

POLITECHNIKA WROCŁAWSKA
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI



Analiza sieci społecznej polskich wykonawców muzycznych

Sprawozdanie z laboratorium

AUTOR

Aleksandra Walczybok

nr albumu: **272454**

kierunek: **Inżynieria systemów**

6 styczeń 2025

1 Wstęp – sformułowanie problemu

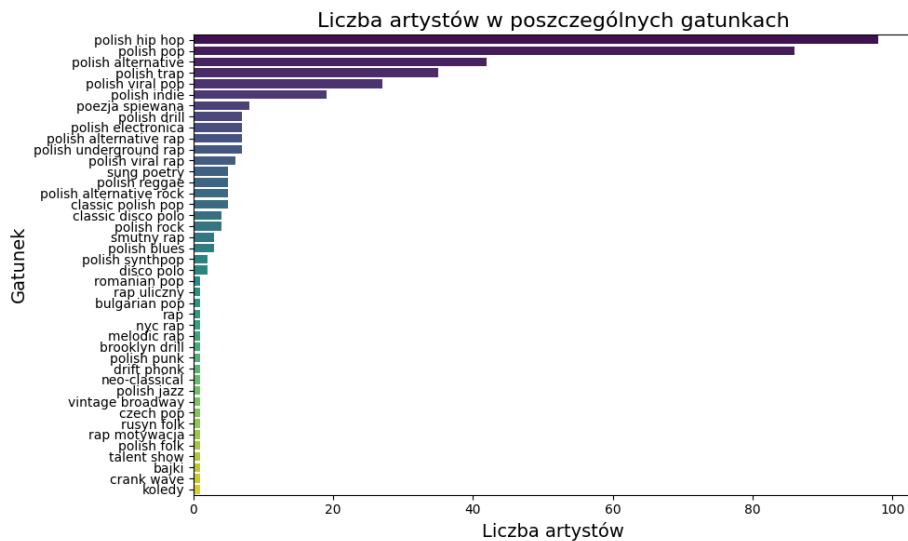
Analizowanym zagadnieniem jest sieć polskich wykonawców muzycznych. Projekt ma na celu zgłębienie współpracy między artystami, powiązań związanych z ich popularnością oraz preferowanymi gatunkami muzycznymi przez Polaków. Głównym celem jest identyfikacja kluczowych postaci na polskiej scenie muzycznej – artystów, którzy odgrywają istotną rolę w kształtowaniu jej charakteru oraz pełnią funkcję łączników między różnymi grupami muzycznymi. Dodatkowo dzięki projektowi będzie można ocenić aktywność działalności i współpracy polskich artystów, co może pomóc w przewidywaniu nowych trendów czy nowych duetów.

2 Opis danych

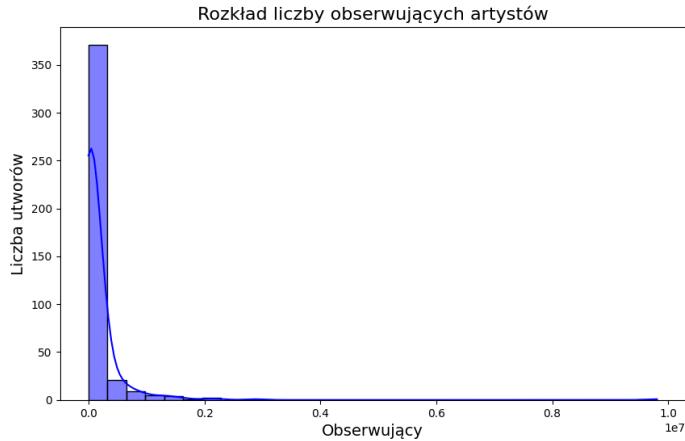
Dane użyte do przeprowadzenia analizy pochodzą z platformy Spotify. Zostały one pobrane poprzez dostępne API tej aplikacji oraz specjalną bibliotekę w Pythonie do wydobywania danych z API Spotify - *spotipy*. Znaleziono 12 playlist złożonych z najpopularniejszych piosenek polskich z lat 2023-2024. (<https://github.com/aleksandra0014/complex-networks/blob/main/data/playlists.txt>).

Playlisty dobrane były w taki sposób, aby dane zawierały różne gatunki muzyczne od hip hopu, poprzez rock, do disco polo. Liczba pobranych utworów to 636, artystów 415.

Z playlist pobrano utwory oraz ich wykonawców, popularność utworu, liczbę obserwujących dla danego artysty oraz gatunki przypisane do każdego artysty. Dane te umieszczone w plikach csv. Dodatkowo stworzono również pliki txt, gdzie zapisywano w odrębnych liniach artystę i wykonywany utwór. Pliki te były potrzebne do stworzenia w dalszym etapie grafu.



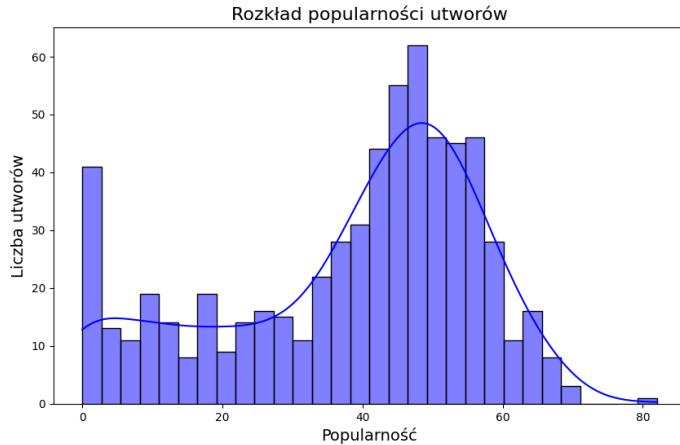
Rysunek 1: Zróżnicowanie gatunków w danych



Rysunek 2: Ilość obserwujących

Tabela 1: Zestawienie artystów, liczby obserwujących i gatunków muzycznych

Artysta	Obserwujący	Gatunek
Lil Tjay	9 806 789	brooklyn drill, melodic rap, nyc rap
sanah	2 881 017	polish pop
Dawid Podsiadło	2 219 112	polish pop, talent show
Mata	2 071 000	polish hip hop, polish trap
Taco Hemingway	1 657 059	polish hip hop, polish trap
Bedoes 2115	1 594 860	polish hip hop, polish trap
Kizo	1 398 440	polish hip hop, polish trap
White 2115	1 360 716	polish hip hop, polish trap
Quebonafide	1 331 902	polish hip hop
Sobel	1 250 738	polish hip hop, polish trap



Rysunek 3: Zróżnicowanie popularności piosenek w danych

Z powyższych wykresów możemy wnioskować, że zbiór danych zawiera szeroki zakres gatunków muzycznych. Warto spojrzeć, że najczęstszym gatunkiem jest hip-hop, co może sugerować aktualne trendy w polskiej muzyce.

