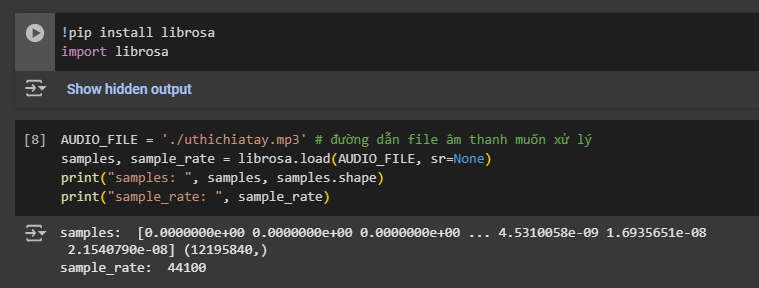
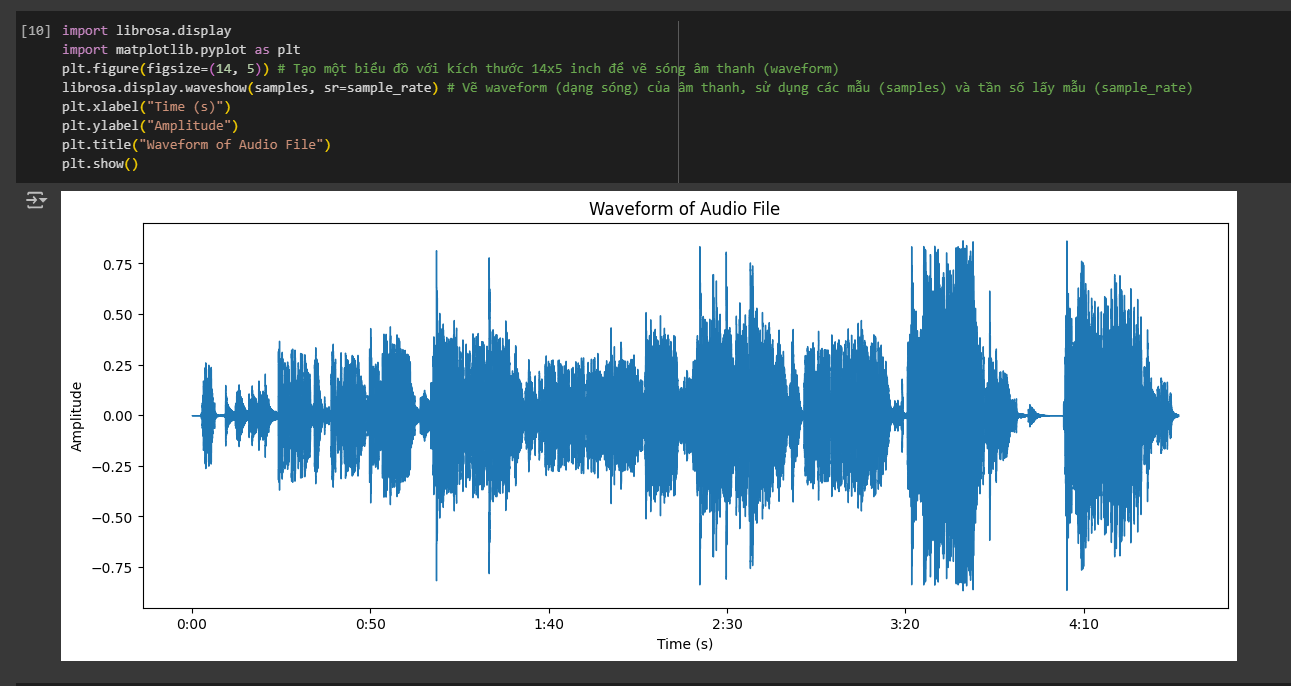
1. **Audio File Formats and Python Libraries**

