

# Techniki wizualizacji danych

## HW8

### Dominik Kędzierski

Celem pracy domowej jest opisanie zbioru danych houses

#### Wstęp

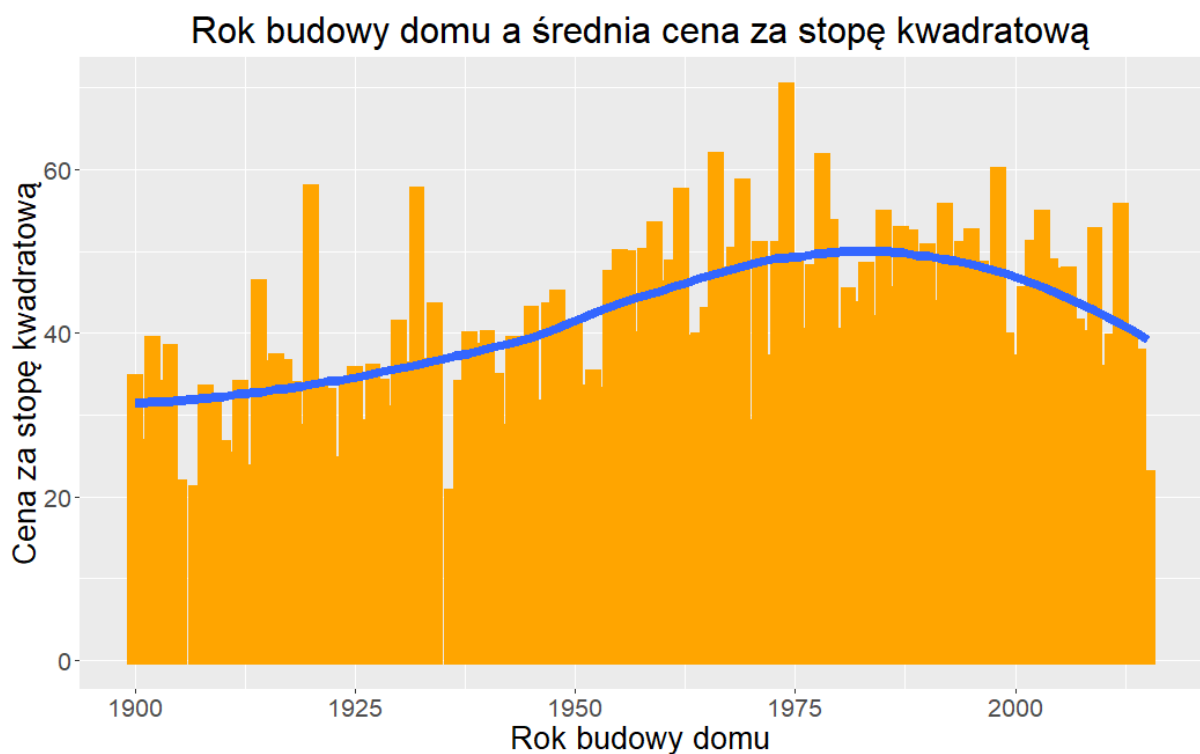
Zestaw houses opisuje informacje na temat mieszkań, zawiera między innymi takie informacje jak:

- 1) Rok budowy
- 2) Ilość sypialni
- 3) Ilość łazienek
- 4) Powierzchnia
- 5) Cena
- 6) Kod pocztowy (ZipCode)
- 7) Ilość pięter

Na podstawie tych informacji można zwizualizować kilka zależności między nimi w celu wyciągnięcia wniosków, które mogą potencjalnie pomóc w szukaniu domu który spełnia nasze oczekiwania.

