* Valence : Is float type of data, a subjective measure from 0.0 to 1.0 describing the musical emotion in two main categories, i.e, high valence means that the song is happy, vibrant etc, and lower values mean that the song is sad, depressing, etc.
  + count 15000.000000 mean 0.436853 std 0.277220 min 0.000000 25% 0.196000 50% 0.416000 75% 0.664000 max 0.995000
* Tempo : Is a float type of data that estimates the BPM tempo of the song, it derives from the average duration so it’s not consistent for a track with variations in tempo, also musically the measurements tend to be discrete while this variable is continuous.
  + count 15000.000000 mean 123.116544 std 31.930958 min 0.000000 25% 99.939000 50% 124.188000 75% 141.986500 max 220.525000
* Features\_duration\_ms : Is a float type of data, it defines the duration of each song in milliseconds, this variable can be easily incorporated in the study since it does not have a subjective component.
  + count 1.500000e+04 mean 2.467947e+05 std 1.279850e+05 min 8.587000e+03 25% 1.800000e+05 50% 2.278185e+05 75% 2.889030e+05 max 4.120258e+06
* Time\_signature : Is an integer type of data, it represents the time signature in a reduced way, it replaces from “3/4" to “7/4” with values from 3 to 7. This metric doesn’t allow the full scope representation of the time signatures, and can be used just as a proxy of time structure in songs, it also has the problem of not being able to catch progressive genres or ideas, in an accurate way. It’s also important to mention that the mean and std don’t serve a good purpose here since we are working with integers that are categorical.
  + count 12938.000000 mean 3.876179 std 0.566491 min 0.000000 25% 4.000000 50% 4.000000 75% 4.000000 max 5.000000
* N\_beats : Is an integer type of data that shows the number of beats the song has. This value is subjective to the time signature of the song and can not be analyzed as the true length of the song, it should be analyzed in conjunction with the time\_signature var. although in this case these are also integer values, since it’s an ordinal var, we can use the mean and the std.
  + count 15000.000000 mean 501.862333 std 280.689804 min 0.000000 25% 327.000000 50% 461.000000 75% 625.000000 max 7348.000000
* N\_bars : Is an integer type of data similar to n\_beats, the difference is that this one contains the number of bars, in this case we will see a similar type of data but with a scale that is smaller since songs have less bars than beats, in the same way as before, this var should be analyzed taking in to consideration the time\_signature var, number of bars depends directly with the time signature of the song.
  + count 15000.00000 mean 128.39340 std 75.11391 min 0.00000 25% 83.00000 50% 117.00000 75% 159.00000 max 2170.00000
* Popularity\_confidence : is a float type of data and explains the level of confidence of another variable called popularity, both variables have a scale from 0-1 or 0-100, this variable gives the level of veracity each obs or pattern has on the popularity var. It’s important to mention that it has a huge number of missing values.
  + count 2217.000000 mean 0.490479 std 0.291330 min 0.000000 25% 0.230000 50% 0.480000 75% 0.739000 max 1.000000
* Genre : is an array of strings type of data, it describes the genre of the song in a string, is not measurable by numbers but can be useful to show characteristics of the songs, association, categorization.
  + count 15000 unique 20 top j-dance freq 750

? processing,