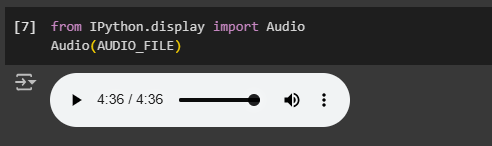
* Âm thanh chúng ta nghe được trong thực tế ở dạng tín hiệu tương tự (analog signal), nó có thẻ số háo và lưu lại theo các định dạng: **.mp3**, **.wav**, **.wma**, **.aac**, **.flac**, ...
* Một số thư viện xử lý dữ liệu Audio:
  + Python: Librosa và Scipy
  + Pytorch: torchaudio
  + Tensorflow: tf.audio
* Đọc file Audio



* Biểu đồ âm thanh



* Có thể nghe được file âm thanh



1. **Audio Signal Data**

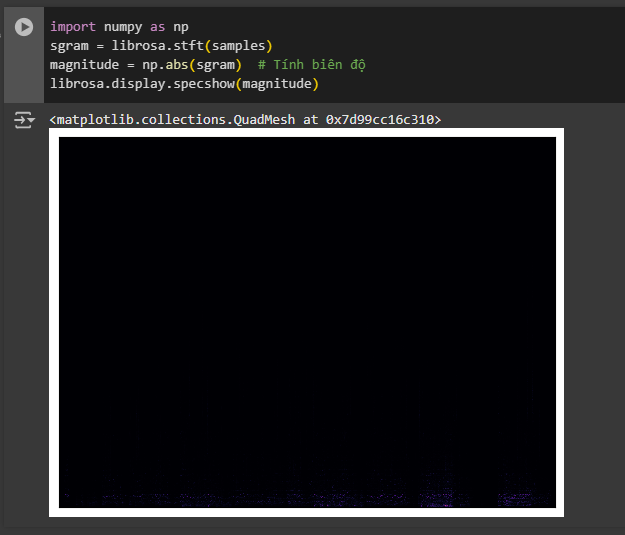
* Audio Data có được bằng cách lấy mẫu Sound Analog Signal theo một chu kỳ thời gian và đi giá trị của iên tại mỗi thời điểm lấy mẫu.
* Audio Data đucợ lưu thành file theo một định dạng nén (.mp3, .wav, ...). Khi đọc lên bằng thư viện xử lý, nó được nén và chuyển thành Numpy Array.
* Audio có thể coi là một chuỗi các giá trị biên độ theo thời gian. Ví dụ như việc đọc file ở trên có sample\_rate = 44100 thì cứ 1s Audio sẽ có 44100 giá trị biên độ.
* Khoảng giá trị của biên độ được quy định bởi thông số bit-length. Ví dụ, bit-length bằng 16 => biên độ có giá trị trong khoảng [0, 2^16 -1].

=> Bit-length càng lớn thì chất lượng của Audio càng tốt

1. **Spectrograms**

Trong Tuần\_1 có nói đến việc chuẩn bị dữ liệu để đưa vào Deep Learning để huấn luyện. Ta phải chuẩn âm thanh sang dạng Spectrogram vì nó thể hiện nhiều đặc trưng của âm thanh.

* Vẽ Spectrogram với âm thanh tại mục 1:

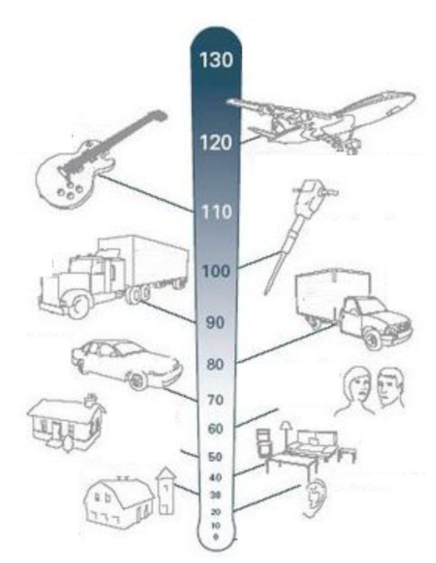


=> Không thể thấy rõ được các thông tin về tần số, biên độ mà Spectrogram thể hiện.