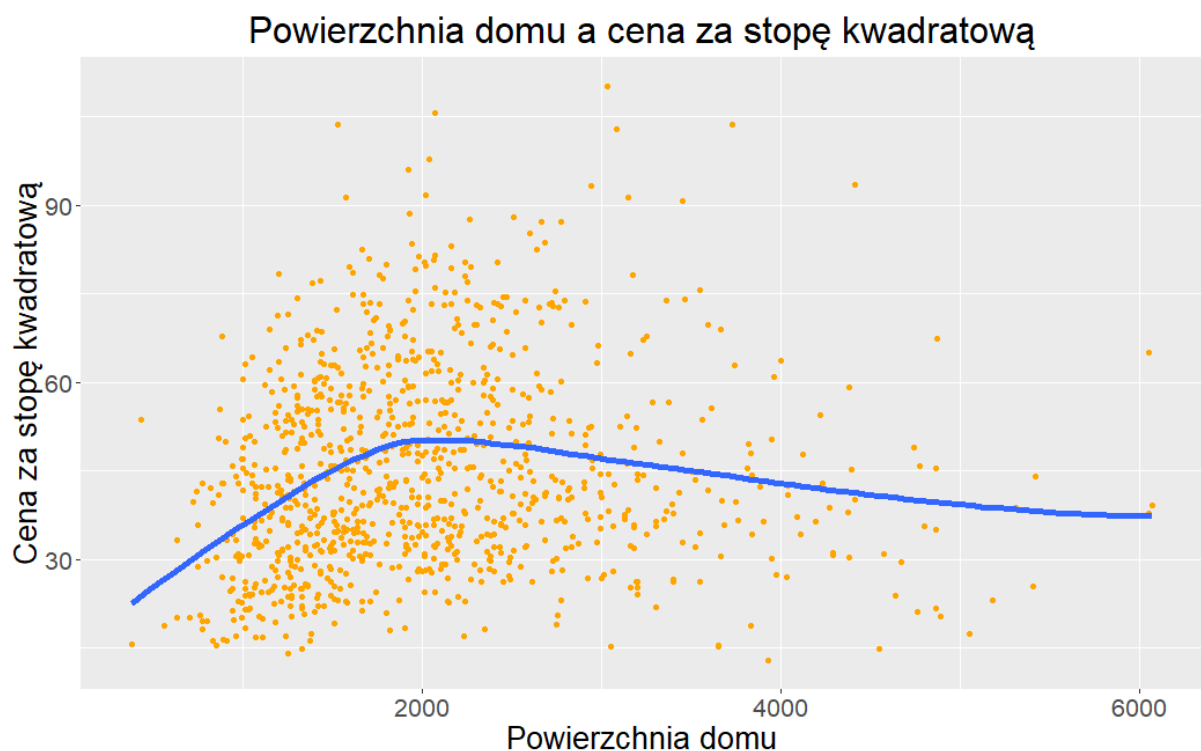
#### Analiza danych

- 1) Pierwszą rzeczą jaką możemy porównać jest zależność między powierzchnią mieszkania a ceną za stopę kwadratową.



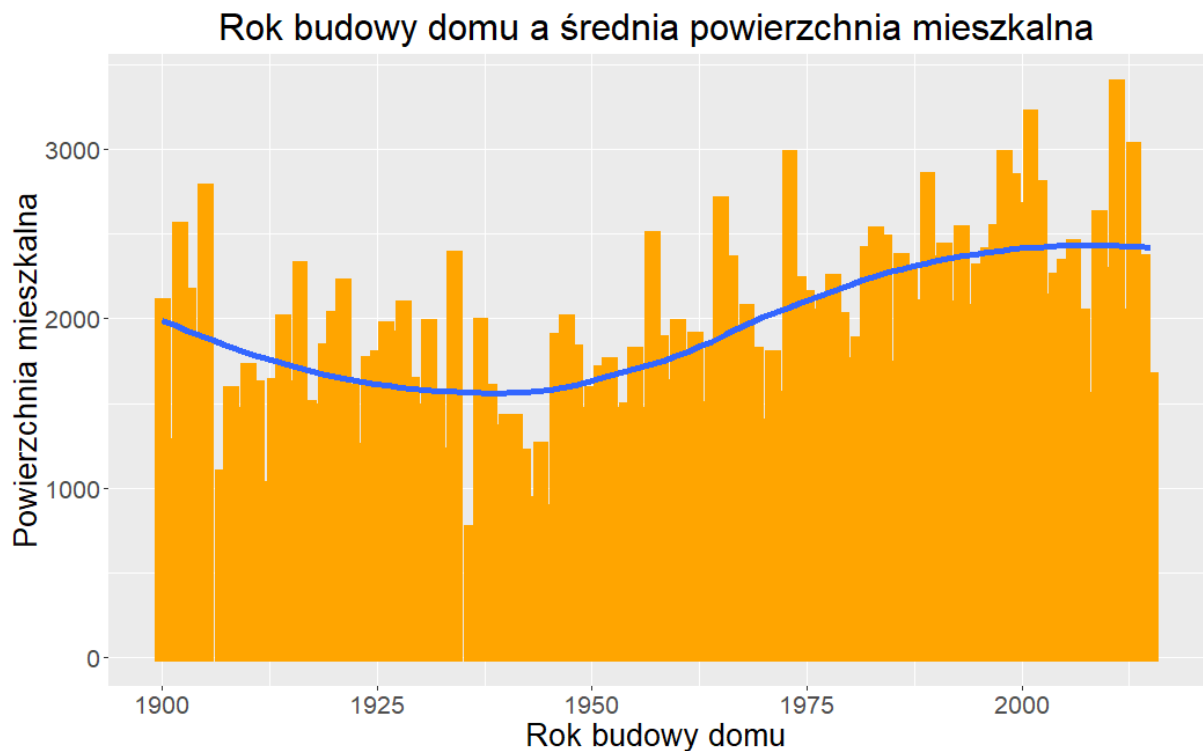
Na podstawie wykresu można zauważyć, że cena za stopę kwadratową jest najniższa dla najstarszych domów, następnie rośnie do lat 80, po czym dosyć maleje. Może być to spowodowane tym, że duża część osób woli wybudować sama swój dom niż kupić taki, który ma dosyć mało lat. Jednak w kwestii opłacalności najlepiej albo kupować stare mieszkania albo bardzo nowe.

