

# STATYSTYCZNE METODY WSPOMAGANIA DECYZJI

## zestaw zadań nr 1

*Cel: statystyka opisowa: metody graficzne prezentacji danych i wyznaczanie podstawowych statystyk próbkowych*

### Zadanie 1

Badania demograficzne przeprowadzone w 1988 roku w USA wykazały, że wśród kobiet (mających 18 i więcej lat) było: 17364 tys. panien, 56128 tys. mężatek, 11239 tys. wdów i 8170 tys. rozwódek.

- Utworzyć wykres kołowy dla stanu cywilnego danej grupy kobiet. Porównać różne formy opisu wykresu.
- Utworzyć wykres słupkowy dla stanu cywilnego danej grupy kobiet. Porównać różne rodzaje wykresów i formy ich opisu.

### Zadanie 2

Koncern paliwowy planuje otworzyć nową stację benzynową w pewnym mieście. Rozważane są cztery możliwe lokalizacje stacji – w południowej, północnej, zachodniej i wschodniej dzielnicy miasta (oznaczenia dzielnic: S – południowa, N – północna, W – zachodnia i E – wschodnia). W ramach badania opinii społecznej odnośnie preferowanej lokalizacji stacji zapytano o to tysiąc kierowców. Ich odpowiedzi znajdują się w pliku **stacje.csv**. Utworzyć wykres słupkowy i wykres kołowy dla badanych preferencji.

### Zadanie 3

Poniższe dane odpowiadają notowaniom pewnej spółki (w PLN) w kolejnych 20 dniach:

23,30	24,50	25,30	25,30	24,30	24,80	25,20	24,50	24,60	24,10
24,30	26,10	23,10	25,50	22,60	24,60	24,30	25,40	25,20	26,80

Utworzyć wykres cen akcji jako funkcję czasu (szereg czasowy).

### Zadanie 4

Wytrzymałość na ciśnienie wewnętrzne szkła butelek jest ich ważną charakterystyką jakościową. W celu zbadania wytrzymałości butelek umieszcza się je w maszynie hydrostatycznej, po czym zwiększa się ciśnienie aż do zniszczenia butelki. Plik **butelki.csv** zawiera dane opisujące graniczną wytrzymałość na ciśnienie wewnętrzne szkła badanej partii butelek (mierzone w psi).

- Utworzyć zmienną o nazwie *cisnienie*, opisującą wytrzymałość na ciśnienie wewnętrzne szkła butelek mierzone w MPa

**Wskazówka:**

1psi = 0,0068947 MPa

- Utworzyć histogram dla danych opisujących wytrzymałość butelek. Prześledzić wpływ liczby klas na kształt histogramu. Porównać różne rodzaje histogramów.
- Utworzyć wykres łamanej licznosci i nałożyć go na wykres histogramu.
- Utworzyć wykres łodygowo-liściowy.
- Utworzyć i zinterpretować wykres skrzynkowy dla wytrzymałości butelek.
- Wyznaczyć i zinterpretować podstawowe statystyki próbkowe dla danych opisujących wytrzymałość butelek.
- Obliczyć i zinterpretować 5, 10, 25, 50, 75, 90 i 95 percentyl dla rozważanych danych.
- Wyznaczyć 10% średnią uciętą dla danych opisujących wytrzymałość butelek. Porównać średnią uciętą ze średnią arytmetyczną i medianą. Prześledzić, jak zmienia się wartość średniej wraz ze zmianą stopnia ucięcia próbki.

### Zadanie 5

Zamieszczone poniżej dane przedstawiają wysokość czynszu płaconego w pewnej spółdzielni mieszkaniowej przez 30 losowo wybranych lokatorów:

334	436	425	398	424	429	392	428	339	389
352	405	392	403	344	400	424	443	378	387
384	498	374	389	367	457	409	454	345	422

Przeprowadzić wstępną analizę statystyczną powyższych danych.

### Zadanie 6

W pliku **samochody.csv** zamieszczono dane dotyczące parametrów samochodów kilku wybranych marek.

- a) Zmienna *mpg* zawiera dane odpowiadające liczbie mil, przejechanych przez dany samochód na galonie paliwa. Utworzyć zmienną *zp* opisującą zużycie paliwa mierzone w litrach na 100 kilometrów.

**Wskazówka:**

1 mila = 1609 m

1 galon (amerykański) = 3,785 l

- b) Utworzyć wykres łądogowo-liściowy dla zużycia paliwa.  
c) Utworzyć histogram dla danych opisujących zużycie paliwa.  
d) Utworzyć wykres skrzynkowy dla zużycia paliwa.  
e) Wyznaczyć i zinterpretować podstawowe statystyki próbkowe dla danych opisujących zużycie paliwa (takie jak: średnia, mediana, wariancja, odchylenie standardowe, rozstęp, kwartyle, rozstęp międzykwartyłowy, wartości ekstremalne, współczynnik asymetrii, kurtoza, współczynnik zmienności).  
f) Obliczyć i zinterpretować 5, 10, 90 i 95 percentyl dla rozważanych danych.  
g) Wyznaczyć 5% średnią uciętą dla danych opisujących zużycie paliwa.

### Zadanie 7

Analizowane w poprzednim zadaniu dane dotyczące zużycia paliwa postanowione poddać kategoryzacji tworząc następujące klasy:

Zużycie paliwa [litry na 100 km]	Kod opisujący kategorię zużycia paliwa
nie więcej niż 7	mało
więcej niż 7 ale nie więcej niż 10	średnio
więcej niż 10	dużo

Utworzyć wykres słupkowy dla wyznaczonych w ten sposób kategorii i wskazać, jaki procent badanych samochodów należy do każdej kategorii.

### Zadanie 8

Przeprowadzić wstępną analizę statystyczną danych dotyczących zużycia paliwa oddzielnie dla samochodów produkowanych w Europie, Ameryce i Japonii (wykorzystać zmienne *producent* i *legenda*). Zestawić wykresy skrzynkowe zużycia paliwa dla samochodów produkowanych w Europie, Ameryce i Japonii.

### Zadanie 9

Przeprowadzić wstępną analizę statystyczną danych dotyczących zużycia paliwa przez samochody o jednakowej liczbie cylindrów (dane dotyczące liczby cylindrów znajdują się w zmiennej *cylindry*).

### Zadanie 10

Przeprowadzić wstępną analizę statystyczną danych dotyczących zużycia paliwa przez samochody mające wyłącznie 6 lub 8 cylindrów.

### Zadanie 11

Porównać przyspieszenie samochodów produkowanych w Ameryce i Japonii (dane dotyczące przyspieszenia znajdują się w zmiennej *przysp*).

### Zadanie 12

Przeprowadzić wstępną analizę statystyczną danych dotyczących zużycia paliwa wyłącznie dla samochodów ważących mniej niż 2500 funtów (wykorzystać zmienną *waga*).