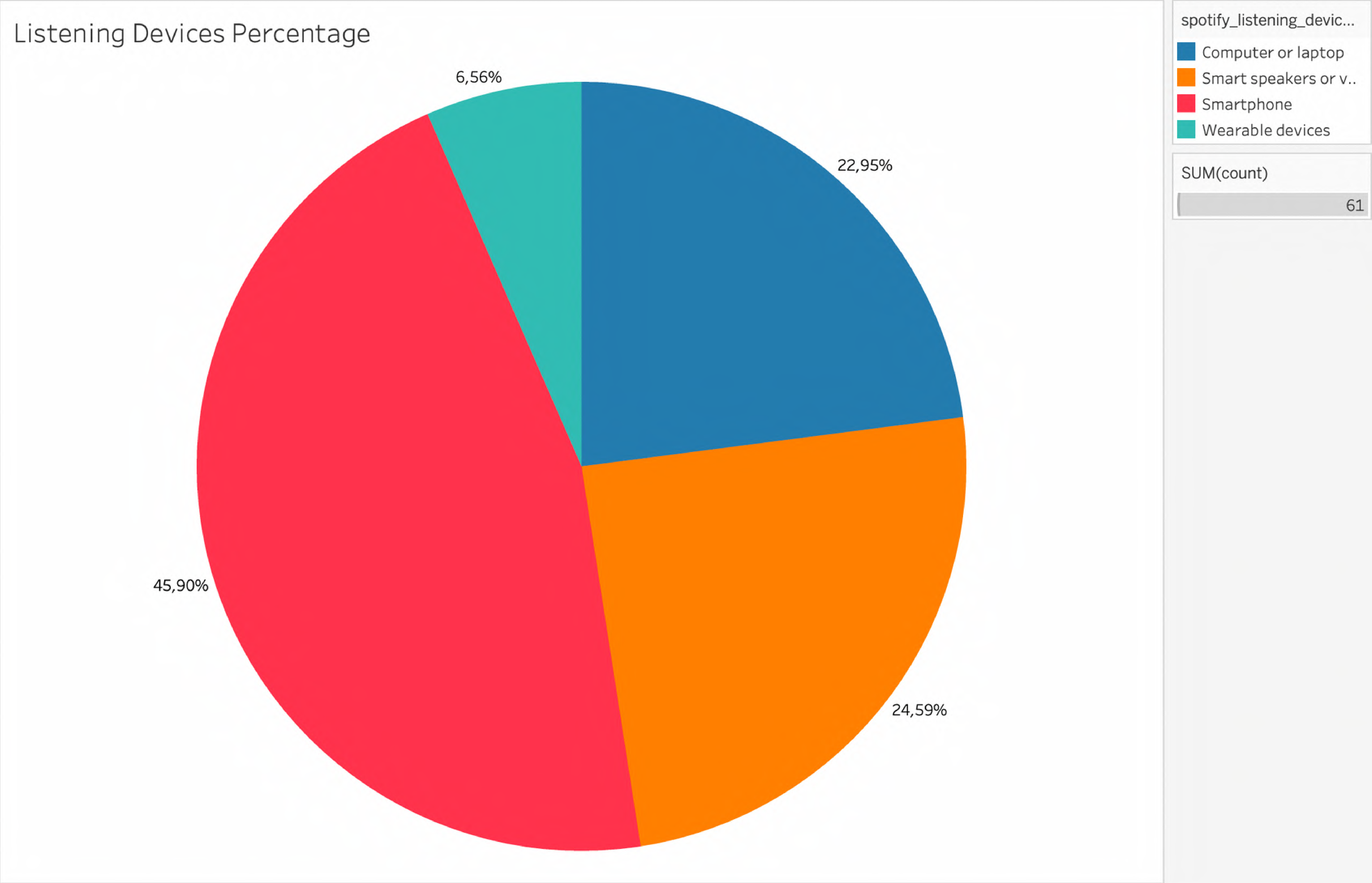
FAVORITE MUSIC GENRE BY AGE



MUSIC TIME SLOT BY AGE



LISTENING DEVICES PERCENTAGE



UNCLEAR DATA ABOUT USERS ON SPOTIFY SYSTEM

	A	B	C	D	E	F	G	H	I	J	
1	Age	Gender	spotify_usage_period	spotify_listening_device	spotify_subscription_plan	premium_	preffered_premium_plan	preferred_listenir	fav_music_genre	music_time_slot	music
2	20-35	Female	More than 2 years	Smart speakers or voice assistants	Free (ad-supported)	Yes	Family Plan-Rs 179/month	Podcast	Melody	Night	Sadne
3	12-20	Male	More than 2 years	Computer or laptop	Free (ad-supported)	Yes	Individual Plan- Rs 119/ month	Podcast	Rap	Afternoon	Socia
4	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	Relax
5	20-35	Female	1 year to 2 years	Smartphone, Smart speakers or voice assistants	Free (ad-supported)	No	None	Music	Melody	Night	Relax
6	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
7	20-35	Male	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Pop	Night	Uplift
8	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Morning	Relax
9	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Socia
10	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	Yes	Individual Plan- Rs 119/ month	Music	Melody	Afternoon	Relax
11	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
12	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Afternoon	Relax
13	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
14	20-35	Female	1 year to 2 years	Smartphone, Computer or laptop	Free (ad-supported)	No	None	Music	Melody	Morning	Relax
15	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
16	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	Yes	None	Music	Pop	Morning	Relax
17	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Pop	Afternoon	Relax
18	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	Yes	Individual Plan- Rs 119/ month	Music	Melody	Afternoon	Relax
19	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Socia
20	20-35	Female	More than 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
21	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	Yes	Individual Plan- Rs 119/ month	Music	Melody	Night	Sadne
22	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
23	20-35	Male	More than 2 years	Smartphone	Free (ad-supported)	Yes	Individual Plan- Rs 119/ month	Music	Classical & melody, danc	Morning	Relax
24	20-35	Female	Less than 6 months	Smartphone	Free (ad-supported)	No	None	Music	Melody	Night	Relax
25	20-35	Female	More than 2 years	Smartphone, Computer or laptop, Smart speakers	Premium (paid subscription)	Yes	Individual Plan- Rs 119/ month	Music	classical	Night	Relax
26	20-35	Male	More than 2 years	Smartphone, Computer or laptop	Premium (paid subscription)	Yes	Family Plan-Rs 179/month	Music	Rock	Night	Relax
27	20-35	Female	1 year to 2 years	Smartphone	Free (ad-supported)	No	None	Music	Pop	Night	Relax
28	12-20	Male	6 months to 1 year	Smartphone, Smart speakers or voice assistants	Premium (paid subscription)	Yes	Student Plan-Rs 59/month	Podcast	Rap	Morning	Relax