W kontekście zróżnicowania ilości obserwujących dla artystów widzimy, że zdecydowana większość ma niedużą liczbę obserwujących. Największą ma Lil Tjay, gdyż jest amerykańskim wykonawcą, który nagrał wspólną piosenkę z jednym z polskich artystów. Reszta wykonawców będących w czołówce pod względem liczby obserwujących ma dość podobne wyniki.

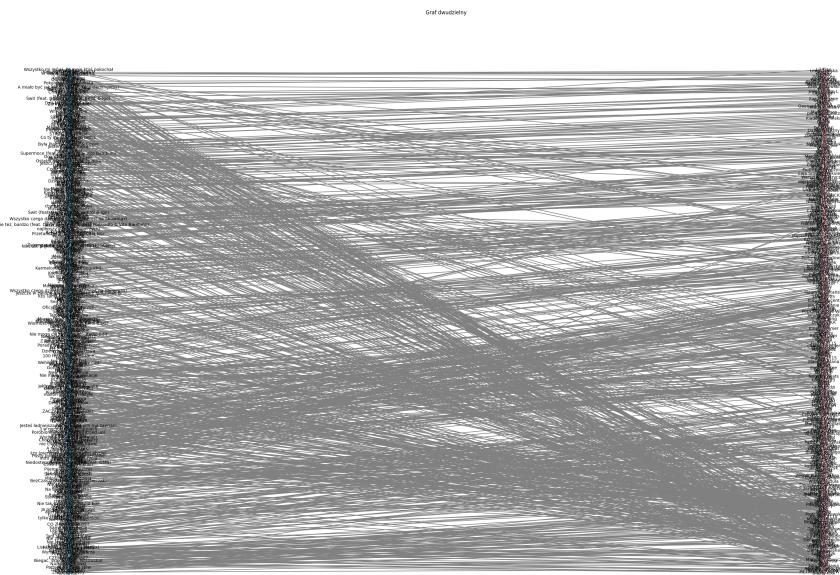
Jeśli chodzi o popularność utworów, jest ona mocno zróżnicowana. W Spotify zakres popularności mierzony jest od 0 do 100, a więc trzeba przyznać, że żaden z utworów w naszej bazie nie jest bliski maksymalnej popularności. Jest to jednak logiczne, gdyż zakres obejmuje także kultowe międzynarodowe hity, a więc ciężko, aby polskie utwory mogły być równie tak samo ocenione.

3 Opis rozwiązania i wyniki

Aby przeprowadzić analizę, stworzono szereg grafów oraz ich projekcji. Obliczono również najważniejsze własności grafów oraz najistotniejszych wierzchołków. Poszczególne implementacje zostały pokazane poniżej.

3.1 Graf dwudzielny

Pierwszą reprezentacją danych jest graf dwudzielny, gdzie wierzchołkami są zarówno artyści, jak i utwory. Artyści połączeni są z utworami, jeśli są jednymi z wykonawców danego utworu. Mamy zatem tzw. węzły dolne - artystów (kolor różowy) i górne - utwory (kolor niebieski).



Rysunek 4: Graf dwudzielny

Stopień grafu: 1015

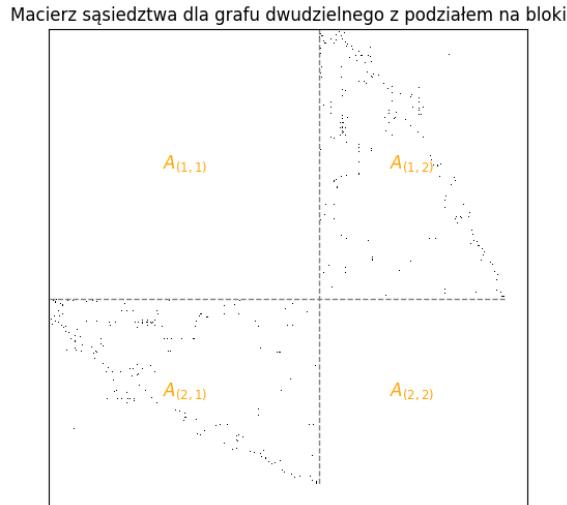
Rozmiar grafu: 1050

Tabela 2: 10 utworów o najwyższym stopniu

Utwór	Stopień
3 w nocy w Międzyzdrojach	10
Bungee	9
A nie pamiętasz jak?	8
PG\$	7
CO Z NAMI BĘDZIE?	6
club2020	6
Pieniądze i sława	6
Landlord	6
Dalej	5
LECE BO CHCE	5

Tabela 3: 10 artystów o najwyższym stopniu

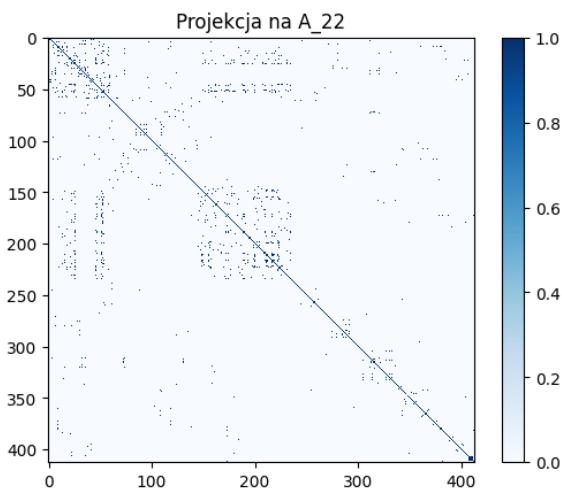
Artysta	Stopień
sanah	30
Oki	27
Dawid Podsiadło	21
Daria Zawiałow	21
Mata	18
Sylwia Grzeszczak	17
Bedoies 2115	16
White 2115	15
bambi	13
francis	13



