Assignment 3

Atharva Jagtap 39283783

2024-10-03

```
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 4.3.3
```

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

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

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

```
options(digits=3)

advice <- matrix(scan("Advice.txt"),ncol=71,nrow=71,byrow=T)
nodes <- read.csv("lawyers.csv", header=T)
head(nodes)
```

```
##   Name Seniority Status Gender City Years Age Practice LawSchool
## 1   V1         1      1      1    1    31  64        1         1
## 2   V2         2      1      1    1    32  62        2         1
## 3   V3         3      1      1    2    13  67        1         1
## 4   V4         4      1      1    1    31  59        2         3
## 5   V5         5      1      1    2    31  59        1         2
## 6   V6         6      1      1    2    29  55        1         1
```

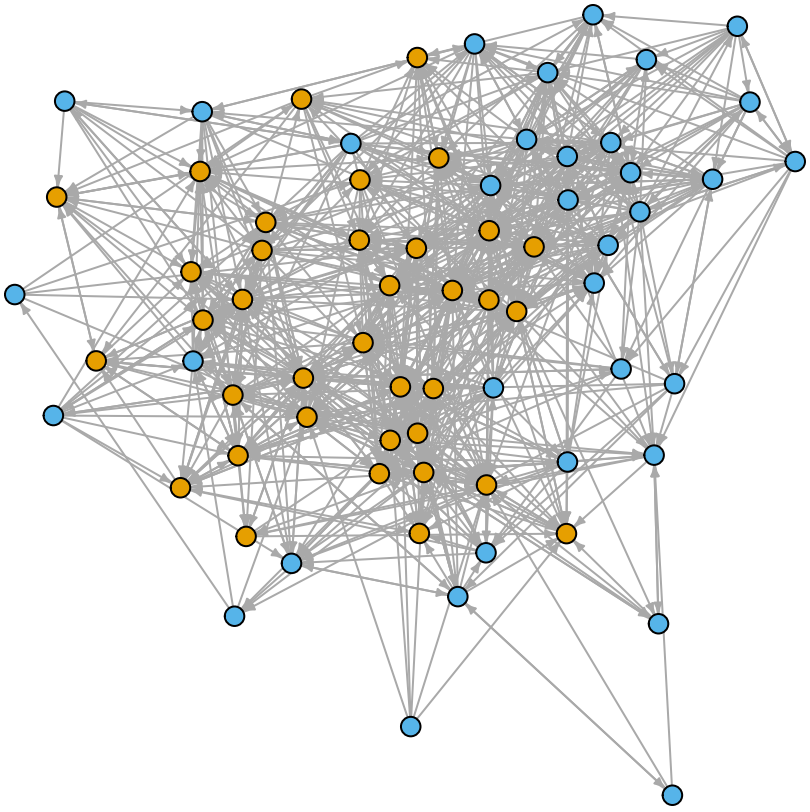
```
graph <- graph_from_adjacency_matrix(advice, mode = "directed")
```

```
#1
status_count <- table(nodes$Status)
partners_count <- status_count[1] # Partners
associates_count <- status_count[2] # Associates

gender_count <- table(nodes$Gender)
men_count <- gender_count[1] # Men
women_count <- gender_count[2] # Women
```

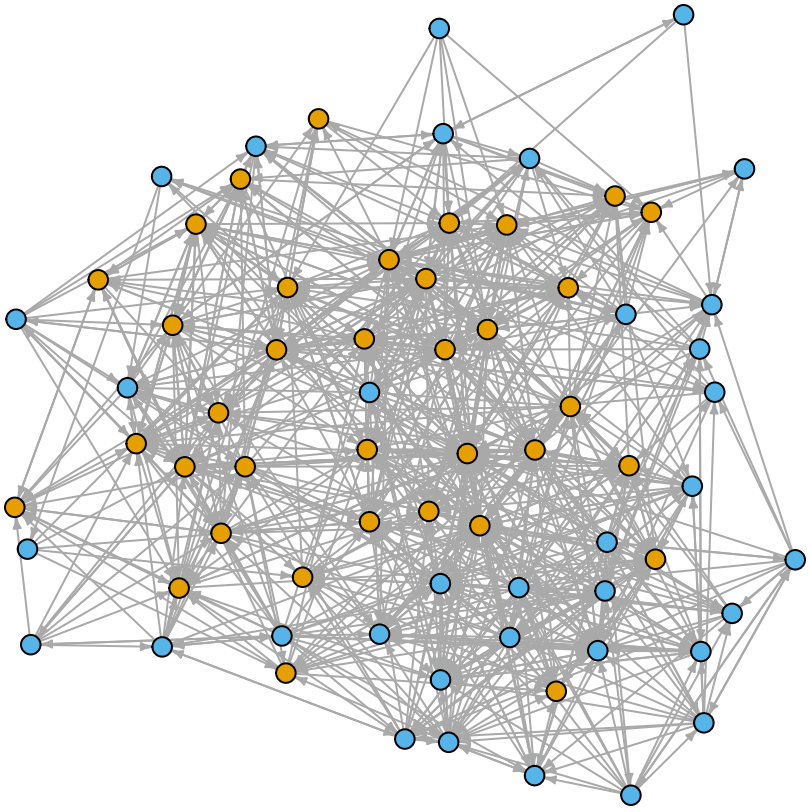
```r
practice_count <- table(nodes$Practice)
litigation_count <- practice_count[1] # Litigation
corporate_count <- practice_count[2] # Corporate

Counts <- data.frame(
  Group = c("Partners", "Associates", "Men", "Women", "Litigation",
            "Corporate"),
  Group_Size = c(partners_count, associates_count, men_count, women_count,
                 litigation_count, corporate_count)
)

Counts
```

```
##        Group Group_Size
## 1   Partners         36
## 2 Associates         35
## 3        Men         53
## 4      Women         18
## 5 Litigation         41
## 6  Corporate         30
```

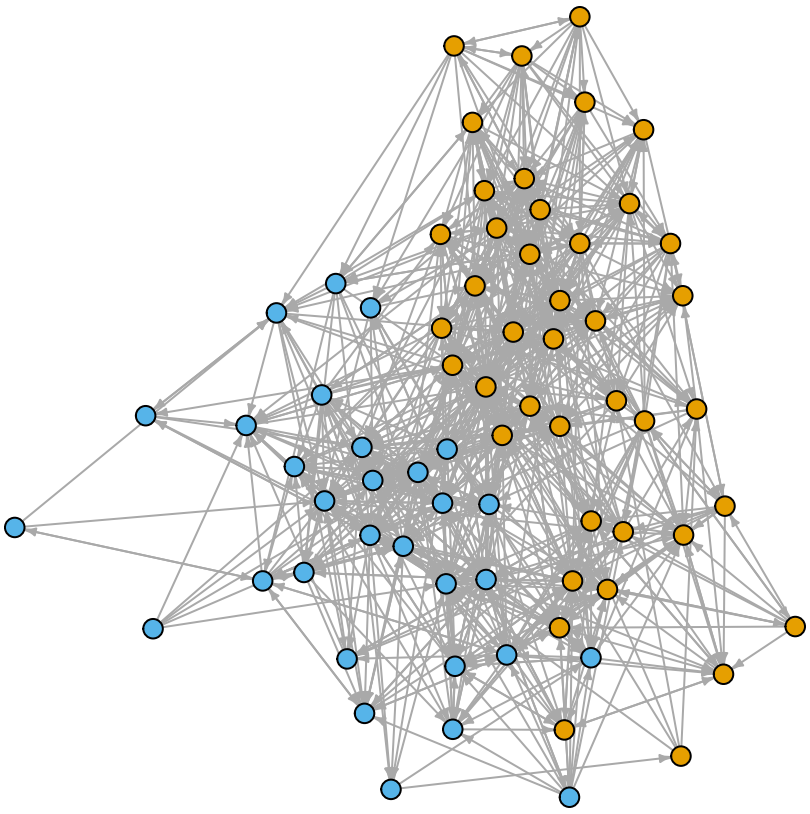
```r
#2
V(graph)$status <- nodes$Status
V(graph)$practice <- nodes$Practice
V(graph)$city <- nodes$City

par(mar = c(0, 0, 0, 0))

#Status network
plot(graph,
     vertex.size=5,
     edge.arrow.size=0.3,
     vertex.color=V(graph)$status,
     vertex.label=NA,
     layout=layout_with_fr)
```



