The datasets used in this project is extracted from: https://github.com/mdeff/fma

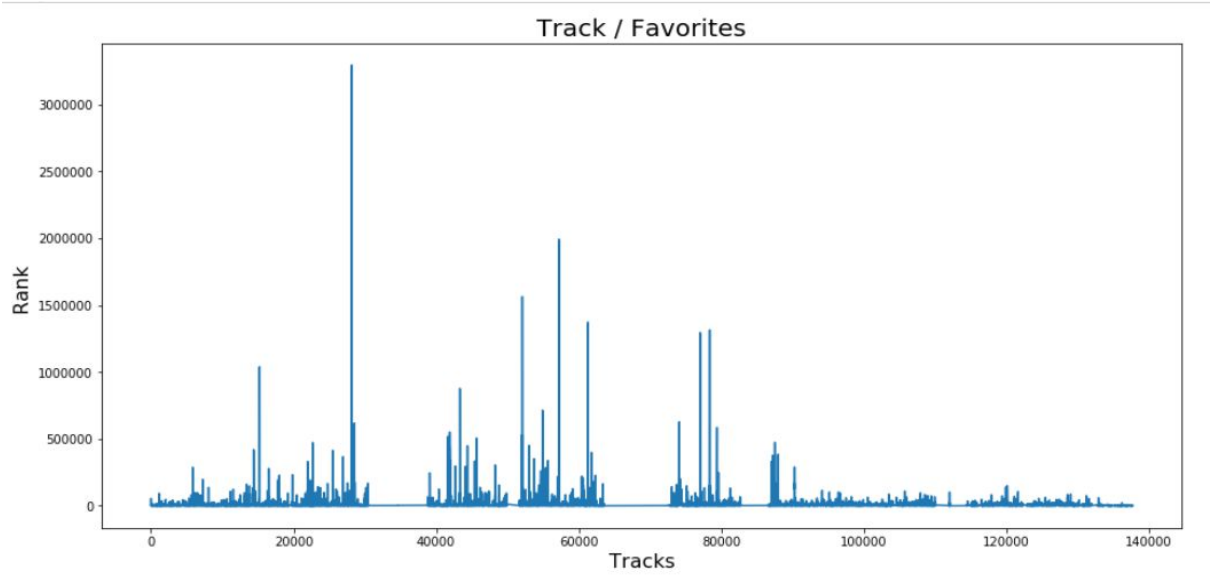
Two datasets are used in this project: tracks.csv and genres.csv. tracks.csv includes information of more than 100,000 audio tracks with almost 50 different features. These features include information about artist, track's specifications, album, genre…

The following specific features are considered for this project:

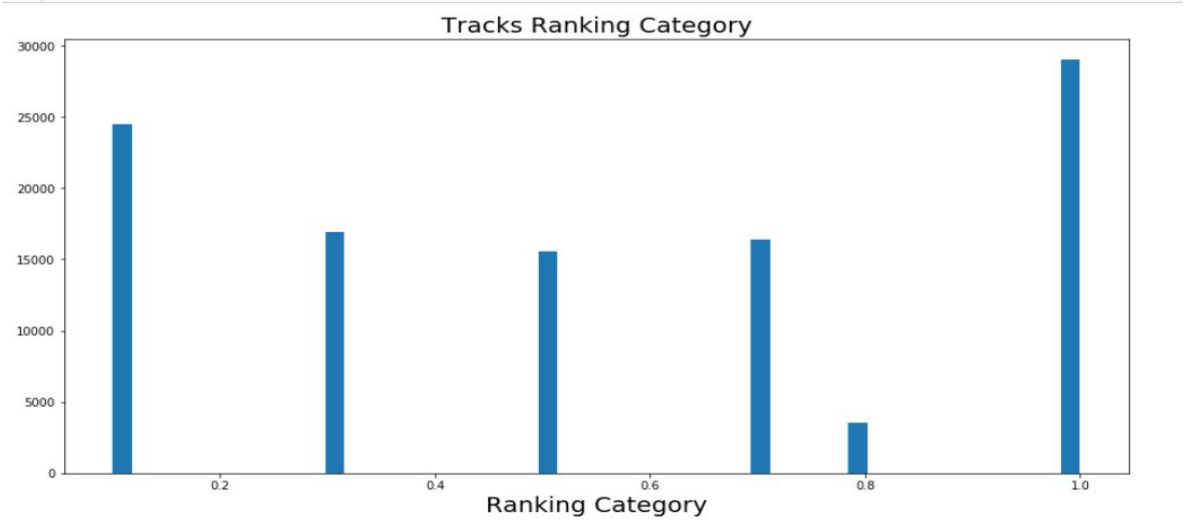
- Track_bit_rate (track quality)
- Track_listens (number of time this track has been listened)
- Track_interest (number of times this track has been favorite)
- Track_date (date of track or album released)
- genre
- Artist

