

ASSIGNMENT 1 - CONTENT-BASED RECOMMENDATIONS

TASK 1

I chose my own messy playlist that contains nearly every track that I have encountered during the last five years of my life. This works well for this task, because the genres are very varied going through rock, pop, heavy metal, techno, house and some country.

The only feature that I couldn't quite place initially was 'key'. From the documentation I see it is the key the track is in, would have absolutely no idea just by looking at the tracks, but maybe someone with more audio knowledge could. I also found it slightly confusing what the exact difference is between tempo and energy for example, but based on the documentation there is also some overlaps, for example is danceability a measure that also considers tempo which is a feature in itself.

Sorting 20 tracks descending based on danceability I get the following scores:

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	duration_ms	time_signature
Ravi Kumar - Breakfast 1pm	0.976	0.573	5	-6.181	0	0.0802	0.320000	0.000175	0.0925	0.757	123.956	161652	4
Princess Nokia - Brujas	0.940	0.718	2	-6.459	1	0.2380	0.068900	0.000133	0.0914	0.676	140.033	178917	4
Ravi Kumar - Office Man	0.926	0.637	0	-5.857	1	0.0756	0.011700	0.000009	0.1280	0.862	116.596	140395	4
Koffee - Toast	0.906	0.638	10	-7.917	0	0.2560	0.034200	0.000017	0.0725	0.769	99.959	151532	4
Naira Marley - Soapy	0.900	0.606	1	-4.816	1	0.2460	0.159300	0.000710	0.0611	0.671	123.036	174043	4
Tierra Whack - Hungry Hippo	0.893	0.510	4	-5.591	0	0.0488	0.022700	0.007440	0.1020	0.778	136.640	60000	4
Father - Heartbreak	0.892	0.658	11	-7.276	0	0.2270	0.024500	0.000004	0.1330	0.636	141.923	138901	4
Nujabes - Luv (sic.) pt3 (feat. Shing02)	0.887	0.619	10	-6.160	0	0.0587	0.419000	0.011100	0.0860	0.963	94.985	374200	4
RMR - DEALER (feat. Future & Lil Baby)	0.885	0.925	1	-5.067	1	0.0750	0.100900	0.000124	0.1320	0.058	119.978	192794	4
Meatz Of Operation - Janet	0.882	0.657	11	-5.159	0	0.1140	0.175000	0.000013	0.3280	0.278	124.998	252283	4
Koffee - W (feat. Gunna)	0.866	0.721	5	-6.813	0	0.0779	0.178900	0.000021	0.0913	0.902	100.980	230151	4
Leikeli47 - Get Blunt	0.836	0.905	1	-4.518	1	0.1890	0.012800	0.000080	0.0586	0.367	93.953	255480	4
AUDREY NUNA - damn Right	0.854	0.671	10	-7.382	0	0.3360	0.116000	0.000000	0.1170	0.662	162.016	164603	4
StarBoy - Socu (feat. Wizkid, Ceeza Mill, Spotless & Tem)	0.841	0.544	9	-3.284	0	0.0830	0.054900	0.004020	0.1420	0.881	108.000	255508	4
21 Savage - a lot	0.837	0.636	1	-7.643	1	0.0860	0.039500	0.001250	0.3420	0.274	145.973	288624	4
Ashnikko - Daisy	0.834	0.603	8	-6.143	1	0.1330	0.230000	0.000000	0.3720	0.774	121.676	146023	4
Luni - Broken	0.830	0.083	0	-2.062	1	0.0412	0.006300	0.000000	0.1090	0.148	134.000	290673	4
SOPHIE - Ponyboy	0.829	0.382	1	-7.370	1	0.0843	0.000498	0.000370	0.1330	0.452	107.990	194438	4
Tierra Whack - Only Child	0.826	0.365	1	-7.524	1	0.0412	0.049000	0.000004	0.1860	0.802	85.244	238024	4
GoldLink - Summertime (feat. Wale & Radiant Children)	0.807	0.689	10	-6.646	0	0.2070	0.249000	0.000059	0.3360	0.677	98.014	226164	4

Initially, I was confused. A track like 'Run with Faith' by Prose scores low on speechiness despite it being mostly rap, maybe being due to it's quite melodic music track. However, a feature like acousticness feels quite spot on scoring quite low on most tracks. The highest scores it gets are on tracks that are based on electronic music, like Nujabes and Mid-Air Thief where the scores are 0.66 and 0.58.

I was intrigued by the feature, danceability as it feels like a very subjective thing. Most people agree that heavy metal is loud, but whether or not you can dance to something is based much more on taste and culture than anything else. For example score 'Ponyboy' by SOPHIE quite high even though I wouldn't categorize her as danceable.

Most tracks that gets high scores in danceability also get quite high scores on features like energy and low on features like instrumentalness and acoustictness which feels very intuitive.

The most danceable track is surprisingly, Breakfast 1pm by Ravi Kuma. Though admittedly upbeat, I would never try to start a dancefloor with it. Toast by Koffee is also one of the most danceable track, but her track, Rapture is also the least danceable track, which feels very off as they are quite similar to listen to.

A lot seems quite random, but I do think that the technicalities of the sounds might also be lost to me who has absolutely no knowledge of music.

TASK 3

The content-based recommender systems operate under the assumption that the greater the similarity between features, the more likely it is that you will enjoy a song based on another. I am curious to see the recommendations generated when I select the most danceable song and prioritize danceability as the main feature. How many additional features I need to include before they become irrelevant in the recommendations.

Recommendations based only on danceability:

['Ravi Kuma - Office Man', 'Tierra Whack - Hungry Hippo', Nujabes - Luv (sic.) pt3 (feat. Shing02)', 'Ashnikko - Daisy', 'Koffee - W (feat. Gunna)', 'StarBoy - Soco (feat. Wizkid, Ceeza Milli, Spotless & Terri)', Greentea Peng - Mr. Sun (miss da sun)', 'Tierra Whack - Only Child', SOPHIE - Immaterial', **Father - Heartthrob']**

The ones in bold are the tracks that are present in the recommendations and on the top scores of danceability. 7/10 were on both lists.

Recommendations based on danceability, energy, loudness, tempo:

['Ravi Kuma - Office Man', 'RMR - DEALER (feat. Future & Lil Baby)', 'Tierra Whack - Hungry Hippo', 'Meanz Of Operation - Janet', 'Father - Heartthrob', 'Ashnikko - Daisy', 'Princess Nokia - Brujas', 'Koffee - Toast', 'Naira Marley - Soapy', 'StarBoy - Soco (feat. Wizkid, Ceeza Milli, Spotless & Terri)'], 10/10 were on both lists.

Recommendations based on danceability, acoustictness and instrumentalness:

['Ashnikko - Daisy', 'GoldLink - Summatime (feat. Wale & Radiant Children)', 'Nujabes - Luv (sic.) pt3 (feat. Shing02)', 'Meanz Of Operation - Janet', 'Naira Marley - Soapy', 'Koffee - W (feat. Gunna)', 'Princess Nokia - Brujas', 'RMR - DEALER (feat. Future & Lil Baby)', 'AUDREY NUNA - damn Right', 'Tierra Whack - CLONES'**]**

7/10 were on both lists

The recommender system produced surprisingly accurate recommendations based on the main feature of danceability. Adding features like energy, loudness, and tempo improved the recommendations, while having features that scored very low surprisingly seemed to not affect the results that much.