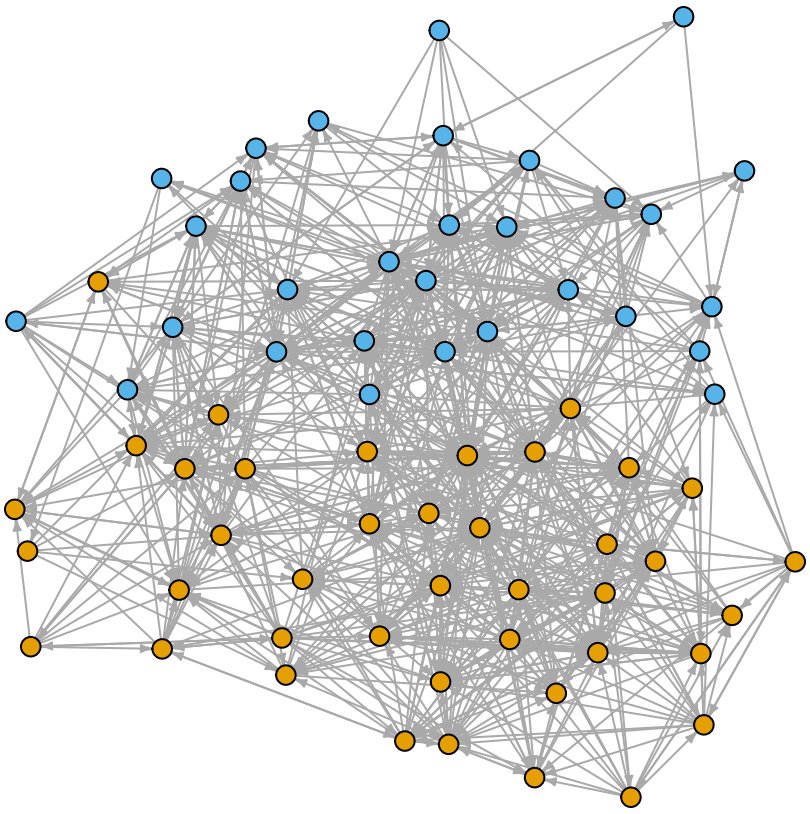
```
#Different Layout
plot(graph,
     vertex.size=5,
     edge.arrow.size=0.3,
     vertex.color=V(graph)$status,
     vertex.label=NA,
     layout=layout_with_kk)
```



```
#Practice Network
plot(graph,
     vertex.size=5,
     edge.arrow.size=0.3,
     vertex.color=V(graph)$practice,
     vertex.label=NA,
     layout=layout_with_fr)
```

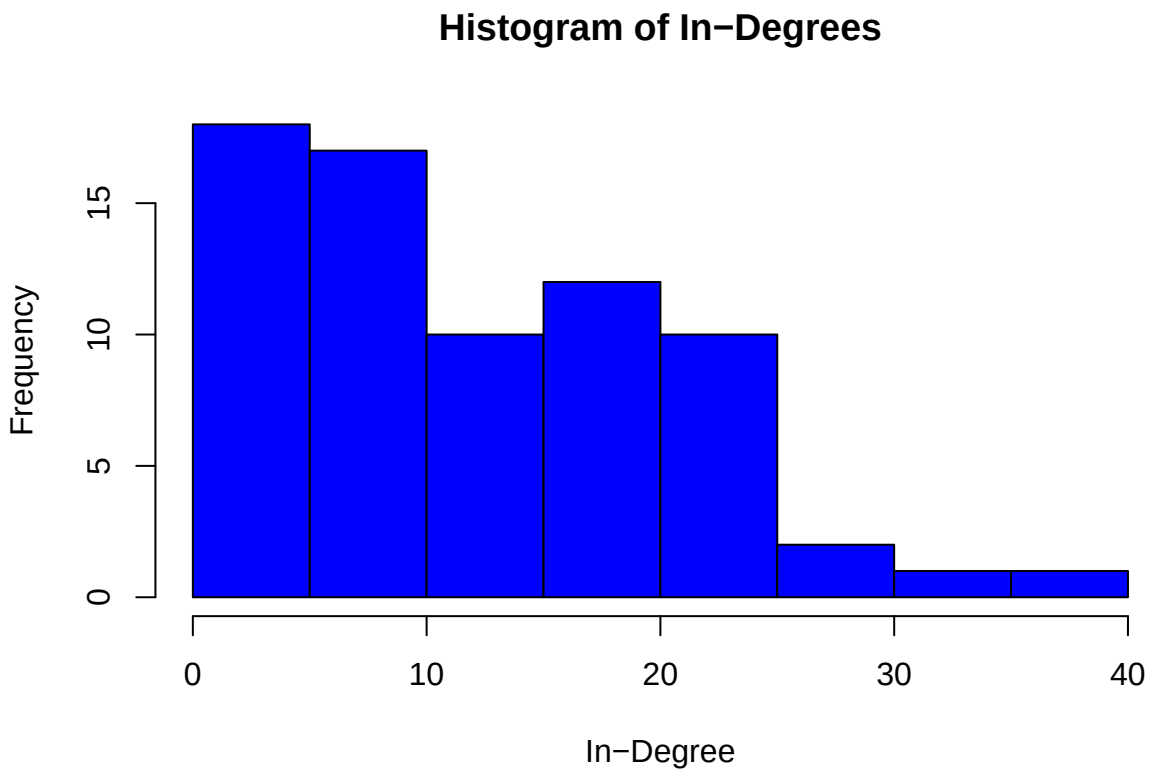


```
#Different Layout
plot(graph,
     vertex.size=5,
     edge.arrow.size=0.3,
     vertex.color=V(graph)$practice,
     vertex.label=NA,
     layout=layout_with_kk)
```

