

Szkolny raport ewaluacyjny

Szkoła Podstawowa nr 92 im. Jana Brzechwy

Żoliborz, ul. Przasnyska 18a

Raport opracowany na podstawie wyników
sprawdzianu w szóstej klasie szkoły podstawowej
z lat 2012-2016

Informacja przygotowana w ramach projektu zleconego przez Biuro Edukacji Urzędu m. st. Warszawa

Autorzy:

Bartosz Kondratek

Filip Kulon

Przemysław Majkut

Karolina Świst

Wprowadzenie

Analizując wyniki egzaminów zewnętrznych można pozyskać dużo informacji pomocnych w opisanu pracy szkoły i nauczycieli, stosowanych przez nich metod dydaktycznych oraz osiągnięć uczniów w nauce. Obecnie dostępne są w Polsce różne wskaźniki, oparte o wyniki uczniów, które pozwalają na lepsze zrozumienie procesu dydaktycznego prowadzonego w szkole, a przez to na lepszą pracę z uczniami.

Raport, który Państwo macie przed sobą, zawiera informacje i wskaźniki oparte o wyniki sprawdzianu uczniów Państwa szkoły. W raporcie nie zostały uwzględnione wyniki z egzaminu z języka obcego. Do obliczenia wyników zastosowano podejście IRT (*item response theory*), dzięki czemu uzyskano wskaźniki umiejętności umożliwiające przeprowadzenie analiz niedostępnych przy wykorzystaniu klasycznych, surowych wyników egzaminacyjnych. Tym samym podejściem posłużono się przygotowując wskaźniki Edukacyjnej Wartości Dodanej dla gimnazjów i szkół ponadgimnazjalnych (www.ewd.edu.pl).

Raport przedstawia rozkłady wyników uzyskanych w szkole, zmiany wyników na przestrzeni lat, a także wyniki szkoły na tle innych warszawskich szkół dla których organem prowadzącym jest m.st. Warszawa¹. Nowością w stosunku do innych raportów tego typu jest informacja o zadaniach, z którymi uczniowie Państwa szkoły radzili sobie najlepiej i najgorzej. Rozwiązywalność każdego zadania w Państwa szkole została porównana z rozwiązywalnością tego zadania przez innych uczniów o zbliżonym poziomie umiejętności. Dzięki takiemu zabiegowi możliwe jest wskazanie zadań egzaminacyjnych, z którymi Państwa uczniowie radzą sobie szczególnie dobrze oraz zadań, z którymi radzą sobie gorzej, niż można oczekiwać na podstawie ich wyniku z całego egzaminu. Taka informacja może być bardzo cenna w analizie procesu dydaktycznego w szkole.

Na końcu zamieszczono „Słownik” pojęć i terminów używanych w raporcie.

W raporcie przedstawiono szereg wskaźników opartych o wyniki egzaminu zewnętrznego. Mogą one pomóc wskazać obszary pracy szkoły, które wymagają udoskonalenia. Ważne jest, by pamiętając o kontekście, w którym szkoła funkcjonuje, w czasie analizy starać się zwracać uwagę na obszary zależne od dyrektora i nauczycieli.

¹ W porównaniu pominięte zostały szkoły dla dorosłych oraz szkoły specjalne. W dalszej części raportu szkoły dla których organem prowadzącym jest m.st. Warszawa określane są jako „szkoły samorządowe”. Wyniki dotyczą jedynie uczniów piszących egzamin standardowy (w tym uczniów z dysleksją) i nie obejmują wyników z arkuszy dostosowanych.

Średnie wyniki z egzaminu

Wstęp – o tym jak obliczono wyniki

Zanim przejdziemy do omówienia wyników Państwa szkoły warto poczynić kilka uwag o tym, w jaki sposób zostały obliczone oraz o skali, na jakiej są przedstawione.

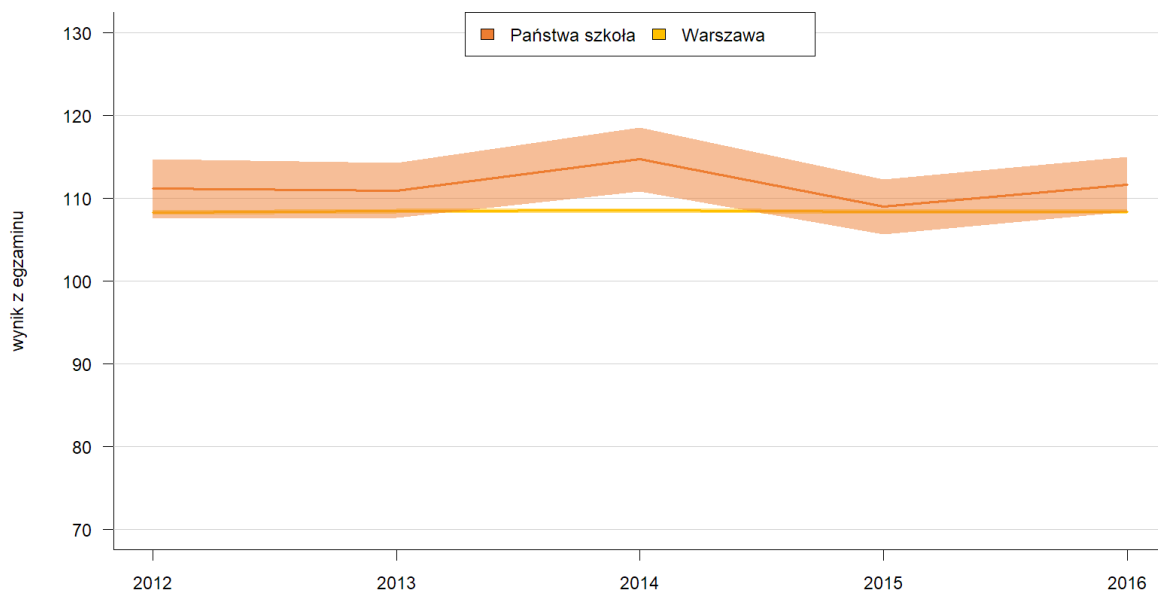
Prezentowane w raporcie ewaluacyjnym wyniki egzaminacyjne zostały obliczone w oparciu o dopasowany do danych egzaminacyjnych model IRT. Jedną z najistotniejszych zalet modelowania IRT jest możliwość oszacowania niepewności związanej z pomiarem uczniowskich umiejętności, zwanego błędem pomiaru. W trakcie egzaminu pojedynczy uczeń rozwiązuje niewielką próbkę zadań z obszaru umiejętności, którego dotyczy egzamin, w związku z czym uzyskany przez niego wynik stanowi jedynie pewną poszlakę do wnioskowania o jego faktycznym poziomie umiejętności. Ponadto, specyficzną właściwością modelowania IRT jest uwzględnianie właściwości każdego zadania, podczas szacowania poziomu umiejętności ucznia. Przykładowo uczeń, który odpowie na 20 najłatwiejszych zadań w egzaminie uzyska przy klasycznym podejściu ten sam wynik co uczeń, który odpowie w tym samym egzaminie poprawnie na 20 zadań najtrudniejszych – przy zastosowanym modelu IRT drugi uczeń uzyskuje wynik wyższy. Uwzględnienie błędu pomiaru oraz „ważenie” wyników poszczególnych zadań ich właściwościami pomiarowymi skutkuje tym, że rozkłady umiejętności uczniów w kraju, mieście, czy nawet szkole mają kształt zbliżony do rozkładu normalnego, co bardzo ułatwia interpretację wyników.

