# In-class practice

**Data Collection**

LexisNexis Academic allows for easy selection and download of files formats suitable for Voyant to analyze. Please follow the steps below to retrieve and properly organize and clean your data for analysis.

**How to search for New York Times, USA Today, and Washington Post news**

1. Once you’ve navigated to LexisNexis Academic from our library’s database page, look for the search window bordered in red. At its upper right-hand corner, you will see a bar that says ‘Search by Subject or Topic.’ Click on it for a pull-down menu to then click on the first link ‘All News’. This will activate the ‘Advanced Option’ pop-up window.
2. When arriving at the ‘Advanced Option’ pop-up window, make sure you’ve selected the following to set your search parameters.
	* Date: From October, 26, 2017 to September 30, 2018
	* Build Your Own Segment Search: HEADLINE(opioid)
	* Source: New York Times (or USA Today, Washington Post)
	* Leave the rest options blank
3. Click on the search button to pull up the search results that look somewhat like the image below. Your search will yield many results. You can sort these results, “newest to oldest.” Also look to the upper right hand corner of the page for ‘Duplicate Options’ and choose ‘On-High similarity’ to avoid overlaps.



1. The result list is organized to show 25 articles per page. Notice how the results have each article numbered with a checkbox proceeding it. Click on the box at the top of all the check boxes to select all 25 stories appearing on the current page. Next, to avoid using opinion and editorial pieces, use ctrl+F (if you’re using a PC) or Command + F (if you’re using a Mac) to search for ‘opinion’ and ‘editorial’ stories. For the highlighted ‘editorial’ and ‘opinion’ pieces, deselect them from your list. To be safe, it’s best that you still read through the title of each article to confirm it is a news story, not an editorial or opinion column. Once you’ve selected the suitable stories, click for the next 25 articles and repeat the select and deselect process detailed earlier. Do so until you’ve exhausted the entire search results. Different publication will yield different numbers of results ranging from the low thirties to low one hundred. Finally, click on the disc icon in the upper right of your screen (it will say ‘download documents’ when you scroll over it).
2. Download the file (it will be saved as a Rich Text Format for now) and label it as ‘Data’ and followed with the publication name (e.g., Data NYT, Data WaPo, Data USA Today).
3. Note: NYTimes dataset also filtered out stories published through NYTimes’ magazine or podcast desk. Quotation of the Day is also filtered out.

**Data Cleaning**

In order to prepare our dataset for Voyant to analyze, you will need to run your downloaded news file through this webpage: “[Remove News Metadata](https://dla.endicott.edu/software/remove-news-metadata/)”.

This website is created by Prof. Hank Feild, Assistant Professor of Computer Science, to help filter out the unnecessary info from dataset so information to be analyzed by Voyant only includes the headline and the body of the news of interest. Some examples of extraneous information from the downloaded dataset include BYLINE, SECTION, LENGTH, LANGUAGE, DOCUMENT-TYPE, PUBLICATION-TYPE, and Copyright information).

# Framing Exercise 2

**Close Reading of Sample Articles**

1. In preparation to engage in a better informed interpretation of the Voyant data output later in the project, the first step is to engage in a close reading of some sample stories. With the different types of frame identified in McGinty et al.’s (2016) article, go through our national dataset and do a close reading of four stories. For the sake of familiarizing yourself with the various context in which frames are built, find four stories that uses distinctively different frames to organize their reporting. Read each headline and the first couple of paragraphs to help you choose the most suitable four.
2. Use the questions listed below to help you annotate and ensure you’ve spend time digesting the information. Most importantly, to engage in a preliminary framing analysis of the individual articles. Your notes should address these questions.

Individual news

* 1. What kinds of experiences are depicted?
	2. Who are the sources the reporters use to verify their information?
	3. Were different angles of the story presented? What are they?
	4. How is the story told in relation to their importance to the entire community or nation?
	5. How complex is the stories?

Comparison between stories

* 1. Are there similarities in the words chosen or images used?
	2. How are the stories treated differently?
	3. Can you make any interpretations or arguments based on the kinds of story covered, sources used, areas covered, or words/images chosen?
1. For the deliverable, upload the annotated PDF document of the four stories along with a typed up .doc file that answers the questions listed above.