

Mario Calito
Dec 15, 2022
DS 210

Final Project Write Up

For my final project I used the data set from Kaggle which includes 4 columns which consist of user id, playlist name, artist song and artist. I decided to do this project because I am very curious to see how many of the same artist are in different playlist. First, since I didn't need two columns from the dataset, I then imported the dataset into python in jupyter notebooks to clean up the data for me to perform the algorithm in rust, I have included the pynb file of my cleaning the data. I then exported the csv file to then import it into visual studio code. I created a read function to read the data into a data frame into rust.

Next since I have many repeat values, I decided to assign those values into integers. The way I did that is by creating a hash table to iterate by all the values and assign each string a key to then turn into a vector as (x,y) In order to perform breath first search algorithm. After, I then wrote a Breath First Search Algorithm to then find out which how many playlists contain similar artist. My results are then shown. I believe that my HashMap function was the most difficult part to code as it did require some thinking on how to code it and what I should include in it.