

Katarzyna Łabuś, Emilia Olech, Agnieszka Peret

# Wpływ sposobu prezentacji danych na ich odbiór

## Wstęp

Dane do analizy pochodzą z wbudowanej w Rstudio bazy danych flightnyc13. Przygotowałyśmy 3 pary wykresów przedstawiających dane, aby zbadać, które metody wizualizacji są dla odbiorców najbardziej czytelne. Ankietowanymi były różne osoby, choć w większości studenci PW.

## Typ wykresu a czytelność danych

Drogi Odbiorco, zapraszamy Cię do wzięcia udziału w krótkim formularzu, który dotyczy oceny czytelności wykresów. Zobaczysz w nim trzy pary wykresów przedstawiających dane związane z odlotami i opóźnieniami lotów na nowojorskich lotniskach.

W każdej parze znajdują się dwa różne sposoby wizualizacji tych samych informacji. Twoim zadaniem jest wybranie z każdej pary tego wykresu, który według Ciebie jest bardziej czytelny i przyjemny w odbiorze.

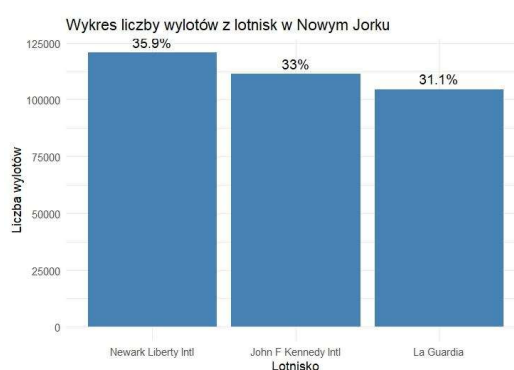
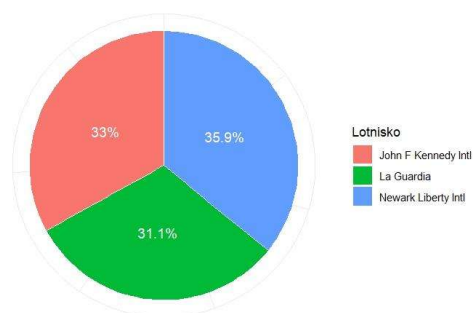
Nie ma dobrych ani złych odpowiedzi – interesuje nas Twoje osobiste odczucie. Formularz ma na celu lepsze zrozumienie, jakie formy prezentacji danych są najbardziej intuicyjne dla odbiorców.

Wypełnienie całości zajmie tylko kilka minut. Dziękujemy za Twój udział!

## 1. Wykres kołowy a słupkowy

Pierwsza para wykresów przedstawia dane informujące o udziale procentowym wylotów konkretnych lotnisk nowojorskich we wszystkich wylotach z miasta. Pierwszy wykres jest kołowy. Każde lotnisko jest reprezentowane przez dany kolor, a na odpowiednich polach są podpisy informujące o odsetku. Drugi wykres przedstawia 3 słupki wraz z podpisanym odsetkiem. Można ponadto odczytać konkretną liczbę odlotów w badanym roku.

Wykres liczby wylotów z lotnisk w Nowym Jorku

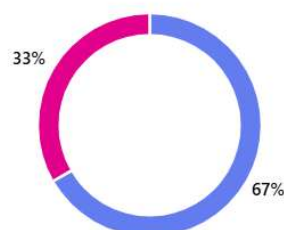


Badanie pokazało, że, w przeciwieństwie do oczekiwań, wykres kołowy jest dla odbiorców bardziej czytelny.  $\frac{2}{3}$  z 45 ankietowanych wybrało pierwszą opcję.

## 1. Który wykres wybierasz?

[Więcej szczegółów](#)

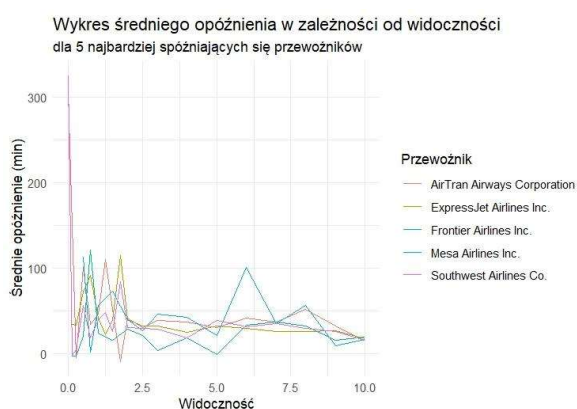
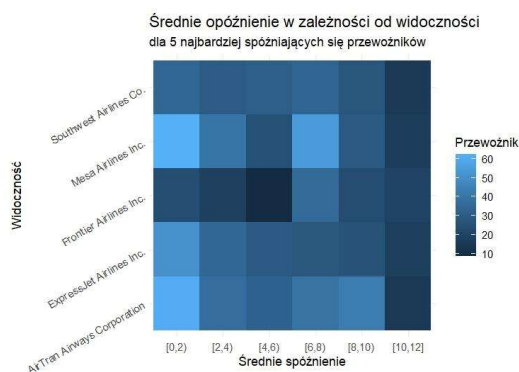
- Opcja 1 30
- Opcja 2 15



Mogło to wynikać z faktu, że słupki obrazowały większą liczbę danych (informacja o konkretnej liczbie), co spowalniało ich analizę. Ponadto wprowadzenie różnych kolorów sprawia, że wykres jest dla odbiorców bardziej przystępny.

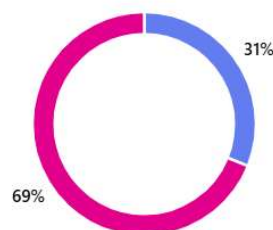
## 2. Błędne podpisy osi oraz nieintuicyjna skala kolorów

Kolejne wykresy obrazują wpływ widoczności na opóźnienia odlotów. Na pierwszym wykresie osie są źle podpisane. Ponadto skala kolorów jest nieintuicyjna (większe wartości średnie opóźnienia są zaznaczane jaśniejszym kolorem niż mniejsze). Na drugim wykresie podpisy są prawidłowe. Problemem natomiast może być 'zlewanie się' linii odpowiadających danym przewoźnikom.



Rozkład głosów był podobny jak w przypadku pierwszego testu. Około  $\frac{2}{3}$  ankietowanych wybrało drugi wykres.

- Opcja 1 14
- Opcja 2 31

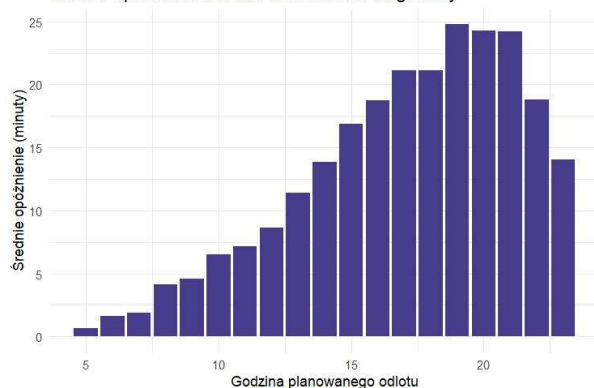


Wybór przez odbiorców drugiego wykresu wynika z błędów na pierwszym. Znacząco utrudniają one odczytanie wykresu i zrozumienie danych, które on przedstawia. Ciekawym jest jednak fakt, że co 3 ankietowany wybrał mimo wszystko wykres kafelkowy. Może to wynikać ze względu na to, że na wykresie liniowym linie są cienkie i się zlewają. Mimo wszystko problemy z podpisami osi oraz kolorystyką na pierwszym wykresie dla większości były większym utrudnieniem niż nachodzące na siebie linie na drugim.

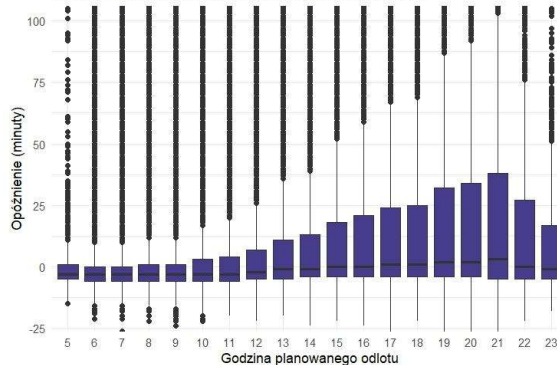
### 3. Ilość przedstawionych danych a czytelność wykresu

Ostatnie wykresy przedstawiają wpływ planowanej godziny odlotu na opóźnienie. Pierwszy wykres to klasyczny wykres słupkowy, z którego można odczytać średnie opóźnienie w zależności od godziny. Drugi wykres to wykres skrzynka - wąsy, z którego możemy odczytać medianę, kwartyle oraz minimum i maksimum zmiennej (średniego opóźnienia).

Średnie opóźnienie odlotów w zależności od godziny



Rozkład opóźnień odlotów w zależności od godziny

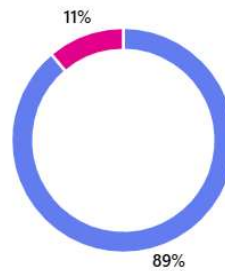


Znacząca większość ankietowanych (około 90%) wybrała pierwszy wykres.

### 3. Który wykres wybierasz?

[Więcej szczegółów](#)

● Opcja 1	40
● Opcja 2	5



Otrzymany wynik może być spowodowany zbyt dużą ilością danych na drugim wykresie. Wykres skrzynka - wąsy jest cenionym narzędziem w statystyce, lecz dla osoby, która nie zajmuje się tą dziedziną, może być on niejasny i trudny w zrozumieniu. W przypadku przygotowywania wizualizacji danych dla przeciętnego odbiorcy lepiej skorzystać z prostszych typów wykresu takich, jak wykres słupkowy.