Exercise 2

Heidi Al Wakeel

2023-03-20

```
options(tinytex.verbose = TRUE)
```

```
library(readr)
library(igraph)
```

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

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

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

```
library(ggraph)
```

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

Create edge list

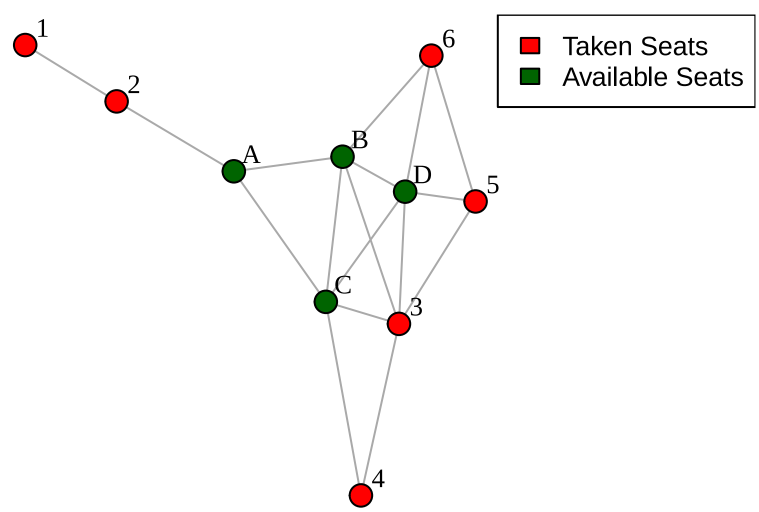
create the edges data frame and plot a undirected graph.

```
from_ <- c("1","2","3","3","3","3","3","4","5","5","6","6","D","D","B","B","A")
to_ <- c("2","A","D","C","B","4","5","C","D","6","D","B","B","C","C","A","C")
egde <- data.frame(from = from_, to=to_)
g <- graph_from_data_frame(egde, directed=FALSE)
g
```

```
## IGRAPH f0f6ef1 UN-- 10 17 --
## + attr: name (v/c)
## + edges from f0f6ef1 (vertex names):
##  [1] 1--2 2--A 3--D 3--C 3--B 3--4 3--5 4--C 5--D 5--6 6--D 6--B D--B D--C B--C
## [16] B--A A--C
```

Plot the network with seat label next to each node.

```
# First, create a vector of colors for each node
colors <- rep("red", vcount(g)) # create a vector of gray colors for all nodes
colors[match(c("A","B","C","D"), V(g)$name)] <- "darkgreen" # assign green color to nodes A, B, C, and

# Plot the graph with the specified colors
plot(g, layout=layout.fruchterman.reingold,
    vertex.size = 10,
    vertex.label = V(g)$name,
    vertex.label.cex = 0.8,
    vertex.label.dist = 1.5,
    vertex.label.color = "black",
    vertex.color = colors)

# Add a legend
legend("topright", legend=c("Taken Seats", "Available Seats"),
       fill=c("red", "darkgreen"), cex=0.8)
```



Betweenness Centrality

Betweenness centrality is a measure of a node's importance in a network. It calculates the number of times a node acts as a bridge along the shortest path between two other nodes. A node with high betweenness centrality has a significant influence on the communication and flow of information within the network, as it is often a key link between different groups or clusters.

```
bc <- betweenness(g)
bc
```

```
##          1          2          3          4          5          6          D
##  0.0000000  8.0000000  4.6333333  0.0000000  0.5333333  0.9333333  3.2666667
##          B          A          C
##  9.0333333 14.0000000  8.6000000
```

Seat A has the highest betweenness centrality value. This suggests that Seat A lies on many of the shortest paths between other seats in the network, making it an important connector in the network. Sitting in Seat A could allow for greater opportunities to interact with other people in the network.

Seat B has the second-highest betweenness centrality value indicating that it also lies on many of the shortest paths in the network. Sitting in Seat B could also provide good opportunities to interact with others in the network.

Seat C has a betweenness centrality value of 8.6, which is also relatively high. This suggests that it could also be a good spot for networking.

