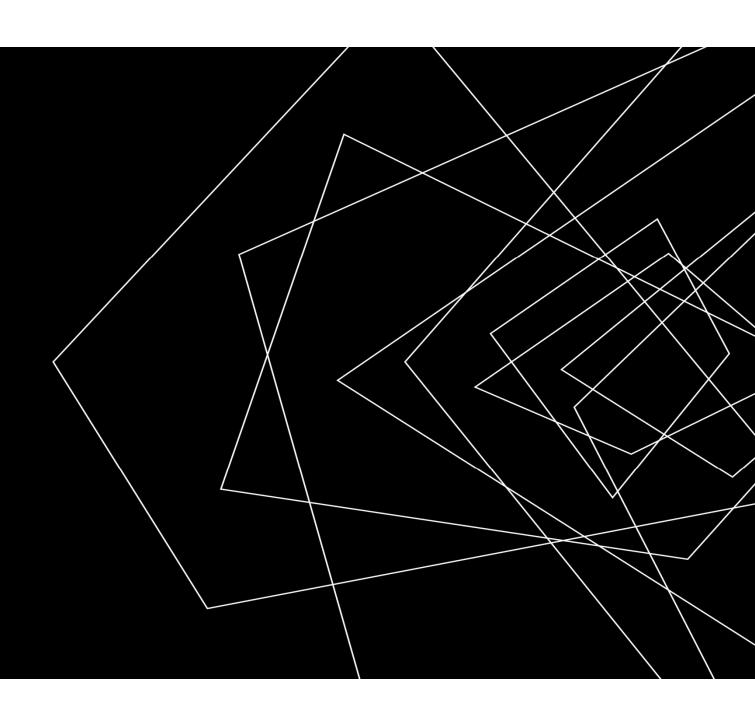
#### PREZENTACJA I WIZUALIZACJA DANYCH W R

Marek Harhala,

Instytut Biologii i Ewolucji Człowieka,

2.42

marek.harhala@amu.edu.pl

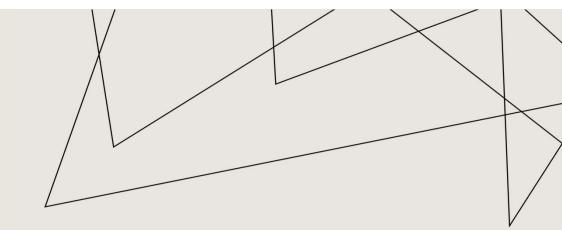






#### 4 pierwsze spotkania

Projekty na zajęciach robione na zaliczenie.

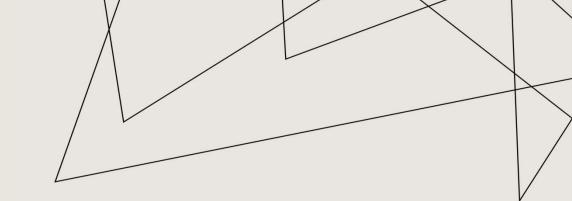


#### Pozostałe spotkania

Projekty na zajęciach robione na ocenę.

2-3 projekty nt. wizualizacji danych.

1 prezentacja pod koniec kursu.



## ORGANIZACJI ZAJĘĆ

#### Pojedyncze zajęcia:

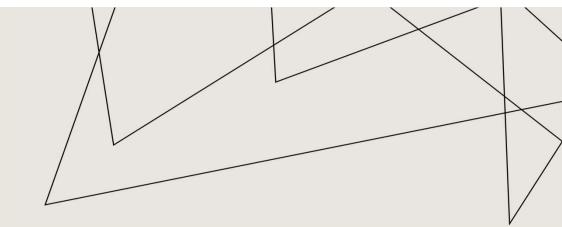
- wstęp do tematu,
- część 'filozoficzna',
- część 'techniczna',
- ćwiczenie na zajęciach,
- projekt na zajęciach na zaliczenie/ocenę.



#### Filozoficzna – DLACZEGO powinno się to robić?

Ma za zadanie pokazać różne sposoby myślenia o wizualizacji i prezentacji danych naukowych.

