

# STATYSTYKA OPISOWA

*Znaczenie podstawowych miar*

*Pytanie wieczoru 1:*

Ile zarabiają  
dyrektorzy w  
działach sprzedaży?

Średnia zarobków dyrektorów sprzedaży  
wynosi

12 161 PLN.

## *PYTANIA:*

- Jak obliczono tę średnią?
- Czy dyrektorzy to generalnie bogaci ludzie?
- Czy można określić ile zarabia jakiś konkretny dyrektor?
- Czy można obliczyć średnią płęć dyrektorów? ;)

# Jak obliczono tę średnią?

Średnią arytmetyczną liczymy z dobrze znanego wzoru:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

*Czyli* dodajemy do siebie wartości wszystkich obserwacji i dzielimy otrzymaną sumę przez liczbę przeprowadzonych obserwacji.

Czy dyrektorzy to generalnie bogaci ludzie?

Trudno powiedzieć...

BO jeśli:

- Zbadano 100 osób
- 10 z nich zarabia 103 610 zł

TO cała reszta zarabia jedynie 2 000 zł

*Inny przykład*

WNIOSEK:

*Średnia arytmetyczna jest bardzo wrażliwa  
na wartości skrajne.*

# Czy można określić ile zarabia jakiś konkretny dyrektor?

**NIE!**

BO jeśli:

- Zbadano 100 osób
- 30 z nich zarabia 8 000 zł
- 30 z nich zarabia 20 000 zł
- 40 z nich zarabia 9 402, 50 zł

TO średnia nadal wynosi 12 161 zł

WNIOSEK:

*Średnia arytmetyczna jest wartością abstrakcyjną;  
może być różna od wszystkich dokonanych  
pomiarów.*

Czy można obliczyć średnią płęć dyrektorów?

**NIE!**

**BO**

*Średnią arytmetyczną można obliczać  
jedynie dla wartości liczbowych*



# ŚREDNIA – podsumowanie

- wartość obliczana ze wzoru:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- Jest wrażliwa na wartości skrajne
- Jest wartością abstrakcyjną
- Może być obliczana tylko dla wartości liczbowych (*w pewnym uproszczeniu*)

Problemy ze średnią pomaga rozwiązywać

# ODCHYLENIE STANDARDOWE.

Obliczane jest z wzoru:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{n}{n - 1} (\overline{x^2} - (\bar{x})^2)}$$

A więc jest (w uproszczeniu):

*Średnią różnicą między średnią arytmetyczną a poszczególnymi wynikami.*

# ODCHYLENIE STANDARDOWE - cd

- Czy odchylenie  $s = 1000$  jest duże?  
gdy  $x_{\text{sr}}=4000$  – TAK, gdy  $x_{\text{sr}}=40$  mln - NIE
- Czy odchylenie  $s = 20$  jest duże?  
gdy  $x_{\text{sr}}= 60$  – TAK, gdy  $x_{\text{sr}}=2000$  - NIE
- Czy odchylenie  $s = 0,04$  jest duże?  
gdy  $x_{\text{sr}}= 0,2$  – TAK, gdy  $x_{\text{sr}}=5$  - NIE

*Odpowiedź:*

Zależy od wysokości średniej

# ODCHYLENIE STANDARDOWE - cd

DLATEGO:

Aby określić, czy odchylenie jest duże

*(czyli, czy uzyskane wartości znajdują się w pobliżu, czy daleko od obliczanej średniej)*

Można posłużyć się współczynnikiem zmienności

$$V = \frac{s}{\bar{x}}$$

Im wyższy współczynnik, tym średnia

**GORZEJ**

określa badaną próbę

Proszę ustawić się wg wzrostu, a  
następnie wybrać osobę, której  
wzrost będzie

**MEDIANA**  
dla grupy.

■ **MEDIANA**  
zarobków wśród  
dyrektorów  
wynosi: 8 250zł

■ **MODA**  
zarobków  
wśród  
dyrektorów  
wynosi: 7 500  
zł.

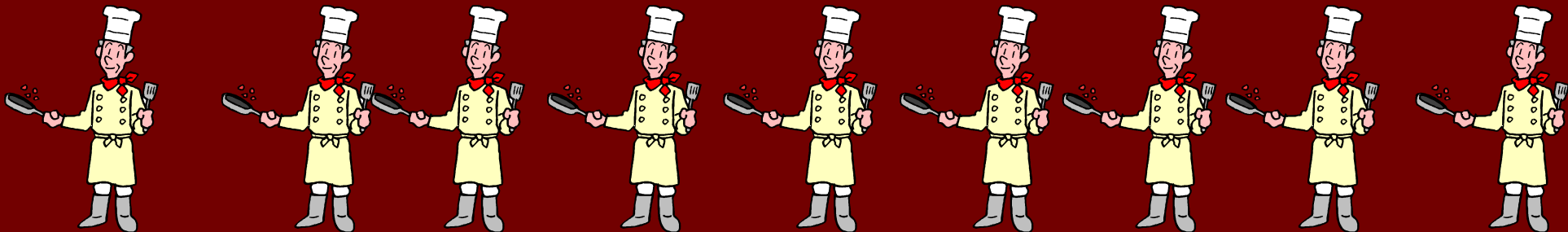
### PYTANIA:

- W jakim przedziale mieszczą się zarobki większości dyrektorów?
- Ile zarabia „przeciętny” dyrektor?
- Czy prawie wszyscy dyrektorzy to bogaci ludzie?
- Czy większość dyrektorów to bogaci ludzie?
- Czy zarobki dyrektorów różnią mocno różnią się od siebie?
- Jakie zarobki są najczęstsze wśród dyrektorów?

