GROUP 4 – SPOTIFY DATA EXPLORATION

# STEPS TO RECREATE ENVIROMENT

|  |  |
| --- | --- |
| Step # | Description |
| 1 | Clone the GitHub repository to your computer (https://github.com/coreymcculley/Rice-Boot-Camp-Project-2.git) |
| 2 | Navigate to <https://www.kaggle.com/yamaerenay/spotify-dataset-19212020-160k-tracks> and download the csv containing the Spotify data. |
| 3 | Save csv file in the main project directory (contains ‘cleaning\_data.ipynb’). |
| 4 | Create “config.py” file in the same directory and assign PostgreSQL login/password in the following format  username = ‘your PostgreSQL user name’  password = ‘your PostgreSQL password’ |
| 5 | Open Postgres SQL Administrator |
| 6 | Create a database called “spotifyDB” in Postgres SQL |
| 7 | Using Jupyter Lab (or Jupyter Notebook) open file ‘cleaning\_data.ipynb’ |
| 8 | While running first lines (loading libraries), make sure all of them installed on your local machine under current virtual environment |
| 9 | Run all lines and wait for them to complete (time varies depends on computer performance) |
| 10 | Pay attention to the last two lines – first one will push our dataframe to database, and second line will call data back for confirmation |
| 11 | Open Postgres SQl Administrator, scroll to ‘spotifyDB > Schemas > Table > songs’ |
| 12 | Right click on ‘songs’ table and go to ‘Properties’, then ‘Constraints’ and assign a primary key to IDs column |
| 13 | Run ‘app.py’ from Terminal/GitBash – that will run Flask app for data distribution from PostgreSQL |
| 14 | Run command ‘python -m http.server’ from the Index.html location |
| 15 | Open your browser and navigate to localhost:8000 to open the landing page of our website, and then enjoy exploring the dataset. |