BUILDING A RECOMMENDATION SYSTEM FOR NTERTANE MUSIC APP

AUTHOR: Ojiambo Ivan, Alli Abdulateef Olamide and Mawoko Tinashe

Introduction:

Ntertane is a musical app that is primarily based in Africa and Nigeria in particular. Although the app has been in the market for two years, it has still failed to attracted many regular users from around the continent.

This failure is attributed to the fact that it can not auto suggest songs to users according to there interests and preferences.

The main aim of this project is to come up with a good recommendation system that can help to auto suggest songs that the user could be interested in listening to.

Data preparation and cleaning.

During data preparation the following actions were performed:

- We had to query and merge the data from different database tables into three data sets (tracks, statistics, users datasets).
- We removed the irrelevant columns from both datasets.
- We corrected some genre values that we felt were wrongly spelt or did not make sense.
- We also removed rows with missing values.

At the end were left with 23,523 fields in the tracks dataset, 76231 fields in statistics datasets and 4886 fields in the user

Software used

We used the following software's:

- R studio 1.2.3
- Tableau 10.4
- Python 3.0

DATA ANALYSIS

Figure 1.0 and figure 1.1 Show the number of users of the musical app in each country

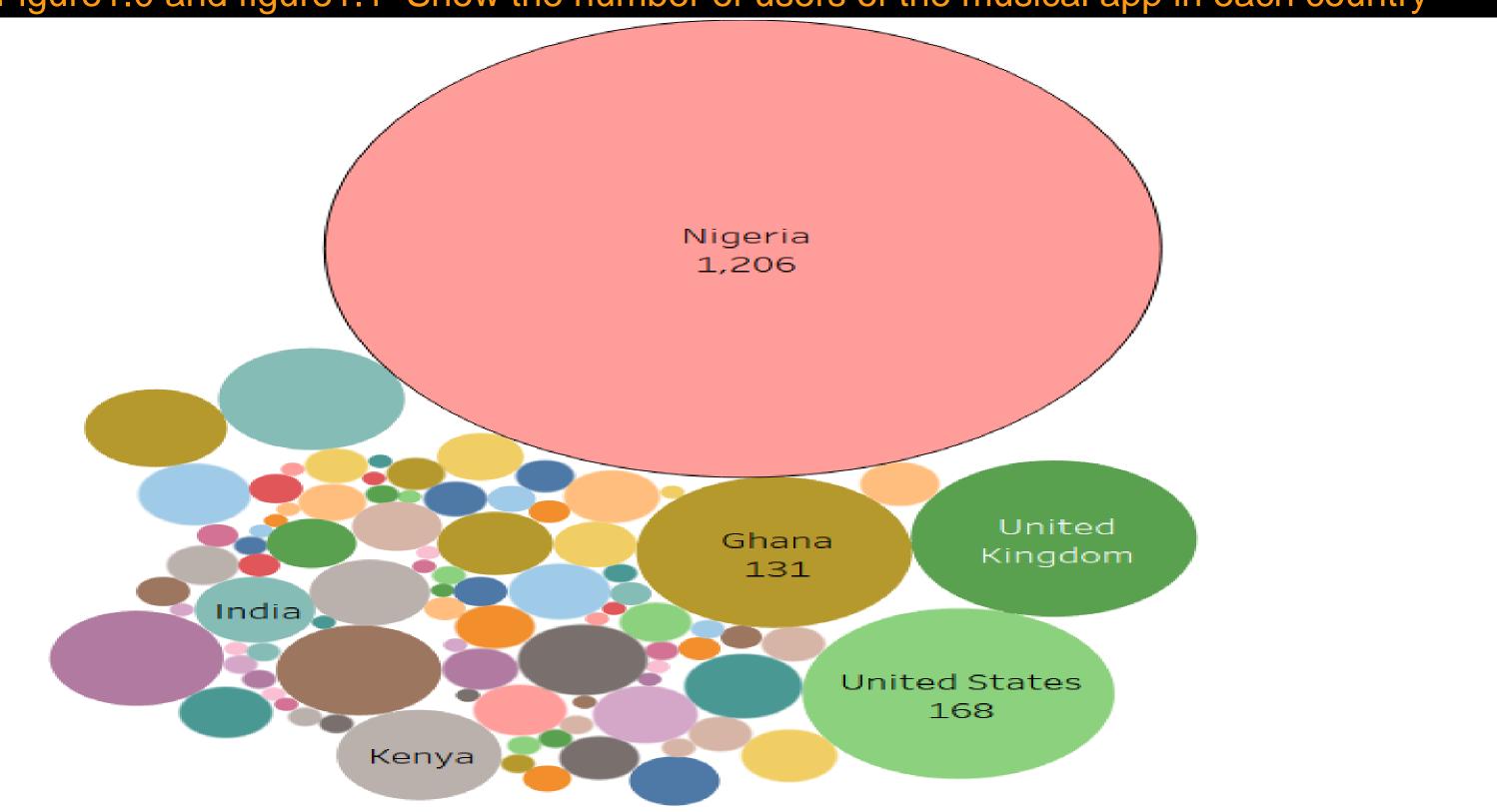


Figure 1.0

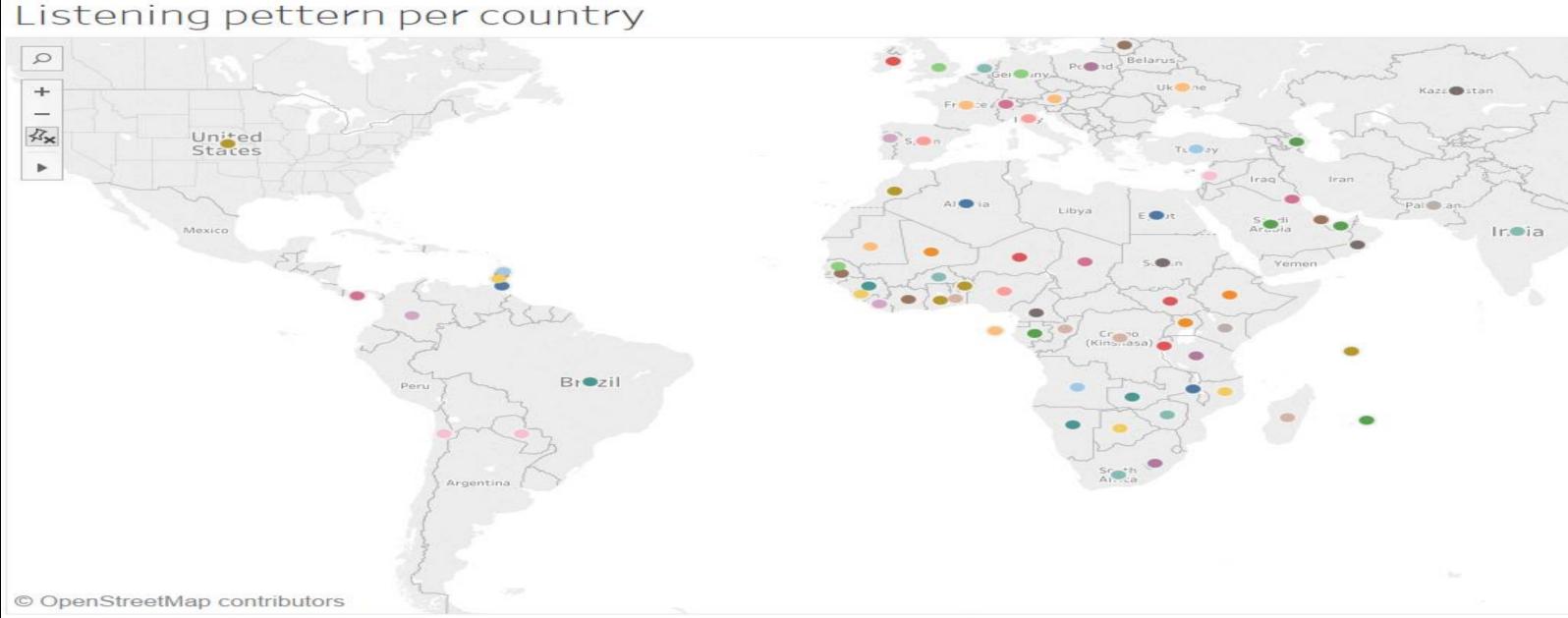


Figure 1.

