

Raport z pracy domowej nr 3

Izabela Telejko, Wiktoria Śliwińska

5/3/2022

1. Wprowadzenie

W ramach pracy domowej nr 3 należało przygotować eksperymenty sprawdzające czy określone problemy z odczytywaniem danych z wykresów wciąż występują oraz czy “dobre praktyki” wizualizacyjne rzeczywiście działają. W celu weryfikacji tego zagadnienia przygotowaliśmy elektroniczną ankietę składającą się z 4 wykresów oraz pytań do nich. Ich dokładna treść ukazana jest w dalszej części raportu. Link do ankiety: <https://docs.google.com/forms/d/1UyVXu3cAVQgRq0n1uSuDo5UmVM9eP3Ct9okhphJjTvs/edit#responses>

2. Omówienie ankiety

Ankietę podzieliliśmy na 2 sekcje - każda z nich dotyczyła dwóch wykresów przedstawiających te same dane, ale na różne sposoby. W każdej z tych sekcji zadałyśmy 5 pytań. 3 z nich weryfikowały stopień dokładności odczytu danych przez wypełniających ankietę - zarówno podawanie dokładnych liczb na podstawie wykresu, jak i porównywanie danych. Pozostałe dwa pytania dotyczyły odczuć ankietowanych względem “poprawnego” i “błędnego” wykresu z pary. Polegały one na wybraniu wykresu, z którego lepiej się odczytywało dane oraz ewentualnym uzasadnieniu dlaczego.

3. Grupa ankietowanych

Grupę ankietowanych stanowiły osoby z naszego najbliższego otoczenia, czyli znajomi, przyjaciele i rodzina, a więc przedział wiekowy ankietowanych był dość duży. W sumie zebraliśmy odpowiedzi od 15 osób.

4. Omówienie wyników

Sekcja 1

W pierwszej sekcji ankietowanym zostały przedstawione dwa wykresy dotyczące wzrostu osobników poszczególnych gatunków ze świata Starwars (ramka danych Starwars). Pierwszy z wykresów był 3D, co już powinno zaburzać zdolność rzetelnego odczytywania danych. Ponadto był on ustawiony w specyficznej perspektywie, która mogła powodować wrażenie, że oś OY nie zaczyna się od 0. Drugi wykres był klasycznym wykresem słupkowym, z danymi uporządkowanymi malejąco oraz informacji o przynależności do gatunku znajdującej się na osi OX.

