

# Praca domowa nr 5

Olga Kachniewska, Julia Pińkowska

## 1. CEL BADANIA

Głównym celem pracy domowej było zbadanie występowania problemów z odczytywaniem danych z wykresów. W ramach naszych badań postanowiliśmy przeanalizować porównawczo dwa aspekty: porównywania kątów i słupków (odczytywanie danych z wykresu kołowego i słupkowego), a także wpływu urozmaiceń graficznych (porównanie dwóch różnie zwizualizowanych wykresów słupkowych). Podstawą naszej analizy jest ankieta przeprowadzona wśród osób z naszego otoczenia

## 2. ANKIETA

### Dane

Wykresy wykorzystane w ankietach zostały stworzone na podstawie dwóch prostych, wymyślonych na potrzeby badania, zbiorów danych, które zawierały informacje o zainteresowaniach [Tabela 1.] oraz ilości opadów z 14 dni [Tabela 2.].

Aktywność	Procent udziału/ liczba obserwacji na 100
Słuchanie muzyki	6
Gry wideo	20
Podróże	9
Oglądanie filmów/seriali	30
Rękodzieło	2
Sport	12
Czytanie książek	18
Gotowanie	3
Spotkania ze znajomymi	10

Tabela 1.

Dzień maja	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Opady [mm]	5	12	3	8	0	2	1	18	9	6	15	4	3	20

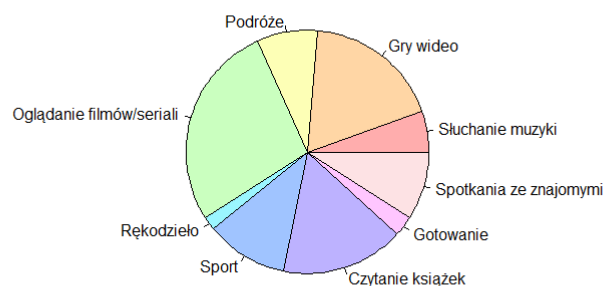
Tabela 2.

### Wykresy

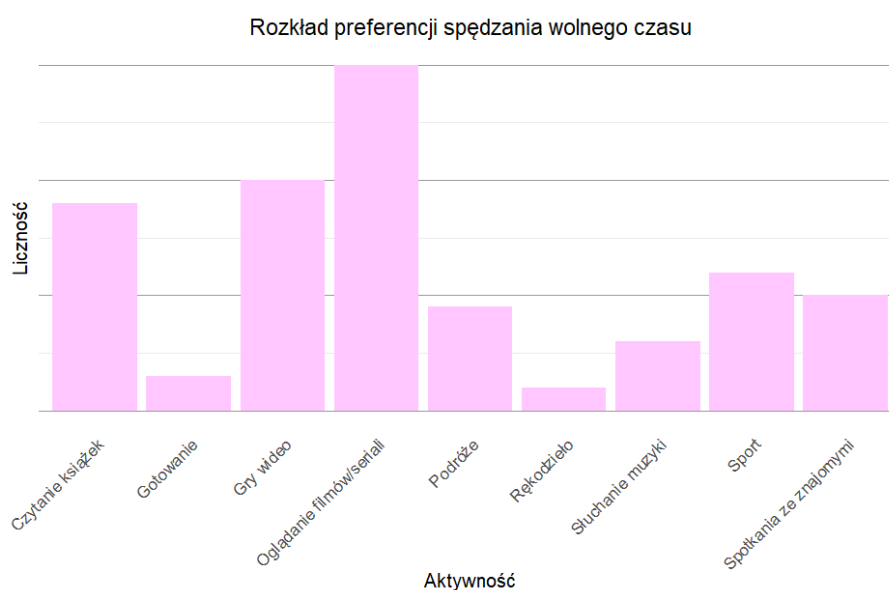
Każdy z badanych problemów został przeanalizowany z wykorzystaniem dwóch wykresów- w obu przypadkach wizualizacją kontrolną był wykres słupkowy. Aby uniknąć zaburzenia badania wynikającego z pokazania jednemu badanemu dwóch wykresów na podstawie tych samych danych, co mogłoby zaburzyć odbiór wykresu z powodu pamiętania danych zauważonym na drugim, ankieta została podzielona na dwie części- A i B. W każdej z nich były zawarte dwa wykresy z różnych badanych wykresów. Ankietowani, którzy odpowiadali w grupie A dostali pytania na podstawie wykresu kołowego z danych z Tabeli 1., a także słupkowego z Tabeli 2., a w grupie B był to analogicznie wykres słupkowy z Tabeli 1. oraz wykres słupkowy 3D z ozdobnikami z tabeli 2. Ten podział sprawił, że każdy z ankietowanych dostał jeden wykres uznany przez nas za prosty do odczytania i jeden trudny.