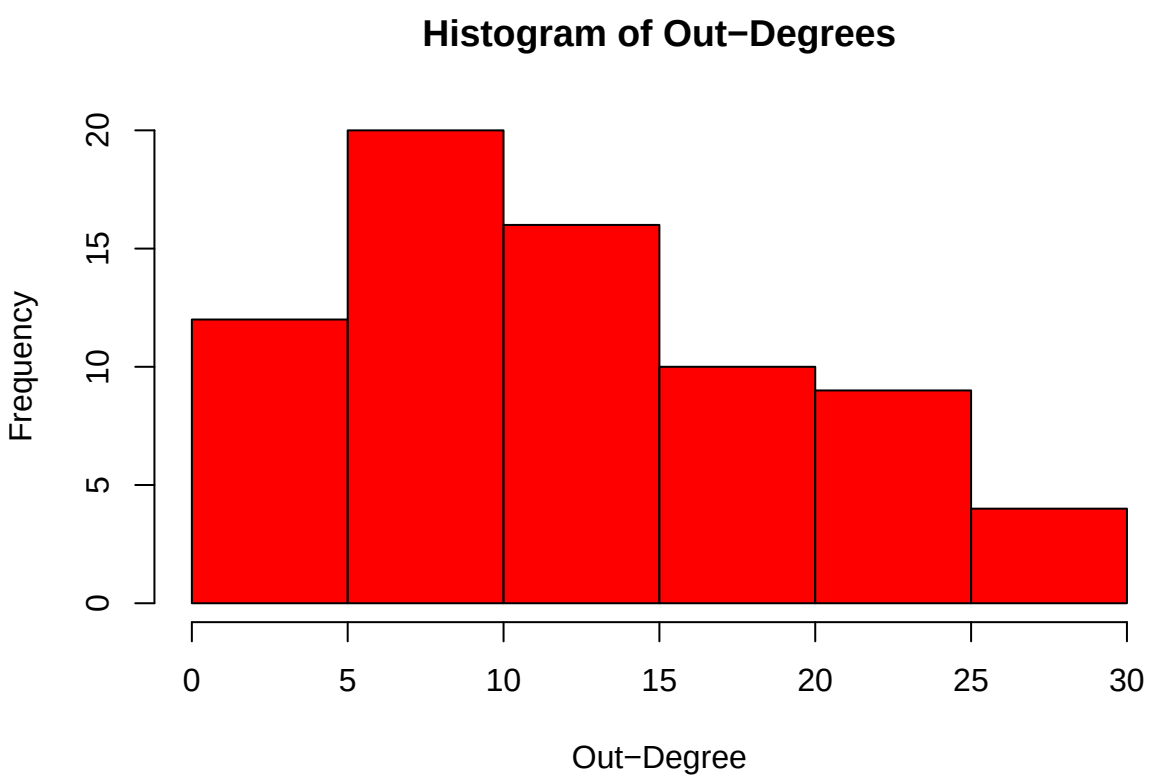


```
## Findings:
##
## In the status plot, you observe that vertices representing partners and
## associates cluster together, indicating that lawyers of similar hierarchical
## levels tend to interact more frequently within the firm. Partners form
## tight-knit groups, reflecting their collaboration on high-level cases,
## while associates group together due to shared training or routine casework.
## While cross-connections between partners and associates can suggest
## hierarchical divides and/or strong mentorship and collaboration.
##
## In the practice plot, clustering highlights how lawyers from the same
## practice group work closely together due to the nature of their specialized
## cases. However, cross-practice ties suggest interdisciplinary collaboration,
## where different practice groups work together on cases requiring diverse legal
## expertise, fostering integration across the firm.
```

```
#3
indegree <- degree(graph, mode = "in")
outdegree <- degree(graph, mode = "out")

# Histogram for in-degrees
hist(indegree,
     main = "Histogram of In-Degrees",
     xlab = "In-Degree",
     col = "blue",
     breaks = 10)
```

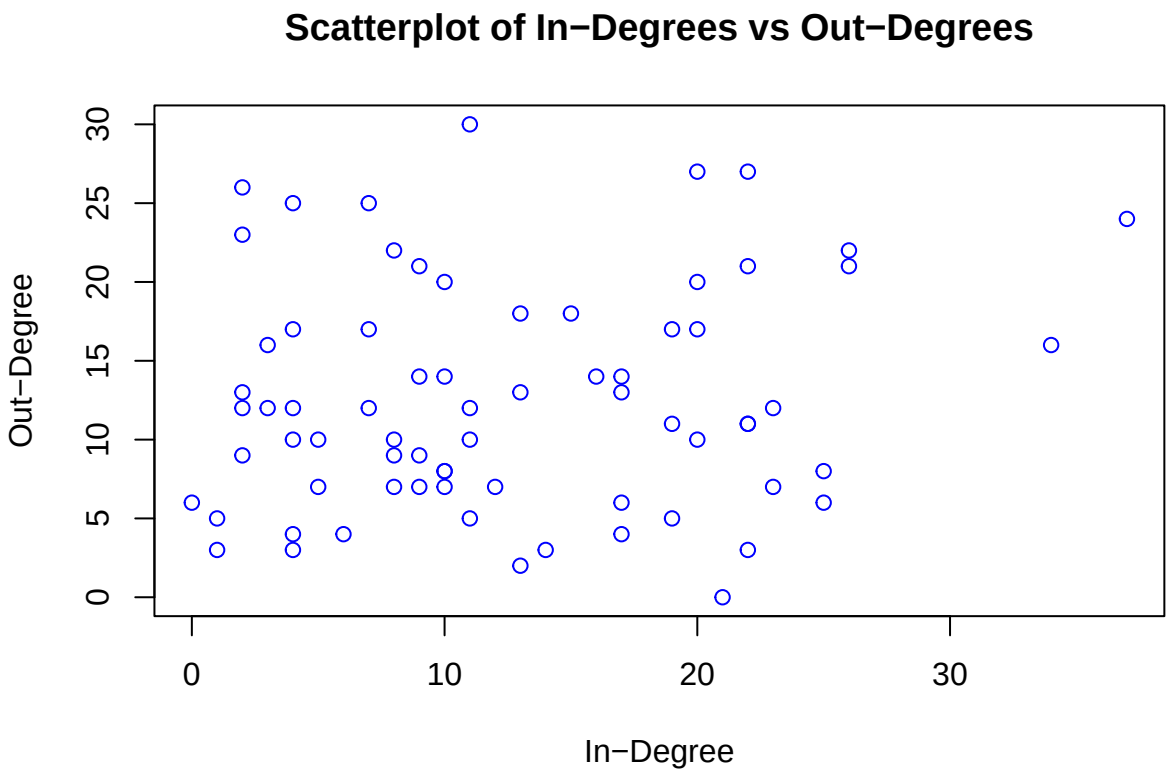



```
# Histogram for out-degrees
hist(outdegree,
     main = "Histogram of Out-Degrees",
     xlab = "Out-Degree",
     col = "red",
     breaks = 10)
```




```
##
## Yes, the two histograms are similar in nature. Both distributions are
## skewed, indicating that a minority of individuals are highly central to the
## advice network, either as recipients or providers of advice.
```

```
#4
plot(indegree, outdegree,
     main = "Scatterplot of In-Degrees vs Out-Degrees",
     xlab = "In-Degree",
     ylab = "Out-Degree",
     col = "blue")
```




```
correlation <- cor(indegree, outdegree)
correlation
```

```
## [1] 0.14
```

