

# Statystyka opisowa

## Wykład 1:

**Informacje organizacyjne  
o statystyce ogólnie**

# Po co nam statystyka???

1. Żyjemy w czasach, kiedy informacja stała się towarem...

Dobrze jest umieć analizować dane, po to żeby lepiej wykorzystywać MOŻLIWOŚCI jakie nam stwarza nowoczesność XXI wieku.

2. Żyjemy w czasach, kiedy informacja stała się narzędziem...

Dobrze jest wiedzieć, jak inni mogą wykorzystywać to narzędzie i umieć ochronić się przed ZAGROŻENIAMI, jakie stwarza nam nowoczesność XXI wieku.

# Najsłynniejsze powiedzenie o statystyce

*Są trzy rodzaje kłamstwa:  
normalne, diabelskie i statystyka*

[Autor niejednoznaczny: M. Twain, B. Disraeli, G.B. Shaw]

**I dopowiedzenie...**

*Ale to nie statystyka kłamie,  
tylko ludzie, którzy się nią posługują!*

# Podstawowe informacje

Wykładowca: **dr Marek Sobolewski**

Strona internetowa: [mareksobolewski.v.prz.edu.pl](http://mareksobolewski.v.prz.edu.pl)

Wykład – **15 godz. (9\*)**

Ćwiczenia – **15 godz. (9\*)**

Laboratorium – **30 godz. (18\*)**


**\* dotyczy studiów niestacjonarnych**

Ocena końcowa: ważona średnia ocen z ćwiczeń,  
laboratoriów oraz **egzaminu**

**Egzamin odbędzie się w sesji, będzie miał formę  
praktyczną, w celu weryfikacji praktycznych  
umiejętności analizy danych.**

# Zakres materiału

- **podstawowe pojęcia związane ze statystyczną analizą danych**
- **konstrukcja arkusza danych statystycznych w programie STATISTICA**
- **sposoby opisu danych statystycznych (sortowanie, rangowanie, grupowanie, statystyki opisowe, prezentacje graficzne)**
- **analiza zależności pomiędzy dwiema cechami statystycznymi (tabele wielodzzielcze, statystyki w grupach, współczynnik korelacji)**
- **elementy wnioskowania statystycznego – praktyczne wykorzystanie testów statystycznych**
- **indeksy statystyczne – analiza dynamiki zjawisk społeczno-gospodarczych**



**Niektóre zagadnienia  
mogą być  
realizowane tylko na  
wykładzie**

# **Program *STATISTICA***

**Większość obliczeń i prezentacji graficznych wykonywać będziemy za pomocą programu *STATISTICA 13* – specjalistycznego narzędzia analizy danych.**

**Program *STATISTICA* jest zainstalowany w pracowniach, ponadto licencja, którą posiada Politechnika Rzeszowska, uprawnia studentów do posługiwania się programem na komputerach domowych.**

**Informacje o sposobie instalacji i wersja instalacyjna programu *STATISTICA* znajdują się na stronie:**

**<http://ci.prz.edu.pl/licencje-oprogramowania>**

**Część obliczeń i analiz wykonywać będziemy także za pomocą arkusza kalkulacyjnego *Excel*.**

# Zbiory danych

Podczas nauki statystyki będziemy najczęściej korzystać z gotowych zbiorów danych – wszystkie pliki są udostępnione do pobrania na [stronie domowej wykładowcy](#) w zakładce [materiały do pobrania](#).

**Uwaga!** Do korzystania z baz danych wymagane jest zainstalowanie programu *STATISTICA*.

# Literatura

1. Hydzik P., Sobolewski M., *Komputerowa analiza danych społeczno-gospodarczych*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2007.
2. Aczel A.D., *Statystyka w zarządzaniu*, PWN, Warszawa 2000 lub 2011.

*Dodatkowe pozycje podawane będą przez osoby prowadzące ćwiczenia i laboratoria*

# CO TO JEST STATYSTYKA?

**STATYSTYKA** – zbiór metod badania zjawisk masowych (zbieranie i przetwarzanie danych):

- *Statystyka opisowa* to narzędzia prezentacji i opisu posiadanych danych.
- *Statystyka matematyczna* to zbiór narzędzi wnioskowania statystycznego, które pozwalają na podstawie posiadanych danych wyciągać wnioski natury ogólniejszej.

## STATYSTYKA / ANALIZA DANYCH a „Zarządzanie”

Znaczenie metod analizy danych w skutecznym zarządzaniu jest oczywiste. Nie sposób podejmować decyzje w jakiegokolwiek dziedzinie, bez dokładnej wiedzy o funkcjonowaniu przedsiębiorstwa w chwili obecnej i przeszłości. Nie sposób podejmować właściwych decyzji bez wiedzy o konkurentach, otoczeniu makroekonomicznym oraz o swoich klientach. Statystyka pozwala na uporządkowanie danych o tych aspektach rzeczywistości w taki sposób, by informację zamienić w wiedzę.

# Jednostka statystyczna, próba i populacja

Przedmiotem analiz statystycznych są tzw. **JEDNOSTKI STATYSTYCZNE**, czyli obiekty podlegające badaniu, których cechy (dane) są analizowane.

Zbiór jednostek statystycznych, o których posiadamy dane będące przedmiotem analizy nazywany jest **PRÓBĄ (STATYSTYCZNĄ)**. Szerszy zbiór jednostek, o którym nie posiadamy danych, lecz chcemy go poznać na podstawie próby jest określany jako **POPULACJA**.