Dla każdego roku przedstawione w raporcie wyniki egzaminu znajdują się na takiej samej skali. Skala dla wszystkich uczniów w kraju ma rozkład w przybliżeniu normalny, o średniej wynoszącej 100 punktów oraz o odchyleniu standardowym równym 15 punktów (tj. ok. 2/3 uczniów uzyskało wyniki od 85 do 115 punktów). Zastosowanie wspólnej skali do każdego roku pozwala na dokonywanie porównań między latami, np. czy uczniowie uzyskują w jednym roku relatywnie wyższe wyniki niż w drugim. Poprzez „relatywność” rozumiemy tu odniesienie do innych uczniów piszących egzamin w danym roku. Porównywanie „surowych” wyników egzaminu (w punktach czy procentach) pomiędzy latami nie jest uprawnione i może prowadzić do błędnych wniosków.

Zmiana wyników w latach 2012-2016

Na wykresach przedstawiono średnie wyniki z lat 2012-2016 wraz z 95% przedziałami ufności, za pomocą linii otoczonej przez jaśniejszy obszar. Przedział ufności jest konieczny, aby uwzględnić niepewność pomiaru poziomu umiejętności uczniów i oznacza, że z 95% prawdopodobieństwem średni wynik szkoły znajduje się w zaznaczonym obszarze.

Jeśli przedziały ufności dla wyników Państwa szkoły i średnich wyników dla całej Warszawy w jakimś roku mają części wspólne (zachodzą na siebie na wykresie) oznacza to, że w tym roku uczniowie Państwa szkoły nie uzyskali wyników różniących się statystycznie od średnich wyników wszystkich uczniów warszawskich szkół samorządowych. W sytuacji, kiedy przedziały ufności nie mają części wspólnych można z 95% prawdopodobieństwem stwierdzić, że wyniki Państwa szkoły są istotnie różne od średniego wyniku wszystkich warszawskich szkół samorządowych. Z uwagi na dużą liczbę uczniów w Warszawie, przedział ufności dla wszystkich uczniów w mieście może być bardzo mały i na wykresach tylko nieznacznie się różnić od średniej.



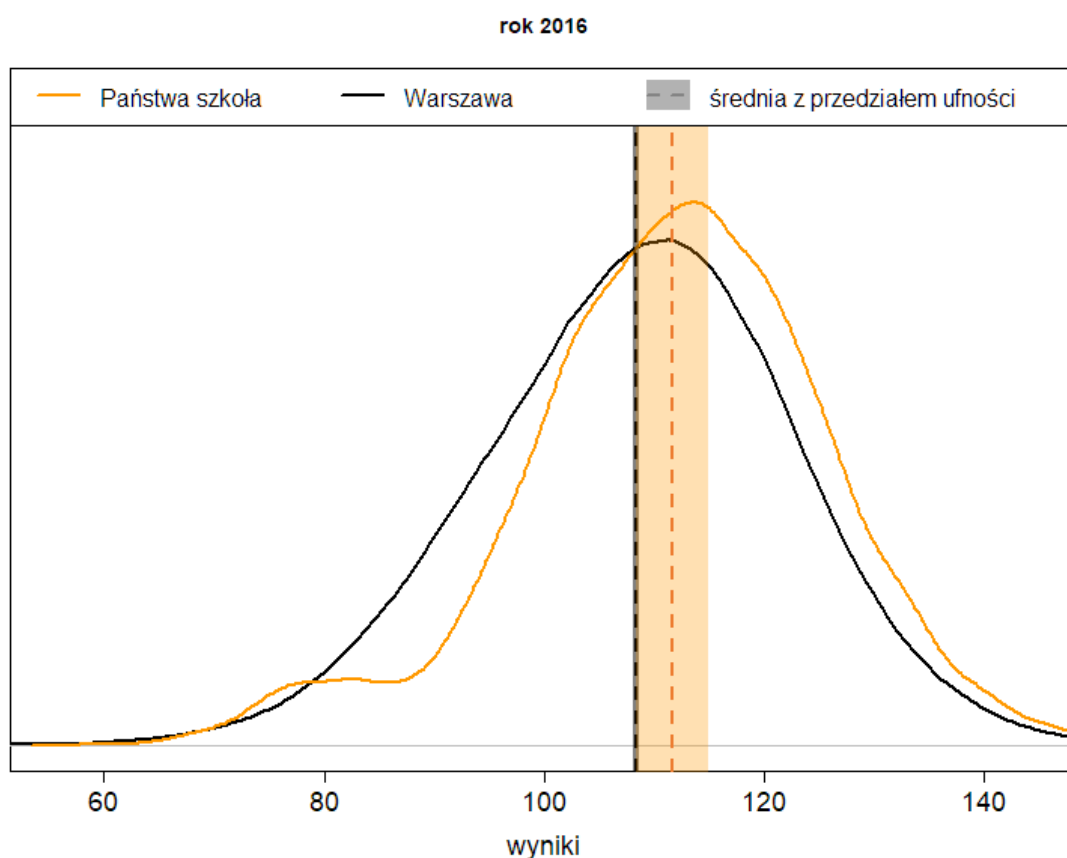
W analizowanym okresie wyniki Państwa szkoły ze sprawdzianu są porównywalne z wynikami uzyskiwanymi przez uczniów ze szkół samorządowych w Warszawie i nie uległy znacznej zmianie, choć pojawiały się ich okresowe wahania.

W tabeli poniżej przedstawiono średni wynik wraz z przedziałem ufności oraz odchylenie standardowe wyników w Państwa szkole oraz w Warszawie w latach 2012-2016. W komórkach minusem (-) oznaczono najniższe, a plusem (+) najwyższe średnie wyniki w rozpatrywanym okresie.

rok		2012	2013	2014	2015	2016
Państwa szkoła	średnia	111,1	110,9	114,7 (+)	109 (-)	111,7
	przedział ufności	(107,6; 114,7)	(107,6; 114,3)	(110,8; 118,6)	(105,6; 112,3)	(108,3; 115)
	odchylenie standardowe	13,4	13,3	13,5	14,6	14
Warszawa	średnia	108,3	108,5	108,5	108,4	108,4
	przedział ufności	(108; 108,6)	(108,2; 108,8)	(108,2; 108,8)	(108,1; 108,6)	(108,1; 108,6)
	odchylenie standardowe	14,2	14,3	14,2	14,6	14,2

