**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU**

**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I INFORMACIJSKIH TEHNOLOGIJA**

**Diplomski studij**

**VIZUALIZACIJA PODATAKA**

**Seminarski rad**

„Vizualizacija podataka o pojedinom filmu redatelja“

**Luka Budak**

**Osijek, 2020.**

SADRŽAJ

[1. UVOD 3](#_Toc45136060)

[1.1. Cilj projektnog zadatka 3](#_Toc45136061)

[2. KORIŠTENE TEHNOLOGIJE 4](#_Toc45136062)

[2.1. Javascript 4](#_Toc45136063)

[2.2. D3.js 4](#_Toc45136064)

[2.3. TopoJSON 5](#_Toc45136065)

[3. PROGRAMSKO RJEŠENJE 7](#_Toc45136066)

[3.1. Vizualizacija programskog rješenja 7](#_Toc45136067)

[4. ZAKLJUČAK 15](#_Toc45136068)

[5. POVEZNICE I LITERATURA 16](#_Toc45136069)

# **1**. UVOD

Današnje tržište je dinamično te ljudi gledaju filmove u kinima, kod kuće ili na mobilnim uređajima. Zbog sve većeg širenja filmske platforme u posljednjem desetljeću filmska industrija je ostvarila sve veći i veći rast prihoda, za primjer se može uzeti 2019. godina kada su ostvarili sveukupno 97 milijardi dolara. Velik razlog tome uz kina koja ostvaruju veći dio profita su i internetske platforme poput Netlixa i Amazon Prime-a koji omogućavaju gledanje sadržaja preko interneta na svome uređaju i to sve uz određenu pretplatu. Do prije je bilo poznato da se većinom američki filmovi nalaze na tržištu, no dolaskom novijih filmova širi se i scena iz ostalih država, pa tako i države poput Australije, Kine i slično prave svoje unikatne filmove koji osvajaju tržišta.

## 1.1. Cilj projektnog zadatka

Projektni zadatak obuhvaća skup podataka koji sadrži informacije o filmovima u pojedinim državama. Time će se vizualizirati vrijeme od 2005. do 2018. kada se primjetno može vidjeti porast broja filmova i ocjena na poznatim stranicama poput Rotten Tomatoes, IMDb ili Metacritic. Za vizualizaciju će se koristiti D3.js tehnologija koja omogućava da se korisnika približi i bolje vizualno osvijesti o porastu broja filmova i uspješnosti na tržištu.

# 2. IZRADA PROJEKTA

Projekt je pisan u JavaScript programskom jeziku uz pomoć d3.js biblioteke. Ova biblioteka ima više verzija, a korištena je verzija 4. Gumbovi, liste na kojima se može odabirati država i godina su pravljeni uz pomoć Bootstrap biblioteke. Izvorno je skinut csv koji je radi preglednosti i lakšeg korištenja prebačen u JSON oblik. Za vizualizaciju karte svijet korišten je topojson s kojim je moguća interakcija.

## 2.1. Json

JSON je lagani format za razmjenu podataka koji olakšava ljudima čitanje i pisanje. Ovaj tekstualni format je potpuno neovisan o jeziku, ali koristi konvencije poznate programerima C-obitelji.

Izgrađen je na dvije strukture:

1. Zbirka parova imena/vrijednosti. U različitim jezicima se to ostvaruje kao objekt, zapis, struktura, rječnik..
2. Uređeni popis vrijednosti koji se u većini jezika ostvaruje kao niz, vektor, popis ili niz

  {

    "Actors": "Kristen Stewart,Sarah Clarke,Matt Bushell,Billy Burke",

    "Awards": "29 wins & 15 nominations.",

    "BoxOffice": "191449475.0",

    "Country": "USA",

    "DVD": "21 Mar 2009",

    "Director": "Catherine Hardwicke",

    "Genre": "Drama,Fantasy,Romance",

    "IMdb\_ID": "tt1099212",

    "IMdb\_score": "52.0",

    "Language": "English",

    "Metascore": "56.0",

    "Plot": "Bella Swan moves to Forks and encounters Edward Cullen,a gorgeous boy with a secret.",

