



Agata Sielska

Szkoła Główna Handlowa w Warszawie
Kolegium Zarządzania i Finansów
Katedra Ekonomii Stosowanej
asiels@sgh.waw.pl

NAKLADY NA OPIEKĘ ZDROWOTNĄ A ROZWÓJ EPIDEMII COVID-19 W WOJEWÓDZTWACH POLSKI W OKRESIE 4 MARCA – 31 MAJA 2020 R.

Streszczenie: Celem artykułu jest zbadanie różnic i podobieństw między województwami w rozwoju epidemii COVID-19 w okresie od 4 marca do 31 maja 2020 r. w kontekście finansowania ochrony zdrowia w 2018 r. W analizie przebiegu epidemii wykorzystano dwa podejścia. Uzyskane wyniki wskazują na istniejące między województwami podobieństwa. W przypadku badania podobieństw województw na podstawie wartości przyjmowanych przez zmienne opisujące rozwój epidemii każdego dnia analizowanego okresu znacząca dla finalnej klasyfikacji okazała się zmienna opisująca liczbę zgonów na 10 tys. osób. Dla alternatywnego podejścia, uwzględniającego dynamikę epidemii, decydującą rolę wydają się mieć wykryte przypadki. Podobieństwa grupowań ze względu na finansowanie oraz przebieg epidemii nie są duże, relatywnie najmniejsze różnice występują w przypadku klasyfikacji uwzględniającej zgony.

Słowa kluczowe: koronawirus, COVID-19, metoda Warda.

JEL Classification: I10, I15.

Wprowadzenie

W pierwszej połowie 2020 r. świat został dotknięty kryzysem wywołanym rozwojem pandemii koronawirusa SARS-CoV-2 powodującego chorobę COVID-19. Brak szczepionki oraz skutecznego lekarstwa (w okresie powstawania tego tekstu dostępne było jedynie leczenie objawowe) spowodował wdrożenie w wielu państwach daleko idących środków, wśród których można wymienić zamknięcie granic, ograniczenie poruszania się, zamknięcie punktów usługowych bądź gastronomicznych czy zamknięcie szkół. Konsekwencje pandemii mają charakter ogólnospołeczny i ogólnogospodarczy. W związku z trudną sytu-

acją, w jakiej znalazła się pod koniec 2019 r. polska ochrona zdrowia [Nojszewska, Sielska, Gołąb-Beltowicz, 2019], istotną wydaje się próba oceny, w jaki sposób epidemia koronawirusa przebiegała w poszczególnych województwach Polski, z uwzględnieniem ich dotychczasowego zróżnicowania pod względem nakładów finansowych, które mają oczywisty wpływ na osiągnięte wyniki.

Celem opracowania jest zbadanie różnic i podobieństw między województwami w rozwoju epidemii COVID-19 w okresie od 4 marca do 31 maja 2020 r. w kontekście finansowania ochrony zdrowia w 2018 r. Artykuł składa się z trzech części. W pierwszej z nich krótko przedstawiono przebieg epidemii w Polsce. Druga część została poświęcona metodom badawczym. Zaprezentowano wykorzystane do analizy zmienne i źródła danych. Uzyskane wyniki przedstawiono i skomentowano w części trzeciej, która składa się z dwóch podrozdziałów. Pierwszy z nich stanowi omówienie wyników analizy różnic i podobieństw między województwami w zakresie finansowania ochrony zdrowia w 2018 r., drugi zaś odnosi się do różnic i podobieństw w rozwoju epidemii COVID-19 w rozpatrywanym okresie. Całość tekstu kończy podsumowanie.

1. Pandemia COVID-19

Pod koniec 2019 r. w Chinach w mieście Wuhan (prowincja Hubei) wykryto przypadki nowej choroby zakaźnej. Koronawirus nazwany SARS-CoV-2, powodujący m.in. groźne zapalenie płuc, szybko rozprzestrzenił się poza kontynent azjatycki, w 2020 r. docierając na wszystkie kontynenty poza Antarktydą. Światowa Organizacja Zdrowia (WHO) 11 marca 2020 r. ogłosiła stan pandemii [Li i in., 2020]. Do dnia, kiedy składano ten artykuł, odnotowano ponad 7 mln przypadków i 400 tys. zgonów. W Polsce jest to odpowiednio blisko 28 tys. przypadków i ok. 1200 zgonów [www 1].

W Polsce pierwsze zachorowanie odnotowano 4 marca [www 2]. Główny Inspektorat Sanitarny (GIS) wydał 8 marca rekomendację odwołania części imprez masowych [Informacja Głównego Inspektora Sanitarnego..., 2020]. Później 11 marca podjęto decyzję o wstrzymaniu zajęć szkolnych oraz prowadzonych na uczelniach wyższych [Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego w sprawie czasowego ograniczenia..., 2020; Rozporządzenie Ministra Zdrowia w sprawie czasowego ograniczenia..., 2020; Rozporządzenie Ministra Edukacji Narodowej w sprawie czasowego ograniczenia..., 2020]. W kolejnych dniach przywrócone zostały kontrole graniczne, a także częściowo zamknięto granice kraju [Rozporządzenie Ministra Spraw Wewnętrznych i Administracji w sprawie

przywrócenia..., 2020; Rozporządzenie Ministra Spraw Wewnętrznych i Administracji w sprawie czasowego zawieszenia..., 2020; Rozporządzenie Rady Ministrów w sprawie zakazów w ruchu lotniczym, 2020]. Stan zagrożenia epidemicznego ogłoszono 13 marca [Rozporządzenie Ministra Zdrowia w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu zagrożenia epidemicznego, 2020], a 20 marca wprowadzony został stan epidemii [Rozporządzenie Ministra Zdrowia w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu epidemii, 2020]. Kolejne ograniczenia wprowadzono już 24 marca [Rozporządzenie Ministra Zdrowia zmieniające rozporządzenie w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu epidemii, 2020]. Obostrzenia wprowadzone zostały ogólnie, w tym samym terminie dla wszystkich województw, również ich późniejsze znoszenie dotyczyło całego kraju.

Na statystyki zgonów i zachorowań w oczywisty sposób wpływa stan sektora ochrony zdrowia w poszczególnych województwach. Dotyczy to posiadanych przez szpitale zasobów, zarówno sprzętowych, jak i kadrowych, oraz dostępnego przez lata finansowania. Istotne jest to nie tylko dla możliwości zwiększenia liczby prowadzonych testów diagnostycznych, a co za tym idzie – wykrywania potencjalnych chorych, jak również dla ogólnego stanu zdrowia ludności i możliwości skutecznego leczenia chorych, co przyczynia się do obniżenia statystyki zgonów.

