

Jacek Stańdo

Jak wykorzystać technologię w zbieraniu, analizie i wizualizacji danych?

- ✓ Procenty, centyle i skala staninowa
- ✓ Wizualizacja danych
- ✓ Manipulacje danymi



Analiza merytoryczna
Elżbieta Miterka

Recenzja
Jolanta Lazar

Redakcja językowa i korekta
Agata Jabłonowska-Turkiewicz

Projekt graficzny, projekt okładki
Wojciech Romerowicz, ORE

Skład i redakcja techniczna
Grzegorz Dębiński

Projekt motywu graficznego „Szkoly ćwiczeń”
Aneta Witecka

ISBN 978-83-65967-00-8 (Zestawy materiałów dla nauczycieli szkół ćwiczeń – matematyka)

ISBN 978-83-65967-28-2 (Zestaw 7. Wykorzystanie technologii informacyjno-komunikacyjnych w edukacji matematycznej w klasach IV–VIII szkoły podstawowej i szkole ponadpodstawowej)

ISBN 978-83-65967-32-9 (Zeszyt 4. Jak wykorzystać technologię w zbieraniu, analizie i wizualizacji danych?)

Warszawa 2017
Ośrodek Rozwoju Edukacji
Aleje Ujazdowskie 28
00-478 Warszawa
www.ore.edu.pl

Publikacja jest rozpowszechniana na zasadach wolnej licencji Creative Commons – Użycie niekomercyjne 3.0 Polska (CC-BY-NC).

Spis treści

Wstęp	4
Zbieranie danych	4
Metody zbierania danych	4
Dlaczego zbieramy dane?	4
Procenty, centyle i skala staninowa	7
Skala staninowa	7
Centyle i procenty	8
Sprawozdanie z egzaminów zewnętrznych	10
Rozkład wyników	10
Podstawowe parametry statystyczne	12
Wizualizacja danych	15
Manipulacje danymi	15
Wykres słupkowy	15
Wykres kołowy	16
Linia trendu	18
Jak przygotować ucznia do prawidłowego czytania i analizowania danych?	20
Edukacyjna wartość dodana	21
Definicja, kryteria	21
Uproszczony model edukacyjnej wartości dodanej	22
Trzyletnie wskaźniki EWD – przykład wizualizacji danych egzaminacyjnych	25



Analiza porównawcza szkół na podstawie wykresu	33
Bibliografia	42



Wstęp

Zbieranie, analiza i wizualizacja danych to zajęcia właściwe wszystkim ludziom, którzy pracują ze zbiorem informacji wymagających uporządkowania. Są to: naukowcy, nauczyciele, uczniowie, bibliotekarze, autorzy różnych publikacji czy pracownicy firm przetwarzających dane.

Firmy komercyjne gromadzą dane dotyczące konsumentów oraz ich zachowań na rynku. Wielu specjalistów, aby posługiwać się statystykami, musi najpierw opracować i zanalizować informacje.

W szkole niejednokrotnie mamy do czynienia z przetwarzaniem danych. Zajmuje się tym każdy nauczyciel prowadzący dziennik klasowy lub dokumentację szkolną. Uczniowie opracowują dane na lekcje z różnych przedmiotów. Na zajęciach z informatyki wyposażamy ich w niezbędne narzędzia i umiejętności, aby potrafili właściwie zająć się danymi w swoim zawodowym (i nie tylko) życiu.

Zbieranie danych

Metody zbierania danych

Stanisław Kaczmarczyk przedstawia następujący podział metod zbierania danych (2011):

- pośrednie metody sondażowe;
- bezpośrednie metody sondażowe;
- metody pozasondażowe (mogą być pośrednie i bezpośrednie);
- metody eksperymentu.

Dlaczego zbieramy dane?

Najważniejszym celem gromadzenia danych jest wykorzystanie ich w celach komercyjnych. Im szybciej dotrzemy do osób, które w danym momencie stoją np. przed wyborem zakupu samochodu, tym bardziej profil użytkownika jest wartościowy.

Największą firmą sprzedającą profil użytkownika jest Google. Około 90% obrotu pochodzi z reklam. Facebook gromadzi dane na zasadzie dobrowolności. Facebook ma jednak bardzo szczegółowe informacje: imię, nazwisko, datę urodzenia.

Każde z państw gromadzi informacje na temat swoich mieszkańców. Dzięki danym statystycznym rząd państwa lub samorząd województwa może podejmować określone decyzje dotyczące polityki społecznej, gospodarki, zmian prawnych itp. W Polsce taką instytucją jest Główny Urząd Statystyczny (GUS) – centralny organ administracji rządowej, podległy Prezesowi Rady Ministrów. GUS zajmuje się zbieraniem i udostępnianiem informacji

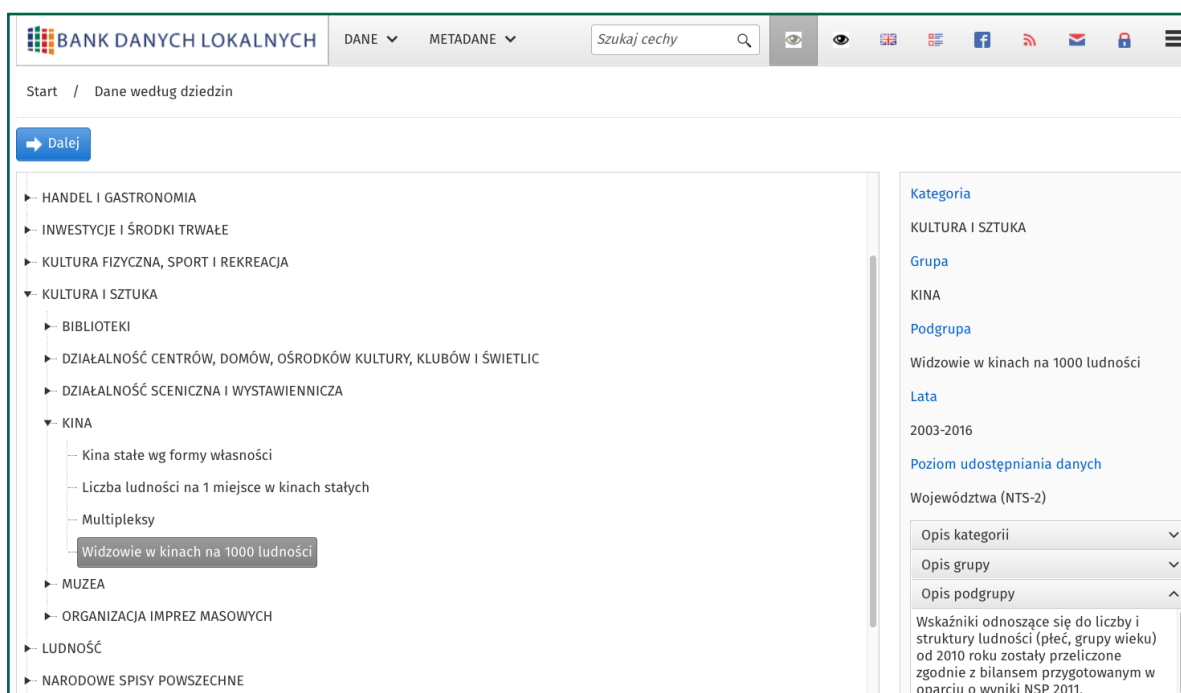


statystycznych na temat większości dziedzin życia publicznego i niektórych stron życia prywatnego Polaków. Przekazywanie danych odbywa się na podstawie przepisów

prawa (Ustawa o statystyce publicznej z dn. 29 czerwca 1995 r., Dz.U. 2016, poz. 1068 oraz ogłaszany corocznie Program Badań Statystycznych). Dane przed udostępnieniem są opracowywane tak, aby uniemożliwić identyfikację pojedynczych respondentów (tajemnica statystyczna).

Część zgromadzonych danych GUS udostępnia bezpłatnie w serwisie internetowym stat.gov.pl. Znajdziemy tutaj różnego rodzaju zestawienia i zbiory danych. Na uwagę zasługują Statystyka Regionalna i Dziedzinowe Bazy Wiedzy. Zgromadzono tu informacje o sytuacji społeczno-gospodarczej, demograficznej, społecznej oraz stanie środowiska

poszczególnych województw, powiatów oraz gmin Polski. Dział Statystyka Międzynarodowa pokazuje sytuację na świecie. Chcąc uzyskać dane dotyczące interesującego nas zakresu, będziemy musieli samodzielnie je wygenerować.



Tak wygenerowane dane możemy obejrzeć przedstawione graficznie w postaci wykresu słupkowego czy mapy. Ale co ważniejsze – możemy je również pobrać (np. w postaci pliku tekstowego CSV) i wykorzystać na lekcjach matematyki, tworząc na podstawie obszernych danych zestawienia w arkuszu kalkulacyjnym.



Kategoria: **KULTURA I SZTUKA**
 Grupa: **KINA**
 Podgrupa: **Widzowie w kinach na 1000 ludności**
 Wymiary: Ogółem; Lata
 Ostatnia aktualizacja: 2017-11-09

Tablica Wykres Mapa

Wybór jednostek terytorialnych Agregaty kod Puste Export Objaśnienia Szukaj w tabeli

Jednostka terytorialna

POLSKA

Export menu:

- XLS - tablica wielowymiarowa
- XLS - tablica relacyjna (zip)
- XLS - tablica przestawna
- CSV - tablica wielowymiarowa
- CSV - tablica relacyjna (zip)

ogółem
2016
[osoba]

13

Jedną z zakładek na stronie GUS-u jest Portal Edukacyjny skierowany do dzieci i młodzieży. W przystępny sposób zostały w nim przedstawione zasady statystyki publicznej, tajemnicy statystycznej, metody zbierania danych, ich zakres, sposoby publikacji oraz przykłady wykorzystania. Znajdziemy tutaj również wytłumaczenie podstawowych pojęć statystycznych: średnia arytmetyczna i średnia ważona.

Uczniowie szkół ponadpodstawowych mogą zadeklarować swój udział w Olimpiadzie Statystycznej, a najmłodsi, grając w gry edukacyjne, zapoznać się z tematyką statystyki.

Portal Edukacyjny

- Olimpiada Statystyczna
- Co to jest statystyka publiczna?
- Historia statystyki
- Statystyka dla smyka
- Co uczeń powinien wiedzieć?
- Edukacja w regionie
- Wydarzenia
- Gry edukacyjne

Gry edukacyjne

MINI SPIS

MAŁY SPIS

PRAWDOPODOBIEŃSTWO I ROZKŁAD NORMALNY

LICZYMY PRAWDOPODOBIEŃSTWO

WYŚCIGI

ŚREDNI WZRÓST

LOSOWANIE

Do gry



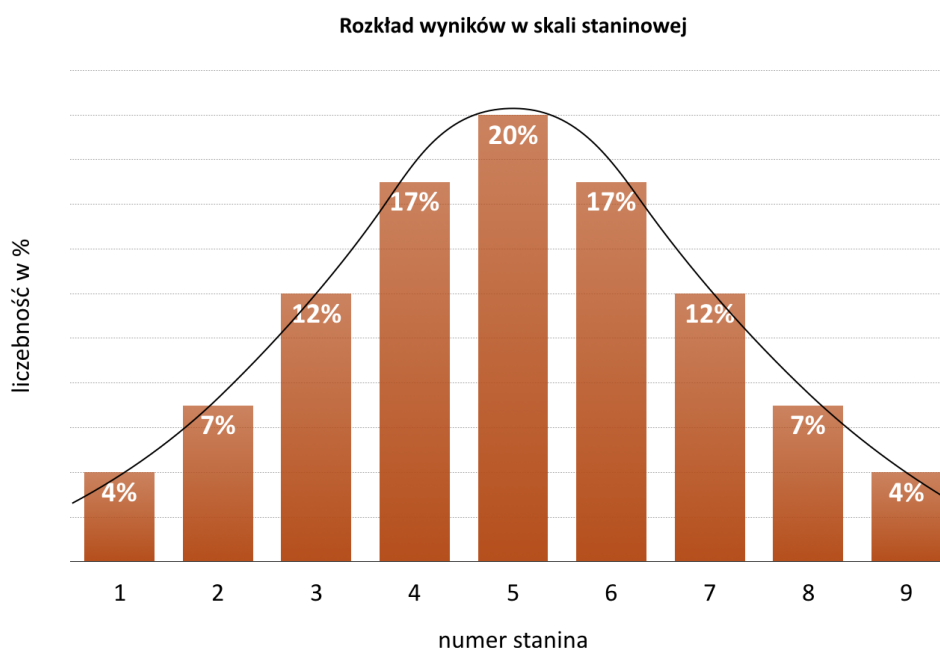
Procenty, centyle i skala staninowa

Skala staninowa

Skala staninowa to dziewięciostopniowa skala testu psychologicznego znormalizowana, gdzie średnia w populacji wynosiła 5, a odchylenie standardowe 2. (Niemierko, 1975: 149). Liczba jednostek skali wynosi 9 staninów, natomiast jej zakres wynosi od 1 do 9 stanina.

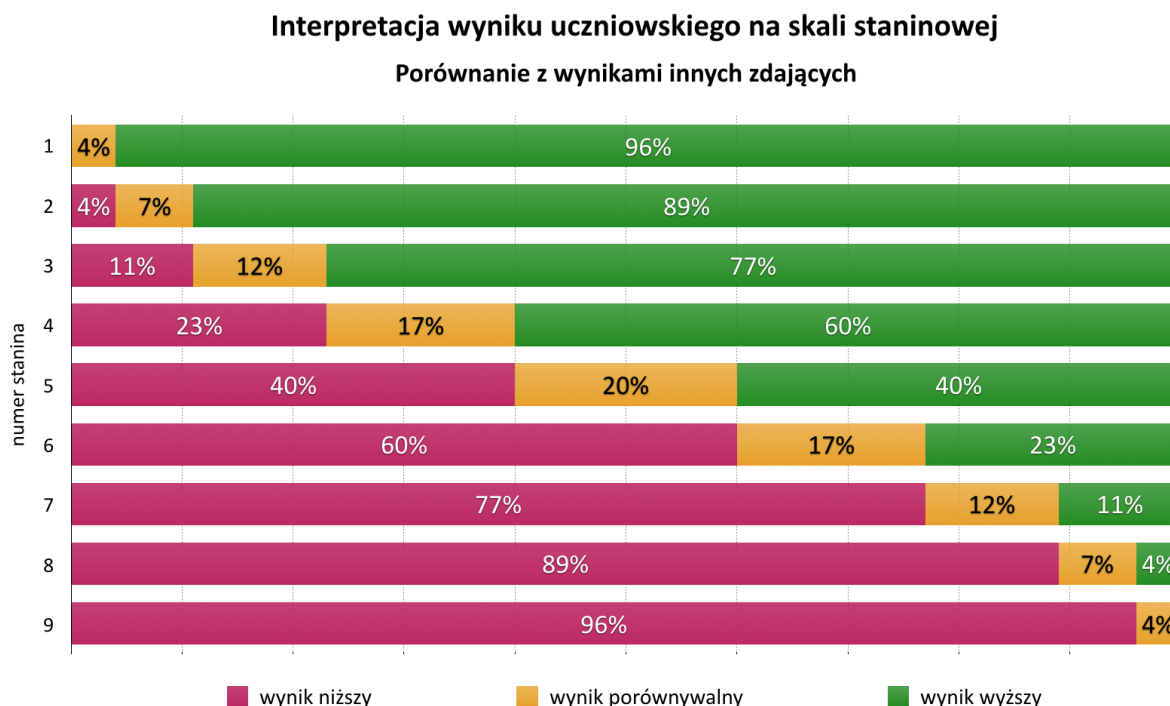
Od wielu lat Centralna Komisja Egzaminacyjna używa skali staninowej do interpretacji wyników egzaminów zewnętrznych. Także w praktyce szkolnej skala ta jest powszechnie stosowana.