- 2) Następną sprawą jest wpływ powierzchni domu na cenę za stopę kwadratową domu.



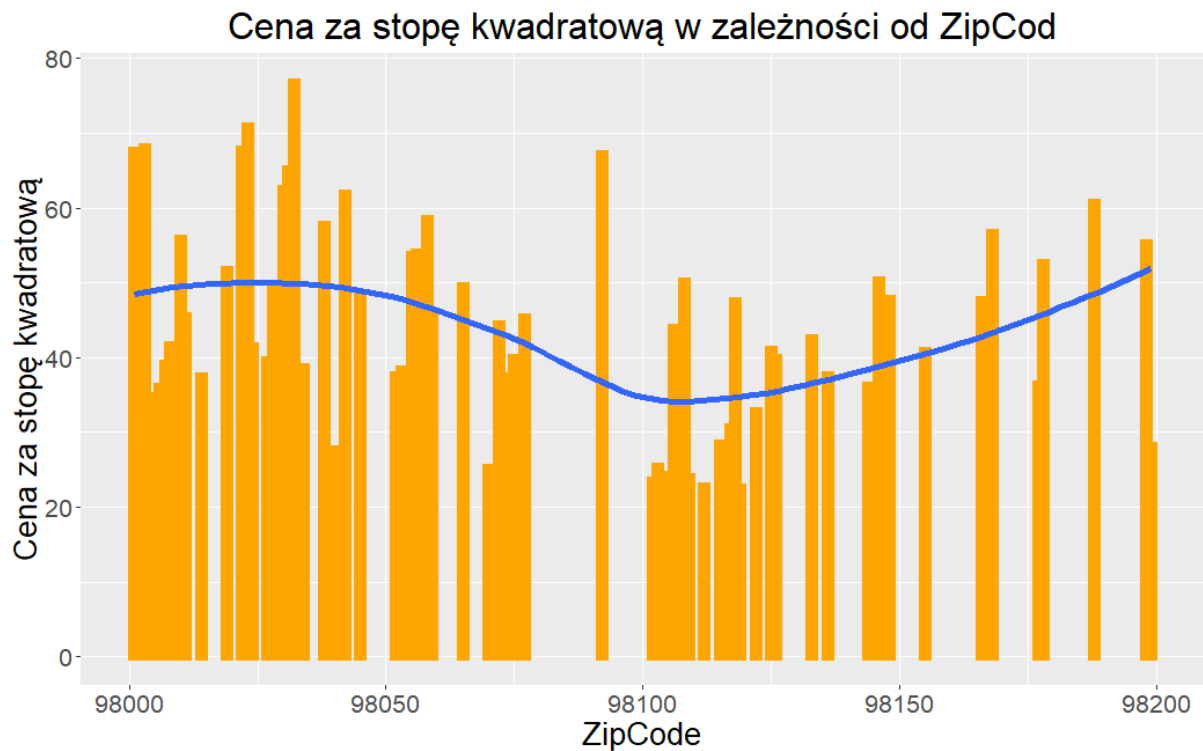
Jak widać najmniejsza cenę za powierzchnię kwadratową mają małe domy, może być to spowodowane małym zainteresowaniem takich domów przez kupujących. Najdrożej wychodzą domy, które mają około 2000 stóp kwadratowych, co można przypuszczać, że są one najczęściej kupowanymi domami. Im większa powierzchnia niż około 2000 stóp kwadratowych, tym średnia cena maleje, jednak nie osiąga ceny porównywalnej dla bardzo małych domów.

- 3) Kolejną analizą jest zależność pomiędzy powierzchnią domu a rokiem jego wybudowania.



