

Cele

1. Estymacja wskaźnika zagrożenia ubóstwem w przekroju podregionów Polski.
2. Zbudowanie dobrego modelu wyjaśniającego poziom zagrożenia ubóstwem.

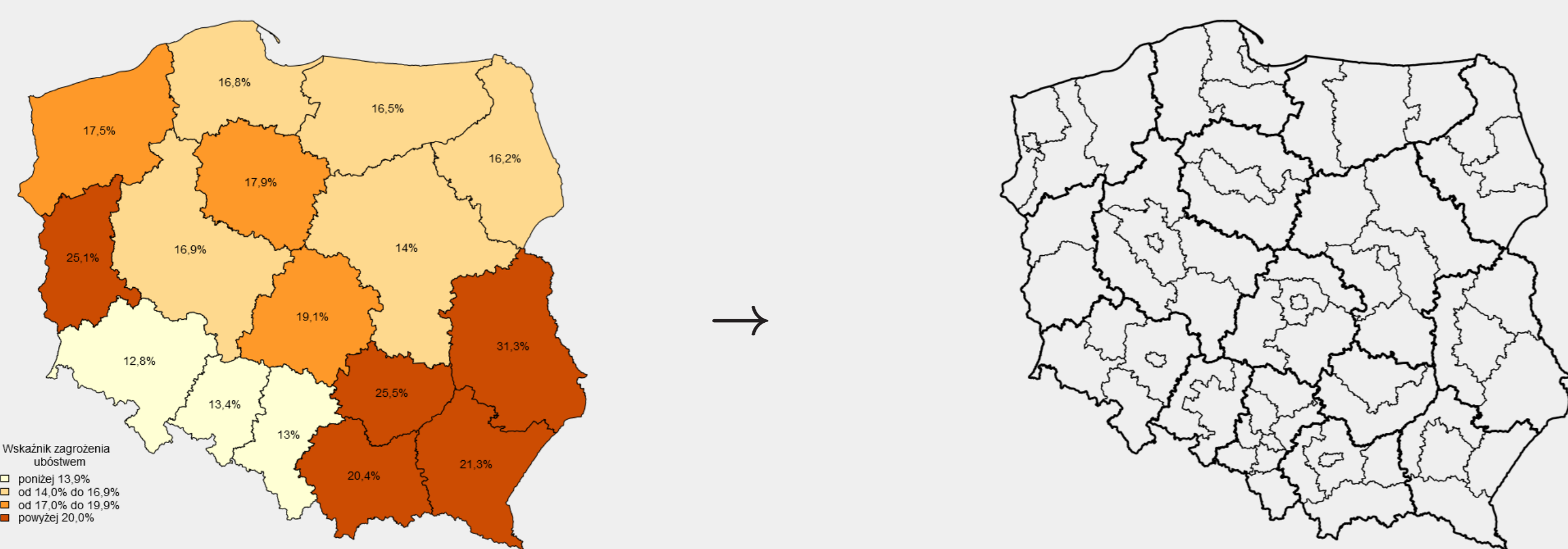
Motywacja

Przedstawione zagadnienia są efektem podjętej współpracy Ośrodka Statystyki Małych Obszarów w Urzędzie Statystycznym w Poznaniu, Departamentu Badań Społecznych i Warunków Życia w Głównym Urzędzie Statystycznym oraz Banku Światowego w ramach projektu *Poverty Mapping in the new EU Member States*.

Wprowadzenie

- ▶ Mapowanie ubóstwa - mapy jako narzędzia do dostarczenia oszacowań wskaźnika zagrożenia ubóstwem w szczegółowych przekrojach geograficznych. Oprócz zwięzłego podsumowania danych pozwala także na interpretację wyników z uwzględnieniem zależności terytorialnych, co nie jest możliwe w przypadku zwykłej tabeli [Bedi i in. 2007].
- ▶ Statystyka małych obszarów - grupa metod pozwalających na wiarygodną estymację na podstawie nielicznej próby przy wykorzystaniu dostępnych informacji dodatkowych [Rao 2003].

Wskaźnik zagrożenia ubóstwem - EU-SILC



Rysunek 1: Przekrój województw

Rysunek 2: Przekrój podregionów

Model Faya-Herriota

- ▶ Model liniowy z efektem losowym na poziomie obszaru został zaproponowany w 1979 roku jako model do estymacji dochodu na osobę w obszarach z populacją mniejszą niż 1000 osób [Hansen i in. 2011]. Przedstawia się następująco:

$$\hat{\theta}_d = \mathbf{x}_d^t \beta + \mathbf{u}_d + \mathbf{e}_d, \quad (1)$$

gdzie: $\hat{\theta}_d$ - zmienna objaśniana dla obszaru d , \mathbf{x}_d^t - wektor zmiennych objaśniających dla obszaru d , β - wektor parametrów, \mathbf{u}_d - efekt obszaru, \mathbf{e}_d - błąd losowy.

