

Praca domowa 3+4

Michał Przerwa

Spis treści

1	Wstęp	2
2	Omówienie wyników ankiety	3
2.1	Uwagi wstępne	3
2.2	Wariant nr 1	3
2.3	Wariant nr 2	4
2.4	Wariant nr 3	5
3	Podsumowanie	6

1 Wstęp

Celem pracy domowej było przeprowadzenie eksperymentu, sprawdzającego czy określone problemy związane z czytaniem danych występują oraz czy sprawdzone praktyki wizualizacji danych działają w społeczeństwie.

Aby poznać odpowiedź, przeprowadziłem ankietę, w której trzeba było ocenić czytelność i przejrzystość dwóch wykresów: kołowego i słupkowego w różnych konfiguracjach (każdy wariant akcentuje inny problem na wykresach). Tematyka tych wykresów to *Liczba klientów kin w roku 2019*. Dane pochodzą z bazy Głównego Urzędu Statystycznego. Odpowiedzi w ankiecie udzieliły osoby z przedziału wiekowego 20-60 lat (rodzina oraz przyjaciele ze studiów, osoby te otrzymały dostęp do ankiety).

Ankieta jest dostępna pod linkiem: <https://forms.gle/yXwsvvXwoCKZe81U7>

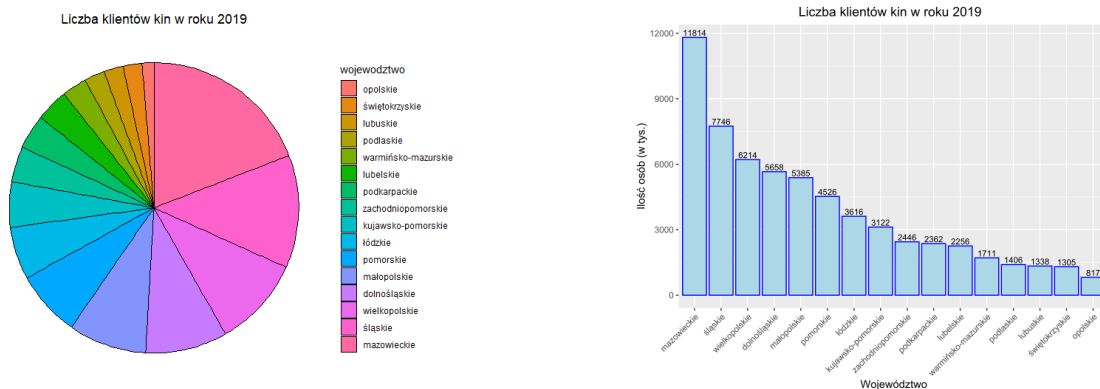
2 Omówienie wyników ankiety

2.1 Uwagi wstępne

Do każdego wariantu został dołączony opis przedstawianych wykresów, moje przewidywane wyniki oraz przedstawienie rzeczywistych rezultatów wraz z uogólnionymi komentarzami ankietowanych.

2.2 Wariant nr 1

Jako pierwsze porównaliśmy wykres kołowy z nieopisanymi wartościami i skalą kolorów dopasowaną automatycznie przez *ggplot* z poprawnie wykonanym wykresem słupkowym. Przewidyuję tutaj, że większość będzie preferować drugi wykres, ponieważ jest zarówno przejrzysty, jak i zawiera wszystkie dane.



(a) Wykres nr 1

(b) Wykres nr 2

Rysunek 1: Wykresy z pierwszego wariantu

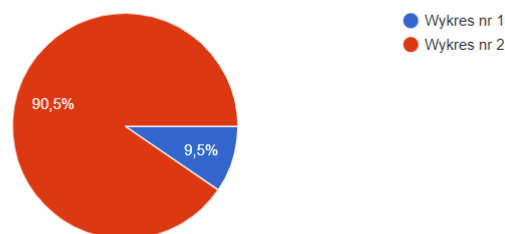
Zgodnie z przewidywaniami, zdecydowana większość osób wskazała wykres słupkowy jako czytelniejszy. Wybór był motywowany poprzez:

- łatwiejszy odczyt proporcji i porównania wartości
- trudny odczyt danych na wykresie kołowym przez zbyt płynnie przechodzące kolory

Okazało się jednak, że dla osób, które zaznaczyły opcję pierwszą, kolory były bardziej przemiawiające i wyraziste.