CODE EXPLAIN

Filter for less popular singers and no music genre

```
import pandas as pd

# Load the data from the Excel file
df = pd.read_excel("Updated_Spotify_data.xlsx")

# # Filter out rows with the word "None"
# df = df[~df.applymap(lambda x: x == "None" if isinstance(x, str) else False).any(axis=1)]

# Filter out rows with the word "None"
df = df[~(df == "None").any(axis=1)]

# Filter out rows with "Others" in the Gender column
df = df[df['Gender'] != "Others"]

# # Extract the month from 'account_creation_date' and 'last_activity_date' columns and convert to numeric format
# df['account_creation_date'] = df['account_creation_date'].astype(str).str.split('-').str[1].astype(int)
# df['last_activity_date'] = df['last_activity_date'].astype(str).str.split('-').str[1].astype(int)

# Extract the day from 'account_creation_date' and 'last_activity_date' columns and convert to numeric format
df['account_creation_date'] = df['account_creation_date'].astype(str).str.split('-').str[2].astype(int)
df['last_activity_date'] = df['last_activity_date'].astype(str).str.split('-').str[2].astype(int)

# Export the cleaned data to an Excel file
df.to_excel("Cleann_Updated_Spotify_data.xlsx", index=False)
```

The provided Python code is designed to preprocess and clean data from a Spotify dataset. Initially, it reads data from an Excel file named "Updated_Spotify_data.xlsx". The primary cleaning steps are:

1. Removal of Incomplete Data: Entire rows containing the placeholder "None" are eliminated. This step ensures that any missing or incomplete data represented by the term "None" does not impact subsequent analyses.
2. Gender Filtering: Any rows with the gender labeled as "Others" are filtered out to focus the dataset solely on "Male" and "Female" categories.
3. Date Transformation: For both the 'account_creation_date' and 'last_activity_date' columns, the day component of the date is extracted. This allows for more granular time-based analysis, concentrating on the specific day of the month when an action occurred.

After these transformations, the cleaned data is saved to a new Excel file named "Clean_Updated_Spotify_data_further_simplified.xlsx". This refined dataset is now primed for further analysis, visualization, or integration into tools like Tableau.

