Experiment No.7
Social Network Analysis using R (for example: Community Detection Algorithm)
Date of Performance:
Date of Submission:

Aim: Social Network Analysis using R (for example: Community Detection Algorithm)

Theory:

Online social platforms have enabled people around the world to interact with each other and build relationships with others they share common interests with. This can be observed in real life — naturally, we tend to develop and maintain relationships with others that are similar to us. People with similar interests tend to gravitate towards each other and become associated in communities — clusters or groups of people that share similar traits with each other. Since people tend to cluster with others similar to them, we can use community detection to identify users with a high number of degrees (connections) and see how far their reach can travel in the network.

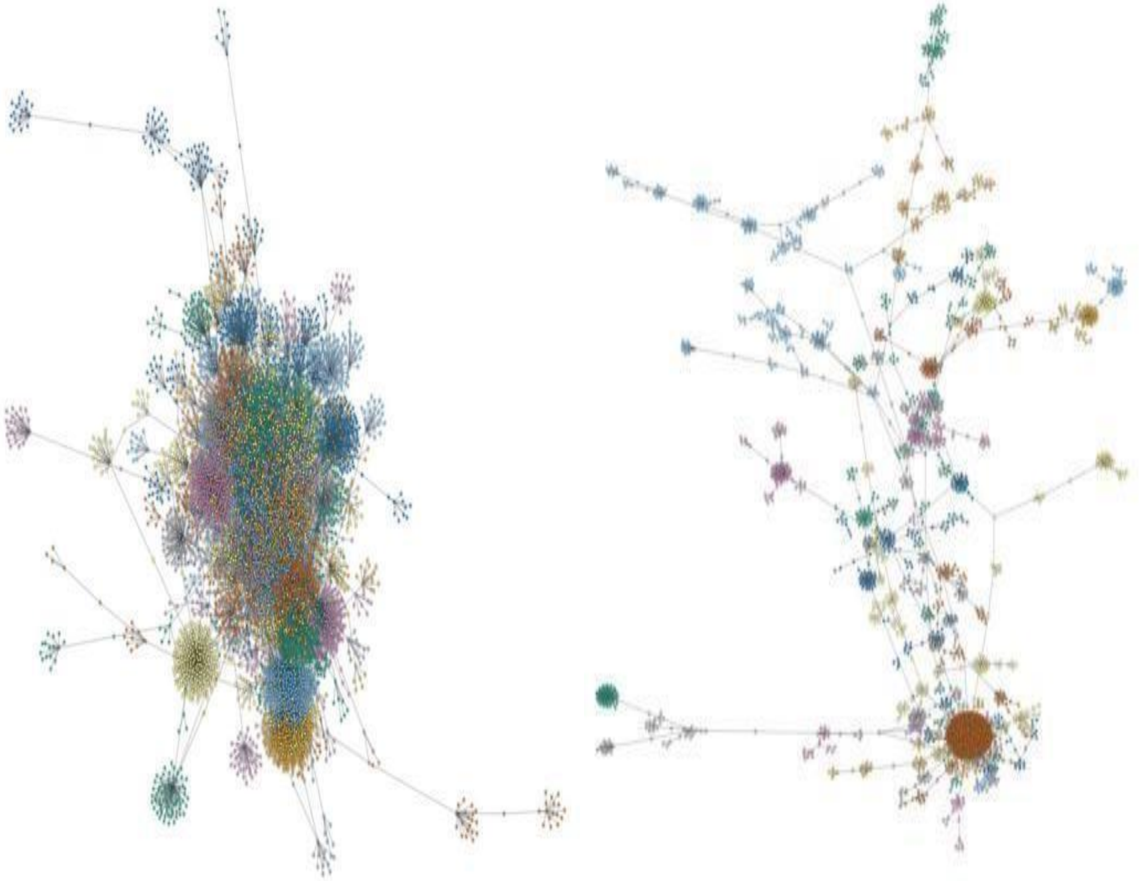
User Data Extraction — Since we are only interested in user data, we will only extract the following variables:

User_id — Yelp user ID; this is needed to make nodes and edges
Name — user's first name
Review count — the number of reviews user has written
Yelping since — date user joined Yelp
Friends — a list containing all of the user's friends by user_id
Fans — number of fans user has
Elite — number of years the user has Elite status
Average stars — user's average rating of all reviews written

CODE:

