

Raport

Czy sposób wizualizacji idzie w parze z czytelnością?

Dominika Filimonow, Dominika Kawczyńska

28 marca 2023

1 Wstęp

1.1 Opis zadania

Zadanie polegało na wykonaniu eksperymentów, sprawdzających, czy określone problemy z czytaniem danych dalej występują. Należało to sprawdzić, tworząc ankietę, zawierającą wizualizacje stworzone niezgodnie z przyjętym kanonem oraz pytania ich dotyczące. Następnie z uzyskanych wyników wyciągnąć wnioski.

1.2 Opis zagadnienia

Tworzenie wykresów nie jest losowym procesem. Wiele lat temu stworzono "dobre praktyki" wizualizacji, aby wykresy były nie tylko czytelne dla każdego, ale również estetyczne i niemylące. Dlatego nie zaleca się używania wykresów kołowych, czy 3D. Także trzeba pamiętać o minimalizmie i zawieraniu najistotniejszych informacji bez zbędnych ozdób.

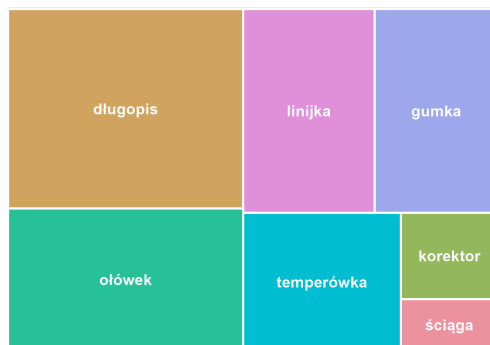
2 Rozwiązanie

Wykonałyśmy ankietę, którą wypełniło 66 osób w wieku od 18 do 60 lat. Każdy z ankietujących na co dzień ma styczność z wykresami, ale żaden z nich nie miał wiedzy na temat poprawnych wizualizacji. Nasz formularz podzieliliśmy na dwa eksperymenty dotyczące wpływu kształtu i koloru na poprawność czytania danych. Wszystkie wykresy tworzyłyśmy w programie *Rstudio* z wykorzystaniem języka *R*, bibliotek *dplyr*, *ggplot*, *treemap*, *fmsb* oraz bibliotek koniecznych do wygenerowania mapy. Ramki danych stworzyłyśmy samodzielnie.

3 Pierwszy eksperyment

W naszym pierwszym eksperymencie skupiliśmy się na wpływie kształtów wykresów. Dane, na których bazowałyśmy, wykonując wykresy to, wyniki ankiety przeprowadzonej wśród uczniów klas maturalnych. Zapytano ich o przybory, które noszą w swoich piórnikach. Dla sprecyzowania każda klasa liczyła trzydziestu uczniów, klasy zaś oznaczyłyśmy kolejnymi literami alfabetu. Do każdego z wykresów zadałyśmy po dwa pytania, których treści umieściłyśmy poniżej. Zaprezentowałyśmy również wyniki ankiet, przy czym zielonym kolorem zaznaczona jest odpowiedź poprawna zaś czerwonym-błędna.

3.1 Wykres 1



Liczba przyborów szkolnych w piórnikach uczniów klasy A

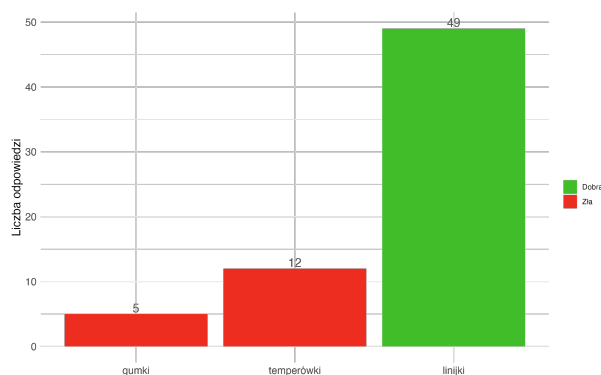
1. Co uczniowie klasy A częściej noszą w piórnikach linijki, gumki czy temperówki?