The goal is determining which category of music track have higher ranking and what contribute to that.

This plot shows the distribution of tracks' ranking.

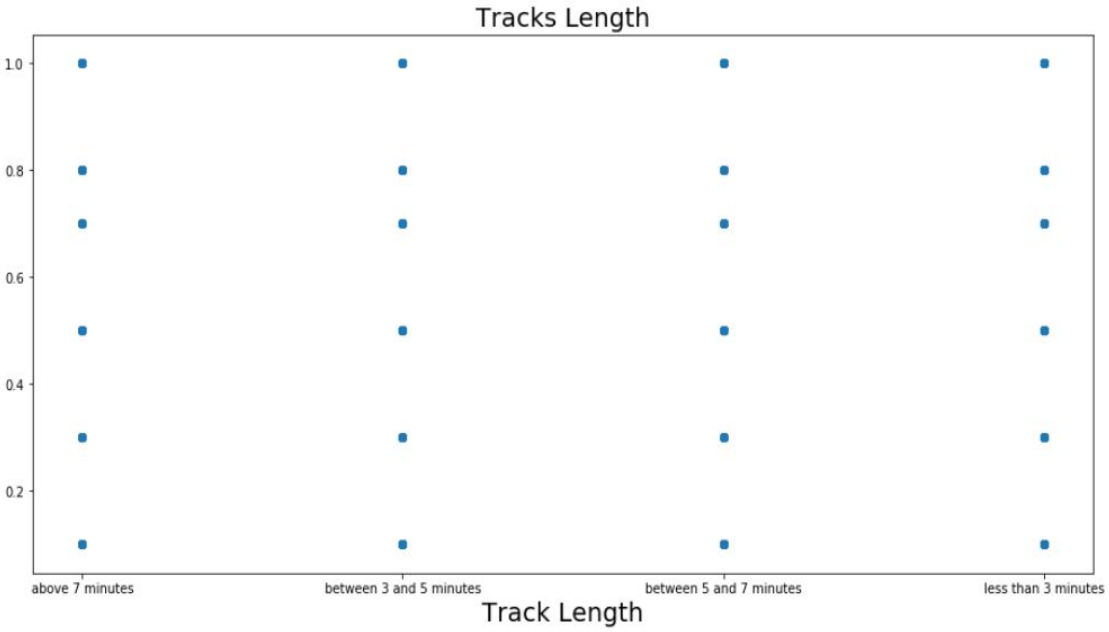


Few tracks have very high ranking and many have very low ranking. These widespread values make it difficult to have a clear result. So tracks were categorized to seven different ranking ranges to have better distribution.