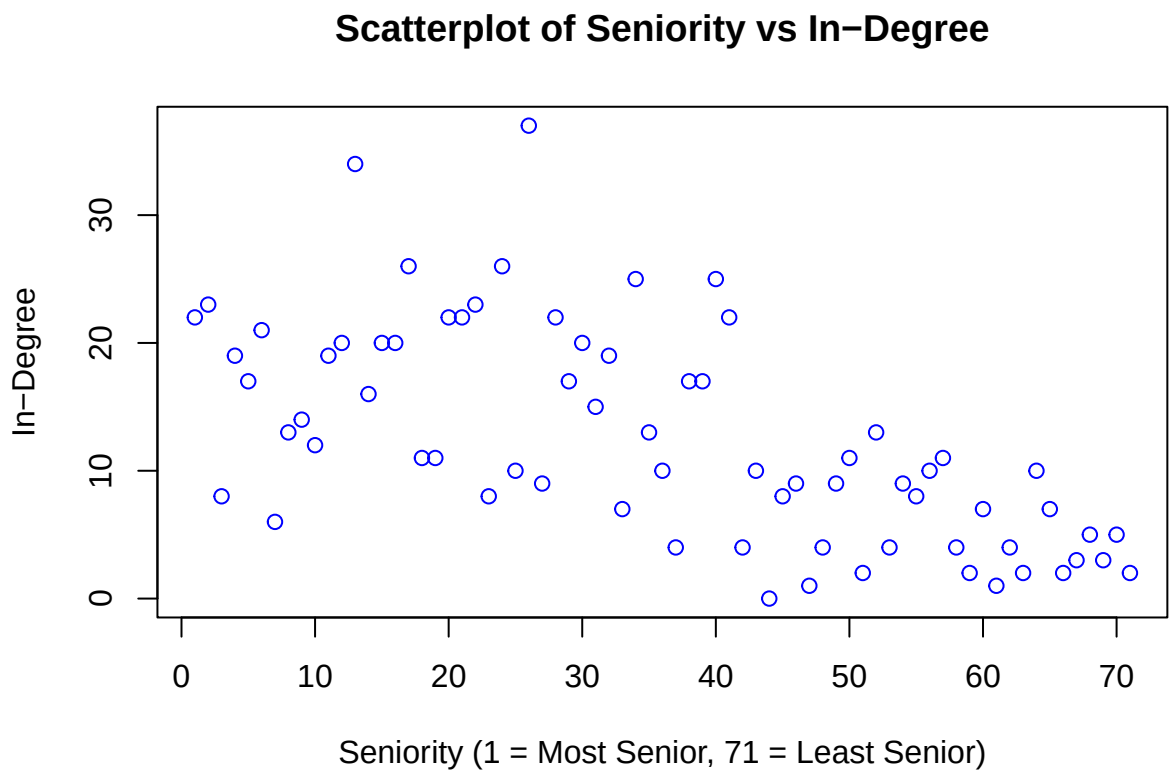
```
##
## With a correlation of 0.14, the scatterplot shows a weak positive relationship
## between in-degrees and out-degrees. This suggests that there is little linear
## relationship between how much advice someone gives and how much advice they
## receive. This could reflect a specialization within the firm, where certain
## individuals are primarily advice-givers (mentors, senior partners), while
## others are primarily advice-receivers (junior associates or those seeking
## guidance).
```

```
#5
nodes[, 10] <- indegree
nodes[, 11] <- outdegree

colnames(nodes)[10:11] <- c("indegree", "outdegree")
head(nodes, 5)
```

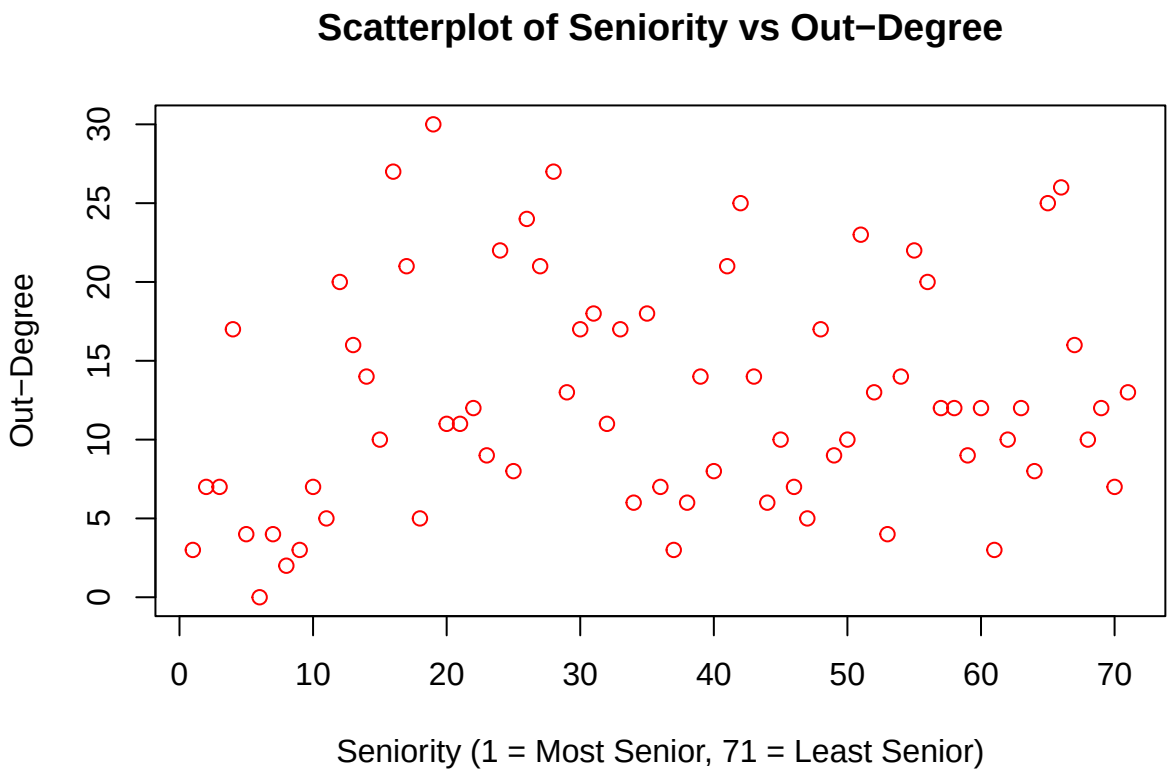
```
##   Name Seniority Status Gender City Years Age Practice LawSchool indegree
## 1   V1         1      1      1    1    31  64        1         1       22
## 2   V2         2      1      1    1    32  62        2         1       23
```

```
## 3   V3         3      1      1    2    13  67        1         1        8
## 4   V4         4      1      1    1    31  59        2         3       19
## 5   V5         5      1      1    2    31  59        1         2       17
##   outdegree
## 1         3
## 2         7
## 3         7
## 4        17
## 5         4
```

```r
#5a
#Seniority vs In-Degree
plot(nodes$Seniority, nodes$indegree,
     main = "Scatterplot of Seniority vs In-Degree",
     xlab = "Seniority (1 = Most Senior, 71 = Least Senior)",
     ylab = "In-Degree",
     col = "blue")
```

