Web Scraping and Web Crawling

**SCRAPY:** Web crawling framework written in python.

Pros: Asynchronous

Cons: Steep Learning Curve

**SELENIUM:** Tool for writing web automated tasks

Used for web scraping because beginner friendly

Easy to extract data from Javascript

Use in pair with Scrapy

Iterate over JS pages with Selenium and extracts data using Scrapy

**VS URLLIB:** Module for opening and reading web pages

Included in Python Library

**VS REQUESTS:** Module for opening and reading web pages

Has great docs

Must install it

**VS BEAUTIFULSOUP:** Module for extracting and scraping

Creates a parse tree that can be used to extract data from HTML

Beginner friendly

**VS LXML:** Module for extracting and scraping

Fast and Memory Efficient

Scrapes HTML and LXML

*For small work use Requests and LXML together*