

Projekt 2

Analiza tweetów

Sandra Leman

Czerwiec 2022

Spis treści

1	Wstęp	2
2	Zliczanie słów i chmury tagów	2
2.1	Disney	2
2.2	Dreamworks	3
2.3	Pixar	4
2.4	Porównanie	5
3	Analiza opinii	5
3.1	Disney	7
3.2	Dreamworks	8
3.3	Pixar	9
4	Analiza fragmentów czasowych	10
4.1	Disney	10
4.2	Dreamworks	12
4.3	Pixar	14

1 Wstęp

W projekcie zostały omówione tweety z hashtagami:

- Disney
- Dreamworks
- Pixar

W pierwszej części brane pod uwagę jest 10 000 najnowszych tweetów z dnia 1 czerwca 2022. Następnie przedstawiona jest analiza czasowa miesięczna z ostatnich 5 miesięcy. Zostało pobrane co najwyżej 1000 tweetów dzienne z okresu styczeń - maj.

2 Zliczanie słów i chmury tagów

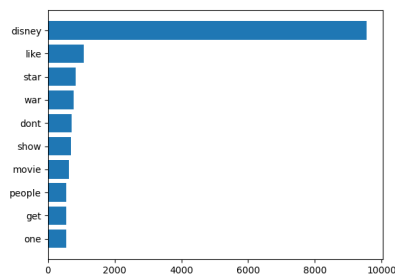
Dla każdego z hashtagów zostały zliczone wszystkie słowa oraz słowa bez powtórzeń dla wersji po:

- tokenizacji
- filtracji
- filtracji i stemmingu
- filtracji i lematyzacji

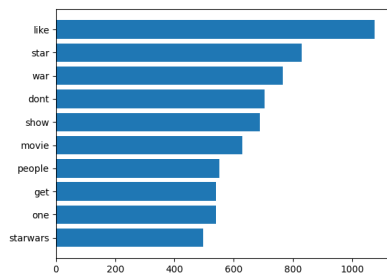
Zostały również przedstawione wykresy słupkowe oraz chmury tagów dla wersji po lematyzacji.

2.1 Disney

Disney	Liczba słów	Liczba słów po usunięciu powtórzeń
Tokenized	222 359	26 964
Filtered	136 750	26 827
Stemmed	136 750	22 712
Lemmatized	136 750	25 325

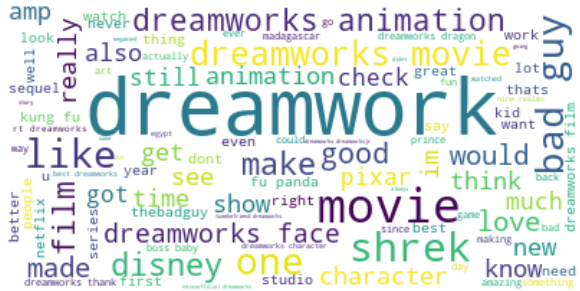
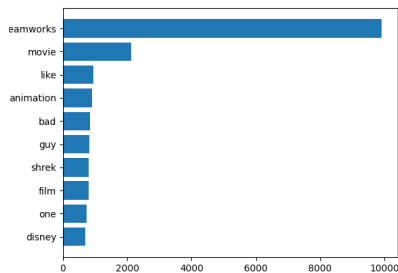


Można zauważyć, że sporą przewagę nad innymi słowami ma samo słowo disney. Żeby dostrzec większą różnicę w pozostałych słowach można dodać hashtag to słów filtrowanych. Wykresy stają się wówczas bardziej przejrzyste.

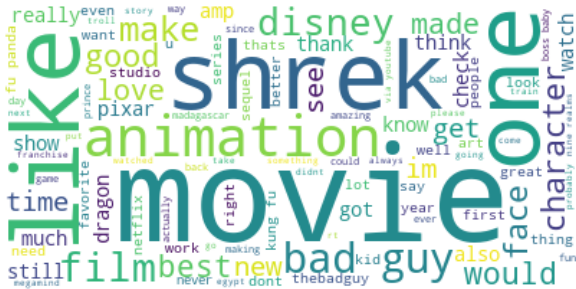
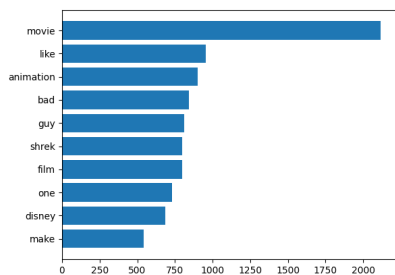


