

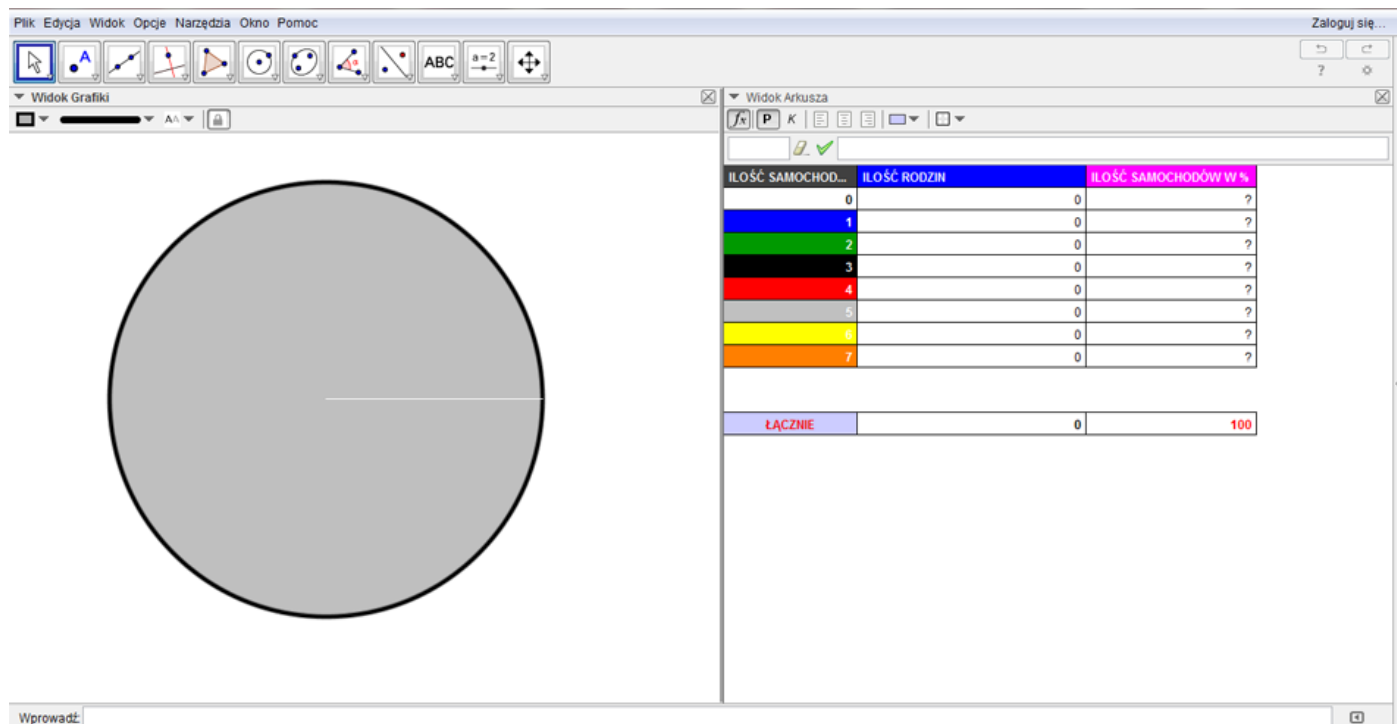


STATYSTYKA W PRAKTYCE – ZBIERAMY I ANALIZUJEMY DANE

CEL PROJEKTU

Celem projektu jest przeprowadzenie badań na 25 rodzinach dotyczące łącznej ilości używanych samochodów przez każdą z tych rodzin. Wybierz 25 rodzin, które znasz – mogą być to rodziny Twoich koleżanek i kolegów czy rodziny Twoich sąsiadów.

WYKORZYSTYWANA PLANSZA



WSTĘP TEORETYCZNY

Dane można analizować poprzez obliczenie i obserwację podstawowych parametrów statystycznych. Do takich parametrów zaliczamy m.in.

- Średnią arytmetyczną n liczb, czyli iloraz sumy tych liczb przez ich liczbę.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

n — ilość wszystkich liczb

$x_1, x_2, + \dots + x_n$ — kolejne liczby

- Medianę. Dany jest zbiór n liczb uporządkowany w kolejności niemalejącej. Medianą nazywamy:
 - wyraz środkowy w przypadku, gdy liczba elementów jest nieparzysta;
 - średnią arytmetyczną dwóch środkowych wyrazów, gdy liczba elementów jest parzysta.
 Mediana może zatem należeć do zbioru danych, ale może także do niego nie należeć.

