Web Scraping Project

- We will be scraping a mere apple products data from the famous Indian shopping website "Flipkart"

In [134]:

```python
import requests
from bs4 import BeautifulSoup as bs
import pandas as pd
```

- Using requests library, we can fetch the content from the URL given
- Beautiful soup library helps to parse,fetch the details the way we want.
- Use requests and Beautiful Soup for scraping and parsing data from the Web.
- You can use a beautiful soup library to fetch data using Html tag, class, id, css selector and many more ways.
- Pandas makes it easy to scrape a table (tag) on a web page.
- After obtaining it as a DataFrame, it is possible to do various processing and save it as an Excel file or csv file.

In [135]:

```python
link="https://www.flipkart.com/search?q=iphone&otracker=search&otracker1=search&marketplace=FLIPKART&as-show=on&as=off"
```

- The above is the url link we are going to fetch for scrapping the required data.

- With the help of requests if we get the response of 200 it means that we can scrap that page.

In [136]:

```python
page = requests.get(link)
```

In [137]:

```python
page
```

Out[137]:

```
<Response [200]>
```

- Beautiful Soup then parses the document using the best available parser.
- The prettify() method will turn a Beautiful Soup parse tree into a nicely formatted Unicode string, with a separate line for each tag and each string

In [90]:

```python
soup = bs(page.text)
```

In [93]:

```python
print(soup.prettify())
```

```
<!DOCTYPE html>
<html lang="en">
 <head>
  <link href="https://rukminim1.flixcart.com" rel="preconnect"/>
  <link href="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app_modules.chunk.905c37.css" rel="style
sheet"/>
  <link href="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app.chunk.104e9a.css" rel="stylesheet"/>
  <meta content="text/html; charset=utf-8" http-equiv="Content-type"/>
  <meta content="IE=Edge" http-equiv="X-UA-Compatible"/>
  <meta content="102988293558" property="fb:page_id"/>
  <meta content="658873552,624500995,100000233612389" property="fb:admins"/>
  <meta content="noodp" name="robots"/>
  <link href="https://static-assets-web.flixcart.com/www/promos/new/20150528-140547-favicon-retina.ico" rel="shortcut ic
on"/>
  <link href="/osdd.xml?v=2" rel="search" type="application/opensearchdescription+xml"/>
  <meta content="website" property="og:type"/>
  <meta content="Flipkart.com" name="og_site_name" property="og:site_name"/>
  <link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes="57x57"/>
  <link href="/apple-touch-icon-72x72.png" rel="apple-touch-icon" sizes="72x72"/>
  <link href="/apple-touch-icon-114x114.png" rel="apple-touch-icon" sizes="114x114"/>
```

- soup.find will find the division tag and class related to required division and store it in info variable