- ☒ linijki
☐ gumki
☐ temperówki

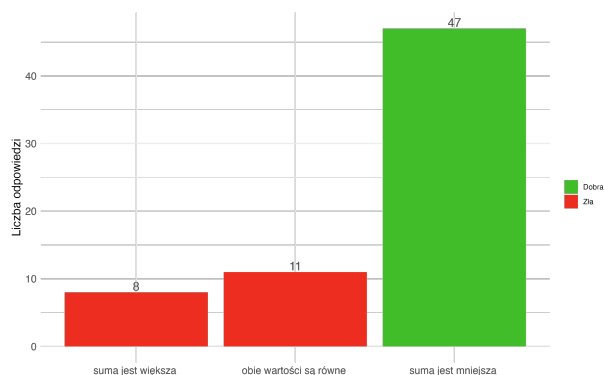
2. Jaka jest relacja pomiędzy sumą ołówków i korektorów a liczbą długopisów w klasie A?

- ☐ suma jest większa
☒ suma jest mniejsza
☐ obie wartości są równe

Wyniki ankiety

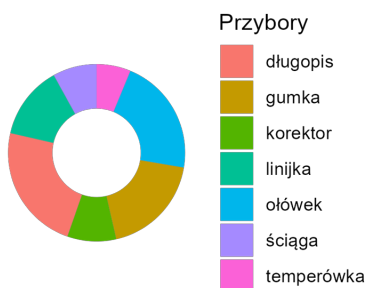


Udzielone odpowiedzi na pytanie 1.



Udzielone odpowiedzi na pytanie 2.

3.2 Wykres 2



Liczba przyborów szkolnych w piórnikach uczniów klasy B

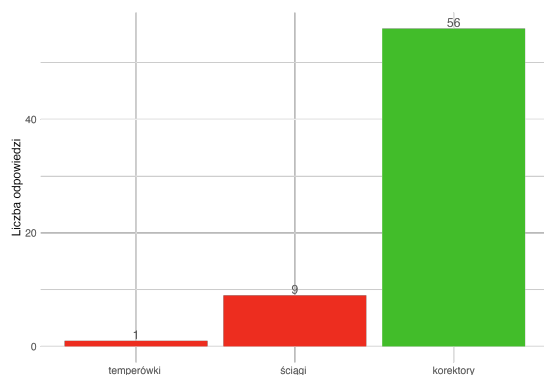
3. Co uczniowie klasy B częściej noszą w piórnikach temperówki, ściagi czy korektory?

- ☐ temperówki
☐ ściagi
☒ korektory

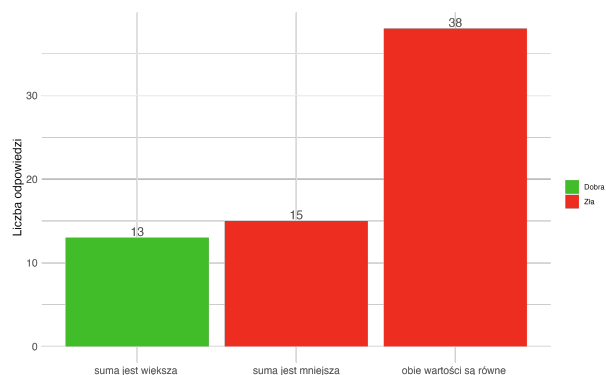
4. Jaka jest relacja pomiędzy sumą temperówek i ściąg a liczbą linijek w klasie B?

- ☒ suma jest większa
☐ suma jest mniejsza
☐ obie wartości są równe

Wyniki ankiety



Udzielone odpowiedzi na pytanie 3.



Udzielone odpowiedzi na pytanie 4.

3.3 Wykres 3



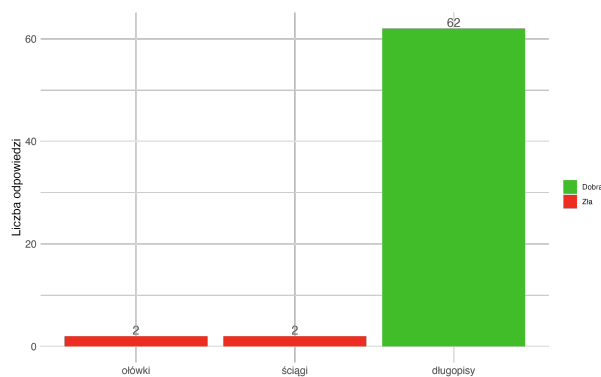
Liczba przyborów szkolnych w piórnikach uczniów klasy C

