

Rafał Roter

## Praca Domowa 2

### Spis treści

<b>1. Wprowadzenie</b>	2
1.1. Treść zadania	2
1.2. Poprawiana wizualizacja	2
<b>2. Rozwiązanie zadania</b>	3
2.1. Kod	3
2.2. Wizualizacja	4
2.3. Uzasadnienie	5

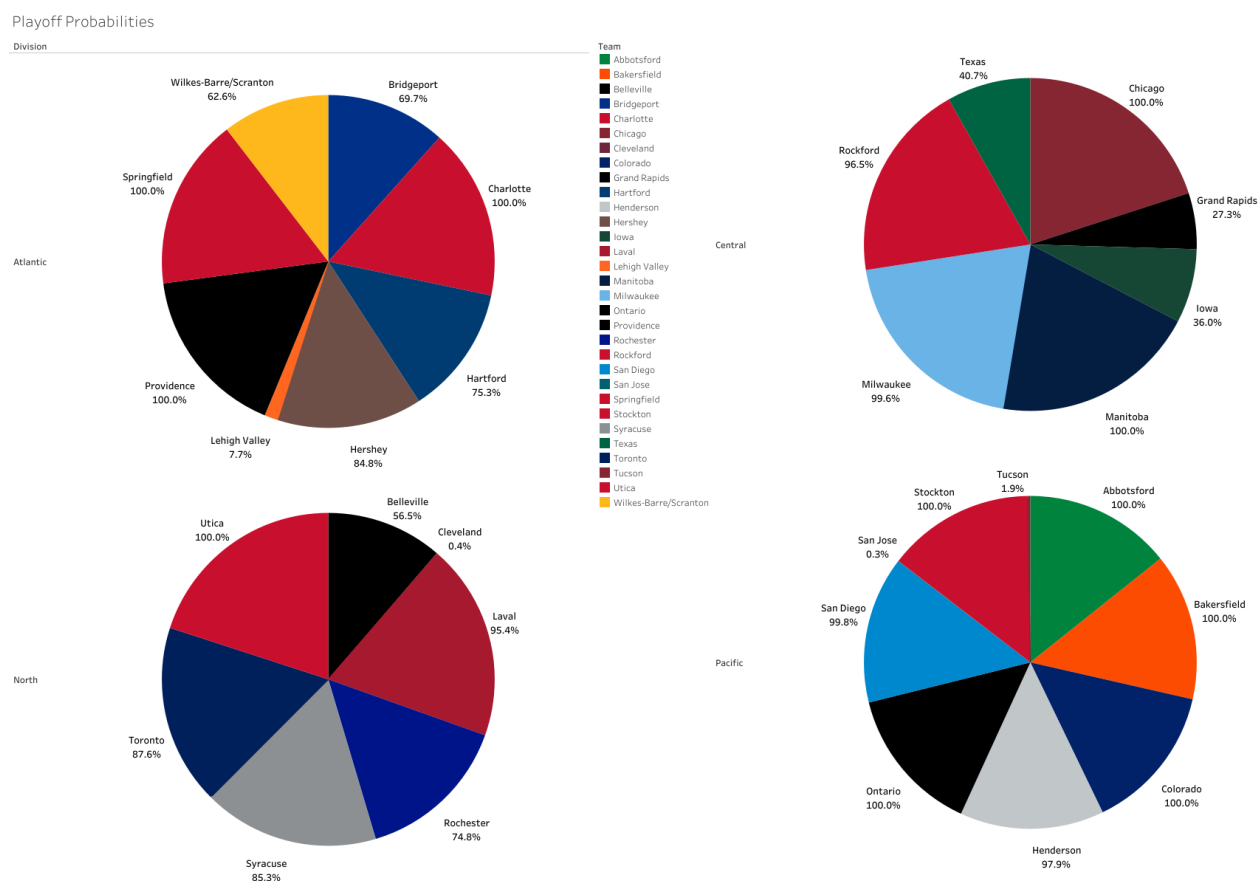
# 1. Wprowadzenie

## 1.1. Treść zadania

W ramach pracy domowej należało przygotować wizualizację z wykorzystaniem pakietu ggplot2. Wizualizacja ta miała zostać wykonana na podstawie innej, niepoprawnej wizualizacji dostępnej w prasie, telewizji lub internecie opublikowanej w ciągu ostatnich 2 tygodni.

