MUSIFEAST-17 Normed Data Description

These datasets comprise two main sections: 1) information collected about the 356 music stimuli included in MUSIFEAST-17 (e.g., clip name in study, genre, familiarity, full track title, artist, release date), and 2) clip responses across all trials from the relevant participant groups, collected from a questionnaire presented to participants after each of their assigned 17 music clips in which they provided responses regarding their experience of the music (e.g., thoughts, familiarity, liking, genre perception, contextual associations).

There are nine MUSIFEAST-17 reference tables: one with all participant responses aggregated into one data frame, and eight for each experiment group referring to participants nationality and age range (UK18to30, UK31to45, UK46to60, UK61to75, US18to30, US31to45, US46to60, US61to75). This enables easy access to aggregated clip rating within and across broad demographic groups.

Column header descriptions

CLIP_NAME:
The name of the music stimulus clip as referred to in study analysis

GENRE:
Preassigned music clip genre as specified by **last.fm** / **IMDb.com** during the stimuli collection process

FAMILIARITY:
Music clip preassigned familiarity label from the stimuli collection process

TRACK_TITLE:
The name of the music piece as collected from Spotify

ARTIST:
The name of the music artist or composer as collected from Spotify

MEDIA_RELATION_or_PERFORMER:
Specific media origin listed if the piece is from a game, movie, or TV show, and Performer listed with relevance to the Classical music genre to indicate the performer or ensemble as collected from Spotify

RELEASE_DATE:
The year the original track was released/composed (approximation for some Classical clips)

VOCALISATION:
Yes/No to indicate the presence of non-lyrical vocalisation in the clip

VOCALISATION_TYPE:
The specific type of non-lyrical vocalisation present in the clip

CLIP_TIME_START:
The start time of the clip from the original track

CLIP_TIME_END:
The end time (30 seconds from the start time of the clip) of the original track

group:
Stimulus group that each clip was assigned to for the experiment; there are 21 stimulus groups in total, with 17 clips per group

style_familiarity_M:
Mean rated familiarity with the style of the music clip, Likert scale 1-5 (1: very unfamiliar, 5: very familiar)

style_familiarity_SD:
Standard deviation of rated familiarity with the style of the music clip

clip_familiarity_M:
Mean rated familiarity with the specific music clip, Likert scale 1-5 (1: very unfamiliar, 5: very familiar)

clip_familiarity_SD:
Standard deviation of rated familiarity with the specific music clip

contrast_M:
Mean rated musical contrast throughout the 30-second clip, Likert scale 1-5 (1: stayed the same, 5: changed a lot)

contrast_SD:
Standard deviation of rated musical contrast throughout the 30-second clip

enjoyment_M:
Mean rated enjoyment of the music clip, Likert scale 1-5 (1: not at all, 5: very much so)

enjoyment_SD:
Standard deviation of rated enjoyment of the music clip

valence_M:
Mean rated valence expressed by the music, Likert scale 1-5 (1: very negative, 5: very positive)

valence_SD:
Standard deviation of rated valence expressed by the music

arousal_M:
Mean rated energy level of the music, Likert scale 1-5 (1: very low energy, 5: very high energy)

arousal_SD:
Standard deviation of rated energy level of the music

cluster:
One of the three identified clusters from the cluster analysis that the clip belongs into

rated_genre:
Most frequently rated genre of the music clip, aggregated over both primary and secondary genre selections from all participants (note: secondary genre selection was optional) (15 options, Ambient – Video game music), multiple listed if there were ties

rated_genre_perc:
Percentage over all trials the clip received the most frequently rated genre label aggregated over both primary and secondary genre selections from all participants. Indicates between-participant agreement.

genre_match:
(*Does not include Sixties and Eighties genres, NA*)
Percentage over all trials the clip's perceived genre matched the preassigned genre label from stimuli collection in either primary or secondary genre selections from all participants

rated_decade:
Most frequently rated decade in which the music clip was created, (Earlier than 1960, 1960s, 1970s, 1980s, 1990s, 2000s, 2010s, 2020s), multiple listed if there were ties

rated_decade_perc:
Percentage over all trials the clip received the most frequently rated decade of creation. Indicates between-participant agreement.

decade_match:
Percentage over all trials the clip's perceived decade of creation matched the clip's original release date

rated_location:
Most frequently rated continent of origin, (Africa, Asia, Australia Europe, Latin/South America, USA/Canada), multiple listed if there were ties

rated_location_perc:
Percentage over all trials the clip received the most frequently rated continent of origin

thought_type:
Most frequently selected thought type over all trials, multiple listed if there were ties

thought_type:
Percentage over all trials the clip received the most frequently selected thought type

freq_context_words:
Lemmatised words that occurred more than three times across all trials in the open-text description of participants' contextual associations with the music clip