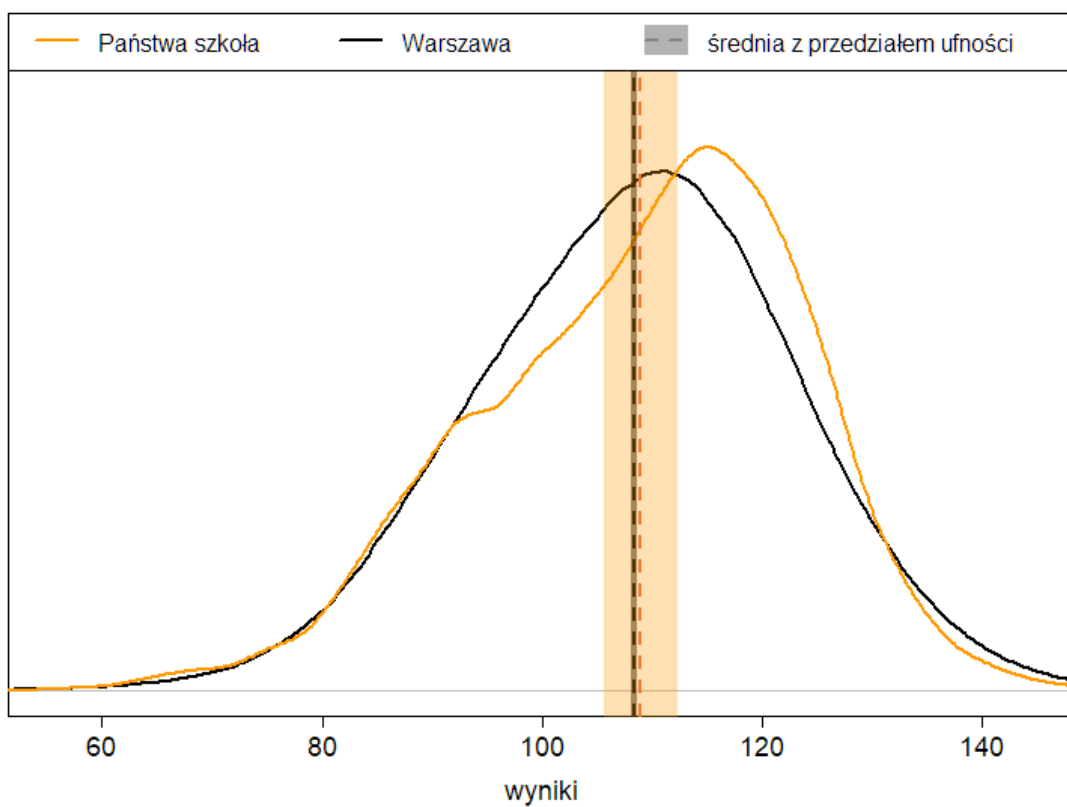
Rozkład wyników w szkole

Wyniki osiągnięte przez uczniów różnią się między sobą, czego nie obrazuje średnia obliczona dla wszystkich uczniów w szkole. Rozkład wyników w szkole został przedstawiony na poniższym wykresie. Widnieją na nim dwie krzywe, jedna pokazuje rozkład wyników uczniów Państwa szkoły, a druga uczniów ze wszystkich szkół samorządowych w Warszawie w 2016 roku. W miejscach, gdzie krzywa jest najwyższa koncentrowały się wyniki osiągnięte przez uczniów. Wypiętrzona krzywa oznacza, że uczniowie osiągnęli bardziej podobne do siebie wyniki, natomiast spłaszczona oznacza duże zróżnicowanie uczniowskich wyników. Jest to związane z miarą zwaną odchyleniem standardowym – im odchylenie standardowe większe, tym bardziej spłaszczona krzywa (a im mniejsze, tym bardziej wypiętrzona). Na wykresie zaznaczono także średni wynik w Państwa szkole i warszawskich szkół samorządowych wraz z 95% przedziałami ufności w postaci zacienionych obszarów. Jeśli przedziały ufności nie zachodzą na siebie, to wyniki Państwa szkoły i wszystkich szkół samorządowych w Warszawie różnią się w istotny statystycznie sposób.

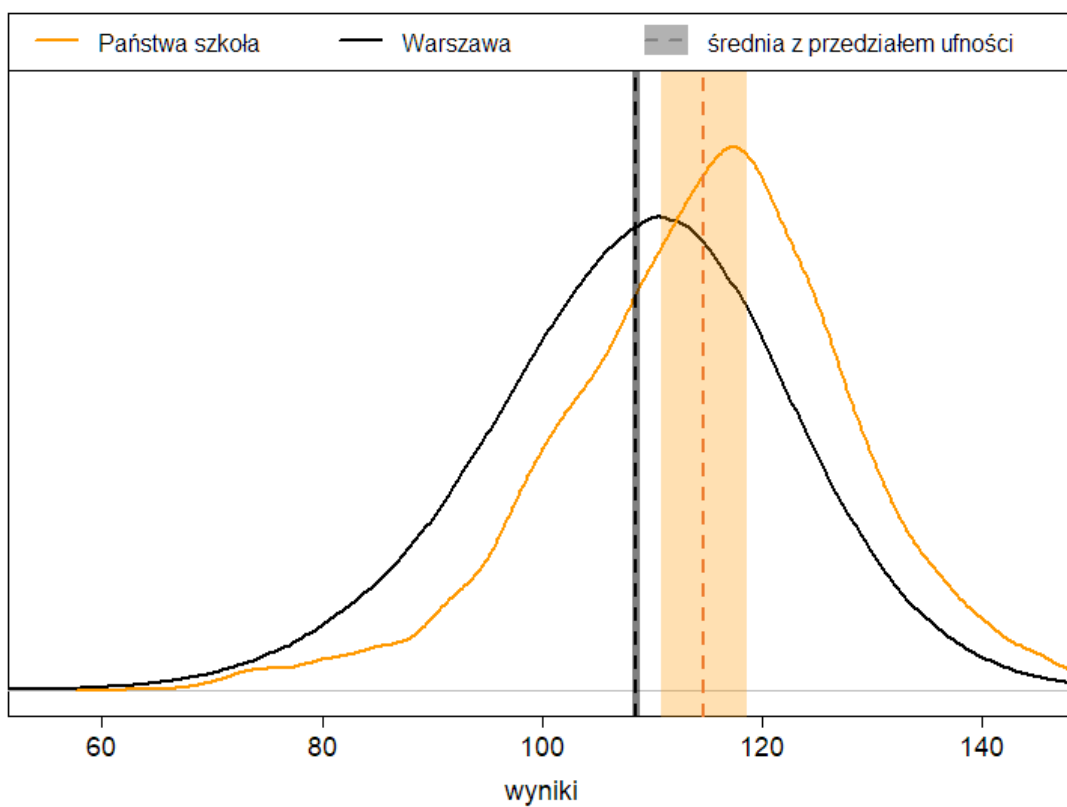


W roku 2016 wyniki Państwa szkoły ze sprawdzianu są porównywalne z wynikami uzyskiwanymi przez uczniów warszawskich szkół samorządowych (ich przedziały ufności nie są rozłączne), a odchylenie standardowe wyników uczniów Państwa szkoły nie różni się znacząco od średniego odchylenia w warszawskich szkołach samorządowych. Poniżej umieszczono analogiczne wykresy dla pozostałych lat z analizowanego okresu. Wartość średniej, przedziałów ufności i odchylenia standardowego przedstawiona została tabelarycznie w poprzedniej części raportu (Zmiana wyników w latach 2012-2016).

