```
import sklearn
           from sklearn.decomposition import TruncatedSVD
           import sklearn
           from sklearn.neighbors import NearestNeighbors
  In [3]: # Load the data
           Tracks = pd.read_csv(r"C:\Users\ahmoh\OneDrive\Desktop\Ahmed Project\ML AHMED\recommndation\Track recomender\tracks.
           csv")
           Tracks.head()
  Out[3]:
              track_id comments date_created date_released engineer favorites id information
                                                                                    listens producer tags
                                                                                                            title tracks type active_year_begin
                                                                                                         AWOL -
                                 2008-11-26
                                             2009-01-05
                           0.0
                                                                              6073.0
            0
                  2.0
                                                         NaN
                                                                  4.0 1.0
                                                                                               NaN
                                                                                                     A Way
                                                                                                                  7.0 Album
                                                                                                                              2006-01-01 0:00
                                     1:44
                                                 0:00
                                                                                                          Of Life
                                                                                                         AWOL -
                                 2008-11-26
                                             2009-01-05
                  3.0
                           0.0
                                                         NaN
                                                                                    6073.0
                                                                                                                   7.0 Album
                                                                                                                              2006-01-01 0:00
                                                                  4.0 1.0
                                                                             NaN
                                                                                                         A Way
                                     1:44
                                                 0:00
                                                                                                          Of Life
                                                                                                         AWOL
                                 2008-11-26
                                             2009-01-05
            2
                           0.0
                                                                                                                              2006-01-01 0:00
                  5.0
                                                         NaN
                                                                  4.0 1.0
                                                                              6073.0
                                                                                               NaN
                                                                                                         A Way
                                                                                                                   7.0 Album
                                     1:44
                                                                                                          Of Life
                                 2008-11-26
                                             2008-02-06
                                                                                                        Constant
                 10.0
                           0.0
                                                         NaN
                                                                  4.0 6.0
                                                                                NaN 47632.0
                                                                                                                   2.0 Album
                                                                                               NaN
                                                                                                        Hitmaker
                                                 0:00
                                     1:45
                                                                           "spiritual
                                 2008-11-26
                                             2009-01-06
                                                                          songs" from
                 20.0
                                                         NaN
                                                                                     2710.0
                                                                                               NaN
                                                                                                                13.0 Album
                                                                                                                              1990-01-01 0:00
                                     1:45
                                                 0:00
                                                                               Nicky
                                                                            Cook
In [204]: #drop the columns we dont need
           Tracks_new = Tracks.drop(['track_id', 'date_released', 'engineer', 'information', 'producer', 'tags', 'active_year_end', 'a
           ssociated_labels','website','wikipedia_page','split','active_year_begin'], axis=1)
In [257]: Tracks_new.head()
Out[257]:
                          date_created favorites id listens
                                                                      title tracks
                                                                                         type
               comments
                     0.0 2008-11-26 1:44
                                         4.0 1.0 6073.0 AWOL - A Way Of Life
                                                                                        Album
                     0.0 2008-11-26 1:45
                                         4.0 6.0 47632.0
                                                            Constant Hitmaker
                                                                            2.0
                                                                                        Album
                     0.0 2008-11-26 1:45
                                         2.0 4.0 2710.0
                                                                                        Album
                                                                     Niris
            10
                    1.0 2008-11-26 1:49
                                         0.0 58.0 3331.0
                                                                     mp3
                                                                            4.0
                                                                                   Single Tracks
                    1.0 2008-11-26 1:49
                                         2.0 59.0 1681.0
                                                               Live at LACE
                                                                          2.0 Live Performance
In [258]: # remove duplicates
           Tracks_new = Tracks_new.drop_duplicates()
In [259]: # drop null values
           Tracks_new=Tracks_new.dropna(axis=0)
In [260]: Tracks_new.info()
           <class 'pandas.core.frame.DataFrame'>
           Int64Index: 13707 entries, 0 to 103504
           Data columns (total 8 columns):
            # Column
                              Non-Null Count Dtype
                               13707 non-null float64
                comments
                date_created 13707 non-null object
                favorites 13707 non-null float64
                               13707 non-null float64
                listens
                               13707 non-null float64
                               13707 non-null object