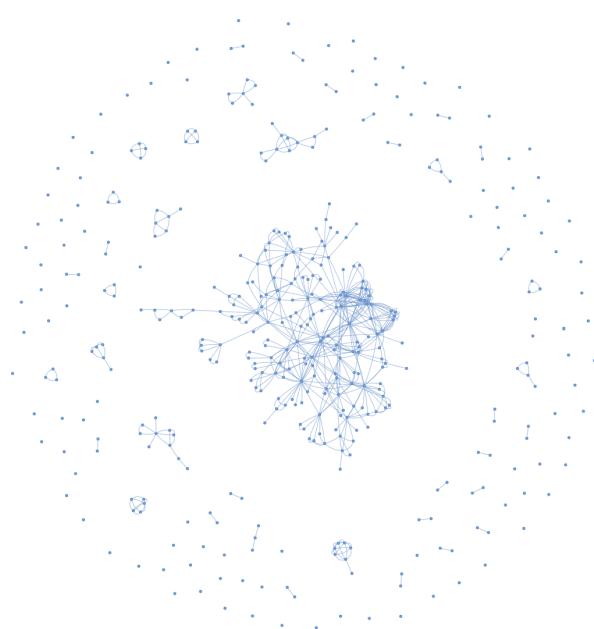
Rysunek 5: Macierz sąsiedztwa

3.2 Projekcja - łączenie artystów

Pierwszą projekcją grafu dwudzielnego będzie projekcja prosta na wierzchołki dolne - artystów. Dwoje artystów jest ze sobą połączonych, jeśli nagrali wspólną piosenkę. Taki graf pozwoli nam zidentyfikować grupy współpracy między artystami oraz ocenić najważniejszych wykonawców, którzy są pośrednikami między wieloma grupami, czy też mają kontakt z największą ilością artystów.



Rysunek 6: Projekcja grafu dwudzielnego na wierzchołki dolne



Rysunek 7: Graf artystów - pyviz

3.2.1 Własności grafu

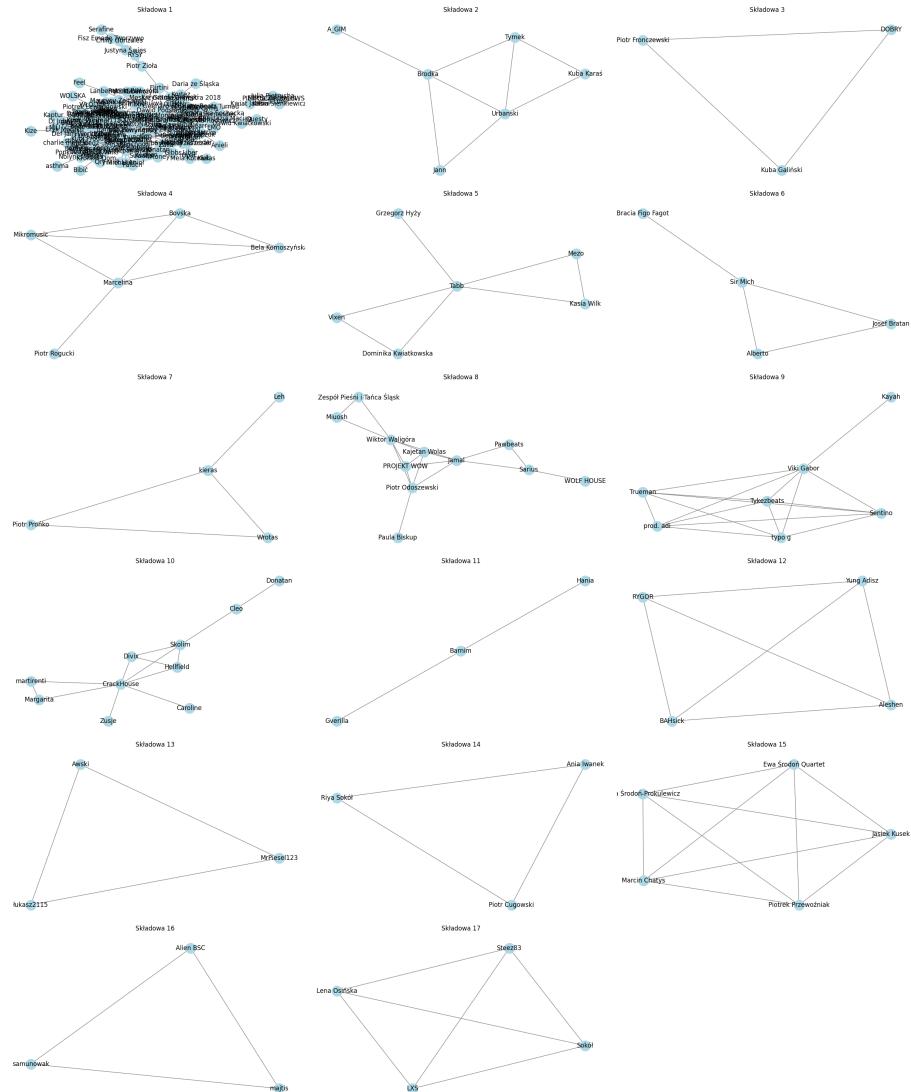
Stopień: 414

Rozmiar: 622

Gęstość: 0.0072756196558702085

3.2.2 Składowe spójne

Łącznie składowych spójnych jest 40, natomiast pomijając składowe złożone z dwóch węzłów, mamy ich 17. Największa składowa spójna ma 183 węzły, czyli to aż 44% wszystkich węzłów.



Rysunek 8: Składowe spójne

3.2.3 Współczynnik gronowania

Średni współczynnik gronowania: 0.3887508705661583

Globalny współczynnik gronowania: 0.40521478948493267

Kilka lokalnych współczynników gronowania zostało wypisanych poniżej:

Artysta	Wsp. gronowania
Błażej Król	0.428571
PIETRUCHA/ZALEWSKA	1.000000
Julia Pietrucha	1.000000
Natalia Szroeder	0.166667
Andrzej Piaseczny	0.000000
Steez83	1.000000
Lena Osińska	1.000000
Lil Tjay	1.000000
PG\$	0.833333
club2020	1.000000

