

Przewodnik po raporcie

Łomża 2010

Spis treści

1. System egzaminów zewnętrznych	3
1.1. Sprawdzian	4
1.2. Egzamin gimnazjalny	5
1.3. Egzamin maturalny	6
1.4. Egzamin zawodowy	7
1.5. Uczniowie o specjalnych potrzebach edukacyjnych	8
2. Podstawowe pojęcia statystyczne	9
2.1. Podstawowe wskaźniki statystyczne	9
2.2. Rozkłady wyników	11
2.3. Przykłady skal pomiarowych	13
2.4. Edukacyjna wartość dodana	18

1. SYSTEM EGZAMINÓW ZEWNĘTRZNYCH

(opracowano na podstawie: www.cke.edu.pl)

Ocenianiem zewnętrznym zajmują się instytucje zewnętrzne wobec szkoły – Centralna Komisja Egzaminacyjna i okręgowe komisje egzaminacyjne.

Ocenianie zewnętrzne pozwala na zdiagnozowanie osiągnięć edukacyjnych uczniów, ocenę jakości oddziaływań edukacyjnych szkół oraz porównywalność zaświadczeń, świadectw i dyplomów bez względu na miejsce ich wystawienia. Istotą oceniania zewnętrznego jest to, że wszyscy zdający rozwiązują te same zadania przygotowane przez komisje egzaminacyjne. Podstawę tych zadań stanowią wymagania egzaminacyjne określone i ogłoszone przez ministra właściwego do spraw oświaty i wychowania. Odpowiedzi zdających sprawdzają i oceniają według jednolitych w całym kraju kryteriów specjalnie przygotowani zewnętrzni egzaminatorzy (wpisani do ewidencji prowadzonych przez okręgowe komisje egzaminacyjne).

Zasady oceniania zewnętrznego określa Rozporządzenie Ministra Edukacji Narodowej z dnia 30 kwietnia 2007 r. w sprawie warunków i sposobu oceniania, klasyfikowania i promowania uczniów i słuchaczy oraz przeprowadzania sprawdzianów i egzaminów w szkołach publicznych (DzU nr 83, poz. 562, z późn. zm.).

Ocenianie zewnętrzne obejmuje:

- sprawdzian w VI klasie szkoły podstawowej
- egzamin gimnazjalny w III klasie gimnazjum
- egzamin zawodowy potwierdzający kwalifikacje zawodowe
- egzamin maturalny
- egzaminy eksternistyczne.

Podstawę do przeprowadzenia egzaminów zewnętrznych stanowią standardy wymagań egzaminacyjnych zawarte w:

- Rozporządzeniu Ministra Edukacji Narodowej z dnia 10 sierpnia 2001 r. w sprawie standardów wymagań będących podstawą przeprowadzania sprawdzianów i egzaminów (DzU nr 92, poz. 1020, z późn. zm.)

- Rozporządzeniu Ministra Edukacji Narodowej z dnia 10 marca 2010 r. w sprawie standardów wymagań będących podstawą przeprowadzania egzaminu potwierdzającego kwalifikacje zawodowe (DzU nr 103, poz. 652)
- Rozporządzeniu Ministra Edukacji Narodowej z dnia 27 września 2007 r. w sprawie standardów wymagań będących podstawą przeprowadzania egzaminów eksternistycznych (DzU nr 184, poz.1309).

1.1. Sprawdzian

- Sprawdzian jest egzaminem przeprowadzonym w szóstej klasie szkoły podstawowej. Jest on powszechny i obowiązkowy, co oznacza, że muszą do niego przystąpić wszyscy uczniowie – zarówno uczniowie szkół podstawowych dla dzieci i młodzieży, jak i słuchacze szkół podstawowych dla dorosłych. Przystąpienie do sprawdzianu jest jednym z warunków ukończenia szkoły.
- Sprawdzianu nie można nie zdać. Wynik ma znaczenie tylko informacyjne i nie powinien być podstawą do prowadzenia jakiegokolwiek selekcji. Każdy uczeń, który ukończył szkołę podstawową, niezależnie od wyników sprawdzianu musi być przyjęty do gimnazjum w swoim rejonie, jeśli nie ukończył 16 roku życia.
- Uczniowie, którzy do sprawdzianu nie przystąpią w danym roku, muszą powtórzyć ostatnią klasę szkoły podstawowej i przystąpić do sprawdzianu w roku następnym.
- Na sprawdzianie badany i oceniany jest poziom osiągnięć uczniów w zakresie opisanym w standardach wymagań egzaminacyjnych, ujętych w pięciu obszarach umiejętności:
 - czytania
 - pisania
 - rozumowania
 - korzystania z informacji
 - wykorzystywania wiedzy w praktyce.
- Na sprawdzianie można uzyskać maksymalnie 40 punktów. Wyniku sprawdzianu nie odnotowuje się na świadectwie szkolnym.

1.2. Egzamin gimnazjalny

- Egzamin gimnazjalny przeprowadzany jest w trzeciej klasie gimnazjum. Jest on powszechny i obowiązkowy, co oznacza, że musi przystąpić do niego każdy uczeń (słuchacz) kończący szkołę. Przystąpienie do egzaminu jest bowiem jednym z warunków jej ukończenia.
- Jeżeli jakiś uczeń (słuchacz) nie przystąpi do egzaminu gimnazjalnego w danym roku, musi powtórzyć ostatnią klasę gimnazjum i przystąpić do niego w roku następnym.
- Liczba punktów uzyskana na egzaminie gimnazjalnym nie ma wpływu na ukończenie szkoły. Wynik egzaminu gimnazjalnego jest jednak brany pod uwagę przez szkoły ponadgimnazjalne przy przyjmowaniu uczniów do tych szkół.
- Egzamin gimnazjalny sprawdza umiejętności międzyprzedmiotowe, uwzględniając powiązania między różnymi dziedzinami wiedzy. Zdawany jest w trzech częściach:
 - *pierwsza* obejmuje wiadomości i umiejętności z zakresu przedmiotów humanistycznych: języka polskiego, historii, wiedzy o społeczeństwie, plastyki, muzyki oraz uwzględnia ścieżki edukacyjne: filozoficzną, regionalną – dziedzictwo kulturowe w regionie, czytelniczą i medialną, europejską, kulturę polską na tle tradycji śródziemnomorskiej;
 - *druga* obejmuje wiadomości i umiejętności z zakresu przedmiotów matematyczno-przyrodniczych: matematyki, biologii, geografii, chemii, fizyki i astronomii oraz uwzględnia ścieżki edukacyjne: filozoficzną, prozdrowotną, ekologiczną, czytelniczą i medialną, regionalną – dziedzictwo kulturowe w regionie, europejską, obronę cywilną;
 - *trzecia* obejmuje wiadomości i umiejętności z zakresu języka obcego nowożytnego (angielski, francuski, hiszpański, niemiecki, rosyjski i włoski), którego uczeń (słuchacz) uczy się w szkole jako przedmiotu obowiązkowego.
- Na egzaminie gimnazjalnym sprawdzane jest opanowanie wiadomości i umiejętności określonych w standardach wymagań egzaminacyjnych będących podstawą przeprowadzania egzaminu w trzeciej klasie gimnazjum, obowiązujących we wszystkich szkołach w Polsce.