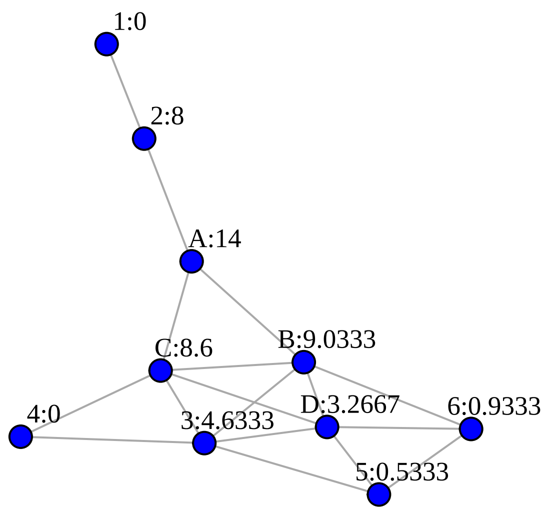
Seat D has the lowest betweenness centrality value of 3.2666667, meaning that it lies on fewer of the shortest paths in the network. Sitting in Seat D may not provide as many opportunities for interaction as the other seats.

Based on the betweenness centrality analysis of the given seat options, seat A has the highest betweenness centrality, which indicates that it may have a greater influence on the flow of information in the network compared to the other seats. Seat D has the lowest betweenness centrality, indicating it may have less influence on the flow of information.

Plot the network graph with labels and betweenness centrality values.

```
V(g)$betweenness <- round(betweenness(g),4)
label1 <- paste(V(g)$name,V(g)$betweenness,sep=":")

plot(g, layout=layout.fruchterman.reingold,
    vertex.size = 10,
    vertex.label = label1, # Set the labels (Node:betweenness)
    vertex.label.cex = 0.8,
    vertex.label.dist = 2,
    vertex.label.color = "black",
    vertex.color = "blue",
    main= "Betweenness Centrality Values")
```

Betweenness Centrality Values



Degree Centrality

Degree centrality is a measure of the centrality or importance of a node (or vertex) in a network based on the number of edges (or links) that are connected to it. Nodes with high degree centrality are more important in the network because they have more connections to other nodes, and therefore may have more influence, control or communication ability in the network. Degree centrality can be calculated by counting the number of edges that are connected to a particular node and then normalizing that count by dividing it by the maximum possible number of edges that a node can have in the network.

```
dc <- degree(g)
dc
```

```
## 1 2 3 4 5 6 D B A C
## 1 2 5 2 3 3 5 5 3 5
```

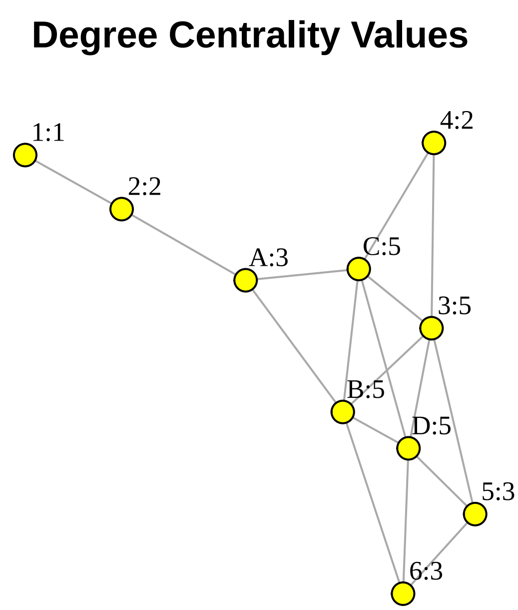
Nodes B, C, and D have the highest degree centrality with 5 connections each. This means that they are the most connected nodes in the network, and are likely to be important in terms of influencing or controlling the flow of information or resources within the network.

Node A has a degree centrality of 3, which is less than the other three nodes. However, it is still a relatively well-connected node and may have some influence or importance in the network.