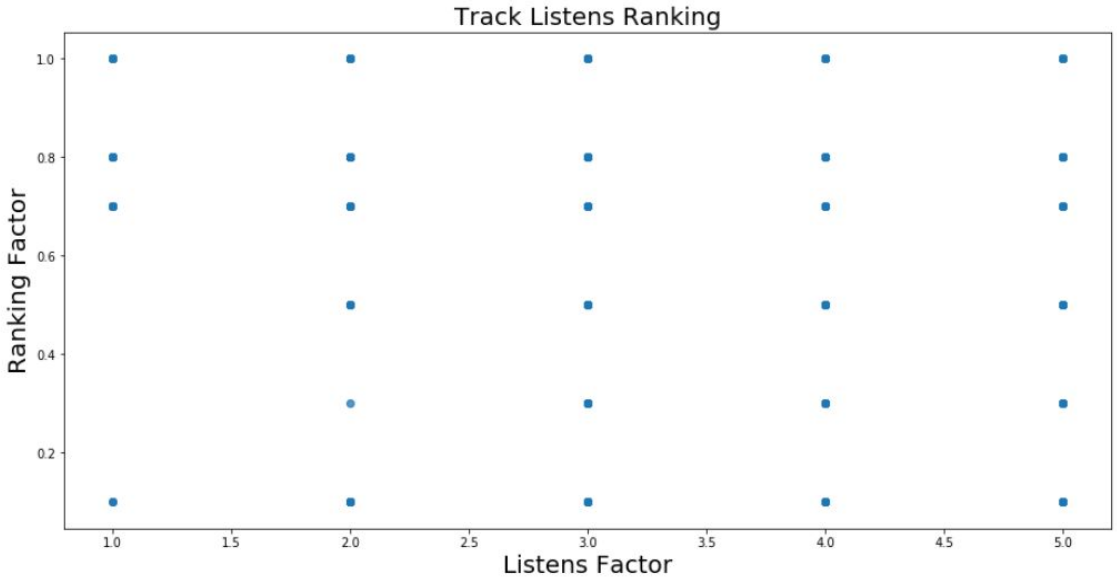


Several basic visualization have been done to find relationship between different features:

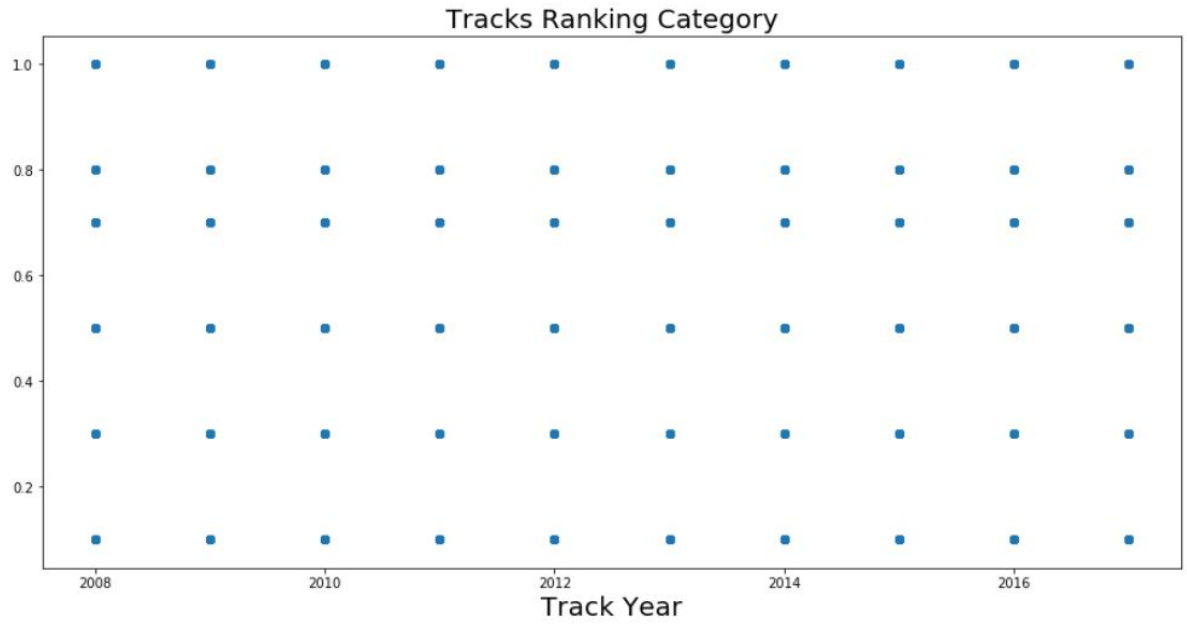
- Track length and Interest factor:



- Track Listens and Interest factor:



- Track released year and interest factor:



as seen in these plots, it is not clear which feature(s) could have more contribution to the ranking value of music tracks. Each category includes records with high ranking value as well as low ones distributed almost equally.