Wzrost przedstawicieli poszczególnych gatunków

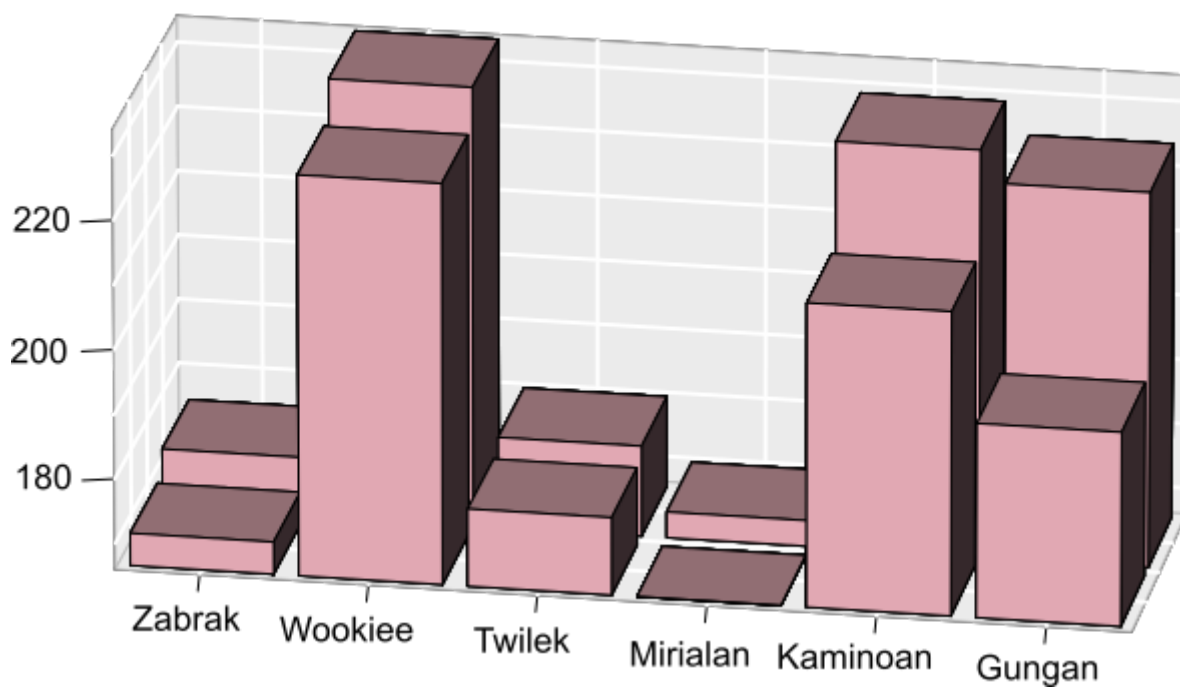


Figure 1: Wykres 1

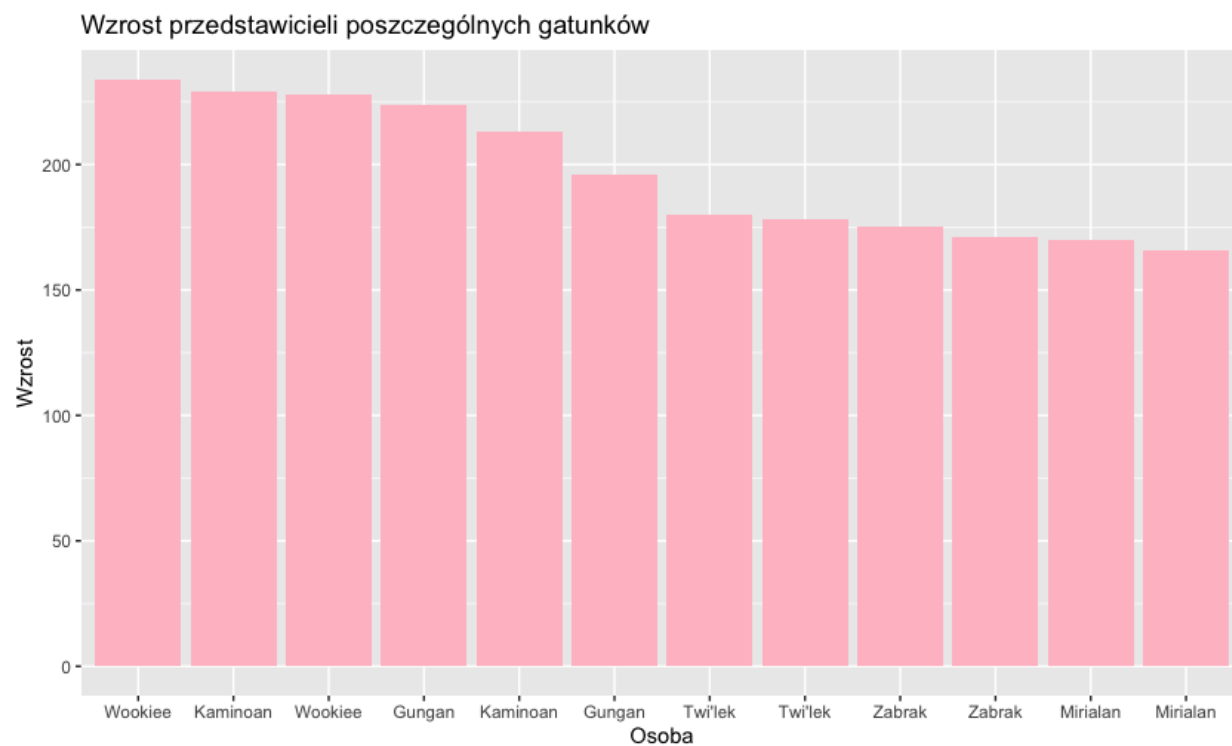


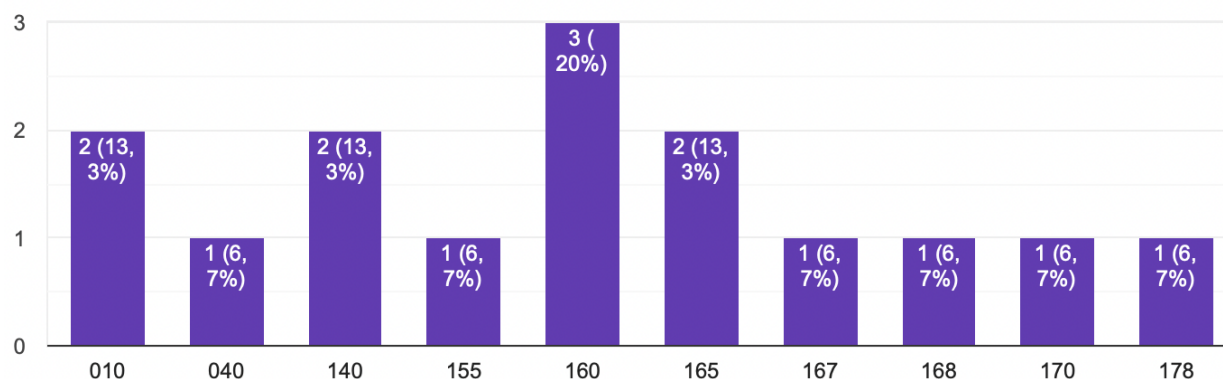
Figure 2: Wykres 2

Porównanie wyników Wykres 1 i Wykres 2

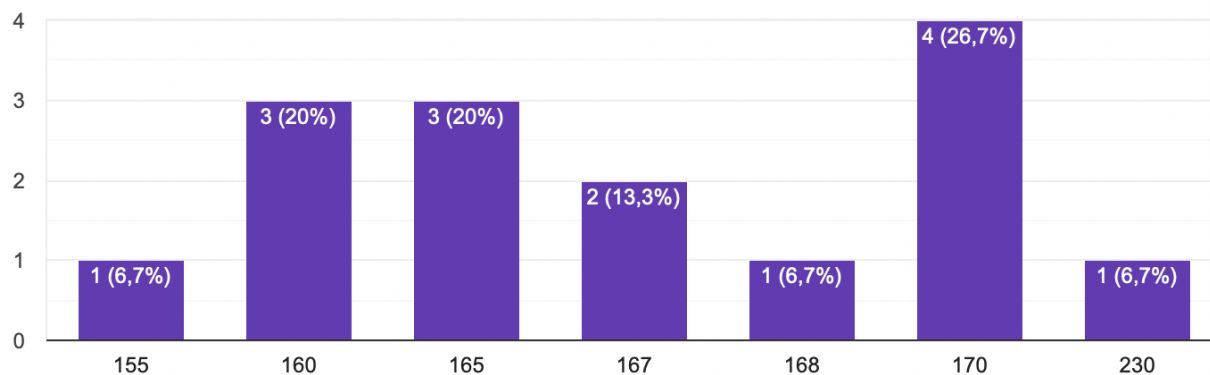
Zadałyśmy 3 pytania sprawdzające jak ankietowani poradzą sobie z odczytywaniem i porównywaniem danych na podstawie danych wykresów.

Pytanie 1 - Jakiego wzrostu jest najniższa osoba?

Wykres odpowiedzi na podstawie Wykresu 1



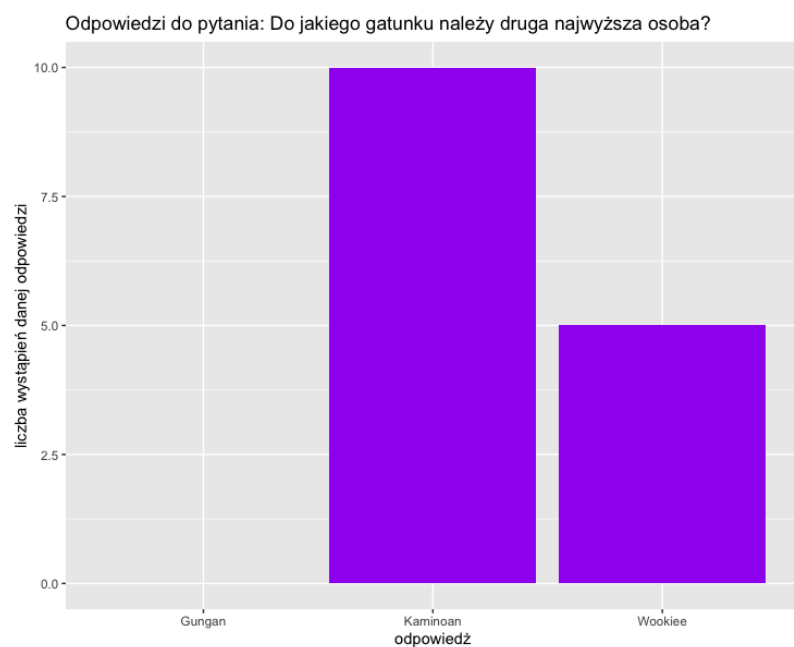
