

# Analiza Preferencji Klientów

Weronika Kłujso, Michał Korzeniewski, Miłosz Malinowski, Piotr Misiejuk

7 kwietnia 2025

## Spis treści

<b>1</b>	<b>Wprowadzenie</b>	<b>3</b>
<b>2</b>	<b>Cel i Zakres Badania</b>	<b>3</b>
<b>3</b>	<b>Przegląd Podobnych Prac</b>	<b>4</b>
<b>4</b>	<b>Słowa Kluczowe</b>	<b>4</b>
<b>5</b>	<b>Zmienne wybrane do analizy oraz ich wizualizacja</b>	<b>5</b>
<b>6</b>	<b>Skalowanie Danych</b>	<b>6</b>
6.1	Standaryzacja . . . . .	7
6.2	Normalizacja Min-Max . . . . .	7
<b>7</b>	<b>Klasteryzacja dla danych po standaryzacji</b>	<b>7</b>
7.1	Klasteryzacja k-średnich . . . . .	9
7.1.1	Wnioski . . . . .	10
7.2	Klasteryzacja GMM . . . . .	14
7.2.1	Wnioski . . . . .	15
<b>8</b>	<b>Klasteryzacja dla danych po normalizacji min-max</b>	<b>19</b>
8.1	Klasteryzacja k-średnich . . . . .	20
8.1.1	Wnioski . . . . .	20
8.2	Klasteryzacja GMM . . . . .	20
8.2.1	Wnioski . . . . .	20

<b>9 Wnioski</b>	<b>20</b>
<b>10 Rekomendacje marketingowe</b>	<b>21</b>
<b>11 Podsumowanie</b>	<b>21</b>

# 1 Wprowadzenie

Współczesne przedsiębiorstwa gromadzą ogromne ilości danych o swoich klientach i stoją przed wyzwaniem ich efektywnego wykorzystania. Jednym z kluczowych zastosowań analizy danych w marketingu jest segmentacja klientów – podział bazy konsumentów na grupy o podobnych cechach. Segmentacja może opierać się na kryteriach demograficznych, behawioralnych czy psychograficznych. Odpowiednio przeprowadzona segmentacja umożliwia skierowanie właściwych ofert i komunikatów do odpowiednich odbiorców, zwiększając skuteczność kampanii marketingowych.

Analiza preferencji klientów stanowi szczegółowe badanie charakterystyki konsumentów. Pozwala ona lepiej zrozumieć zróżnicowanie bazy klientów oraz dostosować produkty i usługi do ich indywidualnych potrzeb. Firma może zidentyfikować segmenty najbardziej skłonne do zakupu danego produktu i skupić na nich swoje działania marketingowe. Tego rodzaju podejście przyczynia się do optymalizacji kosztów promocji oraz zwiększenia zadowolenia klientów poprzez bardziej spersonalizowaną ofertę.

W niniejszej pracy wykorzystano publicznie dostępny zbiór danych Customer Personality Analysis pochodzący z serwisu Kaggle. Dane te dotyczą kampanii marketingowej pewnej firmy i zawierają informacje o klientach, takie jak dane demograficzne oraz najchętniej kupowane produkty. Zbiór obejmuje 2240 klientów, których dane posłużyły do eksploracyjnej analizy danych oraz przeprowadzenia segmentacji z wykorzystaniem metod klasteryzacji. W dalszych rozdziałach przedstawiono proces przygotowania danych, zastosowane metody grupowania klientów oraz analizę otrzymanych wyników.

## 2 Cel i Zakres Badania

Celem projektu jest zastosowanie metod analizy skupień do segmentacji klientów na podstawie ich cech i zachowań zakupowych. Chcemy ustalić, czy możliwe jest wyodrębnienie sensownych grup konsumentów oraz jakimi cechami się one charakteryzują. Projekt zakłada również porównanie kilku metod klasteryzacji (k-średnich oraz GMM) oraz ocenę ich skuteczności.

Zakres badania obejmuje:

- wczytanie i przygotowanie danych,

- skalowanie danych przy użyciu różnych metod (standaryzacja, min-max),
- zastosowanie dwóch metod klasteryzacji: k-średnich oraz GMM,
- ocenę i interpretację otrzymanych klastrów,
- podsumowanie wyników oraz możliwe rekomendacje marketingowe.

### 3 Przegląd Podobnych Prac

Segmentacja klientów za pomocą analizy skupień jest szeroko opisywana w literaturze i stosowana w wielu sektorach. Wśród podobnych badań można wymienić:

- **Customer Segmentation Analysis for Amazon Data** – Karthik Ramireddy. Praca magisterska analizująca dane klientów Amazona i ich segmentację z wykorzystaniem metod nienadzorowanych. Link do pracy
- **Customer personality analysis and clustering for targeted marketing** – D. Acheme, E. Enoyoze. Badanie z wykorzystaniem klasteryzacji k-średnich w celu lepszego targetowania kampanii marketingowych. Link do pracy
- **Customer segmentation: The concepts of trust, commitment and relationships** – B. Hultén. Autor analizuje znaczenie relacji i zaufania jako elementów segmentacji. Link do pracy
- **The role of customer personality in satisfaction, attitude-to-brand & loyalty in mobile services** – T. A. Smith. Badanie powiązań między osobowością klientów a ich lojalnością wobec marki. Link do pracy

### 4 Słowa Kluczowe

**POL:** analiza osobowości klientów, strategia marketingowa, analiza danych, analiza skupień, klastrowanie

**ENG:** customer behaviour analysis, marketing strategy, data analysis, cluster analysis, clustering

## 5 Zmienne wybrane do analizy oraz ich wizualizacja

Zmienne podzielono następująco:

- **Ciągłe:**

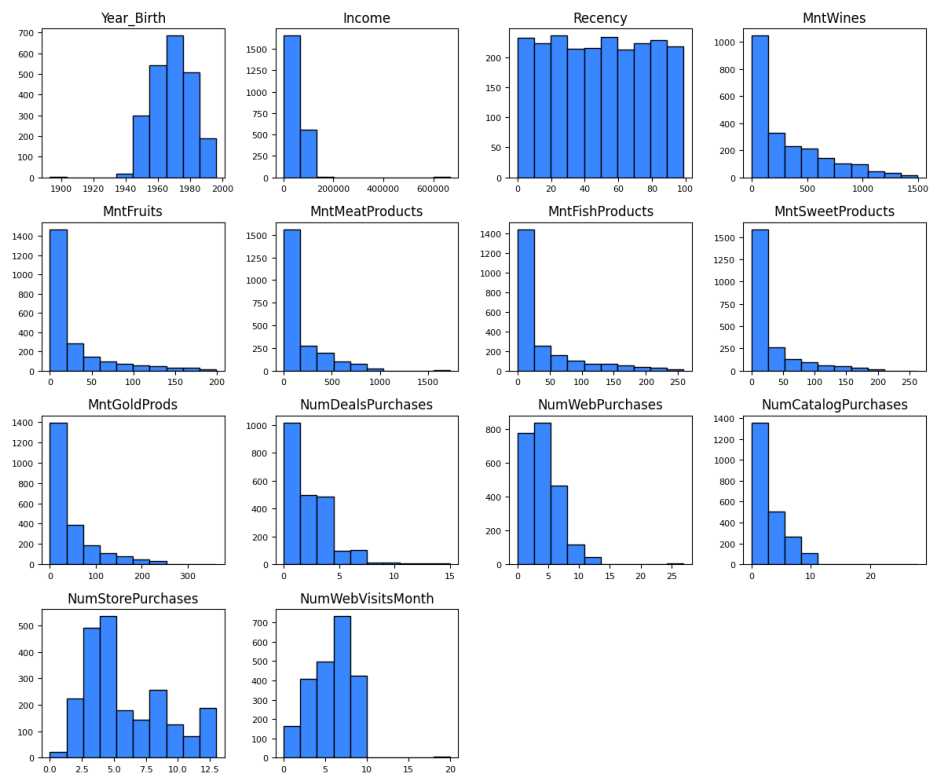
- Rok urodzenia klienta (YearBirth),
- Roczny dochód gospodarstwa domowego klienta (Income),
- Liczba dni od ostatniego zakupu klienta (Recency),
- Kwota wydana na wino w ostatnich 2 latach (MntWines),
- Kwota wydana na owoce w ostatnich 2 latach (MntFruits),
- Kwota wydana na mięso w ostatnich 2 latach (MntMeatProducts),
- Kwota wydana na ryby w ostatnich 2 latach (MntFishProducts),
- Kwota wydana na słodycze w ostatnich 2 latach (MntSweetProducts),
- Kwota wydana na złoto w ostatnich 2 latach (MntGoldProds),
- Liczba zakupów dokonanych z rabatem (NumDealsPurchases),
- Liczba zakupów dokonanych przez stronę internetową (NumWebPurchases),
- Liczba zakupów dokonanych za pomocą katalogu (NumCatalogPurchases),
- Liczba zakupów dokonanych w sklepie (NumStorePurchases):,
- Liczba wizyt na stronie internetowej w ostatnim miesiącu (NumWebVisitsMonth).

- **Jakościowe:**

- Liczba dzieci w gospodarstwie domowym klienta (Kidhome),
- Liczba nastolatków w gospodarstwie domowym klienta (Teenhome),
- Skarga (Complain): wartość = 1, jeśli klient składał skargę w ciągu ostatnich 2 lat, 0 w przeciwnym razie,

- Akceptacja oferty w kampaniach (AcceptedCmp1 - AcceptedCmp5): wartość = 1, jeśli klient zaakceptował ofertę w kampaniach marketingowych,
- Reakcja na ofertę w ostatniej kampanii (Response): wartość = 1, jeśli klient zaakceptował ofertę w ostatniej kampanii.

Stworzono wykresy rozkładu zmiennych. Dla zmiennych ciągłych powstały histogramy oraz wykresy pudełkowe.

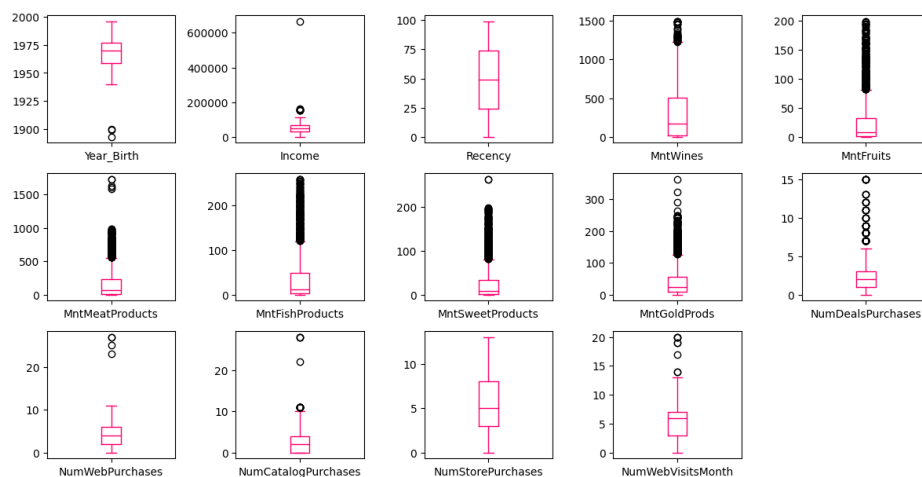


Rys. 1. Histogramy rozkładu zmiennych ciągłych.

