

CZYTANIE DANYCH I STATYSTYKI

# WIZUALIZACJA DANYCH

# Wizualizacja danych

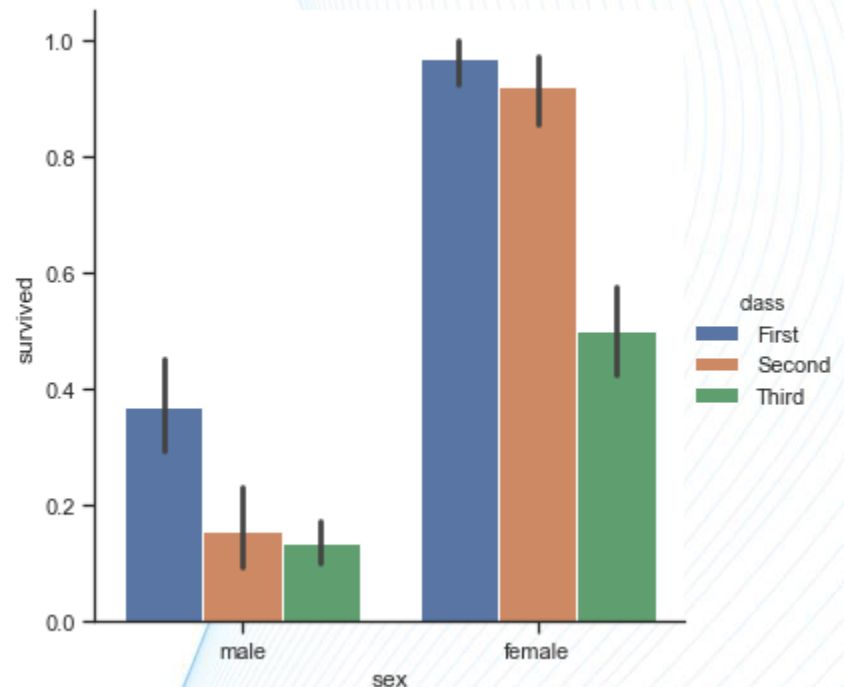
1. Metody wizualizacji
2. Dobór metod wizualizacji: typ grafu a rodzaj zmiennej
3. Najczęstsze błędy w wizualizacji i jak ich uniknąć?
4. Czego uczy nas psychologia poznawcza czyli jak rysować wykresy, które łatwo zinterpretować

# Rodzaje wykresów

- Rodzaje wykresów i ich warianty
- Do czego przydaje się dany wykres
- Kiedy wykres może być problematyczny

# Wykres słupkowy

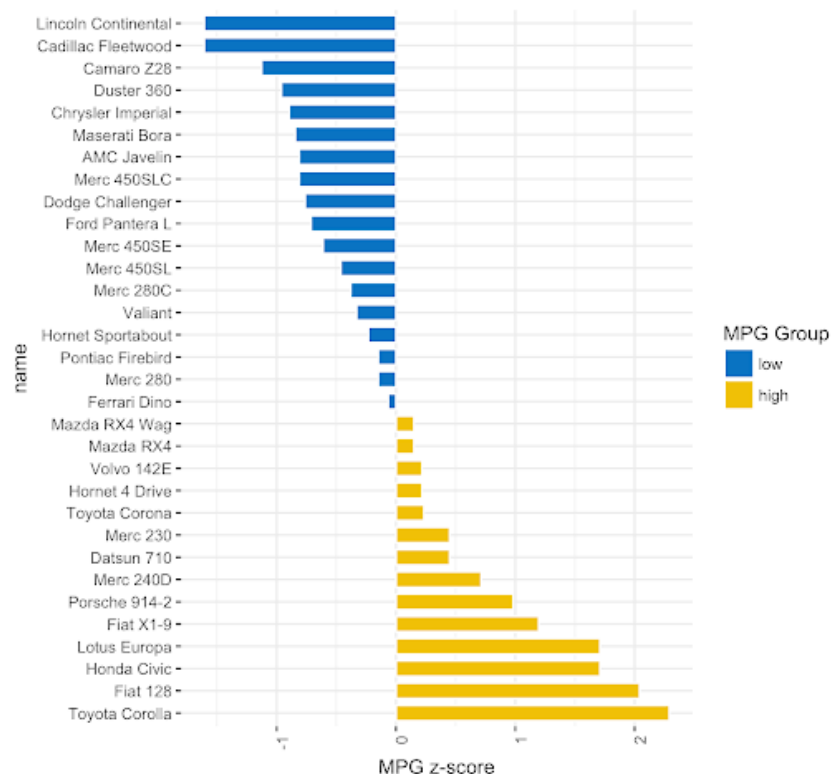
- Jeden z najpopularniejszych wykresów
- Służy do pokazywania danych liczbowych dla poszczególnych kategorii





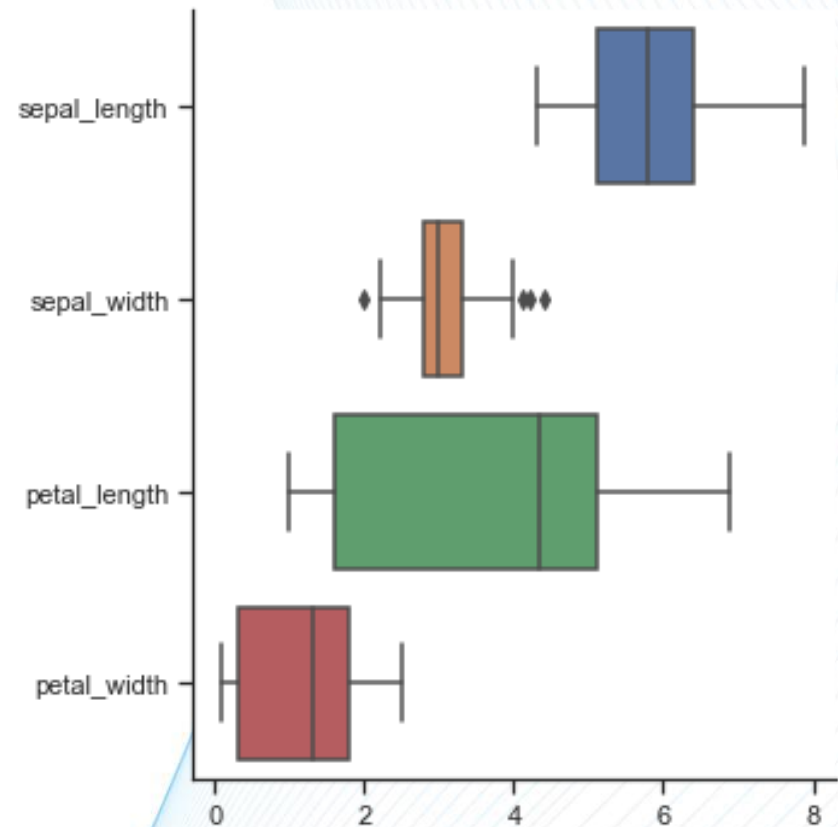
# Wykres słupkowy - horyzontalny

- Stosowany wtedy, gdy nazwy kategorii są długie, ponieważ łatwiej jest czytać poziomy tekst
- Dobry do pokazania odchyleń od 0 lub średniej