2. Metoda badawcza

2.1. Zmienne i źródła danych

W porównaniu województw pod kątem finansowania opieki zdrowotnej uwzględniono środki z budżetów jednostek samorządów terytorialnych (oznaczone dalej jako JST), jak też Narodowego Funduszu Zdrowia (NFZ), dedykowane umowom na leczenie szpitalne i podstawową opiekę zdrowotną (POZ). Wszystkie wartości zostały przedstawione w ujęciu na 10 tys. osób. Zmienne te wybrano, ponieważ odzwierciedlają główne aspekty finansowania ochrony zdrowia.

Badając podobieństwo rozwoju epidemii w poszczególnych województwach, wzięto pod uwagę trzy zmienne: skumulowaną liczbę przypadków na 10 tys. osób, skumulowaną liczbę zgonów na 10 tys. osób oraz skumulowaną liczbę zgonów na jednego chorego.

Dokładną listę cech wraz z oznaczeniami przedstawiono poniżej:

- NFZ_SZPIT – wartość realizacji umów NFZ na leczenie szpitalne,
- NFZ_POZ – wartość realizacji umów NFZ na POZ,
- JST – ogółem wydatki budżetów JST w Dziale 851 – Ochrona zdrowia,
- POST – skumulowana liczba przypadków zachorowań na 10 tys. osób,
- ZOST – skumulowana liczba zgonów na 10 tys. osób,
- ZCOST – skumulowana liczba zgonów na jednego chorego (dla zerowego ZOST przypisano 0).

Źródłem danych dotyczących finansowania przez NFZ jest sprawozdanie z działalności NFZ [Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r., 2019], zaś dane dotyczące zachorowań na COVID-19, potwierdzonych pozytywnymi wynikami testów laboratoryjnych, zaczerpnięto z oficjalnych komunikatów podawanych przez Ministerstwo Zdrowia w serwisie Twitter [www 2]. Źródłem pozostałych danych jest Bank Danych Lokalnych GUS [www 3].

Statystyki opisowe zmiennych dotyczących finansowania przedstawiono w tabeli 1.

Tabela 1. Wartości statystyk opisowych dla zmiennych dotyczących finansowania ochrony zdrowia

Zmienna	Min	Q1	Mediana	Średnia	Q3	Max	SD
NFZ_SZPIT	10 179	10 464	10 875	10 927	11 183	12 174	627,85
NFZ_POZ	2693	2835	2897	2889	2938	3056	86,57
JST	88 595	148 879	226 256	232 955	283 567	511 523	112 109,9

SD – odchylenie standardowe.

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 3].

Analizą objęto okres od 4 marca 2020, tj. od wystąpienia pierwszego przypadku zakażenia na terenie Polski, do 31 maja 2020 r. Ze względu na fakt, że wykorzystane w badaniu dane opisujące rozwój epidemii mają częstotliwość dzienną (zebrano obserwacje dla 89 dni), dla uproszczenia i przejrzystości w tabeli 2 przedstawiono ich statystyki opisowe w ujęciu tygodniowym. Ostatni okres jest nieco krótszy, trwa od 27 do 31 maja.

Tabela 2. Wartości statystyk opisowych dla zmiennych dotyczących rozwoju epidemii COVID-19 w województwach

Data	Zmienna	Min	Q1	Mediana	Średnia	Q3	Max	SD	NA
4.03.2020	POST	0,0000	0,0000	0,0000	0,0006	0,0000	0,0099	0,0025	
	ZOST	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	ZCOST	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	NA	15
11.03.2020	POST	0,0000	0,0038	0,0077	0,0079	0,0105	0,0210	0,0062	
	ZOST	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	ZCOST	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	NA	3
18.03.2020	POST	0,0085	0,0382	0,0535	0,0666	0,0868	0,1703	0,0454	
	ZOST	0,0000	0,0000	0,0000	0,0012	0,0007	0,0069	0,0023	
	ZCOST	0,0000	0,0000	0,0000	0,0153	0,0114	0,0769	0,0282	0
25.03.2020	POST	0,0846	0,1548	0,2045	0,2425	0,2502	0,5839	0,1498	
	ZOST	0,0000	0,0000	0,0000	0,0030	0,0050	0,0138	0,0043	
	ZCOST	0,0000	0,0000	0,0000	0,0129	0,0242	0,0667	0,0194	0
1.04.2020	POST	0,2571	0,4376	0,4861	0,5972	0,6657	1,1326	0,2566	
	ZOST	0,0000	0,0000	0,0050	0,0083	0,0149	0,0236	0,0089	
	ZCOST	0,0000	0,0000	0,0085	0,0130	0,0207	0,0370	0,0144	0
8.04.2020	POST	0,6128	0,9268	1,1467	1,1915	1,3550	2,4318	0,4800	
	ZOST	0,0000	0,0109	0,0238	0,0320	0,0509	0,0794	0,0252	
	ZCOST	0,0000	0,0119	0,0191	0,0254	0,0406	0,0584	0,0187	0
15.04.2020	POST	0,7590	1,2380	1,6900	1,7460	2,2200	3,3160	0,7342	
	ZOST	0,0000	0,0109	0,0457	0,0552	0,0874	0,1555	0,0493	
	ZCOST	0,0000	0,0090	0,0297	0,0277	0,0427	0,0587	0,0204	0
22.04.2020	POST	0,8181	1,5070	2,1684	2,3321	3,1798	3,9397	1,0534	
	ZOST	0,0000	0,0263	0,0572	0,0824	0,1233	0,2424	0,0748	
	ZCOST	0,0000	0,0153	0,0287	0,0307	0,0487	0,0624	0,0211	0
29.04.2020	POST	0,8378	1,7889	2,5817	2,8597	3,9706	5,2357	1,3624	
	ZOST	0,0000	0,0568	0,0981	0,1252	0,1731	0,3646	0,1073	
	ZCOST	0,0000	0,0276	0,0321	0,0385	0,0557	0,0812	0,0242	0
6.05.2020	POST	0,8674	1,9294	2,8350	3,2897	4,5706	6,5004	1,7294	
	ZOST	0,0000	0,0769	0,1058	0,1494	0,2112	0,3979	0,1183	
	ZCOST	0,0000	0,0313	0,0361	0,0406	0,0556	0,0804	0,0225	0
13.05.2020	POST	0,8871	2,0503	3,0209	3,7373	4,9681	9,7032	2,3516	
	ZOST	0,0000	0,0915	0,1281	0,1781	0,2853	0,4220	0,1370	
	ZCOST	0,0000	0,0343	0,0373	0,0441	0,0687	0,0792	0,0251	0
20.05.2020	POST	0,8871	2,1203	3,1679	4,1726	5,5044	13,4155	3,1275	
	ZOST	0,0000	0,1028	0,1604	0,2031	0,3347	0,4479	0,1520	
	ZCOST	0,0000	0,0330	0,0404	0,0470	0,0737	0,0909	0,0267	0
27.05.2020	POST	1,0740	2,3110	3,3850	4,6720	6,0920	16,7420	3,8634	
	ZOST	0,0000	0,1072	0,1668	0,2178	0,3571	0,4764	0,1633	
	ZCOST	0,0000	0,0304	0,0422	0,0464	0,0702	0,0926	0,0271	0
31.05.2020	POST	1,1240	2,3620	3,7070	4,9330	6,3500	18,2020	4,1910	
	ZOST	0,0000	0,1072	0,1708	0,2240	0,3691	0,4886	0,1692	
	ZCOST	0,0000	0,0291	0,0411	0,0455	0,0710	0,0904	0,0268	0

