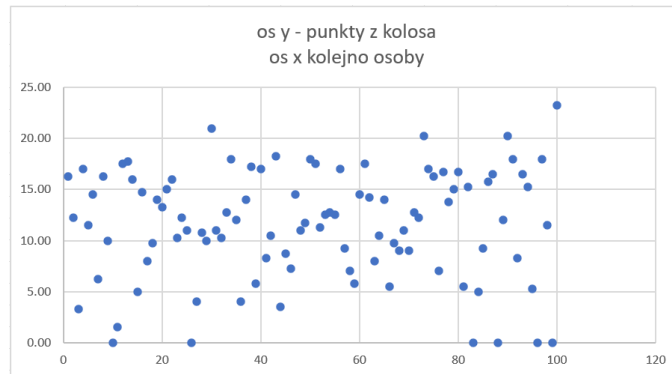


Praca domowa 3

Wstęp do eksploracji danych 2024

Bartosz Pokora

1. Wizualizacja źródłowa



Źródło: grupka na Messengerze kierunku MAD (naszego rocznika)

Data publikacji: 11/11/2023 (dostałem pozwolenie na starszy)

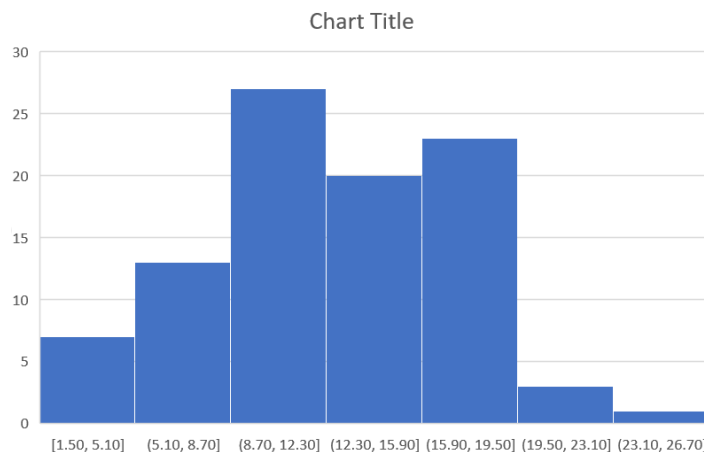
Wykres został stworzony przez dobrego kolegę z kierunku. Miał na celu wizualizację wyników pierwszego kolokwium z AM3, jako że zostały one przekazane nam jako lista w pdf, zamiast być normalnie zamieszczone na USOS-ie.

2. Błędy w wizualizacji

Niestety, nie mieliśmy jeszcze wtedy cudownych zajęć Wstęp do Eksploracji Danych™, więc zostało popełnione parę poważnych błędów:

- Źle dobrany typ wykresu do typu danych.
Wykres miał na celu zobrazować rozkład wyników, czyli jednej zmiennej ilościowej. Zamiast tego został stworzony wykres punktowy, który obrazuje zależność dwóch zmiennych ilościowych (w tym przypadku ilości zdobytych punktów oraz miejsca na liście).
- Osie podpisane (niezrozumiałe) w miejscu na tytuł wykresu oraz brak tytułu wykresu.

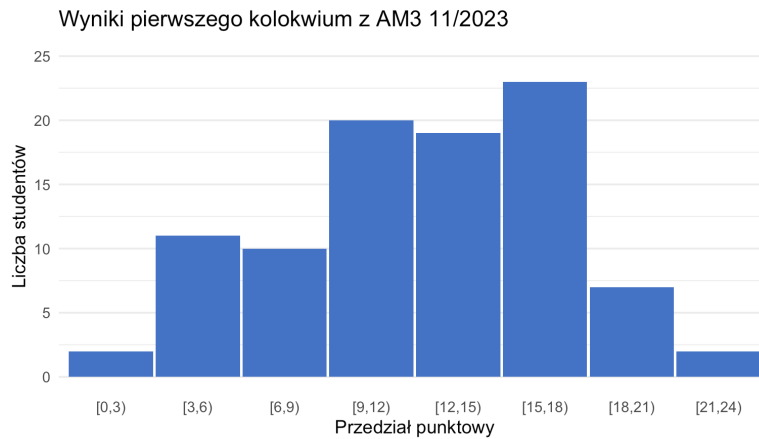
Niedługo później nadesłany został kolejny wykres, tym razem histogram, a więc już dobrze dobranego typu.



Cały czas jednak jest w nim parę rzeczy, które warto by poprawić:

- Brak tytułu („Chart Title”).
- Niepodpisane osie.
- Dziwacznie dobrane przedziały na koszyki, nieporęczne w odczycie.
(np.: wycięte wyniki poniżej 1.5 pkt; maksymalna liczba punktów do uzyskania wynosiła 24, a ostatni przedział zawiera jeszcze liczby do 26.7)

3. Poprawiona wizualizacja



4. Kod do wygenerowania wykresu

```
wyniki %>%  
  mutate(grupa = cut(punkty, breaks = c(seq(0, 30, 3)), right = FALSE)) %>%  
  group_by(grupa) %>%  
  summarise(n = n()) %>%  
  ggplot(aes(x = grupa, y = n)) +  
  geom_col(fill = "#4472C4", width = 0.97) +  
  labs(title = "Wyniki pierwszego kolokwium z AM3 11/2023", x = "Przedział punktowy", y = "Liczba studentów") +  
  theme_minimal() +  
  theme(panel.grid.minor.x = element_blank(), panel.grid.major.x = element_blank()) +  
  ylim(0, 25)
```