

# **RANKING PIOSENEK TAYLOR SWIFT POD WZGLĘDEM CECH HITU RADIOWEGO Z UŻYCIEM RÓŻNYCH METOD PORZĄDKOWANIA LINIOWEGO**

## **AUTORZY:**

- Szymon Mlonek 217441
- Oliwia Strzelec 217450
- Amelia Posiadała 217568
- Tomasz Niewiadomski 217437

## **STRESZCZENIE**

---

Projekt analizy rankingu piosenek Taylor Swift miał na celu porównanie różnych metod porządkowania liniowego oraz zbadanie wpływu wyboru metody na ostateczny ranking. Wykorzystano siedem metod porządkowania liniowego oraz przedstawiono na przykładzie omawianych danych ich zastosowanie krok po kroku.

Po przygotowaniu rankingów stwierdzono silną korelację między nimi, jednak metoda Nowaka wykazała największe odstępstwa od reszty, chociaż nadal rankingi wskazywały silne podobieństwo, co sugerowałoby, że analizując wyniki obliczeń i utwory pozycjonowane na najwyższych miejscach w rankingach, osoba potencjalnie zainteresowana stworzeniem hitu radiowego, mogłaby znacząco ułatwić swoje zadanie. Ostateczne wnioski wskazały na subtelne różnice w rankingu.

## **SŁOWA KLUCZOWE**

---

metody porządkowania liniowego, hity radiowe, Taylor Swift, metoda TOPSIS, metoda unitaryzacji zerowanej, metoda SSW, metoda Strahla, metoda Nowaka, ranking

## **WPROWADZENIE**

---

Przemysł muzyczny stale ewoluuje, a kluczowym aspektem sukcesu jest zrozumienie cech, które czynią utwór atrakcyjnym dla publiczności radiowej. Chociaż sztuka jest pełna odstępów, wyjątków i niespodziewanych sukcesów, mimo upływu lat nadal możemy zaobserwować pewne trendy<sup>1</sup> - większość piosenek odnoszących sukces w radiu jest energiczna, chwytliwa, ma przystępny rytm i bardzo zbliżoną do innych hitów długość utworu. Analizując te oraz kilka innych cech, przy pomocy statystyki możemy określić, które z utworów mają większe szanse na podbicie najpopularniejszych radiostacji, a które najmniejsze.

W ramach tego projektu skupimy się na analizie piosenek, pod kątem spełniania cech hitu radiowego, jednej z najbardziej znanych artystek popowych naszych czasów - Taylor Swift. Ranking został przygotowany na podstawie kilku metod porządkowania liniowego. Bazując na 8 cechach, 4 wpływających pozytywnie na potencjał utworu, 4 wpływających negatywnie, uzyskaliśmy ranking dla każdej z metod.

## **PRZEDMIOT BADANIA**

---

### **Cel**

Celem pracy jest analiza piosenek Taylor Swift pod kątem ich potencjału komercyjnego. Ranking piosenek może być pomocny dla artystów, którzy chcą tworzyć

---

<sup>1</sup> Tauberg, M. (2018). *Anatomy of a Hit Song (2000-2018)*. medium.com.

muzykę o większej szansie na sukces radiowy. W dzisiejszych czasach, gdzie przez różne platformy społecznościowe każdy może uzyskać błyskawiczną popularność jako artysta, znajomość trendów i cech najbardziej poszukiwanych w radiowych hitach może mieć ogromny wpływ na sukces. Wykorzystanie wielu metod rankingowych uczyni ostateczny rezultat bardziej uniwersalnym.

### **Przegląd literatury**

Zastosowanie wielu metod w badaniu pozwala na uzyskanie bardziej uniwersalnych wyników, które mogą też być bardzo pomocne przy wyborze metody porządkowania liniowego do badań własnych. Taki wybór może być znaczący, ponieważ rzutuje on na ranking badanych obiektów. Kukuła [2018] w swoich badaniach porównał wiele metod, w tym te zastosowane w naszej pracy, ostatecznie zauważając ogólne podobieństwa w wyprodukowanych wynikach, ale także subtelne różnice.

W badaniu podobieństwa uzyskanych wyników dużą pomocą może okazać się wykorzystanie macierzy korelacji. Przy porównaniu metod porządkowania liniowego Hellwiga oraz TOPSIS Hwang'a i Yoon'a, Bąk [2016] zauważył bardzo silne podobieństwa w wynikach obu metod, wykorzystując w tym celu współczynniki korelacji, chociaż nadal nie była to całkowita korelacja - metody wyróżniały się między sobą drobnymi wyjątkami. Wykorzystanie macierzy korelacji w naszych badaniach pozwoli na graficzne zilustrowanie różnic między wynikami zastosowanych metod, dzięki czemu możliwe będzie porównanie podobieństwa każdej z nich.

W wielu pracach jako jedna z najistotniejszych czynności w procesie budowy rankingu wymieniane jest dokonanie odpowiedniego wyboru cech diagnostycznych (Kukuła, [2018]). Należy określić te zarówno pożądane, jak i niepożądane. Biorąc pod uwagę obszar naszych badań, opierając się na własnym postrzeganiu hitów radiowych, ale także na wieloletnich analizach utworów podbijających światowe rankingi dokonanych przez osoby specjalizujące się w temacie, zauważyliśmy pewne wnioski. Jak możemy zauważyć w artykule Tauberg'a [2018] oraz w wielu innych pracach dotyczących tej tematyki, piosenki odnoszące największe komercyjne sukcesy są przeważnie pozytywne, energiczne, bardzo łatwo wpadające w ucho i często, poprzez swoją łatwo przyswajalną rytmikę, sprzyjające tańcu. Poza tym nie mogą być zbyt długie i zbyt skomplikowane. Dodatkowo, ze względu na naturę popularnych radiostacji słuchanych przez całe rodziny, takie utwory nie mogą zawierać ofensywnego słownictwa oraz poruszać kontrowersyjnych tematów (Spargo i Mayuri Chakkenchath [2020]).

### **Zmienne wybrane do analizy**

STYMULANTY:

- **danceability (taneczność)**
  - **Opis:** Określa, w jaki sposób utwór nadaje się do tańca w oparciu o kombinację elementów muzycznych, w tym stabilności rytmu, siły rytmu i ogólnej regularności.
  - **Uzasadnienie:** Energetyczny rytm i chwytliwa melodia przyciągną uwagę słuchaczy, zachęcając ich do tańca i ponownego odtwarzania, co sprawia, że staje się popularny wśród różnych grup odbiorców.

- **energy (energiczność)**
  - **Opis:** Reprezentuje percepcyjną miarę intensywności i aktywności. Zazwyczaj energiczne utwory wydają się szybkie i głośne.
  - **Uzasadnienie:** Dynamizm i intensywność przyciągną uwagę słuchaczy, pobudzając ich do działania i zapamiętania utworu
- **valence (pozytywność)**
  - **Opis:** Opisuje pozytywność muzyczną danego utworu.
  - **Uzasadnienie:** Radosna atmosfera i optymistyczne przesłanie sprawią, że słuchacze poczują się dobrze.
- **popularity (popularność)**
  - **Opis:** Określa popularność utworu na rynku muzycznym.
  - **Uzasadnienie:** Szerokie uznanie i rozpoznawalność przyciągną uwagę słuchaczy

#### DESTYMULANTY:

- **speechiness (obecność wypowiedzianych słów)**
  - **Opis:** Wykrywa obecność wypowiedzianych słów w utworze.
  - **Uzasadnienie:** Taki utwór może być trudniejszy do przyswojenia dla słuchaczy oraz mniej chwytliwy, co ogranicza jego potencjał do zapamiętania i powtarzalności
- **liveness (obecność widowni / nagrywania na żywo)**
  - **Opis:** Wykrywa obecność publiczności w nagraniu. Wyższe wartości żywotności oznaczają zwiększone prawdopodobieństwo, że utwór był wykonywany na żywo.
  - **Uzasadnienie:** Nagrania na żywo mogą zawierać niedoskonałości wykonawcze, hałas publiczności lub inne czynniki zakłócające, które mogą obniżyć jakość dźwięku i uniemożliwić komfortowe odsłuchanie
- **duration difference (długość utworu odstająca od oczekiwanej wartości)**
  - **Opis:** Określa, o ile długość utworu odbiega od oczekiwanej wartości (3:30<sup>2</sup>).
  - **Uzasadnienie:** Zbyt długi utwór może zniechęcić słuchaczy do jego regularnego słuchania, zwłaszcza jeśli jest postrzegany jako przeciągający się. Natomiast zbyt krótki utwór nie jest w stanie przekazać pełnej treści muzycznej, co może sprawić, że będzie traktowany jako mniej satysfakcjonujący. Została obliczona jako wartość absolutna różnicy średniej długości piosenek z rankingu Billboard Hot 100 z roku 2018 oraz długości konkretnego utworu.
- **explicit (obecność wulgaryzmów)**
  - **Opis:** Określa czy w utworze występują przekleństwa
  - **Uzasadnienie:** Obecność niestosownych słów w utworze ogranicza jego przeznaczenie tylko dla starszej publiczności, oraz nie jest mile widziane w radiostacjach<sup>3</sup>.

<sup>2</sup> Lindner, J. (2024). *Statistics About The Average Song Length*. Gitnux Market Data Report 2024.

