**Basic Metadata**

**Album Name:** The album/LP this song is officially published on.

**Track Number:** Which number in chronological order the song appears in on the album.

**Track Name:** The official name for the track as listed on the album according to Genius.

**Producers:** The person or people credited with producing the song. This information is also taken from Genius.

**Writers:** All individuals given writing credits on the song according to Genius. These writers are not exclusively credited for the lyrical writing of the songs but rather the entirety of the track (instrumental, lyrics, solos, musical samples, etc), so the writers list is often ‘bloated’.

All producers and writers are listed in alphabetical order rather than the order Genius lists them for consistency.

**Vocals**

**Verse Vocals:** Who is credited with vocals during the designated verses in the song’s structure.

**Chorus & Post-Chorus Vocals:** Who is credited with vocals during the designated Choruses throughout the song’s structure. Post Choruses are seen structurally in music as being a direct extension of the chorus, so post-chorus vocals are also included in this category.

**Pre-Chorus & Bridge Vocals:** Both Pre-Chorus and Bridge serve similar functions in music, so this category lumps these two designated parts of the song structure together.

**Misc Vocals/Samples:** This category is for all vocals that do not fall into a predetermined structural part. This basically serves as a ‘catch-all’ for tracks that do not have normal song structures (usually intermissions or shorter songs leading into another). Intros and outros that have lyrical content or sounds similar or identical to the verse, chorus, pre-chorus, or bridge of the song do not count for this category and are instead lumped into their other respective category. Interludes, intros, and outros that are distinctively different from the lyrical content in the rest of the song count towards this category.

Some sections are labeled differently from this according to Genius (such as ‘Breakdown’). These sections are placed into the categories they are most similar to in song structure (for example, breakdowns are usually similar in content to either the Bridge, Pre-chorus, or Post-chorus). Personal judgment and analysis on my part is used in all outlet cases as to which vocal category they belong in.

All of these vocal categories are labeled in accordance to Genius and their crediting of vocalists on their lyrics for each track. If vocals are not labeled within the lyric page itself, the vocalists will be identified by a combination of my listening to the track and looking at specific vocal credits on the track outside of Genius’s lyrics.

N/A is the default for if a category is not represented in a song. Instrumental tracks, i.e tracks with N/A in all 4 vocal fields, will not be included in the text mining corpus for lyrics but will still have this metadata collected.

**Ratings**

**AOTY Rating:** The user-curated rating for each individual song according to the website albumoftheyear.org. This number is likely to change throughout data collection and will need to be updated with a specific collection date listed for consistency.

**Spotify Listens:** The amount of listens this track has on Spotify according to Spotify’s own platform, rounded to the nearest thousand (Ex. 534,657,890 rounded to 534,658,000). Any drastic shifts in these numbers from collection date will cause me to update all the numbers accordingly and list a collection date for those numbers, but the rounding should help alleviate the need for this.

**All of the AOTY Ratings and Spotify Listens have been collected and updated to be accurate as of February 26, 2024.** The dataset will keep these numbers from this date until otherwise updated (which will also be updated on this document).

---

All of the statistics and metadata for this project is curated for the original run for the album. Any re-released copies such as anniversary editions or deluxe editions likely have slightly different statistics which are not covered in this data set due to this corpus only including the main discography from the band. Within this structure, all ‘bonus tracks’ and singles not connected directly to an album’s track listing are also excluded from this corpus.