

# Slovenská technická univerzita

Fakulta informatiky a informačných technológií

Ilkovičova 2, 812 19 Bratislava

---

## **1. zadanie**

### **Vizualizácia dát**

Cvičiaci:	Ing. Patrik Polatsek
Cvičenie:	štvrtok 17:00 – 18:40
Akademický rok:	2017/2018
Vypracovala:	Hana Kuntová

## 1. Úvod

Na prvé zadanie som si zvolila vizualizovať zisky pre jednotlivé filmy. Ako dataset som využila dáta vo formáte JSON. Na obrázku je možné vidieť časť dát, pričom časti *Worldwide\_Gross* a *Production\_Budget* boli využité pri vizualizácii. Pre získanie zisku pre film som si odpočítala hodnotu *Production\_Budget* od *Worldwide\_Gross*. Pre lepšie zobrazenie ziskových a stratových filmov som využila zelenú a červenú farbu. Kedy zelená farba predstavuje ziskový film a červená farba na grafe stratový film.

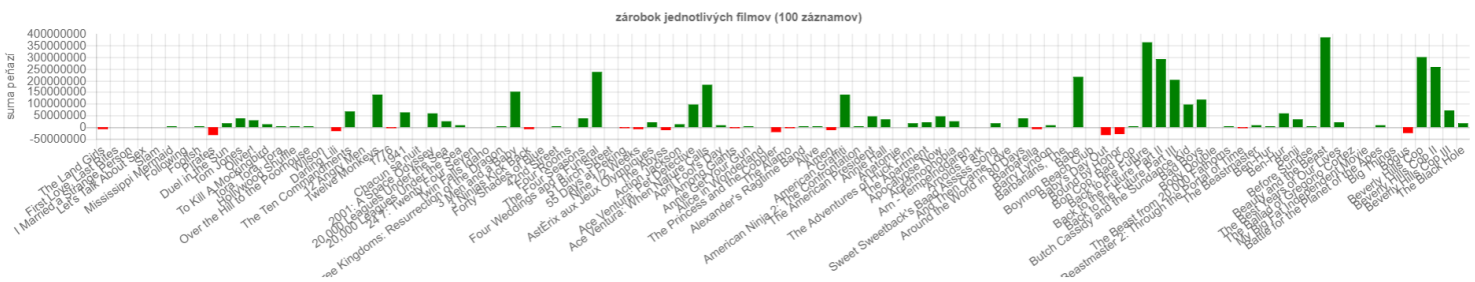
Na zobrazovanie dát som využívala knižnicu *chart.js*, programovací jazyk som použila HTML javascript, pričom ako server som využila *node.js*.

```
1 [
2   {"Title": "The Land Girls", "US_Gross": 146083, "Worldwide_Gross": 146083, "US_DVD_Sales": null, "Production_Budget": 8000000, "Release_Date": "2001-01-01"},
3   {"Title": "First Love, Last Rites", "US_Gross": 10876, "Worldwide_Gross": 10876, "US_DVD_Sales": null, "Production_Budget": 3000, "Release_Date": "2001-01-01"},
4   {"Title": "I Married a Strange Person", "US_Gross": 203134, "Worldwide_Gross": 203134, "US_DVD_Sales": null, "Production_Budget": 3000, "Release_Date": "2001-01-01"},
5   {"Title": "Let's Talk About Sex", "US_Gross": 373615, "Worldwide_Gross": 373615, "US_DVD_Sales": null, "Production_Budget": 3000, "Release_Date": "2001-01-01"},
6   {"Title": "Slam", "US_Gross": 1009819, "Worldwide_Gross": 1087521, "US_DVD_Sales": null, "Production_Budget": 1000000, "Release_Date": "2001-01-01"},
7   {"Title": "Mississippi Mermaid", "US_Gross": 24551, "Worldwide_Gross": 2624551, "US_DVD_Sales": null, "Production_Budget": 16000, "Release_Date": "2001-01-01"},
8   {"Title": "Following", "US_Gross": 44705, "Worldwide_Gross": 44705, "US_DVD_Sales": null, "Production_Budget": 6000, "Release_Date": "2001-01-01"},
9   {"Title": "Foolish", "US_Gross": 6026908, "Worldwide_Gross": 6026908, "US_DVD_Sales": null, "Production_Budget": 1600000, "Release_Date": "2001-01-01"}
10 ]
```

Obrázok 1 príklad dát

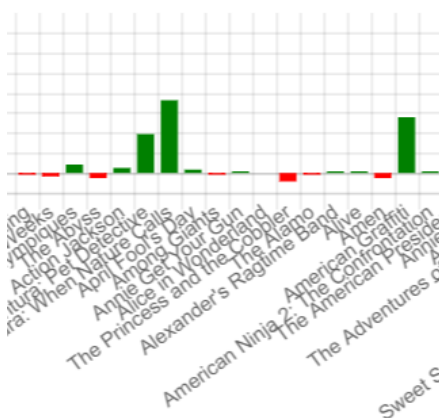
## 2. Zobrazenie dát

Pri zobrazovaní dát, som sa zamerala na trojicu rôzne veľkých záznamov. Na prvom obrázku možno vidieť zobrazenie pre malé dáta (100 záznamov). Jednotlivé dáta sú dobre viditeľné, je možné jasne vidieť, ktoré filmy mali zisk a ktoré naopak boli stratové. Aj keď názvy filmov sa v takomto celistvom zobrazení môžu strácať.