Clustering:

Clustering gave more meaning to the data.

- **Clustering Listen factor versus Ranking**

  Considering the following categories:

  # categorize track_listens to five divisions

  # 1: (Track Listens >= 2000)

  # 2: between 1000 and 2000

  # 3: between 500 and 1000

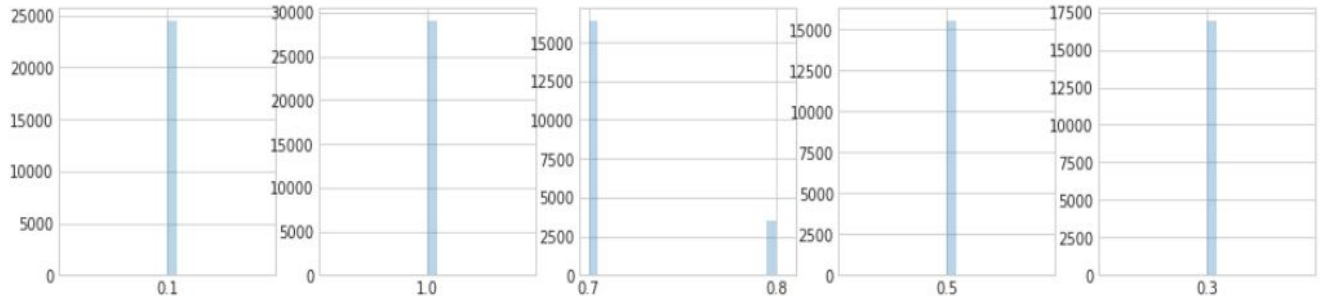
  # 4: between 250 and 500

  # 5: less than 500

  It shows tracks which has been listened more than 1000 times have higher ranking value.

  The first cluster has the lowest ranking which includes tracks being listened less than 1000 times

```
[('3', 11614), ('5', 6203), ('4', 3719), ('2', 2951), ('1', 5)]
[('1', 24720), ('2', 3176), ('3', 919), ('4', 169), ('5', 44)]
[('2', 12294), ('3', 3744), ('1', 2094), ('4', 1332), ('5', 469)]
[('4', 7745), ('3', 3911), ('5', 3865), ('2', 11)]
[('5', 12097), ('4', 4785), ('3', 30), ('2', 2)]
interest_factor
```



- **Clustering genres versus interest factor**

  Rock, Folk, and electronic are always in top three of each cluster. The below graphs shows some have very high ranking and some low.

  Counter({1: 28973, 0: 24284, 3: 16713, 8: 16325, 4: 15399, 5: 3545})

  0: [('Rock', 1562), ('Folk', 1210), ('Electronic', 1100), ('Post-Punk,Rock,Punk', 837), ('Hip-Hop', 741)]

  1: [('Rock', 1824), ('Folk', 1386), ('Electronic', 1304), ('Post-Punk,Rock,Punk', 1226), ('Hip-Hop', 799)]

