

These link are for the datasets used for the project.

Downloading these files locally and replacing all the file paths present in the code.
Then run parts of file in ascending order.

Dataset used for segment pitches : <http://millionsongdataset.com/pages/getting-dataset/#subset>

Dataset used for lyrics:

Train:

http://millionsongdataset.com/sites/default/files/AdditionalFiles/mxm_dataset_train.txt.zip

Test:

http://millionsongdataset.com/sites/default/files/AdditionalFiles/mxm_dataset_test.txt.zip

SQLite database:

http://millionsongdataset.com/sites/default/files/AdditionalFiles/mxm_dataset.db

Python code to get database:

https://github.com/tbertinmahieux/MSongsDB/blob/master/Tasks_Demos/Lyrics/mxm_dataset_to_db.py