➢ **Name-** Subarna Patra

➢ **Mob-** 8945980290

➢ **Email-** subarnapatra07@gmail.com

❖ Steps to complete the given task:

Using Python version 3.9 and Jupyter Notebook. Database using MySQL 8.0.

File Name: amazon_web_scrape.ipynb

● Step 1: Scrap the website

1.To scrap the amazon website:

In first step , Install required packages I.e. BeautifulSoup, requests, Pandas, Numpy library.
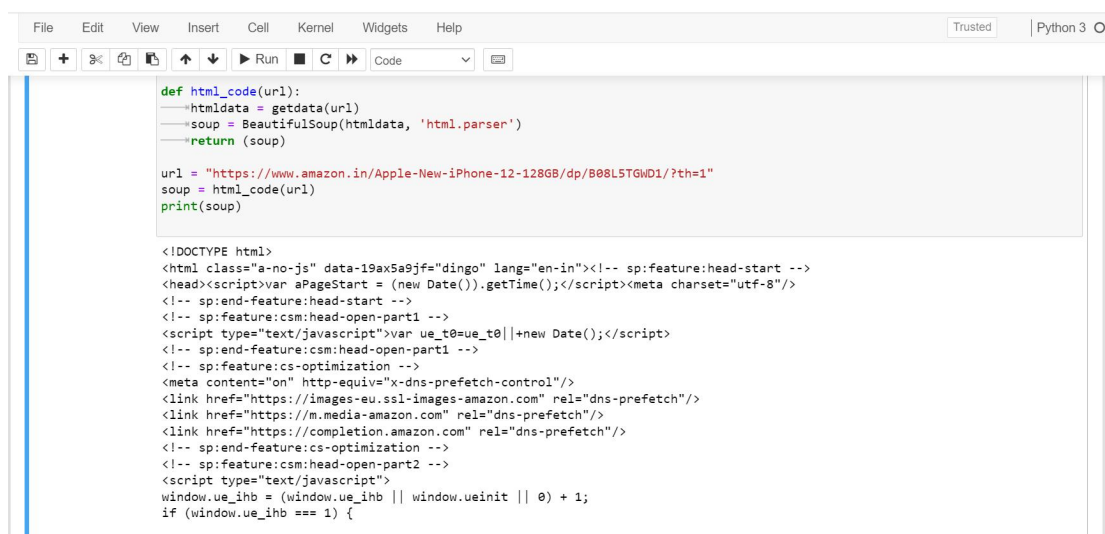
2.Then import all the required packages.

3.Scrap the whole amazon website HTML document by sending request through python code to amazon website and parsing it into the beautifulSoup.

4. We need here amazon web page url and the HTTP header.

5. Make the request to website and get the content from website and convert it to HTML format through Beautifulsoup.

6. Extract all the <a> tag links to get the all information through functions.
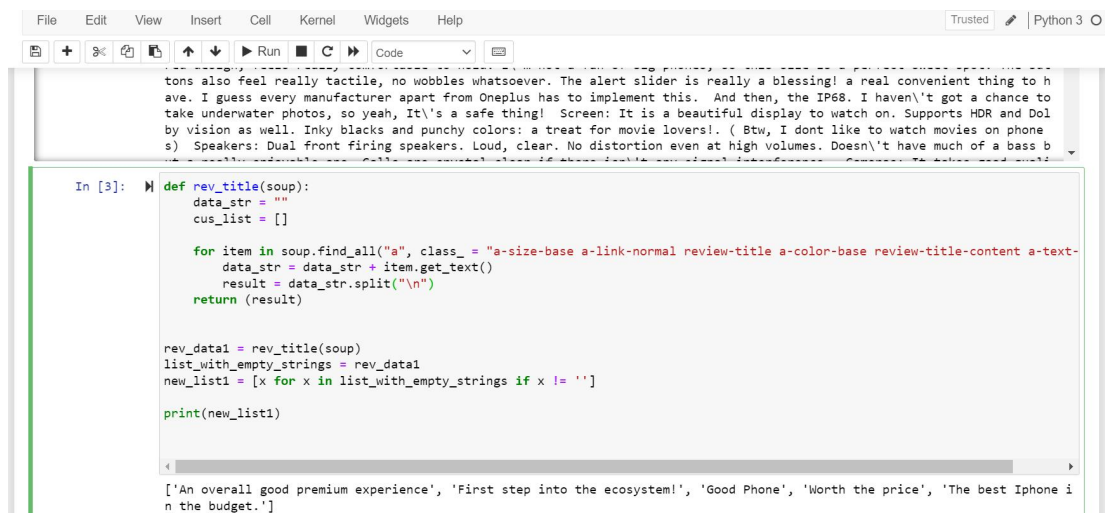
● Step 2: Extract desired information

1. Write the functions to extract all the desired information through the functions one by one. Here we use find() function to get the information through class, id…etc. Tags.

2. def review_title(soup): function get all the reviews titles for the particular given product.

3. def customer_review(soup): function get all the reviews text for the particular given product.

4. def storage_size(soup): function get all storage size purchased for the particular given product.