Wykres odpowiedzi na podstawie Wykresu 2



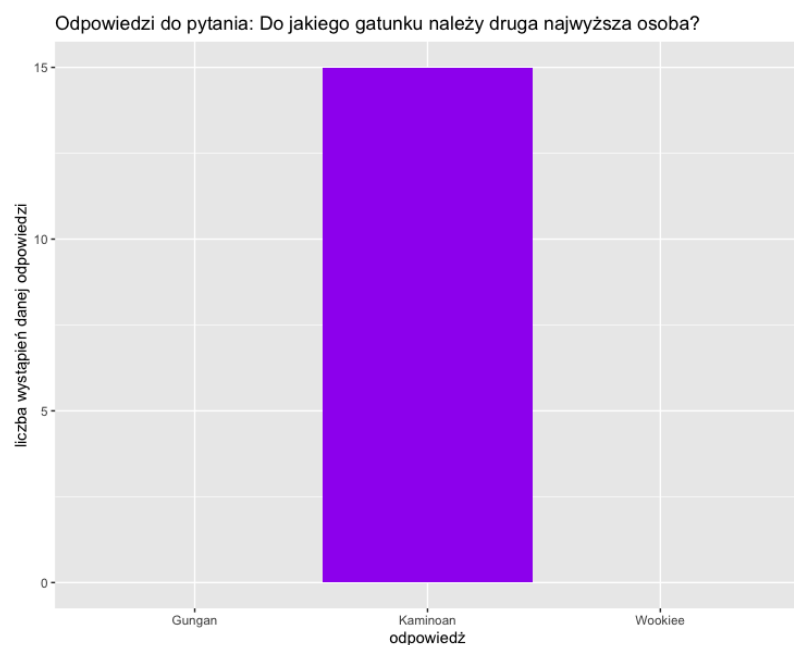
Można zauważyć, że w przypadku odczytywania z Wykresu 2 poza dwoma odstającymi wynikami odpowiedzi mieszczą się w przedziale 160 - 170. Biorąc pod uwagę, że poprawna odpowiedź to 166 to ankietowani byli bliscy właściwej wartości. W przypadku Wykresu 1 wyniki w dużej ilości przypadków są zaniżone, a także zdarzają się wyniki bliskie 0 (aż 20% odpowiedzi), co wskazuje na intuicyjne utożsamianie niskich wysokości słupków z zerowymi wartościami.

