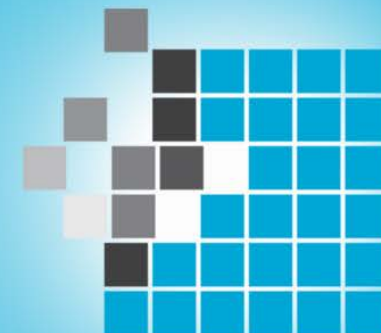




KONFERENCJA BADANIA W EDUKACJI

Warszawa, 19-20 października 2015 r.

Czy polska szkoła uczy coraz skuteczniej?



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE  *entuzjaści
edukacji*

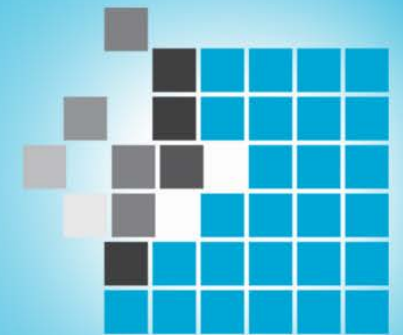
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOLECZNY





KONFERENCJA BADANIA W EDUKACJI

Warszawa, 19-20 października 2015 r.



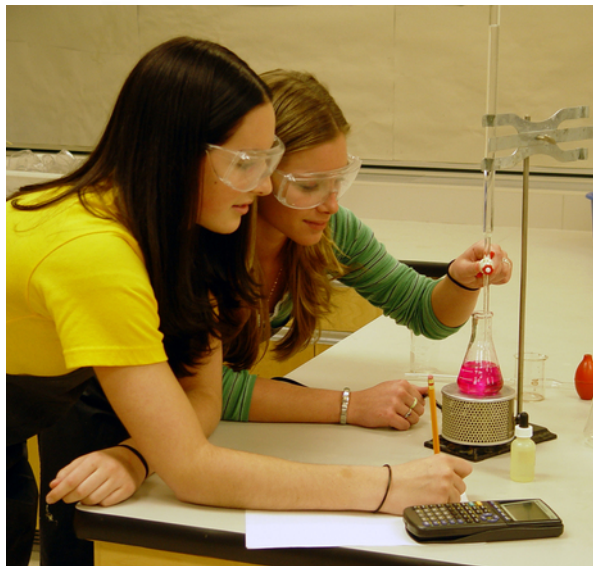
Artur Pokropek Wstęp. Czyli problem porównywalności



Ciekawość



(fot. Artur Pokropek ©)



(fot. mars.nasa.gov)



Chcemy wiedzieć





Chcemy wiedzieć jak działa system edukacyjny

- Popularność międzynarodowych badań edukacyjnych PISA, PIRLS, TIMSS
- W Polsce uzyskiwane co roku przez uczniów wyniki egzaminów są przedmiotem różnorodnych debat i analiz. Często rozmawianiem na temat wyników egzaminacyjnych towarzyszą próby zinterpretowania ich w terminach zmiany w poziomie umiejętności uczniów, poprzez porównywanie średnich wyników między latami, bądź, w przypadku egzaminu maturalnego, odsetka uczniów zdających egzamin.

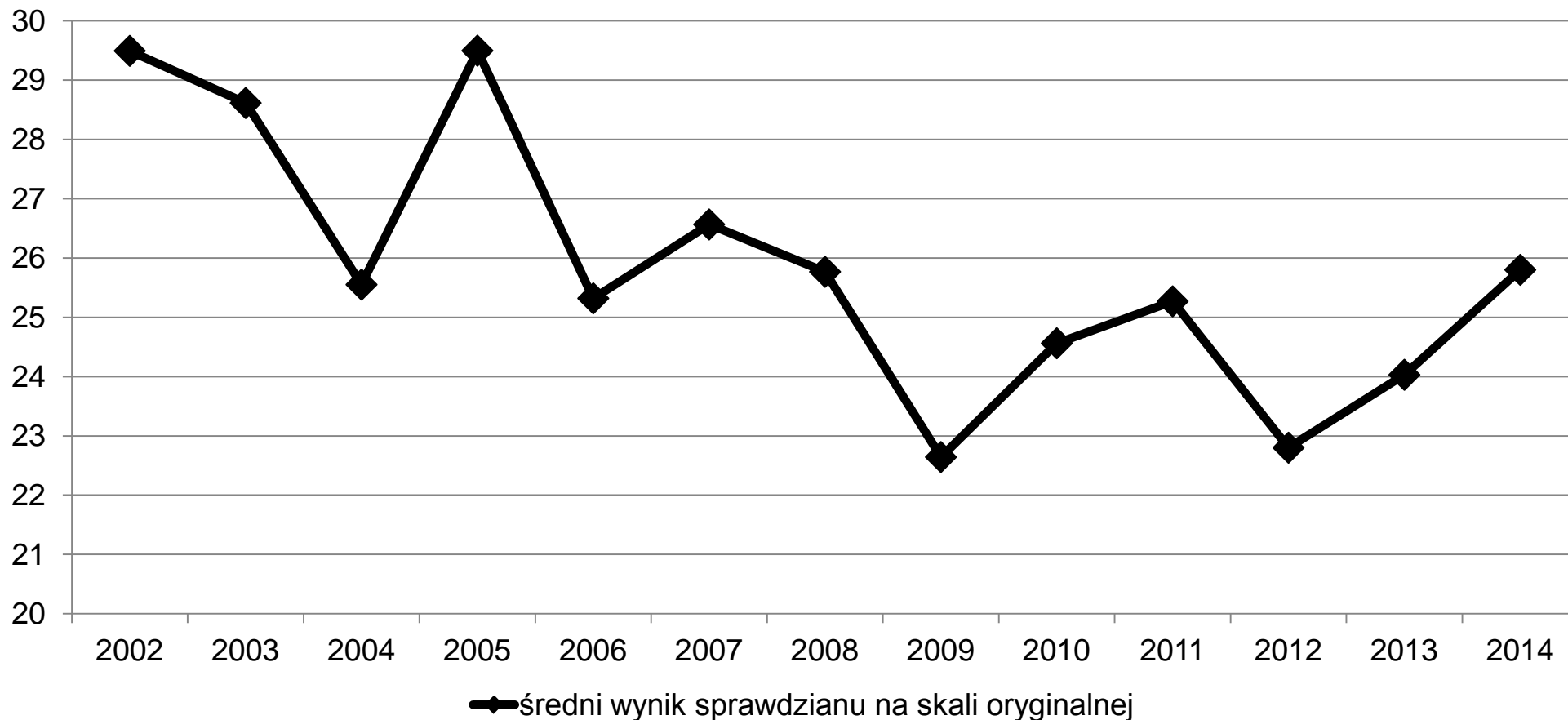


(fot. Artur Pokropek ©)



Wyniki obserwowane – sprawdzian 2002-2014

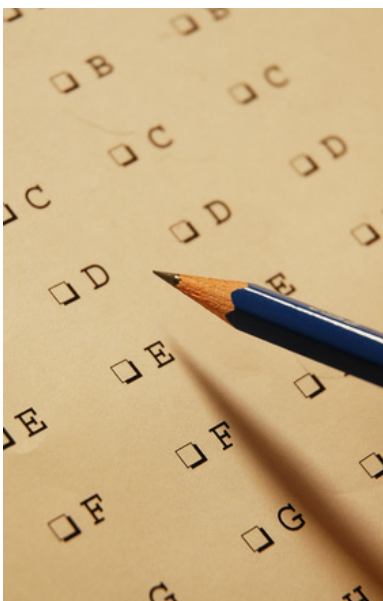
średni wynik sprawdzianu na skali oryginalnej (0-40 pkt.)



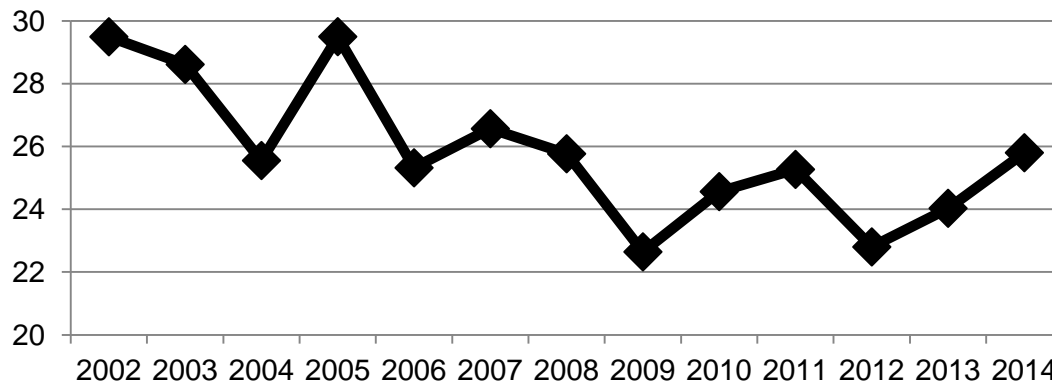


Wyniki obserwowane – sprawdzian 2002-2014

Trudność testu



średni wynik sprawdzianu na skali oryginalnej (0-40 pkt.)



—◆— średni wynik sprawdzianu na skali oryginalnej



Umiejętności uczniów



Potrzebne jest zrównywanie (szerzej linkowanie)

- Łączenie testów, które ma na celu umożliwienie zamiennego, równoważnego posługiwania się wynikami z tych testów.
- Wymaga ono jednoczesnego kontrolowania różnic w poziomie umiejętności osób rozwiązujących różne testy oraz różnic w poziomie trudności tych narzędzi.
- Zadanie to wiąże się z koniecznością zebrania do analizy danych w planie, który dostarcza odpowiedniej informacji łączącej (*link*).



Przykłady zrównywania wyników egzaminacyjnych w innych krajach

Metody zrównywania najwcześniej rozwinęły się w USA

SAT (Scholastic Assessment Test) - Od 1941 roku każda nowa wersja testu zawiera około 20% pytań z poprzedniej edycji. Wyniki kolejnych edycji zrównywane są z poprzedniego roku do roku kolejnego. Średnią skali ustalono na 500 dla roku 1941 (w 1995 skala została ponownie wycentrowana tak, by rokiem bazowym był rok 1995 o średniej wyników 500).

Do zrównywania używa się klasycznych metod zrównywania liniowego i ekwicyntylowego.



Zrównywanie za pomocą planu nierównoważonych grup z testem kotwiczącym

Sesja/ populacja	Zadania testowe		
1	Test A		
2		Test B	
3			test C



Zrównywanie za pomocą planu nierównoważonych grup z testem kotwiczącym

Sesja/ populacja	Zadania testowe		
1	Test A	sekcja zrów.	
2		Test B	sekcja zrów.
3		test C	sekcja zrów.



Przykłady zrównywania wyników egzaminacyjnych w innych krajach

ACT (American College Testing) jest drugim (po SAT) najpopularniejszym w USA testem mierzącym osiągnięcia szkolne uczniów po szkole średniej.

Aby przeprowadzić zrównanie dwóch testów z różnych lat, w przypadku ACT spośród wszystkich uczniów losowana jest reprezentatywna próba (jako próba zrównująca).

Uczniowie należący do wylosowanej próby obok aktualnej edycji testu rozwiązują też nowe zadania, wcześniej nie publikowanych arkuszy egzaminacyjnych.



Zrównywanie za pomocą planu zewnętrznej kotwicy z równoważnymi grupami

Sesja/ populacja	Zadania testowe		
1	Test A		
2		Test B	
3			Test C



Zrównywanie za pomocą planu zewnętrznej kotwicy z równoważnymi grupami

Sesja/ populacja	Zadania testowe		
1	Test A		
2		Test B	
3			Test C
zrównująca	Test A	Test B	Test C



Przykłady zrównywania wyników testowych w innych krajach

	Plan z nierównoważnymi grupami i wewnętrzną kotwicą	Plan z równoważnymi grupami i zewnętrzną kotwicą	Stosowana teoria
USA	SAT	ACT	KTT
Kanada	EQAO (Ontario)		IRT
Izrael	PET		IRT
Australia		NAPLAN	IRT
Szwecja	SweSAT	SweSAT	KTT; IRT
Węgry		Test z matematyki i interpretacji tekstu	IRT

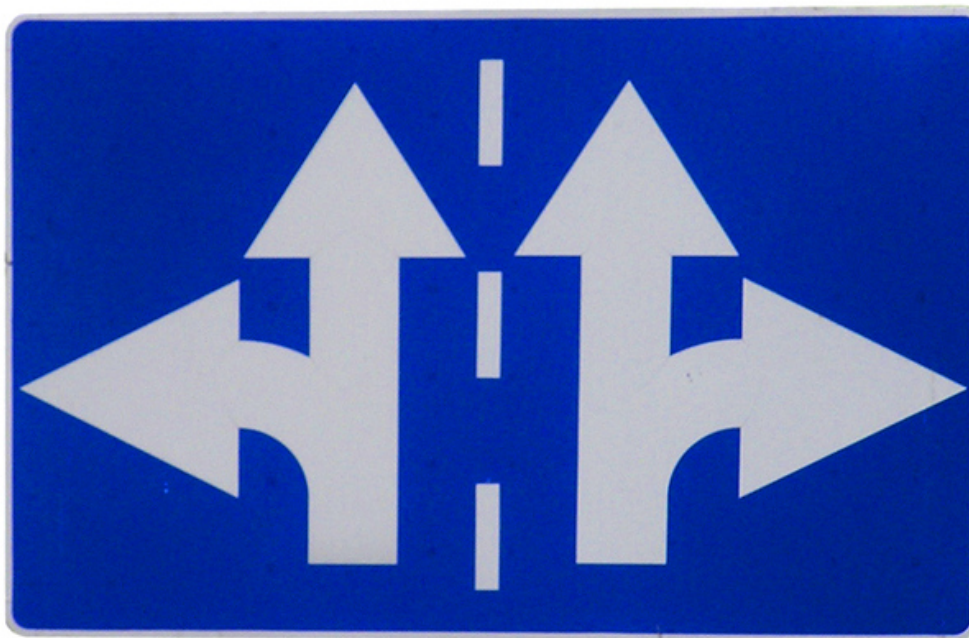
Więcej na ten temat - Pokropek, A., & Kondratek, B. (2012). Zrównywanie wyników testowania. Definicje i przykłady zastosowania. *Edukacja*, 4, 52-71.



Wiele możliwości



Zrównane
egzaminy



Badania
monitorujące



Zrównane egzaminy vs zrównane badania

Zrównane egzaminy

- ✓ Egzaminy mierzą to co dla polskiej szkoły ważne
- ✓ Egzamin doniosły. Brak problemów z motywacją
- ✓ Zrównane wyniki ważne dla uczniów
- ✓ Informacja dla całej populacji
- ✗ Skomplikowane organizacyjnie

☢ Politycznie radioaktywne

Zrównane badania monitorujące

- ✓ **Elastyczność**
- ✓ Pewne aspekty organizacyjne są łatwiejsze
- ✓ Mamy już takie badania
- ✗ Testy niskiej stawki. Problem motywacji
- ✗ Losowa próba
- ✗ Problemy realizacyjne
- ✗ Bezużyteczne dla poszczególnych uczniów



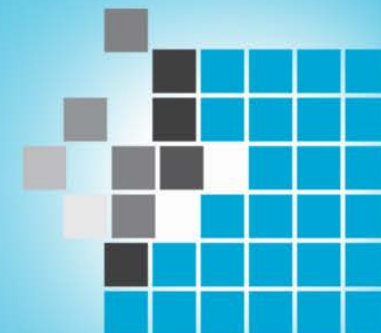
Plan sesji

1. Wstęp. Czyli problem porównywalności *Artur Pokropek*
2. Polskie badanie zrównujące *Henryk Szaleniec*
3. Czy polska szkoła uczy coraz skuteczniej. Wyniki badania zrównującego *Karolina Świst, Paulina Skórska, Filip Kulon*
 - a. Ogólnopolskie zmiany w wynikach egzaminacyjnych
 - b. Zmiany na poziomie szkół (przeglądarka PWE)
 - c. Regionalne zmiany w wynikach egzaminacyjnych
4. Rekomendacje dla systemu *Bartosz Kondratek*
5. Dyskusja



KONFERENCJA BADANIA W EDUKACJI

Warszawa, 19-20 października 2015 r.



Henryk Szaleniec Polskie badanie zrównujące



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE



EDUKACYJNA
WARTOŚĆ
DODANA

IBE



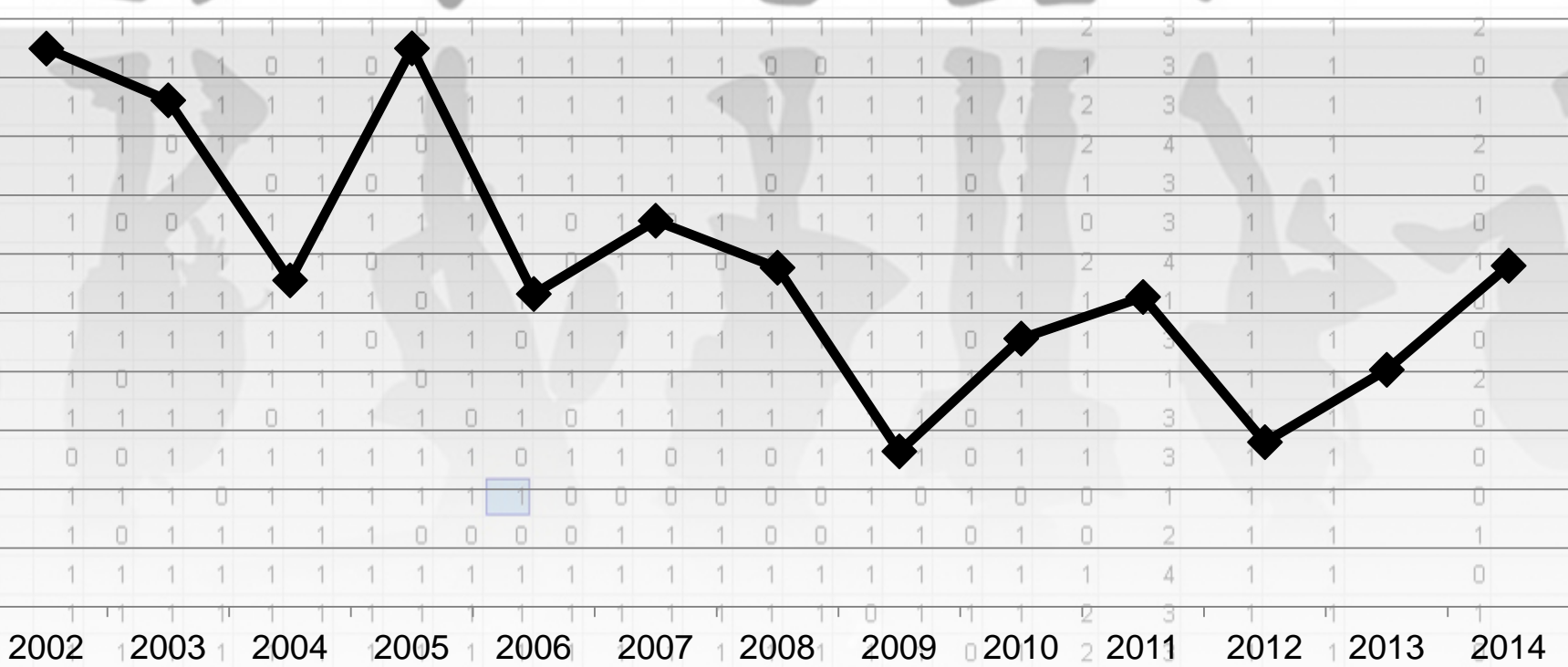
*entuzjaści
edukacji*

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



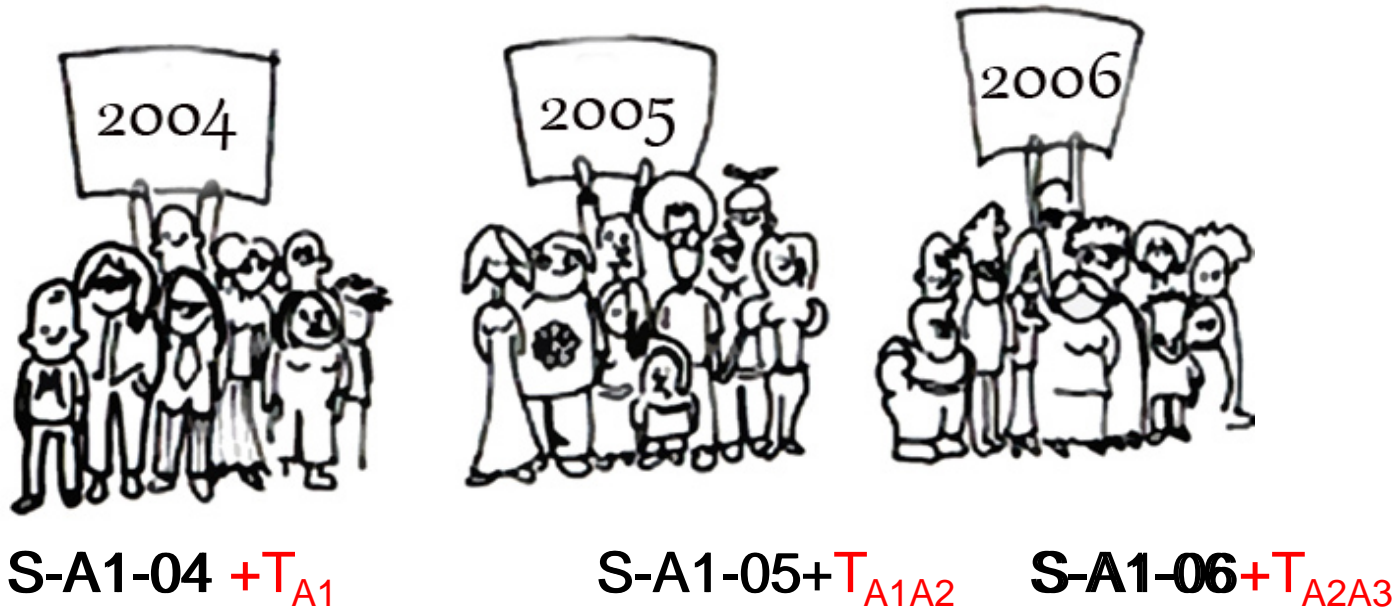
plec	dysl	ark	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_10	z_11	z_12	z_13	z_14	z_15	z_16	z_17	z_18	z_19	z_20	z_21	z_22_1	z_22_2	z_22_3	z_22_4	z_22_5	z_22_6	
K	N	GH-PA1-122	1	0	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	2	4	1	1		2	1	1
K	N	GH-PA1-122	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	0		0	1	0
M	N	GH-PA1-122	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	1	1		2	1	0
M	T	GH-PA1-122	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	0	0	0	0	1	2	1	1		0	1	0	
M	T	GH-PA1-122	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	3	1	1		0	1	0	
M	N	GH-PA1-122	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	4	1	1		0	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	4	1	1		0	1	0	
K	T	GH-PA1-122	1	1	1	0	1	1	1	0	1	0	0	1	1	0	0	1	1	1	0	1	0	4	1	1		0	1	0
M	N	GH-PA1-122	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	1	1	1	2	1	1		2	0	0	
M	T	GH-PA1-122	1	0	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	2	2	1	1		0	1	0	
M	N	GH-PA1-122	1	0	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	3	1	1		0	0	0	
M	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1		1	1	1	
M	N	GH-PA1-122	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	1	1	1		0	1	0	
M	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	4	1	1		0	1	0	
M	T	GH-PA1-122	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1		0	1	0	
M	N	GH-PA1-122	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	1	1		2	0	0	
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	1	1		2	0	0
M	T	GH-PA1-122	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	0	1	1	1	1	3	1	1		0	1	0	
M	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	1	1		1	1	0
K	N	GH-PA1-122	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	3	1	1		0	1	0	
K	N	GH-PA1-122	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	3	1	1		0	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	2	4	1	1		1	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	4	1	1		1	1	0
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		0	0	0	
M	T	GH-PA1-122	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	
M	T	GH-PA1-122	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	3	1	1		0	1	0	
M	N	GH-PA1-122	0	0	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	1	0	3	1	1		0	1	0	
M	N	GH-PA1-122	1	1	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1		0	0	0	
M	N	GH-PA1-122	1	0	1	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	1	0	2	1	1		1	0	0	
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1		0	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	0	
K	N	GH-PA1-122	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	1	1	1	0	1	0	1		0	0	0	
K	N	GH-PA1-122	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	0	
M	N	GH-PA1-122	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	2	3	1	1		1	0	1	

średni wynik sprawdzianu na skali oryginalnej (0-40 pkt.)



◆ średni wynik sprawdzianu na skali oryginalnej

Zrównywanie mające na celu umożliwienie zamiennego, równoważnego posługiwania się wynikami egzaminów, zostało przeprowadzone w IBE post factum i ponadto objęło wyniki z wielu lat.