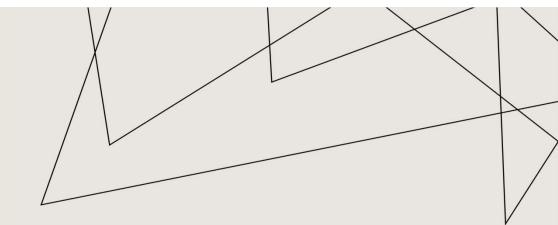
Na zakończenie powinniście zdawać sobie sprawę z podstawowych aspektów prezentacji danych i mieć wyrobione możliwości analizy procesu prezentacji danych pod wieloma względami.

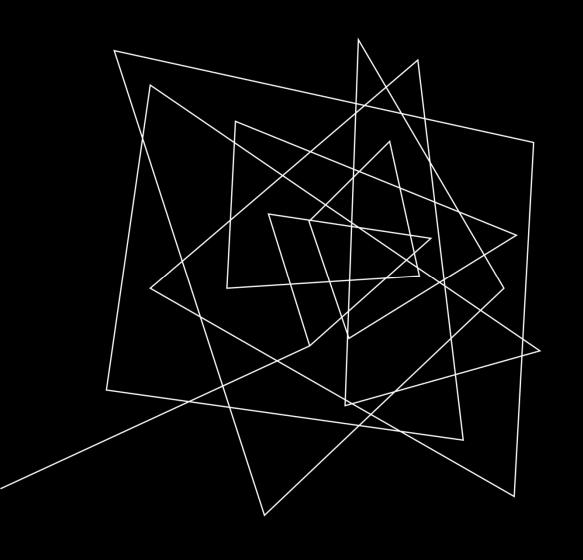




#### Bardziej naukowo-techniczna – JAK to zrobić.

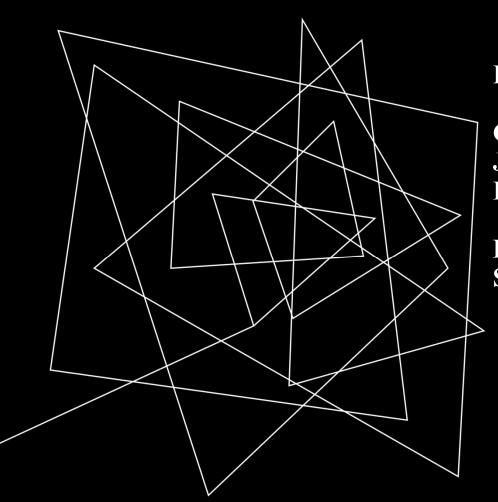
Celem jest pokazanie elementów powiązanych z konkretnymi aspektami i jak wykorzystać je do prezentacji danych.





## PIERWSZE SPOTKANIE

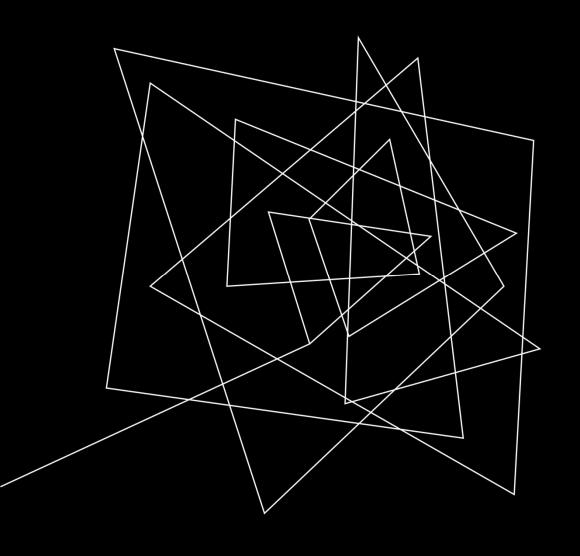
# KOMUNIKACJA!



