

PREZENTACJA I WIZUALIZACJA DANYCH W R

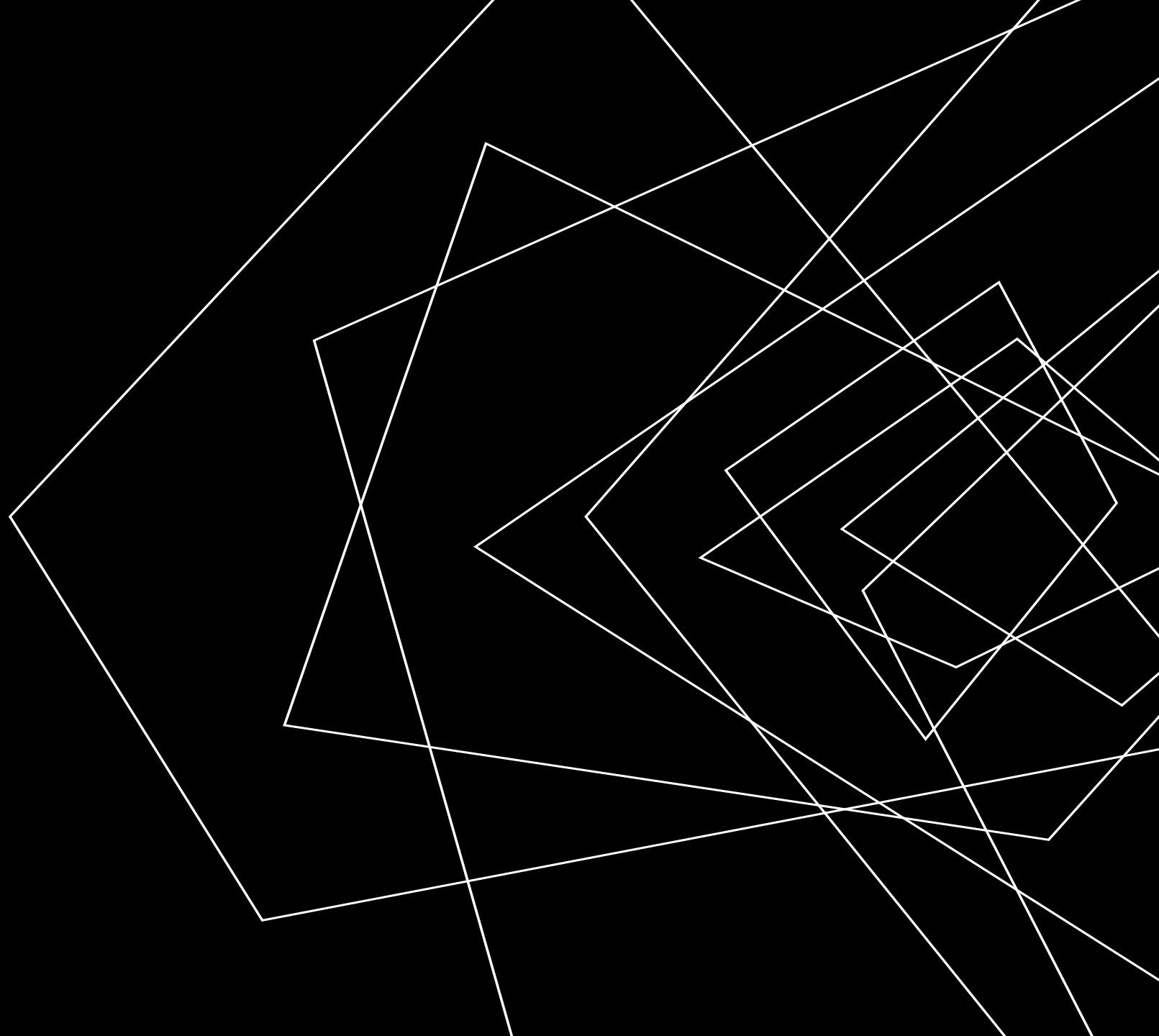
Marek Harhala

Spotkanie nr 2

Instytut Biologii
i Ewolucji Człowieka,

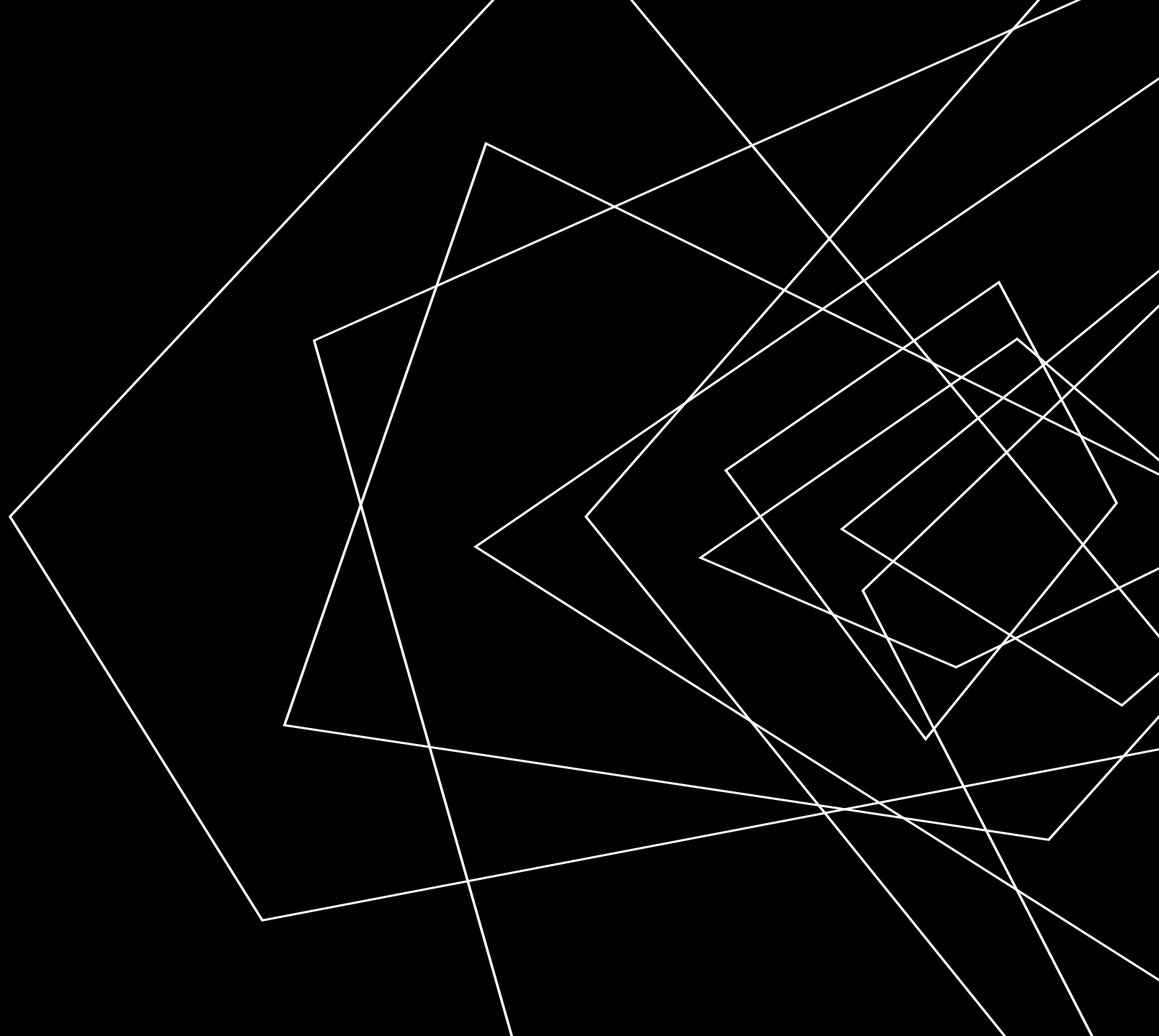
2.42

marek.harhala@amu.edu.pl

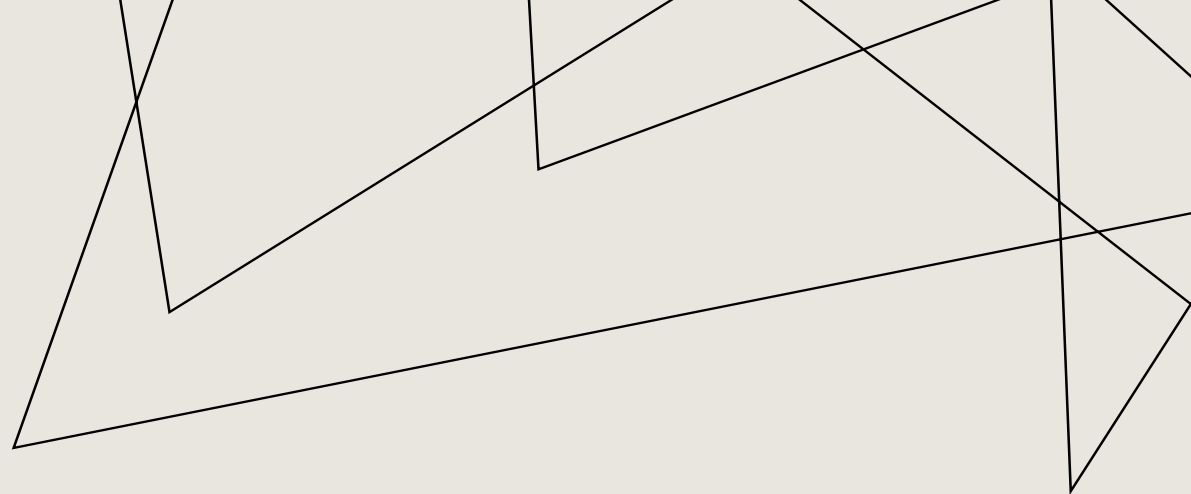


POPRIEDNIE ZAJĘCIA

1. Prezentacja wizualna danych to aspekt KOMUNIKACJI!
2. Prezentacja danych powinna być traktowana jako historia – z kontekstem, doświadczeniem widza i opowieścią jaka jest do tego przyłączona.



ROZRÓŻNIENIE MIĘDZY EKSPLORACJĄ (DANYCH) A WYJAŚNIENIEM:



Podejście skupione na eksploracji:

- skupienie na odkrywaniu danych,
- użytkownik kontroluje interpretację danych,
- wymaga wysokiego zaangażowania,
- wysokie ryzyko przeciążenia widza,
- skierowane do ekspertów w danej konkretnej dziedzinie.