In the context of choosing a seat, these degree centrality values suggest that seats B, C, and D may be more optimal for social networking, as they are more likely to provide opportunities for interacting with a larger number of people. On the other hand, seat A may not be as optimal for social networking, as it is less well-connected.

Plot the network graph with labels and degree centrality values.

```
V(g)$degree <- degree(g)
label2 <- paste(V(g)$name,V(g)$degree,sep=":")

plot(g, layout=layout.fruchterman.reingold,
    vertex.size = 10,
    vertex.label = label2, # Set the labels (Node:degree)
    vertex.label.cex = 0.8,
    vertex.label.dist = 2,
    vertex.label.color = "black",
    vertex.color = "yellow",
    main= "Degree Centrality Values")
```



Closeness Centrality

Closeness centrality is a measure of how quickly and efficiently information can flow from a given node to all other nodes in a network. It is based on the idea that nodes that are closer to all other nodes in the network can more easily access information and have greater influence over the network. Nodes with higher closeness centrality have shorter average distances to all other nodes, indicating that they are more central and have greater potential to influence the flow of information in the network.

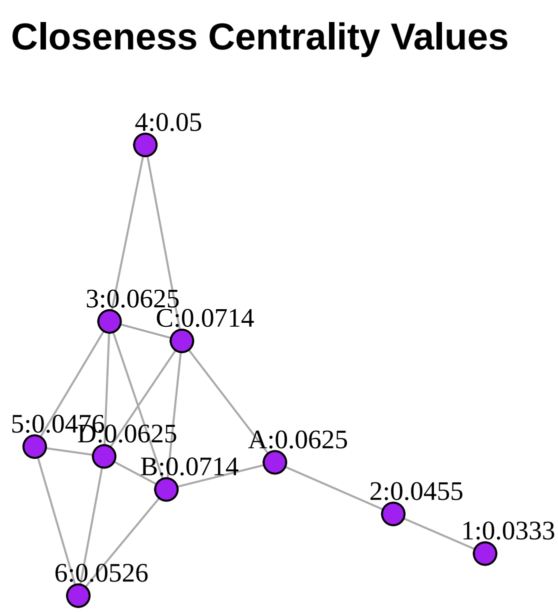
```
cc <- closeness(g)
cc
```

```
##          1          2          3          4          5          6          D
## 0.03333333 0.04545455 0.06250000 0.05000000 0.04761905 0.05263158 0.06250000
##          B          A          C
## 0.07142857 0.06250000 0.07142857
```

Based on these values, seats B and C have the highest closeness centrality, meaning that they are the most optimal for networking with other people in the graph. Seats D and A have lower closeness centrality, meaning that they may not be as good for networking. However, the difference between the closeness centrality values is relatively small, so the choice of seat may not have a significant impact on networking ability.