2.2 Dreamworks

Dreamworks	Liczba słów	Liczba słów po usunięciu powtórzeń
Tokenized	187 795	24 195
Filtered	121 523	24 062
Stemmed	121 523	21 108
Lemmatized	121 523	22 955

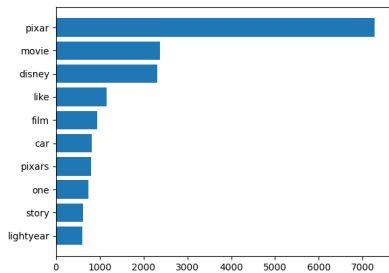


Można zauważyć, że sporą przewagę nad innymi słowami ma samo słowo `dreamworks`. Żeby dostrzec większą różnicę w pozostałych słowach można dodać hashtag to słów filtrowanych. Wykresy stają się wówczas bardziej przejrzyste.



2.3 Pixar

Pixar	Liczba słów	Liczba słów po usunięciu powtórzeń
Tokenized	190 651	25 645
Filtered	122 377	25 509
Stemmed	122 377	22 186
Lemmatized	122 377	24 214



Można zauważyć, że sporą przewagę nad innymi słowami ma samo słowo `pixar`. Żeby dostrzec większą różnicę w pozostałych słowach można dodać hashtag to słów filtrowanych. Wykresy stają się wówczas bardziej przejrzyste.

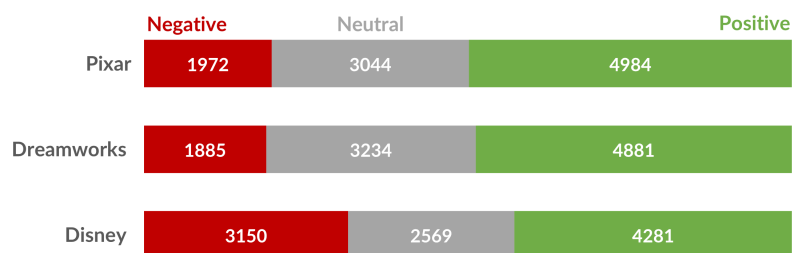
Sentiment	Disney	Dreamworks	Pixar
Negative	0.131	0.074	0.079
Neutral	0.722	0.737	0.749
Positive	0.146	0.189	0.172
Compound	1	1	1

Ocena sentymentu wszystkich tweetów



Liczba tweetów	Disney	Dreamworks	Pixar
Negative	3150	1885	1972
Neutral	2569	3234	3044
Positive	4281	4881	4984

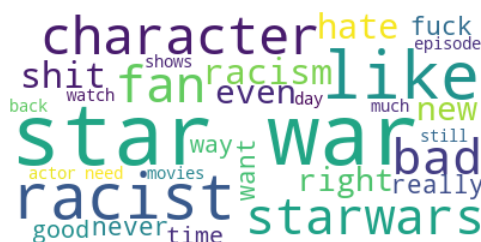
Ilość tweetów według sentymentu



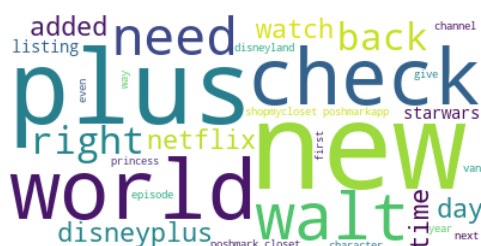
Możemy zauważyć, że liczba tweetów negatywnych jest względnie proporcjonalna do oceny sentymentu, zaś przy tweetach pozytywnych ocena sentymentu jest największa dla hashtagu Dreamworks, a liczba tweetów jest największa dla hashtagu Pixar. Różnica może wynikać ze sposobu liczenia ocen przez narzędzie Vader.

Hashtagi Dreamwork i Pixar mają zbliżone oceny i ilości tweetów w danych kategoriach, podczas gdy hashtag Disney ma widocznie więcej tweetów negatywnych oraz mniej pozytywnych i neutralnych. Jednak w każdym przypadku przeważają tweety pozytywne.