SD – odchylenie standardowe, NA – brakujące obserwacje.

Źródło: Obliczenia własne na podstawie: [www 1; www 2; www 3].

Przeciętna liczba przypadków zachorowań na 10 tys. osób, mierzona medianą, wyniosła 3,71 w ostatnim dniu badanego okresu. Mediana zgonów na 10 tys. osób wykazuje wyraźny trend wzrostowy od kwietnia, ostatniego dnia osiąga wartość 0,17. W przypadku zgonów w ujęciu relatywnym na jedną osobę chorą, po początkowym wzroście na przełomie marca i kwietnia, można zauważyć pewną stabilizację pod koniec badanego okresu.

2.2. Metody analizy

Ponieważ przedstawione zmienne charakteryzują się odmiennymi zakresami wartości, wszystkie zostały poddane normalizacji. Wykorzystano metodę unitaryzacji zerowanej:

$$z_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}$$

gdzie:

x_i – wartość cechy przed normalizacją,

z_i – wartość cechy po dokonaniu normalizacji.

Jak wskazuje się w [Milligan, Cooper, 1988], jest to jeden ze sposobów normalizacji, który pozwala na efektywne odkrycie rzeczywistej struktury danych.

Grupowanie dokonane zostało za pomocą metody Warda [Ward, 1963], która prowadzi do uzyskania przejrzyste wyszczególnionych grup, co wydaje się stanowić jeden z powodów jej popularności [Romesburg, 2004, s. 134]. Wykorzystywana była ona do porównywania jednostek terytorialnych w Polsce w wielu sferach [Kobylińska, 2015; Roszkowska, Filipowicz-Chomko, 2016; Cymerman, Cymerman, 2017].

Analizę przebiegu epidemii przeprowadzono równolegle w dwóch podejściach. W pierwszym, określanym dalej mianem „dziennego”, każde z województw opisano trzema wektorami zmiennych: $\mathbf{td}_{w,i} = [td_{w,i,j}]$, gdzie w oznacza województwo, i – zmienną (POST, ZOST lub ZCOST), j jest datą dzienną należącą do zbioru $J = \{04.03.2020, \dots, 31.05.2020\}$, $td_{w,i,j}$ oznacza wartość zmiennej i w województwie w w dniu j . Każdemu z województw została przyporządkowana macierz $\mathbf{TD}_{3 \times 89} = [\mathbf{td}_{w,POST} \ \mathbf{td}_{w,ZOST} \ \mathbf{td}_{w,ZCOST}]^T$, której poszczególne wiersze odpowiadają wektorom $\mathbf{td}_{w,i}$, a każdej ze zmiennych macierz: $\mathbf{TD}_{16 \times 89}^i = [\mathbf{td}_{1,i} \ \dots \ \mathbf{td}_{16,i}]^T$, $i \in \{POST, ZOST, ZCOST\}$.

Przed dokonaniem grupowania zmienne unormowano metodą unitaryzacji zerowanej. Wynikowe macierze oznaczono jako:

$$TDN_{16 \times 89}^i = [tdn_{w,i,j}]$$

gdzie:

$$tdn_{w,i,j} = \frac{td_{w,i,j} - \min_w(td_{w,i,j})}{\max_w(td_{w,i,j}) - \min_w(td_{w,i,j})}$$

w – województwo,

i – zmienna (POST, ZOST lub ZCOST),

j – data.

Ponieważ rozwój epidemii charakteryzował się w różnych rejonach Polski różną dynamiką, w alternatywnym podejściu, określanym dalej mianem „równomiernego”, zbadano kształtowanie się zmiennych POST, ZOST oraz ZCOST w równych odstępach czasu, z których każdy odpowiadał 1/10 czasu trwania epidemii w danym województwie. W tym celu wyznaczono dla każdego województwa i każdej ze zmiennych datę $tr_{w,i,0}$, dla której odnotowano po raz pierwszy niezerową wartość zmiennej, gdzie w oznacza województwo, zaś i – zmienną (POST, ZOST lub ZCOST). Dla każdego województwa okres od $tr_{w,i,0}$ do 31 maja 2020 r. został następnie podzielony na 10 podokresów równej długości. W konsekwencji uzyskano dla każdego województwa i zmiennej zbiór 11 dat wraz z odpowiadającymi im wartościami zmiennych. Dla przykładu, jeśli w województwie pierwszą niezerową wartość zmiennej odnotowano 4 marca 2020 r. ($tr_{w,i,0}$), to obserwacje uwzględnione w toku dalszej analizy pochodzą z dni: 4.03.2020, 13.03.2020, 21.03.2020, 30.03.2020, 8.04.2020, 16.04.2020, 25.04.2020, 4.05.2020, 13.05.2020, 21.05.2020, 30.05.2020, zaś dla $tr_{w,i,0}$ równego 9 marca 2020 r. będą to obserwacje z dni: 9.03.2020, 17.03.2020, 26.03.2020, 3.04.2020, 11.04.2020, 19.04.2020, 28.04.2020, 6.05.2020, 14.05.2020, 23.05.2020 i 31.05.2020. Kolejne daty oznaczano jako $tr_{w,i,j}$, gdzie $j \in [1,10]$, budując w ten sposób 3 macierze $TR_{16 \times 11}^i = [tr_{w,i,j}]$, $j \in [0,10]$, $i \in \{POST, ZOST, ZCOST\}$, których wymiary odpowiadają liczbie województw i liczbie obserwacji. Elementy macierzy unormowano metodą unitaryzacji zerowanej. Wynikowe macierze oznaczono jako:

$$TRN_{16 \times 11}^i = [tn_{w,i,j}]$$

gdzie: $tn_{w,i,j} = \frac{tr_{w,i,j} - \min_w(tr_{w,i,j})}{\max_w(tr_{w,i,j}) - \min_w(tr_{w,i,j})}$.

