# Exercise 2

Utkarsh Pal

2023-03-21

## The Fakebook Bus

**It takes many rides together to develop an informal connection, so you should pick one seat and take it consistently. Communication is only possible with adjacent seats: side, front, back, diagonal, even across the aisle. The following work demonstrates how to find the best seat to develop informal connections.**

### Loading the Packages:

```
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 4.2.3

##
## Attaching package: 'igraph'

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

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

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

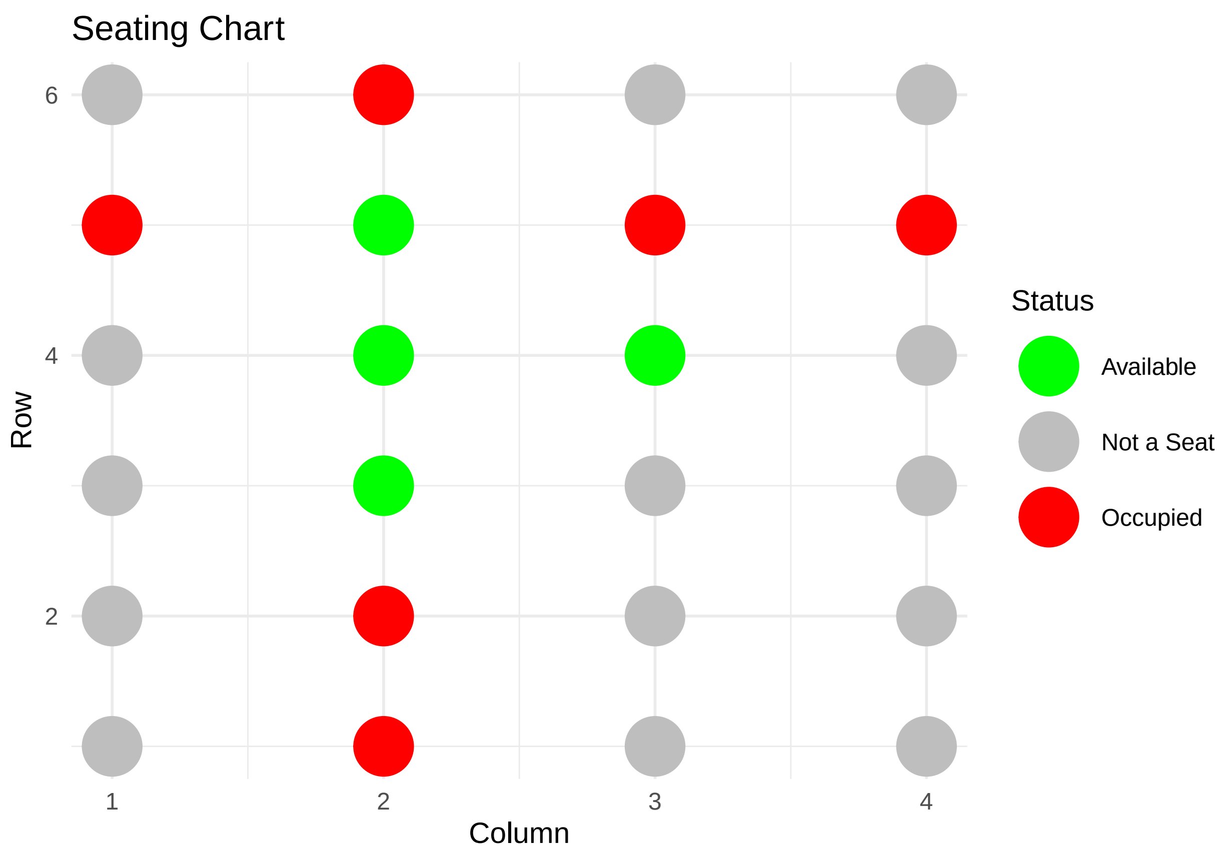
```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.2