5. Co uczniowie klasy C częściej noszą w piórnikach ołówki, długopisy, czy ściagi?

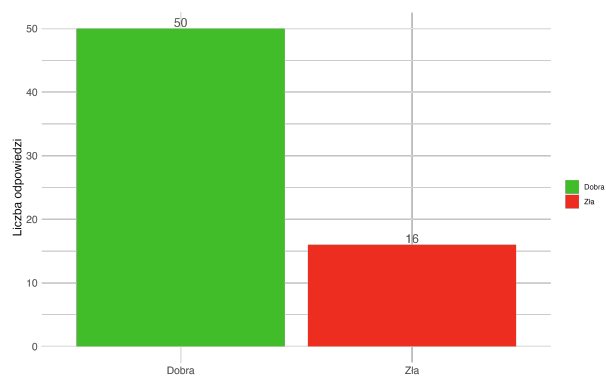
- ☐ ołówki
☒ długopisy
☐ ściagi

6. O ile razy więcej uczniów klasy C nosi linijkę niż korektor?
 Około 2 razy więcej

Wyniki ankiety

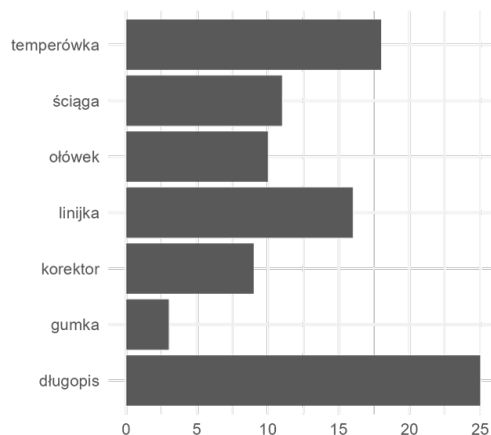


Udzielone odpowiedzi na pytanie 5.



Udzielone odpowiedzi na pytanie 6.

3.4 Wykres 4



Liczba przyborów szkolnych w piórnkach uczniów klasy D

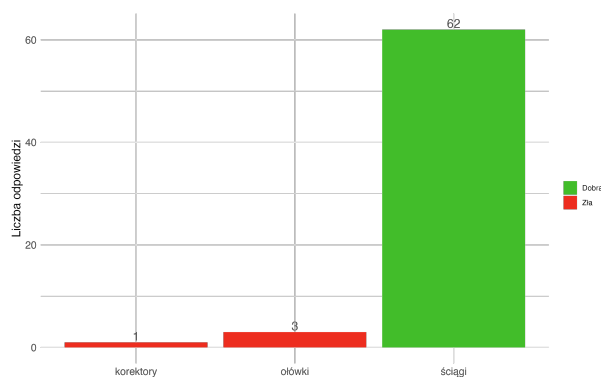
7. Co uczniowie klasy D częściej noszą w piórnkach ołówki, ściągł czy korektory?

- ☐ ołówki
☒ ściągł
☐ korektory

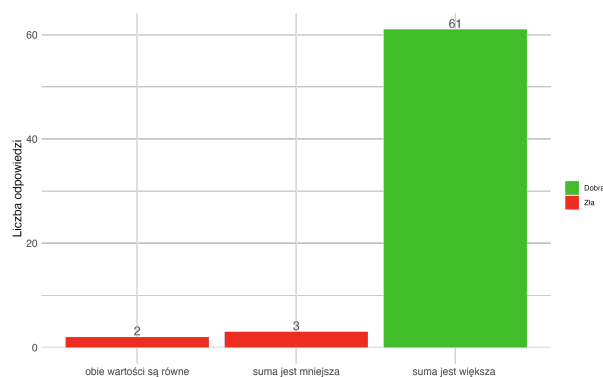
8. Jaka jest relacja pomiędzy sumą linijek i gumek a liczbą korektorów w klasie D?