DATA AFTER BEING FILTERED THROUGH PYTHON CODE

UNIVERSITY OF GREENWICH
Alliance with FPT Education

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Age	Gender	y_usage_p	_listening_	subscriptio	m_sub_wil	d_premiu	_listening_	_music_ge	sic_time_s	Influencial	c_lis_frequ	c_expl_me	ic_recc_ra	_lis_freque	_pod_ger	red_pod_f	host_prefe	ed_pod_d	riety_satis	nt_creatior	_activity_c	playlists_cr
2	20-35	Female	More than	Smart spea	Free (ad-su	Yes	Family Plar	Podcast	Melody	Night	Sadness or	leisure tim	Playlists	3	Daily	Comedy	Interview	Both	Both	Ok	2	5	24
3	20-35	Female	Less than 6	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Afternoon	Relaxation	While Trav	Playlists, R	4	Rarely	Comedy	Story tellin	Well know	Shorter	Satisfied	18	1	47
4	20-35	Female	Less than 6	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Afternoon	Relaxation	While Trav	Playlists, R	1	Daily	Comedy	Educationa	Well know	Longer	Satisfied	2	9	9
5	20-35	Female	More than	Smartphor	Premium (	Yes	Individual	Music	classical	Night	Relaxation	While Trav	recommer	1	Rarely	Comedy	Interview	Both	Shorter	Very Satisf	7	11	44
6	20-35	Male	More than	Smartphor	Premium (	Yes	Family Plar	Music	Rock	Night	Relaxation	Study Hour	recommer	3	Rarely	Lifestyle a	Conversati	Both	Both	Ok	12	29	49
7	12-20	Male	6 months t	Smartphor	Premium (	Yes	Student Pla	Podcast	Rap	Morning	Relaxation	Office hou	recommer	4	Several tin	Lifestyle a	Story tellin	Both	Longer	Satisfied	20	9	42
8	20-35	Female	More than	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Night	Relaxation	While Trav	recommer	3	Rarely	Comedy	Conversati	Well know	Shorter	Satisfied	18	25	40
9	12-20	Female	More than	Smartphor	Free (ad-su	Yes	Student Pla	Music	Pop	Night	Uplifting a	Study Hour	recommer	4	Once a we	Comedy	Conversati	Both	Longer	Very Satisf	10	1	21
10	35-60	Male	1 year to 2	Smartphor	Free (ad-su	Yes	Student Pla	Podcast	Melody	Morning	Sadness or	While Trav	recommer	4	Once a we	Food and c	Interview	Both	Shorter	Satisfied	7	17	49
11	20-35	Male	1 year to 2	Smartphor	Premium (	Yes	Individual	Music	Electronic,	Night	Relaxation	While Trav	recommer	4	Rarely	Comedy	Story tellin	unknown F	Shorter	Very Satisf	14	11	15
12	20-35	Male	More than	Smartphor	Premium (	Yes	Individual	Music	Melody	Night	Relaxation	While Trav	recommer	5	Several tin	Comedy	Conversati	Both	Both	Very Satisf	2	7	0
13	20-35	Male	More than	Smartphor	Premium (	Yes	Individual	Music	Melody	Night	Relaxation	Office hou	recommer	4	Rarely	Comedy	Story tellin	Both	Shorter	Satisfied	27	30	32
14	20-35	Female	More than	Wearable	Free (ad-su	Yes	Duo plan-	Podcast	Rap	Night	Sadness or	Study Hour	recommer	2	Daily	Lifestyle a	Conversati	unknown F	Both	Very Satisf	24	8	33
15	12-20	Female	1 year to 2	Computer	Free (ad-su	Yes	Individual	Podcast	Melody	Afternoon	Sadness or	While Trav	Radio	3	Several tin	Health and	Story tellin	Both	Longer	Very Satisf	31	9	25
16	12-20	Female	1 year to 2	Computer	Free (ad-su	Yes	Duo plan-	Podcast	classical	Afternoon	Uplifting a	Study Hour	Radio	3	Rarely	Lifestyle a	Conversati	unknown F	Longer	Satisfied	4	5	48
17	12-20	Female	Less than 6	Smartphor	Free (ad-su	Yes	Duo plan-	Podcast	Rap	Morning	Relaxation	Study Hour	Radio	3	Once a we	Food and c	Story tellin	Well know	Shorter	Ok	1	17	17
18	20-35	Male	6 months t	Computer	Free (ad-su	Yes	Duo plan-	Podcast	Rap	Morning	Sadness or	While Trav	Playlists	3	Several tin	Sports	Interview	Both	Both	Very Satisf	12	24	13
19	20-35	Female	Less than 6	Smartphor	Free (ad-su	Yes	Individual	Podcast	Electronic,	Night	Relaxation	Office hou	Playlists	5	Several tin	Comedy	Educationa	Well know	Both	Very Satisf	15	8	44
20	20-35	Female	More than	Smartphor	Free (ad-su	Yes	Individual	Music	Melody	Night	Relaxation	While Trav	recommer	5	Rarely	Comedy	Story tellin	Both	Shorter	Satisfied	15	26	46
21	20-35	Female	More than	Smartphor	Premium (	Yes	Individual	Music	Pop	Night	Relaxation	While Trav	recommer	5	Rarely	Comedy	Story tellin	Well know	Shorter	Satisfied	22	15	16
22	20-35	Female	More than	Smartphor	Premium (	Yes	Duo plan-	Music	Pop	Night	Relaxation	While Trav	recommer	5	Rarely	Lifestyle a	Educationa	Both	Shorter	Very Satisf	25	7	9
23	20-35	Male	More than	Smartphor	Free (ad-su	Yes	Duo plan-	Music	Pop	Night	Relaxation	While Trav	Others	3	Rarely	Comedy	Conversati	Well know	Shorter	Ok	7	16	47
24	20-35	Male	1 year to 2	Smartphor	Premium (	Yes	Family Plar	Music	Melody	Night	Relaxation	While Trav	Playlists	5	Daily	Health and	Story tellin	Both	Both	Satisfied	12	24	25
25	20-35	Female	More than	Smartphor	Premium (	Yes	Individual	Music	Melody	Night	Relaxation	While Trav	Radio, Oth	3	Rarely	Lifestyle a	Conversati	Well know	Shorter	Very Satisf	16	1	5
26	20-35	Female	1 year to 2	Computer	Free (ad-su	Yes	Duo plan-	Music	Melody	Night	Uplifting a	Workout s	Radio	3	Rarely	Food and c	Educationa	Well know	Both	Satisfied	2	3	31
27	20-35	Male	More than	Smartphor	Premium (	Yes	Family Plar	Podcast	Pop	Night	Uplifting a	Study Hour	recommer	3	Several tin	Sports	Conversati	Well know	Longer	Very Satisf	26	15	22
28	12-20	Female	1 year to 2	Smart spea	Free (ad-su	Yes	Duo plan-	Podcast	classical	Morning	Sadness or	While Trav	Radio	3	Once a we	Health and	Story tellin	unknown F	Both	Satisfied	17	23	47

DATA PRE-PROCESSING

- **Customer Segmentatio:**
  - Age data will be in column
  - The number of hours of music listening by month is in rows.
  - Gender is colored.

- **Age & Gender Distribution by Subscription Plan:**
  - 'Age' to Columns and 'Subscription Plan' to Rows.
  - 'Engagement Score' to Size and 'Gender' to Color.
  - Select 'Bubble Chart'

- **Operating Systems by Age Group:**
  - Operating System' to Rows.
  - 'Count of Age' to Columns.
  - 'Age' to Color.
  - Choose 'Area'.