Rozkład procentowy w poszczególnych staninach wynika ze znormalizowanego rozkładu normalnego.





Jak należy interpretować wynik staninowy?



Jeśli uczeń jest w staninie 5, oznacza to, że około 20% uczniów zdało egzamin na podobnym poziomie, 40% uczniów uzyskało wynik niższy, a 40% uczniów wynik wyższy.

Jeśli uczeń jest w staninie 8, oznacza to, że około 7% uczniów zdało egzamin na podobnym poziomie, 89% uczniów uzyskało wynik niższy, a 4% uczniów wynik wyższy.

Centyle i procenty

Wyniki z egzaminu gimnazjalnego lub matury otrzymujemy na dokumencie wyrażonym w dwóch postaciach:

- w procentach;
- w skali centylowej.

Wynik wyrażony **w procentach** to odsetek punktów uzyskany za rozwiązanie wszystkich zadań z arkusza egzaminacyjnego.

Przykład

Liczba możliwych do zdobycia punktów na egzaminie wynosiła 80, uczeń uzyskał 44 punkty. Stanowi to 55%.



Wynik centylowy to odsetek uczniów, którzy uzyskali wynik na tym samym poziomie lub niższy. Jeśli wynik centylowy wynosi 65, oznacza to, że 65% uczniów uzyskała taki sam wynik lub niższy. Dzięki temu wynik centylowy pozwala określić położenie ucznia na tle innych. Po raz pierwszy w 2012 r. przedstawiono wyniki egzaminów zewnętrznych z wykorzystaniem skali centylowej.

Można zauważyć, że skala centylowa i skala staninowa są do siebie podobne. Skala staninowa ma 9 poziomów, a skala centylowa 99. Jest zatem dokładniejsza.

Wyniki w raportach Centralnej Komisji Egzaminacyjnej i Okręgowej Komisji Egzaminacyjnej przedstawiane są poniżej:

Część matematyczno-przyrodnicza - matematyka		
wynik procentowy	wartość centyla	stanin
0	1	1
4	1	
7	2	
11	3	
14	7	2
18	11	
21	17	3
25	23	
29	28	4
32	34	
36	40	
39	45	
43	50	5
46	54	
50	59	
54	63	
57	67	6
61	71	
64	74	
68	78	
71	81	7
75	84	
79	87	
82	89	
86	92	9
89	94	
93	97	
96	99	10
100	100	



Sprawozdanie z egzaminów zewnętrznych

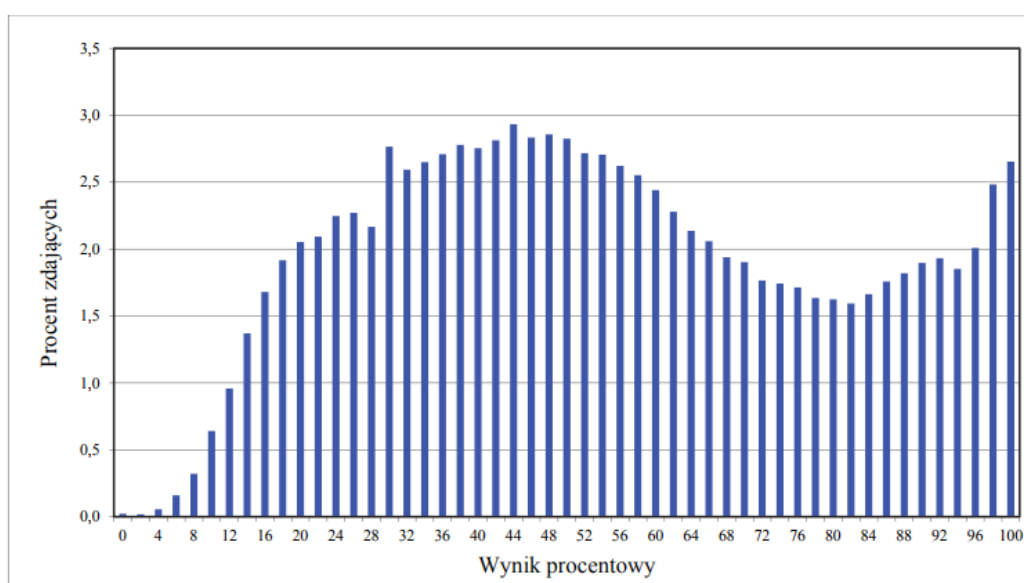
Co roku Centralna Komisja Egzaminacyjna publikuje raporty z przeprowadzonego egzaminu. Znajdują się one na stronie www.cke.edu.pl.

Jakie informacje zawiera raport?

Dane dotyczące populacji zdających

Liczba zdających		261407
Zdający rozwiązujący zadania w arkuszu standardowym	z liceów ogólnokształcących	167806
	z techników	93601
	ze szkół na wsi	10260
	ze szkół w miastach do 20 tys. mieszkańców	49656
	ze szkół w miastach od 20 tys. do 100 tys. mieszkańców	96275
	ze szkół w miastach powyżej 100 tys. mieszkańców	105216
	ze szkół publicznych	242095
	ze szkół niepublicznych	19312
	kobiety	143093
	mężczyźni	118314
	bez dysleksji rozwojowej	240289
	z dysleksją rozwojową	21118

Rozkład wyników



Źródło: www.cke.edu.pl

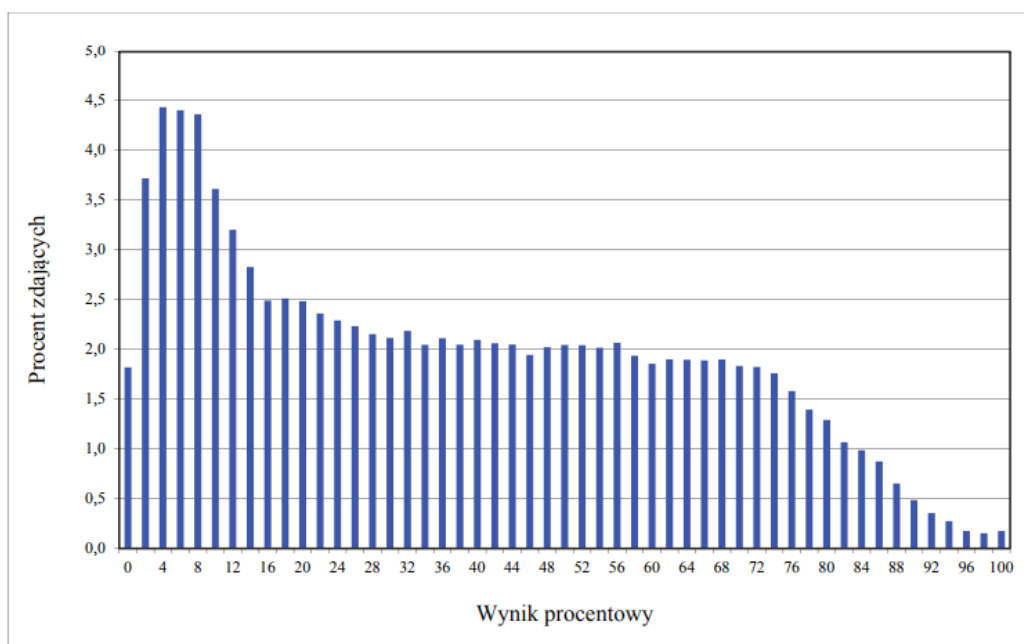


Na osi X prezentowany jest wynik procentowy ucznia. Na osi Y procent uczniów, którzy uzyskali dany wynik. Na tej podstawie tworzony jest wynik centylowy. Powyższy wynik zawiera informacje z przebiegu egzaminu maturalnego z matematyki na poziomie podstawowym w roku 2017.

Wyniki z egzaminu maturalnego z matematyki mają dość często tzw. rozkład dwumodalny. Jest to rozkład, w przypadku którego występują dwie wartości modalne. Znajdują się dwa wyraźne maksima na wykresie funkcji rozkładu.

Jakie są przyczyny dwumodalności rozkładu? Jeśli dana zmienna reprezentuje rozbieżność „opinii”. W naszym przypadku może to świadczyć o występowaniu dwóch grup uczniów, reprezentujących różne poziomy w zakresie matematyki. Prawdopodobnie jest to

też związane z tym, że wszyscy uczniowie uczą się w szkole matematyki na poziomie podstawowym, ale jest też grupa, która uczy się na poziomie rozszerzonym. Poniżej na wykresie przedstawiono charakterystykę rozkładu wyników z matury dla grupy uczniów zdających maturę rozszerzoną.



Źródło: www.cke.edu.pl



Podstawowe parametry statystyczne

Zdający	Liczba zdających	Minimum (%)	Maksimum (%)	Mediana (%)	Modalna (%)	Średnia (%)	Odchylenie	Odsetek sukcesów
ogółem	261 407	0	100	52	44	54	25	83
w tym								
z liceów	167 806	0	100	60	100	60	25	87
z techników	93 601	0	100	42	30	45	21	75
bez dysleksji rozwojowej	240 289	0	100	52	44	54	25	83
z dysleksją rozwojową	21 118	0	100	54	44	56	24	85

Źródło: www.cke.edu.pl

Poziom wykonania poszczególnych zadań

Współczynnik łatwości (rozwiązywalności) zadania obliczamy, dzieląc wynik uzyskany przez uczniów klasy, szkoły, powiatu przez maksymalną liczbę punktów do uzyskania za zadanie.

Podobnie definiujemy łatwość testu. Współczynnik łatwości (rozwiązywalności) testu obliczamy, dzieląc uzyskany wynik z testu ucznia, klasy, szkoły, powiatu przez maksymalną liczbę punktów do uzyskania za test.

Test, zadanie	łatwość
Bardzo trudny	0,00-0,19
Trudny	0,20-0,49
Umiarkowanie trudny	0,50-0,69
łatwy	0,70-0,89
bardzo łatwy	0,90-1,00

**Przykład**

Rozwiązywalność zadań na maturze z matematyki, poziom podstawowy, rok 2017.

Numer zadania	Wymaganie ogólne	Wymaganie szczegółowe	Poziom wykonania zadania (%)
1.	II. Wykorzystanie i interpretowanie reprezentacji.	1. Liczby rzeczywiste. Zdający oblicza potęgi o wykładnikach wymiernych i stosuje prawa działań na potęgach o wykładnikach wymiernych (1.4).	74
2.	II. Wykorzystanie i interpretowanie reprezentacji.	1. Liczby rzeczywiste. Zdający posługuje się w obliczeniach pierwiastkami dowolnego stopnia i stosuje prawa działań na pierwiastkach (1.3).	43
3.	II. Wykorzystanie i interpretowanie reprezentacji.	1. Liczby rzeczywiste. Zdający wykorzystuje definicję logarytmu i stosuje w obliczeniach wzory na logarytm iloczynu, logarytm ilorazu i logarytm potęgi o wykładniku naturalnym (1.6).	72

Źródło: www.cke.edu.pl

Zadanie 17

W ciągu arytmetycznym $a_1 = 5$, $a_2 = 11$. Wybierz prawidłową odpowiedź.

- a) $a_{14} = 71$
- b) $a_{12} = 71$
- c) $a_{11} = 71$
- d) $a_{10} = 71$

Odpowiedź: b)

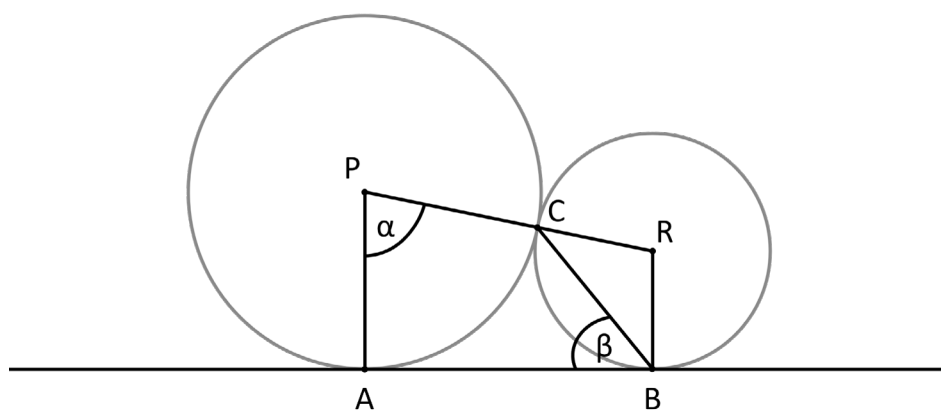
To zadanie było rozwiązywane na poziomie 93%.

Najtrudniejsze zadanie było z geometrii. Poziom rozwiązywalności wynosił 21%.



