

SPOTIFY & GENIUS TRACK

Nefertiti Damnjanović IN38/2020
Sara Stefanović IN55/2020



SADRŽAJ PREZENTACIJE

- 01 - Opis projekta
- 02 - Predstavljanje originalnog skupa podataka
- 03 - Primer prostiranja podataka
- 04 - Izmenjeni skup podataka
- 05 - Predlog logičke šeme
- 06 - Upiti nad šemom podataka



01-OPIS PROJEKTA

- Korišćenjem *MongoDB* baze podataka, istražićemo različite faktore koji utiču na popularnost pesama, korišćenjem 10 upita koji će se fokusirati na različite segmente baze podataka koju istražujemo.
- Cilj analize je identifikovanje i vizualizacija obrazaca i trendova koji doprinose uspehu pesama na muzičkom tržištu, na platformama Spotify i Genius.
- Baza podataka koju koristimo:
<https://www.kaggle.com/datasets/saurabhshahane/spotgen-music-dataset?resource=download>





02- ORIGINALNI SKUP PODATAKA



Spotify albums



Spotify artists



Spotify tracks



Lyrics



TABELA SPOTIFY ALBUMS

- Tabela pod nazivom "spotify_albums" sadrži informacije o albumima dostupnim na Spotify platformi.
- Broj zapisa: 75511
- Broj obeležja: 16

Obeležja u tabeli:

- unnamed - Neimenovani albumi
- id - Id albuma
- name - Naziv albuma
- album_type - Tip albuma
- artist_id - Id izvođača
- available_markets - Dostupna tržišta
- external_urls - Spoljni URL-ovi
- href - Referenca na hipervezu
- images - Slike albuma
- release_date - Datum izdavanja albuma
- release_date_precision - Preciznost datum izdavanja albuma
- total_tracks - Broj pesama na albumu
- track_id - Id pesme na albumu
- track_name_prev - Prethodni naziv pesme
- uri (Uniform Resource Identifier) - Jedinstveni identifikator resursa
- type - Tip



TABELA SPOTIFY ARTISTS

- Tabela pod nazivom "spotify_artists" sadrži informacije o izvođačima dostupnim na Spotify platformi.
- Broj zapisa: 56129
- Broj obeležja: 9

Obeležja u tabeli:

- unnamed - Neimenovani albumi
- artist_popularity - Popularnost izvođača
- followers - Pratioci
- genres - Žanrovi
- id - Id izvođača
- name - Ime izvođača
- track_id - Id pesme
- track_name_prev - Prethodni naziv pesme
- type - Tip



TABELA SPOTIFY TRACKS

- Tabela pod nazivom "spotify_tracks" sadrži informacije o pesmama dostupnim na Spotify platformi.
- Broj zapisa: 101939
- Broj obeležja: 32

Obeležja u tabeli:

- unnamed - Neimenovani albumi
- acousticness - Da li je pesma akustična ili ne
- album_id - Id albuma na kojoj se pesma nalazi
- analysis_url - Url ka analizi pesme
- artists_id - Id izvođača
- available_markets - Dostupna tržišta
- country - Zemlja porekla
- danceability - Pogodnost za ples
- disc_number - Broj diska
- duration_ms - Trajanje pesme (u milisekundama)
- energy - Energija pesme
- href - Referenca na hipervezu
- id - Id pesme
- instrumentalness - Instumentalnost
- key - Tonalitet
- liveness - Uživo izvedba
- lyrics - Tekst pesme
- type - Tip
- valence - Emocionalni ton
- mode - Modalitet (dur ili mol)
- name - Naziv pesme
- playlist - Plejlista kojoj pesma pripada
- popularity - Ocena popularnosti pesme
- preview_url - Url za preview
- speechiness - Udeo teksta u pesmi
- tempo - Tempo
- time_signature - Takt
- track_href - Referenca na hipervezu pesme
- track_name_prev - Prethodni naziv pesme
- track_number - Broj pesme
- uri - Jedinstveni identifikator resursa



TABELA LYRICS FEATURES

- Tabela pod nazivom "lyrics_features" sadrži atribute vezane za tekstove pesama, uključujući dužinu teksta, prosečnu dužinu reči i leksičku raznovrsnost.
- Broj zapisa: 94954
- Broj obeležja: 8

Obeležja u tabeli:

- unnamed - Neimenovani albumi
- mean_syllables_word - Prosečan broj slogova po reči u tekstu pesme
- mean_words_sentence - Prosečan broj reči po rečenici u tekstu pesme
- n_sentences - Ukupan broj rečenica u tekstu pesme
- n_words - Ukupan broj reči u tekstu pesme
- sentence_similarity - Sličnost između rečenica u tekstu
- track_id - Id pesme
- vocabulary_wealth - Bogatstvo vokabulara, range je od -1 do 1

03 - PRIMER PROSTIRANJA: SPOTIFY ALBUMS

- unnamed: 1
- id: 4KfJZV7WfolYlxBzOTo66s
- name: Shostakovich Symphony No.5 - Four Romances
on ...
- album_type: album
- artist_id: 6s1pCNXcbdtQJlsnM1hRIA
- available_markets: 'AD', 'AE', 'AR', 'AT', 'AU', 'BE', 'BG',
'BH...'
- external_urls: <https://open.spotify.com/album/4KfJZV7WfolYlxBzOTo66s>
- href: <https://api.spotify.com/v1/albums/4KfJZV7WfolYlxBzOTo66s>
- images: {'height': 640, 'url': '[https://i.scdn.co/ima...
name Shostakovich Symphony No.5 - Four Romances](https://i.scdn.co/image/4KfJZV7WfolYlxBzOTo66s/640x640.jpg?market=BG&format=auto)
- release_date: 2019-03-01
- release_date_precision: day
- total_tracks: 8
- track_id: 1WQfghEjszJJ4H8MAWrQ2C
- track_name_prev: track_11
- uri (Uniform Resource Identifier):
<spotify:album:4KfJZV7WfolYlxBzOTo66s>
- type: album



03 - PRIMER PROSTIRANJA: SPOTIFY ARTISTS



- unnamed: 2
- artist_popularity: 26
- followers: 1596
- genres: danish pop rock
- id: 6YVY310fjfUzKi8hiqR7iK
- name: Gangway
- track_id: 1bFqWDbvHmZe2f4Nf9qaD8
- track_name_prev: track_38
- type: artist