Tabela 4: Tabela współczynników gronowania

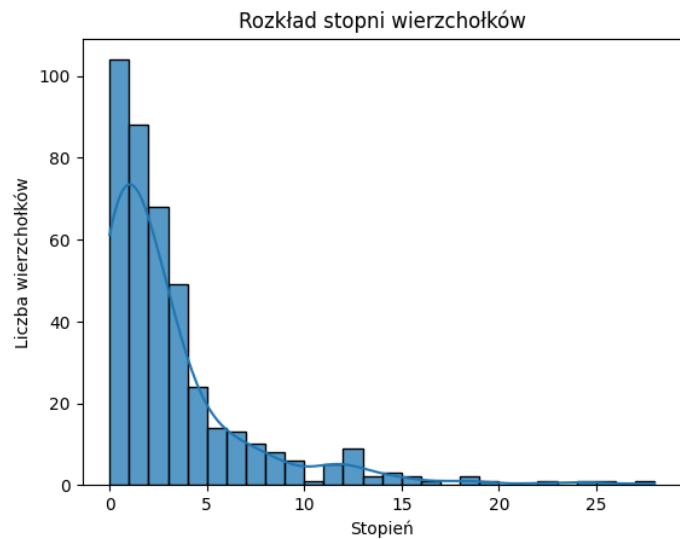
3.2.4 Rozkład wierzchołków

Aby ocenić rozkład wierzchołków grafu, wykonano poniższe wykresy oraz analizę statystyczną z biblioteki *powerlaw*:

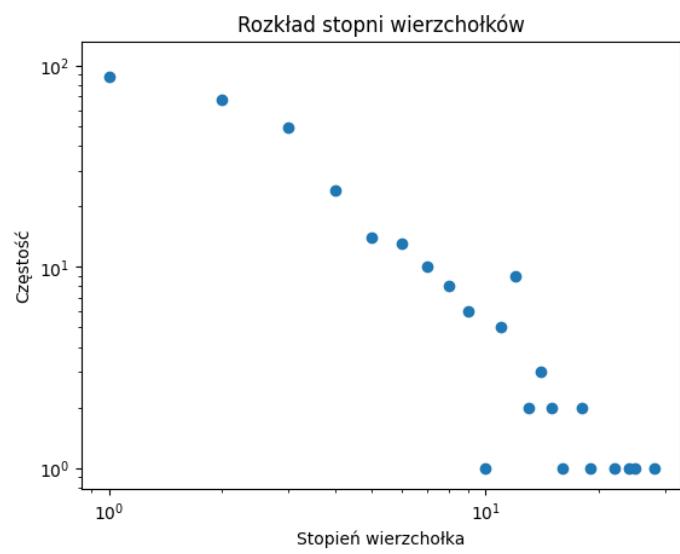
Calculating best minimal value for power law fit

Wyestymowany wykładnik potęgowy : 4.817903264440664

Test porównawczy z wykładniczym: R = 0.901473352629059, p-value = 0.3649724568446543



Rysunek 9: Wykres rozkładu stopni wierzchołków



Rysunek 10: Wykres rozkładu stopni wierzchołków - skala log

Wykresy oraz test potwierdzają, że rozkład w przybliżeniu może być uznany za rozkład potęgowy. Ma on spory współczynnik (> 3), a więc dość rzadko jak na ten rozkład będziemy mieli węzły o dużym stopniu. Test statystyczny wykazał, że rozkład potęgowy jest lepszym opisem niż wykładowczy.

3.2.5 Własności wierzchołków

Artysta	Stopień
White 2115	28
Oki	25
Mata	24
SB Maffija	22
Pedro	19
Taco Hemingway	18
Kubi Producent	18
Bedoies 2115	16
@atutowy	15
Kinny Zimmer	15

Tabela 5: 10 artystów o największym stopniu

Artysta	Bliskość
Sobel	0.171009
@atutowy	0.168495
Taco Hemingway	0.167439
Oki	0.167090
White 2115	0.164351
Pedro	0.160087
Mata	0.159134
Żabson	0.154534
francis	0.152478
Bedoies 2115	0.152189

Tabela 6: 10 artystów o największej bliskości

Artysta	Pośredniość
Oki	0.038045
Sobel	0.031656
Taco Hemingway	0.027977
Dawid Podsiadło	0.027557
Pedro	0.026850
White 2115	0.021997
Mata	0.021377
@atutowy	0.021085
Kubi Producent	0.019296
sanah	0.01788

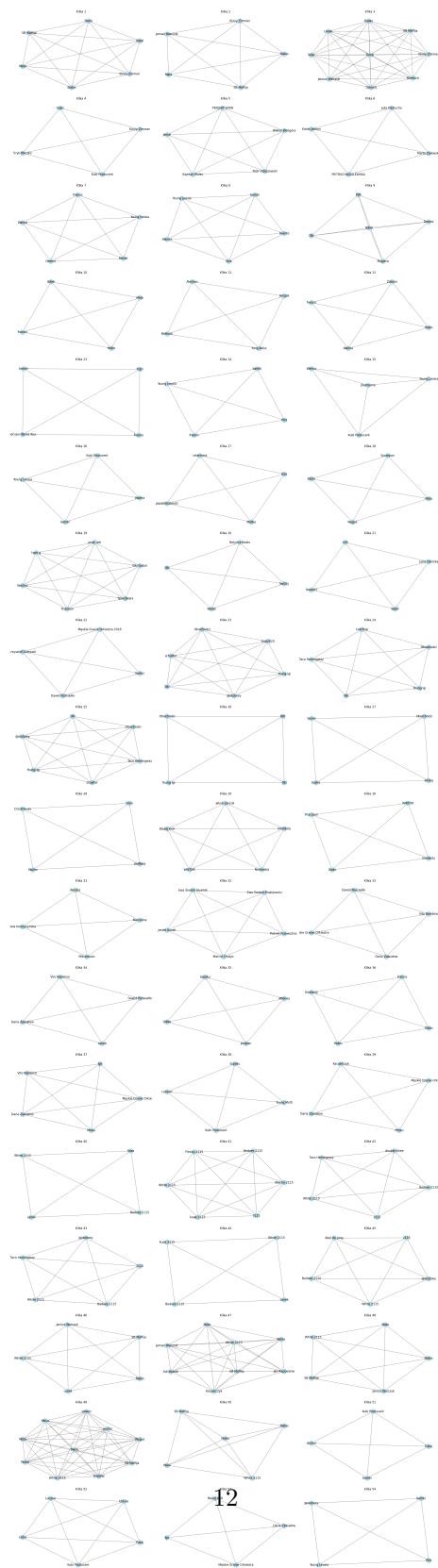
Tabela 7: 10 artystów o największej pośredniości