In [121]:

```python
info = soup.find('div', class_='_2kHMtA')
```

```
1  info
```

```
[<div class="_2kHMtA"><a class="_1fQZEK" href="/apple-iphone-12-mini-black-64-gb/p/itm38b727191eb08?pid=MOBFWBYZXSEGBS6F
&amp;lid=LSTMOBFWBYZXSEGBS6FTONJ6W&amp;marketplace=FLIPKART&amp;q=iphone&amp;store=tyy%2F4io&amp;srno=s_1_1&amp;otracker
=search&amp;otracker1=search&amp;fm=organic&amp;iid=fbc149b5-c25f-4f83-95d0-9a4f65bfb8c0.MOBFWBYZXSEGBS6F.SEARCH&amp;ppt
=None&amp;ppn=None&amp;ssid=400lbqkj7k0000001671600660847&amp;qH=0b3f45b266a97d70" rel="noopener noreferrer" target="_bl
ank"><div class="MIXNux"><div class="_2QcLo-"><div><div class="CXW8mj" style="height:200px;width:200px"><img alt="APPLE
iPhone 12 mini (Black, 64 GB)" class="_396cs4 _3exPp9" src="https://rukminim1.flixcart.com/image/312/312/kg8avm80/mobil
e/n/p/d/apple-iphone-12-mini-dummyapplefsn-original-imafwgbfwkfrerak.jpeg?q=70"/></div></div></div><div class="_3wLduG">
<div class="_3PzNI-"><span class="f3A4_V"><label class="_2iDkf8"><input class="_30VH1S" readonly="" type="checkbox"/><di
v class="_24_Dny"></div></label></span><label class="_6Up2sF"><span>Add to Compare</span></label></div></div><div class
="_2hVSre _3nq8ih"><div class="_36FSn5"><svg class="_1l0elc" height="16" viewbox="0 0 20 16" width="16" xmlns="http://ww
w.w3.org/2000/svg"><path class="eX72wL" d="M8.695 16.682C4.06 12.382 1 9.536 1 6.065 1 3.219 3.178 1 5.95 1c1.566 0 3.06
9.746 4.05 1.915C10.981 1.745 12.484 1 14.05 1 16.822 1 19 3.22 19 6.065c0 3.471-3.06 6.316-7.695 10.617L10 17.897l-1.30
5-1.215z" fill="#2874F0" fill-rule="evenodd" opacity=".9" stroke="#FFF"></path></svg></div></div></div><div class="_3pLy
-c row"><div class="col col-7-12"><div class="_4rR01T">APPLE iPhone 12 mini (Black, 64 GB)</div><div class="gUuXy-"><spa
n class="_1lRcqv" id="productRating_LSTMOBFWBYZXSEGBS6FTONJ6W_MOBFWBYZXSEGBS6F_"><div class="_3LWZlK">4.5<img class="_1w
B99o" src="data:image/svg+xml;base64,PHN2ZyB4bWxucz0iaHR0cDovL3d3dy53My5vcmcvMjAwMC9zdmciIHdpZHRoPSIxMyIgaGVpZ2h0PSIxMiI
+PHBhdGggZmlsbD0iI0ZGRiIgZD0iTTYuNSA5LjQzOWwtMy42NzQgMi4yMy45NC00LjI1LTMuMjEtMi44ODMgNC4yNTQtLjQ1NEw2LjUuMTEbyDEuNjkgNC4
wMSA0LjI1NC40MDtMy4vMSAyLjg4Ljg4MC9sLjI2IIvPiwvc3ZnPg=="/></div></span><span class="_2_R_DZ"><span><span>1,20,982 Rati
```

- Here, We take an empty list and then we use for loop in info which all the div tags are present, and for every div tags we find all class tags also and then we append the list which we took as an empty list.

```
1  name = []
2  for i in info:
3      n = i.find_all ('div', class_='_4rR01T')
4      name.append (n[0].text)
5  name
```

```
['APPLE iPhone 12 mini (Black, 64 GB)',
 'APPLE iPhone 13 (Midnight, 128 GB)',
 'APPLE iPhone 13 (Blue, 128 GB)',
 'APPLE iPhone 11 (Red, 128 GB)',
 'APPLE iPhone 13 (Green, 128 GB)',
 'APPLE iPhone 12 mini (White, 64 GB)',
 'APPLE iPhone 13 ((PRODUCT)RED, 128 GB)',
 'APPLE iPhone 11 (White, 128 GB)',
 'APPLE iPhone 11 (Black, 128 GB)',
 'APPLE iPhone 12 mini (Black, 128 GB)',
 'APPLE iPhone 13 (Midnight, 256 GB)',
 'APPLE iPhone 13 (Blue, 256 GB)',
 'APPLE iPhone 13 (Green, 256 GB)',
 'APPLE iPhone 11 (Black, 64 GB)',
 'APPLE iPhone 11 (White, 64 GB)',
 'APPLE iPhone 12 mini (White, 128 GB)',
 'APPLE iPhone 12 mini (Blue, 64 GB)',
 'APPLE iPhone 11 (Purple, 128 GB)',
 'APPLE iPhone 11 (Yellow, 128 GB)',
 'APPLE iPhone 13 (Starlight, 128 GB)',
 'APPLE iPhone 11 (Purple, 64 GB)',
 'APPLE iPhone 11 (Green, 64 GB)',
 'APPLE iPhone 11 (Red, 64 GB)',
 'APPLE iPhone 11 (Green, 128 GB)']
```

```
1  rating = []
2
3  for i in info:
4      r = i.find_all ('div', class_='_3LWZlK')
5      rating.append (r[0].text)
6  rating
```

```
['4.5',
 '4.7',
 '4.7',
 '4.6',
 '4.7',
 '4.5',
 '4.7',
 '4.6',
 '4.6',
 '4.5',
 '4.7',
 '4.7',
 '4.7',
 '4.6',
 '4.6',
 '4.5',
 '4.5',
 '4.6',
 '4.6',
 '4.7',
 '4.6',
 '4.6',
 '4.6',
 '4.6']
```

```
1  cost = []
2  for i in info:
3      c = i.find_all ('div', class_='_30jeq3 _1_WHN1')
4      cost.append (c[0].text)
5  cost
```

```
['₹37,999',
 '₹62,999',
 '₹62,999',
 '₹43,999',
 '₹62,999',
 '₹37,999',
 '₹62,999',
 '₹45,999',
 '₹45,999',
 '₹43,999',
 '₹69,999',
 '₹69,999',
 '₹69,999',
 '₹40,999',
 '₹40,999',
 '₹43,999',
 '₹37,999',
 '₹45,999',
 '₹45,999',
 '₹62,999',
 '₹40,999',
 '₹40,999',
 '₹40,999',
 '₹45,999']
```

```
1  specs = []
2  for i in info:
3      s = i.find_all ('ul', class_="_1xgFaf")
4      specs.append (s[0].text)
5  specs
```