Negative



Neutral



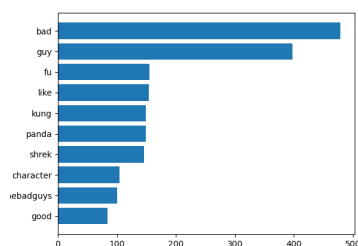
Positive



Dużą część pozytywnych tweetów stanowią słowa typu " like ", " love ", czy " good ". Z ciekawszej strony możemy zauważyć tematykę Pride Month.

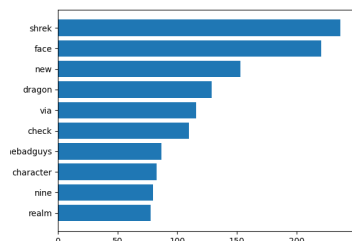
3.2 Dreamworks

Negative



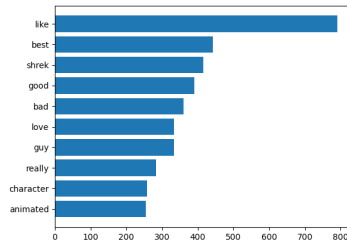
Wśród negatywnych tweetów możemy zauważyć sporo znanych tytułów takich jak Shrek, Kung Fu Panda, czy najnowsze The Bad Guys, gdzie to ostatnie mogło być omylnie uznane za negatywne ze względu na specyficzny tytuł.

Neutral



Neutralne tweety zdominowała tematyka Shreka oraz tzw. Dreamworks Face - charakterystyczny wyraz twarzy, który robi większość postaci Dreamworksa. Możemy też zauważyć wyrazy związane z serią DreamWorks Dragons: The Nine Realms.

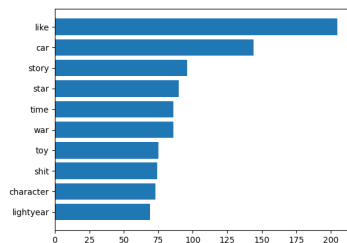
Positive



Jeśli chodzi o tweety pozytywne to sytuacja wygląda analogicznie jak w przypadku tweetów negatywnych - dominują słowa o wydźwięku pozytywnym oraz tytuły znanych produkcji Dreamworksa.

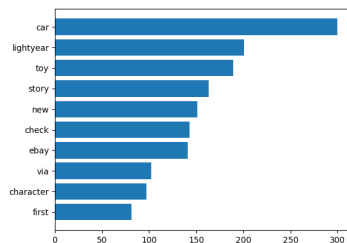
3.3 Pixar

Negative



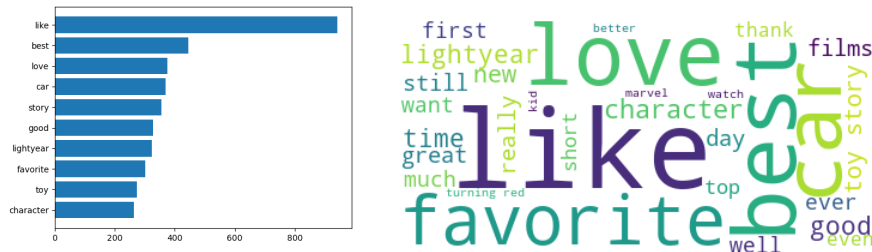
Sytuacja podobna jak w poprzednim przykładzie, dominują słowa o wydźwięku negatywnym oraz tytuły znanych produkcji Pixara, takich jak *Auta* czy *Toy Story*. Możemy też zauważyć małą ilość najczęstszych wyrazów.

Neutral



W neutralnych tweetach również widzimy tytuły różnych produkcji razem z najnowszym Lightyear. Widzimy też słowa związane z zakupami.

