Flipcart Webscraping project using BeautifulSoup and Python

The Flipkart web scraping project aims to extract data from Flipkart's website using Python and the BeautifulSoup library. By scraping Flipkart, you can gather information such as product names, prices, descriptions, ratings, reviews, and more.

The tools uses here are requests, bs4 and re.

the following steps outline the general process:-

Import Required Libraries: Begin by importing the necessary libraries. For Flipkart web scraping, you'll need the requests library to send HTTP requests and the BeautifulSoup library to parse and navigate the HTML content of the website.

Send a GET Request: Use the requests library to send a GET request to the Flipkart website URL. This request fetches the HTML content of the webpage.

Create a BeautifulSoup Object: Create a BeautifulSoup object by passing the response content and specifying the parser (in this case, the 'html.parser' parser). The BeautifulSoup object allows you to navigate and extract data from the HTML structure of the webpage.

Find Elements to Scrape: Use BeautifulSoup's methods to locate specific HTML elements containing the data you want to scrape. This typically involves inspecting the HTML structure of the webpage and identifying relevant HTML tags, attributes, or class names that correspond to the desired data.

Extract Data: Once you've identified the elements to scrape, use BeautifulSoup's methods to extract the data. You can access the text content of an element, extract attributes, navigate through the HTML structure, and more, depending on your specific scraping requirements.

Process and Store the Data: Process the extracted data as needed. You can store it in variables, lists, or data structures, and perform any necessary transformations or cleaning operations.

Analyze or Save the Data: Finally, you can analyze the scraped data or save it for further use. This may involve performing statistical analysis, visualizing the data, or storing it in a structured format such as CSV, JSON, or a database.