A COMPARISON OF USER EXPERIENCES ACROSS ONLINE LEARNING PLATFOEMS [Case Study: EdxOnline, Udemy, Shaw Academy]

BY

SHAW OLUWAGBOUNMI PRECIOUS

STUDENT ID: ST20182956

PROGRAM: Msc. DATA SCIENCE (INTERNSHIP)

COURSE: SOCIAL MEDIA ANALYTICS FOR BUSINESS

Table of Contents

[1.0 BIOGRAPHY 4](#_Toc57988876)

[2.0 ABSTRACT 5](#_Toc57988877)

[2.1 PURPOSE 5](#_Toc57988878)

[2.2 METHODOLOGY 5](#_Toc57988879)

[2.3 FINDINGS 5](#_Toc57988880)

[2.4 SOCIAL IMPLICATION 5](#_Toc57988881)

[2.5 BUSINESS IMPLICATION 5](#_Toc57988882)

[2.6 RESEARCH LIMITATIONS 5](#_Toc57988883)

[3.0 INTRODUCTION 6](#_Toc57988884)

[3.1 AIM AND OBJECTIVES 7](#_Toc57988885)

[4.0 DATA SOURCE 7](#_Toc57988886)

[4.1 Twitter 8](#_Toc57988887)

[4.2 TrustPilot 8](#_Toc57988888)

[5.0 METHOD FOR DATA SCRAPING 9](#_Toc57988889)

[5.1 Twitter API 9](#_Toc57988890)

[5.1.1 Accessing Twitter API in Python 9](#_Toc57988891)

[5.2.2 Test 9](#_Toc57988892)

[5.2 Python 10](#_Toc57988893)

[5.2.1 Import Libraries 10](#_Toc57988894)

[5.2.2 Examine Website 10](#_Toc57988895)

[5.2.3 Build Scrapper With Python 11](#_Toc57988896)

[5.2.4 Test 12](#_Toc57988897)

[6.0 METHODOLOGY FOR SENTIMENT ANALYSIS 14](#_Toc57988898)

[6.1 Python 14](#_Toc57988899)

[6.1.1 Test 14](#_Toc57988900)

[7.0 RESULTS 17](#_Toc57988901)

[7.1 Tweets 17](#_Toc57988902)

[7.1.1 Shaw Academy Tweets 17](#_Toc57988903)

[7.1.2 Udemy Tweets 17](#_Toc57988904)

[7.1.3 EdxOnline Tweets 17](#_Toc57988905)

[7.2 Reviews 17](#_Toc57988906)

[7.2.1 Shaw Academy Reviews 17](#_Toc57988907)

[7.2.2 EdxOnline Reviews 17](#_Toc57988908)

[8.0 VISUALIZATION 17](#_Toc57988909)

[9.0 BUSINESS STORY TELLING 18](#_Toc57988910)

[10.0 CONCLUSION 19](#_Toc57988911)

[10.1 Review Aim and Objectives 19](#_Toc57988912)

[10.2 Recommendations 19](#_Toc57988913)

[REFERENCES 19](#_Toc57988914)

[APPENDIXES 21](#_Toc57988915)

[Appendix A - Twitter API Test 21](#_Toc57988916)

[Appendix B – Web Scrapping Using Python Test 21](#_Toc57988917)

[Appendix C – Sentiment Analysis Test 21](#_Toc57988918)

[Appendix D – Tableau Data Visualisation Test 21](#_Toc57988919)

[Appendix E – Result CSV Files 21](#_Toc57988920)

# 

# 1.0 BIOGRAPHY

Shaw Oluwagbounmi is a data scientist who runs her own consultancy firm, Kay Kay Analytics and Consultancy. Oluwagbounmi believes that data analytics plays a huge role in helping companies and organisations evaluate their progress, make informed decisions, recommend new operational methods and/or amend old ones.

She is of the opinion that data analytics is an integral part of business operations.

Her firm, Kay Kay Analytics and Consultancy has been nominated thrice for the Market and Research Awards (MRS) and has won twice.

Oluwagbounmi holds a degree in BSc. Industrial Mathematics from Covenant University, Nigeria and MSc Data Science from Cardiff Metropolitan University.

# 2.0 ABSTRACT

## 2.1 PURPOSE

The purpose of this study is to provide potential online learning platform users with information needed to select the platform from Shaw Academy, EdxOnline and Udemy.

## 2.2 METHODOLOGY

This study employs the use of Twitter API and Python for data collection, Python for sentiment analysis, Python and Tableau for data visualisation.