['64 GB ROM13.72 cm (5.4 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA14 Bionic Chip with Next Generation Neural Engine ProcessorCeramic ShieldIndustry-leading IP68 Water ResistanceAll Screen OLED Display12MP TrueDepth Front Camera with Night Mode, 4K Dolby Vision HDR RecordingBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '64 GB ROM13.72 cm (5.4 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA14 Bionic Chip with Next Generation Neural Engine ProcessorCeramic ShieldIndustry-leading IP68 Water ResistanceAll Screen OLED Display12MP TrueDepth Front Camera with Night Mode, 4K Dolby Vision HDR RecordingBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM13.72 cm (5.4 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA14 Bionic Chip with Next Generation Neural Engine ProcessorCeramic ShieldIndustry-leading IP68 Water ResistanceAll Screen OLED Display12MP TrueDepth Front Camera with Night Mode, 4K Dolby Vision HDR RecordingBrand Warranty for 1 Year',
 '256 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '256 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '256 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM13.72 cm (5.4 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA14 Bionic Chip with Next Generation Neural Engine ProcessorCeramic ShieldIndustry-leading IP68 Water ResistanceAll Screen OLED Display12MP TrueDepth Front Camera with Night Mode, 4K Dolby Vision HDR RecordingBrand Warranty for 1 Year',
 '64 GB ROM13.72 cm (5.4 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA14 Bionic Chip with Next Generation Neural Engine ProcessorCeramic ShieldIndustry-leading IP68 Water ResistanceAll Screen OLED Display12MP TrueDepth Front Camera with Night Mode, 4K Dolby Vision HDR RecordingBrand Warranty for 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Super Retina XDR Display12MP + 12MP | 12MP Front CameraA15 Bionic Chip ProcessorBrand Warranty for 1 Year',
 '64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year',
 '128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD Display12MP + 12MP | 12MP Front CameraA13 Bionic Chip ProcessorBrand Warranty of 1 Year']

- Now, we created a dictionary to store all the required data.

```
1  data = {
2      'Model Name': names,
3      'Ratings': rating,
4      'Price (INR)': cost,
5      'Specification': specs
6  }
```

```
1 data
```

Out[129]:

```
{'Model Name': ['APPLE iPhone 12 mini (Black, 64 GB)',
  'APPLE iPhone 13 (Midnight, 128 GB)',
  'APPLE iPhone 13 (Blue, 128 GB)',
  'APPLE iPhone 11 (Red, 128 GB)',
  'APPLE iPhone 13 (Green, 128 GB)',
  'APPLE iPhone 12 mini (White, 64 GB)',
  'APPLE iPhone 13 ((PRODUCT)RED, 128 GB)',
  'APPLE iPhone 11 (White, 128 GB)',
  'APPLE iPhone 11 (Black, 128 GB)',
  'APPLE iPhone 12 mini (Black, 128 GB)',
  'APPLE iPhone 13 (Midnight, 256 GB)',
  'APPLE iPhone 13 (Blue, 256 GB)',
  'APPLE iPhone 13 (Green, 256 GB)',
  'APPLE iPhone 11 (Black, 64 GB)',
  'APPLE iPhone 11 (White, 64 GB)',
  'APPLE iPhone 12 mini (White, 128 GB)',
  'APPLE iPhone 12 mini (Blue, 64 GB)',
  'APPLE iPhone 11 (Purple, 128 GB)',
```