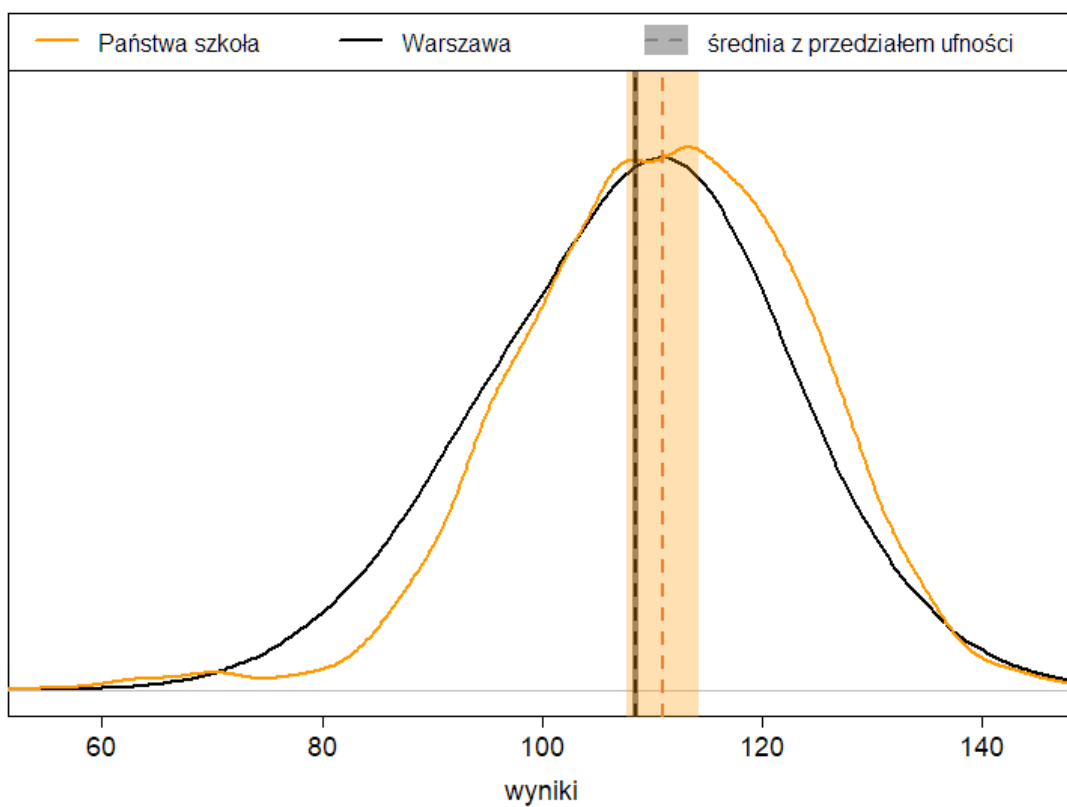
rok 2015



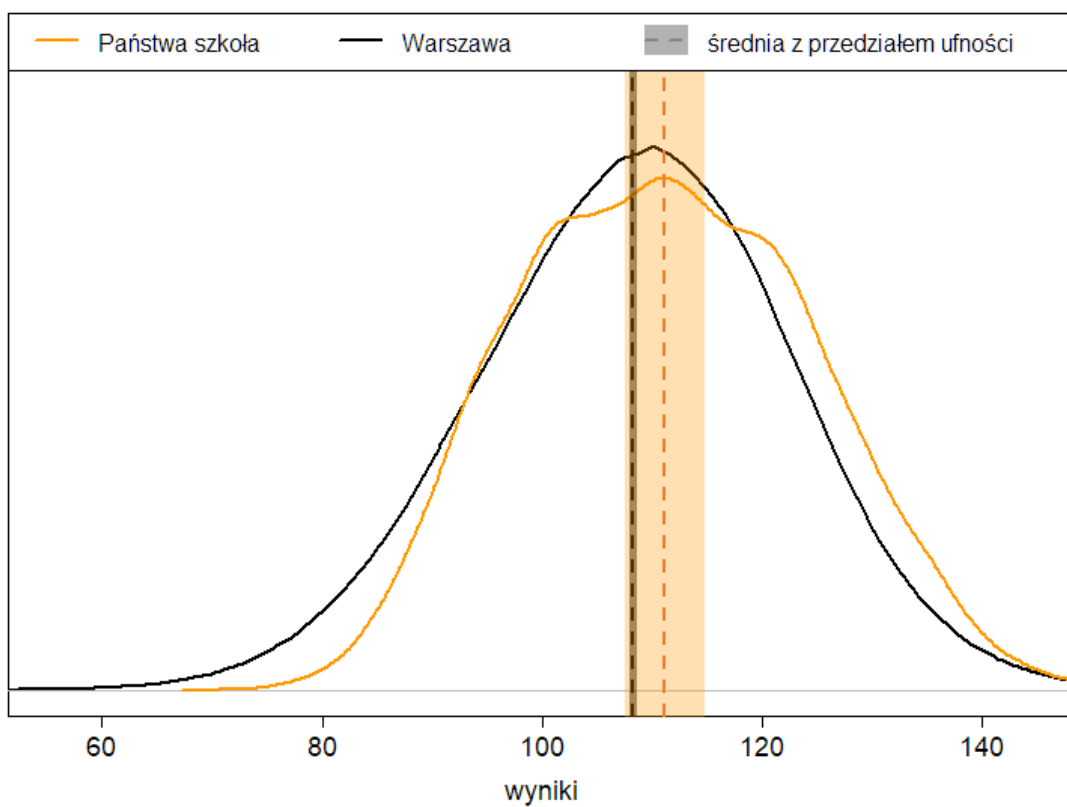
rok 2014



rok 2013

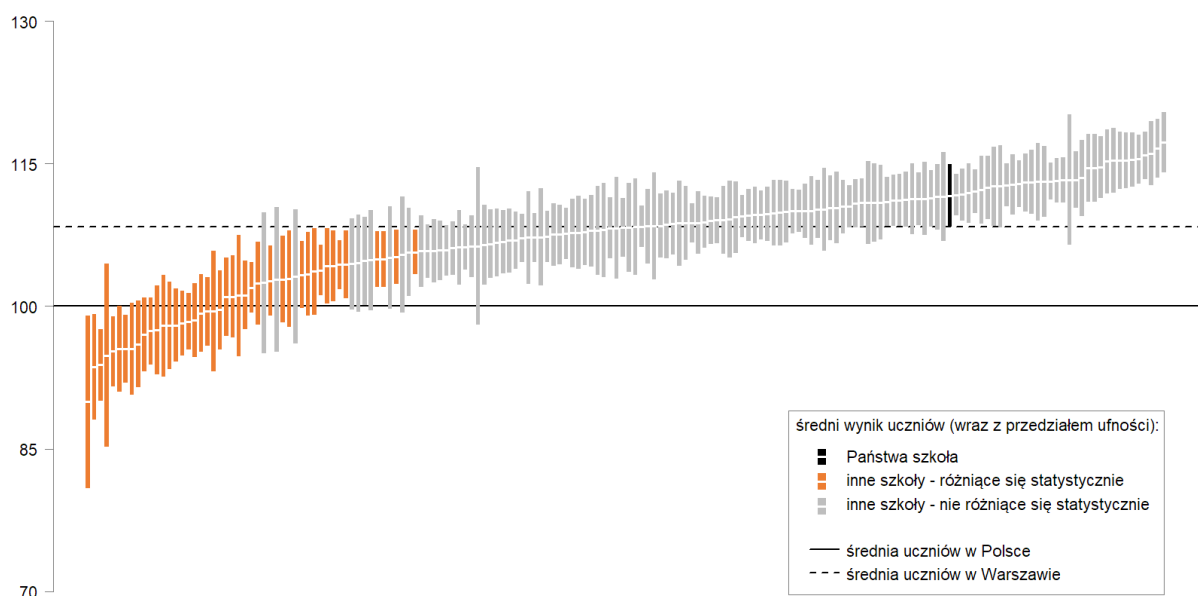


rok 2012



Wyniki szkoły na tle innych warszawskich szkół

Średnie wyniki Państwa szkoły uzyskane na egzaminie z roku 2016 przedstawiono na tle innych warszawskich szkół na wykresie poniżej. Na wykresie szkoły są uporządkowane w kolejności od najniższych do najwyższych średnich wyników na egzaminie. Średnim wynikiom towarzyszy informacja o 95% przedziale ufności wokół średnich, dzięki czemu możliwe jest zidentyfikowanie szkół, których wyniki w sposób istotny statystycznie różnią się od Państwa szkoły (wyróżniono je kolorem pomarańczowym). Długość przedziału ufności jest związana głównie z liczbą uczniów w szkole – dla małych szkół przedział ufności jest większy niż dla dużych szkół.



Dla pozostałych lat pozycja wyników Państwa szkoły na tle innych szkół została przedstawiona w tabeli, w postaci informacji o centylu zajmowanym przez szkołę na tle innych szkół samorządowych w Warszawie oraz wszystkich szkół w Polsce. Wartość centyla wskazuje, że szkoła ma takie same lub wyższe wyniki niż dany procent szkół, na który wskazuje centyl. Przykładowo, wynik Państwa szkoły wśród wszystkich szkół w Polsce w 2016 roku w 97 centylu oznacza, że 96% szkół biorących udział w egzaminie w całym kraju uzyskało wynik niższy od Państwa szkoły, zaś wynik w 81 centylu warszawskich szkół samorządowych oznacza, że w tym roku niższy wynik od Państwa szkoły uzyskało 80% warszawskich szkół samorządowych.

rok	2012	2013	2014	2015	2016
centyl w grupie warszawskich szkół samorządowych	72	70	93	55	81
centyl wśród wszystkich szkół w Polsce	97	96	99	94	97

Informacje o zadaniach, z którymi uczniowie radzili sobie najlepiej i najgorzej

Rozwiązywalność² każdego zadania w Państwa szkole została porównana z rozwiązywalnością tego zadania przez innych uczniów o zbliżonym poziomie umiejętności. Taka analiza, niezależnie od średniego wyniku Państwa szkoły, zawsze wskaże grupę zadań, w której uczniowie poradzili sobie trochę lepiej niż uczniowie o zbliżonych wynikach z egzaminu w całym kraju, oraz trochę gorzej niż uczniowie o zbliżonych wynikach z egzaminu w całym kraju. Wyniki dla poszczególnych lat zebrano na wykresach i w tabelach poniżej.

