CS584 – machine learning

fall 2016

Genrify Project

Group Members:

Henri Toussaint

Victor Saint Guilhem

Benoît Lafon

Table of Contents

[Task 2](#_Toc467172373)

[Dataset 2](#_Toc467172374)

[Data source 2](#_Toc467172375)

[Target variable 2](#_Toc467172376)

[Features 2](#_Toc467172377)

[Data size 2](#_Toc467172378)

[Preprocessing 2](#_Toc467172379)

[Visualization 2](#_Toc467172380)

[Target 2](#_Toc467172381)

[Features 2](#_Toc467172382)

[Evaluation 2](#_Toc467172383)

[Performance Measure 2](#_Toc467172384)

[Classifiers 3](#_Toc467172385)

[Evaluation Strategy 3](#_Toc467172386)

[Performance Results 3](#_Toc467172387)

[Top Features 3](#_Toc467172388)

[Discussion 3](#_Toc467172389)

[Interesting/Unexpected Results 3](#_Toc467172390)

[Contributions of Each Group Member 3](#_Toc467172391)

[Conclusion 3](#_Toc467172392)

[References 3](#_Toc467172393)

Genrify

Group Members: Henri Toussaint, Victor Saint Guilhem, Benoît Lafon

# Task

Given a set of features characterizing a song, we seek to classify it by genre. The task is interesting as it unveils how genres are defined based on attributes that contribute the most to their prediction.

# Dataset

The data set is made of 2000 instances with 13 numerical attributes and a categorical target variable.

Hence, the data set can be resumed into a matrix of 2000 rows and 14 columns.

## Data source

Data set has been collected from the Spotify platform using a Spotify API written in Python. Raw data set was provided in JSON format that we processed to retrieve relevant information to our problem. That is, attributes and target variable described in section ‘Dataset’ as well as the artist and track name in order to briefly interpret the pertinence of some genre classification provided by Spotify.

No further data set from other sources was necessary for our application.

## Target variable

The target variable is categorical comprising 20 distinct genres beforehand handpicked:

‘alternative', 'blues', 'classical', 'country', 'electro', 'folk', 'french', 'hard-rock', 'heavy-metal', 'hip-hop', 'indie', 'jazz', 'pop', 'psych-rock', 'punk-rock', 'r-n-b', 'reggae', 'rock', 'soul', 'techno'

## Features

Input features consist of 13 variables all numerical.

Attributes are chosen such that they describe a given song, those are labelled as:

acousticness, danceability, duration\_ms, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, time\_signature, valence

## Data size

The data set is composed of 2000 instances.

# Preprocessing

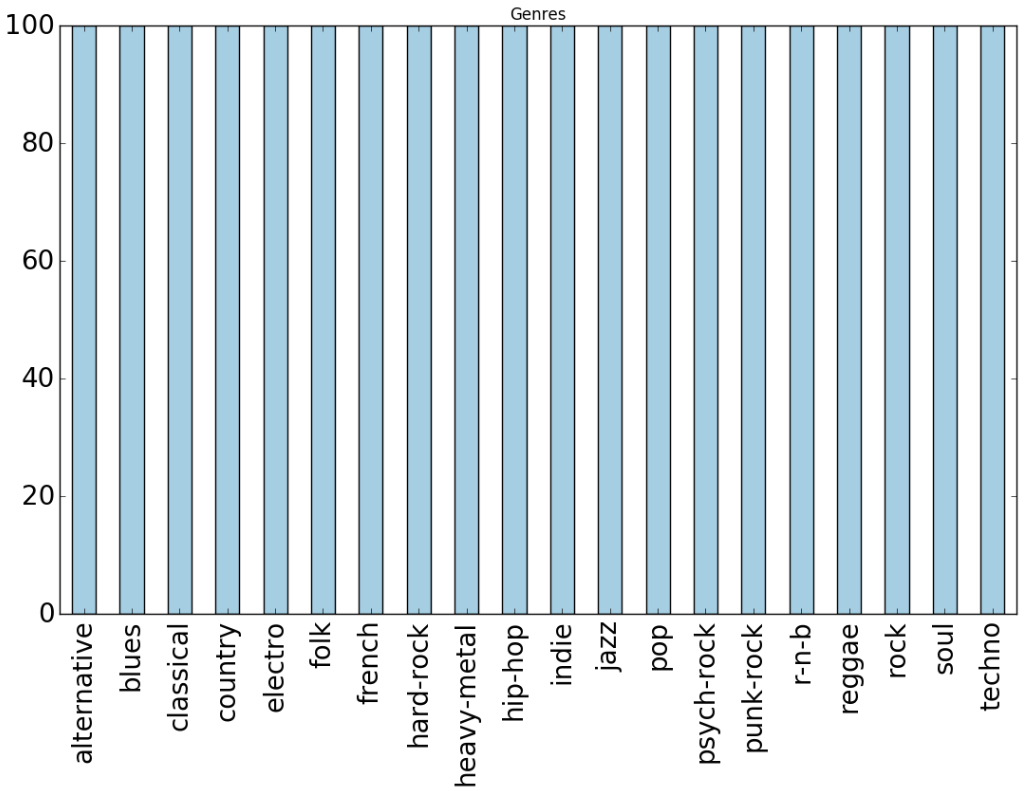
Once we retrieved data from the Spotify API that we stored into dictionaries, we checked for missing or outlier values but none were found as the data was already preprocessed on the Spotify platform.