- ▶ Dla obszarów, które znalazły się w próbie estymator przyjmuje postać:

$$\tilde{\theta}_d = \mathbf{x}_d^t \hat{\beta} + \hat{\mathbf{u}}_d, \quad (2)$$

natomiast dla obszarów, które nie znalazły się w próbie rozważany jest estymator syntetyczny regresyjny:

$$\tilde{\tilde{\theta}}_d = \mathbf{x}_d^t \hat{\beta} \quad (3)$$

Estymator BLUP w modelu Faya-Herriota

- ▶ Estymatorem BLUP jest średnią ważoną estymatora bezpośredniego oraz regresyjnego syntetycznego:

$$\tilde{\theta}_d^{BLUP} = \gamma_d \hat{\theta}_d + (1 - \gamma_d) \mathbf{x}_d^t \hat{\beta}, \quad (4)$$

gdzie: $\gamma_d = \sigma_v^2 / (\sigma_e^2 + \sigma_v^2) \wedge \gamma_d \in \langle 0; 1 \rangle$

- ▶ Waga γ_d mierzy niepewność w modelowaniu θ_d czyli wariancję modelową σ_v^2 względem wariancji całkowitej $\sigma_e^2 + \sigma_v^2$. W związku z tym estymator BLUP dobrze uwzględnia relację między zmiennością w obszarze, a precyzją estymatora bezpośredniego [Rao 2003].

Objaśnienia

¹ Zmienna przyjmująca wartość 1 w przypadku, gdy logarytm wielkości populacji w podregionie jest mniejsza od 33 percentyla wartości logarytmu populacji w podregionach oraz 0 w przeciwnym razie.

Tworzenie modelu

- ▶ Do celów estymacji utworzono zbiór cech dodatkowych dotyczących m.in. demografii, aktywności ekonomicznej, wykształcenia czy poziomu życia na poziomie podregionów (NUTS3) pochodzących przede wszystkim z Banku Danych Lokalnych oraz Narodowego Spisu Powszechnego 2011. Następnie wybrano zmienne wysoko skorelowane z cechą objaśnianą - wskaźnikiem zagrożenia ubóstwem w podregionach Polski. Podczas dalszego etapu prac były one wykorzystywane w różnych układach jako zmienne objaśniające przy konstrukcji modeli statystycznych. Poniżej zamieszczono jedną z roboczych postaci modelu, który został wypracowany wspólnie z ekspertami z Banku Światowego.

Tablica 1: Rezultaty regresji liniowej

	Ocena	Odch.Std.	p-value	
(Stała)	0,8486	0,1957	0,0001	***
Udział osób samotnych (powyżej 25 roku życia) w populacji	1,3100	0,5217	0,0147	*
Zmienna dychotomiczna ¹	0,0347	0,0125	0,0074	**
Liczba pokoi w przeliczeniu na 1 osobę	-0,1678	0,0708	0,0210	*
Udział gosp. dom. z wanną lub prysznicem w ogólnej liczbie gosp. dom.	-0,9022	0,1336	0,0000	***
Udział osób w wieku 0-14 w populacji	0,7826	0,4275	0,0721	.