03 - PRIMER PROSTIRANJA: SPOTIFY TRACKS

- unnamed: 1284
- acousticness: 0.957
- album_id: 5fpbXZI4uOWbhlc63Bzj8A
- analysis_url: <https://api.spotify.com/v1/audio-analysis/5fan>
- artists_id: 4EvJloPitUn0DBAcMhzbRR
- available_markets: ['AD', 'AE', 'AR', 'AT', 'AU', 'BE', 'BG', 'BH...
- country: BE
- danceability: 0.69
- disc_number: 1.0
- duration_ms: 102267.0
- energy: 0.167
- href: <https://api.spotify.com/v1/tracks/5fanHEexUwpw>
- id: 5fanHEexUwpwGqvW50y6Xp
- instrumentalness: 0.000005
- key: 6.0

- liveness: 0.115
- lyrics: \r\nI have seen, desolate one, the voice has i...
- type: track
- valence: 0.236
- mode: 1.0
- name: Song Tournament, New Style
- playlist: Modern Poetry
- popularity: 4.0
- preview_url: <https://p.scdn.co/mp3-preview/99a4a13c8ef5d9e7>
- speechiness: 0.891
- tempo: 139.65
- time_signature: 4.0
- track_href: <https://api.spotify.com/v1/tracks/5fanHEexUwpw>
- track_name_prev: track_69
- track_number: 9.0
- uri: spotify:track:5fanHEexUwpwGqvW50y6Xp



03 - PRIMER PROSTIRANJA: LYRICS FEATURES



- unnamed: 618
- mean_syllables_word: 1.8
- mean_words_sentence: 3.72
- n_sentences: 76
- n_words : 411
- sentence_similarity: 0.152982
- track_id: 6wLsDD4wXkypWRM9fn1ZZ0
- vocabulary_wealth: 0.27



04- IZMENJENI SKUP PODATAKA



Spotify albums:

- Broj zapisa: 75511
- Broj obeležja: 9
- Obrisana obeležja: unnamed, available_markets, external_urls, href, images, uri, type



Spotify artists:

- Broj zapisa: 56129
- Broj obeležja: 6
- Obrisana obeležja: unnamed, track_name_prev, type



Spotify tracks:

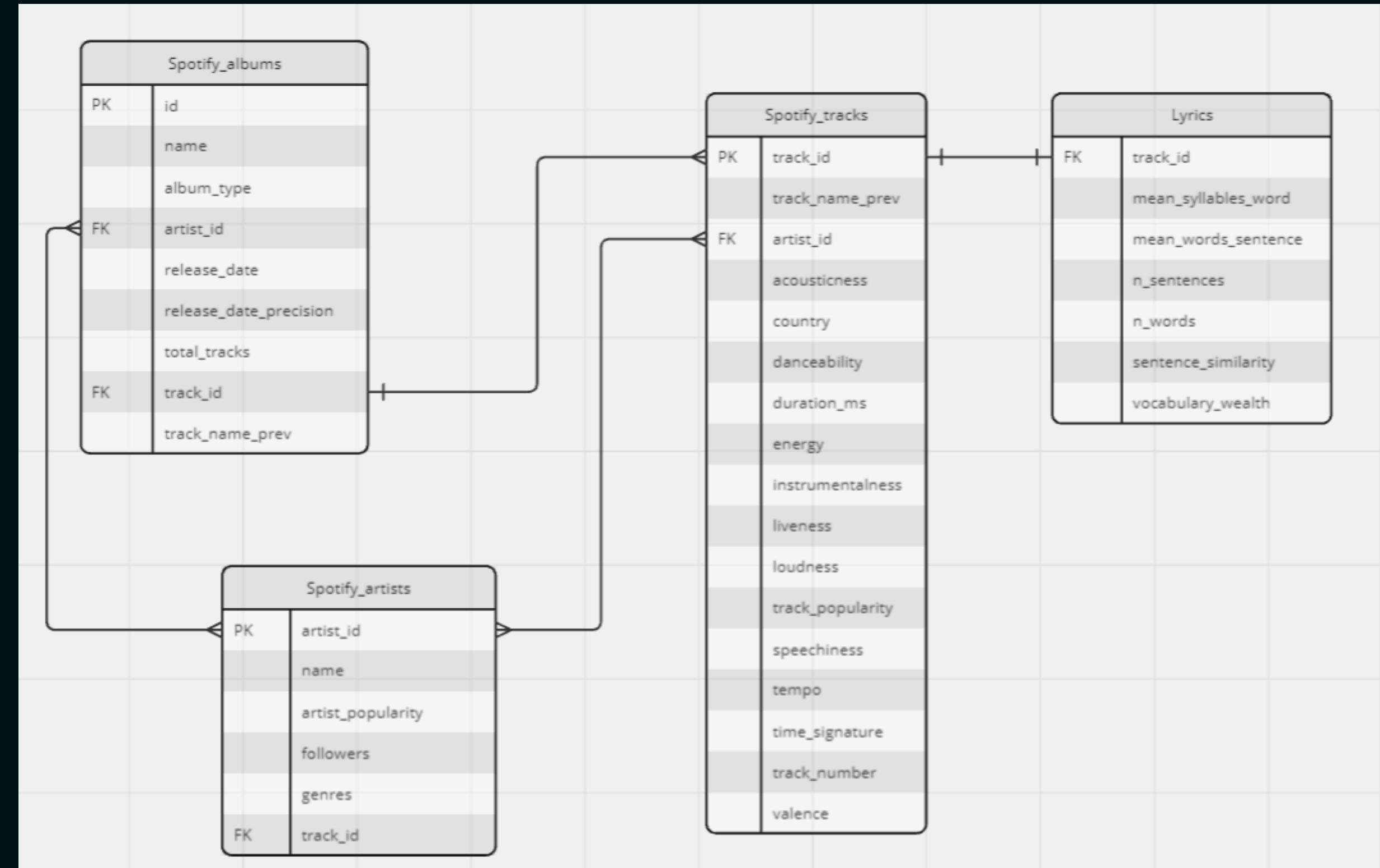
- Broj zapisa: 101939
- Broj obeležja: 17
- Obrisana obeležja: unnamed, album_id, analysis_url, disc_number, playlist, uri, available_markets, href, mode lyrics, name, preview_url, track_href, type



Lyrics:

- Broj zapisa: 94954
- Broj obeležja: 7
- Obrisana obeležja: unnamed

05 - LOGIČKA ŠEMA





06 - UPITI NAD ŠEMOM PODATAKA

Nefertiti:

1. Kako popularnost pesme korelira sa audio karakteristikama poput plesnosti, energije, valence, vocabulary_wealth?
2. Koje su top 3 zemlje sa najvećom prosečnom popularnošću pesama i kako se one razlikuju u audio karakteristikama?
3. Kako akustičnost pesme utiče na njenu popularnost u različitim tržištima?
4. Koji albumi imaju najveću kumulativnu popularnost pesama?
5. Koje su lirske karakteristike najpopularnijih pesama?

