

Raport z trzeciej pracy domowej z przedmiotu Techniki wizualizacji danych

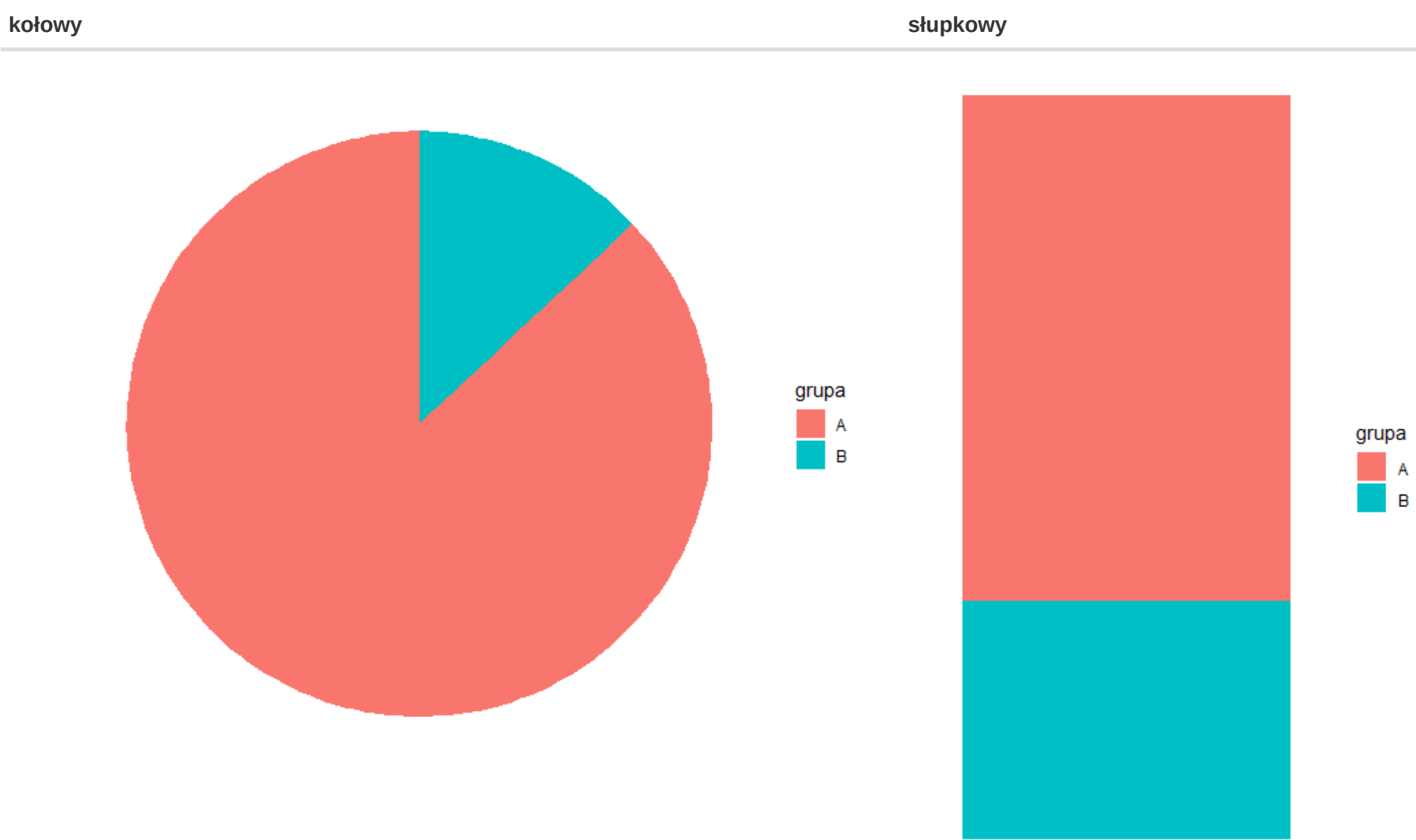
Badanie dokładności odczytywania danych z wykresów kołowych i słupkowych na podstawie proporcji kątów lub wysokości

Alicja Charuza

Cel i opis badania

W ramach tej pracy domowej zdecydowałam się sprawdzić, czy człowiek rzeczywiście mniej dokładnie szacuje wzajemne relacje kątów niż wysokości. Przeprowadziłam ankietę, w której to badani mieli oszacować procentowy udział danej grupy danych na wykresie. Na potrzeby badania wykresy nie były podpisane, jedyne na co mogli liczyć ankietowani był wzajemny stosunek wysokości lub kątów. Na koniec mieli oni stwierdzić, który typ wykresu w ich opinii sprawił im więcej trudności w wykonaniu tego zadania.

