Exercise2

Imports

```
# graph analysis
library(ggraph)
library(igraph)
```

Task Description

```
knitr::include_graphics("img/fakebookBus.PNG")
```

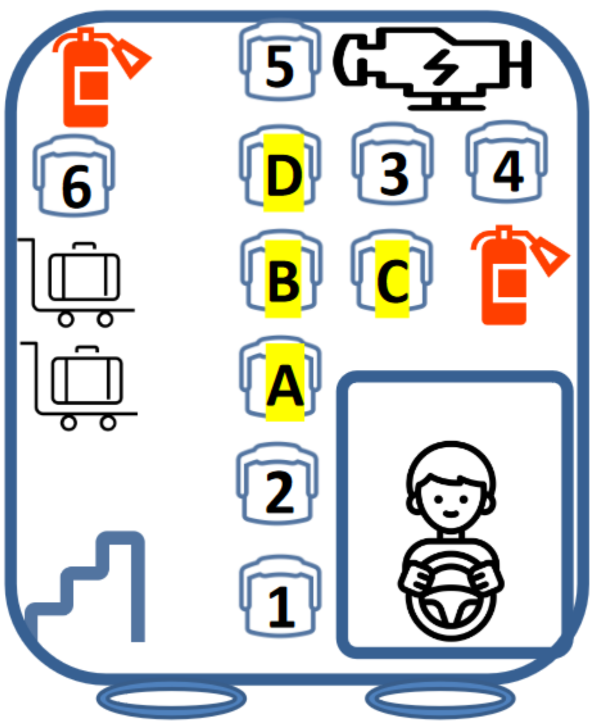

Figure 1: Fakebook Bus. Credit: Roman Galperin

Fakebook bus. - Create a dataset where edges are based on seat adjacency (course document) - For each seat choice, calculate various measures of centrality to discuss possible consequences of each choice. - Visualize the resultant network with centrality measures

Generate edge list

```r
#nodeList = c("A","B","C","D","1","2","3","4","5","6")
busDF = data.frame(seat1=character(), seat2=character(), edge=integer())

# Manually encode adjacencies
# Note: I'm assuming that diaganol adjencies are still valid even when up or down a row
# eg even though 4 cant talk to 5, 4 CAN talk to C, 3 CAN talk to B, etc etc

# Based on this assumption, I should have the following degree for each seat:
# 1: '1'
# 2: '2', '4'
# 3: 'A', '6', '5'
# 4:
# 5: 'B', 'C', 'D', '3'

busDF[nrow(busDF)+1,] = c(seat1='1', seat2='2', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='2', seat2='A', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='A', seat2='B', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='A', seat2='C', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='B', seat2='C', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='B', seat2='D', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='B', seat2='6', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='B', seat2='3', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='D', seat2='6', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='D', seat2='3', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='D', seat2='C', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='D', seat2='5', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='5', seat2='6', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='5', seat2='3', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='C', seat2='3', edge=1)
busDF[nrow(busDF)+1,] = c(seat1='C', seat2='4', edge=1)

busDF[nrow(busDF)+1,] = c(seat1='3', seat2='4', edge=1)

busDF
```

```
##    seat1 seat2 edge
## 1      1     2    1
## 2      2     A    1
## 3      A     B    1
## 4      A     C    1
## 5      B     C    1
## 6      B     D    1
## 7      B     6    1
## 8      B     3    1
## 9      D     6    1
```

```
## 10     D     3    1
## 11     D     C    1
## 12     D     5    1
## 13     5     6    1
## 14     5     3    1
## 15     C     3    1
## 16     C     4    1
## 17     3     4    1
```

Generate Graph Object

```r
fakebookGraph = graph_from_data_frame(d=busDF,vertices=NULL ,directed=FALSE)
fakebookGraph
```

```
## IGRAPH 79942bf UN-- 10 17 --
## + attr: name (v/c), edge (e/c)
## + edges from 79942bf (vertex names):
##  [1] 1--2 2--A A--B A--C B--C B--D B--6 B--3 D--6 D--3 D--C D--5 5--6 5--3 C--3
## [16] C--4 3--4
```

Measures of Centrality

Degree Centrality

This just confirms what we already knew from looking at the image

```r
Degree <- degree(fakebookGraph, v=V(fakebookGraph))
```

Betweenness Centrality

Aka how important each node is in connecting disparate parts of the network

```r
Betweenness <- betweenness(fakebookGraph)
```

Eigenvector Centrality

From Science Direct "Eigenvector centrality measures a node's importance while giving consideration to the importance of its neighbors… It is sometimes used to measure a node's influence in the network"

```r
Eigenvector <- evcent(fakebookGraph)$vector
```

Centrality Discussion

```r
centralities <- cbind(Degree, Eigenvector, Betweenness)
round(centralities,2)
```

```
##    Degree Eigenvector Betweenness
## 1       1        0.03        0.00
## 2       2        0.13        8.00
## A       3        0.49       14.00
## B       5        0.97        9.03
## D       5        1.00        3.27
## 5       3        0.63        0.53
## C       5        0.94        8.60
## 3       5        0.97        4.63
## 6       3        0.63        0.93
## 4       2        0.46        0.00
```

There are several observations to be made based on the above centrality matrix

1) Seat A would be a good choice based on Betweenness centrality alone, but this can be attributed to it being the only seat connecting 1 and 2 with the rest of the bus - if we evaluate the other metrics for seat A it does not seem like as a good of a choice if we wanted to maximize our likelihood of making friends/colleagues on the bus

2) Degree centrality can give us a good intuition of which seats will be good candidates. B, C, and D all have degree 5 which gives plenty of opportunity to make friends with adjacent seats.

