

ICP Technical
Advisory Group
Meeting
February 2020
Washington, DC

PRICE WEB SCRAPING IN WESTERN ASIA



DEVELOPING THE IDEA

- Western Asia's ICP team aimed to modernize price statistics in the region by introducing technology and big data
- The ICP RIA in Western Asia chose 3 pilot countries to launch its newest initiative: the application of web scraping in price data collection
- Bahrain, Kuwait and Qatar will implement web scraping techniques in collecting prices for HHC-FETI in 2020
- Training was conducted for Bahrain and Kuwait in January and February 2020 respectively, and training will be conducted for Qatar in March 2020
- Web scraping represents a complementary data collection methodology to the traditional field collection and does not intend to replace it

BENEFITS

This initiative can be applied in both CPI and ICP price collection

It boosts data collection efficiency by centralizing the collection process

It enables more frequent data collection

It results in more accurate data as the automated process does not allow for data entry errors

It can be used as an additional validation tool for revision of field-collected data

CHALLENGES

Each country provided the regional office with a list of websites for FETI outlets

The regional team worked on coding the web scraping scripts for each website

Each website presented a different case and new challenge

However, the regional team was able to scrape prices from the provided websites, developed a template for each country and trained the national teams on web scraping using python

WEB SCRAPING EXAMPLE

1

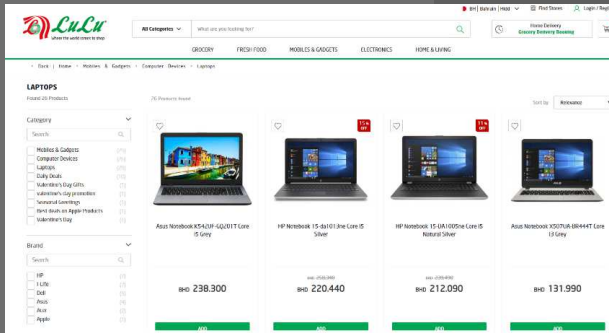
Website is chosen

2

Code is written

3

Data is obtained in excel spreadsheet



```
import requests
from bs4 import BeautifulSoup
import pandas as pd
import re

def scrape():
    headers = requests.utils.default_headers()
    headers.update({
        'User-Agent': 'Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0',
    })
    prod_name_list = []
    prod_price_list = []
    page = requests.get("https://www.luluhypermarket.com/en-bh/laptops-computer-devices-mobiles-gadgets/c/HY00214756*",
                        headers=headers)
    soup = BeautifulSoup(page.text, 'html.parser')
    products = soup.find('div', class_='product_listing product_grid col-xs-12')

    prod_list = products.find_all('div', class_='product-title-main')
    for prod in prod_list:
        prod_name = prod.find('div', class_='p1p-prod-name').text
        price_box = prod.find('div', class_='product-pricing-section')

        try:
            prod_price_new = price_box.find('div', class_='price').text
            prod_price = price_box.find('div', class_='act-price').text
        except:
            prod_price = price_box.find('div', class_='act-price').text
```

	Name	Initial Price	Final Price
1			
2	1 Apple MacBook - Core M3 1.2GHz 8GB 256GB Shared Silver Arabic	BHD593.250	BHD349.990
3	2 Apple MacBook Air (2019) - Core i5 1.6GHz 8GB 128GB Shared 13.3inch Space Grey Arabic	BHD521.745	BHD499.990
4	3 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 1.4GHz 8GB 128GB Shared 13.3inch Space Grey Arabic	BHD607.950	BHD569.990
5	4 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 2.4GHz 8GB 128GB Shared 13.3inch Space Grey Arabic		BHD917.700
6	5 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 1.4GHz 8GB 256GB Shared 13.3inch Space Grey Arabic		BHD702.450
7	6 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 2.4GHz 8GB 256GB Shared 13.3inch Space Grey Arabic		BHD823.200
8	7 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 2.4GHz 8GB 512GB Shared 13.3inch Space Grey Arabic		BHD917.700
9	8 Apple MacBook Pro 13 with Touch Bar (2019) - Core i5 2.4GHz 8GB 256GB Shared 13.3inch Space Grey Arabic		BHD823.200
10	9 Dell Inspiron 15 5584 Laptop - Core i7 2.6GHz 8GB 1TB+128GB 4GB 15.6inch FHD Silver	BHD383.824	BHD362.824
11	10 Dell Inspiron 15 3581 Laptop - Core i3 2.3GHz 4GB 1TB Shared Win10 15.6inch FHD Black		BHD194.562
12	11 Dell Inspiron 14 3480 Laptop - Core i5 1.6GHz 4GB 1TB 2GB Win10 14inch HD Silver		BHD237.822
13	12 Dell G3 15 Gaming Laptop - Core i7 2.6GHz 16GB 1TB+256GB 4GB Win10 15.6inch FHD Black	BHD488.297	BHD472.990
14	13 Dell G5 15 Gaming Laptop - Core i7 2.6GHz 16GB 1TB+256GB 6GB Win10 15.6inch FHD Black		BHD633.218
15	14 Lenovo 300e Convertible Touch Laptop - Celeron 1.1GHz 4GB 128GB Shared Win10 11.6inch HD Black		BHD159.990
16	15 Lenovo IdeaPad 330 Laptop - Celeron 1.8GHz 4GB 500GB Shared Win10 15.6inch HD Onyx Black		BHD109.990
17	16 Lenovo IdeaPad 130 Laptop - Core i3 2.3GHz 4GB 1TB Shared Win10 15.6inch HD Black		BHD139.990
18	17 Lenovo IdeaPad L340 Gaming Laptop - Core i7 2.6GHz 16GB 1TB+128GB 4GB Win10 15.6inch FHD Black		BHD429.990
19	18 Lenovo Legion Y540 Gaming Laptop - Core i7 2.6GHz 16GB 1TB+256GB 6GB Win10 15.6inch FHD Black		BHD599.990
20	19		

THANK YOU