

# GROUP C ASSIGNMENTS

## ASSIGNMENT - 1

### Problem Statement:

Create a review scrapper for any ecommerce website to fetch real time comments, reviews, ratings, comment tags, customer name using Python.

```
In [1]: import requests
from bs4 import BeautifulSoup

HEADERS = ({'User-Agent':
            'Mozilla/5.0 (Windows NT 10.0; Win64; x64) \
            AppleWebKit/537.36 (KHTML, like Gecko) \
            Chrome/90.0.4430.212 Safari/537.36',
            'Accept-Language': 'en-US, en;q=0.5'})

# user define function
# Scrape the data
def getdata(url):
    r = requests.get(url, headers=HEADERS)
    return r.text

def html_code(url):

    # pass the url
    # into getdata function
    htmldata = getdata(url)
    soup = BeautifulSoup(htmldata, 'html.parser')

    # display html code
    return (soup)

url='''https://www.amazon.in/All-new-Echo-Dot-3rd-Gen/dp/B07PDHTHNN/
ref=zg_bs_amazon-renewed_scc1_1/259-2230848-1855512?pd_rd_w=
1tQ6G&content-id=amzn1.sym.7dd29d48-66c1-486c-
967d-2ed40101f2ea&pf_rd_p=7dd29d48-66c1-486c-967d-2ed40101f2ea&pf_rd_r=QHB1XFEH932
&pd_rd_wg=63E32&pd_rd_r=525b9c5f-c04d-
4b5b-9a84-c5adc9abda67&pd_rd_i=B07PDHTHNN&psc=1'''

soup = html_code(url)

def cus_data(soup):
    # find the Html tag
    # with find()
    # and convert into string
    data_str = ""
    cus_list = []

    for item in soup.find_all("span", class_="a-profile-name"):
        data_str = data_str + item.get_text()
        cus_list.append(data_str)
        data_str = ""
```

```
return cus_list
```

```
cus_res = cus_data(soup)  
print(cus_res)
```

```
['Echo Dot (3rd Gen)', 'Amazon Customer', 'Mohd Miraj', 'kailas a kulkarni', 'Vaib  
hav Ratan', 'Prithvi Joshi', 'Pratul Krishna gorain', 'Akshay', 'Neeraj mourya',  
'Preeti M.', 'Preeti M.']
```