

*Grażyna Trzpiot**
*Alicja Ganczarek-Gamrot***

METODY KLASYFIKACYJNE DANYCH W ANALIZACH WIRTUALNYCH ŁAŃCUCHÓW DOSTAW

Streszczenie. W pracy poddano klasyfikacji informacje zebrane za pomocą badania ankietowego przeprowadzonego wśród przedsiębiorstw na terenie Województwa Śląskiego. W grupie tej analizowano niezależnie przedsiębiorstwa produkcyjne, handlowe i usługowe. Głównym celem badania jest klasyfikacja czynników kształtujących wirtualne łańcuchy dostaw w skład których wchodzi: stosowana technologia informacyjno-komunikacyjna, kompetencje logistyczne, branża transportowo-spedycyjno-logistyczna. Ze względu na jakościowy oraz ilościowy charakter zmiennych do klasyfikacji wykorzystano metody oparte na różnych miarach odległości. Wyniki analizy zostaną wykorzystane do pracy z zakresu logistyki na temat „Czynniki rozwoju wirtualnych łańcuchów dostaw.”

Słowa kluczowe: wirtualny łańcuch dostaw, metody aglomeracji, miary odległości.

I. WPROWADZENIE

Zmiany technologiczne i organizacyjne wpłynęły na ukształtowanie się nowych warunków gospodarczych i wzrostu konkurencyjności rynkowej. W warunkach kształtowania się międzynarodowych wpływów, rosnących oczekiwań klientów, szybkiego procesu rozwoju i umierania produktów itp., firmy coraz częściej rezygnują z pionowej struktury organizacji, na którą składa się działalność zaopatrzeniowa, produkcyjna i dystrybucyjna na rzecz struktur, które umacniają funkcjonalność firm w zakresie ich kompetencji. W nowych nieukształtowanych jeszcze warunkach funkcjonowania przedsiębiorstw, wykształciła się teoretyczna koncepcja „organizacji wirtualnej”, która stała się symbolem nowoczesności rozwoju ekonomiczno-społecznego.

Do rozwoju organizacji wirtualnej przyczynił się rozwój informatyki a w szczególności rozwój możliwości symulowania rzeczywistości w postaci „wirtualnej rzeczywistości”. Technika wirtualnej rzeczywistości znalazła zastosowanie między innymi w medycynie, architekturze, budownictwie, w projektowaniu i rozwoju wyrobów.

* Prof. AE, Dr hab., Katedra Statystyki, Akademia Ekonomiczna w Katowicach.

** Dr, Katedra Statystyki, Akademia Ekonomiczna w Katowicach.

Wirtualny, z łacińskiego słowa „virtus” oznacza „mogący zaistnieć”. Generalnie wirtualność postrzegana jest jako zjawisko, którego nie można dostrzec, a jedynie jego skutki. W literaturze przedmiotu organizację wirtualną definiuje się przez wymienienie jej cech funkcjonowania: tymczasowość, koncentrację na kliencie, rozproszenie geograficzne, intensywne wykorzystanie technologii informatycznych, sieciową strukturę organizacyjną, wykorzystanie kluczowych kompetencji swoich uczestników (Adamczyk (2005)).

Organizacja wirtualna jest luźną organizacją podmiotów gospodarczych, powoływaną w celu wykorzystania konkretnej okazji rynkowej, więc z założenia czas jej trwania jest ograniczony. Każdy z podmiotów jest skupiony na jednym rodzaju działalności, w którym jest szczególnie biegły. O przynależności do organizacji nie decyduje położenie geograficzne, lecz posiadanie zasobów (wiedzy, technologii, personelu, majątku). Komunikacja między podmiotami oparta jest na sieciach komputerowych.

Łańcuch dostaw rozumiany jest jako przepływ surowców, materiałów, podzespołów i wyrobów gotowych od momentu pozyskania tych pierwszych, do momentu konsumpcji wyrobu finalnego przez użytkownika końcowego. Łańcuch dostaw można postrzegać jako organizację łączącą przedsiębiorstwa uczestniczące w procesie dostarczania danego produktu na rynek. Wirtualny łańcuch dostaw jest więc organizacją tymczasową powołaną do konkretnego celu, koncentrującą się na kliencie, wykorzystującą technologie informatyczne.

W pracy poddano klasyfikacji informacje zebrane za pomocą badania ankietowego przeprowadzonego wśród przedsiębiorstw na terenie Województwa Śląskiego. W grupie tej analizowano niezależnie 46 przedsiębiorstw produkcyjnych, 26 handlowych i 40 usługowych. Głównym celem badania jest klasyfikacja czynników kształtujących wirtualne łańcuchy dostaw, w skład których wchodzi: stosowana technologia informacyjno-komunikacyjna, kompetencje logistyczne, branża transportowo-spedycyjno-logistyczna.

II. METODY KLASYFIKACJI

Klasyfikacja jest pojęciem wieloznacznym. Celem klasyfikacji może być:

- uzyskanie jednorodnych cech danej rodziny, co może uwydatnić związki przyczynowo skutkowe w rozpatrywanej populacji,
- klasyfikację wykorzystuje się w analizach wielowymiarowych między innymi do zredukowania zbyt dużej liczby informacji, co może uprościć proces wnioskowania (Kolonko (1980), Grabiński (1992), Jajuga (1993)).

