AUDIO FEATURE DICTIONARY

Feature	Definition and details
tempo_bpm	Measures the overall musical tempo in beats per minute (BPM). Estimated by detecting onset/novelty peaks and their periodicity (for example, peaks of 0.5 s apart → 60/0.5 = 120 BPM). Typical ranges: techno >130 BPM; ballads 60–80 BPM. Units: BPM. (See https://librosa.org/doc/main/generated/librosa.beat.beat_track.html)
loudness_mean_db	Average loudness in decibels. Measures mean signal energy on a logarithmic scale and how loud a sound is perceived; more negative = quieter. Not identical to perceived loudness but closer than linear energy. Similar to rms_mean but in dB scale. Two signals can share RMS yet differ in loudness due to human perception. (See https://librosa.org/doc/main/generated/librosa.amplitude_to_db.html#librosa.amplitude_to_db)
rms	Root Mean Square of signal amplitude per frame. Measures average power/energy over time; a stable indicator of perceived intensity. Peak moment vs RMS: peak is instantaneous maximum; RMS reflects sustained energy. Units: same scale as the underlying magnitude (Mel spectrogram units).
rms_mean	Time average of RMS (single global value). Measures overall average loudness/power of the clip. High → generally loud; low → soft or with many silences. RMS vs. RMS_mean → rms is a temporal series (values per each moment) and rms_mean is a single summary number of all those.
rms_dynamic_range	Max(RMS) − Min(RMS) across time. How much volume varies across time. Measures dynamic contrast in level. Low → flat/constant (heavily compressed electronic music); high → wide dynamics (classical music).

spectral_centroid	Weighted mean frequency of the spectrum (magnitudes as weights). Measures spectral "center of mass" or center of "gravity" of the spectrum → perceived brightness vs darkness. Units: Hz. Examples: cymbals/guitars (high centroid), double bass (low). Typical ranges: low (500-1500 Hz), medium (1500-3000 Hz), high (3000-8000+ Hz). (See https://librosa.org/doc/main/generated/librosa.feature.spectral_centroid.html)
spectral_rolloff	Cutoff frequency below which a fixed proportion (typically 85%) of spectral energy lies. Measures spectral spread/brightness; often correlates with centroid. High rolloff → energy extends into high frequencies; low → energy concentrated in lows. Units: Hz. (See https://librosa.org/doc/main/generated/librosa.feature.spectral_rolloff.html)
spectral_flux	The change in spectral distribution between two consecutive frames. Measures how many new strong events occur between consecutive moments. Units: unitless. Interpretation: low → stable/sustained; high → percussive, frequent changes.
richness	Harmonic/timbral complexity. Measures spectral dispersion/structure at a moment. How much variation there is in energy between different frequency bands at a given moment. High → complex timbres (orchestra, synthesizer); low → simple tones (whistle). Units: none (statistic measure).
var_energy	Temporal variance of the total frame energy. Measures fluctuation of overall energy. High → big ups/downs (silences + hits); low → steady/sustained sounds. Units: amplitude².

novelty	Novelty curve. Measures the degree of local change in audio characteristics; used for onset/segment detection. In this dataset it is computed like spectral_flux, so it is effectively equivalent.
onsets_per_sec	Detected onsets per second (density of new sound events: drum hits, note attacks, syllables). Units: onsets/s. Heuristics: 1–2 = slow/ambient; 4–6 = moderate pop/rock/jazz; >8 = very percussive/dense. Difference vs BPM: BPM is underlying metrical pulse; onsets/s is the actual count of new events and can vary strongly over time. (See https://librosa.org/doc/main/generated/librosa.onset.onset_detect.html#librosa.onset.onset_detect)
attack_decay_slope	Sound attack and decay slope. The way in which the amplitude of a sound evolves over time. Measures how quickly energy rises/falls after events (attack/decay sharpness). High → punchy/percussive, abrupt changes; low → smooth/sustained pads/strings. Heuristics: 0.001–0.01 low, 0.01–0.05 medium, >0.05 high.
temporal_centroid_x	Energy weighted time centroid from the Mel spectrogram. Measures when timbre/texture energy concentrates within the clip, when most of the energy is concentrated. It measures whether the song "starts strong and ends softly" or vice versa. Units: seconds Heuristics: (if it is normalized): from 0 to 1 (0 = early, 0.5 = middle, 1 = late). Difference with spectral centroid: spectral centroid focuses in frequencies (Hz) and temporal centroid is focused in time (seconds).
temporal_centroid_y	Energy weighted time centroid from RMS/waveform energy. Measures when overall intensity concentrates. Used when you care about global intensity timing. Units: seconds or frames.
key_estimate	Estimated pitch class (0–11 for C, C#, …, B) via average chroma profile.

	Measures the most prominent note class across the track. Limitation: does not disambiguate major vs minor. (See https://librosa.org/doc/main/generated/librosa.feature.chroma_stft.html)
mfcc_0…mfcc_19	Mel Frequency Cepstral Coefficients (20 values). Compact description of spectral shape/timbre on a perceptual (Mel) scale. "Fingerprint" of the sound tone. These are obtained from how the human ear perceives sound. Heuristics: mfcc_0 = overall energy; mfcc_1–3 = generic characteristics of timbre; mfcc_4–13 = mid detail (brightness/texture); mfcc_14+ = capture finer/less stable details. Units: none. (See https://librosa.org/doc/main/generated/librosa.feature.mfcc.html).