

## Module 4 Faillog

### Exercise1: Social Network Analysis

- Installed R on Mac
  - Installed RStudio on Mac
  - Installed Gephi version 0.9 on Mac OS High Sierra
  - error trying to import my spreadsheet into Gephi. Realised my mistake was i had to files saved and I wasn't loading the right file.
  - Completed the exercise and saved the output as Exercise1Gephi.graphml
  - Saved figure as Exercise1Gephianalysisfile.pdf
- the figure was still a bit confusing, though information is available, I do feel like I could've worked on it a bit more for clearer analysis*

### Exercise 3: Topic Modeling in R

Did the quick tutorial on R as provided by [Dr. Graham](#)

Since I had installed R and RStudio on my Mac, my goal was to use this to [topic model the text of the newspaper articles to see the patterns.](#)

Unfortunately, when i tried to install rJava i'd get this error

```
> library(rJava)
```

*Error: package or namespace load failed for 'rJava':*

*.onLoad failed in loadNamespace() for 'rJava', details:*

*call: dyn.load(file, DLLpath = DLLpath, ...)*

*error: unable to load shared object*

*"/Library/Frameworks/R.framework/Versions/3.5/Resources/library/rJava/lib*

*dlopen(/Library/Frameworks/R.framework/Versions/3.5/Resources/library/rj*

*6): Library not loaded: /Library/Java/JavaVirtualMachines/jdk-*

*9.jdk/Contents/Home/lib/server/libjvm.dylib*

*Referenced from:*

*/Library/Frameworks/R.framework/Versions/3.5/Resources/library/rJava/libs*

*Reason: image not found*

Even when i tried to include ' ' in the library code line, the error was still the same.

That gave me a bit of a headache so I went on to use RStudio in DhBox.

Now as I reflect, I see I must've missed something, and Professor Graham also mentioned in the workbook that if you're using your own machine and using the exercise from Module2, "you would have to tell R to load up those materials from a directory, rather than by reading a CSV file." Which I somehow also missed.

- Instead of using the approach for 20 topics as it was in the workbook, i used 7. I realized my txt data was much less than my ocr.txt pdf data.
- Was it different from the list in the tutorial? I think so, because it's algorithm-based, and my 100 iterations would have been different from Dr. Graham's 1000 iterations, etc.
- ran wordcloud; saved histograms and clustergraph as .pdf
- ran network modelling; saved .graphml file
- Interested in doing more work in R. I like the output (visually), but i'm also intersted in the analysis in R once i figure out exactly the issue with the RStudio on my system

## **Exercise 7: Quick Charts Using RAW**

- Opened Google Sheets and imported the data

- Added the Blank detector
- Tried to fill blank cells with null and wow! it was a mess! It would run till line 350 and give me the error '**run exceeded maximum time**' Tried it about 10times and that keps happening so i uploaded Professor Graham's file
- Tried the Alleviual diagram with different place settings and exported my output as .png. The diagrams were heavily covered in lines which got me really confused and made me realize I might not want to use RAW for my final project.