W teorii mnogościowej definiuje się klasyfikację jako niepustą rodzinę podzbiorów K_i $i = (1, \dots, k)$ określoną na zbiorze obiektów K i spełniającą warunki Grabiński (1992):

$$\mathbf{K}_i \cap \mathbf{K}_j = \emptyset \quad (i \neq j; \quad i, j = 1, \dots, k), \quad (1)$$

$$\bigcup_{i=1}^k \mathbf{K}_i = \mathbf{K}. \quad (2)$$

Najogólniej klasyfikacja traktowana jest jako zbiór klas wyróżnionych z klasyfikowanego zbioru obiektów. W literaturze przedmiotu klasę definiuje się w różny sposób. Między innymi (Gatnar, Walesiak (2004)):

- klasa jest taką zbiorowością obiektów, w której podobieństwo pomiędzy dowolną parą obiektów jest większe niż podobieństwo pomiędzy jakimkolwiek obiektem należącym do klasy a dowolnym obiektem do niej nienależącym,
- klasy tworzą obiekty najbardziej podobne, natomiast w różnych klasach znajdują się obiekty najmniej podobne,
- klasami są takie obszary w przestrzeni m -wymiarowej, które charakteryzują się większą gęstością obiektów i są oddzielone obszarami o mniejszej gęstości obiektów.

Ze względu na charakter zmiennych do klasyfikacji wykorzystano analizę skupień opartą na różnych metodach aglomeracji danych oraz różnych miarach odległości. Wśród metod aglomeracji wykorzystano: metodę pojedynczego wiązania, metodę pełnego wiązania, metodę średnich połączeń, metodę średnich połączeń ważonych, metodę Warda. Do pomiaru odległości między poszczególnymi obiektami zastosowano: odległość euklidesową, odległość euklidesową kwadratową, odległość miejską, odległość Czebyszewa (Gordon (1987), Gatnar (1998), Gatnar, Walesiak (2004)).

III. WYNIKI KLASYFIKACJI PRZEDSIĘBIORSTW

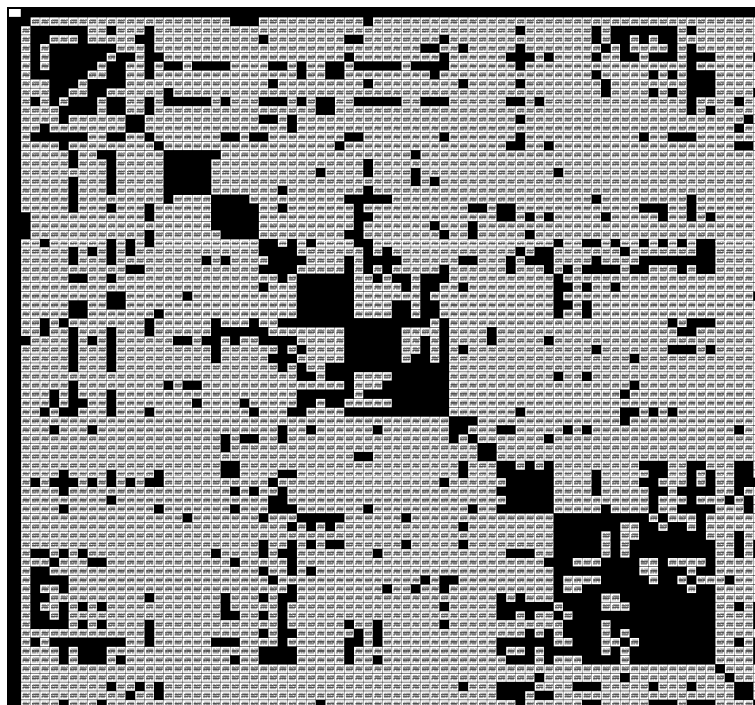
Na podstawie badania ankietowego trzech grup przedsiębiorstw Województwa Śląskiego dysponowano informacjami dotyczącymi kompetencji logistycznych – trzy grupy przedsiębiorstw oraz rozwoju technologii informacyjno-komunikacyjnej (trzy grupy przedsiębiorstw analizowane wspólnie, łącznie 121 przedsiębiorstw). Cztery typy ankiet zostały rozdane respondentom reprezentującym różne przedsiębiorstwa. Otrzymane wyniki ankiet różnią się między sobą zarówno liczbą respondentów, jak również rodzajem pytań. Baza danych zawiera zarówno zmienne: ilościowe przedstawione w skali porządkowej i przedziałowej, zmienne jakościowe przedstawione w skali nominalnej i porządkowej. Do klasyfikacji wykorzystano wyłącznie zmienne przedstawione za pomocą skali przedziałowej i porządkowej.

Na podstawie wstępnej analizy danych oceniono rozkłady analizowanych zmiennych oraz zależności między analizowanymi zmiennymi. Następnie na

bazie wstępnej oceny rozkładów oraz zależności między zmiennymi przeprowadzono ich klasyfikację.

Rozkłady zmiennych charakteryzują się bardzo silną asymetrią (głównie prawostronna). W większości przypadkach również silną koncentracją wartości wokół średniej. Na poziomie istotności 0,05 należy odrzucić hipotezę, że analizowane zmienne mają rozkład normalny.