- **Willingness to Subscribe to Premium by Age and Gender:**
  - 'Age' to Columns and 'Gender' to Rows.
  - 'Engagement Score' to Color.
  - Select 'Heatmap'

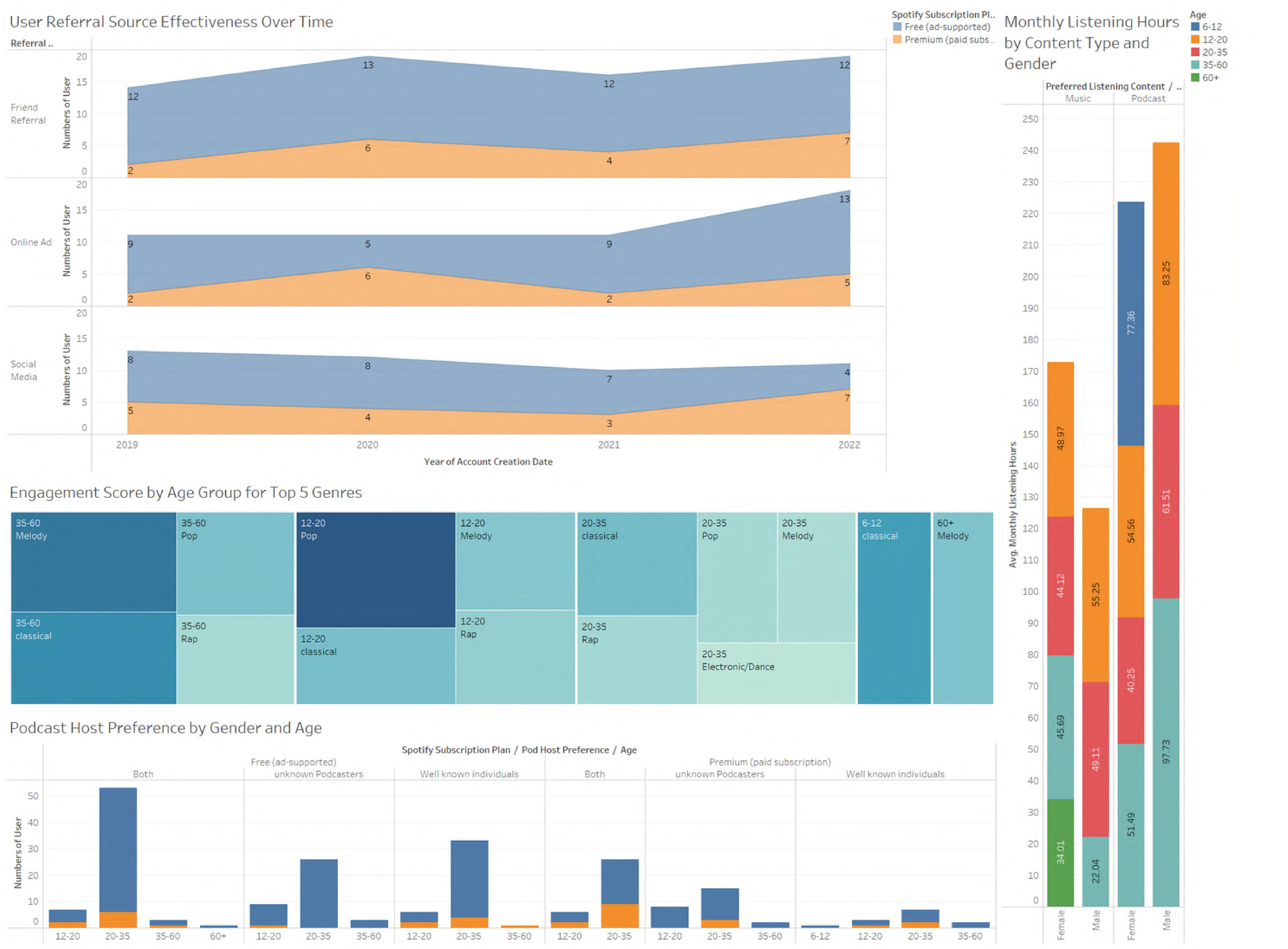
- **Top Music Genres by Engagement Score and Age Group:**
  - 'Music Genre' to Columns and 'Engagement Score' to Rows.
  - 'Age' to the 'Detail' shelf and then to the 'Color' shelf.
  - Choose 'Box Plot'.

DASHBOARD PURPOSE

The dashboard makes it possible for me to see which company Spotify has customer segmentation, age and gender distribution by subscription, willingness to subscribe to paid plans by age and gender, and music genre top by interaction score and age group.