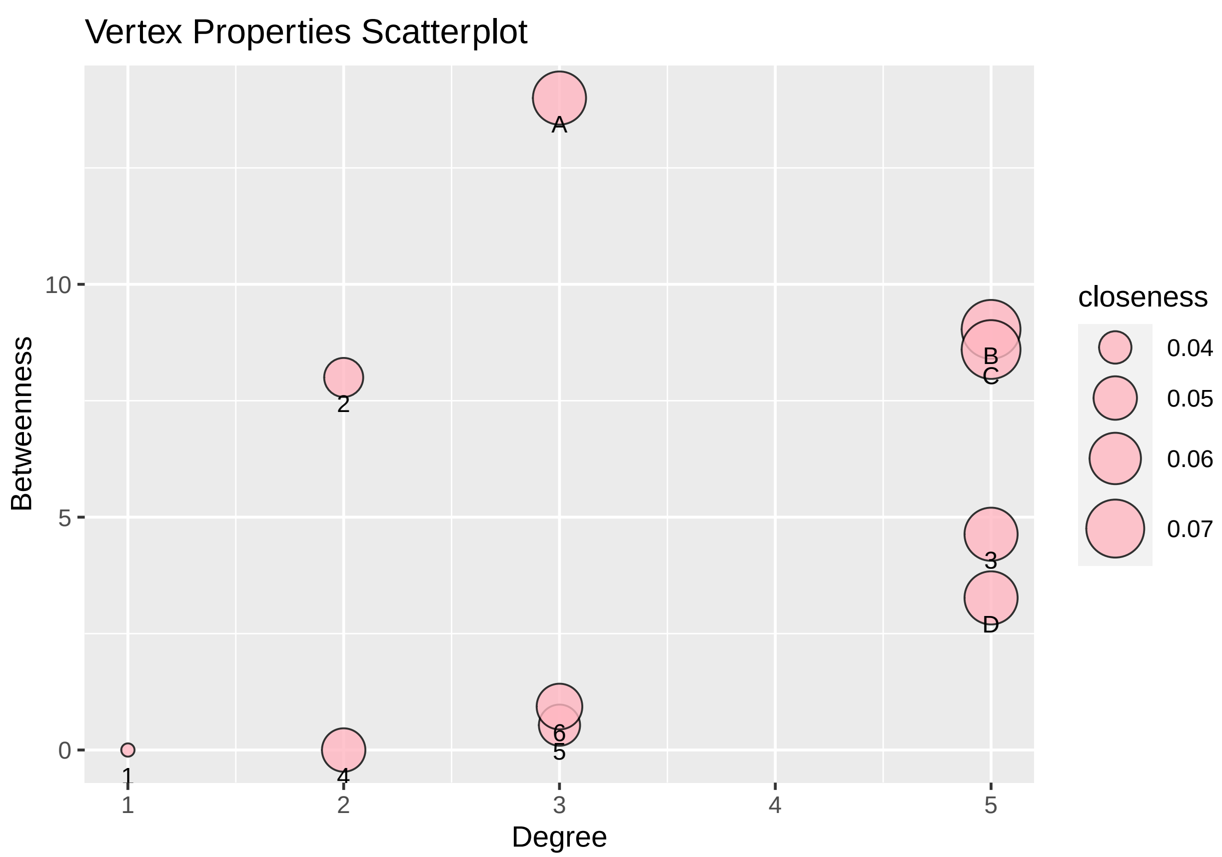
Plot the network graph with labels and closeness centrality values.

```
V(g)$closeness <- round(closeness(g),4)
label3 <- paste(V(g)$name,V(g)$closeness,sep=":")

plot(g, layout=layout.fruchterman.reingold,
    vertex.size = 10,
    vertex.label = label3, # Set the labels (Node:closeness)
    vertex.label.cex = 0.8,
    vertex.label.dist = 2,
    vertex.label.color = "black",
    vertex.color = "purple",
    main= "Closeness Centrality Values")
```



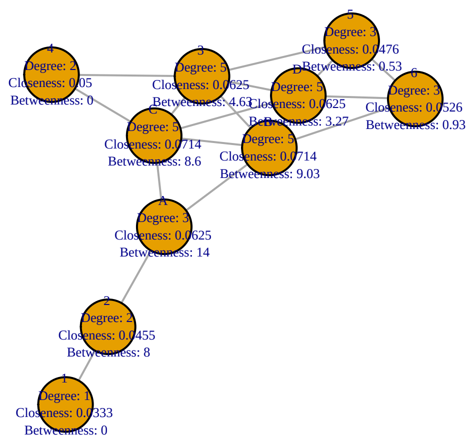
Seat Comparison

```r
#library(ggplot2)

# create a data frame with the vertex properties
df <- data.frame(vertex=V(g)$name,
                 closeness=round(closeness(g),4),
                 degree=degree(g),
                 betweenness=round(betweenness(g),4))

# create the scatterplot
ggplot(df, aes(x=degree, y=betweenness, size=closeness, label=vertex)) +
  geom_point(alpha=0.8, shape=21, fill="lightpink") +
  scale_size_continuous(range=c(2,10)) +
  geom_text(color="black", size=3, vjust=2) +
  xlab("Degree") +
  ylab("Betweenness") +
  ggtitle("Vertex Properties Scatterplot")
```




```r
# Add centrality values to the graph vertices
V(g)$degree <- degree(g)
V(g)$closeness <- closeness(g)
V(g)$betweenness <- betweenness(g)
```

```r
# Customize vertex labels
vertex_labels <- paste(V(g)$name,
                       "\nDegree:", V(g)$degree,
                       "\nCloseness:", round(V(g)$closeness, 4),
                       "\nBetweenness:", round(V(g)$betweenness, 2))

# Plot the network graph with labels and centrality values
plot(g,
     vertex.size = 30,
     vertex.label = vertex_labels,
     vertex.label.cex = 0.45,
     vertex.label.dist = 0.05,
     margin = 0.15)
```




```r
# Create a data frame
my_table <- data.frame(Seat = c("A", "B", "C","D"),
                       Betweeness = c(14, 9.0333333, 8.6,3.2666667),
                       Degree = c(3, 5, 5,5),
                       Closeness = c(0.06250000, 0.07142857, 0.07142857,0.06250000))



# Print the table
my_table
```

```
##   Seat Betweeness Degree  Closeness
## 1    A  14.000000      3 0.06250000
## 2    B   9.033333      5 0.07142857
## 3    C   8.600000      5 0.07142857
## 4    D   3.266667      5 0.06250000
```