Dodatnia różnica między tymi dwoma rozwiązywalnościami wskazuje na zadania, z którymi uczniowie uczęszczający do Państwa szkoły poradzili sobie lepiej niż uczniowie z innych szkół w Polsce o zbliżonych wynikach w całym egzaminie. Pięć zadań o największych dodatnich różnicach zaznaczono kolorem zielonym. Analogicznie – ujemne wartości, wskazują na zadania, z którymi uczniowie Państwa szkoły poradzili sobie gorzej niż można by się spodziewać uwzględniając wynik z całego egzaminu. Pięć zadań o największych ujemnych różnicach zaznaczono kolorem czerwonym.

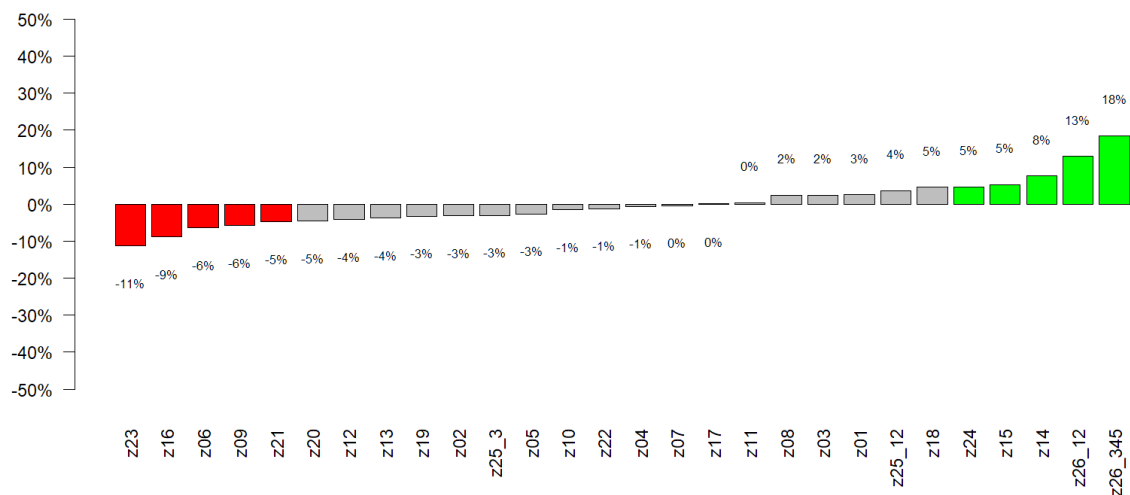
Analiza treści zaznaczonych zadań może pozwolić zidentyfikować specyficzne obszary umiejętności uczniów, które w Państwa szkole są rozwijane najlepiej, a także takie gdzie uczniowie nie uzyskali poziomu osiągniętego przez uczniów o zbliżonych wynikach w całym egzaminie. Umieszczenie w raporcie treści zadań nie było możliwe. Na wykresach i w tabelach zamieszczono identyfikatory zadań³ – ich treści można znaleźć w arkuszach egzaminacyjnych publikowanych na stronie Centralnej Komisji Egzaminacyjnej (www.cke.edu.pl).⁴

² Przez rozwiązywalność zadania dla zadań ocenianych na skali 0-1 należy rozumieć procent poprawnych odpowiedzi w zadaniu. Dla zadań ocenianych wielopunktowo, rozwiązywalność jest to wynik w zadaniu podzielony przez maksymalną liczbę możliwych do uzyskania w zadaniu punktów, wyrażony na skali procentowej.

³ W sprawdzianie znajdowały się zadania otwarte składające się z wielu kryteriów oceny – kryteria oceny zadań są podane po podkreślniku umieszczonym za numerem zadania. Przykładowo, z25_1 oznacza pierwsze kryterium oceny zadania o numerze 25 w arkuszu, natomiast z25_23 oznacza rozwiązywalność dla sumy kryteriów drugiego oraz trzeciego w zadaniu 25 (zastosowana metoda statystyczna wymagała połączenia niektórych kryteriów oceny).

