



Analiza podataka Spotify aplikacije





Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

Uvod



Spotify je najpoznatija i najprestižnija platforma za streaming muzike na svetu, koja okuplja veliki broj korisnika i izvođača iz različitih krajeva sveta.



Analiziranjem podataka o slušanosti i interakcijama na Spotify-u, dobijamo bolji uvid u dinamiku muzičke industrije i kako se ukusi korisnika menjaju.



Ova analiza nam omogućava da otkrijemo određene obrasce, pratimo trendove i identifikujemo ključne faktore koji utiču na to koje pesme postaju hitovi, kako izvođači napreduju i kako korisnici biraju šta će slušati.



Takođe, Spotify koristi ove podatke da poboljša svoje funkcije personalizovanih preporuka, čime korisnicima olakšava pronalaženje muzike koja im se sviđa.



Podaci



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti



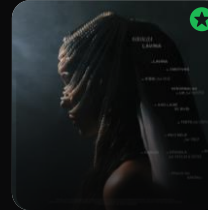
SPOTIFY DATA-SET

Podatke o Spotify-u smo preuzeli sa sajta Kaggle.com i sadrži podatke o izvođačima, pesmama i albumima sa te aplikacije.

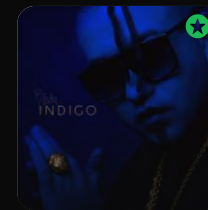
[ADRESA DATA-SETA](#)



Za svaku pesmu postoje podaci o dužini trajanja pesme, koja je na albumu po redu i kom albumu pripada, takođe postoji i naziv pesme i koliko je popularna među korisnicima.



Ovaj dataset sadrži i tabelu o akustičnim osobinama u kojoj su opisane trajanje, energija, glasnoća... Takođe postoje i podaci o izvođačima kao što su broj pratilaca i popularnost među korisnicima, i podaci o albumima kao što je ime, tip, popularnost ...



Dataset poseduje podatke od preko milion izvođača, preko 8 miliona pesama i skoro 5 miliona albuma. Čitav dataset zauzima 5.17GB memorije.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logičke šeme



Upiti



Miljan upiti



Aleksandar upiti

Semantika

Predstavlja podatke o pismama ima kolone: audio_feature_id-Id akustičnih osobina, duration-dužina trajanja pesme popularity-popularnost pesme...

Tracks



Za ovu kolekciju kolone od značaja su: loudness i liveness – glasnoca i živoća pesme kao i energy i tempo – koji predstavljaju energiju i tempo pesme

Audio_features



Predstavlja tabelu za izvođače i sadrži kolone name, popularity i followers koje su tu da predstave ime, popularnost među korisnicima i broj pratilaca izvođača.

Artists



Kolekcije r_albums_artists, r_albums_tracks, r_artist_genre i r_track_artist predstavljaju međutabele.

Ostalo



Albums sadrži polja release_date, name i album_type što predstavlja datum izbacivanja albuma, ime i tip albuma



Međutabele u sebi sadrže samo Id-eve povezanih modela.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

Prostiranje ključeva

Neki od primera prostiranja ključeva su:

track_id i artist_id u međutabeli r_track_artist koja govori koji izvođač izvodi koju pesmu, i po istom principu ostale međutabele.



Audio_feature_id u tracks koja spaja osobine pesama iz audio_features i pesmu.





Prikaz trenutne šeme



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

id	name	album_group	album_type	release_date	popularity
2JkoV1U7VAmExKJ1Jh3w9P	Alkoholik (feat. Erik Sermon, J Ro & Tash)		album	954633600000	0
4JVUDrtPZuVN17FanbHyux	Flava in Ya Ear Remix (feat. Notorious B.I.G., L.L. Cool J, Busta Rhymes, Rampage)		single	757382400000	0
6VjKAKDymlasMqYw731B0w	Bitch Please II		album	959040000000	0
2V1vHjDb4Tyx14A1iCdhAe	Just Dippin'		compilation	1104537600000	0
3UOuBNE1n5peSRqdzv1nWM	Still D.R.E.		album	942710400000	0
2g8HN35AnVGik7B8yMucww	Big Poppa - 2005 Remaster		album	779414400000	0
7i16o9tox1zgHpKUfh9vuC	In Da Club		album	1044489600000	0
6KIKRz9eSTXdnGUnomdtW	Straight Outta Compton		album	587001600000	0
33ZXjLCpiInn8eQIDYEPD	Shook Ones, Pt. II		album	798760000000	0

ALBUMS

name	id	popularity	followers
Xzibit	4tuJQJicOnuZRLiBFdp3Ou	69	1193665
Erik Sermon	2VX0o9LDIVmKigpnwJpOJ	54	142007
J. Ro	3iB0smwGzRKyR8vs2I61xP	45	158
Tash	22qf8cJRzBjIWb2Jc4JeOr	48	3421
Craig Mack	4akj4uteQrrrGxhX9Rjuyf	55	161966
Eminem	7dGJo4pcD2V6cG8kP0tJRR	94	43882754
Dr. Dre	6DPYiyq5kVWQs4RGwxzPC7	83	5591634
Snoop Dogg	7hJcb9fa4a1zc0Q3EaNPoG	87	6831895

ARTISTS

id	disc_number	duration	explicit	audio_feature_id	name	preview_url	track_number	popularity
1dizvxctg9dHEyaYTFuFV1	1	275893	1	1dizvxctg9dHEyaYTFuFV1	Gz And Hustlas (feat. Nancy Fletcher)		12	0
2g8HN35AnVGik7B8yMucww	1	252746	1	2g8HN35AnVGik7B8yMucww	Big Poppa - 2005 Remaster	https://p.scdn.co/mp3-preview/770e023eb8318270ecc5caa818d758e5e0844de97c1dcd021ca5d5e42a8bd440f100...	13	77
49pnyECzcNGCKAqxFTB4JZ	3	315880	0	49pnyECzcNGCKAqxFTB4JZ	You Were Born - Early Version Of "One Of The Three" / Outtake		6	8
4E5IFAXCob6QqZaJMTw5YN	1	248880	1	4E5IFAXCob6QqZaJMTw5YN	Poppin' Them Thangs	https://p.scdn.co/mp3-preview/f3b556ced9657f68b7d2c981814205244daf45487c1dcd021ca5d5e42a8bd440f100...	2	70
1gSt2U1C7mRtJiC5zqkWh	2	203666	0	1gSt2U1C7mRtJiC5zqkWh	It's Hard To Say "I Do", When I Don't		2	50
67IvfvAMYQzJEeHopvMHW	1	173159	1	67IvfvAMYQzJEeHopvMHW	Simon Says	https://p.scdn.co/mp3-preview/e61bbb32e233ab49eda5fcd6964a0ab194ca538c7c1dcd021ca5d5e42a8bd440f100...	1	62

TRACKS





Predlog logičke šeme



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logičke šeme



Upiti



Miljan upiti



Aleksandar upiti

Predlog 1

- ▶ Podatke iz tabele `audio_features` smestićemo u tabelu `tracks` jer se ovi podaci uvek koriste zajedno, preko referencirajućeg polja `audio_feature_id`

Predlog 2

- ▶ Budući da su žanrovi direktno vezani za izvođače, možemo eliminisati tabelu `r_artist_genre` i `genres`. Sve relevantne informacije o žanrovima možemo integrisati direktno u kolekciju `artists`.

Predlog 3

- ▶ Primenićemo šablon proširene reference u cilju optimizacije naše šeme baze podataka kako bismo maksimalno izbegli potrebu za spajanjem tabela i osigurali da dokumenti unapred sadrže sve relevantne informacije



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

Upiti

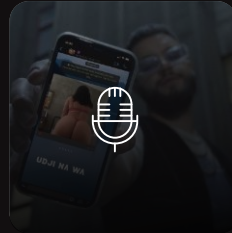
Playlist

Podcast & Shows

Albums

Artist

Downloaded



MILJAN

Naći izvodjače koji su najviše pesama snimali u studiju, takodje prikazati i koliko pesama nije snimio u studiju i odnos studio/van studija.

Pesme kojeg izvodjača su u proseku najtiše?

Koji albumi u proseku imaju najduže pesme po svakoj godini?

Koje su najpopularnije pesme najpopularnijih albuma?

Izlistati godine koje po proseku imaju najenergičnije pesme u opadajućem redosledu.



ALEKSANDAR

Izlistati 10 izvodjaca sa najmanjim prosekom odnosa danceability i energy na njegovim pesmama. Naci sve izvodjace zanra rep, poredjati ih po broju pesama i naci im srednju vrednost spechinessa i najpopularniju pesmu.

Naci najzapracenije izvodjace iz svakog zanra i naci njihove najpopularnije pesme.

Naci top 5 godina sa najmanjim prosekom akusticnosti albuma koji su izašli te godine (akusticnost albuma smatramo prosekom akusticnosti pesama u albumu).

Naci 100 artista sa najvećim odnosom pratioca i broja izbacenih pesama.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

Miljan upiti

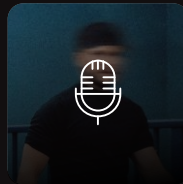
Playlist

Podcast & Shows

Albums

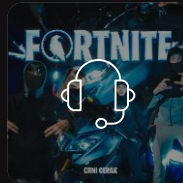
Artist

Downloaded



Naći izvodjače koji su najviše pesama snimali u studiju, takodje prikazati i koliko pesama nije snimio u studiju i odnos studio/van studija.

Umetnici zbog napretka tehnologije sve više biraju da snimaju u studiju. Pomoću ovog upita dobijamo informaciju koji umetnici su najskoliniji tome.



Izlistati jedinstvene izvodjače i njihove najtiše pesme.

Izvođači sve više teže ka glasnim pesmama da bi podstakli adrenalin slušaoca. Zanima nas ko najmanje koristi to u svoju korist.



Koji albumi u proseku imaju najduže pesme po svakoj godini?

Sve je veći trend da su pesme sve kraće što izeljuje manje truda u pisanju samog teksta. Ovim upitom dolazimo do zaključka na kojim albumima je uloženo najviše truda za tekst i muziku.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logičke šeme



Upiti



Miljan upiti



Aleksandar upiti

Miljan upiti

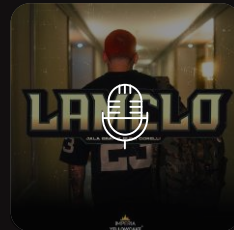
Playlist

Podcast & Shows

Albums

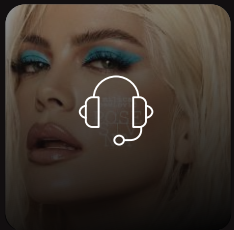
Artist

Downloaded



Koje su najpopularnije pesme najpopularnijih albuma?

Izvođači često traže najpopularnije pesme sa najpopularnijih albuma kako bi ih najviše promovisali, što pomaže u boljoj promociji albuma i povećanju popularnosti izvođača. Cilj ovog upita je obezbeđivanje izvođačima te ključne informacije, uključujući ime albuma, prosečnu popularnost pesama na albumu, broj pesama na albumu, i detalje o izvođačima pesama.



Izlistati godine koje po proseku imaju najenergičnije pesme u opadajućem redosledu.

Trend postaje da su pesme sve više energičnije. Ovim upitom hoćemo da vidimo da li je taj proces bio iz godine u godinu sve vidljiviji ili je bilo drugačijih promena po tom pitanju.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logicke šeme



Upiti



Miljan upiti



Aleksandar upiti

Aleksandar upiti

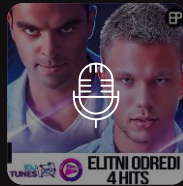
Playlist

Podcast & Shows

Albums

Artist

Downloaded



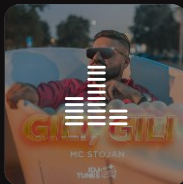
Izlistati 10 izvodjaca sa najmanjim prosekom odnosa danceability i energy na njegovim pesmama.

Ova dva pojma su izuzetno povezana i u današnje vreme je izuzetno retko da je ovaj odnos velik. Ovim upitom možemo zaključiti koji izvodjač je najuspešniji u tome.



Naci sve izvodjace zanra rep, poredjati ih po broju pesama i naci im srednju vrednost spechinessa i najpopularniju pesmu.

Rep je postao sve zastupljeniji žanr među ljudima. Zanima nas ko je najviše pratio taj trend od izvođača i koja pesma mu je najuspesnija.



Naci najzapracenije izvodjace iz svakog zanra i naci njihove najpopularnije pesme.

Cilj je da nadjemo najuspesnije izvodjace iz svakog zanra i njihovu najpoznatiju pesmu sto bi nam reklo najpoznatiju pesmu iz tog zanra takodje.



Uvod



Podaci



Semantika



Prostiranje ključeva



Prikaz trenutne šeme



Predlog logičke šeme



Upiti



Miljan upiti



Aleksandar upiti

Aleksandar upiti

Playlist

Podcast & Shows

Albums

Artist

Downloaded



Naci top 5 godina sa najmanjim prosekom akusticnosti albuma koji su izašli te godine (akusticnost albuma smatramo prosekom akusticnosti pesama u albumu).

Zelimo naci godine gde su izlazile najmanje akusticne pesme jer se aktuelnost akusticnosti menja iz godine u godinu i to u opadajucoj putanji je.



Naci 100 artista sa najvećim odnosom pratioca i broja izbacenih pesama.

Zelimo da nadjemo izvodjace koji su imali najveću uspesnost sa pesmama i za mali broj pesama stekli veliko pracene. Tako mozemo naci i one hit wondere.

Hvala na pažnji

Godinama smo obojica koristili youtube i žalili se na reklame između slušanih pesama kao i na loš kvalitet zvuka. Međutim takođe smo smatrali da spotify nije vredan tih para iako nije skup. Pre godinu dana smo dobili jako dobru preporuku od drugara i isprobali probni period. Spotify se pokazao kao mnogo bolja aplikacija nego što smo mislili i shvatili smo kolko smo pogrešili što ga ranije nismo uzeli i da mu je cena u stvari izuzetno jeftina. Uzimajući ovo u obzir čim smo videli ovaj data-set znali smo da ga moramo odabrati da bismo i drugima predstavili kvalitete ove aplikacije.



<https://github.com/z1cLa/Spotify-Data-Analysis>