Podobieństwo grup województw w poszczególnych latach zbadano za pomocą indeksu Randa [por. Rand, 1971].

Porównania wartości poszczególnych wskaźników w wyodrębnionych grupach oparto na wartościach przeciętnych (medianach). W badaniu podobieństw sytuacji w województwach pod względem przypadków wystąpienia choroby COVID-19 i spowodowanych nią zgonów na podstawie elementów macierzy $TDN_{16 \times 89}^i$ (ujęcie „dziennie”) oraz macierzy $TRN_{16 \times 11}^i$ (ujęcie „równomierne”) obliczono odległości euklidesowe między województwami pod względem przebiegu epidemii COVID-19. Dla przykładu, odległości między województwami opisanymi przez wiersze k i l macierzy TDN^{POST} wyznaczono ze wzoru:

$$d_{k,l}(TDN^{POST}) = \sqrt{\sum_{j=1}^{89} (tdn_{k,POST,j} - tdn_{l,POST,j})^2}$$

zaś odległości między województwami opisanymi przez wiersze k i l macierzy TRN^{POST} wyznaczono ze wzoru:

$$d_{k,l}(TRN^{POST}) = \sqrt{\sum_{j=1}^{11} (tn_{k,POST,j} - tn_{l,POST,j})^2}$$

Analogicznie obliczono odległości dla pozostałych województw i zmiennych (ZOST, ZCOST). Dla każdego województwa, podejścia i zmiennej obliczono następnie medianę jego odległości od pozostałych województw.

Obliczenia wykonano w programie R [R Core Team, 2018] z wykorzystaniem pakietów fossil [Vavrek, 2011] (wyznaczenie wartości indeksu Randa), NbClust [Charrad i in., 2014] (wybór optymalnej liczby grupowań), cluster [Machler i in., 2018] (grupowanie), psych [Revelle, 2019] (zestawienie wyników, dalsza analiza) oraz w arkuszu kalkulacyjnym MS Excel (zestawienie wyników, dalsza analiza).

3. Podobieństwa województw

3.1. Podobieństwa województw pod kątem finansowania opieki zdrowotnej w 2018 r.

Wyboru liczby grup dokonano na podstawie indeksu GAP [Tibshirani, Walther, Hastie, 2001], założono minimalną liczbę grup równą 4 i maksymalną równą 10. Biorąc pod uwagę fakt, że optymalna okazała się najniższa dopuszczona w badaniu liczba grup, można wnioskować, że podobieństwo województw jest na tyle duże, iż różnice między poszczególnymi grupami będą niewielkie. Wyniki grupowania województw pod kątem finansowania przedstawiono w tabeli 3.

Tabela 3. Grupy województw według podobieństw w sferze finansowania i mediany zmiennych

Grupa	NFZ_SZPIT	NFZ_POZ	JST	Województwa
1	10 715,64	2897,42	266 991,8	dolnośląskie, kujawsko-pomorskie, lubuskie, podkarpackie, podlaskie, pomorskie, śląskie, wielkopolskie
2	11 634,16	2930,36	178 502,9	lubelskie, łódzkie, świętokrzyskie
3	10 461,59	2813,54	104 522,7	małopolskie, opolskie, warmińsko-mazurskie, zachodniopomorskie
4	12 173,86	3056,31	511 523,1	mazowieckie

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 3].

Do grupy 4 zaklasyfikowano jedynie województwo mazowieckie, w przypadku którego odnotowano najwyższe wartości wszystkich trzech zmiennych, co szczególnie wyraźnie widać w przypadku zmiennej JST. Z kolei grupa 3 charakteryzowała się najniższymi przeciętnymi wartościami zmiennych. Grupy 1 i 2 cechowały się umiarkowanymi wartościami median, przy czym w przypadku grupy 2 można zauważyć relatywnie wysokie przeciętne finansowanie z NFZ (mediany zmiennych NFZ_SZPIT i NFZ_POZ są niższe tylko od wartości dla grupy 4) i relatywnie niskie finansowanie z JST (mediana wyższa jedynie od grupy 3).

3.2. Podobieństwa województw pod kątem rozwoju epidemii COVID-19

3.2.1. Podejście „dienne”

Wyniki klasyfikacji województw dla rozważanych cech przedstawiono w tabeli 4. Wyboru liczby grup ponownie dokonano na podstawie indeksu GAP [Tibshirani, Walther, Hastie, 2001] przy założeniu minimalnej liczby grup równej 4 i maksymalnej równej 10. Jak można zauważyć na podstawie tabeli 5, najbardziej podobne klasyfikacje uzyskano dzięki wykorzystaniu zmiennej odpowiadającej liczbie zgonów na 10 tys. osób (ZOST) oraz zgonom i przypadkom na 10 tys. osób (POST i ZOST, wartość indeksu Randa wynosi tutaj 0,8083). Wydaje się to prowadzić do wniosku o znaczącej dla finalnej klasyfikacji liczbie zgonów.

Wartości indeksów Randa obliczonych dla grupowań dokonanych w oparciu o przebieg pandemii i o finansowanie w 2018 r., zamieszczone w ostatnim wierszu w tabeli 5, prowadzą do wniosku o niewielkim podobieństwie grupowań na podstawie przebiegu pandemii i finansowania. Wszystkie cztery wyznaczone

wartości indeksów Randa były zbliżone. Relatywnie największa wartość indeksu wystąpiła dla klasyfikacji ze względu na liczbę zgonów na 10 tys. mieszkańców (ZOST).

