

```
# -*- coding: utf-8 -*-  
"""
```

Created on Mon Sep 26 13:54:03 2022

```
@author: akash  
"""
```

```
# importing libraries
```

```
from bs4 import BeautifulSoup  
import requests
```

```
def main(URL):
```

```
    # opening our output file in append mode
```

```
    File = open("out.csv", "a")
```

```
    # specifying user agent, You can use other user agents
```

```
    # available on the internet
```

```
    HEADERS = ({'User-Agent':  
                'Mozilla/5.0 (X11; Linux x86_64)  
                AppleWebKit/537.36 (KHTML, like Gecko)  
                Chrome/44.0.2403.157 Safari/537.36',  
                'Accept-Language': 'en-US, en;q=0.5'})
```

```
    # Making the HTTP Request
```

```
    webpage = requests.get(URL, headers=HEADERS)
```

```
    # Creating the Soup Object containing all data
```

```
    soup = BeautifulSoup(webpage.content, "xml")
```

```
    # retrieving product title
```

```
    try:
```

```
        # Outer Tag Object
```

```
        title = soup.find("span",  
                           attrs={'id': 'productTitle'})
```

```
        # Inner NavigableString Object
```

```
        title_value = title.string
```

```
        # Title as a string value
```

```
        title_string = title_value.strip().replace(',', '')
```

```
    except AttributeError:
```

```
        title_string = "NA"
```

```
    print("product Title = ", title_string)
```

```
    # saving the title in the file
```

```
    File.write(f"{title_string},")
```

```
    # retrieving price
```

```
    try:
```

```
        price = soup.find(  
            "span", attrs={'id': 'priceblock_ourprice'})  
        .string.strip().replace(',', '')
```

```
        # we are omitting unnecessary spaces
```

```
        # and commas form our string
```

```
    except AttributeError:
```

```
        price = "NA"
```

```
    print("Products price = ", price)
```

```
    # saving
```

```
    File.write(f"{price},")
```

```
    # retrieving product rating
```

```
    try:
```

```
        rating = soup.find("i", attrs={  
            'class': 'a-icon a-icon-star a-star-4-5'})
```

```
.string.strip().replace(',', '')
```

```
except AttributeError:
```

```
try:
```

```
rating = soup.find(
    "span", attrs={'class': 'a-icon-alt'})
    .string.strip().replace(',', '')
```

```
except:
```

```
rating = "NA"
```

```
print("Overall rating = ", rating)
```

```
File.write(f"{rating},")
```

```
try:
```

```
review_count = soup.find(
    "span", attrs={'id': 'acrCustomerReviewText'})
    .string.strip().replace(',', '')
```

```
except AttributeError:
```

```
review_count = "NA"
```

```
print("Total reviews = ", review_count)
```

```
File.write(f"{review_count},")
```

```
# print availability status
```

```
try:
```

```
available = soup.find("div", attrs={'id': 'availability'})
available = available.find("span")
    .string.strip().replace(',', '')
```

```
except AttributeError:
```

```
available = "NA"
```

```
print("Availability = ", available)
```

```
# saving the availability and closing the line
```

```
File.write(f"{available},\n")
```

```
# closing the file
```

```
File.close()
```

```
if __name__ == '__main__':
```

```
# opening our url file to access URLs
```

```
file = open("url.txt", "r")
```

```
# iterating over the urls
```

```
for links in file.readlines():
```

```
    main(links)
```