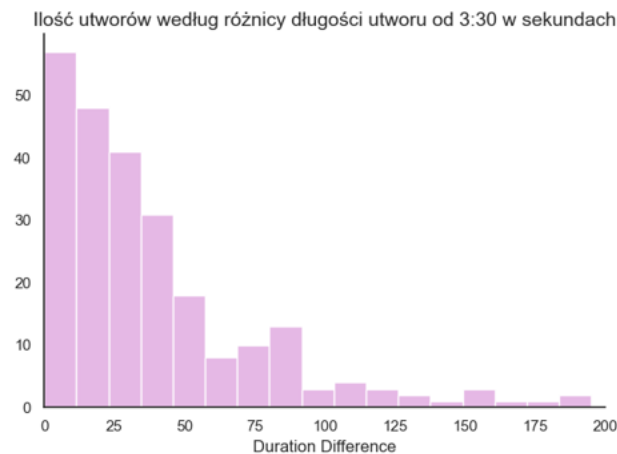
<sup>3</sup> Chakkenchath, A.-M., Spargo, O. (2020). *Explicit and sexual songs leave lasting impacts on the youth of today*, the Wildcat Tribune

## Wstępna analiza danych

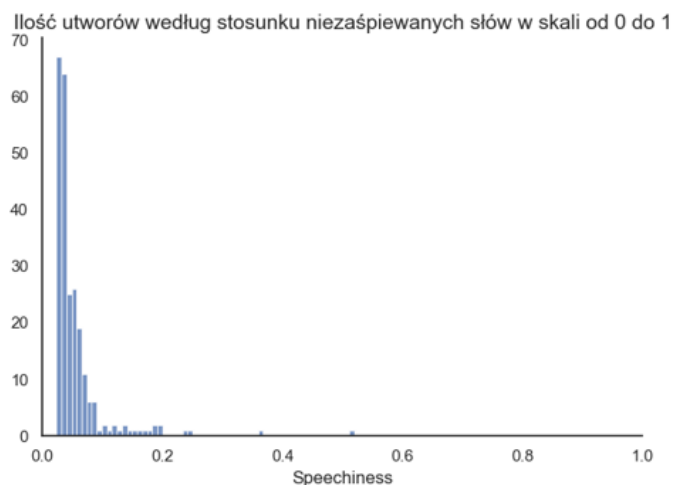
Statystyki opisowe - najważniejsze z charakterystyk liczbowych dla każdej z cech.

Charakterystyki liczbowe	danceability	energy	valence	speechiness	popularity	Duration Diff	liveness	explicit
Średnia	0,571495935	0,573577236	0,3929317	0,054661382	66,743902	39	0,177638	0,093495935
Dominanta	0,511	0,61	0,328	0,0363	70	10	0,108	0
Min	0,243	0,131	0,0382	0,0231	34	0	0,0357	0
Kwartyl 1	0,496	0,4365	0,24975	0,030825	59	13	0,0931	0
Kwartyl 2	0,5815	0,5845	0,3855	0,03845	70	28	0,1165	0
Kwartyl 3	0,651	0,7205	0,5195	0,057325	75	50,75	0,16225	0
Max	0,897	0,949	0,92	0,519	99	195	0,931	1
Rozstęp	0,654	0,818	0,8818	0,4959	65	195	0,8953	1
IQR	0,155	0,284	0,26975	0,0265	16	37,75	0,06915	0
Odchylenie standardowe	0,122760119	0,190669595	0,18728	0,050931048	13,741348	37,24155549	0,177455	0,291126167
Wariancja	0,015131557	0,036503282	0,0352169	0,002604559	189,59537	1392,594408	0,031619	0,085100382
Przedział międzykwartylowy	0,155	0,284	0,26975	0,0265	16	37,75	0,06915	0
Współczynnik zmienności	0,214804885	0,332421831	0,4766222	0,931755582	0,2058817	0,962840005	0,99897	3,113784221
Skośność	-0,202954153	-0,240131907	0,3338169	4,931915032	-0,715555	1,804787705	2,869117	2,792631588

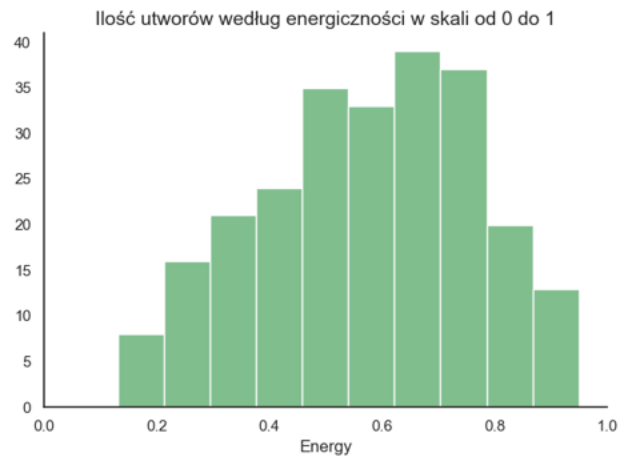
## Graficzne reprezentacje danych utworzone w Jupyter Notebook



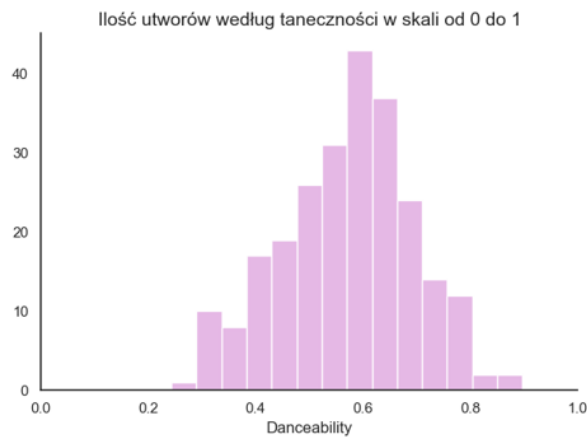
Wykres silnie skośny prawostronnie - większość piosenek jest bliska długości uznanej za idealną.



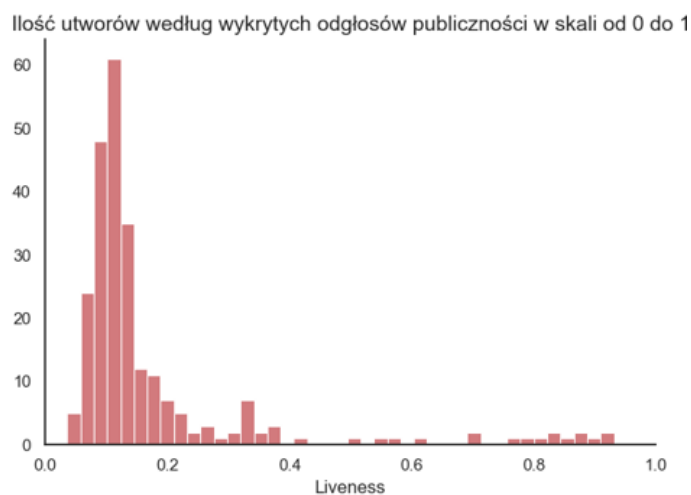
Wykres silnie skośny prawostronnie, ze względu na mały współczynnik “mówionych” słów w piosenkach Taylor Swift - znacząca większość z nich jest śpiewana.



Wykres rozłożony dosyć równomiernie z lekką skośnością lewostronną - energiczność utworów przypomina przesunięty rozkład naturalny.



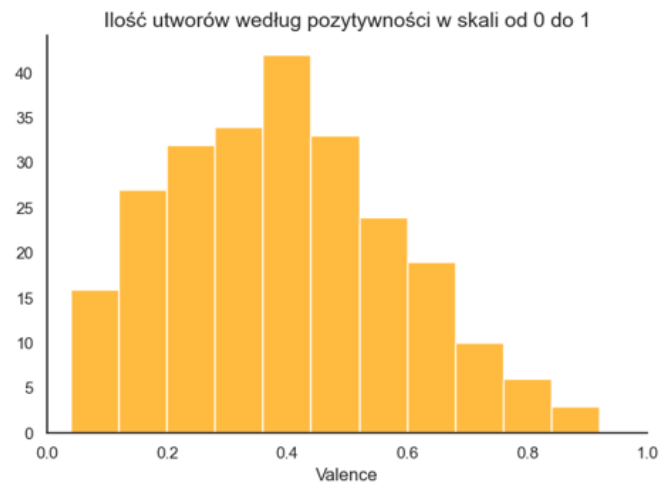
Wykres przypomina rozkład normalny z delikatną skośnością lewostronną.



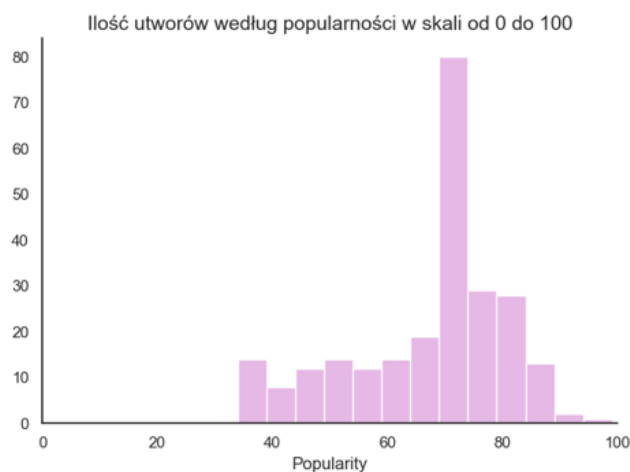
Wykres silnie skośny prawostronnie - większość utworów zawiera mało odgłosów publiczności



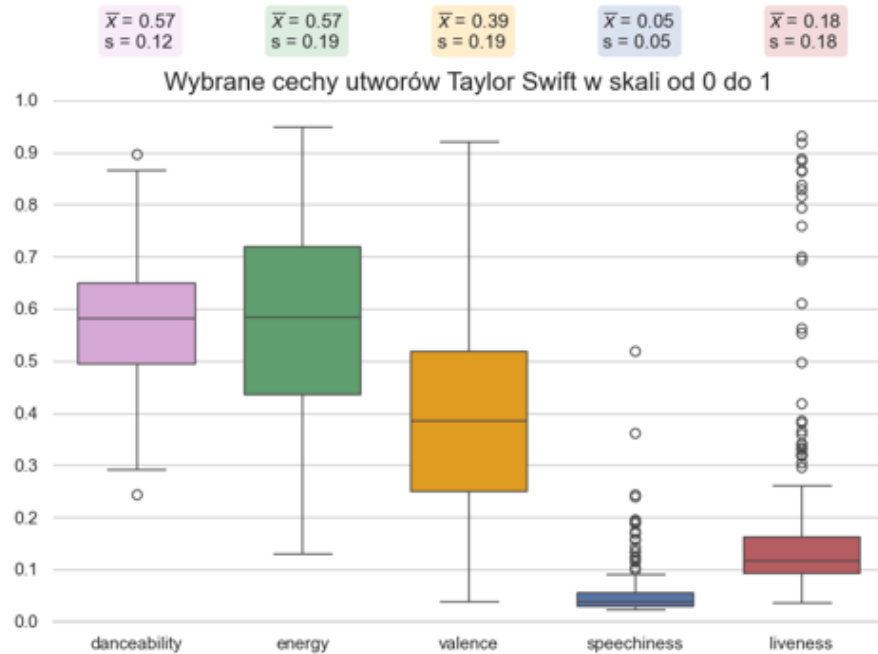
Jak widać na histogramie, duża większość utworów Taylor Swift nie jest uznana za ofensywne, co sprzyja popularności artystki w mediach głównego nurtu.



