Analiza skryptu musicalu „Hamilton” – ocena emocji – Człowiek vs Maszyna

1. **Wstęp**

Mój projekt polega na sprawdzeniu czy program, którego zadaniem jest analiza tekstu, poprawnie (tj. zgodnie z ludzką oceną/analizą) zidentyfikuje emocje danych postaci. Jako bazę tekstową do analizy użyłem skrypt musicalu „Hamilton” z 2015r. Jest to o tyle ciekawa baza, iż zawiera ona dużo słów staroświeckich, czy też skrótowców, których zakładam, że gotowe paczki nie będą w stanie przeanalizować, przez co wyniki w niektórych przypadkach mogą się znacząco różnić od tych zakładanych.

1. **Przygotowanie tekstu**

Na początku musiałem odpowiednio przygotować tekst. Skrypt w pliku .pdf przerzuciłem do pliku .txt, po czym wczytałem go, podzieliłem na osobne słowa, przy okazji wyrzucając niepotrzebne słowa (takie jak zaimki osobowe, znaki interpunkcyjne).

Text

Description automatically generated with medium confidence

Następnie wyciągnąłem z tekstu imiona (tj. te słowa napisane wielką literą), odpowiednio je podzieliłem i za ich pomocą przypisałem odpowiednie słowa wypowiadającym je osobom.

Text

Description automatically generated

Text

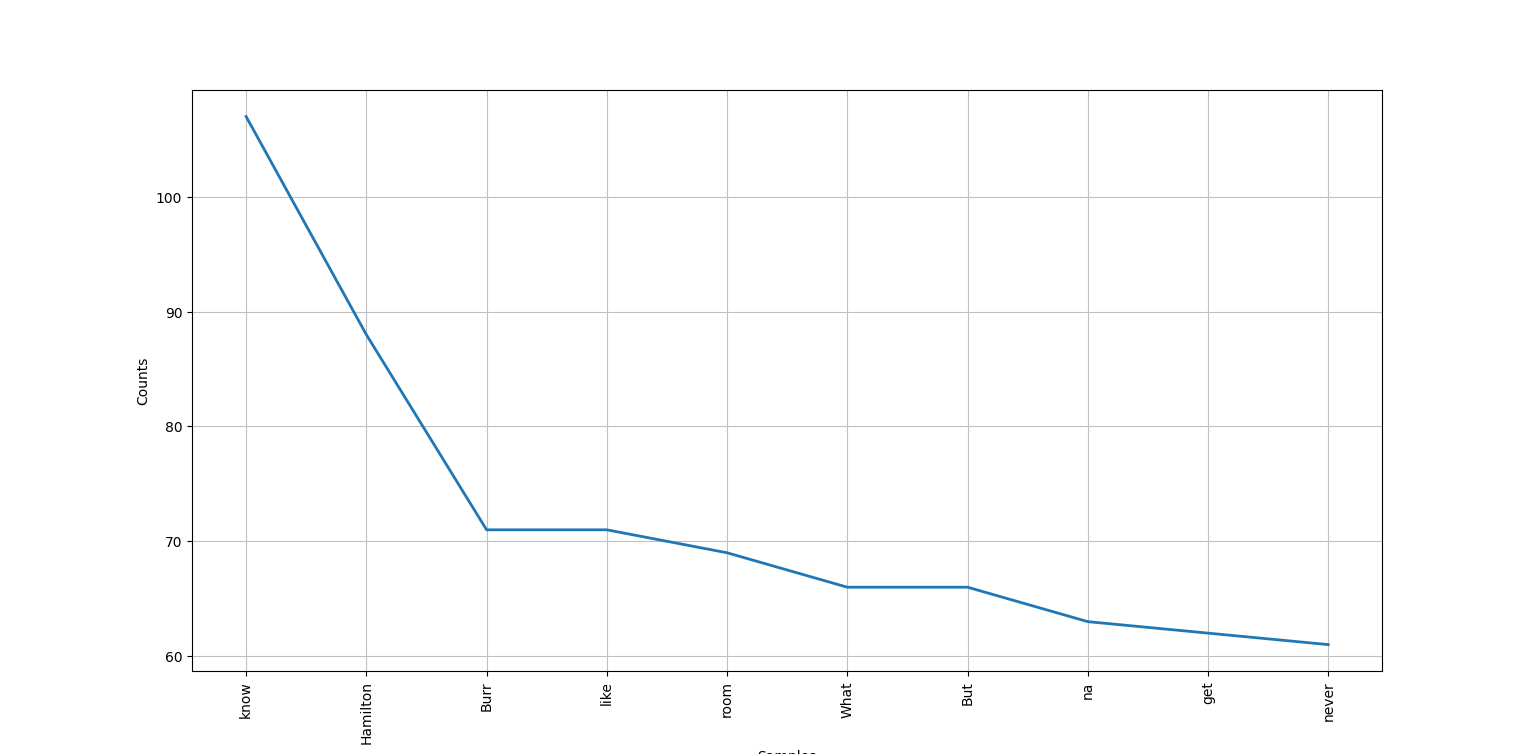
Description automatically generated

Text

Description automatically generated with medium confidence

1. **Najczęściej występujące słowa**

Poniżej przedstawionych jest 10 najpopularniejszych słów w skrypcie. Nie ma w sumie tu nic zaskakującego, najczęściej wymawiane są imiona dwóch głównych postaci oraz typowe czasowniki.



1. **Teza odnośnie emocji danych bohaterów**
2. **Sprawdzenie tezy**
3. **Podsumowanie**