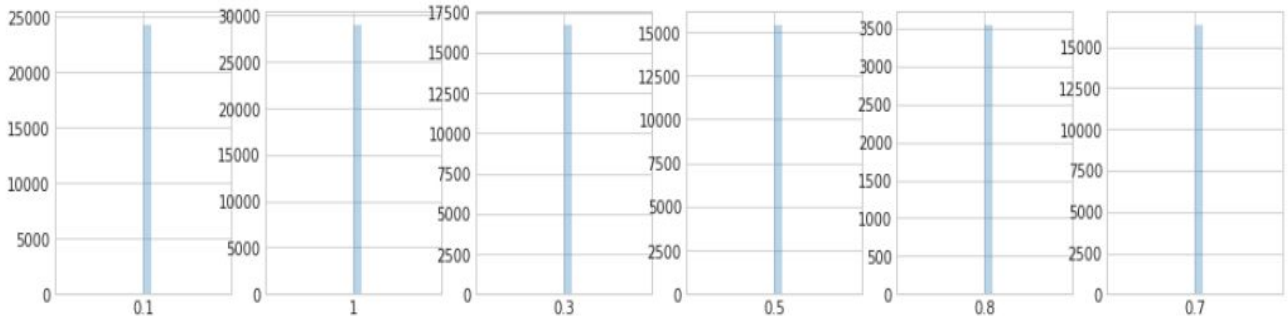
  2: [('Rock', 1220), ('Electronic', 888), ('Folk', 752), ('Post-Punk,Rock,Punk', 583), ('Hip-Hop', 552)]

  3: [('Rock', 1109), ('Folk', 817), ('Electronic', 702), ('Post-Punk,Rock,Punk', 581), ('Hip-Hop', 475)]

  4: [('Rock', 198), ('Electronic', 179), ('Folk', 164), ('Post-Punk,Rock,Punk', 133), ('Hip-Hop', 83)]

  5: [('Rock', 1048), ('Folk', 770), ('Electronic', 730), ('Post-Punk,Rock,Punk', 614), ('Hip-Hop', 470)]

```
interest_factor
```

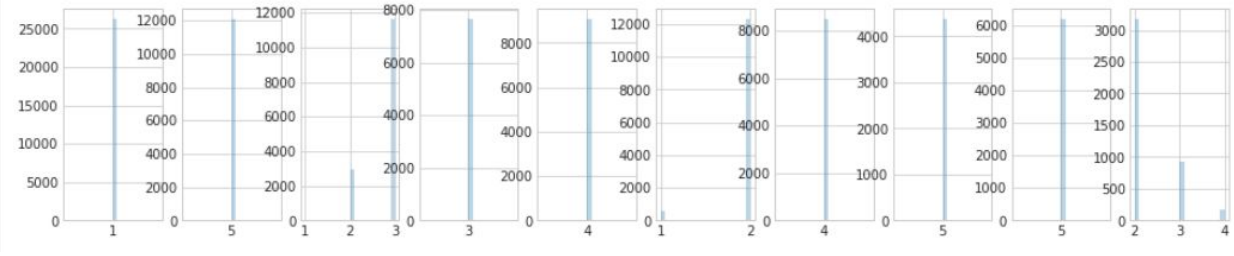


Clusters 1, 4, and 5 have higher ranking values and all have Rock, Folk, and Electronic as their top three genres.

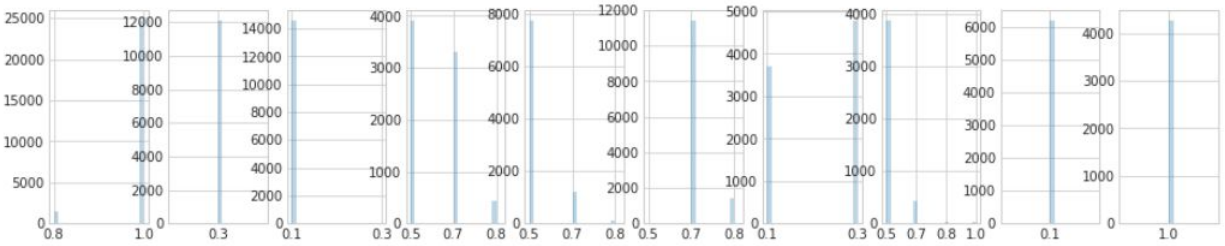
Even clusters with low ranking still have Rock, Folk, and Electronic as their top three.

