Objects to Data

Group Proposal

Martin Roger Ze Schneider, Lars Mangel, Dominique Brouwer, Cassandra Sijtsma, Dianne Ocampo

Question

How often are Hillary Clinton and Donald Trump mentioned on media outlets such as The New York Times or The Guardian? Does the frequency increase as Election Day nears? Who is mentioned more and how are they discussed?

Method

- Scrape data off NYT API and Guardian API
- Filter results according to timeline of August to October
    - August is after both the RNC and DNC, October is just after the first Presidential Debate.
- Filter results according to ranking of keywords in articles
- Filter results according to what section articles appear in on the website
- When there are high spikes in frequency, we can decide whether or not to analyze further and see what context the candidate is discussed in (positive or negative?)
- Use RStudio to plot

Visualisation

- How often is Hillary Clinton mentioned from August to October? Linear distribution
- How often is Donald Trump mentioned? Linear distribution
- Who is mentioned more and when? Shaded range

Data

- NYT API
- Guardian API