Network Visualization

Lihua Deng

August 26, 2019

Group Stratification

```
kable(df,caption= 'Group Size Stratification',
      align = "c") %>%
  kable_styling(c("striped", "bordered"),latex_options = "hold_position",
                full_width = F, fixed_thead = T) %>%
   group_rows("Member", 1,2,
              label_row_css = "background-color: #666; color: #fff;")%>%
  group_rows("Group Pool", 3,4,
             label_row_css = "background-color: #666; color: #fff;")%>%
  group_rows("Photos per Member", 5,6,
             label_row_css = "background-color: #666; color: #fff;")
```

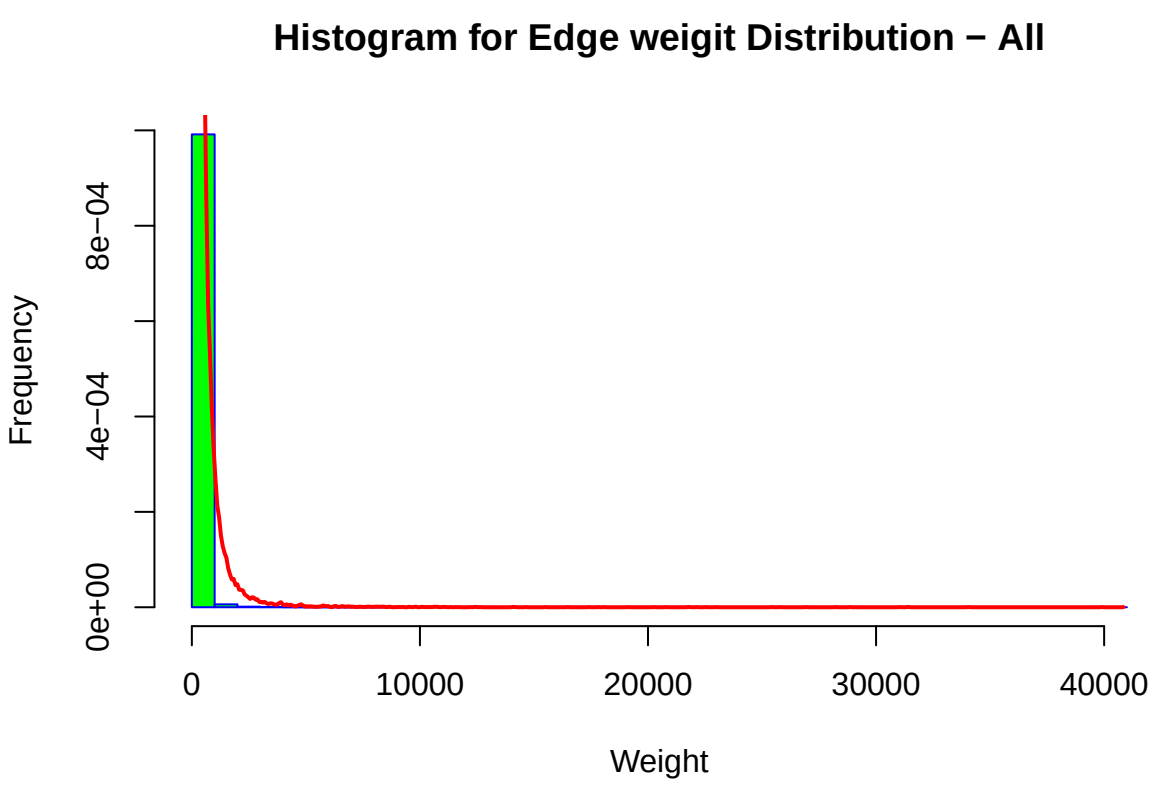
Table 1: Group Size Stratification

Size	Zero	Small	Medium	Large	X-Large
Member					
Range	0	1-30	31-300	301-5,000	> 5,000
Count of Groups	7	1958	1299	1083	152
Group Pool					
Range	0	1-1,000	1,001-10,000	10,001-100,000	> 100,000
Count of Groups	354	2002	1066	822	255
Photos per Member					
Range	0	1-10	11-30	31-100	> 100
Count of Groups	351	1358	1451	1175	155