However, by comparing attribute values between themselves, we observed that variance of ‘duration\_ms’ attribute was far more exceeding any other attribute variance. Since we used Logistic Regression model, we had to standardized our attributes to prevent any variance dominance.

# Visualization

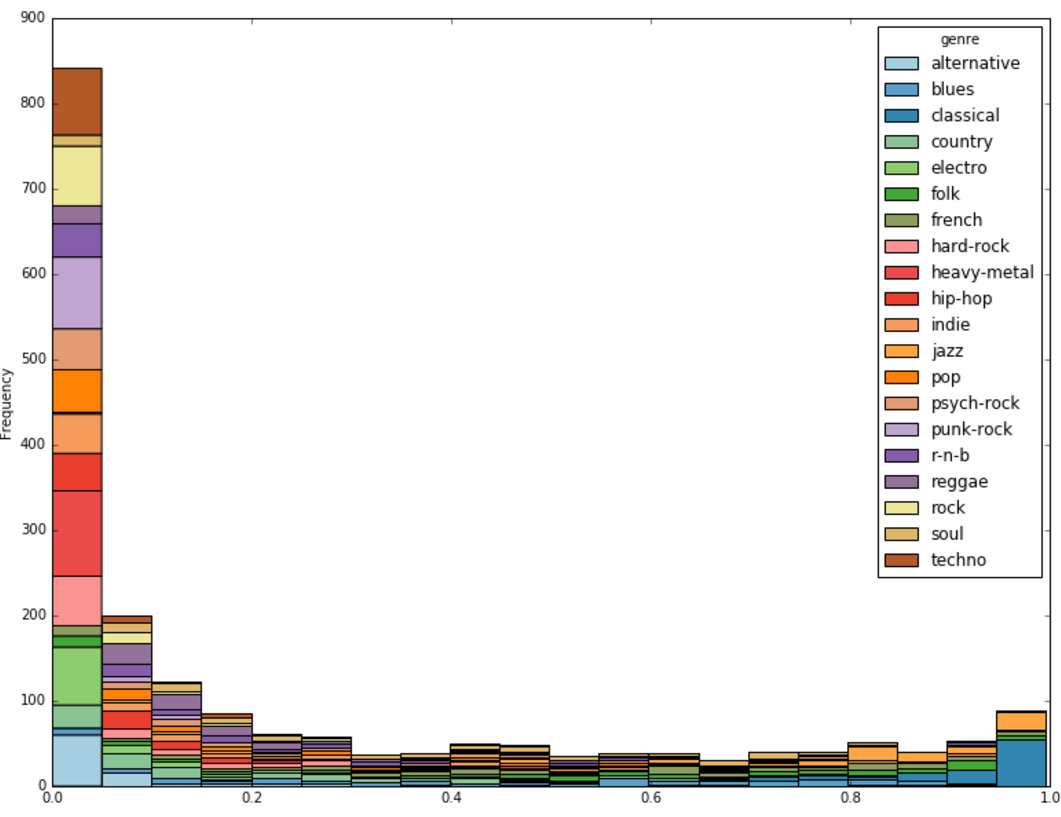
## Target

The target variable consists of 20 categorical genres of which, each of them are represented by 100 instances.