Artysta	Centralność wektora
White 2115	0.067797
Oki	0.060533
Mata	0.058111
SB Maffija	0.053269
Pedro	0.046005
Taco Hemingway	0.043584
Kubi Producent	0.043584
Bedoes 2115	0.038741
@atutowy	0.036320
Kinny Zimmer	0.036320

Tabela 8: 10 artystów o największej centralności

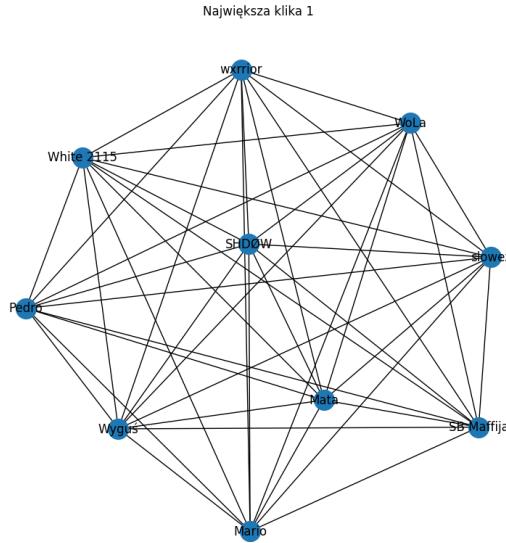
3.2.6 Kliki

Jeśli chodzi o kliki, to jest ich 202, z czego takich, które składają się z więcej niż 2 węzłów, jest 119.



Rysunek 11: Kliki o ilości węzłów większej niż 4

Największą z tych maksymalnych klik jest ta poniżej, złożona z 10 artystów:



Rysunek 12: Największa klika

Następnie, zostawiając ten sam graf, ale zmieniając jedynie labelle na ilości obserwujących za- miast nazw wykonawców, przeranaliżowano, czy popularni artyści częściej wspólnie nagrywają utwory.

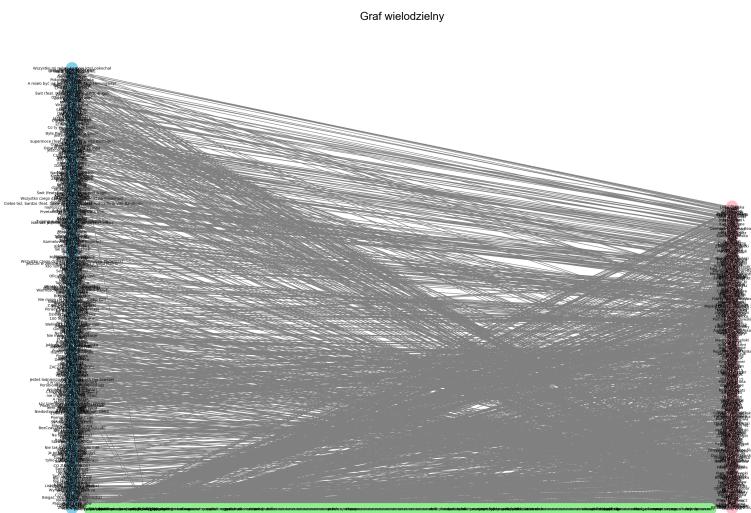
3.3 Graf wielodzielny

W kolejnym kroku dodano do grafu wielodzielnego następną kategorię, czyli wierzchołki gatunków muzycznych artystów. Gatunki zostały połączone z artystami, jeśli dany gatunek przypisany jest w Spotify do danego wykonawcy.

Wierzchołek z grupy górnej o największym stopniu: 3 w nocy w Międzyzdrojach z 10 połączonymi

Wierzchołek z grupy dolnej o największym stopniu: sanah z 31 połączonymi

Wierzchołek z grupy środkowej o największym stopniu: polish hip hop z 98 połączonymi



Rysunek 13: Enter Caption

3.4 Homofilia wierzchołków

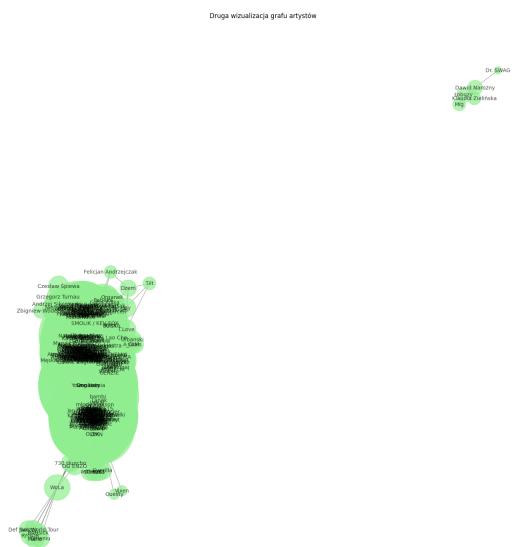
Homofilię wierzchołków w projekcie zastosowano w taki sposób, że jeżeli dwoje artystów ma taki sam gatunek muzyczny, to łączymy ich. Ważne, aby wiedzieć, że w Spotify jeden wykonawca może mieć kilka gatunków.