Zadanie 28 (2 pkt)

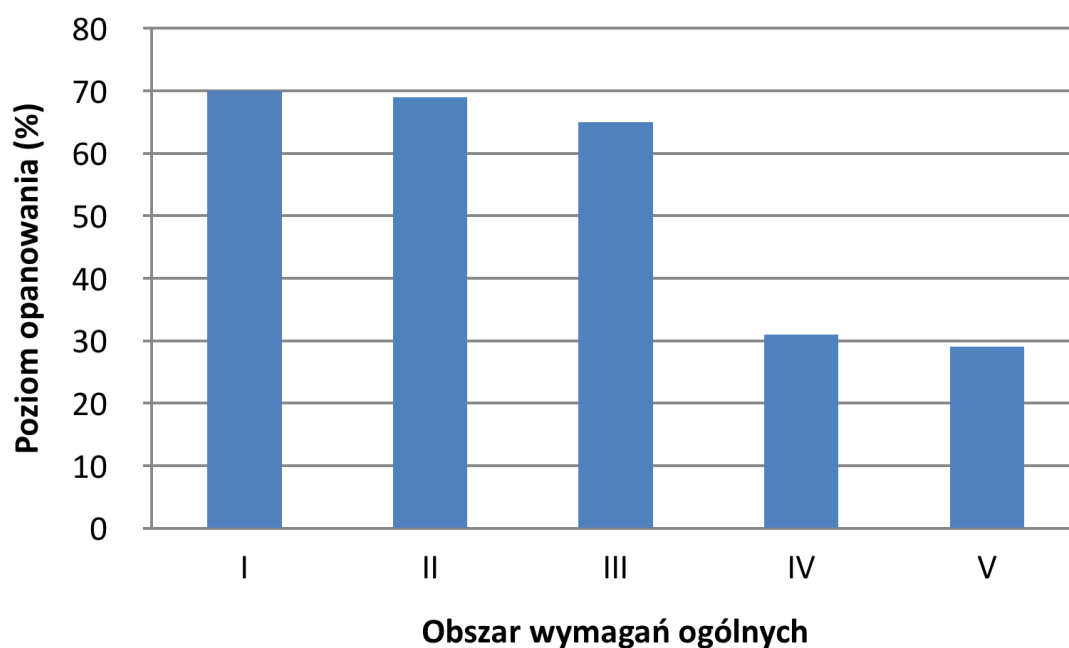
Dane są dwa okręgi o środkach w punktach P i R , styczne zewnętrznie w punkcie C . Prosta AB jest styczna do obu okręgów odpowiednio w punktach A i B oraz $|\sphericalangle APC| = \alpha$ i $|\sphericalangle ABC| = \beta$ (zobacz rysunek). Wykaż, że $\alpha = 180^\circ - 2\beta$.



Poziom wymagań ogólnych

Wymagania ogólne podzielone są na 5 obszarów:

- I. Wykorzystanie i tworzenie informacji;
- II. Wykorzystanie i interpretowanie reprezentacji;
- III. Modelowanie matematyczne;
- IV. Użycie i tworzenie strategii;
- V. Rozumowanie i argumentowanie.



Źródło: www.cke.edu.pl



Wizualizacja danych

Celem wizualizacji danych jest pokazanie informacji w sposób czytelny dla odbiorcy, jednocześnie pozwalający na ich dokładne i efektywne zrozumienie, ale również na ich analizę. Prezentacja graficzna jest dla nas bardziej przystępna. Dzieje się tak dlatego, że ludzie dobrze „skanują”, rozpoznają i zapamiętują przedstawione im obrazy (np. kształt, długość, budowę itp.). Za sprawą wizualizacji możemy łączyć wielkie zbiory danych i pokazać wszystkie informacje jednocześnie w przystępny sposób, co znacznie ułatwia ich analizę. Możemy również stosować porównania wizualne, dzięki którym dużo łatwiej stwierdzić wiele faktów. Kolejną zaletą wizualizacji danych jest możliwość ich analizy na kilku poziomach szczegółowości.

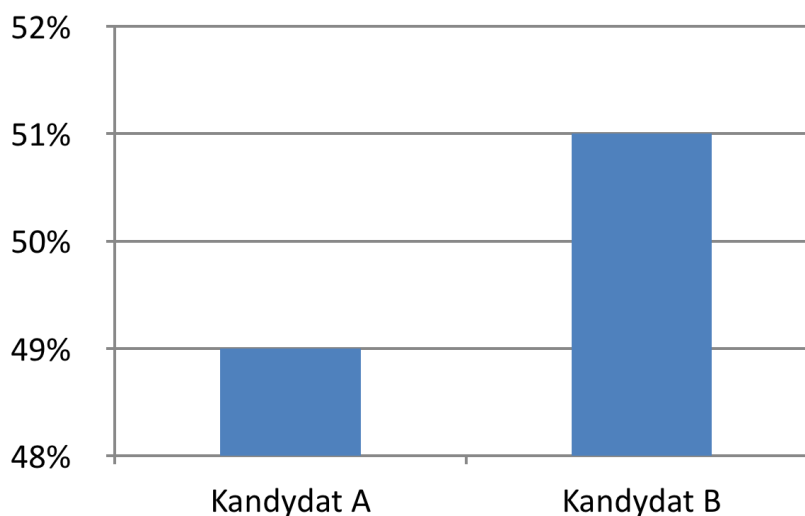
Manipulacje danymi

Używając tych samych danych, możemy je przedstawić na różne sposoby. Manipulacja to kształtowanie poglądów, postaw i zachowań lub emocji bez wiedzy i woli człowieka (Wroński, 2010). Różnorodność sposobów prezentacji danych niekiedy przyczynia się do powstawania manipulacji. Poniżej prezentujemy kilka przykładów wykorzystania różnych prezentacji danych do manipulowania odbiorcą.

Wykres słupkowy

Jak przedstawić korzystniej wyniki naszego kandydata? Wystarczy np. zmienić skalę procentową.

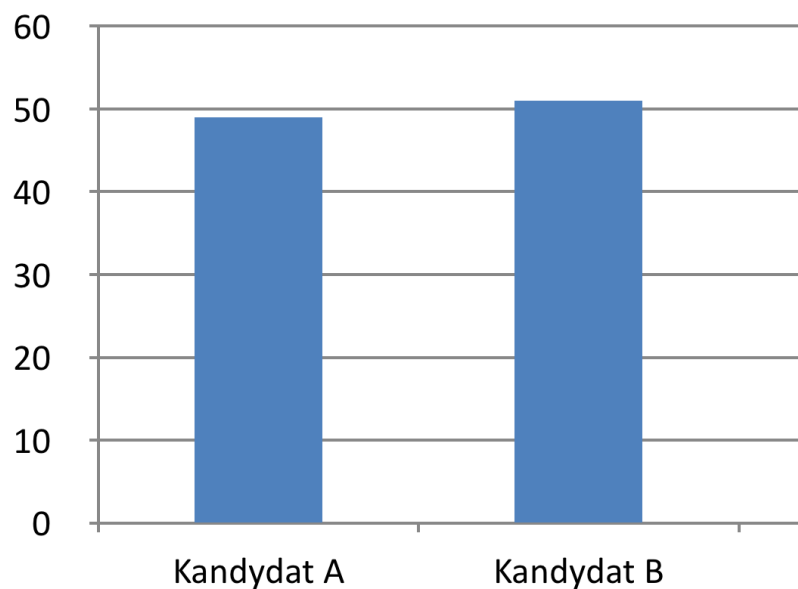
Jeśli naszym kandydatem jest B, to wykres powinien wyglądać jak na rysunku poniżej. Wynika z niego, że kandydat B jest „wyraźnie” lepszy od kandydata A.



Jeśli natomiast naszym kandydatem jest A, to wykres powinien wyglądać jak na rysunku poniżej.



Widać, że różnica między kandydatem A i B jest niewielka. Dodając informację, że błąd statystyczny jest na poziomie 3%, możemy użyć sformułowania, że idą łeb w łeb.



Wykres kołowy

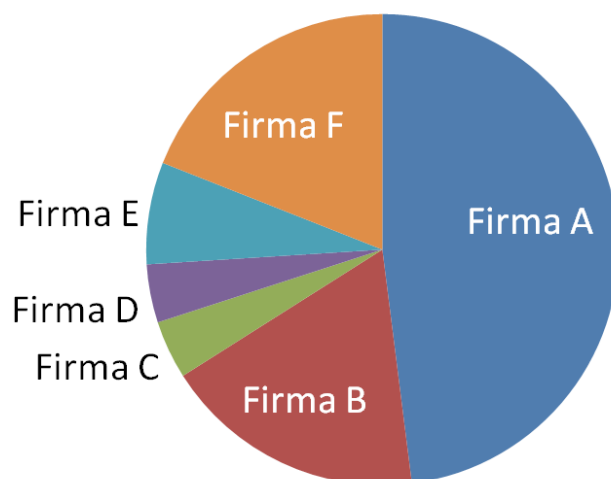
W tabeli znajduje się informacja o sprzedaży produktów firmy A, B, C, D, E.

Firma A	48%
Firma B	18%
Firma C	4%
Firma D	4%
Firma E	7%
Firma F	19%

Założmy, że jesteśmy firmą B.

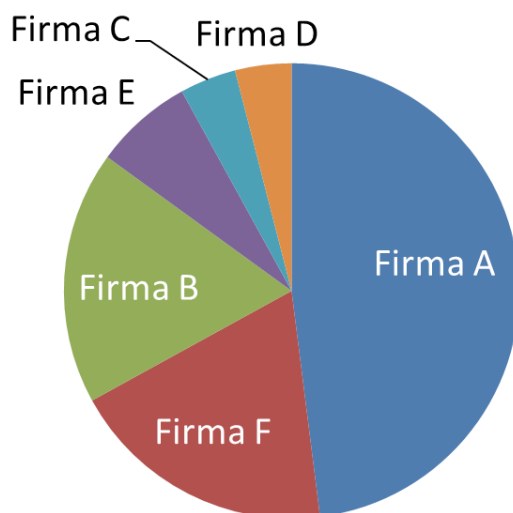


Udział w rynku



Oczywiście firma A jest najlepsza i firma B nie będzie z nią w tym momencie konkurować. Między firmą B i F jest niewielka różnica. Chcemy pokazać, że firma B jest „lepsza” od firmy F. Patrząc na rysunek poniżej, odnosimy wrażenie, że jest to „prawda”.

Udział w rynku



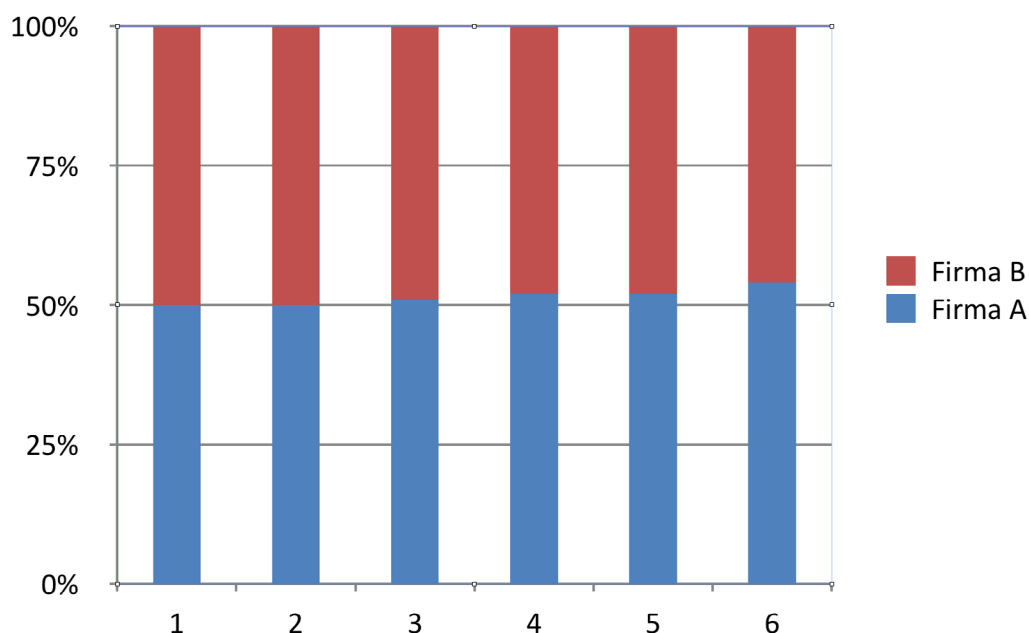
Z jednej strony wizualizacja ułatwia czytanie skomplikowanych danych, z drugiej jednak strony ilustracja powyżej pokazuje, jak wizualizacja danych może być wykorzystywana nieprawidłowo.



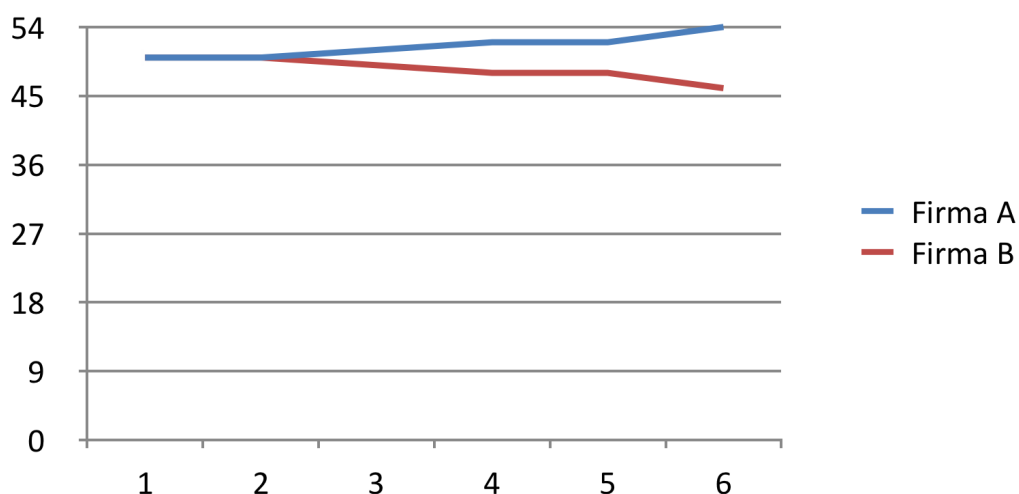
Linia trendu

Zebrano dane sprzedaży firmy A i B w ciągu ostatnich sześciu miesięcy. Jednym z elementów porównania jest tzw. linia trendu.

Jeśli jesteśmy właścicielami firmy A i chcemy ukryć spadek sprzedaży w porównaniu z firmą B, to informację przedstawiamy w formie słupkowej. Wówczas informacja brzmi: firma A i B utrzymują sprzedaż.



Jeśli jesteśmy właścicielami firmy B i chcemy wyraźnie przedstawić informację, że firma A zanotowała spadek sprzedaży w porównaniu z firmą B, to informację przedstawiamy w formie liniowej.

