Untitled

Bottino Manuel & Giuseppe Specio

Today's Date

In our case the weight would be -> weighted, with the weights being the half-weight index (HWI)=the number of times dolphins A and B were seen together divided by the total number of times they were seen together plus half the value of when A was seen without B and B was seen without A, and range from 0 for individuals that were never sighted together in groups to 1 for individuals that were always sighted together

Introduction.

Our network analysis project is based on the research article "The importance of delineating networks by activity type in bottlenose dolphins in Cedar Key, Florida" (Gazda et al, 2015). The goal of this paper was to demonstrate the existence of a fission-fusion structure within these animals.

Fission-Fusion social structure.

The fission-fusion social structure is a pattern where the size and composition of animal groups change frequently and rapidly. Individuals split (fission) and merge (fusion) into groups of varying sizes and compositions based on factors like resource availability, predator presence, and social or reproductive needs. In this case, three main behavioral contexts were analysed: socializing, travelling, foraging. The null hypothesis is that regardless of activity state, the corresponding networks will be similar to each other and to the overall network that does not take activity into account.

Paper's Key findings.

The study's findings reveal that dolphins socialize in large groups with preferential associations, travel in smaller groups with selective partners, and forage in very small, weakly connected groups. There is some overlap between socializing and traveling networks but little overlap with foraging networks, indicating that social bonds are less important during foraging.

Data collection.

Data were collected from 147 resident bottlenose dolphins in Cedar Key, Florida, over two periods in 2008 and 2010. Dolphins were observed for socializing, traveling, and foraging behaviors. Individual dolphins were identified using photographs of their dorsal fins. Group compositions were recorded, excluding sightings of the same group within an hour or on the same day to avoid non-independent sampling.

Network Construction

In this context, the vertices in the network represent the individual dolphins, and the edges represent associations between the corresponding dolphins.

Four main networks were used: an overall network that does not take behavior into account, and the socialize network, the travel network and the forage network that correspond to their respective behaviors. In addition, a comprehensive network was created as well, combining all the other four together.

Each network was simple, undirected, and weighted, with edge weights calculated using the half-weight index (HWI): the number of times dolphins A and B were seen together divided by the total number of times they were seen together plus half the value of when A was seen without B and B was seen without A, and range from 0 for individuals that were never sighted together in groups to 1 for individuals that were always sighted together. This method gives different weights to associations based on their strength, helping to filter out weak associations and reducing sampling bias.

Networks were pruned by removing dolphins sighted fewer than three times to ensure significant patterns of association.

Project Objective

Our project aims to visualize the findings from the paper using the most appropriate graphs. We will employ network analysis tools to graphically represent the different behavioral networks (socializing, traveling, foraging) of the Cedar Key dolphins.

Additionally, we will propose new perspectives beyond those presented in the paper and graphically represent these insights as well. QUESTO INVECE QUANDO FAI summary(overall_dolphins_label) * Summary explanation: + V1 and V2: there are 291 vertices. The 147 (x2 = 294) dolphins of interest were recorded in two distinct periods, and were recorded in the dataset as two different dolphins. The 3 remaining points were caused by the absence of at least three sightings of 3 dolphins. + Type: categorical variable explaining the type of activity (Social, Travel, Forage) the two dolphins were performing at the moment of the sighting. The fourth type, "O", refers to an edge where the activity that was being performed was not of interest at the moment of the sighting. It allows to test the null hypothesis: there is no difference in the networks for specific activities and the overall network, where activity was not of interest. + Sightings: amount of times that two dolphins were seen together. It helps assessing the weight of the edge.

```r
library(graphlayouts)
library(ggforce)
```

```
## Loading required package: ggplot2
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
# install.packages("extrafont",repos = "http://cran.us.r-project.org")
library(ggplot2)
library(igraph)
```

```
##
## Attaching package: 'igraph'

## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union

## The following objects are masked from 'package:stats':
##
##     decompose, spectrum

## The following object is masked from 'package:base':
##
##     union
```

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.3     v tibble    3.2.1
## v purrr     1.0.2     v tidyr     1.3.0
## v readr     2.1.5