                               13707 non-null float64
                               13707 non-null object
               type
           dtypes: float64(5), object(3)
           memory usage: 963.8+ KB
In [261]: Tracks_new.head()
Out[261]:
                          date_created favorites
                                                                      title tracks
                                                                                         type
                     0.0 2008-11-26 1:44
                                          4.0 1.0 6073.0 AWOL - A Way Of Life
                                                                            7.0
                                                                                        Album
                     0.0 2008-11-26 1:45
                                          4.0 6.0 47632.0
                                                                            2.0
                                                                                        Album
                                                            Constant Hitmaker
                     0.0 2008-11-26 1:45
                                          2.0 4.0 2710.0
                                                                           13.0
                                                                                        Album
            10
                     1.0 2008-11-26 1:49
                                         0.0 58.0 3331.0
                                                                     mp3
                                                                            4.0
                                                                                   Single Tracks
            12
                     1.0 2008-11-26 1:49
                                         2.0 59.0 1681.0
                                                               Live at LACE
                                                                            2.0 Live Performance
In [262]: # group by songs
           Tracks_new.groupby('title').sum()
Out[262]:
                                                                    id listens tracks
                                               comments favorites
                    "...Through The Cracks" Mix Vol. 1
                                                            1.0 21245.0 2962.0
            "...when my heart's fire rages out of control..."
                                                             0.0 22482.0 17664.0
                                                                                33.0
                                       "32 Ford"
                                                            2.0 12717.0 11770.0
                                   "A.B.A.Y.A.M"
                                                             0.0 12708.0 2596.0
                                                    0.0
                                                                                1.0
                                   "Aeroplanes"
                                                             1.0 12698.0 1785.0
                                                             2.0 21483.0 42453.0
                                                                                8.0
                                         ロッコ
                                                    0.0
                       章4: 真っ暗闇の人生: Volume 1
                                                             1.0 22049.0 10631.0
                                                                               16.0
                       章4:真っ暗闇の人生: Volume 2
                                                            1.0 22057.0 8136.0
                                                                               15.0
                       章4:真っ暗闇の人生: Volume 3
                                                             0.0 22061.0 20715.0
                                       起死回生
                                                            0.0 17231.0 16443.0
                                                    0.0
           13309 rows × 5 columns
In [263]: Tracks_new.head()
Out[263]:
                          date_created favorites
                                                                      title tracks
               comments
                                               id listens
                                                                                         type
                     0.0 2008-11-26 1:44
                                          4.0 1.0 6073.0 AWOL - A Way Of Life
                                                                                        Album
                     0.0 2008-11-26 1:45
                                          4.0 6.0 47632.0
                                                                            2.0
                                                                                        Album
                                                            Constant Hitmaker
                     0.0 2008-11-26 1:45
                                          2.0 4.0 2710.0
                                                                           13.0
                                                                                        Album
                     1.0 2008-11-26 1:49
            10
                                         0.0 58.0 3331.0
                                                                     mp3
                                                                            4.0
                                                                                   Single Tracks
                     1.0 2008-11-26 1:49
                                         2.0 59.0 1681.0
                                                               Live at LACE
                                                                            2.0 Live Performance
           Recommendation based on total listens
In [264]: # Grouping the data based on the total listen for each song.
           listen_count = pd.DataFrame(Tracks_new.groupby('title')['listens'].sum())
           Most_Listen=listen_count.sort_values('listens', ascending=False).head()
           Most_Listen
Out[264]:
                                             listens
                                       title
                                    Entries 3564243.0
            Nameless: The Hackers RPG Soundtrack 1553769.0
                                   Cylinders 1363291.0
                          Audionautix: Acoustic 1282752.0
                             Classical Sampler 1193803.0
           Recommendation based Similarity (Total Listens)
In [213]: # knowing the top 5 songs with most listens
           Tracks_new.groupby('title')['listens'].sum().sort_values(ascending=False).head()
Out[213]: title
           Entries
                                                      3564243.0
           Nameless: The Hackers RPG Soundtrack
                                                     1553769.0
           Cylinders
                                                     1363291.0
           Audionautix: Acoustic
                                                     1282752.0
           Classical Sampler
                                                     1193803.0
           Name: listens, dtype: float64
In [267]: # creating a pivot to crosstab the songs based on the listens.
           Track_crosstab = Tracks_new.pivot_table(values='listens', index='id', columns='title', fill_value=0)
           Track_crosstab.head()
Out[267]:
                                                                         "After
                                                                          Dark"
                                                                          (Bobb
                                                                                         "Brown
                                                                                  "All or
                 "...Through
                                                                                                                          Serpeggiava
                                                                "Affairs
                                                                       Trimble's
                                                                                           Piss
                                                                                                                    эпоха
                             heart's
                                                                                Nothing"
                                                                                                Flag,
                                                                                                                           Un Freddo