Pytanie 2 - Do jakiego gatunku należy druga najwyższa osoba?

Wykres odpowiedzi na podstawie Wykresu 1



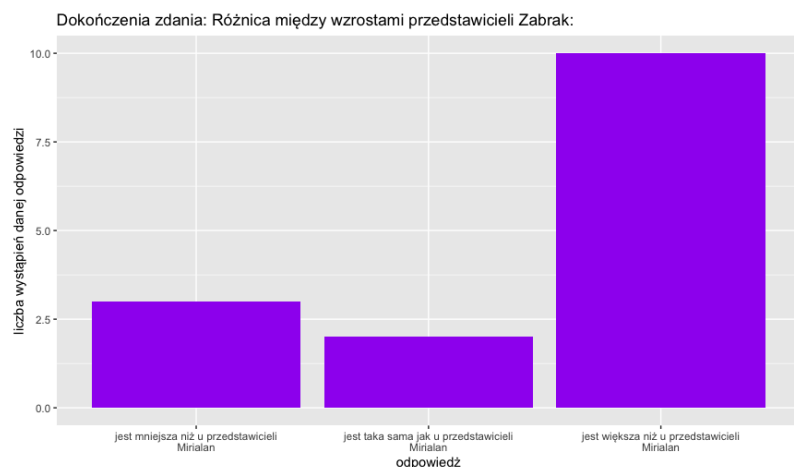
Wykres odpowiedzi na podstawie Wykresu 2



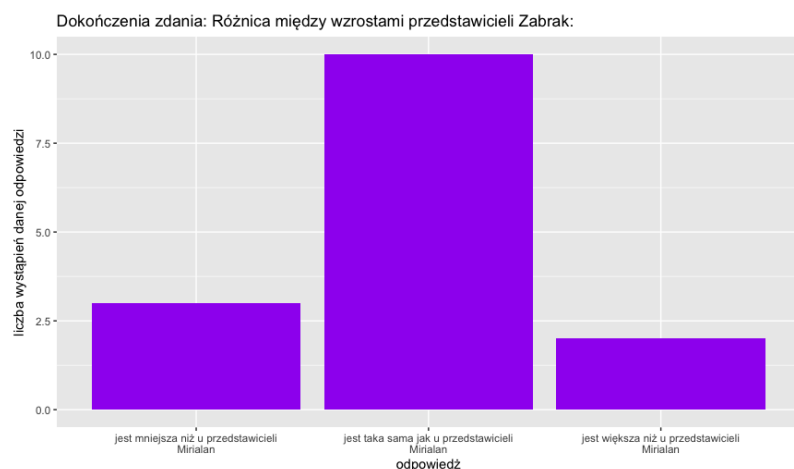
Przy odczytywaniu danych z pierwszego wykresu $\frac{1}{3}$ odpowiedzi była błędna. Natomiast na podstawie drugiego wykresu wszyscy ankietowani udzielili poprawnej odpowiedzi.

Pytanie 3 - Dokończenie zdania: Różnica między wzrostami przedstawicieli Zabrak:

Wykres odpowiedzi na podstawie Wykresu 1



Wykres odpowiedzi na podstawie Wykresu 2



Przy pierwszym wykresie najczęściej udzielaną odpowiedzią było to, że różnica jest większa (67% ankietowanych), a poprawna odpowiedź, czyli, że różnica jest taka sama była wybierana najrzadziej. W przypadku drugiego wykresu większość respondentów udzieliła prawidłowej odpowiedzi