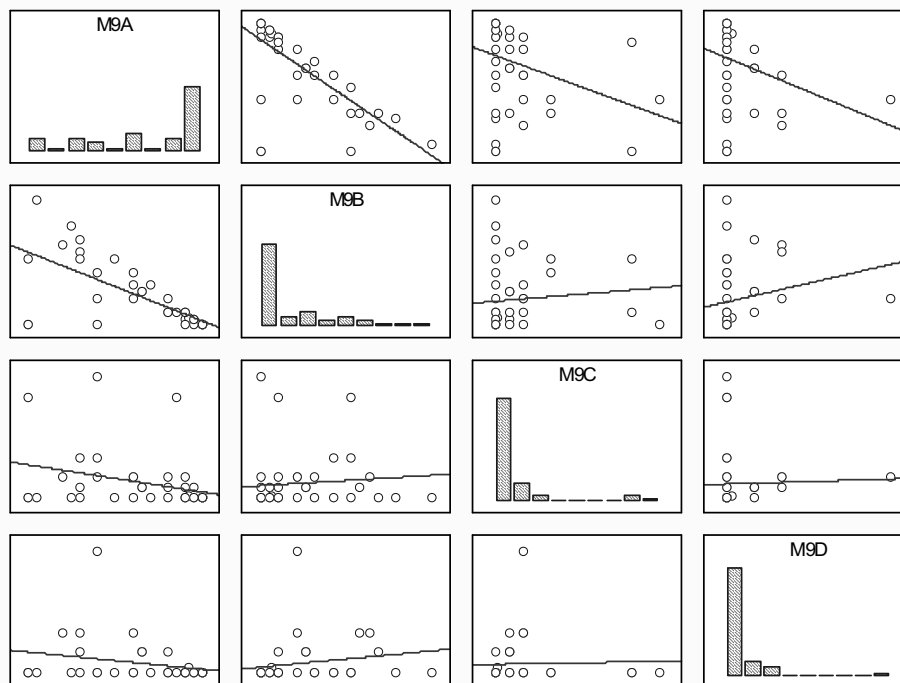
Występujące zależności, są to głównie zależności w obrębie jednego pytania. Na rysunku 1 zaprezentowano macierz korelacji między zmiennymi przedsiębiorstw produkcyjnych. Istotne na poziomie 0,05 współczynniki korelacji Spearmana zaznaczono czarnym kolorem.



Rys. 1. Macierz współczynników korelacji Spearmana przedsiębiorstw produkcyjnych

Na rysunku 2 zaprezentowano korelacyjną macierz rozrzutu odpowiedzi na pytanie: „Proszę określić procent SPRZEDAŻY Państwa firmy wygenerowany w poszczególnych obszarach geograficznych w roku 2005”. Możliwe odpowiedzi:

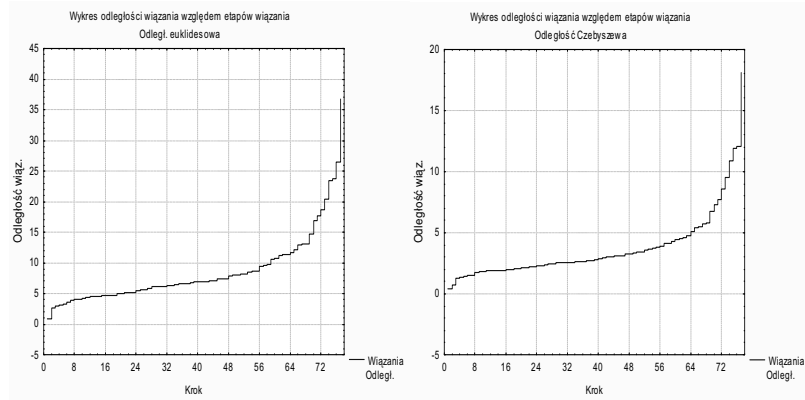
- a) Rynek krajowy,
- b) Poza rynkiem krajowym, ale wewnątrz Unii Europejskiej,
- c) Poza Unią Europejską, ale na terenie Europy,
- d) Poza Europą.



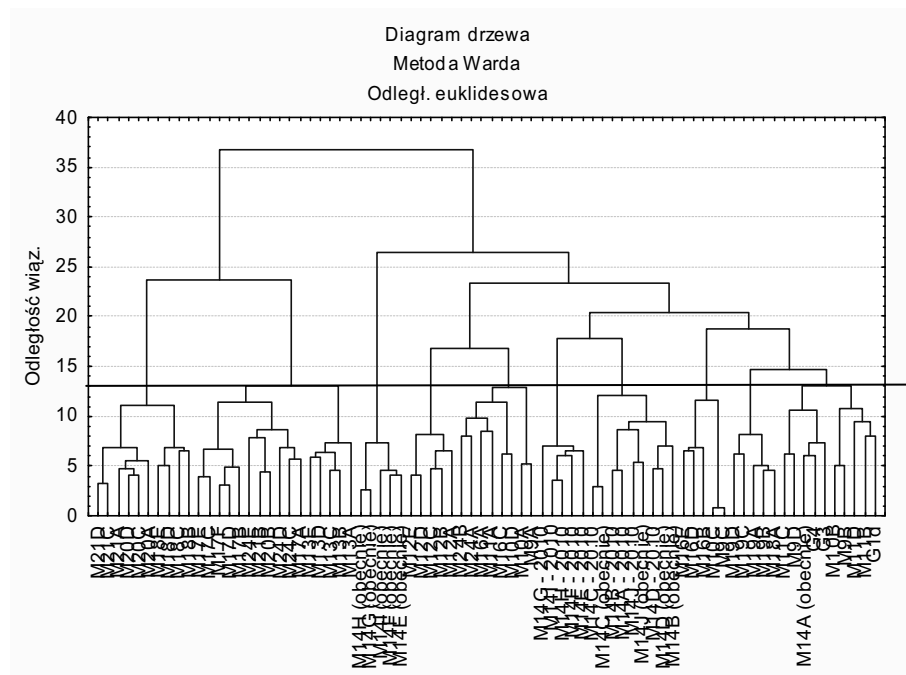
Rys. 2. Zależności między ilością sprzedaży w poszczególnych obszarach geograficznych przedsiębiorstw produkcyjnych Województwa Śląskiego w roku 2005

Ze względu na charakter pytania, odpowiedzi w kategoriach A, B, C, D muszą się sumować do 100%. Naturalnym jest fakt, że jeżeli przedsiębiorstwa w głównej mierze koncentrują się na sprzedaży w kraju, to odsetek sprzedaży poza rynkiem krajowym jest niewielki.

