Application Programming Interface (API) for Web-Scraping & Text Analysis

Individual Class Info

Introduction to Sociology
Laura Senier
Spring 2020
Number of Students: 60

Summary of Module

This module introduces students to Webscraping & APIs for Data Collection, outlines and explores the New York Times API, presents how to use python and jupyter notebooks to make use of the New York Times Article API and collect articles relating to the "Opioid Epidemic". From here, students are exposed to how to go from unstructured data to structured data and begin to ask and answer questions with this data. Students will then have a hands on opporunity to conduct text analysis on Opioid Epidemic articles collected from the New York Times API using the browser based computational text analysis tool Lexos. The module includes a brief discussion on how computational text analysis is useful, the browser tools that can be used to do this method of analysis, and how to interpret the results.

Learning Goals

- Understand the definition and purpose of an API and web-scraping.
- Understand the importance of API documentation.
- Understand the affordances and limitations of using APIs to build a corpus
- Start to understand how to use digital tools to pull out novel insights and findings from text data

Learning Objectives

- Learn about Webscraping & APIs for Data Collection
- Introduce and Explore the New York Times API
- Collect "Opioid Epidemic" News Coverage
- Conduct a Text Analysis of Opioid Epidemic News Coverage in Lexos

Materials for Module

Slides

[Jupyter Notebook](#)

[Data](#)
- New York Times Articles relating to the search term "opioids". Data collected from the New York Times Article API, comprising the 10 most relevant articles per year relating to opioids from 2010 to 2019.

[Stopwords File](#)