Tabela 4. Grupy województw według podobieństw rozwoju epidemii COVID-19 (podejście „dzienne”)

Województwo	Zmienna			
	przypadki na 10 tys. osób	zgony na 10 tys. osób	zgony na jednego chorego	zgony i przypadki na 10 tys. osób
Dolnośląskie	1	1	1	1
Kujawsko-pomorskie	2	2	2	2
Lubelskie	3	1	1	3
Lubuskie	3	3	3	2
Łódzkie	1	2	3	1
Małopolskie	2	2	3	2
Mazowieckie	1	4	2	4
Opolskie	2	4	2	4
Podkarpackie	3	1	4	3
Podlaskie	2	3	3	2
Pomorskie	3	3	3	2
Śląskie	4	4	1	4
Świętokrzyskie	3	2	3	2
Warmińsko-mazurskie	3	3	3	2
Wielkopolskie	2	4	4	4
Zachodniopomorskie	3	2	3	2

Źródło: Obliczenia własne na podstawie: [www 2; www 3].

Tabela 5. Wartości indeksu Randa dla klasyfikacji województw według podobieństw rozwoju epidemii COVID-19 (podejście „dzienne”) i według finansowania w 2018 r.

	POST	ZOST	ZCOST	POST, ZOST	Finansowanie 2018 r.
POST	1,0000				
ZOST	0,6250	1,0000			
ZCOST	0,6250	0,7333	1,0000		
POST, ZOST	0,6833	0,8083	0,7750	1,0000	
Finansowanie 2018 r.	0,5583	0,6000	0,5500	0,5583	1,0000

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 2; www 3].

Mediany odległości euklidesowych od pozostałych województw obliczone dla każdego województwa i grupowań na podstawie indywidualnej zmiennej przedstawiono w tabeli 6. Pod względem liczby zgonów na jednego chorego najbardziej podobne do innych województw są województwa łódzkie i małopolskie, zaś najbardziej odmienne – województwo wielkopolskie. Z kolei pod względem zgonów na 10 tys. mieszkańców najbardziej podobne do innych są województwa małopolskie, pomorskie i zachodniopomorskie, zaś najbardziej odmienne – województwo mazowieckie. Dla liczby odnotowanych przypadków wyraźnie odmienne okazują się województwa dolnośląskie i śląskie. Województwo śląskie jest specyficznym przypadkiem. W maju 2020 r., wraz z wykrytymi nowymi ogniskami zakażeń w śląskich kopalniach, pracujący w nich górnicy oraz ich rodziny objęci zostali prowadzonymi na dużą skalę badaniami przesiewowymi. Zwiększona liczba testów przełożyła się na zwiększoną liczbę wykrytych zachorowań. Jest to wyjątek w badanym okresie, uwzględniamy go jednak dla zachowania ciągłości i kompletności analizy.

Tabela 6. Mediany odległości euklidesowych województw dla 3 zmiennych opisujących rozwój COVID-19 (podejście „dziennie”)

Województwo	Zmienna		
	POST	ZOST	ZCOST
Dolnośląskie	5,178041	4,577803	3,056302
Kujawsko-pomorskie	2,621576	2,848485	3,235605
Lubelskie	3,063918	4,226008	3,603584
Lubuskie	3,404696	2,945902	4,046891
Łódzkie	4,259205	2,781043	2,462505
Małopolskie	2,381445	1,993446	2,552546
Mazowieckie	4,556175	5,471928	3,947769
Opolskie	3,242424	4,447872	3,562793
Podkarpackie	2,584573	3,124999	4,411329
Podlaskie	2,963256	2,042318	3,113261
Pomorskie	2,771282	1,714976	3,08823
Śląskie	5,025013	4,625618	3,648694
Świętokrzyskie	2,304692	2,317457	3,209816
Warmińsko-mazurskie	3,036811	2,759185	3,512903
Wielkopolskie	3,351441	4,692313	4,626683
Zachodniopomorskie	2,928841	1,921602	2,729077

Źródło: Obliczenia własne na podstawie: [www 2; www 3].

3.2.2. Podejście „równomierne”

W ujęciu uwzględniającym częstotliwość równą 1/10 czasu trwania epidemii w poszczególnych województwach uzyskano grupowanie przedstawione w tabeli 7 (podobnie jak poprzednio, wyboru liczby grup ponownie dokonano na podstawie indeksu GAP [Tibshirani, Walther, Hastie, 2001] przy założeniu minimalnej liczby grup równej 4 i maksymalnej równej 10). Grupowanie różni się od uzyskanego wcześniej, a wartości indeksów Randa obliczone dla tej klasyfikacji (tabela 8) prowadzą do wniosku o wyraźnym wpływie na wyniki klasyfikacji liczby potwierdzonych przypadków (wartość indeksu dla klasyfikacji na podstawie POST oraz na podstawie kombinacji POST i ZOST wynosi 0,9167). Na podstawie analizy indeksów Randa obliczonych dla grupowań dokonanych w oparciu o przebieg pandemii i o finansowanie można zauważyć relatywnie największe podobieństwo w przypadku klasyfikacji ze względu na liczbę zgonów na 10 tys. mieszkańców (ZCOST).

Tabela 7. Grupy województw według podobieństw rozwoju epidemii COVID-19 (podejście „równomierne”)

Województwo	Zmienna			
	przypadki na 10 tys. osób	zgony na 10 tys. osób	zgony na jednego chorego	zgony i przypadki na 10 tys. osób
Zachodniopomorskie	1	1	1	1
Wielkopolskie	2	2	2	2
Warmińsko-mazurskie	1	1	3	1
Świętokrzyskie	1	1	1	1
Śląskie	2	3	4	3
Pomorskie	1	1	1	1
Podlaskie	3	1	3	4
Podkarpackie	1	1	2	1
Opolskie	2	4	2	3
Mazowieckie	3	3	2	3
Małopolskie	1	1	1	1
Łódzkie	3	2	1	4
Lubuskie	1	1	3	1
Lubelskie	1	1	4	1
Kujawsko-pomorskie	4	2	2	4
Dolnośląskie	2	2	4	2

Źródło: Obliczenia własne na podstawie: [www 2; www 3].