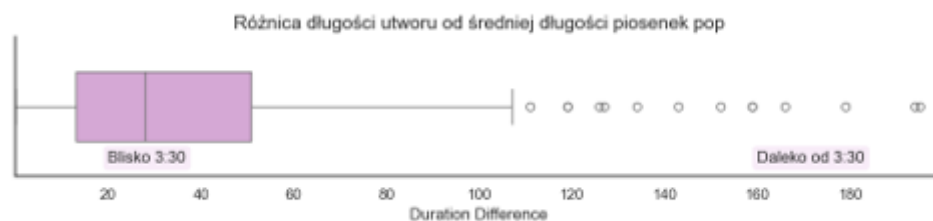
Wykres przypomina rozkład naturalny z lekką skośnością prawostronną.



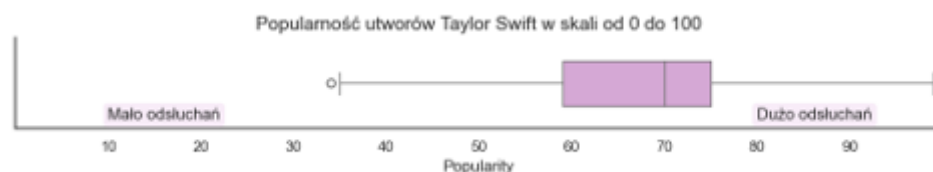
Na wykresie widzimy, że większość utworów Taylor Swift cieszy się ponadprzeciętną popularnością.



Na pierwszej grafice box ploty ukazujące rozmieszczenie danych w cechach opisujących charakter muzyczny utworu. Kolejno od lewej strony taneczność, energia, pozytywność, współczynnik słów “mówionych” oraz współczynnik odgłosów publiczności zawartych w utworze. Najwięcej odstępstw widocznych jest w dwóch ostatnich cechach, za co odpowiadają między innymi nagrania “na żywo” zawarte w zbiorze danych.



Cecha przedstawiona oddzielnie od poprzednich, ze względu na to co opisuje - nie charakter muzyczny, a rzeczywistość, widoczną na pierwszy rzut oka długość utworów. Widoczne jest kilka odstępstw od większości wyników.



Chociaż większość utworów Taylor Swift jest popularne, widoczne jest jedno odstępstwo na niekorzyść utworu - jedna z piosenek jest dużo mniej popularna niż wskazywałaby ogólna statystyka.

## Braki danych

Ze względu na źródło danych - baza danych oparta na statystykach udostępnionych przez Spotify - nie doznaliśmy żadnych braków danych. Za pomiary odpowiedzialne były specjalistyczne algorytmy przygotowane przez jedną z największych platform do streamingu muzyki - Spotify<sup>4</sup>, nie są to subiektywne opinie, ani pomiary narażone na niedokładność związaną z działaniem człowieka. Co za tym idzie, dane były kompletne i nie zaznaliśmy żadnych braków.

## Obserwacje odstające

Pomimo dużej ilości danych odstających (szczególnie w przypadku cech takich jak różnice długości utworów lub obecność odgłosów widowni na nagraniach), zdecydowaliśmy się nie zmieniać wartości tych danych, ze względu na to, że odstępstwa nie wynikają z błędów pomiarów - wartości odzwierciedlają stan rzeczywisty. W takim przypadku odstępstwa od normy widoczne na histogramach są czymś normalnym i spodziewanym, zmiana ich wartości mogłaby naruszyć rzetelność ostatecznego rankingu.<sup>5</sup>

## METODY PORZĄDKOWANIA LINIOWEGO

---

### Metoda TOPSIS

Technika porządkowania liniowego biorąca pod uwagę wiele cech mających wpływ na ostateczny ranking, oryginalnie wynaleziona przez Hwang'a i Yoon'a w 1981 roku. Główną ideą metody TOPSIS jest określenie odległości euklidesowej rozpatrywanych obiektów od rozwiązania idealnego i antyidealnego (wzorca i antywzorca). Posiadając te informacje można uzyskać rezultat w postaci wskaźnika syntetycznego, który jest pewnego rodzaju kompromisem biorącym pod uwagę obie odległości. Najlepszy obiekt ma najmniejszą odległość od wzorca i największą odległość od antywzorca<sup>6</sup>, a co za tym idzie - największą wartość wskaźnika syntetycznego. Jak pisze Adam Witczak w swojej pracy<sup>7</sup>, algorytm TOPSIS jest jedną z najwygodniejszych i najpopularniejszych metod rozstrzygania problemów wielokryterialnych.

W analizie danych operujemy na  $i$ -elementowym zbiorze elementów, gdzie każdy z nich posiada  $j$  cech. W metodzie TOPSIS można dodatkowo określić arbitralne wektory wag.

1. Pierwszym krokiem postępowania w metodzie TOPSIS jest normalizacja danych. W oryginalnej metodzie wykorzystujemy w tym celu następujące przekształcenie:

$$z_{ij} = \frac{x_{ij}}{S_j}$$

---

<sup>4</sup> Dokumentacja Spotify Web Api. *References/Tracks/Get Track's Audio Features*

<sup>5</sup> Trzęsiok, J. (2015). *O odporności na obserwacje odstające wybranych nieparametrycznych modeli regresji*. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach, ISSN 2083-8611, Nr 227 · 2015

<sup>6</sup> Hwang C.-L., Yoon K. (1981) *Multiple Attribute Decision Making*. Lecture Notes in Economics and Mathematical Systems 186, Springer

<sup>7</sup> Witczak, A. (2017). *Metoda TOPSIS w analizie wielokryterialnej*. finweb.pl



gdzie  $x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego elementu, a  $S_j$  oznacza odchylenie standardowe  $j$ -tej cechy. W ten sposób otrzymujemy znormalizowaną macierz decyzyjną.

- Następnie wykorzystując ustalone wagi arbitralne (w naszym przypadku dla każdej z cech o wartości  $\frac{1}{m}$ ), wyznaczamy macierz ocen ważonych o wartościach

$$V_{ij} = w_j * z_{ij}$$

gdzie  $w_j$  oznacza wagę  $j$ -tej cechy, przy czym suma wag wszystkich cech musi być równa 1, a  $m$  jest liczbą cech

- Dla każdej z cech wyznaczamy współrzędne wzorca i antywzorca.

$V_j^+$  - współrzędna wzorca  $j$ -tej cechy (dla stymulanty wartość maksymalna, dla destymulanty wartość minimalna)

$V_j^-$  - współrzędna antywzorca  $j$ -tej cechy (dla stymulanty wartość minimalna, dla destymulanty wartość maksymalna)

- Dla  $i$ -tych obiektów wyznaczamy odległość od rozwiązania idealnego oraz rozwiązania antyidealnego, biorąc kolejno:

$$S_i^+ = \sqrt{\sum_{j=1}^m (V_{ij} - V_j^+)^2}$$

oraz

$$S_i^- = \sqrt{\sum_{j=1}^m (V_{ij} - V_j^-)^2}$$

- Dla wszystkich  $i$ -tych obiektów wyznaczamy wartość wskaźnika syntetycznego:

$$P_i = \frac{S_i^-}{S_i^+ + S_i^-}$$

a następnie porządkujemy obiekty według klucza, w którym obiekt o największej wartości wskaźnika syntetycznego jest najlepszy

### **METODA TOPSIS (standaryzacja danych)**

Dokonując drobnych zmian w schemacie oryginalnej metody TOPSIS, możemy utworzyć kolejne metody porządkowania liniowego, które możemy określić metodami rodziny TOPSIS. Możliwe jest przykładowo wykorzystanie innej metody normalizacji

danych<sup>8</sup>. Pozostawiając kolejne kroki w ich podstawowej formie, jedynym etapem wymagającym zmiany jest krok pierwszy, w którym wykorzystujemy następujące przekształcenie

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}$$

gdzie  $S_j$  jak poprzednio oznacza odchylenie standardowe  $j$ -tej cechy,  $x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego elementu, a nowe oznaczenie  $\bar{x}_j$  symbolizuje średnią wartość  $j$ -tej cechy.

Chociaż wszystkie kolejne kroki metody pozostają w swojej podstawowej formie, to nawet taka drobna zmiana potrafi zmienić znacząco ostateczny ranking.

### **METODA HELLWIGA**

Metoda Hellwiga jest jedną z metod porządkowania liniowego obiektów wielocechowych. Została ona zaproponowana w 1968 roku przez Zbigniewa Hellwiga ([Hellwig, 1968])<sup>9</sup> i od tego czasu jest szeroko stosowana w różnych dziedzinach, takich jak ekonomia, finanse, marketing, zarządzanie, a także nauki społeczne i przyrodnicze.

#### **Zasada działania:**

Metoda Hellwiga opiera się na idei wzorca. Wzorec to obiekt hipotetyczny, który posiada najlepsze wartości wszystkich cech. Odległość każdego obiektu od wzorca jest obliczana za pomocą odległości euklidesowej. Obiekty o mniejszych odległościach od wzorca są uważane za lepsze i zajmują wyższe pozycje w porządku liniowym.

#### **Etapy metody Hellwiga<sup>10</sup>:**

Normalizacja zmiennych (standaryzacja). Standaryzujemy dane za pomocą wzoru:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}$$

gdzie

$z_{ij}$  - wartość unormowana  $j$ -tej cechy dla  $i$ -tego obiektu

$x_{ij}$  - wartość  $j$ -tej cechy dla  $i$ -tego obiektu

$\bar{x}_j, S_j$  to odpowiednio średnia arytmetyczna i odchylenie standardowe  $j$ -tej cechy

#### **Wyznaczanie współrzędnych wzorca:**

Maksymalna wartość dla cech, które są stymulantami i minimalna dla destymulant.

<sup>8</sup> Kukula, K., Luty, L. (2018). O Wyborze metody porządkowania liniowego do oceny gospodarki odpadami w Polsce w ujęciu przestrzennym. Problemy Rolnictwa Światowego tom 18, zeszyt 2, 2018: 183-192

<sup>9</sup> Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologii gospodarstw rolnych. Roczniki Nauk Rolniczych. Przegląd Statystyczny, 4, 307-327.