Obrázok 2 vizualizácia malých dát – celkový pohľad

zárobok jednotlivých filmov (100 záznamov)



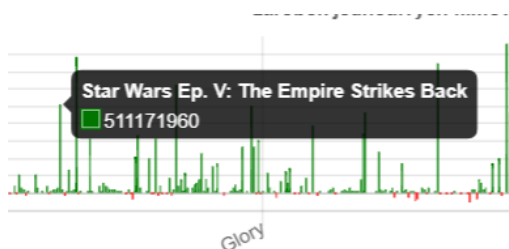
Obrázok 3 vizualizácia malých dát - čiastočný pohľad

Pre ďalšiu vizualizáciu som si zvolila stredne veľké dáta s počtom 1000 záznamov (obrázok 3). Grafová knižnica chart.js ponúka možnosť zobrazenia každého n-tého názvu na x osi. Pre toto zobrazenie som sa rozhodla, kvôli vizualizácii veľkých dát (obrázok 5), kedy nebolo možné prečítať jednotlivé názvy filmov. Dáta sú stále viditeľné, aj keď už nie je možné presne vyčítať zárobok jednotlivých filmov.



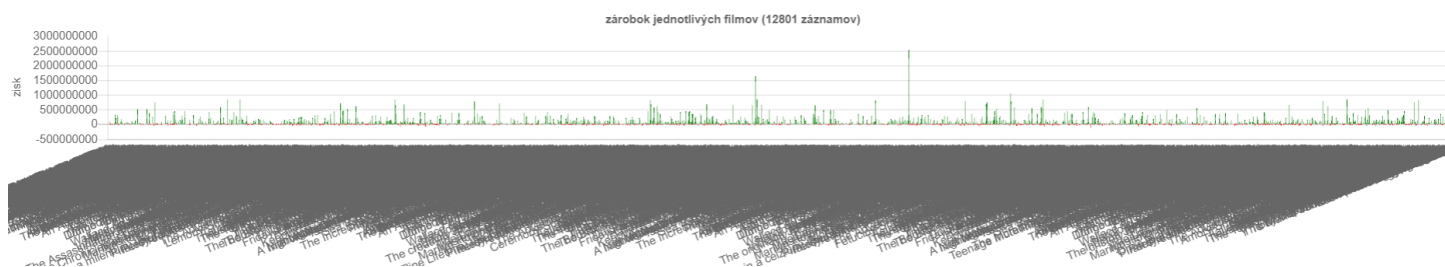
Obrázok 4 vizualizácia stredných dát - celkový pohľad

Ďalším z dôvodov pre zobrazenie názvu každého n-teho filmu je aj, že knižnica ponúka možnosť si zobrazit' informácie o jednotlivom filme priamo na grafe (obrázok 4). Pre stredné dáta je stále ešte aj táto možnosť pomerne prehľadná.



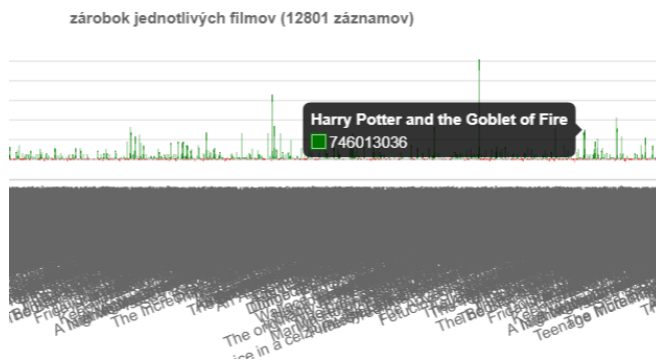
Obrázok 5 vizualizácia stredných dát - zisk pre film

V poslednom rade som si zvolila dáta o veľkosti 12800 záznamov. Ako je možné vidieť na obrázku nižšie, tak názvy filmov na x-vej osi nie sú vôbec viditeľné. Menej zárobokové filmy nie je možné určiť a strácajú sa medzi sebou jednotlivé filmy.



Obrázok 6 vizualizácia veľkých dát - celkový pohľad

Tiež vyššie spomenutá možnosť, pre zobrazenie si konkrétnych informácií o filme, nie je vyhovujúca. Keďže jednotlivé dáta sú veľmi blízko seba, je ťažké si zvoliť konkrétne dáta.



Obrázok 7 vizualizácia veľkých dát - zisk pre film

### 3. Záver

Stĺpcový graf knižnice chart.js nie je vhodný pre zobrazovanie veľkých dát. Ako už bolo spomenuté, jednotlivé názvy filmov nie sú viditeľné a zároveň ani dáta nie je presne určiť. Pomocou vizualizácie veľkých dát sa mohlo zistiť, že pokiaľ máte záujem sa uberať filmovým svetom, tak je šanca, že film bude ziskový.

Naopak vizualizácia malých a prípadne stredných dát môže slúžiť na porovnanie ziskov jednotlivých filmov.