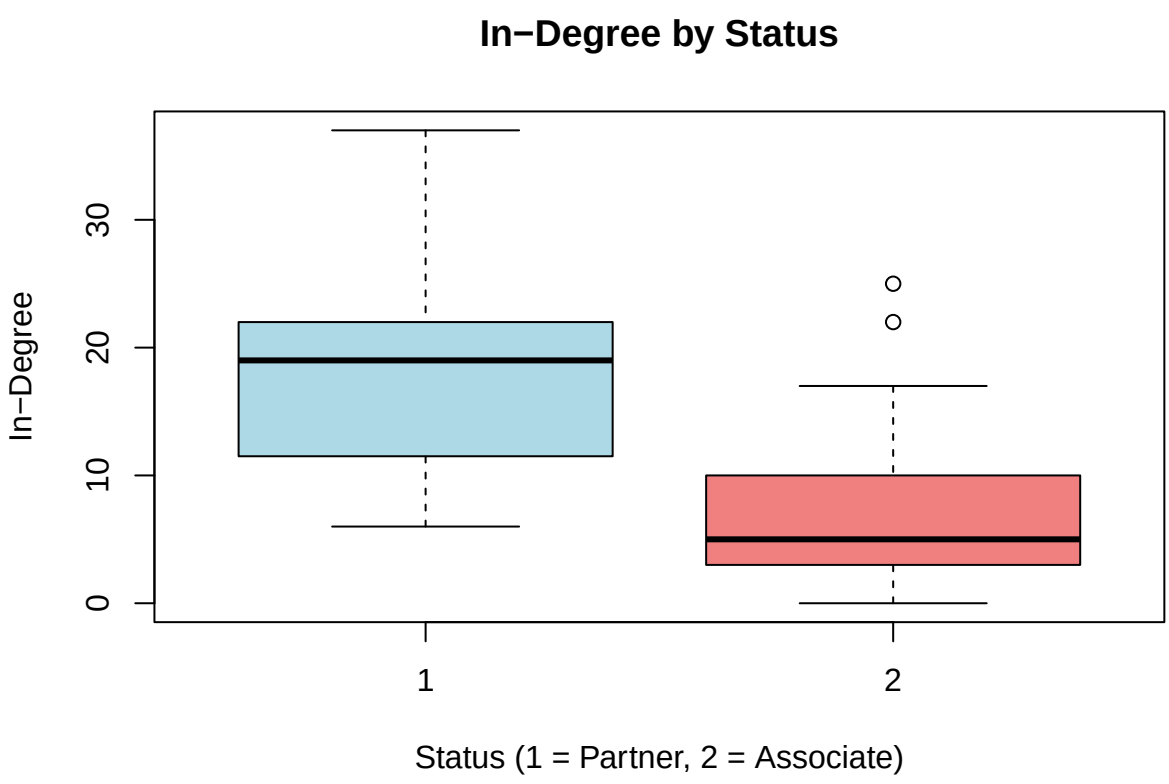



```r
#Seniority vs Out-Degree
plot(nodes$Seniority, nodes$outdegree,
     main = "Scatterplot of Seniority vs Out-Degree",
     xlab = "Seniority (1 = Most Senior, 71 = Least Senior)",
     ylab = "Out-Degree",
     col = "red")
```

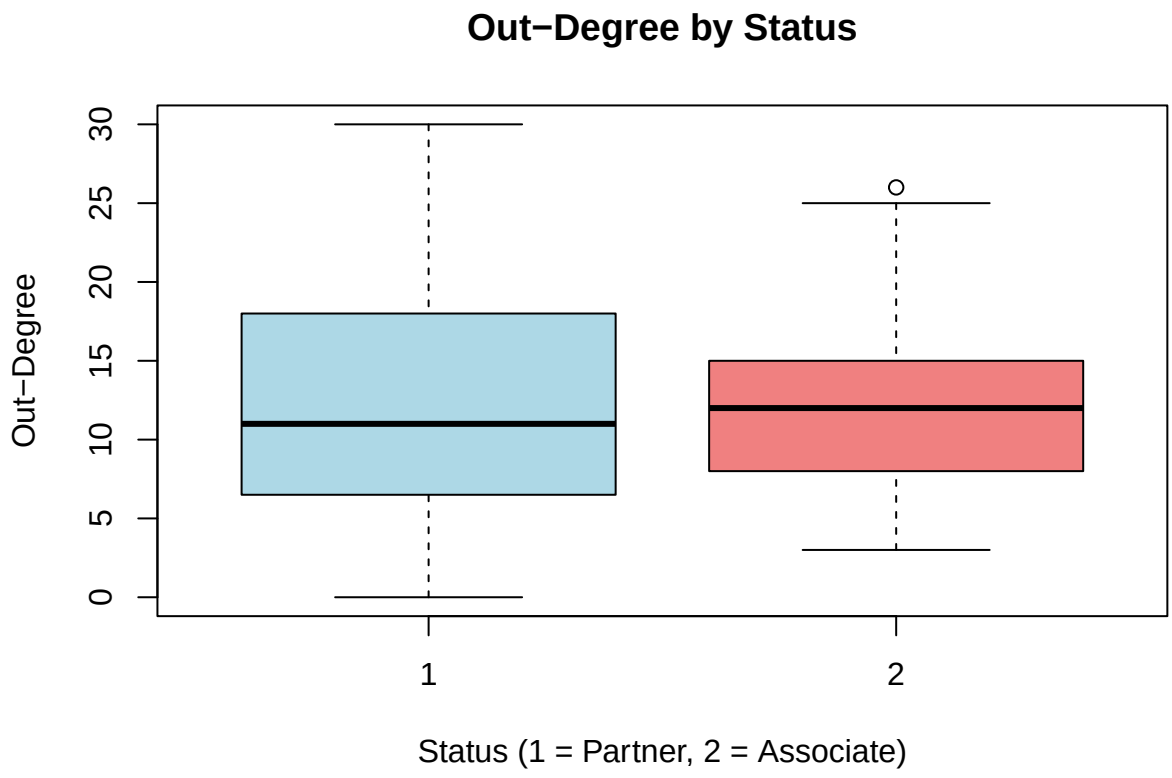



```
## Findings:
##
## Seniority vs In-Degree:
## We notice that senior lawyers, being more experienced, are sought out for
## advice more frequently, leading to higher in-degrees. As we go lower in
## seniority, the in-degrees decrease as junior lawyers are consulted less often.
##
## Seniority vs Out-Degree:
## There is no evident pattern visible but we can deduce that the highest
## seniority lawyers (1-10) have a very low out-degree as they do not need to seek
## advice as much as the rest of the lawyers. Beyond 10, the plot suggests that
## there isn't a strong relationship between seniority and out-degree. Lawyers
## across different seniority levels appear to have a wide range of out-degrees,
## meaning that advice-seeking is somewhat distributed across different levels of
## seniority.
```

```r
#5b
#In-degree by status
boxplot(nodes$indegree ~ nodes$Status,
        main = "In-Degree by Status",
        xlab = "Status (1 = Partner, 2 = Associate)",
        ylab = "In-Degree",
        col = c("lightblue", "lightcoral"))
```




```
#Out-degree by status
boxplot(nodes$outdegree ~ nodes$Status,
        main = "Out-Degree by Status",
        xlab = "Status (1 = Partner, 2 = Associate)",
        ylab = "Out-Degree",
        col = c("lightblue", "lightcoral"))
```