Zadanie to wymagało kontrolowania różnic w poziomie:

A. umiejętności osób przystępujących do egzaminu w różnych latach,

B. trudności zastosowanych arkuszy egzaminacyjnych.

Wiązało się z koniecznością zebrania do analizy danych w planie, który dostarcza odpowiedniej informacji łączącej. Zastosowano plan nierównoważnych grup z testem kotwiczącym (NEAT)

**Wyniki
egzaminacyjne
niezrównane
między latami**

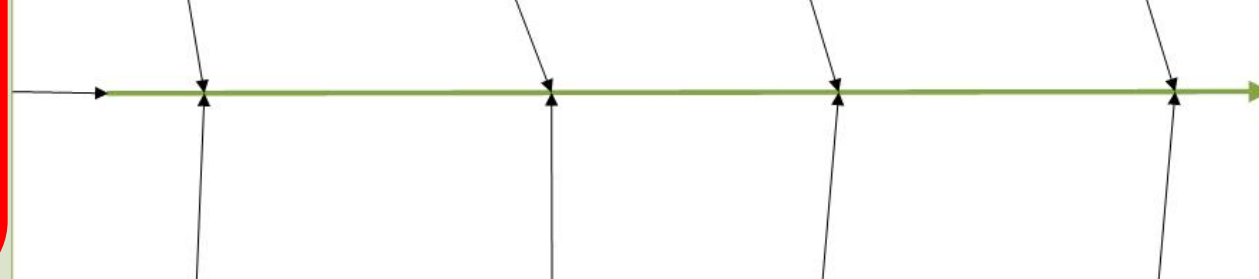
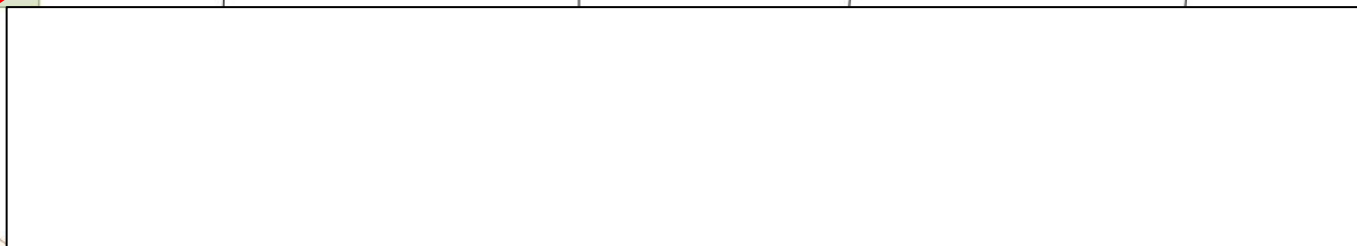
zachowanie
tożsamości
konstruktu,
pomiar tego
samego zakresu
treściowego

sprowadzenie
wymogów
stałej
rzetelności
egzaminów
pomędzy latami

sprowadzenie
wymogów
równoważności
testów
zrównanych

sprowadzenie
wymogów
niezmienniczości
względem
populacji

**Wyniki
porównywalne
między latami**





Proces zrównywania wyników w badaniach IBE

Analiza wymagań egzaminacyjnych, procedur przygotowania i przeprowadzenia egzaminu

Analiza testów egzaminacyjnych i ich wyników

Wybór metody i planu zrównania, dobór próby

Przygotowanie narzędzi

Badania i przygotowanie zbioru danych

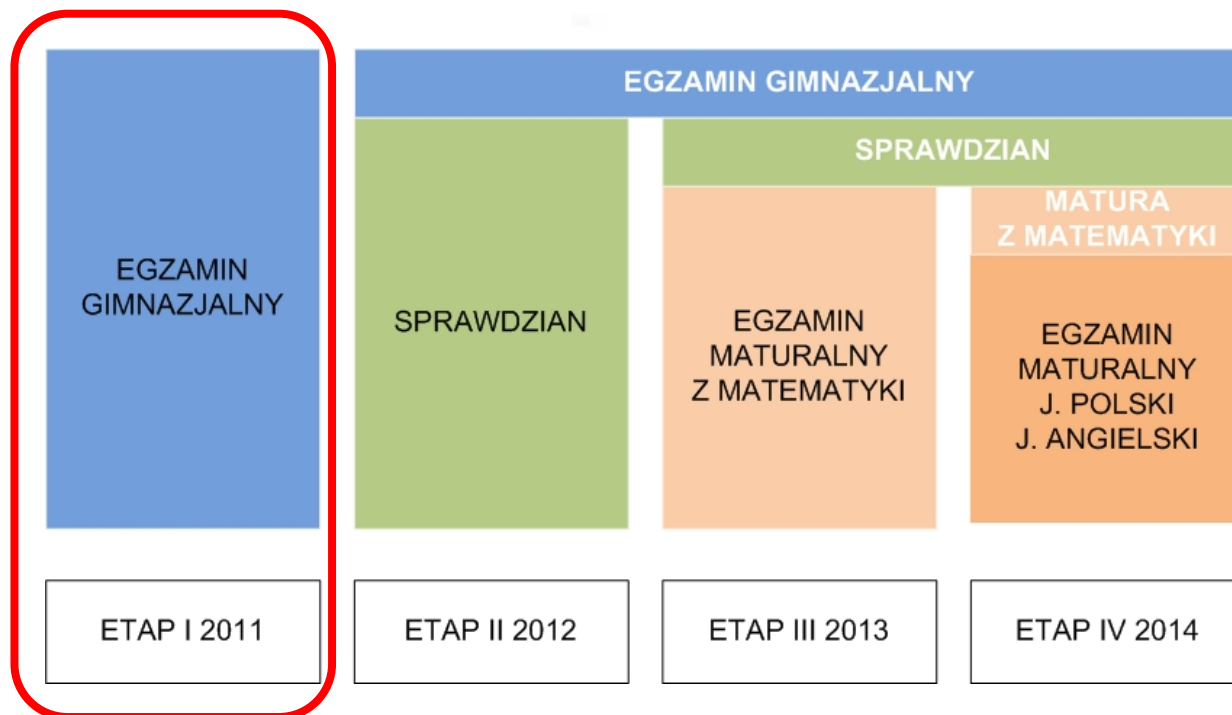
Analizy, dopasowanie modelu, zrównanie

Analizy rezultatów egzaminacyjnych na zrównanych wynikach

Dyskusja porównywalnych pomiędzy latami wyników, komunikowanie



Cztery etapy studium zrównującego wyniki egzaminów



BADANIA TOWARZYSZĄCE:

Efekt oceniania prace z matematyki i języka polskiego

Rok	Typ szkoły (egzamin)	Liczebność próby szkół	Liczebność próby uczniów	Liczebność próby nauczycieli
2011	Egzamin gimnazjalny	440	10398	3118
2012	Sprawdzian	440	9086	1774
	Egzamin gimnazjalny	80	1689	631
2013	Matura z matematyki	200	4530	200
	Sprawdzian	80	1604	315
	Egzamin gimnazjalny	80	1913	618
2014	Matura z języka polskiego	300	6504	309
	Matura z języka angielskiego	360	7312	547
	Matura z matematyki	120	2555	126
	Sprawdzian	80	1551	320
	Egzamin gimnazjalny	80	1784	630



W JAKI SPOSÓB PRZEPROWADZONO ZRÓWNANIE WYNIKÓW EGZAMINU POMIĘDZY LATAMI?



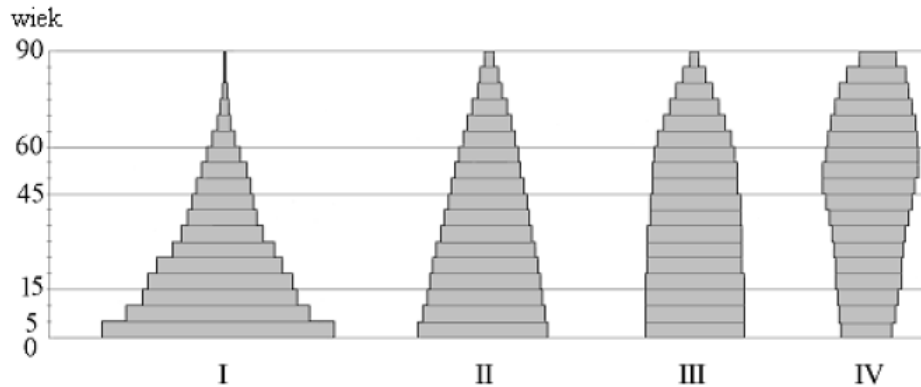
Przykład: zadanie 14 z GMP 2005 oraz z GMP 2008

Łatwość w badaniach 2011: 62,6%

Łatwość na egzaminie 2008: 66,3%

Informacje do zadań 12. – 14.

Poniższe piramidy wieku ilustrują strukturę wiekową czterech populacji. Szerokość poziomów piramid jest proporcjonalna do liczebności danych klas wiekowych.



Na podstawie: *Geografia. Encyklopedia PWN*, Warszawa 2002.

Zadanie 14. (0-1)

W 2000 roku piramida wieku dla światowej populacji była podobna do piramidy I. Jednak tempo wzrostu liczby ludności świata spada i zbliża się do zera. Jedną z prognoz demograficznych przewiduje, że w 2050 roku wszystkie roczniki w wieku poniżej 60 lat będą prawie tak samo liczne. Populacji światowej będzie wtedy odpowiadać piramida typu

A. I

B. II

C. III

D. IV

Łatwość w badaniach 2011: 64,6%

Łatwość na egzaminie 2005: 62,2%

Zadanie 14. (0-1)

Do naczynia o objętości $V = 0,75$ l wlano 0,45 l wody. Jaki procent objętości tego naczynia stanowi objętość wody?

A. 6

B. 16,(6)

C. 33,75

D. 60

Czy zadanie z piramidą wieku jest łatwiejsze od zadania z objętością?

Czy uczniowie w 2008 roku mieli wyższy poziom badanych umiejętności niż uczniowie z 2005?



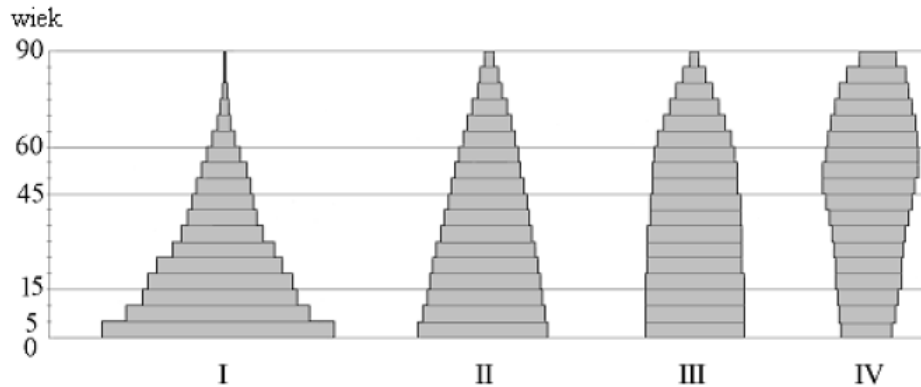
Przykład: zadanie 14 z GMP 2005 oraz z GMP 2008

Łatwość w badaniach 2011: 62,6%

Łatwość na egzaminie 2008: 66,3%

Informacje do zadań 12. – 14.

Poniższe piramidy wieku ilustrują strukturę wiekową czterech populacji. Szerokość poziomów piramid jest proporcjonalna do liczebności danych klas wiekowych.



Na podstawie: *Geografia. Encyklopedia PWN*, Warszawa 2002.

Zadanie 14. (0-1)

W 2000 roku piramida wieku dla światowej populacji była podobna do piramidy I. Jednak tempo wzrostu liczby ludności świata spada i zbliża się do zera. Jedną z prognoz demograficznych przewiduje, że w 2050 roku wszystkie roczniki w wieku poniżej 60 lat będą prawie tak samo liczne. Populacji światowej będzie wtedy odpowiadać piramida typu

A. I

B. II

C. III

D. IV

Łatwość w badaniach 2011: 64,6%

Łatwość na egzaminie 2005: 62,2%

Zadanie 14. (0-1)

Do naczynia o objętości $V = 0,75$ l wlano $0,45$ l wody. Jaki procent objętości tego naczynia stanowi objętość wody?

A. 6

B. 16,(6)

C. 33,75

D. 60

Dane zebrane w przeprowadzonych w 2011 roku badaniach sugerują, że to zadanie z objętością jest łatwiejsze od zadania z piramidą wieku.

Jest to przesłanka do wnioskowania o wyższym poziomie umiejętności uczniów zdających egzamin w 2008 roku od uczniów zdających w 2005 roku.

Plan zrównywania – egzamin gimnazjalny

Źródło	Populacja	Próba	N	Zbiory zadań egzaminacyjnych														Zbiory dodatkowych zadań kotwiczących										
				T ₀₂	T ₀₃	T ₀₄	T ₀₅	T ₀₆	T ₀₇	T ₀₈	T ₀₉	T ₁₀	T ₁₁	T ₁₂	T ₁₃	T ₁₄	C ¹	C ²	C ³	C ⁴	C ⁵	C ⁶	C ⁷	C ⁸	C ⁹	C ¹⁰	C ¹¹	
Dane egzaminacyjne	\mathcal{P}_{02}		$5 \cdot 10^5$	■																								
	\mathcal{P}_{03}		$5 \cdot 10^5$		■																							
	\mathcal{P}_{04}		$5 \cdot 10^5$			■																						
	\mathcal{P}_{05}		$5 \cdot 10^5$				■																					
	\mathcal{P}_{06}		$5 \cdot 10^5$					■																				
	\mathcal{P}_{07}		$5 \cdot 10^5$						■																			
	\mathcal{P}_{08}		$5 \cdot 10^5$							■																		
	\mathcal{P}_{09}		$5 \cdot 10^5$																									
	\mathcal{P}_{10}		$4 \cdot 10^5$																									
	\mathcal{P}_{11}		$4 \cdot 10^5$																									
	\mathcal{P}_{12}		$4 \cdot 10^5$																									
	\mathcal{P}_{13}		$4 \cdot 10^5$																									
	\mathcal{P}_{14}		$4 \cdot 10^5$																									
	Dane z sesji zrównujących	\mathcal{P}_{11}	S_{11}^1	800	▣																							
\mathcal{P}_{11}		S_{11}^2	800																									
\mathcal{P}_{11}		S_{11}^3	800																									
\mathcal{P}_{11}		S_{11}^4	800																									
\mathcal{P}_{11}		S_{11}^5	800																									
\mathcal{P}_{11}		S_{11}^6	800																									
\mathcal{P}_{11}		S_{11}^7	800																									
\mathcal{P}_{11}		S_{11}^8	800																									
\mathcal{P}_{11}		S_{11}^9	800																									
\mathcal{P}_{11}		S_{11}^{10}	800																									
\mathcal{P}_{11}		S_{11}^{11}	800																									
\mathcal{P}_{12}		S_{12}	1600																									
\mathcal{P}_{13}		S_{13}	1600																									
\mathcal{P}_{14}		S_{14}	1600																									

Legenda:

uczniowie nie rozwiązują żadnego zadania

uczniowie rozwiązują wszystkie zadania

uczniowie rozwiązują wszystkie zadania na sesji innego typu

Źródło danych	Populacja	Próba	Liczba obser.	Test lub blok zadań		
				X	Y	Z
Egzamin	P_{01}			■	■	
	P_{02}				■	■
	P_{03}					■
						■
Sesja badawcza		1	800	■	■	
		2	800	■	■	■



Zrównywania wyników dokonano z wykorzystaniem IRT w 3 etapach

1. Umieszczenie na wspólnej skali: rozkładów poziomu umiejętności θ dla wszystkich badanych populacji oraz parametrów zadań dla wszystkich testów, których wyniki były zrównywane.
2. Przełożenie wyrażonych w skali ukrytej zmiennej poziomu umiejętności θ z powrotem na skalę wyników obserwowanych, aby oszacować rozkład wyników zrównywanych testów w populacji docelowej.
3. Zbudowanie różnowartościowej funkcji zrównującej, pozwalającej na przekształcenie wyników z jednego testu na wyniki innego.



Przeprowadzone analizy obejmowały w ogólnym przypadku dwa etapy

1

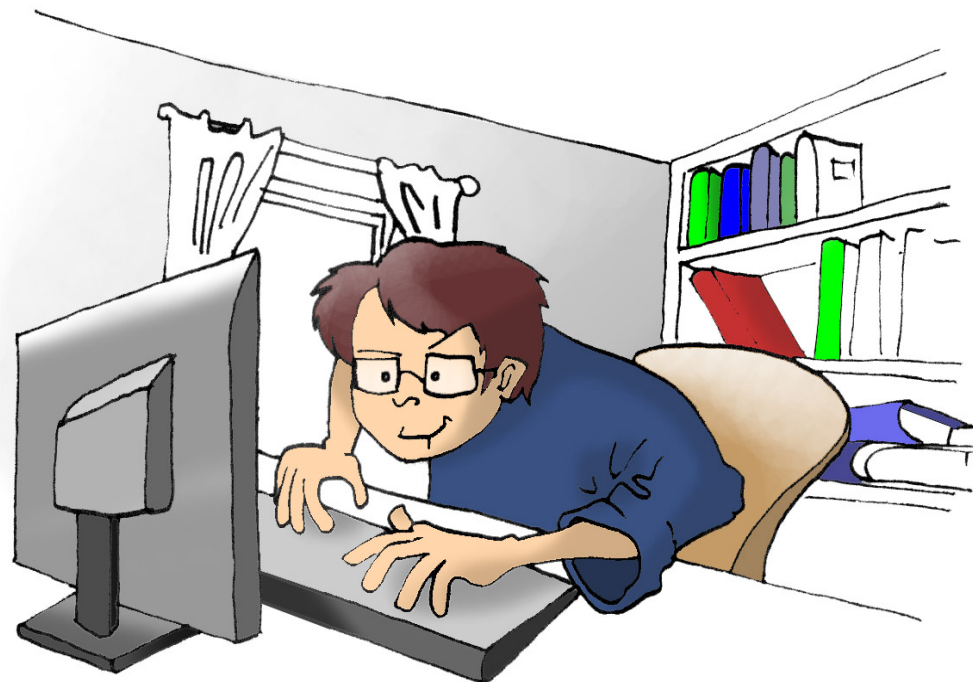
- Dopasowanie do danych jednowymiarowego wielogrupowego modelu IRT.

2

- Oszacowanie rozkładu wyników surowych w egzaminie z 2012 roku (rok referencyjny) w populacjach uczniów zdających dany egzamin w poszczególnych latach.

Pierwszy etap dostarczył informacji o poziomie umiejętności uczniów piszących egzaminy w poszczególnych latach wyrażonym na wspólnej skali zmiennej ukrytej poziomu umiejętności θ .

Drugi etap pozwolił na porównanie trudności egzaminów między latami.



W badaniach zrównujących IBE model IRT został dopasowany do danych za pomocą algorytmu EM (*expectation-maximization algorithm*) do znajdowania oszacowań największej wiarygodności parametrów w modelach ze zmiennymi ukrytymi (latentnymi).

Algorytm EM został zaimplementowany w autorskim programie Bartosza Kondratka irt.ado, funkcjonującym w środowisku Stata.



W przeprowadzonych badaniach zaadaptowano symulacyjną strategię generowania wyrażonych na wspólnej skali wyników obserwowanych roku bazowego.

Za rok bazowy przyjęto 2012.

Do wygenerowania wyników obserwowanych dla danego rocznika na skali z roku bazowego (2012) wykorzystano 15 kompletów wartości potencjalnych/możliwych (*plausible values*, PV) wygenerowanych zgodnie z oszacowaną dla tego rocznika średnią i odchyleniem standardowym rozkładu umiejętności θ oraz przy uwzględnieniu parametrów zadań dla roku bazowego, a także dodatkowych zmiennych warunkujących.

ZRÓWNANIE WYNIKÓW
EGZAMINACYJNYCH,
KONCEPCJA CAŁEGO
STUDIUM

I ETAP
EGZAMIN GIMNAZJALNY

KONCEPCJA I ETAPU,
WYBÓR WYKONAWCY
2011

BADANIA 2011

ANALIZY 2011

PUBLIKACJE 2011

RAPORT 2011

KONSULTACJE
KRAJOWE
I ZAGRANICZNE

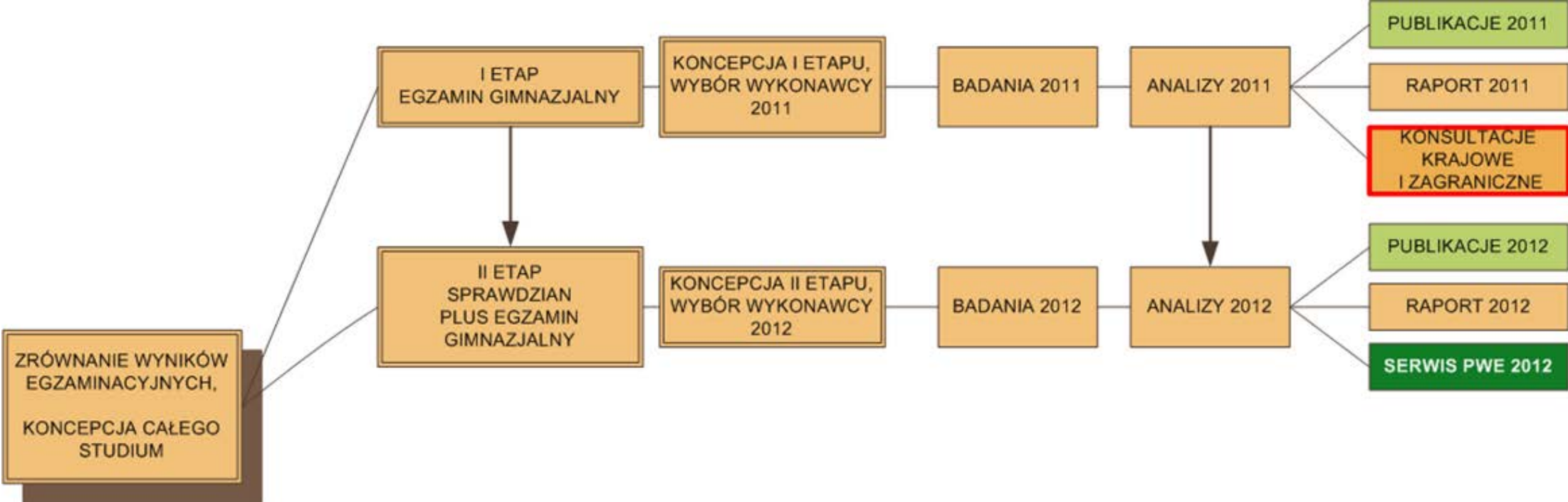


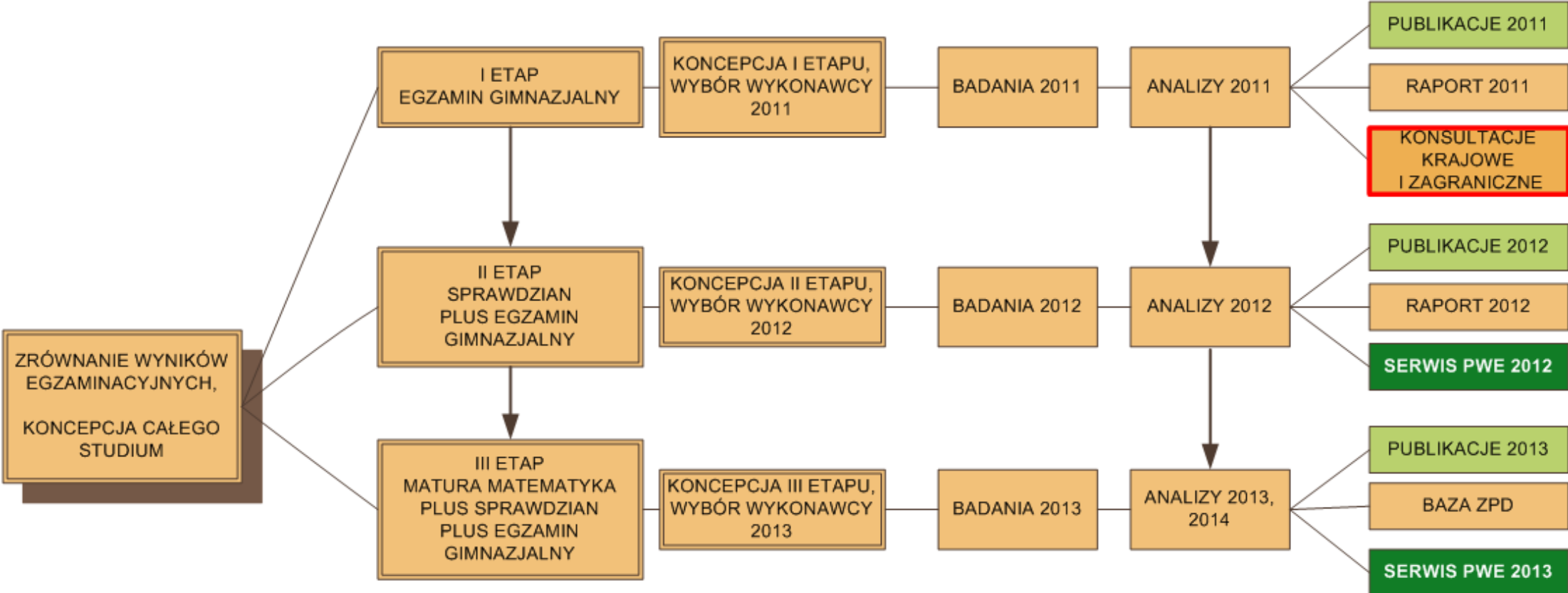
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