3.4.1 Własności grafu

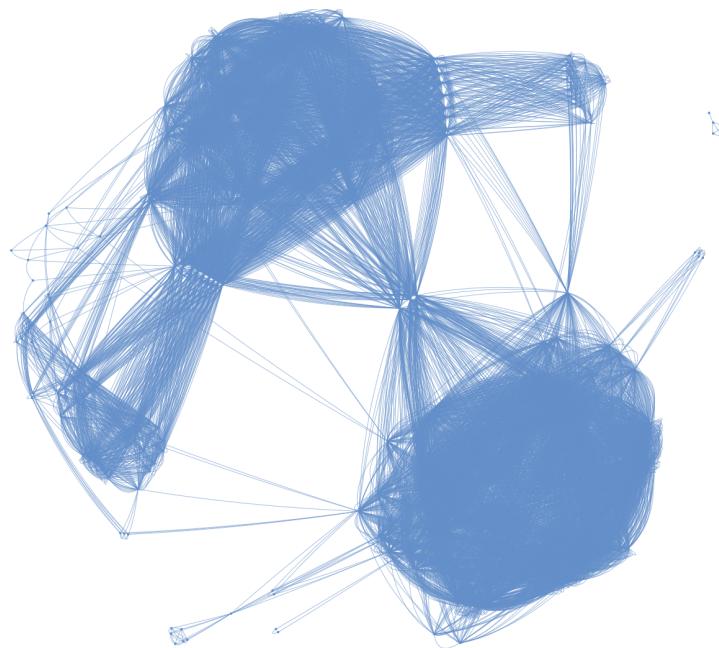
Stopień: 275

Rozmiar: 9632

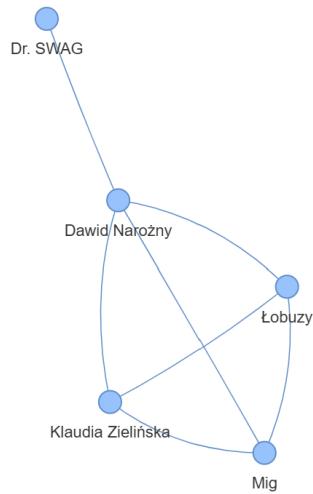
Gęstość: 0.25566025215660254



Rysunek 14: Graf artystów względem gatunku



Rysunek 15: Graf artystów względem gatunku - pyviz



Rysunek 16: Odrębna składowa grafu gatunków

3.4.2 Własności wierzchołków

Artysta	Stopień
Smolasty	179
Young Leosia	177
Donatan	177
Natalia Przybysz	122
Nosowska	120
bambi	120
Vito Bambino	118
Kwiat Jabłoni	118
Bovska	118
Ralph Kaminski	118

Tabela 9: 10 artystów o największym stopniu

Artysta	Bliskość
Smolasty	0.721561
Young Leosia	0.717639
Donatan	0.717639
Natalia Przybysz	0.602948
Nosowska	0.601575
Vito Bambino	0.597491
Kwiat Jabłoni	0.597491
Bovska	0.597491
Ralph Kaminski	0.597491
IGNACY	0.597491

Tabela 10: 10 artystów o największej bliskości

Artysta	Pośredniość
Smolasty	0.124214
Young Leosia	0.117196
Donatan	0.117196
bambi	0.051720
WoLa	0.042191
Nosowska	0.033141
Natalia Przybysz	0.031869
Lanek	0.028705
Vito Bambino	0.018696
Kwiat Jabłoni	0.018696

Tabela 11: 10 artystów o największej pośredniości

Artysta	Centralność wektora
Smolasty	0.653285
Young Leosia	0.645985
Donatan	0.645985
Natalia Przybysz	0.445255
Nosowska	0.437956
bambi	0.437956
Vito Bambino	0.430657
Kwiat Jabłoni	0.430657
Bovska	0.430657
Ralph Kaminski	0.430650

Tabela 12: 10 artystów o największej centralności

4 Analiza wyników

4.1 Graf dwudzielny

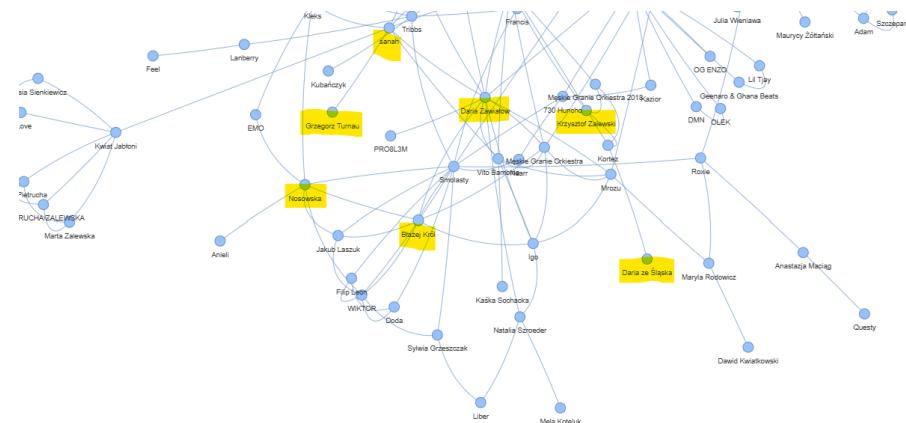
Najważniejsze wnioski z grafu dwudzielnego dotyczą wierzchołków o największym stopniu. Wnioskując z tabeli 2 widać, iż utworem, który zaangażował największą grupę wykonawców (10 osób) był "3 w nocy w Międzyzdrojach". Co ważniejsze, to wszystkie 10 utworów o największym stopniu to utwory hip-hopowe i rapowe, to sugeruje pierwsze wnioski, iż wykonawcy tych gatunków zdecydowanie częściej łączą się w duże grupy nagrywające jeden utwór.

Druga tabela 3 pokazuje zupełnie co innego, ponieważ tutaj jeśli dany artysta ma duży stopień, świadczy to o tym, iż piosenki z jego udziałem najczęściej występowały w pobranych playlistach, czyli były/są popularne, często odsłuchiwane w Polsce. Top 10 artystów z największym stopniem pokazuje zatem 10 najpopularniejszych artystów w latach 23-24. Pierwsze miejsce plasuje się dla sanah. Ciekawym faktem jest to, iż 10. miejsce zajął 'francis', czyli producent muzyczny. Wynika to głównie z tego, iż współpracuje on z wieloma artystami.

4.2 Graf artystów

Graf artystów, jak widać na wizualizacjach 7 składa się z jednej dużej składowej (183 wierzchołki) oraz kilku mniejszych składowych. Spora ilość wierzchołków jest również niepołączona, są one reprezentacją indywidualnych piosenek.

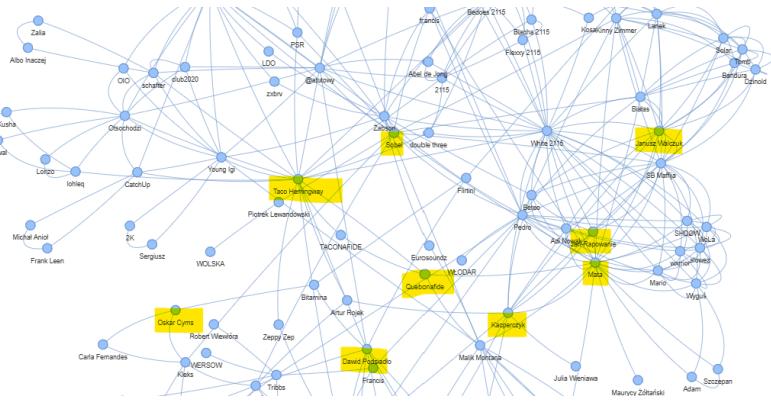
Dużą składową możemy podzielić na 3 etapy w kontekście analizy. Pierwszy etap to artyści starsi na scenie muzycznej, śpiewający utwory popowo/rockowo (np. Grzegorz Turnau, Nosowska, sanah, Zalewski):



