My Soundtrack

January 7, 2021

1 Favorite Songs

1.1 *Data*

1.1.1 Import Libraries

[1]:
```
import spotipy
import pandas as pd
import time
import datetime as dt
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import itertools as it
from os import environ
from spotipy.oauth2 import SpotifyClientCredentials
from scipy.stats import ttest_ind
from sklearn.model_selection import cross_val_score, train_test_split,␣
↪RepeatedStratifiedKFold
from sklearn.neighbors import KNeighborsClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn import metrics
```

1.1.2 Client Credentials Flow

[2]:
```
client_id = environ['SPOTIPY_CLIENT_ID']
secret = environ['SPOTIPY_CLIENT_SECRET']
redirect_uri = environ['SPOTIPY_REDIRECT_URI']
user_id = '12179890696'

scope = 'user-library-read'
train_playlist = '4lexbQSIJBWtJGQS3GQkUC'
test_playlist1 = '37i9dQZF1DWWBHeXOYZf74'
test_playlist2 = '37i9dQZEVXbrG9oAilGQPt'
test_playlist3 = '37i9dQZEVXcEQDzu8ByiUH'

client_credentials_manager = SpotifyClientCredentials(client_id=client_id,␣
↪client_secret=secret)
```

```python
sp = spotipy.Spotify(client_credentials_manager=client_credentials_manager)
```

1.1.3 Authorization Code Flow

[3]:
```python
from spotipy.oauth2 import SpotifyOAuth

sp_auth = spotipy.Spotify(auth_manager=SpotifyOAuth(client_id=client_id,
                                                    client_secret=secret,
                                                    redirect_uri=redirect_uri,
                                                    scope=scope))
```

1.1.4 Compile Dataset

[4]:
```python
# Function to get audio features for a track
def getTrackFeatures(ids):
  meta = sp.tracks(ids)
  features = sp.audio_features(ids)
  tracks = []
  for i in range(len(meta['tracks'])):
      track = meta['tracks'][i]
      # metadata
      name = track['name']
      album = track['album']['name']
      artist = track['album']['artists'][0]['name']
      release_date = track['album']['release_date']
      length = track['duration_ms']
      popularity = track['popularity']
      explicit = track['explicit']

      # audio features
      key = features[i]['key']
      mode = features[i]['mode']
      time_signature = features[i]['time_signature']
      acousticness = features[i]['acousticness']
      danceability = features[i]['danceability']
      energy = features[i]['energy']
      instrumentalness = features[i]['instrumentalness']
      liveness = features[i]['liveness']
      loudness = features[i]['loudness']
      speechiness = features[i]['speechiness']
      tempo = features[i]['tempo']
      valence = features[i]['valence']

      tracks.append((name, album, artist, release_date, length, popularity,␣
↪explicit, key, mode, time_signature, acousticness, danceability, energy,␣
↪instrumentalness, liveness, loudness, speechiness, tempo, valence))
```

```
    return tracks
```

[5]:
```python
# Function to get track ID's from a user's playlist
def getTrackIDs(user, playlist_id):
    ids = []
    playlist = sp.user_playlist(user, playlist_id)
    for item in playlist['tracks']['items']:
        track = item['track']
        ids.append(track['id'])
    return ids
```

[6]:
```python
# Function to get track ID's for a user's saved tracks
def libraryTrackIDs():
    results = sp_auth.current_user_saved_tracks()
    tracks = results['items']
    while results['next']:
        results = sp_auth.next(results)
        tracks.extend(results['items'])
    library_ids = []
    for item in tracks:
        track = item['track']
        library_ids.append(track['id'])
    return library_ids
```

[7]:
```python
# Function to divide list into chunks of 50
def divide_chunks(ids, n):
    id_chunks = []
    for i in range(0, len(ids), n):
        chunk = ids[i:i + n]
        id_chunks.append(chunk)
    return id_chunks
```

Favorite Tracks To make a dataset of my favorite tracks, I am taking all tracks from a playlist I have been curating for years of all the songs I listen to on repeat.

[8]:
```python
ids = getTrackIDs(user_id, train_playlist)
id_chunks = divide_chunks(ids, 50)
```

[9]:
```python
# get audio features for favorite tracks
favorite_tracks = []
for id_chunk in id_chunks:
    for track in getTrackFeatures(id_chunk):
        favorite_tracks.append(track)
```

[10]:
```
```

```python
# create dataframe for favorite tracks
favorites_df = pd.DataFrame(favorite_tracks, columns = ['name', 'album',␣
↪'artist', 'release_date', 'length', 'popularity', 'explicit', 'key', 'mode',␣
↪'time_signature', 'acousticness', 'danceability', 'energy',␣
↪'instrumentalness', 'liveness', 'loudness', 'speechiness', 'tempo',␣
↪'valence'])

# create boolean column to indicate that it is a favorite track
favorites_df['favorite'] = 1
favorites_df
```

[10]:

```
                                  name  \
0          When the Day Met the Night
1    Under Pressure - Remastered 2011
2                     Comfortably Numb
3                   Wish You Were Here
4                       Hero / Heroine
..                                 …
88                                 Boy
89                          Moon River
90                           Good Days
91                           Limitless
92               hot tub DREAM Machine

                                                album             artist  \
0                                       Pretty. Odd.  Panic! At The Disco
1    Queen 40 Limited Edition Collector's Box Set V…              Queen
2                                           The Wall          Pink Floyd
3                                 Wish You Were Here          Pink Floyd
4                                    Boys Like Girls     Boys Like Girls
..                                                …                   …
88             The Orchard (10th Anniversary Edition)          Ra Ra Riot
89                                        Moon River         Frank Ocean
90                                         Good Days                 SZA
91                                            Athena      Sudan Archives
92                             hot tub DREAM Machine            tobi lou

   release_date  length  popularity  explicit  key  mode  time_signature  \
0    2008-03-21  293826          53     False    7     1               4
1    2011-01-01  248440          34     False    2     1               4
2    1979-11-30  382296          75     False   11     0               4
3    1975-09-12  334743          79     False    7     1               4
4    2007-08-31  232240          53     False    0     1               4
..          …       …         …        …   …    …              …
88   2020-08-24  190906          16     False    2     1               4
89   2018-02-14  188323          70     False    0     1               3
90   2020-12-25  279204          84      True    1     0               4
```

```
91   2019-11-01  175601          44      False    8      1               4
92   2020-02-21  205714          60       True   11      1               4

    acousticness  danceability  energy  instrumentalness  liveness  loudness  \
0         0.0618         0.432   0.656          0.000060    0.3910    -6.889
1         0.4220         0.671   0.711          0.000000    0.1040    -7.813
2         0.1500         0.472   0.366          0.308000    0.0837   -12.595
3         0.7350         0.481   0.262          0.011400    0.8320   -15.730
4         0.0200         0.422   0.904          0.000004    0.6860    -4.531
..           …             …     …               …         …         …
88        0.0364         0.579   0.861          0.007960    0.4030    -5.542
89        0.8770         0.240   0.116          0.000920    0.1000   -13.216
90        0.4990         0.436   0.655          0.000008    0.6880    -8.370
91        0.0158         0.642   0.475          0.013700    0.3690    -8.473
92        0.2640         0.677   0.526          0.000003    0.1130    -8.245

    speechiness    tempo  valence  favorite
0        0.0419  130.276   0.1710         1
1        0.0478  113.809   0.4660         1
2        0.0286  127.167   0.1710         1
3        0.0414  122.883   0.3750         1
4        0.1020  163.929   0.3200         1
..          …       …       …       …
88       0.0468  161.954   0.7800         1
89       0.0329   77.349   0.0937         1
90       0.0583  121.002   0.4120         1
91       0.0363   87.890   0.3630         1
92       0.0361  140.082   0.4670         1

[93 rows x 20 columns]
```

All Saved Tracks

[11]:
```
library_ids = libraryTrackIDs()
library_id_chunks = divide_chunks(library_ids, 50)
```

[12]:
```
# get audio features for library tracks
library_tracks = []
for library_id_chunk in library_id_chunks:
    for track in getTrackFeatures(library_id_chunk):
        library_tracks.append(track)
```