## 1.2. Poprawiana wizualizacja

Danymi do poprawienia, które wybrałem były prawdopodobieństwa dostania się do play-offów zespołów z AHL w sezonie 2021-22 na dzień 3.04.2022. Wizualizacja tych danych, na której podstawie utworzyłem ramkę danych znajduje się na tej stronie.



Rysunek 1. Poprawiana wizualizacja

## 2. Rozwiązanie zadania

### 2.1. Kod

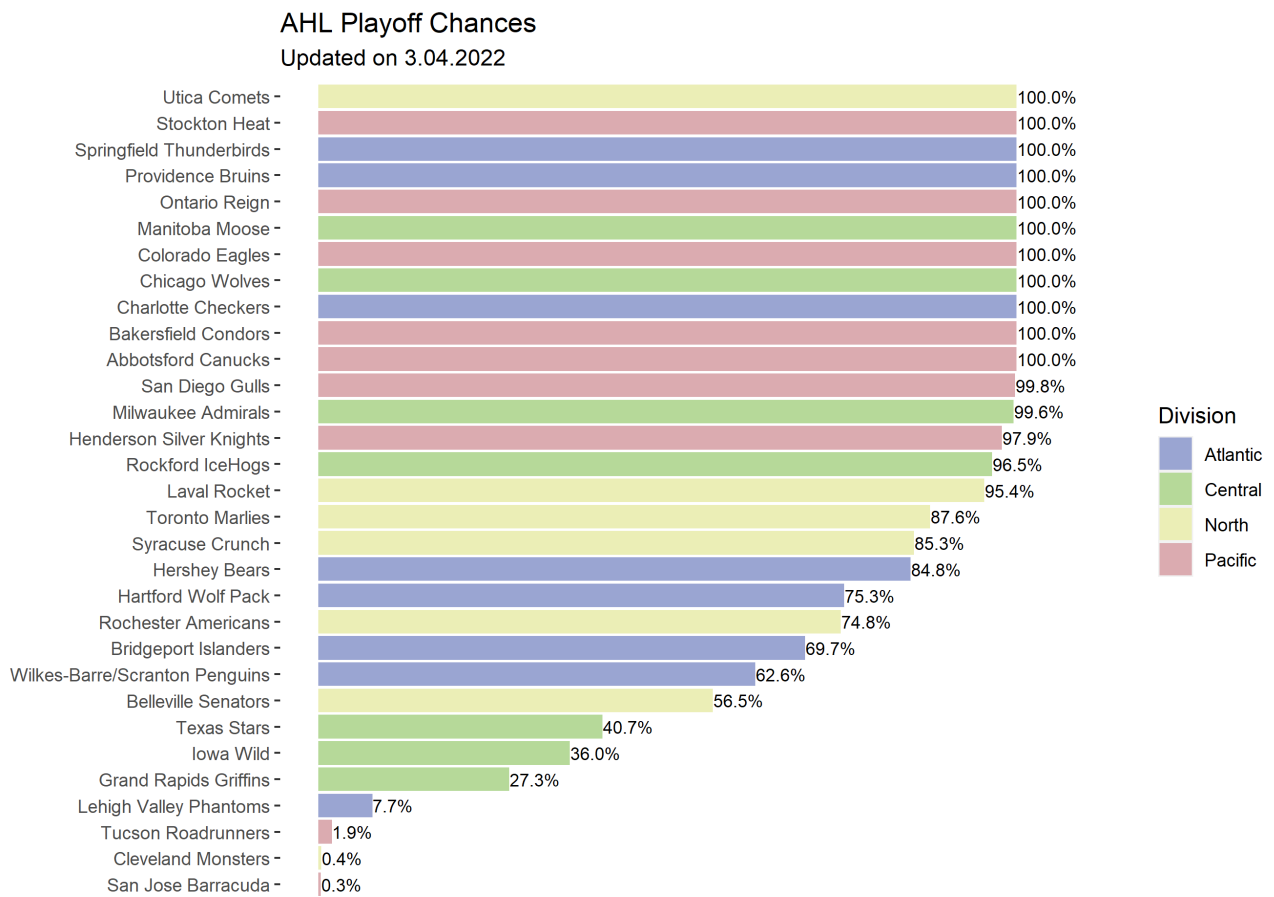
```
Team_Name <- c('Bridgeport_Islanders', 'Charlotte_Checkers',
               'Hartford_Wolf_Pack', 'Hershey_Bears',
               'Lehigh_Valley_Phantoms', 'Springfield_Thunderbirds',
               'Providence_Bruins', 'Wilkes-Barre/Scranton_Penguins',
               'Belleville_Senators', 'Cleveland_Monsters', 'Laval_Rocket',
               'Rochester_Americans', 'Syracuse_Crunch', 'Toronto_Marlies',
               'Utica_Cometes', 'Chicago_Wolves', 'Grand_Rapids_Griffins',
               'Iowa_Wild', 'Manitoba_Moose', 'Milwaukee_Admirals',
               'Rockford_IceHogs', 'Texas_Stars', 'Abbotsford_Canucks',
               'Bakersfield_Condors', 'Colorado_Eagles',
               'Henderson_Silver_Knights', 'Ontario_Reign', 'San_Diego_Gulls',
               'San_Jose_Barracuda', 'Stockton_Heat', 'Tucson_Roadrunners')
Division <- c('Atlantic', 'Atlantic', 'Atlantic', 'Atlantic', 'Atlantic',