Sara:

1. Koja je najpopularnija pesma?
2. U kojoj godini je objavljeno najviše albuma?
3. Odrediti iz koje godine je najduzi album.
4. Koja je prosečna popularnost pesama na albumima iz 2019. godine?
5. Da li postoji korelacija izmedju obeležja vocabulary_wealth i popularity?

INDEKSI

1. db.spotify_tracks.createIndex({ id: 1 });
db.spotify_tracks.createIndex({ popularity: 1, danceability: 1,
energy: 1, valence: 1 });

db.lyrics_features.createIndex({ track_id: 1 });

2. db.spotify_tracks.createIndex({ country: 1, popularity: 1 });

db.spotify_tracks.createIndex({ danceability: 1 });
db.spotify_tracks.createIndex({ energy: 1 });
db.spotify_tracks.createIndex({ valence: 1 });

3. db.spotify_tracks.createIndex({ available_markets: 1 });
db.spotify_tracks.createIndex({ acousticness: 1 });

db.spotify_tracks.createIndex({ available_markets: 1, popularity: 1 });

4. db.spotify_albums.createIndex({ id: 1 });
db.spotify_tracks.createIndex({ album_id: 1 });

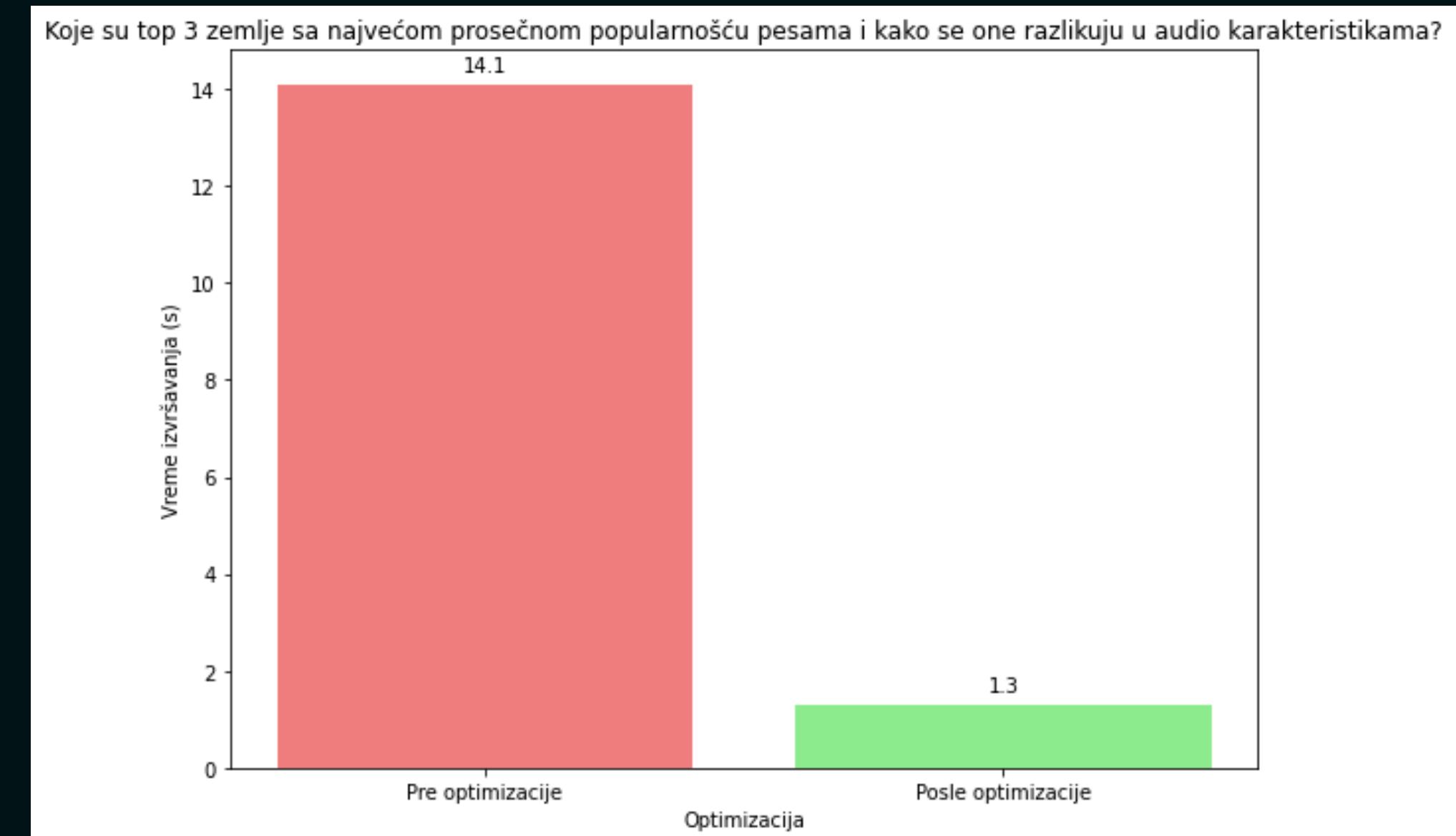
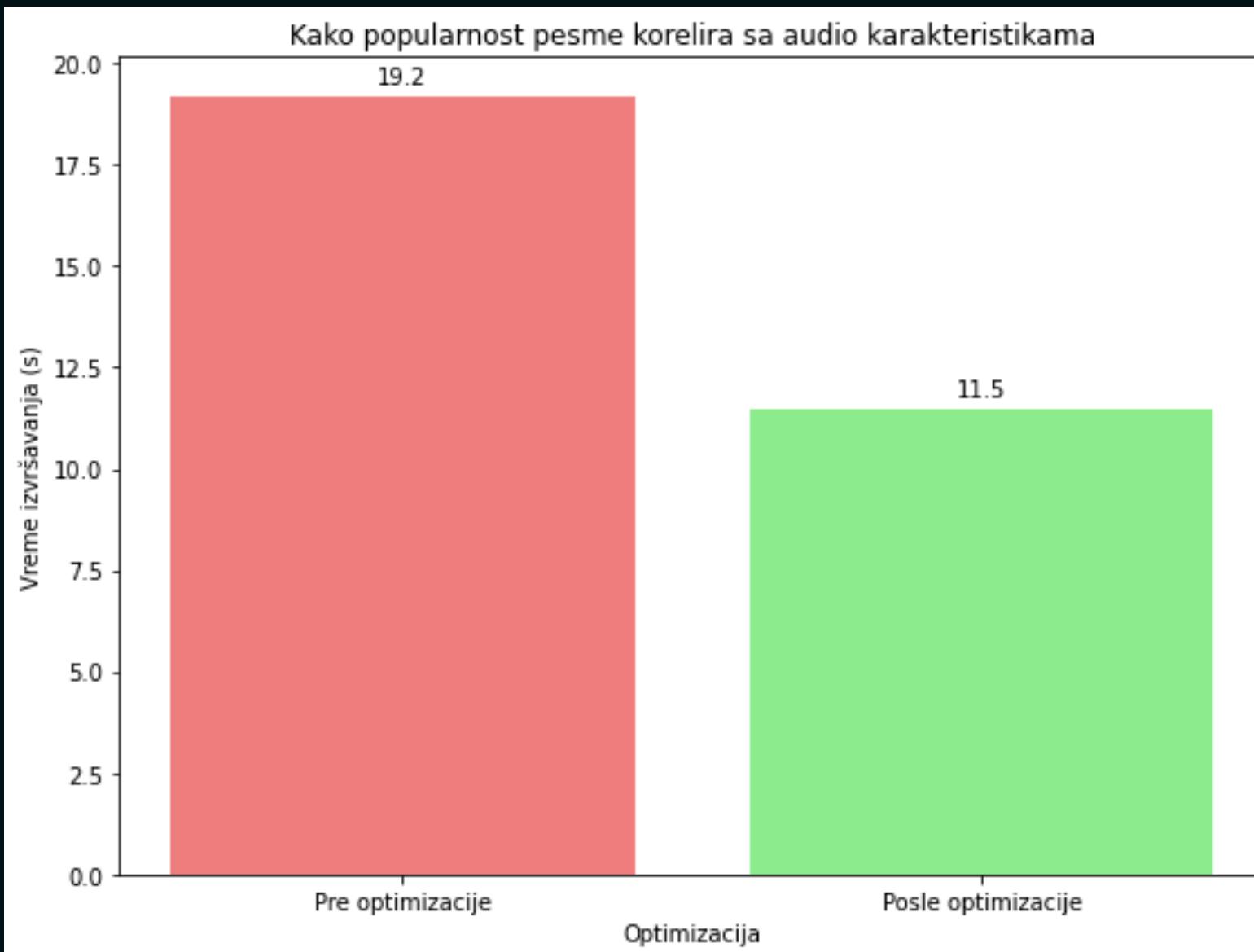
db.spotify_tracks.createIndex({ "album_tracks.popularity": 1 });