Standardy wymagań z zakresu przedmiotów humanistycznych egzaminu ujęte są w dwóch obszarach:

- czytania i odbioru tekstów kultury
- tworzenia własnego tekstu.

Standardy wymagań z zakresu przedmiotów matematyczno-przyrodniczych ujęte są w czterech obszarach:

- umiejętnego stosowania terminów, pojęć i procedur z zakresu przedmiotów matematyczno-przyrodniczych niezbędnych w praktyce życiowej i dalszym kształceniu;
- wyszukiwania i stosowania informacji;
- wskazywania i opisywania faktów, związków i zależności, w szczególności przyczynowo-skutkowych, funkcjonalnych, przestrzennych i czasowych;
- stosowania zintegrowanej wiedzy i umiejętności do rozwiązywania problemów.

Standardy wymagań z zakresu języka obcego nowożytnego ujęte są w trzech obszarach:

- odbioru tekstu słuchanego
- odbioru tekstu czytanego
- reagowania językowego.

W przypadku uczniów niesłyszących i słabo słyszących nie sprawdza się wiadomości i umiejętności z zakresu odbioru tekstu słuchanego.

1.3. Egzamin maturalny

- Egzamin maturalny jest przeprowadzany jeden raz w ciągu roku szkolnego – w okresie od maja do września.
- Egzamin maturalny zdawany jest w macierzystej szkole zdającego.
- Absolwent, przystępując do egzaminu maturalnego, zdaje obowiązkowo *w części ustnej*:
 - język polski,
 - język obcy nowożytny wybrany spośród języków: angielskiego, francuskiego, hiszpańskiego, niemieckiego, rosyjskiego i włoskiego – na poziomie podstawo-

- wym albo – jeżeli wybrany język był drugim językiem nauczania w szkole lub oddziale dwujęzycznym – na poziomie określonym dla szkół dwujęzycznych,
- język mniejszości narodowej, jeżeli jest absolwentem szkoły lub oddziału z językiem nauczania mniejszości narodowej.

w części pisemnej:

- język polski na poziomie podstawowym,
 - matematykę na poziomie podstawowym,
 - język obcy nowożytny, ten sam co w części ustnej, na poziomie podstawowym albo – jeżeli wybrany język był drugim językiem nauczania w szkole lub oddziale dwujęzycznym – na poziomie określonym dla szkół dwujęzycznych,
 - język mniejszości narodowej, jeżeli jest absolwentem szkoły lub oddziału z językiem nauczania mniejszości narodowej na poziomie podstawowym.
- Absolwent może ponadto przystąpić w danym roku do egzaminu maturalnego z nie więcej niż sześciu przedmiotów dodatkowych.
 - Wybór przedmiotów zdawanych na egzaminie maturalnym nie jest zależny od typu szkoły, do której uczęszczał absolwent, ani od przedmiotów nauczanych w tej szkole.

1.4. Egzamin zawodowy

- Egzamin potwierdzający kwalifikacje zawodowe zwany egzaminem zawodowym jest formą oceny poziomu opanowania wiadomości i umiejętności z zakresu danego zawodu, ustalonych w standardach wymagań będących podstawą przeprowadzania egzaminu potwierdzającego kwalifikacje zawodowe.
- Egzamin zawodowy zastępuje dotychczasowy egzamin z nauki zawodu i egzamin z przygotowania zawodowego.
- Egzamin zawodowy przeprowadzany jest dla absolwentów:
 - zasadniczych szkół zawodowych
 - techników
 - techników uzupełniających
 - szkół policealnych.

- Egzamin zawodowy składa się z dwóch etapów:
 - etapu pisemnego (który składa się z dwóch części)
 - etapu praktycznego.
- Absolwent, który zdał egzamin zawodowy, otrzymuje dyplom potwierdzający kwalifikacje zawodowe.
- Na wniosek absolwenta, do dyplomu potwierdzającego kwalifikacje zawodowe dołącza się Europas – Supplement do Dyplomu Potwierdzającego Kwalifikacje Zawodowe, sporządzony na podstawie opisu kwalifikacji absolwenta albo opisu zawodu, określonych w podstawie programowej kształcenia w danym zawodzie. Dokument ten nie zastępuje oryginału dyplomu potwierdzającego kwalifikacje zawodowe, ani też nie uprawnia do formalnego uznania dyplomu przez instytucje w innych krajach, ale ułatwia posiadaczom poruszanie się po europejskim rynku pracy.

1.5. Uczniowie o specjalnych potrzebach edukacyjnych

Około 2% populacji uczniów przystępujących do sprawdzianu w szkole podstawowej i uczniów przystępujących do egzaminu gimnazjalnego stanowią uczniowie o specjalnych potrzebach edukacyjnych. Na egzaminie maturalnym odsetek zdających o specjalnych potrzebach edukacyjnych jest znacznie niższy. Tacy uczniowie/zdający przystępują do sprawdzianu/egzaminu w warunkach dostosowanych i formie dostosowanej do ich potrzeb psychofizycznych i edukacyjnych. Dostosowanie to jest możliwe na podstawie orzeczenia o potrzebie kształcenia specjalnego wydanego przez poradnię psychologiczno-pedagogiczną. Dostosowanie warunków i formy przeprowadzania sprawdzianu/egzaminu uszczegóławiają komunikaty dyrektora CKE i polega m.in. na dostosowaniu arkuszy egzaminacyjnych do rodzaju niepełnosprawności ucznia/zdającego. Analizie statystycznej i interpretacji poddawane są wyniki uczniów/zdających rozwiązujących standardowy zestaw zadań w terminie głównym, ponieważ grupy uczniów/zdających rozwiązujących pozostałe typy zestawów oraz przystępujących do sprawdzianu/egzaminu w terminie dodatkowym są mało liczne i wnioskowanie obciążone byłoby dużym błędem.