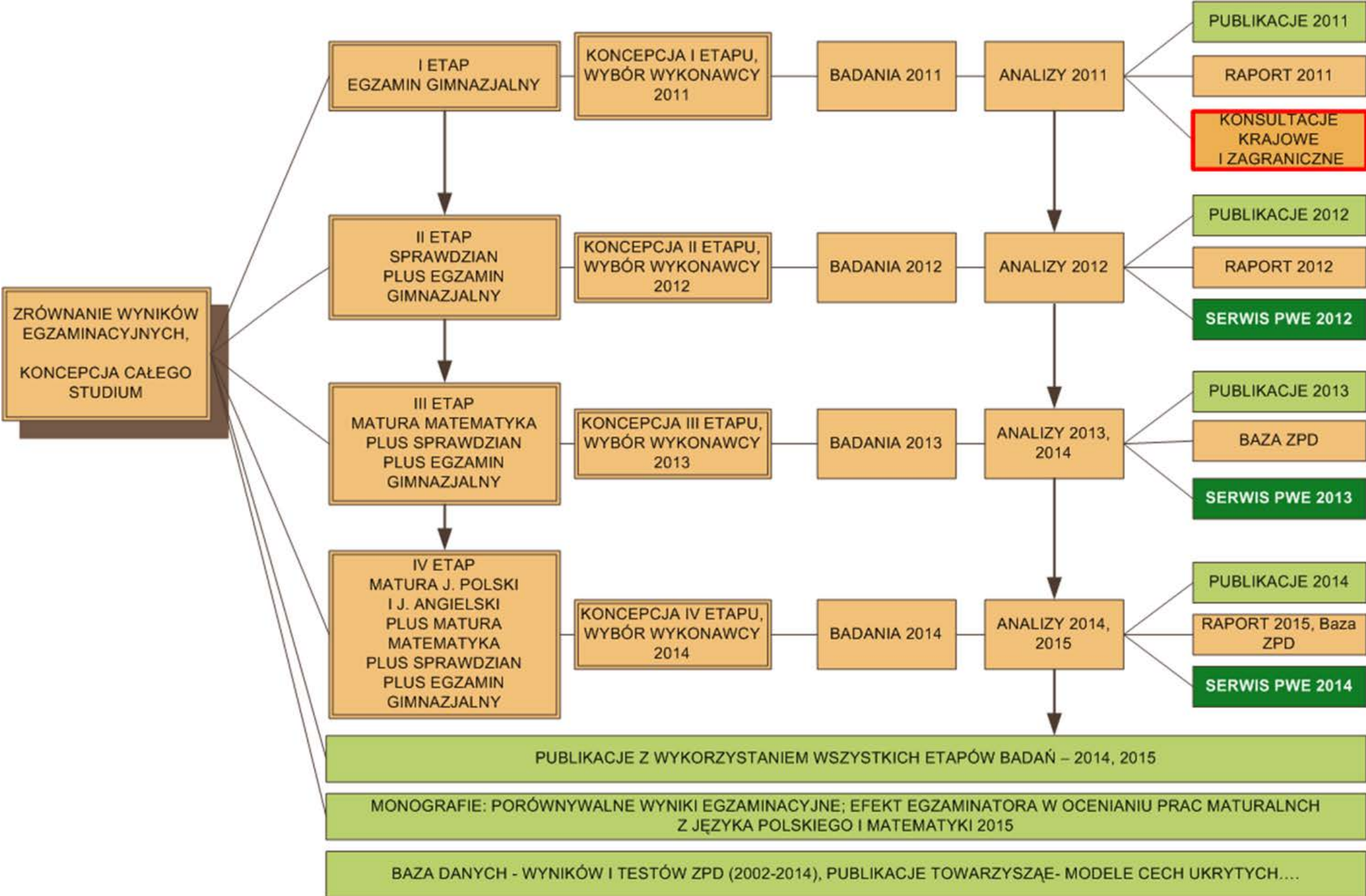
IBE  *entuzjaści
edukacji*

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY





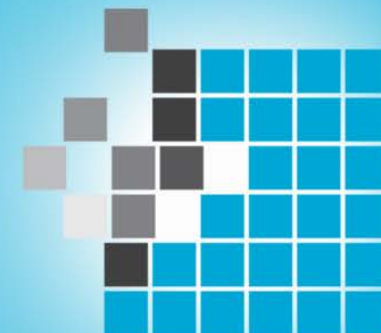




KONFERENCJA BADANIA W EDUKACJI

Warszawa, 19-20 października 2015 r.

Czy polska szkoła uczy coraz skuteczniej? Wyniki badania zrównującego



**Filip Kulon,
Paulina Skórska,
Karolina Świst
Zespół Analiz Osiągnięć Uczniów, IBE**





entuzjaści edukacji

Wyniki badania zrównującego

Karolina Świst

ZESPÓŁ ANALIZ OSIĄGNIĘĆ UCZNIÓW

Instytut Badań Edukacyjnych

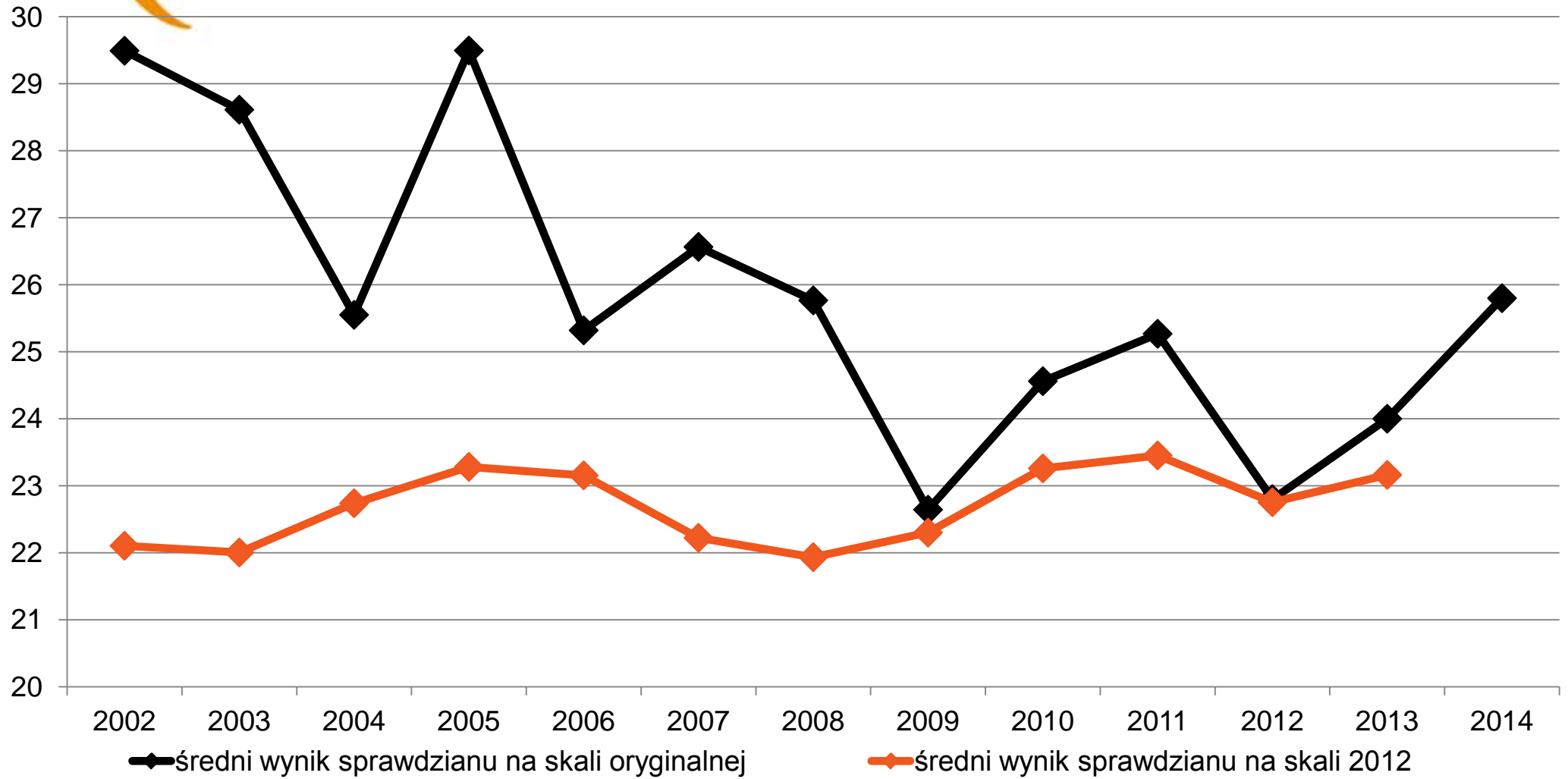


Część pierwsza – najważniejsze zagadnienia

- 1) Co wnoszą informacje z badania zrównującego względem obserwowalnych, surowych danych egzaminacyjnych
- 2) Co można powiedzieć na temat ogólnopolskich zmian w wynikach egzaminacyjnych? (przykład sprawdzianu szóstoklasisty)
- 3) Zastosowanie zrównanych wyników egzaminacyjnych do analizy międzygrupowej – przykład i wnioski.



Wyniki obserwowane vs wyniki zrównane – sprawdzian 2002-2013





Wnioski (sprawdzian):

Zmiany w rozkładzie wyników egzaminów są przede wszystkim **konsekwencją zmian w trudności poszczególnych arkuszy** – w mniejszym stopniu umiejętności uczniów.

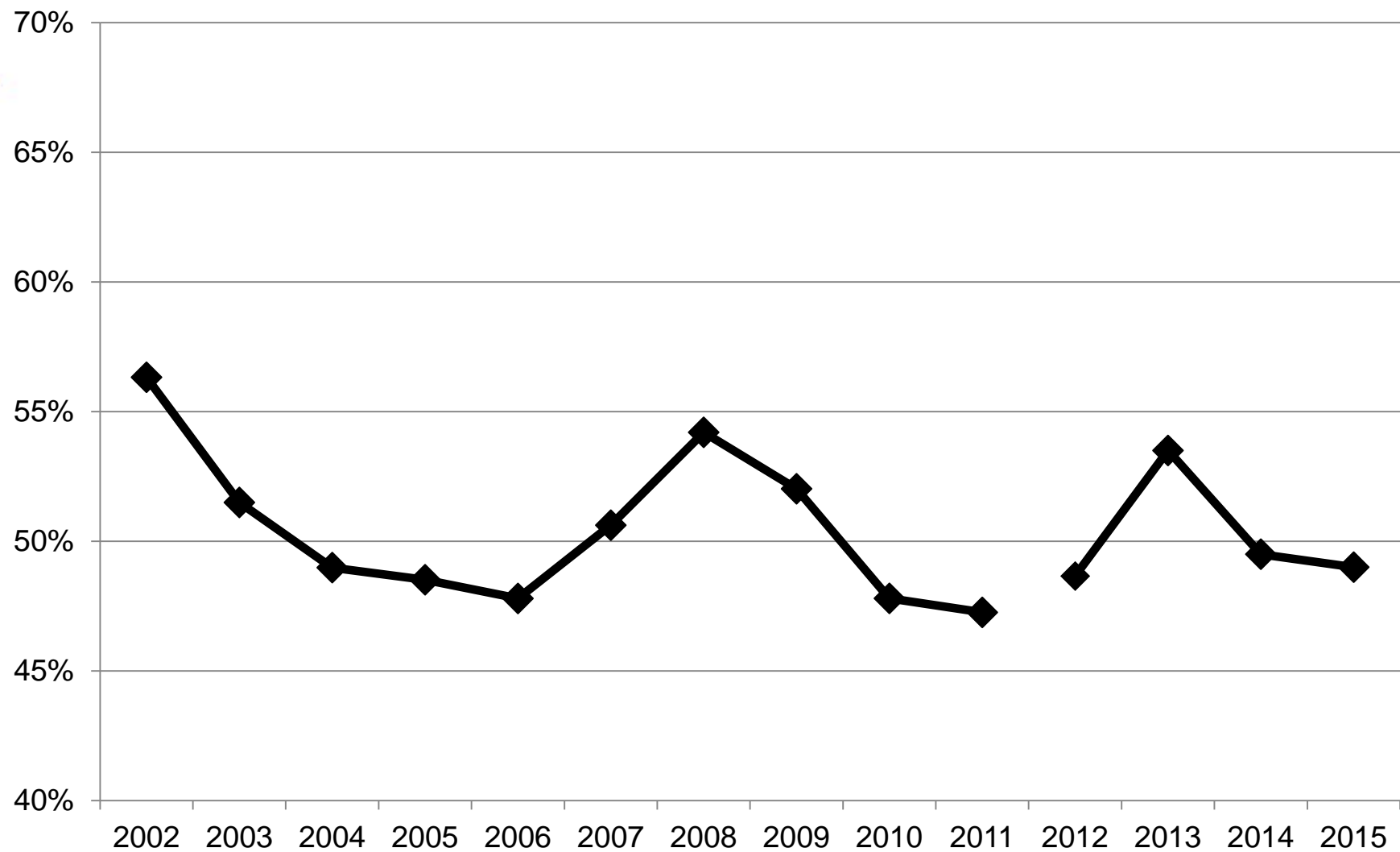
Różnice między wszystkimi analizowanymi edycjami sprawdzianu **nie wykraczają co do absolutnej wartości poza 1,6 punktu.**

Najwyższe wyniki przypadają w okolicach roku 2005 oraz 2011, a **najniższe** w okolicach 2003 oraz 2008.

Zmiany nie są jednak na tyle dramatyczne jak sugerowałyby wyniki surowe.

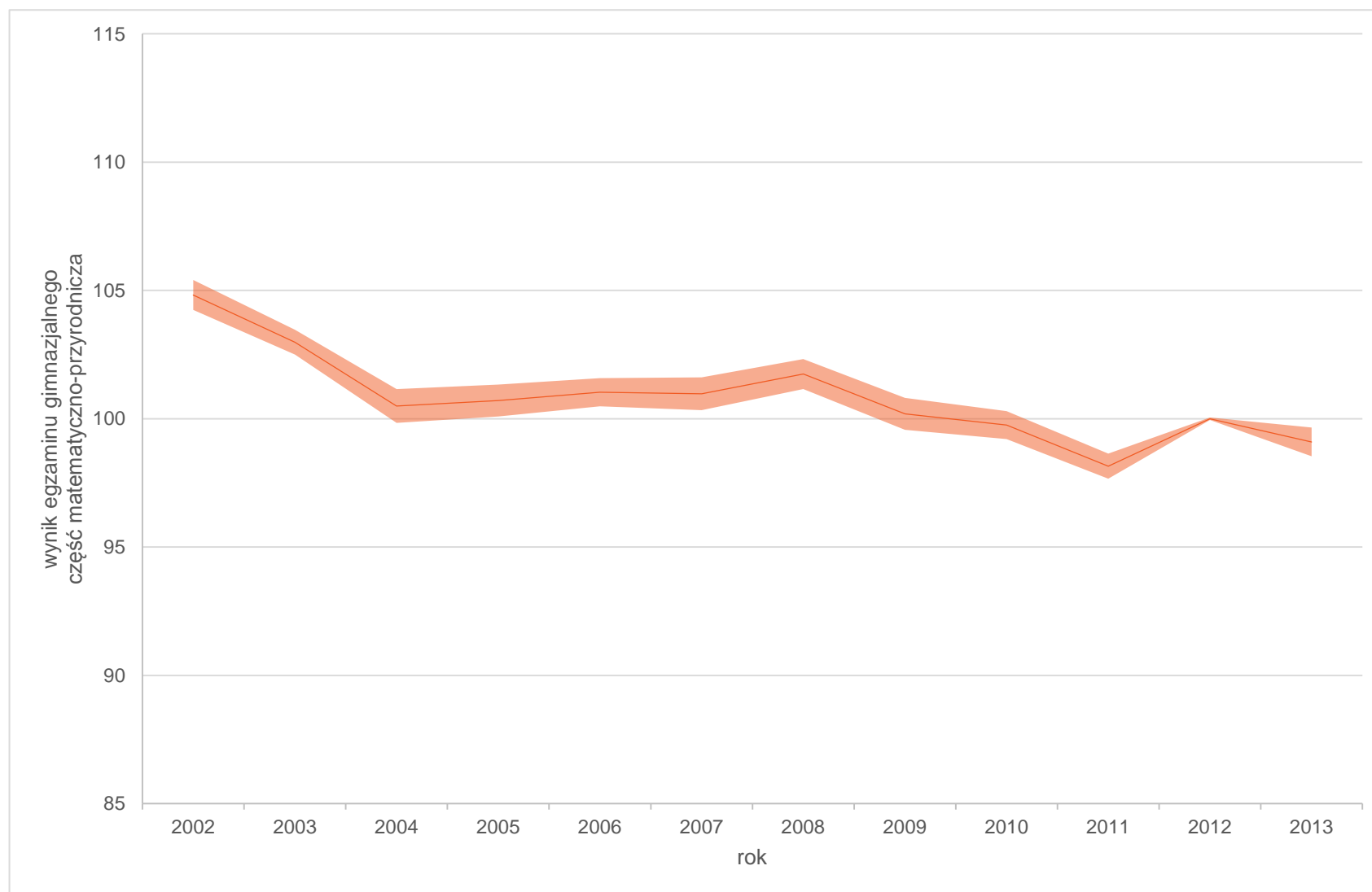


Egzamin gimnazjalny: część matematyczno-przyrodnicza – wyniki obserwowalne





Egzamin gimnazjalny: część matematyczno-przyrodnicza – wyniki zrównane



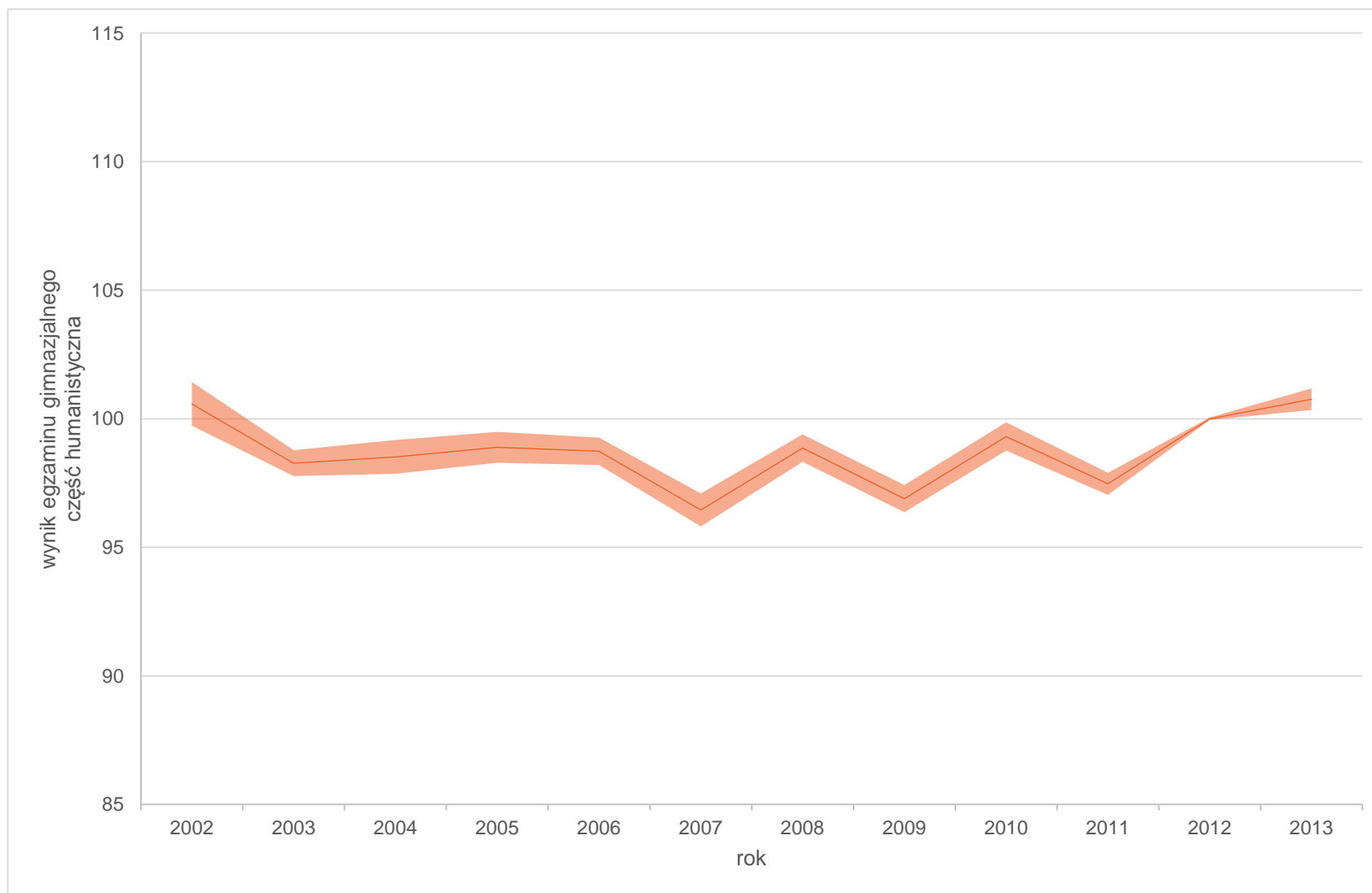


Egzamin gimnazjalny: część humanistyczna – wyniki obserwowalne





Egzamin gimnazjalny: część humanistyczna – wyniki zrównane





Wnioski (egzamin gimnazjalny)

Część matematyczno-przyrodnicza:

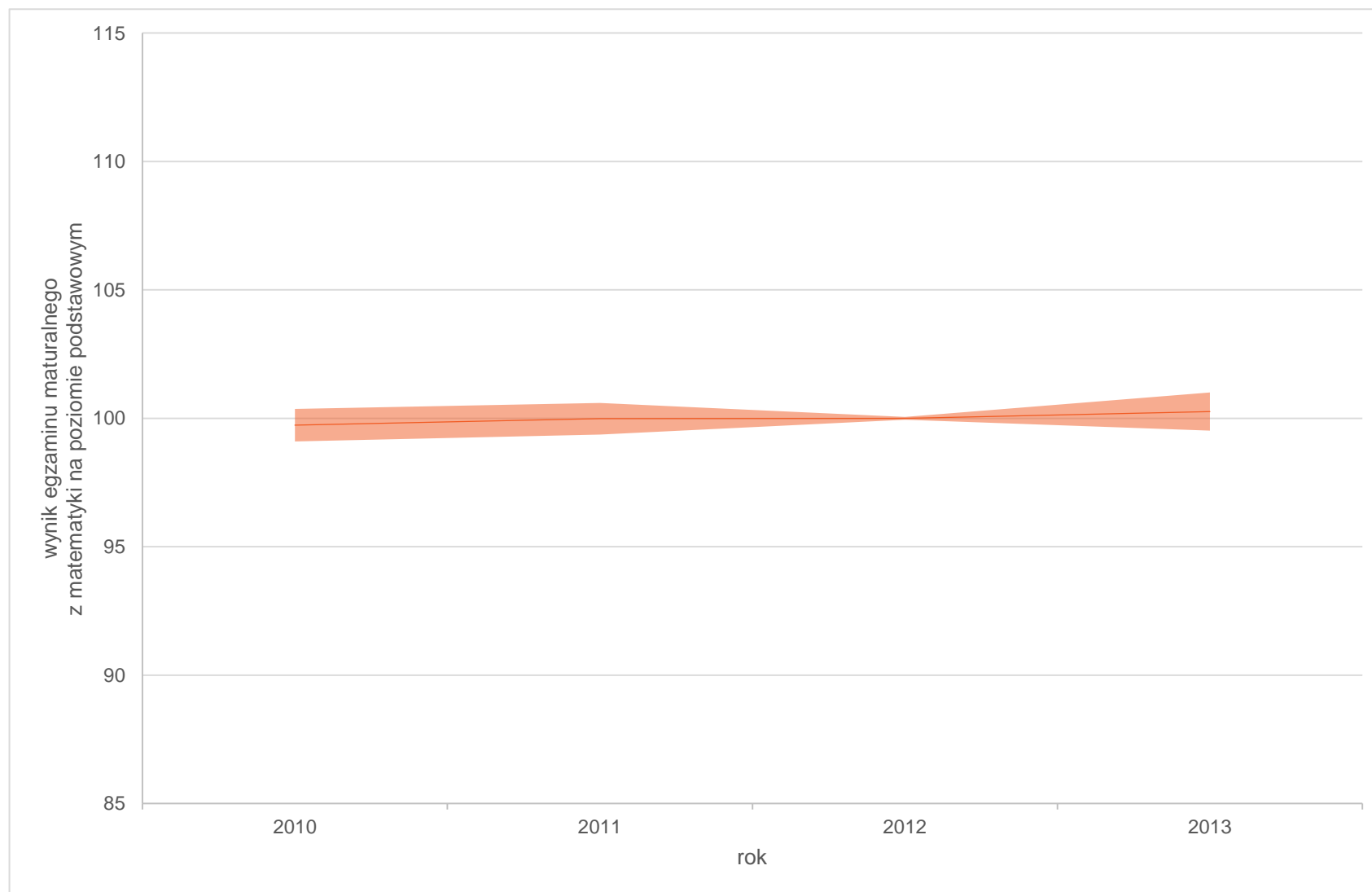
- Trend zmian w poziomie umiejętności: lata 2002-2003 (ale! pierwsze lata funkcjonowania egzaminu) oraz 2008-2011
- Znaczne fluktuacje w surowych średnich wynikach z egzaminu między latami są przede wszystkim **odzwierciedleniem różnic w ich łatwości**