5. def verified_purchase(soup): function get if it is a verified purchase or not for the particular given product.

● Step 3: Store all information into csv file

1. After getting all the desired information through the functions from the amazon website, store all the information by creating a list

2. Then append it into a dictionary one by one into pandas dataframe.

3. Finally store all the data into the csv file through to_csv() function. It will save all information into a csv file in local storage. Here amazon_review.csv file is that file where we store all the data.

● Step 4: Store all information into a database (MySQL)

1. After storing all the data into csv file, now store it into the MySql database.
To do it at first read the csv file.

2. Then import all the packages which required and establish the connection to the database through host, using database name, user id and password.

3. Then create a table named amazon_data into the database 'world' and read csv data and insert it to the database table one by one.


```
        print("Table is created.")
        for i,row in amazon2.iterrows():
            sql = "INSERT INTO world.amazon_data VALUES (%s,%s,%s,%s,%s,%s)"
            cursor.execute(sql, tuple(row))
            print("Record inserted")
            conn.commit()
except Error as e:
        print("Error while connecting to MySQL", e)

You're connected to database:  ('world',)
Table is created.
Record inserted
Record inserted
Record inserted
Record inserted
Record inserted
```

4. In next step see all the store data into the database through the above operations.

```
In [21]:   # Show store data in Database table
           sql = "SELECT * FROM world.amazon_data"
           cursor.execute(sql)
           result = cursor.fetchall()
           for i in result:
               print(i)
```

('0', 'An overall good premium experience', 'SUMMERY: As a first time iPhone user, I would say one should definitely buy iPh
one ONLY for premium built & premium hardware (will talk later) while the whole iOS experience is pretty overrated. iPhones
isn\'t any exceptionally fast with normal apps like YouTube, Insta etc, Android phones are equally snappy with them in 2022.
PROS:1. Impressive hardware, the heptics, the autobrightness & otherwise highly ignored things in Android are way too perfec
tly executed in iPhone and it DOES make a big difference in user experience!!2. The screen is AMAZING.. it doesn\'t have 90
or 120 hz refresh rate but it FEELS better than my Nord with 90 hz!!3. A14 bionic promises years of performance ahead.4. Bat
tery life is somewhere in between good & great.5.6. The dimensions of phone, it\'s smaller than usual android phones and fee
ls much more handy along with less weight.7. Overall build quality is great.8. Best audio output out there.9. Long term supp
ort.CONS:1. SLOW CHARGING. 20 W only, you\'ll be jealous of 1/3rd price phones with 60 W charging. Yes that much high chargi
ng will effect health of battery but 20 W is way way too less, if Apple is such a legendary company they should have created
a battery that can withstand decently fast charging while instead they just use slow charging to keep their battery\'s healt
h.2. Overy exhagerated software experience, reality is that Android is more user friendly with better gesture & commonsense
options & settings & steps to do simple things. iOS looks pretty yes but UX engineers at Apple have really done a crappy jo
b.. there\'s still no notification grouping in correct manner (its executed poorly), no proper homescreen customisation & ju
st so so so many "commonsense" things are missing from iOS I can rant all day about it. Everything you can do in Android wit
h a single click or swipe takes 2 swipes and then a click to do in iOS its SOOOOO ANNOYING. Like you want to clear a notific
ation? Swipe and then click on clear. You want to open a notification? Swipe FIRST and then click open.. like wth????????3.
DON\'T believe Apple fanboys who scream about speed, Android phones are equally snappy in performance today. There are split
second app loading times in iOS also, just like in Android. And yeah iOS DOES HAVE occasional bugs too, apps taking too much
time sometimes etc etc.4. FaceID is nice but an in-fingerprint reader seems the best ID to me atleast, that\'s missing.5. Ri
diculously expensive accessories6. Battery life is above good but not great.In the end I would say iPhone, overall, does pro

● Step 5: Sentiment analysis and keyword frequency

1. In next step parse only all the reviews data to do the sentiment analysis further on it.

2. Used Text Blob to do the sentiment analysis of the review data.

3. After the analysis polarity of each reviews and subjectivity of the review is shown.

```
word_counts = count_words_fast(text)
(num_unique, counts) = word_stats(word_counts)
print("In the reviews, unique keyword number is: ",num_unique," And total frequency of all unique keyword is: ", sum(counts))
```

```
In the reviews, unique keyword number is:  585  And total frequency of all unique keyword is:  1264
```

```
df['polarity'] = df.apply(lambda x: TextBlob(x['review']).sentiment.polarity, axis=1)
df['subjectivity'] = df.apply(lambda x: TextBlob(x['review']).sentiment.subjectivity, axis=1)

print(df)
```

```
                                        review  polarity  subjectivity
0  SUMMERY: As a first time iPhone user, I would ...  0.181879      0.482705
1  I have been a smartphone user since 2015, used...  0.229815      0.478056
2                          Good & value for money  0.700000      0.600000
3  I migrated to iOS after using Android for more...  0.263784      0.556015
4  IPhone 12 is probably the best option to start...  0.403529      0.443922
```

---Thank you---