UNIVERSITY OF GREENWICH
Alliance with FPT Education

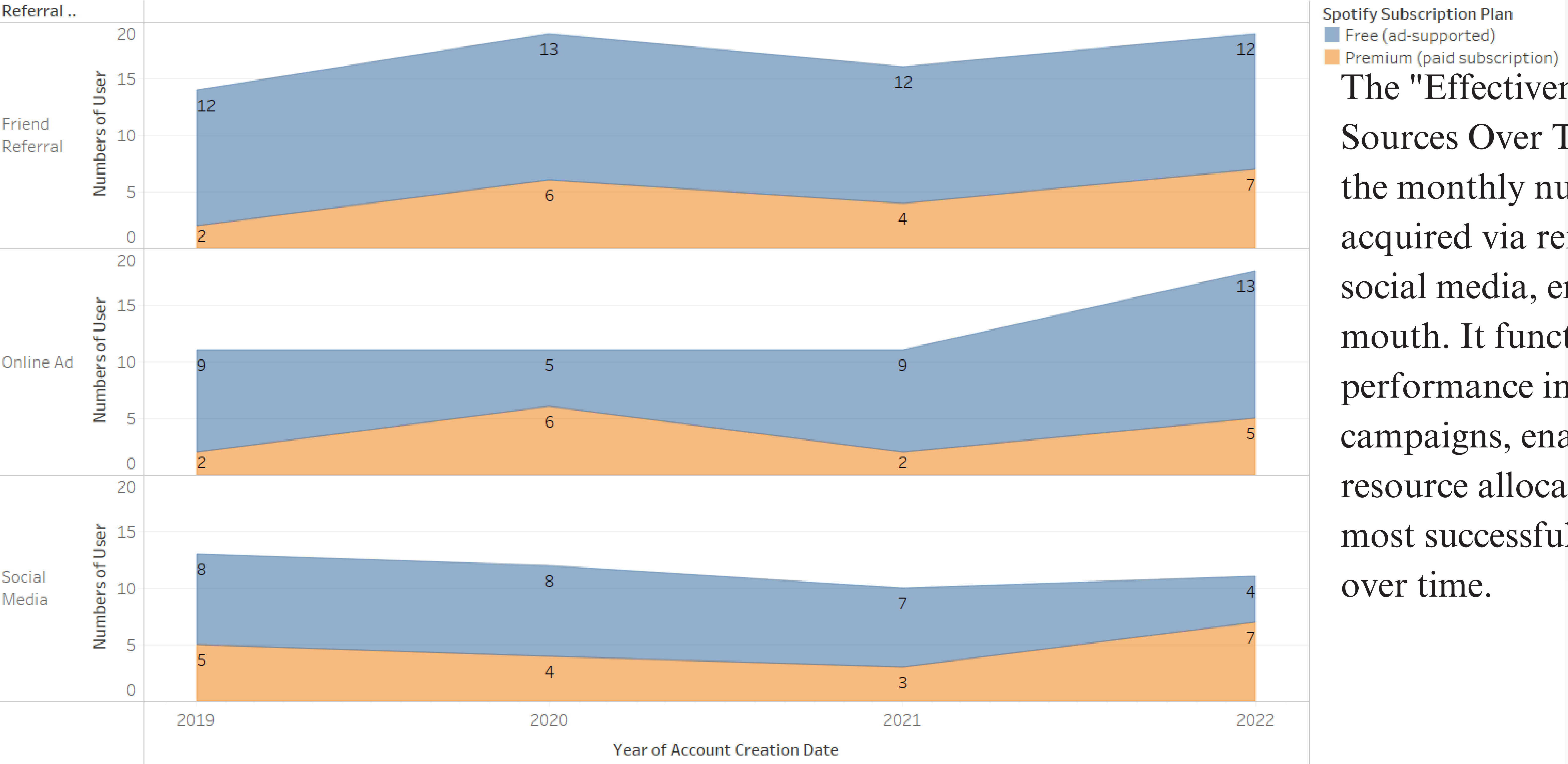
DASHBOARD



DASHBOARD

UNIVERSITY OF GREENWICH
Alliance with FPT Education

User Referral Source Effectiveness Over Time

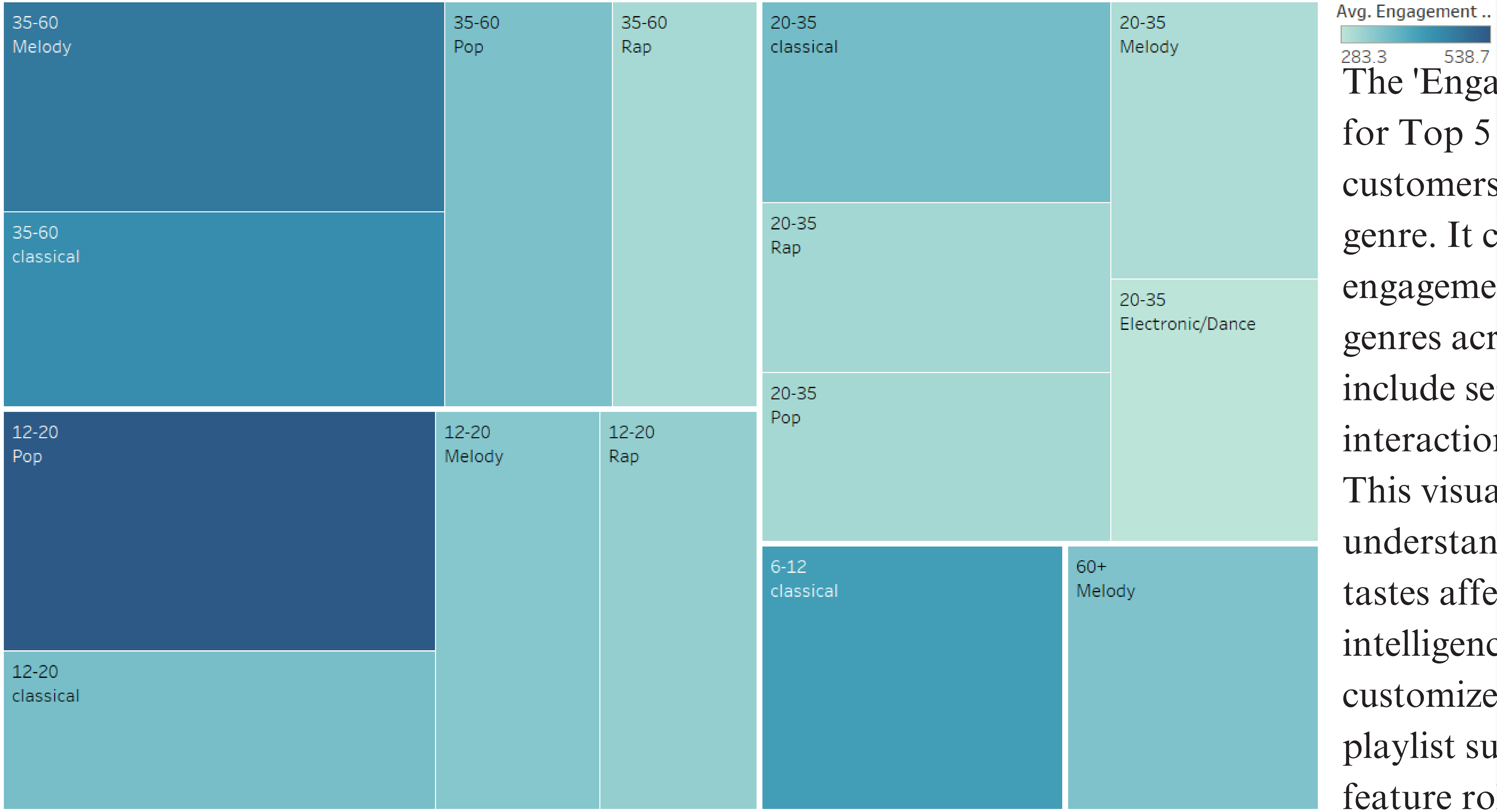


The plot of Numbers of User for Account Creation Date Year broken down by Referral Source. Color shows details about Spotify Subscription Plan. The marks are labeled by Numbers of User. The view is filtered on Referral Source, which keeps Friend Referral, Online Ad and Social Media.