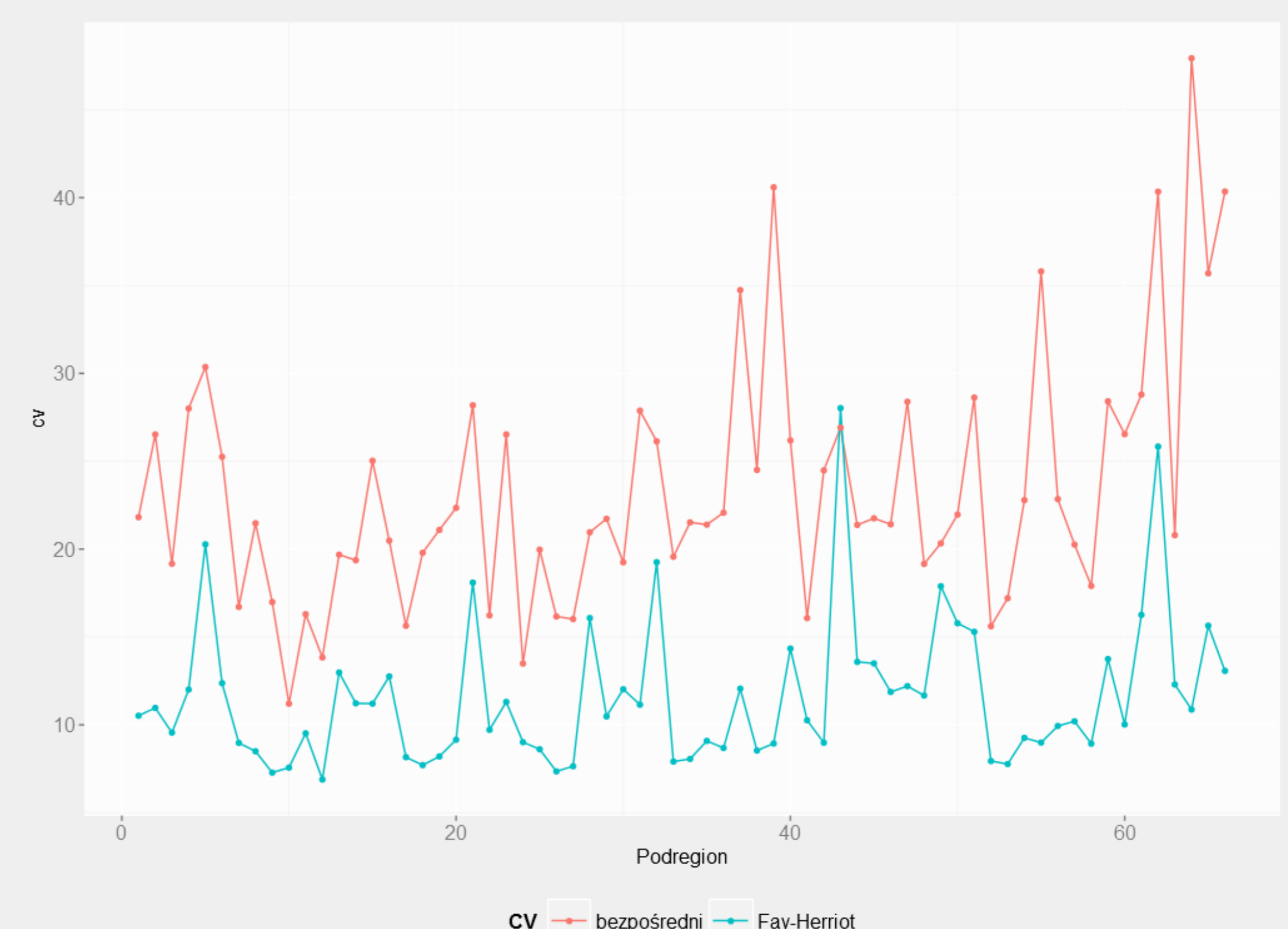
Poziomy istotności: 0 *** 0,001 ** 0,01 * 0,05 .

Skorygowana wartość współczynnika R^2 wyniosła 62%.

- ▶ Trwają prace mające na celu polepszenie dopasowania modelu poprzez włączenie zmiennych specyficznych dla Polski opisujących migracje ludności, kontekst historyczny czy grupowanie podregionów.

Wyniki

- ▶ Zastosowanie powyższego modelu Faya-Herriota pozwoliło na znaczną redukcję błędu standardowego szacunków wskaźnika zagrożenia ubóstwem.



Rysunek 3: Współczynnik zmienności

Wnioski

- ▶ Wykorzystanie metod statystyki małych obszarów pozwoliło na otrzymanie szacunków w dotąd niepublikowanym przekroju terytorialnym.
- ▶ Otrzymane szacunki cechowały się dostateczną precyzją.
- ▶ W dalszych krokach należy poszukiwać zmiennych pomocniczych skorelowanych ze wskaźnikiem zagrożenia ubóstwem.

Literatura

- [1] T. Bedi, A. Coudouel, Simler K. *More Than A Pretty Picture*. World Bank, 2007.
- [2] N. Hansen, V. Quintaes, D. Silva, P. Silva. *A Fay-Herriot model for estimating the proportion of households in poverty in brazilian municipalities*. 58th World Statistical Congress, 2011.
- [3] J.N.K. Rao. *Small Area Estimation*. Wiley and Sons, 2003.