<sup>10</sup> Bąk, A. (2016). Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 426, 22-3

**Obliczenie odległości euklidesowych:**

Dla każdego obiektu obliczana jest odległość euklidesowa od wzorca. Odległość od wzorca liczymy za pomocą wzoru:

$$d_i^+ = \sqrt{\sum_{j=1}^m (z_{ij} - z_j^+)^2}$$

gdzie  $d_i^+$  oznacza odległość obiektu od wzorca,  $m$  jest liczbą cech, a  $z_j^+$  to wartość wzorca dla  $j$  - tej cechy

**Porządkowanie obiektów:**

Obiekty są syntetycznej zmiennej agregowanej  $Q$ . Obiekty o mniejszych odległościach zajmują wyższe pozycje w porządku liniowym. Zmienną  $Q$  obliczamy za pomocą wzorów:

$$S_d = \sqrt{\frac{\sum_{i=1}^n (d_i^+ - \bar{d})^2}{n}}$$

$$d_0 = \bar{d} + 2S_d$$

$$Q_i = 1 - \frac{d_i^+}{d_0}$$

gdzie

$S_d$  – Odchylenie standardowe odległości od wzorca.

$\bar{d}$  – Średnia odległość od wzorca obiektów.

$Q_i$  – wartość cechy syntetycznej dla  $i$ -tego obiektu

Obiekty porządkujemy w kolejności od największej wartości  $Q$ .

W badaniach porównawczych A.Bąk między metodą Hellwiga a metodą TOPSIS, zastosował różne zestawy danych empirycznych. Wyniki tych badań pokazały, że obie metody mogą prowadzić do podobnych lub zgodnych wyników w porządkowaniu obiektów. Współczynniki korelacji liniowej między wynikami uzyskanymi z obu metod były wysokie, co sugeruje, że istnieje silna zależność między wynikami uzyskanymi z metod Hellwiga i TOPSIS.

**METODA UNITARYZACJI ZEROWANEJ**

Metoda ta opiera się na ocenie wielu obiektów przy użyciu określonych kryteriów, które mogą być wyrażone różnymi miarami. Głównym celem metody standaryzacji zerowej jest znormalizowanie tych kryteriów. W tej metodzie uwzględniamy zarówno czynniki, które wykazują dodatnią korelację z analizowaną zmienną (stymulanty), jak i te, które wykazują ujemną korelację (destymulanty).<sup>11</sup>

<sup>11</sup> Kościółek, M. (2015). *Wykorzystanie analizy wielowymiarowej do badania zróżnicowanego potencjału innowacyjnego Polski, Metody Ilościowe w Badaniach Ekonomicznych*. Tom XVI/3, 2015, str. 194-201

Karol Kukuła w swojej pracy<sup>12</sup> wywnioskował, iż najbardziej właściwą metodą normującą cechy ilościowe w korespondencji z metodą normującą cechy jakościowe jest właśnie metoda unitaryzacji zerowanej.

W badaniu danych korzystamy z zestawu  $i$  elementów, z których każdy ma  $j$  cech. W pierwszym kroku normalizuje się zmienne za pomocą odpowiednich wzorów:

a) dla stymulant:

$$z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (i = 1, 2, \dots, r; j = 1, 2, \dots, s)$$

b) dla destymulant:

$$z_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (i = 1, 2, \dots, r; j = 1, 2, \dots, s)$$

gdzie  $x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego obiektu, natomiast  $z_{ij}$  oznacza wartość unormowaną  $j$ -tej cechy dla  $i$ -tego obiektu

W wyniku normalizacji cech według powyższych wzorów tworzona jest macierz. W drugim kroku dla każdego  $i$ -tego obiektu liczona jest średnia z jego  $m$  wartości unormowanych  $z_{ij}$

$$Q_i = \frac{1}{m} \sum_{j=1}^m z_{ij}$$

gdzie  $Q_i$  to wartość cechy syntetycznej dla  $i$ -tego obiektu, a  $m$  jest liczbą cech

Ostatnim krokiem jest uporządkowanie obiektów pod względem wartości cech syntetycznych. Obiekt o największej wartości tej cechy uznawany jest za najlepszy.

## METODA SSW

Metoda ta została zastosowana w pracy J.Perkala już w 1960 roku<sup>13</sup>. Mimo, że porządkowanie liniowe nie jest centralnym tematem tej pracy, możemy w niej znaleźć analizę danych oraz normalizację z użyciem przedstawionych poniżej wzorów. Metoda SSW może być używana w celu porządkowania liniowego.<sup>14</sup>

<sup>12</sup> Kukuła, K. (2012). *Propozycja budowy rankingu obiektów w wykorzystaniu cech ilościowych oraz jakościowych*. Metody Ilościowe w Badaniach Ekonomicznych, Tom XIII/1, 2012, str. 5-16

<sup>13</sup> Perkal, J. (1960). *On the analysis of a set of characteristics*. Applicationes Mathematicae V

<sup>14</sup> Kukuła, K., Luty, L. (2018). *O Wyborze metody porządkowania liniowego do oceny gospodarki odpadami w Polsce w ujęciu przestrzennym*. Problemy Rolnictwa Światowego tom 18, zeszyt 2, 2018: 183-192

W naszym badaniu danych operujemy zbiorem, który zawiera  $n$  elementów, z których każdy jest opisany przez  $j$  cech. Przed przystąpieniem do dalszej analizy, konieczne jest znormalizowanie danych, co można osiągnąć za pomocą poniższego wzoru:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}$$

gdzie:

$x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego obiektu

$z_{ij}$  oznacza wartość unormowaną  $j$ -tej cechy dla  $i$ -tego obiektu,

$\bar{x}_j$  oznacza średnią arytmetyczną  $j$ -tej cechy

$S_j$  oznacza odchylenie standardowe  $j$ -tej cechy

Warto zaznaczyć, że dla destruktywnych cech wartość znormalizowana jest obliczana ze znakiem przeciwnym.

W drugim kroku obliczana jest wartość cechy syntetycznej dla  $i$ -tego obiektu według wzoru:

$$Q_i = \frac{1}{m} \sum_{j=1}^m z_{ij}$$

gdzie  $Q_i$  to wartość cechy syntetycznej dla  $i$ -tego obiektu, a  $m$  jest liczbą cech.

Ostatecznym etapem jest uporządkowanie obiektów według wartości cech syntetycznych, gdzie obiekt o najwyższej wartości cechy uznawany jest za najlepszy.

Zastosowanie opisanych powyżej wzorów możemy znaleźć w pracy Perkal [1960], w której autor dokonuje normalizacji danych w przedstawiony sposób.

## METODA STRAHLA

W swoim artykule z 1978 roku Strahl<sup>15</sup> przedstawia propozycję konstrukcji miary syntetycznej, która może być wykorzystana do porównywania obiektów wielocechowych.

W naszym badaniu danych manipulujemy zbiorem zawierającym  $i$  elementów, z których każdy jest opisany przez  $j$  cech.

Należy znormalizować dane za pomocą wzoru:

$$z_{ij} = \frac{x_{ij}}{\max_i x_{ij}}$$

gdzie

$z_{ij}$  oznacza wartość unormowaną  $j$ -tej cechy dla  $i$ -tego obiektu,

$x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego obiektu.

Należy zaznaczyć, że dla destymulant wartość unormowana brana jest ze znakiem przeciwnym.

---

<sup>15</sup> Strahl, D. (1978). *Propozycja konstrukcji miary syntetycznej*. Przegląd Statystyczny, z. 2, s. 205–215

W efekcie normalizacji cech zgodnie z przedstawionymi wzorami, generowana jest macierz. W kolejnym etapie, dla każdego z obiektów, obliczana jest średnia z jego  $m$  wartości znormalizowanych  $z_{ij}$ . W wersji unormowanej miara syntetyczna jest definiowana jako:

$$Q_i = \frac{1}{m} \sum_{j=1}^m z_{ij}$$

gdzie  $Q_i$  to wartość cechy syntetycznej dla  $i$ -tego obiektu, a  $m$  jest liczbą cech.

Kończącym etapem jest uporządkowanie obiektów według wartości cech syntetycznych, gdzie obiekt najlepszy to obiekt o najwyższej wartości cechy.

## **METODA NOWAKA**

Metoda Nowaka jest jedną z metod porządkowania liniowego obiektów wielocechowych.<sup>16</sup> Metoda ta opiera się na koncepcji syntetycznego miernika, który jest liniową kombinacją wartości cech diagnostycznych. Celem metody jest uporządkowanie obiektów na podstawie wartości tego miernika.

Pierwszym krokiem jest standaryzacja danych diagnostycznych, czyli obliczenie średniej arytmetycznej oraz przekształcenie wartości cech do postaci znormalizowanej. W tym celu korzysta się ze wzoru:

$$z_{ij} = \frac{x_{ij}}{\bar{x}_j}$$

gdzie:

$z_{ij}$  oznacza wartość unormowaną  $j$ -tej cechy dla  $i$ -tego obiektu,

$x_{ij}$  oznacza wartość  $j$ -tej cechy dla  $i$ -tego obiektu.

$\bar{x}_j$  oznacza średnią arytmetyczną  $j$ -tej cechy

Należy zwrócić uwagę, że w przypadku destymulantów wartości unormowane są brane z odwrotnym znakiem.

Następnie dla każdego obiektu liczymy średnią z  $m$  wartości, które zostały wcześniej przekształcone do postaci unormowanej

$$Q_i = \frac{1}{m} \sum_{j=1}^m z_{ij}$$

gdzie  $Q_i$  to wartość cechy syntetycznej dla  $i$ -tego obiektu, a  $m$  jest liczbą cech.

Kończącym działaniem jest przeprowadzenie porządkowania obiektów na podstawie wartości syntetycznej.

---

<sup>16</sup> Nowak, E. (1977). *Syntetyczne mierniki plonów w krajach europejskich (Synthetic measure of crops in European countries)*. Wiadomości Statystyczne, 10, 19-22.