##
## Attaching package: 'dplyr'

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

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

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

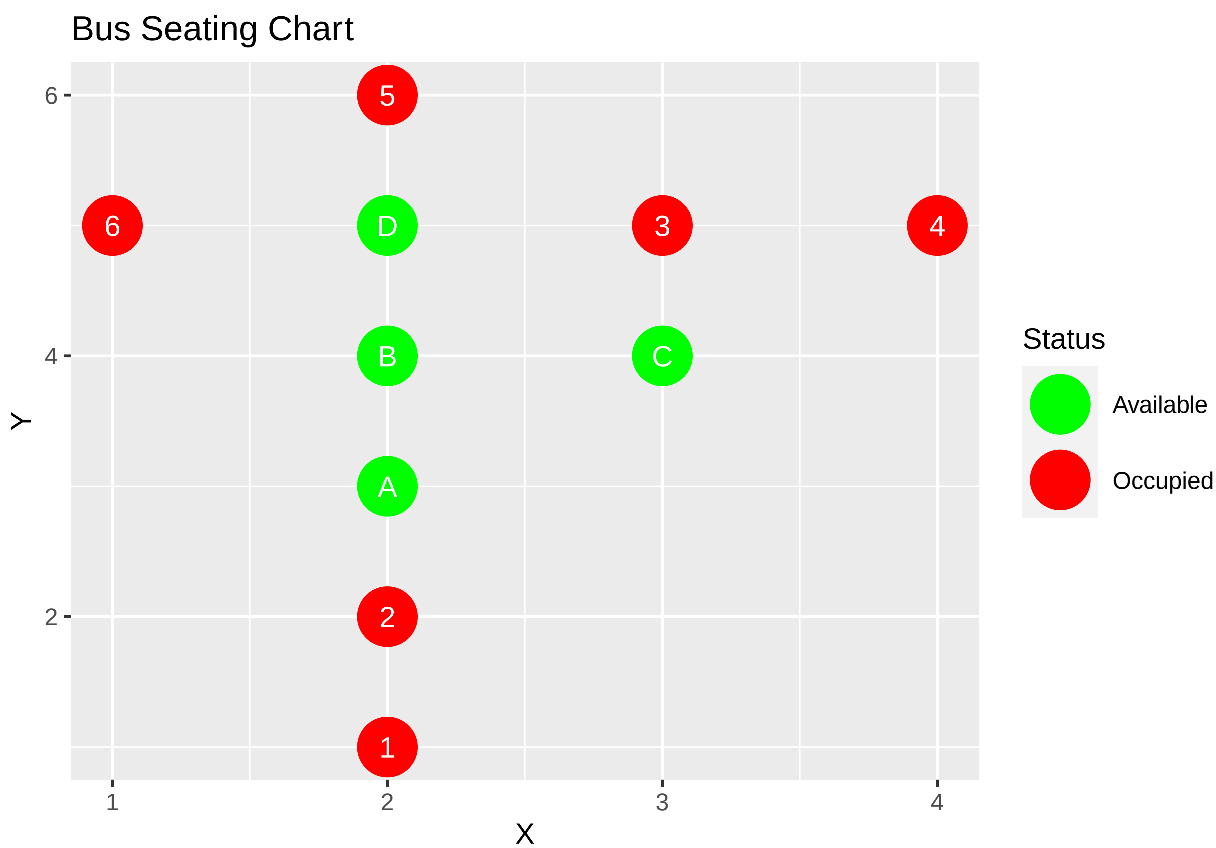
### Plotting the Bus Seating Chart including 'Not a Seat'

```r
# Define the dimensions of the bus
x_dimention <- 4
y_dimention <- 6

# Create a data frame with seat information
seats <- data.frame(
  Column = rep(c(1:x_dimention), times = y_dimention),
  Row = rep(c(1:y_dimention), each = x_dimention),
  Status = rep("Available", x_dimention * y_dimention)
)

# Set some conditions for unavailable seats
unavailable <- "Not a Seat"

# Remove non-existent seats
seats$Status[seats$Row != 5 & seats$Column == 1] <- unavailable
seats$Status[seats$Row != 5 & seats$Column == 4] <- unavailable
seats$Status[seats$Row < 4 & seats$Column > 2] <- unavailable
seats$Status[seats$Row == 6 & seats$Column > 2] <- unavailable

# Mark occupied seats
taken <- "Occupied"
seats$Status[(seats$Row < 3 | seats$Row > 5) & seats$Status == "Available"] <- taken
seats$Status[seats$Column < 2 & seats$Status == "Available"] <- taken
seats$Status[seats$Column > 2 & seats$Status == "Available" & seats$Row > 4] <- taken

# Define color mapping for different seat statuses
colors <- c("Available" = "green", "Occupied" = "red", "Not a Seat" = "gray")

# Create the plot
ggplot(seats, aes(x = Column, y = Row, color = Status)) +
  geom_point(size = 10) +
  scale_color_manual(values = colors) +
  labs(title = "Seating Chart") +
  theme_minimal()
```