[13]:

```
# create dataframe for all tracks
library_df = pd.DataFrame(library_tracks, columns = ['name', 'album', 'artist',␣
 ↪'release_date', 'length', 'popularity', 'explicit', 'key', 'mode',␣
 ↪'time_signature', 'acousticness', 'danceability', 'energy',␣
 ↪'instrumentalness', 'liveness', 'loudness', 'speechiness', 'tempo',␣
 ↪'valence'])
# create boolean column to indicate that it is not a favorite track
library_df['favorite'] = 0
library_df
```

```
[13]:                   name                               album  \
0              Good Days                           Good Days
1                Darlin'                             Darlin'
2               Paradise                            Paradise
3          Wouldn't Leave                                  ye
4      Pink Skies (Demo)                            Demo 001
…                      …                                   …
1984         Cough Syrup  Young The Giant (Special Edition)
1985           Tightrope                       Walk The Moon
1986    Different Colors                    TALKING IS HARD
1987      Work This Body                    TALKING IS HARD
1988            Anna Sun                       Walk The Moon

                   artist release_date  length  popularity  explicit  key  \
0                     SZA   2020-12-25  279204          84      True    1
1                tobi lou   2018-04-23  205090          69      True    9
2                   Bazzi   2019-04-04  169038           3      True   11
3              Kanye West   2018-06-01  205546           0      True    3
4      Wiley from Atlanta   2018-10-26  223101          60      True    9
…                       …            …       …           …         …    …
1984      Young the Giant         2011  249520          73     False   11
1985        WALK THE MOON   2012-06-19  209186          54     False    1
1986        WALK THE MOON   2014-12-02  222053          52     False    0
1987        WALK THE MOON   2014-12-02  175906          59     False    4
1988        WALK THE MOON   2012-06-19  321280          66     False   10

      mode  time_signature  acousticness  danceability  energy  \
0        0               4      0.499000         0.436   0.655
1        1               4      0.571000         0.866   0.388
2        0               4      0.082800         0.844   0.644
3        1               4      0.494000         0.555   0.433
4        0               4      0.485000         0.565   0.566
…        …               …             …             …       …
1984     0               3      0.034300         0.534   0.721
1985     0               4      0.000084         0.467   0.794
1986     1               4      0.000797         0.480   0.826
1987     1               4      0.028300         0.421   0.831
```

```
1988     1              4      0.001730         0.472   0.844

      instrumentalness  liveness  loudness  speechiness    tempo  valence  \
0             0.000008     0.688    -8.370       0.0583  121.002    0.412
1             0.000000     0.100   -11.009       0.3200  109.976    0.556
2             0.000000     0.113    -6.273       0.0479  122.061    0.591
3             0.000000     0.313    -8.559       0.5460  164.236    0.352
4             0.000007     0.104    -8.737       0.1410  138.836    0.435
…                  …         …        …            …        …        …
1984          0.000006     0.115    -7.307       0.0417  128.978    0.225
1985          0.002040     0.103    -6.174       0.0349  162.435    0.589
1986          0.000001     0.125    -4.602       0.0397   96.000    0.687
1987          0.000000     0.464    -5.128       0.1070  134.027    0.488
1988          0.000000     0.240    -6.578       0.0540  140.034    0.340

      favorite
0            0
1            0
2            0
3            0
4            0
…          …
1984         0
1985         0
1986         0
1987         0
1988         0

[1989 rows x 20 columns]
```

Merge Datasets into Training Data

[14]:
```python
library_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1989 entries, 0 to 1988
Data columns (total 20 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   name              1989 non-null   object
 1   album             1989 non-null   object
 2   artist            1989 non-null   object
 3   release_date      1989 non-null   object
 4   length            1989 non-null   int64
 5   popularity        1989 non-null   int64
 6   explicit          1989 non-null   bool
 7   key               1989 non-null   int64
 8   mode              1989 non-null   int64
```

```
 9   time_signature    1989 non-null   int64
 10  acousticness      1989 non-null   float64
 11  danceability      1989 non-null   float64
 12  energy            1989 non-null   float64
 13  instrumentalness  1989 non-null   float64
 14  liveness          1989 non-null   float64
 15  loudness          1989 non-null   float64
 16  speechiness       1989 non-null   float64
 17  tempo             1989 non-null   float64
 18  valence           1989 non-null   float64
 19  favorite          1989 non-null   int64
dtypes: bool(1), float64(9), int64(6), object(4)
memory usage: 297.3+ KB
```

[15]:
```python
# sort by release date ascending and remove duplicates saved in library
favorites_df['release_date'] = pd.to_datetime(favorites_df['release_date'],
 ↪format='%Y-%m-%d')
library_df['release_date'] = pd.to_datetime(library_df['release_date'],
 ↪format='%Y-%m-%d')

favorites_df.sort_values(by=['release_date'])
library_df.sort_values(by=['release_date'])

alltracks_df = pd.concat([favorites_df,library_df]).
 ↪drop_duplicates(subset=['name','artist']).reset_index(drop=True)
```

Test Data To compile a test dataset, I have taken all the songs from playlists that I frequently listen to.

[16]:
```python
ids1 = getTrackIDs(user_id, test_playlist1)
ids2 = getTrackIDs(user_id, test_playlist2)
ids3 = getTrackIDs(user_id, test_playlist3)
all_ids = list(it.chain(ids1, ids2, ids3))
id_chunks = divide_chunks(all_ids, 50)
```

[17]:
```python
# get audio features for test data tracks
test_tracks = []
for id_chunk in id_chunks:
    for track in getTrackFeatures(id_chunk):
        test_tracks.append(track)
```

[18]:
```python
# create dataframe for favorite tracks
testtracks_df = pd.DataFrame(test_tracks, columns = ['name', 'album', 'artist',
 ↪'release_date', 'length', 'popularity', 'explicit', 'key', 'mode',
 ↪'time_signature', 'acousticness', 'danceability', 'energy',
 ↪'instrumentalness', 'liveness', 'loudness', 'speechiness', 'tempo',
 ↪'valence'])
```

```
testtracks_df
```

[18]:

```
                            name                          album  \
0                      Good Days                      Good Days
1       MAZZA (feat. A$AP Rocky)       MAZZA (feat. A$AP Rocky)
2                        Sourire                Coast / Sourire
3                 Out The Window                 Out The Window
4    Having a Good Time, Sometimes  Having a Good Time, Sometimes
..                           …                              …
155                 Bad Behavior                   Bad Behavior
156                     hey girl                   Wachito Rico
157            Back to the Start              Back to the Start
158                          Sun                         Texoma
159                          $10                  Prize//Reward

               artist release_date  length  popularity  explicit  key  mode  \
0                 SZA   2020-12-25  279204          84      True    1     0
1            slowthai   2021-01-05  171866           0      True    4     0
2           bad tuner   2020-12-08  223173          38     False    0     1
3          Lo Village   2020-12-01  232600          53      True    1     1
4               Bakar   2020-12-24  177073          57     False    5     1
..                …            …       …           …         …  …    …
155      Austin Millz   2019-11-14  202222          53     False    0     0
156         boy pablo   2020-10-23  187000          61     False    6     1
157              KALI   2020-11-12  203261          56     False    5     1
158  Herrick & Hooley   2016-05-04  277340          46      True    1     1
159      Good Morning   2018-05-11   89508          56     False    4     1

     time_signature  acousticness  danceability  energy  instrumentalness  \
0                 4       0.49900         0.436   0.655          0.000008
1                 4       0.07910         0.672   0.630          0.000000
2                 4       0.02080         0.914   0.344          0.005860
3                 4       0.28400         0.674   0.822          0.000025
4                 4       0.80200         0.685   0.631          0.024600
..              …             …           …      …               …
155               4       0.01610         0.581   0.522          0.000001
156               4       0.00464         0.585   0.487          0.060500
157               4       0.24000         0.651   0.628          0.000392
158               4       0.68200         0.577   0.519          0.001300
159               4       0.47500         0.624   0.596          0.203000

     liveness  loudness  speechiness    tempo  valence
0      0.6880    -8.370       0.0583  121.002    0.412
1      0.1670    -6.841       0.2850   75.286    0.354
2      0.1070    -9.431       0.3110  119.962    0.622
3      0.2650    -6.384       0.2000   97.987    0.706
4      0.1070    -7.354       0.0675   82.000    0.460
```

```
..        …       …          …       …       …
155     0.4330   -6.707      0.0718  107.962   0.499
156     0.0868  -10.038      0.0395   99.933   0.890
157     0.1090   -5.042      0.0335  111.256   0.468
158     0.0558   -8.167      0.0545   89.980   0.279
159     0.1190   -9.804      0.0314  120.969   0.896

[160 rows x 19 columns]
```

1.1.5 Inspect

[19]:
```python
pd.set_option('display.max_columns',10)
alltracks_df.sample(10)
```

[19]:
```
                                        name  \
1481                         Bad Intentions
6                                   Forever
745                              Be Alright
1756                          Jilted Lovers
1214                     Electric Relaxation
1382                           Make It To Me
1783  Angel of Small Death and the Codeine Scene
457                         WISH FEAT. KIDDO MARV
1655                              Lingering
1677                                 Up 2 U

                                         album                artist  \
1481                            Bad Intentions          Niykee Heaton
6                                      Forever            Chris Brown
745                                 Be Alright             Jada Facer
1756                  Passive Me, Aggressive You  The Naked And Famous
1214                            The Anthology  A Tribe Called Quest
1382  In The Lonely Hour (Drowning Shadows Edition)         Sam Smith
1783                                    Hozier                Hozier
457                                        ZUU          Denzel Curry
1655                                Bombs Away              Sheppard
1677                           TALKING IS HARD         WALK THE MOON

     release_date  length  …  loudness  speechiness    tempo  valence  \
1481   2014-09-09  198080  …    -4.128       0.0568  123.861    0.190
6      2007-11-02  278035  …    -4.457       0.0463  120.013    0.446
745    2018-10-01  180620  …   -11.554       0.0443  113.748    0.461
1756   2010-01-01  195773  …    -7.825       0.0394  104.940    0.347
1214   1999-10-26  226133  …    -9.201       0.2290   98.243    0.841
1382   2015-11-06  162732  …    -8.573       0.0685  149.967    0.225
1783   2014-10-07  219213  …    -5.761       0.0432   94.078    0.389
457    2019-05-31  192013  …    -6.746       0.0736   95.017    0.622
```

```
1655   2015-03-10  232853  …    -8.830       0.0279  124.004    0.543
1677   2014-12-02  201840  …    -5.140       0.0318   97.984    0.495

      favorite
1481         0
6            1
745          0
1756         0
1214         0
1382         0
1783         0
457          0
1655         0
1677         0

[10 rows x 20 columns]
```

