

Text Analysis 1

Dani Madrid-Morales

22-November-2023

10:00am - 10:10am : Setting up, R reminder, housekeeping & learning objectives

10:10am - 10:40am : Overview of computational text analysis & bag of words (BoW) approach.

10:40am - 10:55am : `quanteda` demo #1

- Installing packages
- Loading data to `quanteda` using `readtext`
- Creating a corpus on `quanteda`
- Corpus level metadata (or docvars)

10:55am - 11:00am : Break

11:00am - 11:15am : Common steps in text pre-processing

11:15am - 11:35am : `quanteda` demo #2

- Creating a tokens object in `quanteda`
- Adjusting pre-processing options in `quanteda`
- Generating a DFM (document feature matrix)

11:30am - 11:40am : Using dictionaries (or lexicons) for text analysis

11:40am - 11:50am : dictionary methods demo

- Applying a dictionary using `quanteda`
- Visualizing the results of a dictionary analysis
- Using newly create data to answer RQs

11:50am - 11:55am : Recap and what's next?