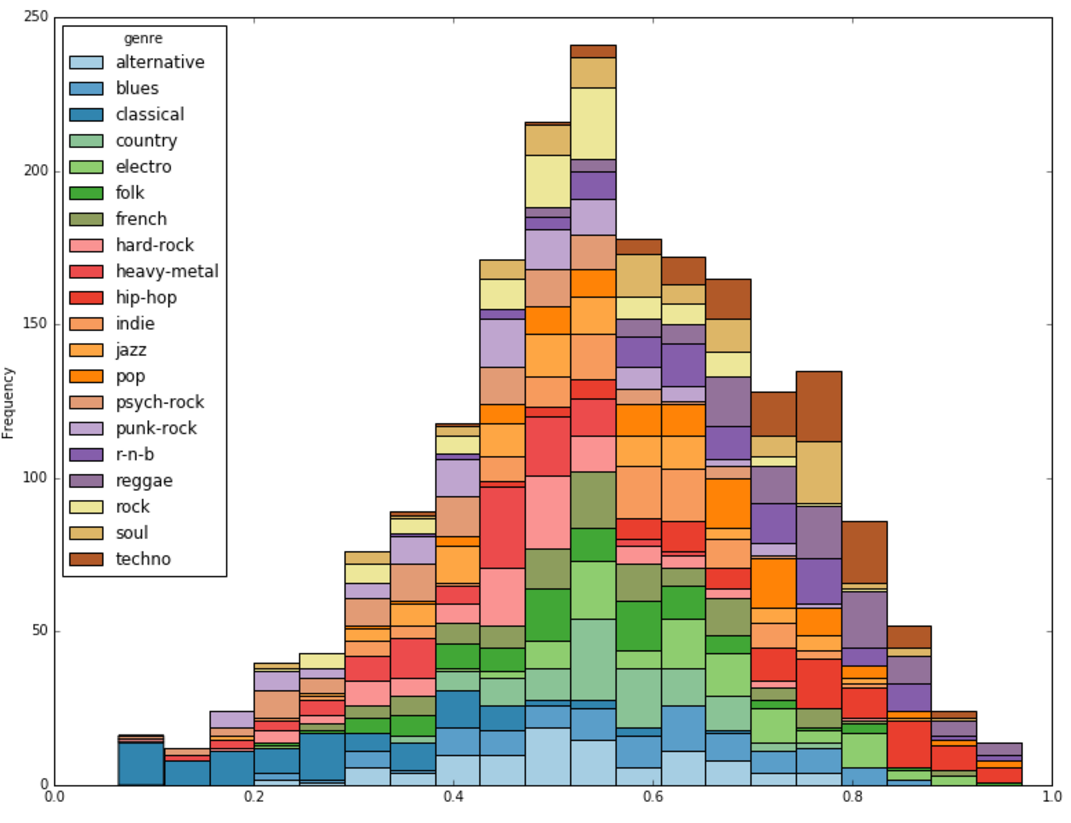
## Features

Acousticness attribute

A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic, i.e. not electronically amplified.