[20]:
```
alltracks_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1941 entries, 0 to 1940
Data columns (total 20 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   name              1941 non-null   object
 1   album             1941 non-null   object
 2   artist            1941 non-null   object
 3   release_date      1941 non-null   datetime64[ns]
 4   length            1941 non-null   int64
 5   popularity        1941 non-null   int64
 6   explicit          1941 non-null   bool
 7   key               1941 non-null   int64
 8   mode              1941 non-null   int64
 9   time_signature    1941 non-null   int64
 10  acousticness      1941 non-null   float64
 11  danceability      1941 non-null   float64
 12  energy            1941 non-null   float64
 13  instrumentalness  1941 non-null   float64
 14  liveness          1941 non-null   float64
 15  loudness          1941 non-null   float64
 16  speechiness       1941 non-null   float64
 17  tempo             1941 non-null   float64
 18  valence           1941 non-null   float64
 19  favorite          1941 non-null   int64
dtypes: bool(1), datetime64[ns](1), float64(9), int64(6), object(3)
memory usage: 290.1+ KB
```

[21]:
```
alltracks_df.describe().T
```

[21]:
```
                   count           mean           std           min  \
length            1941.0  223305.038125  66547.018692  35093.000000
popularity        1941.0      44.986605     25.347601      0.000000
key               1941.0       5.160742      3.658196      0.000000
mode              1941.0       0.648635      0.477520      0.000000
time_signature    1941.0       3.961875      0.317187      1.000000
acousticness      1941.0       0.259322      0.263781      0.000041
danceability      1941.0       0.634393      0.152177      0.128000
energy            1941.0       0.609189      0.180040      0.031600
instrumentalness  1941.0       0.043435      0.156638      0.000000
liveness          1941.0       0.182678      0.142884      0.021100
loudness          1941.0      -7.345796      2.880246    -32.031000
speechiness       1941.0       0.136002      0.127786      0.023800
tempo             1941.0     119.182115     29.675970     46.489000
valence           1941.0       0.469787      0.227946      0.030400
favorite          1941.0       0.047913      0.213638      0.000000

                        25%            50%            75%         max
length            185080.0000  215306.000000  249333.000000  929346.000
popularity            31.0000      51.000000      64.000000      94.000
key                    2.0000       5.000000       8.000000      11.000
mode                   0.0000       1.000000       1.000000       1.000
time_signature         4.0000       4.000000       4.000000       5.000
acousticness           0.0379       0.163000       0.418000       0.988
danceability           0.5270       0.641000       0.748000       0.979
energy                 0.4890       0.618000       0.744000       0.983
instrumentalness       0.0000       0.000003       0.000687       0.959
liveness               0.1010       0.124000       0.213000       0.966
loudness              -8.5910      -6.862000      -5.437000      -1.304
speechiness            0.0410       0.072600       0.208000       0.856
tempo                 94.9510     117.987000     140.003000     207.969
valence                0.2920       0.458000       0.636000       0.974
favorite               0.0000       0.000000       0.000000       1.000
```

1.2 *Analysis*

1.2.1 Exploration

Distribution Comparison

[22]:
```
audio_features = alltracks_df.drop(columns =␣
↪['name','album','artist','release_date','length','popularity','explicit'])
audio_features_plot = alltracks_df.drop(columns =␣
↪['name','album','artist','release_date','length','popularity','explicit','key','time_signat
```

[23]:
```python
# Compare means of audio feature columns between favorites sample and library␣
↪sample

column_list = [x for x in audio_features.columns if x != 'favorite']
t_test_results = {}
for column in column_list:
    favorites = audio_features.where(audio_features.favorite == 1).
↪dropna()[column]
    library = audio_features.where(audio_features.favorite == 0).
↪dropna()[column]
    t_test_results[column] = ttest_ind(favorites,library, equal_var=False)
ttest_results = pd.DataFrame.from_dict(t_test_results,orient='Index')
ttest_results.columns = ['T Statistic','P-value']
ttest_results
```

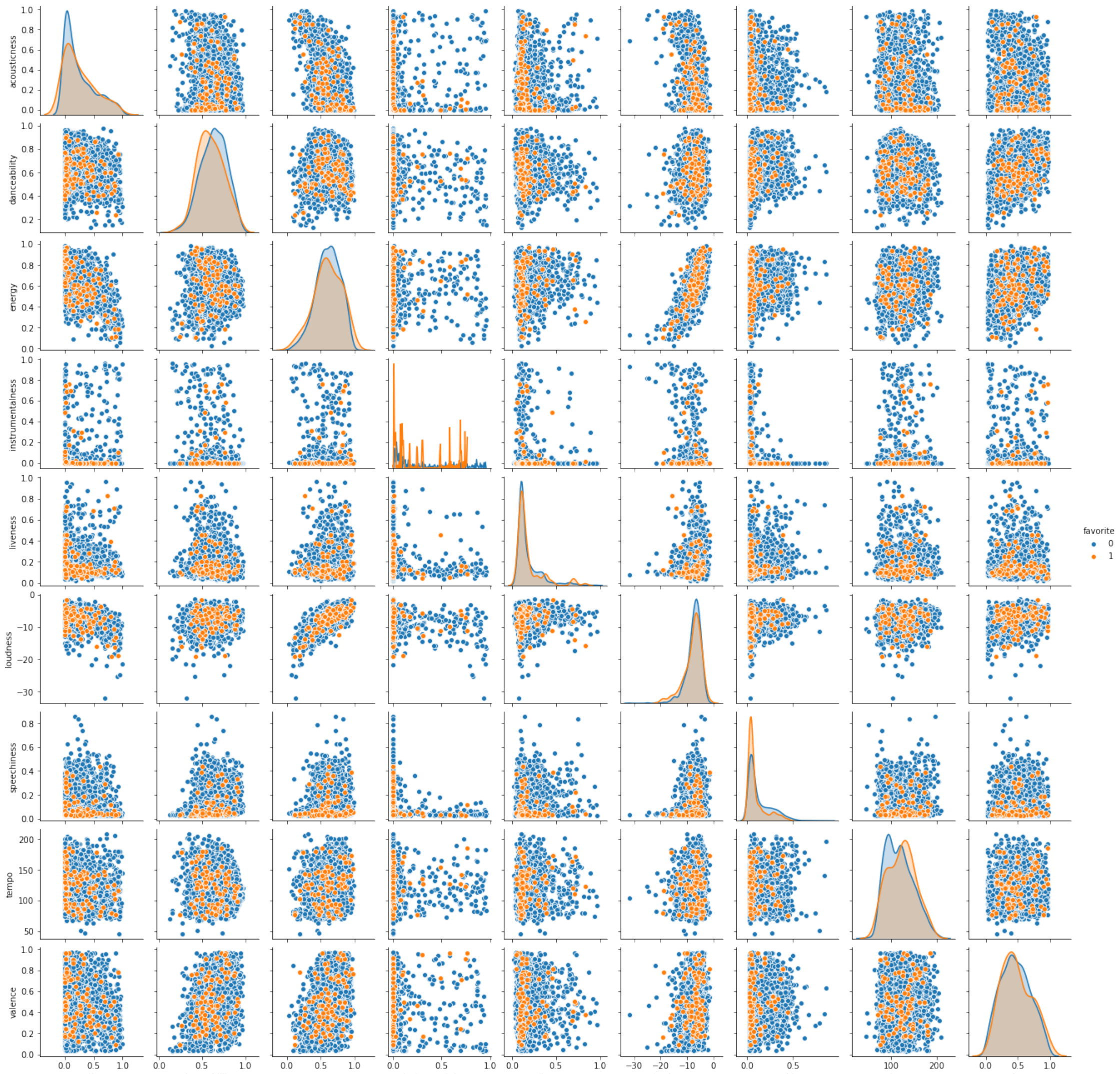
[23]:

	T Statistic	P-value
key	-0.249913	0.803164
mode	3.234475	0.001630
time_signature	-0.418374	0.676583
acousticness	0.047592	0.962135
danceability	-2.080762	0.039962
energy	-0.174840	0.861561
instrumentalness	0.913259	0.363305
liveness	0.359488	0.719993
loudness	-1.225972	0.223129
speechiness	-3.932405	0.000149
tempo	1.457504	0.148051
valence	0.804868	0.422777

It seems that the features that show a significant difference in means between favorite and non-favorite songs are *Mode*, *Danceability*, and *Speechiness*.