2. PODSTAWOWE POJĘCIA STATYSTYCZNE

2.1. Podstawowe wskaźniki statystyczne

Wynik ucznia/zdającego – liczba punktów lub odsetek punktów uzyskanych przez ucznia/zdającego za wszystkie zadania z zestawu.

Wynik średni – średnia arytmetyczna wyników uzyskanych przez wszystkich zdających z badanej grupy (np. wszystkich uczniów z danej szkoły czy gminy przystępujących do sprawdzianu/egzaminu). Obliczany jest jako iloraz sumy wyników tych zdających przez liczbę zdających:

$$\text{wynik średni} = \frac{\text{suma punktów uzyskanych przez wszystkich zdających}}{\text{liczba zdających}}$$

Średni wynik może być wyrażony w procentach.

Średni wynik, żeby był wiarygodny, musi być wyliczony z odpowiednio dużego zbioru wyników. Im większy zbiór wyników, tym bardziej godny zaufania jest wynik średni. Natomiast im mniejszy zbiór, tym większy udział przypadku w powstawaniu wyniku średniego – nawet jeden skrajny wynik może znacząco wpłynąć na wartość średniej arytmetycznej. Dlatego wynik średni nie powinien być jedyną miarą efektywności nauczania.

Wynik minimalny – najniższy wynik spośród uzyskanych przez uczniów danej grupy (np. uczniów danej szkoły lub gminy).

Wynik maksymalny – najwyższy wynik spośród uzyskanych przez zdających z danej grupy (np. uczniów danej szkoły lub gminy).

Modalna – wynik najczęściej występujący w badanej grupie.

Mediana – wynik środkowy w badanej grupie zdających (np. uczniów danej szkoły lub gminy); dzieli zdających na dwie równoliczne grupy: w jednej zdający mają wynik niższy od mediany lub równy medianie, w drugiej – wynik wyższy od mediany lub jej równy.

Odchylenie standardowe – wyrażona w punktach miara rozrzutu wyników zdających z danej grupy (np. uczniów danej szkoły lub gminy):

$$\text{odchylenie standardowe} = \sqrt{\frac{\text{suma kwadratów odchyleń od wyniku średniego}}{\text{liczba uczniów}}}$$

Informację tę wykorzystujemy dla określenia stopnia skupienia wyników wokół wyniku średniego: im wyższa wartość odchylenia standardowego, tym większe zróżnicowanie wyników. Na przykład:

Przyjmijmy, że średni wynik uczniów na sprawdzianie wynosi 20,5 punktów a odchylenie standardowe 6,5. Oznacza, że około 70% wyników uczniów uzyskało wyniki w przedziale od 14,0 do 27,0 punktów. Granice tego przedziału obliczamy, odejmując od wyniku średniego i dodając do wyniku średniego wartość odchylenia standardowego (20,5-6,5 i 20,5+6,5).

Współczynnik zmienności – miara zróżnicowania rozkładu cechy. Wielkość niemianowana obliczana następująco:

$$\text{współczynnik zmienności} = \frac{\text{odchylenie standardowe}}{\text{wynik średni}}$$

Stosowany jest najczęściej przy porównywaniu zróżnicowania cechy w dwóch różnych rozkładach (np. wyników chłopców i dziewcząt).

Łatwość zadania (umiejętności) obliczana jest następująco:

$$\text{łatwość} = \frac{\text{suma punktów uzyskanych przez uczniów za dane zadanie (umiejętność)}}{\text{suma punktów możliwych do uzyskania przez uczniów za to zadanie (umiejętność)}}$$

Zgodnie z interpretacją prof. B. Niemierki przyjmuje się, że zadanie jest:

<i>bardzo trudne</i>	gdy ma łatwość:	0,00 – 0,19,
<i>trudne</i>	gdy ma łatwość:	0,20 – 0,49,
<i>umiarkowanie trudne</i>	gdy ma łatwość:	0,50 – 0,69,
<i>łatwe</i>	gdy ma łatwość:	0,70 – 0,89,
<i>bardzo łatwe</i>	gdy ma łatwość:	0,90 – 1,00.

Wskaźnik łatwości może służyć analizie porównawczej wyników szkoły. Porównując wartości wskaźnika w szkole z wartościami krajowymi i wojewódzkimi można określić mocne i słabe strony zdających.

Poziom opanowania umiejętności określonych standardami (osiągnięcia w obszarach umiejętności) obliczany jest analogicznie jak łatwość zadania jako iloraz sumy punktów uzyskanych przez zdających za wszystkie zadania badające daną umiejętność przez sumę punktów możliwych do uzyskania za te zadania.

Średnie wyniki uzyskane przez szkoły w poszczególnych obszarach umiejętności pozwalają na pełniejszą analizę osiągnięć szkoły i ocenę efektywności kształcenia.

2.2. Rozkłady wyników

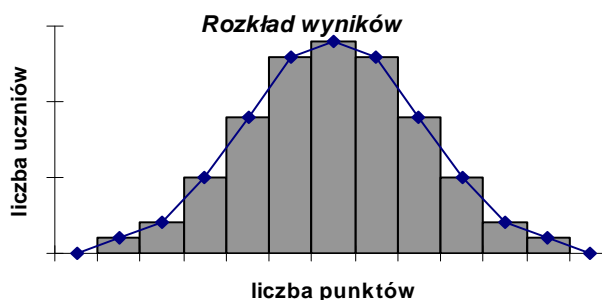
Tabelaryczny rozkład wyników jest zestawieniem, które obejmuje wynik surowy (liczbę lub procent punktów) oraz liczbę zdających (liczebność) i/lub procent zdających, którzy dany wynik uzyskali.

W postaci tabelarycznej można także przedstawić łatwość zadań w podtestach np. według standardów, rozkłady wyników wyrażonych w skalach znormalizowanych. Można również zestawiać wyniki różnych grup zdających (np. szkół).

Graficzny rozkład wyników to wykres (np. słupkowy). Na osi poziomej umieszczona jest liczba uzyskanych punktów. Mogą one być podawane co 1 punkt lub pogrupowane w klasy po kilka punktów (np. co 2 lub 5 punktów). Na osi pionowej podany jest procent lub liczba zdających, którzy dany wynik uzyskali. Rozkłady wyników w postaci graficznej stosuje się do prezentacji surowych wyników dla dużych populacji.