              'Atlantic', 'Atlantic', 'Atlantic', 'North', 'North', 'North',
              'North', 'North', 'North', 'North', 'Central', 'Central',
              'Central', 'Central', 'Central', 'Central', 'Central',
              'Pacific', 'Pacific', 'Pacific', 'Pacific', 'Pacific',
              'Pacific', 'Pacific', 'Pacific', 'Pacific')
Probability <- c(.697, 1.00, .753, .848, .077, 1.00, 1.00, .626, .565,
                 .004, .954, .748, .853, .876, 1.00, 1.00, .273, .36,
                 1.00, .996, .965, .407, 1.00, 1.00, 1.00, .979, 1.00,
                 .998, .003, 1.00, .019)

playoffs <- data.frame(Team_Name, Division, Probability)

playoff_plot <- ggplot(playoffs, aes(x = reorder(Team_Name, Probability),
                                         y = Probability, fill = Division)) +
  geom_col() +
  coord_flip() +
  scale_fill_manual(values = c('#9aa5d2', '#b6d999', '#ebee6', '#dbabb0')) +
  geom_text(aes(label = scales::percent(Probability, accuracy = .1),
                size = 3, hjust = -0.01)) +
  labs(title = 'AHL_Playoff_Chances',
        subtitle = 'Updated_on_3.04.2022',
        ) +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        axis.title.y = element_blank(),
        panel.background = element_blank()) +
  expand_limits(y = 1.1)

ggsave('playoff_plot.png')
```