Dla zmiennych jakościowych powstały wykresy kołowe.

## 6 Skalowanie Danych

Skalowanie danych jest jednym z kluczowych etapów przygotowania danych do analizy. Jego celem jest zapewnienie, że wszystkie cechy będą miały tę



Rys. 2. Wykresy pudełkowe dla zmiennych ciągłych.

samą wagę i przyczynią się równomiernie do wyników modelu.

W tym projekcie zastosowano dwie popularne metody skalowania: **standaryzację** oraz **skalowanie min-max**.

## 6.1 Standaryzacja

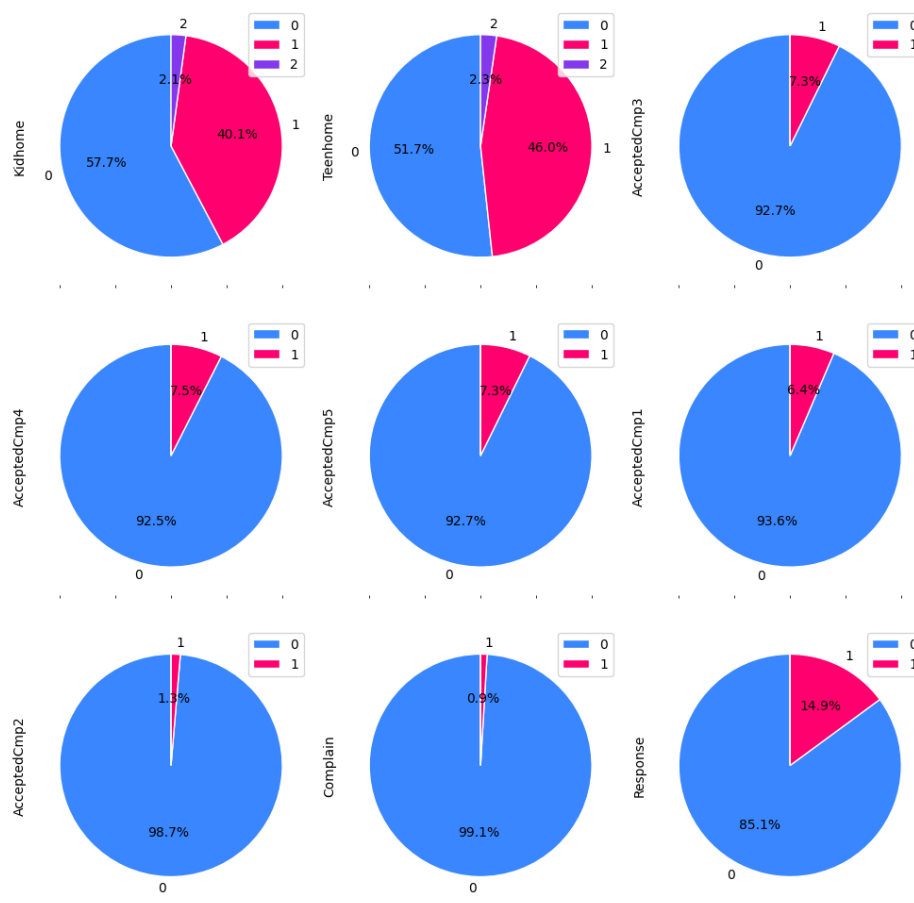
Standaryzacja (inaczej normalizacja z wykorzystaniem średniej i odchylenia standardowego) polega na przekształceniu danych w taki sposób, że każda cecha ma średnią wartość 0 oraz odchylenie standardowe 1. Dzięki temu zmienne z różnych zakresów wartości stają się porównywalne.

## 6.2 Normalizacja Min-Max

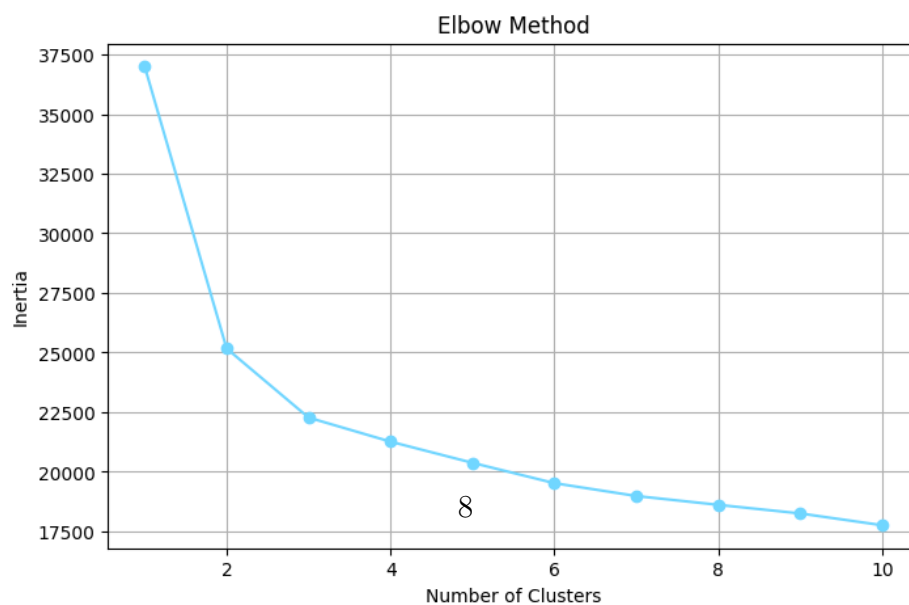
Metoda min-max polega na przekształceniu danych w taki sposób, aby mieściły się one w określonym przedziale, zazwyczaj od 0 do 1. Jest użyteczna, gdy zależy nam na zachowaniu rozkładu wartości w tych samych proporcjach. Jest mniej oporna na wartości odstające.