Tabela 8. Wartości indeksu Randa dla klasyfikacji województw według podobieństw rozwoju epidemii COVID-19 (podejście „równomierne”)

	POST	ZOST	ZCOST	POST, ZOST	Finansowanie 2018 r.
POST	1,0000				
ZOST	0,8167	1,0000			
ZCOST	0,6250	0,5917	1,0000		
POST, ZOST	0,9167	0,8667	0,6250	1,0000	
Finansowanie 2018 r.	0,5500	0,5500	0,5917	0,5500	1,0000

Źródło: Obliczenia własne na podstawie: [www 2; www 3].

Oceniając na podstawie median odległości euklidesowych przedstawionych w tabeli 9, można uznać, że najmniej typowymi, biorąc pod uwagę przypadki na 10 tys. osób, okazało się województwo śląskie, dla zgonów w ujęciu na 10 tys. osób – województwa opolskie, mazowieckie i śląskie, zaś dla zgonów w stosunku do chorych – województwa lubuskie i wielkopolskie.

Tabela 9. Mediany odległości euklidesowych województw dla 3 zmiennych opisujących rozwój COVID-19 (podejście „równomierne”)

Województwo	Zmienna		
	POST	ZOST	ZCOST
Dolnośląskie	1,2860274	0,9292214	1,0909265
Kujawsko-pomorskie	1,3708932	1,1255357	1,1591852
Lubelskie	0,9946615	0,6291762	1,1027701
Lubuskie	1,2468355	1,047872	1,6070058
Łódzkie	1,3990068	0,8509134	0,9183785
Małopolskie	0,8476092	0,7424336	0,8740503
Mazowieckie	1,4607806	1,9644456	1,3796187
Opolskie	1,0865219	2,2506668	1,4290133
Podkarpackie	0,9214852	0,6317107	1,3902248
Podlaskie	1,0617029	0,8356059	1,0662022
Pomorskie	0,9076777	0,7193117	1,0686304
Śląskie	1,7696869	1,6145994	1,3355121
Świętokrzyskie	0,7939065	0,7519121	0,9978002
Warmińsko-mazurskie	1,1259917	0,860645	1,2805601
Wielkopolskie	1,1270381	1,3083365	1,5385928
Zachodniopomorskie	0,8932459	0,672046	0,8788132

Źródło: Obliczenia własne na podstawie: [www 2; www 3].

3.2.3. Porównanie podejść „dziennego” i „równomiernego”

Wartości indeksu Randa dla klasyfikacji województw według podobieństw rozwoju epidemii COVID-19 i obu podejść przedstawiono w tabeli 10. Jak można zauważyć, podobieństwo jest dość wysokie, relatywnie najbardziej zbliżone grupowania występują w przypadku zgonów na jednego chorego.

Tabela 10. Wartości indeksu Randa dla klasyfikacji województw według podobieństw rozwoju epidemii COVID-19 oraz podejść „dziennego” i „równomiernego”

Podejście „równomierne”	Podejście „dzienne”			
	POST	ZOST	ZCOST	POST, ZOST
POST	0,7917	0,6500	0,6833	0,7083
ZOST	0,7583	0,6333	0,7000	0,7417
ZCOST	0,6167	0,7917	0,8250	0,6833
POST, ZOST	0,7917	0,6833	0,7000	0,7417

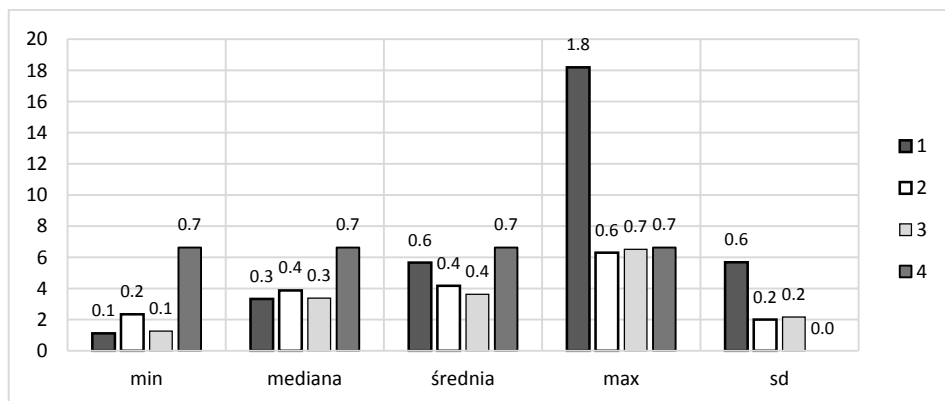
Źródło: Obliczenia własne na podstawie: [www 2; www 3].

3.3. Finansowanie ochrony zdrowia w województwach a sytuacja epidemiczna

W dalszej kolejności zbadano kształtowanie się statystyk opisowych liczby przypadków COVID-19 na 10 tys. mieszkańców (obliczone dla danej grupy województw dla dnia 31 maja 2020 r.) w zależności od tego, do jakiej grupy zaklasyfikowano województwo w 2018 r. pod kątem finansowania ochrony zdrowia. Wyniki przedstawione na rysunku 1 wskazują, że czwarta grupa cechuje się najwyższym współczynnikiem przypadków. Wszystkie statystyki zmiennej POST, z wyłączeniem maksymalnej liczby przypadków, były w jej wypadku najwyższe. Grupa ta była jednoelementowa, należało do niej tylko województwo mazowieckie. Relatywnie wysoką średnią liczbą przypadków cechuje się grupa 1, do której należały województwa śląskie i dolnośląskie. Zróznicowanie liczby przypadków, mierzone odchyleniem standardowym, kształtuje się w niej na najwyższym poziomie, na najniższym zaś dla grupy drugiej.

