This dataset was created from the RAVDESS dataset

(https://es.sonicurlprotectionfra.com/click?PV=2&MSGID=202209141411100289250&URLID=1&ESV=10.0.18.7423&IV=CDD893A9D96AB6D6469DFBAF03B52C7A&TT=1663164671774&ESN=SJ44mpH7HPEgdwadIHqGKE9aHAKB%2FGfCfiESv9Jnnj4%3D&KV=1536961729280&B64\_ENCODED\_URL=aHR0cHM6Ly96ZW5vZG8ub3JnL3JlY29yZC8xMTg4OTc2KSw&HK=3428F8A9C48043894424396B826370722E127A5AEC482B778236DC3B3D0A300B

extracting basic statistics (mean, std, min, max, etc.) from the original audio data and after transforming it using: zero-crossing rate, Mel-Frequency Cepstral Coefficients, spectral centroid, and the stft chromagram.

Features were extracted from the 2452 wav files.

Main Features:

* modality (audio-only)
* vocal\_channel (speech, song)
* emotion (neutral, calm, happy, sad, angry, fearful, disgust, surprised)
* emotional\_intensity (normal, strong). NOTE: There is no strong intensity for the 'neutral' emotion
* statement ("Kids are talking by the door", "Dogs are sitting by the door")
* repetition (1st repetition, 2nd repetition)
* actor (01 to 24)
* sex (M, F)
* channels (number of channels; 1 for mono, 2 for stereo audio)
* sample\_width (number of bytes per sample; 1 means 8-bit, 2 means 16-bit)
* frame\_rate (frequency of samples used (in Hertz))
* frame\_width (Number of bytes for each âframeâ. One frame contains a sample for each channel.)
* length\_ms (audio file length (in milliseconds))
* frame\_count (the number of frames from the sample)
* intensity (loudness in dBFS (dB relative to the maximum possible loudness))

Features Extracted:

Skew: Skewness is a measure of symmetry, or more precisely, the lack of symmetry.

A distribution, or data set, is symmetric if it looks the same to the left and right of the center point.

Kur: Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution.That is, data sets with high kurtosis tend to have heavy tails, or outliers.

Data sets with low kurtosis tend to have light tails, or lack of outliers.

In other words, tails represent how often outliers occur.

* (sum of the zero-crossing rate): zero\_crossings\_sum (sum of the zero-crossing rate)

The zero-crossing rate is the rate at which a signal changes from positive to zero to negative or from negative to zero to positive.

* (statistics of the original audio signal): 'mean', 'std', 'min', 'max', 'kur', 'skew'
* (statistics of the Mel-Frequency Cepstral Coefficients) : mfcc\_ 'mean', 'std', 'min', 'max'

Mel Frequency Cepstral Coefficents (MFCCs or MFCC) are a feature widely used in automatic speech and speaker recognition.

* (statistics of the spectral centroid) : sc\_ 'mean', 'std', 'min', 'max', 'kur', 'skew'

A sound spectrum is **a representation of a sound – usually a short sample of a sound – in terms of the amount of vibration at each individual frequency**. The center of gravity of the spectrum.

* (statistics of the stft chromagram) : stft\_ 'mean', 'std', 'min', 'max', 'kur', 'skew'

 A 12-element representation of the spectral energy where the bins represent the 12 equal-tempered pitch classes of western-type music (semitone spacing).