[24]:
```python
sns.pairplot(audio_features_plot, hue="favorite", height=2)
```

[24]: <seaborn.axisgrid.PairGrid at 0x102c07550>



Feature Correlation

[25]:
```
corr_matrix = alltracks_df.corr()
corr_matrix
```

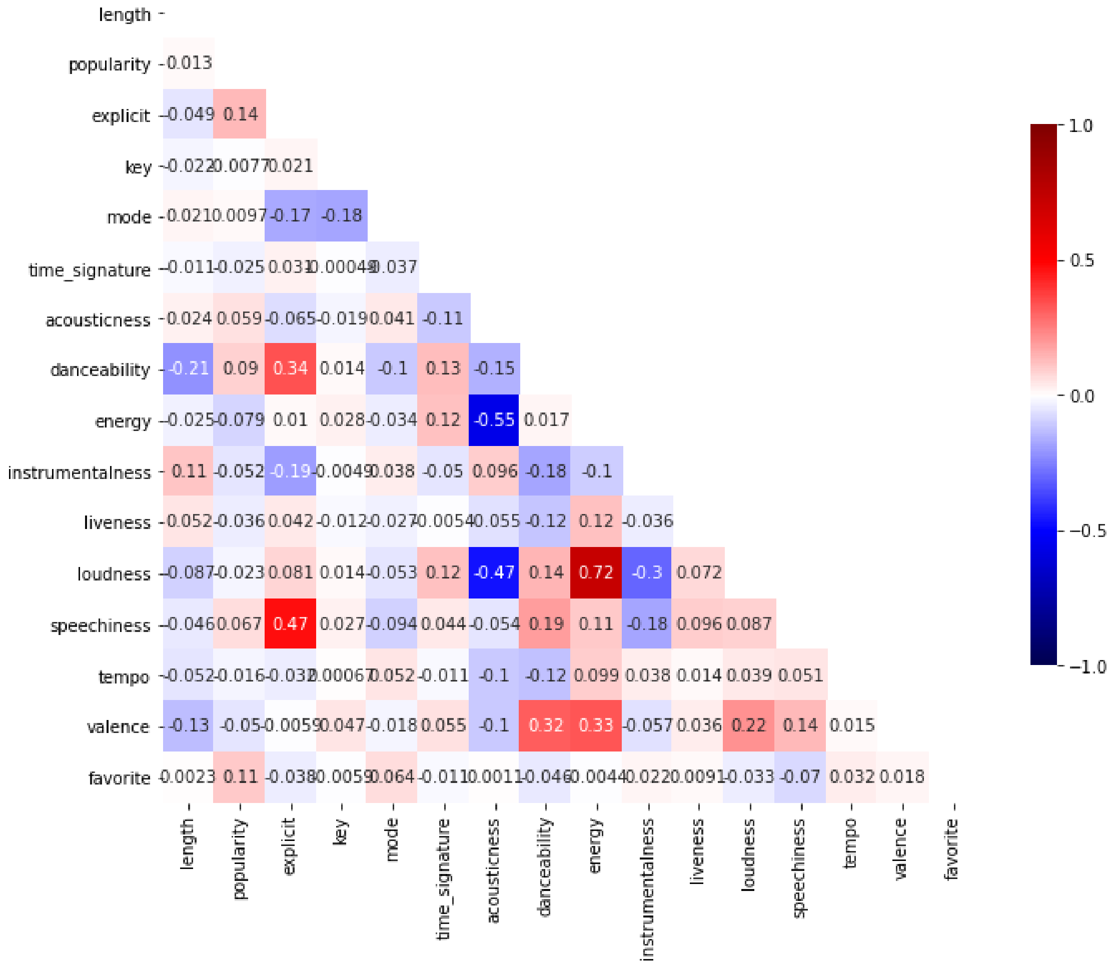
[25]:

	length	popularity	explicit	key	mode	…	\
length	1.000000	0.013001	-0.048938	-0.021743	0.020654	…	
popularity	0.013001	1.000000	0.139278	-0.007676	0.009704	…	
explicit	-0.048938	0.139278	1.000000	0.020767	-0.170378	…	
key	-0.021743	-0.007676	0.020767	1.000000	-0.179519	…	
mode	0.020654	0.009704	-0.170378	-0.179519	1.000000	…	
time_signature	-0.011455	-0.025004	0.030560	-0.000491	-0.037439	…	
acousticness	0.024063	0.058874	-0.065370	-0.018600	0.040966	…	
danceability	-0.212044	0.089505	0.337765	0.013917	-0.102734	…	

```
energy           -0.025001   -0.078822  0.010161  0.028198 -0.034319  …
instrumentalness  0.111517   -0.052496 -0.191586 -0.004891  0.038459  …
liveness          0.051562   -0.036122  0.041566 -0.011829 -0.026817  …
loudness         -0.087345   -0.023003  0.080548  0.013815 -0.052607  …
speechiness      -0.046448    0.066778  0.471524  0.027432 -0.093531  …
tempo            -0.051714   -0.016458 -0.032465  0.000674  0.052393  …
valence          -0.127759   -0.049774 -0.005945  0.047277 -0.017936  …
favorite          0.002267    0.108347 -0.038416 -0.005902  0.064054  …

                  loudness  speechiness     tempo   valence  favorite
length           -0.087345    -0.046448 -0.051714 -0.127759  0.002267
popularity       -0.023003     0.066778 -0.016458 -0.049774  0.108347
explicit          0.080548     0.471524 -0.032465 -0.005945 -0.038416
key               0.013815     0.027432  0.000674  0.047277 -0.005902
mode             -0.052607    -0.093531  0.052393 -0.017936  0.064054
time_signature    0.124790     0.044349 -0.011462  0.055343 -0.011063
acousticness     -0.467036    -0.053512 -0.101821 -0.102604  0.001083
danceability      0.141597     0.194620 -0.117213  0.317126 -0.046163
energy            0.719125     0.109306  0.098640  0.333666 -0.004430
instrumentalness -0.301785    -0.175834  0.038330 -0.057475  0.022470
liveness          0.072248     0.096280  0.014200  0.035691  0.009097
loudness          1.000000     0.086965  0.039076  0.218073 -0.032682
speechiness       0.086965     1.000000  0.051221  0.138434 -0.070464
tempo             0.039076     0.051221  1.000000  0.015114  0.032410
valence           0.218073     0.138434  0.015114  1.000000  0.018274
favorite         -0.032682    -0.070464  0.032410  0.018274  1.000000

[16 rows x 16 columns]
```

[26]:
```
mask = np.zeros_like(corr_matrix, dtype=np.bool)
mask[np.triu_indices_from(mask)]= True
```

[27]:
```
f, ax = plt.subplots(figsize=(11, 15))
heatmap = sns.heatmap(
    corr_matrix,
    mask = mask,
    square = True,
    cmap = 'seismic',
    cbar_kws = {'shrink': .4, 'ticks' : [-1, -.5, 0, 0.5, 1]},
    vmin = -1,
    vmax = 1,
    annot = True,
    annot_kws = {'size' : 10})
#add the column names as labels
ax.set_yticklabels(corr_matrix.columns, rotation = 0)
ax.set_xticklabels(corr_matrix.columns)
sns.set_style({'xtick.bottom': True}, {'ytick.left': True})
```



Loudness and *Energy* show a strong positive correlation, as does *Speechiness* and *Explicit*. *Acousticness* has a fairly strongly negative correlation with both *Energy* and *Loudness*.

1.2.2 Model Selection

[28]:
```
audio_features = audio_features.drop(columns=['favorite'])
target = alltracks_df.favorite
```

[29]:
```
# Need normalized data for distance based KNN model
audio_features_normalized = (audio_features - audio_features.mean()) /␣
 ↪audio_features.std()
audio_features_normalized
```

[29]:
```
        key      mode  time_signature  acousticness  danceability  …  \
0  0.502777  0.735813        0.120196     -0.748812     -1.329982  …
1 -0.864017  0.735813        0.120196      0.616718      0.240554  …
2  1.596213 -1.358341        0.120196     -0.414443     -1.067131  …
```

```
3     0.502777  0.735813        0.120196      1.803310     -1.007989  ...
4    -1.410734  0.735813        0.120196     -0.907277     -1.395695  ...
...        ...       ...             ...           ...           ...  ...
1936  1.596213 -1.358341       -3.032518     -0.853065     -0.659711  ...
1937 -1.137375 -1.358341        0.120196     -0.982779     -1.099987  ...
1938 -1.410734  0.735813        0.120196     -0.980076     -1.014560  ...
1939 -0.317299  0.735813        0.120196     -0.875811     -1.402266  ...
1940  1.322854  0.735813        0.120196     -0.976539     -1.067131  ...

      liveness  loudness  speechiness     tempo   valence
0     1.457981  0.158596    -0.736404  0.373834 -1.310783
1    -0.550646 -0.162210    -0.690233 -0.181059 -0.016616
2    -0.692720 -1.822484    -0.840484  0.269069 -1.310783
3     4.544408 -2.910933    -0.740317  0.124710 -0.415833
4     3.522598  0.977276    -0.266086  1.507849 -0.657119
...        ...       ...          ...       ...       ...
1936 -0.473660  0.013470    -0.737969  0.330095 -1.073885
1937 -0.557645  0.406839    -0.791183  1.457505  0.522986
1938 -0.403673  0.952625    -0.753620 -0.781175  0.952913
1939  1.968886  0.770002    -0.226958  0.500233  0.079898
1940  0.401177  0.266573    -0.641714  0.702652 -0.569379

[1941 rows x 12 columns]
```

[30]:
```
audio_features_train, audio_features_test, target_train, target_test =␣
↪train_test_split(audio_features_normalized, target, test_size = 0.25)
```

[31]:
```
audio_features_train.head()
```

[31]:
```
           key      mode  time_signature  acousticness  danceability  ...  \
1650 -1.410734 -1.358341        0.120196      0.495405      0.470549  ...
1847 -1.137375 -1.358341        0.120196     -0.735164      0.194555  ...
188   0.776136  0.735813        0.120196     -0.592621      0.332552  ...
759   0.776136  0.735813        0.120196     -0.982793     -1.927968  ...
1247 -1.137375  0.735813        0.120196      0.101137     -0.068296  ...

      liveness  loudness  speechiness     tempo   valence
1650  1.569960  0.388785    -0.830311 -0.779288 -0.455316
1847 -0.473660  0.824164    -0.765359  0.095528  0.031641
188   0.583144 -0.303170     0.125194  1.812237  1.396001
759  -0.864888 -1.096158    -0.733274  0.073423  0.071124
1247 -0.067736 -0.309419     0.813846 -1.432510  0.501051

[5 rows x 12 columns]
```

[32]:
```
target_train.head()
```

[32]:
```
1650    0
1847    0
188     0
759     0
1247    0
Name: favorite, dtype: int64
```