## ZASTOSOWANIE METOD DLA DANYCH

### Metoda TOPSIS

Pierwszym krokiem jest utworzenie macierzy decyzyjnej znormalizowanej, która będzie podstawą w kolejnych obliczeniach.

track_name	danceability	energy	valence	speechiness	popularity	duration diff	liveness	explicit
All Too Well	0,06566299	0,063817	0,0462861	0,027905613	0,039296656	0,141305155	0,0931905	0
Blank Space (Taylor's Version)	0,07995178	0,077319	0,1026789	0,057176639	0,067365695	0,024936204	0,0426594	0
Forever & Always (Piano Version) (Taylor's Version)	0,059554804	0,028797	0,0396947	0,033879292	0,063623157	0,067683982	0,0228279	0
Begin Again (Taylor's Version)	0,056609787	0,055589	0,0391088	0,038146205	0,06830133	0,033248272	0,0977612	0
This Is Why We Can't Have Nice Things	0,061845374	0,083226	0,064156	0,027052231	0,072979503	0,003562315	0,0332642	0
the 1 - the long pond studio sessions	0,077224911	0,026371	0,0228501	0,062211597	0,057073714	0,010686944	0,0911591	0,2085144
Drops Of Jupiter - Live/2011	0,04079395	0,053163	0,0395482	0,037463499	0,049588637	0,124681019	0,0174447	0
"Slut!" (Taylor's Version) (From The Vault)	0,064463167	0,043459	0,0480438	0,165556238	0,069236965	0,035623148	0,0774472	0
Girl At Home (Taylor's Version)	0,075370641	0,077635	0,0896426	0,055981903	0,063623157	0,011874383	0,0284396	0
the lakes - bonus track	0,034140392	0,027214	0,0392553	0,043949208	0,069236965	0,001187438	0,0238436	0

Dane po pierwszym przekształceniu - normalizacji.

Następnym krokiem jest przekształcenie danych zgodnie z wektorem wag cech oraz dobranie współrzędnych rozwiązania idealnego i antyidealnego, spośród wartości uzyskanych po przekształceniach.

Charakterystyki liczbowe	danceability	energy	valence	speechiness	popularity	Duration Diff	liveness	explicit
Współrzędne wzorca	0,012230004	0,012512862	0,016844614	0,002464142	0,011578479	0	0,00113314	0
Współrzędne antywzorca	0,003313145	0,001727276	0,000699418	0,055363201	0,003976447	0,028943808	0,02955053	0,026064302

Ostatnie etapy to obliczenie dystansów obiektów od współrzędnych wzorca i antywzorca, a następnie wyliczenie na ich podstawie wartości wskaźnika syntetycznego, w ostateczności dają nam następujący ranking (biorąc pod uwagę pierwsze 10 miejsc)

Miejsce	album_name	track_name	P	W
1	Lover	Paper Rings	0,94041	
2	Lover	ME! (feat. Brendon Urie of Panic! At The Disco)	0,905231	
3	evermore (deluxe version)	closure	0,901961	
4	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)	0,897827	
5	Lover	The Man	0,897085	
6	Taylor Swift	Picture To Burn	0,896187	
7	1989 (Taylor's Version) [Deluxe]	I Know Places (Taylor's Version)	0,895984	
8	Fearless (Taylor's Version)	Forever & Always (Taylor's Version)	0,895454	
9	1989 (Taylor's Version) [Deluxe]	Welcome To New York (Taylor's Version)	0,895445	
10	Red (Taylor's Version)	22 (Taylor's Version)	0,894445	

### Metoda TOPSIS (standaryzacja danych)

Główną różnicę od metody TOPSIS widzimy w pierwszym kroku, mającym największe znaczenie na ostateczny wynik. Tak wygląda macierz uzyskana po pierwszych przekształceniach.

track_name	danceability	energy	valence	speechiness	popularity	duration diff [sekundy]	liveness	explicit
All Too Well	0,247979573	0,164467	-0,409949	-0,430321005	-1,797025873	2,152372976	1,0649292	-0,320499
Blank Space (Taylor's Version)	1,312930227	0,834419	1,6416144	0,241768319	0,381724076	-0,473742096	-0,0542008	-0,320499
Forever & Always (Piano Version) (Taylor's Version)	-0,207266509	-1,573222	-0,649742	-0,293159918	0,091224083	0,490953237	-0,4934171	-0,320499
Begin Again (Taylor's Version)	-0,426760155	-0,243785	-0,671057	-0,195187714	0,454349075	-0,286162448	1,166157	-0,320499
This Is Why We Can't Have Nice Things	-0,036549228	1,127524	0,240157	-0,449915446	0,817474066	-0,956089762	-0,2622802	-0,320499
the 1 - the long pond studio sessions	1,109695369	-1,693605	-1,262546	0,357375521	-0,417150905	-0,795307207	1,019939	3,1074489
Drops Of Jupiter - Live/2011	-1,60552233	-0,364168	-0,655071	-0,210863266	-0,998150891	1,77721368	-0,612641	-0,320499
"Slut!" (Taylor's Version) (From The Vault)	0,158556235	-0,845696	-0,346004	2,730262319	0,526974073	-0,232568263	0,7162555	-0,320499
Girl At Home (Taylor's Version)	0,971495666	0,850121	1,167357	0,214336102	0,091224083	-0,768510114	-0,3691318	-0,320499
the lakes - bonus track	-2,101415382	-1,651732	-0,665728	-0,061945515	0,526974073	-1,009683947	-0,470922	-0,320499

Podobnie jak w oryginalnej metodzie najpierw bierzemy pod uwagę wagi arbitralne, a następnie wybieramy współrzędne wzorca i antywzorca, tym razem prezentujące się w następujący sposób

Charakterystyki liczbowe	danceability	energy	valence	speechiness	popularity	Duration Diff	liveness	explicit
Współrzędne wzorca	0,330768862	0,245620528	0,351075934	-0,077303455	0,292824879	-0,12956013	-0,0997782	-0,040062402
Współrzędne antywzorca	-0,333809123	-0,289556374	-0,236283926	1,137306953	-0,297253232	0,523619001	0,52959184	0,388431118

Po obliczeniu kolejno odległości od wzorca, od antywzorca i wskaźnika syntetycznego, tym razem uzyskujemy ranking różniący się od poprzedniego.

## METODA HELLWIGA

### Normalizacja zmiennych

track_name	danceability	energy	valence	speechiness	popularity	duration diff [s]	Duration Diff2	liveness	explicit
All Too Well	0,247979573	0,164466872	-0,409948712	-0,430321005	-1,797025873	2,152372976	2,152372976	1,064929177	-0,320499219
Blank Space (Taylor's Version)	1,312930227	0,834419375	1,641614417	0,241768319	0,381724076	-0,473742096	-0,473742096	-0,054200785	-0,320499219
Forever & Always (Piano Version)	-0,207266509	-1,573222434	-0,649741805	-0,293159918	0,091224083	0,490953237	0,490953237	-0,493417117	-0,320499219
Begin Again (Taylor's Version)	-0,426760155	-0,243785435	-0,671056747	-0,195187714	0,454349075	-0,286162448	-0,286162448	1,166157013	-0,320499219
This Is Why We Can't Have Nice	-0,036549228	1,127523595	0,240157007	-0,449915446	0,817474066	-0,956089762	-0,956089762	-0,262280225	-0,320499219
the 1 - the long pond studio ses	1,109695369	-1,693604525	-1,262546376	0,357375521	-0,417150905	-0,795307207	-0,795307207	1,019939028	3,107448946
Drops Of Jupiter - Live/2011	-1,60552233	-0,364167526	-0,655070541	-0,210863266	-0,998150891	1,77721368	1,77721368	-0,612641012	-0,320499219
"Slut!" (Taylor's Version) (From T	0,158556235	-0,845695888	-0,346003887	2,730262319	0,526974073	-0,232568263	-0,232568263	0,716255521	-0,320499219

### Obliczenie współrzędnych wzorca

Charakterystyki liczbowe	danceability	energy	valence	speechiness	popularity	Duration Diff	Duration Diff2	liveness	explicit
Współrzędne wzorca	2,646150894	1,964964224	2,808607469	-0,618427638	2,342599031	-1,03648104	-1,03648104	-0,798225378	-0,320499219

### Obliczenie odległości od wzorca dla każdego obiektu

track_name	danceability	energy	valence	speechiness	popularity	duration c	Duration D	liveness	explicit	Odległość
All Too Well	0,247979573	0,164466872	-0,409948712	-0,430321	-1,7970259	2,15237298	2,15237298	1,06493	-0,320499	7,767427084
Blank Space (Taylor's Version)	1,312930227	0,834419375	1,641614417	0,24176832	0,3817241	-0,4737421	-0,4737421	-0,0542	-0,320499	3,192079619
Forever & Always (Piano Version)	-0,20726651	-1,57322243	-0,649741805	-0,2931599	0,0912241	0,49095324	0,49095324	-0,4934	-0,320499	6,523376631
Begin Again (Taylor's Version)	-0,42676015	-0,24378544	-0,671056747	-0,1951877	0,4543491	-0,2861624	-0,28616245	1,16616	-0,320499	5,929485181
This Is Why We Can't Have Nice	-0,03654923	1,127523595	0,240157007	-0,4499154	0,8174741	-0,9560898	-0,95608976	-0,2623	-0,320499	4,141218392
the 1 - the long pond studio ses	1,109695369	-1,69360452	-1,262546376	0,35737552	-0,4171509	-0,7953072	-0,79530721	1,01994	3,107449	7,487416516
Drops Of Jupiter - Live/2011	-1,60552233	-0,36416753	-0,655070541	-0,2108633	-0,9981509	1,77721368	1,77721368	-0,6126	-0,320499	7,917926593
"Slut!" (Taylor's Version) (From T	0,158556235	-0,84569589	-0,346003887	2,73026232	0,5269741	-0,2325683	-0,23256826	0,71626	-0,320499	6,491219652

### Obliczenie średniej oraz odchylenia standardowego odległości od wzorca

Średnia Odległość od wzorca	Odchylenie
5,851386137	1,49799333

### Obliczenie zmiennej syntetycznej Q

track_name	danceability	energy	valence	speechiness	popularity	duration c	Duration D	liveness	explicit	Odległość	Q
All Too Well	0,247979573	0,164466872	-0,409948712	-0,430321	-1,7970259	2,15237298	2,15237298	1,06493	-0,320499	7,767427084	0,122064001
Blank Space (Taylor's Version)	1,312930227	0,834419375	1,641614417	0,24176832	0,3817241	-0,4737421	-0,4737421	-0,0542	-0,320499	3,192079619	0,639205933
Forever & Always (Piano Version)	-0,20726651	-1,57322243	-0,649741805	-0,2931599	0,0912241	0,49095324	0,49095324	-0,4934	-0,320499	6,523376631	0,262676415
Begin Again (Taylor's Version)	-0,42676015	-0,24378544	-0,671056747	-0,1951877	0,4543491	-0,2861624	-0,28616245	1,16616	-0,320499	5,929485181	0,329802721
This Is Why We Can't Have Nice	-0,03654923	1,127523595	0,240157007	-0,4499154	0,8174741	-0,9560898	-0,95608976	-0,2623	-0,320499	4,141218392	0,531926767
the 1 - the long pond studio ses	1,109695369	-1,69360452	-1,262546376	0,35737552	-0,4171509	-0,7953072	-0,79530721	1,01994	3,107449	7,487416516	0,153713008
Drops Of Jupiter - Live/2011	-1,60552233	-0,36416753	-0,655070541	-0,2108633	-0,9981509	1,77721368	1,77721368	-0,6126	-0,320499	7,917926593	0,105053357
"Slut!" (Taylor's Version) (From T	0,158556235	-0,84569589	-0,346003887	2,73026232	0,5269741	-0,2325683	-0,23256826	0,71626	-0,320499	6,491219652	0,266311051