                                         "A.B.A.Y.A.M" "Aeroplanes"
                                                                         Flying
                                                                                         Junkie"
                                                                                                        шизнь
                                                                                                                  лишних
                                                                                                                                    SD
                          fire rages Ford"
                                                                                                                            Viscido Di
                   Cracks"
                                                                                  Action
                                                                                                 Put
                                                                                                               музыкантов
                                                                 Heart"
                                                                        Spiders
                                                                                           split
                  Mix Vol. 1
                             out of
                                                                               Orchestra
                                                                                                Out A
                                                                                                                              Ombra
                                                                         Live in
                                                                                        cassette
                          control..."
                                                                                                                            Perenne'
                                                                       Brooklyn
                                                                        6/25/09)
             id
             1.0
                                                                                                                                  0 0
             4.0
                                      0
                                                                                                                                  0 0
                                                                                                                                  0 0
            58.0
                                0
                                      0
                                                  0
                                                                    0
                                                                                     0
                                                                                             0
                                                                                                            0
                                                                                                                       0
                                                                                                                                  0 0
                                                                                                                                  0 0
           5 rows × 13309 columns
In [268]: Track_crosstab.shape
Out[268]: (13707, 13309)
In [269]: X = Track_crosstab.T
           X.shape
Out[269]: (13309, 13707)
In [270]: # using svd to transform the data.
           SVD = TruncatedSVD(n_components=12, random_state=17)
           Track_matrix = SVD.fit_transform(X)
           Track_matrix.shape
Out[270]: (13309, 12)
In [276]: # obtaning the index of the cylinders song
           Track_names = Track_crosstab.columns
           Track_list = list(Track_names)
           Cylinders = Track_list.index('Cylinders')
           Cylinders
Out[276]: 2247
In [295]: corr_Cylinders = corr_mat[2247]
           corr_Cylinders.shape
Out[295]: (13309,)
In [296]: # finding simailr tracks to cylunder based on0.27 correlation
           list(Track_names[(corr_Cylinders<1.0) & (corr_Cylinders >0.27)])
Out[296]: ['Alice In Wonderland (1903)',
            "J'irai Molarder sur vos Urnes",
            "Live on WFMU's Infinite Distortion: May 27, 2015",
            "Live on WFMU's Three Chord Monte with Joe Belock - February 05, 2013",
            'Rodopted',
            'Root',
            'in beach\u200b,\u200bside ill\u200b-\u200bspot']
In [297]: Track_names = Track_crosstab.columns
           Track_list = list(Track_names)
           Track = Track_list.index('Classical Sampler')
           Track
Out[297]: 1948
In [298]: Track_corr = corr_mat[1948]
           Track_corr.shape
Out[298]: (13309,)
In [299]: list(Track_names[(Track_corr<1.0) & (Track_corr >0.4)])
Out[299]: ['"...when my heart\'s fire rages out of control..."',
             'AWOL - A Way Of Life',
            'Niris',
            'mp3']
In [300]: Track_names = Track_crosstab.columns
           Track_list = list(Track_names)
           Track = Track_list.index('Tour CD 2006')
           Track
Out[300]: 11287
In [301]: Track_corr = corr_mat[11287]
           Track_corr.shape
Out[301]: (13309,)
In [302]: list(Track_names[(Track_corr<1.0) & (Track_corr >0.985)])
Out[302]: ['Bad Panda #39',
             'Campfire on Bubble Mountain',
            'Jekyll Electric Effect',
            'Opus Oratorium',
             'Sit Up and Shut Down',
             'Waves of Grey - Single']
           Content based recommendation system
In [304]: Tracks_new.head()
Out[304]:
                          date_created favorites
                                             id listens
                                                                      title tracks
                                                                                         type
                     0.0 2008-11-26 1:44
                                         4.0 1.0 6073.0 AWOL - A Way Of Life
                                                                            7.0
                                                                                        Album
                                                            Constant Hitmaker
                     0.0 2008-11-26 1:45
                                         4.0 6.0 47632.0
                                                                                        Album
                     0.0 2008-11-26 1:45
                                         2.0 4.0 2710.0
                                                                     Niris
                                                                           13.0
                                                                                        Album
            10
                    1.0 2008-11-26 1:49
                                         0.0 58.0 3331.0
                                                                                   Single Tracks
                                                                     mp3
            12
                    1.0 2008-11-26 1:49
                                         2.0 59.0 1681.0
                                                               Live at LACE
                                                                          2.0 Live Performance
In [318]: # creating a test set to obtain songs that are simialar based on favorites number, listens and number of tracks
           test = [4, 3000, 2]
           X = Tracks_new[['favorites','listens','tracks']].values
           X[0:5]
Out[318]: array([[4.0000e+00, 6.0730e+03, 7.0000e+00],
                   [4.0000e+00, 4.7632e+04, 2.0000e+00],
                   [2.0000e+00, 2.7100e+03, 1.3000e+01],
                  [0.0000e+00, 3.3310e+03, 4.0000e+00],
                  [2.0000e+00, 1.6810e+03, 2.0000e+00]])
In [319]: # using nearestneighbors to fit the data
           nbrs = NearestNeighbors(n_neighbors=1).fit(X)
In [321]: # obtaining the index of song closest to the test set
           print(nbrs.kneighbors([test]))
           (array([[4.12310563]]), array([[2588]], dtype=int64))
In [323]: # obtaining the info of the closest song to the test set.
           Tracks_new.iloc[2588]
Out[323]: comments
                           2010-02-05 21:05
           date_created
           favorites
                                         5447
           id
                                         3000
           listens
```

title

type

In []:

tracks

Name: 15482, dtype: object

Wortsport

1 Album

In [2]: import pandas as pd

import numpy as np

Nan