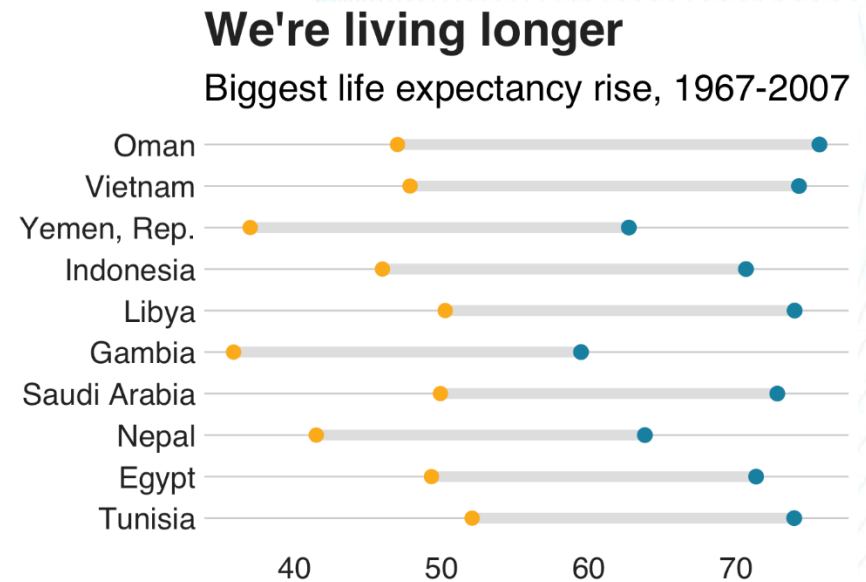
# Wykres pudełkowy

- Zaawansowana wersja wykresu słupkowego, daje nam informacje o rozkładzie zmiennej liczbowej.



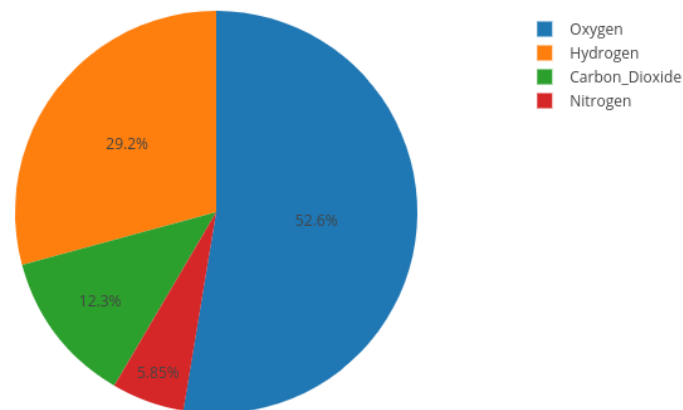
# Wykres dumbbell

- Wykres służy do pokazania zmiany w czasie lub różnicy między kategoriami (kiedyś/dziś, kobiety/mężczyźni)



# Wykres kołowy

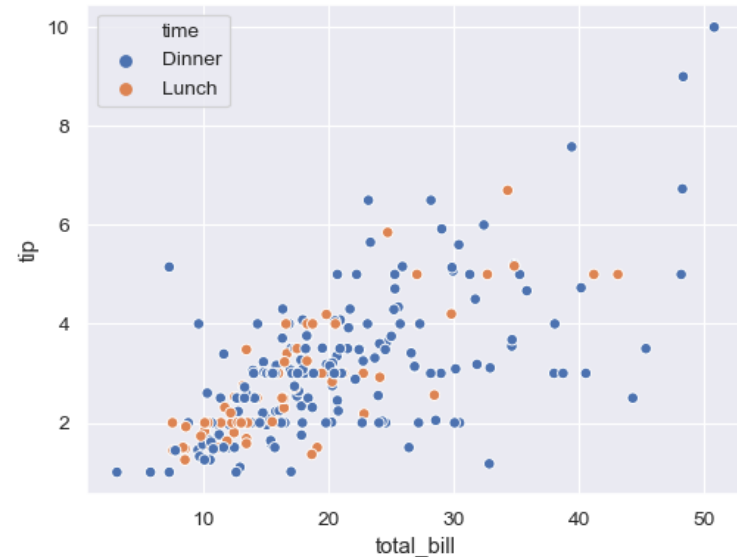
- Wykorzystywany do pokazania udziałów poszczególnych kategorii w całości
- Musi sumować się do 100%
- Jest tylko kilka uzasadnionych przypadków użycia wykresu kołowego





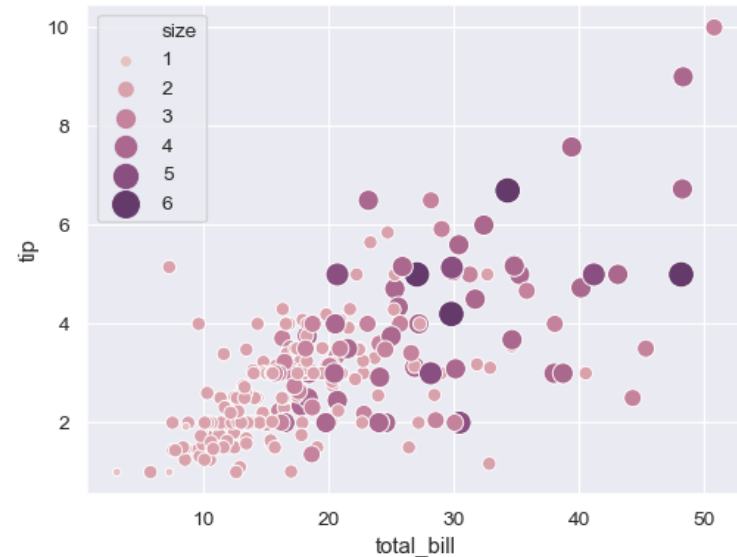
# Wykres punktowy (scatterplot)

- Wykres pokazujący zależności między dwoma zmiennymi liczbowymi.
- Można wprowadzić symbole lub kolory żeby rozróżnić kategorie obserwacji lub kolejną zmienną liczbową



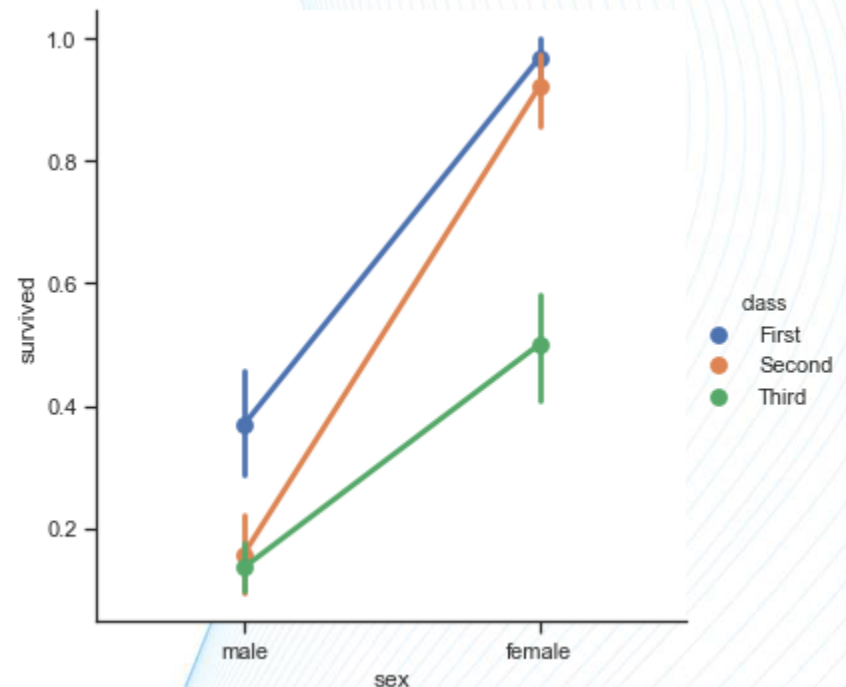
# Wykres bąbelkowy

- Wykres punktowy z dodanym rozmiarem punktów reprezentującym np. populację lub udział dla danego punktu



