# „Analiza wielowymiarowa jakości służby zdrowia w wybranych krajach OECD”

# Cel Badania:

Celem przeprowadzonej analizy jest dokonanie klasyfikacji wybranych państw pod względem jakości służby zdrowia oraz podział na grupy z wyszczególnieniem cech charakterystycznych dla każdej z nich.

# Patryk Szymkowiak

# Krzysztof Ignasiak

# Wprowadzenie

W dobie ogólnoświatowej pandemii bardziej niż kiedykolwiek zainteresowanie budzi poziom oraz efektywność służby zdrowia. W naszym projekcie podjęliśmy próbę porównanie jakości opieki medycznej w różnych krajach. Dane zostały pozyskane z oficjalnych zasobów OECD(Organization for Economic Co-Operation and Development) międzynarodowej organizacji wspierającej państwa członkowskie w osiągnięciu jak najwyższego poziomu wzrostu gospodarczego i stopy życiowej obywateli. Ze względu na źródło danych wszystkie analizowane państwa należą do wyżej wskazanej organizacji, ponadto zważywszy na kompletność danych do analizy zostało włączone 21 państw tj. Korea Południowa, Meksyk, Republika Czeska, Estonia, Węgry, Litwa, Łotwa, Polska, Słowacja, Austria, Belgia, Finlandia, Francja, Irlandia, Islandia, Izrael, Włochy, Luksemburg, Słowenia, Hiszpania oraz Holandia.

# Wybór i charakter zmiennych

Określiliśmy sześć zmiennych objaśniających:

1. Liczba lekarzy na tysiąc mieszkańców(stymulanta)
2. Liczba szpitali na milion mieszkańców( stymulanta)
3. Zgony, którym można było zapobiec w przeliczeniu na 100 tysięcy osób (destymulanta)
4. Liczba łóżek szpitalnych na 1000 mieszkańców(stymulanta)
5. Ilość urządzeń służących do rezonansu magnetycznego na milion mieszkańców(stymulanta)
6. Śmiertelność noworodków na tysiąc porodów.(destymulanta)

Zmienne zostały podzielone ze względu na charakter na stymulanty i destymulanty, przy czym pierwsza z grup zawiera charakterystyki o pozytywnej interpretacji, których wysoki poziom świadczy o jakości służby zdrowia. Natomiast destymulanty posiadają negatywną interpretacje, a ich wysokie wartości świadczą o problemach w opiece medycznej.

Pierwsza ze zmiennych określa liczbę aktywnie pracujących w swoim zawodzie przedstawicieli wszystkich specjalności lekarskich w przeliczeniu na tysiąc mieszkańców danego państwa. Uważamy to, za niezwykle istotną stymulantę, ponieważ każdemu choremu przede wszystkim potrzebny jest specjalista, który będzie wstanie zdiagnozować jego przypadłość i zastosować odpowiednią kurację. Znaczna liczba aktywnych zawodowo lekarzy implikuje ich wysoką dostępność oraz świadczy o tym, że są wstanie więcej czasu poświęcić konkretnemu pacjentowi, co jest szczególnie istotne w kontekście trudnych i rzadkich przypadków.

Drugą i czwartą zmienną również w oczywisty sposób należy zaliczyć do stymulant. Każda osoba cierpiąca na określone schorzenie powinna mieć możliwość bezproblemowego dotarcia do miejsca, gdzie może liczyć na pomoc specjalistów oraz gdzie będzie mogła zostać, jeżeli będzie wymagała opieki. Zmienne te choć na pierwszy rzut oka bardzo podobne uznajemy, za wzajemnie uzupełniające się. Mianowice: Liczba szpitali warunkuje dostępność miejsc, w których można liczyć na specjalistyczną pomoc, im ich więcej tym należy uznać, że statystycznie łatwiej dostać się konkretnemu obywatelowi do takiego punku lub zostać do niego przetransportowanym. Z kolei liczba łóżek szpitalnych świadczy o przepustowości opieki zdrowotnej. Jest to szczególnie istotne w przypadku pandemii lub kataklizmu. Innymi słowy nieuwzględnienie liczby szpitali dawało by informacje niepełną, przykładowo można sobie wyobrazić państwo, które posiada odpowiedne środki oraz specjalistów, lecz te zasoby są zlokalizowane w przeważającej większości w rozbudowanych szpitalach w najbardziej istotnych ośrodkach miejskich, a mieszkańcy terenów rzadziej zaludnionych mają problemy z uzyskaniem profesjonalne pomocy.