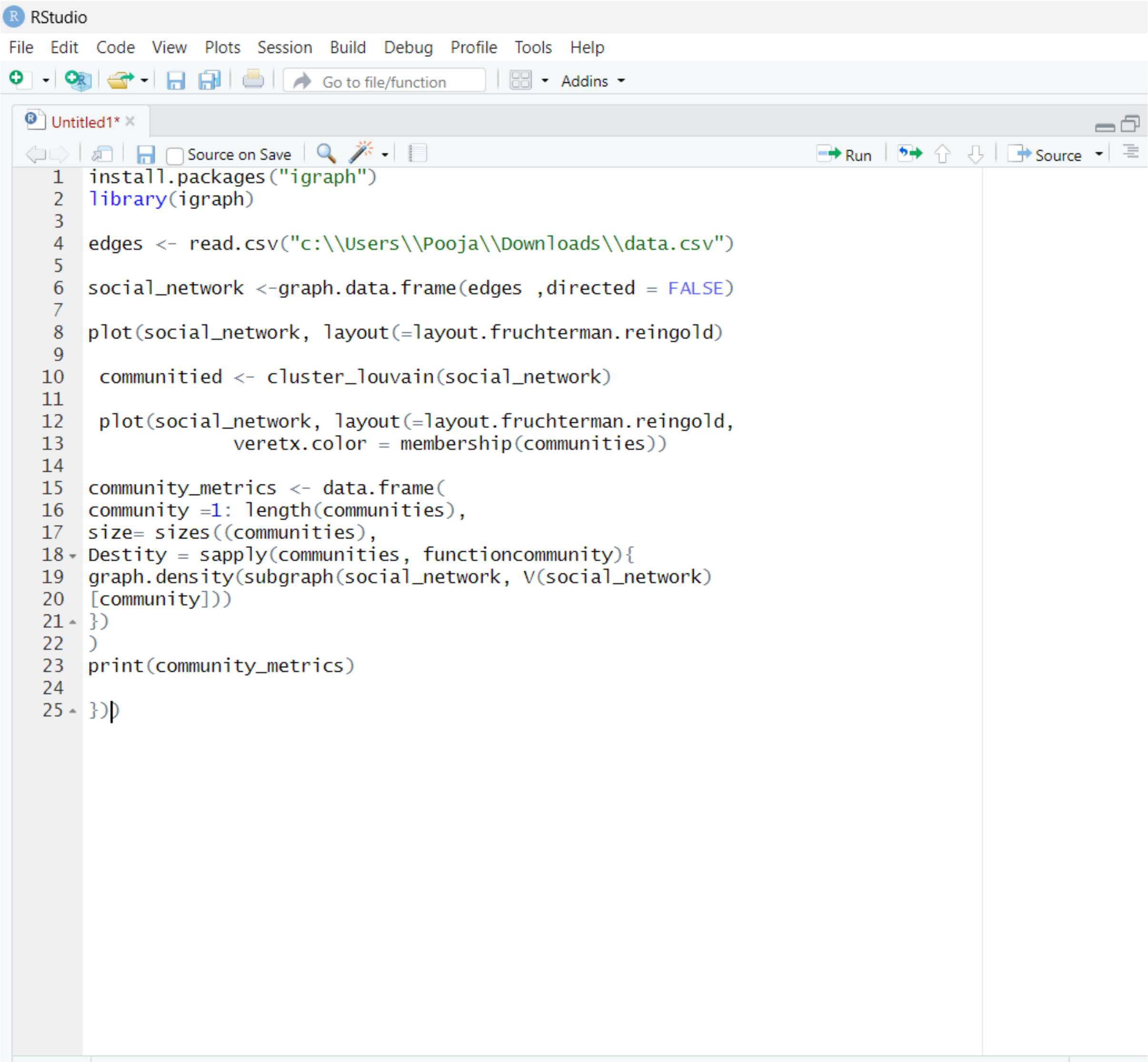
```
 #remove users with no friends

sample <- subset(user_df, friends != "None")

#make a subset; we only need to retain data of users with some social
activity sub <- subset(sample, year == 2005 & review_count >= 2 &
no_of_friends >= 2)
```

```
#make links (nodes and edges)
sample_friends <- sub %>%
select(user_id, friends) sample_users <-
strsplit(sample_friends$friends, split =,")
sample_dat<-data.frame(user_id=rep(sample_friends$user_id,

sapply(sample_users, length)), friends = unlist(sample_users))

#network is still too big, take a random sample of
100k nodes samp_net <- sample_n(sample_dat,
100000)
#make network

network <- graph.data.frame(samp_net) network_s <- simplify(network)
net_deg <- degree(network_s) all_degree <- degree(network, mode = 'all')
#graph user with max degrees sub_all <- subcomponent(network_s,
which(all_degree == max(all_degree)), 'all') g_sub <-
induced_subgraph(network_s, sub_all)

#communities

graph.com <- fastgreedy.community(as.undirected(g_sub))

V(g_sub)$color <- graph.com$membership + 1

#create pdf graph for high resolution (try zooming in!)
pdf("communities2005.pdf", 10,10)
plot(g_sub, vertex.color = V(g_sub)$color, vertex.size = 1,
vertex.label = NA, vertex.frame.color = adjustcolor("#41424c",
alpha.f = 0.25), edge.arrow.size = 0.1, edge.color =
adjustcolor("#41424c", alpha.f = 0.20), edge.width = 1.5,
edge.arrow.mode=0, layout=layout_with_lgl, asp = 0.9, dpi=300
)
dev.off()
```