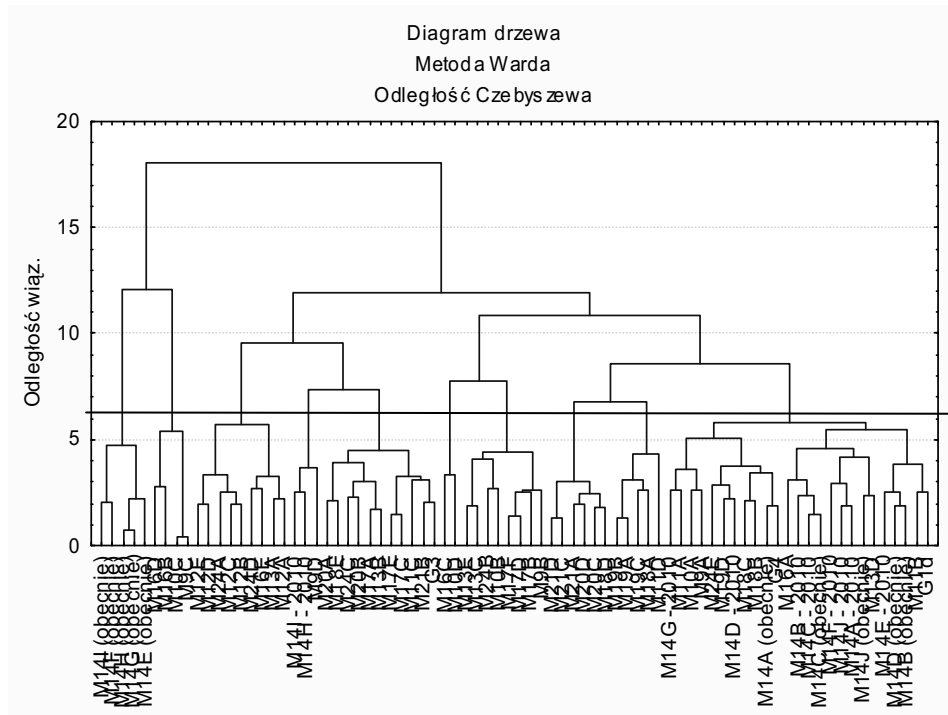
Niezależnie w każdej grupie przedsiębiorstw na podstawie zestandaryzowanych wyników odpowiedzi przeprowadzono klasyfikację analizowanych zmiennych. Z wykorzystanych metod aglomeracji najlepsze wyniki dała metoda Warda z odległością euklidesową, miejską i Czebyszewa. Na rysunkach (3–14) przedstawiono wyniki klasyfikacji dla poszczególnych grup przedsiębiorstw uzyskane metodą aglomeracji Warda z odległością euklidesową i Czebyszewa. Jako kryterium wyboru klasyfikacji posłużono się wykresem przebiegu aglomeracji.



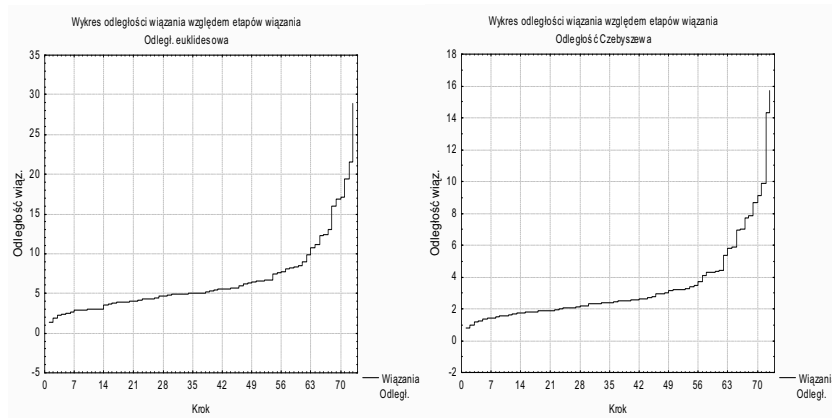
Rys. 3. Wykresy przebiegu aglomeracji przedsiębiorstw produkcyjnych