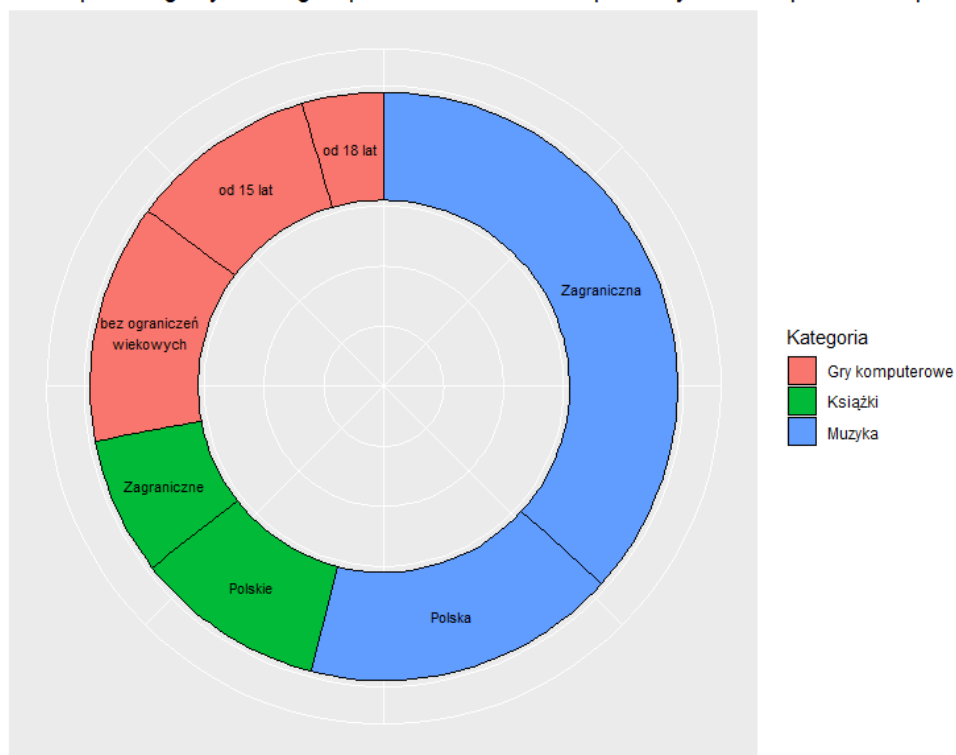
Wnioski na podstawie odpowiedzi na pytania do Wykresu 1 i Wykresu 2

Na podstawie analizy odpowiedzi na 3 zadane przez nas pytania można zauważyć, że zastosowanie na wykresie 3D osi nie zaczynającej się w 0 było mylące. Ponadto ankietowani lepiej odczytywali i porównywali wartości z Wykresu 2. Przewagę wykresu drugiego nad pierwszym potwierdziło 80% respondentów, odpowiedź uzasadniając tym, że Wykres 2 jest “bardziej przejrzysty, lepiej widać hierarchię wysokości i różnice między wzrostami poszczególnych osobników, a oś zaczyna się od zera”

