Network Analytics
Project: Shiny App

INSTRUCTIONS

- This is a group assignment. Each team must have at most 5 students.

  - **One element of each group should communicate the group composition via email to `44067@novasbe.pt` until February 27.**

- Submit your answer digitally as a single **zip file** named `NA_Project_Group_XX.zip` through Moodle. See detailed instructions on how to create this file below.

- Follow the Style Guide (available on Moodle)

- Questions regarding the assignment should be posted exclusively on the respective discussion forum on Moodle.

Warning: The detection of any form of plagiarism in your work means the assignment will be graded with ZERO points.

Project: Shiny App

For this project you will apply knowledge acquired during this class to create a web application that explores a real-world network data set. Your group will choose the data set and main question to answer, and will develop the web application using the Shiny framework (`http://shiny.rstudio.com`). You can choose one of the data sets suggested in class or any other data set. In case you have difficulties in finding a suitable data set please contact Rodrigo.

App Structure

The app should be appealing to the user and should reveal interesting information from the data set. First, the app should provide some basic plots and <u>descriptive statistics</u> on the data. Second the app should perform <u>some exploratory analyses on the network</u>, preferably taking user input into account by filtering data accordingly. Third the app should perform at least one more <u>elaborate analysis</u>, either one of the analyses covered in class or another analysis you find interesting.

The app should provide adequate (dynamic) textual descriptions of the results. Imagine you present your shiny app to a manager in a company or in a government institution interested in your analysis. He/she did not take this elective but will eventually give the yes/no decision to buy your app or not. Describe what the graphs and analysis are saying about the networks and how he/she can read the numeric results. This should be done in the app and in the shortest amount of lines possible.

You can find more details on the three main components of the app in the following sections.

Descriptive Statistics

The app should show some basic descriptive statistics on your data before looking explicitly at network-related characteristics. The app should report available variables, number of observations, number of entities (e.g., students, employees, companies), averages, standard deviations, etc.

Network Exploration

The app should also perform some exploratory analyses on your network, such as:

1. Calculations of basic descriptive statistics: average path length, average clustering coefficient, diameter

2. The degree distribution, plotted as a histogram

3. Several different visualizations of all or part of the network. For each visualization, say what features of the network you are highlighting or emphasizing.

Network Analysis

The app should perform at least one more elaborate analysis. For example, you can answer a question similar to the ones below.

1. What is the relation between the different centrality measures and other node characteristics?

2. Are two nodes more likely to be connected if they have the same characteristics (homophily)?

3. Which links do you predict will be formed in the future?

Deliverables

Your project will be comprised of two deliverables for grading:

1. A working app hosted at `http://shinyapps.io`. Your app will be tested using this version. You will learn how to deploy apps to `http://shinyapps.io` in the session of week 4.

2. All the code required to run the app, distributed by the following files:

   setup.R not strictly part of the app, includes all the code used to transform the data from the original source into the format used by the app. The last instruction of this file should be the saving of the prepared data set to be loaded by the app.

   global.R includes code required to run but that does not belong to the user interface or the server. This file should contain at least the instruction that loads the (prepared) data set into memory.

   app.R includes both the user interface and server side of the app. Alternatively you can use:

   ui.R includes the user interface of the app.

   server.R includes the server side of the app.

   You need to create a **zip file** containing these four files and upload it to Moodle.

Grading Criteria

Your project will be graded by the criteria in the table below.

Completeness	50%
- Descriptive Statistics: Does the app provide basic descriptives on the data (static analyses)?	10%
- Network Exploration: Does the app provide exploratory analyses on your network (dynamic analyses)?	20%
- Network Analysis: Does the app provide at least one more elaborate analysis on the network (dynamic analyses)?	20%

Readability	20%
- Does the code comply with the style guide?	10%
- Can the user easily understand each component of the app (graphs, inputs and results)?	10%

Business / Societal Relevance	20%
- Does the app present results that are relevant from a business/ societal point of view?	10%
- Does the app provide a textual interpretation of the results with business/societal relevancy?	10%

Robustness	10%
- Does the app behave correctly in all situations, including displaying the right messages when users provide unexpected inputs?	10%