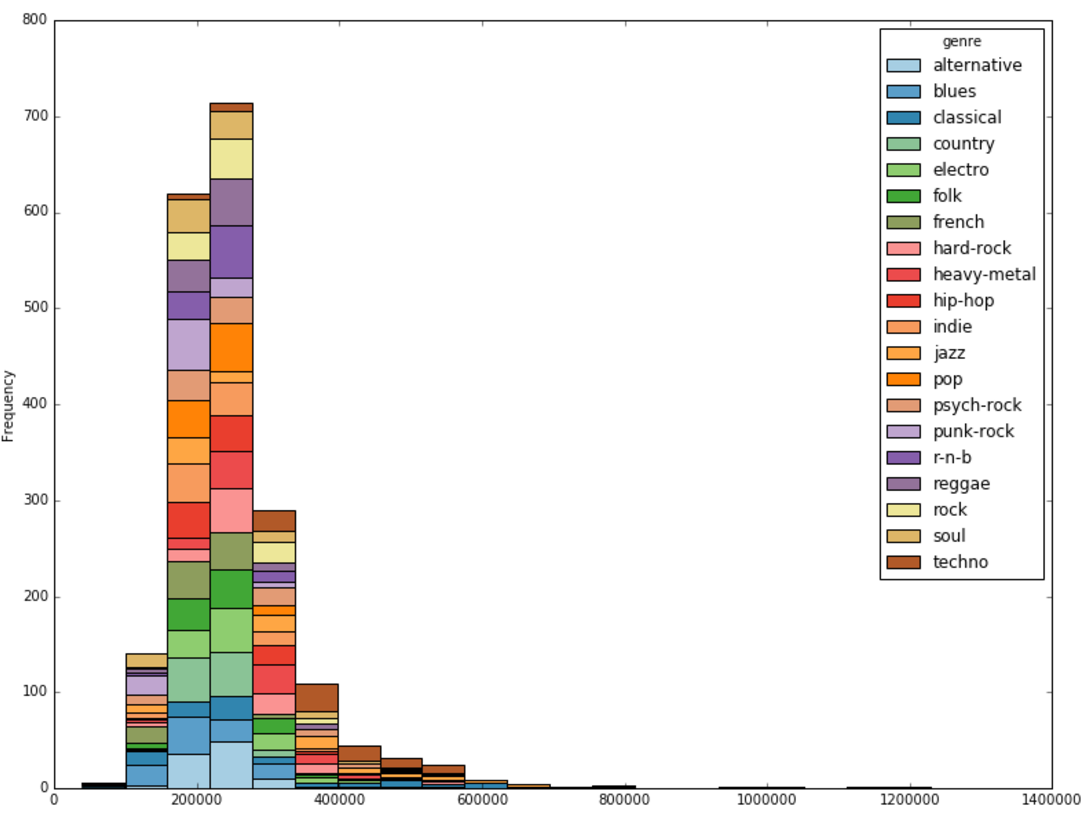
Danceability attribute

Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.



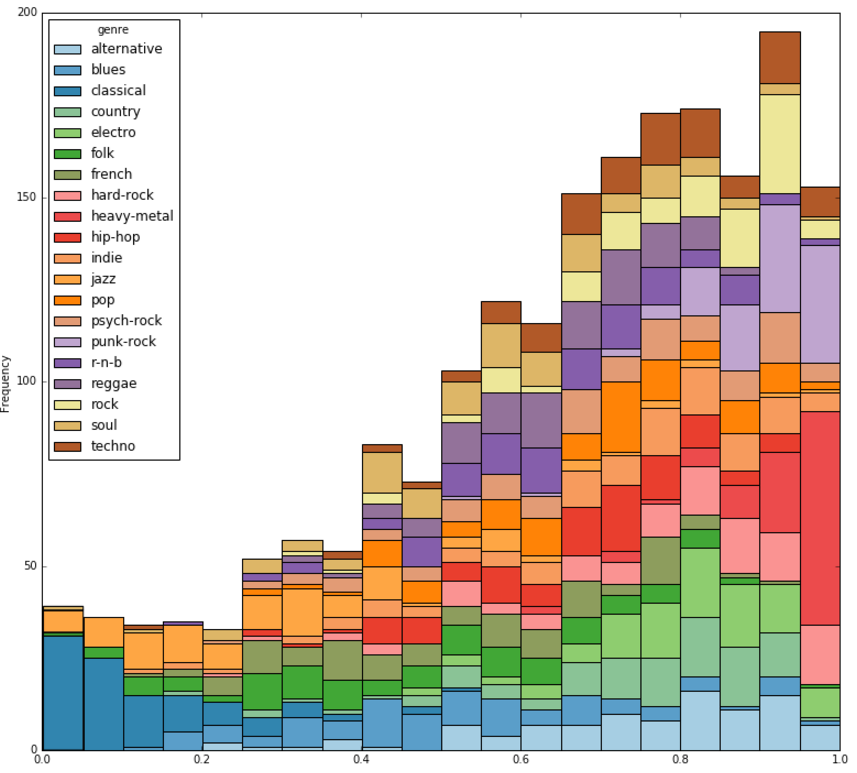
Duration attribute

The duration of the track in milliseconds.



