

Raport - praca domowa 3+4

Adam Frej, Piotr Marciniak

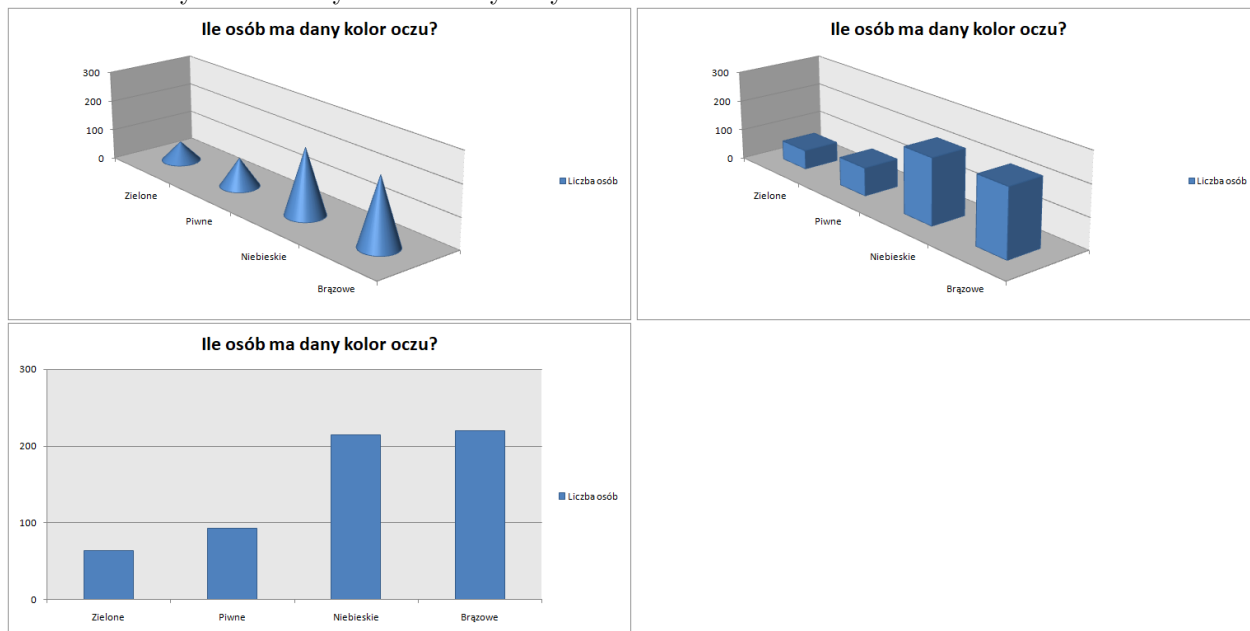
Wstęp

W celu sprawdzenia jakie wykresy są czytelne w wizualizacji danych przeprowadziliśmy ankietę zawierającą różne wykresy. Skupiliśmy się głównie na porównaniu wykresów trójwymiarowych do płaskich. W ramach zadania zrobiliśmy dwa eksperymenty. W ankiecie wzięło udział 15 osób.

Link do ankiety: <https://forms.gle/a9g4PdDVLB9yZeYg7>

Eksperyment 1 - wykresy słupkowe

Przedstawiliśmy ankietowanym 3 różne wykresy:



Pod każdym wykresem zadaliśmy dwa pytania:

- Ile osób ma niebieskie oczy? - poprawna odpowiedź: 215
- Jaka jest różnica między liczbą osób z zielonymi oczami a z piwnymi? - poprawna odpowiedź: 29

Chcieliśmy sprawdzić jak zmienia się trudność odczytywania wysokości słupków i porównywania ich między sobą.

Odczytywanie wysokości słupków - pierwsze pytanie

Przy wykresach 3d ankietowani zdecydowanie zaniżali odpowiedzi. Perspektywa znacznie utrudniała odczytanie dokładnej wysokości słupka. Powodowała, że słupki były niższe niż w rzeczywistości i wydawały się mniejsze. W dodatku w pierwszych wykresach odpowiedzi były bardziej rozstrzelone. Mniej więcej połowa osób zdawała sobie sprawę, że wykresy zniekształcają dane i próbowała strzelać wyższe wartości. Jednak wciąż nie byli

oni w stanie określić jak wielką poprawkę trzeba nałożyć.

Dodatkowo pierwszy wykres miał słupki zastąpione stożkami. To sprawiało, że dane wydawały się jeszcze mniejsze. Przy nim było więcej małych odpowiedzi.

W pierwszych dwóch wykresach mniej więcej 1/3 osób dała odpowiedź w okolicy 200. Reszta raczej zaniżała odpowiedzi. W ostatnim wykresie 2d wszystkie odpowiedzi były w okolicy dobrego wyniku.

Porównywanie słupków między sobą - drugie pytanie

Przy wykresach 3d rozstrzał odpowiedzi był ogromny. Od 10 do 50. Część osób wyraźnie zaniżała wynik, a część zawyżała. Przy wykresie stożkowym 3d 13% dało odpowiedź 30 - praktycznie dobrą, a przy słupkowym 3d 7% osób. Przy obu wykresach większość zaniżała odpowiedzi. Trójwymiarowość zacierza różnicę między wielkością słupków. Ciężiej jest je "odjemować" od siebie.