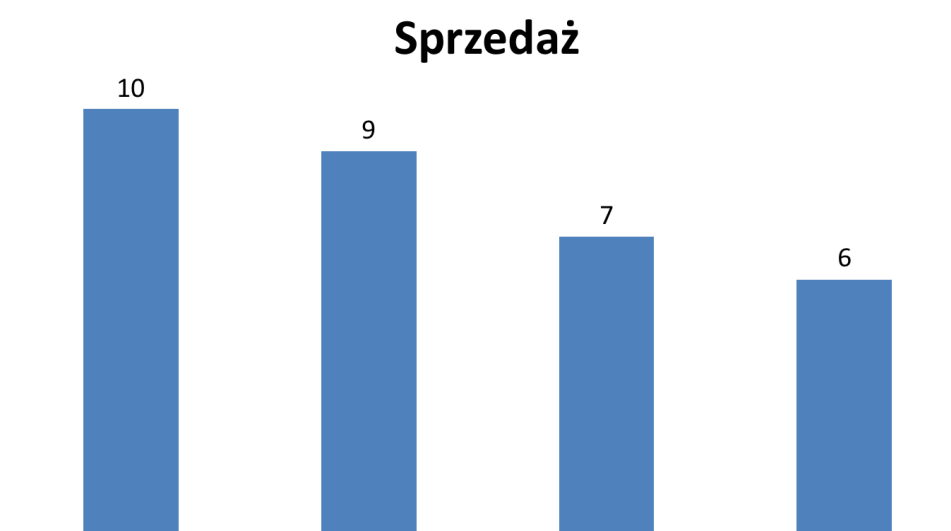




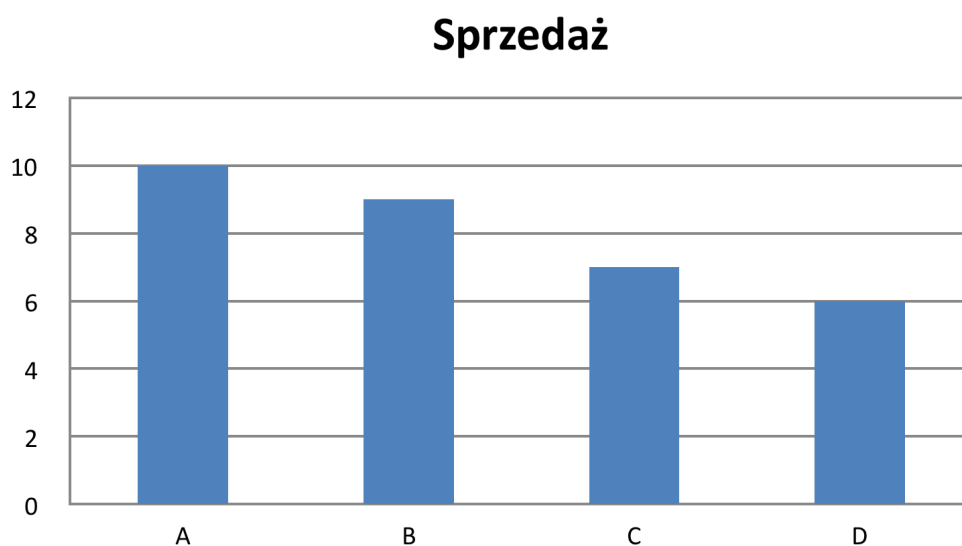
Poniżej przedstawiamy wzmocnienie prezentowanej różnicy między najlepszą firmą i firmą osiągnącą podobne wyniki.

Naszym celem jest wyraźne pokazanie, że różnica między firmą A i B jest „duża”. Jak tego dokonać?

a) Tak wygląda standardowa prezentacja danych. Tu różnica nie wydaje się duża.

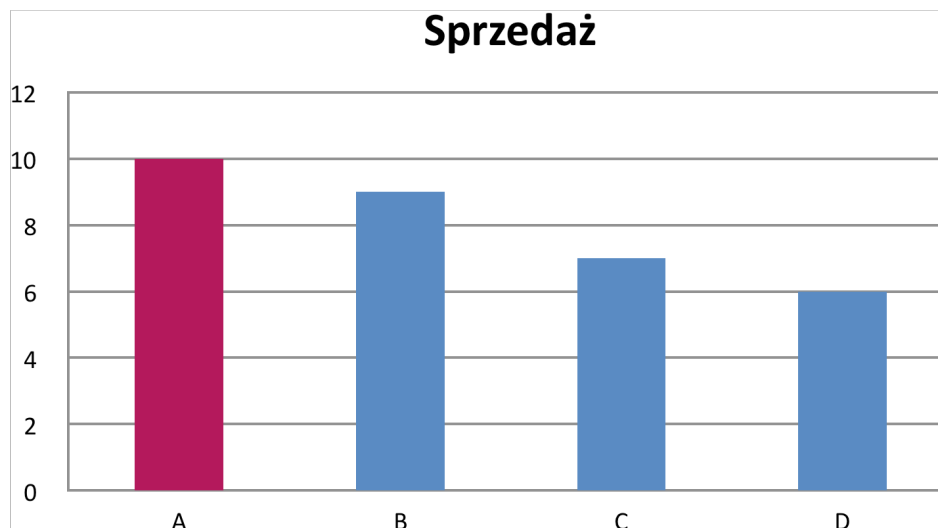


b) Wprowadzenie linii i wydłużenie wykresu sprawia, że różnica staje się bardziej widoczna.





c) Różnicę między firmami możemy podkreślić, wprowadzając manipulację kolorami.



Jeśli porównamy prezentowane wykresy, to możemy stwierdzić, że w ostatnim wykresie różnica między firmą A i B jest widoczna najwyraźniej.

Jak przygotować ucznia do prawidłowego czytania i analizowania danych?

W codziennym życiu często używamy przybliżeń. Zamiast rzeczywistej wartości bardziej akceptowalne jest jej przybliżenie. Oto przykład: Uczeń w sklepie za zakupy zapłacił 39,83 zł. Jeśli mama zapytała go, ile wydał pieniędzy, to odpowiedź 40 zł jest do zaakceptowania. Ale jeśli powie, że to było 50 zł, to raczej zgodzimy się, że ta kwota nie jest do zaakceptowania.

Analizując dane, nauczyciel powinien w sposób naturalny wprowadzić definicję błędu bezwzględnego i błędu względnego.

Błąd bezwzględny to moduł różnicy pomiędzy wartością dokładną a wartością przybliżoną, $E_b = |x - y|$.

$$E_b = |x - y| = |39,83 - 40| = 0,17$$

x – wartość dokładna

y – wartość przybliżona

E_w – błąd względny

Błąd względny to iloraz błędu bezwzględnego: $E_w = \frac{|x - y|}{x}$.

$$E_w = \frac{|x - y|}{x} \cdot 100\% = \frac{|39,83 - 40|}{39,83} \cdot 100\% = 0,004268$$



x – wartość dokładna

y – wartość przybliżona

E_w – błąd względny Czasami błąd względny wyraża się w procentach, czyli:

$$E_w = \frac{|x - y|}{x} \cdot 100\% = \frac{|39,83 - 40|}{39,83} \cdot 100\% = 0,004268 \cdot 100\% = 0,4268\%$$

Jeśli za przybliżenie przyjmiemy 50 zł, to błąd będzie równy:

$$E_w = \frac{|x - y|}{x} \cdot 100\% = \frac{|39,83 - 50|}{39,83} \cdot 100\% = 0,2553 \cdot 100\% = 25,53\%$$

Edukacyjna wartość dodana

Definicja, kryteria

Edukacyjna wartość dodana jako metoda oceny pracy szkoły była rozwijana w Polsce od 2005 r., a zyskała na popularności w ostatnich kilku latach wraz z upowszechnieniem ewaluacji jako formy nadzoru pedagogicznego. Metoda pozwala szacować nie tylko efekty kształcenia, ale przede wszystkim efektywność nauczania w zakresie sprawdzanym egzaminem zewnętrznym.

Edukacyjna wartość dodana (EWD) to zestaw technik statystycznych pozwalających określić wkład szkoły w wyniki nauczania. Na wyniki egzaminacyjne uczniów wpływa wiele czynników niezależnych od szkoły, dlatego ważne, by miara wkładu szkoły była od tych czynników wolna. Uwzględniając uprzednie osiągnięcia uczniów, można zbudować miarę postępu edukacyjnego uczniów, która jest jednocześnie dobrą miarą jakości pracy szkoły (trafność).

Publicznie udostępnione są trzyletnie wskaźniki [EWD](#) dla 6 tysięcy gimnazjów i ponad 4 tysięcy liceów i techników. Pozwalają one śledzić zmiany w czasie oraz dokonywać porównań szkół ze względu na efekty i efektywność nauczania w zakresie sprawdzanym egzaminem zewnętrznym. Do pogłębionych analiz wewnątrzszkolnych służą jednoroczne wskaźniki EWD wyliczane przez szkoły z użyciem corocznie aktualizowanego Kalkulatora EWD 100.

Wiele lat temu ranking szkół i ich ocenę dokonywano na podstawie liczby olimpijczyków. Na przykład w 2005 r. zaproponowano ranking szkół, w którym decydowała liczba laureatów i finalistów olimpiad przedmiotowych.



Od wielu lat miesięcznik „Perspektywy” każdego roku przygotowuje ranking szkół średnich. Metodologia Ogólnopolskiego Rankingu Szkół Ponadgimnazjalnych postanowieniem Kapituły ocenia w Polsce szkoły za pomocą czterech kryteriów:

- sukcesy szkoły w olimpiadach przedmiotowych i zawodowych;
- wyniki matury z przedmiotów obowiązkowych;
- wyniki matury z przedmiotów dodatkowych;
- ocena szkoły przez kadrę akademicką.

Do sporządzenia danych o szkołach wykorzystuje się informacje od komitetów głównych olimpiad przedmiotowych i okręgowych komisji egzaminacyjnych oraz badanie ankietowe przeprowadzone przez Fundację Edukacyjną Perspektywy wśród kadry akademickiej w całej Polsce.

Uproszczony model edukacyjnej wartości dodanej

Obecnie stosowane w Polsce modele edukacyjnej wartości dodanej wykorzystują zaawansowane metody statystyczne (Żółtak, 2015). Naszym celem nie jest jednak przedstawienie najnowszych modeli EWD, a jedynie odniesienie się do idei EWD jako przykładu analizy wyników egzaminacyjnych, dlatego odwołamy się tutaj do uproszczonego jednorocznego modelu edukacyjnej wartości dodanej. Badamy związek wyniku np. egzaminu gimnazjalnego z części matematyczno-przyrodniczej i matury z matematyki, stosując metodę regresji. W praktyce sprowadza się to do zbudowania odpowiedniej funkcji. Funkcja może być opisana za pomocą tabeli, wzoru, algorytmu lub sieci neuronowej. Taką funkcję będziemy nazywać też modelem edukacyjnej wartości dodanej. Im większą pulą danych dysponujemy, tym precyzyjniej można taką funkcję wyznaczyć. Przykładowo, maturalne wskaźniki EWD budowane są na bazie wyników ponad 300 tysięcy maturzystów z danego rocznika.

- Mając zbudowaną funkcję, na podstawie wyniku ucznia na egzaminie gimnazjalnym z części matematyczno-przyrodniczej obliczamy przewidywany z modelu (statystycznie najbardziej prawdopodobny) wynik z egzaminu maturalnego z matematyki.
- Następnie obliczamy różnicę punktów między rzeczywistym wynikiem ucznia na egzaminie a jego wynikiem przewidywanym na podstawie funkcji regresji (inaczej na podstawie modelu EWD).
- Dla danej szkoły uśredniamy wszystkie różnice i tę wartość nazywamy EWD szkoły.

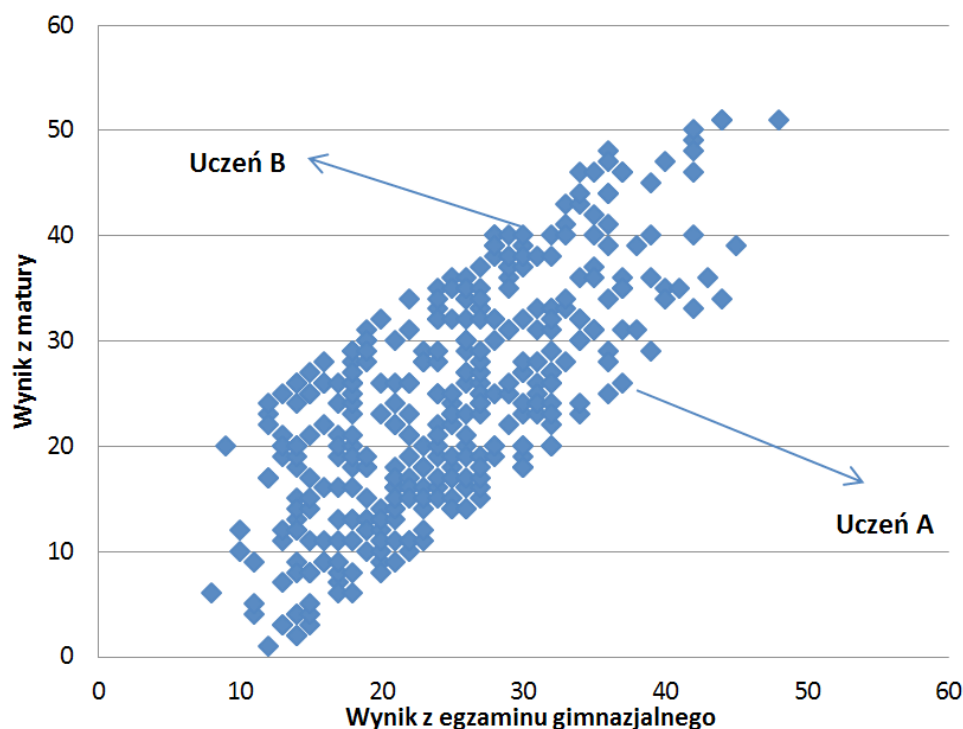
Przykład 1

Prosta metoda obliczania EWD.

- a) Dane wejściowe:
Wynik z egzaminu gimnazjalnego z części matematyczno-przyrodniczej, wynik z matury z matematyki.



- b) Wyznamy funkcję regresji – możemy w tym celu użyć programu Excel. Kolumna A – wynik z części matematyczno-przyrodniczej, kolumna B – liczba punktów z matury.



Obliczenia

Uczeń A i uczeń B otrzymali na egzaminie gimnazjalnym w części matematyczno-przyrodniczej po 30 punktów. Przewidywany wynik dla tych uczniów to 27,5 punktu ($y = 0,98 \times 30 - 1,89 \approx 27,5$). Uczeń A otrzymał na maturze z matematyki 19 punktów, czyli różnica między wynikiem uzyskanym i przewidywanym (jego EWD) wynosi minus 8,5 punktu. Uczeń B otrzymał 42 punkty, zatem w tym przypadku różnica między wynikiem uzyskanym a wynikiem przewidywanym wynosi plus 14,5 punktu.