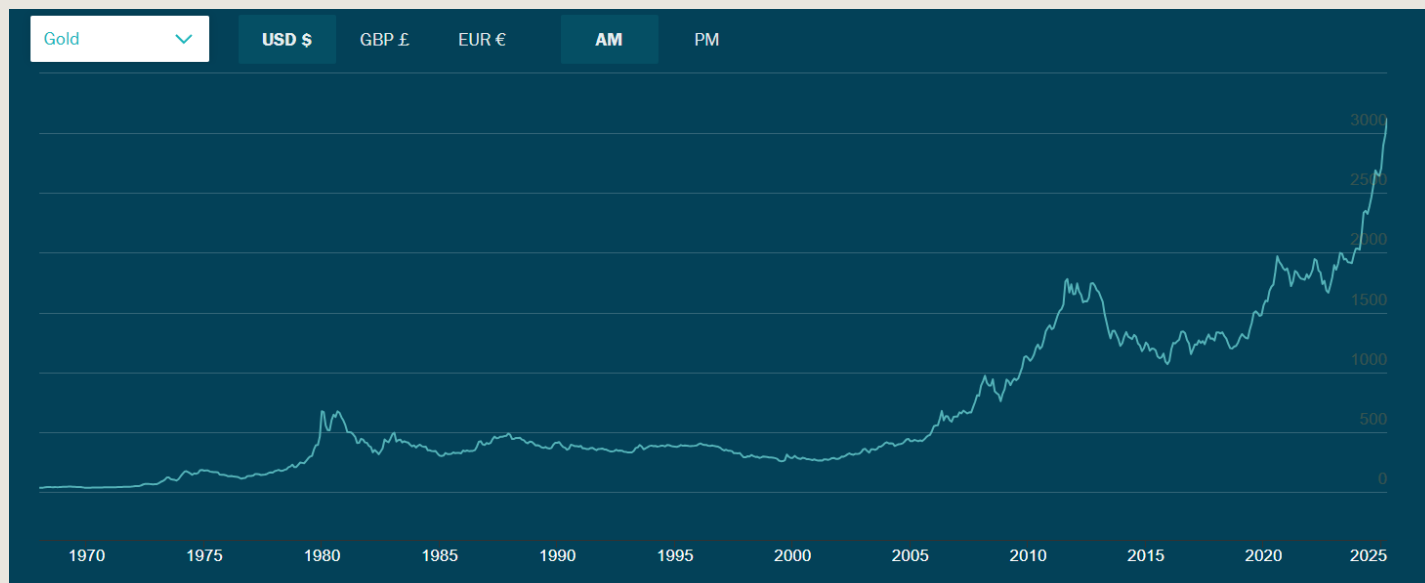
Podejście skupione na wyjaśnieniu:

- skupienie na przekazaniu wniosków,
- autor kontroluje interpretację danych,
- wymaga niskiego poziomu zaangażowania (przez krótki czas),
- niskie ryzyko przeciążenia widza,
- skierowane do szerokiej publiczności, która nie ma doświadczenia i szkolenia w skonfrontowaniu sposobu interpretacji danych.

WIZUALIZACJA OPARTA NA EKSPLORACJI DANYCH

Wskazane jest zaangażowanie widza poprzez danie mu kontroli nad dostępem do danych.

1. Udostępnienie materiałów do **samodzielnego** przeglądania: wydrukowanie wykresów i wręczenie ich widzom, udostępnienie filmów, plików, wykresów, danych.
2. Umożliwienie lub wręcz wymuszenie interakcji widza z danymi poprzez, np.: omawianie rzeczy do których widz musi 'sam się dostać'.



WIZUALIZACJA OPARTA NA WYJAŚNIANIU DANYCH

Celem jest przekazanie 'narracji autora', czyli przekazanie widzom określonych wniosków.

1. Każdy element prezentacji i wizualizacji ma budować określoną narrację prowadzącą do określonych wniosków.
2. Uwagę widza trzeba skierować na kilka konkretnych aspektów budujących narrację prowadzącą do określonego wniosku.