### Plotting the Bus Seating Chart without 'Not a Seat'

```
Seats_Update <- seats %>%
  filter(Status != unavailable) %>%
  rename("Y" = "Row", "X"= "Column")
Seats_Update$id <- c("1", "2", "A", "B", "C", "6", "D", "3", "4", "5")

ggplot(data=Seats_Update,aes(x=X,y=Y,color=Status,label=id))+
  geom_point(size=10)+
  geom_text(hjust=.5, vjust=0.5,color="white")+
  scale_color_manual(values=c("green", "red"))+
  labs(title="Bus Seating Chart")
```




### Adjacency Matrix of the Seating Chart

```r
# Creating an adjacency matrix
from_node <- c("1","2","3","3","3","3","3","4","5","5","6","6","D","D","B","B","A")
to_node <- c("2","A","D","C","B","4","5","C","D","6","D","B","B","C","C","A","C")
edge <- data.frame(from=from_node, to=to_node)
g <- graph_from_data_frame(edge, directed=FALSE)

# The adjacency matrix of the seating chart
adj_mat <- get.adjacency(g, sparse = FALSE)
adj_mat
```

```
##   1 2 3 4 5 6 D B A C
## 1 0 1 0 0 0 0 0 0 0 0
## 2 1 0 0 0 0 0 0 0 1 0
## 3 0 0 0 1 1 0 1 1 0 1
## 4 0 0 1 0 0 0 0 0 0 1
## 5 0 0 1 0 0 1 1 0 0 0
## 6 0 0 0 0 1 0 1 1 0 0
## D 0 0 1 0 1 1 0 1 0 1
## B 0 0 1 0 0 1 1 0 1 1
## A 0 1 0 0 0 0 0 1 0 1
## C 0 0 1 1 0 0 1 1 1 0
```

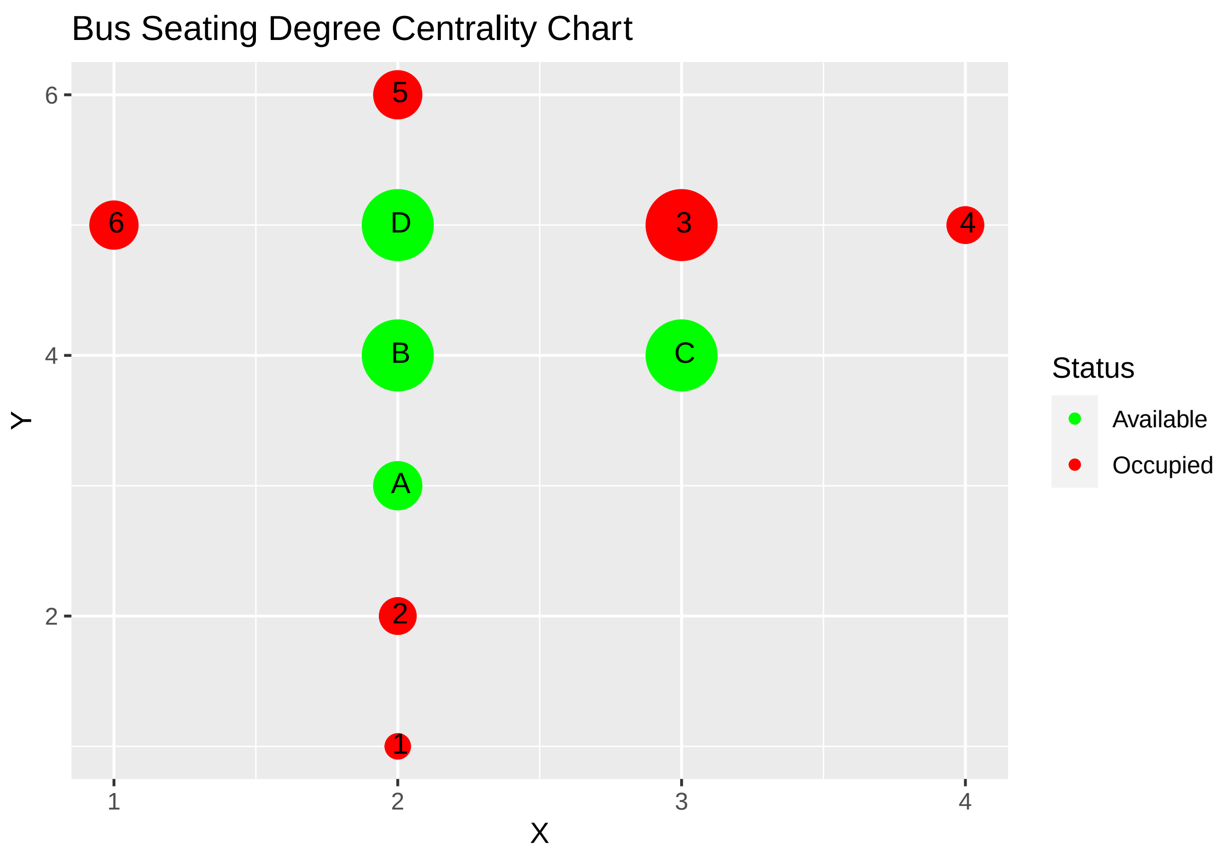
## Calculating and Plotting Degree Centrality