Ćwiczenie – EWD klasy

O edukacyjnej wartości dodanej mówimy w kontekście grupy uczniowskiej, a nie pojedynczego ucznia. Dodatnie EWD dla grupy oznacza, że większość uczniów tej grupy uzyskuje wyniki powyżej przeciętnej dla innych maturzystów o takich samych zasobach na wejściu (uprzednie osiągnięcia na egzaminie gimnazjalnym). Ujemne EWD dla grupy wskazuje na to, że w tej grupie udało się uzyskać mniej niż w innych grupach z uczniami o podobnych zasobach na wejściu. EWD zerowe świadczy o przeciętnej efektywności nauczania.



Obliczenia

Linia regresji zadana jest wzorem $y = 0,98x - 1,89$. Uzupełnij tabelę i oblicz EWD klasy jako średnią z różnic dla poszczególnych uczniów.

Uczeń	Wynik GMP	Matura	Wynik przewidywany	Różnica
U1	30	19	27,5	-8,5
U2	30	42	27,5	14,5
U3	40			
U4	20			
U5	50			
U6	30			
U7	40			
U8	25			
U9	50			
U10	40			
EWD klasy				

Linia regresji w jednorocznych modelach EWD nie jest zależnością liniową głównie ze względu na nieliniowy charakter zależności w strefie bardzo niskich i bardzo wysokich wyników (Żółtak, Rappe, 2015: 5). Co roku linię przewidywanego wyniku dopasowuje się do realnych wyników maturzystów i implementuje do Kalkulatora EWD 100 (<http://ewd.edu.pl/pobierz>). Kalkulator EWD 100 jest narzędziem, które pozwala analizować wyniki egzaminacyjne z wykorzystaniem metody EWD. Dane przedstawiane są na różnorodnych wykresach (wykresy rozrzutu, wykresy z przedziałami ufności, rozkłady wyników, wykresy skumulowane procentowe) i tabelach, w podziale na grupy zdefiniowane przez użytkownika, co pozwala prowadzić eksploracyjną analizę danych.

Ćwiczenie

Przeanalizuj wyniki maturalne z matematyki z lat 2015, 2016, 2017 dla pewnego liceum. Co wzbudziło twoją ciekawość?

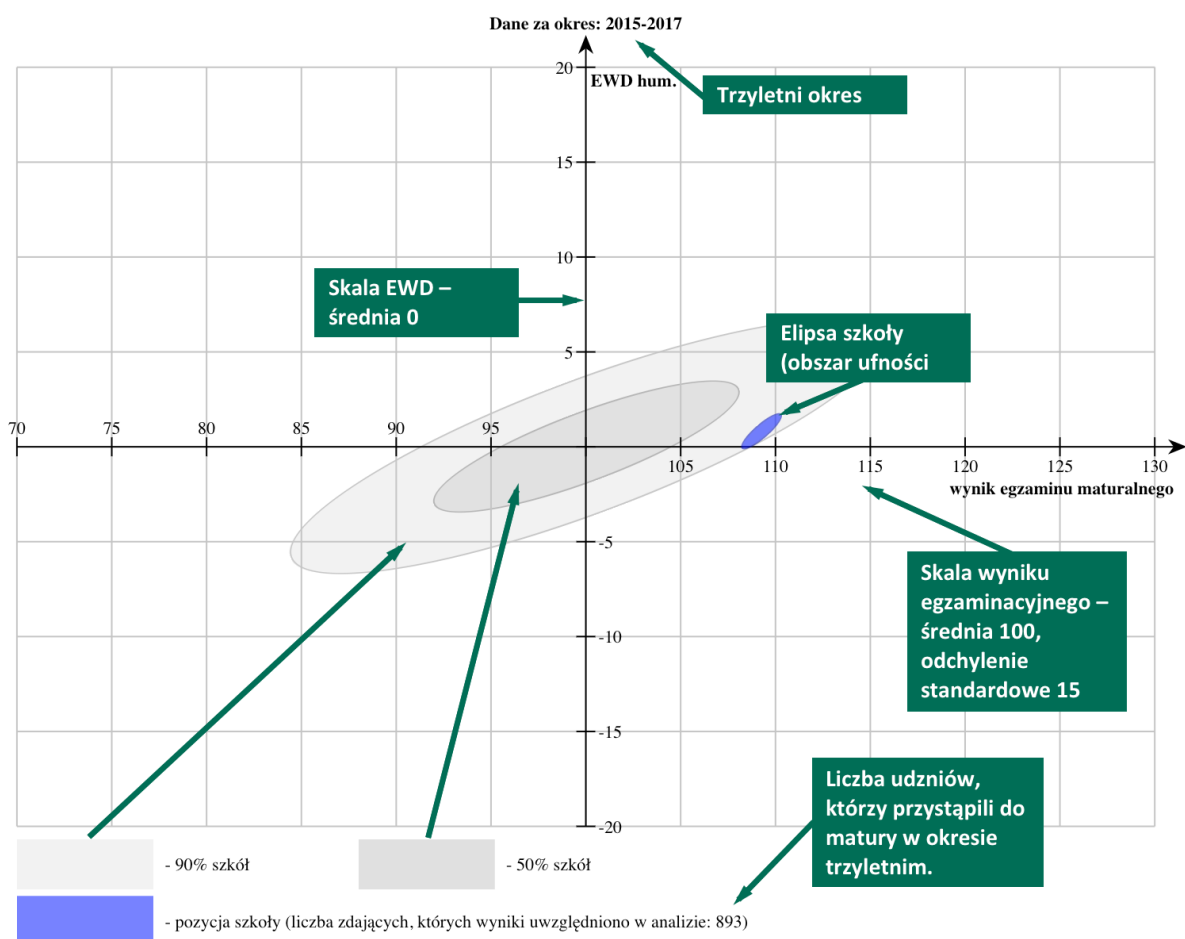
Przykładowe pytania pomocnicze:

- Jak zmienia się EWD szkoły w poszczególnych latach?
- Jak wyjaśnisz zróżnicowanie EWD w oddziałach klasowych?
- Czy zmienia się jakościowo (ze względu na uprzednie osiągnięcia) nabór do szkoły?
- Czy w szkole uczy się więcej chłopców, czy dziewcząt? Która grupa uzyskuje wyższe wyniki, a która wyższe EWD?

Trzyletnie wskaźniki EWD – przykład wizualizacji danych egzaminacyjnych

Wybór sposobu prezentacji danych powinien być podporządkowany celowi, w jakim te dane będą wykorzystywane. Trzyletnie wskaźniki EWD zostały pomyślane jako wskaźniki ewaluacyjne, stosowane do ewaluacji zewnętrznej prowadzonej przez nadzór pedagogiczny, a także do wykorzystania przez organy prowadzące do oceny pracy szkół.

Zatem sposób prezentacji powinien zawierać jak najwięcej informacji i dawać możliwość poprawnego statystycznie porównywania szkół. Zespół pracujący nad modelami EWD w Polsce zdecydował się na graficzną formę prezentacji danych, pokazując w układzie współrzędnych zarówno wynik egzaminacyjny, jak i EWD szkół. Oba wyniki podane są z przedziałem ufności (wynik szkoły na wykresie jest obszarem ufności – elipsą). Poszczególne elementy prezentacji trzyletnich wskaźników pokazuje rysunek poniżej.

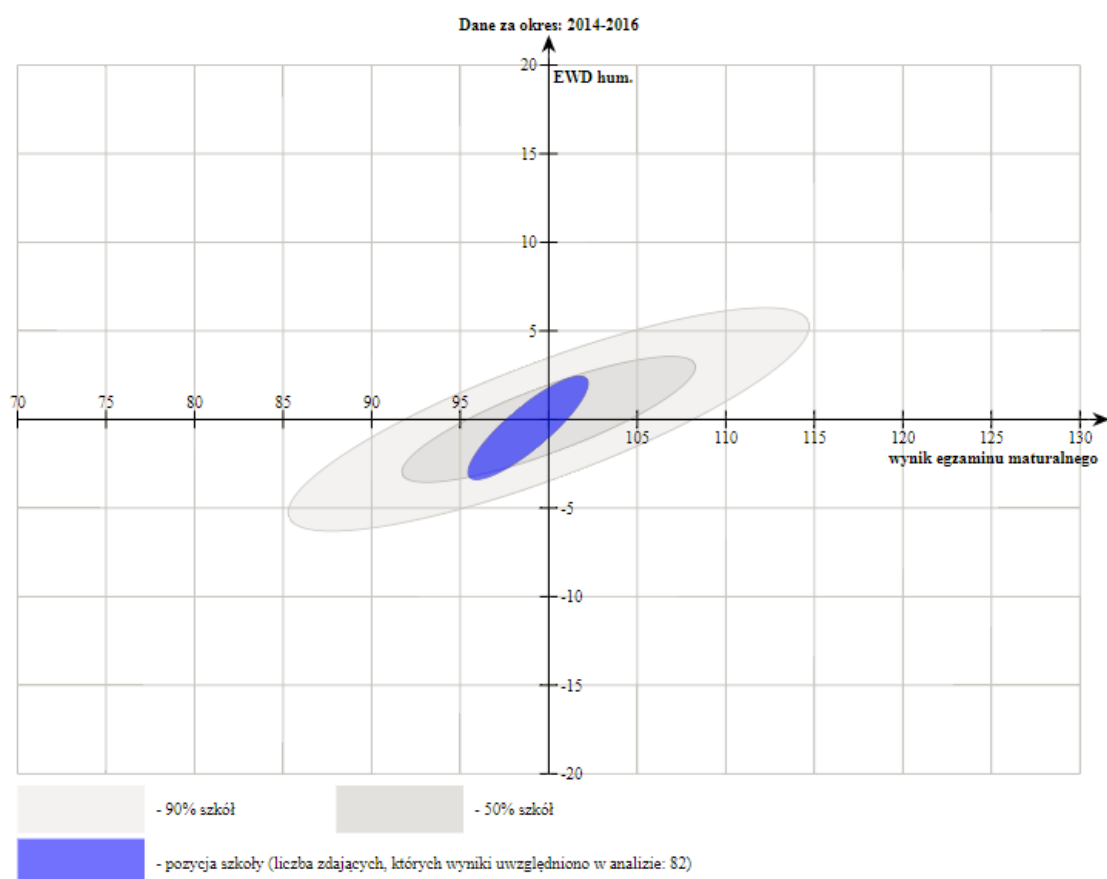


Źródło: Żółtak, Rappe, 2015



Ze względu na położenie elipsy w układzie współrzędnych możemy wyróżnić:

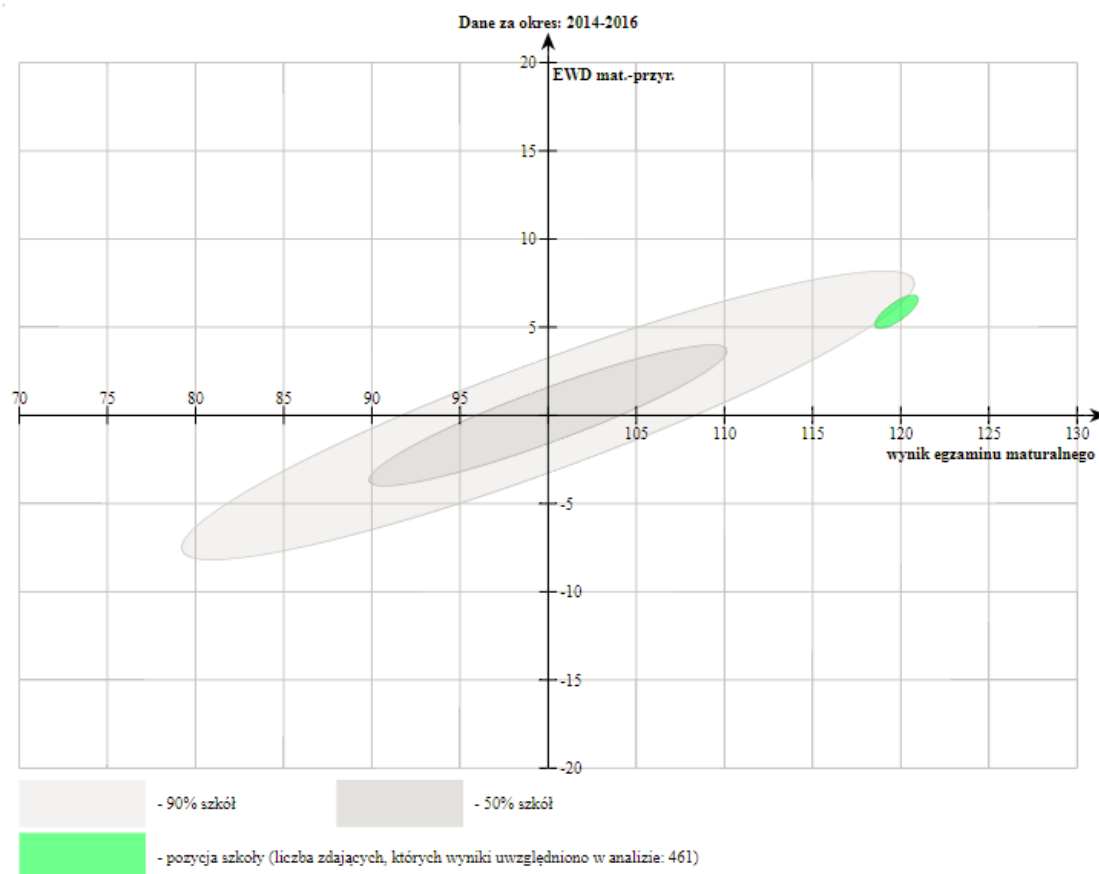
- Szkoły, w których notujemy zarówno średni w skali kraju poziom wyników egzaminacyjnych, jak i przeciętną efektywność.



