explore_data

February 19, 2022

```
                                       playlist_uri  \
0   spotify:user:36069af6af076ccd9e597184a67b68c9:…
1   spotify:user:d1144a65b1c31c5f9f56b94f831124d5:…
2   spotify:user:6b7fbed9edd6418ddd3b555bba441536:…
3   spotify:user:580b98725077a94c3c8d01d07390426b:…
4   spotify:user:1305d39070c95d161cc502e15014897d:…


                               owner  streams  stream30s  dau  wau  mau  \
0   36069af6af076ccd9e597184a67b68c9       27         27    1    1    3
1   d1144a65b1c31c5f9f56b94f831124d5        0          0    0    1    2
2   6b7fbed9edd6418ddd3b555bba441536        4          2    1    1    7
3   580b98725077a94c3c8d01d07390426b       12         12    1    1    4
4   1305d39070c95d161cc502e15014897d       20          4    1    1    2


   mau_previous_month  mau_both_months  users  skippers owner_country  \
0                   3                0      8         0            US
1                   1                1      3         0            US
2                   5                0     15         0            US
3                   6                1     10         0            US
4                   1                1      2         1            US


   n_tracks  n_local_tracks  n_artists  n_albums  monthly_stream30s  \
0        52               0          4         7                 30
1       131               0        112       113                112
2        43               0         35        36                 63
3        27               1         27        26                154
4        52               0         47        51                230


   monthly_owner_stream30s                                            tokens  \
0                       27  ["ambient", "music", "therapy", "binaural", "b…
1                       94                                  ["good", "living"]
2                        0                                    ["norte\u00f1a"]
3                      108                                                  []
4                        0                                  ["cheesy", "pants"]


         genre_1      genre_2          genre_3    mood_1      mood_2    mood_3
0   Dance & House     New Age  Country & Folk  Peaceful    Romantic    Somber
```

1

```
1           Pop    Indie Rock      Alternative   Excited     Yearning    Defiant
2         Latin            -                -     Lively       Upbeat   Romantic
3  Dance & House  Electronica              Pop   Excited   Aggressive    Defiant
4     Indie Rock  Alternative      Electronica   Excited      Defiant   Yearning

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 403366 entries, 0 to 403365
Data columns (total 25 columns):
 #   Column                 Non-Null Count   Dtype
---  ------                 --------------   -----
 0   playlist_uri           403366 non-null  object
 1   owner                  403366 non-null  object
 2   streams                403366 non-null  int64
 3   stream30s              403366 non-null  int64
 4   dau                    403366 non-null  int64
 5   wau                    403366 non-null  int64
 6   mau                    403366 non-null  int64
 7   mau_previous_month     403366 non-null  int64
 8   mau_both_months        403366 non-null  int64
 9   users                  403366 non-null  int64
 10  skippers               403366 non-null  int64
 11  owner_country          403366 non-null  object
 12  n_tracks               403366 non-null  int64
 13  n_local_tracks         403366 non-null  int64
 14  n_artists              403366 non-null  int64
 15  n_albums               403366 non-null  int64
 16  monthly_stream30s      403366 non-null  int64
 17  monthly_owner_stream30s  403366 non-null  int64
 18  tokens                 403366 non-null  object
 19  genre_1                403366 non-null  object
 20  genre_2                403366 non-null  object
 21  genre_3                403366 non-null  object
 22  mood_1                 403366 non-null  object
 23  mood_2                 403366 non-null  object
 24  mood_3                 403366 non-null  object
dtypes: int64(15), object(10)
memory usage: 76.9+ MB

             streams      stream30s            dau             wau            mau  \
count  4.033660e+05   4.033660e+05  403366.000000  403366.000000  4.033660e+05
mean   7.101375e+01   4.283333e+01       4.418265      21.784446  6.614290e+01
std    6.492014e+03   3.772412e+03     358.855685    1614.650805  4.732580e+03
min    0.000000e+00   0.000000e+00       0.000000       0.000000  2.000000e+00
25%    0.000000e+00   0.000000e+00       0.000000       1.000000  2.000000e+00
50%    1.000000e+00   0.000000e+00       0.000000       1.000000  2.000000e+00
75%    8.000000e+00   5.000000e+00       1.000000       2.000000  4.000000e+00
max    2.629715e+06   1.513237e+06  152929.000000  669966.000000  1.944150e+06
```

