

[PD3] Frelik Stanislaw Poglód Jan

Stanisław Frelik, Jan Poglód

28 marca 2023

1 Ankieta

Czy wykresy kołowe mają miejsce w wizualizacji danych? Postanowiliśmy to sprawdzić! W szczególności chcemy sprawdzić poziom czytelności wykresów kołowych poprzez zebranie odpowiedzi w prostej [ankiecie](#) porównującej je z wykresem słupkowym. Postanowiliśmy zadać trzy pytania dotyczące:

1. Porównania liczności różnych klas.
2. Porównania stosunku różnych klas do siebie.

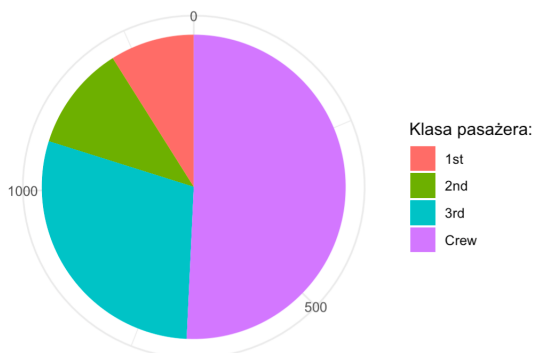
Poniżej przedstawiamy wyniki przeprowadzonej analizy dla zbioru danych *Titanic*, w szczególności liczby pasażerów, którzy przetrwali lub nie, tę katastrofę, z podziałem na klasy pasażerów i załogę.

2 Kółka i prostokąty

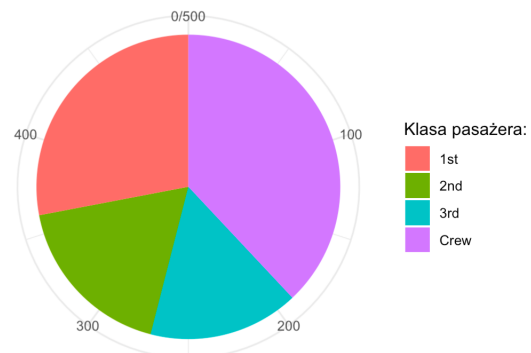
2.1 Liczność - którego więcej?

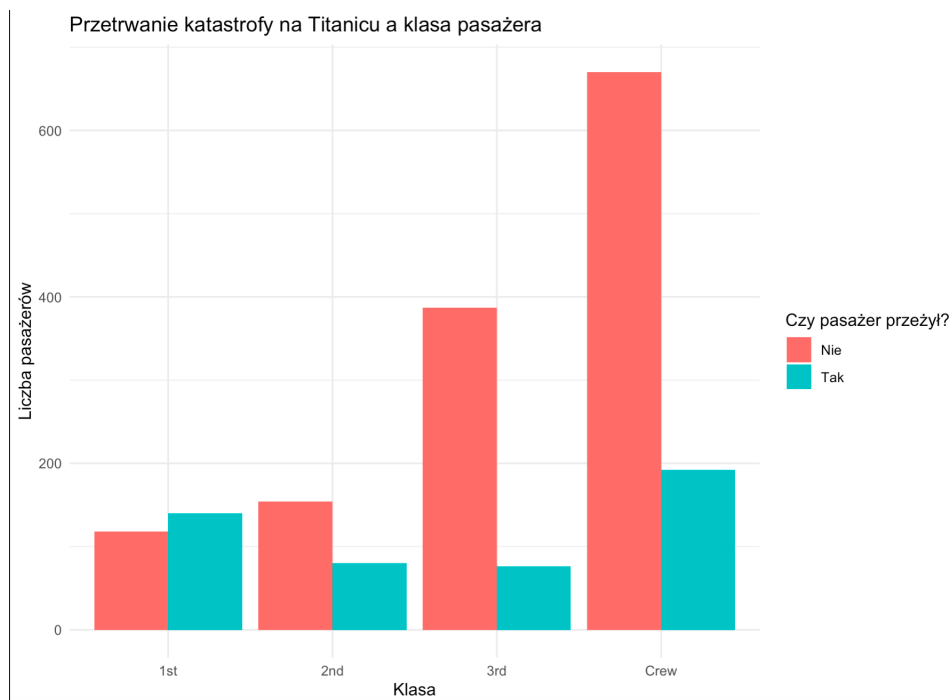
Rozważymy najpierw pierwsze dwa z zadanych przez nas pytań, które polegały na wskazaniu poprzez analizę wykresu, która klasa jest bardziej liczna.