Wykresy na podstawie zainteresowań wyglądały następująco:

**Rozkład preferencji spędzania wolnego czasu**



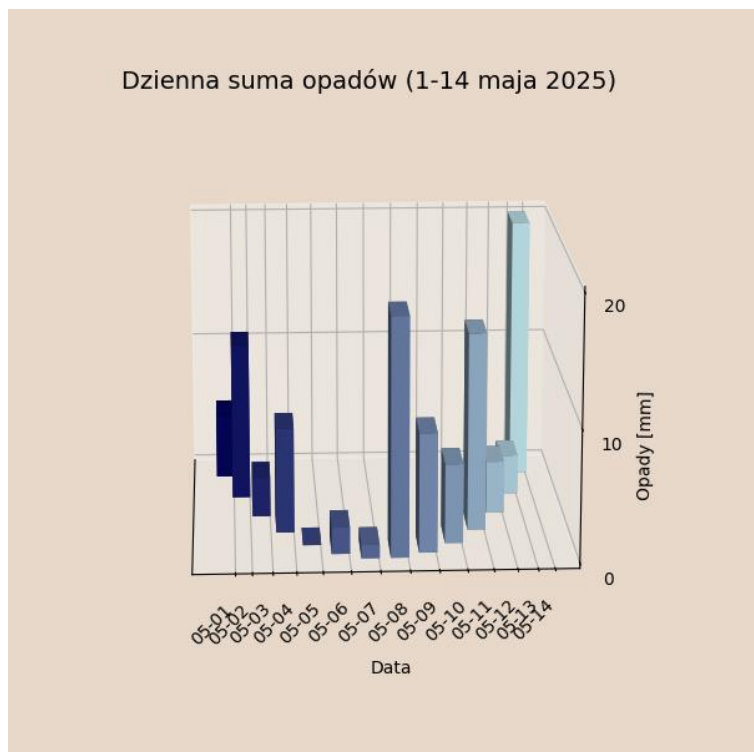
Wykres 1.1.



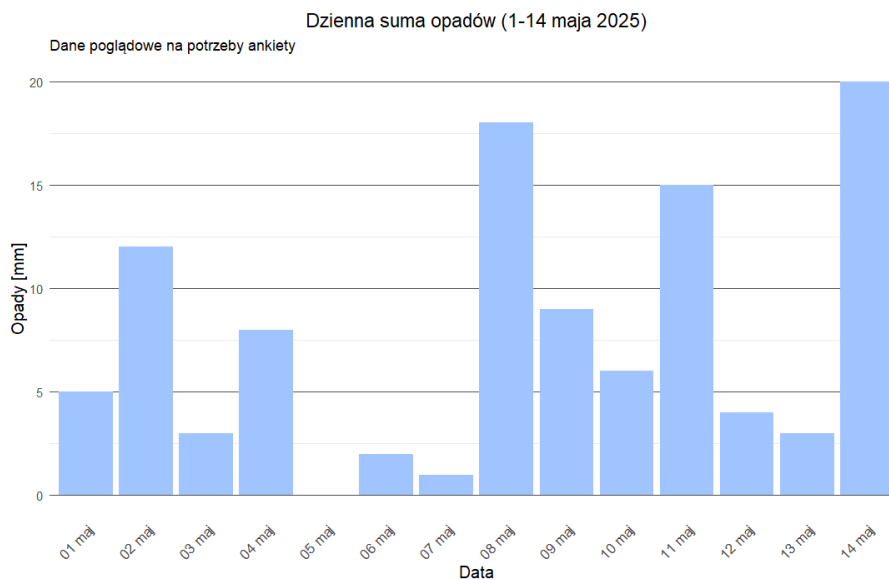
Wykres 1.2.

Łatwo jest zauważyć, że Wykres 1.2. nie jest w pełni poprawnym wykresem, jednak brak wartości na osi licznosci jest zabiegiem celowym. Na wykresie kołowym również nie występują żadne liczby, zatem postanowiliśmy sprawdzić umiejętność porównywania samych kształtów (słupków i kątów) z jedną informacją, że suma stanowi całość danych.

Wykresy na podstawie danych pogodowych wyglądały następująco:



Wykres 2.1.



Wykres 2.2.