⁴ W czasie tworzenia raportu arkusze z 2016 i 2015 dostępne były pod adresem www.cke.edu.pl/sprawdzian/arkusze/. Arkusze z 2014 roku i starsze dostępne były pod adresem www.cke.edu.pl/sprawdzian/archiwum/arkusze-archiwum/.

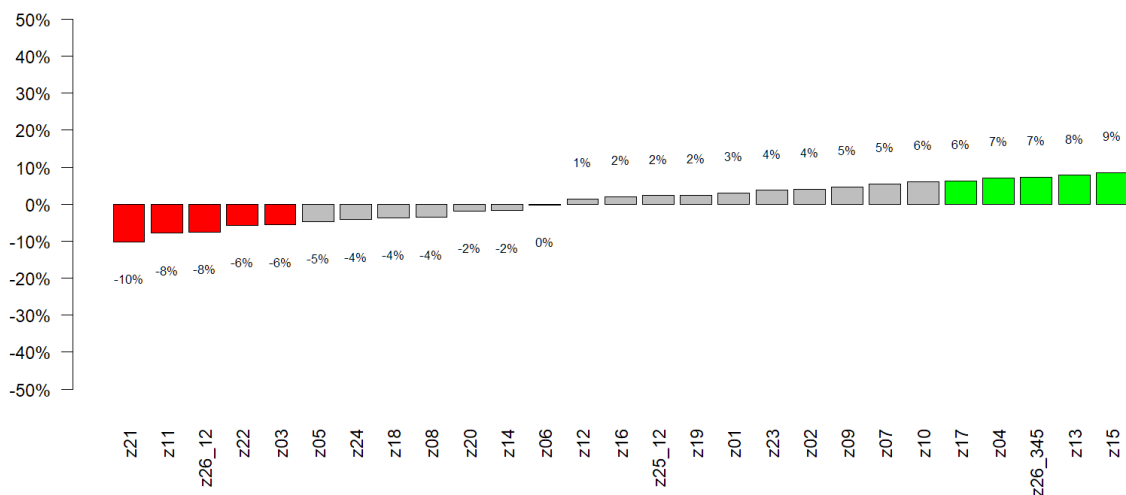
Różnice w rozwiązywalności zadań w roku 2012



Zadania o największych różnicach w rozwiązywalności w roku 2012

Numer zadania	Rozwiązywalność		Różnica w rozwiązywalności
	Państwa szkoła	Uczniowie o podobnych wynikach w egzaminie	
z26_345	77%	58%	18%
z26_12	74%	61%	13%
z14	48%	40%	8%
z15	63%	58%	5%
z24	37%	32%	5%
z21	67%	72%	-5%
z09	73%	79%	-6%
z06	30%	37%	-6%
z16	59%	68%	-9%
z23	40%	52%	-11%

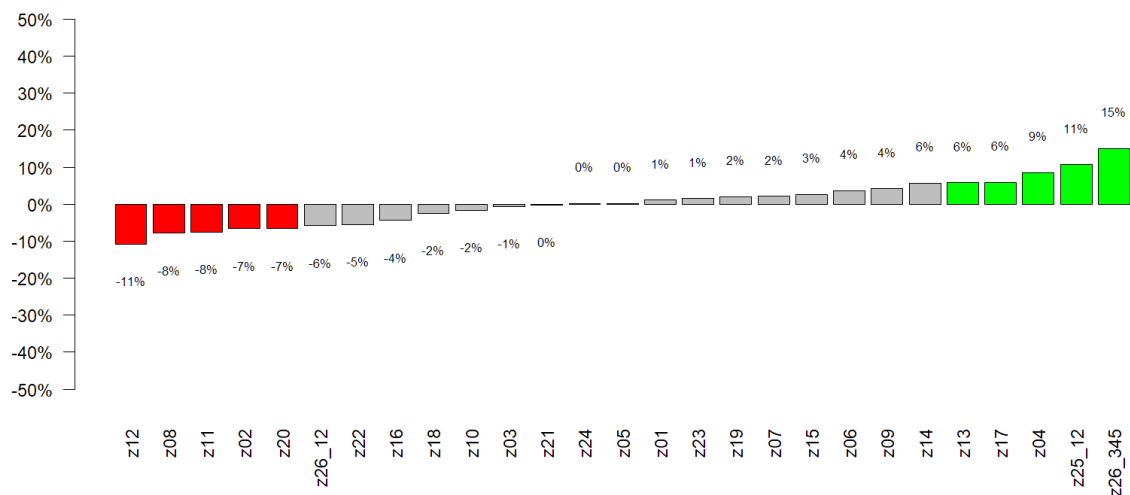
Różnice w rozwiązywalności zadań w roku 2013



Zadania o największych różnicach w rozwiązywalności w roku 2013

Numer zadania	Rozwiązywalność		Różnica w rozwiązywalności
	Państwa szkoła	Uczniowie o podobnych wynikach w egzaminie	
z15	73%	65%	9%
z13	73%	65%	8%
z26_345	73%	66%	7%
z04	96%	89%	7%
z17	63%	57%	6%
z03	83%	89%	-6%
z22	62%	67%	-6%
z26_12	68%	76%	-8%
z11	58%	66%	-8%
z21	68%	78%	-10%

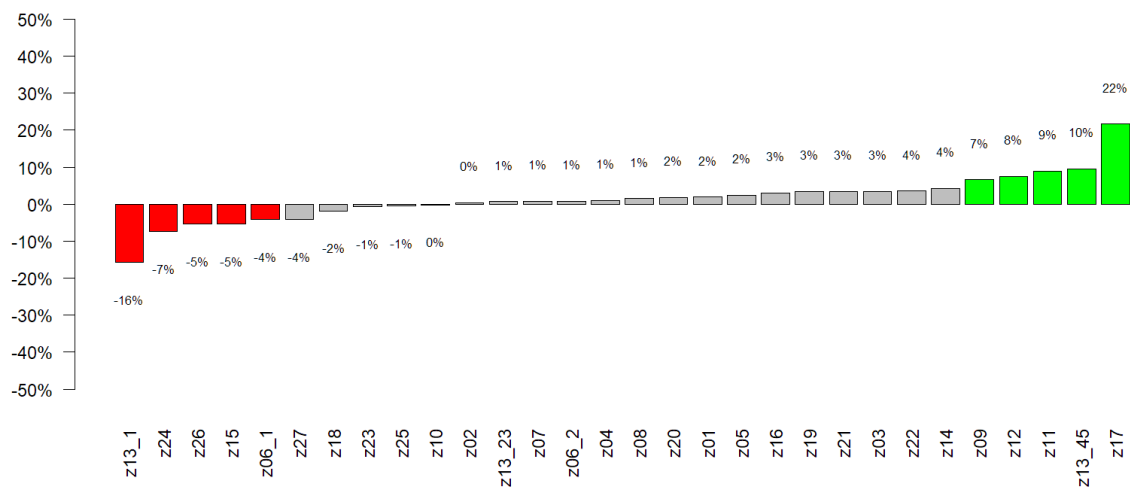
Różnice w rozwiązywalności zadań w roku 2014