# Wykres slopegraph

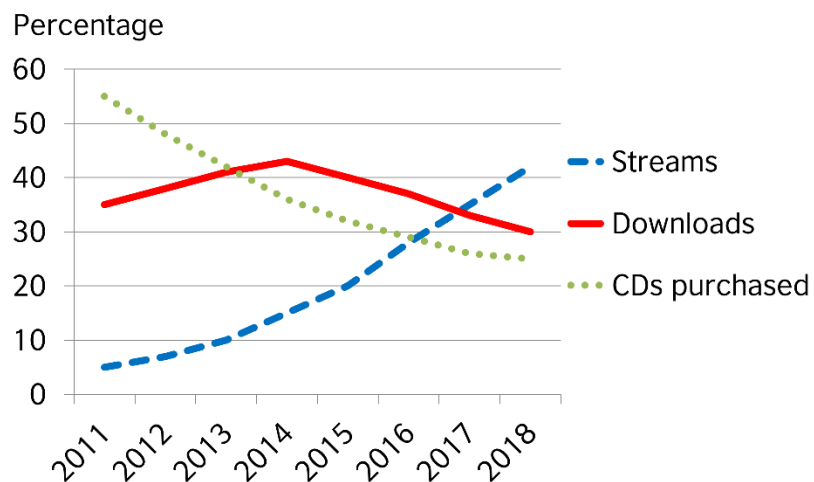
- Wykres pokazujący zmianę w dwóch okresach czasu lub dla dwóch różnych kategorii
- Wygodny sposób do porównania danych



# Wykres liniowy

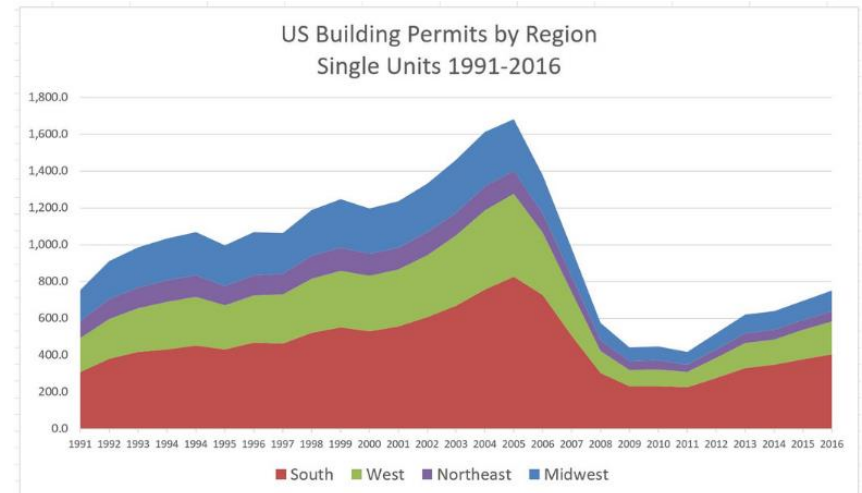
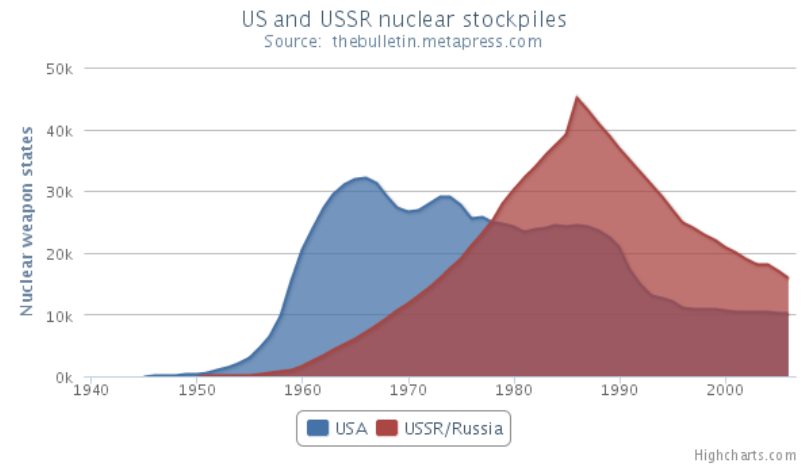
- Wykorzystywany do pokazania zmian w czasie.
- Żeby wykres miał sens na osi X muszą być postępujące po sobie okresy

Percentage of total music sales by method



# Wykres warstwowy

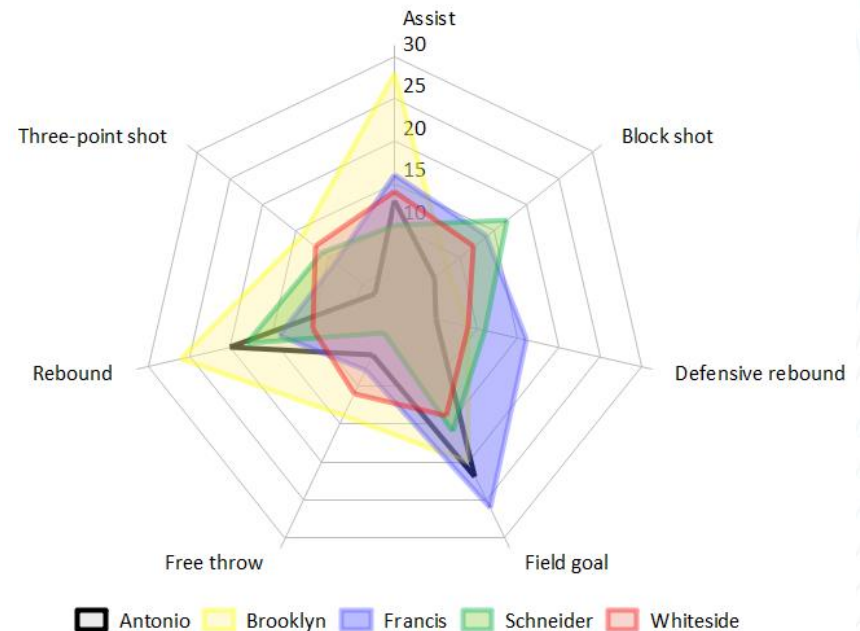
- Warstwy mogą być na sobie lub pod sobą
- Wykres z warstwami pod sobą jest podobny do liniowego
- Wykres z warstwami na sobie jest trudny w interpretacji