    "Production": "Summit Entertainment",

    "RT\_score": "49.0",

    "Rated": "PG-13",

    "Released": "21 Nov 2008",

    "Runtime": "122.0",

    "TMdb\_id": "[8966,17031,181898,17005,209861,38519,209860,260401,229876,52021,195781,46089]",

    "Title": "Twilight",

    "Type": "movie",

    "Website": "http://twilightthemovie.com/",

    "Year": "2008",

    "actor\_1": "Kristen Stewart",

    "actor\_2": "Sarah Clarke",

    "awards": "29.0",

    "bo\_year": "2008.0",

    "bo\_year\_rank": "13.0",

    "director\_1": "Catherine Hardwicke",

    "director\_2": "",

    "domestic-gross": "192800000.0",

    "domestic-pct": "49.0",

    "imdbVotes": "383325.0",

    "itunes\_id": "351883169",

    "logBoxOffice": "8.282054179746567",

    "mojo\_title": "Twilight",

    "nominations": "15.0",

"num\_actor": "4",

    "num\_director": "1",

    "num\_genre": "3",

    "num\_lang": "1.0",

    "oscar\_noms": "0.0",

    "oscar\_wins": "0.0",

    "overseas-gross": "200800000.0",

    "overseas-pct": "51.0",

    "search\_strs": "Twilight",

    "studio": "Sum.",

    "titles": "Twilight (2008)",

    "worldwide-gross": "393600000.0",

    "years": "2008"

  }

**Slika 2.1.:** Primjer JSON formata iz movies.jsona

Slika 2.1. prikazuje dio json-a koji je korišten za vizualizaciju ovog projekta. Jasno se vide parovi imena i vrijednosti, s obzirom na to da je godina u obliku stringa potrebnu ju je najprije pretvoriti u integer vrijednost i zatim koristiti kao na slici 2.2.

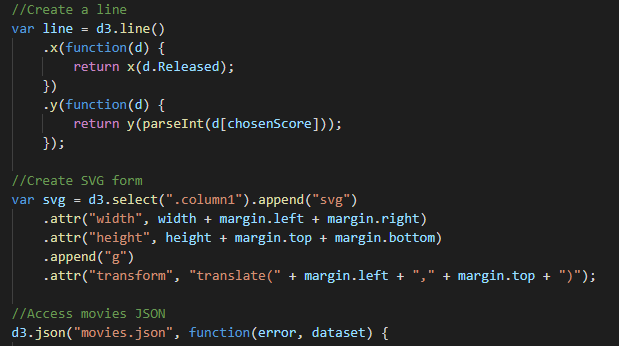


**Slika 2.2.:** Primjer JSON formata iz movies.jona

Uz to ova slika prikazuje i max funkciju koja služi za pronalazak najveće vrijednosti u dobivenom polju. U ovom slučaju radi se o polju IMDb rezultata koji se izvlače iz filmova za točno odabranu državu koji su prethodno u kodu filtrirani iz čitavog json-a.

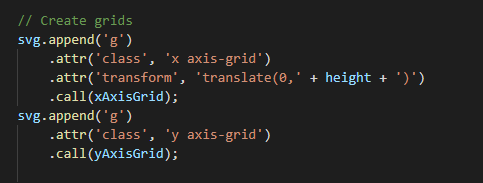
## 2.2. D3.js

D3 omogućuje povezivanje proizvoljnih podataka s objektnim modelom dokumenta, a zatim na tom dokumentu se primjenjuje određena transformacija dobivenih podataka. Zbog toga se najčešće koristi za generiranje HTML tablica iz niza brojeva ili koristi za izradu interaktivnog SVG grafikona s glatkim prijelazima i interakcijom.



