Learning activity: Automated text analysis

Today, our goal is to understand computational possibilities for text analysis. For this purpose, we analyze Trump’s tweets, available here: <https://helsinkifi-my.sharepoint.com/:x:/g/personal/mnelimar_ad_helsinki_fi/EaNe6xeTqJpAkQOGowwyVU4BLkTO7O1imG6UrdMKaC5V2g?e=DtdOiC>

In groups

* Determine what type of research question you are interested, what are your key concepts you operate with. For example, the affectivity of a message could be one, as could the framing occurring in the messages.
* Conduct a manual text analysis on the data sample, that is, use some traditional qualitative analysis methods on a random (smallish) sample of the tweets. You can try grounded theory, frame analysis, thematic classification, …  
  Main idea is that you familiarize yourself with the data and get a feeling what traditional social science researchers would do with the data.
* Now, try at least two automated analysis methods for computational data analysis. See the following code examples:
  + <https://compsocialscience.github.io/summer-institute/2019/materials/day3-text-analysis/basic-text-analysis/rmarkdown/Basic_Text_Analysis_in_R.html>
  + <https://compsocialscience.github.io/summer-institute/2019/materials/day3-text-analysis/dictionary-methods/rmarkdown/Dictionary-Based_Text_Analysis.html>
  + <https://compsocialscience.github.io/summer-institute/2019/materials/day3-text-analysis/topic-modeling/rmarkdown/Topic_Modeling.html>
  + <https://compsocialscience.github.io/summer-institute/2019/materials/day3-text-analysis/text-networks/rmarkdown/Text_Networks.html>
* Compare your findings. What differences did you observe? What similarities did the manual and automated text analysis techniques demonstrate?