## Uporządkowanie obiektów na podstawie zmiennej

Miejsce	album_name	track_name
1	Lover	Paper Rings
2	Lover	The Man
3	reputation	Look What You Made Me Do
4	Lover	ME! (feat, Brendon Urie of Panic! At
5	1989 (Taylor's Version) [Deluxe]	Blank Space (Taylor's Version)
6	1989 (Taylor's Version) [Deluxe]	Welcome To New York (Taylor's Vers
7	Red (Taylor's Version)	22 (Taylor's Version)
8	1989 (Taylor's Version) [Deluxe]	Shake It Off (Taylor's Version)
9	evermore (deluxe version)	closure
10	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)

## METODA UNITARYZACJI ZEROWANEJ

Normalizacja zmiennych - stworzenie macierzy.

track_name	zi1	zi2	zi3	zi4	zi5	zi6	zi7	zi8
All Too Well	0,548929664	0,579462103	0,315037423	0,980641258	0,123076923	0,38974359	0,629956439	1
Blank Space (Taylor's Version)	0,749235474	0,73594132	0,751644364	0,911474088	0,584615385	0,892307692	0,852228303	1
Forever & Always (Piano Version) (Taylor's Version)	0,463302752	0,173594132	0,264005443	0,966525509	0,523076923	0,707692308	0,939461633	1
Begin Again (Taylor's Version)	0,422018349	0,484107579	0,259469267	0,956442831	0,6	0,856410256	0,609851446	1
This Is Why We Can't Have Nice Things	0,495412844	0,804400978	0,453390792	0,982657794	0,676923077	0,984615385	0,893555233	1
the 1 - the long pond studio sessions	0,711009174	0,145476773	0,133590383	0,899576528	0,415384615	0,953846154	0,638891992	0
Drops Of Jupiter - Live/2011	0,20030581	0,45599022	0,262871399	0,95805606	0,292307692	0,461538462	0,963140847	1
"Slut!" (Taylor's Version) (From The Vault)	0,532110092	0,343520782	0,328645951	0,655374067	0,615384615	0,846153846	0,69920697	1
Girl At Home (Taylor's Version)	0,685015291	0,739608802	0,650714448	0,914297237	0,523076923	0,948717949	0,91477717	1
the lakes - bonus track	0,107033639	0,155256724	0,260603311	0,942730389	0,615384615	0,994871795	0,934993857	1

Obliczenie średniej wartości z  $m$  unormowanych zmiennych  $z_{ij}$  dla  $i$ -tego obiektu

track_name	zi1	zi2	zi3	zi4	zi5	zi6	zi7	zi8	Qi
All Too Well	0,548929664	0,579462103	0,315037423	0,980641258	0,123076923	0,38974359	0,629956439	1	0,570855925
Blank Space (Taylor's Version)	0,749235474	0,73594132	0,751644364	0,911474088	0,584615385	0,892307692	0,852228303	1	0,809680828
Forever & Always (Piano Version) (Taylor's Version)	0,463302752	0,173594132	0,264005443	0,966525509	0,523076923	0,707692308	0,939461633	1	0,629707338
Begin Again (Taylor's Version)	0,422018349	0,484107579	0,259469267	0,956442831	0,6	0,856410256	0,609851446	1	0,648537466
This Is Why We Can't Have Nice Things	0,495412844	0,804400978	0,453390792	0,982657794	0,676923077	0,984615385	0,893555233	1	0,786369513
the 1 - the long pond studio sessions	0,711009174	0,145476773	0,133590383	0,899576528	0,415384615	0,953846154	0,638891992	0	0,487221952
Drops Of Jupiter - Live/2011	0,20030581	0,45599022	0,262871399	0,95805606	0,292307692	0,461538462	0,963140847	1	0,574276311
"Slut!" (Taylor's Version) (From The Vault)	0,532110092	0,343520782	0,328645951	0,655374067	0,615384615	0,846153846	0,69920697	1	0,627549541
Girl At Home (Taylor's Version)	0,685015291	0,739608802	0,650714448	0,914297237	0,523076923	0,948717949	0,91477717	1	0,797025977
the lakes - bonus track	0,107033639	0,155256724	0,260603311	0,942730389	0,615384615	0,994871795	0,934993857	1	0,626359291

Uporządkowanie obiektów pod względem wartości cech syntetycznych (malejąco)

miejsce	album_name	track_name	Q i
1	Lover	Paper Rings	0,898473106
2	reputation	Look What You Made Me Do	0,847858169
3	1989 (Taylor's Version) [Deluxe]	Shake It Off (Taylor's Version)	0,844025632
4	Lover	The Man	0,840902112
5	Lover	ME! (feat, Brendon Urie of Panic! At The Disco)	0,837759111
6	Taylor Swift	Picture To Burn	0,830781703
7	evermore (deluxe version)	closure	0,83074519
8	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)	0,824774455
9	1989 (Taylor's Version) [Deluxe]	I Wish You Would (Taylor's Version)	0,817365782
10	1989 (Taylor's Version) [Deluxe]	Welcome To New York (Taylor's Version)	0,815207391

## METODA SSW

Normalizacja zmiennych - stworzenie macierzy.

track_name	zi1	zi2	zi3	zi4	zi5	zi6	zi7	zi8
All Too Well	0,248485138	0,164802176	-0,41078449	0,431198317	-1,800689538	-2,1567611	-1,067100289	0,321152633
Blank Space (Taylor's Version)	1,315606948	0,836120538	1,644961239	-0,242261222	0,382502312	0,474707932	0,054311286	0,321152633
Forever & Always (Piano Version) (Taylor's Version)	-0,207689071	-1,576429823	-0,651066459	0,293757595	0,091410065	-0,491954162	0,494423065	0,321152633
Begin Again (Taylor's Version)	-0,427630207	-0,24428245	-0,672424856	0,19558565	0,455275374	0,286745858	-1,168534502	0,321152633
This Is Why We Can't Have Nice Things	-0,036623742	1,129822321	0,240646624	0,450832706	0,819140682	0,958038979	0,262814946	0,321152633
the 1 - the long pond studio sessions	1,111957748	-1,697057341	-1,265120378	-0,358104116	-0,418001366	0,79692863	-1,022018417	-3,113784221
Drops Of Jupiter - Live/2011	-1,608795569	-0,364909968	-0,656406058	0,211293161	-1,00018586	-1,780836953	0,613890027	0,321152633
"Slut!" (Taylor's Version) (From The Vault)	0,15887949	-0,84742004	-0,346709299	-2,735828609	0,528048435	0,233042409	-0,717715778	0,321152633
Girl At Home (Taylor's Version)	0,973476291	0,851854562	1,169736901	-0,214773077	0,091410065	0,770076905	0,369884393	0,321152633
the lakes - bonus track	-2,105699618	-1,655099943	-0,667085256	0,062071806	0,528048435	1,011742428	0,471882129	0,321152633

Obliczenie średniej wartości z  $m$  unormowanych zmiennych  $z_{ij}$  dla  $i$ -tego obiektu

track_name	zi1	zi2	zi3	zi4	zi5	zi6	zi7	zi8	Qi
All Too Well	0,248485138	0,164802176	-0,41078449	0,431198317	-1,800689538	-2,1567611	-1,067100289	0,321152633	-0,533712144
Blank Space (Taylor's Version)	1,315606948	0,836120538	1,644961239	-0,242261222	0,382502312	0,474707932	0,054311286	0,321152633	0,598387708
Forever & Always (Piano Version) (Taylor's Version)	-0,207689071	-1,576429823	-0,651066459	0,293757595	0,091410065	-0,491954162	0,494423065	0,321152633	-0,215799519
Begin Again (Taylor's Version)	-0,427630207	-0,24428245	-0,672424856	0,19558565	0,455275374	0,286745858	-1,168534502	0,321152633	-0,156764062
This Is Why We Can't Have Nice Things	-0,036623742	1,129822321	0,240646624	0,450832706	0,819140682	0,958038979	0,262814946	0,321152633	0,518228143
the 1 - the long pond studio sessions	1,111957748	-1,697057341	-1,265120378	-0,358104116	-0,418001366	0,79692863	-1,022018417	-3,113784221	-0,745649933
Drops Of Jupiter - Live/2011	-1,608795569	-0,364909968	-0,656406058	0,211293161	-1,00018586	-1,780836953	0,613890027	0,321152633	-0,533099823
"Slut!" (Taylor's Version) (From The Vault)	0,15887949	-0,84742004	-0,346709299	-2,735828609	0,528048435	0,233042409	-0,717715778	0,321152633	-0,425818845
Girl At Home (Taylor's Version)	0,973476291	0,851854562	1,169736901	-0,214773077	0,091410065	0,770076905	0,369884393	0,321152633	0,541602334
the lakes - bonus track	-2,105699618	-1,655099943	-0,667085256	0,062071806	0,528048435	1,011742428	0,471882129	0,321152633	-0,254123423

Uporządkowanie obiektów pod względem wartości cech syntetycznych (malejąco)

Miejsce	album_name	track_name	Q i
1	Lover	Paper Rings	1,085937846
2	reputation	Look What You Made Me Do	0,835220313
3	Lover	The Man	0,799413106
4	evermore (deluxe version)	closure	0,746879039
5	Lover	ME! (feat, Brendon Urie of Panic! At The Disco)	0,737966044
6	1989 (Taylor's Version) [Deluxe]	Shake It Off (Taylor's Version)	0,732289833
7	Taylor Swift	Picture To Burn	0,71666001
8	1989 (Taylor's Version) [Deluxe]	I Wish You Would (Taylor's Version)	0,673427207
9	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)	0,672281582
10	1989 (Taylor's Version) [Deluxe]	Welcome To New York (Taylor's Version)	0,66650129