Część humanistyczna:

- Fluktuacje również spowodowane różnicami w łatwości
- Trend zmian: lata 2003-2006, 2007-2012 oraz 2011-2013



Egzamin maturalny – matematyka, poziom podstawowy (wyniki zrównane)





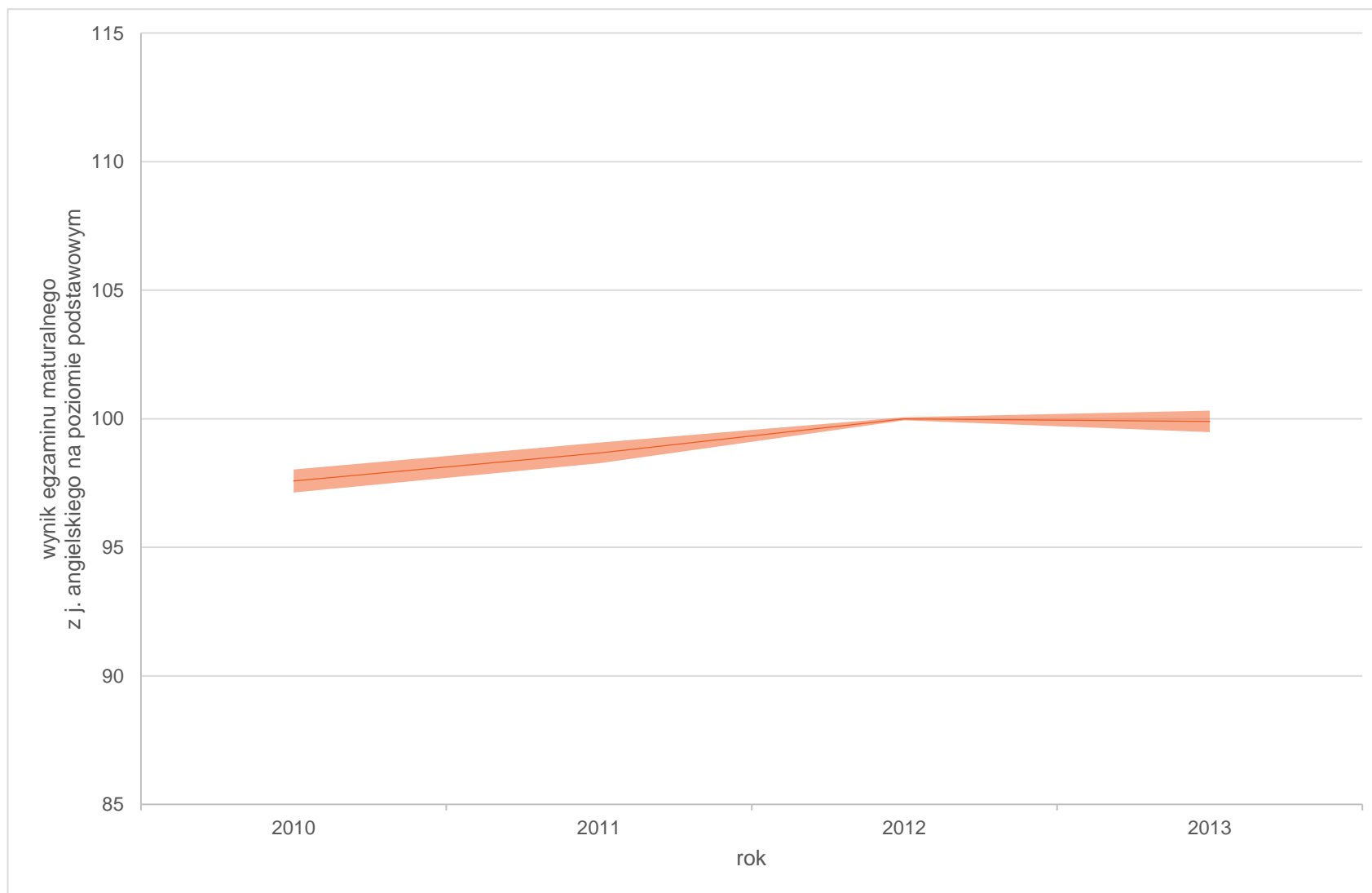
Wnioski (egzamin maturalny z matematyki – poziom podstawowy)

- Fluktuacje średniego wyniku -> głównie konsekwencje zmian w trudności
- **Zdawalność wyliczona przy uwzględnieniu wyników zrównanych jest o wiele bardziej stabilna między latami, niż ta faktycznie zaobserwowana w danych rocznikach**

Rok egzaminu	Procent uczniów uzyskujących 15 punktów lub więcej	
	niezrównane egzaminacyjne	wyniki na skali egzaminu z 2012 roku
2010	90,3%	84,4%
2011	82,1%	87,2%
2012	88,8%	88,8%
2013	88,2%	87,2%



Egzamin maturalny – j. angielski, poziom podstawowy (wyniki zrównane)



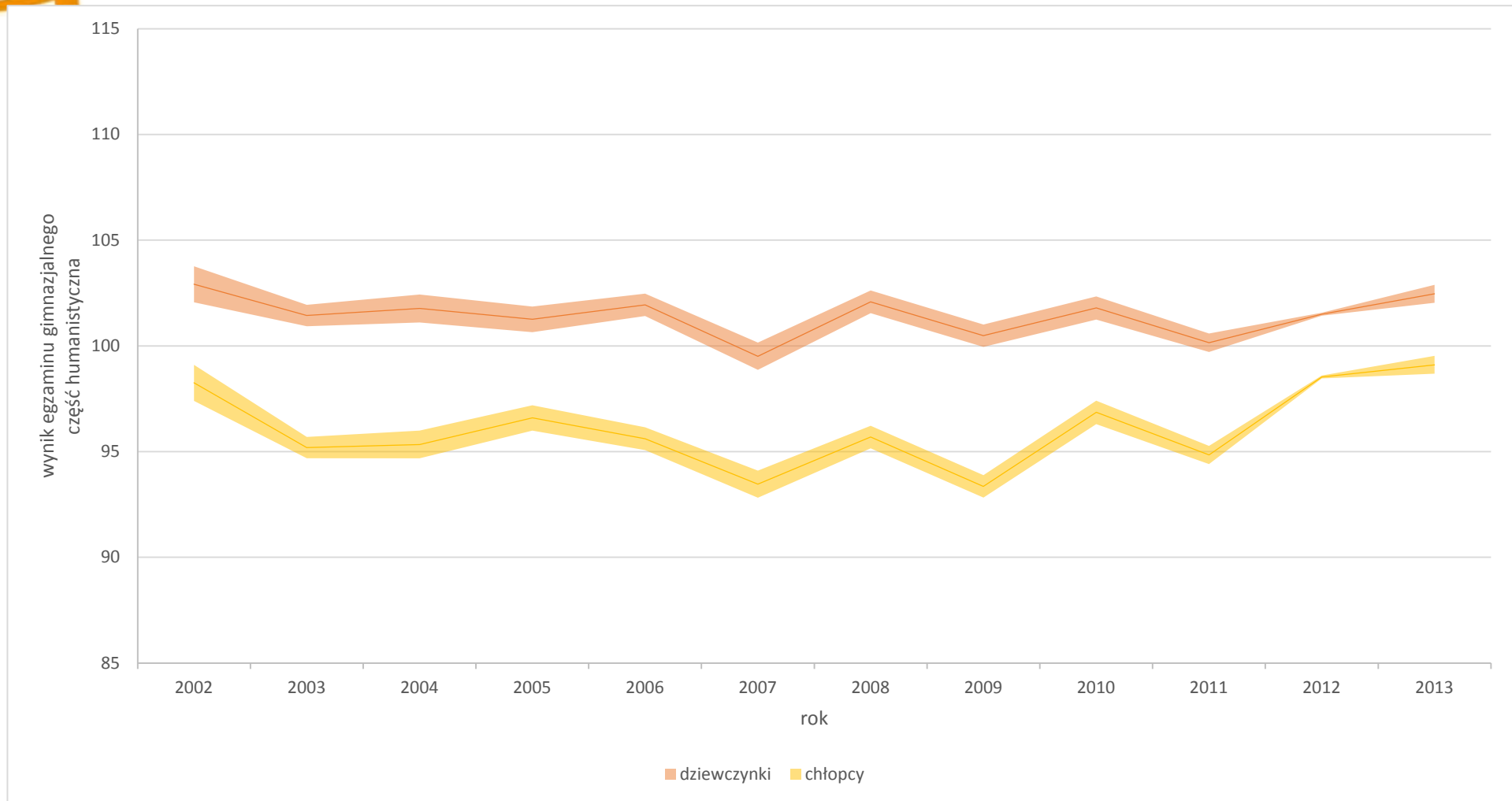


Wnioski (egzamin maturalny z j. angielskiego – poziom podstawowy)

- Fluktuacje średniego wyniku -> głównie konsekwencje zmian w trudności
- W porównaniu z rokiem 2012, próg zdawalności był zauważalnie zawyżony w edycji egzaminu z 2010 roku, a w latach 2011 oraz 2013 był nieznacznie zaniżony.
- **Wzrost umiejętności w latach 2010-2012**

Rok egzaminu	Procent uczniów uzyskujących 15 punktów lub więcej		
	niezrównane egzaminacyjne	wyniki	wyniki na skali egzaminu z 2012 roku
2010	91,1%		93,2%
2011	95,5%		94,5%
2012	94,6%		94,6%
2013	94,4%		94,1%

Analiza międzygrupowa – przykład (płeć a wyniki w części humanistycznej)





Wnioski:

- **Sprawdzian, obie części egzaminu gimnazjalnego, matura z matematyki – niewielkie fluktuacje** wyników, ale w dłuższej perspektywie czasowej wyniki tych egzaminów utrzymują się na **podobnym poziomie**
- Dla matury z języka angielskiego obserwujemy **niewielki wzrost wyników w latach 2010–2013**
- Porównania wyników między płciami wskazały na **przewagę dziewcząt w sprawdzianie i części humanistycznej egzaminu gimnazjalnego**, natomiast pozostałe egzaminy nie różnicowały obu płci.
- **Diagnoza dysleksji (lub jej brak) jest zmienną, która nie różnicuje w większości przypadków wyników egzaminacyjnych (poza maturą z j. angielskiego)**



Wnioski:

- Wyniki dla sprawdzianu i egzaminu gimnazjalnego **wyższe dla szkół niepublicznych niż publicznych**
- Dla sprawdzianu i egzaminu gimnazjalnego można zaobserwować różnice na **korzyść gmin miejskich**, natomiast nie występują różnice pomiędzy gminami wiejskimi i miejsko-wiejskimi
- Najwyższe wyniki matury osiągają uczniowie **liceów ogólnokształcących**. Niższe wyniki uzyskują uczniowie techników (którzy są drugą pod względem liczebności grupą), a najniższe uczniowie liceów profilowanych.





entuzjaści edukacji

Porównywalne Wyniki Egzaminacyjne – zrób to sam!

Paulina Skórska

ZESPÓŁ ANALIZ OSIĄGNIĘĆ UCZNIÓW

Instytut Badań Edukacyjnych



Gdzie ich szukać?

<http://pwe.ibe.edu.pl/>

Czym są Porównywalne Wyniki Egzaminacyjne (PWE)?

Zespół Analiz Osiągnięć Uczniów (ZAOU) działający w Instytucie Badań Edukacyjnych podjął w ramach projektu „Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego” prace nad zrównywaniem wyników egzaminacyjnych, w celu osiągnięcia porównywalności wyników między latami.

Wynik egzaminu zależy od trudności arkusza testowego oraz umiejętności uczniów. Gdyby umiejętności uczniów w krajowej populacji nie zmieniały się, to średni krajowy wynik egzaminacyjny byłby informacją o trudności testu. Gdyby trudność arkuszy testowych w kolejnych latach była podobna, to na podstawie wyników egzaminacyjnych można byłoby wnioskować o tym, jak zmienia się w czasie poziom umiejętności uczniów. Skonstruowanie arkuszy testowych o podobnych parametrach jest jednak zadaniem bardzo trudnym i nie zawsze zakończonym powodzeniem. Porównywalność wyników między latami można osiągnąć kontrolując trudność arkusza (procedury zrównywania wyników) lub zakładając stały poziom umiejętności w populacji (normalizacja i standaryzacja rozkładów, np. do skali staninowej). W pierwszym przypadku uzyskujemy bezwzględną porównywalność wyników i możliwość wnioskowania o zmianie umiejętności w populacji. W drugim przypadku uzyskujemy względną porównywalność w stosunku do średniej w danym roku.

W latach 2011-2014 Zespół Analiz Osiągnięć Uczniów przeprowadził badania zrównujące w różnych typach szkół, co po zastosowaniu złożonej procedury zrównywania wyników pozwoliło uzyskać porównywalność wyników sprawdzianu i egzaminu gimnazjalnego od 2002 roku, a matury z j. polskiego, angielskiego oraz matematyki od 2010 roku. Za rok bazowy dla wszystkich egzaminów przyjęto rok 2012 (krajowa średnia dla 2012 roku wynosi 100), a wyniki danego egzaminu w pozostałych latach należy rozumieć jako wzrost/spadek poziomu umiejętności uczniów w porównaniu do 2012 roku.




Zrównane wyniki egzaminacyjne zapewniają porównywalność rezultatów egzaminacyjnych pomiędzy latami, dlatego też nazwaliśmy je Porównywalnymi Wynikami Egzaminacyjnymi (PWE). Mogą być one wykorzystywane przede wszystkim do monitorowania poziomu osiągnięć uczniów (sprawdzanych w trakcie egzaminu) w skali kraju. Mogą też być przydatne do monitorowania w wieloletnich przedziałach czasowych procesów edukacyjnych w skali województwa i powiatu.

W tym serwisie internetowym można zapoznać się ze zrównanymi wynikami egzaminacyjnymi z lat 2002-2013 dla poszczególnych szkół podstawowych/gimnazjów/liceów/techników, gmin, powiatów, województw i kraju, jak również dowolnie zestawiać je ze sobą.

Wyszukiwarka

Parametry wykresu

Opisy

-  Czym są PWE?
-  Procedura zrównywania
-  Jak odczytywać wykresy?

Publikacje

Opisy

i Czym są PWE?

⚙️ Procedura zrównywania

📊 Jak odczytywać wykresy?



- efekt badań zrównujących Zespołu Analiz Osiągnięć Uczniów w latach 2011-2014 w różnych typach szkół
- uzyskano porównywalność wyników sprawdzianu i egzaminu gimnazjalnego **od 2002 roku**, a matury z j. polskiego, angielskiego oraz matematyki **od 2010 roku**
- rok bazowy dla wszystkich egzaminów to **rok 2012** (krajowa **średnia** dla 2012 roku wynosi **100**)
- serwis zawiera zrównane wyniki egzaminacyjne z lat **2002-2013** dla poszczególnych szkół podstawowych/gimnazjów/liceów/techników, gmin, powiatów, województw i kraju.

Opisy

i Czym są PWE?

⚙ Procedura zrównywania

📈 Jak odczytywać wykresy?



- pozwala skorygować wyniki różnych edycji tego samego egzaminu (np. w kolejnych latach) w taki sposób, aby wyeliminować z nich czynnik wynikający z **różnej trudności arkuszy** egzaminacyjnych zastosowanych w poszczególnych latach.

2011

2012


2013


2014

egzaminu g1n z języka polskiego dla uczniów z obcego państwa i niepełnosprawności

Opisy

 Czym są PWE?

 Procedura zrównywania


 Jak odczytywać wykresy?





- Jak odczytywać wykres podstawowy?
- Jak odczytywać wykres zaawansowany?
- Przykłady




Publikacje

 *Porównywalne wyniki egzaminacyjne (2015)*

 *Analiza porównawcza wyników egzaminów zewnętrznych - sprawdzian w szóstej klasie szkoły podstawowej i egzamin gimnazjalny (2012)*

 *Modele cech ukrytych w badaniach edukacyjnych (2015)*

 *Efekt egzaminatora w ocenianiu prac maturalnych z języka polskiego i matematyki (2015)*



Wpisz nazwę lub adres

tylko całe wyrazy

Typ szkoły / JST:

- Wybierz -

Położenie szkoły / JST:

- Wybierz województwo -

- Wybierz powiat -

- Wybierz gminę -

Z wynikami za lata:

- 2013
- 2011
- 2009
- 2007
- 2005
- 2003
- 2012
- 2010
- 2008
- 2006
- 2004
- 2002

Szukaj

Wyniki wyszukiwania (27)

Szkoły (27):

wszystkie +

Gimnazjum im. św. Jadwigi Królowej w Kielcach
(Słowackiego 5, 25-365 Kielce)



Gimnazjum Integracyjne nr 4 im. Jana Karłowicza w Zespole Szkół
Ogólnokształcących Integracyjnych nr 4 w Kielcach
(Jasna 20/22, 25-523 Kielce)



Gimnazjum Językowe w Kielcach
(Karczówkowska 27, 25-713 Kielce)



Gimnazjum Niepubliczne w Kielcach
(Helenówek 4, 25-661 Kielce)



Gminy (0):

wszystkie +

Powiaty (0):

wszystkie +

Województwa (0):

wszystkie +

Wpisz nazwę lub adres

tylko całe wyrazy

Typ szkoły / JST:

Gimnazjum

Położenie szkoły / JST:

świętokrzyskie

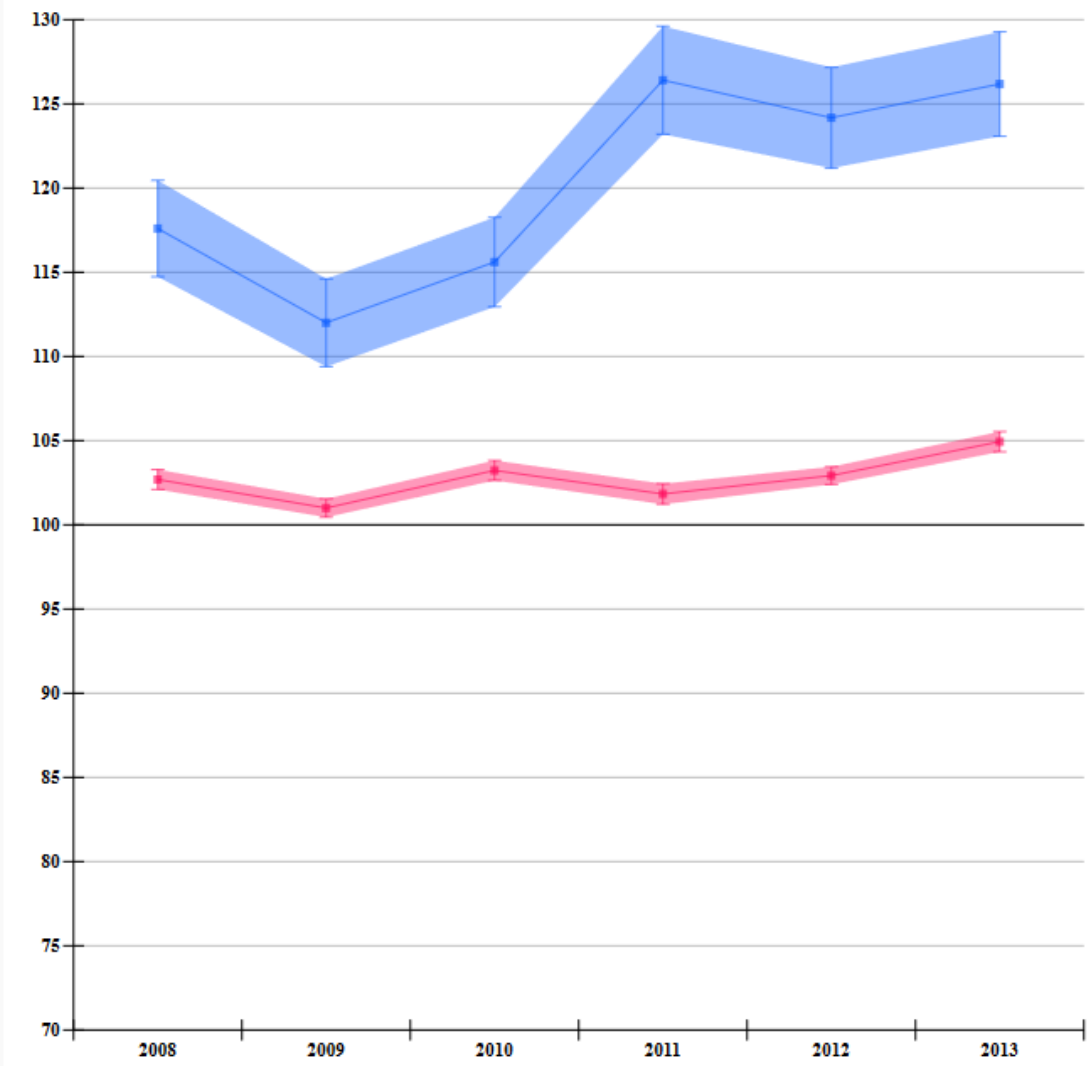
Kielce

Kielce

Z wynikami za lata:

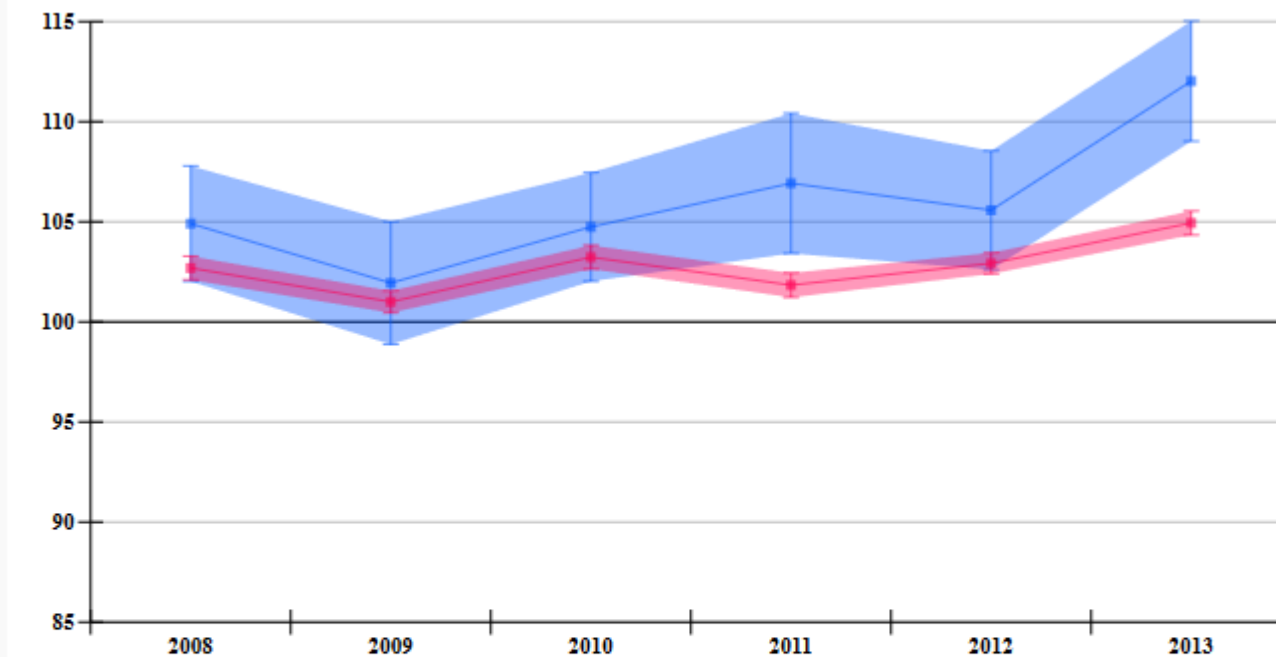
- 2013
- 2011
- 2009
- 2007
- 2005
- 2003
- 2012
- 2010
- 2008
- 2006
- 2004
- 2002