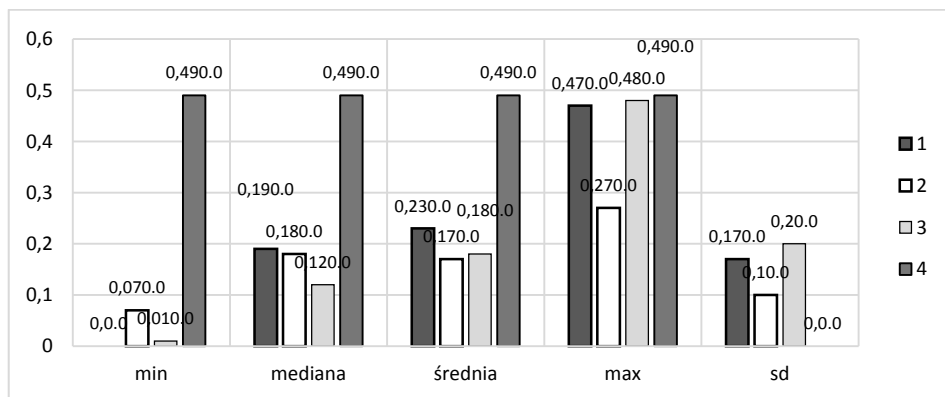
W przypadku zgonów na 10 tys. osób, podobnie jak dla liczby przypadków na 10 tys. osób, wykorzystano wartości statystyk opisowych wyznaczone dla danej grupy województw dla dnia 31 maja 2020 r. Ponownie najwyższe wartości odnotowano dla 4 grupy (rysunek 2). Najniższe przeciętne liczby zgonów, mierzone medianą, wystąpiły dla grupy 3. Najmniej zróżnicowana wewnętrznie była grupa 2, relatywnie najbardziej – 3, do której należały województwa małopol-

skie, opolskie, warmińsko-mazurskie i zachodniopomorskie. W grupie tej odnotowano relatywnie niskie wartości wszystkich zmiennych odzwierciedlających finansowanie ochrony zdrowia.



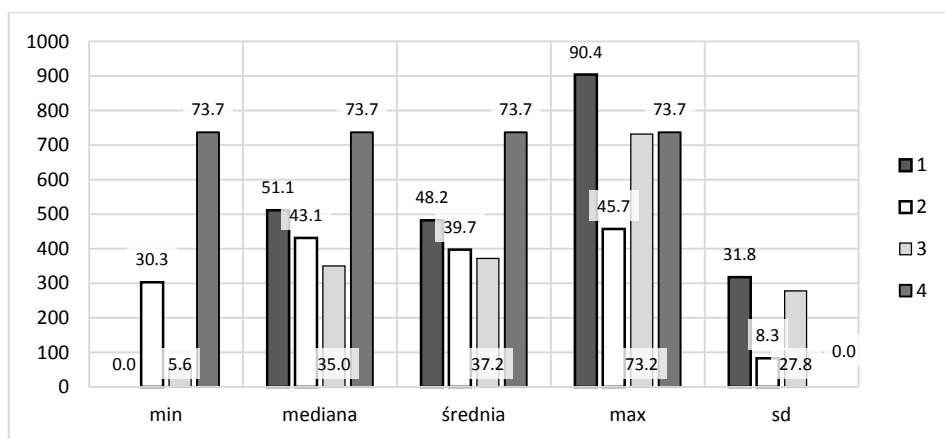
Rys. 1. Statystyki opisowe liczby przypadków na 10 tys. osób w dniu 31.05.2020 r. według grupy województw wyznaczonej na podstawie finansowania w 2018 r.

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 2; www 3].



Rys. 2. Statystyki opisowe liczby zgonów z powodu COVID-19 na 10 tys. osób w dniu 31.05.2020 r. według grupy województw wyznaczonej na podstawie finansowania w 2018 r.

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 2; www 3].



Rys. 3. Statystyki opisowe liczby zgonów z powodu COVID-19 na 10 tys. chorych w dniu 31.05.2020 r. według grupy województw wyznaczonej na podstawie finansowania w 2018 r.

Źródło: Obliczenia własne na podstawie: Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. [2019]; [www 2; www 3].

Na rysunku 3 przedstawiono statystyki opisowe liczby zgonów z powodu COVID-19 na 10 tys. chorych (poziom odniesienia zmieniono dla czytelności wykresu), obliczone dla danej grupy województw dla dnia 31 maja 2020 r. Można zauważyć, że stosunek ten po raz kolejny był najwyższy dla województwa mazowieckiego, które stanowiło jednoelementową grupę 4. Wartości maksymalne wyraźnie wyróżniały się w przypadku grupy pierwszej. Również mediany i średnie tej zmiennej kształtują się w grupie pierwszej na relatywnie wysokim poziomie. Występują tu również województwa o niskich wartościach, o czym świadczy zarówno zerowa wartość minimalna, jak i relatywnie wysokie w porównaniu z pozostałymi dwiema grupami odchylenie standardowe. Najmniejsze zróżnicowanie występowało ponownie wewnątrz grupy 2.

Podsumowanie

W artykule podjęto próbę analizy różnic i podobieństw między województwami w rozwoju epidemii COVID-19 w okresie od 4 marca do 31 maja 2020 r. w kontekście finansowania ochrony zdrowia w 2018 r.

Uzyskane wyniki, w szczególności niska liczba grup (minimalna założona w opracowaniu), wskazują na istniejące między województwami podobieństwa. Można zauważyć, że dla podejścia „dziennego”, w którym podobieństwo woje-

wództw badano na podstawie wartości przyjmowanych przez zmienne opisujące rozwój epidemii każdego dnia analizowanego okresu, znacząca dla finalnej klasyfikacji okazała się zmienna opisująca liczbę zgonów na 10 tys. osób. Z kolei w przypadku alternatywnego podejścia, w którym uwzględniono fakt, że pandemia zaczęła się w różnych województwach w różnych momentach (podejście „równomierne”), decydującą rolę wydają się mieć nie zgony, a wykryte przypadki. Podobieństwa między dwoma zaproponowanymi podejściami mierzone za pomocą indeksu Randa są wysokie.

Na podstawie analizy indeksów Randa obliczonych dla grupowań dokonanych w oparciu o przebieg pandemii i o finansowanie najwyższe wartości indeksów Randa uzyskano dla grupowań opartych na zmiennych obrazujących liczbę zgonów: na 10 tys. mieszkańców (dla podejścia „dziennego”) i na chorego (dla podejścia „równomiernego”), co może być przejawem związku umiERALNOŚCI z finansowaniem. Badanie wartości statystyk opisowych zmiennych ZOST i ZCOST w zależności od grupy nie jest jednak wystarczające do potwierdzenia tego wniosku. Bez wątplenia wskazuje to na celowość dalszej analizy rozprzestrzeniania się COVID-19 w Polsce, zwłaszcza wobec rozwoju epidemii i jej coraz wyraźniejszego wpływu na gospodarkę oraz funkcjonowanie ochrony zdrowia.

