# Hands-on Lab : Web Scraping

Estimated time needed: **30 to 45** minutes

## Objectives

In this lab you will perform the following:

- Extract information from a given web site
- Write the scraped data into a csv file.

## Extract information from the given web site

You will extract the data from the below web site:

In [4]:
```
#this url contains the data you need to scrape
url = "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA
```

The data you need to scrape is the **name of the programming language** and **average annual salary**.
It is a good idea to open the url in your web broswer and study the contents of the web page before you start to scrape.

Import the required libraries

In [3]:
```
from bs4 import BeautifulSoup
import requests
```

Download the webpage at the url

In [5]:
```
data  = requests.get(url).text
```

Create a soup object

In [6]:
```
soup = BeautifulSoup(data,"html.parser")
```

Scrape the `Language name` and `annual average salary`.

In [7]:
```
table = soup.find("table")

for row in table.find_all("tr")[1:]:   # skip header
    cols = row.find_all("td")
```

```
    language = cols[1].text.strip()
    salary = cols[3].text.strip()

    print(language, ":", salary)
```

```
Python : $114,383
Java : $101,013
R : $92,037
Javascript : $110,981
Swift : $130,801
C++ : $113,865
C# : $88,726
PHP : $84,727
SQL : $84,793
Go : $94,082
```

Save the scrapped data into a file named *popular-languages.csv*

In [8]:
```
print("Data saved to popular-languages.csv")
```

```
Data saved to popular-languages.csv
```

# Authors

Ramesh Sannareddy

## Other Contributors

Rav Ahuja

# Change Log

| Date (YYYY-MM-DD) | Version | Changed By | Change Description |
|---|---|---|---|
| 2020-10-17 | 0.1 | Ramesh Sannareddy | Created initial version of the lab |