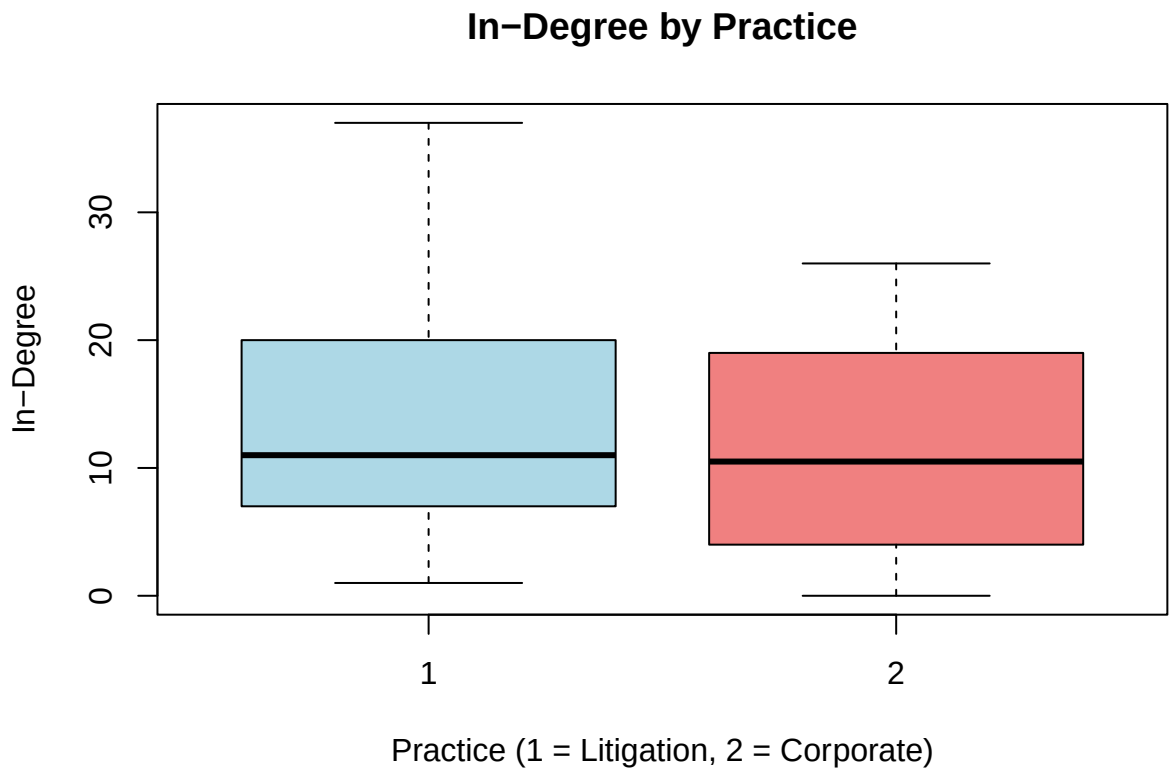



```
mean_indegree_by_status <- tapply(nodes$indegree, nodes$Status, mean)

mean_outdegree_by_status <- tapply(nodes$outdegree, nodes$Status, mean)

df <- data.frame(
  Status = c("Partner (1)", "Associate (2)"),
  Indegree = mean_indegree_by_status,
  Outdegree = mean_outdegree_by_status
)

df
```

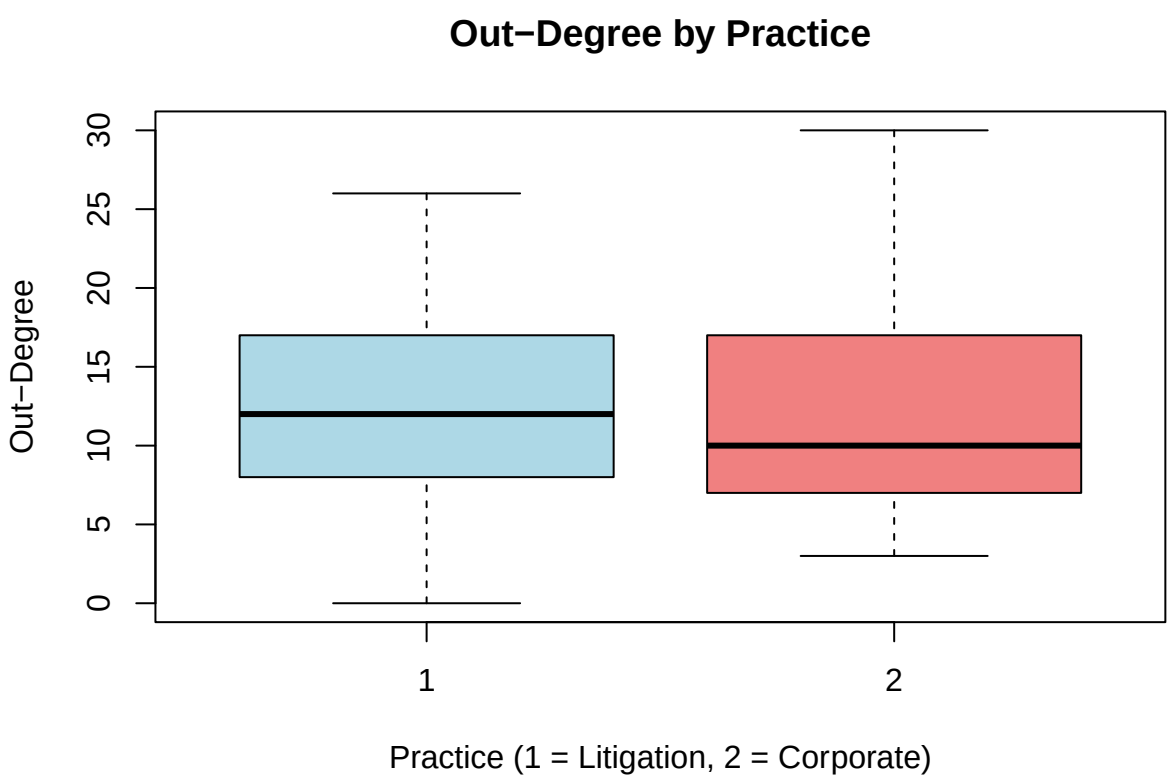
```
##          Status Indegree Outdegree
## 1   Partner (1)    17.69      12.6
## 2 Associate (2)     7.29      12.5
```

```
##
## The boxplots and mean values indicate that partners are generally more central
## in the advice network in terms of being sought for advice (higher in-degrees),
## while both partners and associates are active in seeking advice (similar
## out-degrees). This reflects the hierarchical structure of the firm, where
## partners are seen as key sources of expertise.
```

```r
#5c
#In-degree by practice
boxplot(nodes$indegree ~ factor(nodes$Practice),
        main = "In-Degree by Practice",
        xlab = "Practice (1 = Litigation, 2 = Corporate)",
        ylab = "In-Degree",
        col = c("lightblue", "lightcoral"))
```




```r
#Out-degree by practice
boxplot(nodes$outdegree ~ factor(nodes$Practice),
        main = "Out-Degree by Practice",
        xlab = "Practice (1 = Litigation, 2 = Corporate)",
        ylab = "Out-Degree",
        col = c("lightblue", "lightcoral"))
```




```r
mean_indegree_by_practice <- tapply(nodes$indegree, nodes$Practice, mean)

mean_outdegree_by_practice <- tapply(nodes$outdegree, nodes$Practice, mean)

df <- data.frame(
  Practice = c("Litigation (1)", "Corporate (2)"),
  Indegree = mean_indegree_by_practice,
  Outdegree = mean_outdegree_by_practice
)

df
```

```
##         Practice Indegree Outdegree
## 1 Litigation (1)     13.3      12.8
## 2  Corporate (2)     11.6      12.2
```

```
##
## The analysis suggests that litigation lawyers may be slightly more (13.3
## vs 11.6) central in the advice network in terms of being consulted for advice
## (in-degrees), but both litigation and corporate lawyers are equally active in
## lookout for advice (out-degrees). This could reflect areas where litigation
## often involves more consultation and strategic discussions, whereas corporate
## practice might have a more collaborative structure where advice-giving/seeking
## is spread evenly across colleagues.
```

```r
#6
eigen.cent <- eigen_centrality(graph)
eigen.cent
```

```
## $vector
##  [1] 0.4422 0.5168 0.2371 0.6680 0.2658 0.3304 0.1524 0.2575 0.3171 0.3169
## [11] 0.4039 0.6891 0.7934 0.4916 0.5015 0.7196 0.8010 0.1764 0.6526 0.5987
## [21] 0.5865 0.5627 0.2277 0.8022 0.2952 1.0000 0.5048 0.7293 0.5646 0.5307
## [31] 0.5062 0.4371 0.3392 0.4915 0.4521 0.2739 0.0952 0.4019 0.4706 0.5203
## [41] 0.6566 0.5215 0.4167 0.1080 0.2927 0.2623 0.0707 0.3580 0.2689 0.3047
## [51] 0.3058 0.3787 0.1354 0.3066 0.4503 0.4328 0.3328 0.1947 0.1266 0.2791
## [61] 0.0414 0.1798 0.2027 0.2340 0.4218 0.3792 0.2345 0.2366 0.1585 0.1921
## [71] 0.1871
##
## $value
## [1] 31
##
## $options
## $options$bmat
## [1] "I"
##
## $options$n
## [1] 71
##
## $options$which
## [1] "LA"
##
## $options$nev
## [1] 1
##
## $options$tol
## [1] 0
##
## $options$ncv
## [1] 0
##
## $options$ldv
## [1] 0
##
## $options$ishift
## [1] 1
##
## $options$maxiter
## [1] 3000
##
## $options$nb
## [1] 1
##
## $options$mode
## [1] 1
##
## $options$start
## [1] 1
```

```
##
## $options$sigma
## [1] 0
##
## $options$sigmai
## [1] 0
##
## $options$info
## [1] 0
##
## $options$iter
## [1] 1
##
## $options$nconv
## [1] 1
##
## $options$numop
## [1] 20
##
## $options$numopb
## [1] 0
##
## $options$numreo
## [1] 13
```

```
nodes[, 12] <- eigen.cent$vector
colnames(nodes)[12] <- "eigenvector_centrality"
```