The "Effectiveness of User Referral Sources Over Time" graph illustrates the monthly number of new users acquired via referral sources such as social media, email, and word-of-mouth. It functions as a key performance indicator for marketing campaigns, enabling more efficient resource allocation by identifying the most successful referral methods over time.

CHART

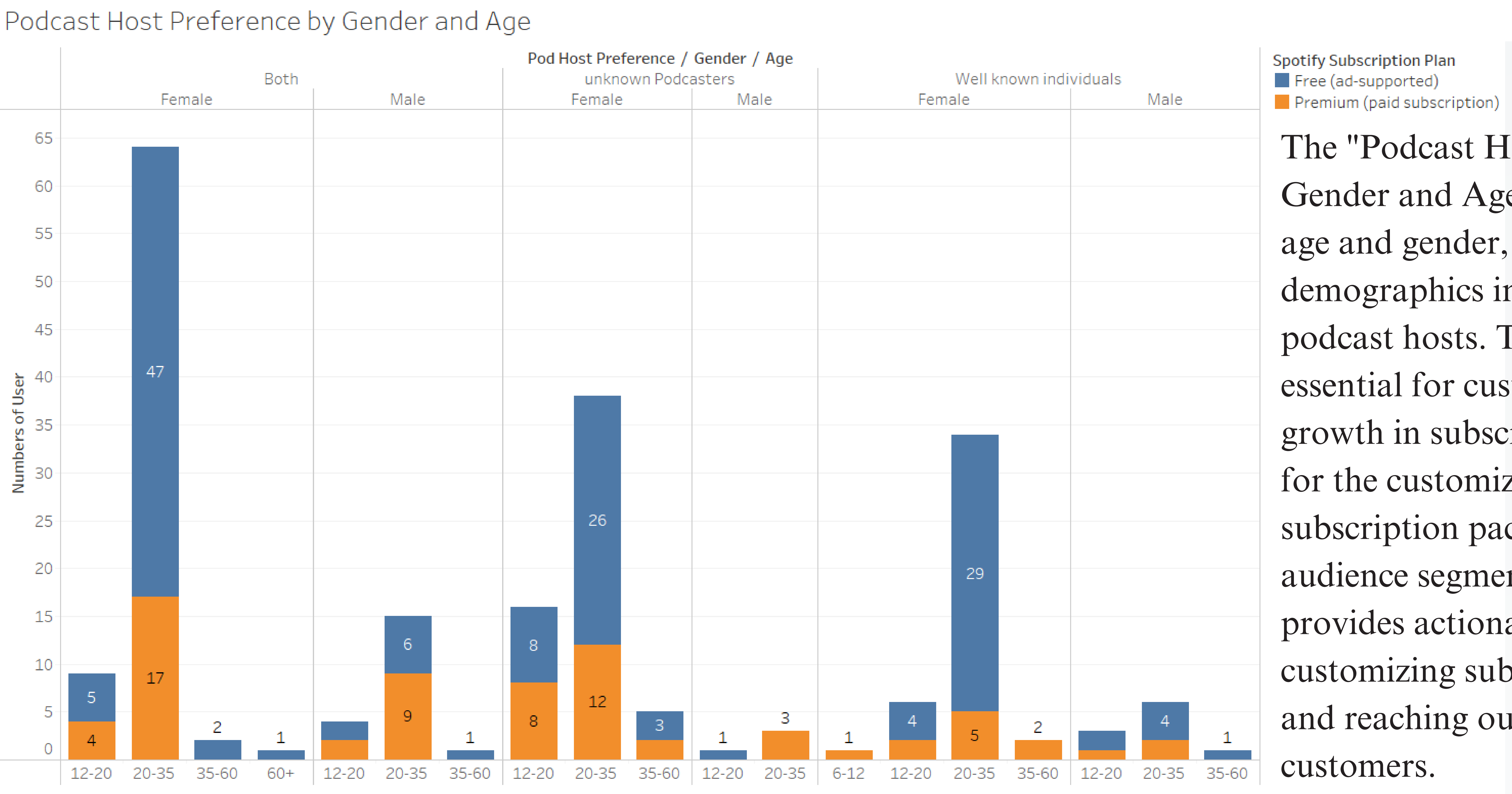
Engagement Score by Age Group for Top 5 Genres



Age and Fav Music Genre. Color shows average of Engagement Score. Size shows average of Engagement Score. The marks are labeled by Age and Fav Music Genre. The view is filtered on Fav Music Genre, which keeps classical, Electronic/Dance, Melody, Pop and Rap.