Zadania o największych różnicach w rozwiązywalności w roku 2014

Numer zadania	Rozwiązywalność		Różnica w rozwiązywalności
	Państwa szkoła	Uczniowie o podobnych wynikach w egzaminie	
z26_345	86%	71%	15%
z25_12	88%	77%	11%
z04	89%	81%	9%
z17	86%	80%	6%
z13	96%	91%	6%
z20	55%	62%	-7%
z02	82%	89%	-7%
z11	38%	45%	-8%
z08	88%	95%	-8%
z12	77%	88%	-11%

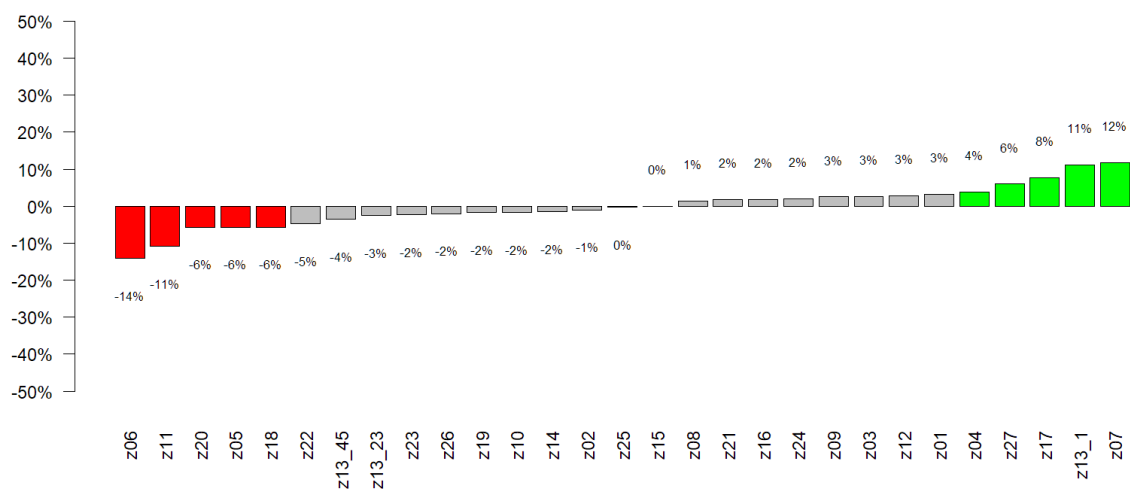
Różnice w rozwiązywalności zadań w roku 2015



Zadania o największych różnicach w rozwiązywalności w roku 2015

Numer zadania	Rozwiązywalność		Różnica w rozwiązywalności
	Państwa szkoła	Uczniowie o podobnych wynikach w egzaminie	
z17	89%	68%	22%
z13_45	69%	59%	10%
z11	77%	68%	9%
z12	78%	71%	8%
z09	82%	75%	7%
z06_1	81%	85%	-4%
z15	42%	48%	-5%
z26	68%	74%	-5%
z24	67%	75%	-7%
z13_1	63%	79%	-16%

Różnice w rozwiązywalności zadań w roku 2016



Zadania o największych różnicach w rozwiązywalności w roku 2016

Numer zadania	Rozwiązywalność		Różnica w rozwiązywalności
	Państwa szkoła	Uczniowie o podobnych wynikach w egzaminie	
z07	80%	68%	12%
z13_1	88%	77%	11%
z17	89%	81%	8%
z27	70%	64%	6%
z04	95%	91%	4%
z18	71%	77%	-6%
z05	82%	88%	-6%
z20	66%	72%	-6%
z11	33%	44%	-11%
z06	59%	74%	-14%

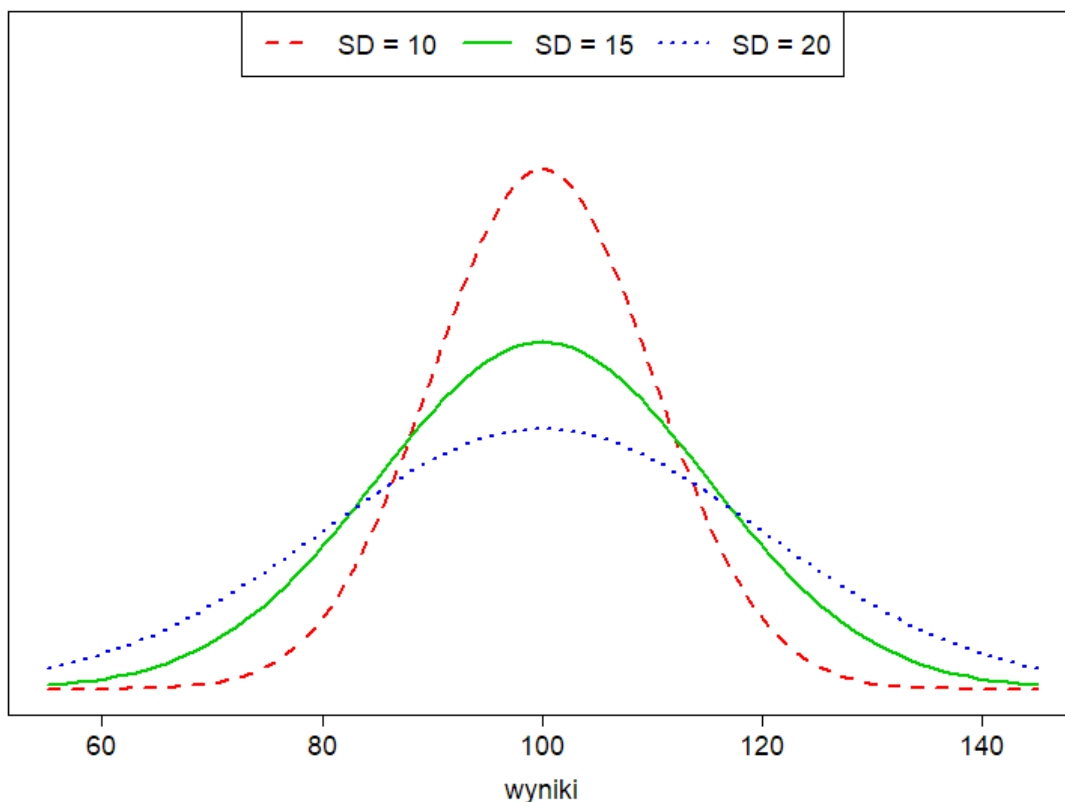
Słownik

Rozkład normalny

Rozkład prawdopodobieństwa obrazujący rozkład częstości występowania wartości dowolnej zmiennej. Posiada dwa parametry – średnią oraz odchylenie standardowe. W uproszczeniu – obrazuje sytuację, w której większość obserwacji ma wyniki bardzo bliski średniej, natomiast im dalej od wartości średniej tym mniej obserwacji uzyskuje dany wynik (jednym z klasycznych przykładów tego rodzaju zależności jest rozkład inteligencji w populacji). Około 68% wyników leży w odległości jednego odchylenia standardowego od średniej, około 95% wyników w okolicy dwóch odchyleń standardowych od średniej, natomiast 99% w okolicy trzech odchyleń standardowych od średniej. Rozkład normalny o średniej 100 i odchyleniu standardowym 15 został przedstawiony na poniższym rysunku (dla hasła „Odchylenie standardowe”) przy pomocy zielonej linii ciągłej. Przerwanymi liniami zaznaczono rozkłady normalne o mniejszym (kolor niebieski) oraz większym (kolor czerwony) skoncentrowaniu wyników wokół wartości średniej.