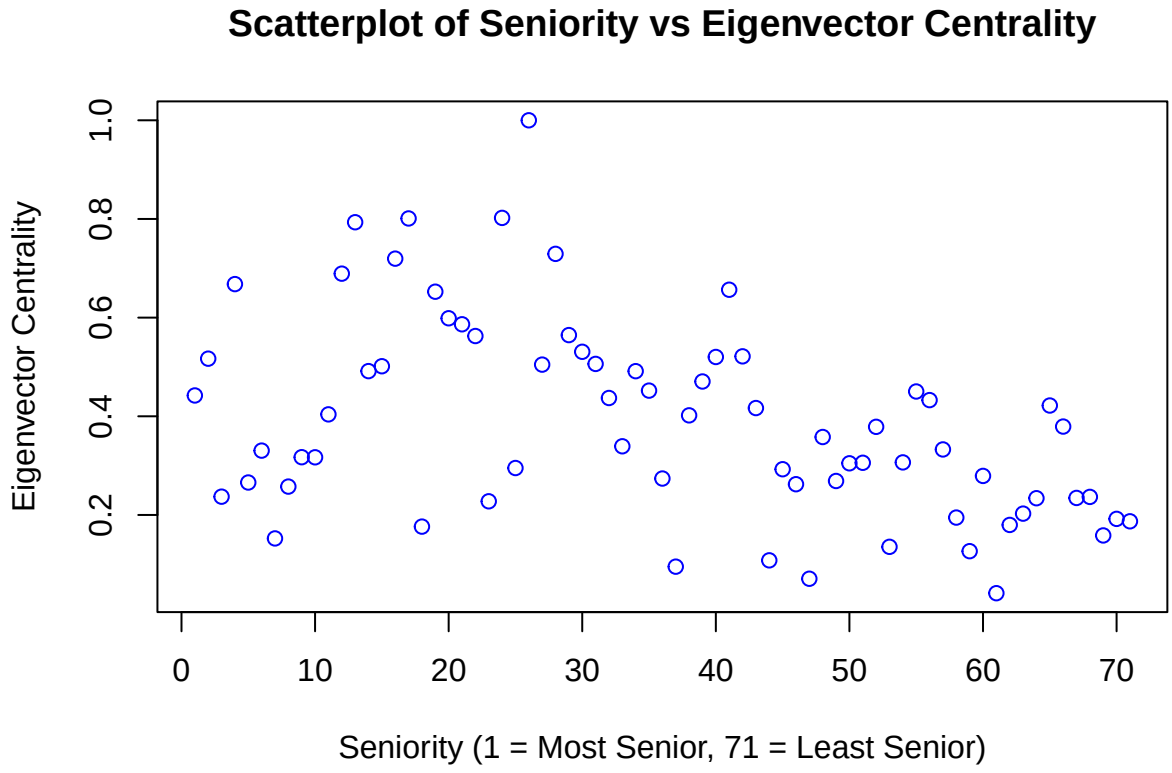
```
#6a
head(nodes, 5)
```

```
##   Name Seniority Status Gender City Years Age Practice LawSchool indegree
## 1   V1         1      1      1    1    31  64        1         1       22
## 2   V2         2      1      1    1    32  62        2         1       23
## 3   V3         3      1      1    2    13  67        1         1        8
## 4   V4         4      1      1    1    31  59        2         3       19
## 5   V5         5      1      1    2    31  59        1         2       17
##   outdegree eigenvector_centrality
## 1         3                 0.442
## 2         7                 0.517
## 3         7                 0.237
## 4        17                 0.668
## 5         4                 0.266
```

```
#6b
max_centrality <- max(nodes$eigenvector_centrality)

most_central_lawyers <- nodes[nodes$eigenvector_centrality == max_centrality, ]
print(most_central_lawyers)
```

```
##    Name Seniority Status Gender City Years Age Practice LawSchool indegree
## 26  V26        26      1      1    1    15  41        1         3       37
##    outdegree eigenvector_centrality
## 26        24                     1
```

```r
#6c
plot(nodes$Seniority, nodes$eigenvector_centrality,
     main = "Scatterplot of Seniority vs Eigenvector Centrality",
     xlab = "Seniority (1 = Most Senior, 71 = Least Senior)",
     ylab = "Eigenvector Centrality",
     col = "blue")
```