Choice of seat

We will make decision based on the three measure of centrality we calculated.

Seat A:

Seat A has the highest betweenness centrality, which indicates that it is a key player in connecting other nodes in the network. However, it has a relatively low degree centrality and closeness centrality, which means it is not directly connected to as many other nodes and may not be the most efficient choice for quickly reaching other nodes.

Seat B:

Seat B has a high degree centrality, indicating that it is directly connected to many other nodes. It also has a relatively high betweenness centrality, suggesting that it plays an important role in connecting other nodes. Its closeness centrality is also relatively high, indicating that it can quickly reach other nodes. Therefore, Seat B could be a good choice for networking with many people.

Seat C:

Seat C has the same degree centrality and closeness centrality as Seat B, but a slightly lower betweenness centrality. This means that it is also well-connected and can quickly reach other nodes, but it may not play as critical a role in connecting other nodes as Seat B. However, it is still a good choice for networking with multiple people.

Seat D:

Seat D has the lowest betweenness centrality, indicating that it is not as critical in connecting other nodes as Seats A, B, and C. However, it still has a high degree centrality and can quickly reach other nodes with a relatively high closeness centrality. Therefore, Seat D could be a good choice for networking with a few people who are directly connected to it, but it may not be the best choice for reaching a large number of people.

Conclusion:

Looking at the degree centrality measure, seats B, C, and D have the highest degree, which means that they have the most direct connections to other nodes. However, seat A has a relatively low degree, which means it has fewer direct connections.

Examining the closeness centrality measure, seats B and C have the highest values, which means they are the most central nodes in terms of how quickly they can access other nodes in the network. This implies that sitting in these seats would allow us to reach out to other people in the network quickly and efficiently.

Finally, looking at the betweenness centrality measure, seat A has the highest value, which means it has the most potential to act as a bridge between different parts of the network. This suggests that sitting in seat A would provide opportunities to connect people from different parts of the network and potentially open up new job opportunities.

Based on the analysis of the centrality measures, the seat "B" appears to be the most optimal for networking. It has the highest degree centrality and second-highest closeness centrality, which suggests that it is well-connected to other seats in the network and has a relatively short path to other seats. Additionally, it has a high betweenness centrality, which means that it is located on many of the shortest paths between other pairs of seats, making it a potential hub for information flow and networking opportunities.However, it's important to consider that the betweenness centrality score for seat A is the highest, meaning that people in that seat may be a bridge between different groups or individuals in the room. Sitting in seat A may provide unique opportunities to connect with diverse groups of people, but it may also come with the responsibility of bridging different networks and potentially facing challenges in managing those connections. On the other hand, seats C and D have relatively lower scores in all centrality measures, indicating that sitting in these seats may not provide as many networking opportunities as seats B and A. Hence the seat "A" could also be a good option for networking.