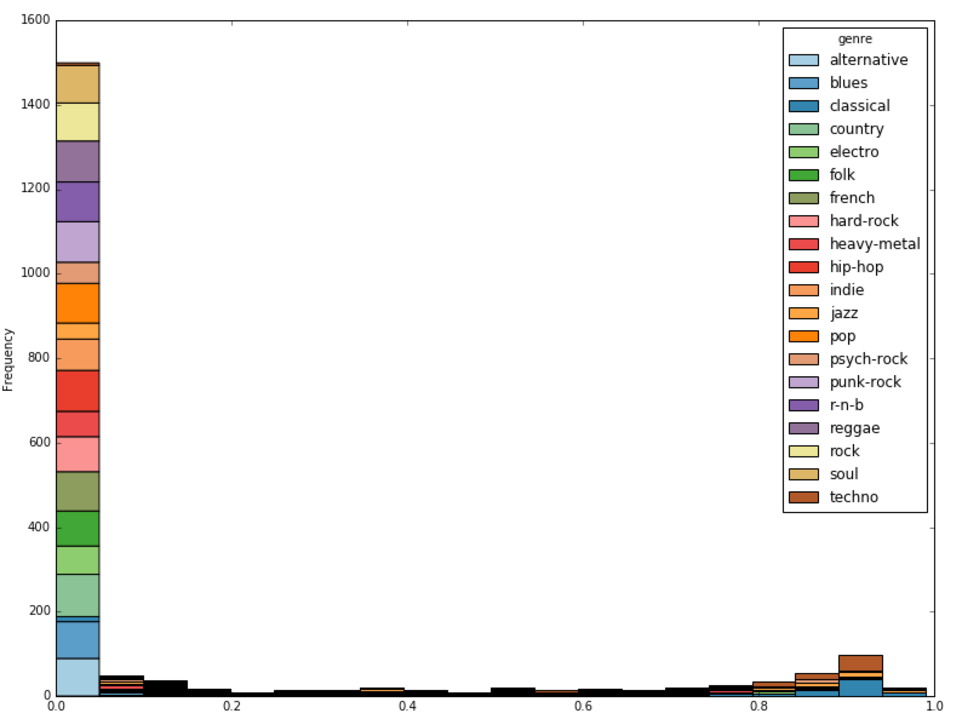


Energy attribute

Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.



Instrumentalness attribute

Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

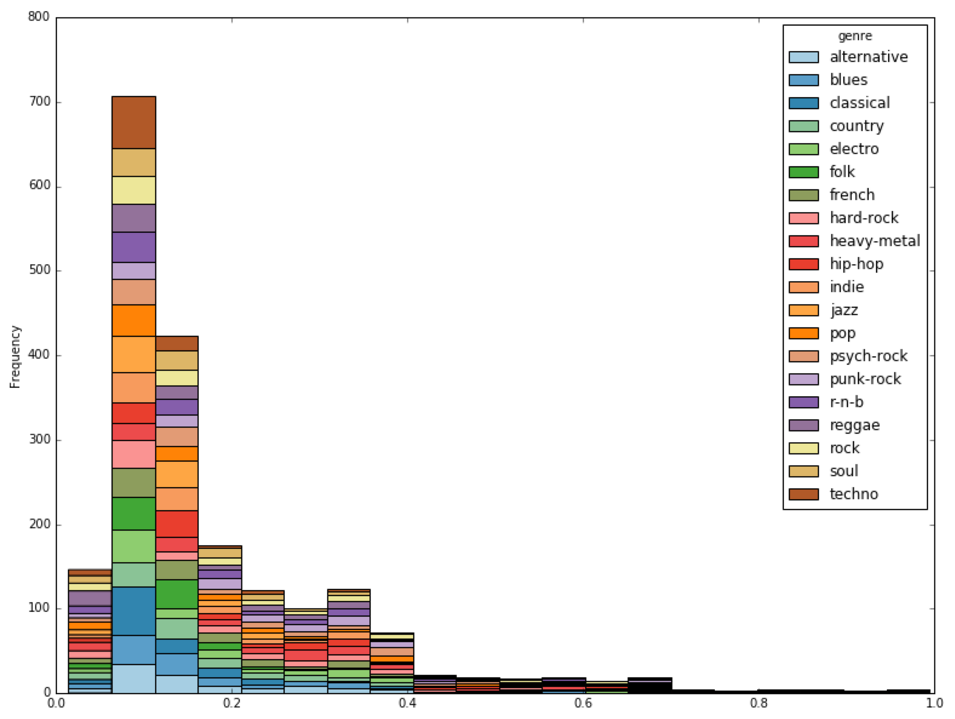


Key attribute

The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on.