Rysunek 17: Grupa 1

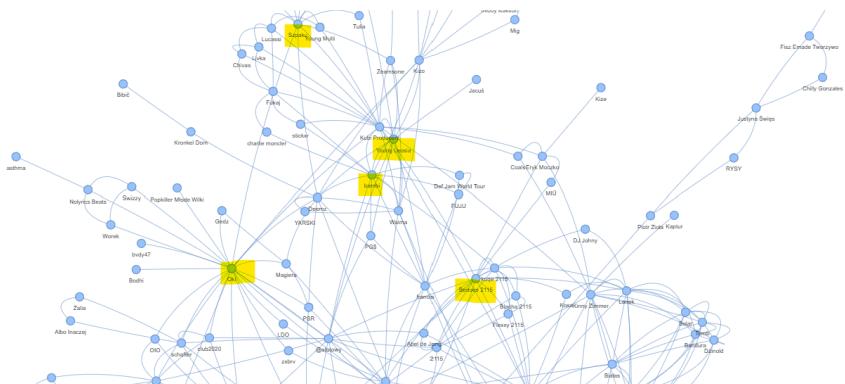
Następną grupą to młodzi ludzie śpiewający coś bliżej rapu, ale jest to jednak "ładniejszy" typ muzyki, często uważany za polską alternatywną muzykę. Takimi wykonawcami są np. Taco Hemingway, Dawid Podsiadło, Jan Rapowanie, Kacperczyk, Mata.

4.2 Graf artystów



Rysunek 18: Grupa 2

Ostatnią grupą są osoby typowo działające w rapie i hip hopie. Ich piosenki są już dość mocniejsze, można by rzec, "mniej grzeczne". Artyści w tej grupie to np. bambi, OKI, Young Leosia.

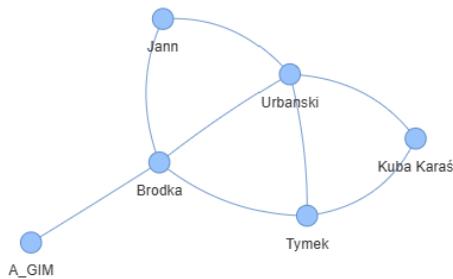


Rysunek 19: Grupa 3

To pokazuje zależności hierarchiczne na polskiej scenie muzycznej. Osoby rockowe współpracują z młodymi alternatywnymi, a ci częściej tworzą coś wspólnego z raperami.

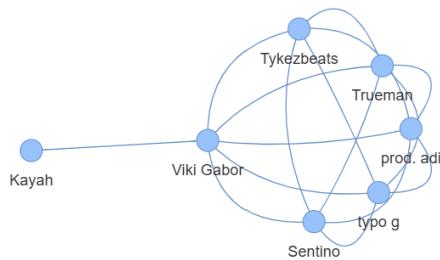
Mimo wszystko największą grupę połączeń rejestruje się w grupie ostatniej, czyli rapowo/hip hopowe środowisko najczęściej tworzy wspólne utwory.

Ciekawe wnioski wynikają również z analizy mniejszych składowych, gdyż na poniższym obrazku widać wykonawców, którzy pasują do wymienionej powyżej grupy 1, ale przez brak popularności wspólnych piosenek lub przestój artystyczny w 2024 roku nie znaleźli się oni w dużej składowej.



Rysunek 20: Wykonawcy pasujący do grupy 1

Jeśli natomiast w danym roku artysta wykonał wspólny utwór z drugim, który za to stworzył coś z zupełnie innym gatunkiem, możemy mieć dość niespodziewane bliskie połączenie typowo popowej artystki Kayah z raperem Sentino, a więc typowo dla sieci społecznych mamy tu zjawisko podobne do małych światów, czyli przypuszczalnie odrębne byty w muzyce mają jednak bliski kontakt (krótką ścieżkę).



Rysunek 21: Różnorodne powiązania

Współczynnik gronowania

Graf jest dość spójny, ale nie w pełni.

Istnieją w nim pewne wierzchołki (np. "Andrzej Piaseczny"), które mają bardzo niski współczynnik gronowania 4, co może sugerować, że te osoby nie są mocno powiązane z innymi w sieci (brak sąsiednich połączeń).

Dla "PIETRUCHA/ZALEWSKA" współczynnik wynosi 1, co oznacza, że wszyscy jej sąsiedzi są wzajemnie połączeni. Tacy artyści współpracują w bardzo ścisłym środowisku, często wydając więcej utworów razem niż oddzielnie.

Kliki

Duża ilość klik w grafie świadczy o ścisłych współpracych artystów, często nawet w obrębie jednego utworu. Są to grupy bardziej "zaprzyjaźnione" złożone z osób wspólnie pracujących np. w jednej wytwórni. To grupy zamknięte, mogące współpracować nad pewnymi nagraniami, np. wspólniej płyty.

Analiza wierzchołków

1. Stopień

Z tabeli 5 widać, iż największy stopień ma raper White 2115. Jest to zatem wykonawca bardzo popularny w omawianych latach oraz osoba, która często zapraszała/była zapraszana do tworzenia wspólnych utworów.

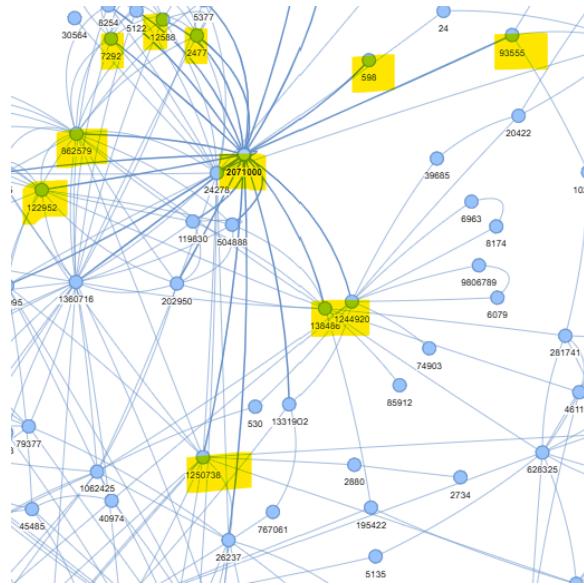
2. Bliskość

Największą bliskość ma za to Sobel - 6. Ma on "krótką ścieżkę" do przeciętnego wykonawcy, niezależnie od gatunku muzycznego. To może świadczyć o wspólnych piosenkach z kluczowymi artystami, którzy mają spory stopień. Drugi za nim plasuje się producent muzyczny, co jest zrozumiałe, ponieważ tworzy on z wieloma artystami.