## 2.3 FINDINGS

Most of these online learning platforms are so not different each other in terms of user experience and reviews.

## 2.4 SOCIAL IMPLICATION

The social implication of this research is that it points potential users in the right direction, prevent them from falling victims to fraudulent activities.

## 2.5 BUSINESS IMPLICATION

The business implication of this study that it might affect the opinion people have about the considered online platforms and may deter them using the platform.

## 2.6 RESEARCH LIMITATIONS

The limitation of this research is that it is based off the opinion and experiences of other people. The amount of data collected is only a tiny fraction to the amount of data available for reviews.

# 3.0 INTRODUCTION

Over the years, there has been constant improvement and increase in the use of technology which spun across various ministries, industries and businesses such as finance, agriculture, education, auditing, security, health, manufacturing and automotive sectors etc. This has caused a lot of businesses to go digital, this means that a lot of businesses now operate online and education is not left out.

As important as gaining knowledge is, factors like time and money come into play as to why a lot of people do not further their education beyond high school or university degree. Hence, online learning platforms were birthed such as Ruzuku, Udemy, EdxOnline, Shaw Academy, Skillshare, Treehouse etc. as solutions to problem of ‘why can’t we learn and get certificates without going to school?’

Some advantages of online learning platforms are:

* It provides flexibility: this means that most courses available are self-paced i.e. each enrolled student can start and continue as they please.
* Enables career advancements: online platforms help members of the working population get certified in career advancing courses such as project management and the likes.
* Helps develop new hobbies and imbibes self-disciple and responsibility.

The National Centre for Education Statistics estimates the number of people that participates in at least one online course to about 5.4 million per year. The implication of this is that an average member of the society has enrolled for an online course. This then yields the following questions for individuals that have interest in enrolling for online courses:

* How was learning online?
* Were the courses detailed enough for self-learning?
* What is the price-range for the platform?
* Is the user interface friendly?

Individuals that are interested in enrolling for online courses tend to seek answers to the above question in order to be able to determine which platform is best suitable for them.

This study considers and compares user experiences across Shaw Academy, Udemy and EdxOnline to determine how users feel about using them and how they rate them.

## AIM AND OBJECTIVES

The aim of this study is to analyse data given in form of tweets and reviews, compare the opinion of Twitter users to that of reviews given on a review website and determine if there is a similarity.

The objectives of this study are:

* analysing the data by sentiments and charts (determine if how the users feel).
* determine which of the three platforms has the best user experience

# 4.0 DATA SOURCE

For the purpose of this study, data was scrapped from two sources, namely:

## 4.1 Twitter

Twitter is a form of social media site that allows its users share their experiences, thoughts and ideals with other people. Twitter has gone beyond being just a social platform to being a community that caters for individuals, companies and organisations.

Twitter is a great source of data because millions of people have access to Twitter, relevant topics and events are discussed and dissected there. Topics including: politics, science, religion, sports, computing, television shows and more.

There are various reasons why Twitter is a data source such as:

* **Data from differing sources**: Twitter is accessible by anybody, hence the information could be include official, non-official or any form of media. Mixed origin of information gives multifaceted view on the effect of a particular event and measures taken to deal with the event.
* **Embedded content**: Twitter users are permitted to insert pictures, videos and other elements of disturbance and quantitatively and visually.

From Twitter, by using “shawacademy”, “udemy”, and “edx” as the search variables, 100 tweets were scrapped from Twitter for each variable.

## 4.2 TrustPilot

TrustPilot is an open online review platform where any user, costumer or consumer can write a review about any company they have or have had purchasing or service encounter with. TrustPilot has over 1.1 trillion reviews and ratings for over 400,000 thousand businesses.

This website was selected as a data source because it is very rich in data with users stating their experiences with a company and even rating them.

From TrustPilot, the first 60 reviews was scrapped for each platform Shaw Academy, Udemy and EdxOnline as at the time of the scrapping.

# 5.0 METHOD FOR DATA SCRAPING

## 5.1 Twitter API

Using Twitter API to get data from Twitter is the best option as it allows you read and write to Twitter. It permits you to read profiles and tweets, access your follower’s data and high quantity of tweets on a particular event or subject.

To access Twitter API:

* You must have an active Twitter account
* Apply for developer access
* Once the access has been granted, save your keys and tokens.