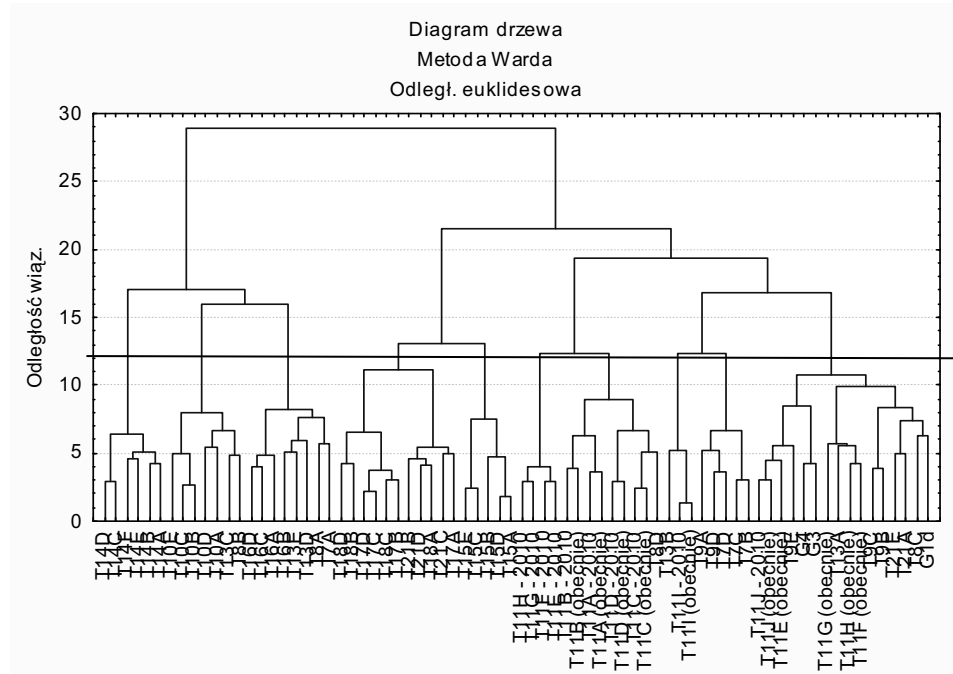
Rys. 4. Klasyfikacja przedsiębiorstw produkcyjnych w oparciu o odległość euklidesową



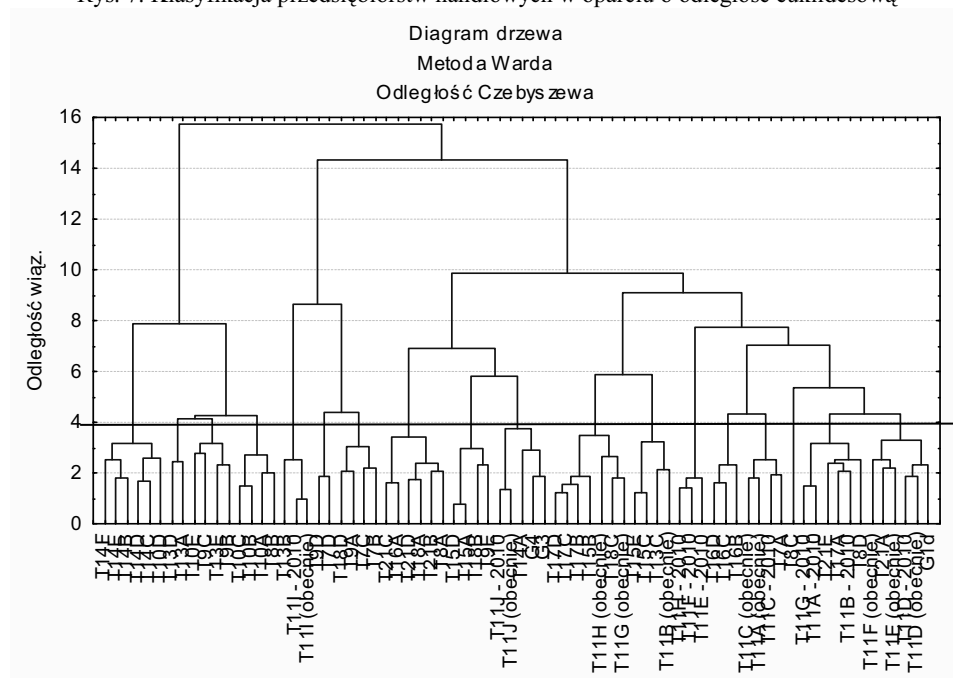
Rys. 5. Klasyfikacja przedsiębiorstw produkcyjnych w oparciu o odległość Czebyszewa



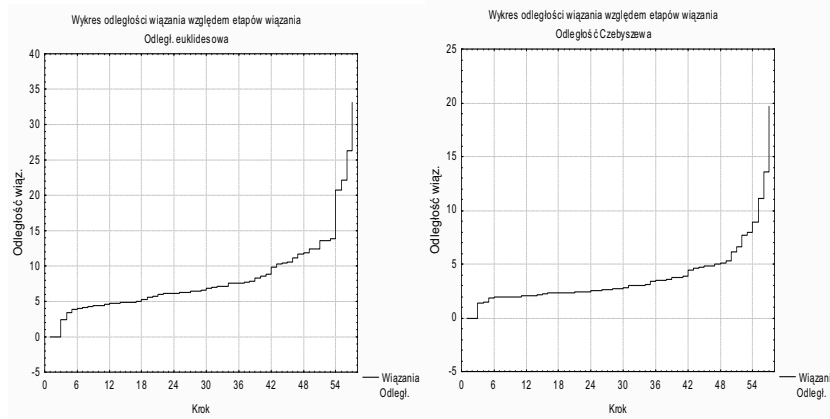
Rys. 6. Wykresy przebiegu aglomeracji przedsiębiorstw handlowych



