Elgamiel, Laila

PSTAT131 FINAL PROJECT
Codebook

The columns of my final dataset are described as follows:

instance_id [dropped]: a numerical index for track identification
artist_name [dropped]: character variable of name of artist of specified track
track_name [dropped]: character variable of name of specified track
popularity: numerical variable from 0 to 100 that describes how popular track is on Spotify (closer to 100= more popular)
acousticness: numerical variable from 0 to 1 that describes how acoustic track is (closer to 1=more acoustic)
danceability: numerical variable from 0 to 1 that describes how suitable a track is for dancing (closer to 1= more danceable)
duration_ms: duration of track in milliseconds
energy: numerical variable from 0 to 1 that describes how energetic a track is (closer to 1= more energetic)
instrumentalness [dropped]: numerical variable from 0 to 1 that describes levels of vocals on track (closer to 1= more vocals)
key [One Hot Encoded]: character variable describing which system of related musical chords track falls under
liveness: numerical variable from 0 to 1 describing probability that track was recorded with live audience present
loudness: numerical variable indicating how loud a song is
mode [One Hot Encoded]: character variable describing type of musical scale coupled with a set of characteristic melodic and harmonic behaviors track belongs to
speechiness: numerical variable from 0 to 1 describing presence of speech on track (closer to 1= more speech)
tempo [converted to numeric]: variable describing perceived speed of song
obtained_date [dropped]: date that information about track was obtained from spotify
valence: numerical variable from 0 to 1 describing the musical positiveness conveyed by a track (closer to 1= more positive)
Aname_length: length of string variable artist name
Tname_length: length of string variable track name
music_genre: genre which specified track was labeled under on Spotify API