Przygotowałam 8 wykresów przedstawiających losowo wygenerowane dane (każdy inne) - 4 słupkowe skumulowane oraz 4 kołowe, przy czym każda para słupkowy-kołowy miała inną liczbę grup danych na nich przedstawioną (od 2 grup do 5). Przykładowa taka para wygląda następująco:

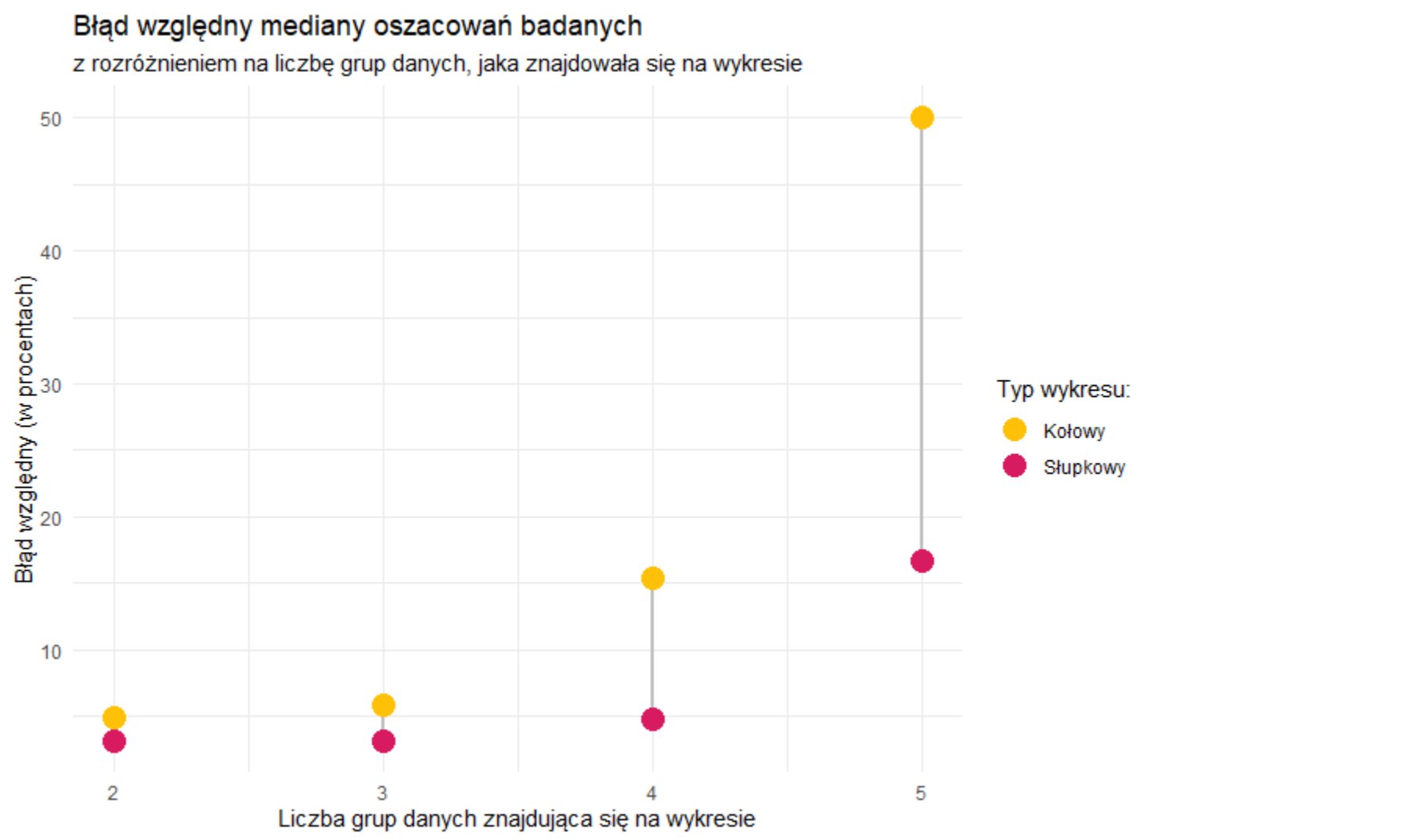


Otrzymałam 54 odpowiedzi na moją ankietę. Link do nieprzetworzonych wyników pobranych z Google Forms: https://drive.google.com/file/d/1ZHGxqNNCTaiv8BXSGGFOMPO_iAv223dQ/view?usp=sharing. Badani