

Wstęp do eksploracji danych. Praca domowa 4

Aleksandra Uznanska

April 2024

1 Konkurencja między wybranymi gatunkami zwierząt mięsożernych o pokarm

Ramka danych, którą się zainteresowałam zawierała wiele informacji na temat wybranych zwierząt. Skupiłam się na kolumnie "Predators" opisującą jakie gatunki polują na zwierzę opisywane w danym wierszu oraz przekształciłam tak moje dane aby posiadać informację, które zwierzęta konkurują ze sobą o pokarm co w tym przypadku oznaczało.

2 Ramka danych

Zbiór danych obejmuje różnorodną gamę atrybutów odnoszących się do różnych gatunków zwierząt na całym świecie. Zawiera kolumny takie jak zwierzę, wzrost (cm), waga (kg), kolor, długość życia (lata), dieta, siedlisko, drapieżniki, średnia prędkość (km/h), znalezione kraje, status ochrony, rodzina, okres ciąży (dni), maksymalna prędkość (km/h), struktura społeczna i potomstwo na narodziny. Kolumny te pozwalają bardzo szeroką analizę.

3 Możliwe wnioski wyciągnięte z grafu

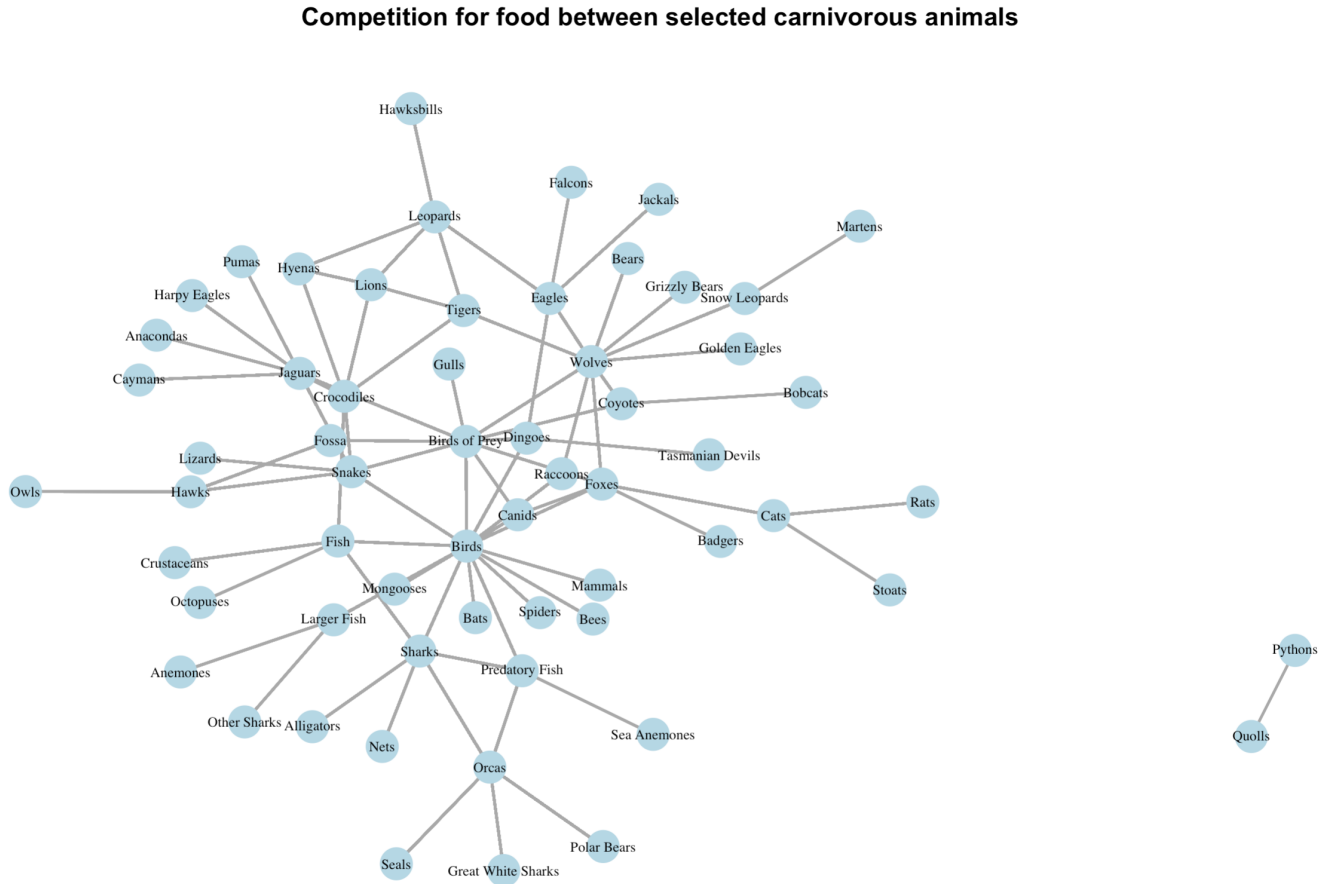
Najważniejszym wnioskiem z mojej pracy była możliwość łatwego zaobserwowania tego, że osobniki żyjące w tym samym środowisku znacznie częściej konkurują ze sobą o pokarm (np. Wilki, niedźwiedzie oraz lisy). Widać też jak środowiska te na siebie w pewien sposób "nachodzą". (np. Orki konkurują z rekinami, te zaś z ptakami, a te ostatnie z nietoperzami oraz pajakami). Pojawiło się także kilka połączeń, które faktycznie mnie zaskoczyły. Jednymi z nich były:

Alligators - Sharks

Pythons - Quolls

Rats - Cats

4 Finalny Efekt



In the above graph, two vertices have been connected by an edge if they have at least one food in common.

Rysunek 1: Enter Caption

5 Kod w R studio

```

1 library(igraph)
2 library(tidygraph)
3 library(ggraph)

```

```
4 library(dplyr)
5 #install.packages("tidyr")
6 library(tidyr)
7
8 df <- read.csv("Animal_dataset.csv")
9 df <- na.omit(df)
10
11 split_predators <- function(predators) {
12   return(unlist(strsplit(as.character(predators), ", ")))
13 }
14 df$Predators <- lapply(df$Predators, split_predators)
15
16 predators_df <- df[,c("Predators","Animal")]
17 head(predators_df,30)
18
19 predators_df <- predators_df %>%
20   unnest(Predators)
21
22 predator_pairs <- expand.grid(Predator1 = predators_df$Predators,
23   Predator2 = predators_df$Predators)
24 predator_pairs <- predator_pairs %>%
25   filter(Predator1 != Predator2)
26
27 predator_pairs <- predator_pairs %>%
28   inner_join(predators_df, by = c("Predator1" = "Predators")) %>%
29   inner_join(predators_df, by = c("Predator2" = "Predators")) %>%
30   filter(Animal.x == Animal.y) %>%
31   select(-Animal.x, -Animal.y)
32
33 predators_df_filtered <- predator_pairs %>%
34   filter(Predator1 != "Humans" & Predator1 != "Not Applicable" &
35     Predator2 != "Humans" & Predator2 != "Not Applicable" )
36 predators_df_filtered
37
38 graph <- predators_df_filtered %>%
39   graph_from_data_frame(directed = FALSE)
40 ?plot.igraph
41 plot.igraph(graph,
42   vertex.size = 5,
43   vertex.color = 'lightblue',
44   vertex.frame.color = 'lightblue',
45   vertex.frame.width = 1,
46   vertex.shape = 'circle',
47   vertex.label.cex = 0.67,
48   width = 0.8,
49   vertex.label.color = 'black',
```

```
48         vertex.label.dist = 0,
49         edge.color = 'darkgrey',
50         edge.width = 2,
51         edge.lty = 1,
52         edge.curved = FALSE,
53         layout = layout_with_kk,    #layout_with_kk
54         margin = 0,
55         rescale = TRUE,
56         asp = 0.6,
57         frame = FALSE,
58         main = 'Competition for food between selected
              carnivorous animals',
59         edge.arrow.size=0.5,
60         vertex.size=5,
61         vertex.label.cex=0.8)
62 ?plot.igraph
63 title(sub = "In the above graph, two vertices have been connected by
              an edge if they have at least one food in common. ")
```

6 Skąd wzięłam dane ?

<https://www.kaggle.com/datasets/iamsouravbanerjee/animal-information-dataset/data>