Trzecie zmienna tj. Zgony którym można było zapobiec w przeliczeniu na 100 tysięcy osób ( ang. Avoidable mortality) definiowana jest przez OECD w sposób następujący: "Causes of death that can be mainly avoided through effective public health and primary prevention interventions", co w wolnym tłumaczeniu oznacza te przypadki zgonów, których w większości dałoby się uniknąć, gdyby tylko zastosowana odpowiednie środki zapobiegawcze/ interwencyjne. Za pomocą takich informacji można w bezpośredni sposób wnioskować o jakości służby zdrowia. Niska wartość tego współczynnika świadczy zarówno o skuteczności podejmowanych działań ratunkowych jak i prewencyjnych, dlatego uznajemy tę zmienną za destymulantę.

Podobnie traktujemy zmienną szóstą, która określa śmiertelność noworodków w przeliczeniu na tysiąc porodów. Uznajemy, iż odsetek przypadków, w których czynniki zewnętrzne takie jak choroby genetyczne, spożywanie alkoholu lub innych używek w trakcie ciąży, urazy i wypadki doznane przez matki nie różni się w sposób statystycznie istotny pomiędzy poszczególnymi krajami. W konsekwencji przyjmujemy, że główną przyczyną warunkującą śmierci noworodków jest jakość opieki prenatalnej. Niewielka wartość tego współczynnika oznacza, że w danych kraju wypadki medyczne dotyczące noworodków zdarzają się relatywnie rzadko, a kobietom ciężarnych zapewniane jest odpowiednia pomoc, gdy tylko jest to konieczne.

Natomiast zmienna piąta określająca ilość sprzętu do rezonansu magnetycznego w przeliczeniu na milion mieszkańcu stanowi uproszczoną miarę określającą dostępność nowoczesnych sprzętów wspomagających proces diagnozowania oraz leczenia.

# Metodyka

1. Mierniki syntetyczne

W celu skutecznego wyłonienia państwa o najlepszym/ najgorszym ogólnym poziomie służby zdrowia skonstruowano dwa mierniki syntetyczne : miernik bezwzorcowy ze zróżnicowanymi wagami oraz TOPSIS. Za ich pomocą możliwe jest stworzenie rankingu obejmującego wiele zmiennych różniących się charakterem, poziomem wielkości i zmiennością. Aby to osiągnąć, konstruuje się zmienną syntetyczną złożoną ze wszystkich pozostałych. W przypadku obydwu metod na początku konieczne jest ujednolicenie wszystkich zmiennych do stymulant. Następnie przeprowadza się normalizacje, której celem jest ustanowienie jednolitego poziomu wielkości( aby te czynniki, które są wyrażone w tysiącach nie miały większego wpływu od tych wyrażonych w jednostkach). Następnie w przypadku miernika bezwzorcowego oblicza się średnią ważoną z tak zmodyfikowanych zmiennych, przy czym wagi to ilorazy średniej z wartości danego czynnika i jego odchylenia standardowego.

Natomiast w przypadku TOPSIS najpierw z każdej zmiennej wybiera się wzorzec oraz antywzorzec i na tej podstawie zostaje skonstruowany miernik.