Edge Weight Distribution

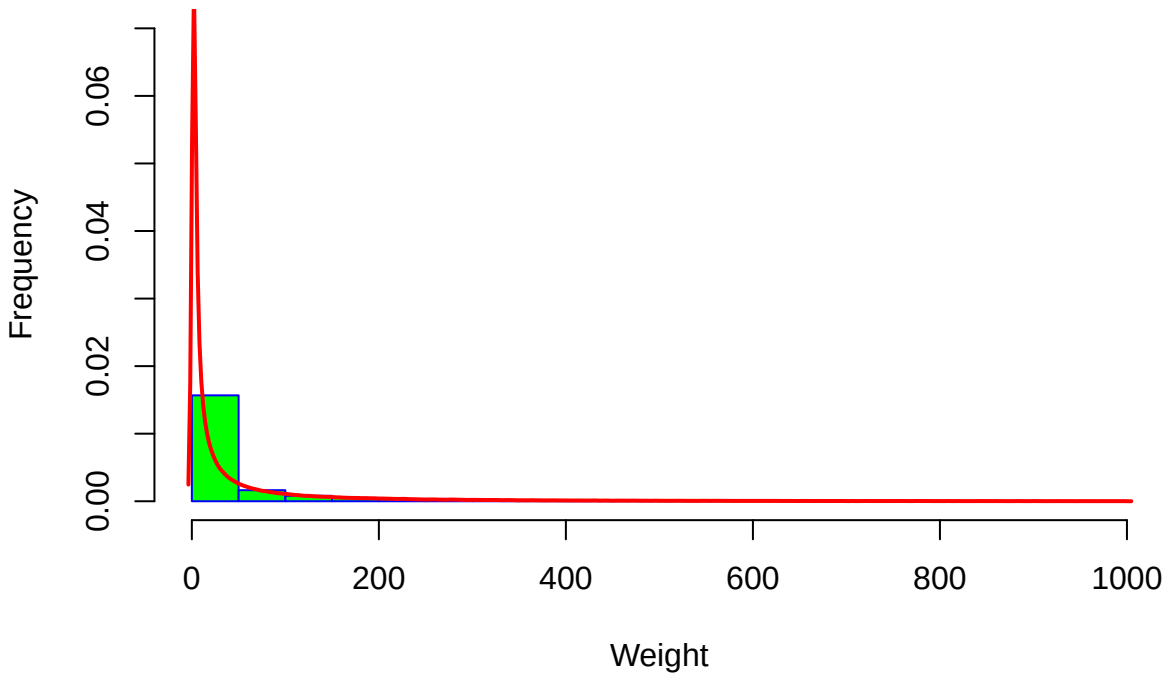
```
#Edge weights
ed_weig = read.csv("C:\\Users\\lunad\\Desktop\\flickr\\co-mem1.csv", sep = ",")
hist(ed_weig$number, breaks=40,
main="Histogram for Edge weigit Distribution - All",
     xlab="Weight", ylab="Frequency",
     border="blue",
     col="green",
prob = TRUE)

lines(density(ed_weig$number), lwd = 2, col = 'red')
```



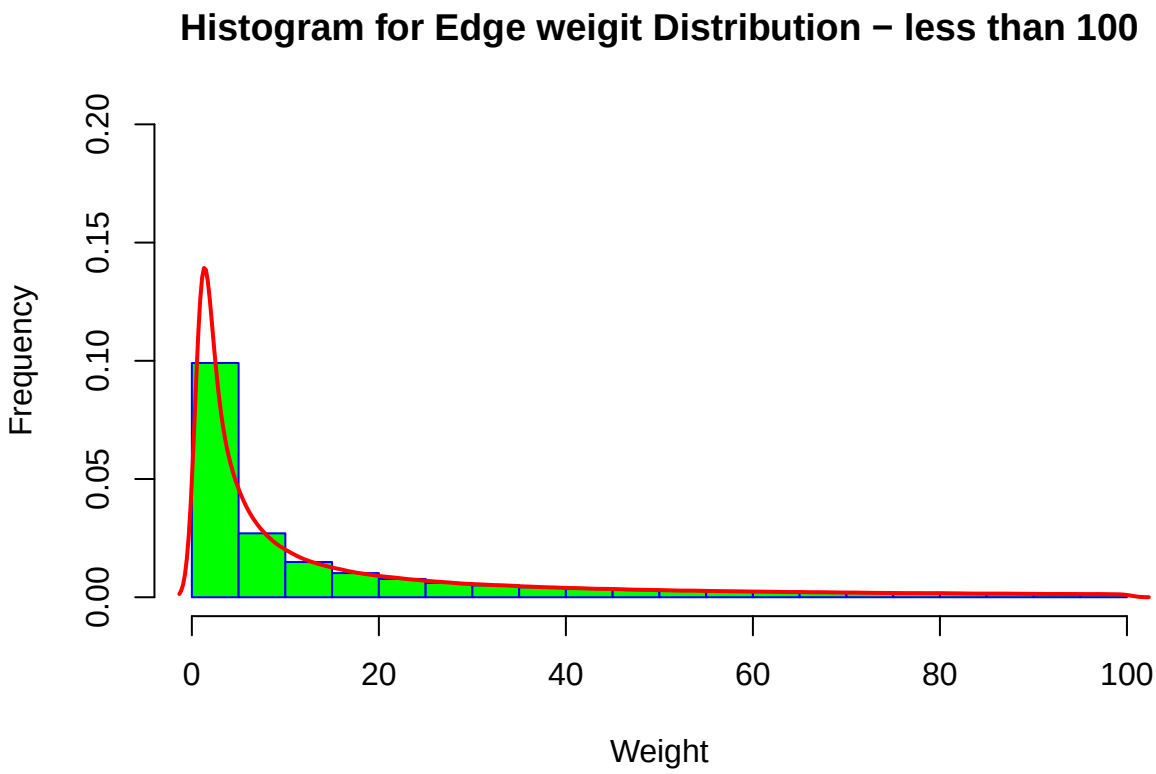

```
#Edge weight under 1000
subed_weig1 = read.csv("C:\\Users\\lunad\\Desktop\\flickr\\1000.csv", sep = ",")
hist(subed_weig1$number, ylim=c(0,0.07),
main="Histogram for Edge weigit Distribution - less than 1000",
     xlab="Weight", ylab="Frequency",
     border="blue",
     col="green",
prob = TRUE)

lines(density(subed_weig1$number), lwd = 2, col = 'red')
```




```
#Edge weight under 100
subed_weig2 = read.csv("C:\\Users\\lunad\\Desktop\\100.csv", sep = ",")
hist(subed_weig2$number, xlim=c(0,100), ylim=c(0,0.2), breaks = 25,
main="Histogram for Edge weigit Distribution - less than 100",
     xlab="Weight", ylab="Frequency",
     border="blue",
     col="green",
prob = TRUE)

lines(density(subed_weig2$number), lwd = 2, col = 'red')
```