#### PIERWSZY PUNKT WIDZENIA:

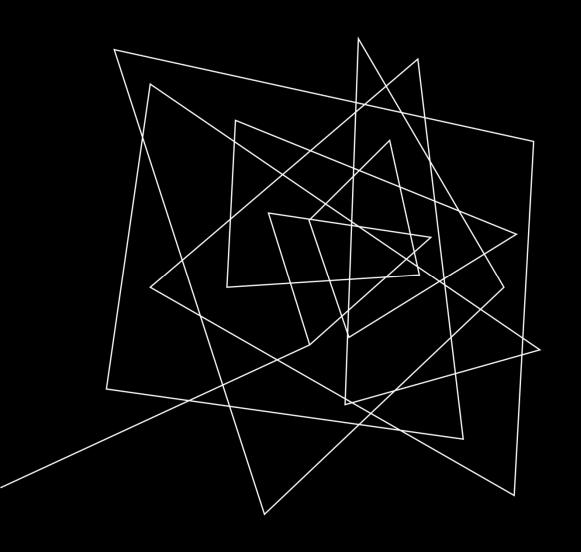
GRAFICZNA PREZENTACJA DANYCH JEST DLA LUDZI W CELACH KOMUNIKACJI.

DO CELÓW NAUKOWYCH JEST STATYSTYKA.

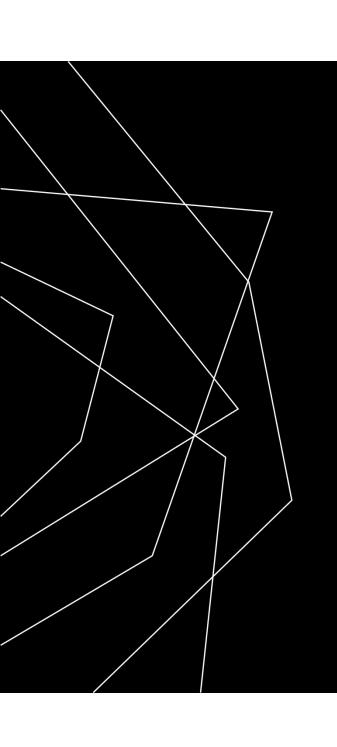


DRUGI PUNKT WIDZENIA:

EXPLANATORY VS. EXPLORATORY



TRZECI PUNKT WIDZENIA
'STORY DRIVEN DATA
VISUALIZATION'
LUB 'DATA-DRIVEN STORIES'



WSPÓLNE ELEMENTY?
WIZUALNE ASPEKTY
SKUPIAJĄ NASZĄ PERCEPCJĘ
NA WZORCACH W DANCYH

Zanim pojawią się opowieści, są wydarzenia, które dają im początek To właśnie od nich zaczynamy: od postrzegania i rozumienia zdarzeń.

## DOŚWIADCZENIE PRZED ZROZUMIENIEM?

Zazwyczaj tak, ale doświadczenie wizualne może być poprzedzone planowaniem, opowieściami, wprowadzeniem teoretycznym, nastrojem, itp.

Elementy poprzedzające (głównie) i poprzedzające prezentację danych mają wpływ na wyciągnięte wnioski (m. in. błędy poznawcze).

## JAK LUDZIE DOŚWIADCZAJĄ?

'Breakpoints' - kiedy ludzie uznają, że coś się nowego zaczęło? Na generalnym (wysokim, dużym) poziomie zaczyna się coś nowego, kiedy zmienia się obiekt, z którym *coś* się dzieje.

Na szczegółowym (niskim, małym, chwilowym) poziomie coś nowego się zaczyna, gdy czynność wykonywana na jakimś obiekcie się zmienia.

## PORYWAJĄCA OPOWIEŚĆ

42% opowieści jest niedokładna wg. opowiadających te opowieści.

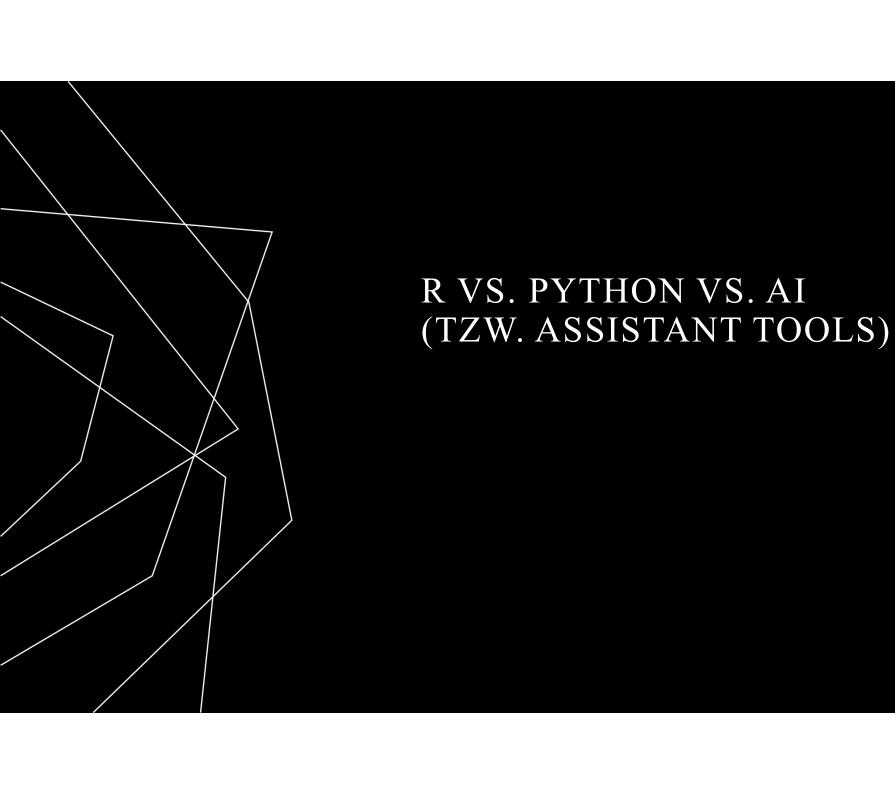
'Data-driven storytelling' wymaga wyższego podobieństwa do prawdy.

### INNE SPOSOBY PRZEKAZYWANIA DANYCH NAUKOWYCH (POZA 'STORY')

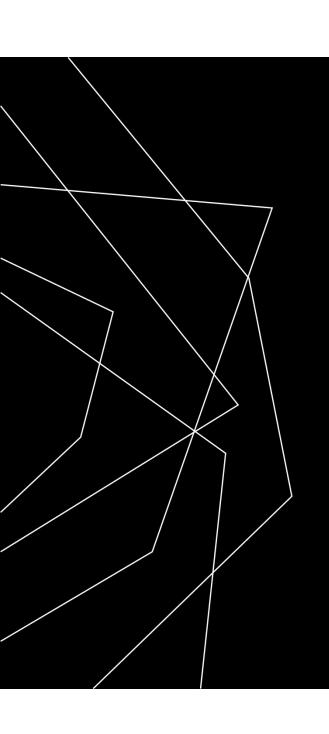
twarzą-w-twarz: komunikacja o zabarwieniu naukowym często odbywa się twarzą-w-twarz, głównie podczas rozmów lub prezentacji – jest wtedy wzbogacana o elementy jak, gestykulacja, emocje, ton głosu, itp.

#### VISUAL STORYTELLING

Dane są prezentowane na płaszczyźnie 2D, rzadziej na 1D, o wiele rzadziej prezentuje się je na płaszczyźnie 3D.

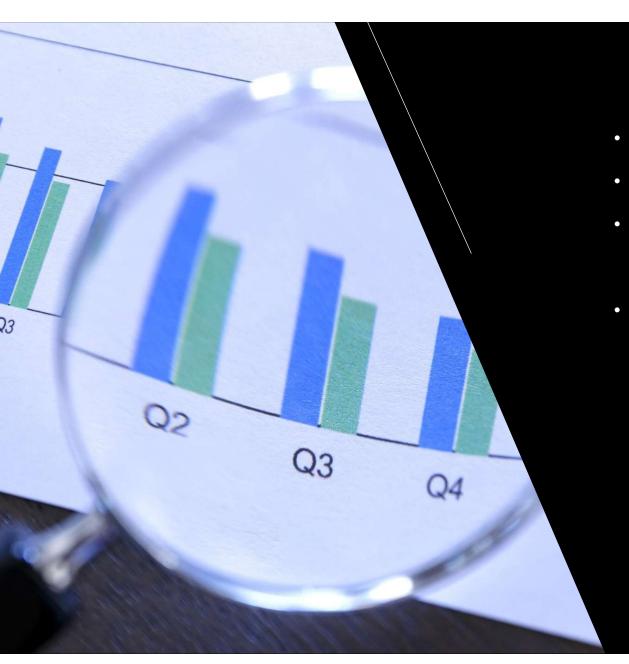




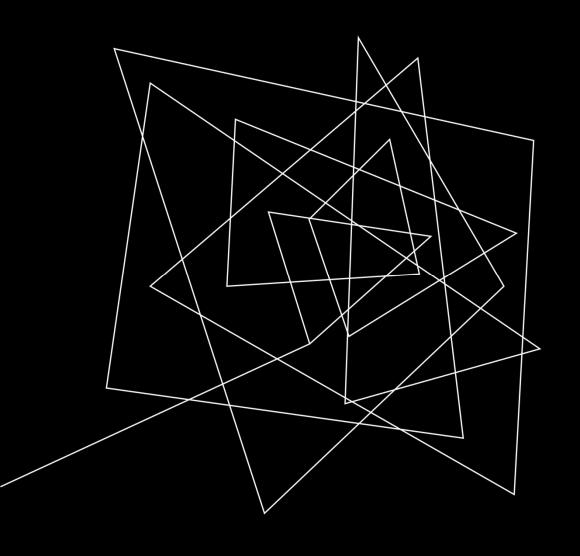


## RODZAJE DANYCH

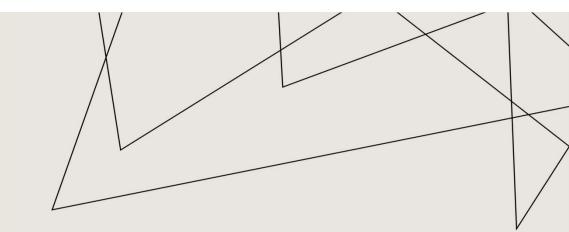
- binarne/dychotomiczne (binary)
- kategoryczne (categorical),
- porządkowe (ordinal),
- ciągłe (continouuuous).



- rzeczywiste wartości takie, które istnieją niezależnie od pomiaru.
   wartości danych zarejestrowane
- wartości danych zarejestrowane pomiary rzeczywistych wartości.
   wartości danych w pliku
- wartości danych w pliku komputerowym – wartości przechowywane w formie cyfrowej w pliku.
- wartości danych w systemie analitycznym (np. r) wartości wczytane i używane w środowisku analiz.



## DATA AS FACTORS IN R

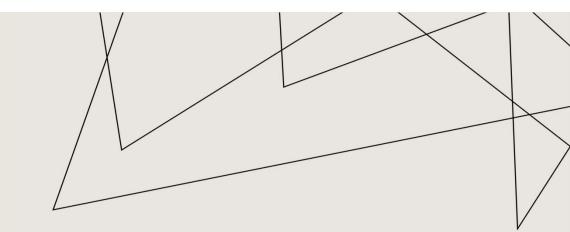


#### WCZYTYWANIE I WSTĘPNE EKSPLOROWANIE DANYCH

- WCZYTANIE DANYCH.
  - Załaduj wbudowany zbiór danych (np. iris).
- SPRAWDZENIE UMIEJĘTNOŚCI
  - Wyświetl pierwsze kilka wierszy zbioru danych (np. funkcja do wyświetlania początkowych rekordów).
  - Sprawdź strukturę zbioru danych (informacje o typach zmiennych).
  - Wykonaj statystyki opisowe, aby zapoznać się z podstawowymi cechami zbioru (liczba obserwacji, liczba zmiennych, minimalne i maksymalne wartości itp.).

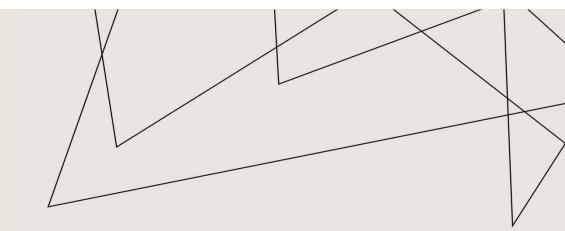
#### ZADANIA

- Odczytaj liczbę wierszy (obserwacji) i kolumn (zmiennych).
- Napisz, które kolumny mają jaki typ danych.
- Zwróć uwagę, czy w danych występują wartości brakujące.

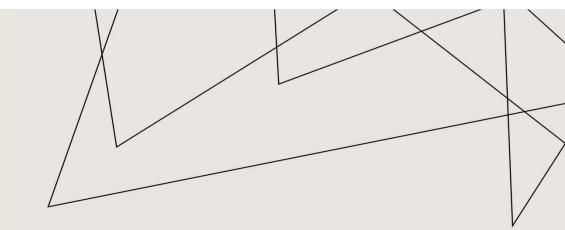


#### 2. Podstawowe wizualizacje (histogramy, wykresy słupkowe)

- · Histogram zmiennej liczbowej
  - Wybierz jedną zmienną liczbową i narysuj jej histogram.
- Wykres słupkowy (barplot) zmiennej kategorycznej
  - Wybierz zmienną kategoryczną i narysuj wykres słupkowy.
  - Napisz, która kategoria pojawia się najczęściej.
- Wykres gęstości (opcjonalnie)
  - Dla tej samej zmiennej liczbowej co w histogramie możesz również narysować wykres gęstości, aby porównać kształt rozkładu z histogramem.

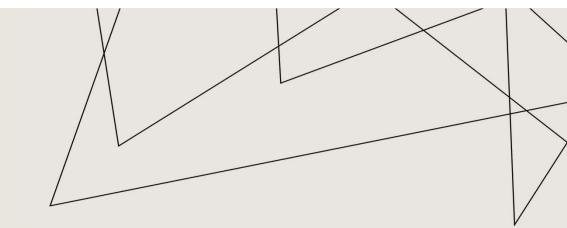


- 3. Wizualizacja rozkładów (box plot) i porównania grup
- · Box plot dla jednej zmiennej liczbowej
  - Narysuj wykres pudełkowy (box plot) wybranej zmiennej liczbowej, aby zaobserwować medianę, kwartyle i ew. wartości odstające.
- Box plot w podziale na grupy
  - Podziel dane na kategorie (np. płeć lub gatunek) i porównaj wartości zmiennej liczbowej między tymi grupami.
  - Zastanów się, która grupa ma wyższą medianę i czy rozrzut wartości w różnych grupach jest podobny.



#### 4. Wykresy punktowe (scatter plots) i relacje między zmiennymi

- Podstawowy wykres punktowy
  - Wybierz dwie zmienne liczbowe i narysuj wykres punktowy, aby sprawdzić ewentualną korelację między nimi.
  - Oceń, czy zależność jest dodatnia, ujemna czy raczej brak wyraźnego związku ('na oko')
- · Użycie koloru lub symboli do wyróżnienia kategorii
  - Jeśli masz w danych zmienną kategoryczną użyj kolorów lub punktów, by zaznaczyć przynależność do danej kategorii.



#### 5. Podsumowania i dostosowanie wykresów

- · Obliczanie statystyk opisowych w podziale na grupy
  - Skorzystaj z funkcji do obliczania średniej, mediany lub innej statystyki
  - Wprowadź dane na wybrany wykres, dlaczego w taki sposób?
- · Modyfikacja wizualna wykresów
  - Dostosuj tytuł, etykiety osi i ewentualnie legendę w swoich wykresach, tak aby były bardziej czytelne i opisowe.