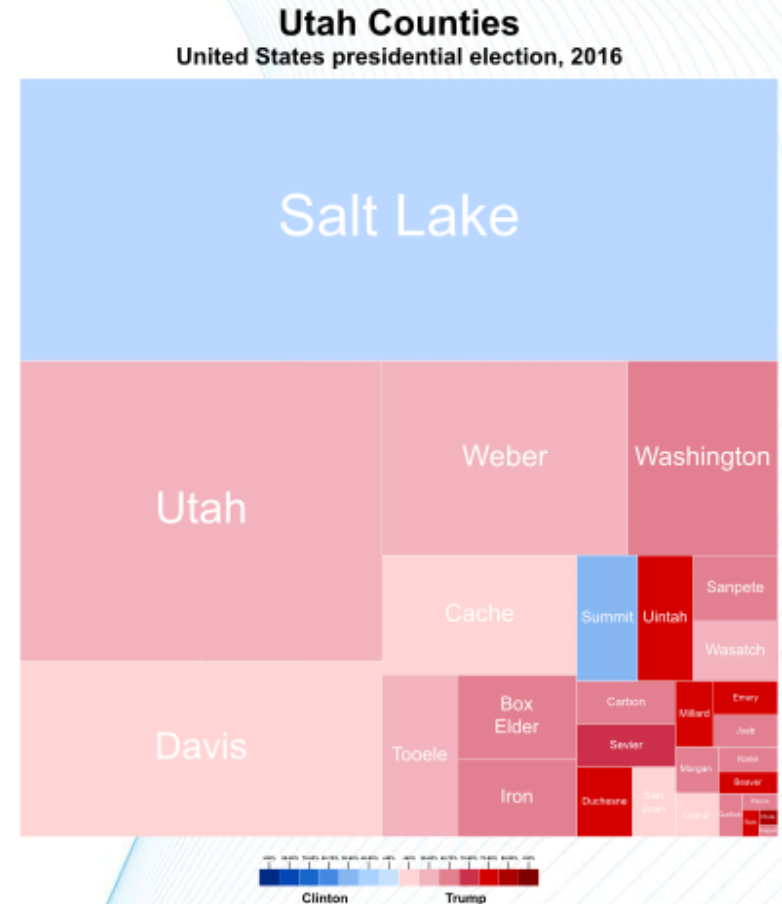
# Wykres radarowy

- Mało kto umie go narysować
- Nikt nie umie go odczytać 😊
- Dane liczbowe dla kilku kategorii
- Wykres jest trudny w interpretacji



# Wykres treemap

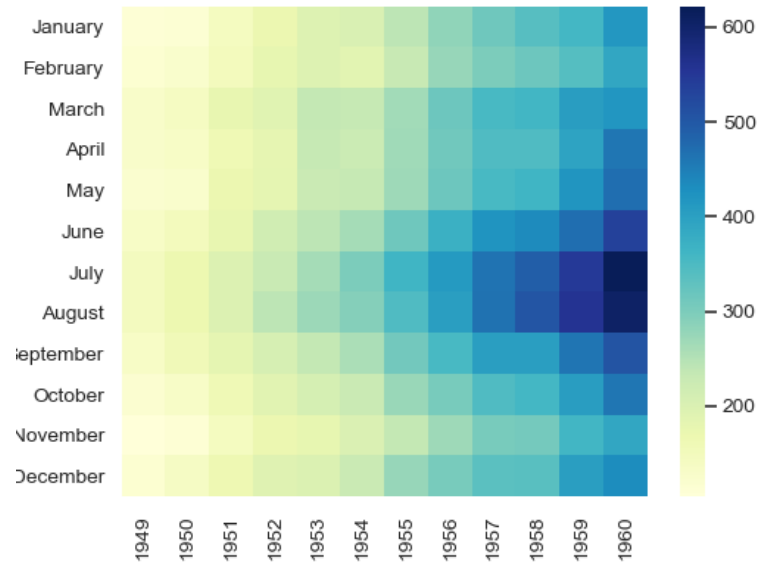
- Służy do pokazywania zależności między zmienną kategorii a inną zmienną z uwzględnieniem udziałów poszczególnych kategorii i ewentualnym grupowaniem





# Mapa cieplna

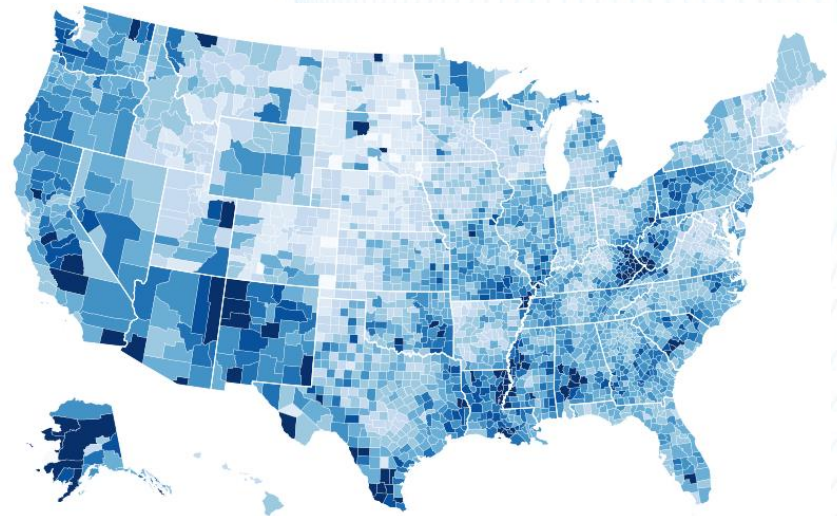
- Służy do pokazywania zależności między dwoma zmiennymi kategoriami (i jedną liczbową)
- Często wykorzystywana do pokazania korelacji





# Mapy

- Świetne kiedy chcemy zwrócić uwagę na różnice i podobieństwa między sąsiadującymi regionami.
- Trudno jest dokładnie porównywać odległe regiony





# Dobra wizualizacja

To taka z której już nie możemy usunąć żadnego komponentu bez utraty informacji. Jednocześnie jej treść i oznaczenia osi są zrozumiałe.