- **Clustering Track Year versus interest & listens**
  First cluster has highest ranking and listen factor. Mostly includes tracks created after 2012.
  Second cluster has low ranking and listen factor. Mostly tracks released before 2012

```
[('2016', 5268), ('2015', 4833), ('2013', 3773), ('2014', 3189), ('2012', 3020)]
[('2010', 2146), ('2012', 1603), ('2013', 1561), ('2011', 1525), ('2009', 1192)]
[('2016', 2363), ('2015', 2221), ('2013', 1921), ('2014', 1623), ('2010', 1534)]
[('2009', 1221), ('2010', 1006), ('2012', 966), ('2013', 903), ('2016', 891)]
[('2010', 1480), ('2011', 1301), ('2013', 1198), ('2012', 1144), ('2009', 1098)]
[('2016', 2249), ('2015', 2143), ('2013', 1761), ('2014', 1564), ('2012', 1475)]
[('2009', 1335), ('2010', 1261), ('2013', 1032), ('2011', 1021), ('2012', 890)]
[('2010', 944), ('2009', 853), ('2011', 731), ('2013', 571), ('2012', 547)]
[('2009', 1049), ('2017', 1001), ('2016', 869), ('2010', 849), ('2011', 693)]
[('2009', 882), ('2011', 795), ('2010', 792), ('2012', 775), ('2013', 471)]
listens_factor
```



```
interest_factor
```



- **Clustering-Track Year versus Genres**

  Shows which genres have been more released each year

  2010 the worst year (many electronics, hip-hop, experimental created)
  After 2012 highest rank (mostly rock, folk, electronics)

  Counter({4: 14471, 3: 13682, 1: 12919, 5: 12717, 6: 12346, 0: 11219, 2: 11030, 7: 9769, 8: 4233, 9: 2853})

  0: [('Post-Punk,Rock,Punk', 969), ('Folk', 769), ('Rock', 635), ('Electronic', 436), ('Indie-Rock,Lo-Fi,Rock', 423), ('Avant-Garde,Experimental', 354), ('Punk,Rock', 350), ('Noise,Experimental', 308), ('Lo-Fi,Rock', 236), ('Folk,Pop,Experimental Pop,Singer-Songwriter', 216)]

  1: [('Rock', 1042), ('Electronic', 626), ('Folk', 546), ('Hip-Hop', 377), ('Punk,Rock', 289), ('Noise,Experimental', 274), ('Experimental', 222), ('Indie-Rock,Rock', 219), ('Ambient Electronic,Electronic', 207), ('Post-Punk,Rock,Punk', 193)]

  2: [('Hip-Hop', 441), ('Electronic', 429), ('Folk', 334), ('Rock', 311), ('Old-Time / Historic', 303), ('Punk,Rock', 276), ('Experimental', 221), ('Ambient Electronic,Electronic', 193), ('Indie-Rock,Rock', 181), ('Noise,Experimental,Electronic', 174)]

3: [('Rock', 1034), ('Electronic', 692), ('Folk', 563), ('Hip-Hop', 489), ('Punk,Rock', 374), ('Old-Time / Historic', 303), ('Avant-Garde,Experimental', 295), ('Indie-Rock,Rock', 250), ('Experimental', 232), ('Post-Punk,Rock,Punk', 227)]

4: [('Rock', 1656), ('Electronic', 901), ('Folk', 733), ('Hip-Hop', 541), ('Punk,Rock', 461), ('Indie-Rock,Rock', 368), ('Noise,Experimental', 350), ('Ambient Electronic,Electronic', 340), ('Rock,Garage', 285), ('Experimental,Radio,Spoken,Field Recordings', 274)]

5: [('Electronic', 360), ('Hip-Hop', 313), ('Experimental', 292), ('Pop,Experimental Pop', 223), ('Chip Music,Electronic', 157), ('Chiptune,Chip Music,Electronic', 154), ('Rock', 138), ('Rock,Garage', 129), ('Indie-Rock,Rock', 117), ('Noise,Lo-Fi,Rock,Experimental', 108)]

