DATA SCIENCE PROJECT

Song Recommendation System with Spotify Data Set

B Venkata Sravan Kumar (PES1UG19ME046)

Introduction:

Recommendation systems are designed to recommend objects to the user based on many different factors. These systems predict the most likely products that users are most likely to purchase and are of interest to them. Companies, such as Netflix and Amazon, use recommendation systems to help users identify the correct product or movie.

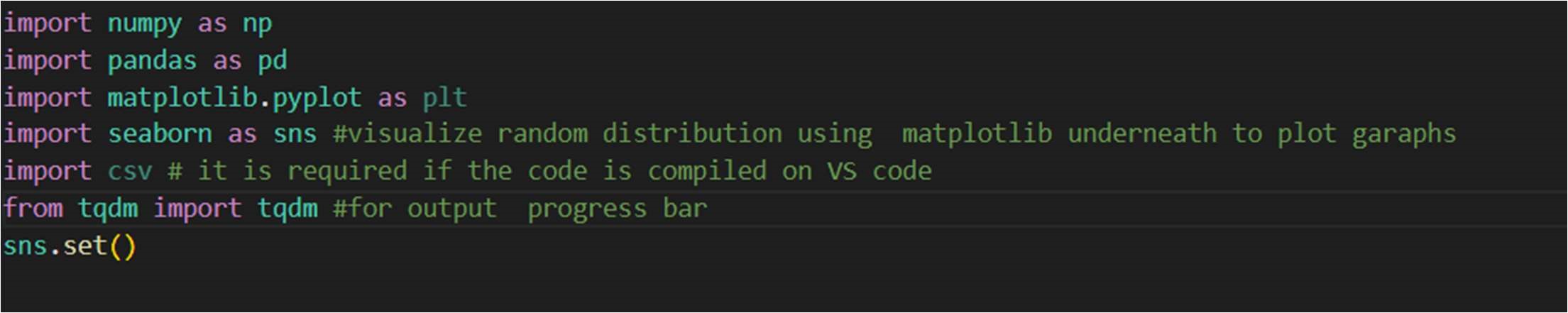
The recommendation system deals with a large volume of information present by filtering the most important information based on the data provided by a user and other factors that consider the user’s preferences and interests. It determines the match between the user and item, and correlates the similarities between users and items for recommendation. Both users and services provided have benefited from these types of systems. Quality and decision-making processes have also improved through these types of systems.

The Spotify Recommendation system recommends songs based on their different properties, such as acoustiness, energy, loudness, and tempo.

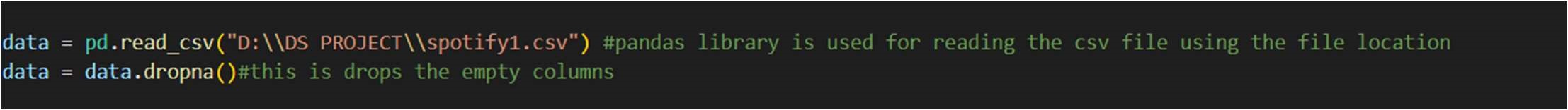
Goals:

The Purpose of this project was to create a recommendation system using the Spotify dataset. To suggest songs to users, the spotify recommendation algorithm uses collaborative filtering. To improve the user experience, collaborative filtering promotes products or services based on the similarities between users and the mentioned products or services.

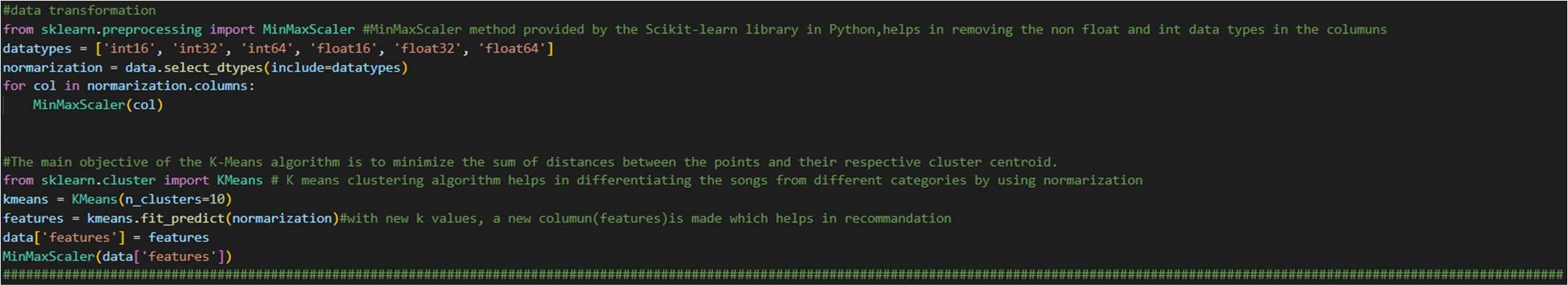
Code:



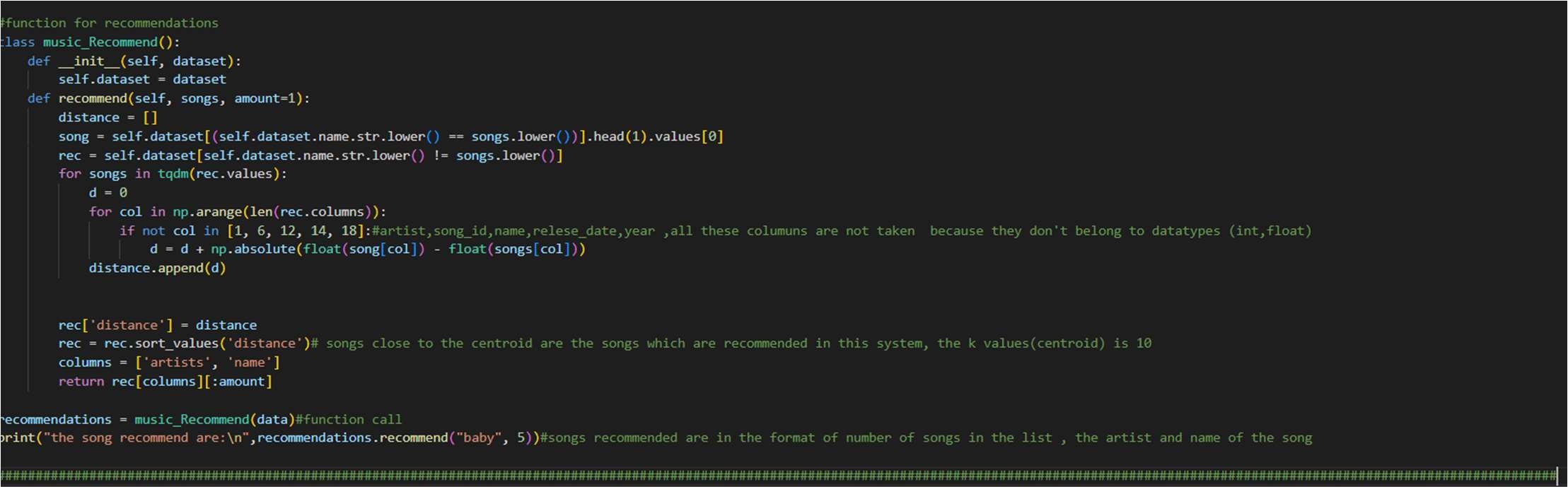
The program is written in Python. Various libraries are used to analyze and operate the datasets, such as NumPy, Pandas, Matplot, and Sklearn.



Pandas.read\_csv was used to read the csv file from the file location.



The MinMaxScaler was imported from the Sklearn Library. The MinMaxScaler method provided by the Scikit-learn library in Python helps remove the non-float and int data types in the columns. The K-means clustering algorithm helps in differentiating songs from different categories by using the normalization of data. The main objective of the K-means algorithm is to minimize the sum of the distances between the points and their respective cluster centroids.