Positive



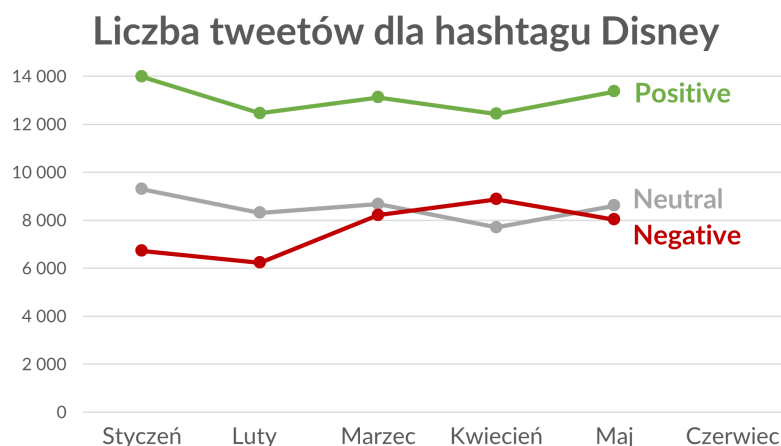
W tym przypadku również sytuacja typowa - dominują słowa o wydźwięku pozytywnym oraz tytuły znanych produkcji Pixara.

4 Analiza fragmentów czasowych

Podobnie jak w dziale poprzednim zliczamy tweety o wydźwięku negatywnym, neutralnym i pozytywnym z każdego miesiąca od stycznia do maja dla każdego z tagów.

4.1 Disney

Disney	Styczeń	Luty	Marzec	Kwiecień	Maj
Negative	6 720	6 230	8 217	8 870	8 030
Neutral	9 299	8 308	8 666	7 698	8 607
Positive	13 981	12 462	13 117	12 432	13 363



Możemy dostrzec największe wahania na przestrzeni marca i kwietnia, gdy ilość negatywnych tweetów znacząco wzrosła kosztem tweetów neutralnych. Chmury tagów z tego okresu wyglądają następująco:

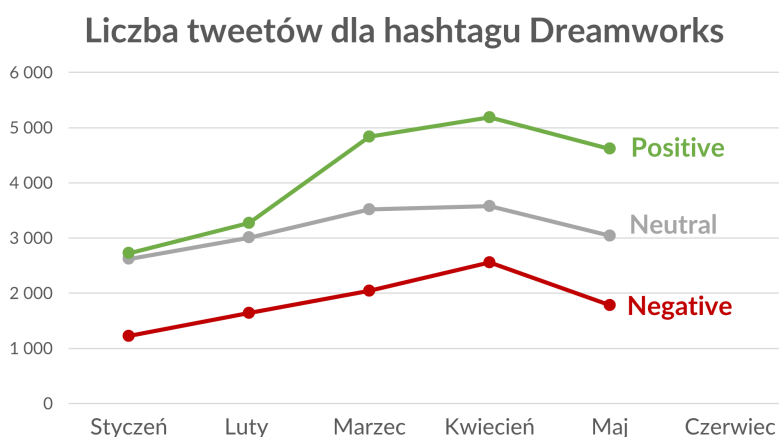


Znalezienie przyczyn wahań nie było łatwe, ponieważ nie dotyczyły one produkcji Disneya ani samego Disney World we Florydzie. Początek miały za to w polityce, gdy gubernator Florydy Ron DeSantis podpisał kontrowersyjny projekt ustawy "Parental Rights in Education", który przeciwnicy nazywają "Don't Say Gay". The Walt Disney Company otwarcie skrytykowała projekt, co skutkowało agresywną reakcją ze strony DeSantisa, która może doprowadzić do znacznych zmian w zobowiązaniach podatkowych Disneya. Fragment artykułu na ten temat:

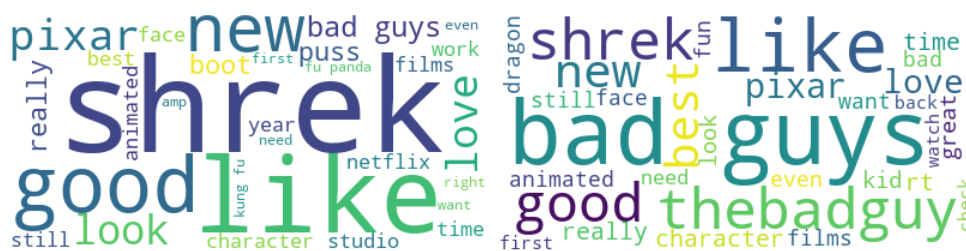
A handful of high-profile Republicans have expressed their unease with the [punitive actions](#) [DeSantis took](#) against the Walt Disney Company after its CEO [spoke out](#) against a new Florida measure restricting certain classroom instruction about sexual orientation and gender identity. A fellow Republican governor, an ally of former President Donald Trump and the single largest donor to DeSantis' reelection campaign are among those now questioning DeSantis for retaliating against a private enterprise.