# 7 Klasteryzacja dla danych po standaryzacji

Zastosowano metodę łokcia do wyznaczenia optymalnej liczby klastrów.



Rys. 3. Wykresy kołowe dla zmiennych jakościowych.



Rys. 4. Metoda łokcia w celu wyznaczenia optymalnej liczby klastrów.

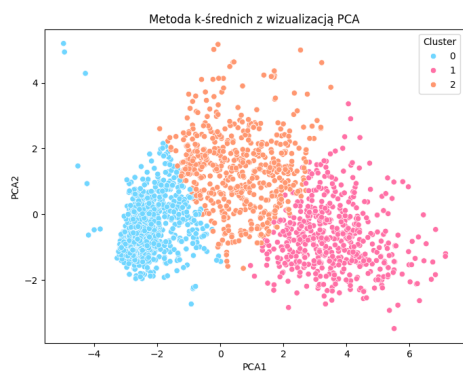


Na podstawie otrzymanych wyników zdecydowano dokonać klasteryzacji dla 3 oraz 4 klastrów.

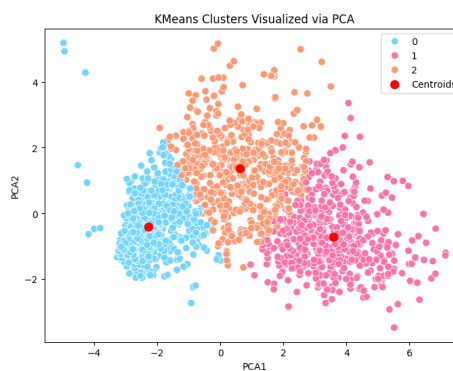
## 7.1 Klasteryzacja k-średnich

Jako pierwszą metodę wybrano klasteryzację k-średnich. Jest to jeden z algorytmów najczęściej stosowanych w analizie segmentacji klientów, ponieważ:

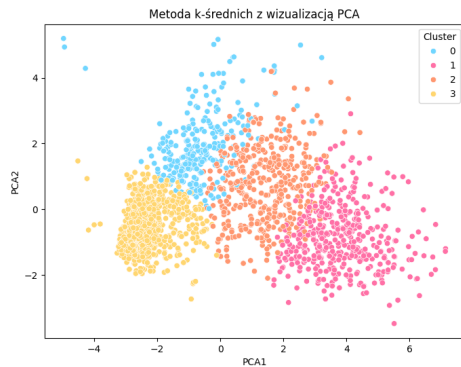
- Ma niską złożoność obliczeniową, co sprawia, że działa efektywnie nawet w przypadku dużych zbiorów danych, a z takim pracujemy przy projekcie.
- Algorytm jest szczególnie skuteczny w przypadku zmiennych ciągłych, jak w naszym przypadku, gdzie analizujemy dochody, wydatki oraz liczbę zakupów.



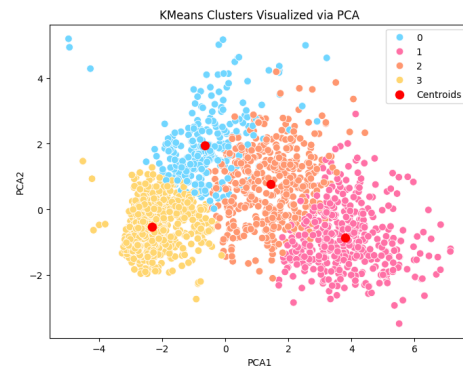
Rys. 5. Wyniki klasteryzacji k-średnich dla 3 klastrów.



Rys. 6. Wizualizacja wyników klasteryzacji k-średnich dla 3 klastrów.



Rys. 7. Wyniki klasteryzacji k-średnich dla 4 klastrów.



Rys. 8. Wizualizacja wyników klasteryzacji k-średnich dla 4 klastrów.

### 7.1.1 Wnioski

#### • Dla 3 klastrów

##### – **Klaster 0: Oszczędne rodziny z dziećmi (1047 osób)**

- \* Niski dochód, najniższy w porównaniu do innych klastrów
- \* Duża liczba dzieci (średnio 1,23 dziecka)
- \* Aktywność zakupowa: kupują zarówno stacjonarnie, jak i on-line
- \* Reagują umiarkowanie na promocje i nie zależą od kampanii reklamowych
- \* Edukacja wyższa (licencjat)
- \* Składają najwięcej reklamacji

Podsumowanie: Są to rodziny z dziećmi, które preferują oszczędne zakupy i są aktywne w poszukiwaniach promocji.

##### – **Klaster 1: Bogate pary kupujące towary wyższej klasy (564 osoby)**

- \* Najwyższy dochód w porównaniu do innych klastrów
- \* Niska liczba dzieci i mała liczba nastolatków
- \* Wydają najwięcej na wino, mięso i ryby
- \* Najczęściej kupują stacjonarnie i przez katalog

- \* Najlepiej reagują na kampanie marketingowe
- \* Edukacja wyższa (licencjat)
- \* Rzadko składają reklamacje

Podsumowanie: Klienci premium, skłonni do zakupu drogich produktów, z lojalnością wobec marek wysokiej jakości.