Przykłady rozkładów graficznych

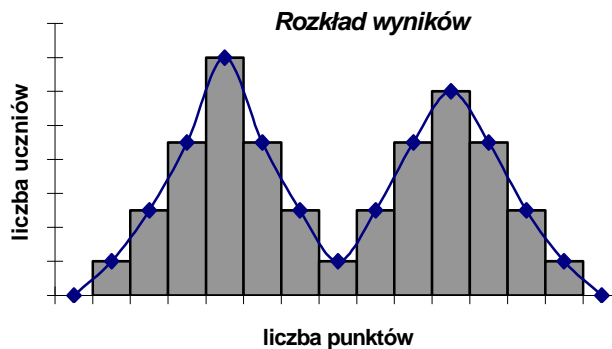
Rozkład normalny. Jest to rozkład wzorcowy. Przykładem zmiennej o rozkładzie normalnym jest wzrost.



Rozkład zbliżony do normalnego może oznaczać, że:

- *zadania w teście dobrze reprezentowały zróżnicowany poziom wiedzy i umiejętności uczniów*
- *w teście były zadania zarówno dla uczniów bardzo słabych, dobrych jak i bardzo dobrych*
- *maksymalnych wyników jest niewiele, co może oznaczać także, że czas przeznaczony na jego rozwiązanie sprzyjał bardzo dobrym uczniom.*

Rozkład dwumodalny



Rozkład taki może oznaczać, że:

- zestaw zadań był średnio trudny
- mało jest wyników przeciętnych, a dużo wyników słabych oraz dużo wyników dobrych
- zdający zostali wyraźnie zróżnicowani na dwie grupy, co może być korzystne w przypadku konkursu: zdający z jednej grupy przechodzą do kolejnego etapu, z drugiej – nie przechodzą.

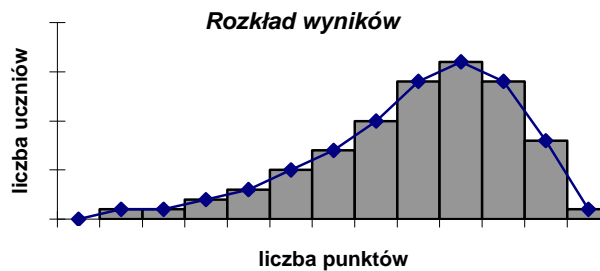
Rozkład dodatnio skośny (prawoskośny)



Rozkład taki oznaczać może, że:

- w teście było wiele zadań, których nie rozwiązał prawie nikt spośród badanych
- zdający w badanej grupie mają niskie osiągnięcia
- zestaw zadań był zbyt trudny dla badanej grupy.

Rozkład ujemnie skośny (lewoskośny)



Rozkład taki może oznaczać, że:

- zdający w badanej grupie osiągają wysokie wyniki
- zestaw zadań był zbyt łatwy
- test miał głównie zadanie sprawdzające a nie różnicujące
- taki test nie nadaje się na konkurs ani do selekcji uczniów np. na studia.

2.3. Przykłady skal pomiarowych

Skala dziewięciostopniowa (staninowa)

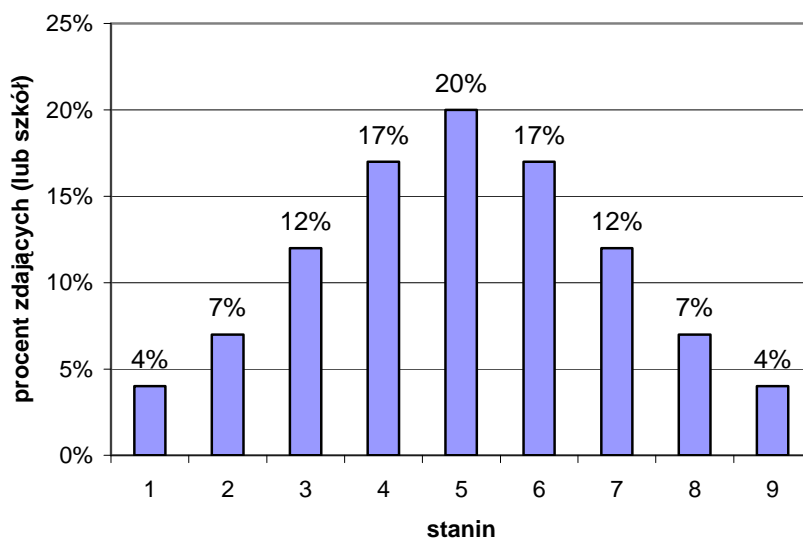
Skala staninowa – dziewięciostopniowa znormalizowana skala wyników. Nazwa pochodzi od ang. *standard nine* czyli „standardowa dziewiątka”. Pozwala pozycjonować wynik ucznia lub też wynik szkoły (na przykład średni wynik lub wskaźnik EWD) i porównywać z wynikami innych uczniów (innych szkół). Skala ta jest przydatna do opisu wyników egzaminów, jeżeli nie są one wyrażone na standaryzowanej skali, tak jak w przypadku polskich egzaminów. Ze względu jednak na to, że zawiera tylko 9 kategorii, jest mało dokładna i może jedynie służyć jako jeden ze sposobów komunikowania wyników egzaminów.

Skala staninowa opracowywana jest co roku przez Centralną Komisję Egzaminacyjną w Warszawie dla konkretnego egzaminu: sprawdzianu i poszczególnych części egzaminu gimnazjalnego (humanistycznej, matematyczno-przyrodniczej i z języków obcych nowożytnych: angielskiego, niemieckiego, francuskiego i rosyjskiego). Dla każdego z tych egzaminów opracowywane są dwie krajowe skale: jedna dla wszystkich uczniów rozwiązujących standardowy zestaw zadań, druga dla średnich wyników wszystkich szkół.

Skala ta umożliwia także porównywanie wyników szkół w dłuższym przedziale czasowym.