**The degree centrality of a node is simply the number of edges it has. The higher the degree, the more central the node is.**

```
degree_centrality <- degree(g, v = c("1", "2", "A", "B", "C", "6", "D", "3", "4", "5"), mode = "all")
print(degree_centrality)
```

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

```
# define the range of node sizes
min_size <- 4
max_size <- 12

# normalize the degree centrality values to range between 0 and 1
dc_normalized <- (degree_centrality - min(degree_centrality)) / (max(degree_centrality) - min(degree_cen

# map the normalized degree centrality values to the desired range of sizes
node_sizes <- dc_normalized * (max_size - min_size) + min_size

# add the node sizes to the Seats_Update data frame
Seats_Update$size <- node_sizes

# plot the graph with node size based on degree centrality
ggplot(data = Seats_Update, aes(x = X, y = Y, color = Status, label = id)) +
  geom_point(aes(size = size)) +
  geom_text(hjust = 0.35, vjust = 0.35, color = "black") +
  scale_color_manual(values = c("green", "red")) +
  scale_size_identity() +
  labs(title = "Bus Seating Degree Centrality Chart")
```




### Calculating and Plotting Closeness Centrality

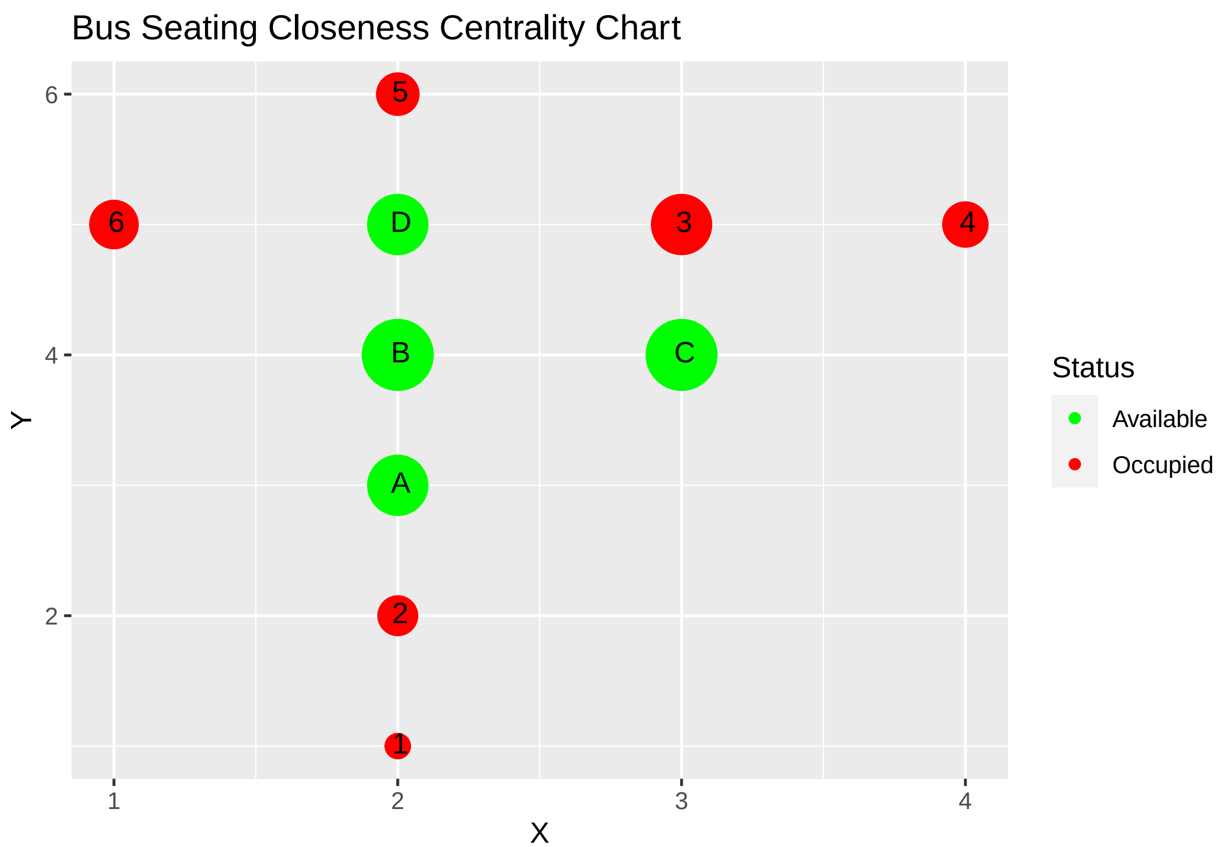
**Closeness centrality measures how short the shortest paths are from node i to all nodes.**

```
closeness_centrality <- closeness(g, v = c("1", "2", "A", "B", "C", "6", "D", "3", "4", "5"), mode = "al
print(closeness_centrality)
```

```
##         1         2         A         B         C         6         D         3
## 0.3000000 0.4090909 0.5625000 0.6428571 0.6428571 0.4736842 0.5625000 0.5625000
##         4         5
## 0.4500000 0.4285714
```

```
dc_normalized <- (closeness_centrality - min(closeness_centrality)) / (max(closeness_centrality) - min(c

# map the normalized degree centrality values to the desired range of sizes
node_sizes <- dc_normalized * (max_size - min_size) + min_size

# add the node sizes to the Seats_Update data frame
Seats_Update$size <- node_sizes

# plot the graph with node size based on degree centrality
ggplot(data = Seats_Update, aes(x = X, y = Y, color = Status, label = id)) +
  geom_point(aes(size = size)) +
```

```
  geom_text(hjust = 0.35, vjust = 0.35, color = "black") +
  scale_color_manual(values = c("green", "red")) +
  scale_size_identity() +
  labs(title = "Bus Seating Closeness Centrality Chart")
```