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x lubridate::%--%()       masks igraph::%--%()
## x tibble::as_data_frame() masks igraph::as_data_frame(), dplyr::as_data_frame()
## x purrr::compose()        masks igraph::compose()
## x tidyr::crossing()       masks igraph::crossing()
## x dplyr::filter()         masks stats::filter()
## x dplyr::lag()            masks stats::lag()
## x purrr::simplify()       masks igraph::simplify()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

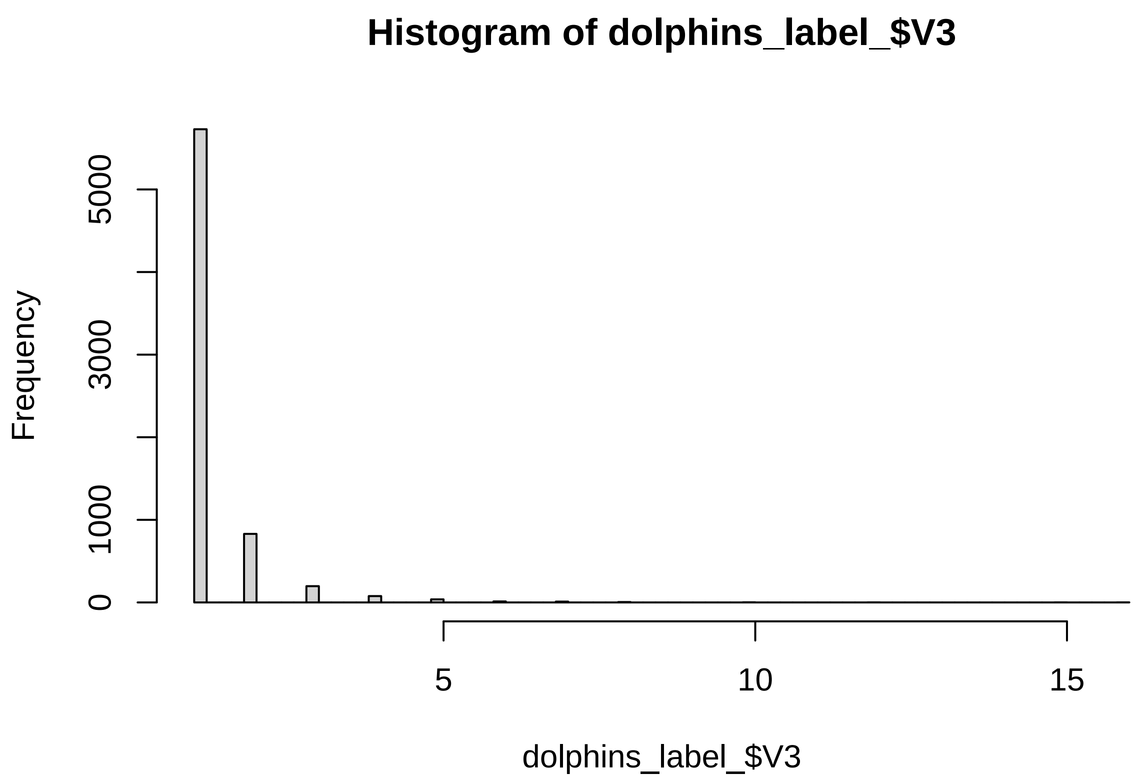
```r
library(ggraph)
library(ggrepel)
library(networkdata)
```

```
##
## Attaching package: 'networkdata'
##
## The following object is masked from 'package:dplyr':
##
##     starwars
```

```r
library(knitr)
library(extrafont)
```

```
## Registering fonts with R
```

```r
# Import the data
# Overall dolphin social network
overall_dolphins <- read.table("Data/Overall/mammalia-dolphin-florida-overall.edges")
# Forage dolphins social network
forage_dolphins <- read.table("Data/Forage/mammalia-dolphin-florida-forage.edges")
# Social dolphins social network
social_dolphins <- read.table("Data/Social/mammalia-dolphin-florida-social.edges")
# Travel dolphins social network
travel_dolphins <- read.table("Data/Travel/mammalia-dolphin-floridatravel.edges")

# Create a new dataset where we can identify all of the different relations among delphins
forage_dolphins$type <- "F"
social_dolphins$type <- "S"
travel_dolphins$type <- "T"
overall_dolphins$type <- "O"
dolphins_label_ <- rbind(overall_dolphins, forage_dolphins, social_dolphins, travel_dolphins)
dolphins_label_$type <- as.factor(dolphins_label_$type)
hist(dolphins_label_$V3, breaks = 100)
```




```r
dolphins_label_ = mutate(dolphins_label_, sightings = case_when(
    V3 == 1 ~ "1",
    V3 == 2 ~ "2",
```

```
    V3 > 2 & V3 <= 4 ~ "3-4",
    V3 > 4 ~ ">4",
))
dolphins_label = dolphins_label_[, c(1, 2, 3, 4, 5)]
head(dolphins_label)
```

```
##   V1  V2 V3 type sightings
## 1  1 133  1    0         1
## 2  1 265  1    0         1
## 3  1  10  1    0         1
## 4  1 267  1    0         1
## 5  1  12  1    0         1
## 6  1 142  1    0         1
```