Kolejne stopnie tej skali zawierają odpowiednio 4%, 7%, 12%, 17%, 20%, 17%, 12%, 7%, 4% wyników (uczniów lub szkół) uporządkowanych rosnąco. Podział danej grupy przedstawiono w tabeli i na wykresie poniżej:

Stopień skali (stanin)	Opis wyniku	Procent zdających (lub szkół)
1	najniższy	4%
2	bardzo niski	7%
3	niski	12%
4	niżej średni	17%
5	średni	20%
6	wyżej średni	17%
7	wysoki	12%
8	bardzo wysoki	7%
9	najwyższy	4%

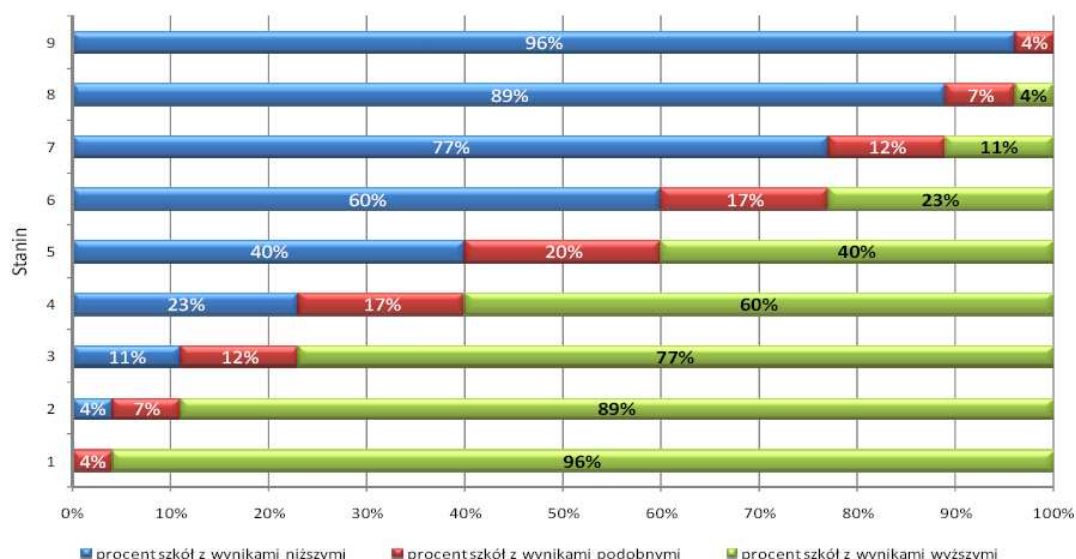


Średnia dla tej skali wynosi 5, a odchylenie standardowe 2.

Dla danego egzaminu należy określić granicę (czyli wynik) między stopniami. W tabeli poniżej przedstawiono takie przedziały dla średnich wyników szkół na *Sprawdzianie 2010* i na *Sprawdzianie 2009*:

Stopień skali (stanin)	Sprawdzian 2009	Sprawdzian 2010
1	7,9–17,0	7,6–18,5
2	17,1–18,7	18,6–20,3
3	18,8–20,0	20,4–21,8
4	20,1–21,3	21,9–23,2
5	21,4–22,7	23,3–24,8
6	22,8–24,1	24,9–26,2
7	24,2–25,6	26,3–27,8
8	25,7–27,7	27,9–29,9
9	27,8–35,7	30,0–37,7

Skala staninowa dla średnich wyników szkół pozwala na porównanie wyników danej szkoły do osiągnięć wszystkich szkół w kraju. Poniżej przedstawiono graficzną ilustrację procentowego udziału szkół według wyników w standardowej skali dziewięciostopniowej:



Jeżeli średni wynik z egzaminu w części humanistycznej w danym gimnazjum plasuje tę szkołę w ósmym staninie (jest określany jako wynik „bardzo wysoki”), to można stwierdzić, że:

- *wynik podobny do wyniku tej szkoły (kolor czerwony na wykresie) uzyskało 7% gimnazjów w Polsce*
- *wynik niższy (kolor niebieski) – 89% gimnazjów*
- *wynik wyższy (kolor zielony) – 4% gimnazjów.*

Analizując wyniki szkoły w skali staninowej w dłuższym przedziale czasowym (szkoła podstawowa od roku 2002, gimnazjum od roku 2003), można określić tendencję rozwojową szkoły (np. szkoła systematycznie osiąga coraz lepsze wyniki w stosunku do innych szkół w kraju lub wyniki szkoły ustabilizowały się na poziomie średnim itd.).

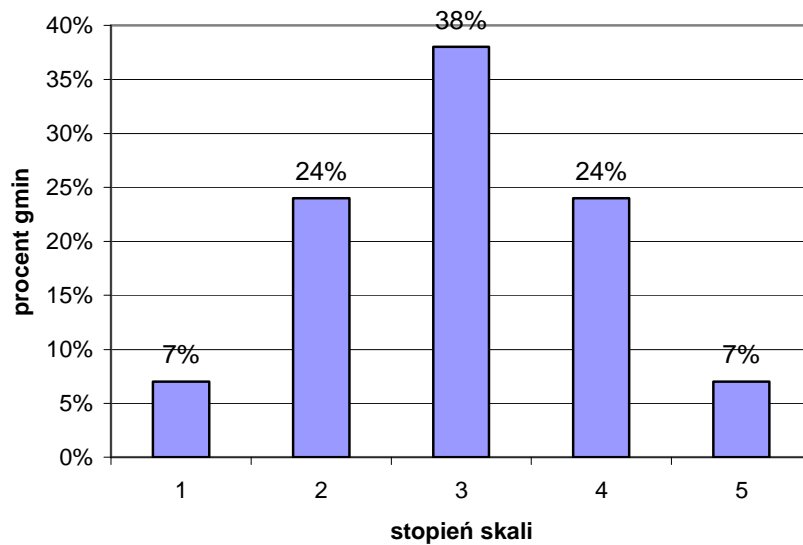
Skalę staninową wyników uczniów opracowaną dla sprawdzianu lub egzaminu gimnazjalnego można wykorzystać do podzielenia uczniów na trzy grupy:

- uczniowie, których wynik plasuje się w 1., 2. lub 3. staninie, to uczniowie, którzy w kolejnym etapie kształcenia wymagają dodatkowej pomocy, gdyż zagrożeni są niskimi osiągnięciami
- uczniowie, których wynik plasuje się w 4., 5. lub 6. staninie, to uczniowie o przeciętnych osiągnięciach
- uczniowie, których wynik plasuje się w 7., 8. lub 9. staninie, to uczniowie o znacznym potencjale, który warto wykorzystać w kolejnym etapie kształcenia.