Network Visualization

Filter edges with weight greater than 100

```
filter_mem <- as.matrix(filter_mem)
net_m <- network(filter_mem,directed=FALSE) # Create a network object
Isolated1 = which(degree(net_m)==0)      #Delete isolated nodes
net_m1 = delete.vertices(net_m, Isolated1)
net_m1  # Get a quick description of the data
```

```
##  Network attributes:
##   vertices = 1311
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 96994
##     missing edges= 0
##     non-missing edges= 96994
##
##  Vertex attribute names:
##     vertex.names
##
```

```
##  Edge attribute names not shown
```

```
#visualization
#gplot(net_m1, gmode='graph', label = NULL, usearrows = FALSE,
#vertex.cex = 1, vertex.col='red', edge.col='black')
```

Filter edges with weight greater than 500

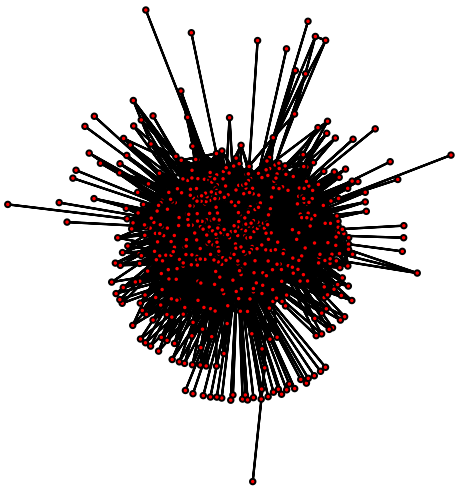
```
filter_mem4 <- as.matrix(filter_mem4)
net_m5 <- network(filter_mem4,directed=FALSE) # Create a network object
net_m5 # Get a quick description of the data
```

```
##  Network attributes:
##   vertices = 3813
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 13861
##     missing edges= 0
##     non-missing edges= 13861
##
##  Vertex attribute names:
##     vertex.names
##
##  Edge attribute names not shown
```

```
#remove isolated nodes
Isolated5 = which(degree(net_m5)==0)
new_net_m5 = delete.vertices(net_m5, Isolated5)
new_net_m5
```

```
##  Network attributes:
##   vertices = 570
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 13861
##     missing edges= 0
##     non-missing edges= 13861
##
##  Vertex attribute names:
##     vertex.names
##
##  Edge attribute names not shown
```

```
#visualization
par(mfrow=c(1,1))
g1 <- gplot(new_net_m5, gmode='graph', usearrows = FALSE, vertex.cex = 1,
            vertex.col='red', edge.col='black', main='Edge Weight > 500')
```



Filter edges with weight greater than 1000

```
filter_mem1 <- as.matrix(filter_mem1)
ncol(filter_mem1)
```