5. db.spotify_tracks.createIndex({ id: 1 });

db.lyrics_features.createIndex({ track_id: 1 });

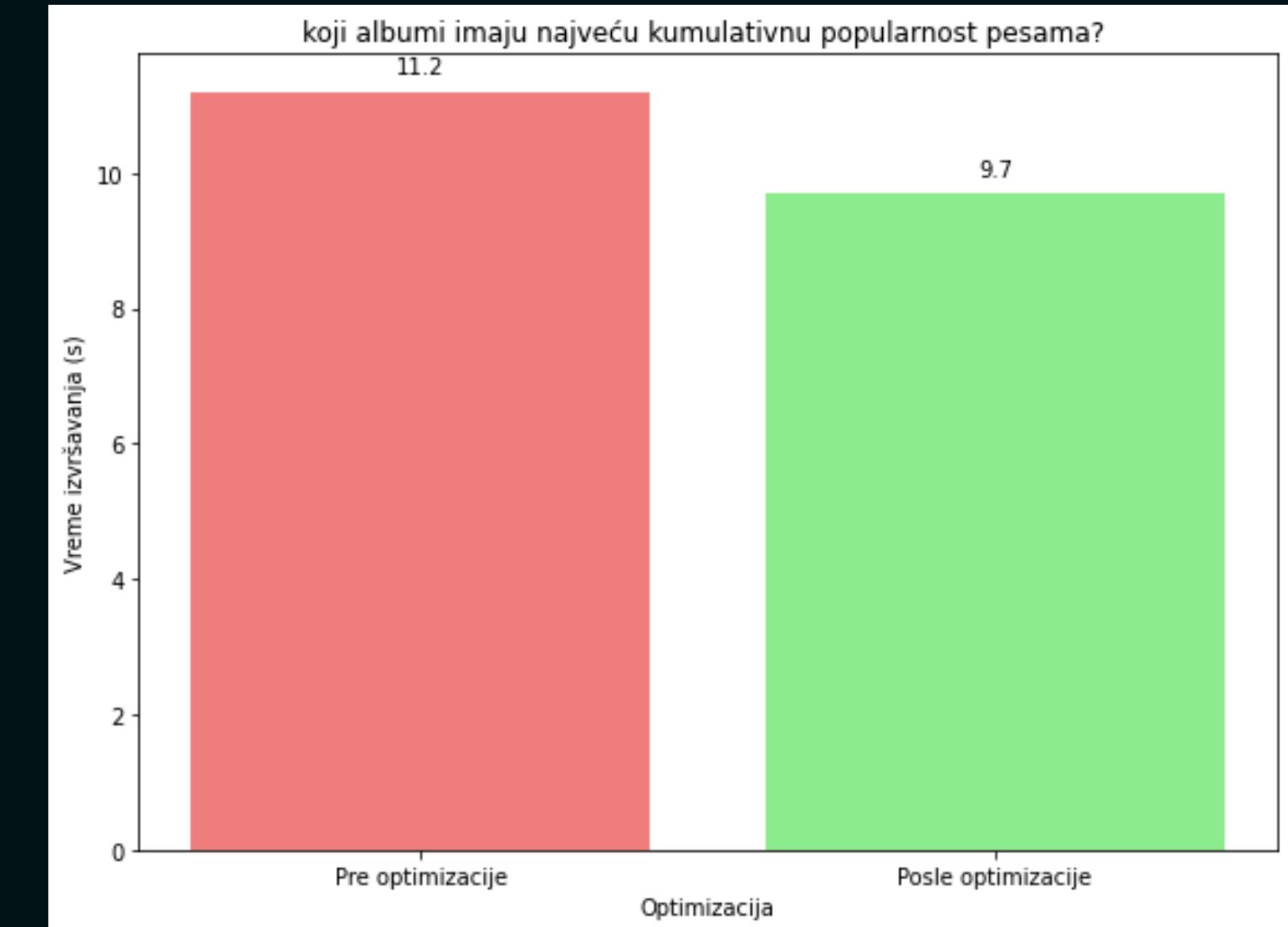
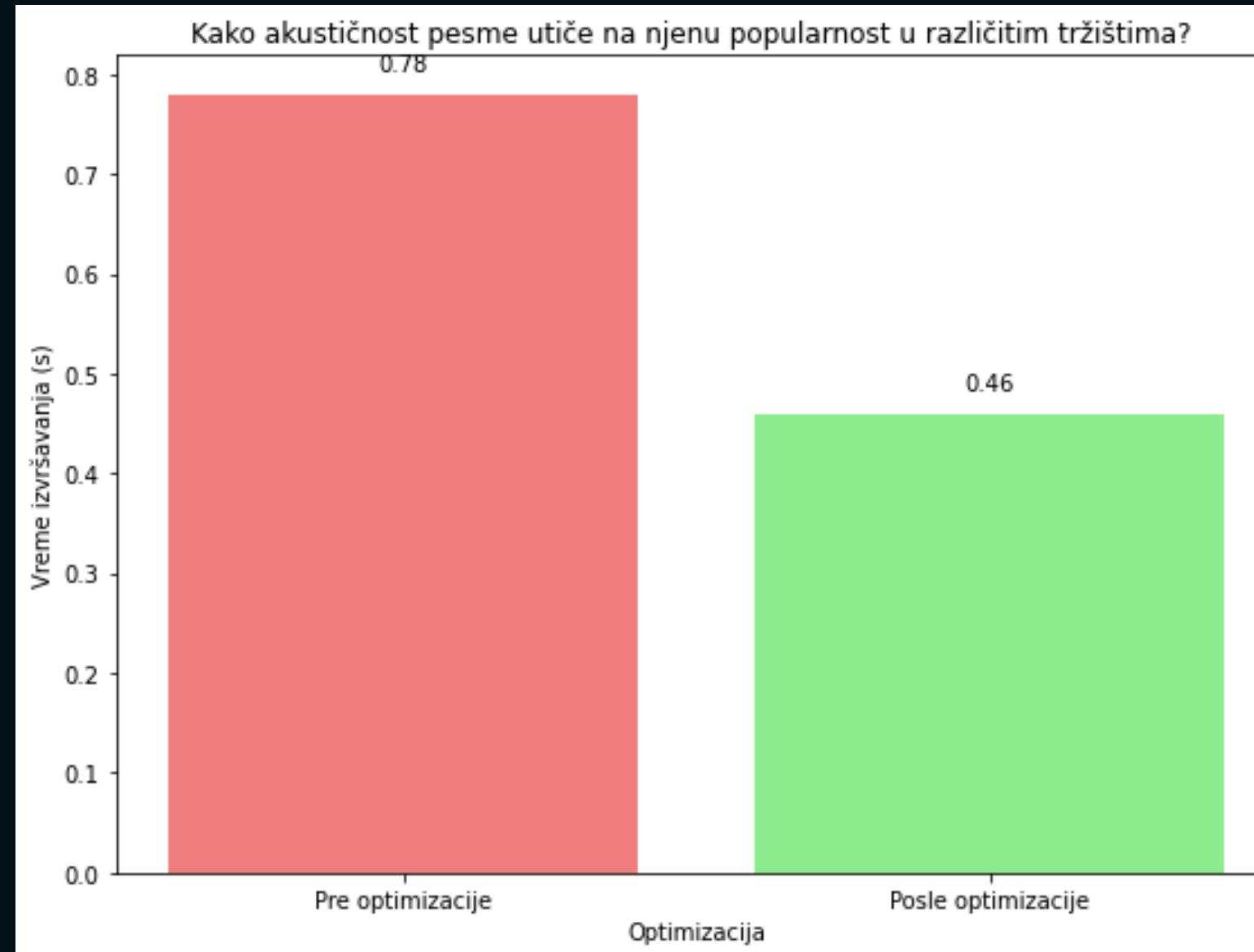
Grafikon koji prikazuje razliku u vremenima izvršavanja pre i posle optimizacije