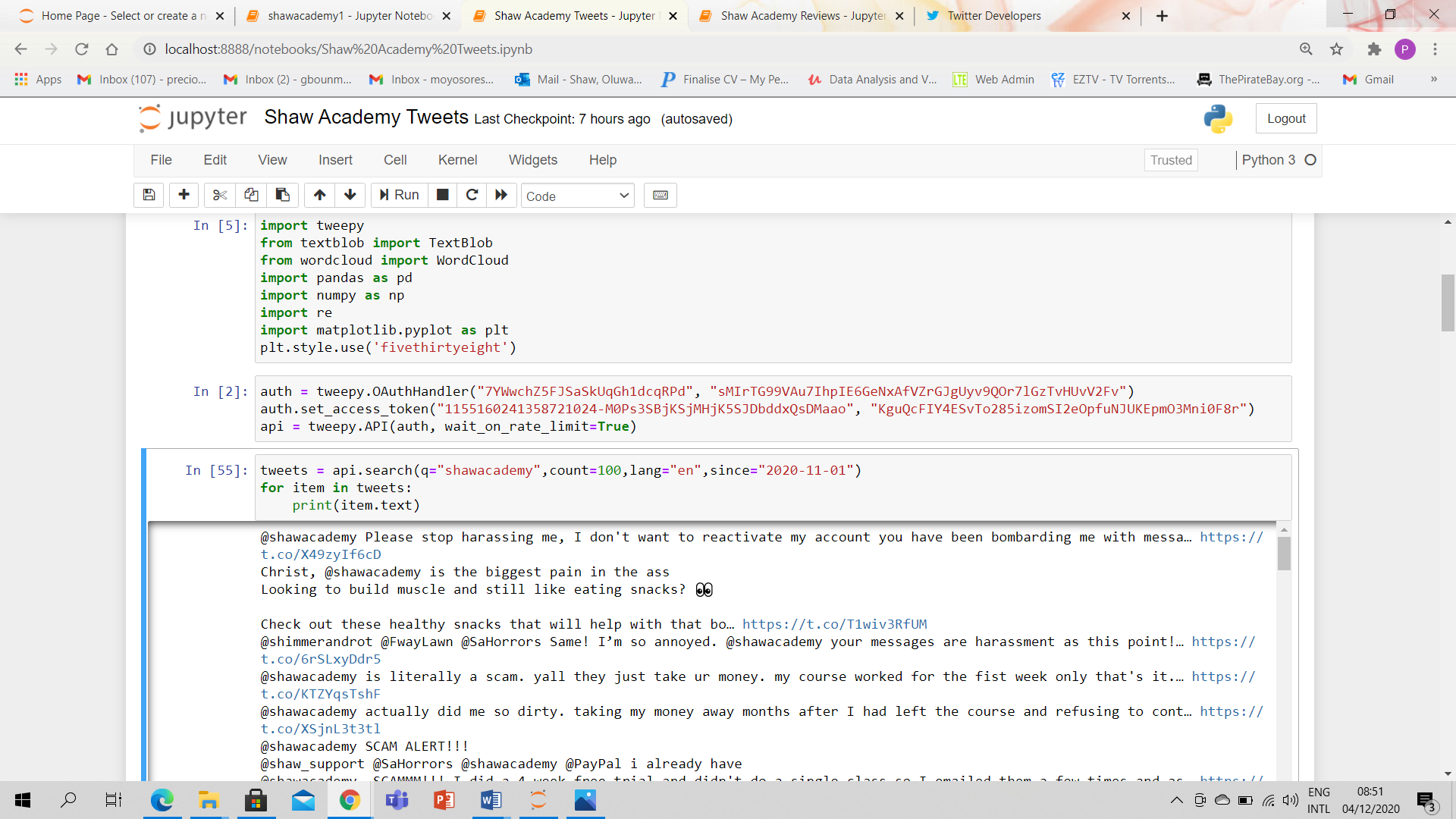
### 5.1.1 Accessing Twitter API in Python

Python is a programming language that is easy to use and has an extensive amount available libraries, resources and immerse community participation. Python enables developers rollout programs and run prototypes which makes development very fast. Python allows the user to create CSV files which makes the output simple and efficient for reading the data.

To access Twitter API using Python:

* You need consumer key, consumer secret key, access token key, access token secret key.
* Install tweepy on your terminal.