1. Metody aglomeracyjne

Poza uszeregowaniem przypadków od najlepszych do najgorszych, zależało nam także na dokonaniu podziału wszystkich obserwacji na zróżnicowane grupy. W celu dokonania wstępnego podziału oraz ustalenia liczby skupień wykorzystano będącą jedną z typów grupowania hierarchicznego metodę aglomeracyjną. Algorytm rozpoczyna się od potraktowania każdej obserwacji jako jednoelementowego klastra. Następnie w każdej iteracji dwa najbardziej zbliżone klastry są łączone w grupę. Procedura kończy się w momencie, gdy wszystkie obserwacje znajdują się w ramach jednej grupy. Aby możliwe było użycie wyżej opisanej metody konieczne jest zdefiniowanie metryki odległości oaz sposobu łączenia klastrów. Pierwszy z parametrów określa w jaki sposób definiowana jest odległość pomiędzy poszczególnymi grupami. Postanowiliśmy wykorzystać odległość euklidesową. Natomiast w kontekście metody połączeń postanowiliśmy trzy razy dokonać grupowania z wykorzystaniem różnych sposobów połączeń: metody Warda, pełnych wiązań oraz średnich połączeń. W przypadku metody pełnych wiązań odległość pomiędzy grupami równa jest odległości pomiędzy dwoma najdalszymi obiektami znajdującymi się w dwóch różnych klastrach. Z kolei w przypadku średnich połączeń odległość pomiędzy grupami warunkowana jest przez średnie arytmetyczne w poszczególnych klastrach. Natomiast w przypadku metody Warda odległość między grupami jest sumą kwadratów odchyleń obserwacji od centroidów skupień.

1. Metoda K-średnich

W celu polepszenia jakości grupowania zastosowano algorytm k-średnich. Procedura polega na przenoszeniu obiektów z jednego skupienia do innego tak długo aż podobieństwo wewnątrz skupień będzie jak największe, a między skupieniami jak najmniejsze. Postanowiono wykorzystać trzy wersje metody k-średniej, które różnią się tym, w jakim sposób zostały ustalone wstępne centra skupień. Pierwszy sposób polegał na wyborze obserwacji tak aby zmaksymalizować odległości między skupieniami, drugi polegał na losowaniu co n-tej obserwacji z posortowanego zbioru, a ostatnia z metod po prostu przyjmowała n pierwszych obserwacji za wstępne centra skupień, gdzie n oznacza liczbę grup.

1. ANOVA

Jako ostatni element analizy zostanie przeprowadzona ANOVA, która pozwala stwierdzić, czy różnice pomiędzy średnimi poziomami danej zmiennej w grupach są różne od siebie w sposób istotnie statystyczny. Należy podkreślić, że zostanie wykorzystana jednoczynnikowa analiza wariancji dla każdej zmiennej z osobna co oznacza przyjęcie niejawnego założenia, iż poszczególne zmienne są od siebie niezależne.

# Rezultaty

1. Mierniki Syntetyczne

TOPSIS

|  |  |
| --- | --- |
| Korea | 0,743 |
| Austria | 0,522 |
| Finland | 0,515 |
| France | 0,460 |
| Italy | 0,443 |
| Iceland | 0,440 |
| Estonia | 0,405 |
| Lithuania | 0,392 |
| Czech Republic | 0,383 |
| Slovenia | 0,379 |
| Spain | 0,365 |
| Latvia | 0,352 |
| Poland | 0,352 |
| Netherlands | 0,351 |
| Ireland | 0,325 |
| Belgium | 0,310 |
| Israel | 0,310 |
| Hungary | 0,300 |
| Luxembourg | 0,296 |
| Slovak Republic | 0,293 |
| Mexico | 0,199 |

Po ustaleniu klasyfikacji został dokonany podział analizowanych krajów na cztery grupy. Kolor niebieski oznacza bardzo wysoki poziom służby zdrowia, żółty(pomarańczowy) świadczy o lepszej( gorszej) niż przeciętna jakości opieki medycznej, a czerwony wskazuje na zły poziom służby zdrowia.