– **Klaster 2: Średniozamożne starsze małżeństwa (605 osób)**

- \* Średni dochód i średnie wydatki
- \* Najwięcej nastolatków, umiarkowana liczba dzieci
- \* Wydają głównie na wino, mięso i złoto
- \* Duża aktywność zakupowa, zarówno online, jak i stacjonarnie
- \* Skłonność do składania umiarkowanej liczby reklamacji
- \* Najwyższy poziom wykształcenia (magister)

Podsumowanie: Starsze małżeństwa z preferencjami do bardziej premium produktów, z wykształceniem wyższym.

Dla 3 klastrów (raczej poprawne): Klaster 0: - niski (najniższy) dochód - najwięcej dzieci (średnio 1,23 dziecka) - ważne są dla nich promocje - kupują zarówno stacjonarnie jak i online - najwięcej wizyt na stronie (przeglądają, ale stosunkowo mało kupują) - ich zakupy nie zależą od kampanii reklamowych - są stosunkowo najmłodszą grupą - większość jest w związku - niskie wydatki - są grupą, która składa najwięcej reklamacji - kupują głównie produkty podstawowe - edukacja wyższa (licencjat) - liczność: 1047 Podsumowanie: Oszczędne rodziny z dziećmi

Klaster 1: - najwyższy dochód - bardzo mała liczba dzieci - najwyższe wydatki, największa liczba transakcji - najwięcej wydają na wino, mięso i ryby - najczęściej kupują stacjonarnie i przez katalog - rzadko odwiedzają stronę internetową - najlepiej reagują na kampanie reklamowe - większość z nich jest w związku - składają najmniej reklamacji - średni wiek 53 lata - edukacja wyższa (licencjat) - liczność: 564 Podsumowanie: Bogate pary kupujące towary wyższej klasy

Klaster 2: - średni dochód i średnie wydatki - największa liczba nastolatków, umiarkowana dzieci - najkrótszy średni czas od ostatnich zakupów - najwięcej wydają na wino, mięso i złoto - większość w związku - najstarsza grupa - najwięcej zakupów dokonują stacjonarnie i online - średnia skłonność do reklamacji - grupa z najwyższym poziomem

edukacji - wyższy (magister) - liczność: 605 Podsumowanie: Średnioza-  
możne starsze małżeństwa

- **Dla 4 klastrów**

- **Klaster 0: Stabilne rodziny z nastolatkami (289 osób)**

- \* Dochód średni, około 49 261
    - \* Umiarkowana liczba dzieci i najwięcej nastolatków
    - \* Średnie wydatki, najwięcej na wino i mięso
    - \* Często korzystają z promocji
    - \* Równowaga między zakupami online i stacjonarnymi
    - \* Umiarkowana reakcja na kampanie
    - \* Średni wiek 55 lat
    - \* Większość w związku

Podsumowanie: Stabilne rodziny z dziećmi, umiarkowane dochody,  
wrażliwe na promocje.

- **Klaster 1: Zamożni koneserzy (473 osoby)**

- \* Najwyższy dochód ( 78 511)
    - \* Brak dzieci, bardzo mało nastolatków
    - \* Najwyższe wydatki, szczególnie na wino i produkty premium
    - \* Preferują zakupy katalogowe i stacjonarne
    - \* Najlepsza reakcja na kampanie
    - \* Średni wiek 52 lata
    - \* Większość w związku

Podsumowanie: Zamożni klienci preferujący wysokiej jakości pro-  
dukty, lojalni wobec marek.

- **Klaster 2: Dojrzałe pary lubiące wino (468 osób)**

- \* Wyższy średni dochód ( 64 346)
    - \* Mało dzieci, sporo nastolatków
    - \* Wysokie wydatki na wino i mięso
    - \* Częste zakupy stacjonarne i online
    - \* Umiarkowana reakcja na kampanie
    - \* Najstarsza grupa (średni wiek 56)

- \* Wykształcenie magisterskie

Podsumowanie: Dojrzałe pary z wykształceniem wyższym, o ugruntowanych preferencjach zakupowych, lubiące wino i tradycyjne zakupy.

– **Klaster 3: Oszczędne rodziny z małymi dziećmi (986 osób)**

- \* Najniższy dochód ( 34 781)
- \* Najwięcej dzieci, umiarkowana liczba nastolatków
- \* Najniższe wydatki, głównie zakupy stacjonarne
- \* Najwięcej wizyt na stronie, ale mało zakupów online
- \* Brak reakcji na kampanie
- \* Najmłodsza grupa (średni wiek 50)
- \* Wykształcenie wyższe

Podsumowanie: Rodziny z małymi dziećmi, skoncentrowane na podstawowych potrzebach, z niewielkim budżetem.

Dla 4 klastrów (3 chyba lepsze wg. metody łokcia): Klaster 0: "Stabilne rodziny z nastolatkami"(289 osób) Dochód: Średni ( 49 261) Gospodarstwo domowe: Umiarkowana liczba dzieci (0.57) Najwięcej nastolatków (0.91) Zachowania zakupowe: Średnie wydatki (484), głównie na wino (302) i mięso (91) Często korzystają z promocji (5.5 transakcji) Równowaga między zakupami online (6.02) i stacjonarnymi (5.92) Reakcja na marketing: Umiarkowana reakcja na kampanie (17.6) Demografia: Średni wiek (55 lat) Wykształcenie wyższe (licencjat) Większość w związkach Podsumowanie: Rodziny z nastolatkami, stabilne dochody, umiarkowane wydatki. Wrażliwi na promocje.