4.2 Dreamworks

Dreamworks	Styczeń	Luty	Marzec	Kwiecień	Maj
Negative	1 222	1 638	2 040	2 554	1 780
Neutral	2 619	3 005	3 516	3 575	3 041
Positive	2 722	3 271	4 834	5 183	4 616



Na przestrzeni marca i kwietnia widzimy stopniowy wzrost liczby pozytywnych tweetów. W kwietniu możemy też dostrzec wzrost tweetów negatywnych. Chmury tagów z tych miesięcy prezentują się następująco:



Jeśli chodzi o marzec zdecydowanie najpopularniejszym tematem by Shrek - kultowy film Dreamworksa. Mamy informacje na temat nowego filmu z tego uniwersum, który nie jest związany z samym Shrekiem. Jest za to sequelem spin-offu "Kot w butach" z 2011. Z pewnością miało to wpływ na popularność tematu.



Innym aspektem wartym uwagi może być popularność serii na serwisie Netflix. Po tym gdy film pojawił się na serwisie długo utrzymywał się w top 10 najchętniej oglądanych filmów danego dnia. Fragment artykułu na ten temat:

At the start of March, *Shrek* and *Shrek 2* made their way to [Netflix](#) and instantly became [some of the most popular movies on the entire streaming service](#). One of the most beloved animated films in history, *Shrek* has been a force for Netflix, as has its sequel, consistent staples of the daily [Netflix Top 10](#) list. Subscribers are loving being able to watch the franchise's early entries on Netflix, and now they'll have the chance to check out even more. Additional films from the *Shrek* franchise are making their way to Netflix in April.

Fani musicali z pewnością nie pominęli również faktu, że w marcu wróciła musicalowa wersja tego kultowego filmu.

SHREK THE MUSICAL

The Green Ogre With The Big Heart Is Back! St Agnes Choral Society presents Shrek the Musical.

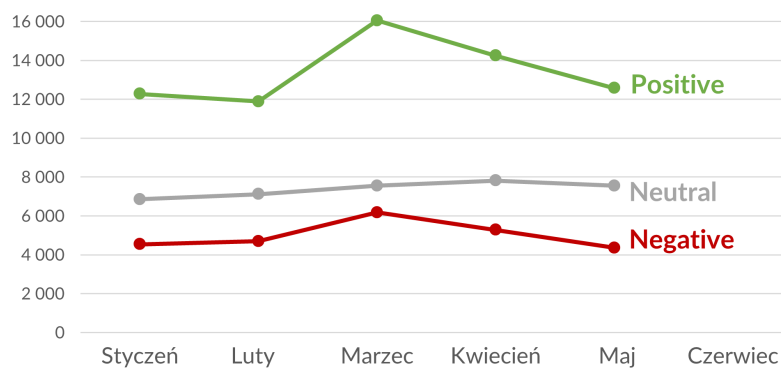
Jeśli chodzi o kwiecień uwagę przykuł zupełnie nowy film Dreamworksa - The Bad Guys. Ze względu na swój specyficzny tytuł mógł on też być powodem wzrostu tweetów uznanych za negatywne.



4.3 Pixar

Pixar	Styczeń	Luty	Marzec	Kwiecień	Maj
Negative	4 541	4 695	6 181	5 274	4 354
Neutral	6 852	7 117	7 554	7 818	7 552
Positive	12 274	11 881	16 045	14 241	12 570

Liczba tweetów dla hashtagu Pixar



W tym przypadku od razu rzuca nam się w oczy marzec, który przyniósł nagły wzrost pozytywnych tweetów wraz z lekkim wzrostem negatywnych. Chmura tagów wygląda następująco:



Powód wahań jest tym razem oczywisty - nowy film, który mimo niskich oczekiwań spotkał się z bardzo dobrym odzewem. Co ciekawe "Turning Red" jest pierwszym filmem Pixara wyreżyserowanym wyłącznie przez kobietę, pierwszym którego akcja rozgrywa się w Kanadzie oraz drugim z azjatyckim głównym bohaterem po " Odłot ".