Szukaj



[Pokaż tabelę danych](#) [Pobierz jako obraz PNG](#) [Pobierz jako dane CSV](#) [Wyczyść analizę](#)

- Gimnazjum im. św. Jadwigi Królowej w Kielcach (Słowackiego 5, 25-365 Kielce) G P W ✕
- powiat Kielce (powiat Kielce, województwo świętokrzyskie) W ✕



Pokaż tabelę danych

Pobierz jako obraz PNG

Pobierz jako dane CSV

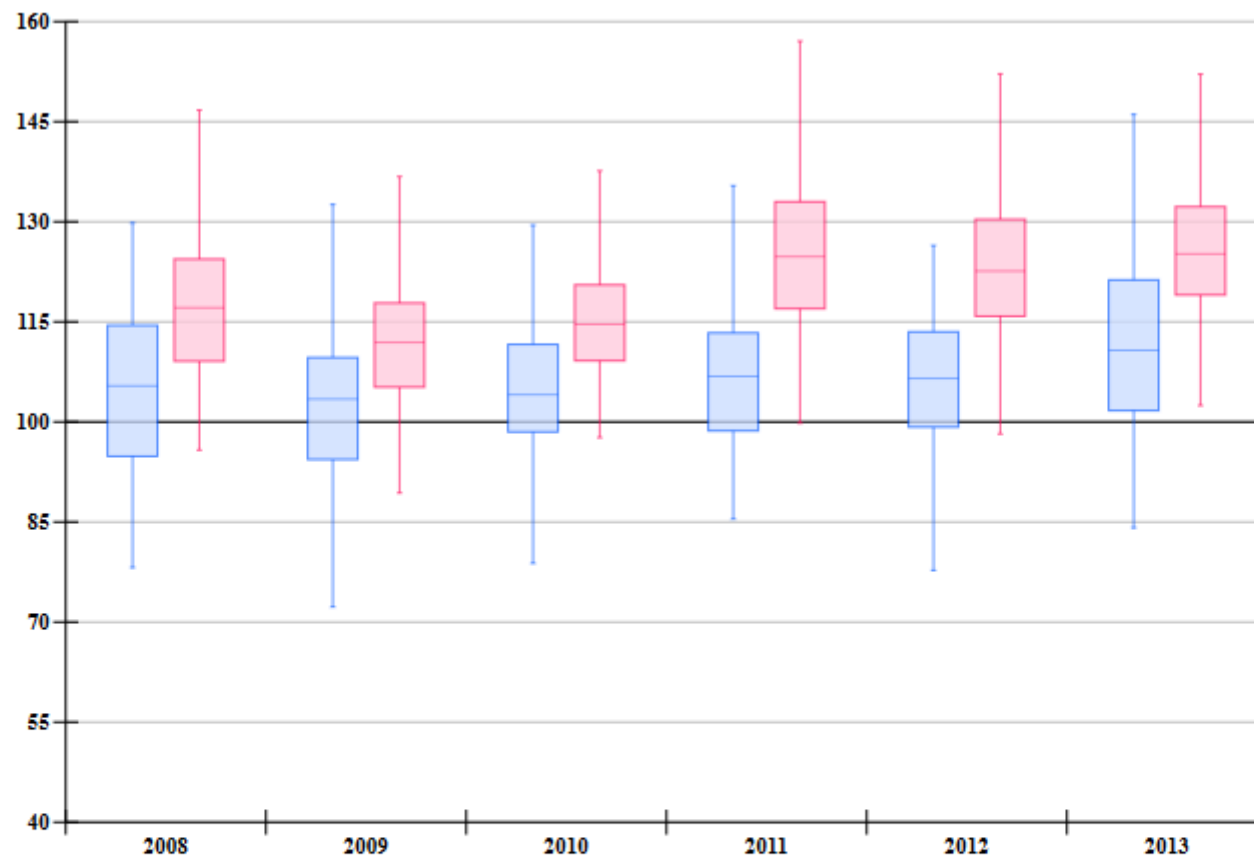
Wyczyść analizę

Gimnazjum nr 26 w Kielcach
(Dygasińskiego 6, 25-120 Kielce)

G P W ✕

powiat Kielce
(powiat Kielce, województwo świętokrzyskie)

W ✕



Pokaż tabelę danych

Pobierz jako obraz PNG

Pobierz jako dane CSV

Wyczyść analizę



Gimnazjum nr 26 w Kielcach
(Dygasińskiego 6, 25-120 Kielce)

G

P

W



Gimnazjum im. św. Jadwigi Królowej w Kielcach
(Słowackiego 5, 25-365 Kielce)

G

P

W



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE  entuzjaści
edukacji

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY





entuzjaści edukacji

Regionalne zmiany w wynikach egzaminacyjnych

Filip Kulon

ZESPÓŁ ANALIZ OSIĄGNIĘĆ UCZNIÓW

Instytut Badań Edukacyjnych



Możliwości analizy terytorialnej wyników egzaminów

- Różne poziomy podziału terytorialnego:
 - regiony,
 - województwa,
 - powiaty,
 - gminy.

- Różne egzaminy:
 - sprawdzian w klasie szóstej szkoły podstawowej,
 - egzamin gimnazjalny (część humanistyczna i matematyczno-przyrodnicza),
 - egzamin maturalny (matematyka, język angielski).



Podział kraju na regiony wg klasyfikacji zgodnej z NTS

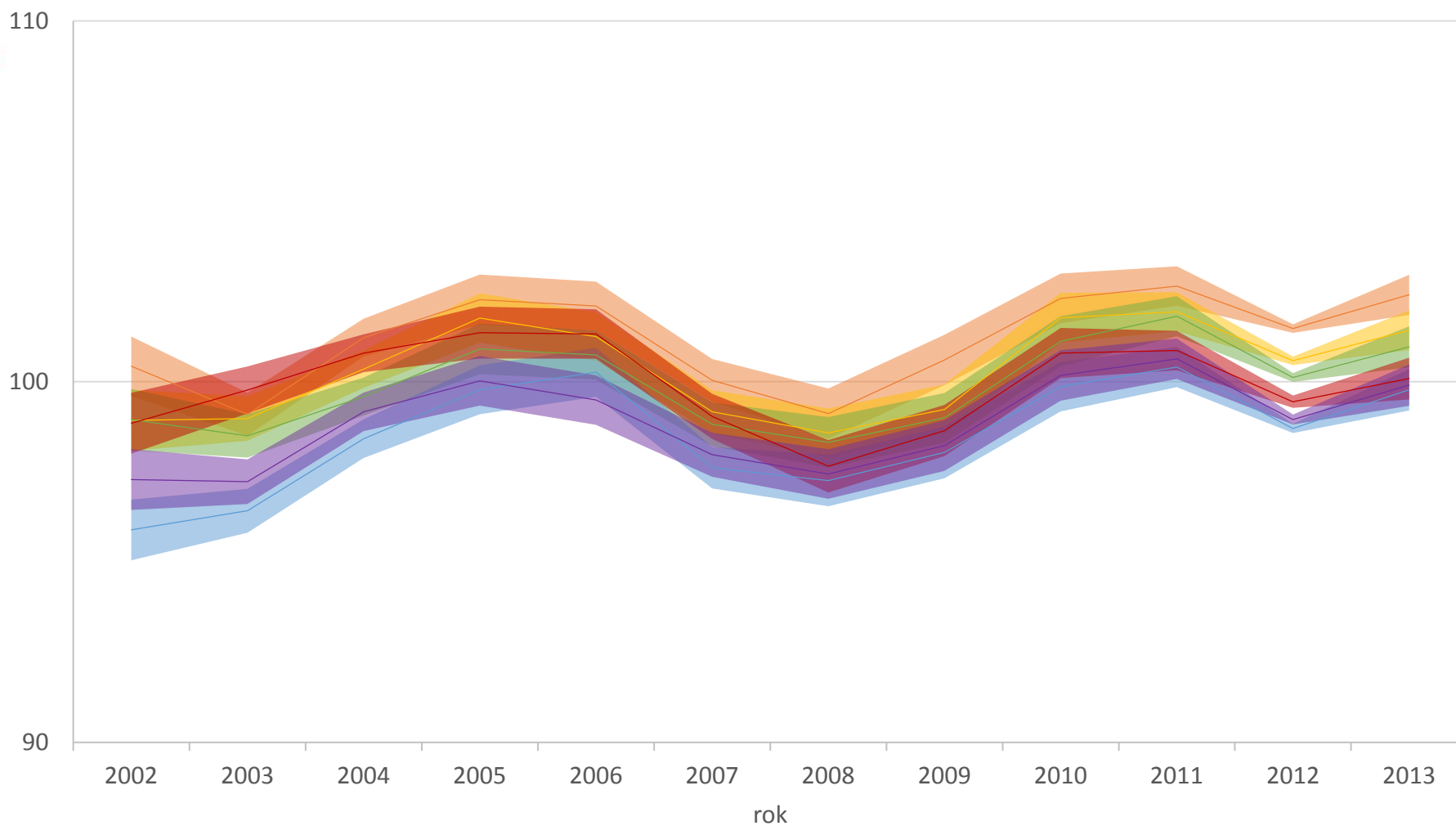
Nomenklatura Jednostek Terytorialnych do Celów Statystycznych
(Dz. U. z 2000 r. Nr 58, poz. 685)



Kod	Region	Województwa
1	centralny	łódzkie i mazowieckie
2	południowy	małopolskie i śląskie
3	wschodni	lubelskie, podkarpackie, świętokrzyskie i podlaskie
4	północno-zachodni	wielkopolskie, zachodniopomorskie i lubuskie
5	południowo-zachodni	dolnośląskie i opolskie
6	północny	kujawsko-pomorskie, warmińsko-mazurskie i pomorskie



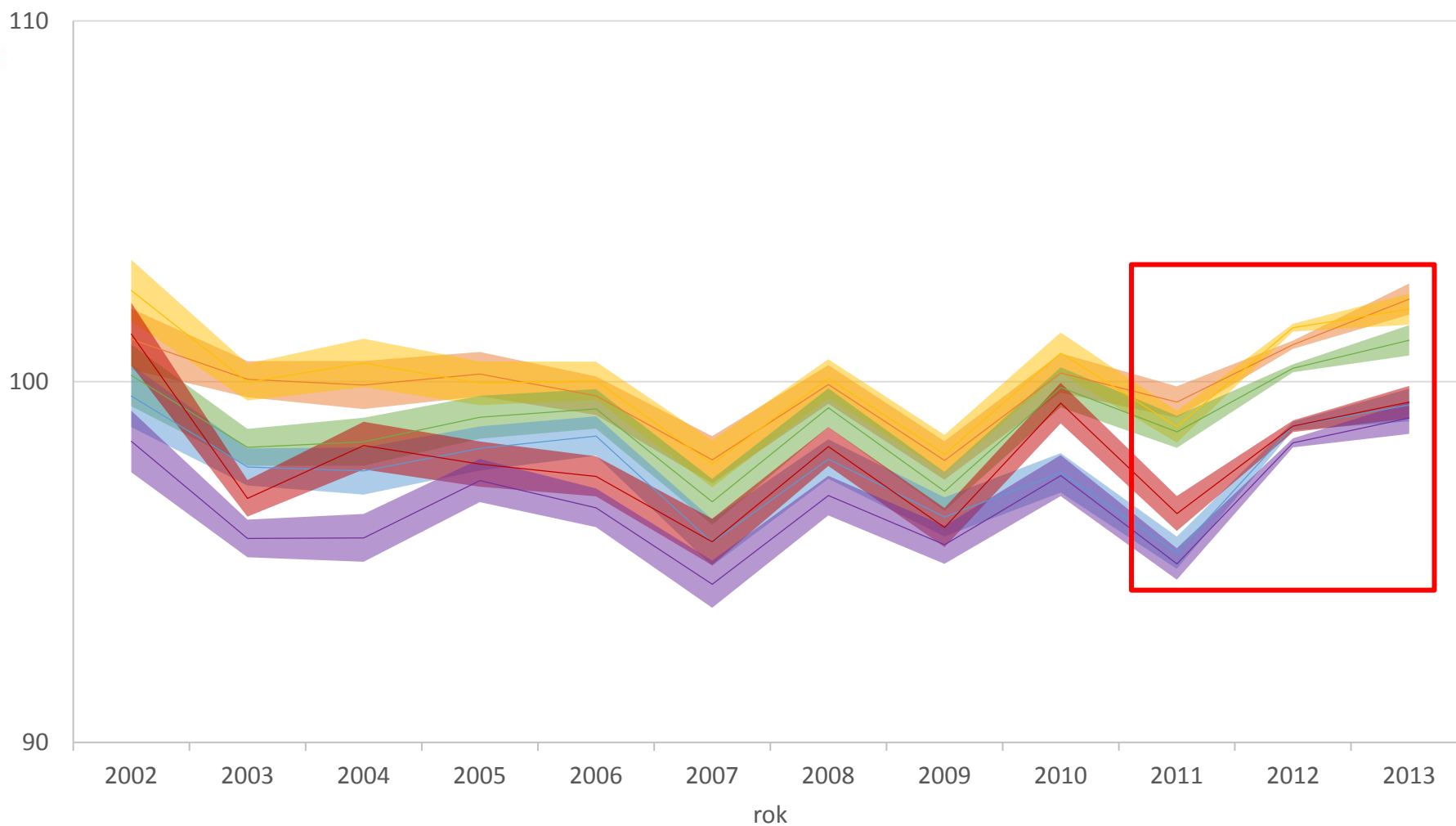
Różnice regionalne: sprawdzian



regiony: centralny południowy wschodni północno-zachodni południowo-zachodni północny



Różnice regionalne: egzamin gimnazjalny, część humanistyczna



regiony: centralny południowy wschodni północno-zachodni południowo-zachodni północny



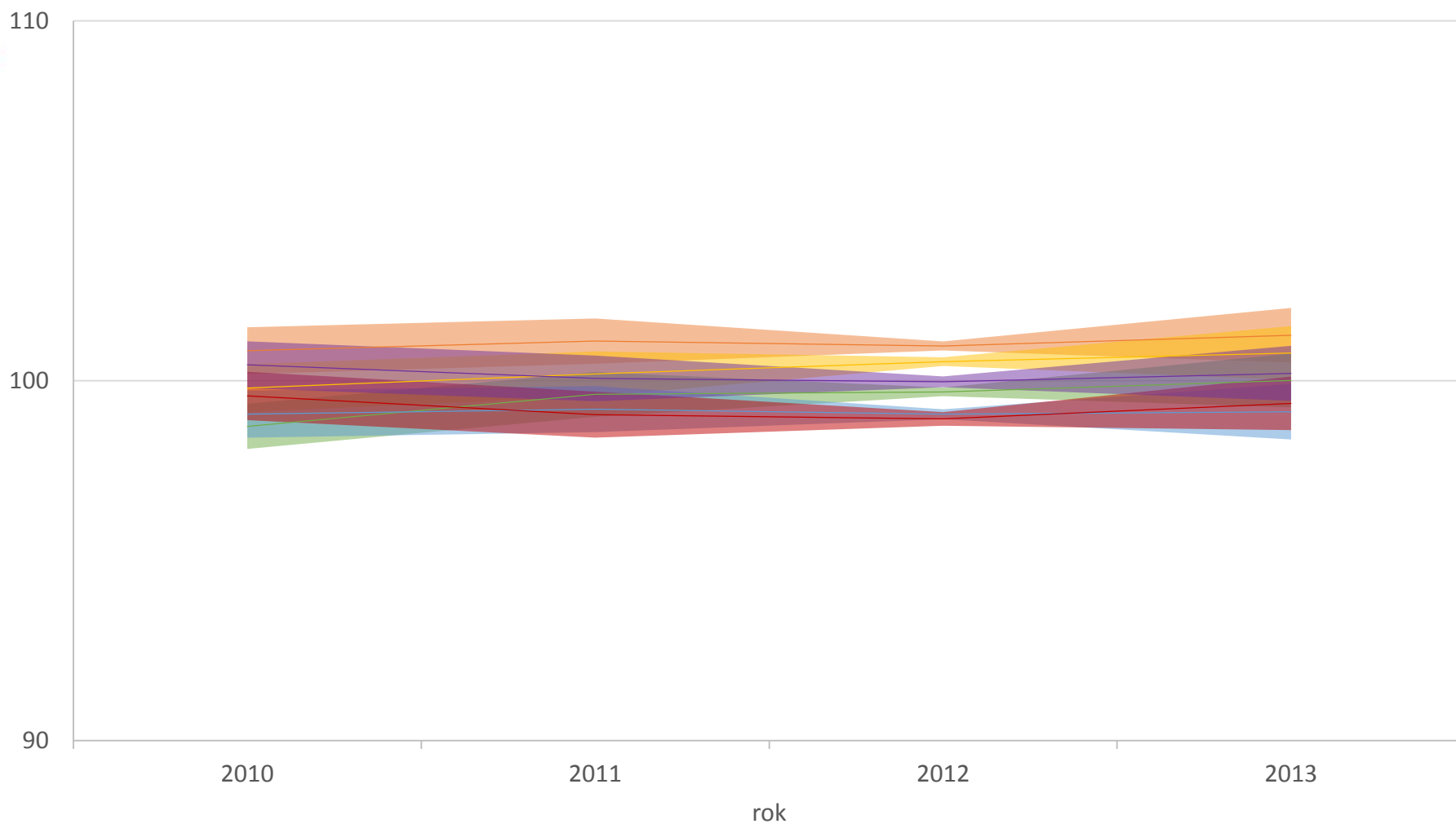
Różnice regionalne: egzamin gimnazjalny, część matematyczno-przyrodnicza



regiony: centralny południowy wschodni północno-zachodni południowo-zachodni północny



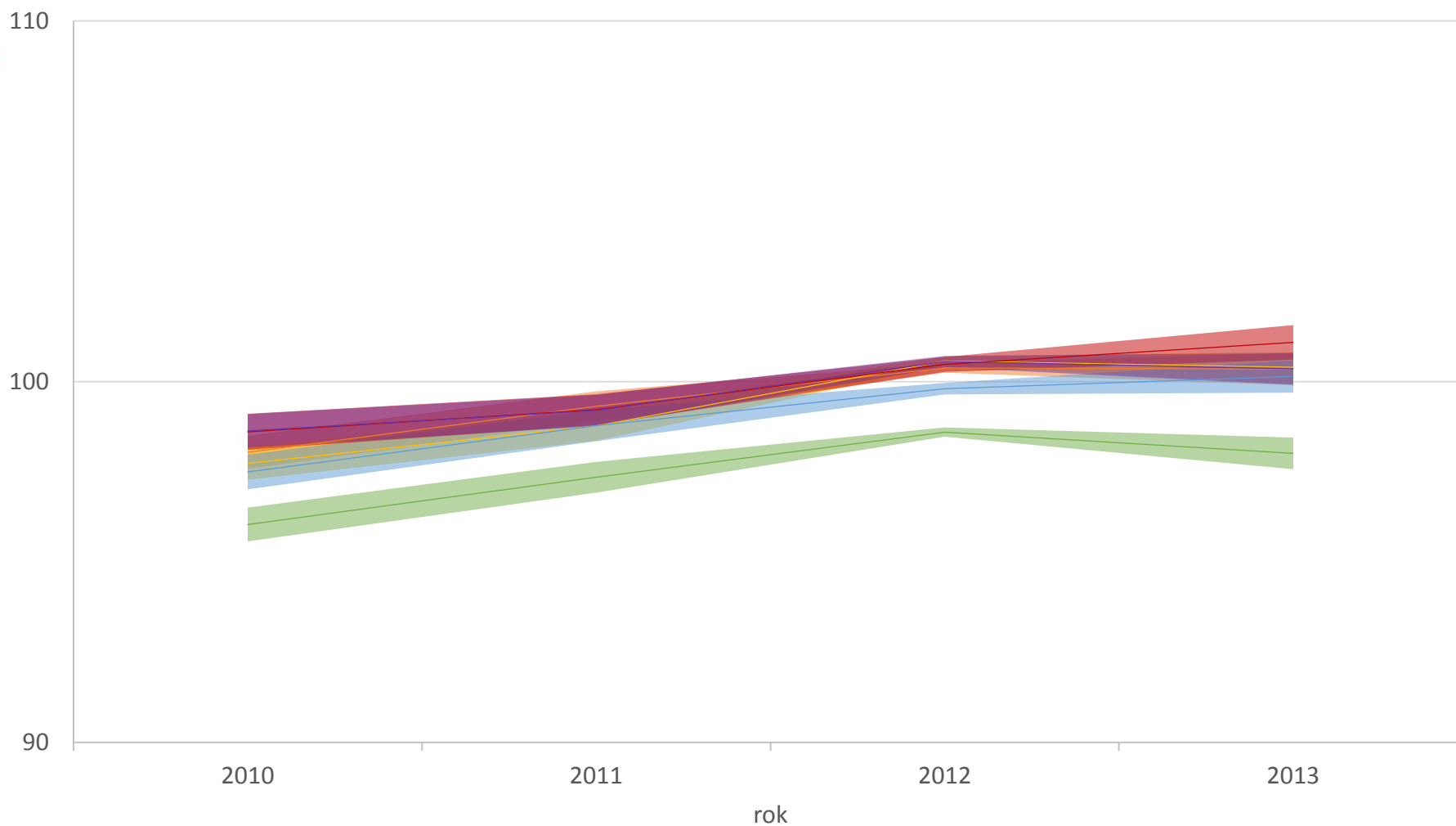
Różnice regionalne: matura, matematyka, poziom podstawowy



regiony: centralny południowy wschodni północno-zachodni południowo-zachodni północny



Różnice regionalne: matura, język angielski, poziom podstawowy



regiony: centralny południowy wschodni północno-zachodni południowo-zachodni północny

Czym są Porównywalne Wyniki Egzaminacyjne (PWE)?

Zespół Analiz Osiągnięć Uczniów (ZAOU) działający w Instytucie Badań Edukacyjnych podjął w ramach projektu „Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego” prace nad zrównywaniem wyników egzaminacyjnych, w celu osiągnięcia porównywalności wyników między latami.

Wynik egzaminu zależy od trudności arkusza testowego oraz umiejętności uczniów. Gdyby umiejętności uczniów w krajowej populacji nie zmieniały się, to średni krajowy wynik egzaminacyjny byłby informacją o trudności testu. Gdyby trudność arkuszy testowych w kolejnych latach była podobna, to na podstawie wyników egzaminacyjnych można byłoby wnioskować o tym, jak zmienia się w czasie poziom umiejętności uczniów. Skonstruowanie arkuszy testowych o podobnych parametrach jest jednak zadaniem bardzo trudnym i nie zawsze zakończonym powodzeniem. Porównywalność wyników między latami można osiągnąć kontrolując trudność arkusza (procedury zrównywania wyników) lub zakładając stały poziom umiejętności w populacji (normalizacja i standaryzacja rozkładów, np. do skali staninowej). W pierwszym przypadku uzyskujemy bezwzględną porównywalność wyników i możliwość wnioskowania o zmianie umiejętności w populacji. W drugim przypadku uzyskujemy względną porównywalność w stosunku do średniej w danym roku.

W latach 2011-2014 Zespół Analiz Osiągnięć Uczniów przeprowadził badania zrównujące w różnych typach szkół, co po zastosowaniu złożonej procedury zrównywania wyników pozwoliło uzyskać porównywalność wyników sprawdzianu i egzaminu gimnazjalnego od 2002 roku, a matury z j. polskiego, angielskiego oraz matematyki od 2010 roku. Za rok bazowy dla wszystkich egzaminów przyjęto rok 2012 (krajowa średnia dla 2012 roku wynosi 100), a wyniki danego egzaminu w pozostałych latach należy rozumieć jako wzrost/spadek poziomu umiejętności uczniów w porównaniu do 2012 roku.




Zrównane wyniki egzaminacyjne zapewniają porównywalność rezultatów egzaminacyjnych pomiędzy latami, dlatego też nazwaliśmy je Porównywalnymi Wynikami Egzaminacyjnymi (PWE). Mogą być one wykorzystywane przede wszystkim do monitorowania poziomu osiągnięć uczniów (sprawdzanych w trakcie egzaminu) w skali kraju. Mogą też być przydatne do monitorowania w wieloletnich przedziałach czasowych procesów edukacyjnych w skali województwa i powiatu.

W tym serwisie internetowym można zapoznać się ze zrównanymi wynikami egzaminacyjnymi z lat 2002-2013 dla poszczególnych szkół podstawowych/gimnazjów/liceów/techników, gmin, powiatów, województw i kraju, jak również dowolnie zestawiać je ze sobą.