Literatura

- Charrad M., Ghazzali N., Boiteau V., Niknafs A. (2014), *NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set*, “Journal of Statistical Software”, Vol. 61(6), s. 1-36.
- Cymerman J., Cymerman W. (2017), *Zastosowanie analizy skupień do klasyfikacji województw według rozwoju rynków nieruchomości rolnych*, „Świat Nieruchomości”, nr 3(101), s. 55-62.
- Informacja Głównego Inspektora Sanitarnego w sprawie imprez masowych (2020).
- Kobylińska M. (2015), *Statystyczna analiza rozwoju społeczeństwa informacyjnego województw Polski w latach 2008 i 2012*, „Roczniki Kolegium Analiz Ekonomicznych”, nr 36, s. 427-439.
- Li L., Yang Z., Dang Z., Meng C., Huang J., Meng H., Wang D., Chen G., Zhang J., Peng H., Shao Y. (2020), *Propagation Analysis and Prediction of the COVID-19*, “Infectious Disease Modelling”, Vol. 5, s. 282-292.
- Maechler M., Rousseeuw P., Struyf A., Hubert M., Hornik K. (2018), *Cluster: Cluster Analysis Basics and Extensions*, R package ver. 2.0.7-1.
- Milligan G.W., Cooper M.C. (1988), *A Study of Standardization of Variables in Cluster Analysis*, “Journal of Classification”, Vol. 5, s. 181-204.

- Nojszewska E., Sielska A., Gołąb-Beltowicz D. (2019), *Raport z badania sytuacji finansowej szpitali powiatowych – Szklarska Poręba '19*, <https://zpp.pl/artukul/1591-raport-z-badania-sytuacji-finansowej-szpitali-powiatowych> (dostęp: 20.04.2020).
- R Core Team (2018), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna.
- Rand W.M. (1971), *Objective Criteria for the Evaluation of Clustering Methods*, "Journal of the American Statistical Association", Vol. 66, s. 846-850.
- Revelle W. (2019), *Psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois.
- Romesburg C. (2004), *Cluster Analysis for Researchers*, Lulu Press.
- Roszkowska E., Filipowicz-Chomko M. (2016), *Ocena realizacji koncepcji zrównoważonego rozwoju województw Polski w zakresie poziomu rozwoju społecznego w roku 2005 oraz 2013*, „Optimum. Studia Ekonomiczne”, nr 3(81), s. 94-118.
- Rozporządzenie Ministra Edukacji Narodowej z dnia 11.03.2020 r. w sprawie czasowego ograniczenia funkcjonowania jednostek systemu oświaty w związku z zapobieganiem, przeciwdziałaniem i zwalczaniem COVID-19, Dz.U. 2020 r., poz. 410.
- Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 11.03.2020 r. w sprawie czasowego ograniczenia funkcjonowania niektórych podmiotów systemu szkolnictwa wyższego i nauki w związku z zapobieganiem, przeciwdziałaniem i zwalczaniem COVID-19, Dz.U. 2020 r., poz. 405.
- Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 13.03.2020 r. w sprawie przywrócenia tymczasowo kontroli granicznej osób przekraczających granicę państwową stanowiącą granicę wewnętrzną, Dz.U. 2020 r., poz. 434.
- Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 13.03.2020 r. w sprawie czasowego zawieszenia lub ograniczenia ruchu granicznego na określonych przejściach granicznych, Dz.U. 2020 r., poz. 435.
- Rozporządzenie Ministra Zdrowia z dnia 11.03.2020 r. w sprawie czasowego ograniczenia funkcjonowania uczelni medycznych, Dz.U. 2020 r., poz. 406.
- Rozporządzenie Ministra Zdrowia z dnia 13.03.2020 r. w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu zagrożenia epidemicznego, Dz.U. 2020 r., poz. 433.
- Rozporządzenie Ministra Zdrowia z dnia 20.03.2020 r. w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu epidemii, Dz.U. 2020 r., poz. 491.
- Rozporządzenie Ministra Zdrowia z dnia 24.03.2020 r. zmieniające rozporządzenie w sprawie ogłoszenia na obszarze Rzeczypospolitej Polskiej stanu epidemii, Dz.U. 2020 r., poz. 522.
- Rozporządzenie Rady Ministrów z dnia 13.03.2020 r. w sprawie zakazów w ruchu lotniczym, Dz.U. 2020 r., poz. 436.

- Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2018 r. (2019), Warszawa, https://www.nfz.gov.pl/gfx/nfz/userfiles/_public/bip/uchwaly_rady/sprawozdanie_z_dzialalnosci_nfz_za_2019_rok.pdf (dostęp: 12.06.2020).
- Tibshirani R., Walther G., Hastie T. (2001), *Estimating the Number of Clusters in a Data Set via the Gap Statistic*, "Journal of the Royal Statistical Society B", Vol. 63(2), s. 411-423.
- Vavrek M.J. (2011), *Fossil: Palaeoecological and Palaeogeographical Analysis Tools*, "Palaeontologia Electronica", No. 14:1T, http://palaeo-electronica.org/2011_1/238/index.html (dostęp: 12.06.2020).
- Ward J.H. (1963), *Hierarchical Grouping to Optimize an Objective Function*, "Journal of the American Statistical Association", Vol. 58(301), s. 236-244.
- [www 1] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200611-covid-19-sitrep-143.pdf?sfvrsn=2adbe568_4 (dostęp: 12.06.2020).
- [www 2] https://twitter.com/MZ_GOV_PL (dostęp: 1.06.2020).
- [www 3] <https://bdl.stat.gov.pl/BDL/dane/podgrup/temat> (dostęp: 20.04.2020).

HEALTHCARE EXPENDITURE AND THE COURSE COVID-19 EPIDEMIC IN POLISH VOIVODESHIPS IN THE PERIOD MARCH 4 – MAY 31, 2019

Summary: The goal of the paper is to examine the differences and similarities between Polish voivodships in the field the development of the COVID-19 epidemic in the period from March 4 to May 31, 2020. Healthcare financing in 2018 was taken into account. Two approaches were used to study the development of the epidemic. Results indicate similarities between voivodships. For analysis based on the values of variables describing the epidemic development on each day of the studied period, the number of deaths per 10 thousand people turned out to be important for the final classification. In the alternative approach, dynamics was taken into account. The number of cases per 10 thousand people was a decisive factor in the classification. The similarities of groupings in terms of financing and the course of the epidemic are not large, the relatively smallest differences occur in the case of the classification taking into account deaths.

Keywords: coronavirus, COVID-19, Ward's method.