Skala pięciostopniowa

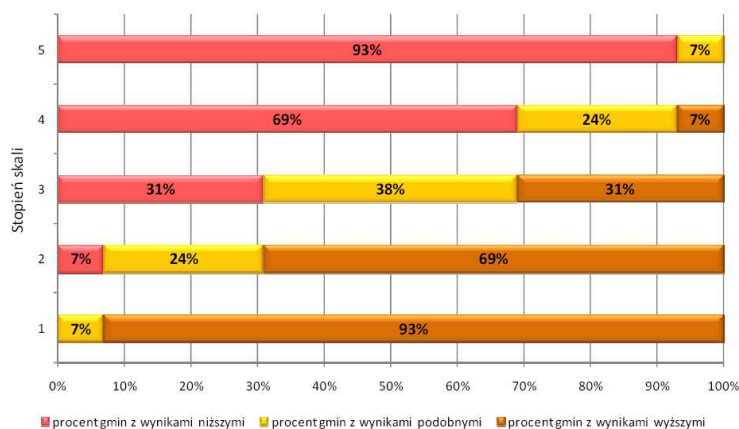
Znormalizowaną skalą standardową, którą wykorzystuje się do mniejszych zbiorów wyników (takich jak zbiór średnich wyników gmin w danym województwie) jest skala pięciostopniowa. Opracowuje się ją podobnie jak skalę dziewięciostopniową w oparciu o rozkład normalny średnich wyników gmin danego województwa. Po uporządkowaniu surowych wyników od najniższego do najwyższego, wyróżnia się 5 przedziałów:

Stopień skali	Opis wyniku	Procent gmin
1	niski	7%
2	niżej średni	24%
3	średni	38%
4	wyżej średni	24%
5	wysoki	7%



Średnia w skali pięciostopniowej wynosi 3, a odchylenie standardowe 1.

Skala pięciostopniowa pozwala na porównanie wyniku danej gminy z wynikami wszystkich gmin województwa. Graficzną ilustracją procentowego udziału gmin według wyników w standardowej skali pięciostopniowej przedstawia poniższy wykres:



Jeśli średni wynik uczniów gminy X na egzaminie gimnazjalnym w części humanistycznej odpowiada 4 przedziałowi skali pięciostopniowej (wyznaczonej dla odpowiedniego województwa), to wniosek jest następujący:

- 7% gmin w województwie ma wynik wyższy (kolor brązowy na wykresie)
- 24% – wynik zbliżony (kolor żółty)
- 69% – wynik niższy (kolor czerwony).

Gmina może również (na podstawie sprawozdań OKE) porównać swoją pozycję na skali pięciostopniowej w poszczególnych latach.

Skala centylowa

Jest to skala pozycyjna wyników uporządkowanych rosnąco: aby ustalić wartość centyli, porządkuje się wyniki od najniższych do najwyższych, dzieli ich zbiór na 100 równych części, a następnie określa położenie granic między tymi częściami na skali wyników i oznacza się je jako centyle C1, C2, C3, C4, C5,... C99. Liczba centyli wynosi 99. Grup centylowych, czyli wyników leżących między centylami, jest 100. Numer kolejny najbliższego centyla stanowi rangę centylową danego wyniku testowania, to jest procent wyników niższych od danego. Na przykład wynik szkoły w 45. centylu oznacza, że 44% szkół uzyskało wynik niższy od tej szkoły.

2.4. Edukacyjna wartość dodana

(opracowano na podstawie materiałów zespołu EWD – <http://www.ewd.edu.pl> i <http://cke.scholaris.pl/>)

Wynik egzaminu zewnętrznego to informacja o końcowym poziomie osiągnięć uczniów w zakresie sprawdzanym testem egzaminacyjnym, a nie informacja o efektywności nauczania w szkole. Średni wynik egzaminu zewnętrznego może być i niejednokrotnie jest fałszywym sygnałem efektywności pracy szkoły. Metoda edukacyjnej wartości dodanej jest alternatywnym sposobem wykorzystania wyników egzaminacyjnych i może posłużyć do oszacowania na podstawie wyników egzaminacyjnych miary efektywności pracy szkoły w rozumieniu efektywności wykorzystania zasobów na wejściu, czyli innymi słowy pozwala ocenić wkład szkoły w końcowy wynik egzaminacyjny. Taki alternatywny sposób wykorzystania wyników egzaminacyjnych pozwala porównać szkoły nie tylko ze względu na osiągnięte przez uczniów końcowe wyniki egzaminacyjne, ale również ze względu na wkład szkoły w uzyskanie tych wyników.

Wskaźnik edukacyjnej wartości dodanej jest potrzebny polskiej szkole do dokonywania pełniejszej ewaluacji nauczania, powinien uzupełniać inne wskaźniki obrazujące pracę szkoły.

Termin *edukacyjna wartość dodana* oznacza zarówno metodę, jak i wskaźnik liczbowy wyliczony zgodnie z tą metodą. Metody EWD to metody statystyczne (najczęściej wykorzystywane są modele regresji liniowej lub modele wielopoziomowe) pozwalające na podstawie zasobów na wejściu (np. wynik na egzaminie poprzedniego etapu kształcenia, informacja o zasobach materialnych szkoły, środowisku, nakładach na szkołę) oraz na wyjściu (np. wynik na egzaminie końcowym dla danego etapu kształcenia) oszacować efektywność nauczania, czyli wkład danej szkoły w końcowy poziom wiedzy uczniów na danym etapie kształcenia.

Model szacowania EWD dla potrzeb polskiego systemu egzaminów zewnętrznych opracowany został przez Zespół EWD w 2006 roku i aktualizowany w latach następnych. Pozwala on szacować wskaźniki EWD dla gimnazjum na podstawie wyników jednej edycji egzaminów zewnętrznych (sprawdzian po VI klasie szkoły podstawowej i mający miejsce 3 lata później egzamin gimnazjalny). Model ten wykorzystuje technikę regresji liniowej. W modelu uwzględnia się wynik sprawdzianu (wejście), wynik egzaminu gimnazjalnego (wyjście) oraz dodatkowe informacje o uczniach (płeć, dysleksja) jako zmienne kontrolne. Na podstawie modelu wyznacza się empirycznie (na podstawie wyników wszystkich polskich gimnazjalistów przystępujących w danym roku do egzaminu) krzywą przewidywanego wyniku. W uproszczeniu przewidywany wynik można wyobrazić sobie jako pewną uśrednioną wartość wyników gimnazjalnych wszystkich uczniów w Polsce, którzy uzyskali dany wynik na sprawdzianie (np. w 2008 roku wynik przewidywany w części humanistycznej egzaminu gimnazjalnego dla uczniów, którzy uzyskali 20 punktów na sprawdzianie wynosił 19,9 punktów, a w części matematyczno-przyrodniczej – 17,7 punktów). Dla każdego ucznia wyliczana jest reszta – różnica między wynikiem rzeczywiście uzyskanym na egzaminie a wynikiem przewidywanym. Dodatnia reszta oznacza, że uczeń przez 3 lata nauki w gimnazjum zrobił relatywny postęp, czyli uzyskał wynik wyższy niż przeciętnie jego rówieśnicy z takim samym wynikiem na sprawdzianie. Jeśli reszta jest ujemna – uczeń zrobił relatywnie mniejsze postępy niż jego rówieśnicy z takim samym wynikiem na sprawdzianie. Średnia z indywidualnych reszt dla grupy uczniów (szkoła, klasa, uczniowie uczeni przez tego samego nauczyciela itp.) jest **wskaźnikiem edukacyjnej wartości dodanej** dla tej grupy. Linia przewidywanego wyniku jest poprowadzona tak, żeby średnie reszt dla grup uczniów