Wyszukiwarka

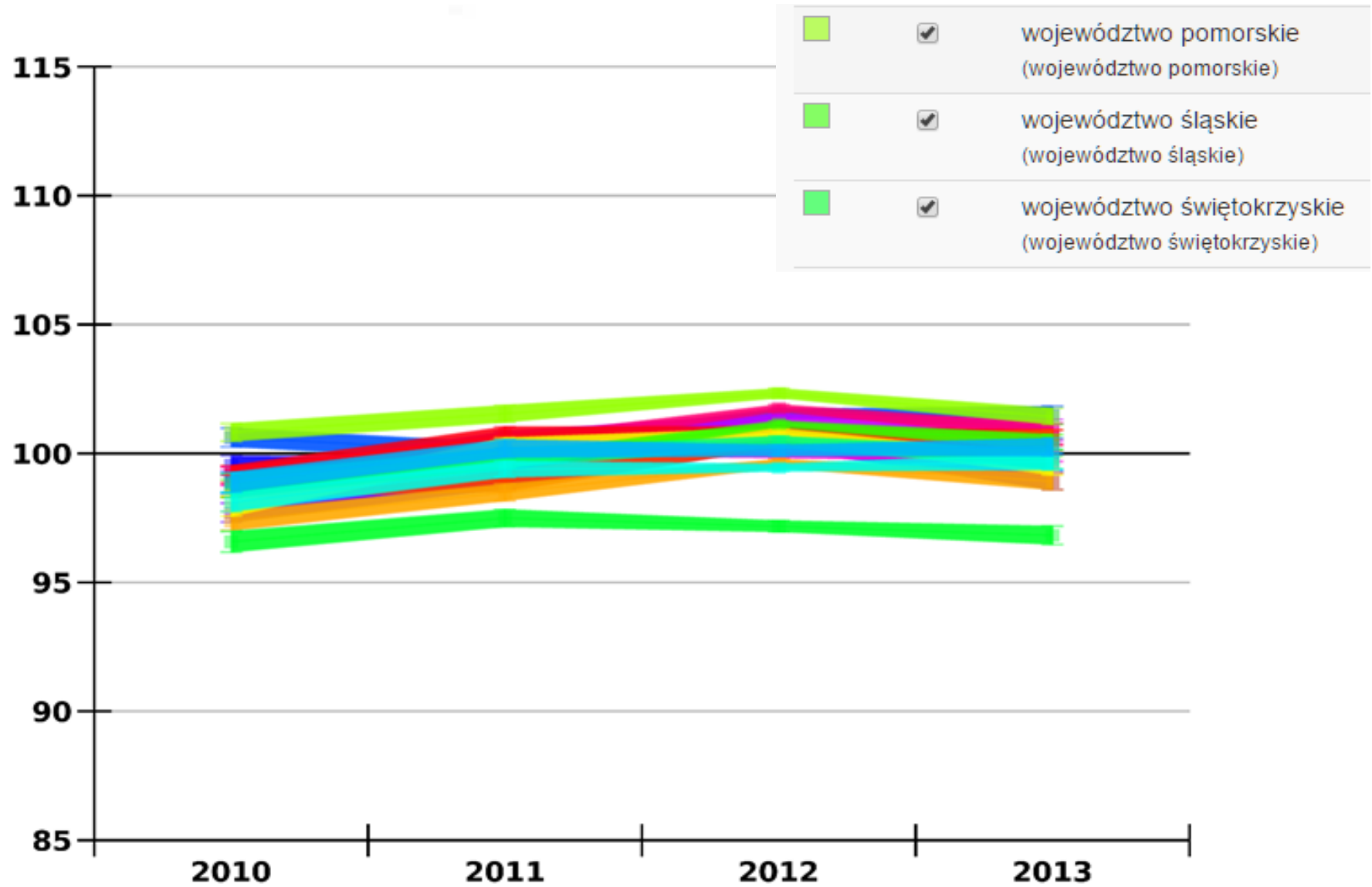
Parametry wykresu

Opisy

-  Czym są PWE?
-  Procedura zrównywania
-  Jak odczytywać wykresy?

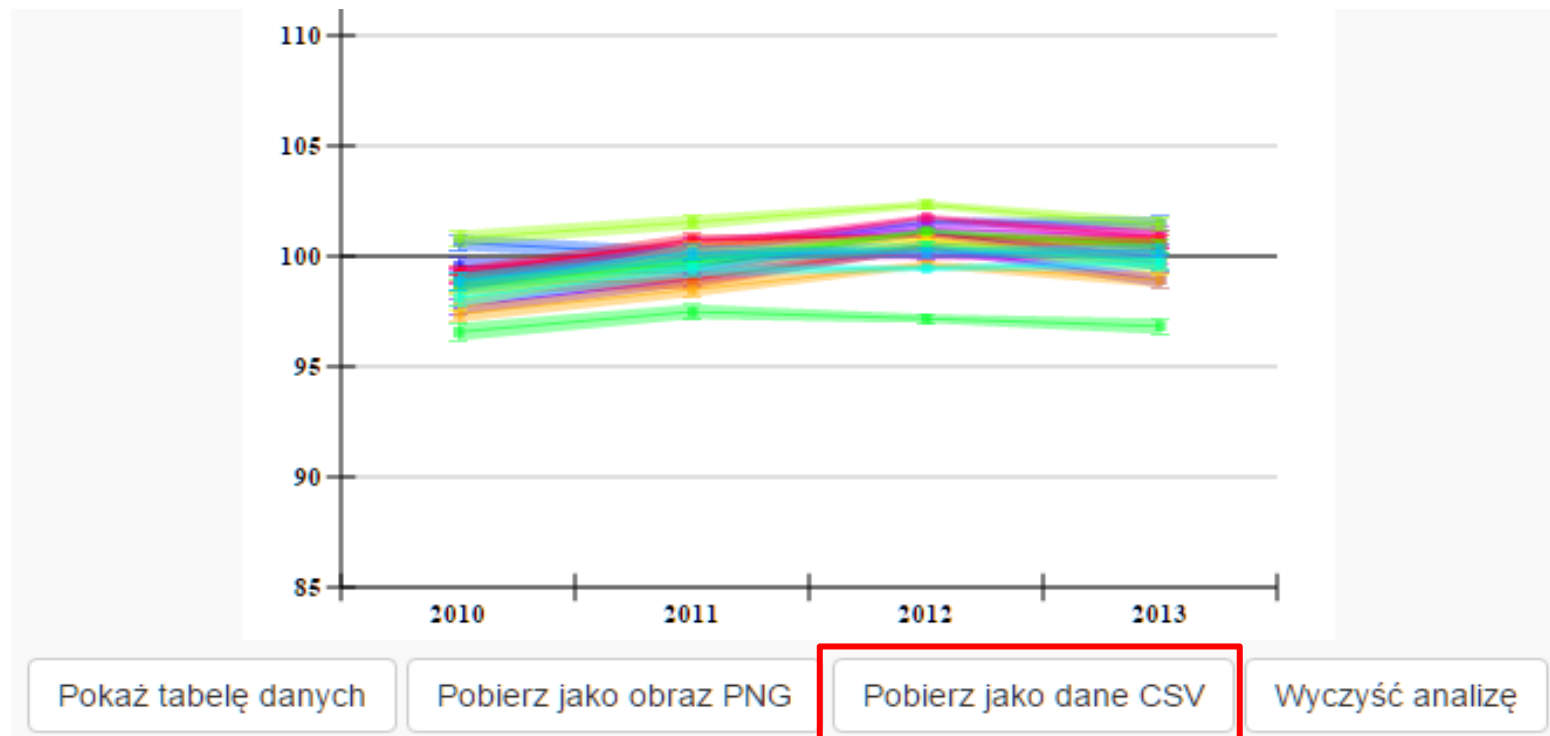
Publikacje

J. angielski – strona PWE, województwa





A gdyby przedstawić dane w inny sposób?



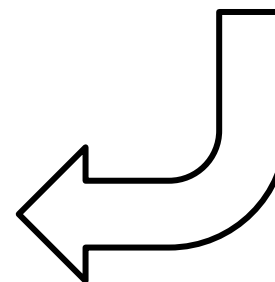
Microsoft Excel - danePWE

WYKŁAD STRONY FORMUŁY DANE RECENZJA WIDOK

Wklej Wytnij Kopiaj Malarz formatów Schowek Czcionka Wyrównanie Liczba Formuła

A1 id_województwa

	A	B	C	D	E	F	G	H	I	J	K
1	id_województwa	nazwa_jed	statystyka	2010	2011	2012	2013				
2	20000	województw max		161,727	157,717	159,357	162,867				
3	20000	województw q3		109,078	109,479	110,376	111,083				
4	20000	województw ggpu		100,9775	100,4074	101,716	101,8222				
5	20000	województw srednia		100,615	100,07	101,545	101,486				
6	20000	województw mediana		101,815	101,082	102,681	103,101				
7	20000	województw dgpu		100,2525	99,73258	101,374	101,1498				
8	20000	województw q1		93,865	92,4584	94,4709	93,8661				
9	20000	województw min		30,3285	36,2856	37,4001	26,2565				
10	20000	województw lu		14790	17225	16052	16697				
11	20000	województw bs		0,257769	0,239953	0,121631	0,239102				
12	20000	województw teryt		20000	20000	20000	20000				
13	40000	województw max		159,529	154,505	158,272	165,507				
14	40000	województw q3		107,752	108,035	109,499	108,801				
15	40000	województw ggpu		99,90707	99,84637	101,2844	100,0013				
16	40000	województw srednia		99,5352	99,5129	101,093	99,6629				
17	40000	województw mediana		100,6	99,9815	101,578	100,864				





http://zpd.ibe.edu.pl/doku.php?id=generator_map

	A	B	C
1	id_województwa	rok	kolor
2	2	2010	#ECE583
3	4	2010	#FEDF82
4	6	2010	#FCAF79
5	8	2010	#FDCE7E
6	10	2010	#FED480
7	12	2010	#FED580
8	14	2010	#FED680
9	16	2010	#FDCA7E
10	18	2010	#FBA777
11	20	2010	#FCB67A
12	22	2010	#E6E483
13	24	2010	#FDC87D
14	26	2010	#FA9273
15	28	2010	#FDCE7E
16	30	2010	#FCBA7B
17	32	2010	#FDCE7E
18	2	2011	#FDEA84
19	4	2011	#FEDE82

Generator map

Przygotuj plik danych CSV (separator pola - średnik, separator tekstu - cudzysłów) zawierający kolumny:

- rok - rok, z którego ma pochodzić podkład dla danego województwa/powiatu/gminy - obsługiwany jest zakres od 2002 do 2014 (tak, tak, od czasu do czasu granice JST się zmieniają)
- id_województwa - kod TERYT województwa
- id_powiatu [gdy mapa powiatów lub gmin] - kod TERYT powiatu w ramach województwa
- id_gminy [gdy mapa gmin] - kod TERYT gminy w ramach powiatu
- kolor - kolor w zapisie HTML (np. #FE059A).

Prześlij go korzystając z formularza poniżej, aby otrzymać mapę Polski w formacie PNG pokolorowaną zgodnie z przygotowanym plikiem danych (przykładowy plik danych dla województw).

Generowanie map w R

[Patrz tutaj](#)

Formularz

Szer. mapy: px (100-2000)
Wys. mapy: px (100-2000)

plik danych: Nie wybrano pliku

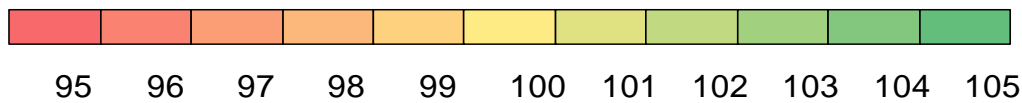
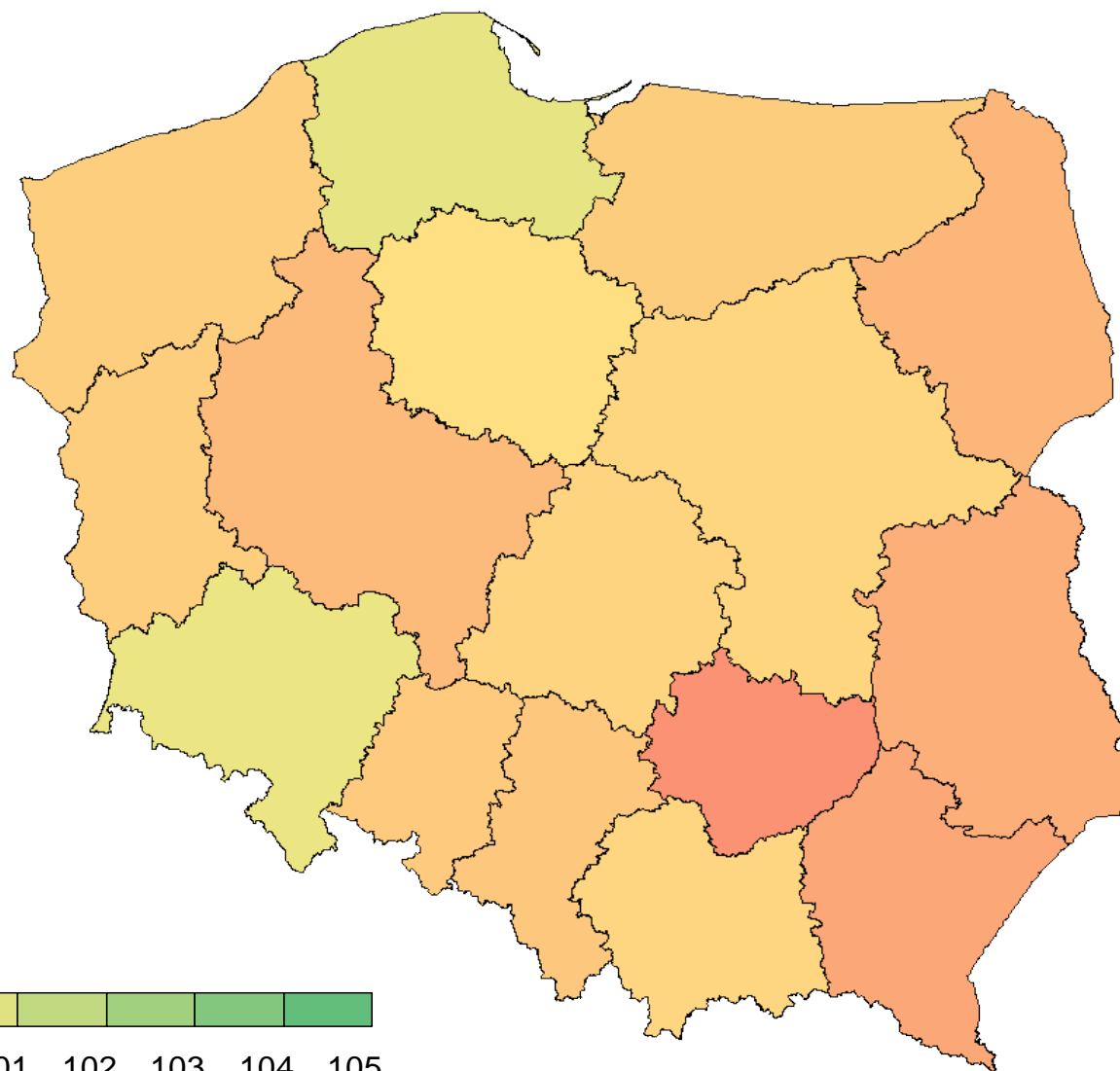
generator_map.txt · ostatnio zmienione: 2015/10/14 18:30 przez zozlak

Wszystkie treści w tym wiki, którym nie przyporządkowano licencji, podlegają licencji: CC Attribution-Share Alike 3.0 Unported

BY-SA DONATE PHP POWERED W3C HTML5 W3C CSS DOKUWIKI

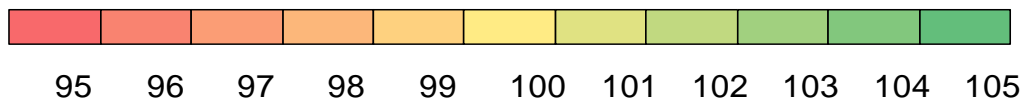
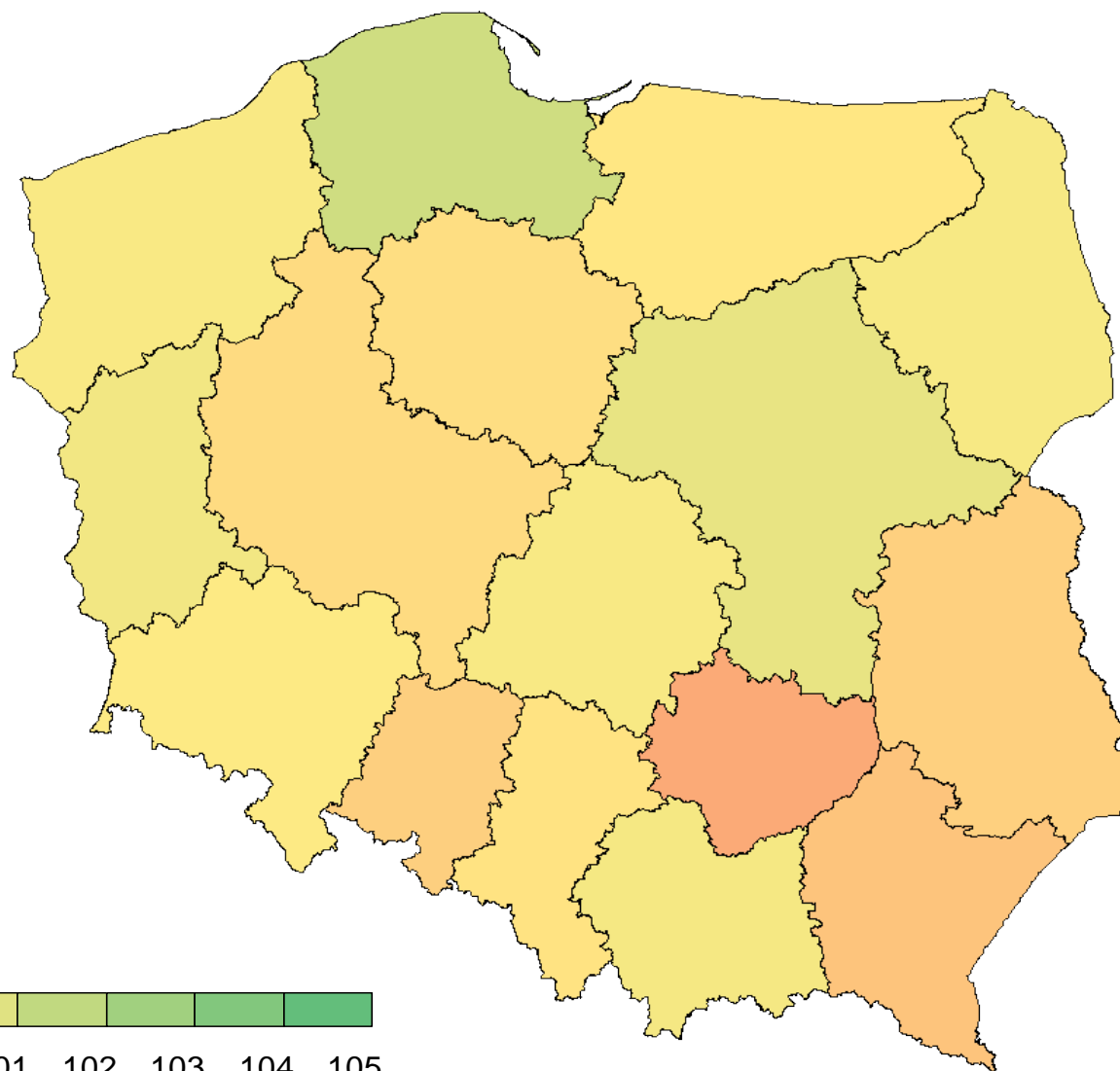


J. angielski 2010



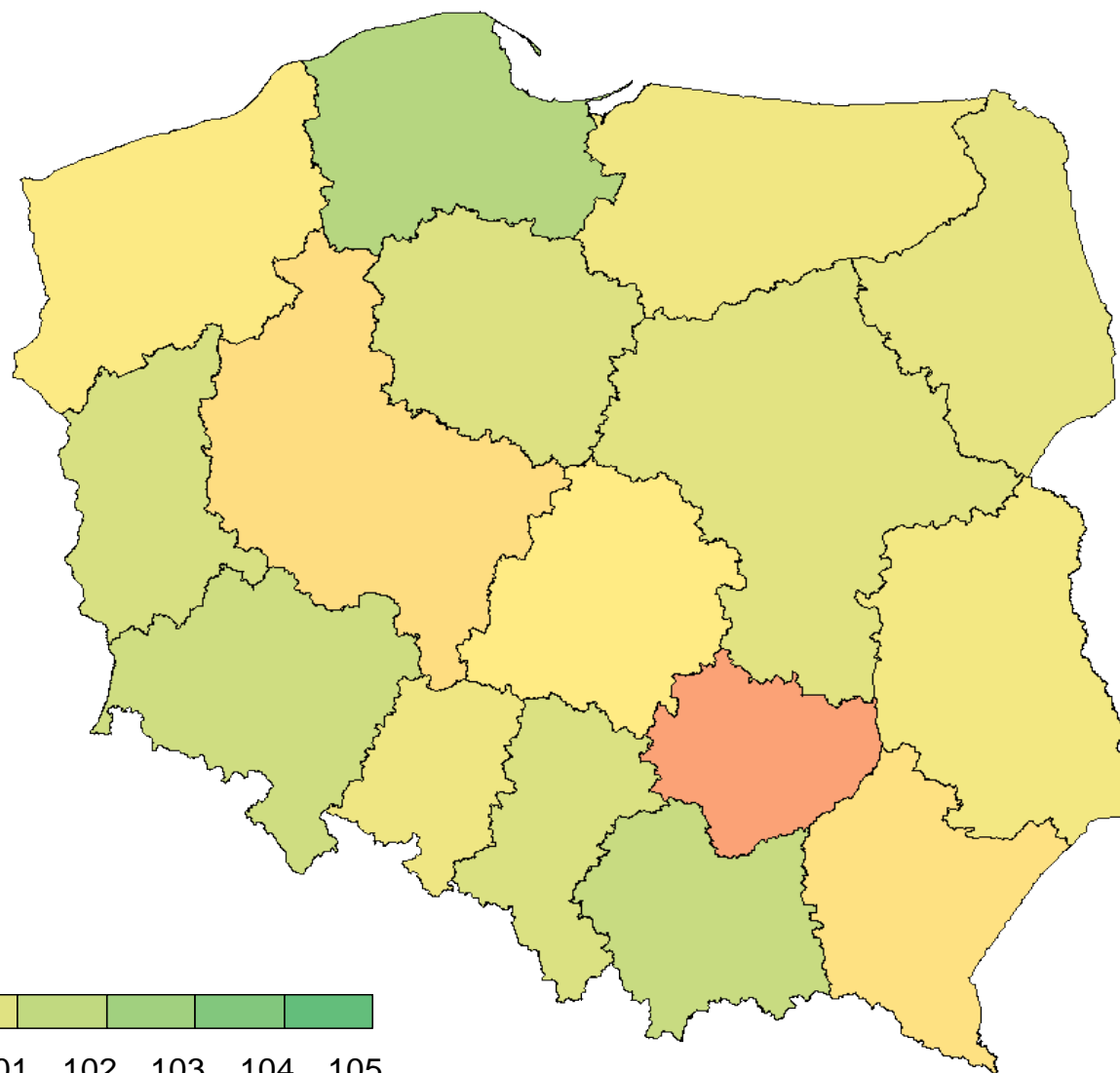


J. angielski 2011





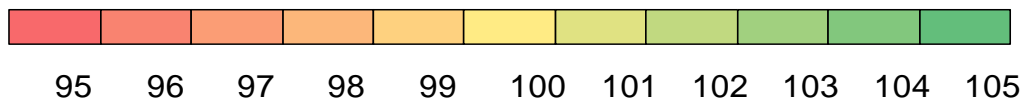
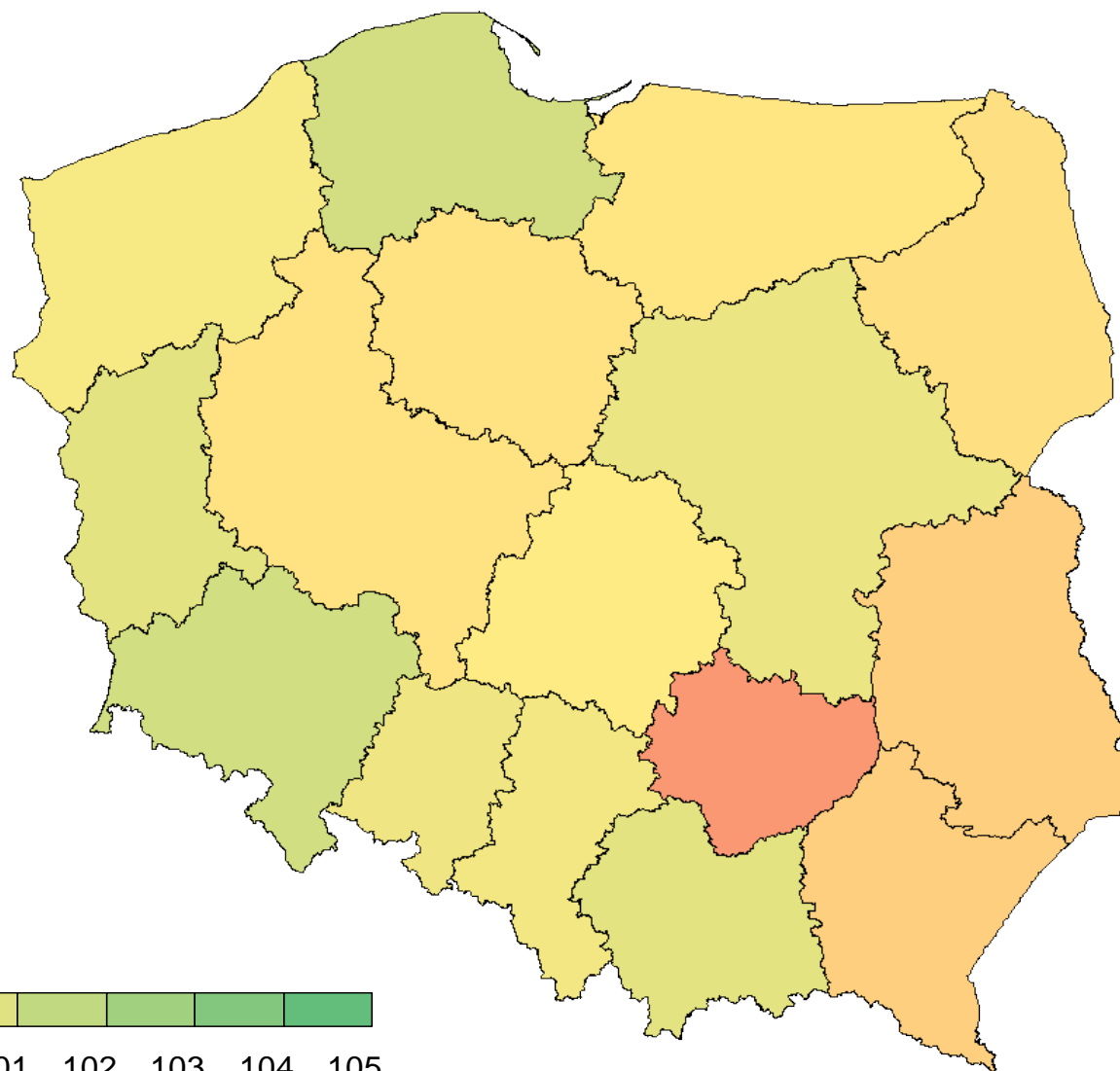
J. angielski 2012