Źródło: www.ewd.edu.pl



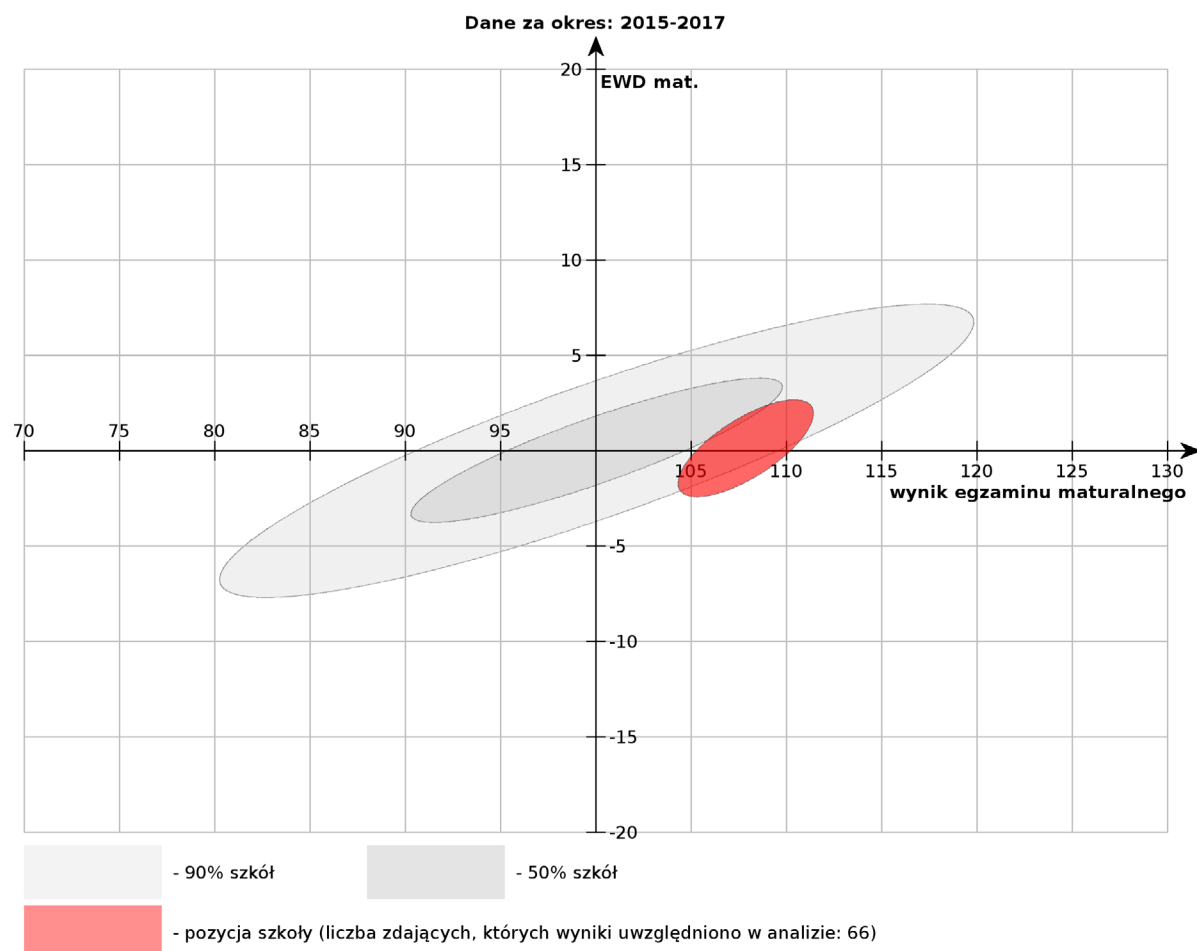
- Szkoły o wysokich wynikach egzaminacyjnych i wysokiej efektywności nauczania.



Źródło: www.ewd.edu.pl

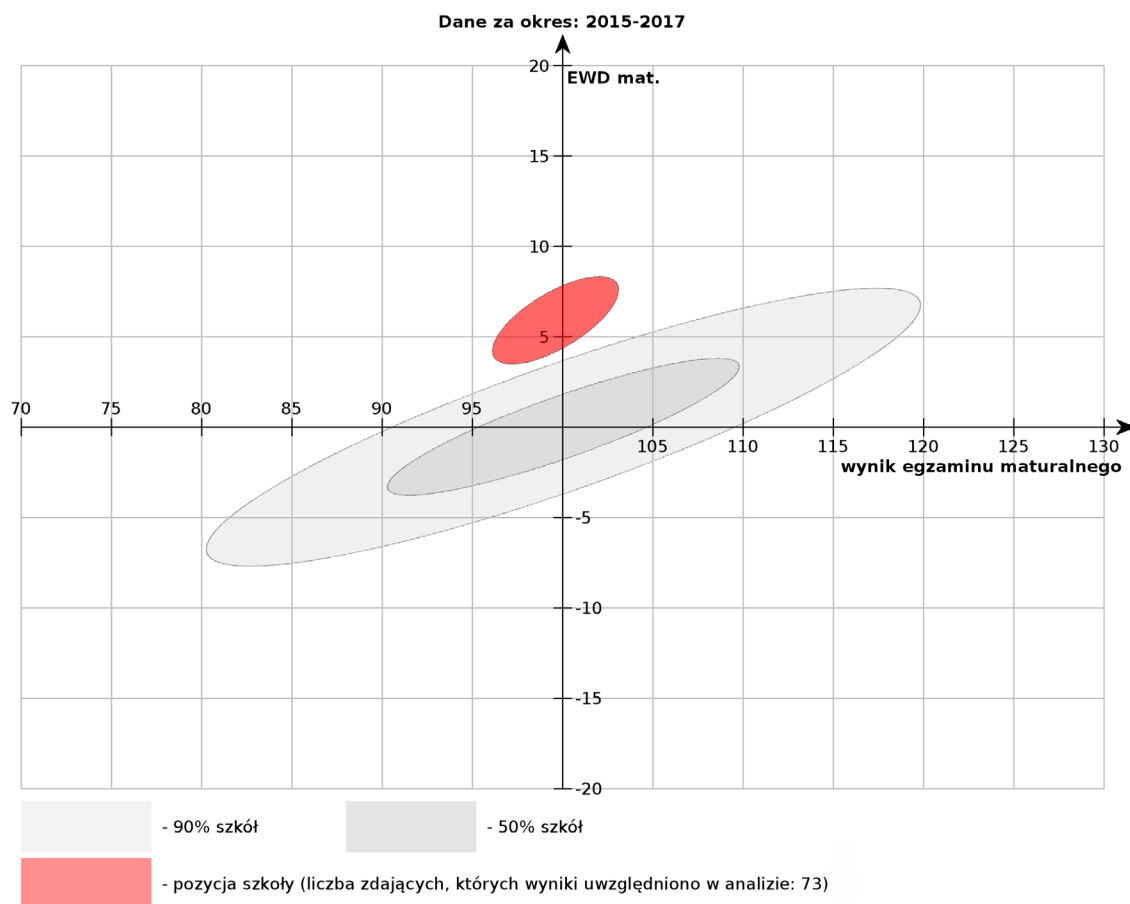


- Szkoły o przeciętnej efektywności nauczania mierzonej EWD i wysokich wynikach egzaminacyjnych.



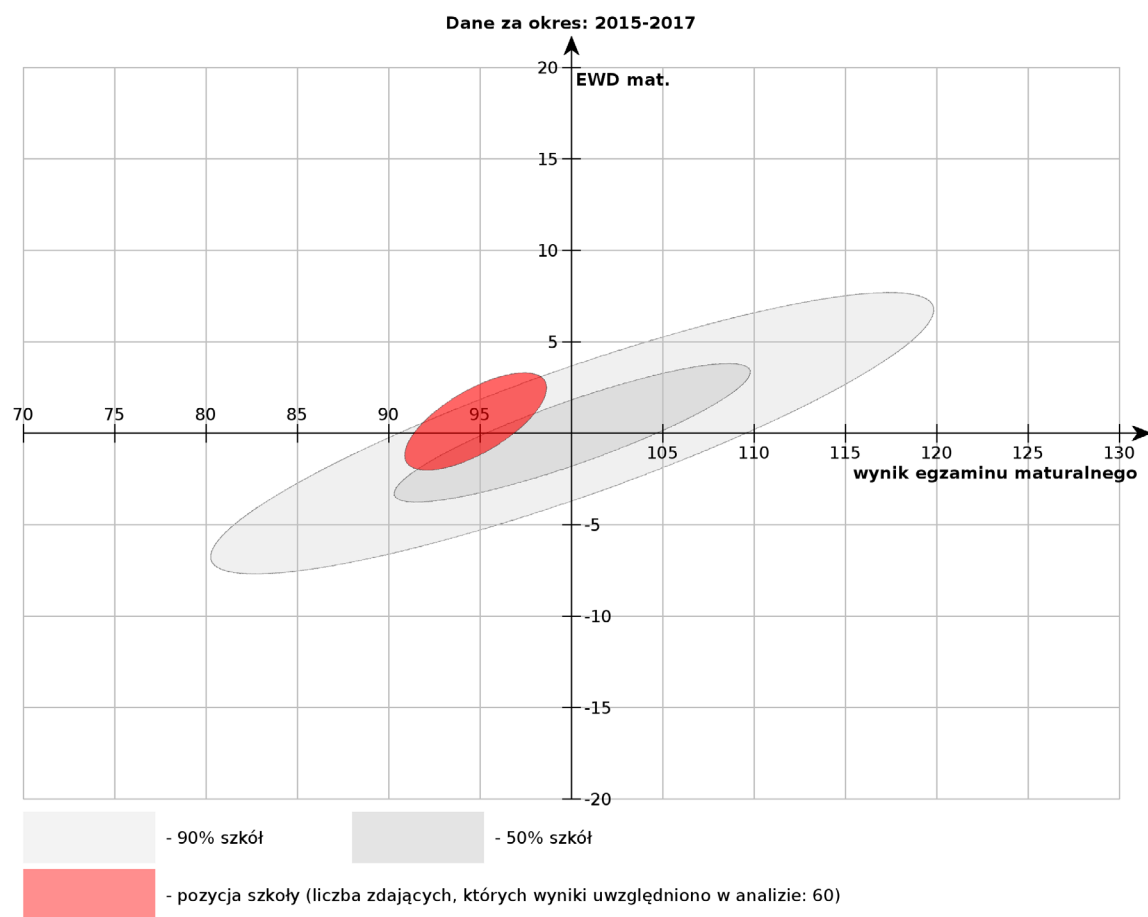


- Szkoły o niskich wynikach egzaminacyjnych, ale wysokiej efektywności.

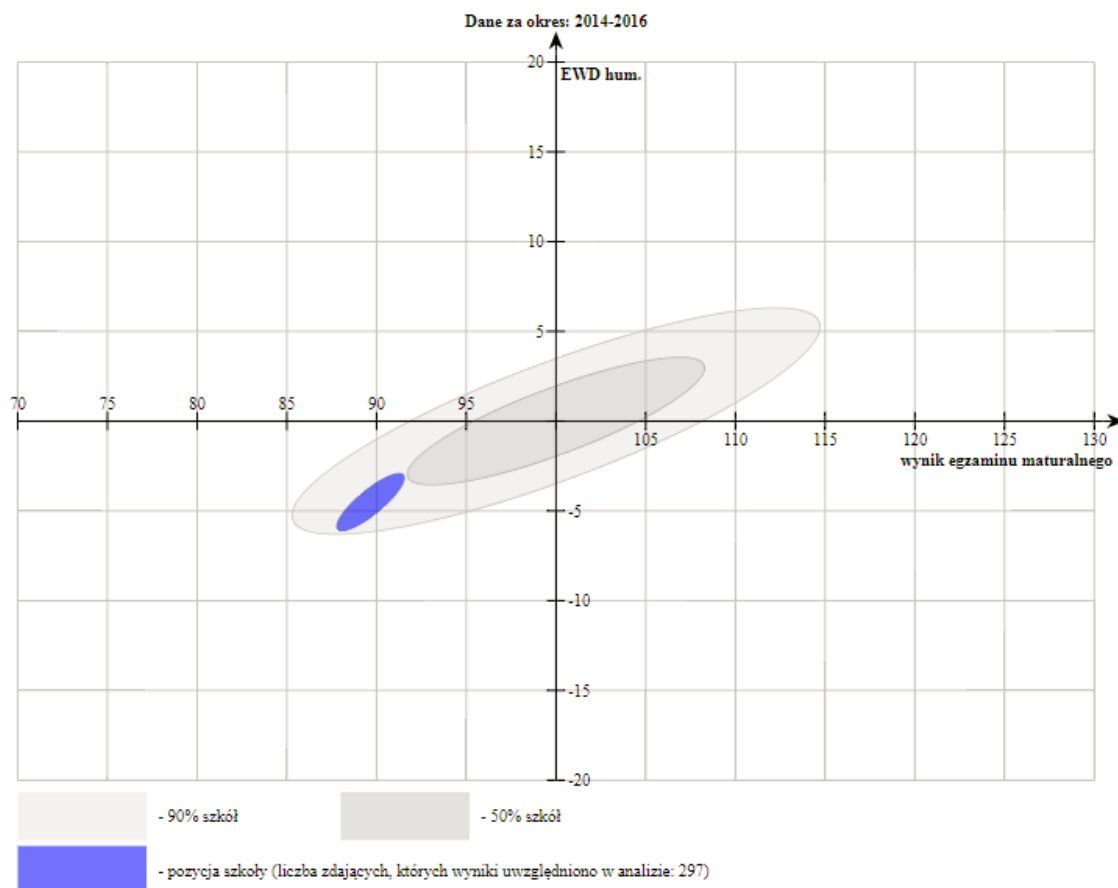




- Szkoły o przeciętnej efektywności i niskich wynikach egzaminacyjnych.

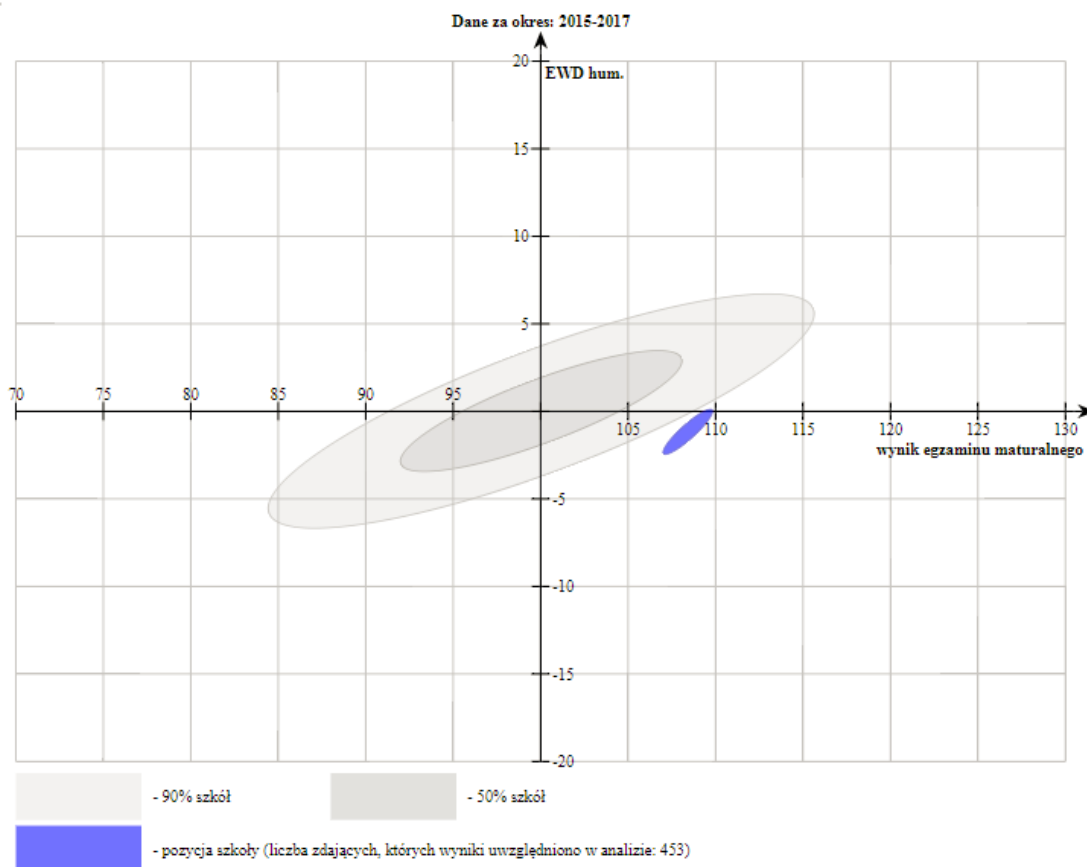


- Szkoły o niskich wynikach egzaminacyjnych i niskiej efektywności nauczania.