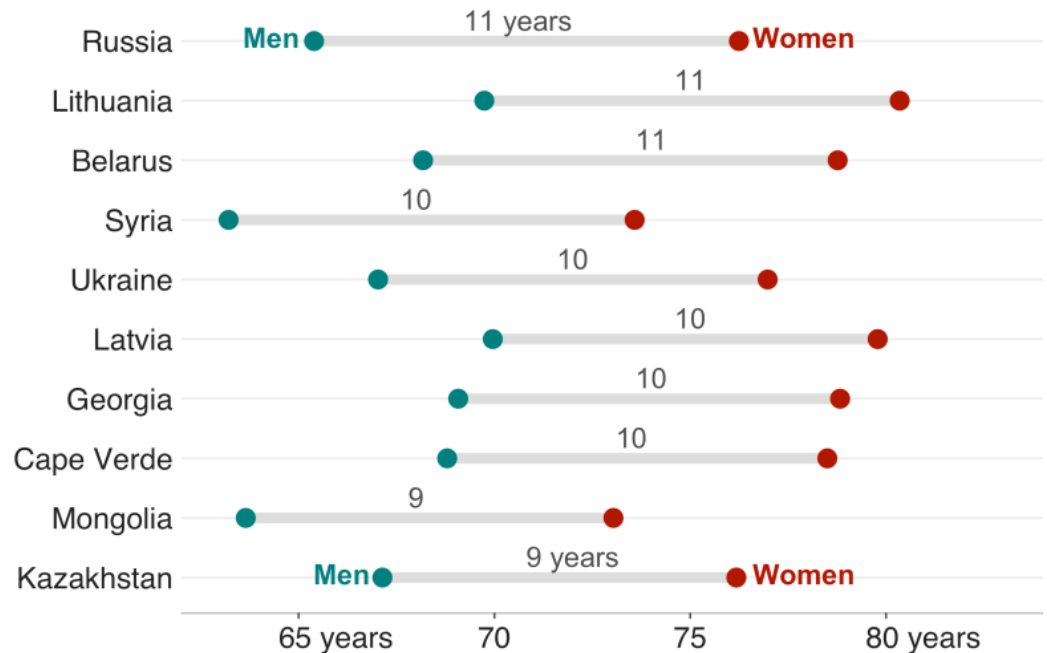
Wizualizacja powinna być zrozumiała nawet gdy wytniemy ją z raportu i otaczającego ją tekstu.

# Elementy wykresu

- Tytuł
- Podtytuł
- Opisy osi
- Skala osi
- Legenda
- Kolory
- Autor i źródło danych

## Eastern European countries have large gender differences

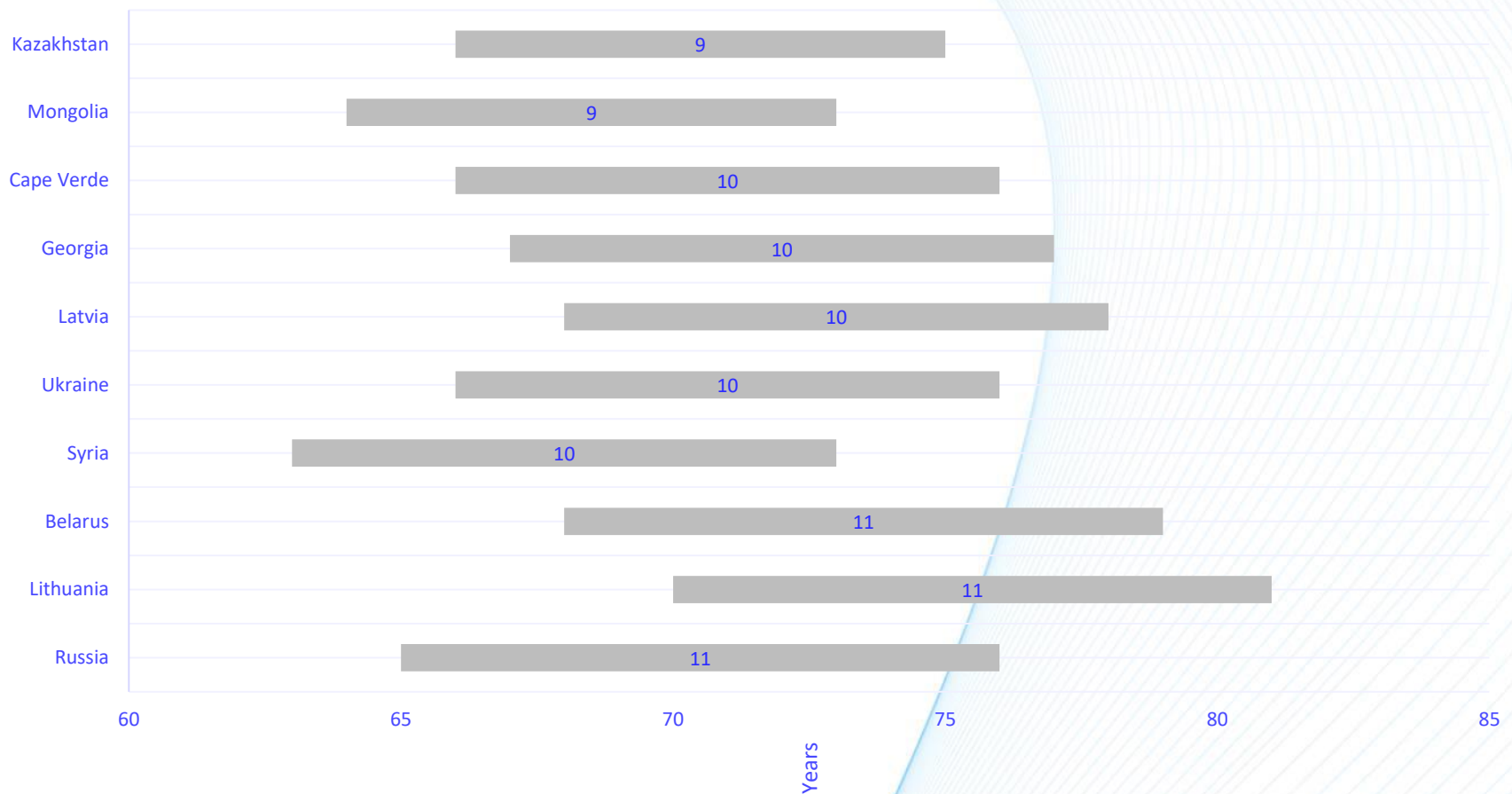
Gap in life expectancy between men and women



Source: IHME

BBC

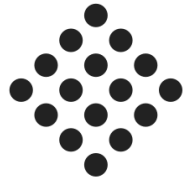
# Ten sam wykres, ale byle jak



# Zasady grupowania

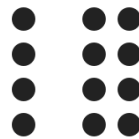
## Gestalt Principles

UX CHEAT.com



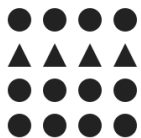
### Good Figure

Objects grouped together tend to be perceived as a single figure. Tendency to simplify.



### Proximity

Objects tend to be grouped together if they are close to each other.



### Similarity

Objects tend to be grouped together if they are similar.



### Continuation

When there is an intersection between two or more objects, people tend to perceive each object as a single uninterrupted object.



### Closure

Visual connection or continuity between sets of elements which do not actually touch each other in a composition.



### Symmetry

The object tend to be perceived as symmetrical shapes that form around their center.



# Zasady grupowania (Gestalt principles)

## Market size over time





# Błędy w wykresach

Niestety często wykresy potrafią nas oszukiwać i wprowadzać w błąd.

Blog smarterpoland co roku ogłasza [plebiscyt](#) na najgorszy polski wykres.