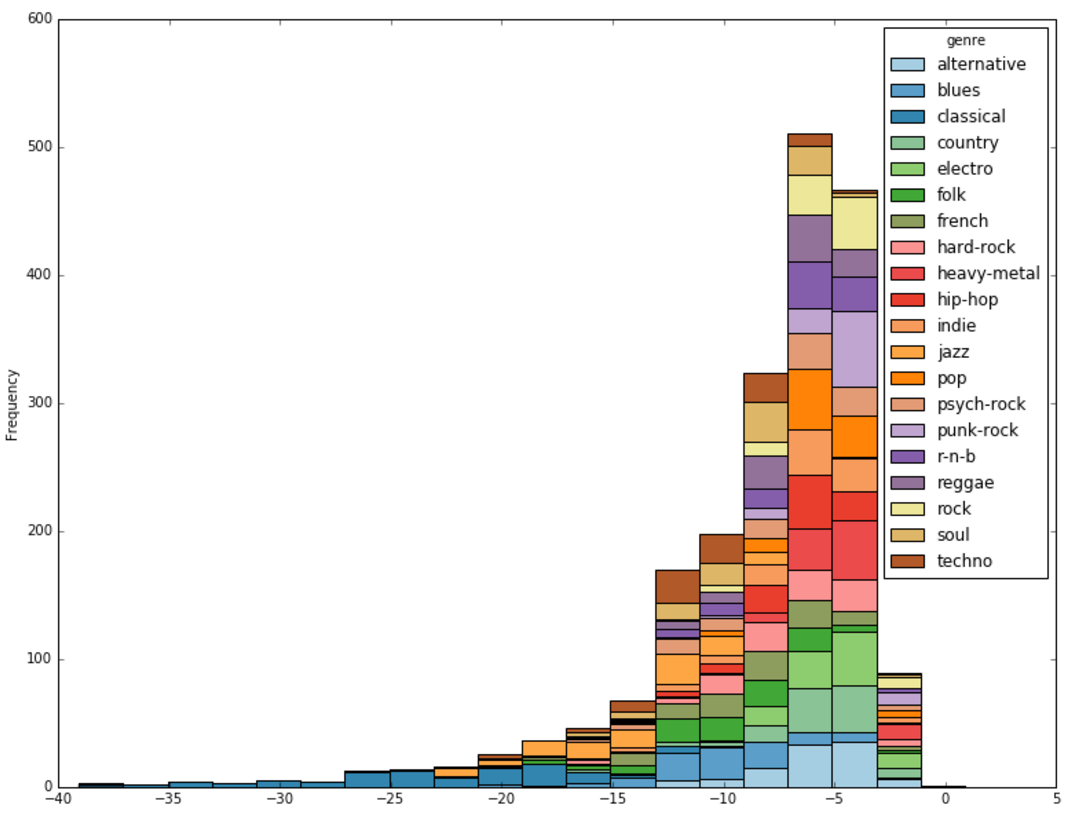


Liveness attribute

Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.



Loudness attribute

The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.