*Aby móc wnioskować o populacji na podstawie próby należy zastosować metody **WNISKOWANIA STATYSTYCZNEGO** – niezbędnym jednak warunkiem jest to, by próba była **REPREZENTATYWNA** dla całej populacji.*

# Przykłady badań statystycznych

| Temat badania   | Jednostka statystyczna            | Próba  | Populacja  |
|---|-----------------------------------|--|--|
| <i>Sondaż wyborczy</i>  | <i>Dorosły mieszkaniec Polski</i> | <i>Losowy podzbiór mieszkańców Polski (zwykle ok. 1000 osób)</i> | <i>Wszyscy mieszkańcy Polski (BADANIE CZĘŚCIOWE)</i>   |
| <i>Wykorzystanie środków unijnych w gminach woj. podkarpackiego</i>         | <i>Gmina</i>                      | <i>Gminy woj. podkarpackiego (N = 160)</i>                       | <i>Tak jak próba (BADANIE PEŁNE)</i>                   |
| <i>Poziom życia w Polsce na tle innych państw Unii Europejskiej</i>         | <i>Państwo</i>                    | <i>Państwa należące do UE (N = 27)</i>                           | <i>Tak jak próba (BADANIE PEŁNE)</i>                   |
| <i>Zadowolenie mieszkańców Rzeszowa z rozwoju miasta w ostatnich latach</i> | <i>Mieszkaniec Rzeszowa</i>       | <i>Losowo wybrani mieszkańcy Rzeszowa</i>                        | <i>Wszyscy mieszkańcy Rzeszowa (BADANIE CZĘŚCIOWE)</i> |

# Rodzaje badań statystycznych

Badanie statystyczne może mieć charakter **PEŁNY** lub **NIEPEŁNY** (badanie częściowe).

Wykonanie jedynie badanie częściowego może być spowodowane:

- kosztami i czasochłonnością wykonania badania pełnego (sondaż wyborczy, badanie opinii publicznej);
- niemożnością wykonania badania pełnego, z powodu otwartego charakteru populacji (badanie skuteczności leków, jakości nauczania statystyki);
- niszczeniem elementów podlegających badaniu (kontrola jakości produktów spożywczych).

# Cechy statystyczne i ich rodzaje

Dobór metody analizy statystycznej zależy od charakteru **CECH** podlegających badaniu. Z praktycznego punktu widzenia istotna jest umiejętność rozróżnienia trzech typów cech statystycznych:

- cechy mierzalne (ilościowe, liczbowe)

*Są to cechy, których wartości są wyrażane za pomocą liczb: wiek, dochody, ocena ze statystyki, stopa bezrobocia, środki unijne pozyskane przez gminę w roku 2009, itd.*

- cechy porządkowe

*Są to cechy, których wartości nie są wyrażane za pomocą liczb ale dają się logicznie uszeregować: poziom wykształcenie, stosunek emocjonalny do reklamy, itd.*

- cechy nominalne (jakościowe)

*Są to cechy, których warianty określają tylko i wyłącznie przynależność do pewnej grupy: płeć, miejsce zamieszkania, kierunek studiów, posiadanie prawa jazdy, itd.*

# Typy danych statystycznych

Ze względu na charakter jednostek statystycznych podlegających badaniu wyodrębniamy trzy podstawowe typy zbiorów danych:

- **dane ankietowe (kwestionariuszowe)** – charakteryzują się losowym doбором próby z większej populacji, jednostki podlegające badaniu są nierozróżnialne;
- **dane przekrojowe** – zwykle są efektem przeprowadzenia badania pełnego, jednostki podlegające badaniu są rozróżnialne i mogą to być: państwa, regiony, miasta itp.;
- **dane czasowe** – jednostki są tutaj kolejnymi momentami (okresami) czasu, w których dokonywano pomiaru – mogą to być sesje giełdowe, dni, tygodnie, lata. Oczywiście są one rozróżnialne, co więcej istotna jest ich kolejność.

# Analitik danych na rynku pracy

Ranking najlepszych zawodów wg raportu „Jobs Rated Almanac” 2015 przygotowanego przez serwis CareerCast \*

1. aktuariusz – ok. 94 tys. dol. rocznie;
2. audiolog – 71 tys. dol.;
3. matematyk (ale taki od biznesu i przemysłu) – 102 tys. dol.;
4. statystyk analityk (prowadzi badania i analizuje dane) – 79 tys. dol.;
5. inżynier biomedyczny – 89 tys. dol.;
6. naukowiec (chodzi o kogoś, kto jest jednocześnie informatykiem, statystyko-analitikiem i jeszcze ma wiedzę z konkretnej dziedziny) – 89 tys. dol.;
7. higienistka stomatologiczna – 71 tys. dol.;
8. software inżynier – 93 tys. dol.;
9. terapeuta (zarówno od ciała, jak i od duszy) – 77 tys. dol.;
10. analityk systemów komputerowych – 81 tys. dol.;
11. patolog mowy (coś takiego jak logopeda, bo zajmuje się zaburzeniami mowy i słuchu, czyli osobami z problemami z komunikacją) – 71 tys. dol.;
12. dietetyk – 56 tys. dol.

\*Przy tworzeniu zestawienia brano pod uwagę nie tylko zarobki, lecz także stopień niebezpieczeństwa profesji czy wysiłek konieczny do wykonania obowiązków, stres towarzyszący wykonywaniu pracy oraz łatwość znalezienia zatrudnienia.