Jakość służby zdrowia okazała się być najlepsza w Korei, ze względu na to, że ten azjatycki kraj jest liderem pod względem dostępności szpitali oraz łóżek szpitalnych a także wyposażenia w sprzęt do rezonansu magnetycznego (dokładniej zostało to omówione przy przedstawieniu wyników grupowania).

Z kolei Austria posiada najwięcej praktykujących lekarzy w przeliczeniu na tysiąc mieszkańców, co w połączeniu z jednym z najniższych współczynników śmiertelności noworodków zapewniło jej drugie miejsce.

Warto też podkreślić zasadność zaklasyfikowania Meksyku do osobnej kategorii ze względu na niezwykle wysoką śmiertelność noworodków oraz niewielką ilość sprzętu do rezonansu magnetycznego.

Miernik Bezwzorcowy

|  |  |
| --- | --- |
| Korea | 0,741 |
| Finland | 0,559 |
| Austria | 0,557 |
| Iceland | 0,533 |
| Italy | 0,509 |
| Estonia | 0,447 |
| France | 0,444 |
| Slovenia | 0,436 |
| Spain | 0,425 |
| Czech Republic | 0,396 |
| Netherlands | 0,381 |
| Lithuania | 0,367 |
| Ireland | 0,364 |
| Latvia | 0,327 |
| Israel | 0,325 |
| Belgium | 0,315 |
| Poland | 0,308 |
| Luxembourg | 0,308 |
| Hungary | 0,252 |
| Slovak Republic | 0,251 |
| Mexico | 0,094 |

Należy zauważyć, że również przy użyciu miernika bezwzorcowego za państwa z najlepszą służbą zdrowia zostały uznane Korea Południowa, Finlandia i Austria.

Natomiast poza Meksykiem do najgorszej kategorii zostały zaliczone także Węgry i Słowacja ze względu na odpowiednio bardzo dużą liczbę zgonów, których można było uniknąć, oraz wysoki odsetek śmiertelności noworodków.

Wysoki wskaźnik avoidable mortality był również przyczyną tego, że w tej klasyfikacji Litwa została uznana za kraj o poniżej przeciętnej jakości opieki medycznej. A Słowenia została zakwalifikowana do drugiej najwyższej grupy ze względu na najniższą śmiertelność niemowląt wśród analizowanych państw (ex aequo z Islandią).

1. Metody aglomeracyjne

W przypadku każdego z wariantów za miejsce odpowiednie do „cięcia drzewa” uznaliśmy łączenia z którymi związane były największe przyrosty względne i bezwzględne, które w przypadku analizowanych przez nas danych wskazywały te same miejsca. W rezultacie otrzymaliśmy cztery grupy przy użyciu metody Warda oraz trzy przy wykorzystaniu pozostałych metod, jednakże w takim przypadku otrzymalibyśmy jedno skupienie zawierające 19 z 21 obserwacji oraz dwa klastry jednoelementowe. Uznaliśmy taką sytuację za niepożądaną i dokonaliśmy podziału na cztery grupy zgodnie z wynikami metody Warda.

Otrzymany podział prezentuje się następująco:

Grupa 1

* Korea Południowa

Grupa 2

* Meksyk

Grupa 3

* Słowacja
* Polska
* Litwa
* Łotwa
* Węgry
* Estonia
* Słowenia
* Republika Czeska
* Austria

Grupa 4

* Włochy
* Hiszpania
* Irlandia
* Francja
* Finlandia
* Holandia
* Islandia
* Izrael
* Luksemburg
* Belgia.

1. Metody K-średnich

Z pomocą miernika jakości grupowania wybrano ostateczny podział.