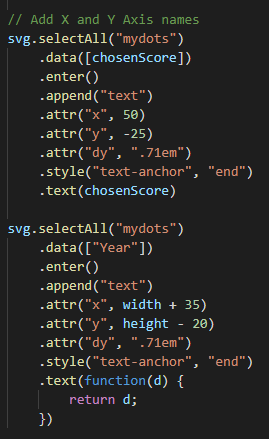
**Slika 2.3.**: Funkcija za kreiranje linije, dodavanje SVG prostora i pristup movies JSON-u

Na slici 2.3. prikazana je funkcija za kreiranje linije u koju se šalje film te se iz njega uzimaju podatci o datumu izlaska i odabranom rezultatu koji se želi prikazati. Slika 2.4. sadrži dodavanje g elementa koji će predstavljati grid linije na grafu. Ove linije služe za bolje očitavanje s grafa pa se tako vrijednosti ne moraju uvećavati već se iščitaju s osi.



**Slika 2.4.**: Dodavanje grid linija

Nakon što su se postavile X, Y osi i grid, dodaju se oznake na x i y os. Na x osi se nalazi godina tj. datum, a na Y rezultat ovisno o tome odabere li se IMDb, Metascore ili RTscore.



**Slika 2.5.**: Dodavanje oznaka na x i y os

Boje linija koje će prikazivati odabranu državu trebaju biti različite i zbog toga se koristi ordinalna skala sa schemom kategorije 10. Ova kategorija sadrži predefinirane boje koje će se samo uzeti u trenutku poziva tako da se napiše kao funkcija ime varijable i u zagradi broj.

var color = d3.scaleOrdinal(d3.schemeCategory10);

**Slika 2.6.**: Ordinalna skala s bojama

## 2.3. TopoJSON

TopoJSON ekstenzija se koristi za prikaz karte dobivene iz brojnih točaka i kordinata iz json-a. Nema ponavljanja dijelova pa se još učinkovitije učitava. S obzirom da se koristila karta svijeta potrebno je najprije učitavanje koje uzrokuje čekanje od pola sekunde i zatim ju je moguće koristiti. Slika 2.6. prikazuje učitavanja karte svijeta u obliku 1 : 50 milijuna. Uz to dostupni su i drugi omjeri i prikazi karte svijeta na stranici (https://github.com/topojson/world-atlas).



**Slika 2.7.**: Dodavanje karte svijeta

# 3. IZGLED KONAČNE APLIKACIJE

U nastavku je prikazana realizacija web aplikacije. Slika 2.8. prikazuje izgled glavne stranice na kojoj korisnik može odabrati usporedbu između dvije države. Ako odabere samo jednu prikazat će se pravilno naslov grafa i osi ovisno o odabiru. Zbog toga se mogu pratiti rezultati pojedine zemlje ili raditi usporedbe. S obzirom da za svaku godinu ima određen broj filmova, a neke zemlje nemaju više od 2 filma moguće je da će u nekim trenutcima pokazivati prazan graf, tj. graf bez linije prikaza no zemlje poput UK, USA, Kina, Australija imaju puno više filmova i zato će se za njih prikazati linija grafa. Uz to s desne strane grafa se nalaze podatci o usporedbi, pa tako iako nema nacrtane linije i dalje će prikazivati glavne podatke o toj zemlji.

Chart

Description automatically generated

**Slika 2.8.**: Grafički prikaz rezultata IMDb-a za UK

Prva odabrana zemlja će biti plave boje, a druga narančaste što se može vidjeti na slici 2.9. Rezultat na y-osi je prikazan od minimuma 0 do maximuma 100. Jedan od načina na koji se također može ovo prikazati je taj da se pronađu min i max vrijednost iz obje zemlje i time još zornije prikaže na y-osi.

