The Music Streaming Sessions Dataset - Schema

The schema for the session logs is given below. Each row corresponds to the playback of one track, and has the following fields, with corresponding values

Field	**Values**
session_id	E.g. 65_283174c5-551c-4c1b-954b-cb60ffcc2aec - unique identifier for the session that this row is a part of
session_position	{1-20} - position of row within session
session_length	{10-20} - number of rows in session
track_id_clean	E.g. t_13d34e4b-dc9b-4535-963d-419afa8332ec - unique identifier for the track played. This is linked with track_id in the track features and metadata table.
skip_1	Boolean indicating if the track was only played very briefly
skip_2	Boolean indicating if the track was only played briefly
skip_3	Boolean indicating if most of the track was played
not_skipped	Boolean indicating that the track was played in its entirety
context_switch	Boolean indicating if the user changed context between the previous row and the current row. This could for example occur if the user switched from one playlist to another.
no_pause_before_play	Boolean indicating if there was no pause between playback of the previous track and this track

short_pause_before_play	Boolean indicating if there was a short pause between playback of the previous track and this track
long_pause_before_play	Boolean indicating if there was a long pause between playback of the previous track and this track
hist_user_behavior_n_seekfwd	Number of times the user did a seek forward within track
hist_user_behavior_n_seekback	Number of times the user did a seek back within track
hist_user_behavior_is_shuffle	Boolean indicating if the user encountered this track while shuffle mode was activated
hour_of_day	{0-23} - The hour of day
date	E.g. 2018-09-18 - The date
premium	Boolean indicating if the user was on premium or not. This has potential implications for skipping behavior.
context_type	E.g. editorial playlist - what type of context the playback occurred within
hist_user_behavior_reason_start	E.g. fwdbtn - the user action which led to the current track being played
hist_user_behavior_reason_end	E.g. trackdone - the user action which led to the current track playback ending

The schema for the track metadata and features is given below, each row has the following fields, with corresponding values

Field	**Values**
track_id	E.g. t_13d34e4b-dc9b-4535-963d-419afa8332ec - unique identifier for the track played. This is linked with track_id_clean in the session logs
duration	Length of track in seconds
release_year	Estimate of year the track was released
us_popularity_estimate	Estimate of the US popularity percentile of

	the track as of October 12, 2018
acousticness	See https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/
beat_strength	See acousticness
bounciness	See acousticness
danceability	See acousticness
dyn_range_mean	See acousticness
energy	See acousticness
flatness	See acousticness
instrumentalness	See acousticness
key	See acousticness
liveness	See acousticness
loudness	See acousticness
mechanism	See acousticness
mode	See acousticness
organism	See acousticness
speechiness	See acousticness
tempo	See acousticness
time_signature	See acousticness
valence	See acousticness
acoustic_vector_0	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_1	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_2	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_3	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_4	See http://benanne.github.io/2014/08/05/spotify-cnns.html and

	http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_5	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_6	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-
acoustic_vector_7	See http://benanne.github.io/2014/08/05/spotify-cnns.html and http://papers.nips.cc/paper/5004-deep-content-based-