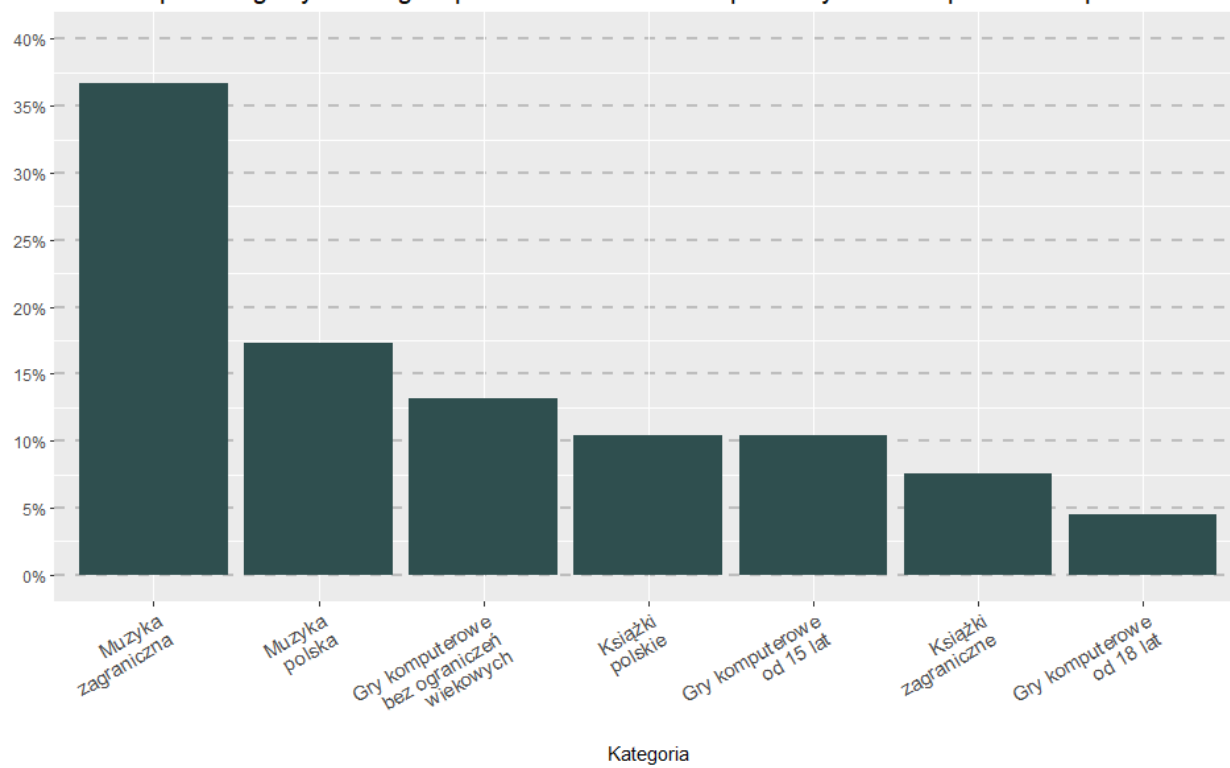
Sekcja 2

W drugiej sekcji ankietowanym zostały przedstawione dwa wykresy dotyczące udziału poszczególnych kategorii produktów w wartości sprzedaży sieci sklepów Kulturopol w 2017 r. Pierwszym z wykresów w tej parze był wykres pączkowy pokolorowanych według kategorii produktów, gdzie podkategorie były tylko podpisane bezpośrednio na wykresie. Drugim wykresem ponownie był klasyczny wykres słupkowy, jednokolorowy, posortowany malejąco z informacją o kategorii i podkategorii umieszczoną na osi OX.

Udział poszczególnych kategorii produktów w wartości sprzedaży sieci sklepów Kulturopol w 2017

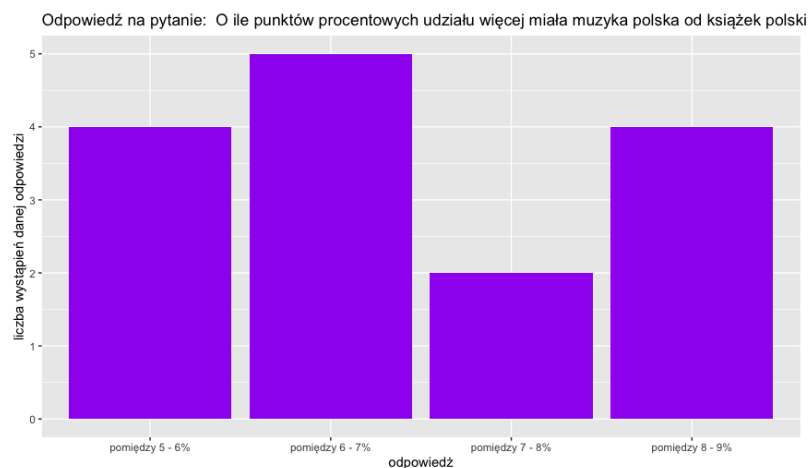


Udział poszczególnych kategorii produktów w wartości sprzedaży sieci sklepów Kulturopol w 2017

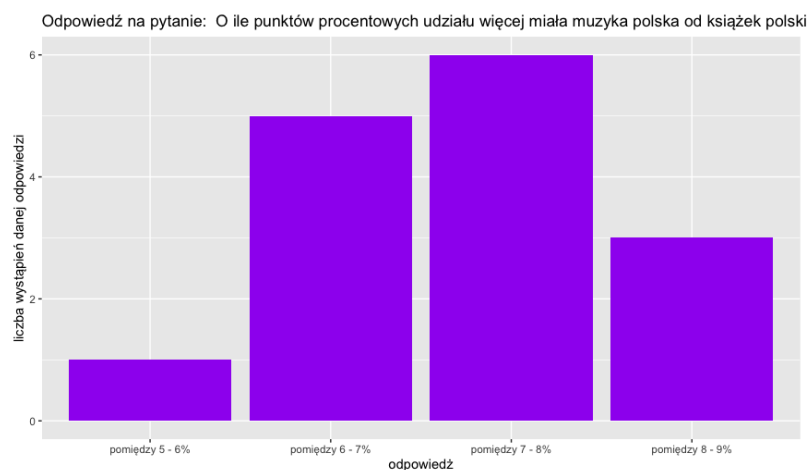


Pytanie 1 - O ile punktów procentowych udziału więcej miała muzyka polska od książek polskich?

Wykres odpowiedzi na podstawie Wykresu 3



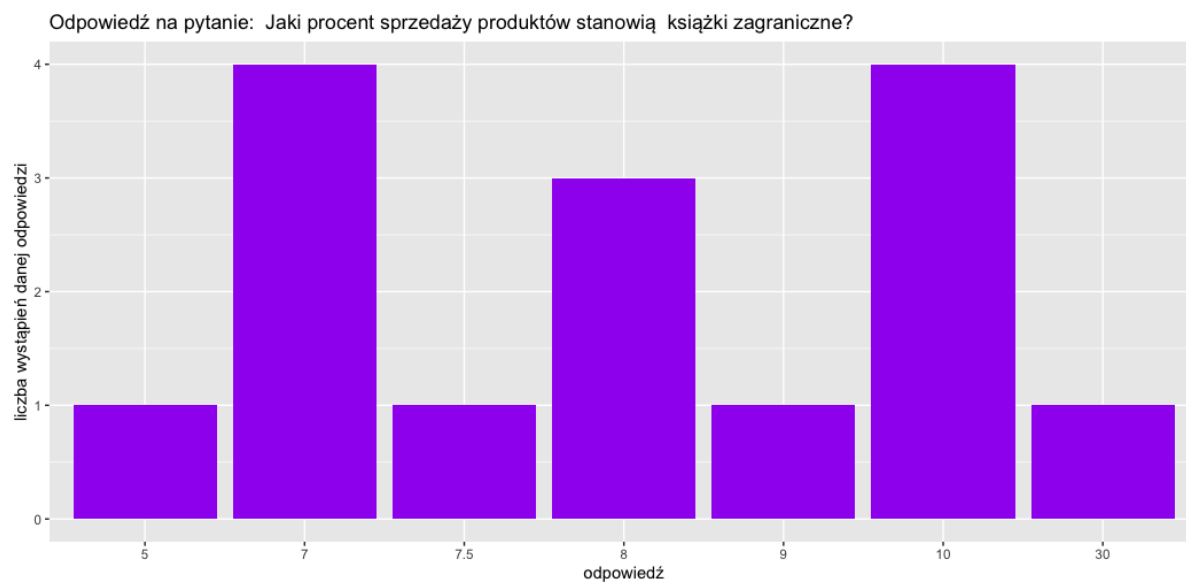
Wykres odpowiedzi na podstawie Wykresu 2



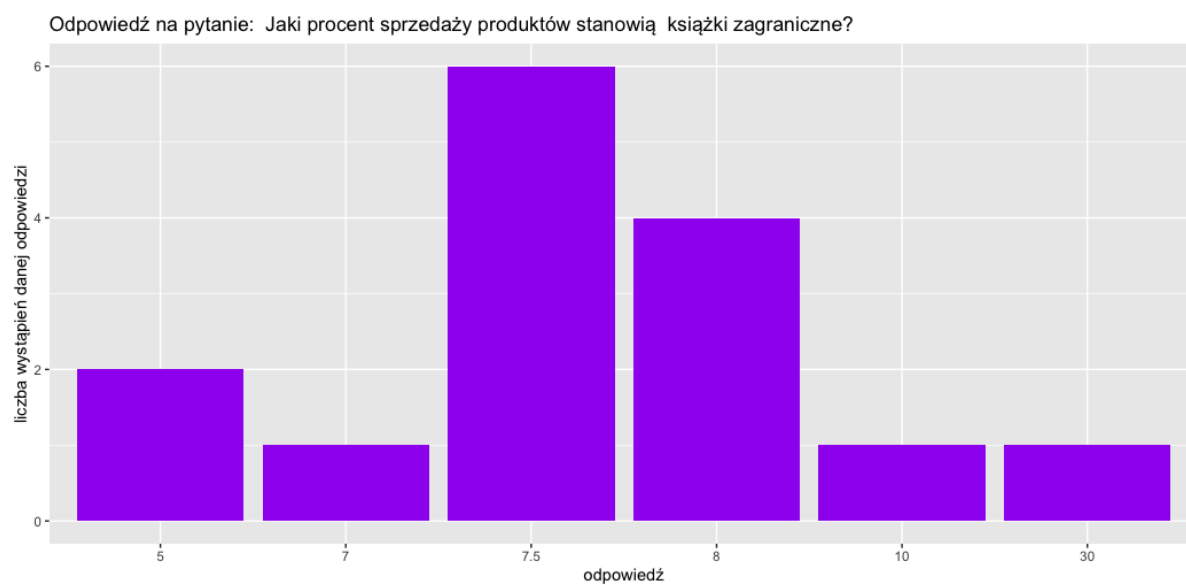
Zarówno w przypadku trzeciego, jak i czwartego wykresu tyle samo osób udzieliło poprawną odpowiedź, czyli pomiędzy 6 - 7%. Jednak w przypadku Wykresu 4 respondenci bardziej przeszacowywali wynik.

Pytanie 2 - Jaki procent sprzedaży produktów stanowią książki zagraniczne?

Wykres odpowiedzi na podstawie Wykresu 3



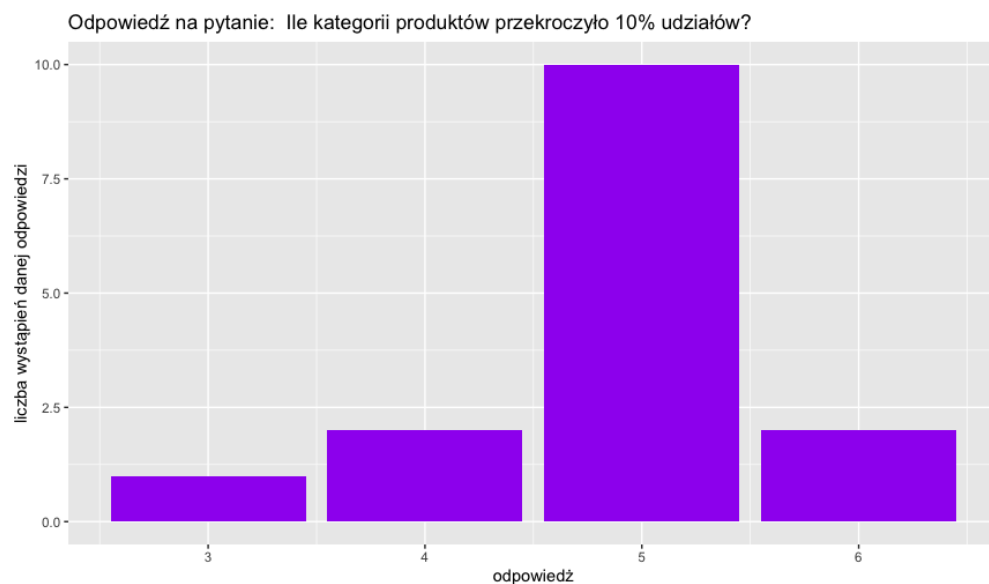
Wykres odpowiedzi na podstawie Wykresu 4