byli głównie w wieku około 20 lat z nielicznymi wyjątkami. Zwizualizowane wyniki znajdują się poniżej.

Otrzymane wyniki

Rzeczywiste wyniki szacowania

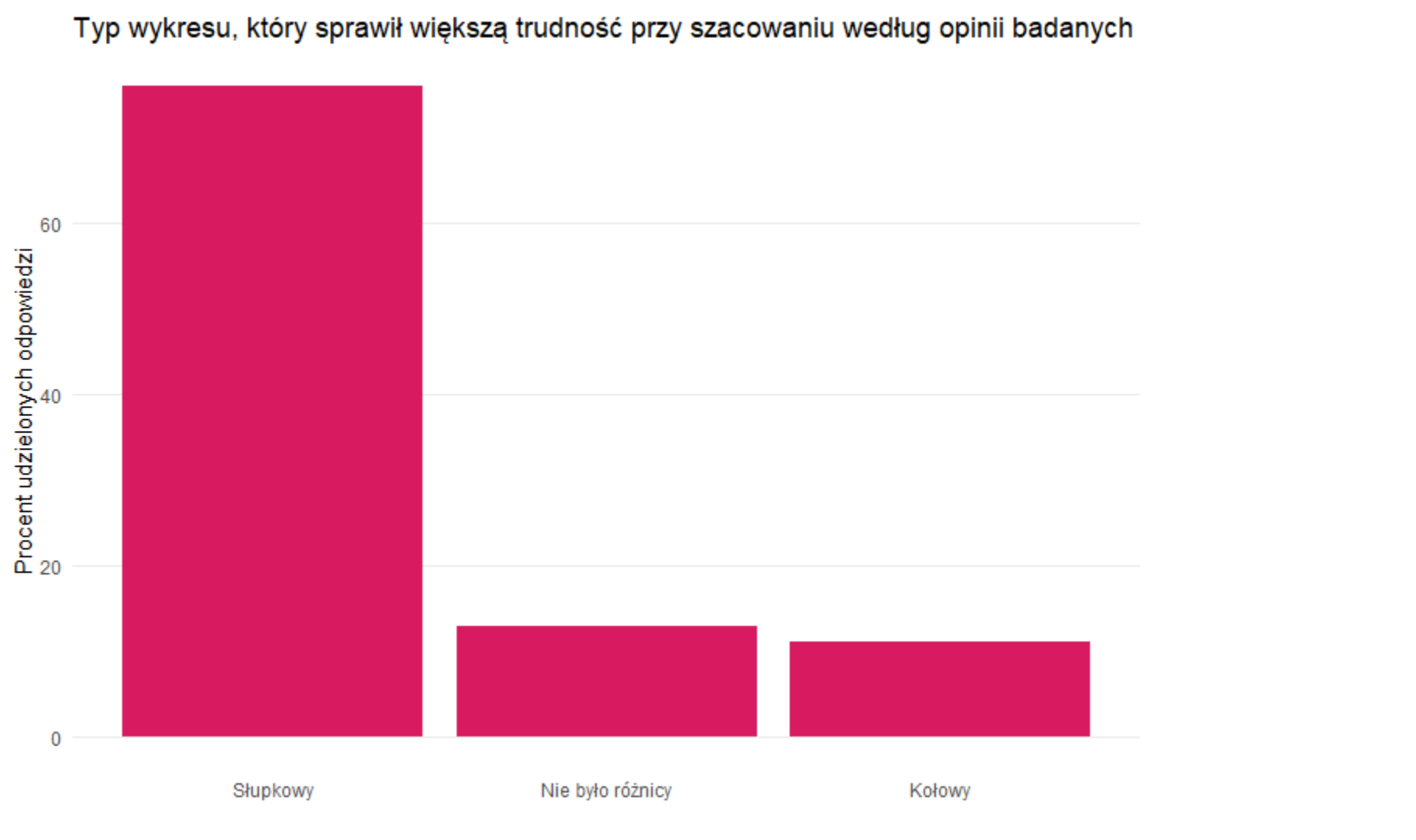
Pierwszym zadaniem badanych było oszacowanie udziału procentowego danej grupy danych na wykresie kołowym lub słupkowym. Przeanalizowałam powstałe dane, licząc medianę odpowiedzi, a następnie jej błąd względny. Poniższy wykres lizakowy przedstawia uzyskane przeze mnie wyniki.