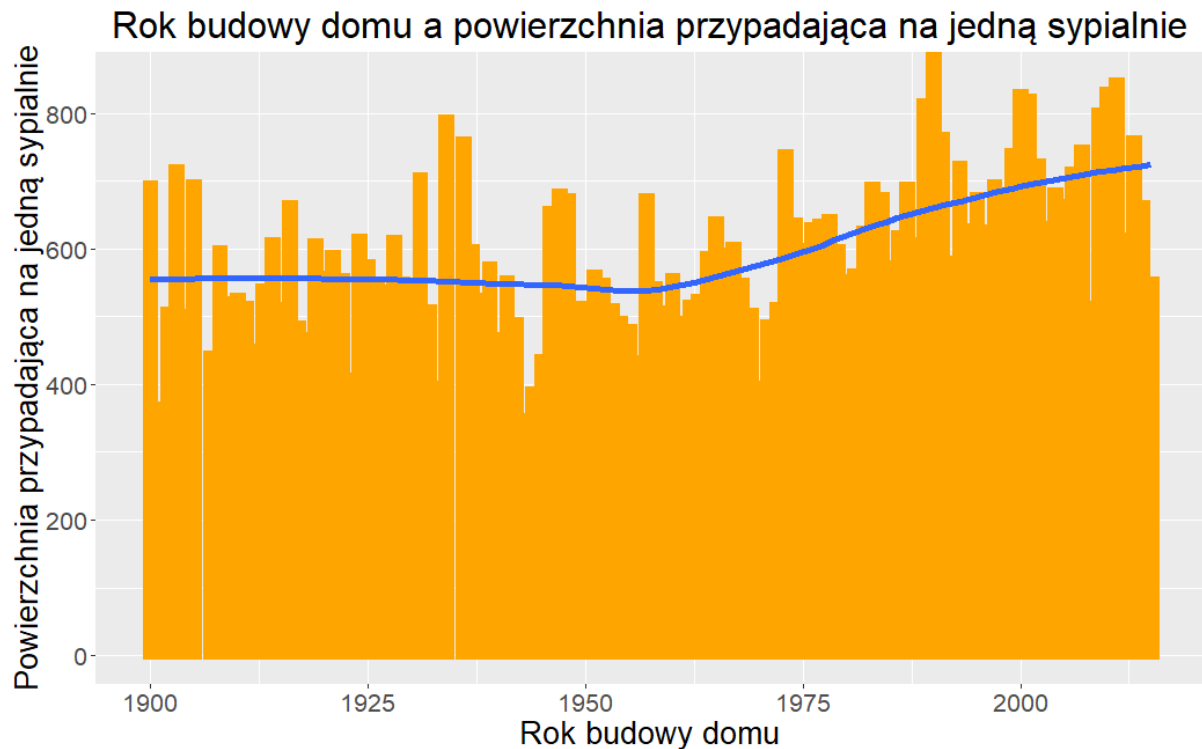
Jak widać bardzo stare domy mają dosyć dużą powierzchnię, następnie do około lat 40 średnia powierzchnia domów maleje po czym zaczyna rosnąć aż do obecnych czasów. Trend ten poza niektórymi domami wydaje się być dalej aktualny co oznacza, że nowo budowane domy są coraz większe.

- 4) Czwartym aspektem jest wpływ miejsca zamieszkania na średnią cenę powierzchni. Można to porównać na podstawie kodu pocztowego danego domu.



Z wykresu widać zależność pomiędzy kodem pocztowym a ceną. Nie mamy danych co oznaczają dokładnie te kody. Które z nich to są duże miasta, a które to tereny poza miejskie, jednak widać że największa cena jest tam gdzie, ostatnie trzy cyfry kodu pocztowego mają wartość najmniejszą. Można przypuszczać że jest to teren bliski jakiemuś dużemu miastu.

- 5) Ostatnia analiza dotyczy tego jak zmieniła się powierzchnia przypadająca na jedną sypialnię w zależności od roku budowy domu.



Wnioskiem z tego wykresu jest to że, do lat 60 mniej więcej powierzchnia przypadająca na jedną sypialnię była dosyć równa, jednak od tych lat powierzchnia ta systematycznie rośnie co jest analogiczne do rośnięcia ogólnej powierzchni domów w obecnych czasach

## Podsumowanie

Z powyższego zbioru danych możemy wiele dowiedzieć się na temat średnich parametrów domów. Może to pomóc w określeniu jakiego typu domu powinniśmy szukać. Np. Jeżeli szukamy dużego domu w którym powierzchnia na sypialnię będzie duża to należy szukać domów które są dosyć młode, ponieważ mają one najlepszy stosunek ceny spełniający nasze kryteria.