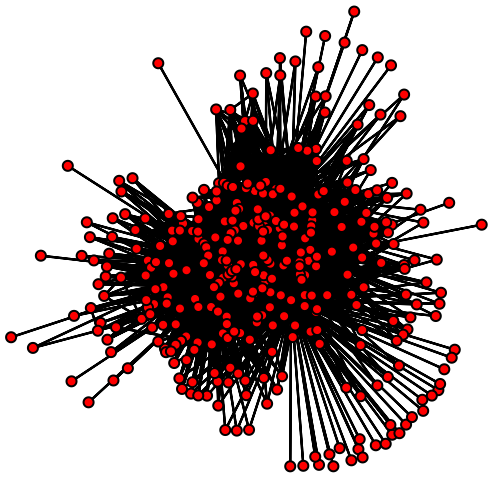
```
## [1] 3813
```

```
net_m2 <- network(filter_mem1,directed=FALSE) # Create a network object
#remove isolated nodes
Isolated2 = which(degree(net_m2)==0)
net_m2 = delete.vertices(net_m2, Isolated2)
net_m2 # Get a quick description of the data
```

```
##  Network attributes:
##   vertices = 347
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 4398
##     missing edges= 0
##     non-missing edges= 4398
##
##  Vertex attribute names:
```

```
##     vertex.names
##
##  Edge attribute names not shown
```

```
#visualization
par(mfrow=c(1,1))
g2 <- gplot(net_m2, gmode='graph', label = NULL, usearrows = FALSE,
            vertex.cex = 1.5, vertex.col='red', edge.col='black',
            main='Edge Weight > 1000')
```



Filter edges with weight greater than 5000

```
filter_mem3 <- as.matrix(filter_mem3)
ncol(filter_mem3)
```

```
## [1] 3813
```

```
net_m4 <- network(filter_mem3,directed=FALSE) # Create a network object
net_m4 # Get a quick description of the data
```

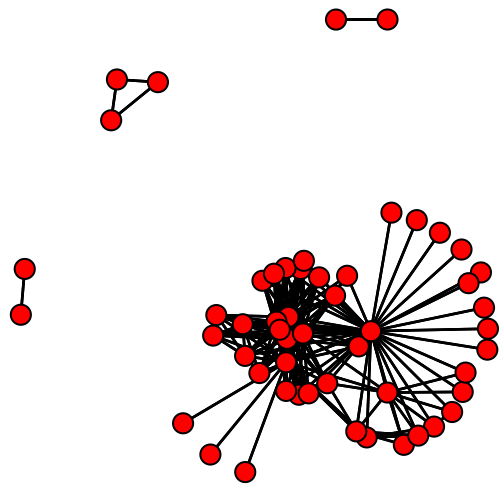
```
##  Network attributes:
##   vertices = 3813
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
```

```
##   bipartite = FALSE
##   total edges= 195
##     missing edges= 0
##     non-missing edges= 195
##
##  Vertex attribute names:
##     vertex.names
##
## No edge attributes
```

```
#remove isolated nodes
Isolated4 = which(degree(net_m4)==0)
new_net_m4 = delete.vertices(net_m4, Isolated4)
new_net_m4
```

```
##  Network attributes:
##   vertices = 53
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 195
##     missing edges= 0
##     non-missing edges= 195
##
##  Vertex attribute names:
##     vertex.names
##
## No edge attributes
```

```
#visualization
par(mfrow=c(1,1))
g3 <- gplot(new_net_m4, gmode='graph', usearrows = FALSE,
            vertex.cex = 1.5, vertex.col='red', edge.col='black',
            main='Edge Weight > 5000')
```



Filter edges with weight greater than 10000

```
filter_mem2 <- as.matrix(filter_mem2)
net_m3 <- network(filter_mem2,directed=FALSE) # Create a network object
net_m3 # Get a quick description of the data
```

```
##  Network attributes:
##   vertices = 3813
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 53
##     missing edges= 0
##     non-missing edges= 53
##
##  Vertex attribute names:
##     vertex.names
##
## No edge attributes
```

```
#remove isolated nodes
Isolated3 = which(degree(net_m3)==0)
new_net_m3 = delete.vertices(net_m3, Isolated3)
```

```
new_net_m3 # Get a quick description of the data
```

```
##  Network attributes:
##   vertices = 23
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 53
##     missing edges= 0
##     non-missing edges= 53
##
##  Vertex attribute names:
##     vertex.names
##
## No edge attributes
```

```
#visualization
par(mfrow=c(1,1))
g4 <- gplot(new_net_m3, gmode='graph', usearrows = FALSE,
            vertex.cex = 1.5, vertex.col='red', edge.col='black',
            main='Edge Weight > 10000')
```