The 'Engagement Score by Age Group for Top 5 Genres' chart shows Spotify customers' interactions by age and genre. It calculates the average engagement ratings for the top 5 music genres across age groups, which may include session duration, frequency of interactions, and feature utilization. This visualisation is essential for understanding how age and musical tastes affect user engagement. This intelligence may help businesses customize tailored marketing efforts, playlist suggestions, and premium feature rollouts to particular user demographics.

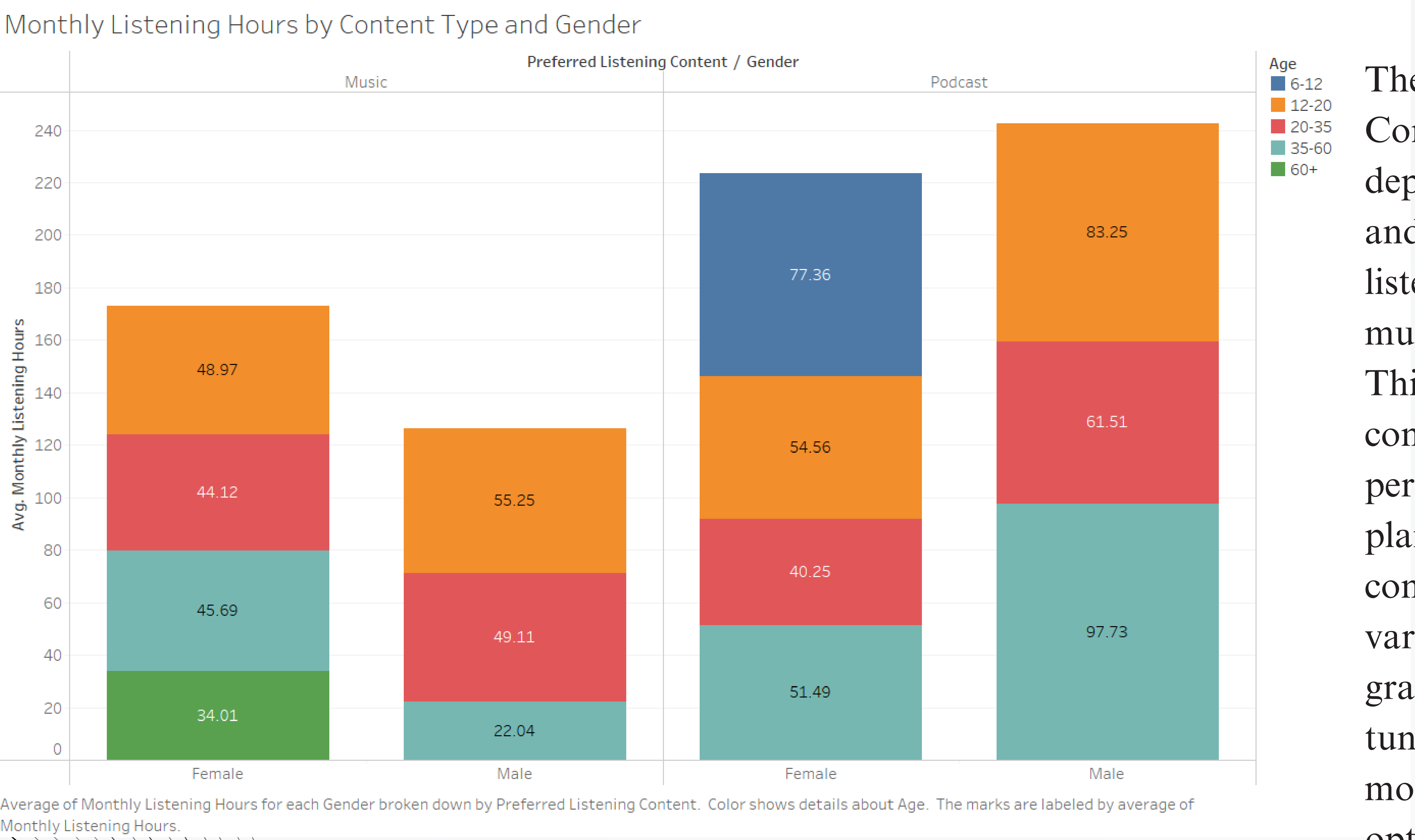
DASHBOARD



Numbers of User for each Age broken down by Pod Host Preference and Gender. Color shows details about Spotify Subscription Plan. The marks are labeled by Numbers of User.

The "Podcast Host Preference by Gender and Age" chart illustrates, by age and gender, how various demographics interact with various podcast hosts. This information is essential for customer retention and growth in subscriptions, as it allows for the customization of podcast subscription packages to specific audience segments. The chart provides actionable insights for customizing subscription bundles and reaching out to specific customers.