wyróżnionych ze względu na wynik na sprawdzanie w całej populacji uczniów były równe 0. Oznacza to, że szansa szkoły na uzyskanie wysokiego lub niskiego EWD nie zależy od tego, jakie wyniki na sprawdzanie mają poszczególni uczniowie danego gimnazjum.

Ze sposobu obliczania edukacyjnej wartości dodanej wynika, że EWD jest miarą względną i pozwala porównać efektywność nauczania w zakresie badanym egzaminem zewnętrznym. Dodatnia wartość EWD informuje o ponadprzeciętnej efektywności, ujemna wartość EWD o efektywności niższej przeciętnej. Wskaźniki EWD są podawane w punktach egzaminu gimnazjalnego dla danego roku. Zatem $EWD = 0,4$ oznacza, że średnio uczeń w szkole uzyskał wynik o 0,4 punktu wyższy od wyniku przewidywanego. Podobnie jak dla średnich wyników egzaminacyjnych szkół, tak i dla wskaźników EWD dla szkół buduje się krajową skalę staninową.

Wyznaczona wartość wskaźnika EWD jest obarczona błędem (błędem pomiaru na sprawdzanie i egzaminie gimnazjalnym oraz błędem próbkowania). Dlatego też należy EWD dla analizowanej grupy uczniów komunikować wraz z przedziałem ufności. **Przedział ufności**, to przedział, który z określonym **współczynnikiem ufności** (czyli z określonym prawdopodobieństwem) zawiera prawdziwą wartość interesującego nas parametru. Wyznaczone dla EWD przedziały ufności możemy traktować jako regułę decyzyjną, niezbędną do odpowiedzialnego sformułowania oceny, która szkoła charakteryzuje się wyższą efektywnością. Jeżeli przedziały ufności EWD dla porównywanych szkół (grup) są rozłączne, to możemy wnioskować, że szkoły (grupy) zdecydowanie różnią się efektywnością nauczania. W przypadku, gdy przedziały częściowo pokrywają się, oceny należy formułować z większą ostrożnością. W przypadku EWD dla szkoły przyjmujemy 95% współczynnik ufności, a dla klas lub innych grup uczniów porównywanych w obrębie jednej szkoły można przyjąć niższy współczynnik ufności, np. 90%. Długość przedziału ufności zależy od liczby uczniów w szkole oraz od zróżnicowania reszt. Im większa liczba uczniów tym węższy przedział ufności, im mniejsze zróżnicowanie reszt tym węższy przedział ufności, czyli wyznaczona wartość EWD precyzyjniej charakteryzuje szkołę.

Przedział ufności dla szkoły możemy komunikować również w skali centylowej. Umieszczenie dolnego i górnego przedziału ufności na skali centylowej pozwala ocenić pozycję

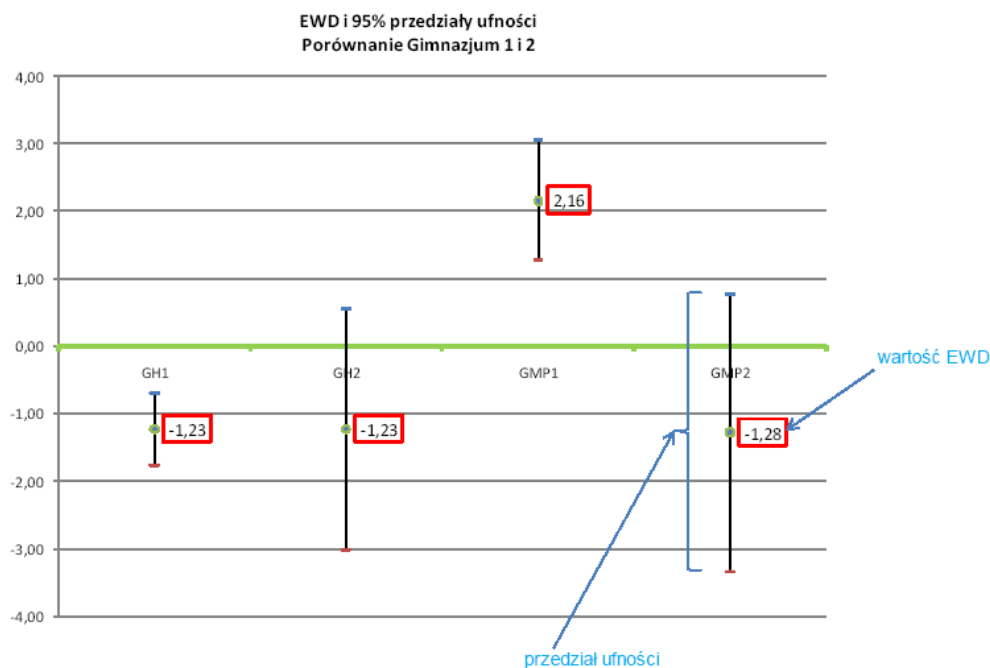
szkoły w porównaniu z innymi. Skala centylowa budowana jest – tak jak staninowa – na podstawie wyników EWD wszystkich gimnazjów w Polsce.

Przykład 1.

Jeżeli EWD dla danego gimnazjum wyniosła +3,5 pkt to oznacza, że uczniowie tej szkoły uzyskali średnio o 3,5 pkt więcej na egzaminie gimnazjalnym w porównaniu z innymi szkołami w kraju (*o analogicznym składzie ze względu na wyniki na sprawdzianie, płeć i dysleksję*).

Przykład 2.

