









WEB SCRAPPING – iPhone online store

Web scrapping is one of method to get data from website. At this project, the data will be scrapped is iPhone and its price from the local website.

The first thing to scrape is knowing the local website. The website i scrapped is

<https://www.bhinneka.com/jual?cari=iphone> . This website contain about the iPhone type and its price.

Here it's the example of the website.

				
Ferrari - Power Case iPhone 11 Pro Max 6.5...	Ferrari - Power Case iPhone 11 Pro Max 6.5...	CASEOLOGY Parallax for iPhone X/ iPhone Xs...	CASEOLOGY Waterfall Series for iPhone 7/...	PROTEGO Back Protector for iPhone 7 / iPhone 8...
Cicilan Rp 82.500/bln	Cicilan Rp 82.500/bln	Cicilan Rp 9.984/bln	Cicilan Rp 4.150/bln	Cicilan Rp 10.834/bln
Rp 990.000	Rp 990.000	Rp 119.800	Rp 49.800	Rp 130.000
				

To scrape the web, we need some package such as BeautifulSoup, requests, etc.

Here the code i used for scrapping. I used the selenium package for parsing html code due to different the code of inspect html and the code requested. I have to know the path of chromedriver.exe first to assist this selenium package. The code i writte shown below.

```
from selenium import webdriver
import time
from bs4 import BeautifulSoup

driver = webdriver.Chrome("C:/Users/ACER/anaconda3/chromedriver.exe")
url= "https://www.bhinneka.com/jual?cari=iphone"
driver.maximize_window()
driver.get(url)

time.sleep(5)
content = driver.page_source.encode('utf-8').strip()
soup = BeautifulSoup(content,"html.parser")
```

After request success. I able to parsing html by print the “soup” variabel above. Here the result

```
print(soup)

<html lang="id-ID"><head><script async="" src="https://www.google-analytics.com/plugins/ua/ec.js" type="text/javascript"></script><script async="" src="//www.googleadservices.com/pagead/conversion_async.js" type="text/javascript"></script><script async="" src="https://www.google-analytics.com/analytics.js" type="text/javascript"></script><script async="" src="https://connect.facebook.net/signals/config/1452556695066117?v=2.9.40&r=stable"></script><script async="" src="//connect.facebook.net/en_US/fbevents.js"></script><script async="" src="https://www.googletagmanager.com/gtm.js?id=GTM-PXD88J&gtm_auth=&gtm_preview=&gtm_cookies_win=x"></script><script>
(function(w,d,s,l,i){w[l]=w[l]||[];
w[l].push({'gtm.start': new Date().getTime(),event:'gtm.js', });
var f=d.getElementsByTagName(s)[0],j=d.createElement(s),dl=l!='dataLayer'?'&l='+l:'';
j.async=true;j.src='https://www.googletagmanager.com/gtm.js?id='+i+dl+'&gtm_auth=&gtm_preview=&gtm_cookies_win=x';
f.parentNode.insertBefore(j,f);
})(window,document,'script','dataLayer','GTM-PXD88J');</script><script>
window.dataLayer = window.dataLayer || [];
window.dataLayer.push({'event':"Next.js-hydration","data":{"event_category":"Next.js custom metric","value":14950,"event_label":"1622786278789-8845381061050","non_interaction":true}})</script><meta charset="utf-8"/><link href="https://www.bhinneka.com/jual" rel="canonical"/><meta content="initial-scale=1.0, width=device-width" name="viewport"/><meta content="Temukan pilihan produk B2B terlengkap untuk UMKM, Korporasi & Pemerintah dengan pilihan produk maupun jasa terlengkap, Asli Bikin Tenang." name="description"/><link href="//static.bmdstatic.com/sf/assets/img/favicons/favicon-152.png?v=2.6.0" rel="apple-touch-icon-precomposed"/><link href="//static.bmdstatic.com/sf/assets/img/favicons/favicon-152.png?v=2.6.0" rel="apple-touch-icon-precomposed" sizes="152x152"/><link href="//static.bmdstatic.com/sf/assets/ime/favicons/favicon-144.png?v=2.6.0" rel="apple-touch-icon-precomposed" sizes="144x144"/></head><body>
```

