README

LE: 31 July 2023

# Purpose

The purpose of this project is to automate web scraping various websites where economics working papers (i.e., new papers that are not yet peer-reviewed) are published. The basic project runs several Python scripts that each scrape a specific website and extract data: The title of the paper, its author(s), the date it was posted, a link to the text, the number (using the numbering system of the publishing website), and an abstract. Data from all sources are then combined and compared to previous scrapes. Those papers that haven’t been seen before – the newest papers – are outputted to a .csv file which can be easily opened in Excel.

Websites that are scraped for data, as of July 2023, are:

|  |  |
| --- | --- |
| Name of website | Name of script |
| [Bureau of Economic Analysis](https://www.bea.gov/) | roundup\_scripts/scrapers/BEA.py |
| Becker Friedman Institute (at the University of Chicago) |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

This web scraper was originally developed for use at the [Brookings Institution Hutchins Center on Fiscal and Monetary Policy](https://www.brookings.edu/centers/the-hutchins-center-on-fiscal-and-monetary-policy/the-hutchins-roundup/). The center publishes a weekly newsletter called the [Hutchins Roundup](https://www.brookings.edu/centers/the-hutchins-center-on-fiscal-and-monetary-policy/the-hutchins-roundup/) which summarizes three new economics papers, as well as featuring a chart and recent quote from a famous macroeconomist.