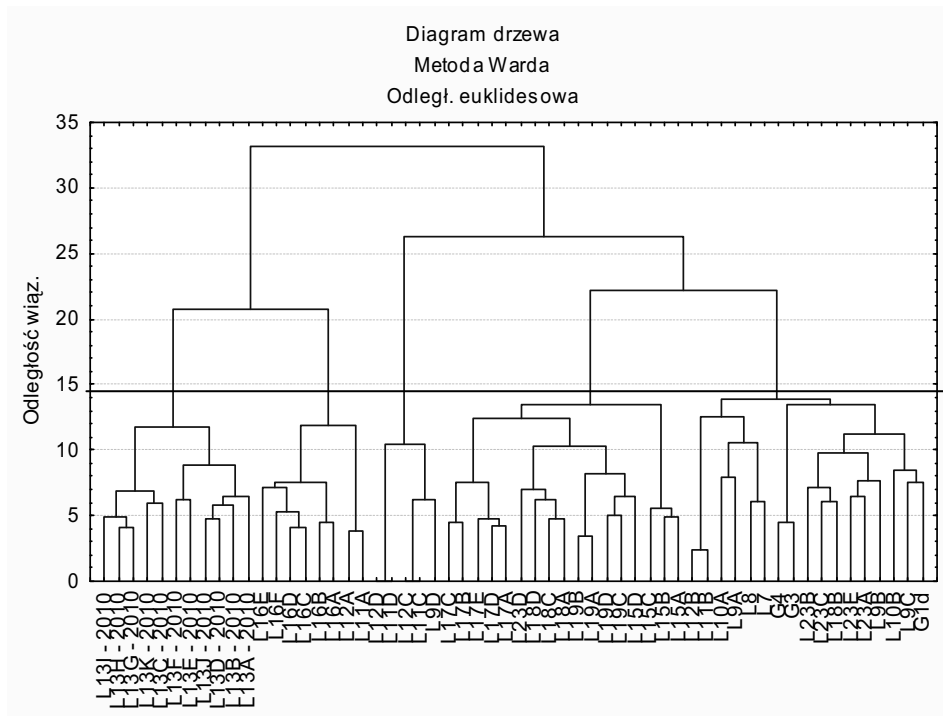
Rys. 7. Klasyfikacja przedsiębiorstw handlowych w oparciu o odległość euklidesową



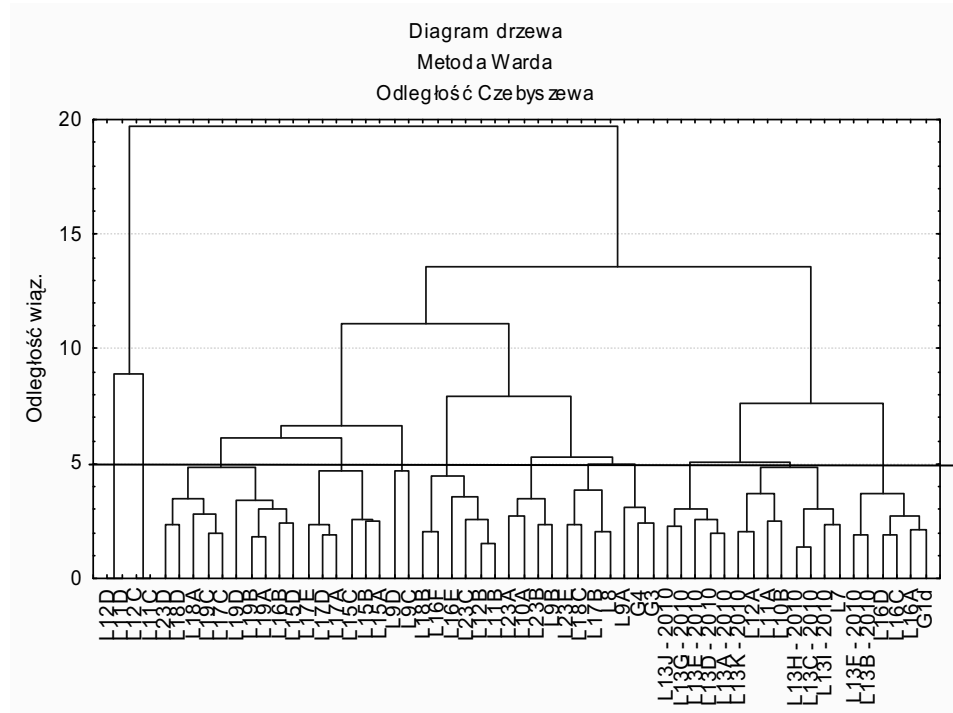
Rys. 8. Klasyfikacja przedsiębiorstw handlowych w oparciu o odległość Czebyszeva



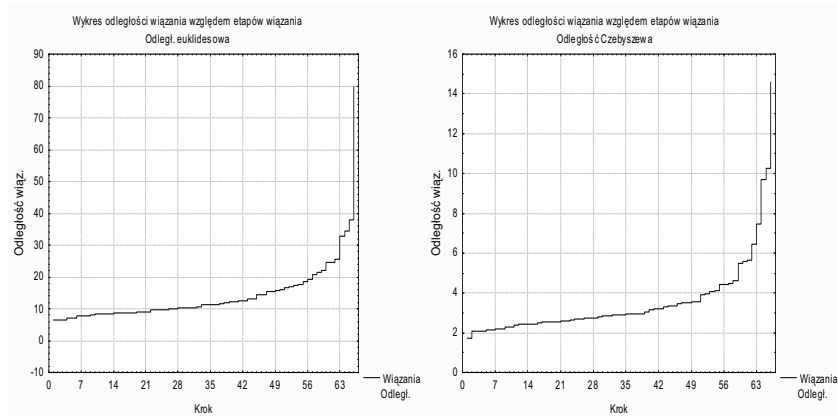
Rys. 9. Wykresy przebiegu aglomeracji przedsiębiorstw usługowych