From figure 1.0 and figure 1.1 we can clearly see that the majority of the listeners are from Nigeria and very few from other countries.

Figure 2.0 shows a tree map with the genre distribution of all the songs in the dataset

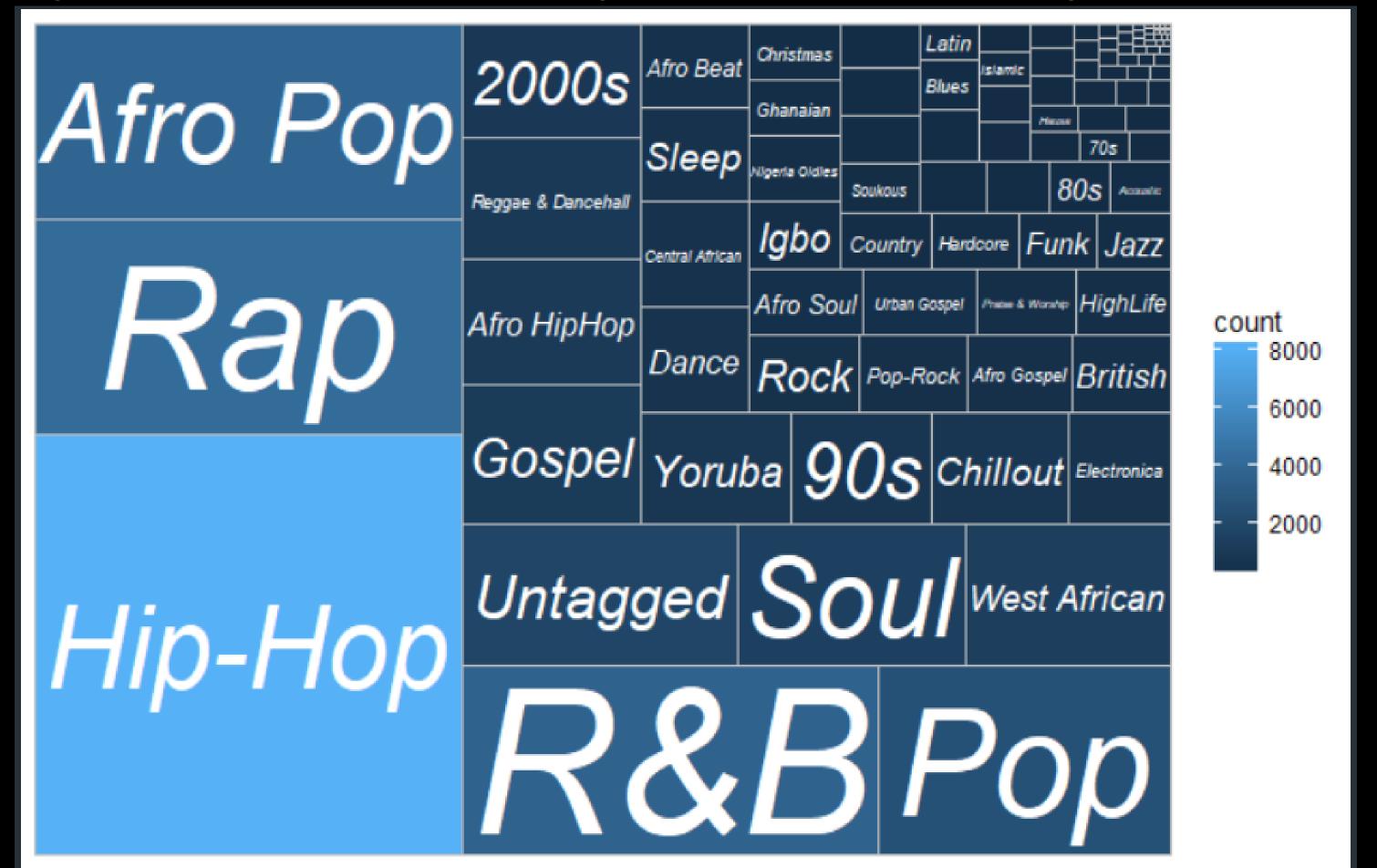


Figure2.0

From figure 2.0 we can see that majority of the songs that we have hip-Hop, Rap and Afro pop.