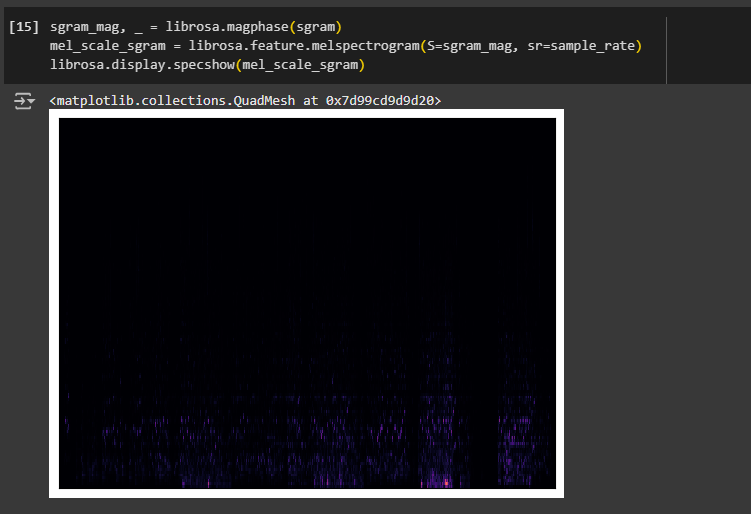
**Lý do:** Hầu hết âm thanh mà chúng ta nghe được tập trung xung quanh một dải tần số và biên độ khá hẹp. Để giải quyết vấn đề này, Spectrogram đucợ chuyển sang một dạng mới, gọi là **Mel Spectrogram** mà tại đó:

* Tần số được thay thế bằng giá trị Logarithmic, goi là Mel Scale.
* Biên độ được thay thế bằng giá trị Logarithmic, gọi là Decibel Scale.

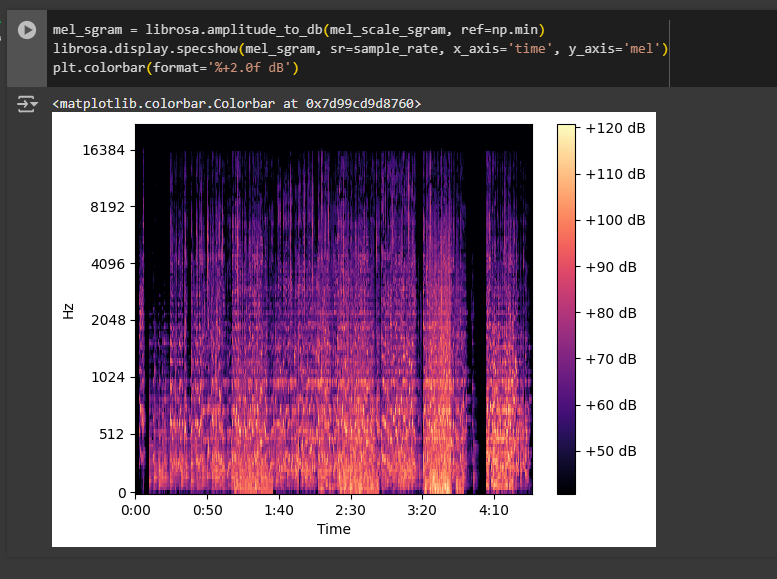
Hình minh họa các laoij âm thay theo thang Decibel Scale:



* Vẫn với file âm thanh trên nhưng đucợ thẻ hiện với Mel Spetrogram:
  + Biểu đồ tần số - thời gian:



* + Biểu đồ biên độ - thời gian:



=> Thông tin âm thanh được thể hiện rõ ràng trên hình ảnh Mel Spectrogram.

**Mã nguồn thực hành:** [Link google colab thực hiện](https://colab.research.google.com/drive/1RueWT0FVKdg32jZS-wObLwEOkNonn6wG?usp=sharing)