- ☒ suma jest większa
☐ suma jest mniejsza
☐ obie wartości są równe

Wyniki ankiety

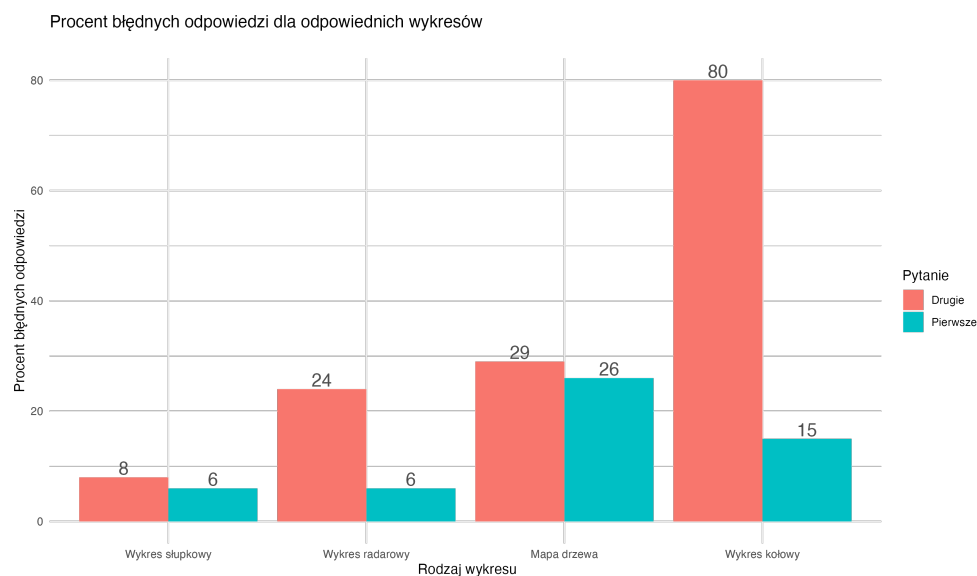


Udzielone odpowiedzi na pytanie 7.



Udzielone odpowiedzi na pytanie 8.

3.5 Podsumowanie pierwszego eksperymentu

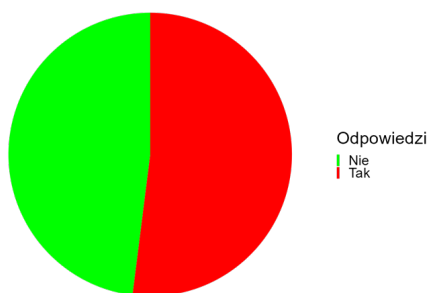


Jak możemy zauważyć, odczytując dane z wykresu słupkowego, mimo braku szczegółowej skali, mylimy się rzadziej. Jesteśmy w stanie lepiej oszacowywać ilość, ale również porównywać między sobą wartości. Wykres radarowy i mapa-drzewo są na podobnym poziomie. Ankietujący mieli trudności z oszacowywaniem sumy oraz porównywaniem jej. Dodatkowo problematyczne było oszacowywaniem ilości, studiując wykres typu mapa-drzewo. Najbardziej wyróżnia się wykres kołowy. Na pytanie dotyczące porównania sumy dwóch przedmiotów większość odpowiedziała błędnie. Wiąże się to prawdopodobnie z kształtem wykresu i brakiem intuicji porównywania obszarów bez dodatkowej skali.

4 Drugi eksperyment

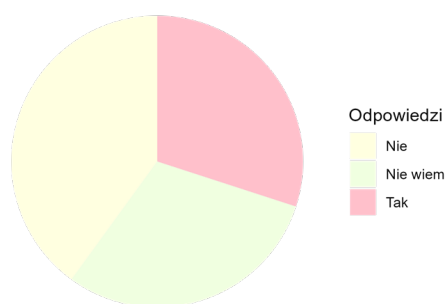
W tej części skupiliśmy się na roli kolorów w poprawnym odczytywaniu wykresów. Użyliśmy barw monochromatycznych, gdzie ukryła się iluzja. Zastosowaliśmy również nieintuicyjne barwy do odpowiedzi, jak i odwróciłyśmy kolorystyczną skalę.