Figure 3.0 shows a bar graph with song genre which are most rated

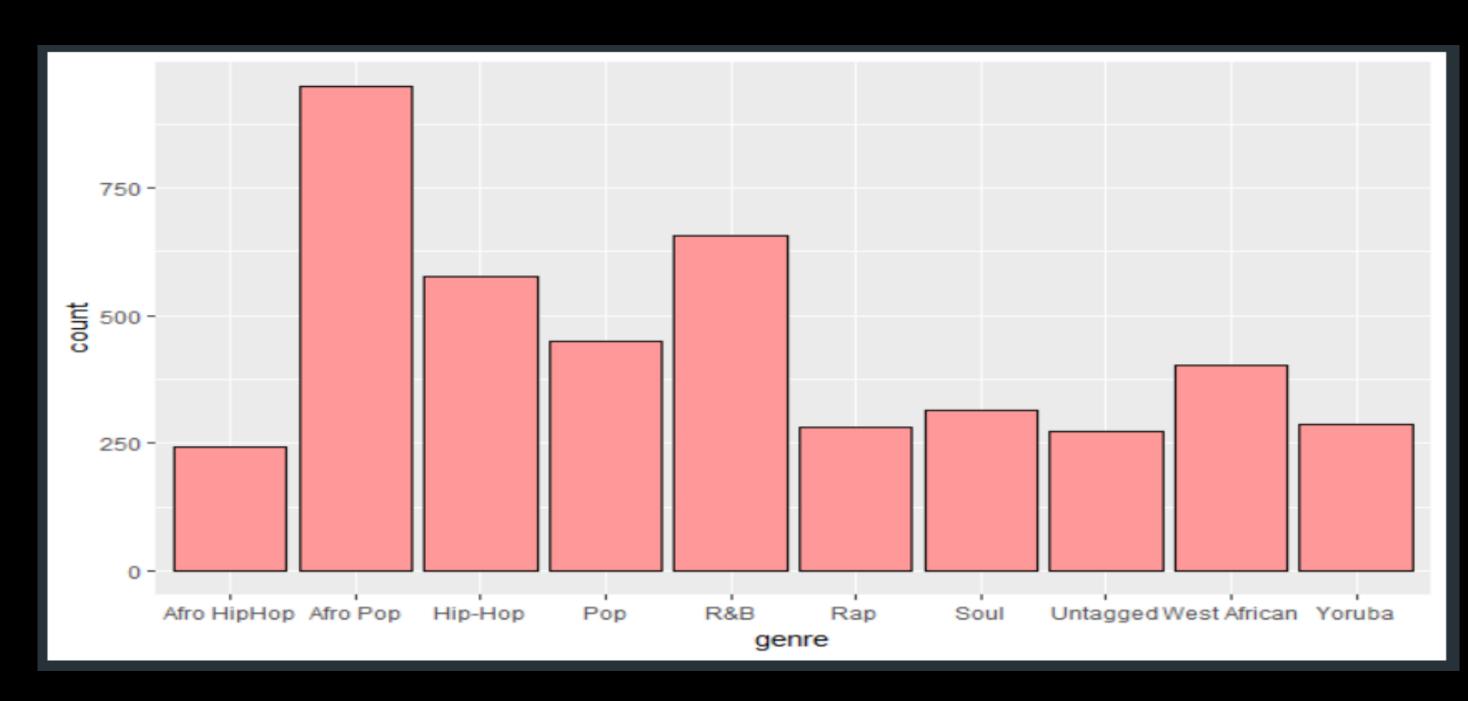


Figure 3.0

From the bar graph we see that songs with genre Afro pop are the most ranked songs followed by R&B songs

METHODOLOGY

We have done our recommendation using Matrix factorization which has been identified as one of the methods of Collaborative filtering.