UPIT 1&2



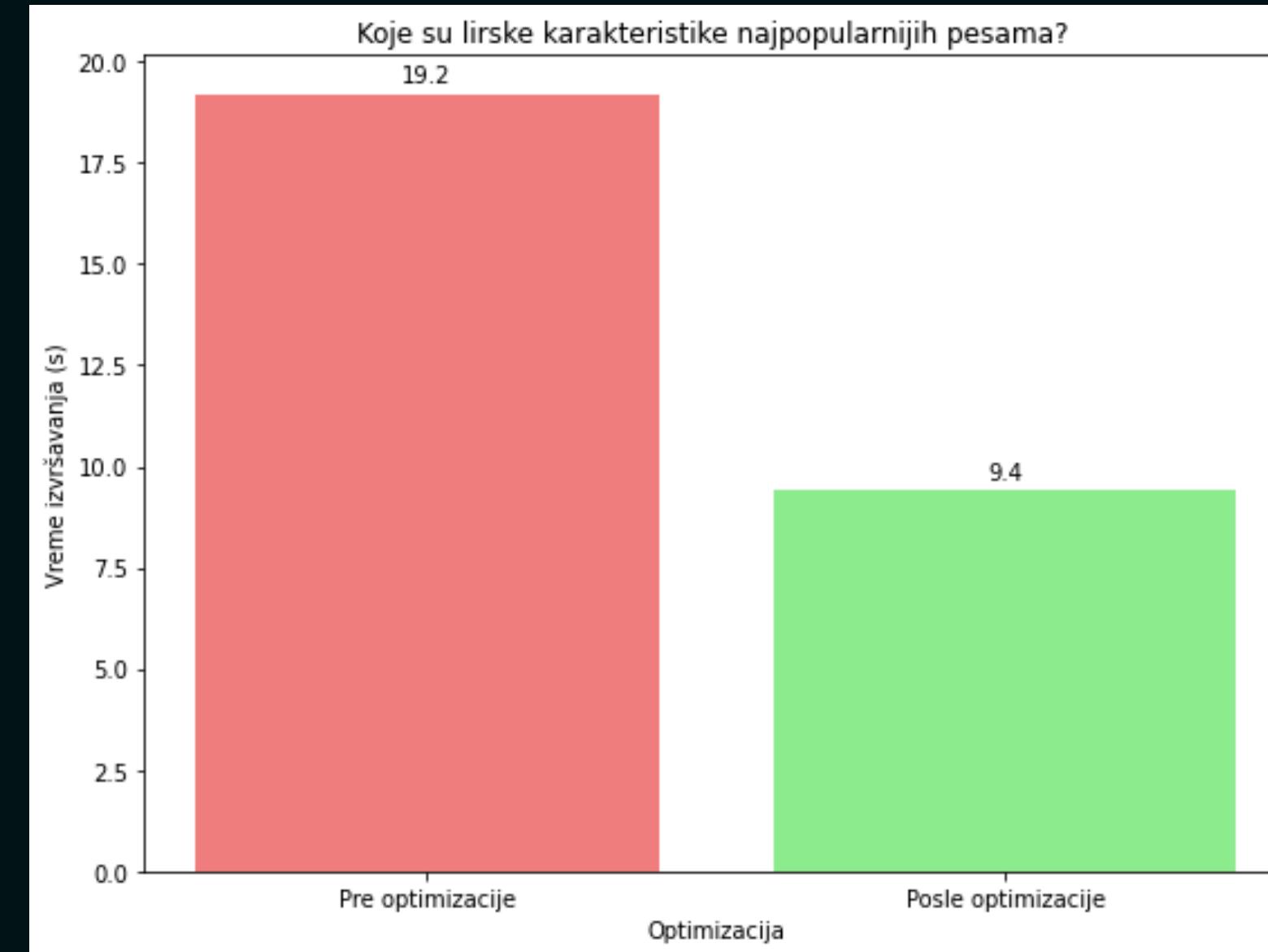
Grafikon koji prikazuje razliku u vremenima izvršavanja pre i posle optimizacije