Jak łatwo można się domyśleć dokładność szacunków zarówno w przypadku wykresu kołowego, jak i słupkowego spadała wraz ze wzrostem liczby obecnych grup danych. Jednak przy wszystkich liczbach grup danych, błąd względny mediany odpowiedzi badanych jest większy w przypadku wykresu kołowego. W dodatku w przypadku wykresu kołowego rośnie on szybciej wraz ze wzrostem liczby grup danych. Wyniki potwierdzają tezę, że człowiek mniej dokładnie odczytuje z wykresu kąty niż wysokości.

Opinia badanych odnośnie poziomu trudności szacowania

Na koniec badani byli zapytani o to, który typ wykresu sprawił im więcej trudności przy szacowaniu. Stwierdziłam, że uzyskanie subiektywnej opinii ankietowanych na ten temat również dobrze pokaże nastawienie społeczeństwa do dobrej praktyki, jaką jest stosowanie wykresów słupkowych zamiast kołowych.



Zdecydowana większość badanych, bo aż 75,9%, odpowiedziała, że większą trudność przy szacowaniu sprawił im wykres słupkowy, natomiast pozostali podzieliili się mniej więcej po równo między brakiem znaczenia, a kołowym. Jest to dość interesujące, że mimo bardziej trafnych odpowiedzi w przypadku wykresów słupkowych, badani uznali szacowanie za łatwiejsze w przypadku wykresu kołowego albo też stwierdzili, że nie było różnicy w trudności wykonywanych zadań.

Wnioski z badania i podsumowanie

Przed wszystkim wyniki potwierdzają tezę, że człowiek lepiej odczytuje wysokości i ich wzajemne relacje z wykresów, niż kąty. W każdej parze wykresów szacowania badanych okazały się bardziej trafne dla wykresu słupkowego skumulowanego. Nie jest to dobra wiadomość dla zapalonych miłośników wykresów kołowych,

Szczerze mówiąc uznałam kontrast między rzeczywistymi wynikami szacowania, a subiektywną opinią badanych za dość zabawny. Bardzo prawdopodobne jest to, że przez to, iż nasze społeczeństwo jest wystawione częściej na kontakt z wykresami kołowymi, niż z wykresami słupkowymi skumulowanymi, ankietowani uznali za łatwiejsze to, co wydawało się im bardziej znajome. Byłaby to naturalna reakcja mózgu człowieka (miałam nawet takie podejrzenia przy tworzeniu ankiety, jednak stwierdziłam, że nieopisane zwykłe wykresy słupkowe byłyby jeszcze trudniejsze do analizy, gdyż badani nie mieliby jednoznacznej wysokości oznaczającej "100%"). Jednak mogę jedynie spekulować w tej kwestii.

Interesujące było również to, jak bardzo liczba grup danych wpływała na wzrost błędu względnego w przypadku wykresu kołowego. Jako, że były to jedynie 4 różne liczby, był to być może przypadek spowodowany wygenerowanymi losowo danymi (np nie pojawiły się do oszacowania wartości 25%, 50%, 75%, które ze względu na kąt prosty dawałaby pewną "taryfę ulgową" wykresom kołowym), jednak również pokazywałoby to, jak relacje między kątami są w stanie oszukiwać nasze oczy.

Na sam koniec chcę powiedzieć, że przeprowadzenie tego eksperymentu stanowiło naprawdę ciekawe doświadczenie. Także badani, jak informowali mnie po wypełnieniu ankiety, mieli przyjemność z szacowania, a następnie porównywania swoich wyników do rzeczywistej wartości. Pokazało to mi, a także przy okazji ankietowanym, realny obraz różnicy pomiędzy tymi dwoma typami wykresów.