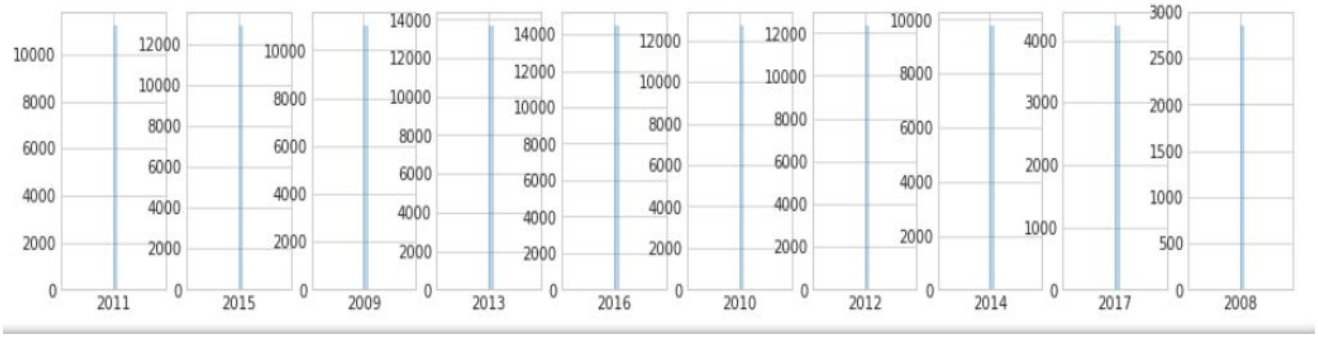
6: [('Post-Punk,Rock,Punk', 2188), ('Folk', 1700), ('Rock', 1258), ('Indie-Rock,Lo-Fi,Rock', 928), ('Avant-Garde,Experimental', 806), ('Electronic', 800), ('Punk,Rock', 625), ('Lo-Fi,Rock', 516), ('Experimental,Field Recordings', 429), ('Noise,Experimental', 429)]

7: [('Electronic', 309), ('Hip-Hop', 281), ('Pop,Experimental Pop', 208), ('Experimental', 192), ('Chiptune,Chip Music,Electronic', 138), ('Chip Music,Electronic', 134), ('Rock,Garage', 128), ('Noise,Experimental,Electronic', 118), ('Rock', 93), ('Indie-Rock,Rock', 92)]

8: [('Old-Time / Historic', 229), ('Hip-Hop', 159), ('Avant-Garde,Experimental', 146), ('Electronic', 105), ('Punk,Rock', 104), ('Rock', 102), ('Dance,Pop,Electronic', 89), ('Experimental', 87), ('Techno,Electronic', 72), ('Folk', 72)]

9: [('Rock', 692), ('Electronic', 245), ('Folk', 206), ('Post-Punk,Rock,Punk', 108), ('Avant-Garde,Experimental', 103), ('Experimental,Field Recordings', 102), ('Punk,Rock', 96), ('Noise,Experimental', 96), ('Rock,Garage', 89), ('Spoken Weird,Spoken', 75)]

track_year_created



2008: variety of genres was released such as punk,post-punk, avant-garde, experimental. Also includes Rock, Electronic and Folk which have highest ranking compare to other genres in 2008

2009: hip-hop is leading in this group but still number of listened is less than 500. And it has close competition with electronic, Folk, and Rock

2010: variety of genres such as hip-hop, electronic, chip music were released but not been listened very much. Few Rock, no Folk

2011 and 2012: genres has moved toward punk, post-punk but still many Rock and Folk been released

2013: Rock, Electronic, and Folk lead the competition

2014: Electronic, Hip-Hop, Pop, and Experimental Pop are still on top of the selection although number of tracks released in that year has decreased.

2015: Rock, Electronic, Folk are still in top of the chart

2016: Rock, Electronic, Folk

2017: Old-Time / Historic, Hip-Hop, Avant-Garde, Experimental

Summary:

These models shows some facts of what contribute to a music track ranking.

First, the number of times a track has been listened could have positive effect on its ranking.

Second, tracks released after 2012 have higher ranking than the ones released in previous years.

Third, genres such as Rock, Folk, and Electronics are always on top of the chart

Genres after 2012 includes: Rock, Folk, Electronic