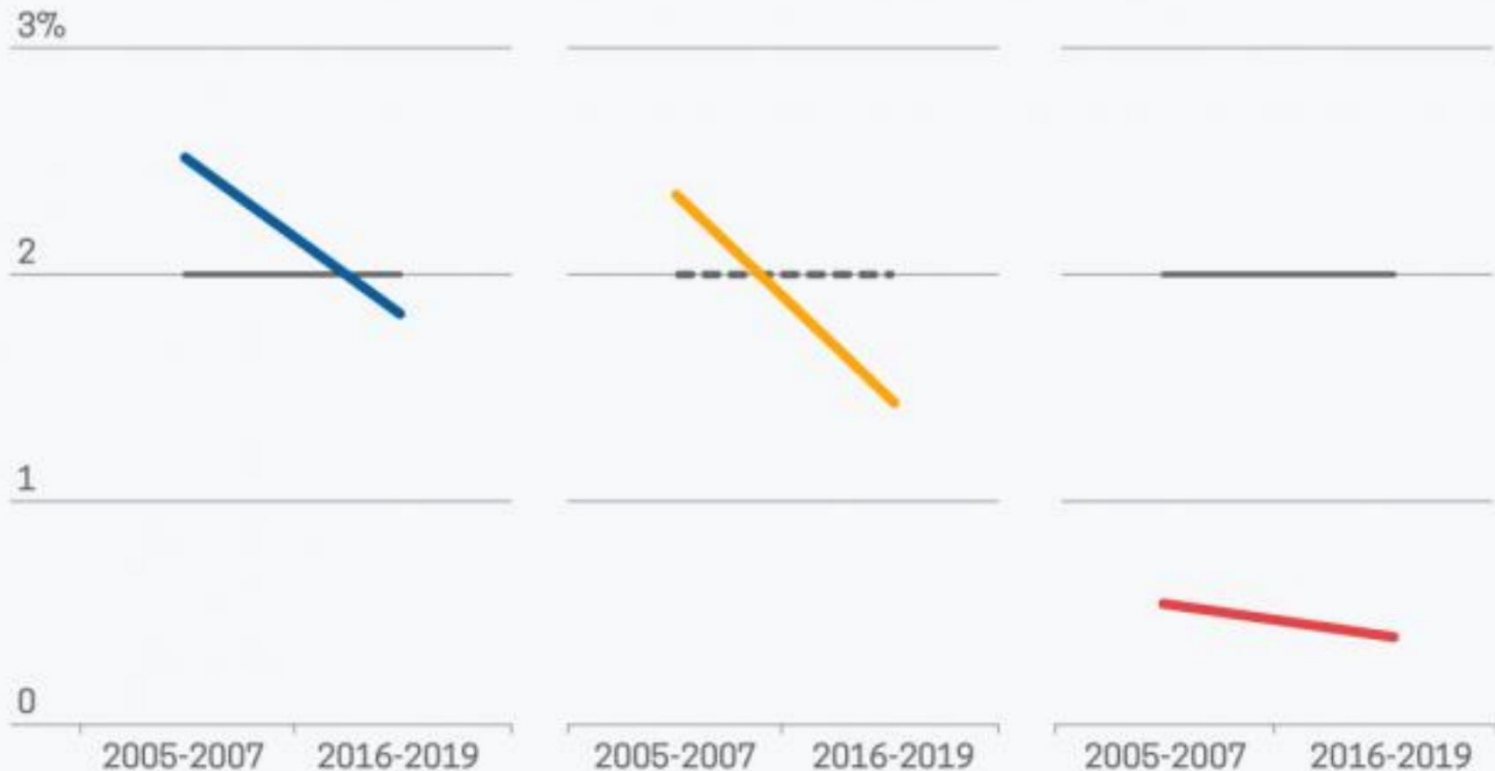
Przyjrzyjmy się kilku przypadkom.

# Co tu jest nie tak?



# Co tu jest nie tak?

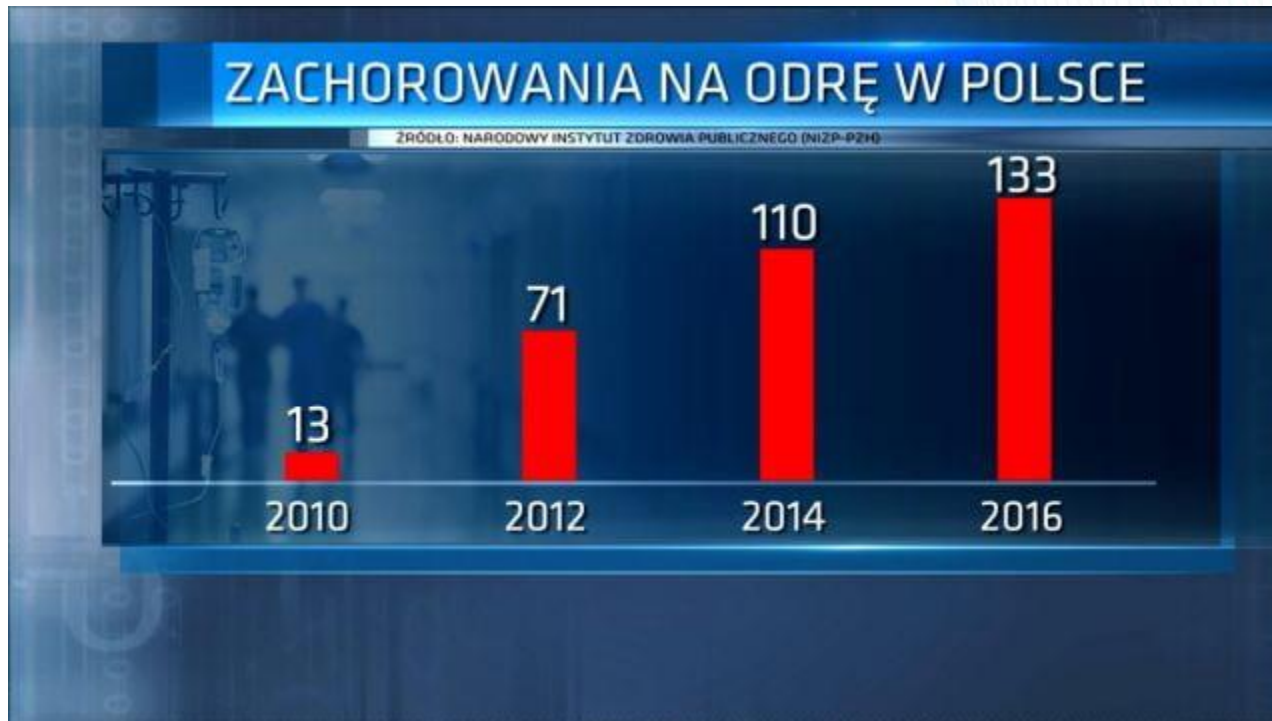
źródło: **obserwatorfinansowy.pl**



Oczekiwania inflacyjne spadają (graf Obserwator Finansowy)

