Project 2

Logan Bolton

2025-02-08

Setup

```
print("Data structure:")
```

```
## [1] "Data structure:"
```

```
str(D)
```

```
## 'data.frame':    13288 obs. of  1 variable:
##  $ X0.4...weight...0.002105263157894737.: chr  "0 12 {'weight': 0.002105263157894737}" "0 18 {'weigh
```

```
print("First few rows:")
```

```
## [1] "First few rows:"
```

```
head(D)
```

```
##   X0.4...weight...0.002105263157894737.
## 1 0 12 {'weight': 0.002105263157894737}
## 2 0 18 {'weight': 0.002105263157894737}
## 3 0 25 {'weight': 0.004210526315789474}
## 4 0 30 {'weight': 0.002105263157894737}
## 5  0 46 {'weight': 0.00631578947368421}
## 6 0 55 {'weight': 0.002105263157894737}
```

```
# Load the library
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(stringr)
```

```
# Format the data
edges_df <- data.frame(
  from = as.numeric(sub("^(\\d+).*", "\\1", D$X0.4...weight...0.002105263157894737.)),
  to = as.numeric(sub("^\\d+\\s+(\\d+).*", "\\1", D$X0.4...weight...0.002105263157894737.)),
  weight = as.numeric(sub(".*'weight':\\s*([0-9.]+).*", "\\1", D$X0.4...weight...0.002105263157894737.))
)
```

```r
# Create the graph
g <- graph_from_data_frame(edges_df, directed = TRUE)
```

Graph Characteristics

Network Understanding

```r
print("Network Order (number of vertices):")
```

```
## [1] "Network Order (number of vertices):"
```

```r
vcount(g)
```

```
## [1] 475
```

```r
# Network size (number of edges)
print("Network Size (number of edges):")
```

```
## [1] "Network Size (number of edges):"
```

```r
ecount(g)
```

```
## [1] 13288
```

```r
# Network density
print("Network Density:")
```

```
## [1] "Network Density:"
```

```r
edge_density(g)
```

```
## [1] 0.05901843
```

```r
# Check strong connectivity
components <- components(g, mode="strong")
cat("\nNumber of strongly connected components:", components$no, "\n")
```

```
##
## Number of strongly connected components: 7
```

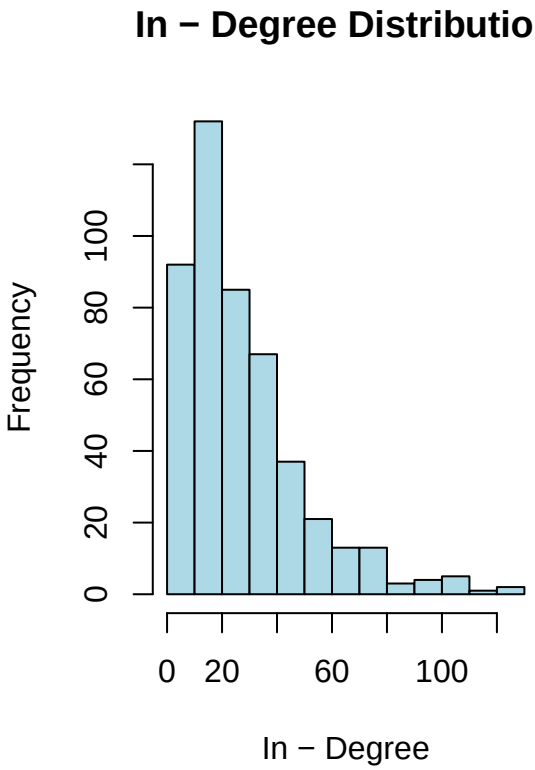
```r
cat("Size of largest strongly connected component:", max(components$csize), "\n")
```

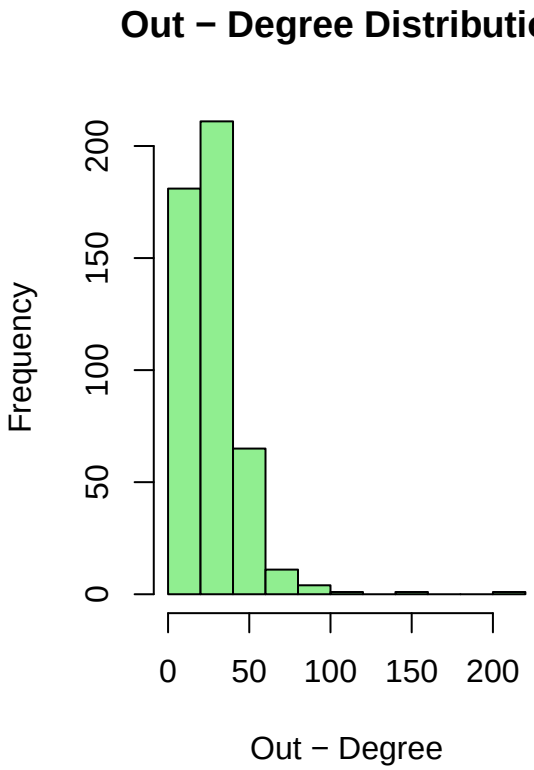
```
## Size of largest strongly connected component: 469
```

Degree Distribution

```r
# Calculate different degree measures
in_deg <- degree(g, mode="in")
out_deg <- degree(g, mode="out")
total_deg <- degree(g, mode="total")

par ( mfrow = c (1 ,2))
hist ( in_deg , main = " In - Degree Distribution " ,
xlab = " In - Degree " , ylab = " Frequency " , col = " lightblue " )
hist ( out_deg , main = " Out - Degree Distribution " ,
xlab = " Out - Degree " , ylab = " Frequency " , col = " lightgreen " )
```





PageRank

Hub and Authority Scores

Closeness Centrality

Betweenness Centrality

Nodes

Justification