Wykres 2.2 jest klasycznym wykresem słupkowym, natomiast przy tworzeniu Wykresu 2.1 zastosowaliśmy urozmaicenia wizualne takie jak zmiana koloru tła, gradientowe słupki, ułożenie słupków na półokręgu oraz widok perspektywiczny, aby jak najbardziej przekształcić bazowy wykres za pomocą efektów, które zwykły użytkownik może początkowo uznać za ciekawy dodatek.

## Pytania

Do każdego z czterech wykresów zamieściliśmy analogiczne zestawy pytań (identyczne dla danych zestawów danych) skupiające się głównie na porównywaniu danych, a także postrzeganiu wartości. Pięć pytań merytorycznych do wykresów na podstawie tabeli 1:

1. Ustaw podane kategorie od najmniej do najbardziej popularnej:  
[Sport, Spotkania ze znajomymi, Podróże]
2. Która aktywność jest najpopularniejsza?  
[Czytanie książek, oglądanie filmów/seriali, Gry wideo, Sport]
3. Jaki jest jej przybliżony udział procentowy?  
[26, 28, 30, 32, 34, 36, 40]
4. Jaki jest jej przybliżony udział procentowy?  
[Rękodzieło, Gotowanie, Słuchanie muzyki, Podróże]
5. Czy aktywność „Sport” ma większy udział niż suma udziałów „Gotowanie” i „Słuchanie muzyki”?  
[Tak, Nie]

Odpowiednio do drugiego zestawu danych:

1. Ustaw podane dni od najmniejszej do największej sumy opadów  
[10, 4, 9, 2]
2. Który dzień miał największą sumę opadów?  
[14, 11, 8]
3. Jaka była wartość opadów w tym dniu?  
[20, 19, 18, 17, 16, 15]
4. W których dwóch dniach suma opadów jest identyczna?  
[3 i 13, 3 i 12, 4 i 9, 1 i 10]
5. Czy suma opadów z 3 i 6 maja jest wyższa niż ta z 10 maja?  
[Tak, Nie, Jest taka sama]

A także po każdym wykresie umieszczone było pytanie: „Jak oceniasz wygodę odczytywania danych z tego wykresu?”

Na zakończenie całej ankiety dodałyśmy pytania, które mogłyby urozmaicić wnioski z odpowiedzi do ankiet, w których zawarłyśmy pytanie o porównanie wykresów w ramach zestawu pytań, na który właśnie udzieliło się odpowiedzi (czyli porównanie wykresów zrobionych na podstawie różnych danych), swobodną wypowiedź na temat przedstawionych wykresów, a także możliwość zidentyfikowania się jako student Politechniki Warszawskiej.

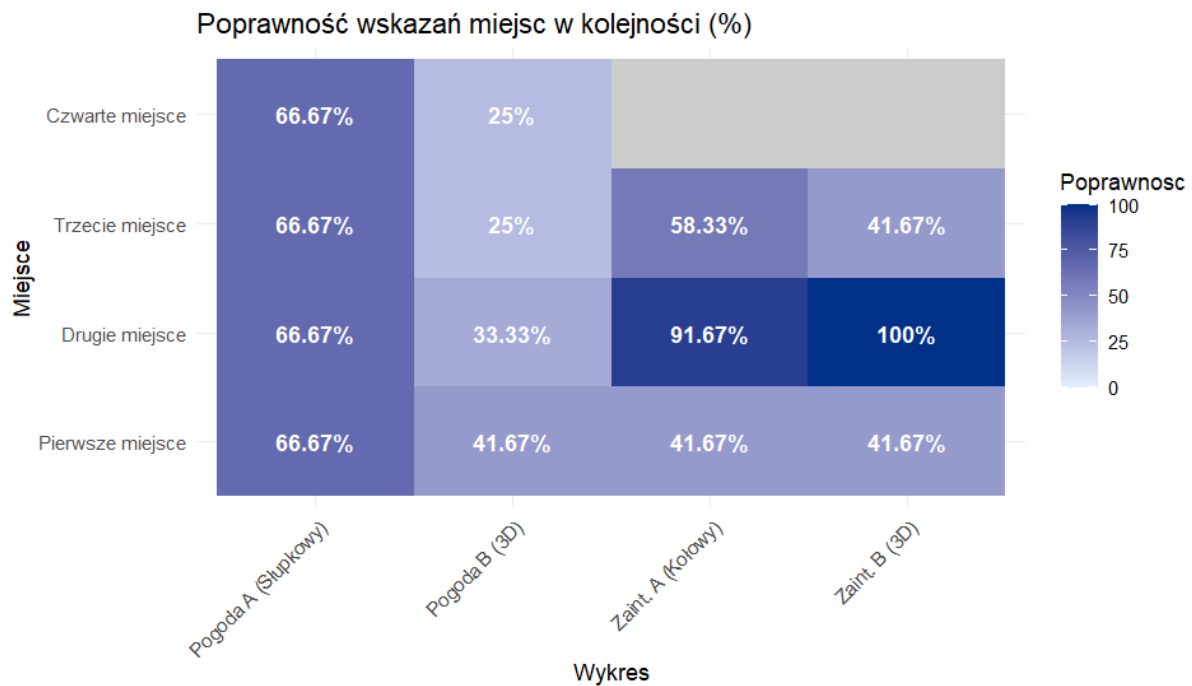
Ankieta jest nadal dostępna - <https://forms.cloud.microsoft/e/Q3WZmSUfda>