Przy wykresie 2d 50% dało dobrą odpowiedź, a 25% bliską odpowiedź: 35.

Wnioski

Wykres 2d daje świetne rezultaty. Pomimo niedokładnej podziałki i niepodpisanych słupków ankietowani dawali bardzo dokładne odpowiedzi.

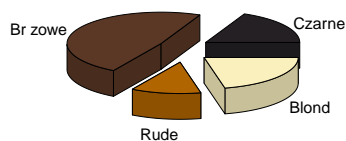
Oba wykresy 3d pozwalają na ogromną manipulację danymi. Są słabe zarówno do odczytywania słupków jak i do porównywania ich między sobą.

Eksperyment 2

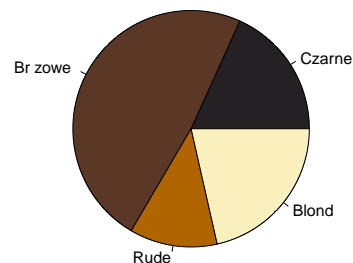
Chcieliśmy sprawdzić jak dane, które są częścią pewnej większej całości, mogą zostać odczytane w zależności od ich przedstawienia. Do tego celu użyliśmy 3 typów wykresów

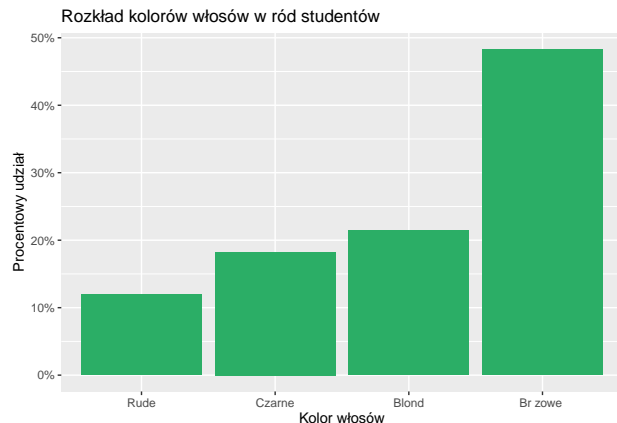
- kołowy 3d
- kołowy 2d
- słupkowy

Rozkład kolorów włosów w ród studentów



Rozkład kolorów włosów w ród studentów





W celu sprawdzenia zadaliśmy dwa pytania o relacje między danymi.

- Która część przedstawiająca kolor włosów jest większa? (odpowiedziami były część blond oraz czarna) Odpowiedzią na to pytanie był kolor blond, różnica między wynikami wynosiła około 3%.
- Ile razy więcej jest studentów o brązowym kolorze włosów niż o blond? Odpowiedzią na to pytanie było około 2,25. W ankiecie poprawną odpowiedź na to pytanie było pomiędzy 2 a 3 razy, ponieważ chcieliśmy sprawdzić generalną relację, a nie dokładny stosunek.

Porównanie wyników

Wykres 3D kołowy

- Tylko 20% osób wskazało prawidłową część włosów (blond), reszta udzieliła błędnej odpowiedzi na to pytanie. Błędne odpowiedzi wynikają z faktu, że osoby te skupiły się na porównaniu objętości figur a nie kątów koła, co przy zadanej perspektywie było bardzo mylące.
- Z kolei na drugie pytanie, aż 93,3% osób odpowiedziało, że pomiędzy 2 a 3 razy jest więcej studentów o brązowym kolorze włosów niż o blond. Tylko 1 osoba wskazała, że jest więcej niż 3 razy.

Wykres 2D kołowy

- Tutaj z kolei proporcje się odwracają i aż 93,3% wskazuje, że osób o blond kolorze włosów jest więcej niż czarnym. Dobre wyniki sugerują, że porównywanie kątów ostrych między sobą wcale nie należy do najtrudniejszych.
- Na drugie pytanie 20% osób wskazało błędną odpowiedź, że osób o brązowym kolorze włosów jest 2 razy mniej niż o blond. Reszta wskazała poprawną odpowiedź. Błędy w odpowiedziach zapewne wynikają z trudności odczytania kąta rozwartego, to spowodowało zaniżenie wartości tego kąta.

Wykres słupkowy

- Wszyscy wskazali poprawną odpowiedź na pierwsze pytanie.
- Tylko jedna osoba wskazała, że stosunek ludzi o brązowym kolorze do blondynów jest mniejszy niż 2 razy. Reszta wskazała poprawną odpowiedź

Wnioski

- Wykres kołowy 3D jest w miarę łatwo zmanipulować, ponieważ ludzie zwracają większą uwagę na objętości niż na kąty.
- Wykres kołowy 2D pokazuje nam, że osoby dosyć dobrze radzą sobie z odczytywaniem kątów ostrych, natomiast mają problemy z kątami rozwartymi.
- Wykres słupkowy nie sprawiał prawie żadnych problemów z odczytaniem wartości.

Ogólne wnioski

W obu eksperymentach wykresy 2d dają najlepsze rezultaty. Trójwymiarowość zdecydowanie zmniejsza czytelność wykresów. Jeżeli jest dodatkowo połączona z nietypowymi słupkami lub wykresem kołowym to takie wizualizacje stają się praktycznie bezużyteczne.