3) Eigenvector is perhaps most revealing, since it takes into account not only a seat's individual importance, but also the importance of surrounding seats (analagous to google search algorithm PageRank). The eigenvector centrality would tell us seats 3, B, and D are the top important seats on the bus, with D grading out as the absolute most important. This result somewhat matches the "eye test" as 3 D and B are all more or less in the "center" of the seating arrangement.

In the end, B or D are good choices for trying to make friends on the bus. Betweenness could be seen as the differentiating factor between them since their eigenvector centrality are quite similar. Perhaps if one cared solely about being in the most "important" seat, they could pick D. Otherwise, if one cared about being connected to all the different parts of the bus, they could pick B.

Visualize

We will size by degree, then display betweenness and eigenvector separately

```
V(fakebookGraph)$size = Degree
V(fakebookGraph)$eig = round(Eigenvector,2)
V(fakebookGraph)$bet = round(Betweenness,2)

ggraph(fakebookGraph, layout="kk") +
  geom_edge_link() +
  geom_node_point(aes(size=size, color="red"), show.legend=F)+scale_size(range=c(2,30))+
  geom_node_text(aes(label=paste(name,paste("\nE=",eig),paste("\nB=",bet))))
```

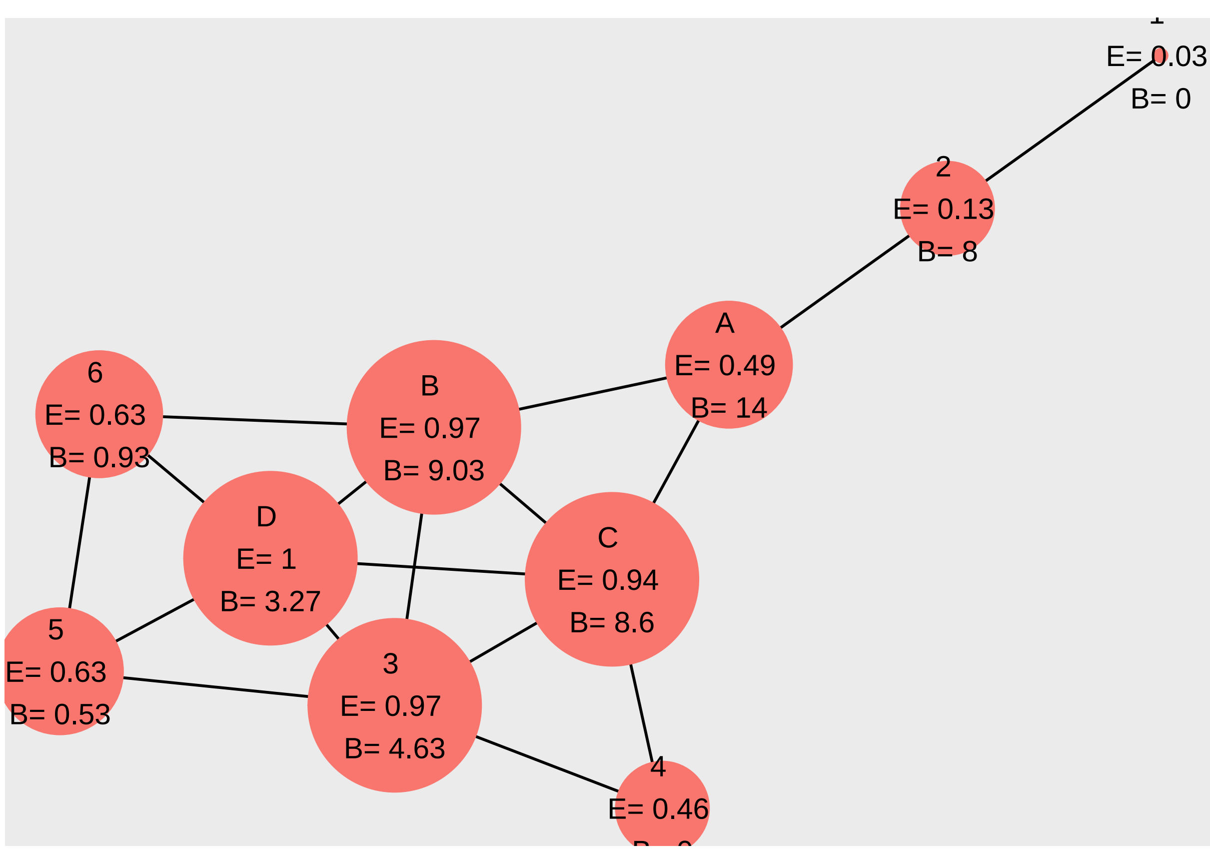

Figure 2: Fakebook Network. Node sizes proportional to their degree centrality; eigenvector and betweenness centralities are provided within node labels.