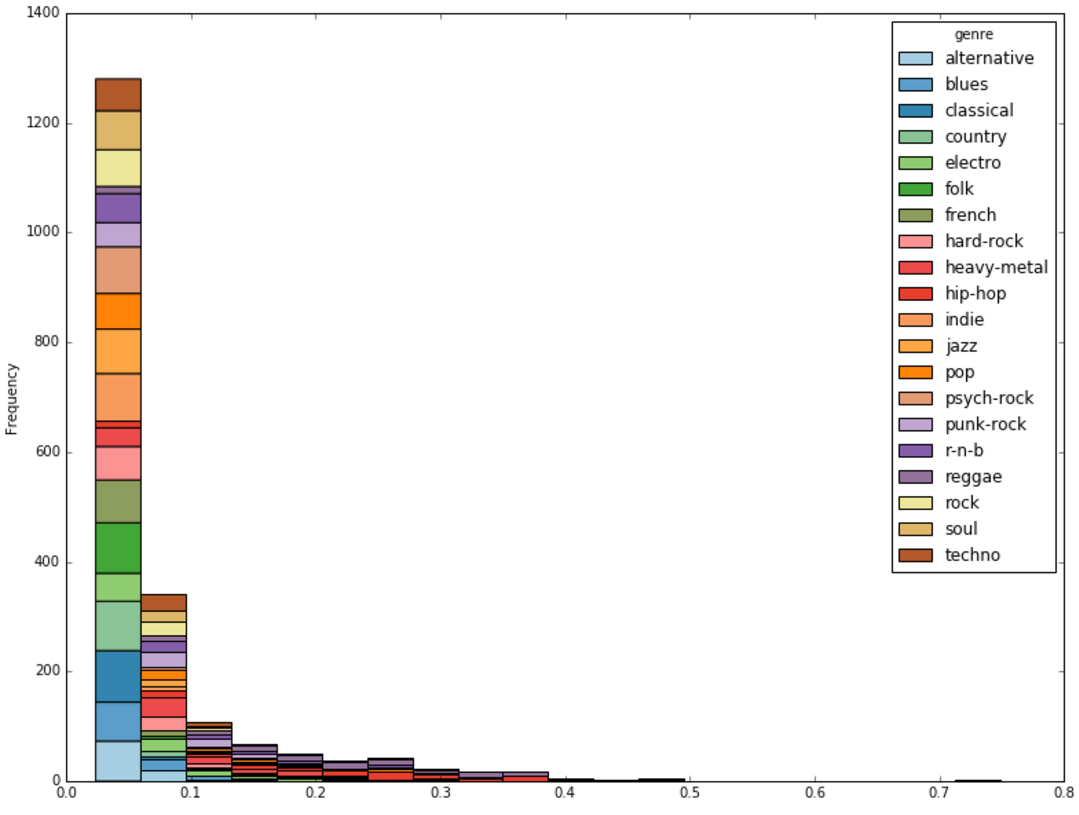


Mode attribute

Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.

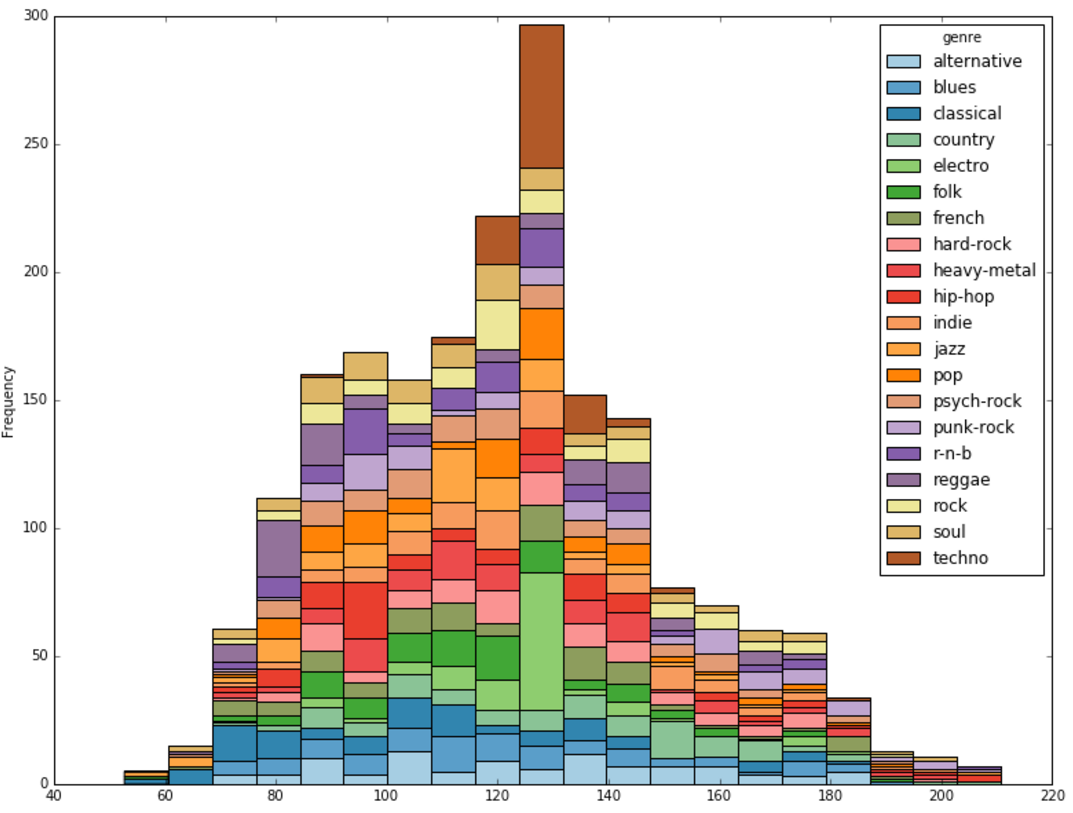


Speechiness attribute

Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.



