

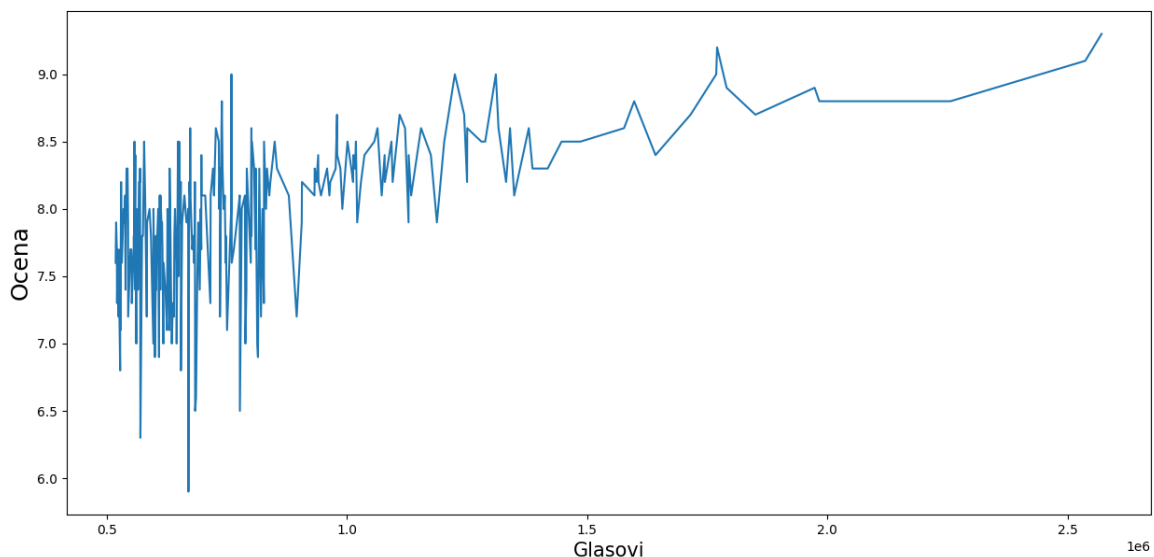
PREDSTAVITEV REZULTRATOV ANALIZE FILMOV IZ IMDb

UVOD

Glavni cilj najine analize je bila analiza nekaj filmov iz IMDb in na podlagi pridobljenih podatkov ugotoviti morebitno povezavo med njimi. S pomočjo te analize bi lahko bolje napovedali kako uspešen bo nek film, ali pa bi nov film prilagodili glede na rezultate, ter tako izboljšali verjetnost njegovega uspeha. Pri analizi sva pridobila podatke s spletne strani IMDb: (https://www.imdb.com/search/title/?title_type=feature&sort=num_votes,desc&count=250) . Za analizo sva vzela prvih 250 filmov, ki imajo na strani prejetih največ glasov in je njihova ocena tako najbolj verodostojna. Pridobila sva podatke o imenu in letu izida filma, ocene, številu glasov in podatke o zasluhu. Za boljšo preglednost, sva se odločila, da narediva dva python programa. Eden nam omogoči, da prebrane podatke zapiše v excel tabelo, kar je uporabniku prijazno in bistveno bolj pregledno, drugi pa nam pomaga z analiziranjem podatkov (pridobivanju nekaterih ključnih informacij in risanju grafov). Funkcije za risanje grafov so pripravljene tako, da se posamezen graf prikaže, hkrati pa se shrani v isto mapo v pdf formatu.