95 96 97 98 99 100 101 102 103 104 105

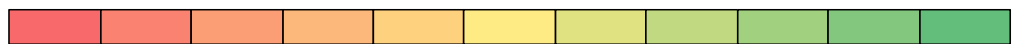
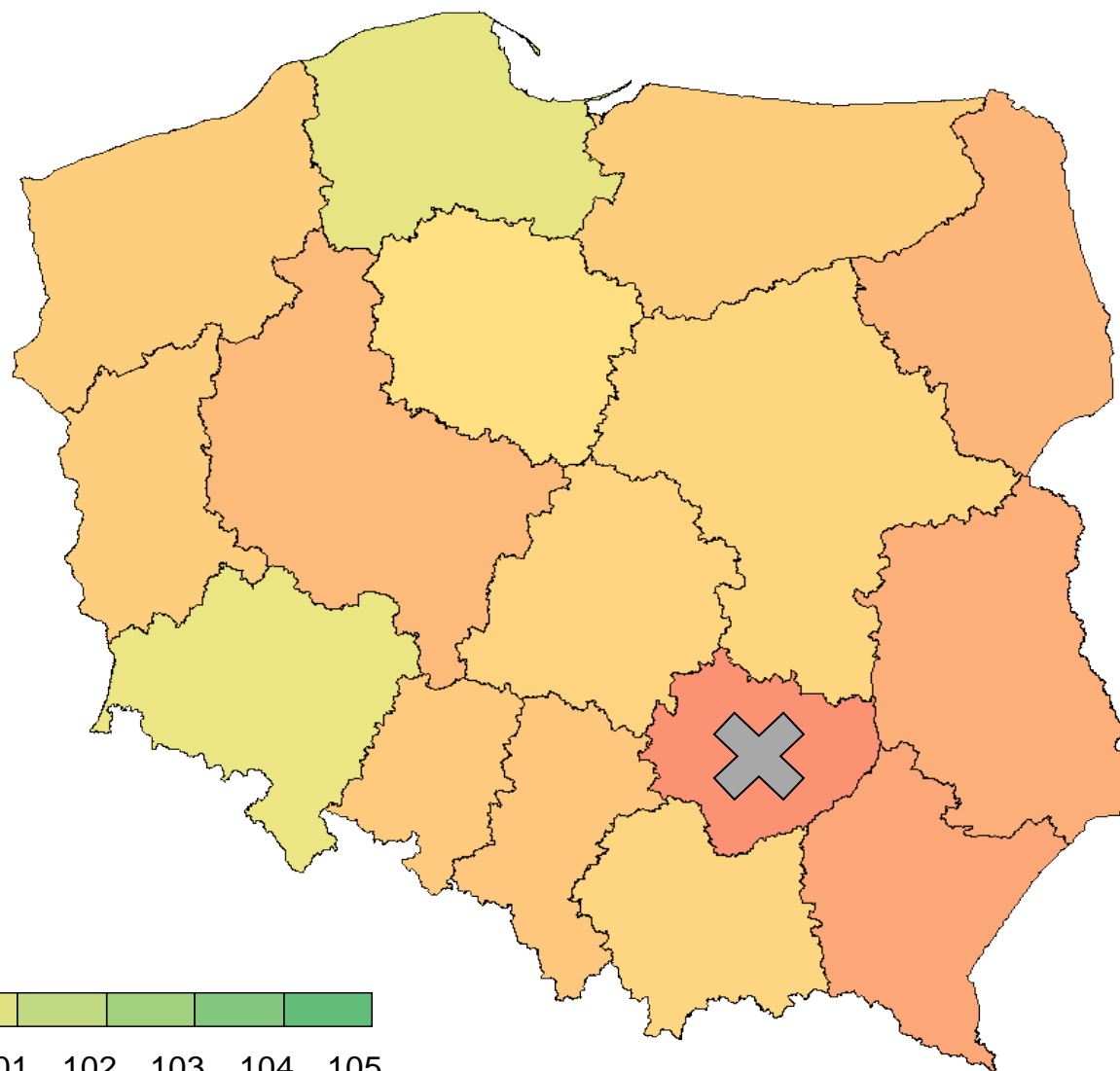


J. angielski 2013





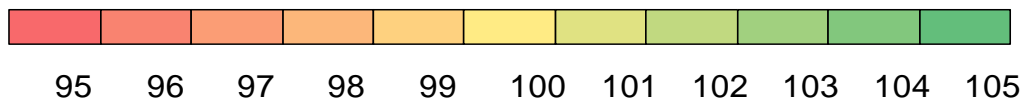
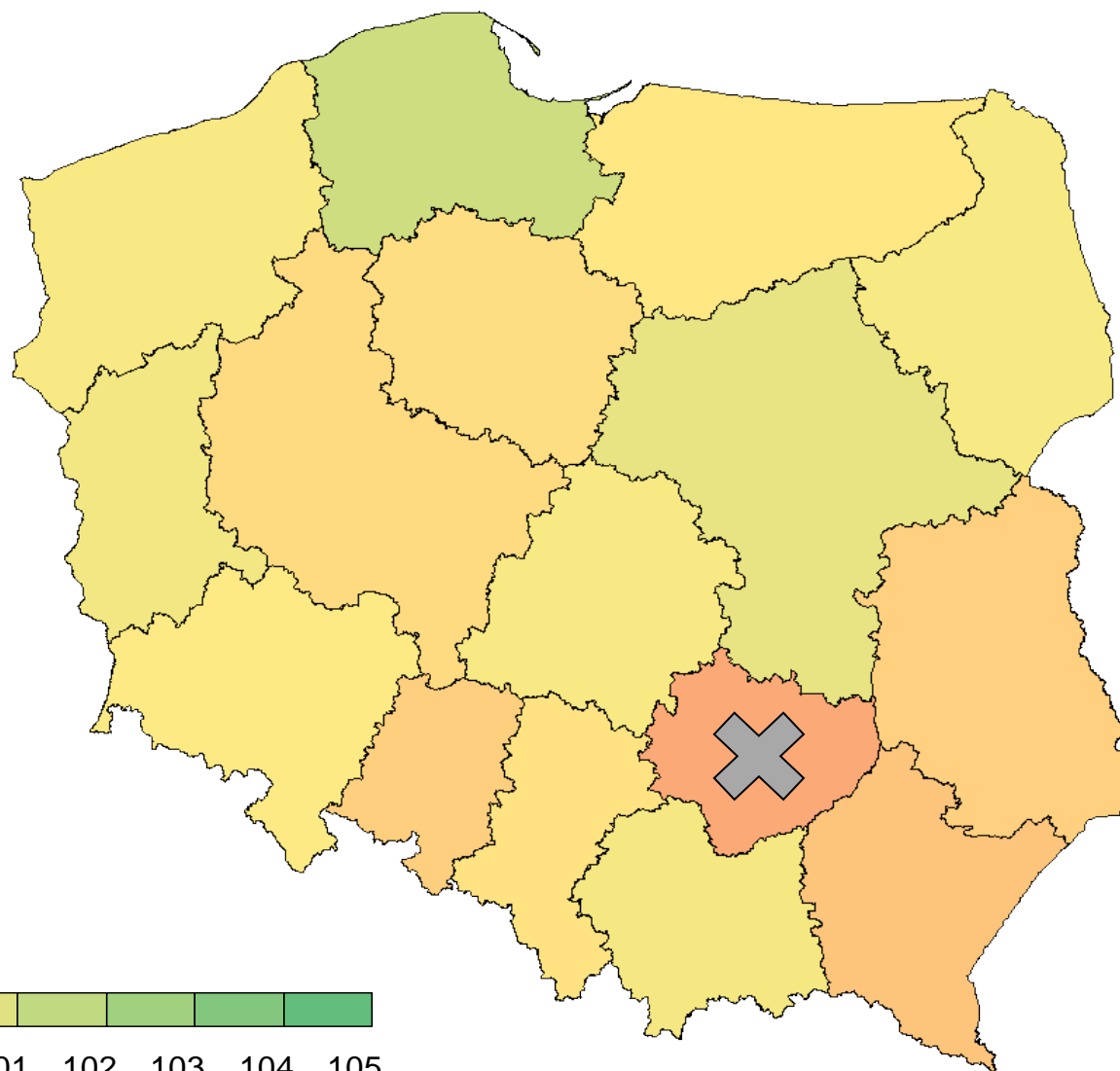
J. angielski 2010



95 96 97 98 99 100 101 102 103 104 105

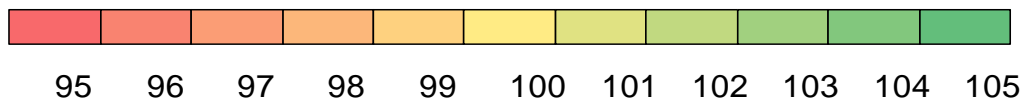
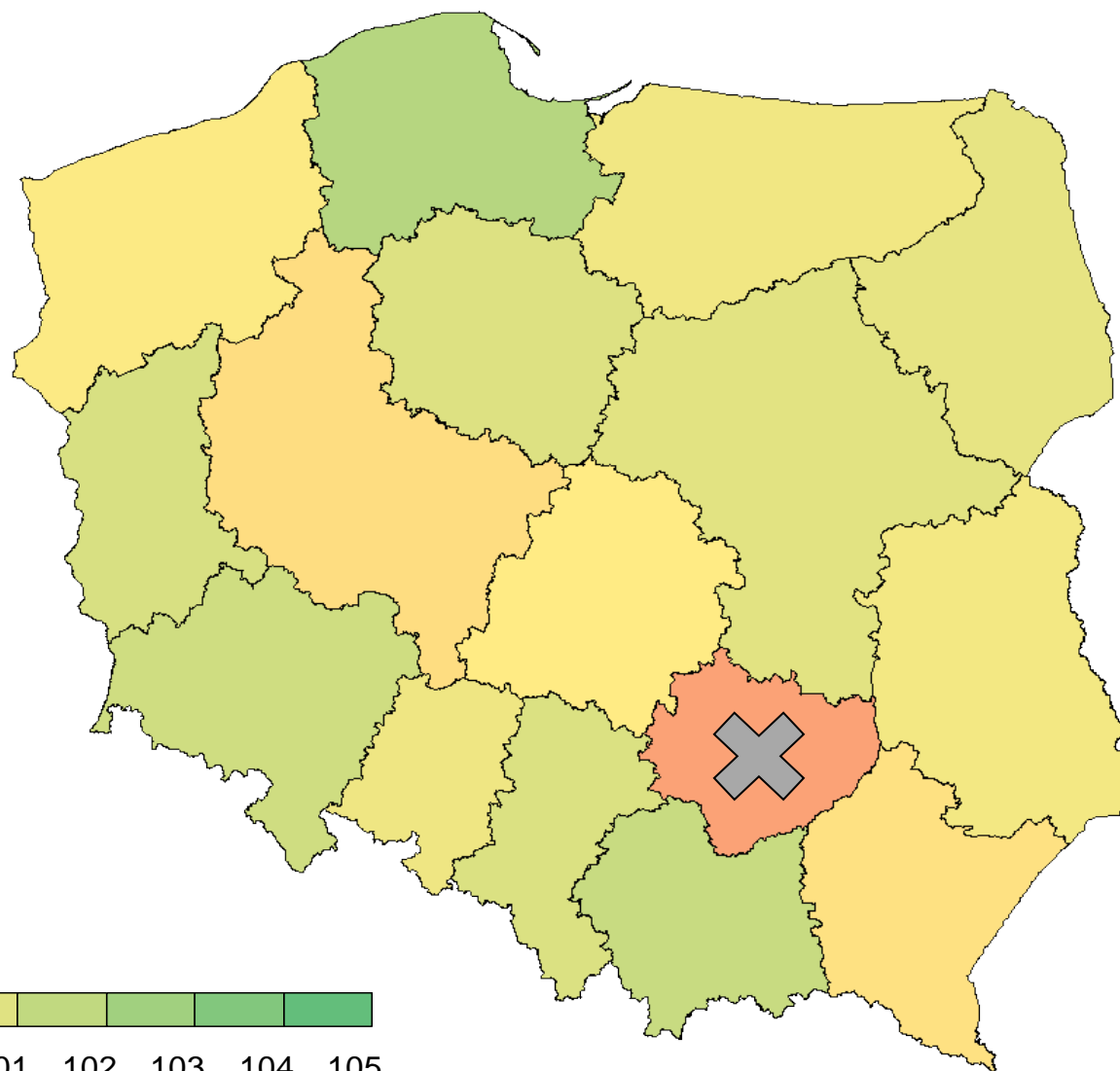


J. angielski 2011



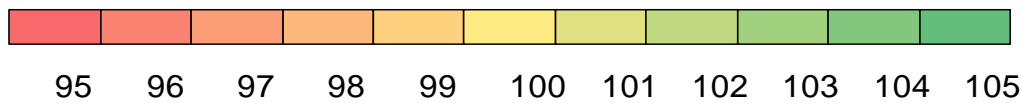
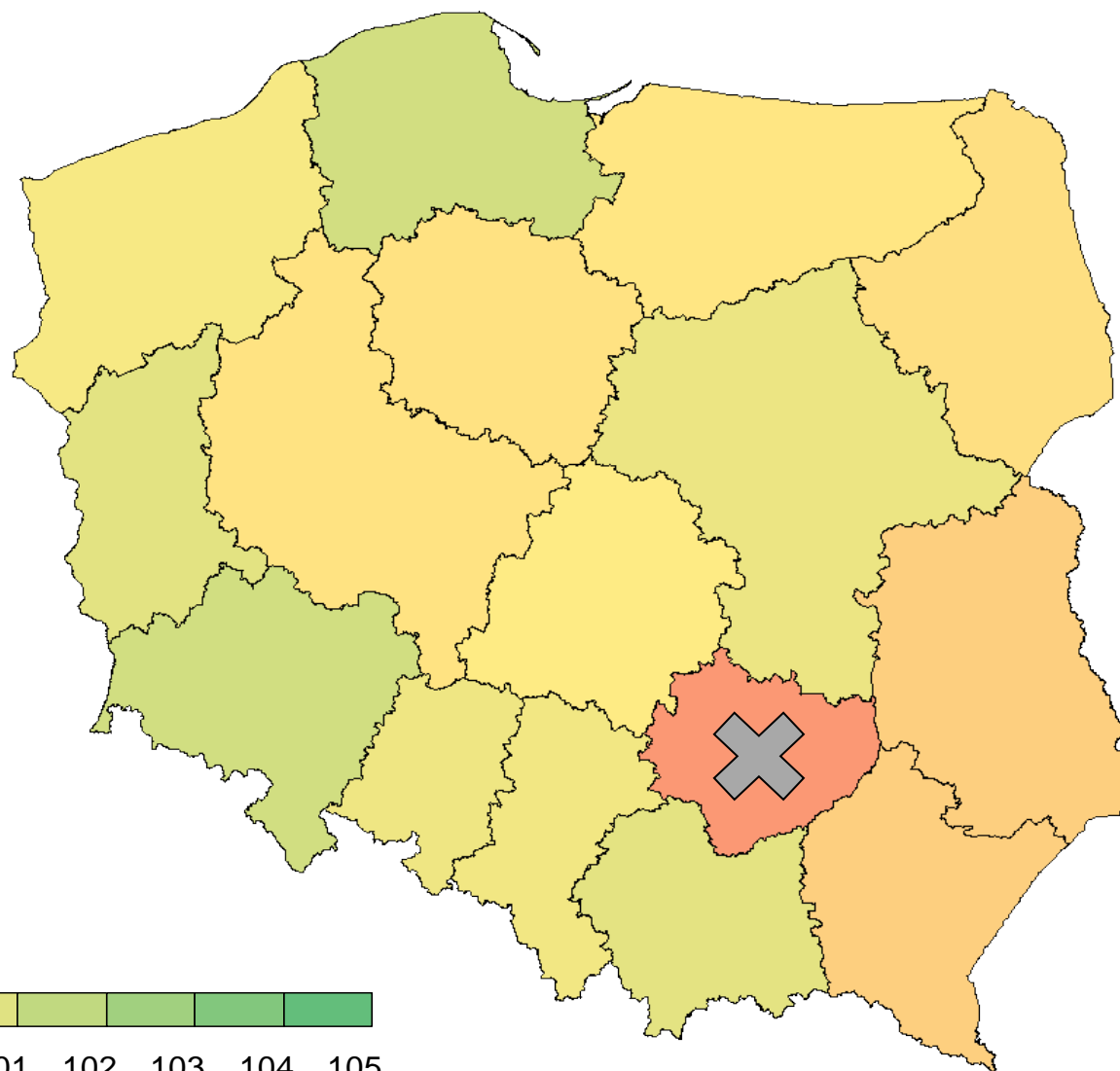


J. angielski 2012





J. angielski 2013



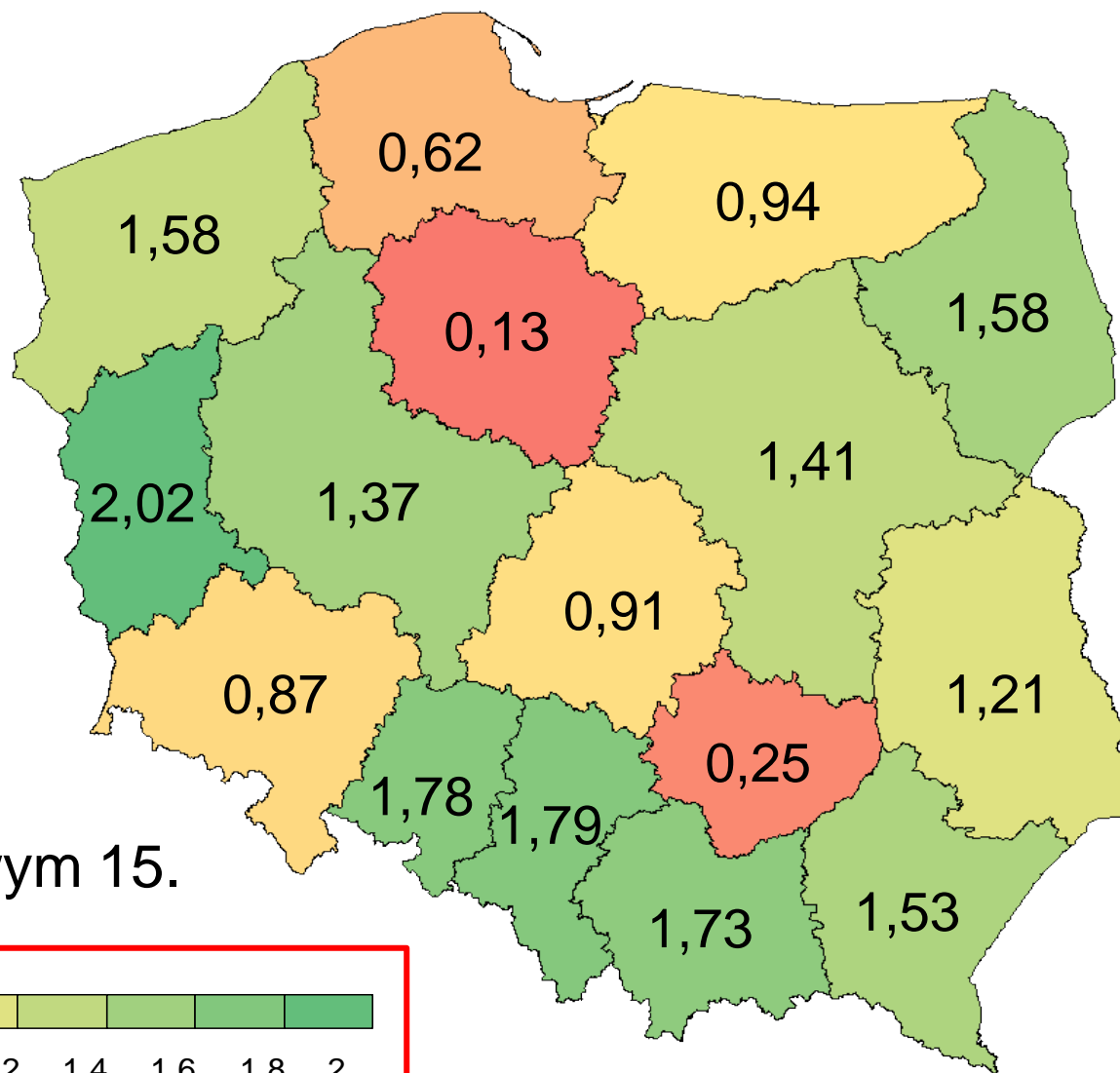


Porównywalne Wyniki Egzaminacyjne

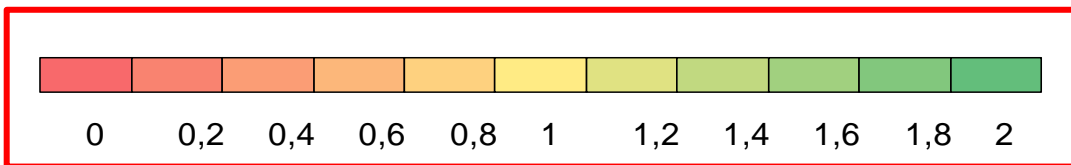
- Możliwość dokonywania przekształceń:
 - średnia,
 - odchylenie standardowe → rozproszenie,
 - różnica → przyrost.