UPIT 3&4



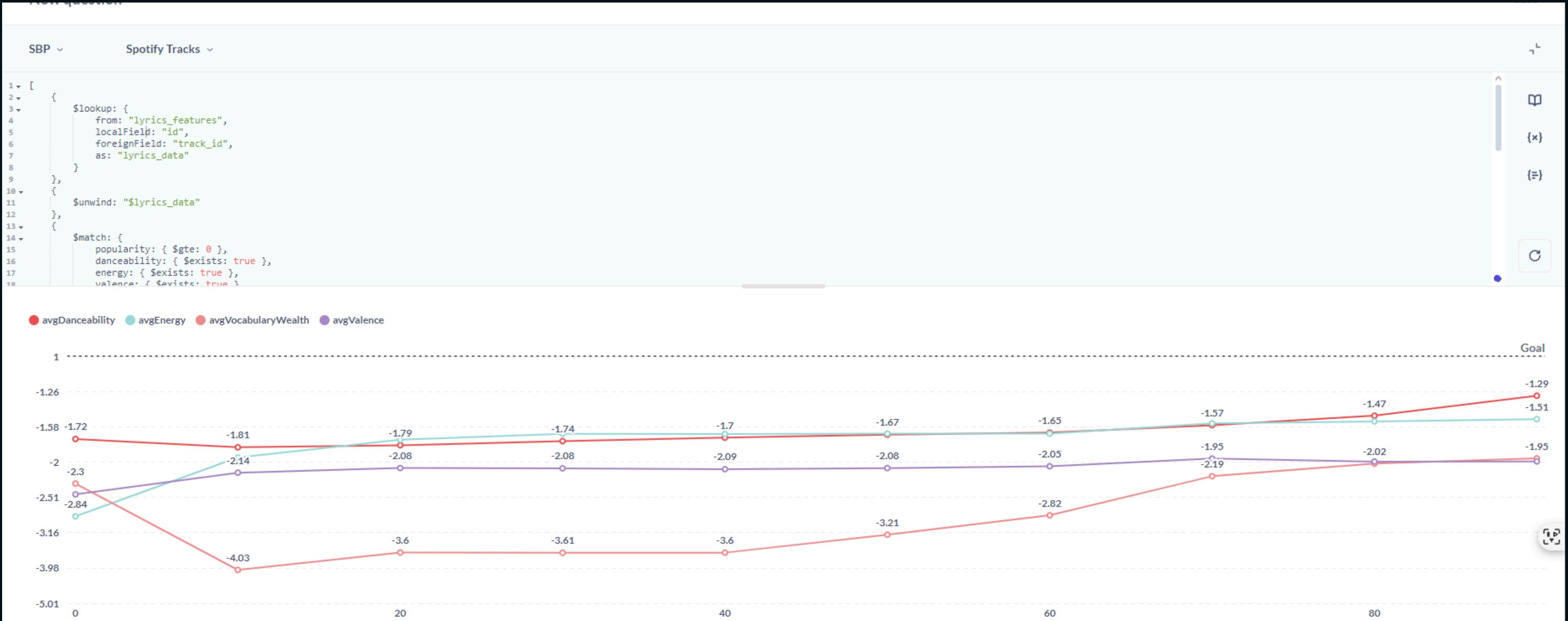
Grafikon koji prikazuje razliku u vremenima izvršavanja pre i posle optimizacije

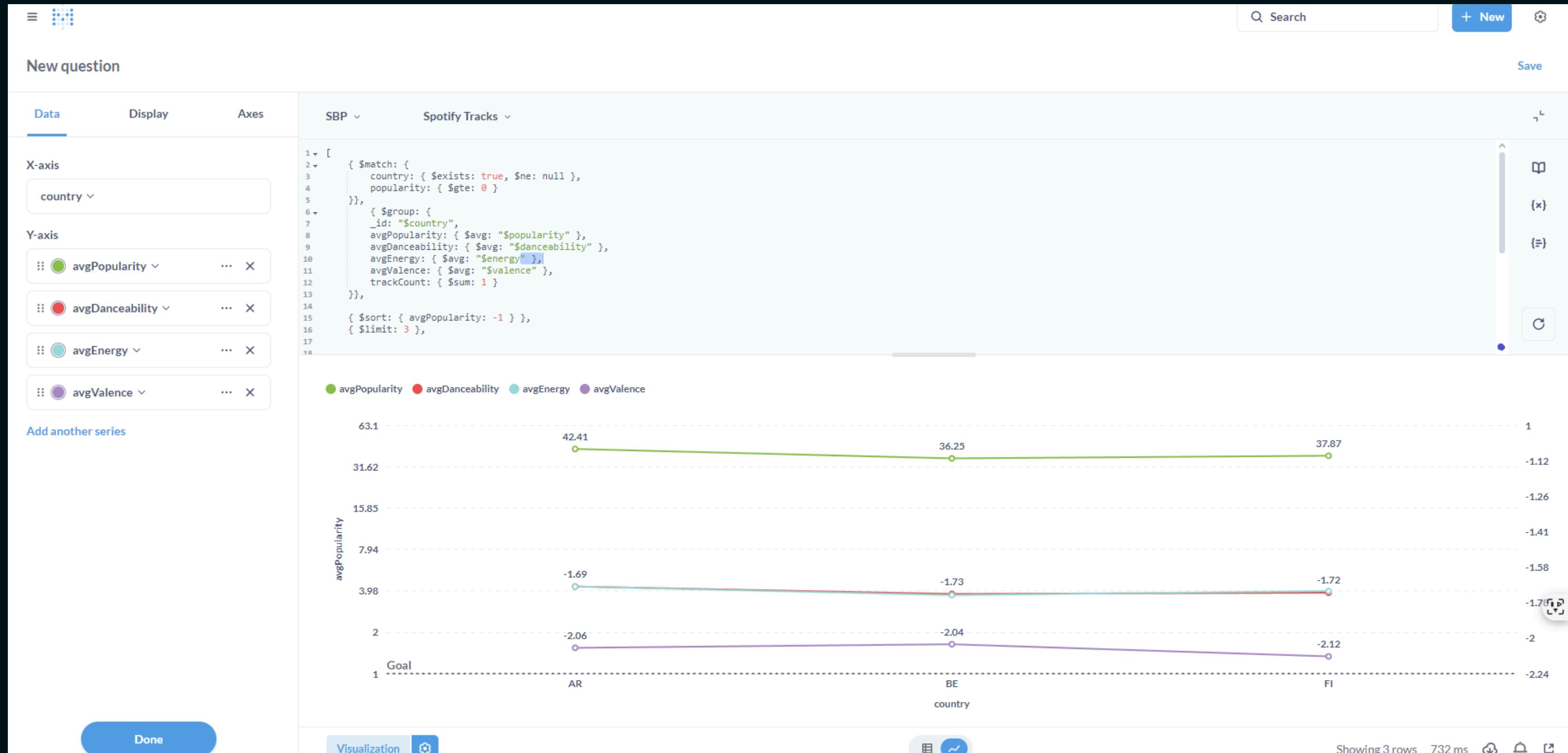
UPIT 5



METABASE

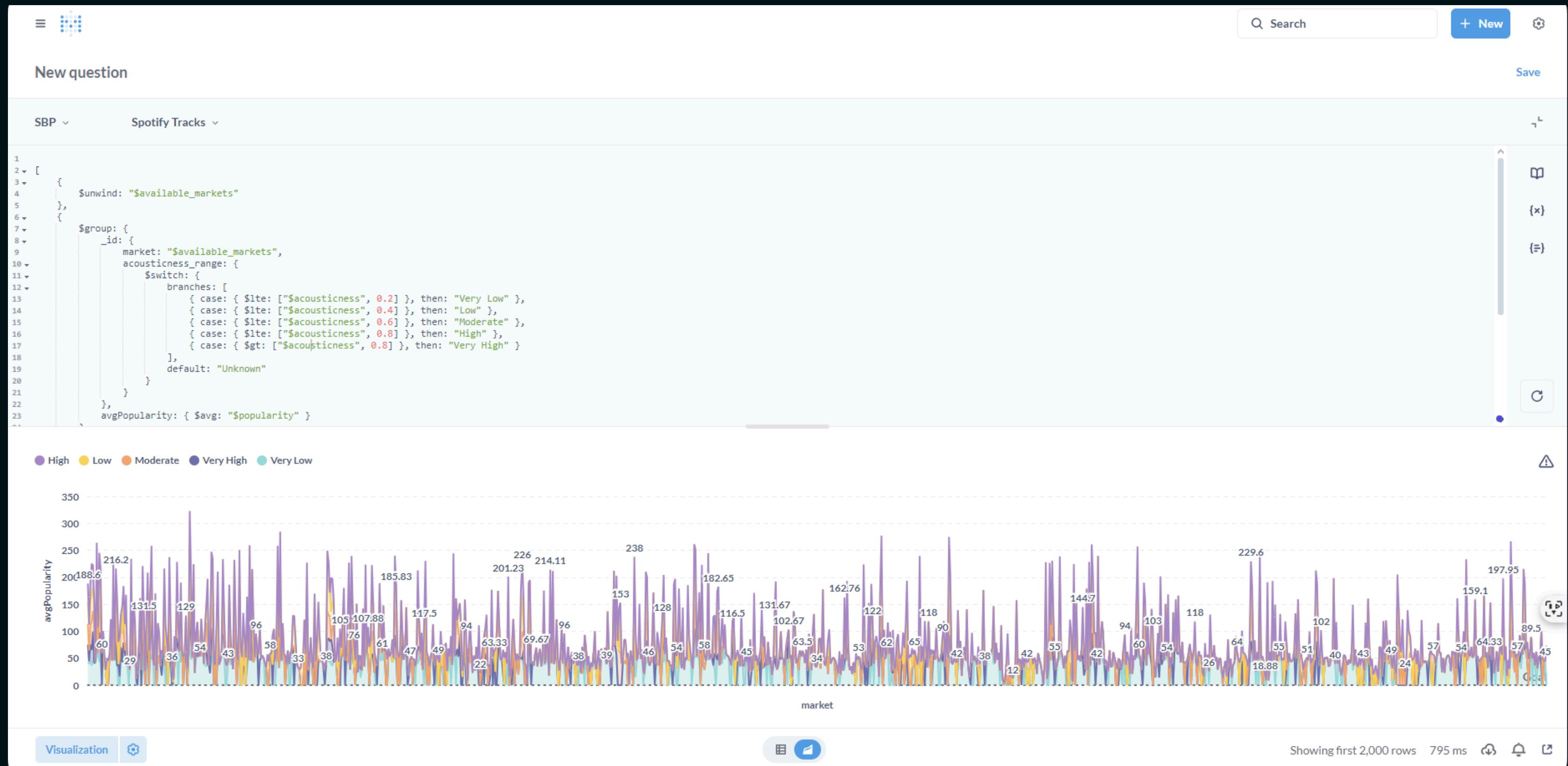
UPIT 1

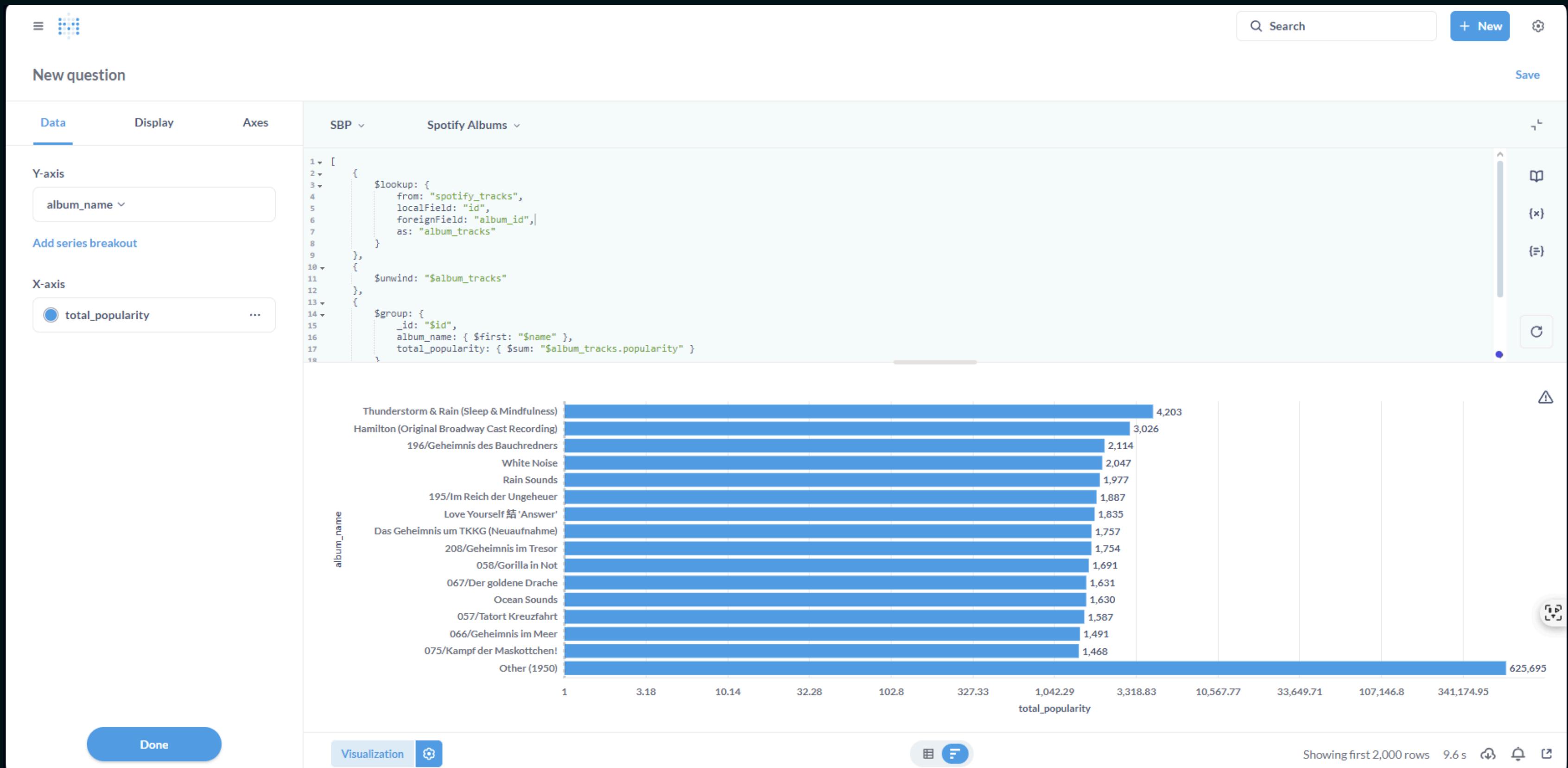


METABASE
UPIT 2

METABASE

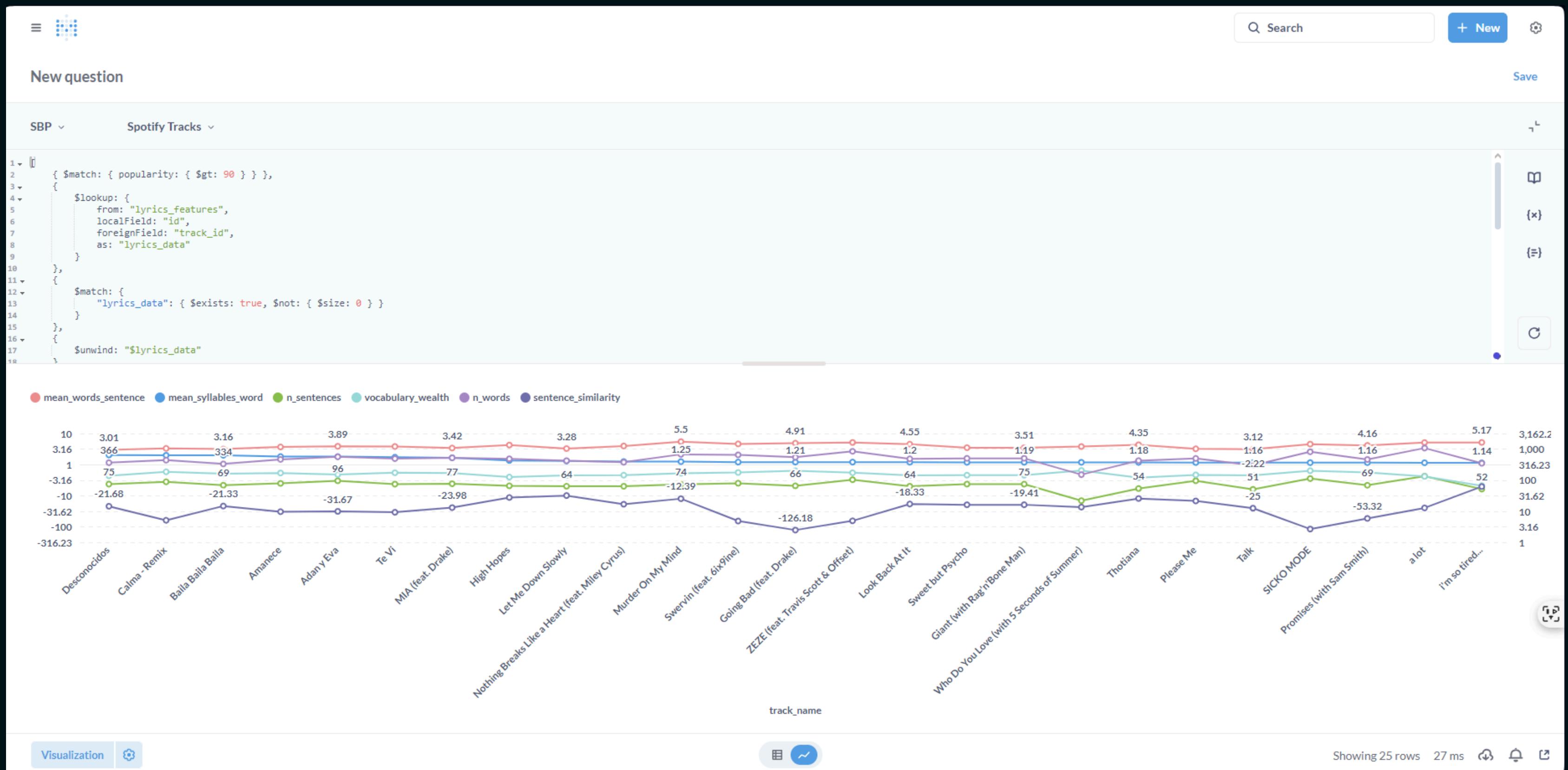
UPIT 3



METABASE
UPIT 4

METABASE

UPIT 5



A photograph of a large crowd at a concert. The people are silhouetted against the bright stage lights, with many of their hands raised in the air. The scene is filled with warm, golden light from the stage spotlights.

HUĀLĀ NAPĀŽNU