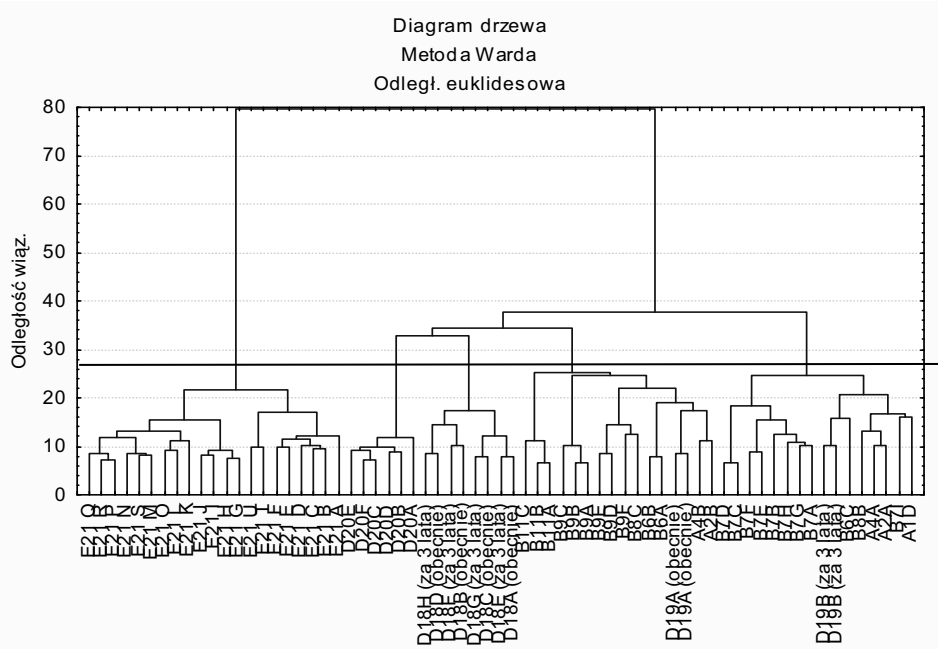
Rys. 10. Klasyfikacja przedsiębiorstw usługowych w oparciu o odległość euklidesową



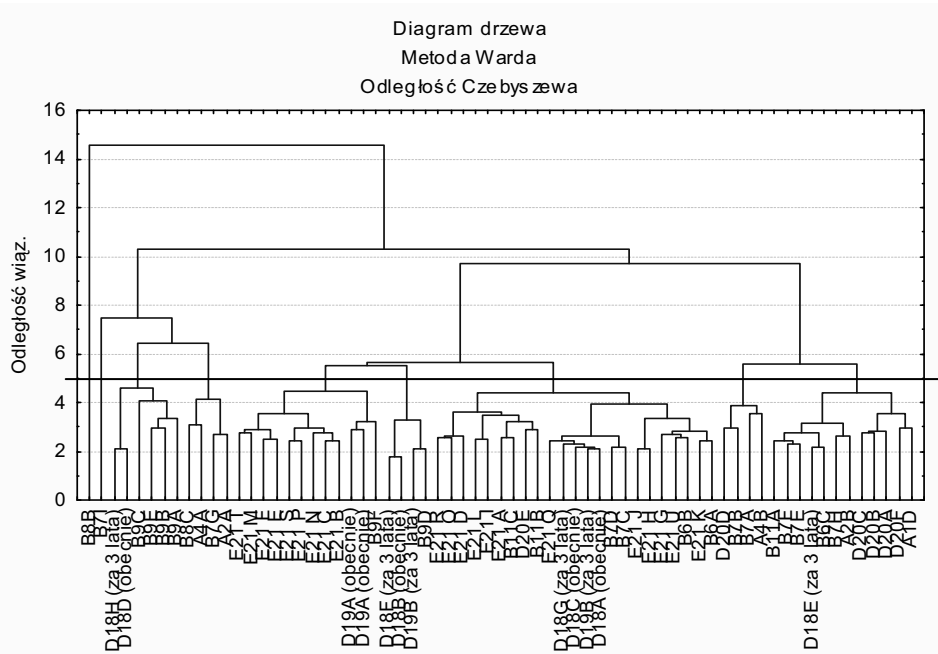
Rys. 11. Klasyfikacja przedsiębiorstw usługowych w oparciu o odległość Czebyszewa



Rys. 12. Wykresy przebiegu aglomeracji przedsiębiorstw Województwa śląskiego



Rys. 13. Klasyfikacja przedsiębiorstw Województwa Śląskiego w oparciu o odległość euklidesową



Rys. 14. Klasyfikacja przedsiębiorstw Województwa Śląskiego w oparciu o odległość Czebyszewa

Niezależnie od rodzaju przedsiębiorstwa, oraz wyboru miary odległości, w drodze hierarchicznej aglomeracji zmiennych otrzymano skupiska odpowiadające tym samym kategoriom pytań związanym z konstrukcją ankiety. Można scharakteryzować poszczególne skupiska i przyporządkować im określoną funkcję w kształtowaniu się wirtualnego łańcucha dostaw. W tabeli 1 przedstawiono szczegółowo wyniki klasyfikacji wszystkich przedsiębiorstw Województwa Śląskiego, które wzięły udział w ankiecie (wyniki te są przedstawione na rysunku 13). Analizując tematykę pogrupowanych w ten sposób zmiennych wyodrębniono pięć czynników kształtujących wirtualny łańcuch dostaw.