Przetrwanie katastrofy na Titanicu a klasa pasażera
W przypadku, gdy pasażer nie przeżył:

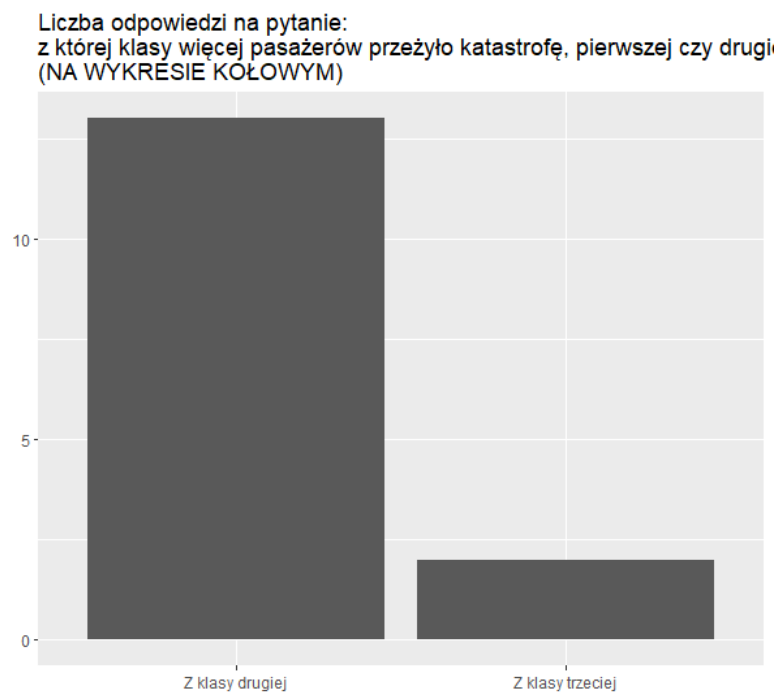


W przypadku, gdy pasażer przeżył:

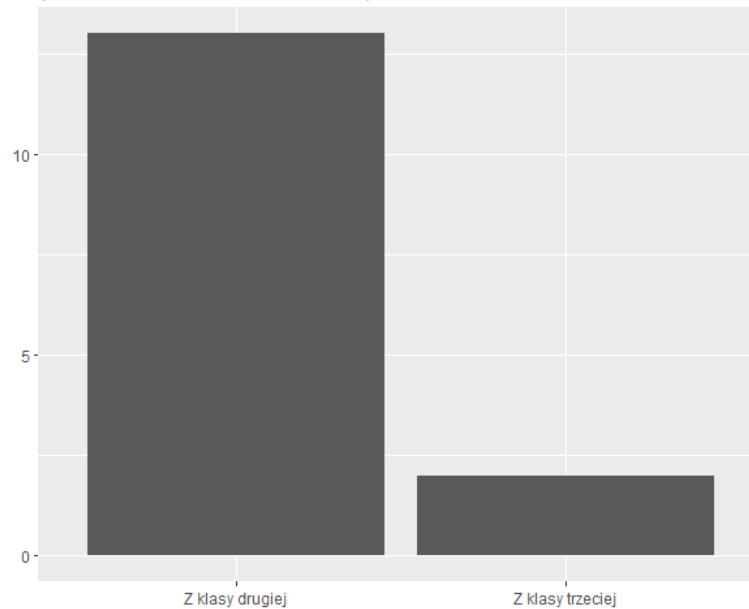




Interesowała nas umiejętność ankietowanych w porównaniu liczności klasy drugiej i trzeciej oraz to, czy "łatwiej" było to zauważyć na wykresie słupkowym.



Liczba odpowiedzi na pytanie:
z której klasy więcej pasażerów przeżyło katastrofę, pierwszej czy drugiej
(NA WYKRESIE SŁUPKOWYM)