### Calculating and Plotting Betweenness Centrality

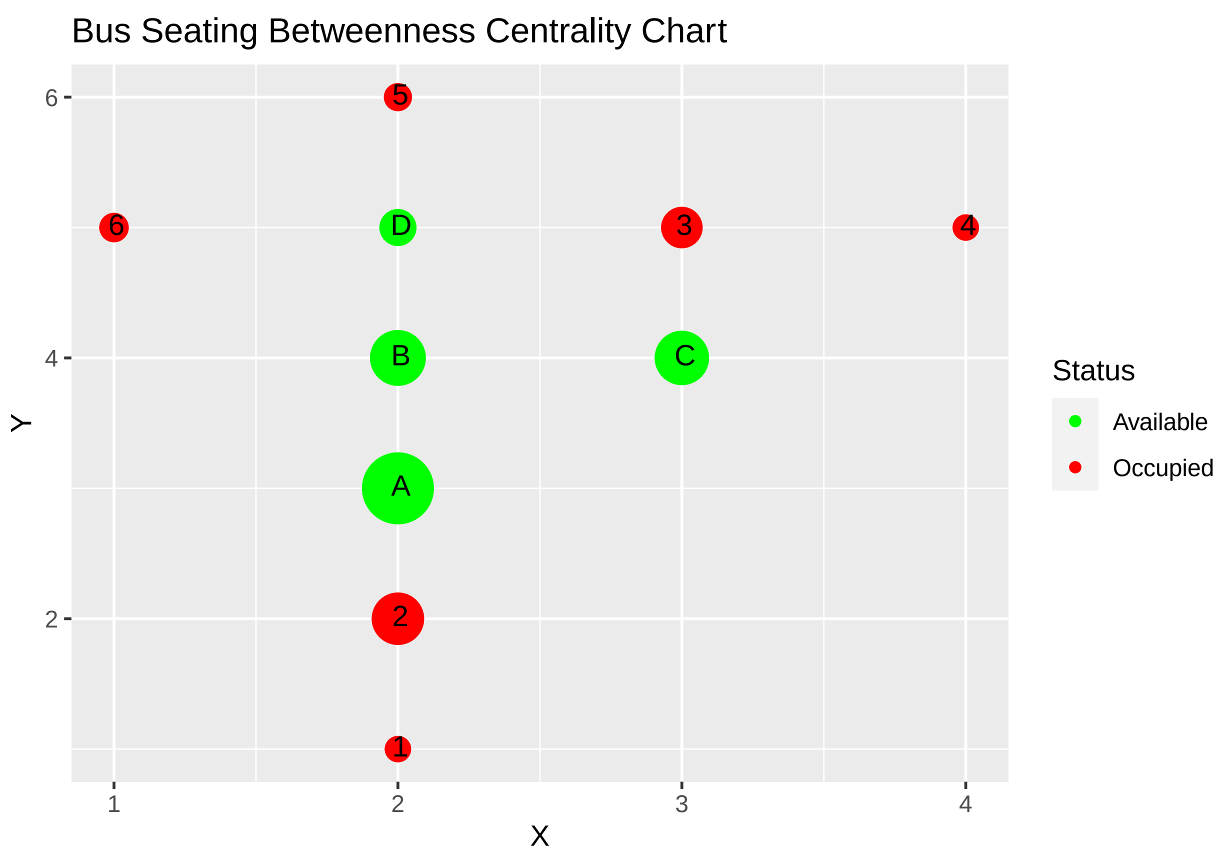
**Betweenness centrality is a way of detecting the amount of influence a node has over the flow of information in a graph. It is often used to find nodes that serve as a bridge from one part of a graph to another.**

```
betweenness_centrality <- betweenness(g, v = c("1", "2", "A", "B", "C", "6", "D", "3", "4", "5"))
print(betweenness_centrality)
```

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

```
dc_normalized <- (betweenness_centrality - min(betweenness_centrality)) / (max(betweenness_centrality) -