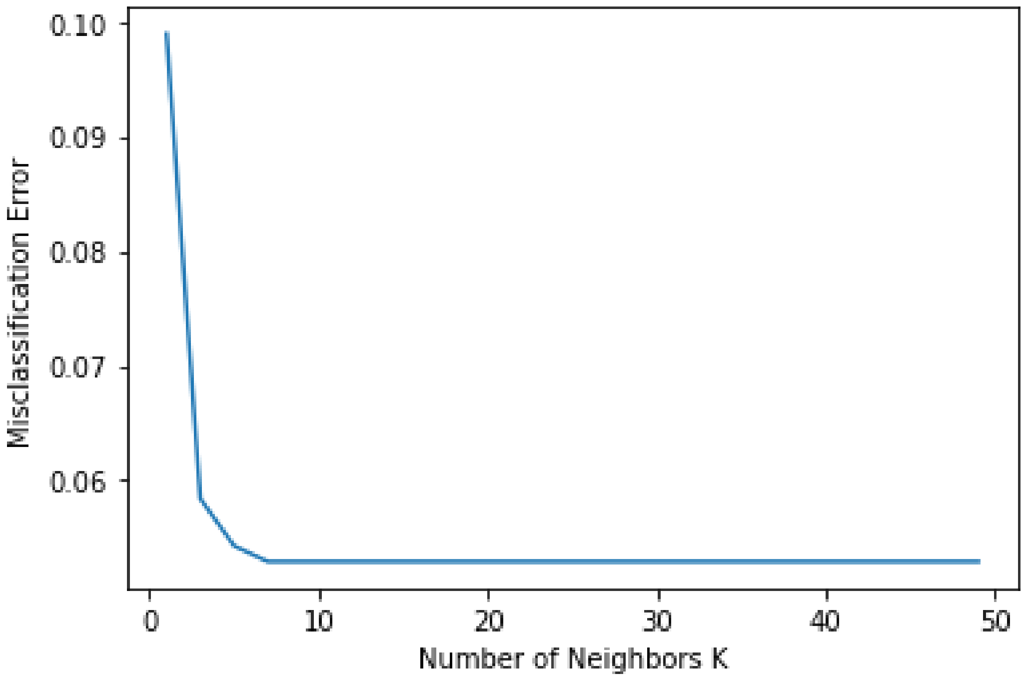
[33]:
```
print(audio_features_train.shape, audio_features_test.shape)
print(target_train.shape, target_test.shape)
```

```
(1455, 12) (486, 12)
(1455,) (486,)
```

K-Nearest Neighbors

[34]:
```
# Finding the optimal k value for the KNN model

neighbors = list(range(1, 50, 2))
cv_scores = []

for k in neighbors:
    knn = KNeighborsClassifier(n_neighbors=k)
    scores = cross_val_score(knn, audio_features_train, target_train, cv=10,␣
 ↪scoring='accuracy')
    cv_scores.append(scores.mean())

mse = [1 - x for x in cv_scores]

optimal_k = neighbors[mse.index(min(mse))]
print("The optimal number of neighbors is {}".format(optimal_k))

# plot misclassification error vs k
plt.plot(neighbors, mse)
plt.xlabel("Number of Neighbors K")
plt.ylabel("Misclassification Error")
plt.show()
```

```
The optimal number of neighbors is 7
```



[35]:
```
# Nearest Neighbors
knneighbors = KNeighborsClassifier(n_neighbors=optimal_k, weights='distance')
knneighbors.fit(audio_features_train, target_train)
```

[35]:
```
KNeighborsClassifier(n_neighbors=7, weights='distance')
```

[36]:
```
print(knneighbors.predict(audio_features_test))
print(target_test.values)
```

```
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

```
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1
 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0
 0 0 0 0 0]
```

[37]:
```python
knneighbors.predict_proba(audio_features_test)
```

[37]:
```
array([[0.73026587, 0.26973413],
       [0.55769197, 0.44230803],
       [1.        , 0.        ],
       [0.85822539, 0.14177461],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.87785398, 0.12214602],
       [0.86959838, 0.13040162],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.85920433, 0.14079567],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.86761414, 0.13238586],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.86966365, 0.13033635],
       [1.        , 0.        ],
```

```
[0.84694929, 0.15305071],
[0.86852237, 0.13147763],
[0.71975776, 0.28024224],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.81485708, 0.18514292],
[1.        , 0.        ],
[1.        , 0.        ],
[0.72825762, 0.27174238],
[0.84207679, 0.15792321],
[1.        , 0.        ],
[0.86424294, 0.13575706],
[1.        , 0.        ],
[1.        , 0.        ],
[0.73750349, 0.26249651],
[1.        , 0.        ],
[1.        , 0.        ],
[0.88231642, 0.11768358],
[1.        , 0.        ],
[0.88154589, 0.11845411],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87611065, 0.12388935],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.7040379 , 0.2959621 ],
[0.87746846, 0.12253154],
[0.74159344, 0.25840656],
[1.        , 0.        ],
[0.70319467, 0.29680533],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.88618822, 0.11381178],
[0.84260031, 0.15739969],
[1.        , 0.        ],
[1.        , 0.        ],
[0.75365152, 0.24634848],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
```

```
[0.86064204, 0.13935796],
[1.        , 0.        ],
[0.82553843, 0.17446157],
[0.84961966, 0.15038034],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.88239008, 0.11760992],
[0.59421059, 0.40578941],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87851183, 0.12148817],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86166514, 0.13833486],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.88204501, 0.11795499],
[0.88222744, 0.11777256],
[1.        , 0.        ],
[0.86691177, 0.13308823],
[1.        , 0.        ],
[0.818587  , 0.181413  ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86980643, 0.13019357],
[1.        , 0.        ],
[1.        , 0.        ],
[0.84803785, 0.15196215],
[0.82319149, 0.17680851],
[0.83202888, 0.16797112],
[1.        , 0.        ],
[0.73580813, 0.26419187],
[1.        , 0.        ],
[1.        , 0.        ],
```

```
[1.        , 0.        ],
[0.75073907, 0.24926093],
[1.        , 0.        ],
[0.        , 1.        ],
[0.86466402, 0.13533598],
[0.86895193, 0.13104807],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86511875, 0.13488125],
[0.80221948, 0.19778052],
[0.84809542, 0.15190458],
[1.        , 0.        ],
[0.71655134, 0.28344866],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.82801524, 0.17198476],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86414056, 0.13585944],
[0.85451685, 0.14548315],
[1.        , 0.        ],
[0.68609367, 0.31390633],
[1.        , 0.        ],
[0.80895448, 0.19104552],
[0.83850048, 0.16149952],
[0.8710946 , 0.1289054 ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87047555, 0.12952445],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
```

```
[0.85717846, 0.14282154],
[1.        , 0.        ],
[0.86076767, 0.13923233],
[1.        , 0.        ],
[1.        , 0.        ],
[0.85430581, 0.14569419],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86865278, 0.13134722],
[1.        , 0.        ],
[1.        , 0.        ],
[0.8562165 , 0.1437835 ],
[0.85057207, 0.14942793],
[0.86845525, 0.13154475],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.53587963, 0.46412037],
[1.        , 0.        ],
[0.87620494, 0.12379506],
[0.74488894, 0.25511106],
[1.        , 0.        ],
[0.85965522, 0.14034478],
[0.56328125, 0.43671875],
[1.        , 0.        ],
[0.841978  , 0.158022  ],
[0.80042676, 0.19957324],
[1.        , 0.        ],
[1.        , 0.        ],
[0.8690147 , 0.1309853 ],
[1.        , 0.        ],
[0.68999795, 0.31000205],
[0.88065784, 0.11934216],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86276504, 0.13723496],
[0.84604173, 0.15395827],
[0.76680167, 0.23319833],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
```

```
[1.        , 0.        ],
[1.        , 0.        ],
[0.8605379 , 0.1394621 ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.        , 1.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.85490225, 0.14509775],
[1.        , 0.        ],
[0.86198199, 0.13801801],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.82905372, 0.17094628],
[1.        , 0.        ],
[1.        , 0.        ],
[0.72263709, 0.27736291],
[1.        , 0.        ],
[0.84130241, 0.15869759],
[0.85487633, 0.14512367],
[1.        , 0.        ],
[0.8678847 , 0.1321153 ],
[0.86403817, 0.13596183],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86163232, 0.13836768],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.8132598 , 0.1867402 ],
[1.        , 0.        ],
[0.88018861, 0.11981139],
[1.        , 0.        ],
[0.85166048, 0.14833952],
[0.84750117, 0.15249883],
```

```
[1.        , 0.        ],
[0.80178786, 0.19821214],
[1.        , 0.        ],
[0.71661998, 0.28338002],
[1.        , 0.        ],
[1.        , 0.        ],
[0.88108622, 0.11891378],
[1.        , 0.        ],
[0.86021647, 0.13978353],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86664234, 0.13335766],
[1.        , 0.        ],
[0.86613289, 0.13386711],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.75127715, 0.24872285],
[1.        , 0.        ],
[0.85093304, 0.14906696],
[1.        , 0.        ],
[0.84736237, 0.15263763],
[1.        , 0.        ],
[0.80590298, 0.19409702],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87322601, 0.12677399],
[0.86456255, 0.13543745],
[1.        , 0.        ],
```