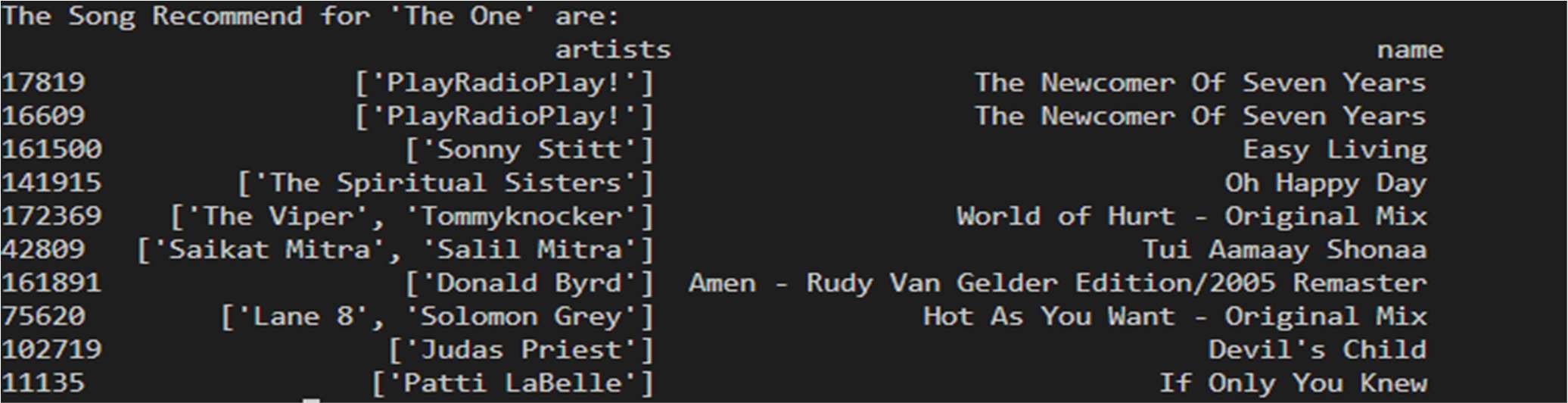


Output:

The output of the program displays five to ten songs and artists that are similar to the song input data. The songs recommended are in the format of place of songs in the list, the artist, and name of the song, and the code was tested with different songs:

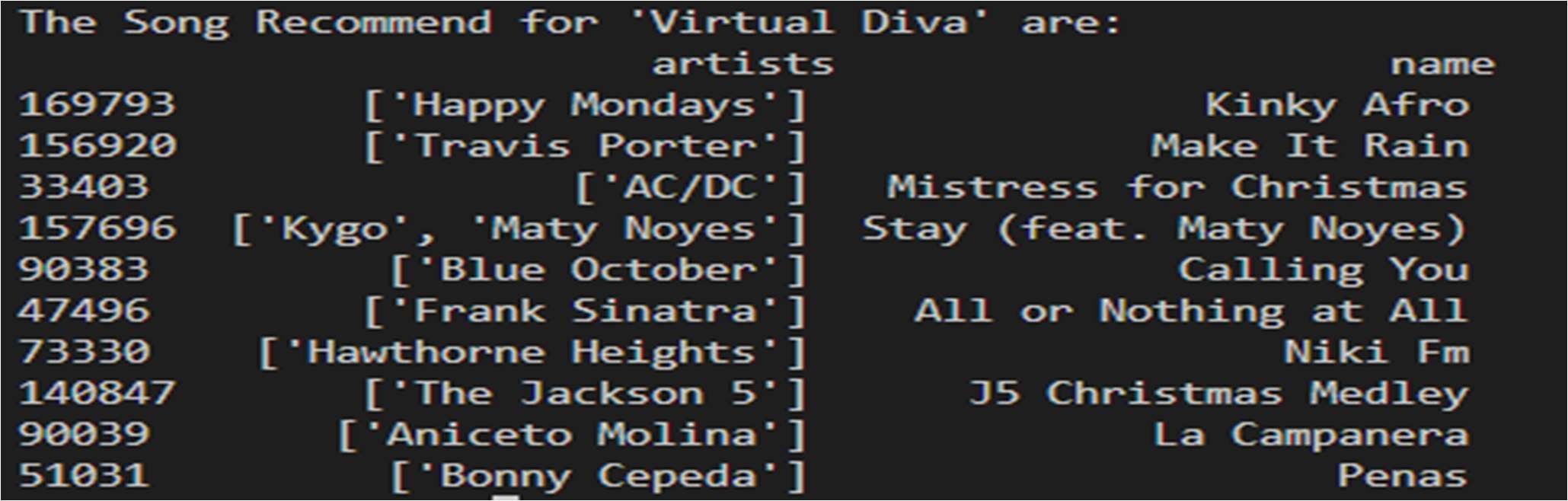
Example 1:

The Program recommends certain song for the input ‘The One’



Example 2:

The Program recommends certain song for the input ‘Virtual Diva’



Example 3:

The Program recommends certain song for the input ‘A Little More’