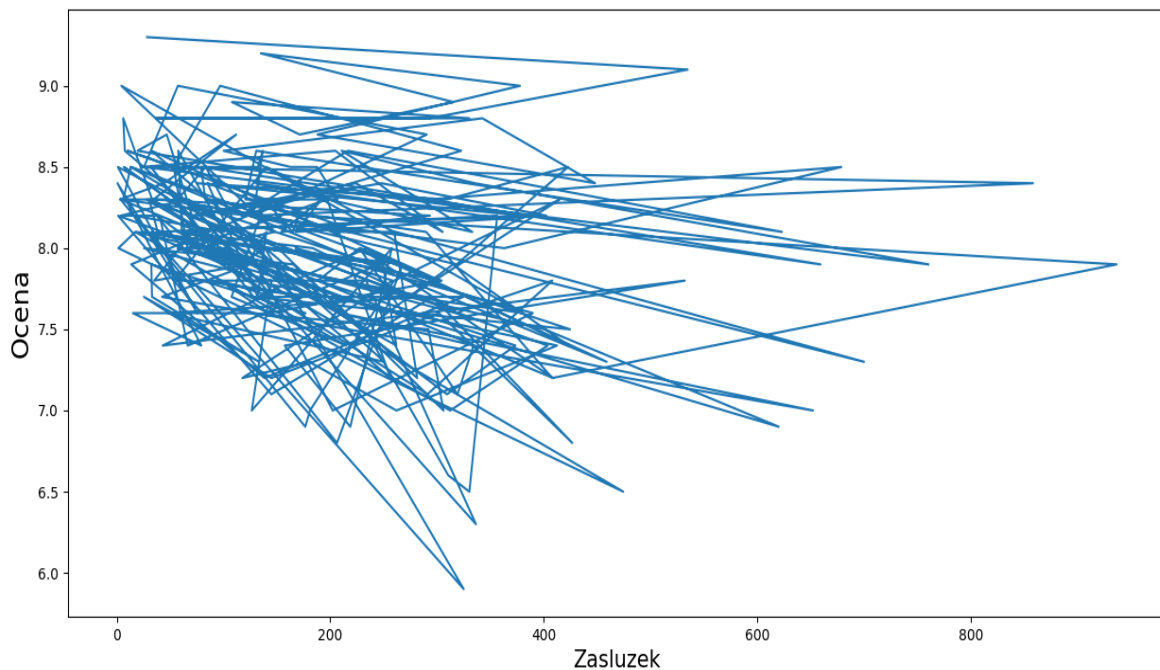
KRATKA ANALIZA

Pogledali si bomo nekaj grafov, ki jih lahko dobimo z našim programom. Zaradi vrste podatkov, ki jih analiziramo, je težko pričakovati da bo pri katerem grafu prišlo do nekega lepega sorazmerja, vendar lahko vseeno poskušamo opaziti kakšno informacijo ali vzorec, s katerim bi si lahko pomagali.



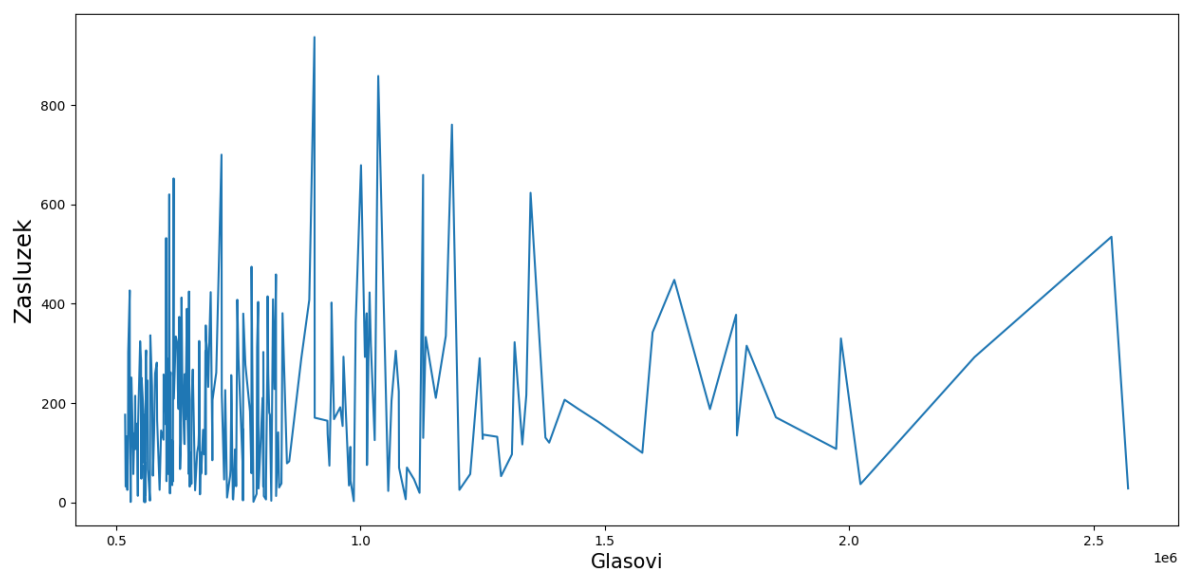
Slika 1 Graf ocene v odvisnosti od števila glasov

V sliki 1 imamo graf, ki prikazuje oceno v odvisnosti od števila glasov nekega filma. Če si pri interpretaciji tega grafa pomagamo z ukazoma `podrobnosti('Glasovi')` in `podrobnosti('Ocena')` v našem programu lahko ugotovimo, da ima kar 75% filmov manj kot 960000 glasov in da je povprečna ocena filma 7.97. Če pogledamo na graf vidimo, da ima večina filmov ki imajo več kot 1 milijon glasov oceno boljšo od povprečja, torej lahko sklepamo, da če nam uspe film predstaviti več ljudem bo imel verjetno tudi boljšo oceno. Še ena možna interpretacija pa je, da večina ljudi pač glasuje za filme, ki so jim všeč.