## METODA STRAHLA

Normalizacja zmiennych - stworzenie macierzy.

track_name	zi1	zi2	zi3	zi4	zi5	zi6	zi7	zi8
All Too Well	0,671125975	0,637513172	0,343478261	-0,06300578	0,424242424	-0,61025641	-0,394199785	0
Blank Space (Taylor's Version)	0,817168339	0,772391992	0,761956522	-0,129094412	0,727272727	-0,107692308	-0,180451128	0
Forever & Always (Piano Version) (Taylor's Version)	0,608695652	0,287671233	0,294565217	-0,076493256	0,686868687	-0,292307692	-0,096562836	0
Begin Again (Taylor's Version)	0,578595318	0,555321391	0,290217391	-0,086127168	0,737373737	-0,143589744	-0,413533835	0
This Is Why We Can't Have Nice Things	0,632107023	0,831401475	0,476086957	-0,061078998	0,787878788	-0,015384615	-0,140708915	0
the 1 - the long pond studio sessions	0,789297659	0,263435195	0,169565217	-0,140462428	0,616161616	-0,046153846	-0,385606874	-1
Drops Of Jupiter - Live/2011	0,416945373	0,531085353	0,293478261	-0,084585742	0,535353535	-0,538461538	-0,073791622	0
"Slut!" (Taylor's Version) (From The Vault)	0,658862876	0,434141201	0,356521739	-0,373795761	0,747474747	-0,153846154	-0,327604726	0
Girl At Home (Taylor's Version)	0,770345596	0,77553214	0,665217391	-0,126396917	0,686868687	-0,051282051	-0,120300752	0
the lakes - bonus track	0,348940914	0,271865121	0,291304348	-0,099229287	0,747474747	-0,005128205	-0,100859291	0

Obliczenie średniej wartości z  $m$  unormowanych zmiennych  $z_{ij}$  dla  $i$ -tego obiektu

track_name	z1	z2	z3	z4	z5	z6	z7	z8	Q <sub>i</sub>
All Too Well	0,671125975	0,637513172	0,343478261	-0,06300578	0,424242424	-0,61025641	-0,394199785	0	0,126112232
Blank Space (Taylor's Version)	0,817168339	0,772391992	0,761956522	-0,129094412	0,727272727	-0,107692308	-0,180451128	0	0,332693966
Forever & Always (Piano Version) (Taylor's Version)	0,608695652	0,287671233	0,294565217	-0,076493256	0,686868687	-0,292307692	-0,096562836	0	0,176554626
Begin Again (Taylor's Version)	0,578595318	0,555321391	0,290217391	-0,086127168	0,737373737	-0,143589744	-0,413533835	0	0,189782136
This Is Why We Can't Have Nice Things	0,632107023	0,831401475	0,476086957	-0,061078998	0,787878788	-0,015384615	-0,140708915	0	0,313787714
the 1 - the long pond studio sessions	0,789297659	0,263435195	0,169565217	-0,140462428	0,616161616	-0,046153846	-0,385606874	-1	0,033279567
Drops Of Jupiter - Live/2011	0,416945373	0,531085353	0,293478261	-0,084585742	0,535353535	-0,538461538	-0,073791622	0	0,135002953
"Slut!" (Taylor's Version) (From The Vault)	0,658862876	0,434141201	0,356521739	-0,373795761	0,747474747	-0,153846154	-0,327604726	0	0,16771924
Girl At Home (Taylor's Version)	0,770345596	0,775553214	0,665217391	-0,126396917	0,686868687	-0,051282051	-0,120300752	0	0,325000646
the lakes - bonus track	0,348940914	0,271865121	0,291304348	-0,099229287	0,747474747	-0,005128205	-0,100859291	0	0,181796043

Uporządkowanie obiektów pod względem wartości cech syntetycznych (malejąco)

Miejsce	album_name	track_name	Q <sub>i</sub>
1	Lover	Paper Rings	0,406696849
2	1989 (Taylor's Version) [Deluxe]	Shake It Off (Taylor's Version)	0,370285358
3	reputation	Look What You Made Me Do	0,359448839
4	Lover	ME! (feat, Brendon Urie of Panic! At The Disco)	0,358461945
5	evermore (deluxe version)	closure	0,35730899
6	Taylor Swift	Picture To Burn	0,356563076
7	Lover	The Man	0,353109259
8	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)	0,350402212
9	1989 (Taylor's Version) [Deluxe]	I Wish You Would (Taylor's Version)	0,343200112
10	Speak Now (Taylor's Version)	Better Than Revenge (Taylor's Version)	0,34043141

## METODA NOWAKA

Normalizacja zmiennych - stworzenie macierzy.

track_name	z1	z2	z3	z4	z5	z6	z7	z8
All Too Well	1,053376	1,054784	0,804211	-0,00049	0,629271	-3,07662	-2,066	0
Blank Space (Taylor's Version)	1,282599	1,277945	1,784025	-0,001	1,07875	-0,54293	-0,94574	0
Forever & Always (Piano Version) (Taylor's Version)	0,955387	0,47596	0,689687	-0,00059	1,01882	-1,47367	-0,50609	0
Begin Again (Taylor's Version)	0,908143	0,918795	0,679507	-0,00067	1,093733	-0,72391	-2,16733	0
This Is Why We Can't Have Nice Things	0,992133	1,375578	1,114698	-0,00047	1,168646	-0,07756	-0,73746	0
the 1 - the long pond studio sessions	1,238854	0,435861	0,397016	-0,00109	0,913941	-0,23269	-2,02097	-10,6957
Drops Of Jupiter - Live/2011	0,654423	0,878696	0,687142	-0,00066	0,79408	-2,71466	-0,38674	0
"Slut!" (Taylor's Version) (From The Vault)	1,034128	0,718299	0,834751	-0,00291	1,108716	-0,77562	-1,71698	0
Girl At Home (Taylor's Version)	1,209107	1,283175	1,557523	-0,00098	1,01882	-0,25854	-0,6305	0
the lakes - bonus track	0,547685	0,449809	0,682052	-0,00077	1,108716	-0,02585	-0,5286	0

Obliczenie średniej wartości z  $m$  unormowanych zmiennych  $z_{ij}$  dla  $i$ -tego obiektu

track_name	z1	z2	z3	z4	z5	z6	z7	z8	Q <sub>i</sub>
All Too Well	1,053376	1,054784	0,804211	-0,00049	0,629271	-3,07662	-2,066	0	-0,20018323
Blank Space (Taylor's Version)	1,282599	1,277945	1,784025	-0,001	1,07875	-0,54293	-0,94574	0	0,491704762
Forever & Always (Piano Version) (Taylor's Version)	0,955387	0,47596	0,689687	-0,00059	1,01882	-1,47367	-0,50609	0	0,144937587
Begin Again (Taylor's Version)	0,908143	0,918795	0,679507	-0,00067	1,093733	-0,72391	-2,16733	0	0,08853345
This Is Why We Can't Have Nice Things	0,992133	1,375578	1,114698	-0,00047	1,168646	-0,07756	-0,73746	0	0,47944524
the 1 - the long pond studio sessions	1,238854	0,435861	0,397016	-0,00109	0,913941	-0,23269	-2,02097	-10,6957	-1,24559051
Drops Of Jupiter - Live/2011	0,654423	0,878696	0,687142	-0,00066	0,79408	-2,71466	-0,38674	0	-0,010964956
"Slut!" (Taylor's Version) (From The Vault)	1,034128	0,718299	0,834751	-0,00291	1,108716	-0,77562	-1,71698	0	0,15004905
Girl At Home (Taylor's Version)	1,209107	1,283175	1,557523	-0,00098	1,01882	-0,25854	-0,6305	0	0,522325789
the lakes - bonus track	0,547685	0,449809	0,682052	-0,00077	1,108716	-0,02585	-0,5286	0	0,279129093

## Uporządkowanie obiektów pod względem wartości cech syntetycznych (malejąco)

Miejsce	album_name	track_name	Q i
1	Lover	Paper Rings	0,661893174
2	1989 (Taylor's Version) [Deluxe]	Shake It Off (Taylor's Version)	0,644808121
3	Speak Now (Taylor's Version)	Better Than Revenge (Taylor's Version)	0,585597386
4	reputation	Look What You Made Me Do	0,584653267
5	1989 (Taylor's Version) [Deluxe]	New Romantics (Taylor's Version)	0,56990543
6	Lover	ME! (feat, Brendon Urie of Panic! At The Disco)	0,566568232
7	Red (Taylor's Version)	Stay Stay Stay (Taylor's Version)	0,56538133
8	1989 (Taylor's Version) [Deluxe]	Welcome To New York (Taylor's Version)	0,561213607
9	1989 (Taylor's Version) [Deluxe]	I Wish You Would (Taylor's Version)	0,550582426
10	Red (Taylor's Version)	I Knew You Were Trouble (Taylor's Version)	0,540977025