# map the normalized degree centrality values to the desired range of sizes
```

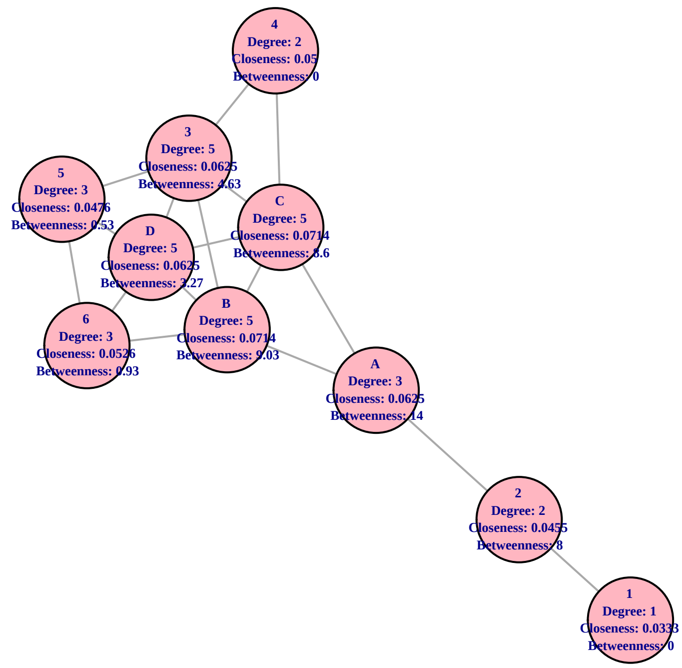
```
node_sizes <- dc_normalized * (max_size - min_size) + min_size

# add the node sizes to the Seats_Update data frame
Seats_Update$size <- node_sizes

# plot the graph with node size based on degree centrality
ggplot(data = Seats_Update, aes(x = X, y = Y, color = Status, label = id)) +
  geom_point(aes(size = size)) +
  geom_text(hjust = 0.35, vjust = 0.35, color = "black") +
  scale_color_manual(values = c("green", "red")) +
  scale_size_identity() +
  labs(title = "Bus Seating Betweenness Centrality Chart")
```