```
[1.        , 0.        ],
[1.        , 0.        ],
[0.74259487, 0.25740513],
[1.        , 0.        ],
[1.        , 0.        ],
[0.8056994 , 0.1943006 ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.83816047, 0.16183953],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86932804, 0.13067196],
[0.86088756, 0.13911244],
[1.        , 0.        ],
[0.86705579, 0.13294421],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86808076, 0.13191924],
[0.856413  , 0.143587  ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87632307, 0.12367693],
[1.        , 0.        ],
[0.85789467, 0.14210533],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86784822, 0.13215178],
[1.        , 0.        ],
[0.73449293, 0.26550707],
[0.72074926, 0.27925074],
[0.8784556 , 0.1215444 ],
[0.85515561, 0.14484439],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.57560855, 0.42439145],
```

```
[1.        , 0.        ],
[1.        , 0.        ],
[0.56782757, 0.43217243],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.8623355 , 0.1376645 ],
[0.71219187, 0.28780813],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87143326, 0.12856674],
[0.85764498, 0.14235502],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.85597841, 0.14402159],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.85784848, 0.14215152],
[1.        , 0.        ],
[0.84390508, 0.15609492],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.66121661, 0.33878339],
[1.        , 0.        ],
[0.85975373, 0.14024627],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.83997636, 0.16002364],
[0.85408259, 0.14591741],
[0.86484784, 0.13515216],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.73273609, 0.26726391],
[0.85973011, 0.14026989],
```

```
[0.89744289, 0.10255711],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.82938616, 0.17061384],
[0.73092992, 0.26907008],
[0.85563769, 0.14436231],
[1.        , 0.        ],
[0.85710579, 0.14289421],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86926508, 0.13073492],
[0.87203967, 0.12796033],
[1.        , 0.        ],
[1.        , 0.        ],
[0.83345802, 0.16654198],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86750867, 0.13249133],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[1.        , 0.        ],
[0.87438089, 0.12561911],
[1.        , 0.        ],
[1.        , 0.        ],
[0.83954346, 0.16045654],
[0.87126562, 0.12873438],
[1.        , 0.        ],
[0.87220677, 0.12779323],
[1.        , 0.        ],
[0.87170775, 0.12829225],
[0.87274323, 0.12725677],
[0.81602305, 0.18397695],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86209301, 0.13790699],
[1.        , 0.        ],
[0.86728091, 0.13271909],
[1.        , 0.        ],
[1.        , 0.        ],
[0.86396022, 0.13603978],
[1.        , 0.        ],
[1.        , 0.        ],
```

```
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.83532725, 0.16467275],
       [0.87139728, 0.12860272],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.88007799, 0.11992201],
       [0.86907673, 0.13092327],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.7121798 , 0.2878202 ],
       [1.        , 0.        ],
       [0.86025808, 0.13974192],
       [0.8474394 , 0.1525606 ],
       [1.        , 0.        ],
       [0.87283771, 0.12716229],
       [0.81651271, 0.18348729],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [0.87733128, 0.12266872],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ],
       [1.        , 0.        ]])
```

[38]:
```
knneighbors.score(audio_features_test, target_test)
```

[38]:
```
0.9629629629629629
```

[39]:
```
# Repeated Stratified K Fold Cross Validation
results_rstratifiedk = cross_val_score(knneighbors, audio_features, target,␣
 ↪cv=RepeatedStratifiedKFold(n_splits=5, n_repeats=10))
results_rstratifiedk.mean()
```

[39]:
```
0.9505420984284315
```

Random Forest

All Audio Features

[40]:
```
# Features
features =␣
↪alltracks_df[['key','mode','time_signature','acousticness','danceability','energy','instrum
# Target
target = alltracks_df['favorite']
```

[41]:
```
# Split into training and testing
features_train, features_test, target_train, target_test =␣
↪train_test_split(features, target, test_size = 0.3)
```

[42]:
```
rf = RandomForestClassifier(n_estimators=100)
rf.fit(features_train, target_train)
```

[42]:
```
RandomForestClassifier()
```

[43]:
```
target_predict = rf.predict(features_test)
print("Accuracy:",metrics.accuracy_score(target_test, target_predict))
```

```
Accuracy: 0.9502572898799314
```

Feature Importance

[44]:
```
feature_imp = pd.Series(rf.feature_importances_,index=features.columns).
↪sort_values(ascending=False)
feature_imp
```

[44]:
```
valence             0.116661
tempo               0.115147
acousticness        0.113149
energy              0.107687
danceability        0.101360
speechiness         0.100204
liveness            0.098245
loudness            0.097038
instrumentalness    0.071664
key                 0.054248
time_signature      0.012556
mode                0.012040
dtype: float64
```

[45]:
```
results_rf = cross_val_score(rf, features, target,␣
↪cv=RepeatedStratifiedKFold(n_splits=5, n_repeats=10))
results_rf.mean()
```

[45]:
```
0.9511594625394215
```

Below, I am testing the model using different subsets of audio features based on my own estimation of most influential features, feature importance as stated by the model, weak correlation, and statistically significant differences between favorite and non-favorite populations.

Subset 1 : My Most Influential Features

[46]:
```python
audio_features1 = features[['valence','energy','key']]
```

[47]:
```python
rf.fit(audio_features1, target)
```

[47]:
```
RandomForestClassifier()
```

[48]:
```python
print(audio_features1.columns)
rf.feature_importances_
```

```
Index(['valence', 'energy', 'key'], dtype='object')
```

[48]:
```
array([0.46562529, 0.45133648, 0.08303823])
```

[49]:
```python
results1_rf = cross_val_score(rf, audio_features1, target,␣
 ↪cv=RepeatedStratifiedKFold(n_splits=5, n_repeats=10))
results1_rf.mean()
```

[49]:
```
0.9495115681233932
```

Subset 2 : Features with Significantly Different Means Between Favorite and Non-Favorite Tracks

[50]:
```python
audio_features2 = features[['mode','danceability','speechiness']]
```

[51]:
```python
rf.fit(audio_features2, target)
```

[51]:
```
RandomForestClassifier()
```

[52]:
```python
print(audio_features2.columns)
rf.feature_importances_
```

```
Index(['mode', 'danceability', 'speechiness'], dtype='object')
```

[52]:
```
array([0.00873677, 0.47624507, 0.51501816])
```

[53]:
```python
results2_rf = cross_val_score(rf, audio_features2, target,␣
 ↪cv=RepeatedStratifiedKFold(n_splits=5, n_repeats=10))
results2_rf.mean()
```

[53]:
```
0.9468311557522594
```

Subset 3 : Most Important Features from Previous Models

[54]:
```python
audio_features3 =␣
 ↪features[['danceability','speechiness','valence','energy','loudness','tempo']]
```

[55]:
```
rf.fit(audio_features3, target)
```

[55]:
```
RandomForestClassifier()
```

[56]:
```
print(audio_features3.columns)
rf.feature_importances_
```

```
Index(['danceability', 'speechiness', 'valence', 'energy', 'loudness',
       'tempo'],
      dtype='object')
```

[56]:
```
array([0.14928249, 0.1545399 , 0.15738248, 0.17424129, 0.18579441,
       0.17875944])
```

[57]:
```
results3_rf = cross_val_score(rf, audio_features3, target,␣
 ↪cv=RepeatedStratifiedKFold(n_splits=5, n_repeats=10))
results3_rf.mean()
```

[57]:
```
0.9509015980706542
```

I will proceed with subset 3 because it yielded the highest average score, although all of the scores were very close.

1.2.3 Prediction

[58]:
```
features_rf = audio_features3
rf_final = rf.fit(features_rf, target)
```

[59]:
```
features_test_rf =␣
 ↪testtracks_df[['danceability','speechiness','valence','energy','loudness','tempo']]
features_test_knn = testtracks_df.drop(columns =␣
 ↪['name','album','artist','release_date','length','popularity','explicit'])
```

[60]:
```
features_knn = (features - features.mean()) / features.std()
knn_final = knneighbors.fit(features_knn, target)
predictions_knn = knneighbors.predict(features_test_knn)
predictions_knn
```

[60]:
```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0])
```

[61]:
```
predictions_rf = rf.predict(features_test_rf)
predictions_rf
```

[61]:
```
array([1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0])
```

[62]:
```
testtracks_df['rf_favorite'] = predictions_rf
testtracks_df['knn_favorite'] = predictions_knn
```

1.3 *Results*

Likelihood of Each Song Becoming a Favorite

[63]:
```
rf_likelihood = rf_final.predict_proba(features_test_rf)
knn_likelihood = knn_final.predict_proba(features_test_knn)
testtracks_df['knn_likelihood'] = knn_likelihood[:,1]
testtracks_df['rf_likelihood'] = rf_likelihood[:,1]
```

[64]:
```
testtracks_df
```

[64]:
```
                              name                           album  \
0                        Good Days                       Good Days
1         MAZZA (feat. A$AP Rocky)        MAZZA (feat. A$AP Rocky)
2                           Sourire                  Coast / Sourire
3                    Out The Window                   Out The Window
4     Having a Good Time, Sometimes  Having a Good Time, Sometimes
..                             …                               …
155                   Bad Behavior                    Bad Behavior
156                       hey girl                    Wachito Rico
157              Back to the Start              Back to the Start
158                            Sun                          Texoma
159                            $10                   Prize//Reward

              artist release_date  length  …  valence  rf_favorite  \
0                SZA   2020-12-25  279204  …    0.412            1
1           slowthai   2021-01-05  171866  …    0.354            0
2          bad tuner   2020-12-08  223173  …    0.622            0
3         Lo Village   2020-12-01  232600  …    0.706            0
4              Bakar   2020-12-24  177073  …    0.460            0
..               …          …      …  …      …          …
155      Austin Millz   2019-11-14  202222  …    0.499            0
156        boy pablo   2020-10-23  187000  …    0.890            0
```