Poniżej przedstawiony wykres obrazuje oszacowania EWD wraz z przedziałami ufności dla dwóch przykładowych gimnazjów i dwóch części egzaminu gimnazjalnego.



Interpretacja wykresu:

- Obie szkoły mają tę samą wartość EWD (*dla części humanistycznej egzaminu*), a przedziały ufności pokrywają się – w tej części egzaminu obydwa gimnazja charakteryzuje podobna efektywność nauczania.
- W części matematyczno-przyrodniczej Gimnazjum 1 ma wyższą EWD niż Gimnazjum 2, a przedziały ufności tych dwóch szkół nie mają wspólnej części – w pełni uprawniony statystycznie sposób można wnioskować, że efektywność na-

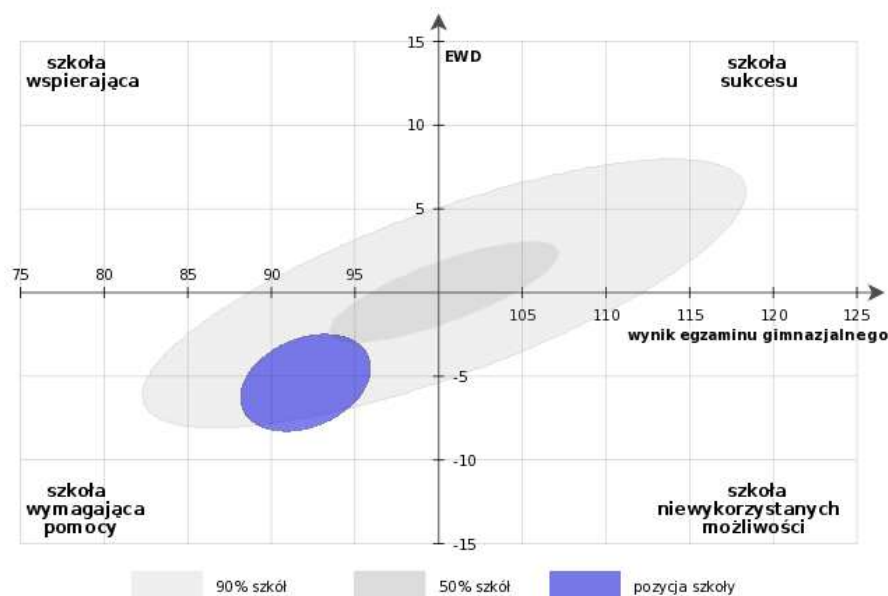
uczania w tych dwóch szkołach jest różna (w obszarze sprawdzanym przez egzamin matematyczno-przyrodniczy).

Jeśli przedziały ufności częściowo by się pokrywały, to takie wnioskowanie byłoby obarczone dużym ryzykiem popełnienia błędu.

- Szersze przedziały ufności dla Gimnazjum 2 mogą wynikać z mniejszej liczby uczniów w tej szkole lub (jeżeli liczby uczniów w gimnazjach są zbliżone) z większego zróżnicowania efektów pracy w Gimnazjum 2, co jest dodatkową, interesującą informacją o tej szkole.

Trzyletni wskaźnik EWD lub inaczej trzyletni wskaźnik egzaminacyjny to łączna prezentacja wyniku egzaminacyjnego i edukacyjnej wartości dodanej szkoły za okres trzech lat. Po raz pierwszy został opublikowany przez Zespół EWD w lipcu 2009 roku i obejmował lata 2006-2009. Forma graficzna, w jakiej przedstawiany jest wskaźnik, powoduje, że jego interpretacja jest intuicyjna, nie sprawia większych problemów i nie wymaga specjalistycznej wiedzy statystycznej.

Przykład 3.

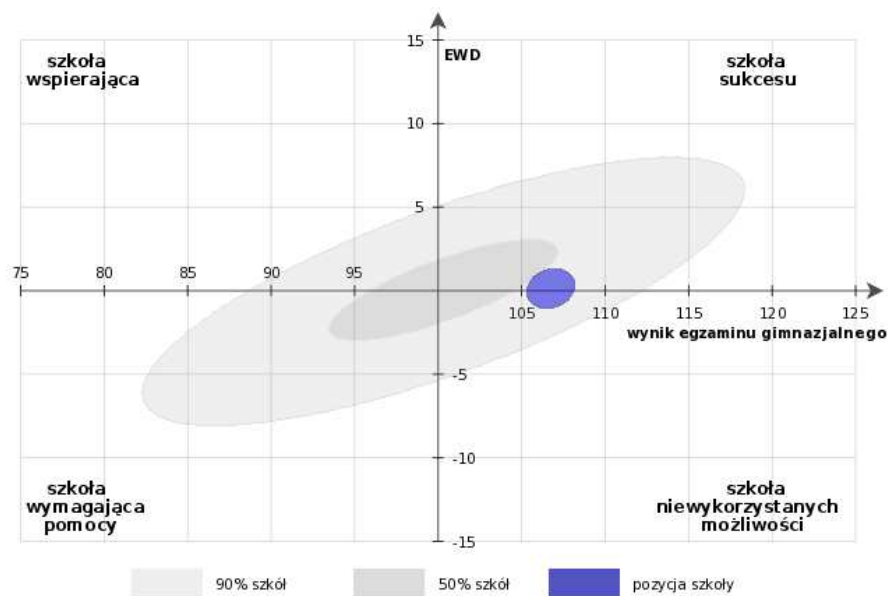


Liczba wyników egzaminacyjnych uwzględniona w analizie: **51**.
95% powierzchnia ufności dla łącznego oszacowania wskaźników.

Interpretacja wykresu:

- gimnazjum przez 3 lata uzyskiwało niskie wyniki egzaminu gimnazjalnego w części, której dotyczy wykres
- gimnazjum przez 3 lata w tej części egzaminu uzyskiwało niską EWD
- rozrzut jest duży, bo badana grupa jest mała (z każdego roku uwzględniono wyniki średnio 17 uczniów).

Przykład 4.



Liczba wyników egzaminacyjnych uwzględniona w analizie: **209**.
95% powierzchnia ufności dla łącznego oszacowania wskaźników.

Interpretacja wykresu:

- gimnazjum przez 3 lata uzyskiwało wysokie wyniki egzaminu gimnazjalnego w części, której dotyczy wykres
- gimnazjum przez 3 lata w tej części egzaminu uzyskiwało przeciętną EWD
- rozrzut jest mały, bo badana grupa jest duża (z każdego roku uwzględniono wyniki średnio 70 uczniów).