Klaster 1: "Zamożni koneserzy"(473 osoby) Dochód: Najwyższy ( 78 511) Gospodarstwo domowe: Praktycznie brak dzieci (0.03) Niewiele nastolatków (0.18) Zachowania zakupowe: Najwyższe wydatki (1451), głównie na wino (607), mięso (499) i produkty premium Preferują zakupy katalogowe (6.3) i stacjonarne (8.29) Rzadko odwiedzają stronę (2.7 wizyt/miesiąc) Reakcja na marketing: Najlepsza reakcja na kampanie (28.8) Demografia: Nieco młodszy (52 lata) Wykształcenie wyższe (licencjat) Większość w związkach Podsumowanie: Zamożni klienci premium, lojalni wobec marek wysokiej jakości.

Klaster 2: "Dojrzałe pary lubiące wino"(468 osób) Dochód: Wyższy średni ( 64 346) Gospodarstwo domowe: Mało dzieci (0.1) Sporo nastolatków (0.72) Zachowania zakupowe: Wysokie wydatki (920), szczególnie na wino (566) i

mięso (186) Częste zakupy stacjonarne (8.74) i online (6.28) Umiarkowane korzystanie z katalogów (3.93) Reakcja na marketing: Najlepiej reagują na Cmp4 (15) Umiarkowany response rate (13) Demografia: Najstarsi (średnio 56 lat) Wykształcenie magisterskie Większość w związkach Podsumowanie: Dojrzałe pary o ugruntowanych preferencjach, lubiące wino i tradycyjne zakupy.

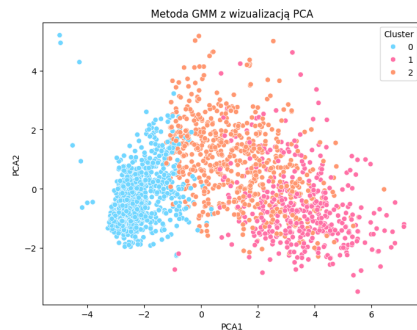
Klaster 3: "Oszczędne rodziny z małymi dziećmi" (986 osób) Dochód: Najniższy (34 781) Gospodarstwo domowe: Najwięcej dzieci (0.77) Umiarkowana liczba nastolatków (0.44) Zachowania zakupowe: Najniższe wydatki (90) Głównie zakupy stacjonarne (3.18) Najwięcej wizyt na stronie (6.35), ale mało zakupów online Reakcja na marketing: Praktycznie brak reakcji na kampanie Najwięcej reklamacji (1.2) Demografia: Najmłodsi (średnio 50 lat) Wykształcenie wyższe (licencjat) Większość w związkach Podsumowanie: Rodziny z małymi dziećmi, niskie wydatki, skupione na podstawowych potrzebach.

Według metody łokcia oraz otrzymanych wyników preferowane jest wykorzystanie 3 klastrów.

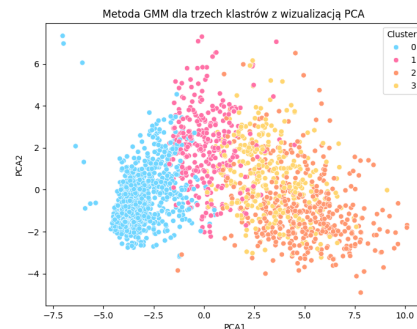
## 7.2 Klasteryzacja GMM

Metoda mieszanki gaussowskiej (GMM) jest probabilistycznym podejściem do klasteryzacji, które zakłada, że dane pochodzą z mieszanki rozkładów normalnych (Gaussa). Każdy klaster jest reprezentowany przez rozkład normalny, który ma określoną średnią oraz kowariancję. W przeciwieństwie do algorytmu k-średnich, który przypisuje punktom jednoznacznie przynależność do jednego klastra, GMM przypisuje prawdopodobieństwo przynależności punktu do każdego z klastrów.

- GMM jest bardziej elastyczne niż k-średnie, ponieważ może wykrywać klastry o nieregularnych kształtach, takich jak elipsy.
- Każdemu punktowi przypisywane jest prawdopodobieństwo przynależności do każdego klastra.
- Liczba klastrów ( $k$ ) musi być określona przed rozpoczęciem procesu klasteryzacji.



Rys. 9. Wizualizacja wyników klasteryzacji GMM dla 3 klastrów.



Rys. 10. Wizualizacja wyników klasteryzacji GMM dla 4 klastrów (błąd w podpisie!!!).

### 7.2.1 Wnioski

#### • Dla 3 klastrów

##### – Klaster 0 (1096 osób)

- \* Największe zarobki ze wszystkich
- \* Średnio najwięcej małych dzieci, i dzieci w ogóle
- \* Mają kilka razy niższe zarobki ale nawet kilkanaście razy niższe wydatki
- \* W miarę zainteresowani promocjami, ale nie najbardziej ze wszystkich grup.
- \* Kupują głównie przez internet i stacjonarnie, praktycznie nie korzystają z katalogu, ale bardzo często odwiedzają stronę internetową
- \* Co ciekawe, rzadko korzystają z kampanii promocyjnych, oraz rzadko narzekają
- \* Najmłodsza grupa ze wszystkich
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku
- \* Składają średnio najwięcej skarg
- \* Wydają relatywnie najmniej w sklepie, średnio 0.3 procent zarobków

**Podsumowanie:** Skrajnie oszczędni wcześni rodzice trudni do manipulacji poprzez akcje promocyjne oraz skorzy do skarżenia się na niedogodności