źródło: Inne

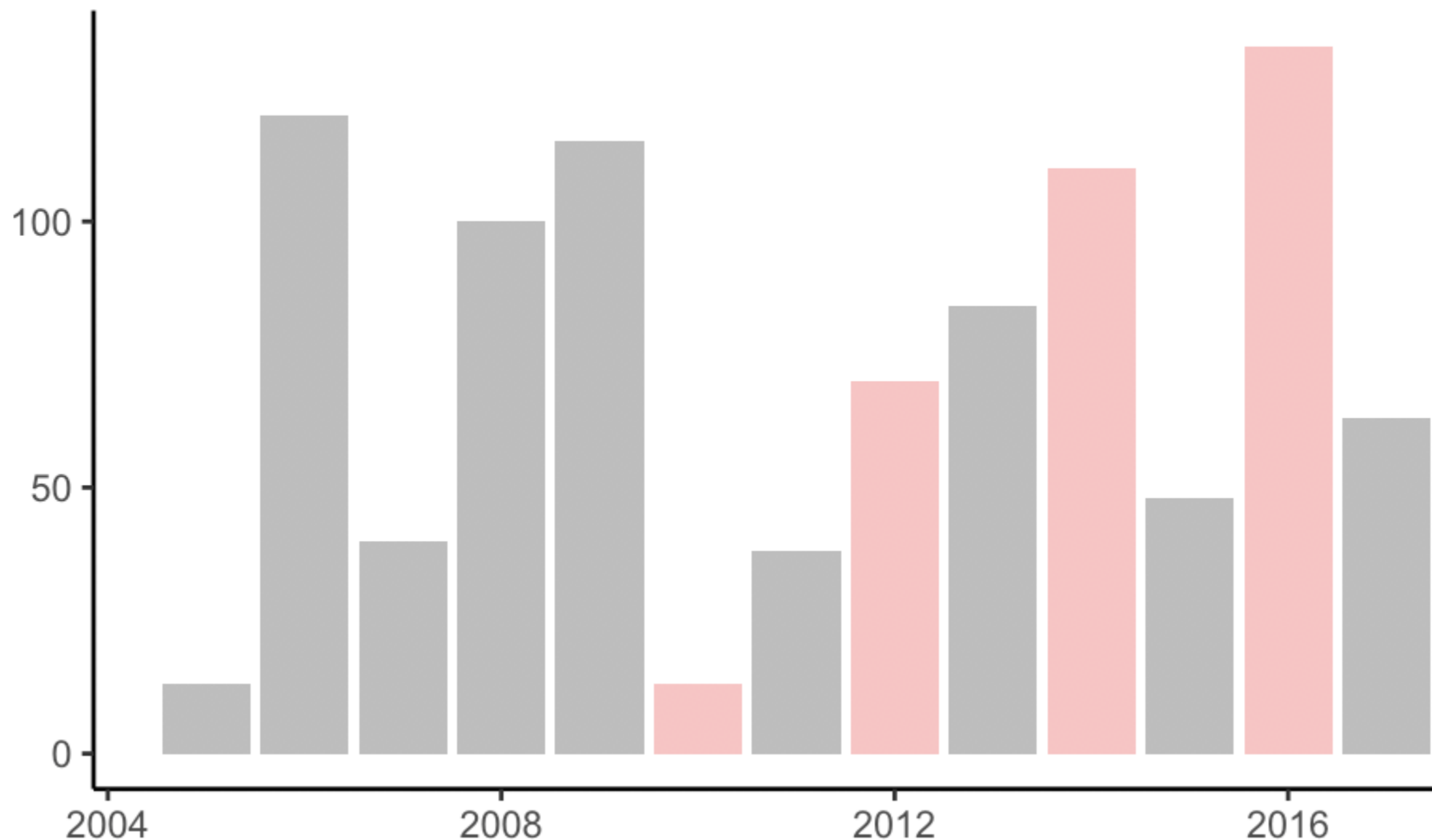
# Co tu jest nie tak?





# Zachorowania na odrę w Polsce

Czerwone słupki to dane pokazane na wykresie TVN



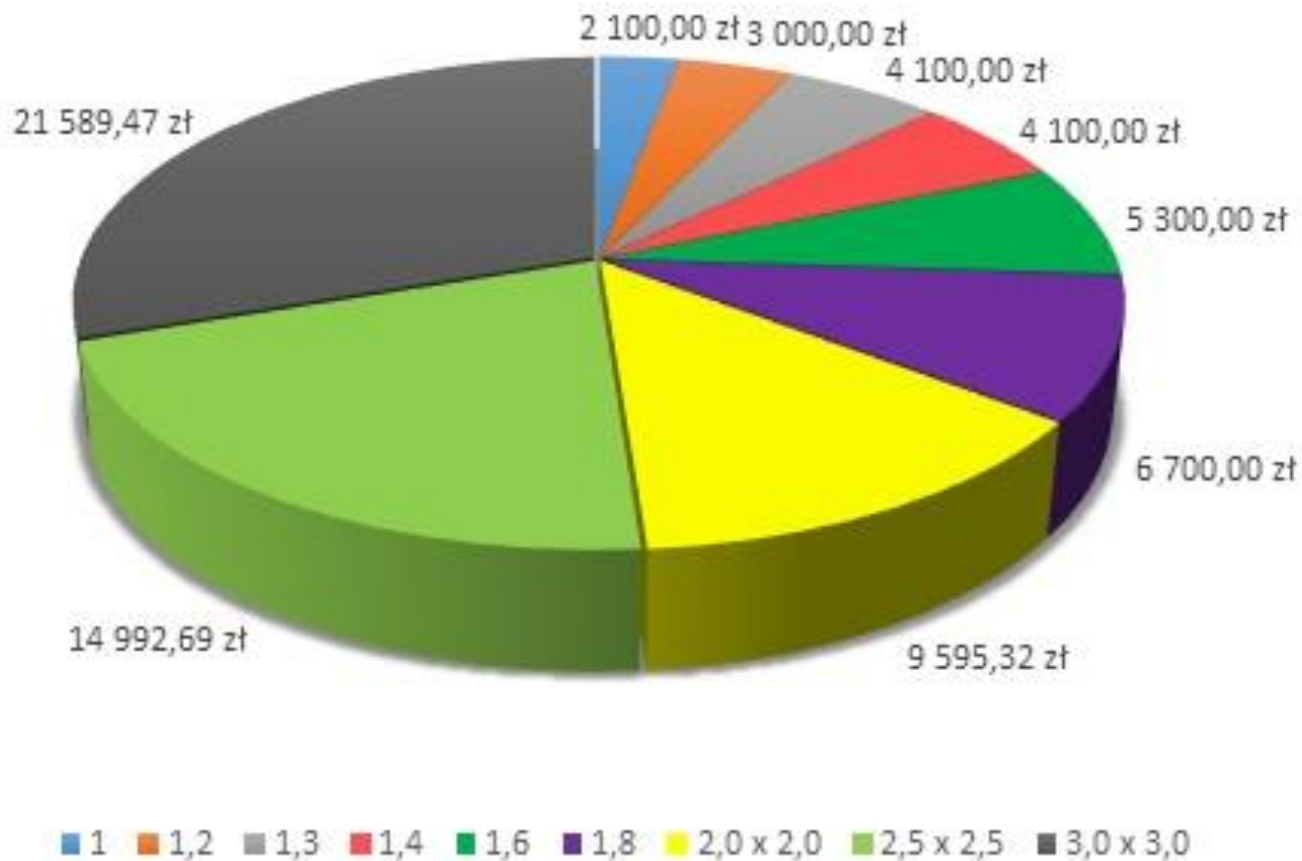


# Co tu jest nie tak?



# Co tu jest nie tak?

Cena netto 1 mb konstrukcji przepustu o określonym przekroju  $\emptyset$  [m]



