

## Analiza Danych Przestrzennych w programie GeoDa

### Instalacja programu GeoDa

Pobierz program GeoDa ( <https://geodacenter.github.io/> ) i zainstaluj postępując zgodnie z instrukcją.

1. Pobierz dane ze strony: <https://geodacenter.github.io/data-and-lab/ohiolung/>. Folder ten zawiera informacje dotyczące zachorowań na raka płuc w Ohio w latach: 1968, 1978, 1988. Zbiór danych **OHIO LUNG** zawiera plik tekstowy **ohioutmbnd.txt** ze współrzędnymi granic 88 hrabstw Ohio projektowanych przy użyciu UTM-17. Wczytaj utworzony na tej podstawie **polygon shape file 'ohlung.shp'**.
2. Wyświetl 3 mapy kwantylowe , aby zbadać jak zmieniała się śmiertelność na raka płuc białych kobiet w 1968, 1978 i 1988 roku. Musisz wybrać ilość kategorii dla **quantile map**- w tym ćwiczeniu utwórz mapę kwantylową (4 grupy). Wczytaj odpowiednio zmienne: **LFW68, LFW78, LFW88**.
3. W tym ćwiczeniu chcemy przedstawić mapę chorób jako rate maps, które są szczególnym przypadkiem map Choropleta. Wybierz z menu funkcję **Map-> Rates-Calculated Map-> Raw Rate**. Pojawi się okno dialogowe, w którym należy wybrać zmienną zdarzeń (**Event variables**) oraz zmienną bazową (**Base variables**).

Wybierz LFW68 jako zdarzenie (całkowita liczba zgonów białych kobiet na raka płuc w 1968 roku) i POPFW68 jako populację zagrożoną (całkowita liczba białych kobiet w 1968 roku). Następnie należy wybrać odpowiedni typ mapy z rozwijanej listy. Domyślnie jest to mapa percentylowa, ale nie byłoby to właściwe w tym przykładzie (Ohio ma 88 hrabstw, czyli mniej niż 100 wymaganych dla mapy percentylowej). Zamiast tego należy wybrać opcję **Box Map** z wartością **Hinge= 1,5**. Trzy hrabstwa pojawiają się jako górne wartości odstające o podwyższonej śmiertelności.

Zapisz obliczone wskaźniki śmiertelności, aby można było je wykorzystać do dalszych analiz. Kliknij prawym przyciskiem myszy na mapę i wybierz opcję **Save Rates** do utworzenia zmiennej. Zapisz wynik do domyślnej zmiennej o nazwie **R\_RAW\_RT**.

4. Na podstawie utworzonej zmiennej wyświetl **Box Plot (Explore-> Box Plot)**. Znajdź na wykresie 3 górne wartości odstające i klikając na te punkty, sprawdź które to hrabstwa (Podświetlą się na mapie bazowej).
5. Powszechnie stosowanym pojęciem w analizie zdrowia publicznego jest pojęcie znormalizowanego współczynnika umieralności (**SMR- standardized mortality rate**), czyli stosunek obserwowanego współczynnika umieralności do standardu krajowego (lub regionalnego).

Formalnie można to wyrazić wzorem:

$$\tilde{\pi} = \frac{\sum_{i=1}^n O_i}{\sum_{i=1}^n P_i},$$

co daje wartość oczekiwaną zdarzeń dla każdego obszaru:

$$E_i = \tilde{\pi} P_i$$

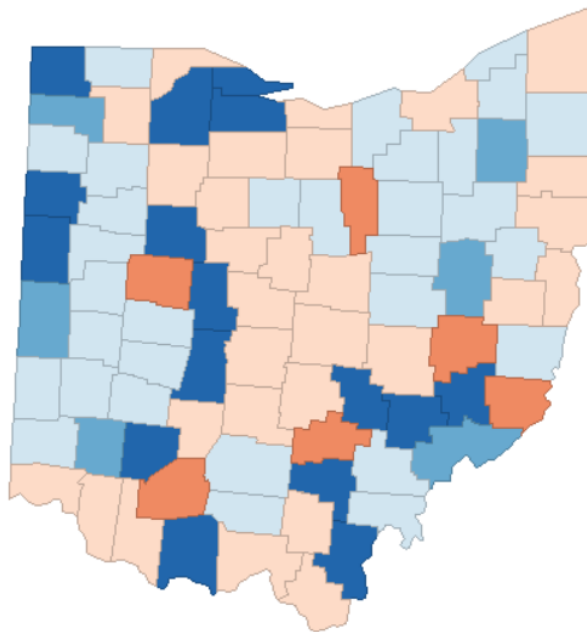
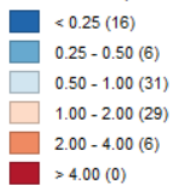
Względne ryzyko jest proporcjonalne do stosunku zaobserwowanej liczby zdarzeń (np. przypadków nowotworów) do wartości oczekiwanej:

$$SMR_i = \frac{O_i}{E_i}$$

GeoDa realizuje to w formie mapy nadmiernego ryzyka (Excess Risk Map). Funkcję tę można znaleźć w opcji **Map-> Rates-Calculated Map-> Excess Risk**.

Nadmierne ryzyko to stosunek obserwowanej stopy do średniej stopy obliczonej dla wszystkich danych. Jest on obliczany jako stosunek całkowitej sumy wszystkich zdarzeń do całkowitej sumy wszystkich zagrożonych populacji (np. w naszym przykładzie wszystkie zgony białych kobiet w stanie nad populacją białych kobiet w stanie). Wybierz LFW68 jako zdarzenie (całkowita liczba zgonów białych kobiet na raka płuc w 1968 roku) i POPFW68 jako populację zagrożoną (całkowita liczba białych kobiet w 1968 roku).

Excess Risk Map: LFW68 over POPFW68



Zapisz obliczone wartości dla hrabstw pod nazwą R\_EXCESS.