Pogrupowanie niewiele różni się od otrzymanego z pomocą metod aglomeracyjnych, jedynie Austria i Słowenia została przeniesiona z grupy trzeciej do czwartej. Zmiana ta była zdecydowanie uzasadniona ze względu na odpowiednio najwyższy odsetek lekarzy i najniższy współczynnik śmiertelności wśród noworodków( ex aequo z Islandią). Ostateczny skład grup prezentuje się następująco:

Grupa 1

* Korea Południowa(Lider pod względem infrastruktury i sprzętu)

Grupa 2

* Meksyk (Antywzorzec)

Grupa 3 (Kraje rozwijające się)

* Słowacja
* Polska
* Litwa
* Łotwa
* Węgry
* Estonia
* Republika Czeska

Grupa 4 (Kraje rozwinięte)

* Słowenia
* Włochy
* Hiszpania
* Irlandia
* Francja
* Finlandia
* Holandia
* Islandia
* Izrael
* Luksemburg
* Austria
* Belgia

Nie trudno było wyszczególnić konkretne cechy, które posiadają państwa należące do poszczególnych klastrów.

Koreę jako jedynego przedstawiciela grupy pierwszej należy uznać za lidera pod względem infrastruktury oraz sprzętu. Kraj ten deklasuje w sposób znaczący wszystkie pozostałe pod względem liczby szpitali, dostępności łóżek szpitalnych oraz ilości sprzętu do rezonansu magnetycznego. Widać to w szczególności w zestawieniu z drugimi w kolejności państwami w tych kategoriach. W Korei na milion mieszkańców przypada prawie 78 szpitali, podczas gdy we Francji tylko około 45 tego typu placówek. Na każdy tysiąc obywateli przygotowanych jest ponad 12 łóżek szpitalnych, a w Austrii zaledwie około 7. Na wyposażeniu służby zdrowia jest średnio 30 urządzeń do rezonansu magnetycznego, podczas gdy najwyższa średnia z grupy czwartej wynosi niecałe 17, a wicelider pod tym względem tj. Włochy posiadają takich urządzeń około 29. Szczególnie w przypadku dwóch pierwszych miar widać prawdziwą przepaść dzielącą to azjatyckie państwo od innych krajów, w tym wysoko rozwiniętych. To sprawia, że ilość śmierci, którym w znacznym stopniu dałoby się zapobiec oraz śmiertelność noworodków utrzymuje się na relatywnie niskim (choć zdecydowanie nie najniższym poziomie) i to mimo posiadania najmniejszej liczby lekarzy na tysiąc mieszkańców.

Z kolei jedynego przedstawiciela grupy drugiej tj. Meksyk można uznać za antywzorzec. Przede wszystkim w tym południowoamerykańskim kraju obserwuje się niezwykle wysoką śmiertelność noworodków. Prawie 13 na 1000 dzieci umiera w wieku niemowlęcym. W Słowacji, która jako druga najgorzej radzi sobie z tą kwestią, zgonów tego typu jest około pięciu. Kraj ten również posiada największe braki w kontekście sprzętu do rezonansu magnetycznego. Na milion mieszkańców przypadają niecałe trzy tego typu urządzenia. W porównaniu na Węgrzech( które są drugim z państw najgorzej wyposażonych w tego typu sprzęt) jest ich prawie 5. Co również warte podkreślenia mimo posiadania relatywnie wysokiej liczby szpitali, dostępność łóżek jest również najgorsza w ramach analizowanych państw. Na tysiąc obywateli ledwie jeden może liczyć na miejsce w odpowiedniej jednostce.

Na wysokim poziomie utrzymuje się także ilość zgonów, którym można było zapobiec( porównywalnie do grupy trzeciej), a liczba lekarzy również jest bardzo niska i niewiele przekracza tą dotyczącą Korei Południowej.

Podstawą kwestią rozróżniającą trzecią i czwartą grupę jest ilość zgonów którym można było zapobiec. W krajach rozwijających się śmierci tego typu jest średnio ponad dwukrotnie więcej niż w krajach rozwiniętych. Mało tego w grupie trzeciej są kraje, które pod tym względem mają nawet większy problem niż Meksyk. Na Łotwie, Litwie oraz na Węgrzech było odpowiednio 269, 267 oraz 249 zgonów którym można było zapobiec na 100 tysięcy mieszkańców w porównaniu do 213 w Meksyku.