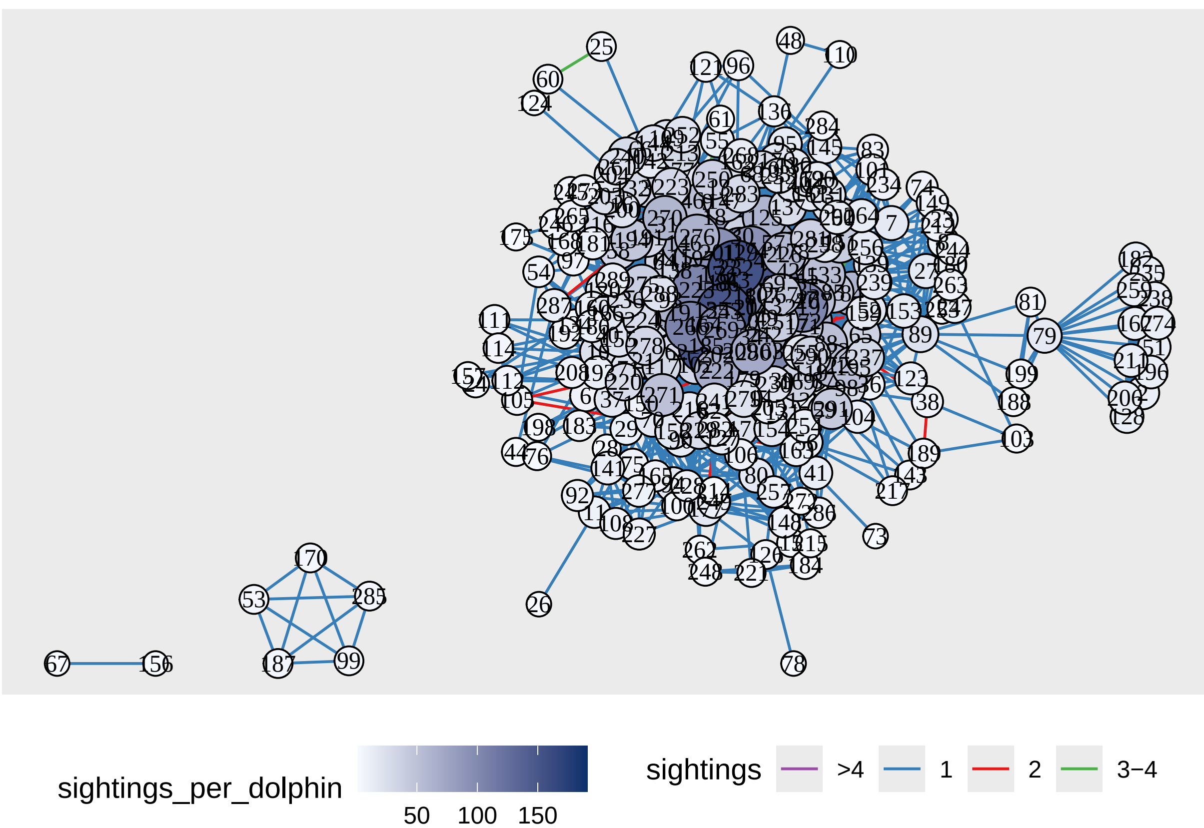
Data preprocessing and network creation

```
# Function to create a network whatever the dataset
net_dolphins = function(df){

  sightings_per_dolphin_1 = group_by(df, V1) %>% summarise(sightings = sum(V3))
  sightings_per_dolphin_2 = group_by(df, V2) %>% summarise(sightings = sum(V3))

  # Merge the datasets based on the "name" column
  merged_data <- merge(sightings_per_dolphin_1, sightings_per_dolphin_2, by.x = "V1", by.y = "V2", all =
  head(merged_data)
  # Replace missing values with 0
  merged_data$sightings.x[is.na(merged_data$sightings.x)] <- 0
  merged_data$sightings.y[is.na(merged_data$sightings.y)] <- 0

  # Create a new column with the sum of occurrences
  merged_data$sightings <- merged_data$sightings.x + merged_data$sightings.y
  head(merged_data)

  # Remove unnecessary columns
  merged_data <- select(merged_data, V1, sightings)

  # Create our network
  net_dolphins_label <- graph_from_data_frame(df, directed = FALSE) #from the first 250 rows of the com
  # Add the sightings column to the network
  names_net <- as.numeric(V(net_dolphins_label)$name)
  merged_data = merged_data[order(match(merged_data$V1, names_net)),]
  net_dolphins_label=set_vertex_attr(net_dolphins_label, "sightings_per_dolphin", index = V(net_dolphins
  return(net_dolphins_label)
}

dolphins_label_F = dolphins_label[dolphins_label$type == "F",]
dolphins_label_S = dolphins_label[dolphins_label$type == "S",]
dolphins_label_T = dolphins_label[dolphins_label$type == "T",]
dolphins_label_O = dolphins_label[dolphins_label$type == "O",]

net_forage = net_dolphins(dolphins_label_F)
net_social = net_dolphins(dolphins_label_S)
net_travel = net_dolphins(dolphins_label_T)
```

```
net_overall = net_dolphins(dolphins_label_0)
net = net_dolphins(dolphins_label)
```