The result above is about html code that similar to inspect of its website. Here the html code in website can be shown after inspected.

```
<!DOCTYPE html>
<html lang="id-ID">
  <head>...</head>
  <body style="overflow: inherit;"> == $0
    <noscript>...</noscript>
    <div id="__next">...</div>
    <script id="__NEXT_DATA__" type="application/json">...</script>
    <script nomodule src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/polyfills-0d978c7...js"></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/webpack-4a76ff5...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/framework.7ddfe22...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/commons.a33d746...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/3a8ad4b...0c59127...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/main-96ce595...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/75fc9c18.daa6e84...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/0f06aff...831a97a...js" async></script>
    <script src="https://static.bmdstatic.com/gk/production/assets/next/static/chunks/de20fa5...9830fa5...js" async></script>
```

I have to look forward about the title of iPhone type and its price in html code. Here the example the html code for showing the product in header of website.

```
<div class="col css-mz9zn6">
  <div class="css-i0xmc f eluu5jkk7"> flex
    <a href="/ferrari-power-case-iphone-11-pro-max-65-full-cover-powerbank-4000-mah-black-sku0014861759" data-testid="product-click" class="product-wrapper css-lbx1jd">
      
      <div class="css-1pthsci">
        <div class="product-title">
          <p font-size="14" class="css-194yrqz">
            "Ferrari - Power Case iPhone 11 Pro Max 6.5 Full Cover - Powerbank 4000 mAh Black"
          </p>
        </div>
        <span class="extra">Cicilan Rp 82.500/bln</span>
      </div>
    <div class="css-1qhv185"> flex == $0
      <div></div>
      <div class="price">Rp 990.000</div>
    </div>
```



Ferrari - Power Case
iPhone 11 Pro Max 6.5...
Cicilan Rp 82,500/bln

Rp 990.000

0% interest
Dibayar Maksimal

After i get the html code and its pattern, i have to find this keyword “div”, “col css-mz9zn6”, “product-title”, and “price”. The “div” means about division in html code and the “col css-mz9zn6” is class name which are contain products of iPhone. This class load the data about the iPhone type and its price.

By using findAll function, we can seek this keyword. Here the code for parsing.

```
produk = soup.findAll('div', {'class' : 'col css-mz9zn6'})
p = 1
hp = []
harga = []
for p in produk:
    sell = p.find('div', {'class' : 'product-title'}).get_text()
    price = p.find('div', {'class' : 'price'}).get_text()
    hp.append(sell)
    harga.append(price)
```

After parsed the data, i get type of iPhone and its price. The type of iPhone named as “hp”, and the price named as “harga”. The next step is to change its data become a dataframe by using code shown below.

```
import pandas as pd
column = ['hp']
unit1 = pd.DataFrame(hp, columns=column)
column1 = ['harga']
unit2 = pd.DataFrame(harga, columns=column1)
barang = pd.concat([unit1, unit2], axis=1).reindex(unit1.index)
print(barang)
```

Here the result in dataframe type

	hp	harga
0	Ferrari - Power Case iPhone 11 Pro Max 6.5 Ful...	Rp 990.000
1	Ferrari - Power Case iPhone 11 Pro Max 6.5 Ful...	Rp 990.000
2	CASEOLOGY Parallax for iPhone X/ iPhone Xs Cha...	Rp 119.800
3	CASEOLOGY Waterfall Series for iPhone 7/ iPhone 8	Rp 49.800
4	PROTEGO Back Protector for iPhone 7 / iPhone 8...	Rp 130.000
5	PROTEGO Tempered Glass for Apple iPhone 7 / iP...	Rp 130.000
6	SPIGEN iPhone Stand S315	Rp 69.800
7	VIVAN Paket Fast Charging iPhone VIVAN Charger...	Rp 110.000
8	APPLE iPhone 11 128GB - Black	Rp 14.999.000
9	APPLE iPhone 12 64GB - White	Rp 13.999.000
10	APPLE iPhone 12 64GB - Red	Rp 14.999.000
11	APPLE iPhone 12 128GB - Blue	Rp 16.499.000
12	APPLE iPhone 12 128GB - Red	Rp 16.499.000
13	APPLE iPhone 12 128GB - Black	Rp 16.499.000
14	APPLE iPhone 12 64GB - Black	Rp 14.999.000
15	APPLE iPhone 12 256GB - White	Rp 17.999.000
16	APPLE iPhone 12 128GB - White	Rp 16.499.000
17	APPLE iPhone 12 256GB - Red	Rp 17.999.000
18	APPLE iPhone 12 256GB - Black	Rp 16.999.000
19	APPLE iPhone SE 2020 64GB - Black	Rp 7.499.000
20	APPLE iPhone SE 2020 128GB - Black	Rp 8.799.000

The dataframe above show about the type of iPhone and its price.