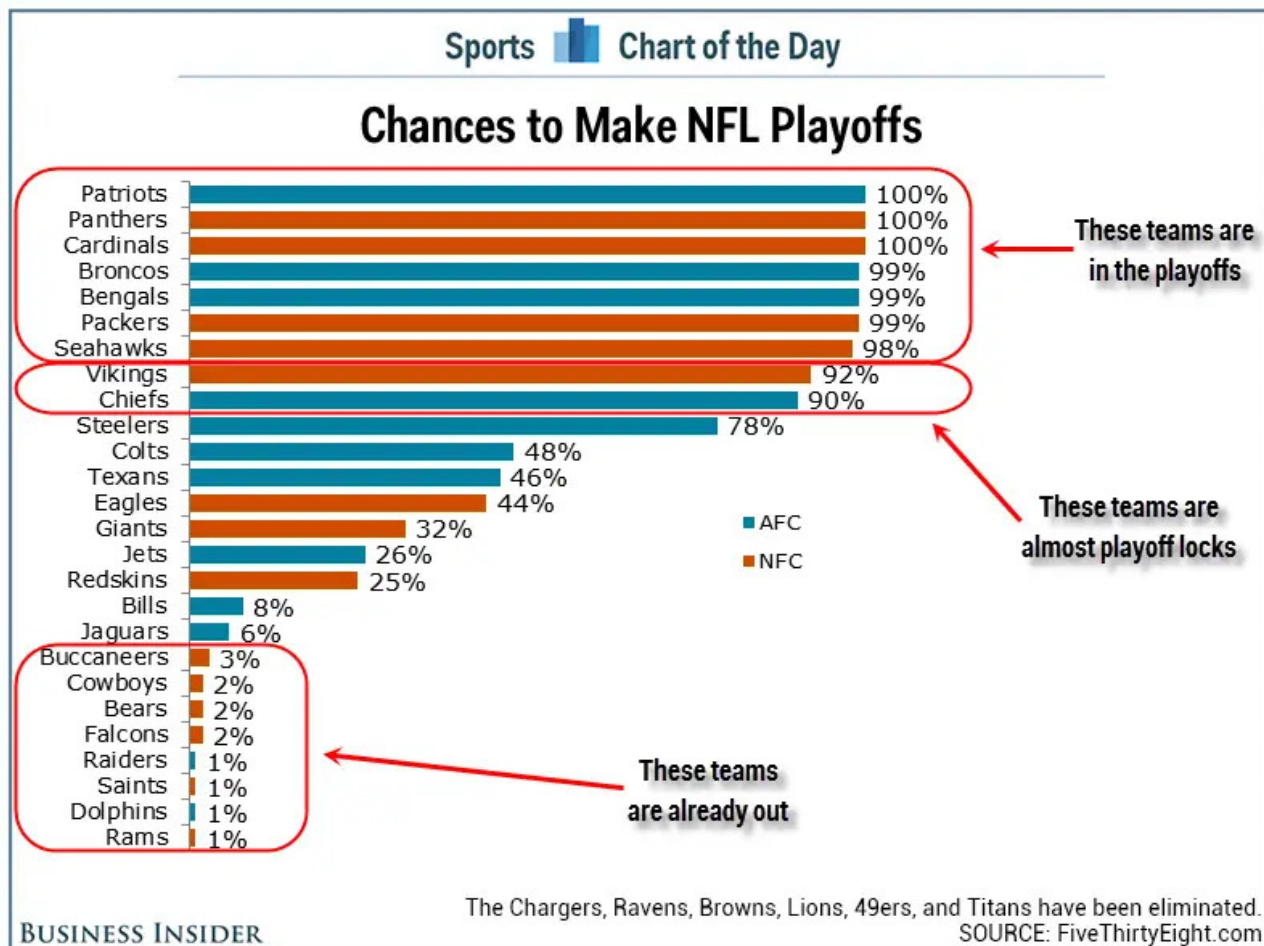
## 2.2. Wizualizacja



Rysunek 2. Poprawiona wizualizacja

### 2.3. Uzasadnienie

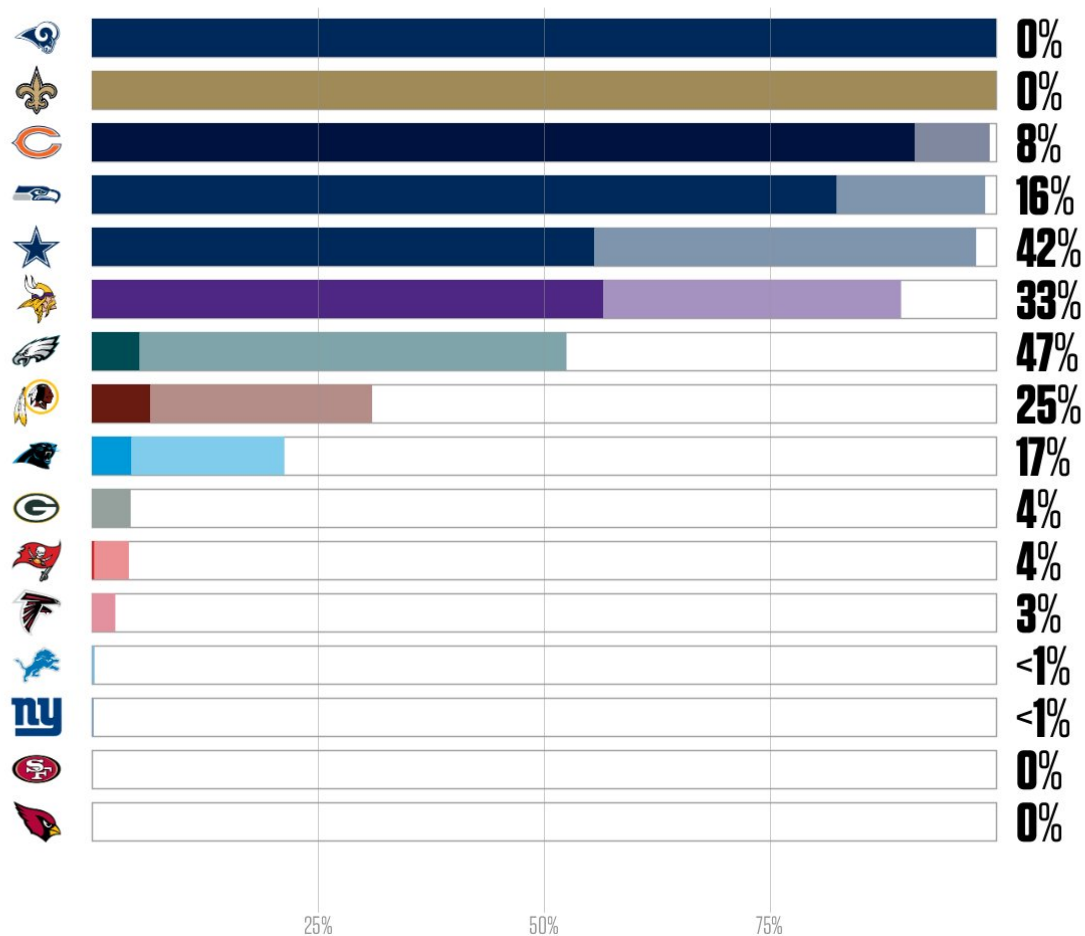
Wykres kołowy, szczególnie taki na którym prezentujemy więcej niż 100% jest zdecydowanie złym sposobem na wizualizację danych. Wykres zaprezentowany przeze mnie wydaje mi się być zdecydowanie bardziej przejrzysty i umożliwia przedstawienie tych danych na 1 wykresie zamiast na 4 oddzielnych wykresach. Format wykresu bazowałem na wizualizacjach tego typu danych przygotowywanych dla Business Insider oraz ESPN:



Rysunek 3. Przykład wykresu przygotowanego dla Business Insider

# Playoff Probability Leverage

NFC Week 14



Rysunek 4. Przykład wykresu przygotowanego dla ESPN