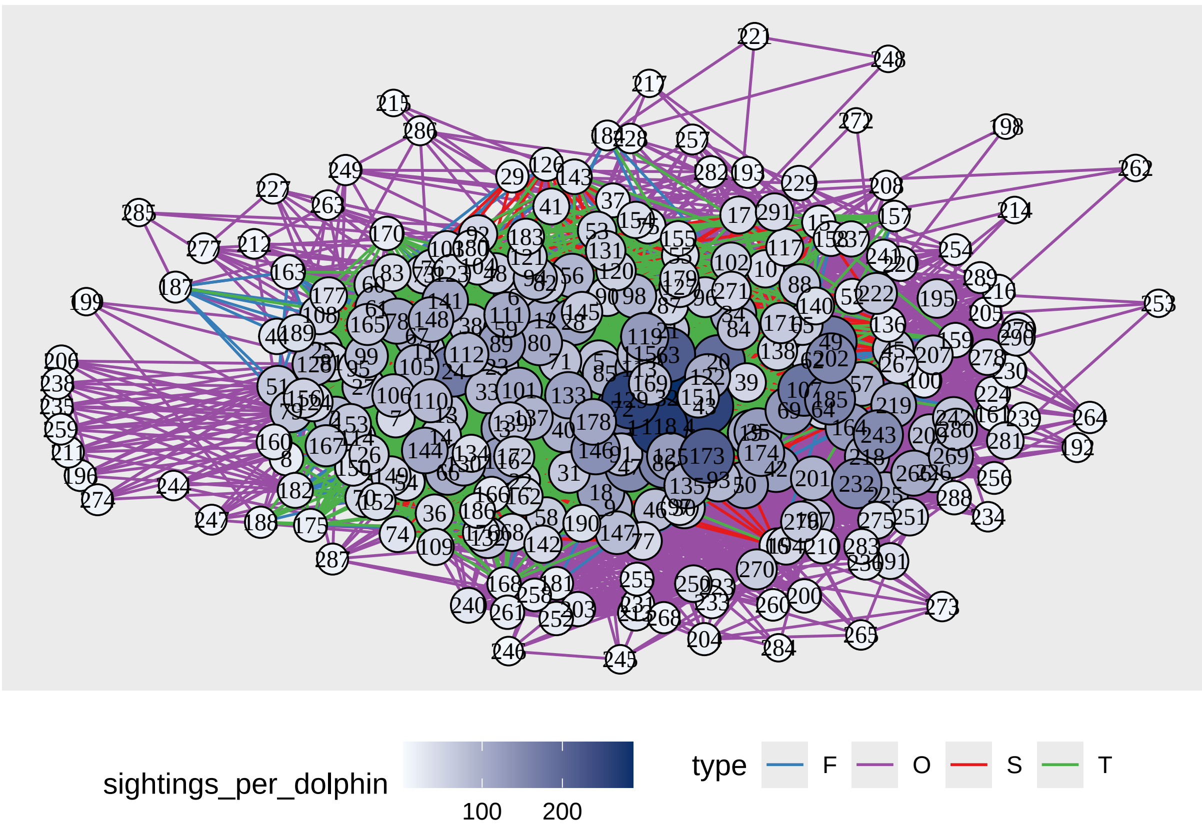
Network visualization

```
# Suggested colors for edge categories and node gradient
edge_colors <- c("1" = "#377eb8", "2" = "#e41a1c", "3-4" = "#4daf4a", ">4" = "#984ea3")
node_gradient <- c("#f7fbff", "#08306b")  # From light blue to dark blue
edge_colors_general <- c("F" = "#377eb8", "S" = "#e41a1c", "T" = "#4daf4a", "O" = "#984ea3")

# Plot the network
ggraph(net_overall, layout = "stress", bbox = 15) +
  geom_edge_link2(aes(edge_colour = sightings), edge_linewidth = 0.5) +
  scale_edge_colour_manual(values = edge_colors) +
  geom_node_point(aes(fill = sightings_per_dolphin, size = sightings_per_dolphin), shape=21) +
  scale_fill_gradient(low = node_gradient[1], high = node_gradient[2]) +
  geom_node_text(aes(label = name), vjust = 0.5, hjust = 0.5,
                 family = "serif", size = 3) +  # Adjust font size as needed
  scale_size(range = c(4, 10), guide = "none") +
  theme(legend.position = "bottom")
```