- **Dominantę (modę) zbioru danych**, czyli taką wartość, która w tym zbiorze występuje najczęściej.
Jeśli w zbiorze kilka wartości występuje z tą samą (najwyższą) częstością, to każda z tych wartości jest dominantą. Jeśli wszystkie wartości w zbiorze występują z tą samą częstością, to przyjmuje się, że zbiór danych nie ma dominanty.
- **Odchylenie standardowe**. Danych jest n dowolnych liczb $(x_1, x_2, x_3, \dots, x_n)$. Niech \bar{x} oznacza średnią arytmetyczną tych liczb. Wówczas odchylenie standardowe obliczamy według wzoru:

$$\sigma = \sqrt{\frac{k_1(x_1 - \bar{x})^2 + k_2(x_2 - \bar{x})^2 + k_3(x_3 - \bar{x})^2 + \dots + k_p(x_n - \bar{x})^2}{n}}$$

n — ilość wszystkich liczb

x_1, x_2, \dots, x_n — kolejne liczby

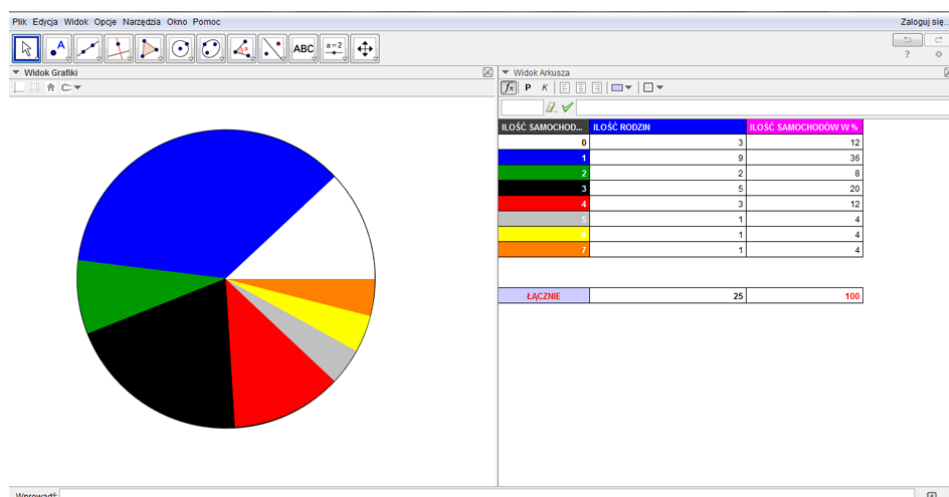
k_1, k_2, \dots, k_p — powtarzające się kwadraty różnicy

Odchylenie standardowe jest parametrem statystycznym informującym o rozproszeniu danych względem średniej arytmetycznej. Im większe odchylenie standardowe, tym dane są bardziej odległe od siebie i od średniej arytmetycznej.

Dobrze również analizować rozkład procentowy poszczególnych wyników. Rozkład może być przedstawiony w postaci tabelarycznej lub przy pomocy diagramu

ETAP 1 – ZEBRANIE DANYCH DO ANALIZY

Otwórz plik Badanie statystyczne – ilość samochodów w rodzinie. Wypisz znane Ci rodziny z ilością samochodów. Wpisz do tabeli, ile rodzin z analizowanych przez Ciebie nie używa samochodu, ile używa 1 samochodu, 2 samochodów, 3 samochodów, ... itd. W pliku masz przygotowaną tabelę z potencjalną ilością samochodów. Tabela przewiduje, że ilość samochodów będzie się wahać od 0 do 7. W tabeli otrzymasz również wyniki procentowe. 25 rodzin odpowiada 100%, więc 1 rodzina to 4%. Po lewej stronie wyniki zostaną przedstawione na diagramie kołowym. Przykładowe zestawienie znajduje się poniżej.





ETAP 2 – OBLICZANIE PODSTAWOWYCH PARAMETRÓW STATYSTYCZNYCH

Średnia arytmetyczna

Oblicz łączną ilość samochodów we wszystkich rodzinach i podziel przez ich ilość. Zapisz otrzymaną średnią.

$\bar{x} =$

Mediana

Uporządkuj wszystkie rodziny wg ilości posiadanych samochodów w kolejności rosnącej. Jeżeli, któryś z wyników się powtarza np. 4 razy to piszemy go 4 razy. Medianą będzie środkowy wyraz uporządkowanego zbioru wyników.

Zapisz wartość mediany.

$m_e =$

Dominanta

Dominanta to wynik występujący najczęściej. Zapisz wartość dominanty

$D =$

Odchylenie standardowe

Oblicz wartość odchylenia standardowego.

$\sigma =$

ETAP 3 – ANALIZA DANYCH

Odpowiedz na pytania:

1. Co możesz ogólnie stwierdzić na podstawie otrzymanych parametrów statystycznych?
2. Czy wartość mediany jest zbliżona do wartości średniej arytmetycznej
3. Jaka wartość (ilość samochodów) występuje najczęściej ?
4. Jaka wartość (ilość samochodów) występuje najrzadziej ?
5. Ile co najmniej samochodów ma więcej niż 50% rodzin?
6. Ile procent rodzin ma 3 lub 4 samochody?
7. Czy rozproszenie wyników jest duże czy małe (wykorzystaj informacje dotyczącą odchylenia standardowego)?
8. Ile procent rodzin ma więcej samochodów niż średnia, a ile mniej niż średnia? Czy są to takie same ilości rodzin?



9. Czy mediana jest większa od dominanty czy mniejsza od dominanty? Czy coś może z tego wynikać?

10. Zaznacz prawidłową odpowiedź.

Gdyby każdej rodzinie przybył jeden samochód to:

- a. mediana zwiększyłaby się o 1.
- b. średnia arytmetyczna wzrosłaby o 1.
- c. dominanta zwiększyłaby się o 1.