DASHBOARD

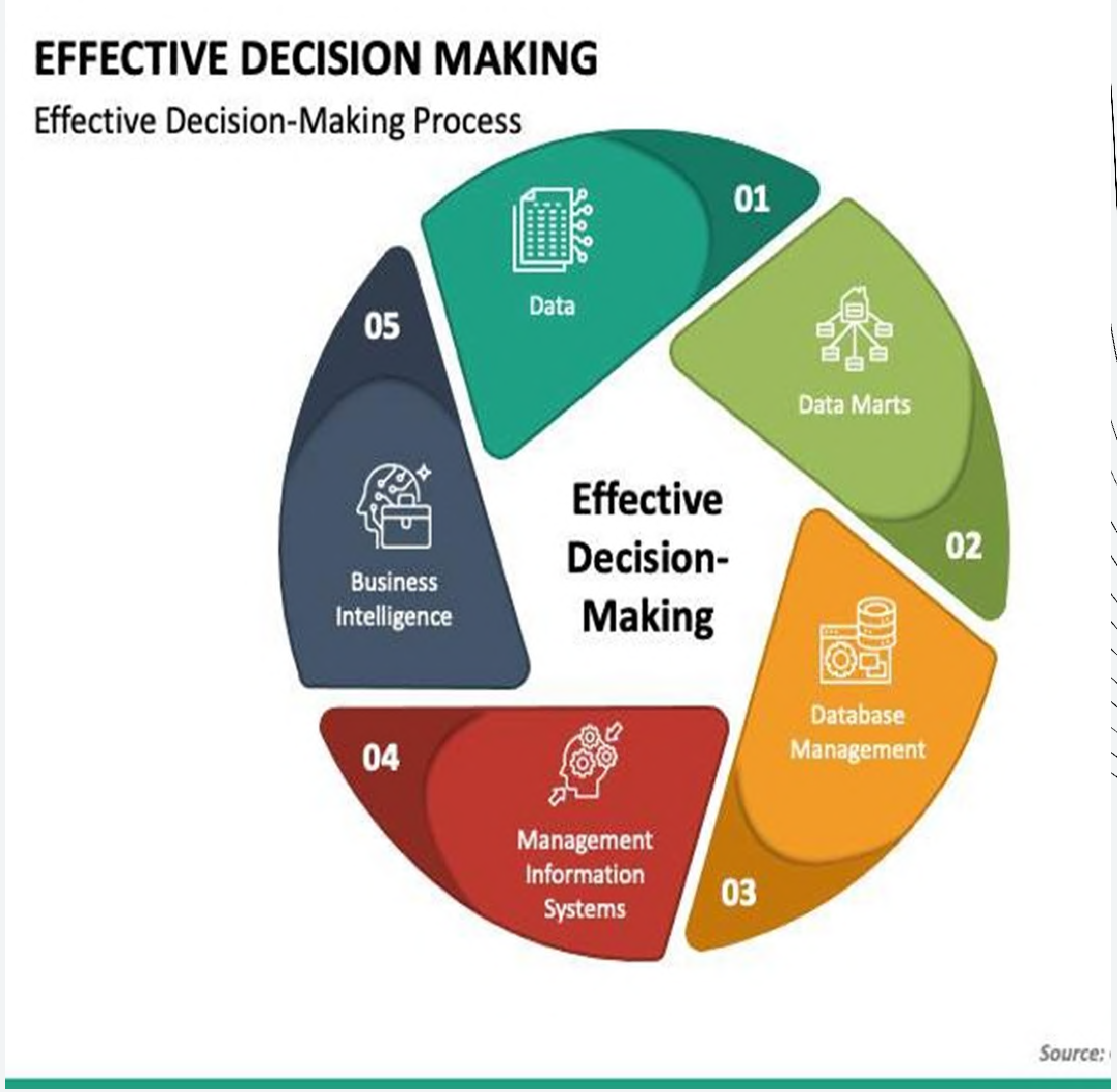


The "Monthly Listening Hours by Content Type and Gender" chart depicts the number of hours male and female users spend each month listening to content types such as music, podcasts, and audiobooks. This information is essential for consumer segmentation and personalization of subscription plans, as it identifies the categories of content that are most engaging for various demographic groups. The graph enables the company to fine-tune its content strategy and provide more customized subscription options based on the listening patterns of its customers.

HOW BUSINESS INTELLIGENCE TOOLS CAN CONTRIBUTE TO EFFECTIVEDECISION-MAKING

The company has access to a diverse range of resources that can be employed to support the development of well-informed business choices. Certain, such as gathering data using dashboards and reports, creating and analyzing through interactive and data mining, making selections and implementing strategies via ad hoc queries, what-if scenarios, and instruments forecasting, and outcomes analysis with dashboards and reports —distinct components that are not closely linked—enhance the coherence among these tools, enabling effective decision-making for the organization.

These tools aid in identifying vulnerabilities in the system and embrace data that can be utilized to determine the organization's requirements. They contribute to the process of identifying suitable growth and profitable strategies to be implemented within the company, facilitating the achievement of its objectives.



LEGAL ISSUES RELATED TO USER DATA MINING

1. When accumulating, retaining, and processing user data, businesses are required to conform with data protection laws. These laws, such as the General Data Protection Regulation (GDPR) in the European Union and the California Consumer Privacy Act (CCPA) in the United States, require businesses to obtain user consent prior to acquiring and utilizing their personal information.

2. When acquiring and utilizing user data for BI purposes, companies must uphold user privacy. Users have a reasonable expectation of privacy, and businesses must ensure that their data is collected and used only for legitimate BI purposes, and not for other purposes, such as profiling or advertising, without their consent.

3. Access levels and authorizations must be thoroughly examined: More businesses are turning to business intelligence (BI), an easy-to-use technology for analyzing data and providing prospective insights, in order to better comprehend customer behavior. consumers, but it will be exceedingly negative when such data is analyzed by individuals with nefarious intent, when they can predict information about user behavior and use it to make money. just for my own benefit

4. When accumulating user data for BI purposes, organizations must clarify whose data it is. In some instances, the user may own the data and have the right to control its use, whereas in others, the business may own the data and have the right to utilize it for BI purposes.

THANK'S FOR WATCHING

UNIVERSITY OF GREENWICH

Alliance with FPT Education