```
157             KALI   2020-11-12  203261  …    0.468            0
158  Herrick & Hooley   2016-05-04  277340  …    0.279            0
159      Good Morning   2018-05-11   89508  …    0.896            0

     knn_favorite  knn_likelihood  rf_likelihood
0               0             0.0           0.65
1               0             0.0           0.06
2               0             0.0           0.06
3               0             0.0           0.01
4               0             0.0           0.01
..            …             …            …
155             0             0.0           0.08
156             0             0.0           0.02
157             0             0.0           0.03
158             0             0.0           0.02
159             0             0.0           0.08

[160 rows x 23 columns]
```

[65]:
```python
pd.set_option('display.max_columns',200)
pd.set_option('display.max_rows',200)
likelihood =␣
↪testtracks_df[['name','album','artist','release_date','rf_likelihood']]
likelihood.sort_values(by=['rf_likelihood'], ascending=False)
```

[65]:
```
                                                  name  \
42                                    Dead Man Walking
0                                            Good Days
61   Take Care In Your Dreaming (feat. Denzel Curry…
126  The Girl On Angel Pavement - Arranged Band Ver…
125                                       Morning Sail
108                       Strange (feat. Hillary Smith)
27                                   Back on the Fence
98                                  snakes x elephants
30                                              CRISIS
17   The Chocolate Conquistadors (From Grand Theft …
66                                          Never Know
150                                Dominic's Interlude
23                                        Liberty Bell
134                                      What Once Was
136                                     Need Your Love
130                                        Malibu 1992
128            Adderall (Corvette Corvette) - Remix
86                                        Day Dreaming
33                              We Will Always Love You
135  The Stranger (feat. Sachi, Dan Reeder, Tobias …
112  You're the One for Me - Digital Farm Animals R…
```

48	CLOUDS
45	Walking Flames
120	Sorrow, Tears And Blood
131	Sunflower, Vol. 6
21	Sister
121	Moonwalking in Calabasas - YG Remix
124	Mercy Mercy Me
145	Kingston
37	BITE
52	Her Revolution
16	Atomic Vomit
79	Bheka Mina
67	Alone
7	Wildfires
60	I Don't Wanna Feel No More
155	Bad Behavior
41	Désolé
114	Night Life
159	$10
13	Play (feat. Lancey Foux)
18	2hrs
5	Delete Forever - Channel Tres Remix
28	Lovezone
44	take your time (feat. Tinashe)
15	how 2 find hope
58	Dunya (feat. LukeyWorld)
97	Interstellar Love (feat. Leon Bridges)
96	Amber
1	MAZZA (feat. A$AP Rocky)
85	J'adore
127	Dedication
59	CAUSEWAY
139	Show Me How
73	JEWELZ
2	Sourire
152	Stone
148	Cool to You
69	Can We (with Kacy Hill)
109	Boink Boink (feat. Rich The Kid, VV$ Ken)
25	3am - Toro Y Moi Remix
92	Gas
39	MANGO (feat. Adeline)
100	Anyone
104	MMXX - XII - Kölsch Remix
90	1st Time
24	Lay Up
83	Fair Chance - Floating Points Remix

118	cheers (with Wiz Khalifa)
6	Take Me Where Your Heart Is
93	raise it up!
117	My Play
8	Talk (feat. Saba)
65	fue mejor
70	Before
140	I Keep Calling
146	Easter Sunday
153	1:45AM (feat. Bearface)
141	Juno
64	MODUS
56	Door - Oklou Remix
71	So and So
137	Ode to a Conversation Stuck in Your Throat
149	Dionne (feat. Justin Vernon)
55	Rain On Me
22	1491
46	GSG
14	Parallel 4
105	Blue Lights X 216 - Machiavelli Sessions
144	Feel That Again
88	FOR THE REST OF MY LIFE
157	Back to the Start
35	Morning Bells
10	TRUST FUND BABY
84	Playin Too Much
132	Female Energy, Part 2
29	Opiate
26	The Other Lover (Little Dragon & Moses Sumney)
111	Maybe The Day Has Come
20	fuego (feat. Tyler, The Creator)
102	i don't want another sorry
156	hey girl
158	Sun
147	GOOD
154	Small Talk
123	Oreo - Mura Masa eternal mix
151	INC.
12	Road Of The Lonely Ones
142	Blinding My Vision
11	Feels Right
9	Infrunami
133	Funny Thing
19	It's a good day (to fight the system)
116	Russian Anthem
40	The Divine Chord

74	Watchem
53	Infrastructure - ESTA. Remix
75	Cherries
76	VINYL
77	Track 6 (feat. Kanye West, Anderson .Paak & Th…
32	Golden pt. 2 (feat. Mereba)
81	Fluff
82	What Moves - Yuno Remix
87	Dora
57	Feel Good
54	Sex Wax
91	Green Eyes
49	Buzzin (with Unknown Mortal Orchestra)
110	GOOD
143	Wait
63	Spells
3	Out The Window
34	feel good
51	Moments / Tides
47	HIT EM WHERE IT HURTS
36	la luz(Fín) - Buscabulla Remix
43	Garage Rooftop
38	Lifetime - Jayda G Baleen Mix
4	Having a Good Time, Sometimes
31	Breathless
115	Zaybo Just To Win
68	Peng Black Girls Remix
122	450
72	I Feel Fantastic
138	Miss Summer
113	Nada - Remix
78	do you come here often?
107	Driftn
119	Andele Andele
94	Energy (feat. Mahalia)
106	TRICKS N KIDS
99	SULA (Hardcover)
103	Love Not War (The Tampa Beat) (Show N Prove Re…
95	We're All Gonna Be Killed
50	Honey
101	Part Of The Game (feat. NLE Choppa & Rileyy La…
62	OKAY (feat. Dreezy)
89	Columbia
129	Screwed Up (Eric Hudson Remix) feat. A Boogie …
80	Julia (Deep Diving)

album \

42	Dead Man Walking
0	Good Days
61	Music Makes Me High / Take Care In Your Dreaming
126	Jewel Box
125	Morning Sail
108	Bridgerton (Covers from the Netflix Original S…
27	Back on the Fence
98	snakes x elephants
30	CRISIS
17	The Chocolate Conquistadors (From Grand Theft …
66	Never Know
150	Manic
23	Liberty Bell
134	Songs of Her's
136	Swimmer
130	How Will You Know If You Never Try
128	Adderall (Corvette Corvette) [Remix]
86	Day Dreaming
33	We Will Always Love You
135	The Stranger (feat. Sachi, Dan Reeder, Tobias …
112	You're the One for Me
48	CLOUDS
45	Walking Flames
120	Sorrow, Tears And Blood
131	Fine Line
21	Sister
121	Moonwalking in Calabasas (YG Remix)
124	Mercy Mercy Me
145	Atlanta Millionaires Club
37	BIRDSONGS, Vol. 2
52	Her Revolution / His Rope
16	The Lo-Fis
79	Off The Meds
67	Alone
7	Untitled (Black Is)
60	I Don't Wanna Feel No More
155	Bad Behavior
41	Désolé
114	Night Life
159	Prize//Reward
13	Play (feat. Lancey Foux)
18	2hrs
5	Miss Anthropocene (Rave Edition)
28	Lovezone
44	i can't go outside
15	how 2 find hope
58	Dunya (feat. LukeyWorld)

97	Interstellar Love (feat. Leon Bridges)
96	Amber
1	MAZZA (feat. A$AP Rocky)
85	A Butterfly In-between Time
127	Dedication
59	CAUSEWAY
139	Show Me How
73	JEWELZ
2	Coast / Sourire
152	Stone
148	Cool to You
69	Can We (with Kacy Hill)
109	Loner
25	3am (Toro Y Moi Remix)
92	Gas
39	MANGO (feat. Adeline)
100	Anyone
104	MMXX - XII (Kölsch Remix)
90	1st Time
24	Hues
83	Fair Chance (Floating Points Remix)
118	cheers (with Wiz Khalifa)
6	Take Me Where Your Heart Is
93	there goes the neighborhood.
117	My Play
8	Limbo (Deluxe)
65	Sin Miedo (del Amor y Otros Demonios) ∞
70	Before
140	Before
146	Dyn-O-Mite
153	1:45AM (feat. Bearface)
141	Honeybloom
64	Nectar
56	Door (Oklou Remix)
71	So and So / Areyoudown? Pt. 2
137	Ode to a Conversation Stuck in Your Throat
149	Chewing Cotton Wool