# MEDIANA

*Definicja:*

„Jest to wartość cechy poniżej i powyżej, której znajduje się dokładnie połowa obserwacji”  
(Sambor 1990)



# Na jakie pytania można odpowiedzieć na podstawie MEDIANY:

- W jakim przedziale mieszczą się zarobki większości dyrektorów?

*Powyżej 8 250zł*

- Czy większość dyrektorów to bogaci ludzie?

*Jeśli 8 250zł oznacza człowieka bogatego, to TAK*

CZYLI można powiedzieć coś o połowie  
badanej grupy

I wyciągnąć z tego wnioski



# Czego NIE MOŻNA stwierdzić na podstawie MEDIANY:

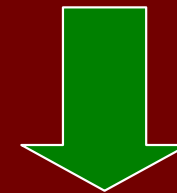
- Ile zarabia „przeciętny” dyrektor?  
*BO z bycia w środku nie wynika bycie „przeciętnym”*
- Czy prawie wszyscy dyrektorzy to bogaci ludzie?  
*BO nie wiemy NIC o całości badanej grupy*
- Czy zarobki dyrektorów różnią mocno różnią się od siebie?  
*BO nie wiemy NIC o całości badanej grupy*
- Jakie zarobki są najczęstsze wśród dyrektorów?  
*BO to co jest w środku nie musi być najbardziej popularne*

# MODA (dominanta)

*Definicja:*

„Jest to najczęstsza wartość cechy (...) jest to ta wartość obserwacji, której odpowiada maksymalna liczebność”

(Sambor 1990)



Na jakie pytania można odpowiedzieć na podstawie DOMINANTY:

- Ile zarabia „przeciętny” dyrektor?

7 500 zł

BO „przeciętny” to znaczy taki, jaki występuje najczęściej

- Jakie zarobki są najczęstsze wśród dyrektorów?

7 500 zł

BO dominanta oznacza właśnie wartość występującą najczęściej

# Czego NIE MOŻNA stwierdzić na podstawie DOMINANTY:

- W jakim przedziale mieszczą się zarobki większości dyrektorów?

- Czy większość dyrektorów to bogaci ludzie?

*BO najczęstsza wartość wcale nie musi dotyczyć większości*

- Czy prawie wszyscy dyrektorzy to bogaci ludzie?

*BO jeśli nie wiemy nic o większości, to tym bardziej o prawie wszystkich*

- Czy zarobki dyrektorów różnią mocno różnią się od siebie?

*BO nie wiemy nic o całości grupy*

*Pytanie wieczoru 2:*

Jak oceniana jest  
subiektywna częstość  
występowania słów  
*PISTOLET*, MAMA I DBAĆ  
w języku polskim?

(Imiołczyk 1984)

## KTO WIE CO:

- Każda grupa losuje dwie wartości – jedną „swoją” – na jej podstawie odpowiada na pytania, drugą „cudzą” do której zadaje pytania
- Zadaniem grupy jest sformułowanie jak największej liczby trafnych pytań, na które można odpowiedzieć na podstawie „cudziej” wartości
- Za każde dobre pytanie grupa dostaje **1 pkt**
- Za każde błędne pytanie - **-1pkt**
- Grupa otrzymująca pytanie – jak odpowie dobrze – **1 pkt**; jak odpowie na pytanie, na które nie ma informacji w ich wartości - **-1pkt**
- Pytania zadawane są po kolei
- Wygrywa grupa, która dostanie najwięcej punktów

Każda z badanych osób ocenia częstość występowania podanego wyrazu w mowie w skali od 1 do 7 gdzie:

- 1 – wyraz nie używany nigdy lub nieznan
- 2 – wyraz używany bardzo rzadko
- 3 – wyraz używany rzadko
- 4 – wyraz używany niezbyt często / niezbyt rzadko
- 5 – wyraz używany często
- 6 – wyraz używany bardzo często
- 7 – wyraz używany stale

# Uzyskane wyniki:

<i>słowo</i>	$X_{\text{śr}}$	Mo	Me
MAMA	6,37	7	7
DBAĆ	4,12	4	4
PISTOLET	2,83	2	3

1 – wyraz nie używany nigdy lub nieznan

2 – wyraz używany bardzo rzadko

3 – wyraz używany rzadko

4 – wyraz używany niezbyt często / niezbyt rzadko

5 – wyraz używany często

6 – wyraz używany bardzo często

7 – wyraz używany stale



# PODSUMOWANIE

- **ŚREDNIA** – obejmuje wszystkie pomiary, jest wartością abstrakcyjną.

*Jaka jest długość (iloczas) głoski A w języku polskim?*

- **MEDIANA** – wartość środkowa, oznacza wartość, która znajduje się dokładnie w środku wszystkich pomiarów.

*Jaką ilość słów opanowuje się w ciągu pierwszego roku nauki języka obcego?*

- **DOMINANTA** (wartość modalna) – oznacza wartość, na którą przypada najwięcej obserwacji.

*Jakie znaczenie jest najczęściej kojarzone ze słowem SIATKA?*

# BIBLIOGRAFIA:

Imiołczyk, J. 1984, *Prawdopodobieństwo subiektywne wyrazów. Podstawowy słownik frekwencyjny języka polskiego*. Poznań.

Han, J. 2006, *Wynagrodzenia w działach sprzedaży w 2006 roku*. [www.pracuj.pl](http://www.pracuj.pl)

Hammerl, R., Sambor, J., *Statystyka dla językoznawców*, Warszawa 1990