4.1 Część pierwsza



9. Jakiej odpowiedzi udzieliła większość studentów?

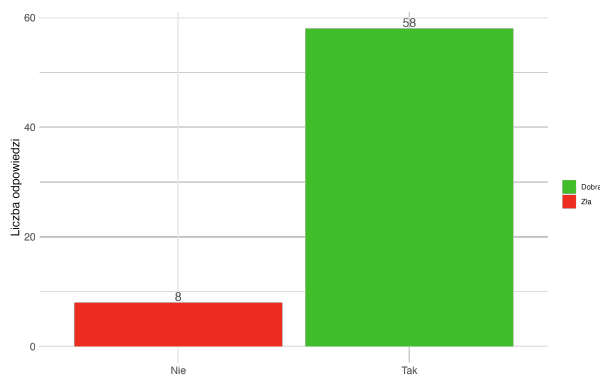
- ☐ Nie
☒ Tak



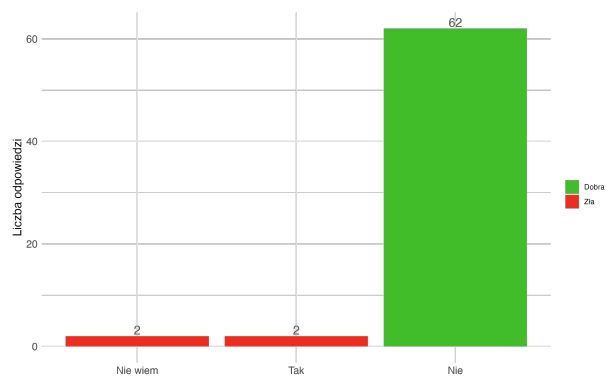
10. Jakiej odpowiedzi udzieliła większość studentów?

- ☐ Tak
☒ Nie
☐ Nie wiem

Wyniki ankiety

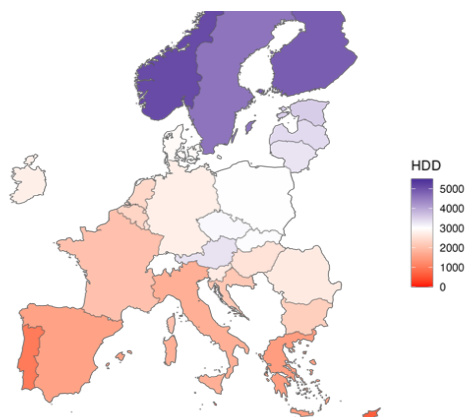


Udzielone odpowiedzi na pytanie 9.



Udzielone odpowiedzi na pytanie 10.