### Plotting the Degree, Closeness, and Betweenness Centrality and Showing the Link Between the Different Seats

```
V(g)$degree <- degree(g)
V(g)$closeness <- closeness(g)
V(g)$betweenness <- betweenness(g)

vertex_labels <- paste(V(g)$name,
                       "\nDegree:", V(g)$degree,
                       "\nCloseness:", round(V(g)$closeness, 4),
```

```r
                         "\nBetweenness:", round(V(g)$betweenness, 2))

plot(g,
     vertex.size = 30,
     vertex.label = vertex_labels,
     vertex.label.cex = 0.45,
     vertex.label.dist = 0.05,
     margin = 0.15,
     vertex.color = "light pink",
     vertex.label.font = 2
     )
```



## Possible Consequences for the Choice of Seat

**Seat A:**

**According to the Degree centrality, Seat A has a moderate level of connectivity with other seats on the bus, as it has a degree of 3, which means that there are three other seats that are directly connected to Seat A. However, in terms of Closeness Centrality, Seat A has a high value of 0.56, indicating that it is relatively closer to most other seats on the bus. Additionally, the Betweenness Centrality of Seat A is 14, which suggests that it may be a significant seat in terms of connecting other passengers on the bus.**

**Seat B:**

**Seat B has a degree centrality of 5, indicating that it is directly connected to five other seats on the bus. Moreover, the Closeness Centrality of Seat B is relatively high at 0.64, suggesting that it is closer to most other seats on the bus than many other seats. Furthermore, Seat B has a Betweenness Centrality of 9.03, which indicates that it may play a vital role in connecting passengers on the bus.**

**Seat C:**

**Seat C has the highest Degree Centrality of 6, indicating that it is directly connected to six other seats on the bus. Moreover, the Closeness Centrality of Seat C is 0.64, similar to Seat B, indicating that it is relatively closer to most other seats on the bus than many other seats. Additionally, the Betweenness Centrality of Seat C is 8.6, which implies that it may also be an essential seat for connecting other passengers on the bus.**

**Seat D:**

**Seat D has a Degree Centrality of 5, which means it is directly connected to five other seats on the bus. The Closeness Centrality of Seat D is 0.56, indicating that it is relatively closer to most other seats on the bus. Moreover, the Betweenness Centrality of Seat D is 3.26, which suggests that it may play a moderate role in connecting passengers on the bus.**

## Conclusion

**The best seat to make connections would depend on the specific preferences and goals of the individual.**

**i) For someone who wants to meet and interact with a large number of people, seats B,C, and D can be the ideal choice because they have the highest degree centrality. This means that they are connected to the largest number of other seats, making them a hub for social activity.**

**ii) For someone who values efficient communication and wants to be able to reach other seats quickly, seats B and C might be the best choice because they have the highest closeness centrality. This means that they are physically close to most other seats, making them ideal for quick and easy communication.**

**iii) For someone who wants to have the most influence over communication and interactions between other seats, seat A might be the best choice because it has the highest betweenness centrality. This means that it is a key mediator in the communication network, connecting otherwise disconnected seats and having the potential to control the flow of information.**

**If the individual wants a balance of all the different centrality measures, then I would recommend Seat B because it tops the measures for degree and closeness, and comes in second for the betweenness metric.**