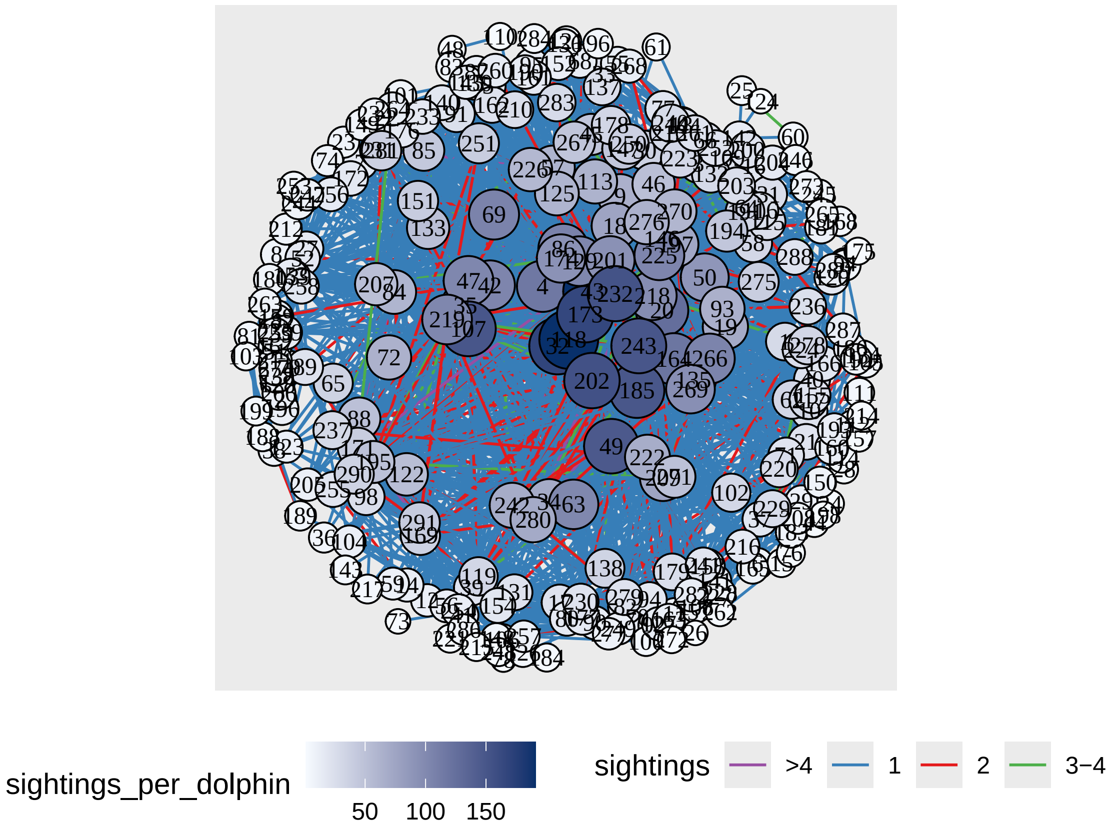

```
# general one (total confusion)
ggraph(net, layout = "stress", bbox = 15) +
```

```
  geom_edge_link2(aes(edge_colour = type), edge_linewidth = 0.5) +
  scale_edge_colour_manual(values = edge_colors_general) +
  geom_node_point(aes(fill = sightings_per_dolphin, size = sightings_per_dolphin), shape=21) +
  scale_fill_gradient(low = node_gradient[1], high = node_gradient[2]) +
  geom_node_text(aes(label = name), vjust = 0.5, hjust = 0.5,
                 family = "serif", size = 3) +  # Adjust font size as needed
  scale_size(range = c(4, 10), guide = "none") +
  theme(legend.position = "bottom")
```




```
comps <- components(net_overall)
net_O_restricted <- delete_vertices(net_overall, which(comps$membership == which.min(comps$csize)))

comps <- components(net_O_restricted)
net_O_restricted <- delete_vertices(net_O_restricted, which(comps$membership == which.min(comps$csize)))



ggraph(net_O_restricted, layout = "centrality", cent = graph.strength(net_O_restricted)) +
  geom_edge_link2(aes(edge_colour = sightings), edge_linewidth = 0.5) +
  scale_edge_colour_manual(values = edge_colors) +
  geom_node_point(aes(fill = sightings_per_dolphin, size = sightings_per_dolphin), shape=21) +
  scale_fill_gradient(low = node_gradient[1], high = node_gradient[2]) +
  geom_node_text(aes(label = name), vjust = 0.5, hjust = 0.5,
                 family = "serif", size = 3) +  # Adjust font size as needed
  scale_size(range = c(4, 10), guide = "none") +
  coord_fixed() +
  theme(legend.position = "bottom")
```

```
## Warning: ‘graph.strength()‘ was deprecated in igraph 2.0.0.
## i Please use ‘strength()‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



Research Findings + Graphical Representation

Here we report the research study findings, we explain what does it mean in terms of graph construction and then proceed to propose the appropraté graphs.