## WYNIKI BADAŃ

### Top 10 dla każdej z siedmiu metod

Miejsce	HELLWIG	TOPSIS_u	TOPSIS_n	SSW	MUZ	STRAHL	NOWAK
1	Paper Rings	Paper Rings	Paper Rings	Paper Rings	Paper Rings	Paper Rings	Paper Rings
2	The Man	The Man	ME! (feat, Brendon Urie of Panic! At The Disco)	Look What You Made Me Do	Look What You Made Me Do	Shake It Off (Taylor's Version)	Shake It Off (Taylor's Version)
3	Look What You Made Me Do	Look What You Made Me Do	closure	The Man	Shake It Off (Taylor's Version)	Look What You Made Me Do	Better Than Revenge (Taylor's Version)
4	ME! (feat, Brendon Urie of Panic! At The Disco)	ME! (feat, Brendon Urie of Panic! At The Disco)	New Romantics (Taylor's Version)	closure	The Man	ME! (feat, Brendon Urie of Panic! At The Disco)	Look What You Made Me Do
5	Blank Space (Taylor's Version)	closure	The Man	ME! (feat, Brendon Urie of Panic! At The Disco)	ME! (feat, Brendon Urie of Panic! At The Disco)	closure	New Romantics (Taylor's Version)
6	Welcome To New York (Taylor's Version)	Blank Space (Taylor's Version)	Picture To Burn	Shake It Off (Taylor's Version)	Picture To Burn	Picture To Burn	ME! (feat, Brendon Urie of Panic! At The Disco)
7	22 (Taylor's Version)	Picture To Burn	I Know Places (Taylor's Version)	Picture To Burn	closure	The Man	Stay Stay Stay (Taylor's Version)
8	Shake It Off (Taylor's Version)	Welcome To New York (Taylor's Version)	Forever & Always (Taylor's Version)	I Wish You Would (Taylor's Version)	New Romantics (Taylor's Version)	New Romantics (Taylor's Version)	Welcome To New York (Taylor's Version)
9	closure	22 (Taylor's Version)	Welcome To New York (Taylor's Version)	New Romantics (Taylor's Version)	I Wish You Would (Taylor's Version)	I Wish You Would (Taylor's Version)	I Wish You Would (Taylor's Version)
10	New Romantics (Taylor's Version)	New Romantics (Taylor's Version)	22 (Taylor's Version)	Welcome To New York (Taylor's Version)	Welcome To New York (Taylor's Version)	Better Than Revenge (Taylor's Version)	I Knew You Were Trouble (Taylor's Version)

## ANALIZA WYNIKÓW

Piosenki, które we wszystkich rankingach znajdowały się na pierwszych 10 miejscach:

Utwór	Najwyższe miejsce	Liczba wystąpień na podium
Paper Rings	1	7
Shake It Off (Taylor's Version)	2	3



ME! (feat. Brendon Urie of Panic! At The Disco)	2	1
New Romantics (Taylor's Version)	4	0

Wyniki pokazują, że utwór "Paper Rings" konsekwentnie zajmował pierwsze miejsce we wszystkich rankingach i pojawił się na podium 7 razy.

Należy również zauważyć, że na podium pojawiło się kilka innych piosenek Taylor Swift, takich jak "Look What You Made Me Do", "The Man" i "Better Than Revenge (Taylor's Version)", "closure".

- Look What You Made Me Do (5 wystąpień)
- The Man (3 wystąpienia)
- Better Than Revenge (Taylor's Version) (1 wystąpienie)
- closure (1 wystąpienie)

### Macierz korelacji wyników

Macierz korelacji jest użytecznym narzędziem, które pozwala porównać i znaleźć powiązanie między otrzymanymi rankingami.<sup>17</sup>

Do obliczenia współczynnika korelacji użyto wbudowanej funkcji excel'a, która używa współczynnika korelacji Pearsona do obliczenia zależności liniowej między dwoma zestawami danych.

$$Correl(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

gdzie

- X to tablica1 - pierwszy zakres wartości komórek,
- Y to tablica2 - drugi zakres wartości komórek,
- $\bar{x}$  i  $\bar{y}$  są wynikami funkcji ŚREDNIA(tablica1) i ŚREDNIA(tablica2).

Współczynnik ten może przyjmować wartości od -1 do 1, gdzie:

- -1 oznacza idealną korelację ujemną (gdy jedna zmienna rośnie, druga maleje).
- 0 oznacza brak korelacji (zmienne nie są ze sobą powiązane).
- 1 oznacza idealną korelację dodatnią (obie zmienne rosną lub maleją razem).

<sup>17</sup> Bąk, A. (2016). *Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza*. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 426, 22-3

	TOPSIS (s)	TOPSIS (n)	Hellwig	MUS	SSW	STRAHL	NOWAK
TOPSIS (s)	1,00000	0,96758	0,99425	0,99118	0,99300	0,98905	0,86307
TOPSIS (n)	0,96758	1,00000	0,97858	0,98097	0,97704	0,98598	0,91622
Hellwig	0,99425	0,97858	1,00000	0,99461	0,99510	0,99480	0,89920
MUS	0,99118	0,98097	0,99461	1,00000	0,99901	0,99891	0,89905
SSW	0,99300	0,97704	0,99510	0,99901	1,00000	0,99783	0,89204
STRAHL	0,98905	0,98598	0,99480	0,99891	0,99783	1,00000	0,90927
NOWAK	0,86307	0,91622	0,89905	0,89905	0,89204	0,90927	1,00000

Macierz korelacji ilustruje silną zależność między różnymi metodami rankingu piosenek Taylor Swift, gdzie wszystkie współczynniki korelacji przekraczają 0,86. To sugeruje, że wybór konkretnej metody porządkowania liniowego nie ma znaczącego wpływu na ostateczny ranking.

Niemniej jednak, warto zauważyć, że metoda Nowaka wykazuje największe rozbieżności w stosunku do pozostałych metod. Współczynniki korelacji między metodą Nowaka a innymi metodami są niższe, oscylując między 0,86 a 0,92, podczas gdy współczynniki korelacji między innymi metodami są wyższe niż 0,96. To sugeruje, że metoda Nowaka może generować wyniki bardziej odstające od ogólnej tendencji rankingu.

Warto zwrócić uwagę na sposób wcześniejszej obróbki danych na przykładzie metody TOPSIS. Metoda ta została użyta dwukrotnie, raz z zastosowaniem standaryzacji danych, a drugi raz z zastosowaniem normalizacji danych poprzez przekształcenie ilorazowe. Współczynnik korelacji między tymi metodami wynosi 0,96, co jest najniższym wynikiem, nie biorąc pod uwagę metody Nowaka. Co za tym idzie, nawet metody z tej samej rodziny mogą zwrócić wyniki, które ostatecznie będą się od siebie różnić.

Najbardziej zbliżone do siebie wyniki zwróciły metody MUS, SSW oraz Strahla, gdzie współczynniki korelacji przekraczają 0,99783.

## PODSUMOWANIE

Analiza wykazała, że choć wszystkie metody porządkowania liniowego są silnie skorelowane ze sobą, istnieją subtelne różnice w ostatecznym rankingu piosenek Taylor Swift. Metoda Nowaka wykazywała największe odstępstwa od reszty metod, co sugeruje, że może generować wyniki bardziej odstające od ogólnej tendencji rankingu. Analizując każdy z rankingów można dostrzec pewne powtarzające się zależności, przykładowo silne znaczenie cechy Explicit w każdym z zestawień - dla wszystkich metod utwory uznane jako ofensywne są skupione wokół ostatnich miejsc rankingu, a w metodzie NOWAK wszystkie z 22 utworów ofensywnych zajmują ostatnie 22 miejsca tabeli.

Chociaż popularność utworów jest zależna od aktualnie panujących trendów i wydarzeń, to wysokie pozycje hitów takich jak *Shake It Off (Taylor's Version)*, *ME!* z gościnnym występem niezwykle popularnego Brendona'a Urie, *Look What You Made Me Do* okupującego niegdyś radiostację, a także bardzo tanecznego i energicznego *Paper Rings*, pokazują że wykazane w badaniu zależności potrafią wyłonić prawdziwe hity spośród ponad 200 przeróżnych utworów, które powstały na przestrzeni prawie 20 lat.

Wnioski te mogą być istotne dla osób zajmujących się analizą danych dotyczących muzyki lub podejmujących decyzje w zakresie marketingu muzycznego, a nawet dla artystów pragnących podbić scenę.

## BIBLIOGRAFIA

- Bąk, A. (2016). Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 426, 22-3.
- Chakkenchath, A.-M., Spargo, O. (2020). Explicit and sexual songs leave lasting impacts on the youth of today. *The Wildcat Tribune*.
- Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologii gospodarstw rolnych. *Roczniki Nauk Rolniczych*.
- Hwang, C.-L., Yoon, K. (1981). Multiple Attribute Decision Making. *Lecture Notes in Economics and Mathematical Systems* 186, Springer.
- Kościółek, M. (2015). Wykorzystanie analizy wielowymiarowej do badania zróżnicowanego potencjału innowacyjnego Polski. *Metody Ilościowe w Badaniach Ekonomicznych*, Tom XVI/3, 194-201.
- Kukuła, K. (2012). Propozycja budowy rankingu obiektów w wykorzystaniem cech ilościowych oraz jakościowych. *Metody Ilościowe w Badaniach Ekonomicznych*, Tom XIII/1, 5-16.
- Kukuła, K., Luty, L. (2018). O Wyborze metody porządkowania liniowego do oceny gospodarki odpadami w Polsce w ujęciu przestrzennym. *Problemy Rolnictwa Światowego* tom 18, zeszyt 2, 183-192.
- Lindner, J. (2024). Statistics About The Average Song Length. *Gitnux Market Data Report 2024*.
- Nowak, E. (1977). Syntetyczne mierniki plonów w krajach europejskich (Synthetic measure of crops in European countries). *Wiadomości Statystyczne*, 10, 19-22.
- Perkal, J. (1960). On the analysis of a set of characteristics. *Applicationes Mathematicae* V.
- Strahl, D. (1978). Propozycja konstrukcji miary syntetycznej (A proposal for the construction of a synthetic measure). *Przegląd Statystyczny*, 25(2), 205-215.
- Tauberg, M. (2018). Anatomy of a Hit Song (2000-2018). *medium.com*.
- Trzęsiok, J. (2015). "O odporności na obserwacje odstające wybranych nieparametrycznych modeli regresji." *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, ISSN 2083-8611, Nr 227 · 2015.
- Spotify Web API Documentation, References/Tracks/Get Track's Audio Features.

## **ZAŁĄCZNIKI**

### **Załącznik 1:** Taylor.ipynb

Kod Jupyter Notebook wykorzystany do utworzenia graficznych reprezentacji danych

### **Załącznik 2:** Taylor.xlsx

Arkusz Microsoft Excel na którym dokonane zostały wszystkie z obliczeń wykorzystanych w badaniach

### **Załącznik 3:** images.zip

Pliki bitmapowe zawierające graficzne reprezentacje danych wykorzystane w sekcji *Wstępna analiza danych*

### **Załącznik 4:** taylorswift-Features.csv

Jeden z plików zawierających oryginalne dane, w tym konkretnym cechy opisujące charakter muzyczny utworów. Znajdują się w nim między innymi wykorzystane przez nas cechy (danceability, energy, valence, speechiness, liveness)

### **Załącznik 5:** taylorswift-Tracks.csv

Drugi z plików zawierających oryginalne dane, w nim znajdują się trzy wykorzystane przez nas cechy - długość utworu, popularność i oznaczenie utworu jako ofensywne lub nie