

This code is a web scraper that is used to extract articles from the Business sections of two different websites: The Guardian and NBC News. It uses the Python library BeautifulSoup to parse HTML data and the requests library to send HTTP requests to the websites.

The code first sends a request to the Business section of The Guardian's website and extracts a list of links to articles on the page using the BeautifulSoup object. It then iterates through these links, sending requests to each one and using BeautifulSoup to extract the main body of text from each article. It stores these bodies of text in a list called "articles1".

Next, the code sends a request to the Business section of NBC News' website and extracts two lists of links to articles on the page. It combines these lists into one and iterates through the links, sending requests to each one and using BeautifulSoup to extract the main body of text from each article. It stores these bodies of text in a list called "articles2".

Finally, the code combines the lists of articles from both sources into one list called "articles". The code then uses the Natural Language Toolkit (nltk) library to perform various operations on the text in the articles list, including tokenization and part-of-speech tagging.

The code also uses pickle, a library for serializing and deserializing Python objects, to save the list of articles to a file for later use.