How to populate the databases

1. Set up scraper3.py by editing it and following the comments. The comments should guide you to the end of the file to where all the directories are listed. Change the directories to match your system environment.
2. Run scraper3.py from the command prompt
3. After Scraper3.py has finished, open the “Topic Modeling” R Project
4. Run “LoadData.R” this will load data from MongoDB and store it to a .RData file
5. Run HarmonicMean.R This will take the data and find the optimal number of topics to generate.
6. Edit ModelingData.R and set k equal to the number the harmonicmean program gave us
7. Run the ModelingData.R script and it will generate CSV files with the information
8. Edit RelatesTo.py and TopicImporting.py so anywhere a file is mentioned make sure it is correct and accurate
9. Run TopicImporting.py then RelatesTo.py