- Now finally we covert this raw extracted data into tabular form using DataFrame function from Pandas Library

In [130]:

```
1 Apple_Phones = pd.DataFrame(data)
```

- To save this tabular data into a csv file (comma seperated values) we use to_csv('filename') and also we are giving index=False to remove the unrequired numbering in the table.

In [131]:

```
1 Apple_Phones.to_csv('Apple Phones.csv',index=False)
```

- Now to print the data from the saved csv file we use read_csv function using Pandas library.

In [132]:

```
1 pd.read_csv('Apple Phones.csv')
```

Out[132]:

	Model Name	Ratings	Price (INR)	Specification
0	APPLE iPhone 12 mini (Black, 64 GB)	4.5	₹37,999	64 GB ROM13.72 cm (5.4 inch) Super Retina XDR ...
1	APPLE iPhone 13 (Midnight, 128 GB)	4.7	₹62,999	128 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
2	APPLE iPhone 13 (Blue, 128 GB)	4.7	₹62,999	128 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
3	APPLE iPhone 11 (Red, 128 GB)	4.6	₹43,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...
4	APPLE iPhone 13 (Green, 128 GB)	4.7	₹62,999	128 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
5	APPLE iPhone 12 mini (White, 64 GB)	4.5	₹37,999	64 GB ROM13.72 cm (5.4 inch) Super Retina XDR ...
6	APPLE iPhone 13 ((PRODUCT)RED, 128 GB)	4.7	₹62,999	128 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
7	APPLE iPhone 11 (White, 128 GB)	4.6	₹45,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...
8	APPLE iPhone 11 (Black, 128 GB)	4.6	₹45,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...
9	APPLE iPhone 12 mini (Black, 128 GB)	4.5	₹43,999	128 GB ROM13.72 cm (5.4 inch) Super Retina XDR...
10	APPLE iPhone 13 (Midnight, 256 GB)	4.7	₹69,999	256 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
11	APPLE iPhone 13 (Blue, 256 GB)	4.7	₹69,999	256 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
12	APPLE iPhone 13 (Green, 256 GB)	4.7	₹69,999	256 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
13	APPLE iPhone 11 (Black, 64 GB)	4.6	₹40,999	64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD ...
14	APPLE iPhone 11 (White, 64 GB)	4.6	₹40,999	64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD ...
15	APPLE iPhone 12 mini (White, 128 GB)	4.5	₹43,999	128 GB ROM13.72 cm (5.4 inch) Super Retina XDR...
16	APPLE iPhone 12 mini (Blue, 64 GB)	4.5	₹37,999	64 GB ROM13.72 cm (5.4 inch) Super Retina XDR ...
17	APPLE iPhone 11 (Purple, 128 GB)	4.6	₹45,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...
18	APPLE iPhone 11 (Yellow, 128 GB)	4.6	₹45,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...
19	APPLE iPhone 13 (Starlight, 128 GB)	4.7	₹62,999	128 GB ROM15.49 cm (6.1 inch) Super Retina XDR...
20	APPLE iPhone 11 (Purple, 64 GB)	4.6	₹40,999	64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD ...
21	APPLE iPhone 11 (Green, 64 GB)	4.6	₹40,999	64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD ...
22	APPLE iPhone 11 (Red, 64 GB)	4.6	₹40,999	64 GB ROM15.49 cm (6.1 inch) Liquid Retina HD ...
23	APPLE iPhone 11 (Green, 128 GB)	4.6	₹45,999	128 GB ROM15.49 cm (6.1 inch) Liquid Retina HD...

In [ ]: