

TELIA Lietuva | Data Science internship

Duomenys

Jūs gavote TV filmų rinkinį iš MovieLens duomenų.

- movies_metadata.csv: duomenys apie filmus.
- keywords.csv: filmų raktiniai žodžiai
- ratings.csv: 700 žiūrovų vertinimai suteikti 9000 filmų skalėje nuo 1 iki 5.

Užduotis

- Paanalizuoti duomenis ir pateikti keletą aprašomosios statistikos išvadų
- Pateikti keletą skirtingų grafikų įžvalgų / išvadų pagrindimui.
- Ką duomenys sako apie filmų populiarumą? Ar galime sakyti, kad egzistuoja populiarumą lemiantys faktoriai?

Rezultatas

1. Padarykite aiškų pamatytų įžvalgų ir pasiūlymų pristatymą
2. Failas su naudotais kodais (geriausia Python) (gali būti apjungta į vieną notebook'ą)

Siūlome užduočiai skirti ne daugiau kaip 2-4 valandas. Nesitikim išspręstos užduoties iki galo, labiau tikimės suprasti mąstymo procesą, ir nuo ko ir koku būdu imtumėtės spęsti problemą.