This data set contains songs from Spotify. A start-up that plans to develop a product that generates music is interested in using the dataset to understand consumer preferences (e.g. www. HookGen.com).

Use clustering to profile the songs.

1. How many clusters are there in the data?
2. Describe the characteristics of the clusters you identified.
3. What product features should be added to the product when it has to generate songs.

**Attribute Information:**

|  |  |  |
| --- | --- | --- |
|  | acousticness | (Ranges from 0 to 1). A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic. |
|  | artists | Track Artists |
|  | danceability | (Ranges from 0 to 1). Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable. |
|  | duration\_ms | The duration of the track in milliseconds. (Integer typically ranging from 200k to 300k) |
|  | energy | Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. |
|  | song\_id | The Spotify ID for the track. |
|  | instrumentalness | Predicts whether a track contains no vocals. Ranges from 0.0 to 1.0. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. |
|  | loudness | The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Values typical range between -60 and 0 db. |
|  | name | Track name |
|  | popularity | The popularity of the song lately, default country = US. Higher score indicates higher popularity |
|  | tempo | The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration. |
|  | valence | A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry). (Ranges from 0 to 1) |