Module 4: Analytics RStudio

Goals for Module 4: Analytics RStudio

~This lab introduces participants to the RStudio commands in igraph that produce the social network analytics from this module. Participants will prepare networks for the appropriate analytics, conduct analytics, and incorporate analytics into network visualization.

Lab 4 Files

The `Lab 4 Files` folder contains all of the files needed to complete the Module 4: Analytics RStudio Lab. There are three files in this folder: two csv data files and one script file.

The three files included in the `Lab 4 Files` folder are:

`edgelist.csv`

`attributes.csv`

`analytics.R`

If you need a refresher of what these files look like outside of RStudio, you can view the csv files in either a spreadsheet software program or a simple text editing program. You can view the `analytics.R` script file in a text editing program.

Analytics.R in RStudio

A. Open RStudio on your computer.

B. Set the working directory to the `Lab 4 Files` folder.

1. Select the Session menu located toward the middle of the menu bar at the very top of the RStudio program.
2. Select Set Working Directory.
3. Select Choose Directory.
4. Locate the `Lab 4 Files` folder on your computer.
5. Click open.

C. Open the `analytics.R` script.

1. In RStudio click on the `Open` icon shaped like a folder located in the upper-left pane of the RStudio window
2. Check that the folder is `Lab 4 Files`
3. Select the `analytics.R` script
4. The `analytics.R` script should open in the top-left pane. It will look like a text file with the file name at the top
5. If the text in the script is running off the pane and requires a left to right scroll bar, select the Tools menu at the top of the RStudio window, select Global Options from Tools, select the Code Editing option from the left menu, check the box for "Soft wrap R source files," and then click Apply and OK. This should set text wrap as your default in RStudio.

D. Complete Module 4: Analytics RStudio Lab by following the `analytics.R` script in RStudio. Once you have completed the script return to this document for the final review section of this Lab. Tasks in the RStudio portion of Module 4: Analytics Lab include:

Setting up RStudio Session
Importing Edgelist and Attribute Data
Setting Up the Social Network
Density
Components
Degree
Cut-points
K-core
Distance
Brokerage/ Betweenness
Neighborhoods
Compiling Network Statistics
Review

Review of Module 4: Analytics RStudio

This lab introduced participants to the RStudio commands in igraph that produce the social network analytics from this module. Participants reviewed loading edgelist and attribute data and converting the data into a network. Participants worked through a series of social network analytics using igraph commands: density, components, degree, cut-points, k-core, distances, betweenness, and neighborhoods. Participants altered the networks as necessary for proper calculations. Participants incorporated some of the analytics into network visualizations. Participants compiled tables of the various work from the lab and exported those tables as csv files that are compatible with software other than RStudio.

Below is a briefly annotated list of RStudio commands used in the Module 4: Analytics RStudio Lab:

```
dir()                                   # Check the contents of the working directory
library(igraph)                         # Load the igraph library
e<-read.csv("edgelist.csv",header=TRUE)     # Load the edgelist.csv file
e                                       # Inspect object "e"
a<-read.csv("attributes.csv",header=TRUE)  # Load the attribute file
a                                       # Inspect object "a"
g<-graph.data.frame(e,directed=FALSE,vertices=a) # Make a network named "g"

graph.density(g)                        # Calculate the density of network "g"
clusters(g)$no                          # Calculate the number of components in network "g"
clusters(g)$csize                       # Calculate the size of each component in network "g"
clusters(g)$membership                  # Generate a list of component membership for all nodes in "g"
degree(g)                               # Generate a list of each node's degree score for network "g"
articulation.points(g)                  # Identify which nodes are the cut-points in network "g"
graph.coreness(g)                       # Generate a list of each node's k-core in network "g"

shortest.paths(g4)                      # Generate a matrix of geodesic distances within a component
betweenness(g4)                         # Generate a list of each node's betweenness score within a component

neighborhood.size(g, 2, 15)             # Calculate the number of nodes in node #15's 2nd order neighborhood
neighborhood(g, 2, 15)                  # Identify the nodes in node #15's 2nd order neighborhood
```

That completes Module 4: Analytics. Participants should move onto the final module when they are ready. Module 5: Direction explores data, visualization, and analytics for networks with directed ties.