Kraje wchodzące w skład grupy trzeciej odnotowywały także większy odsetek śmierci niemowląt i posiadały średnio mniejszą ilość sprzętu do rezonansu magnetycznego.

Co ciekawe to właśnie w krajach rozwijających się jest więcej łóżek szpitalnych w przeliczeniu na 1000 mieszkańców, a także w niewielkim stopniu większa jest średnia liczba szpitali. W ramach czwartej grupy znalazł się także Izrael, który mimo posiadania najmniejszej liczby szpitali na milion mieszkańców wśród wszystkich badanych państw oraz prawie najmniejszej( po Meksyku i Węgrzech) ilości sprzętu do rezonansu magnetycznego, posiadał jednocześnie najmniejszy współczynnik zgonów, którym można było zapobiec.

Warto też podkreślić, że dwa spośród krajów rozwiniętych tj. Słowenia i Irlandia osiągnęły niezwykle wysoki poziom opieki prenatalnej, a niewielki odsetek śmierci wśród niemowląt wyróżnia je nie tylko w porównaniu do wszystkich analizowanych krajów, ale także do tych wchodzących w skład tego samego klastra.

1. ANOVA

Różnice w poziomach wszystkich zmiennych w poszczególnych grupach okazały się różne od siebie w sposób statystycznie istotny, z wyjątkiem liczby lekarzy na tysiąc mieszkańców.

Jednakże po wielokrotnym przeprowadzeniu testu Tukey’a ustalono, że istotne różnice występują tylko pomiędzy poszczególnymi parami grup.

Przede wszystkim Odsetek śmiertelności noworodków w Meksyku istotnie różnił się od pozostałych grup. Jest to oczywisty rezultat mając na względzie, że w pozostałych grupa umiera średnio około 3 ze 100 noworodków, a w Meksyku prawie 13.

Nie dziwi również fakt, że ilość sprzętu do rezonansu magnetycznego istotnie różni się w Korei Południowej oraz Meksyku, gdyż pierwszy z tych krajów jest najlepiej wyposażony pod tym względem, a drugi najgorzej.

Dodatkowo liczba łóżek szpitalnych w przeliczeniu na tysiąc mieszkańców w Korei jest zauważalnie większa w porównaniu do pozostałych grup, podobna sytuacja ma miejsce w kontekście ilości szpitali na milion mieszkańców, jednakże w tym przypadku różnica pomiędzy Koreą i Meksykiem była minimalnie poniżej wartości krytycznej.

Ponad to, jak już wcześniej wspominano odsetek śmierci, którym można było zapobiec stanowi główną cechę rozróżniającą kraje rozwijające się od rozwiniętych, co również znalazło potwierdzenie w wynikach testu Tukey’a.

# Podsumowanie

Wyniki przeprowadzonej analizy pokazują jak złożonym i skomplikowanym zjawiskiem jest jakość opieki zdrowotnej. Możemy zauważyć strukturalne różnice w jakości opieki medycznej, które są skorelowane z położeniem geograficznym. W szczególności możemy zaobserwować różnice pomiędzy krajami Europy Zachodniej oraz Środkowo-Wschodniej. Zaskakiwać mogą zróżnicowane wyniki dotyczące śmierci, którym można było zapobiec oraz ilości sprzętu do rezonansu magnetycznego, a także współczynnika śmiertelności noworodków.

Widzimy potrzebę dalszych badań tego typu, gdyż uważamy, że rezultaty takich badań mogą stanowić podstawę do bardziej szczegółowych analiz, z pomocą których możliwe będzie tworzenie rekomendacji i zaleceń mających na celu poprawę jakości służby zdrowia w poszczególnych krajach.