4.2 Część druga



11. Wskaźnik HDD jest wyższy w Szwecji czy Hiszpanii?

☐ w Hiszpanii

☒ w Szwecji

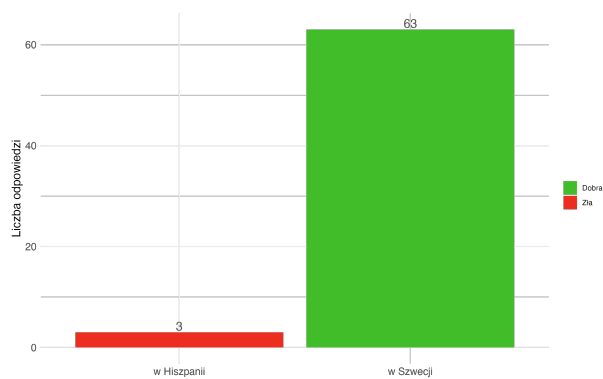


12. Wskaźnik CDD jest wyższy w Niemczech czy w Słowacji?

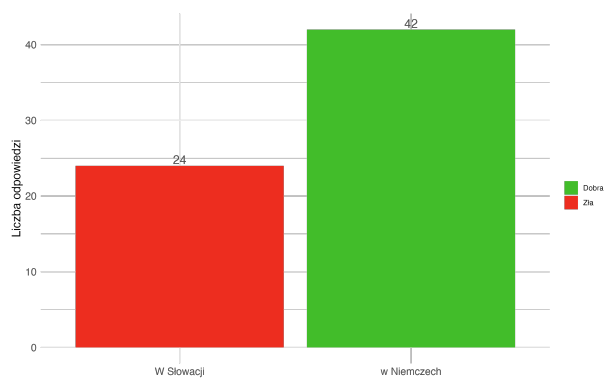
☒ w Niemczech

☐ w Słowacji

Wyniki ankiety

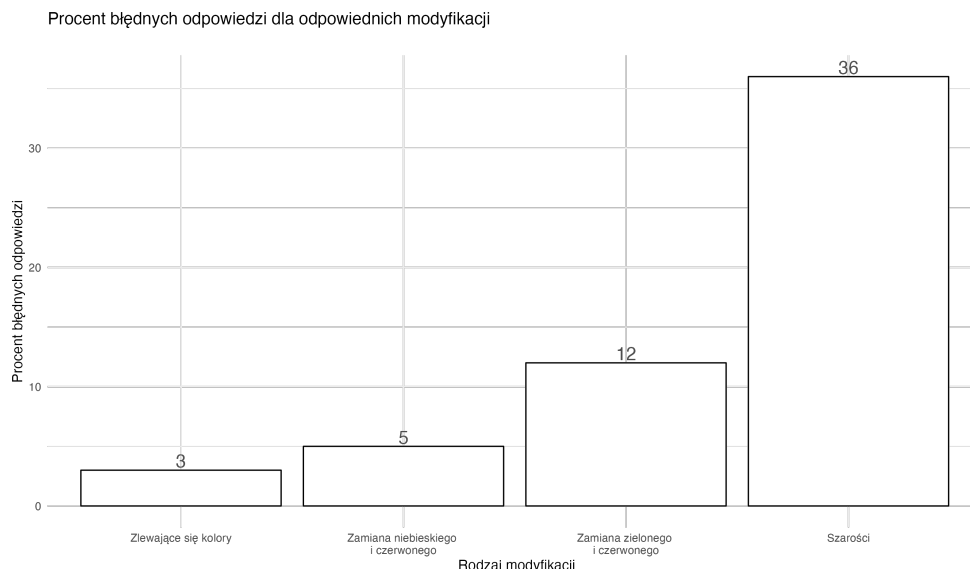


Udzielone odpowiedzi na pytanie 11.



Udzielone odpowiedzi na pytanie 12.

4.3 Podsumowanie drugiego eksperymentu



Jak można zauważyć, najbardziej mylący okazał się wykres, na którym zastosowaliśmy skalę szarości. Zamiana kolorów na nieintuicyjne oraz wybranie palety bardzo zbliżonych do siebie odcieni miała wpływ na odczytywanie wykresu, jednakże nie w tak dużym stopniu. Na szczególną uwagę zasługuje zamiana kolorów zielonego i czerwonego, powszechnie utożsamianych z odpowiedziami „tak” i „nie”, gdzie co dziesiąty ankietowany wybrał niepoprawną odpowiedź.

5 Wnioski

Wykresy z wyraźnymi i czytelnymi osiami wydają się skuteczniejsze, gdy potrzebujemy odczytać z nich dane. W sytuacji, gdy musimy bazować wyłącznie na naszej intuicji, o wiele częściej zdarzają się pomyłki. Tak było w przypadku wykresów typu kołowego oraz mapy drzewa, gdzie pojawiały się większe i mniejsze problemy z porównywaniem kilku wartości obok siebie. Jedynie 20% ankietowanych było w stanie poprawnie porównać sumę dwóch wielkości do trzeciej, odczytując dane z wykresu kołowego, co czwarty ankietowany popełnił błąd przy wskazywaniu wartości największej z trzech podanych zaprezentowanych na wykresie typu mapy drzewa. W przypadku wykresu radarowego i słupkowego ankietowani odpowiadali w większości poprawnie, z wyjątkiem porównywaniem krotności liczby przedmiotów. W reszcie przypadków oba wykresy okazały się dobrymi narzędziami do wizualizacji danych.

Nie tylko kształt wykresów jest istotny. Kolor odgrywa ważną rolę w odbiorze wykresu, starając się nie przesadzić z jego wielobarwnością, należy pamiętać, że brak koloru nie zawsze jest najlepszym rozwiązaniem. Jak pokazał nasz eksperyment, w przypadku konieczności odróżnienia od siebie dwóch bliskich odcieni szarości, nie wszyscy są w stanie sobie z tym poradzić. Warto również trzymać się intuicyjnych kolorów, ponieważ choć większość ankietowanych była zdolna podać poprawną odpowiedź, zdarzyły się jednostki, które zasugerowały się użytymi kolorami na wykresie. Zaburza to odbiór wykresu i może być formą manipulacji danymi. Zamiana koloru niebieskiego na czerwonego, jak i użycie zlewających się kolorów nie zmyliła ankietowanych, którzy byli w stanie poprawnie odczytać dane z wykresu.

Można zatem stwierdzić, że wykresy słupkowe, jak i radarowe są lepszym sposobem do ukazywania danych niż wykresy typu mapy drzewa czy typu kołowego. Dodatkowo nie należy kolorów traktować jako obowiązkowe ozdoby wykresów, a używać ich, gdy ukazywane dane tego wymagają.