Tempo attribute

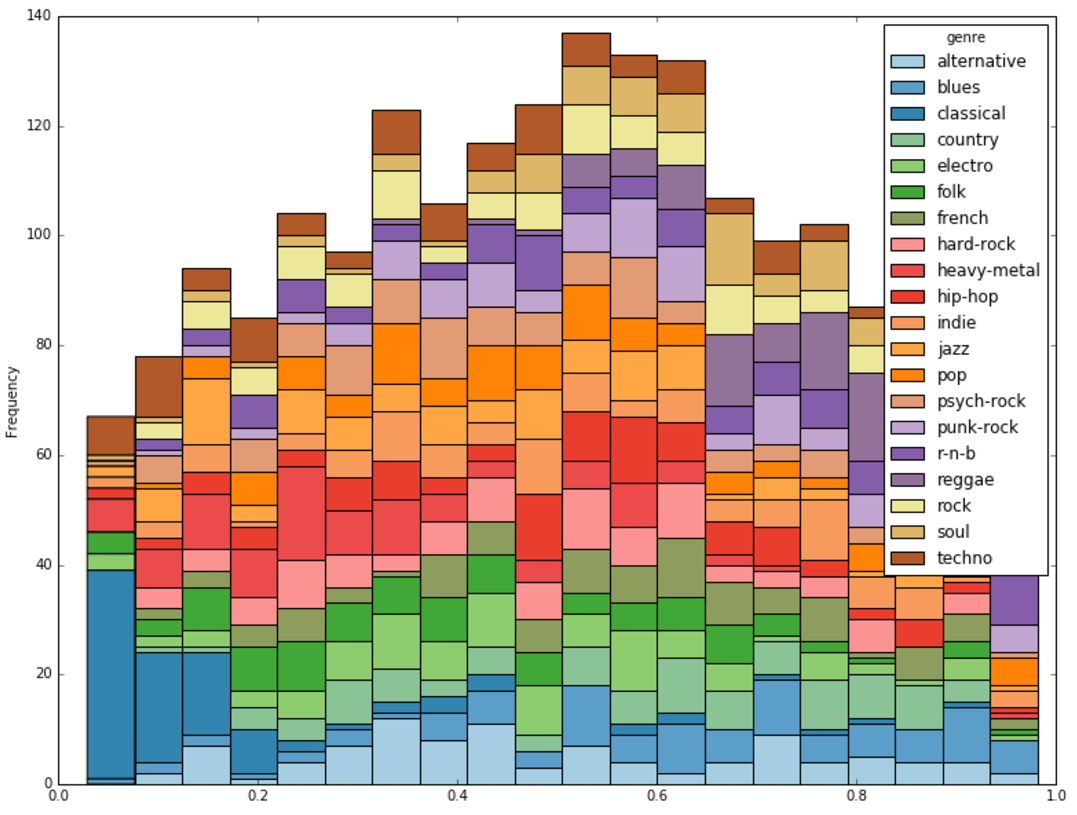
The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.



Time Signature attribute

An estimated overall time signature of a track. The time signature (meter) is a notational convention to specify how many beats are in each bar (or measure).

Valence attribute

A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).



# Evaluation

## Performance Measure

<Which performance measure you chose and why?>

## Classifiers

<Which classifiers and parameter settings did you try and why?>

## Evaluation Strategy

<Did you do train-test split or cross-validation and why?>

## Performance Results

<Report your results, including baselines, using a table similar to the one on slide 7 of the project presentation template file.>

## Top Features

<Present the top features with respect to your model.>

## Discussion

<Briefly discuss your results. Did the best classifier perform as well as you expected? If things did not work out as well, why do you think they did not work? Did one classifier perform much better (or worse) than others? And so on.>

# Interesting/Unexpected Results

<Discuss a few interesting/unexpected cases. See slide 9 of the project presentation template file.>

# Contributions of Each Group Member

<If you are working in a group, please discuss in detail what each member did for this project.>

# Conclusion

<Provide concluding remarks.>

# References

<Provide references if you have any.>

Notes (erase these notes before you save and submit):

1. Once you are done editing, this file, update Table of Contents; you can simply click on Table of Contents and click on the button Update Table once it appears. Then, save this file as a pdf file.
2. Do not copy text from any website/paper. All the wording has to be your own.
3. Delete instructions (the text that appear inside <> above) and these notes before you save and submit.
4. Anywhere it says “discuss”, one or two paragraphs should be sufficient. Do not write anything unnecessary, but also do not omit any necessary details.