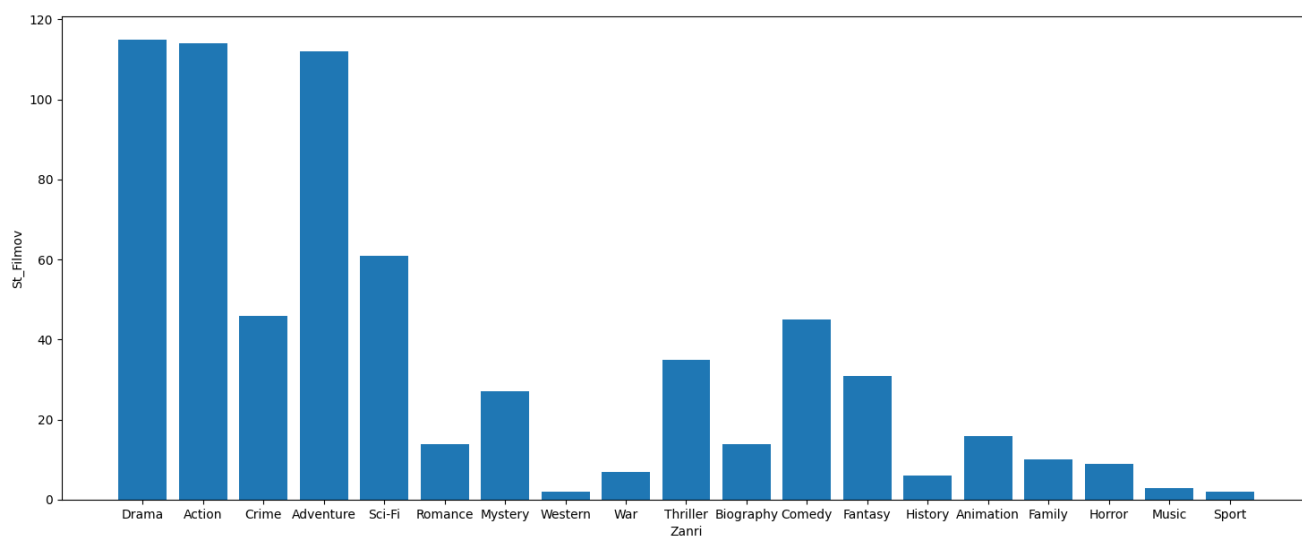
Slika 2 Graf ocene v odvisnosti od zaslužka

Čeprav na videz kaotičen graf na sliki 2, lahko razberemo nekaj zanimivih podatkov. Intuitivno bi pomislili, da bodo filmi z največjimi ocenami imele tudi večji zaslužek, vendar iz grafa vidimo, da temu ni tako. Če pogledamo filma z največjim zaslužkom (Vojna zvezd: Sila se prebuja in Maščevalci: zaključek) imata oba dokaj povprečno oceno (7.9 in 8.4) vendar sta po zaslužku bistveno bolj uspešna kot recimo Kaznilnica odrešitve, ki ima pri oceni 9.3 in več kot 2 milijona glasovih, le nekaj več kot 28 milijonov dolarjev zaslužka. Z uporabo funkcije podrobnosti('Zasluzek') opazimo da ima 75% filmov manj kot 292 milijonov dolarjev zaslužka, pri zaslužku pa opazimo tudi precej visoko standardno deviacijo in sicer 162.24 milijonov. Iz tega grafa je torej težko sklepati, da obstaja kakšno posebno razmerje med oceno filma in zaslužkom.



Slika 3 Graf zaslužka v odvisnosti od števila glasov

Sedaj pogledimo graf, ki prikazuje graf zaslužka v odvisnosti od števila glasov. Tudi tokrat lahko na grafu opazimo, da imajo filmi s številom glasov večjim od 960000, v povprečju večji zaslužek od preostalih filmov, kljub nekaterim izjemam. Že prej smo ugotovili, da ima 75% filmov manj kot 960000 glasov in če pogledamo na graf opazimo, da zelo redki filmi znotraj teh 75% dosegajo zaslužek, ki bi bil primerljiv s filmi z več glasovi. Sklepamo lahko, da večje število glasov pomeni tudi večji zaslužek.



Slika 4 Število filmov določenega žanra

Na grafu številka 4 je očitno kateri žanri prevladujejo. Daleč največ je drame, akcijskih in pustolovskih filmov. Naslednji žanr (Sci-Fi) pa ima že skoraj polovico manj filmov izmed teh, ki smo jih analizirali. Najmanj predstavnikov pa je športnih, glasbenih in western filmov. Verjeten razlog je to, da ljudje ki največ gledajo filme, po navadi ne gledajo teh žanrov.

POVZETEK IN KOMENTAR

Na podlagi opravljene analize lahko sklepamo, da je največji in najpomembnejši dejavnik pri uspehu filma na IMDb predvsem število glasov. Videli smo, da imajo filmi z veliko glasovi v povprečju tako boljšo oceno, kot tudi večji zaslužek. Od žanrov pa smo ugotovili, da nam največjo možnost velikega števila glasov zagotavljajo drame, akcije in pustolovščine. Analizo bi lahko izboljšali, tako da bi vključili večje število filmov in dodali še kakšno funkcijo za dodatno analizo. Smiselno bi bilo analizirati kakšen vpliv na uspeh filma imajo igralci in režiserji in morda celo opis filma.