- Szkoły o wysokich wynikach egzaminacyjnych oraz niskiej efektywności nauczania.



Dla każdej szkoły obliczane są wskaźniki z zakresu czterech obszarów nauczania:

Obszar 1 – humanistyczny (obejmuje język polski, historię i wiedzę o społeczeństwie).

Obszar 2 – obejmuje wyniki egzaminu z języka polskiego.

Obszar 3 – matematyczno-przyrodniczy (obejmuje matematykę, informatykę, fizykę, chemię, biologię i geografę).

Obszar 4 – matematyka.

Wyniki egzaminacyjne przedstawione są na standardowej skali o średniej 100 i odchyleniu standardowym 15. Skalowanie wyników przeprowadzane jest z uwzględnieniem trudności poszczególnych zadań.



Analiza porównawcza szkół na podstawie wykresu

Wskaźniki EWD mają charakter względny i mogą być wykorzystane do porównywania szkół. Przy dokonywaniu porównania należy wziąć pod uwagę następujące aspekty:

- Szkoły różnią się istotnie statystycznie wynikami, jeśli ich obszary ufności (elipsy) są rozłączne.
- Szkoły różnią się istotnie statystycznie ze względu na wyniki maturalne, jeśli rzuty elips na oś wyniku maturalnego są rozłączne.
- Szkoły różnią się istotnie statystycznie ze względu na EWD, jeśli rzuty ich elips na oś EWD są rozłączne.

Przykład 1

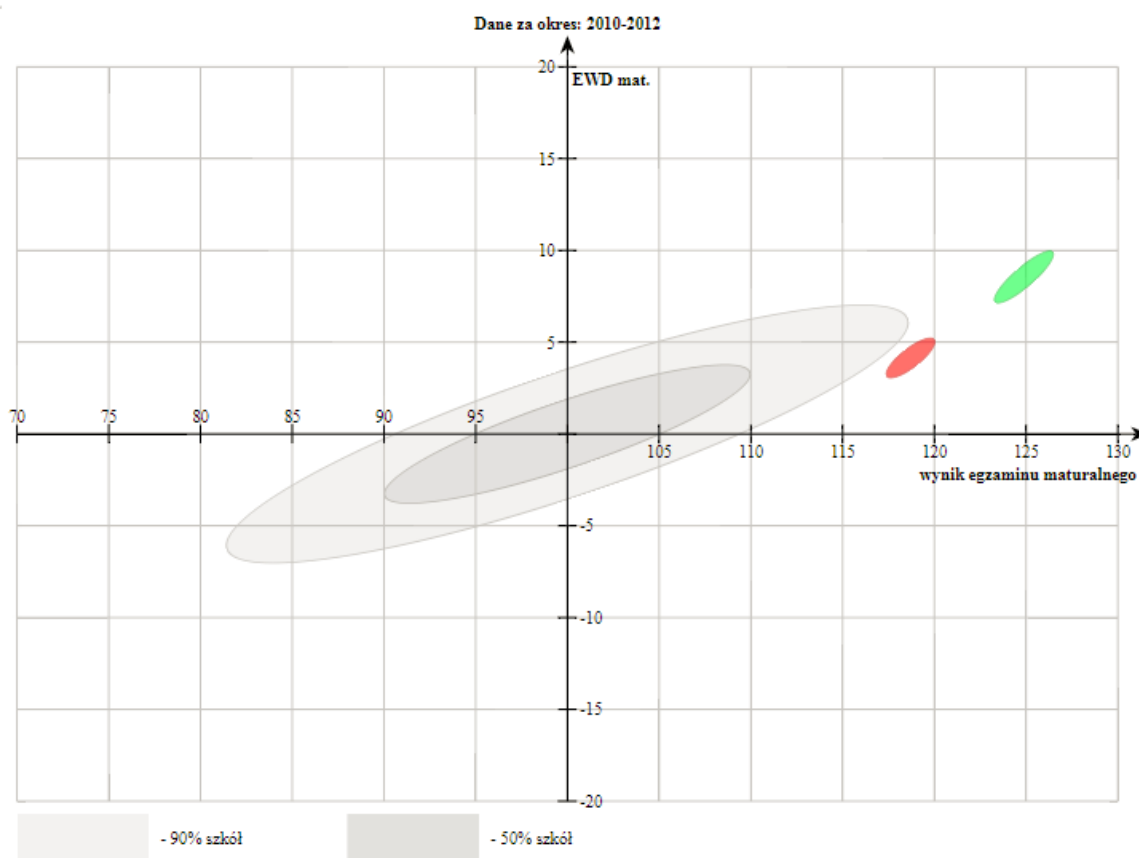
Porównajmy zatem dwie szkoły A i B w zakresie matematyki, języka polskiego i przedmiotów przyrodniczych. W każdym przypadku są te same szkoły, tylko wykresy zmieniają kolor.



Matematyka

Szkoła A (kolor zielony) i szkoła B (kolor czerwony).

Przedstawione na wykresie szkoły różnią się ze względu na wyniki maturalne i EWD. Wyższa efektywność nauczania w szkole A pozwala uzyskiwać uczniom wyższe wyniki na maturze. Obie szkoły charakteryzuje ponadprzeciętna efektywność nauczania matematyki i bardzo wysokie (powyżej jednego odchylenia od średniej) wyniki maturalne.



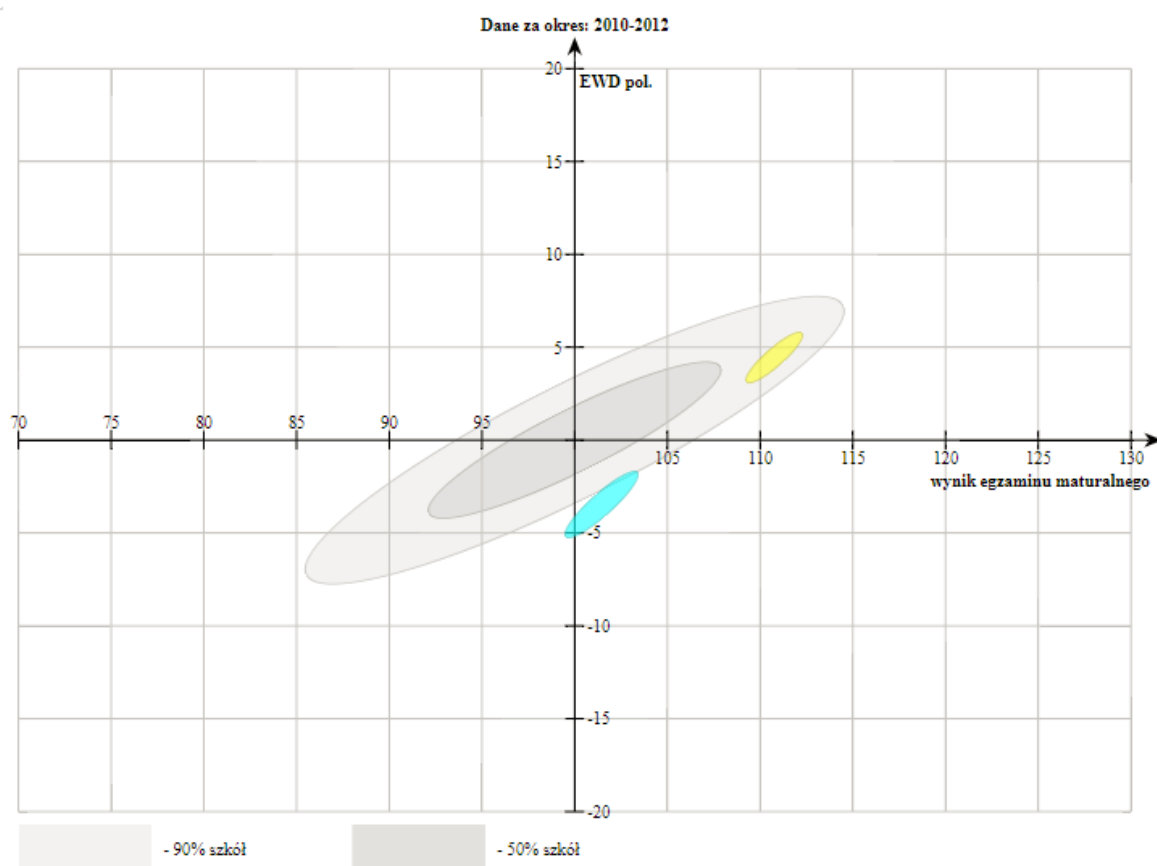
Źródło: www.ewd.edu.pl



Język polski

Szkoła A (kolor niebieski) i szkoła B (kolor żółty).

Przedstawione na wykresie szkoły różnią się statystycznie ze względu na wyniki maturalne, jak i efektywność nauczania (EWD). Wyższa efektywność nauczania cechuje szkołę B, która jest szkołą o wysokich wynikach i ponadprzeciętnej efektywności nauczania. Natomiast szkoła A uzyskuje wyniki nieznacznie wyższe od średniej krajowej przy ujemnej edukacyjnej wartości dodanej.



Źródło: www.ewd.edu.pl

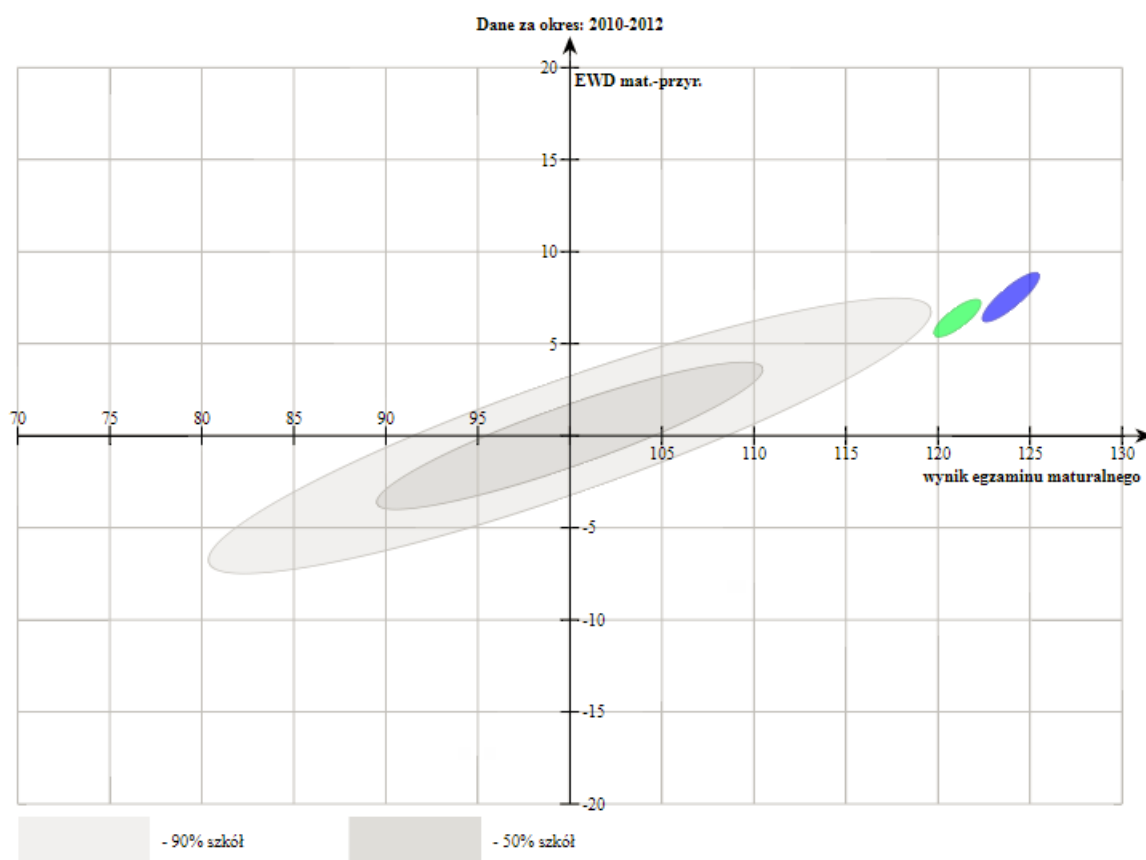
Zauważmy na tym przykładzie, że wskaźniki EWD pokazują, jak jest, ale nie tłumaczą, dlaczego tak jest. Stawianie hipotez i dociekanie przyczyn jest zadaniem badawczym dla samej szkoły, podobnie jak odpowiedź na pytanie, jak powinno być. Czy szkoła, uzyskując wysokie wyniki i efektywność nauczania w jednym obszarze, powinna uzyskiwać porównywalne wyniki w innych obszarach, w szczególności gdy mówimy o wysokoprofilowanych szkołach?



Przedmioty przyrodnicze

Szkoła A (kolor niebieski) i szkoła B (kolor zielony).