– **Klaster 1 (446 osób)**

- \* Najrzadziej mają dzieci, a jeżeli już to starsze
- \* Wydają najwięcej ze wszystkich grup na wszystkie produkty, ulubionym produktem jest wino, za nim ryby (w co może wliczać się kawior) i złoto
- \* Kupują stacjonarnie, z taką samą częstotliwością korzystają z katalogu i sklepu internetowego.
- \* Bardzo mało wchodzi na stronę internetową
- \* Bardzo chętni do brania udziału w kampaniach promocyjnych
- \* Wydają zdecydowanie najwięcej ze wszystkich grup, zarówno bezwzględnie jak i procentowo (prawie 2 procent przychodów)
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku

**Podsumowanie:** bogaci i/lub rozrzutni zblazowani bezdzietni (lub z już dorosłymi dziećmi) łatwi do zmanipulowania przez kampanie promocyjne

– **Klaster 2**

- \* Średnio najwięcej nastoletnich dzieci
- \* Przychody, wydatki na poszczególne produkty jak i chęć korzystania z promocji pomiędzy pozostałymi klastrami
- \* Wydatki ogólne zarówno bezwzględnie jak i procentowo pomiędzy dwoma pozostałymi klastrami
- \* Wydają najwięcej na wino, mięso i złoto
- \* Średnio najstarsi
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku

**Podsumowanie:** Nieco starsi, ale poza tym bardzo "średni"klienci. Brak znaczących znaków szczególnych

**Podsumowanie całości:** Sklep powinien najbardziej skupić się i obserwować zachowania klientów z klastru numer 1, równocześnie mając



na uwagę posiadanie stabilnej bazy klientów z klastru nr.2, zaś klienci z klastru numer 0 są dosyć trudni w utrzymaniu i mało dochodowi więc można ich w dużej mierze ignorować

- **Dla 4 klastrów**

- **Klaster 0 (1096 osób):**

- \* Najmniejsze zarobki ze wszystkich
- \* Średnio najwięcej małych dzieci, i dzieci w ogóle
- \* Mają kilka razy niższe zarobki ale nawet kilkanaście razy niższe wydatki
- \* W miarę zainteresowani promocjami, ale nie najbardziej ze wszystkich grup.
- \* Kupują głównie przez internet i stacjonarnie, praktycznie nie korzystają z katalogu, ale bardzo często odwiedzają stronę internetową
- \* Co ciekawe, rzadko korzystają z kampanii promocyjnych, oraz rzadko narzekają
- \* Najmłodsza grupa ze wszystkich
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku
- \* Składają średnio najwięcej skarg
- \* Wydają relatywnie najmniej w sklepie, średnio 0.3 procent zarobków

**Podsumowanie:** Skrajnie oszczędni wcześni rodzice trudni do manipulacji poprzez akcje promocyjne oraz skorzy do skarżenia się na niedogodności (wszystkie dane dokładnie tak samo jak w wypadku trzech klastrów)

- **Klaster 1 (311 osób):**

- \* Najczęściej ze wszystkich grup mają nastoletnie dzieci
- \* Najwięcej wydają na wino, mięso i złoto
- \* Najczęściej ze wszystkich grup korzystają z promocji i sklepu internetowego
- \* Druga najczęściej wchodząca na stronę internetową grupa
- \* Bardzo chętnie korzystają z kampanii promocyjnych

- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku

**Podsumowanie:** Średnio zamożni i chętnie korzystający z promocji starsi rodzice (z pewnym potencjałem na dołączenie pod względem przyzwyczajień do klastru 20

– **Klaster 2:**

- \* Najrzadziej mają dzieci, a jeżeli już to starsze
- \* Wydają najwięcej ze wszystkich grup na wszystkie produkty, ulubionym produktem jest wino, za nim ryby (w co może wliczać się kawior) i złoto
- \* Kupują stacjonarnie, z taką samą częstotliwością korzystają z katalogu i sklepu internetowego.
- \* Bardzo mało wchodzi na stronę internetową
- \* Bardzo chętni do brania udziału w kampaniach promocyjnych
- \* Wydają zdecydowanie najwięcej ze wszystkich grup, zarówno bezwzględnie jak i procentowo (prawie 2 procent przychodów)
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku

**Podsumowanie:** bogaci i/lub rozrzutni zblazowani bezdzietni (lub z już dorosłymi dziećmi) łatwi do zmanipulowania przez kampanie promocyjne

– **Klaster 3:**

- \* Mają mniej dzieci niż klaster 1, z tą samą tendencją (dzieci raczej w wieku nastoletnim)
- \* Składają średnio najwięcej zamówień (łącznie ze wszystkich źródeł)
- \* Najwyższe zainteresowanie niektórymi promocjami, znikome innymi
- \* Druga najwyższa średnia ilość skarg
- \* Druga najwyższa proporcja wydatków do zarobków
- \* Tak jak pozostałe grupy, najczęściej mają za sobą edukację na poziomie licencjatu, oraz są w związku