Zmiany w czasie – matura, język angielski różnica pomiędzy 2010 i 2013

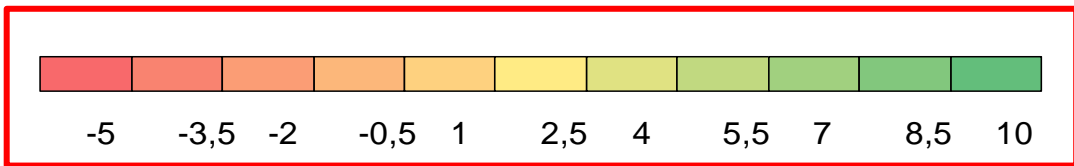
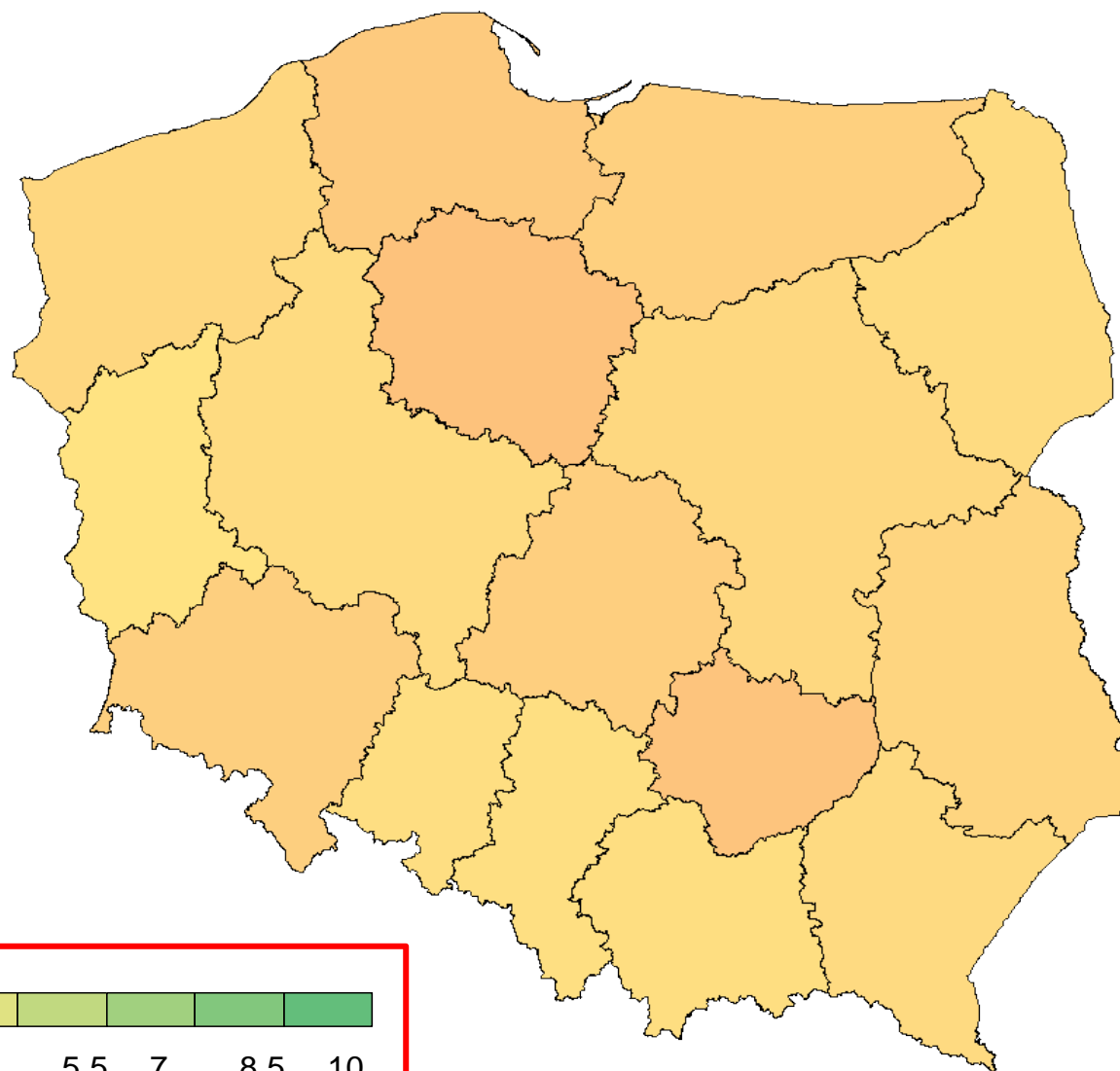


Skala o średniej 100
i odchyleniu standardowym 15.

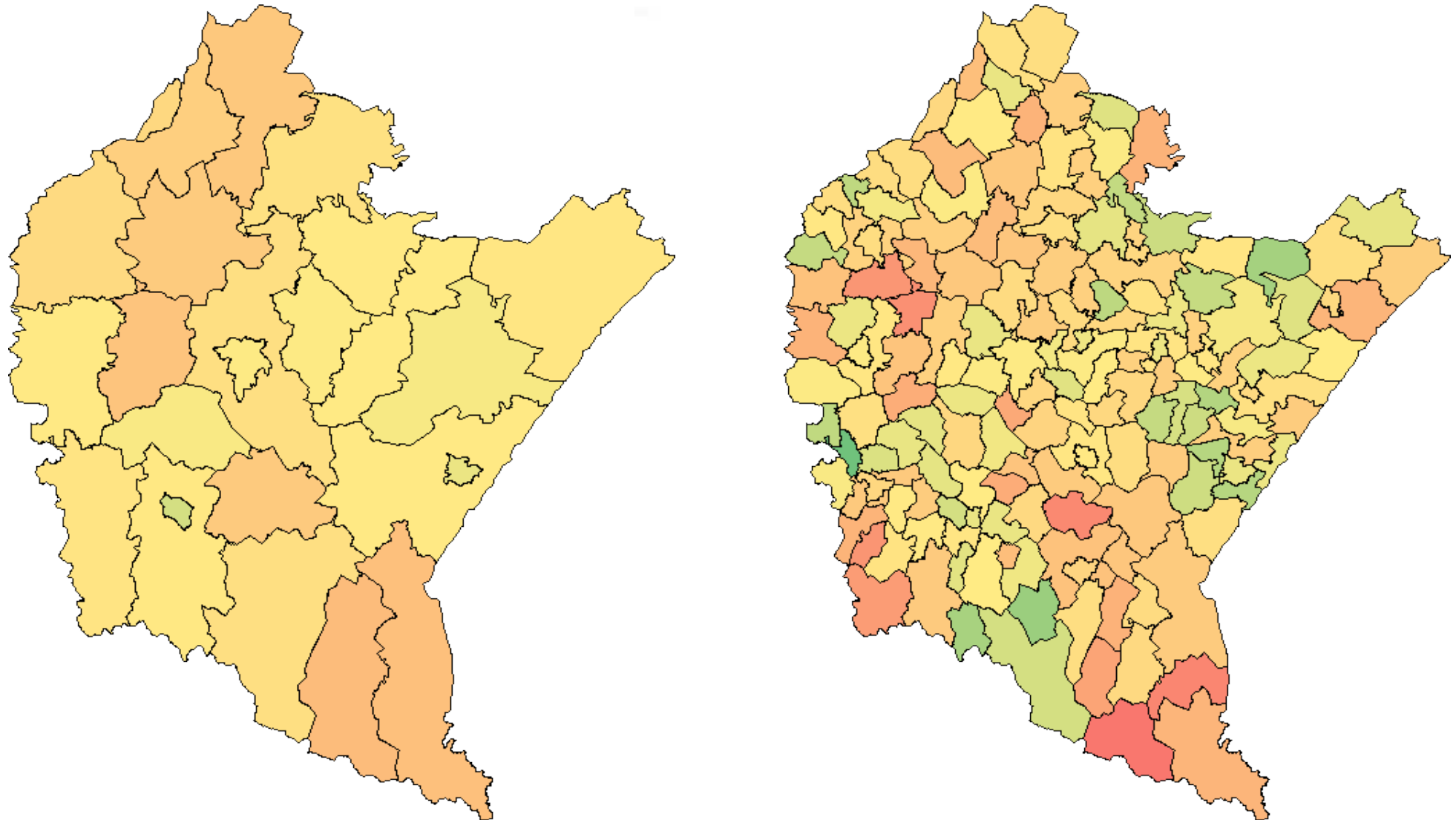




Różnice nie są duże – dobór kolorów jest istotny



Sprawdzian – województwo podkarpackie (różnica pomiędzy 2010 i 2013)



-10 -8 -6 -4 -2 0 2 4 6 8 10



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE



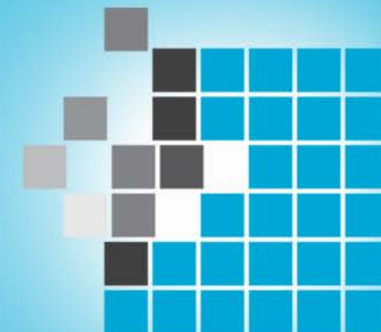
entuzjaści
edukacji

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



KONFERENCJA BADANIA W EDUKACJI

Warszawa, 19-20 października 2015 r.



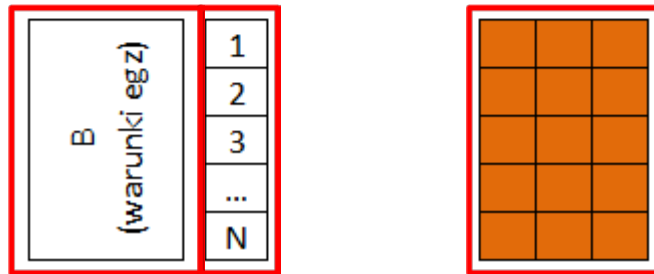
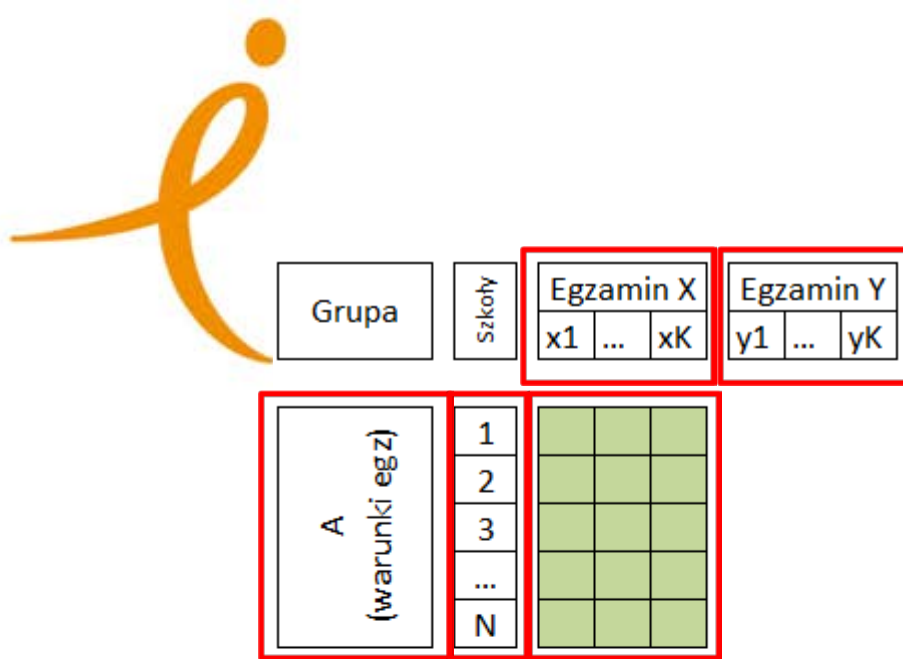
Bartosz Kondratek Rekomendacje dla systemu



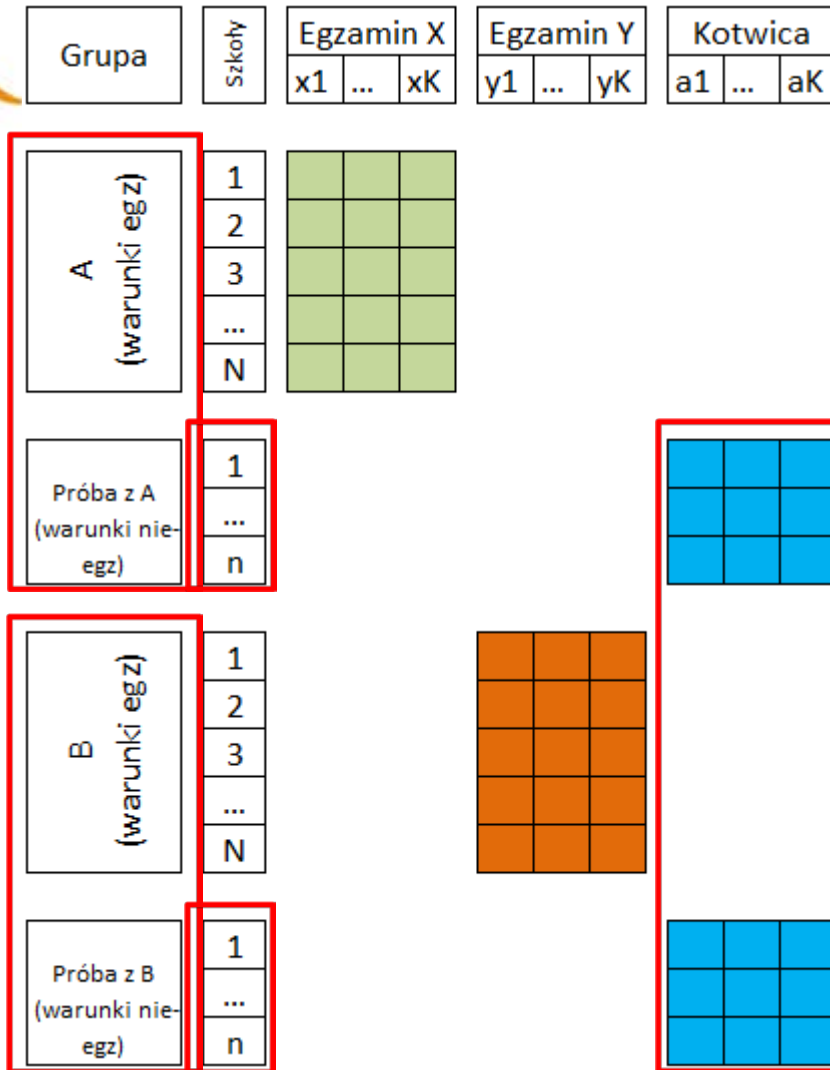
Decyzje

- 1) Schemat zbierania informacji umożliwiającej zapewnienie porównywalności
- 2) Komunikacja wyników zrównanych (skala)

Problem



Dodatkowe badania na reprezentatywnej próbie (C)



Zalety:

- Nie ingeruje w system egzaminacyjny

Wady:

- Niekontrolowany czynnik motywacji
- Niska moc statystyczna
(2x2500 uczniów, $m=0.1 \rightarrow se=0.144$)
- Bardzo wrażliwa na dobór próby oraz zachowanie jednakowych warunków badania
- Możliwe zastosowanie jedynie dla powszechnych części egzaminów



Jednoczesna standaryzacja zadań z wielu edycji egzaminu (B)

Zalety:

- Wyższa moc wnioskowania statystycznego niż C (5000 uczniów, $m=0.1 \rightarrow se=0.038$)
- Czynniki reprezentatywności nie jest tak istotny
- Wymaga jedynie zmian w organizacji badań standaryzacyjnych
- Ma zastosowanie do egzaminów rozwiązywanych przez podpopulacje uczniów (poprawki, poziomy rozszerzone)

Wady:

- Niekontrolowany czynnik motywacji
- Niebezpieczeństwo ujawnienia zadań (przed egzaminem) lub efekt pamięci (po egzaminie)
- Efekty kontekstualne oraz kolejności
- Wymaga niewprowadzania zmian w zadaniach po statecznej standaryzacji

Grupa	szkoły	Egzamin X			Egzamin Y		
		x1	...	xK	y1	...	yK

A (warunki egz)	s1	■	■	■
	s2	■	■	■
	...	■	■	■
	...	■	■	■
	sN	■	■	■

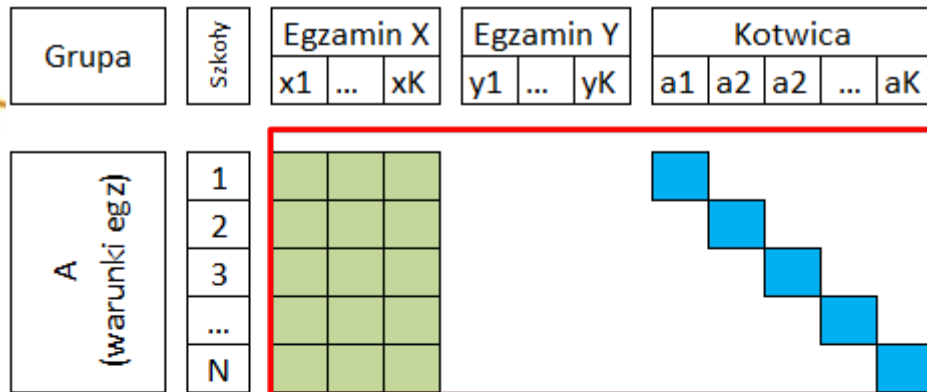
Próba uczniów (warunki nie-egz)	wewnątrz szkół	■	■	■
		■	■	■

B (warunki egz)	1
	2
	3
	...
	N

■	■	■
■	■	■
■	■	■
■	■	■
■	■	■

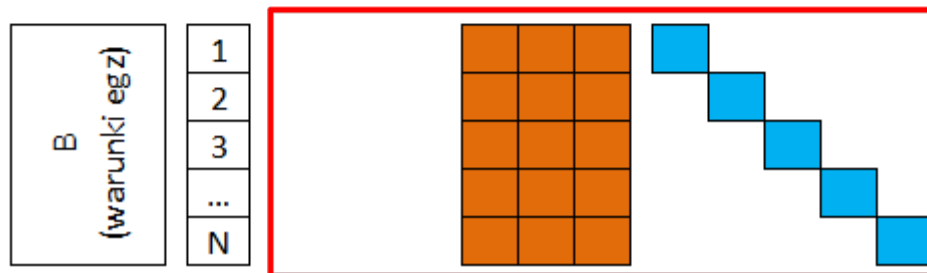


Włączenie do egzaminów niejawnych zadań kotwiczących (A)



Zalety:

- Najwyższa moc statystyczna (200 zadań wiel. wyb., pojedyncze zadanie randomizowane po szkołach, $\mu=0.1 \rightarrow se=0.014$)
- Brak efektów związanych z motywacją/pamięcią itp.



Wady:

- Zadania kotwiczące muszą być niejawne (...)
- Druk, logistyka itp.



Wybór skali do komunikowania wyników

Postulat: komunikujemy wyniki zaokrąglone co części całkowitej

1) Skala wyników surowych

obserwowane -> obserwowane		
X	X->Y	zaokr(X->Y)
...
30	31,536	32
31	32,490	32
32	33,407	33
33	34,333	34
34	35,232	35
35	36,111	36
36	37,000	37
37	37,995	38
38	38,862	39
39	39,725	40
40	40,548	41
41	41,386	41
...

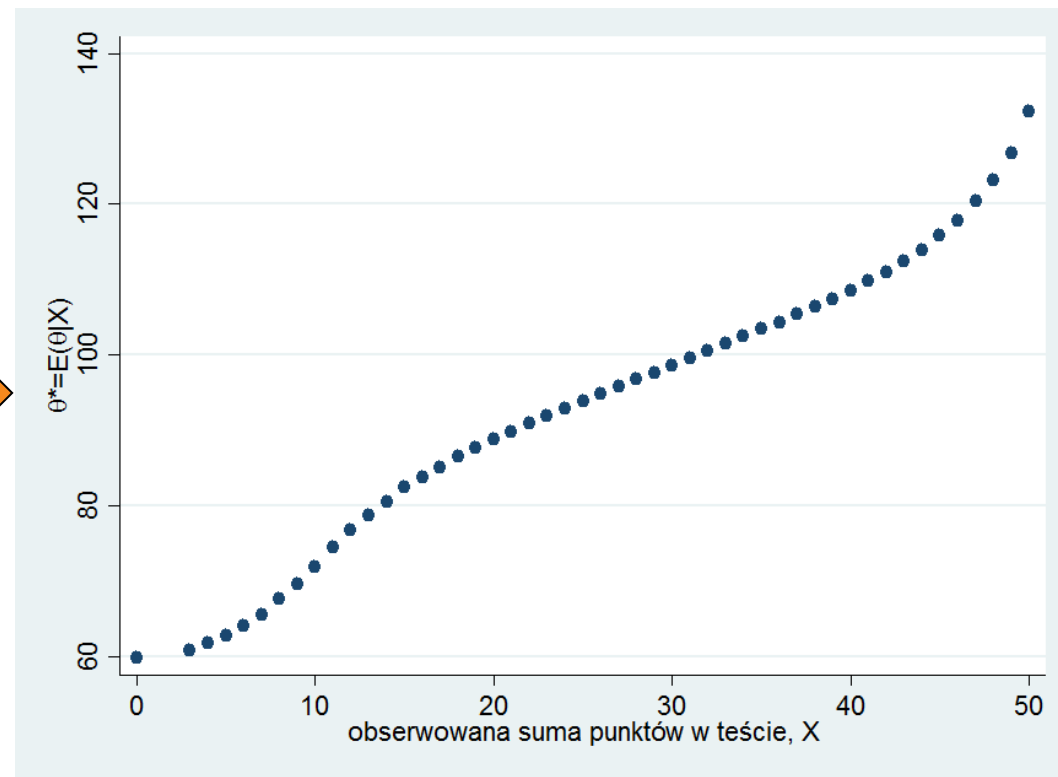
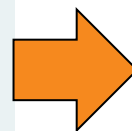
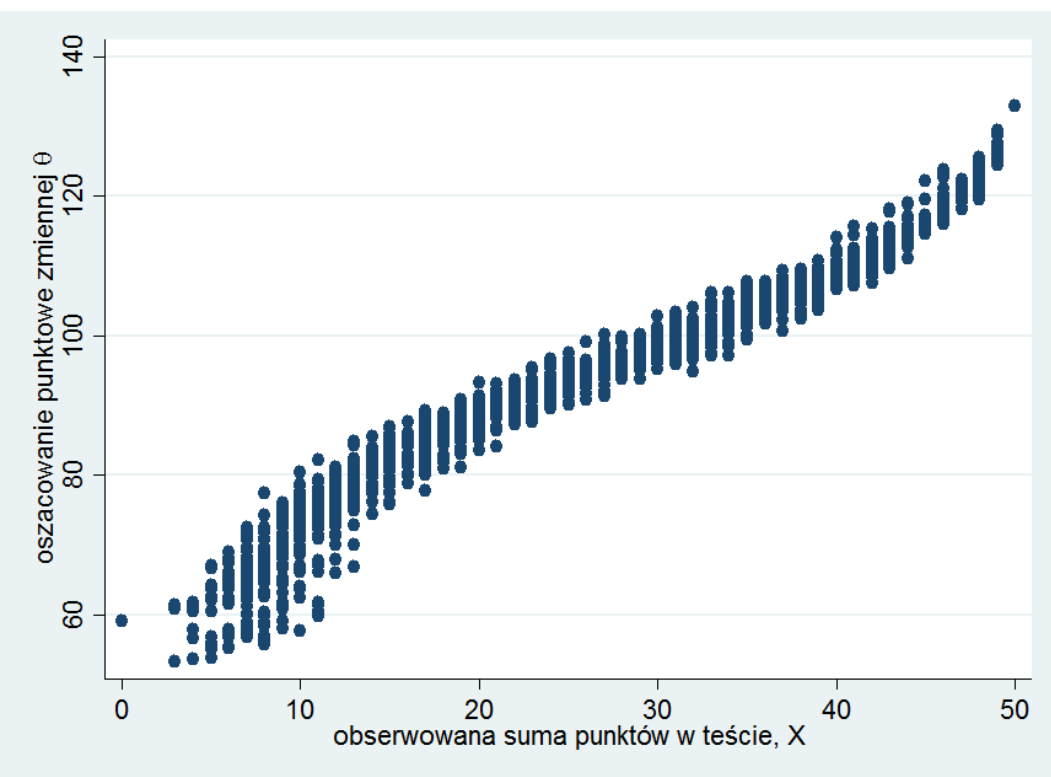
2) Skala zmiennej ukrytej

obserwowane -> zmienna ukryta				
X	m=100;sd=15		m=500;sd=100	
	X-> θ	zaokr(X-> θ)	X-> θ	zaokr(X-> θ)
...
20	90,888	91	439,255	439
21	91,911	92	446,076	446
22	92,919	93	452,794	453
23	93,866	94	459,104	459
24	94,773	95	465,151	465
25	95,694	96	471,294	471
26	96,623	97	477,484	477
27	97,514	98	483,424	483
28	98,398	98	489,321	489
29	99,307	99	495,381	495
30	100,207	100	501,379	501
31	101,098	101	507,322	507
...

(...) obserwowane->zmienna ukryta IRT

Postulat: relacja między zmienną obserwowaną, a ukrytą powinna być 1-1.

- 1) Model Rascha – relacja 1-1 jest właściwością modelu.
- 2) Bardziej skomplikowane modele IRT , a następnie: $\theta^* = E(\theta|X)$





Planowany schemat standaryzacji zadań w CKE

Rok 2017, I standaryzacja zadań z egzaminu S_2019	Arkusze	Bloki	5x bloki A + B + C_kotw	5x bloki A + C + B_kotw	5x bloki B + C + A_kotw	
	Ark_18_1	18_1A				
		18_1B				
		18_1C				
	Ark_18_2	18_2A				
		18_2B				
		18_2C				
	Ark_18_3	18_3A				
		18_3B				
		18_3C				
	Ark_18_4	18_4A				
		18_4B				
		18_4C				
	Ark_18_5	18_5A				
		18_5B				
		18_5C				
	zadania wybrane z lat wcześniejszych	A_kotw				
		B_kotw				
		C_kotw				

statystyka	Błąd przewidywania		
	N/zad=200	N/zad=300	N/zad=350
średnia	0,211	0,168	0,146
sd	0,134	0,117	0,117
rzetelność	0,014	0,013	0,013



DZIĘKUJEMY ZA UWAGĘ