# Co tu jest nie tak?

rowadzenie postępowań przygotowawczych w sprawach ka  
rokuraturze Krajowej



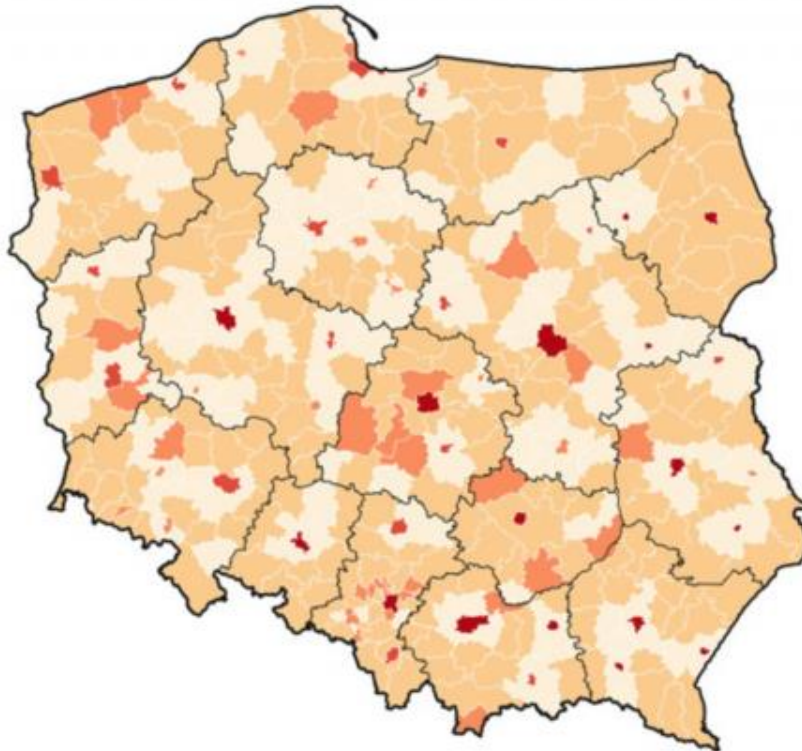


# Co tu jest nie tak?

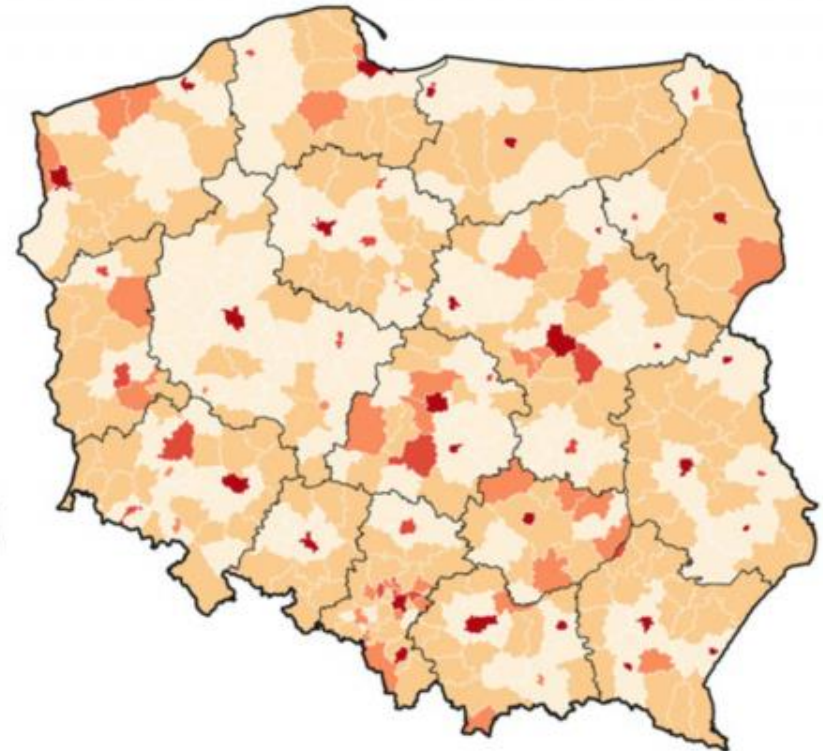
Liczba lekarzy na 1000 mieszkańców  
wg głównego miejsca zatrudnienia



2006



2015

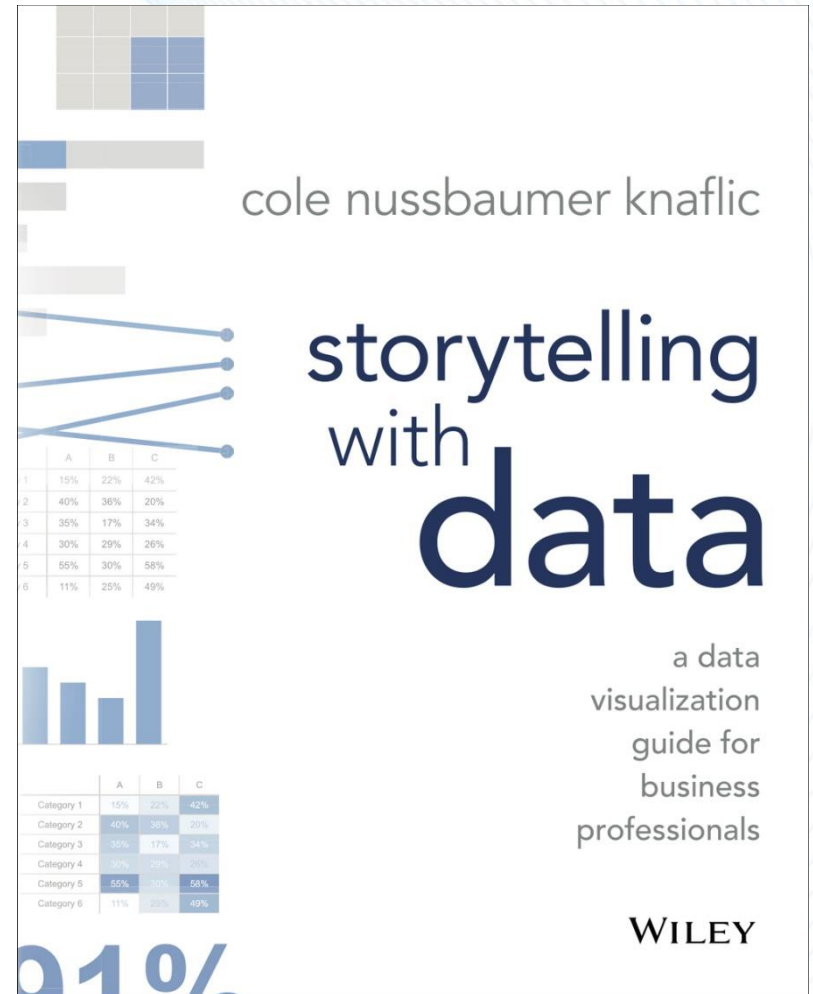




# Więcej na ten temat

## Storytelling with data

Autor Cole Nussbaumer Knaflic



# Ćwiczenie

- W folderze ćwiczenia znajduje się plik wykresy. Zobaczmy jak możemy efektywnie zwizualizować różne rodzaje danych.