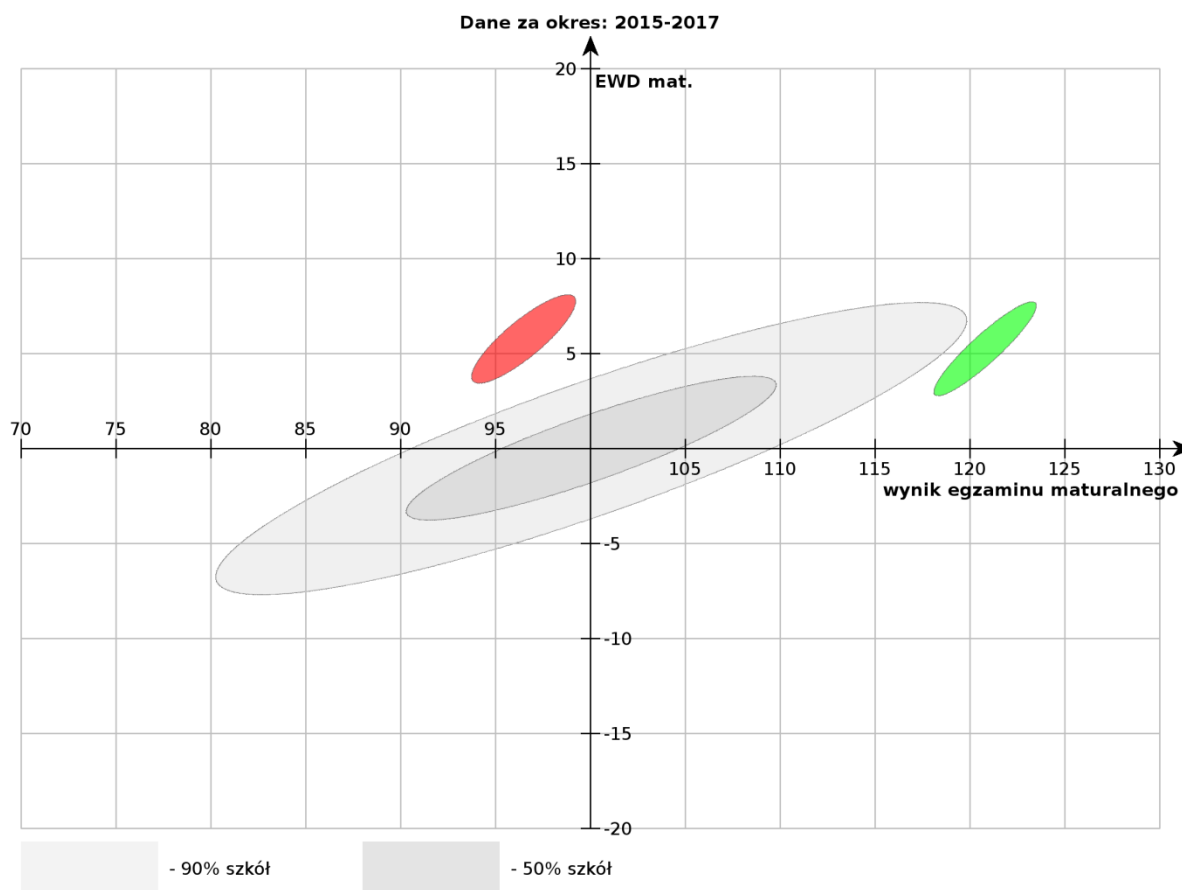
Obie szkoły charakteryzują się wysokimi wynikami maturalnymi i ponadprzeciętną efektywnością nauczania w zakresie przedmiotów przyrodniczych. Szkoły A i B nie różnią się statystycznie ze względu na EWD – efektywność nauczania jest podobna, wyniki maturalne z przedmiotów przyrodniczych wyższe są w szkole B.



Źródło: www.ewd.edu.pl

Przykład 2

Porównanie szkół o podobnym EWD i różnych wynikach egzaminacyjnych.



Te dwa licea ogólnokształcące charakteryzują się podobną efektywnością nauczania matematyki, choć pracują z uczniami o odmiennych uprzednich osiągnięciach. W szkole z wynikami poniżej średniej krajowej uczyli się uczniowie, którzy osiągnęli na egzaminie gimnazjalnym w części matematyczno-przyrodniczej niskie wyniki. Ponadprzeciętna praca uczniów i nauczycieli pozwoliła im uzyskać wyniki maturalne znacznie wyższe niż innym uczniom o podobnych niskich wynikach gimnazjalnych.

Zupełnie inna jest szkoła, którą reprezentuje kolor zielony. Tu pracowano z uczniami, których wynik na egzaminie gimnazjalnym był wysoki, a ponadprzeciętna praca uczniów i nauczycieli dała efekt w postaci bardzo wysokich wyników maturalnych (znacząco powyżej jednego odchylenia od średniej krajowej). Zauważmy, że wyniki obu szkół są dość nietypowe, plasują się poza obszarem szarej elipsy, w którym mieści się 90% wyników szkół. Podkreślimy także, że EWD pokazuje względny postęp ucznia, a nie to, ile on umie. Nie ulega wątpliwości, że poziom wiedzy matematycznej jest wyższy w przypadku uczniów szkoły o wysokich wynikach egzaminacyjnych. Wysoki postęp względny uczniów w liceum, które prezentuje czerwona elipsa, zapewne jest niewystarczający, żeby ubiegać się o miejsce na renomowanej uczelni.



Jak przygotować własny pomiar edukacyjnej wartości dodanej?

Modele EWD to wszelkiego rodzaju modele mierzące efektywność typu wejście–wyjście. W zależności od tego, jak zdefiniujemy zasoby na wejściu i opiszemy efekty na wyjściu, możemy budować modele na własny, szkolny użytek. Pamiętać jednak przy tym należy, że model statystyczny będzie tym lepszy, im więcej wyników w nim uwzględnimy. Oto kilka przykładów:

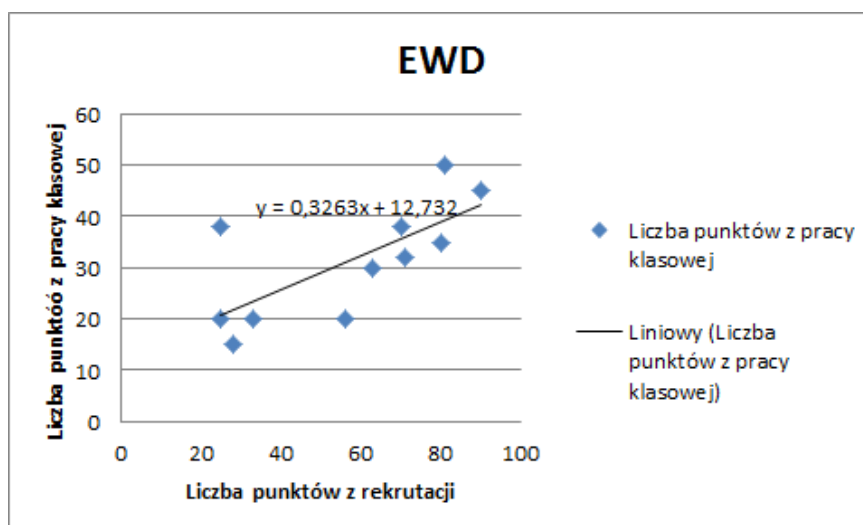
1. W szkole przeprowadzany jest egzamin próbny z matematyki dla wszystkich klas trzecich (około 200 osób). Można zbudować model typu EWD, gdzie na wejściu uwzględnimy wyniki na egzaminie gimnazjalnym, a na wyjściu wyniki egzaminu próbnego, a następnie będziemy analizować EWD wg oddziałów.
2. Egzamin próbny z matematyki przeprowadzamy w klasie II i III dla wszystkich uczniów (około 200 osób), a w modelu EWD porównujemy wyniki z klasy II i klasy III.
3. Model typu EWD zbudowano na podstawie wyników maturalnych z matematyki studentów I roku wszystkich wydziałów Politechniki Łódzkiej (PŁ) i wyników centralnego egzaminu z algebry dla studentów PŁ (ponad 2 tysiące wyników). Wyliczone wskaźniki EWD mogą służyć do oceny postępów względnych studentów na poszczególnych kierunkach lub służyć ocenie efektywności pracy nauczyciela akademickiego.

Ćwiczenie

- Przygotuj dane wejściowe, np. wynik rekrutacji, ocenę z matematyki z poprzedniej szkoły.
- Przygotuj dane wyjściowe, np. oceny z klasówki, punkty z klasówki.

Identyfikator ucznia	Dane wejściowe	Dane wyjściowe
1	23	78

- Wpisz dane do Excela.
- Wykonaj wykres z regresją liniową.



- Wyznacz edukacyjną wartość dodaną dla każdego ucznia.

x – liczba punktów z rekrutacji

y – liczba punktów z pracy klasowej

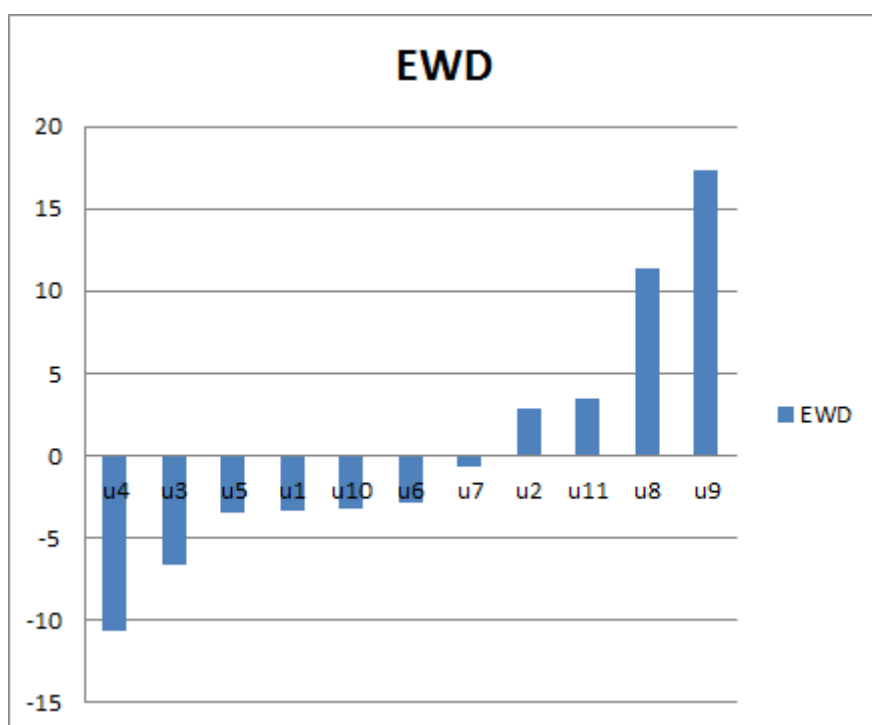
$ax + b$ – wzór regresji liniowej

Edukacyjną wartość dodaną (EWD) obliczamy ze wzoru: $y - (ax + b)$

Identyfikator ucznia	Dane wejściowe	Dane wyjściowe	EWD
1	80	35	-3,3
2	70	38	2,9

- Przygotuj raport.

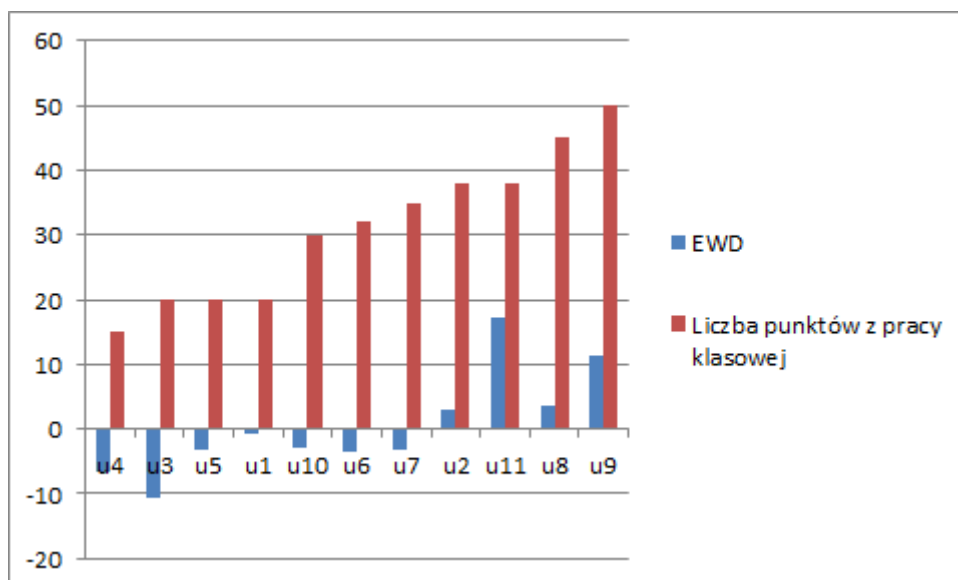
Posortujmy dane według edukacyjnej wartości dodanej i przedstawmy graficzną prezentację danych.



Wnioski:

Uczeń u4 napisał pracę klasową poniżej swoich możliwości. Uczniowie u8 i u9 napisali znacznie powyżej swoich możliwości.

Posortujmy dane według liczby punktów uzyskanych ze sprawdzianu.





Jak wynika z rysunku, uczniowie słabi nie dają sobie rady z zadaniami.

Jakie zalecenia można poczynić na podstawie powyższych danych?

- Powinno się bardziej zróżnicować zadania.



Bibliografia

Kaczmarczyk S., (2011), *Metody zbierania danych ze źródeł pierwotnych przez internet w badaniach marketingowych*, w: „Nierówności społeczne a wzrost gospodarczy”, z. 45(1).

Niemierko B., (1975), *Testy osiągnięć szkolnych podstawowe pojęcia i techniki obliczeniowe*, Warszawa: WSiP.

Raport o stanie edukacji 2014. *Egzaminy zewnętrzne w polityce i praktyce edukacyjnej*, (2015), Dolata R., Sitek M. (red.). Warszawa: Instytut Badań Edukacyjnych.

Włodarczyk P., (2007), *Wizualizacja danych*, pod kierunkiem prof. Włodzisława Ducha.

Wroński Ryszard, (2010) , *Edukacja behawiorystyczna i konstruktywistyczna w aspekcie manipulacji*, [w:] Manipulacja – pedagogiczno-społeczne aspekty, Część II Komunikacja, dydaktyka, wychowanie a manipulacja, Aksman J. (red.), Kraków: Krakowska Akademia im. Andrzeja Frycza Modrzewskiego.

Żółtak T., Rappe A., (2015), *Maturalne wskaźniki edukacyjnej wartości dodanej. Metoda, narzędzia, interpretacje*, Warszawa: Instytut Badań Edukacyjnych

Żółtak T., (2015), *Statystyczne modelowanie wskaźników edukacyjnej wartości dodanej – podsumowanie polskich doświadczeń 2005-2015*, Warszawa: Instytut Badań Edukacyjnych.