55	Let's Go Out
22	Song of Sage: Post Panic!
46	GSG
14	Parallel
105	Blue Lights X 216 (Machiavelli Sessions)
144	Feel That Again
88	FOR THE REST OF MY LIFE
157	Back to the Start
35	WULM (acoustic) / Morning Bells
10	THE ANGEL YOU DON'T KNOW

84	Playin Too Much
132	WILLOW
29	Opiate
26	The Other Lover (Little Dragon & Moses Sumney)
111	Maybe The Day Has Come
20	i can't go outside
102	i don't want another sorry
156	Wachito Rico
158	Texoma
147	GOOD
154	Indiana
123	Oreo (Mura Masa eternal mix)
151	INC.
12	Road Of The Lonely Ones
142	K. Roosevelt
11	Feels Right
9	The Lo-Fis
133	It Is What It Is
19	I (motsi)
116	Russian Anthem
40	We Will Always Love You
74	Full Circle (Deluxe)
53	Infrastructure (ESTA. Remix)
75	Girl Eats Sun
76	VINYL
77	Featuring Ty Dolla $ign
32	Golden pt. 2 (feat. Mereba)
81	Friend Goals
82	What Moves (Yuno Remix)
87	Dora
57	Feel Good
54	Sex Wax
91	Green Eyes
49	Limbo (Deluxe)
110	GOOD
143	Wait
63	Spells
3	Out The Window
34	feel good
51	Moments / Tides
47	HIT EM WHERE IT HURTS
36	la luz(Fín) [Buscabulla Remix]
43	The Shave Experiment
38	Lifetime Remixes
4	Having a Good Time, Sometimes
31	Breathless
115	Zaybo Just To Win

68	Peng Black Girls Remix
122	450
72	The Leo Sun Sets
138	Miss Summer
113	Nada (Remix)
78	do you come here often?
107	Driftn
119	Andele Andele
94	Send Them To Coventry
106	A.I. (All + In)
99	SULA
103	Love Not War (The Tampa Beat) (Show N Prove Re…
95	We're All Gonna Be Killed
50	Condition
101	Part Of The Game (feat. NLE Choppa & Rileyy La…
62	OKAY (feat. Dreezy)
89	Columbia
129	Screwed Up (Remixes)
80	Julia (Deep Diving)

	artist	release_date	rf_likelihood
42	Brent Faiyaz	2020-09-18	0.670000
0	SZA	2020-12-25	0.650000
61	The Avalanches	2020-09-14	0.300000
126	Elton John	2020-12-17	0.260000
125	Gary Franks	2020-12-16	0.252000
108	Vitamin String Quartet	2020-12-25	0.210000
27	Healy	2020-12-09	0.190000
98	Fana Hues	2020-11-20	0.170000
30	Sam Ezeh	2020-10-16	0.160000
17	BADBADNOTGOOD	2020-12-18	0.152500
66	Sports	2020-12-04	0.150000
150	Halsey	2020-01-17	0.140000
23	DARKSIDE	2020-12-21	0.140000
134	Her's	2017-05-12	0.140000
136	Tennis	2020-02-14	0.130000
130	COIN	2017-04-21	0.130000
128	Popp Hunna	2020-12-18	0.120000
86	Brijean	2020-11-11	0.120000
33	The Avalanches	2020-12-11	0.120000
135	Dijon	2020-12-18	0.118333
112	Great Good Fine Ok	2020-12-25	0.110000
48	Park Hye Jin	2020-12-09	0.110000
45	Actress	2020-09-01	0.110000
120	GoldLink	2020-12-04	0.110000
131	Harry Styles	2019-12-13	0.110000
21	TSHA	2020-08-18	0.100000

121	DDG	2020-12-18	0.100000
124	Masego	2020-12-04	0.100000
145	Faye Webster	2019-05-24	0.100000
37	Baird	2020-10-20	0.100000
52	Burial	2020-12-11	0.092500
16	Steve Lacy	2020-12-04	0.090000
79	Off The Meds	2020-11-20	0.090000
67	Q	2020-11-27	0.090000
7	SAULT	2020-06-19	0.090000
60	reggie	2020-12-02	0.090000
155	Austin Millz	2019-11-14	0.080000
41	808vic	2020-12-05	0.080000
114	Nosleepnodrugs	2020-12-25	0.080000
159	Good Morning	2018-05-11	0.080000
13	Bakar	2020-12-09	0.080000
18	tobi lou	2020-12-18	0.080000
5	Grimes	2021-01-01	0.070000
28	Rome Fortune	2020-12-18	0.070000
44	Channel Tres	2020-12-10	0.070000
15	redveil	2020-12-31	0.070000
58	GoldLink	2020-12-04	0.070000
97	The Avalanches	2020-10-29	0.070000
96	Unusual Demont	2020-08-11	0.060000
1	slowthai	2021-01-05	0.060000
85	jamesjamesjames	2020-11-20	0.060000
127	Yung Tripp	2020-12-14	0.060000
59	Zack Villere	2020-10-14	0.060000
139	Men I Trust	2018-02-28	0.060000
73	Anderson .Paak	2020-10-06	0.060000
2	bad tuner	2020-12-08	0.060000
152	Collard	2019-11-06	0.060000
148	Teenage Priest	2019-09-06	0.050000
69	Jim-E Stack	2020-10-27	0.050000
109	Tory Lanez	2020-12-22	0.050000
25	HAIM	2020-12-18	0.050000
92	Gianni Lee	2020-08-14	0.050000
39	KAMAUU	2020-09-04	0.050000
100	Justin Bieber	2021-01-01	0.050000
104	Diplo	2020-12-28	0.040000
90	Bakar	2020-10-29	0.040000
24	Fana Hues	2020-12-11	0.040000
83	Thundercat	2020-11-11	0.040000
118	blackbear	2020-12-25	0.040000
6	Q	2020-10-09	0.040000
93	grouptherapy.	2020-10-30	0.040000
117	AJR	2020-12-22	0.040000
8	Aminé	2020-12-04	0.040000

65	Kali Uchis	2020-11-18	0.040000
70	James Blake	2020-10-14	0.040000
140	James Blake	2020-10-14	0.040000
146	Zelooperz	2019-09-16	0.040000
153	No Rome	2020-07-30	0.040000
141	Choker	2018-08-03	0.040000
64	Joji	2020-09-25	0.040000
56	Caroline Polachek	2020-12-08	0.040000
71	Saba	2020-11-24	0.030000
137	Del Water Gap	2020-05-01	0.030000
149	The Japanese House	2020-08-12	0.030000
55	Bella Boo	2020-12-04	0.030000
22	Navy Blue	2020-12-22	0.030000
46	Leah Dou	2020-10-30	0.030000
14	Four Tet	2020-12-25	0.030000
105	Jorja Smith	2020-12-29	0.030000
144	Hello Yello	2018-11-08	0.030000
88	Zack Villere	2020-11-11	0.030000
157	KALI	2020-11-12	0.030000
35	Hether	2020-12-23	0.030000
10	Amaarae	2020-11-12	0.030000
84	Lo Knowles	2020-10-23	0.030000
132	WILLOW	2019-07-19	0.030000
29	Puma Blue	2020-11-17	0.030000
26	Little Dragon	2020-12-14	0.030000
111	Profit Knowledge	2020-12-25	0.030000
20	Channel Tres	2020-12-10	0.030000
102	Dax	2020-12-30	0.030000
156	boy pablo	2020-10-23	0.020000
158	Herrick & Hooley	2016-05-04	0.020000
147	Ihaterare	2020-12-21	0.020000
154	Briston Maroney	2019-05-17	0.020000
123	Tohji	2020-12-16	0.020000
151	Dori Valentine	2018-10-05	0.020000
12	Madlib	2020-12-14	0.020000
142	K. Roosevelt	2018-07-27	0.020000
11	Biig Piig	2020-11-17	0.020000
9	Steve Lacy	2020-12-04	0.020000
133	Thundercat	2020-04-03	0.020000
19	Shungudzo	2020-10-30	0.020000
116	Ski Blxst	2020-12-27	0.020000
40	The Avalanches	2020-12-11	0.020000
74	Nocturnal Sunshine	2020-12-11	0.020000
53	St. Panther	2020-12-08	0.020000
75	Hope Tala	2020-11-13	0.020000
76	BERWYN	2020-11-25	0.020000
77	Ty Dolla $ign	2020-10-23	0.020000

32	Berhana	2020-11-11	0.020000
81	Tank and The Bangas	2020-11-20	0.020000
82	LA Priest	2020-11-18	0.020000
87	Tierra Whack	2020-10-30	0.020000
57	Polo & Pan	2020-07-03	0.020000
54	Hether	2020-12-09	0.020000
91	Arlo Parks	2020-10-20	0.020000
49	Aminé	2020-12-04	0.020000
110	Ihaterare	2020-12-21	0.020000
143	Billy Lemos	2019-01-25	0.017500
63	Greentea Peng	2020-11-30	0.010000
3	Lo Village	2020-12-01	0.010000
34	Tierra Whack	2020-11-18	0.010000
51	Goth Babe	2020-08-05	0.010000
47	PawPaw Rod	2020-09-18	0.010000
36	Kali Uchis	2020-12-15	0.010000
43	Q	2020-12-11	0.010000
38	Romy	2020-12-02	0.010000
4	Bakar	2020-12-24	0.010000
31	Caroline Polachek	2020-12-17	0.010000
115	BonafideBros	2020-12-18	0.010000
68	ENNY	2020-12-01	0.010000
122	Felly	2020-12-18	0.010000
72	Serena Isioma	2020-12-02	0.010000
138	ODIE	2020-10-23	0.010000
113	Cali Y El Dandee	2020-12-17	0.010000
78	Nina Cobham	2020-11-25	0.010000
107	Disposable Impressions	2020-12-17	0.010000
119	Uk Drill	2020-12-23	0.010000
94	Pa Salieu	2020-11-13	0.010000
106	Levi Carter	2020-12-31	0.010000
99	Jamila Woods	2020-09-18	0.010000
103	Jason Derulo	2020-11-19	0.010000
95	Terrell Hines	2020-10-21	0.000000
50	salute	2019-09-23	0.000000
101	50 Cent	2020-12-30	0.000000
62	tobi lou	2020-11-10	0.000000
89	AG Club	2020-11-06	0.000000
129	Nevaeh Jolie	2020-12-18	0.000000
80	Fred again..	2020-11-23	0.000000