

Lab #9: Text Networks, Pt. I

Due in Sakai April 5, 2019 (by 10:00a)

Introduction

For this lab, you will create a text network. You will then analyze the formal properties of this network in lab #10.

Getting Started

First, set your working directory. Open up a fresh RStudio session (if you open RStudio and a previous work session is loaded, be sure you save any R and/or RData files before you close them out). Then set your working directory:

```
setwd("working_directory_here")
```

Then open up a fresh R script. Save it using your first initial, full last name, and then “_lab9.” So my script would be titled **MTaylor_lab9.R**.

You should always strive to keep your scripts tidy. At the top of your R script, type this (substituting in your first initial and last name):

```
#####  
## MTaylor_lab9.R  
## Note: Code for lab assignment #9  
## Author: Marshall A. Taylor  
#####  
  
###BEGIN###
```

You are ready to begin your code. Be sure to include all the code necessary for me to check your work. Document your code thoroughly (using “#”).

Remember that saving your R work is a two-step process. You save your R script using Cmd+Enter (Mac) or Cntrl+Enter (Windows). You save your R data objects like this:

```
save.image("MTaylor_lab9.RData")
```

Lastly, prep a Word, Pages, or L^AT_EX document that has the same title structure: e.g., **MTaylor_lab9.docx**. This is where you put your write-ups and visualizations.

You should turn in three documents to Sakai: your **R script** that shows the code you used, **RData file** that provides the data your scraped, and **Word document** (or whatever text processor you choose to use) with your write-ups.

Assignment

For this lab, we are going to use a dataset of 5,011 New Years’ resolution tweets from 2015 to create a text network (Figure Eight 2015). The column of texts (called “text”) is already cleaned.

The dataset is called “nyr_data.rds,” and you can read it in using the **readRDS()** function.

Once the data are loaded, do the following:

1. Create the text network. Group the texts by the resolution topic (hint: look at the actual data frame to see what I mean).
2. Visualize the "document-level" projection of the network. What are the "documents," in this case?
3. Copy and paste an image of the network into your Word document.
4. What is this network telling us? That is, what does an edge between any two vertices tell us? What do the automatically-generated clusters indicate?
5. The **textnets** package (Bail 2018) takes care of two important problems when converting texts into networks. What are they, and why are they problematic?

Hints

The R scripts for the text network slides (pres111.R and pres112.R) might be very helpful. Like, *really really helpful*.

References

Bail, Christopher A. 2018. “**textnets**: Graph-Based Automated Text Analysis.” Retrieved March 28, 2019 ([link](#)).

Figure Eight. 2015. “2015 New Year’s Resolutions.” Retrieved March 28, 2019 ([link](#)).