```r
knitr::include_graphics("Screenshot 2024-06-10 at 22.56.24.png")
```

		association indices		
		real	**random**	***p*-value**
overall network	mean	0.03312	0.03558	*0.00055*
	s.d.	0.08157	0.06967	0.99999
	CV	2.46307	1.95818	1
forage network	mean	0.04186	0.04181	0.53489
	s.d.	0.11154	0.09057	0.99999
	CV	2.66463	2.16634	0.99999
socialize network	mean	0.18217	0.18762	*0.03554*
	s.d.	0.20699	0.17423	1
	CV	1.13628	0.92884	0.99999
travel network	mean	0.0596	0.06184	*0.0033*
	s.d.	0.1317	0.11237	0.99999
	CV	2.20953	1.81724	1

#####

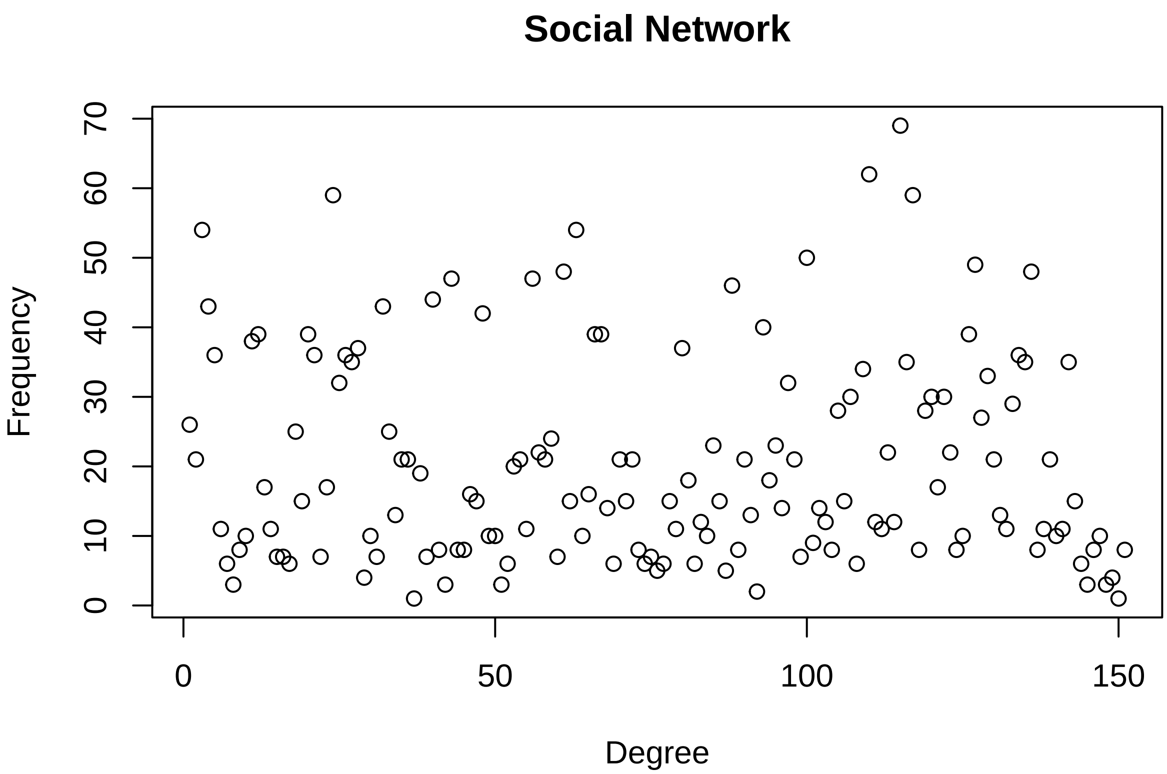
Aggiungi Table2 e Table3

- There are preferential associations between individuals in the overall network, the socialize network and the travel network, but not in the forage