**Kulvir Singh**

**19BCE2074**

**Experiment 2 (15.02.2020)**

*Question 1*

***Problem Statement :***

Use scrapy to crawl any one of the E-commerce websites of your choice and perform the same. The following information needs to be extracted from the page: (Choose any one product: e.g. laptop, smartphone … etc.)

a)Product name

b)Product price

c)Product discount

d)Product image

***Procedure :***

Tools Used: Scrapy packages, python, ide to run the program, terminal/cmd (scrappy shell) for execution of various commands

Steps for Implementation:

1. Install the scrapy packages (pip install scrappy)

2. Create a new scrapy project (scrapy startproject [projectname])

3. Generate a spider to crawl through the web pages

4. Configure/edit the settings.py and the spider files as per requirement.

5. Run the program using ide and cmd to obtain results.

***Code :***

***items.py***

import scrapy

class ProductsItem(scrapy.Item):

# define the fields for your item here like:

product\_name = scrapy.Field()

product\_price = scrapy.Field()

product\_discount = scrapy.Field()

product\_image = scrapy.Field()

***flipkart\_spider.py***

import scrapy

from ..items import ProductsItem

class pro(scrapy.Spider):

name = "pro"

start\_urls = ["https://www.flipkart.com/samsung-galaxy-m21-midnight-blue-128-gb/p/itm0cec19c31b3cb?pid=MOBFSF85ZMVH3ZMG&lid=LSTMOBFSF85ZMVH3ZMGHVRMDR&marketplace=FLIPKART&fm=neo%2Fmerchandising&iid=M\_ab883bcf-9673-4bf2-adc9-fccdcf6198a4\_1\_1BUWY8OBA8L9\_MC.MOBFSF85ZMVH3ZMG&ppt=clp&ppn=samsung-mobile-store&ssid=0lzd5i7qgw0000001596104962790&otracker=clp\_pmu\_v2\_Latest%2BSamsung%2Bmobiles%2B\_1\_1.productCard.PMU\_V2\_Latest%2BSamsung%2Bmobiles%2B\_samsung-mobile-store\_MOBFSF85ZMVH3ZMG\_neo%2Fmerchandising\_0&otracker1=clp\_pmu\_v2\_PINNED\_neo%2Fmerchandising\_Latest%2BSamsung%2Bmobiles%2B\_LIST\_productCard\_cc\_1\_NA\_view-all&cid=MOBFSF85ZMVH3ZMG"]

def parse(self,response):

product\_name=response.css("span.B\_NuCI::text").extract()

product\_price=response.css("div.\_30jeq3.\_16Jk6d::text").extract()

product\_discount=response.css("div.\_3Ay6Sb.\_31Dcoz").css("span::text").extract()

product\_image=response.css("div.q6DClP::attr(style)").extract()

items=ProductsItem()

items["product\_name"]=product\_name

items["product\_price"]=product\_price

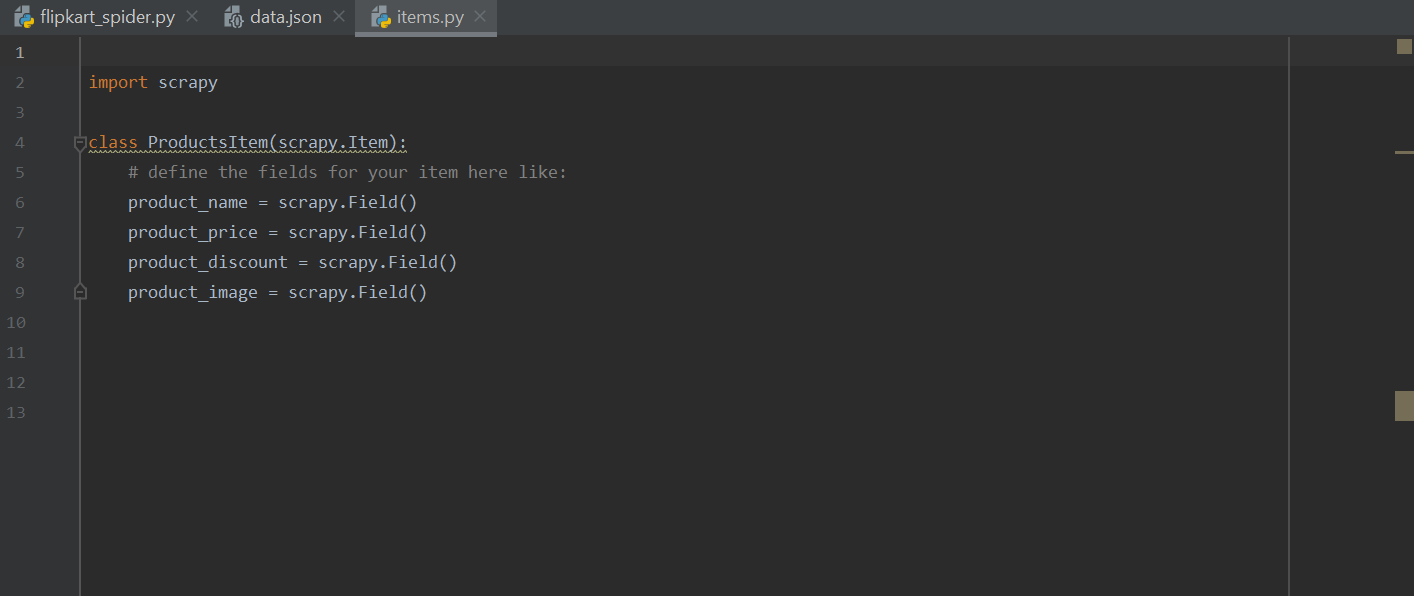
items["product\_discount"]=product\_discount