Który wykres jest czytelniejszy?

21 odpowiedzi

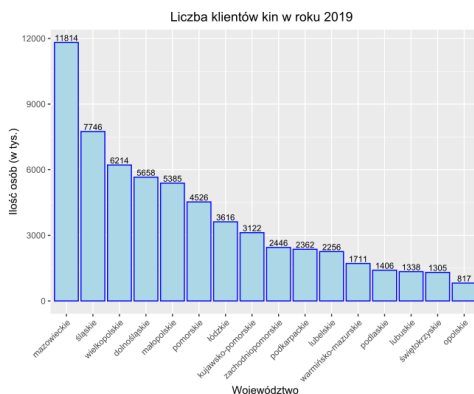


Rysunek 2: Wyniki wariantu nr 1

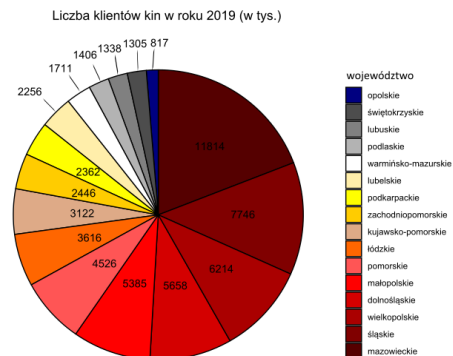
Ciekawe jest to, że Google prezentuje te dane automatycznie na wykresie kołowym. ☺

2.3 Wariant nr 2

W tym wariantcie zmieniony został wykres kołowy. Dodano do niego wartości liczbowe oraz inaczej dopasowano skalę kolorów. Pomimo zwiększenia jakości wizualizacji tego wykresu, to wykres słupkowy nadal powinien być preferowany przez większą ilość osób.