network (table 2). This is not an artefact of sample size: the number of sightings in the forage network (153) is greater than that in the travel network (77) and socialize network (38).
  - Node degree distribution plots helps to visualize preferential associations. Nodes with higher degrees have more connections, indicating preferential associations.

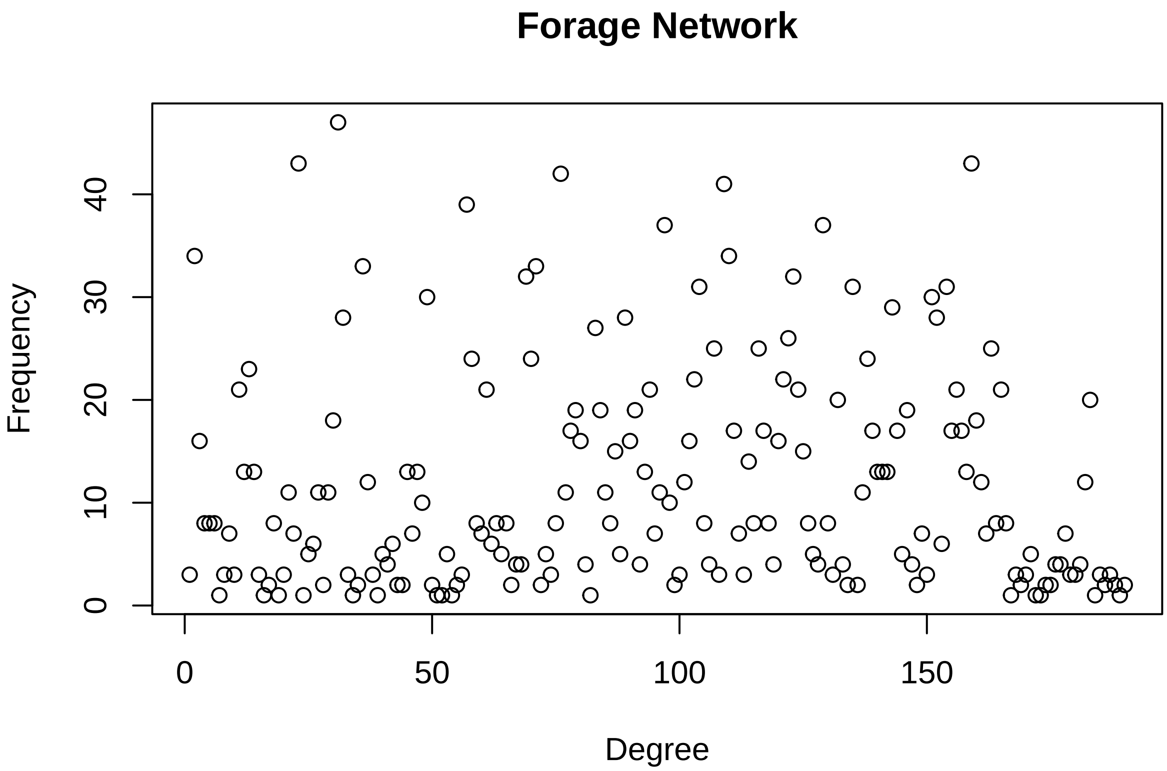
Social Network: Lower degree distribution in the social network indicates that dolphins in this network have fewer connections or associations compared to the forage network.

Forage Network: Higher degree distribution in the forage network suggests that dolphins in this network have more connections or associations compared to the social network.