	mau_previous_month	mau_both_months	users	skippers
count	4.033660e+05	403366.000000	4.033660e+05	403366.000000
mean	5.819009e+01	12.937065	1.493085e+02	2.827749
std	3.827248e+03	1240.912979	9.247484e+03	205.059728
min	0.000000e+00	0.000000	2.000000e+00	0.000000
25%	1.000000e+00	1.000000	2.000000e+00	0.000000
50%	2.000000e+00	1.000000	3.000000e+00	0.000000
75%	3.000000e+00	2.000000	7.000000e+00	0.000000
max	1.478684e+06	578391.000000	3.455406e+06	86162.000000

	n_tracks	n_local_tracks	n_artists	n_albums
count	403366.000000	403366.000000	403366.000000	403366.000000
mean	201.483432	3.084035	83.852050	88.224250
std	584.077765	40.330266	128.152488	133.193118
min	1.000000	0.000000	1.000000	1.000000
25%	38.000000	0.000000	18.000000	19.000000
50%	84.000000	0.000000	46.000000	48.000000
75%	192.000000	0.000000	100.000000	106.000000
max	79984.000000	9117.000000	5226.000000	6397.000000

	monthly_stream30s	monthly_owner_stream30s
count	4.033660e+05	403366.000000
mean	1.260489e+03	93.556621
std	1.062463e+05	226.250189
min	2.000000e+00	0.000000
25%	3.100000e+01	6.000000
50%	7.900000e+01	30.000000
75%	1.930000e+02	96.000000
max	4.249733e+07	25904.000000

playlist_uri	403366
owner	314899
streams	1639
stream30s	1329
dau	566
wau	1019
mau	1597
mau_previous_month	1552
mau_both_months	765
users	2293
skippers	522
owner_country	1
n_tracks	4134
n_local_tracks	576
n_artists	1560
n_albums	1621

```
monthly_stream30s           6094
monthly_owner_stream30s     2418
tokens                    192107
genre_1                       26
genre_2                       26
genre_3                       26
mood_1                        27
mood_2                        27
mood_3                        27
dtype: int64
```

Quick summary of data: There are no missing values for any of the columns. Each row is uniquely indexed by playlist_uri. There are 26 distinct values of the genre variables and 27 distinct values of the mood variables. There is only one distinct value of owner_country, so we will drop it here. We should then mention that this analysis only applies to US-created playlists and may or may not generalize to other country's playlists.

```
                                        playlist_uri  \
0  spotify:user:36069af6af076ccd9e597184a67b68c9:…
1  spotify:user:d1144a65b1c31c5f9f56b94f831124d5:…
2  spotify:user:6b7fbed9edd6418ddd3b555bba441536:…
3  spotify:user:580b98725077a94c3c8d01d07390426b:…
4  spotify:user:1305d39070c95d161cc502e15014897d:…


                                owner  streams  stream30s  dau  wau  mau  \
0  36069af6af076ccd9e597184a67b68c9        27         27    1    1    3
1  d1144a65b1c31c5f9f56b94f831124d5         0          0    0    1    2
2  6b7fbed9edd6418ddd3b555bba441536         4          2    1    1    7
3  580b98725077a94c3c8d01d07390426b        12         12    1    1    4
4  1305d39070c95d161cc502e15014897d        20          4    1    1    2


   mau_previous_month  mau_both_months  users  skippers  n_tracks  \
0                   3                0      8         0        52
1                   1                1      3         0       131
2                   5                0     15         0        43
3                   6                1     10         0        27
4                   1                1      2         1        52


   n_local_tracks  n_artists  n_albums  monthly_stream30s  \
0               0          4         7                 30
1               0        112       113                112
2               0         35        36                 63
3               1         27        26                154
4               0         47        51                230


   monthly_owner_stream30s                                            tokens  \
0                       27  ["ambient", "music", "therapy", "binaural", "b…
```

```
1                      94                          ["good", "living"]
2                       0                             ["norte\u00f1a"]
3                     108                                            []
4                       0                          ["cheesy", "pants"]

        genre_1         genre_2          genre_3    mood_1      mood_2     mood_3
0  Dance & House        New Age  Country & Folk  Peaceful    Romantic     Somber
1            Pop     Indie Rock     Alternative   Excited    Yearning    Defiant
2          Latin              -               -    Lively      Upbeat   Romantic
3  Dance & House    Electronica             Pop   Excited  Aggressive    Defiant
4     Indie Rock    Alternative     Electronica   Excited     Defiant   Yearning
```