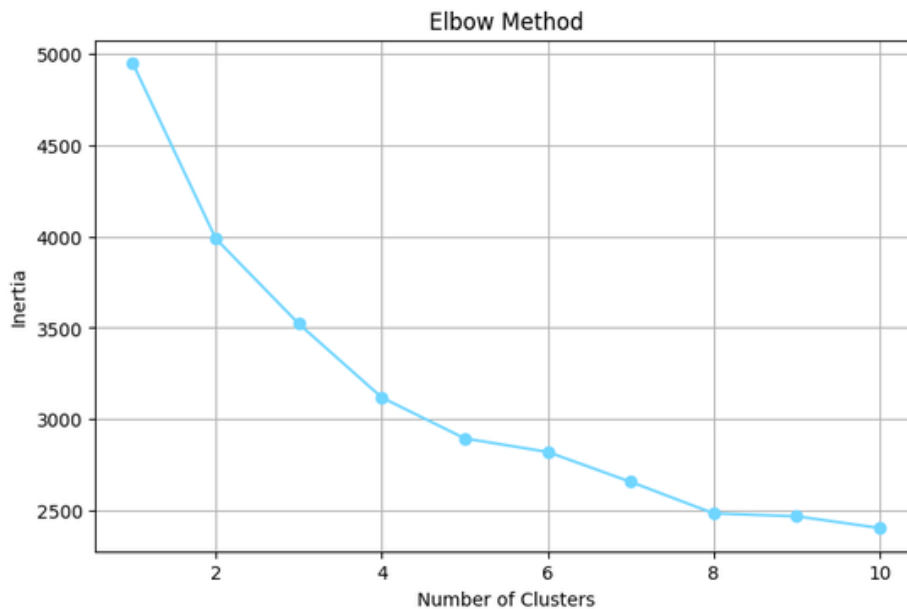
**Podsumowanie:** Części zakupowicze o wyrobionych preferencjach (więc nie łapią byle promocji, ale jeżeli promocja dotyczy produktu który i tak kupują to wezmą w niej bardzo czynny udział),

mają potencjał na bycie bardzo dochodowymi jeżeli będzie się organizować

**Podsumowanie całości:** Sklep powinien pielęgnować relację z bardzo dochodowymi klientami z klastru nr.2 oraz bardzo skupiać się na preferencjach klientów z klastru nr.3 i monitorować zachowania klientów z klastru nr.1 kierując ich w kierunku zachowań klientów z klastru nr.2. Klienci z klastru numer 0 są dosyć trudni w utrzymaniu i mało dochodowi więc można ich w dużej mierze ignorować

## 8 Klasteryzacja dla danych po normalizacji min-max

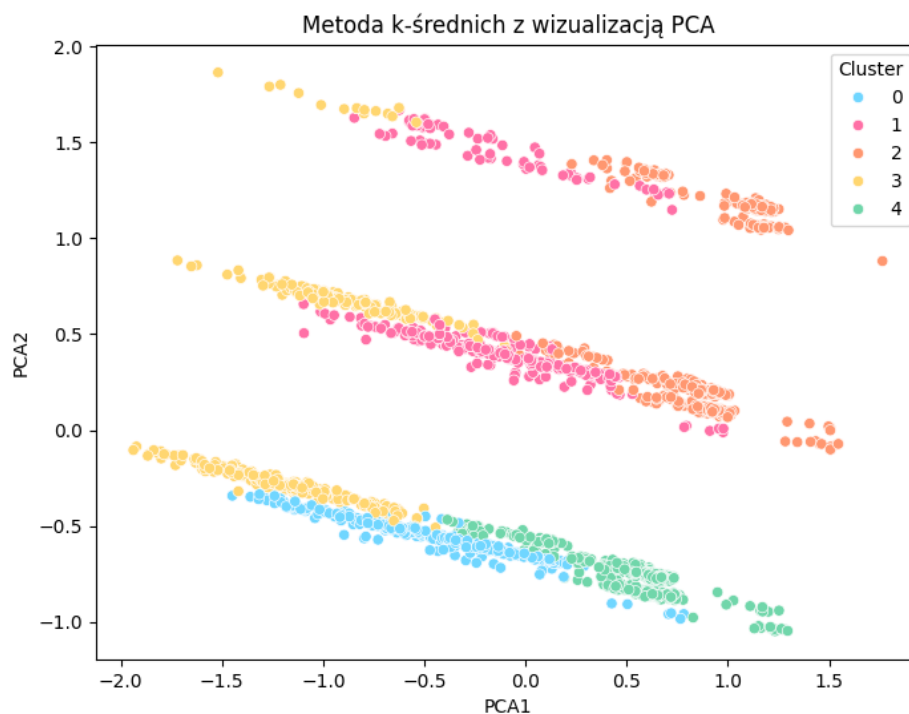
Zastosowano metodę łokcia do wyznaczenia optymalnej liczby klastrów.



Rys. 11. Metoda łokcia w celu wyznaczenia optymalnej liczby klastrów.

Na podstawie otrzymanych wyników zdecydowano się na wykonanie klasteryzacji dla 5 klastrów.

## 8.1 Klasteryzacja k-średnich



Rys. 12. Wyniki klasteryzacji k-średnich dla 5 klastrów.

### 8.1.1 Wnioski

## 8.2 Klasteryzacja GMM

### 8.2.1 Wnioski

## 9 Wnioski

Analiza segmentacji klientów przeprowadzona za pomocą metod klasteryzacji (k-średnich i GMM) ujawnia wyraźne różnice między grupami klientów, które mogą być pomocne w tworzeniu spersonalizowanych strategii marketingowych. Na podstawie wyników klasteryzacji udało się wyróżnić kilka istotnych grup konsumentów, które charakteryzują się odmiennymi preferencjami i zachowaniami zakupowymi.

napisz co wyróżniono

## 10 Rekomendacje marketingowe

## 11 Podsumowanie

Analiza segmentacji klientów przeprowadzona za pomocą metod klasteryzacji (k-średnich i GMM) pozwoliła na wyodrębnienie kilku grup o odmiennych preferencjach i zachowaniach zakupowych. Na podstawie wyników klasteryzacji opracowane zostały rekomendacje marketingowe, które uwzględniają specyfikę każdej z grup. Kluczowym elementem w przyszłości będzie testowanie tych rekomendacji oraz monitorowanie efektywności działań marketingowych skierowanych do każdej z grup.