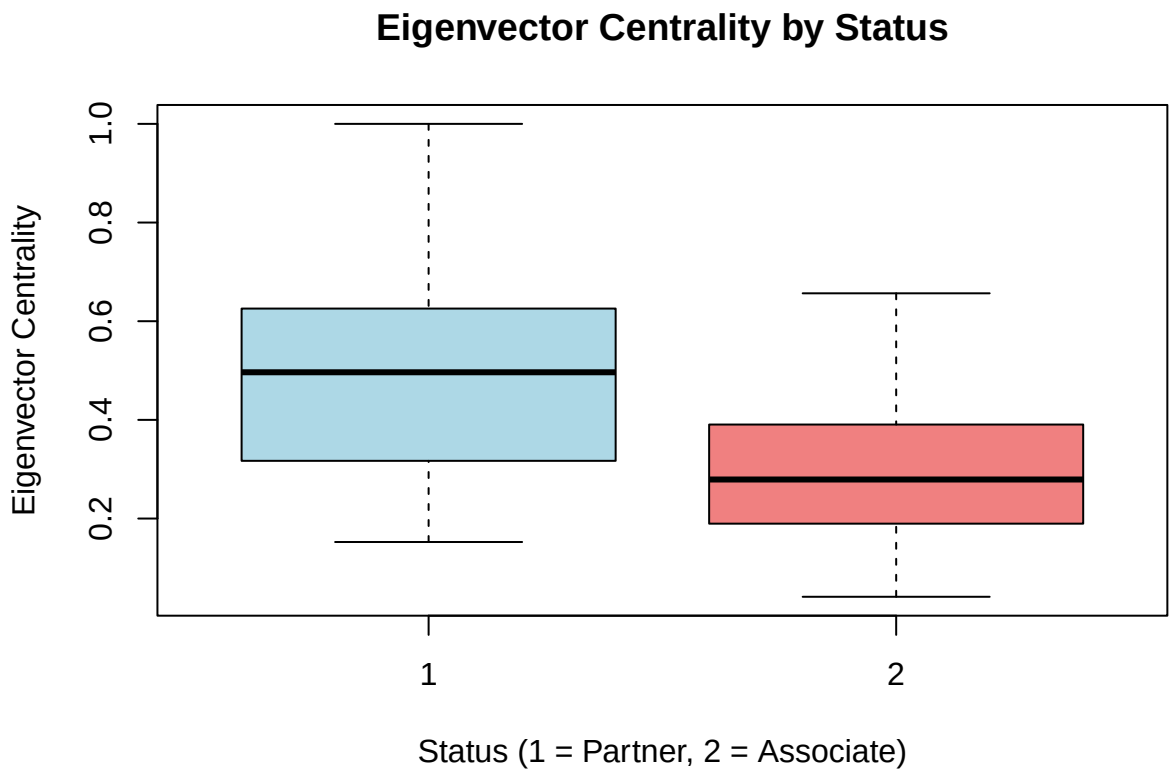

```r
correlation <- cor(nodes$Seniority, nodes$eigenvector_centrality)
correlation
```

```
## [1] -0.451
```

```
##
## The negative correlation (-0.451) suggests that seniority has a substantial
## impact on a lawyer's centrality in the firm's advice network. Senior lawyers,
## due to their experience and authority are more central in the network and are
## more likely to be connected to other highly central lawyers. Junior lawyers,
## on the other hand, are less central. This is typical in hierarchical
## structures like law firms, where seniority often correlates with influence
## and network centrality.
```

```r
#6d
boxplot(nodes$eigenvector_centrality ~ factor(nodes$Status),
        main = "Eigenvector Centrality by Status",
```

```
        xlab = "Status (1 = Partner, 2 = Associate)",
        ylab = "Eigenvector Centrality",
        col = c("lightblue", "lightcoral"))
```

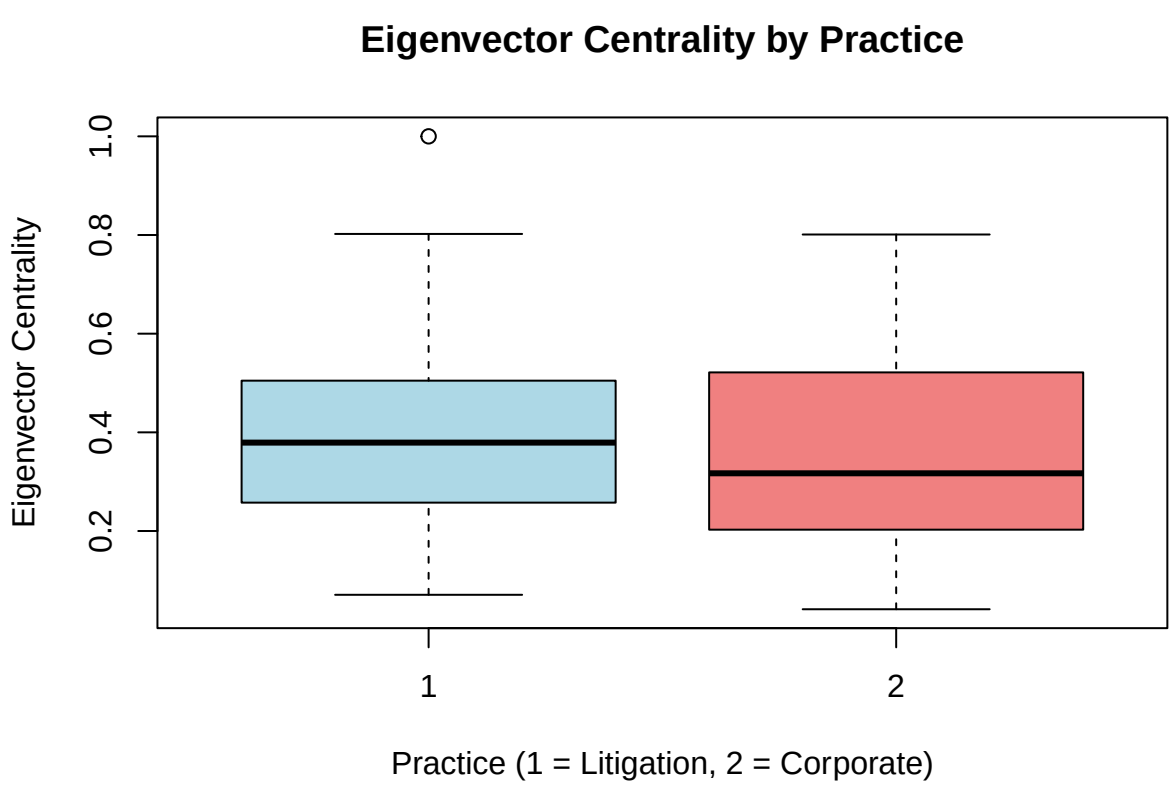



```
mean_eigenvector_by_status <- aggregate(eigenvector_centrality ~ Status, data = nodes, mean)
mean_eigenvector_by_status
```

```
##   Status eigenvector_centrality
## 1      1                   0.49
## 2      2                   0.29
```

```
##
## The plots and computation shows that partners tend to occupy more central and
## powerful positions within the firm's advice network compared to associates.
## This again reflects the nature of law firms, where position often correlates
## with centrality in networks.
```

```
#6e
boxplot(nodes$eigenvector_centrality ~ factor(nodes$Practice),
        main = "Eigenvector Centrality by Practice",
        xlab = "Practice (1 = Litigation, 2 = Corporate)",
        ylab = "Eigenvector Centrality",
        col = c("lightblue", "lightcoral"))
```




```
mean_eigenvector_by_practice <- aggregate(eigenvector_centrality ~ Practice, data = nodes, mean)
mean_eigenvector_by_practice
```

```
##   Practice eigenvector_centrality
## 1        1                 0.396
## 2        2                 0.386
```

```
##
## The scores for litigation and corporate lawyers are quite similar. This
## suggests that both groups are well-integrated into the firm's advice structure,
## with no significant dominance of one practice over the other in terms of
## influence and connectivity within the network. There is presence of an outlier
## in the litigation practice which could reflect the importance of certain
## individuals who are influential within that practice area.
```