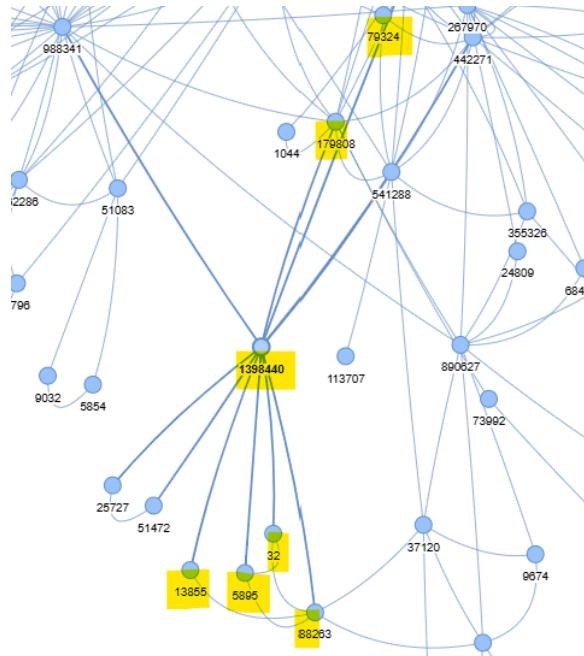
3. Pośredniość

Najlepszym pośrednikami są za to OKI, Taco, Sobel, Dawid Podsiadło - 7. Są to kluczowi artyści, którzy otwierają się na nowe grupy i połączenia z innymi artystami, łącząc grupy i gatunki "z których pochodzą". Ich znaknięcie ze sceny muzycznej mogłoby spowodować większą hermetyczność.

Sprawdzono również, czy popularni artyści częściej nagrywają z innymi popularnymi. Dwa pośrednie fragmenty grafu pokazują, że nie jest to reguła. Popularni artyści nie zawsze mają największą liczbę powiązań oraz łączą się również z mało popularnymi osobami (choć mogą to być osoby takie jak producenci).



Rysunek 22: Graf popularności urywek 1



Rysunek 23: Graf popularności urywek 2

4.3 Graf wielodzielny i graf gatunków

Po dołączeniu do grafu dwudzielnego wierzchołków z gatunkami oraz zrobieniu projekcji prostej, otrzymaliśmy graf, gdzie artyści łączą się, jeśli mają ten sam gatunek 15. Warto zaznaczyć, że w Spotify artyści mają zazwyczaj przypisanych kilka gatunków.

Analizując wizualizacje 15 widać, iż ten graf jest zdecydowanie gęstszy od wcześniejszego, ma on jedynie dwie składowe spójne. Mniejsza z nich - 16 prezentuje grupę zaledwie 5 węzłów. Są to artyści z gatunku disco polo. Z analizy wynika zatem, że jest to zupełnie odrębny byt, który nie mieszka gatunków, tzn. jeśli dana osoba jest przypisana do gatunku disco polo, to jest to jedyne jej przypisane. Są to artyści nieróżnorodni.

Zarówno w stopniu, bliskości, czy pośredniości, główną i najważniejszą postacią jest Smolasty, który ma dużą gamę przypisanych do siebie gatunków, przez co jest najlepszym łącznikiem między grupami, a jego odległość do przeciętnego wykonawcy jest krótka. Co ciekawe, nie ma on jednak największej ilości przypisanych gatunków, ale przypisane do niego gatunki są najbardziej popularne, czyli takie gatunki, jak on ma większość węzłów.

Artysta	Obserwujący	Gatunek	Liczba Gatunków
Lil Tjay	9,806,789	brooklyn drill, melodic rap, nyc rap, rap	4
Niemen	173,100	classic polish pop, poezja spiewana, sung poetry	3
Brodka	237,576	crank wave, polish indie, polish pop	3
Cyrko	49,352	polish pop, polish viral pop, polish viral rap	3
Smolasty	789,492	polish hip hop, polish pop, polish trap	3

Tabela 13: Liczba gatunków artystów

Dodatkowo widać, że oprócz gatunku disco polo polscy wykonawcy odnajdują się w wielu gatunkach i przez co tworzą jedną wspólną sieć. To może również wskazywać, że polscy artyści nie mają zerojedynkowego stylu piosenek i często, w zależności od wydanej płyty, mogą być przypisywani do różnych grup.

5 Wnioski

Analiza grafów polskich wykonawców muzycznych pozwoliła zidentyfikować hierarchiczne zależności na scenie muzycznej, wyróżniając trzy główne grupy artystów: starszych wykonawców popowo-rockowych, młodszych alternatywnych oraz środowisko rapowo-hiphopowe, które okazało się naj-

bardziej dynamiczne i skłonne do współpracy. Kluczowymi postaciami okazały się artyści pełniący rolę łączników między grupami, jak Smolasty, wpływający na strukturę sieci i kierunki współpracy.

Popularność artysty nie zawsze przekłada się na liczbę współprac – mniej znani wykonawcy, w tym producenci, często odgrywają kluczową rolę w budowaniu połączeń. Gatunek muzyczny istotnie wpływa na strukturę sieci, czego przykładem jest zamknięta grupa disco polo, pozbawiona powiązań z innymi gatunkami. Rap i hip-hop charakteryzują się natomiast dużą skłonnością do tworzenia wspólnych utworów, co wyróżnia te gatunki na tle innych.

Sieć wykazuje zjawisko „małych światów”, gdzie pozornie odrębne grupy są łączone przez pojedynczych artystów współpracujących międzygatunkowo. Analizy te pokazały, że metody grafowe dostarczają cennych wniosków o współpracy, popularności i preferencjach artystów, mogąc wspierać prognozowanie trendów oraz planowanie działań promocyjnych w branży muzycznej.

A Dodatek

Kody źródłowe umieszczone zostały w repozytorium GitHub:

<https://github.com/aleksandra0014/complex-networks>.