We considered all users' interaction(Listen, Offline, Likes) with each track

We use R Studio to carry out data cleaning and data munging.

We decided to use python as our choice of implementation language for the actual prediction system because it seemed more promising for us to integrate it with the existing system.

It was also quite easy to get support online on how to go about creating our recommender system with Python using available packages.

We used the pandas, numpy and scipy packages from python to complete our task.

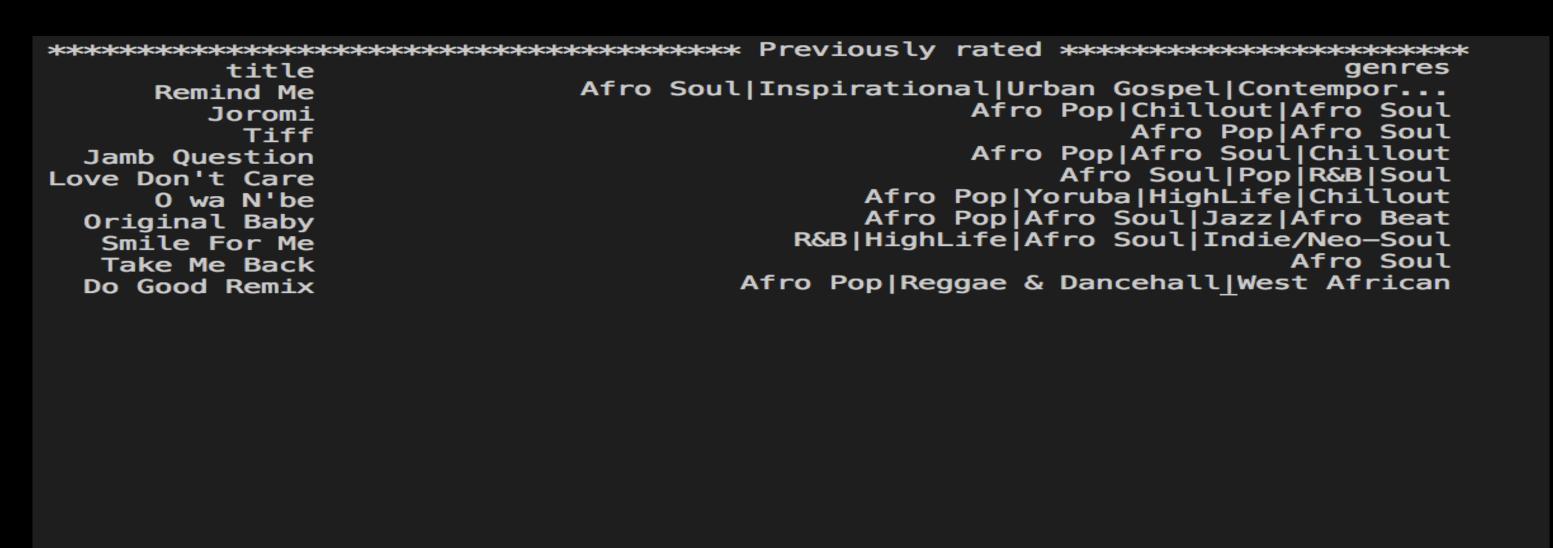
The methods that helped us achieve our task in particular are;

mean from numpy package

svds from scipy package

RESULTS

We tested our recommendation system by testing it on a user who had previously liked the following songs below



Our recommendation system predicited the following songs to the above user

	10 predicted ratings tracks not already rated.	

title	genres	
African Bad Gyal	Afro Pop West African Pop Dance	
Olowogbogboro	Gospel Urban Gospel Afro Gospel	
If (Remix)	Afro Pop West African	
Baby Na Yoka	Afro Pop HighLife Igbo Reggae & Dancehall	
Hold My Hand	Afro Pop West African	
We Go Party	Afro Pop	
TGIF	Afro Pop Yoruba	
Halleluyah	Gospel	
On Fire	Afro HipHop	
I've Come To Worship	Gospel Afro Gospel Praise & Worship	