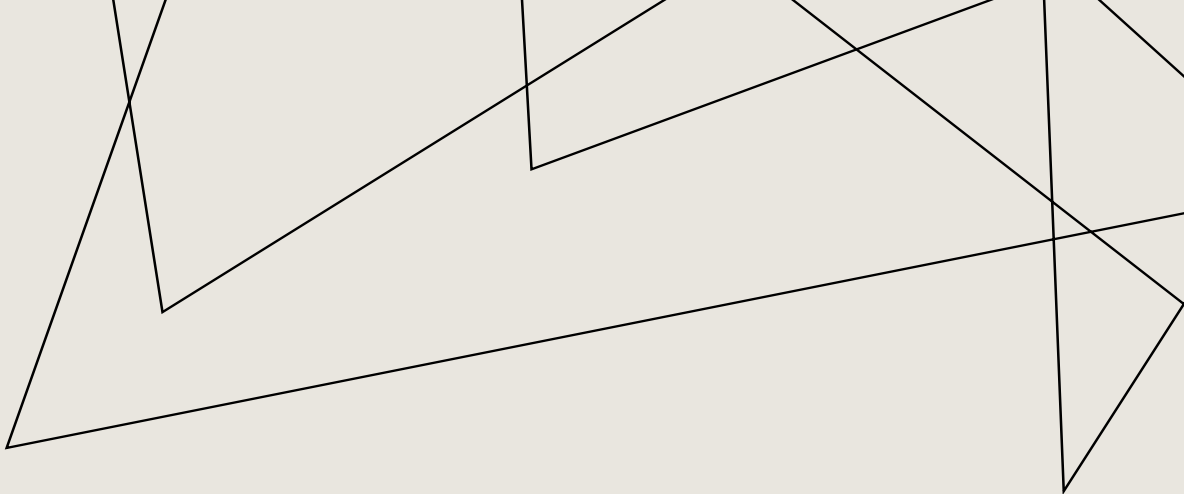
PODEJŚCIE 'MIESZANE', ALE NIE WPROST

Zawiera eksplorację danych, ale jednocześnie proponuje określony sposób interpretacji danych zamiast budowania konkretnej narracji.

1. Prezentacja danych jest dokonywana w stylu 'exploratory', ale jednocześnie jest dostarczany opis interpretacji tych danych.
2. Różnica pomiędzy narracją a interpretacją (w tym kontekście): narracja przekazuje wnioski, a interpretacja danych przekazuje sposób analizy/perspektywę patrzenia na prezentowane informacje (który może prowadzić do wniosków).



ZAANGAŻOWANIE UŻYTKOWNIKA A KONTROLA



Eksploracja danych

- większe zaangażowanie widza,
- większe wymagania dot. uwagi i skupienia (zaangażujcie widza),
- wysokie ryzyko przeciążenia widza.

Wyjaśnienie danych

- niskie zaangażowanie widza,
- wymaga 'przebicia' się do widza,
- niskie ryzyko przeciążenia widza,
- łatwo 'zgubić' widza w związku z niskim zaangażowaniem.



TZW. KONTEKST NARRACYJNY

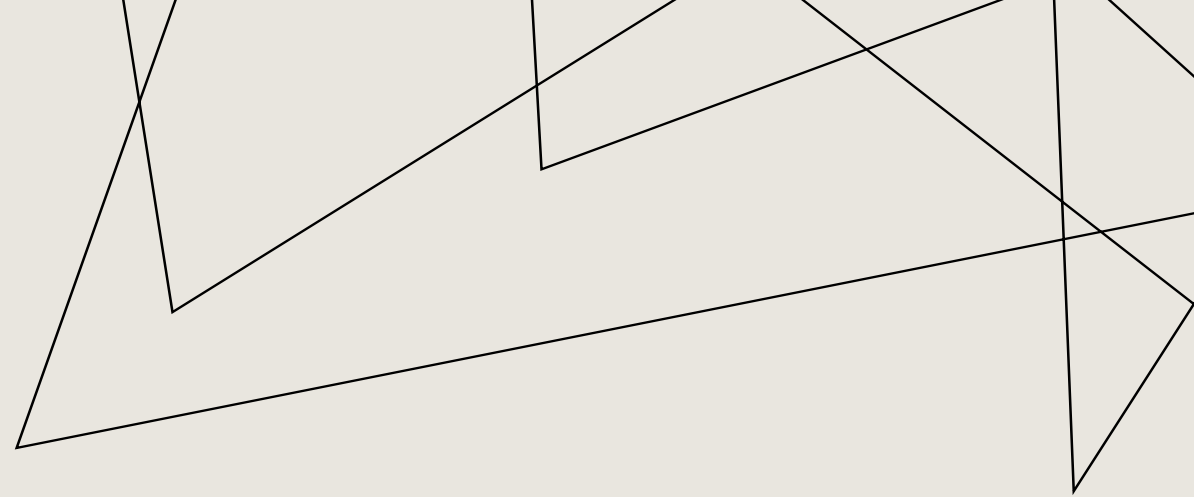
Prezentacje oparte na wyjaśnieniach danych zapewniają, że autor kontroluje kontekst prezentacji danych, więc upewnia widzów w sposobie zrozumienia tych danych. Te same dane przedstawione w eksploracyjny sposób mogłyby być kompletnie niezrozumiałe dla widza.

ROZWAŻ NARZĘDZIA

Rozważ dostępne narzędzia i sposoby prezentacji danych – ich obecność może decydować o sposobie przedstawiania danych (np.: obecność wskaźnika laserowego).

OBCIĄŻENIE TZW. POZNAWCZE (COGNITIVE LOAD)

Oceńcie które elementy prezentacji i wizualizacji są bardziej wymagające od widza. Zbyt wymagające elementy podzielcie na mniejsze etapy.

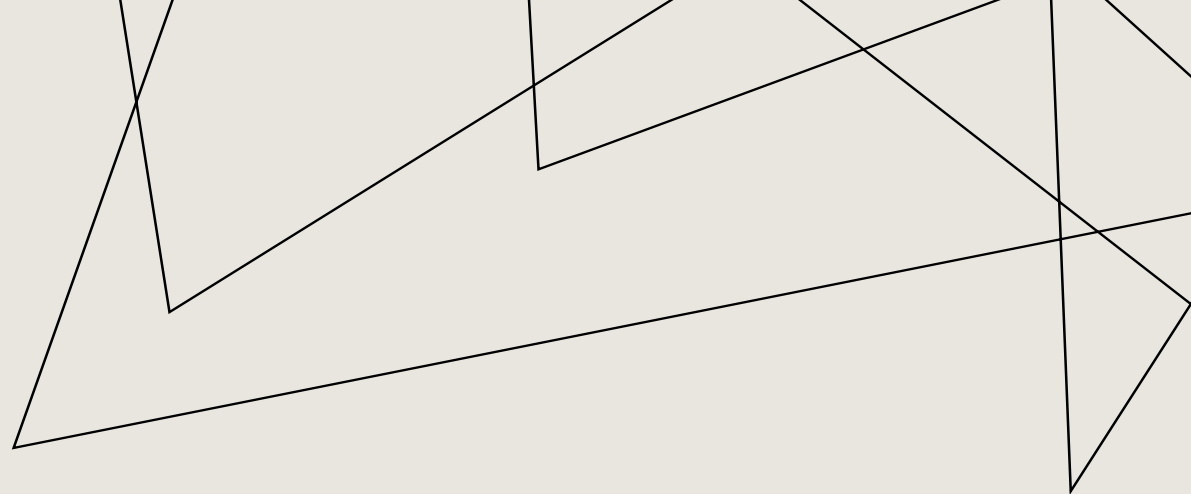


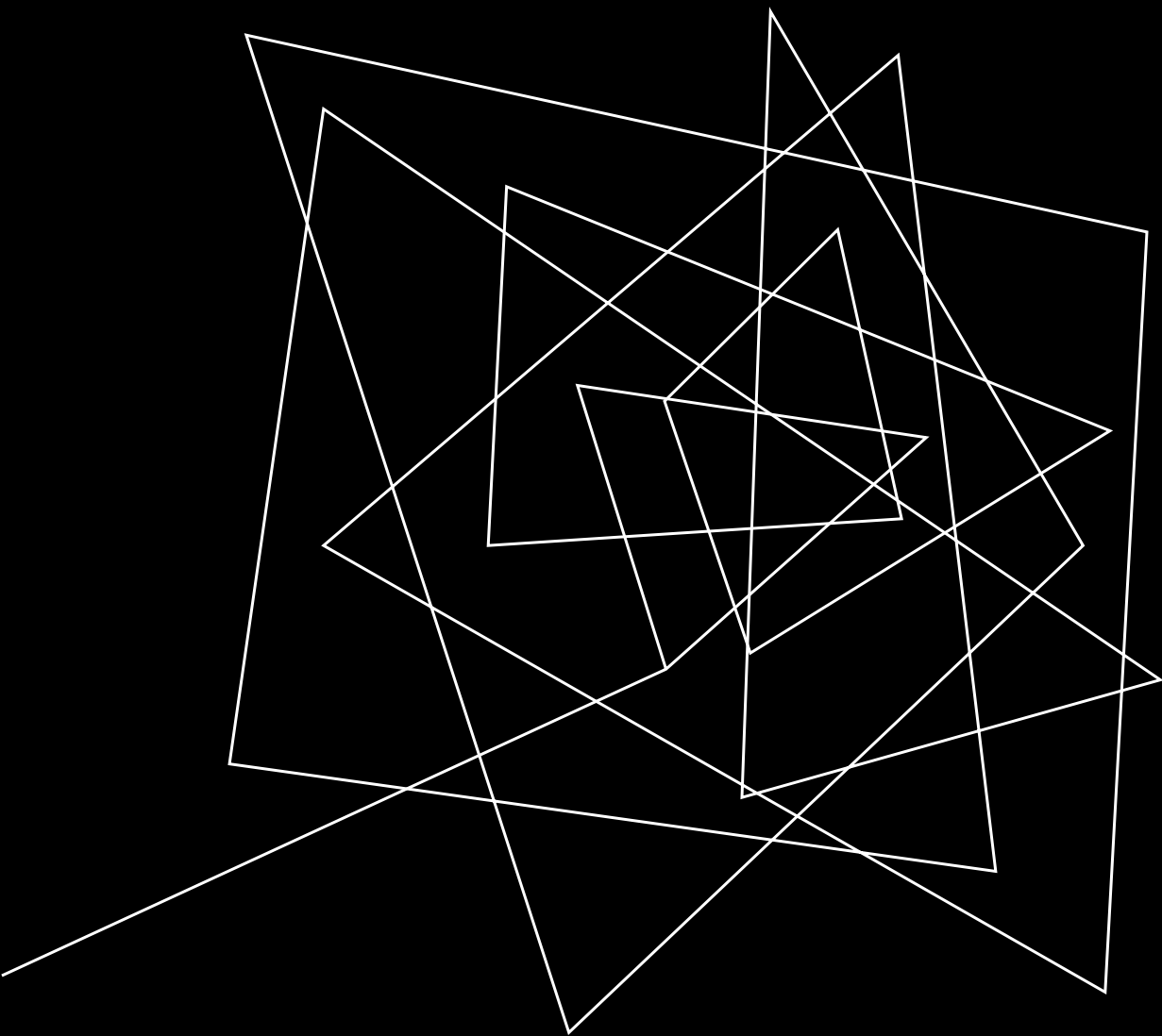
MOTYWACJA WIDZA

Oceńcie, czy widzowie są gotowi na eksplorację danych, czy ich wyjaśnienie. Dopasujcie się do widza.

PŁYNNOŚĆ NARRACJI VS. ODKRYWANIE DANYCH:

W praktyce potrzebne są często oba te elementy, ale nie są naturalnie kompatybilne – albo zaplanujcie prezentację bez mieszania ich, albo się skupcie na ich zgodnej koegzystencji.





DRUGIE SPOTKANIE - ZADANIA

A series of white, overlapping geometric lines and polygons on a black background, located on the left side of the slide.

RODZAJE DANYCH

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



ĆWICZENIE NR 1

Wczytaj i zbadaj zbiór danych

WCZYTANIE DANYCH.

- Załaduj zbiór danych z <https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv> .

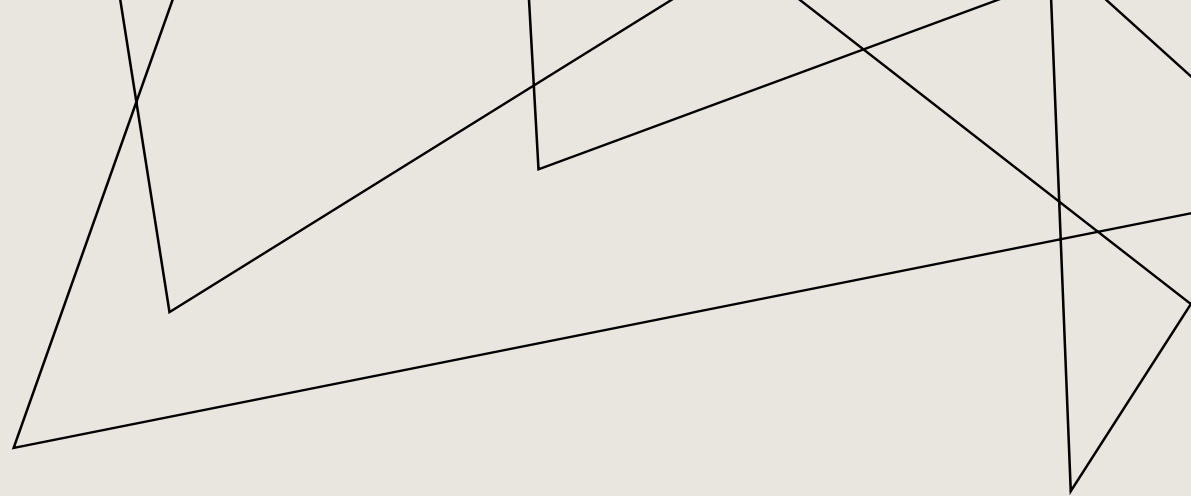
ZADANIA

- Wczytaj dane.
- Opisz strukturę danych.
- Znajdź liczbę unikalnych wartości w kolumnie day.

ĆWICZENIE NR 2

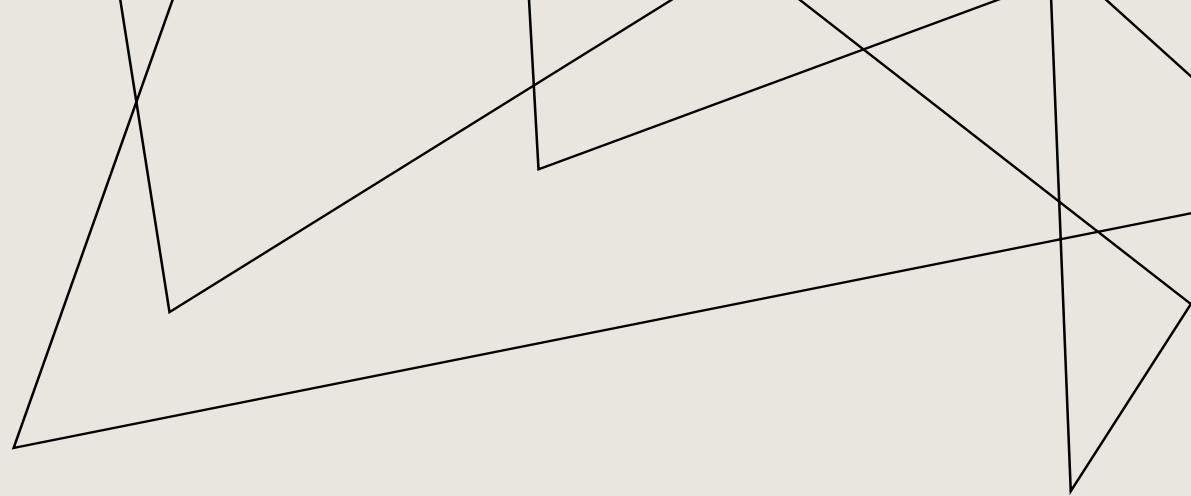
Tworzenie prostego wykresu z ggplot2

- **Wstęp do wizualizacji.**
- Napisz, w który dzień tygodnia napiwki są najwyższe.
- Stwórz wykres i zmień typ wykresu na `geom_violin()`.



ĆWICZENIE NR 3

- **Interaktywny wykres z plotly**
- Stwórz interaktywny wykres, który pozwala analizować zależność między rachunkiem a napiwkiem w zależności od płci.



ĆWICZENIE NR 4

- Wybierz/stwórz dowolny zestaw danych ze zmienną zależną i niezależną typu binarnego.

DZIĘKUJĘ ZA
UWAGĘ