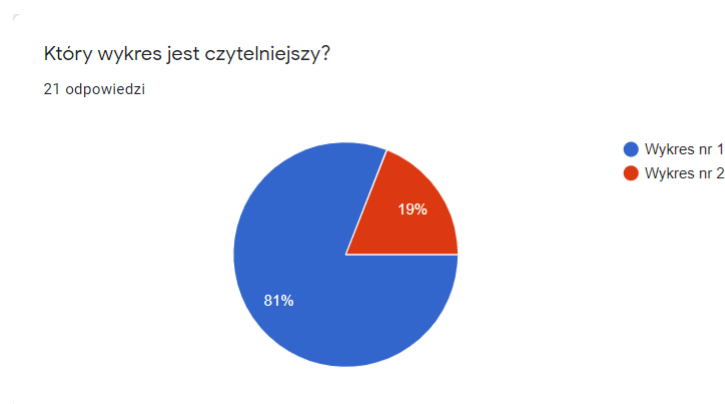
(a) Wykres nr 1



(b) Wykres nr 2

Rysunek 3: Wykresy z drugiego wariantu

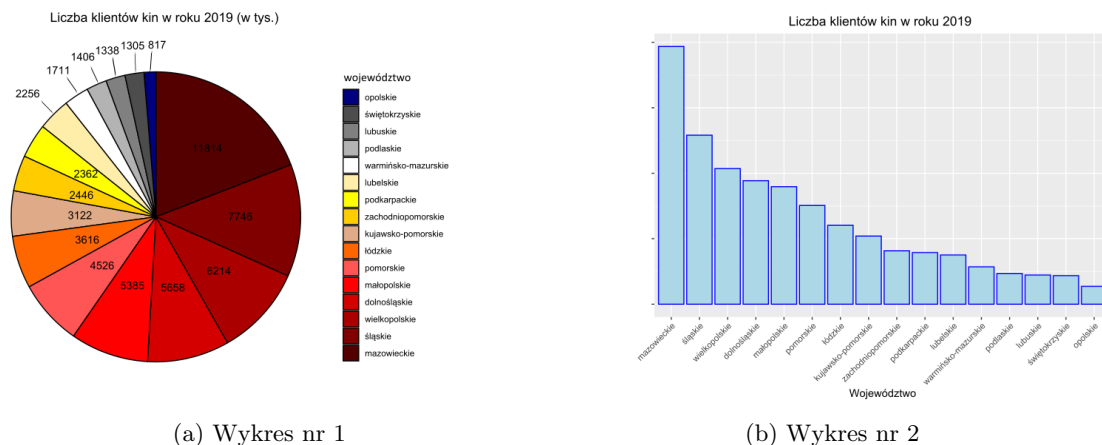
Również tutaj przewidywania zgodziły się z wynikami ankiety, przy czym różnica w głosach jest nieznacznie mniejsza niż poprzednio. Ankietowani zwrócili uwagę na lepszą przejrzystość wykresu słupkowego oraz na to, że część danych na wykresie kołowym widać mało wyraźnie na ciemnym tle. Cały czas jednak są osoby, które uważają wykres kołowy za bardziej przejrzysty dzięki różnorodności kolorów.



Rysunek 4: Wyniki wariantu nr 2

2.4 Wariant nr 3

W ostatnim wariantcie zmianie uległ wykres słupkowy. Została zmniejszona jakość tego wykresu poprzez usunięcie konkretnych wartości ze słupków oraz osi wartości liczbowych. Jako, że oba wykresy z teoretycznego punktu widzenia są źle skonstruowane (bądź cały wykres jest nieodpowiedni do przedstawiania tego typu danych), ciężko przewidzieć, który wykres będzie częściej wybierany. Moim zdaniem, byłby to jednak wykres kołowy, ponieważ zawiera on wartości liczbowe, do których można się szybciej odnieść.

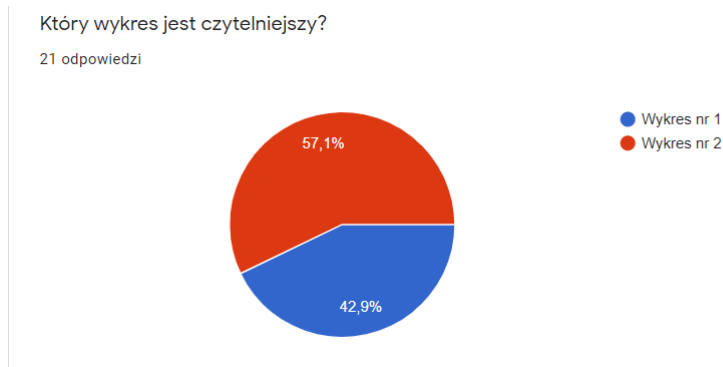


(a) Wykres nr 1

(b) Wykres nr 2

Rysunek 5: Wykresy z trzeciego wariantu

Okazuje się, że częściej wybierany jest wykres słupkowy, jednak różnica głosów nie jest duża. Osoby wybierające wykres kołowy zwróciły uwagę na występowanie bardziej precyzyjnych danych, niż na słupkowym. Zwolennicy drugiej opcji z kolei wspominali o lepszej przejrzystości i czytelności (zwłaszcza dla wielu danych). Jedna osoba słusznie zauważyła, że dla zliczeń, gdzie chcemy odczytać konkretne dane, lepszy będzie wykres kołowy, a słupkowy będzie preferowany, gdy chcemy dowiedzieć się, w którym województwie było więcej klientów bez poznawania dokładnych wartości.



Rysunek 6: Wyniki wariantu nr 3

3 Podsumowanie

Analizując na wyniki eksperymentu, można powiedzieć, że wykres słupkowy jest przystępniejszy w odbiorze, niż kołowy. Podobnie możemy powiedzieć o stosowaniu kolorów na wykresach, gdyż bardziej preferowana jest jednolitość, choć nie można zapomnieć o osobach, dla których kolory są istotną częścią. Zaskakujący natomiast jest mniejszy wpływ występowania dokładniejszych danych na odbieranie wykresu jako bardziej czytelnego. Uogólniając, ludzie szybciej zwracają uwagę na wygląd wykresu, niż na zawarte w nim dane.