## 3. WYNIKI

Ostatecznie odpowiedzi udzieliły 24 osoby- po 12 w każdej z grup A i B.

### Ustawianie w kolejności (porównanie wielu wartości)

Pierwsze pytanie sprawdzające ustawienie wartości w kolejności rosnącej można jednocześnie porównywać dla wszystkich czterech wykresów, a poprawność udzielanych odpowiedzi można przedstawić za pomocą heatmapy poprawnych odpowiedzi:

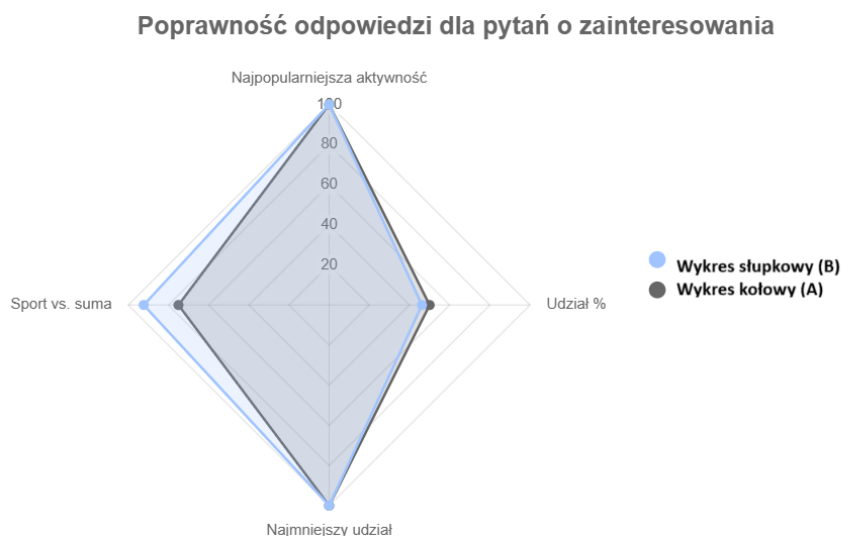


Wykres 3.1.

Niezależnie od zestawu danych bardziej poprawne dopasowanie wystąpiło dla standardowych wykresów słupkowych. Porównanie wyników z opcji na trzy i cztery miejsca (a także ręczne sprawdzenie odpowiedzi) sugeruje, że wyniki mogą być zaniżone przez błąd niewynikający z wykresów, a jedynie interpretacji polecenia- duża grupa osób ułożyła odpowiedzi w dobrej, ale odwrotnej kolejności (wszystkie niepoprawne odpowiedzi dla wykresu słupkowego o pogodzie miały taką postać). To pytanie obrazuje zatem dwa zjawiska- zdecydowaną porażkę wykresu 3D oraz nieuwagę w czytaniu pytań.

## Zainteresowania - porównanie wykresu kołowego i słupkowego

Poprawne odpowiedzi dla pozostały pytań padły w rozkładzie:

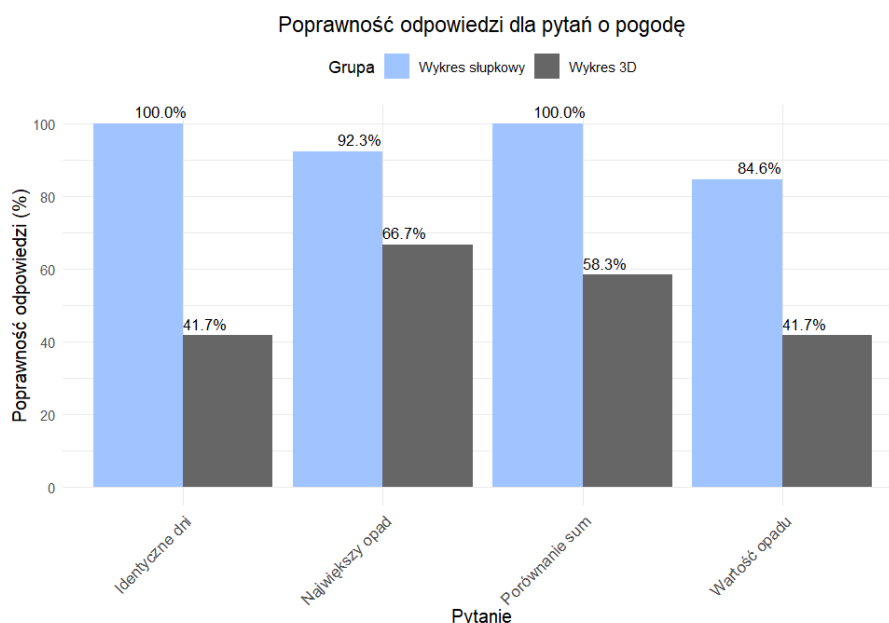


Wykres 3.2.

Wskazanie najpopularniejszej i najmniej popularnej aktywności wykazało stuprocentową poprawność dla obu wykresów, zatem dla tak zróżnicowanych danych porównywanie kątów nie sprawiło żadnych trudności, a nawet nieznacznie ułatwiło wskazywanie procentowego udziału aktywności o największym udziale. Jednocześnie sumowanie wycinków koła wypadło gorzej niż dodawanie słupków (tożsamy z odcinkami).

## Opady- porównanie wykresu standardowego i 3D

Drugie badanie porównawcze wskazuje na zdecydowaną przewagę standardowego wykresu słupkowego, w prostym porównaniu za pomocą takiego wykresu:

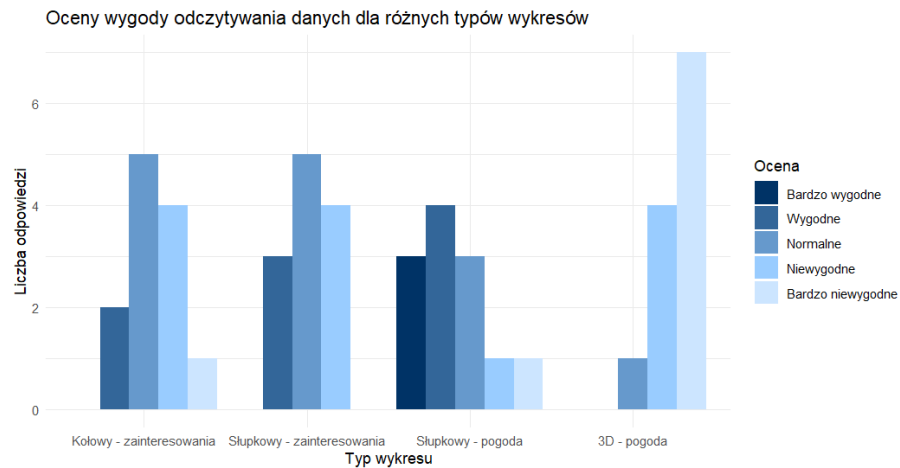


Wykres 3.4.

można zauważyć znacznie niższą poprawność dla wykresu 3D. Potwierdza to założenie, że zniekształcenie wywołane perspektywą i ułożeniem na krzywej znacząco zmienia postrzeganie wielkości.

## Opinie ankietowanych

Wśród swobodnych wypowiedzi ankietowanych głównie pojawia się preferencja wykresu słupkowego, a także niezadowolenie z wykresu 3D („*Drugi jest okropny*”). Naszym zdaniem na szczególną uwagę zasługuje również opinia: „*Odczytanie było łatwiejsze na słupkowym ale porównywanie między kilkoma wydawało się łatwiejsze na kołowym*”, która wyraża zalety wykresu słupkowego, ale też kołowego. Wśród odpowiedzi na pytanie „*Jak oceniasz wygodę odczytywania danych z tego wykresu?*” jedynie wykres 3D jest jednoznacznie negatywny, a w przypadku porównywania wykresów związanych z zainteresowaniami nie są one aż tak rozbieżne.



Wykres 3.4.

#### 4. PODSUMOWANIE

Porównywanie postrzegania kątów i słupków wykazało przewagę wykresu słupkowego, jednak różnica nie jest aż tak znacząca, jak na początku zakładaliśmy. W przypadku drugiego problemu badawczego wynik jest jednoznaczny- urozmaicenia wykresu związane z perspektywą są znacząco niekorzystne i zaburzają odczytywanie wartości, zatem należy ich unikać.