Odchylenie standardowe

Odchylenie standardowe (SD) jest miarą zróżnicowania wyników. Jego wartość zależy od tego jak bardzo wyniki pojedynczych uczniów różnią się od średniej w analizowanej grupie. Czym mniejsza jest wartość SD, tym bardziej wyniki uczniów są skupione wokół średniej, tym bardziej uczniowie są do siebie podobni pod względem badanych umiejętności. Ze wzrostem SD zróżnicowanie wyników uczniów jest coraz większe, rośnie udział uczniów o najniższych i najwyższych wynikach, a maleje udział uczniów o wynikach przeciętnych. Na wykresie poniżej zilustrowano te zależności ukazując trzy rozkłady normalne dla tej samej średniej (100) oraz trzech różnych odchyleń standardowych (10, 15, 20). Rozkład normalny o średniej 100 oraz odchyleniu standardowym 15 jest układem odniesienia dla uczniowskich wyników w niniejszym raporcie, tj. rozkładem wyników w ogólnopolskiej populacji uczniów zdających egzamin w danym roku.



IRT

Pod nazwą *item response theory* (IRT) kryje się rodzina narzędzi statystycznych wykorzystywanych do modelowania odpowiedzi na rozwiązywane zadania oraz umiejętności uczniów. Dzięki modelom IRT możemy określić zarówno właściwości zadań (np. ich parametry trudności) oraz rozkład poziomu umiejętności uczniów. Innymi słowy, oszacowanie poziomu umiejętności ucznia, który rozwiązał poprawnie 20 najłatwiejszych zadań w egzaminie jest niższe od oszacowania poziomu umiejętności ucznia, który rozwiązał poprawnie 20 zadań najtrudniejszych. IRT pozwala na modelowanie zadań ocenianych dychotomicznie (na skali 0-1 punktów) i politomicznie (np. 0-1-2 punktów). Modele IRT uwzględniają zarówno błąd pomiaru oraz „ważą” wyniki poszczególnych zadań ich właściwościami pomiarowymi, co skutkuje tym, że rozkłady umiejętności uczniów w kraju, mieście, czy nawet szkole mają kształt zbliżony do rozkładu normalnego, co bardzo ułatwia interpretację wyników.

Przedział ufności

Zarówno średni wynik z danego egzaminu w szkole jak i wskaźniki EWD są obarczone dwojakim błędem: po pierwsze, błędem pomiaru na egzaminie zewnętrznym; po drugie, błędem próbkowania (absolwenci danego rocznika są tylko próbą uczniów szkoły, na podstawie której wypowiadamy się o jakości nauczania w całej szkole). To powoduje konieczność wyznaczenia przedziału ufności oszacowanej wartości dla analizowanej grupy uczniów. Przedział ufności, to przedział, który z założonym współczynnikiem ufności zawiera prawdziwą wartość interesującego nas parametru. W raporcie wykorzystywane zostały przedziały ufności z 95% prawdopodobieństwem, tj. dla średniej 100 z przedziałem ufności +/- 7 można wnioskować z 95% prawdopodobieństwem, że średnia znajduje się między 93 a 107.

Przedziały ufności możemy traktować jako regułę decyzyjną służącą do rozstrzygnięcia czy obserwowane wskaźniki różnią się od pewnej wartości (np. średnia w Polsce) lub między sobą w sposób istotny statystycznie (nieprzypadkowo, systematycznie). Przy formułowaniu ocen typu szkoła A uzyskała wyższy średni wynik w zakresie sprawdzanym egzaminem od szkoły B, musimy uwzględnić niepewność w oszacowaniu obu średnich, którą podano za pomocą przedziałów ufności. Jeżeli 95% przedziały ufności dla porównywanych szkół są rozłączne, to przyjmujemy, że mamy wystarczającą pewność do uznania że średnie szkół są różne. Gdy przedziały częściowo pokrywają się, formułowanie oceny staje się bardziej ryzykowne – i nawet, jeżeli jest to nieintuicyjne – należy stwierdzić, że porównywane szkoły nie różnią się od siebie w zakresie danego wskaźnika.

Skala centylowa

Skala pozycyjna wyników uporządkowanych rosnąco. Aby ustalić wartości centyli, porządkujemy wyniki od najniższych do najwyższych, dzielimy ich zbiór na 100 równych pod względem liczby obserwacji części, a następnie określamy położenie granic między tymi częściami na skali wyników i oznaczamy je jako centyle C1, C2, C3, C4, C5,... C99. Liczba centyli wynosi dziewięćdziesiąt dziewięć. Numer kolejny najbliższego centyla stanowi rangę centylową danego wyniku testowania, to jest procent wyników niższych od danego centyla. Np. wynik szkoły w 45 centylu oznacza, że 44% szkół uzyskało wynik niższy od tej szkoły.

Rzeczonym przykładem skali centylowej spoza obszaru edukacji są siatki centylowe wzrostu/wagi dzieci.

Zróznicowane funkcjonowanie zadań (*differential item functioning, DIF*)

Ze zróznicowanym funkcjonowaniem zadania mamy do czynienia, gdy osoby o tym samym poziomie mierzonej przez test umiejętności, ale należące do różnych grup (np. chłopcy i dziewczęta, badana szkoła i reszta uczniów w kraju), charakteryzują się różnym rozkładem odpowiedzi na to zadanie. Ustalenie czy w przypadku danego zadania występuje DIF, wymaga analizy międzygrupowych różnic w rozkładzie odpowiedzi na to zadanie, przy kontroli poziomu umiejętności osób rozwiązujących test.

Przedstawione w raporcie wyniki ograniczają się jedynie do analizy wielkości różnicy w rozwiązywalności zadań, gdyż jest to najistotniejszy aspekt oceny efektu DIF. Odpowiadamy

na pytanie na ile zadanie byłoby łatwiejsze, lub trudniejsze, w analizowanej szkole, gdyby uczniowie w tej szkole rozwiązywali to zadanie w zależności od swojego poziomu umiejętności zgodnie z tym jak zadanie funkcjonuje w grupie wszystkich uczniów w kraju. Raportowane efekty DIF zostały wyliczone w trzech krokach: (1) skonstruowano nieparametryczne funkcje opisujące prawdopodobieństwo uzyskania różnych wyników w zadaniu w zależności od poziomu umiejętności uczniów w grupie ogólnopolskiej; (2) wyliczono rozwiązywalność zadań w szkole, jakiej spodziewalibyśmy się, gdyby zadania funkcjonowały w szkole w taki sam sposób, jak w całym kraju, korzystając z oszacowań umiejętności uczniów w szkole oraz funkcji wyznaczonych w kroku 1; (3) od faktycznej rozwiązywalności zadania odjęto rozwiązywalność zadania obliczoną w kroku 2.