Tabela 1. Wyniki klasyfikacji czynników kształtujących wirtualne łańcuchy dostaw

Odległość euklidesowa	11,68138	17,26794	21,52996	24,60771	25,42024
Zmienne x	Znaczenie handlu elektronicznego	Procesy transportowe	Ocena wykorzystania rozwiązań informacyjno-komunikacyjnych	Planowanie zastosowania technologii informacyjno-komunikacyjnych w różnych obszarach działalności firmy	Wykorzystane do tej pory Rozwiązania informacyjno-komunikacyjne
Obj. Nr 1	D20A	D18A (obecnie)	E21 A	A1D	A2B
Obj. Nr 2	D20B	D18E (za 3 lata)	E21 B	B7I	A4B
Obj. Nr 3	D20D	D18C (obecnie)	E21 C	A2A	D19A (obecnie)
Obj. Nr 4	D20C	D18G(za 3 lata)	E21 D	A4A	D19A (obecnie)
Obj. Nr 5	D20F	D18B (obecnie)	E21 E	B8B	B6A
Obj. Nr 6	D20E	D18F (za 3 lata)	E21 F	B6C	B6B
Obj. Nr 7		D18D (obecnie)	E21 T	D19B(za 3 lata)	B8C
Obj. Nr 8		D18H(za 3 lata)	E21 U	D19B(za 3 lata)	B9F
Obj. Nr 9			E21 G	B7A	B9D
Obj. Nr 10			E21 H	B7G	B9E
Obj. Nr 11			E21 I	B7H	B9A
Obj. Nr 12			E21 J	B7B	B9B
Obj. Nr 13			E21 K	B7E	B9C
Obj. Nr 14			E21 L	B7F	B11A
Obj. Nr 15			E21 O	B7C	B11B
Obj. Nr 16			E21 M	B7D	B11C
Obj. Nr 17			E21 S		
Obj. Nr 18			E21 N		
Obj. Nr 19			E21 P		
Obj. Nr 20			E21 R		
Obj. Nr 21			E21 Q		

IV. PODSUMOWANIE

Ze względu na charakter zmiennych oraz niewielką liczebność prób w klasyfikacji czynników kształtujących wirtualne łańcuchy dostaw wykorzystano hierarchiczną analizę skupień. Wśród wykorzystanych metod aglomeracji oraz miar odległości najlepsze wyniki otrzymano na podstawie aglomeracji Warda z odległością euklidesową oraz odległością Czebyszewa. Wśród wyodrębnionych czynników można wyróżnić:

- znaczenie handlu elektronicznego,
- procesy transportowe,
- ocenę wykorzystania rozwiązań informacyjno-komunikacyjnych,
- planowanie zastosowania technologii informacyjno-komunikacyjnych w różnych obszarach działalności firmy,

- wykorzystane do tej pory Rozwiązania informacyjno-komunikacyjne,

Uzyskane wyniki klasyfikacji zmiennych są pierwszym etapem analizy przedsiębiorstw Województwa Śląskiego. Znając czynniki kształtujące wirtualne łańcuchy dostaw w dalszych etapach analiz można próbować określić ich rolę w kształtowaniu się wirtualnych łańcuchów dostaw.

BIBLIOGRAFIA

- Adamczyk M. (2005), *Charakterystyka organizacji wirtualnej*, „Gazeta IT”, 2005, nr 9 (39).
- Ganczarek A (2003), *Klasyfikacja polskiego rynku energii*, Inżynieria Ekonomiczna w Badaniach Społeczno-Gospodarczych, Politechnika Rzeszowska, 51–66.
- Ganczarek A. (2005), *APT Model for Electricity Prices on the Day Ahead Market of the Polish Power Exchange*, Acta Universitatis Lodzianis Folia Oeconomica, 194, 259–268.
- Gordon A. D. (1987), *A review of hierarchical classification*, Journal of the Royal Statistical Society, ser. A, 119–137.
- Gatnar E. (1998), *Symboliczne metody klasyfikacji danych*, PWN, Warszawa.
- Gatnar E., Walesiak M. (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, AE, Wrocław.
- Grabiński T. (1992), *Metody taksonometrii*, AE, Kraków.
- Jajuga K. (1993), *Statystyczna Analiza wielowymiarowa*, PWN, Warszawa.
- Kolonko J. (1980), *Analiza dyskryminacyjna i jej zastosowanie w ekonomii*, PWN, Warszawa.
- Trzpiot G., Ganczarek A. *The classification of risk on the Polish Power Exchange*, Zeszyty Naukowe Katedry Ekonometrii, red. nauk. Józef Dziechciarz, Wydawnictwo AE, Wrocław, w druku.

Grażyna Trzpiot, Alicja Ganczarek

**CLASSIFICATION METHODS IN DATA ANALYSIS
OF VIRTUAL SUPPLY CHAINS**

Abstract

In this paper the classification of data obtained in the survey of Silesian businesses is presented. The survey was aimed at production, commerce and service enterprises, The main objective was to classify the factors determining the virtual supply chain consisting of IT and communication systems logistic competences and the transport-spedition-logistic branch. Due to the qualitative and quantitative character of variables, methods based on varying similarity measures were used.