```
plot(degree(net_social), main="Social Network", xlab="Degree", ylab="Frequency")
```




```
plot(degree(net_forage), main="Forage Network", xlab="Degree", ylab="Frequency")
```



Socialise network * Individuals have strong and repeated connections to many other individuals. + Highest average degree, Highest average strength, Highest average edge weight.

- Socializing happens in large groups and these groups are not exclusive.
  - Largest group size per sighting, Highest size per community, Least number of communities, Fewest connected components.
- Dolphins that are connected to a particular dolphin are more likely to be connected to one another.
  - Highest clustering coefficient.

Travel network * Dolphins in this network do not have strong and repeated connections to many others except their preferential associates + Lower average degree, Lower average strength, Lower average edge weight).

- Travelling happens in smaller groups than socializing.
  - Smaller group sizes per sighting, Smaller community size, Larger number of communities, Larger number of connected components.
- Dolphins that are connected to a particular dolphin are less likely to be connected to each other.
  - Smaller clustering coefficient.
- Though dolphins do have preferential associations while travelling, they do not travel in groups as large as those they socialize in or as small as they forage in.

Forage network * The dolphins in the forage network have the weakest and least repeated connections to other individuals. + Lower average degree, Lower average strength, Lowest average edge weight).

- Foraging happens in smaller groups than any other activity and these groups are exclusive with fewer links to other foraging groups or they are more likely in groups that never forage together.

    - Smallest group sizes per sighting, Smallest community size; Highest number of communities or Highest number of connected components.
- Foraging dolphins that are connected to other foraging dolphins are not as likely to be connected to each other as they are in the socialize and forage networks.
    - Lowest clustering coefficient)

Community Overlap ### AGGIUNGI TABLE4 A large community overlap between two networks means dolphins that tend to associate closely with each other in one network also associate closely in the other. * The socialize network and the travel network have the most substantial community structure overlap, the travel network and the forage network have the least and the socialize network and the forage network have an intermediate value. The overlap between the overall network and each activity network is less than that of the activity networks to each other.

Descriptive statistics