A picture containing chart

Description automatically generated

**Slika 2.9.**: Grafički prikaz rezultata IMDb-a za usporedbu UK s Australijom

Nakon pregleda stanja filmova, najboljih rezultata za pojedinu državu, korisnik klikom na „World Map“ otvara drugu stranicu s prikazom karte svijeta (Slika 2.10.). Pravilan prikaz mape sadržava i legendu. Legenda za kartu svijeta se nalazi s desne strane i prikazuje koja nijansa boje označava koliko filmova je ta država izbacila filmova u zadanoj godini. Tamnije nijanse označavaju veće brojke, a svijetlije manje. Oznaka 0 do 3 označava brojeve 0, 1 i 2 ne uključujući 3.

Chart

Description automatically generated

**Slika 2.10.**: Prikaz druge stranice aplikacije

Ova karta omogućuje interakciju približavanja, zumiranja i odabira pojedine države. Klikom na pojedinu državu otvaraju se podatci o njoj i približava na karti kao na slici. 2.11.. Države poput Rusije imaju puno veće područje i moguće je da se dio preslikava i na lijevoj strani karte te se one udalje sa zumiranjem da se prikaže čitava država.

A picture containing map

Description automatically generated

**Slika 2.11.**: Prikaz odabrane države Kine i njenih podataka

Slika prikazuje Kinu koja se približila i obojala u tamniju nijansu narančaste, zatim su se prikazali podatci o filmskoj industriji u Kinu. Ti podatci obuhvaćaju ime države, broja filmova za zadanu godinu, broj direktora tijekom godine koji su sudjelovali u izradi filmova te najbolje ocjene na Imdb, Metacritic i Rotten Tomatoes. Na slici 2.12. nalazi se izbornik sa godinama koje je moguće izabrati. Klikom na pojedinu se neće obrisati karta već će se samo zamijeniti nijanse boja po karti što ubrzava sam proces i smanjuje potrebno vrijeme čekanja. Na sličan način radi i animacija koja se pokreće klikom na „Start Animation“ gumb i zaustavlja klikom na „Stop“. Animacija će prikazati promjenu stanja karte od 2005. godine do 2018. Moguće je u svakom trenutku zaustaviti animaciju pri čemu će se resetirati karta na onu prethodnu koju je sam korisnik odabrao. S kartom se može interaktirati i tijekom prikaza animacije, no pri prikazu svake godine će se resetirati karta na sredinu i maknuti podatci o odabranoj državi.

Map

Description automatically generated

**Slika 2.12.:** Izbornik sa mogućim godinama za prikaz

# 4. ZAKLJUČAK

Porast filmske industrije ne mora, ali i može značiti da se kvaliteta filmova poboljšala, no glavni razlog je sve veća digitalizacija sustava zbog koje ljudi provode više vremena na internetu gledajući filmove i opuštajući se kod kuće ili u kinu.

Projekt je moguće unaprijediti još boljom optimizacijom i podrškom za mobitele što bi se riješilo naprednijim uređivanjem elemenata i njihov prikaz na ekranu ovisno o samoj veličini ekrana. Kako bi se još bolje vizualizirala promjena među državama i njihovi podatci bilo bi potrebno proširiti skup podataka za pojedine godine što uključuje spajanje sa drugim skupovima podataka i filtriranje duplikata. Za lakše razumijevanje projekta dodani su komentari.

# 5. POVEZNICE I LITERATURA

1. Priručnik za laboratorijske vježbe iz Vizualizacije podataka
2. JavaScript, https://hr.wikipedia.org/wiki/JavaScript
3. JSON, <https://www.json.org/json-en.html>
4. TopoJSON karte svijeta, https://github.com/topojson/world-atlas
5. Skup podataka filmova, https://github.com/zacharyang/movies-project/blob/master/data/clean/OMdb\_mojo\_clean.csv
6. Linijski grafovi, <https://www.d3-graph-gallery.com/index.html>

**RJEŠENJE**

GitHub link:

<https://github.com/lbudak/MovieDirectorRatings>

Web aplikacija:  
<https://movie-director-ratings.herokuapp.com/>