**Python Libraries**

**Webscraping**

1. **Scrapy** - Scrapy helps to build crawling programs (spider bots) that can retrieve structured data from the web – for example, URLs or contact info. It's a great tool for scraping data used in, for example, Python machine learning models.
2. **BeautifulSoup** - BeautifulSoup is another really popular library for web crawling and data scraping. If you want to collect data that’s available on some website but not via a proper CSV or API, BeautifulSoup can help you scrape it and arrange it into the format you need.

**Data Processing and Modeling**

1. **Numpy** - The library offers many handy features performing operations on n-arrays and matrices in Python. It helps to process arrays that store values of the same data type and makes performing math operations on arrays (and their vectorization) easier. In fact, the vectorization of mathematical operations on the NumPy array type increases performance and accelerates the execution time.
2. **Pandas** - Pandas is a library created to help developers work with "labeled" and "relational" data intuitively. It's based on two main data structures: "Series" (one-dimensional, like a list of items) and "Data Frames" (two-dimensional, like a table with multiple columns). Pandas allows converting data structures to DataFrame objects, handling missing data, and adding/deleting columns from DataFrame, imputing missing files, and plotting data with histogram or plot box. It’s a must-have for data wrangling, manipulation, and visualization.
3. **Scikit-Learn** - This is an industry-standard for data science projects based in Python. Scikits is a group of packages in the SciPy Stack that were created for specific functionalities – for example, image processing. Scikit-learn uses the math operations of SciPy to expose a concise interface to the most common machine learning algorithms.
4. **PyTorch** - PyTorch is a framework that is perfect for data scientists who want to perform deep learning tasks easily. The tool allows performing tensor computations with GPU acceleration. It's also used for other tasks – for example, for creating dynamic computational graphs and calculating gradients automatically. PyTorch is based on Torch, which is an open-source deep learning library implemented in C, with a wrapper in Lua.

**Data Visualisation**

1. **Matplotlib** - This is a standard data science library that helps to generate data visualizations such as two-dimensional diagrams and graphs (histograms, scatterplots, non-Cartesian coordinates graphs). Matplotlib is one of those plotting libraries that are really useful in data science projects — it  provides an object-oriented API for embedding plots into applications.
2. **Seaborn** - Seaborn is based on Matplotlib and serves as a useful Python machine learning tool for visualizing statistical models – heatmaps and other types of visualizations that summarize data and depict the overall distributions. When using this library, you get to benefit from an extensive gallery of visualizations (including complex ones like time series, joint plots, and violin diagrams).

**Steps to install a library**

Let’s say if we want to install scikit-learn:

1. Install 64bit version of Python3
2. If using Conda environment,

conda install scikit-learn