items["product\_image"]=product\_image

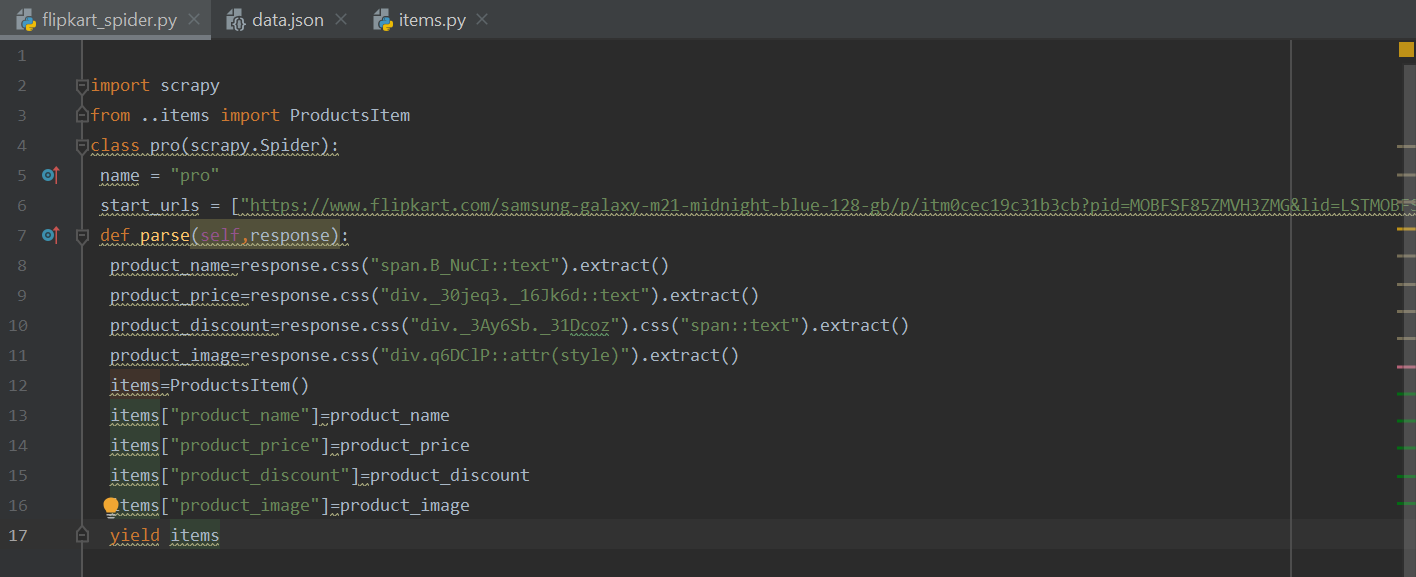
yield items

***Code Screenshot:***

***items.py***

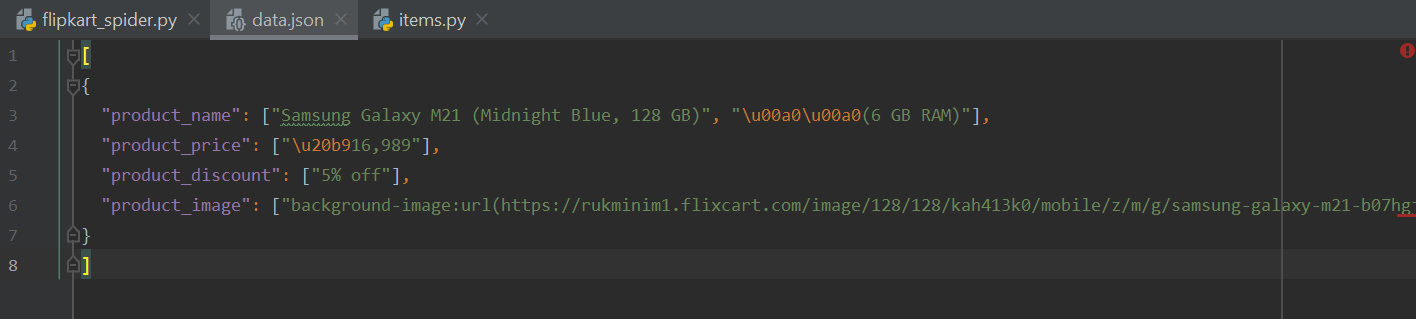


***flipkart\_spider.py***



***Output Screenshots :***

data.json



***data.json:***

[

{

"product\_name": ["Samsung Galaxy M21 (Midnight Blue, 128 GB)", "\u00a0\u00a0(6 GB RAM)"],

"product\_price": ["\u20b916,989"],

"product\_discount": ["5% off"],

"product\_image": ["background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfewggf3dqwc.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfewptm9vmsb.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfewxhmhptvw.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfewyghchvfy.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfgtjwa5vmgf.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/a/w/samsung-galaxy-m21-b07hgjj559-original-imafsfgtgxhpzpfr.jpeg?q=70)", "background-image:url(https://rukminim1.flixcart.com/image/128/128/kah413k0/mobile/z/m/g/samsung-galaxy-m21-b07hgjj55il-original-imafsfgtx2gty9tu.jpeg?q=70)"],

}

]