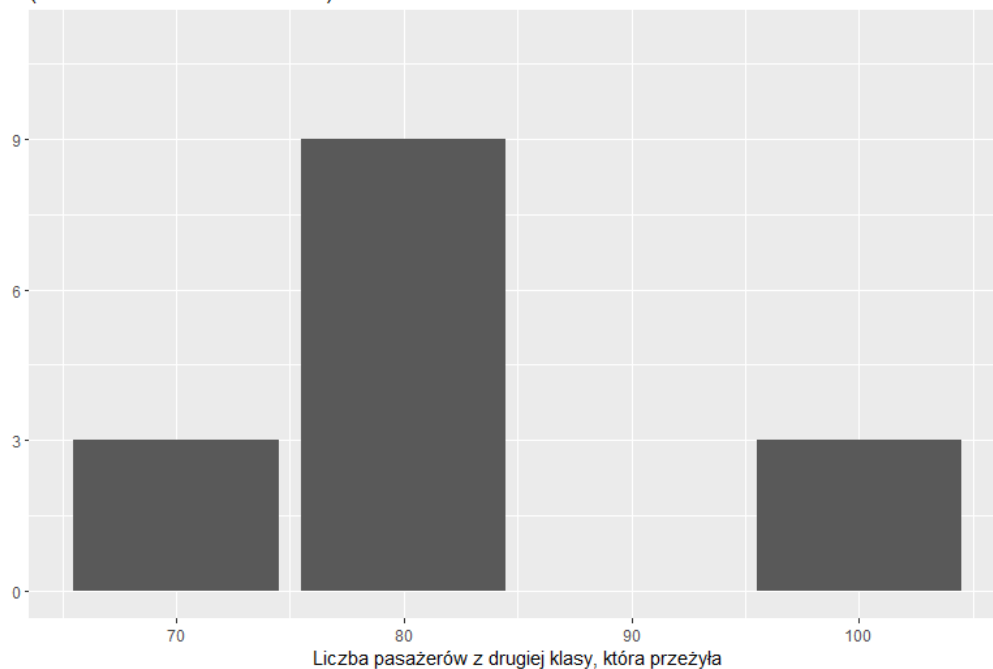


Możemy zauważyć, iż ankietowani poradzili sobie podobnie w odczytywaniu danych z obu rodzajów wykresów. Ciekawym eksperymentem byłoby zmierzenie faktycznego czasu potrzebnego na odczytanie danych z wykresu. Spodziewamy się, że czas na dokładne odczytanie danych z wykresu kołowego byłby dłuższy.

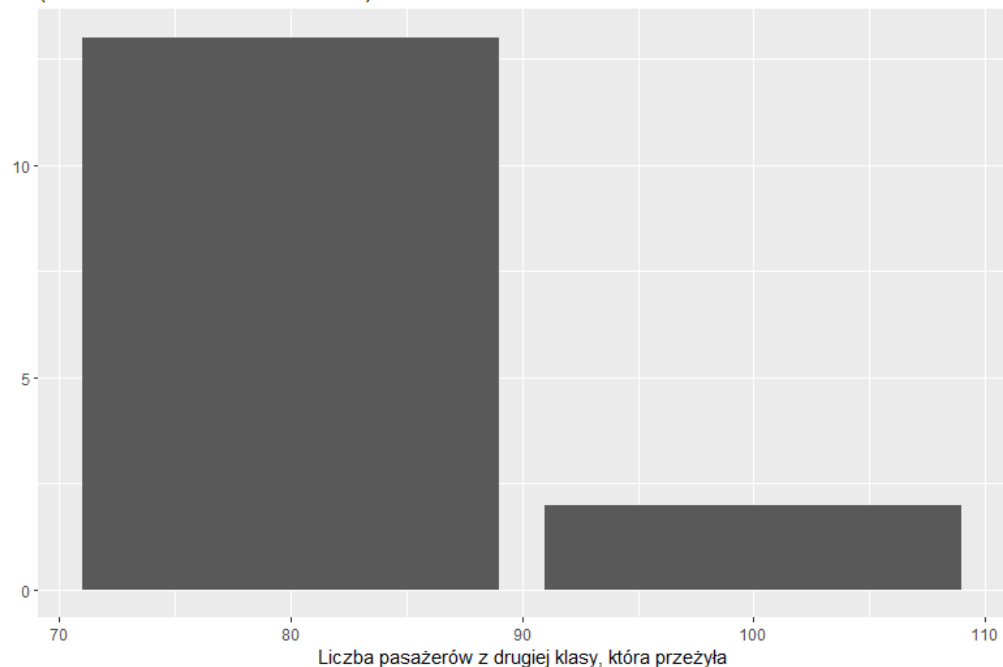
2.2 Liczność - dokładna wartość

Dalej zadaliśmy ankietowanym pytanie (dotyczące tych samych wykresów) o liczbę pasażerów z klasy drugiej, którzy przetrwali katastrofę, przy czym do wyboru były odpowiedzi: 50, 70, 80 i 100 pasażerów.

Liczba odpowiedzi na pytanie:
 ilu pasażerów z drugiej klasy przeżyło katastrofę?
 (NA WYKRESIE KOŁOWYM)



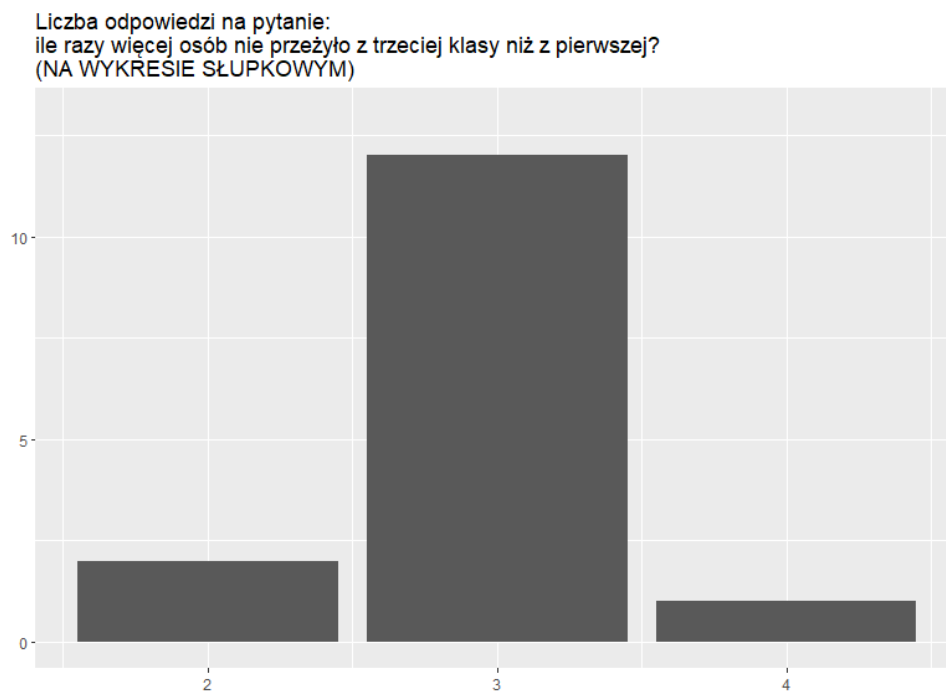
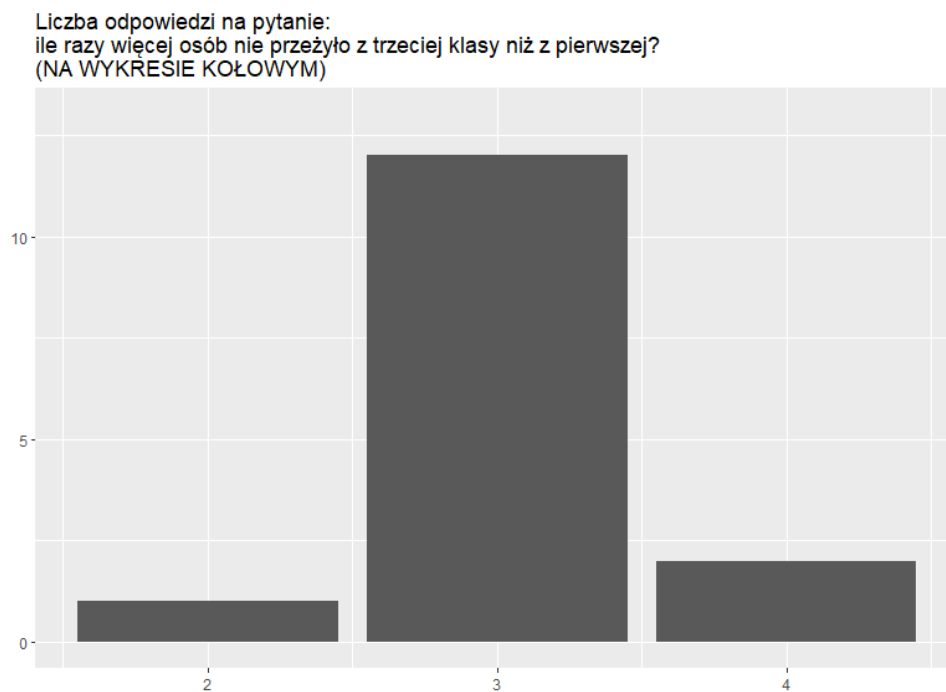
Liczba odpowiedzi na pytanie:
 ilu pasażerów z drugiej klasy przeżyło katastrofę?
 (NA WYKRESIE SŁUPKOWYM)



Widzimy, że większość odpowiedzi była poprawna (80 pasażerów). Jednak zarówno część poprawnych odpowiedzi jak i ich rozrzut były "lepsze" dla wykresu słupkowego.

2.3 Proporcje

Na koniec postanowiliśmy sprawdzić, czy na wykresie kołowym możemy porównywać iloczyny różnych klas. Zbadaliśmy poprawność odpowiedzi na pytanie: ile razy więcej pasażerów nie przeżyło z trzeciej klasy niż z pierwszej. Możliwymi odpowiedziami były: 2, 3 lub 4 razy (więcej). Wyniki dla obu typów wykresów prezentują się następująco:



Co ciekawe przy tym pytaniu ankietowani poradzili sobie równie dobrze w przypadku obu wykresów.

3 Wnioski i podsumowanie

Jest rzeczą bardzo trudną, wręcz niegodną aspirującego *Data Scientista* wyciąganie pochopnych wniosków na podstawie krótkiej ankiety z niewieloma odpowiedziami. Niemniej z ogólnej wiedzy, delikatnie podpartej powyższymi rozważaniami, iż wykresy kołowe, pomimo oczywistych walorów estetycznych, nie stanowią najlepszego narzędzia do wizualizacji danych.

Dziękujemy za uwagę!