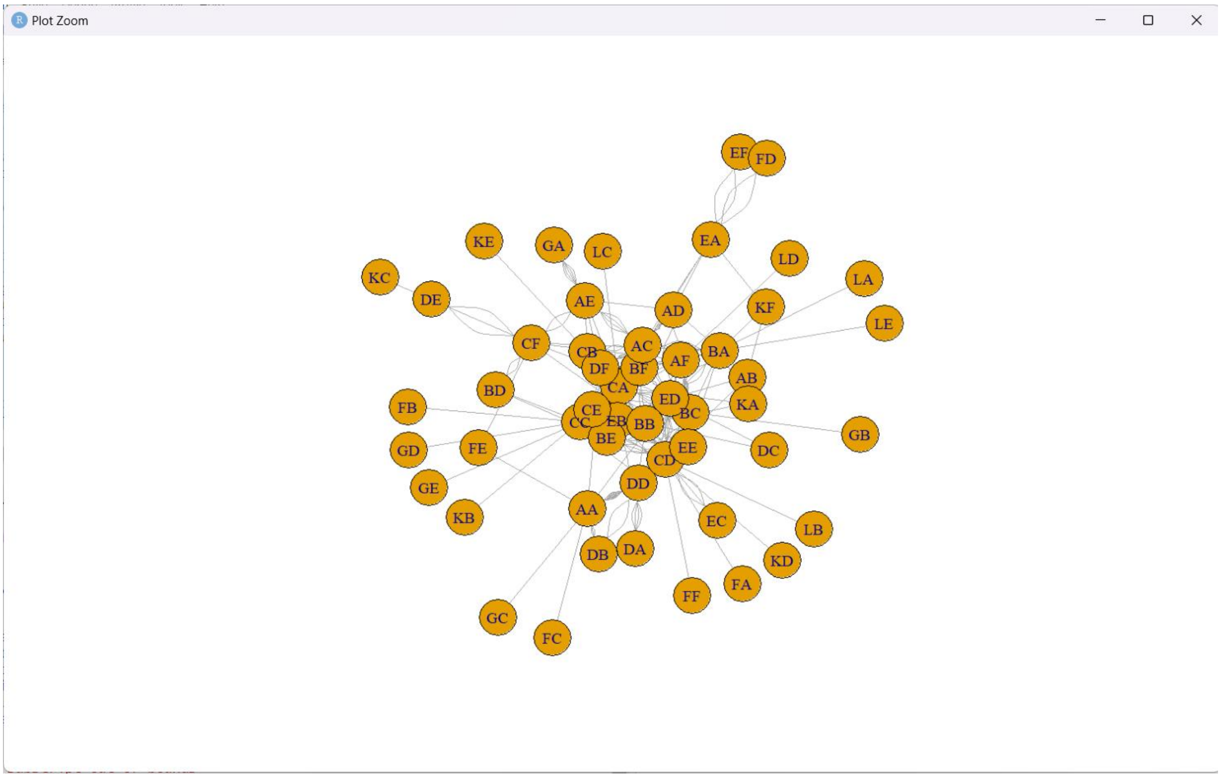
Code:

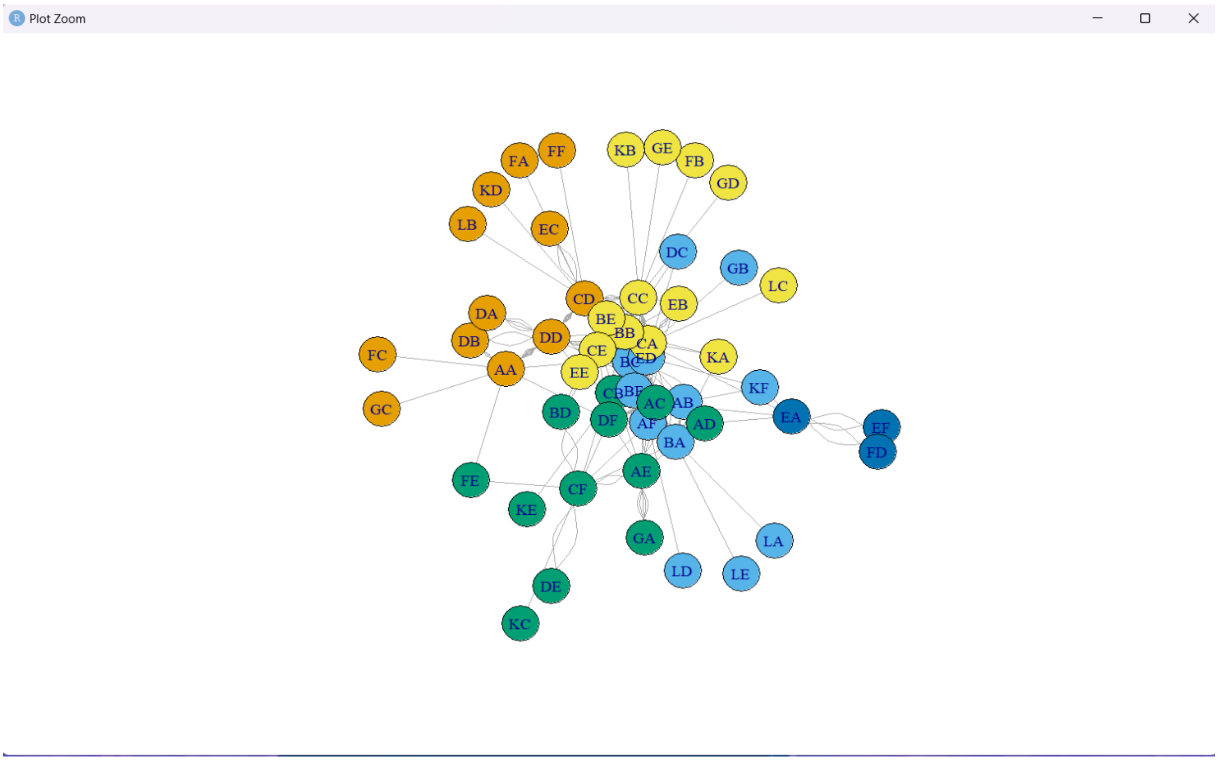
"



```
install.packages("igraph")
library(igraph)

edges <- read.csv("c:\\Users\\Pooja\\Downloads\\data.csv")

social_network <-graph.data.frame(edges ,directed = FALSE)

plot(social_network, layout(=layout.fruchterman.reingold)

 communitied <- cluster_louvain(social_network)

 plot(social_network, layout(=layout.fruchterman.reingold,
             veretx.color = membership(communities))

community_metrics <- data.frame(
community =1: length(communities),
size= sizes((communities),
Destity = sapply(communities, functioncommunity){
graph.density(subgraph(social_network, V(social_network)
[community]))
})
)
print(community_metrics)

})
```

Output:


Plot Zoom
EF
FD
KE
GA
LC
EA
LD
LA
KC
DE
AE
AD
KF
LE
CF
CB
AC
AF
BA
DF
BF
AB
CA
BD
ED
BC
KA
FB
CE
CC
EB
BB
BE
GB
GD
FE
EE
DC
CD
GE
DD
KB
AA
EC
LB
DB
DA
KD
FA
FF
GC
FC



Conclusion:

An technique to understanding the structure and relationships inside a social network is to use R and community detection algorithms for social network analysis (SNA). By exposing patterns of linkage that might not be obvious at first look, community detection assists in locating hidden subgroups within a network.It offers information on how the network is organized and structured, including which nodes communicate with one another most frequently. The structure of a social network can be better understood by researchers and analysts through the visualization of its communities. Usually, nodes in the same community have similar colors to make them easy to identify.

We successfully use the community detection approach to examine social networks in this experiment.