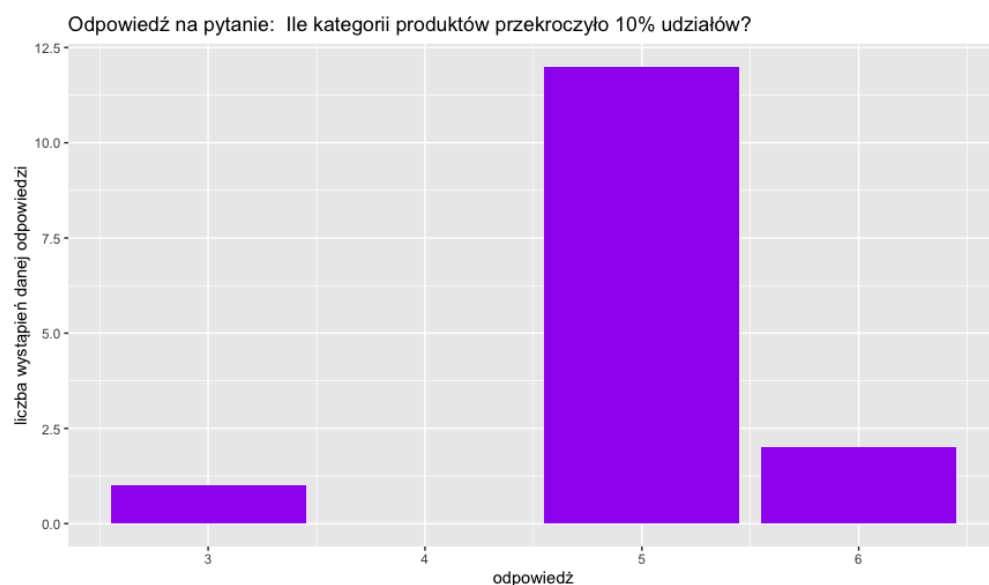
Na podstawie Wykresu 4 jest zdecydowanie więcej poprawnych odpowiedzi, czyli 7.5, przy zdecydowanie niższym udziale odpowiedzi dość odległych jak 5 czy 10, które były dość często wybierane na podstawie Wykresu 3.

Pytanie 3 - Ile kategorii produktów przekroczyło 10% udziałów?

Wykres odpowiedzi na podstawie Wykresu 3



Wykres odpowiedzi na podstawie Wykresu 4



Widać, że w przypadku tego pytania większość ankietowanych udzieliła poprawnej odpowiedzi (5) niezależnie od wykresu, z którego te dane odczytywała, z nieznaczną przewagą (2 osoby) dla Wykresu 4.

Wnioski na podstawie odpowiedzi na pytania do Wykresu 3 i Wykresu 4

W przypadku obu wykresów liczba poprawnych odpowiedzi na pytania była duża. Trochę więcej było ich w pytaniach pod czwartym wykresem. Można zatem stwierdzić, że respondenci dość dobrze radzą sobie z wykresami kołowymi/pączkowymi. Natomiast wykres słupkowy jest jeszcze czytelniejszy. Warto dodać, że ankietowani jednogłośnie wskazali Wykres 4 jako ten przyjemniejszy do odczytywania i porównywania danych.

5. Podsumowanie

Przeprowadzone przez nas eksperymenty wskazują na to, że wykresy słupkowe 2D są łatwiejsze do odczytania wartości, które chcemy przekazać za pomocą wykresu, co potwierdza, że “dobre praktyki” wizualizacji danych są faktycznie skuteczne. Ankieta pokazała również, że wprowadzenie do wykresu 3 wymiaru bardzo zaburza poprawne postrzeganie. Podobnym efektem skutkuje wprowadzenie osi nie zaczynającej się od 0. Z wykresów kołowych również ciężiej jest odczytać dane. Jednak poprawność wyników uzyskiwanych z nich jest dość wysoka, co może sugerować, że społeczeństwo przyzwyczyło się do tego typu wykresów, a w rezultacie dobrze sobie z nimi radzi.