### 5.2.2 Test



## 5.2 Python

### 5.2.1 Import Libraries

Before data can be scrapped from a website using Python, the necessary libraries must be imported as seen below:

import requests

from requests import get

from bs4 import BeautifulSoup as bs

import pandas as pd

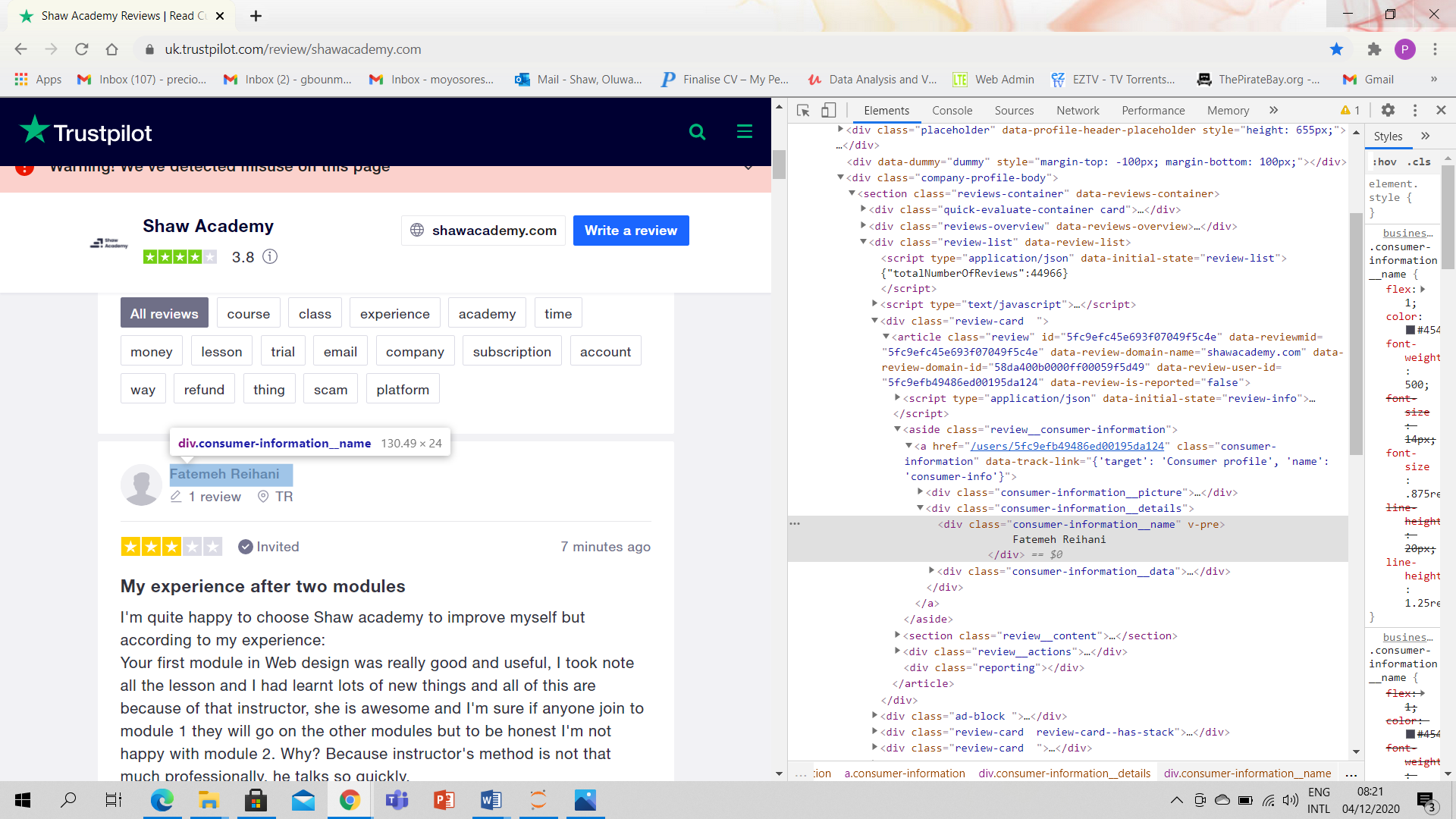
import numpy as np

### 5.2.2 Examine Website

To scrap data from the website, we must study the pagination to know how the increment of the page works. To do this, click on the next button and study how the link to know how the first page is different from the current one.

In this case, the first page is <https://uk.trustpilot.com/review/shawacademy.com?page=1> and the next page is <https://uk.trustpilot.com/review/shawacademy.com?page=2> . It can be observed that the page increment is in [page=‘number of page’].

Next, we inspect the data value of the elements we want to scrap from the page. This is done by right clicking on the element and clicking on inspect. The source code of the page opens up at the right corner of the screen. In the source code, we can see the html tags and class of the element we want to scrape.

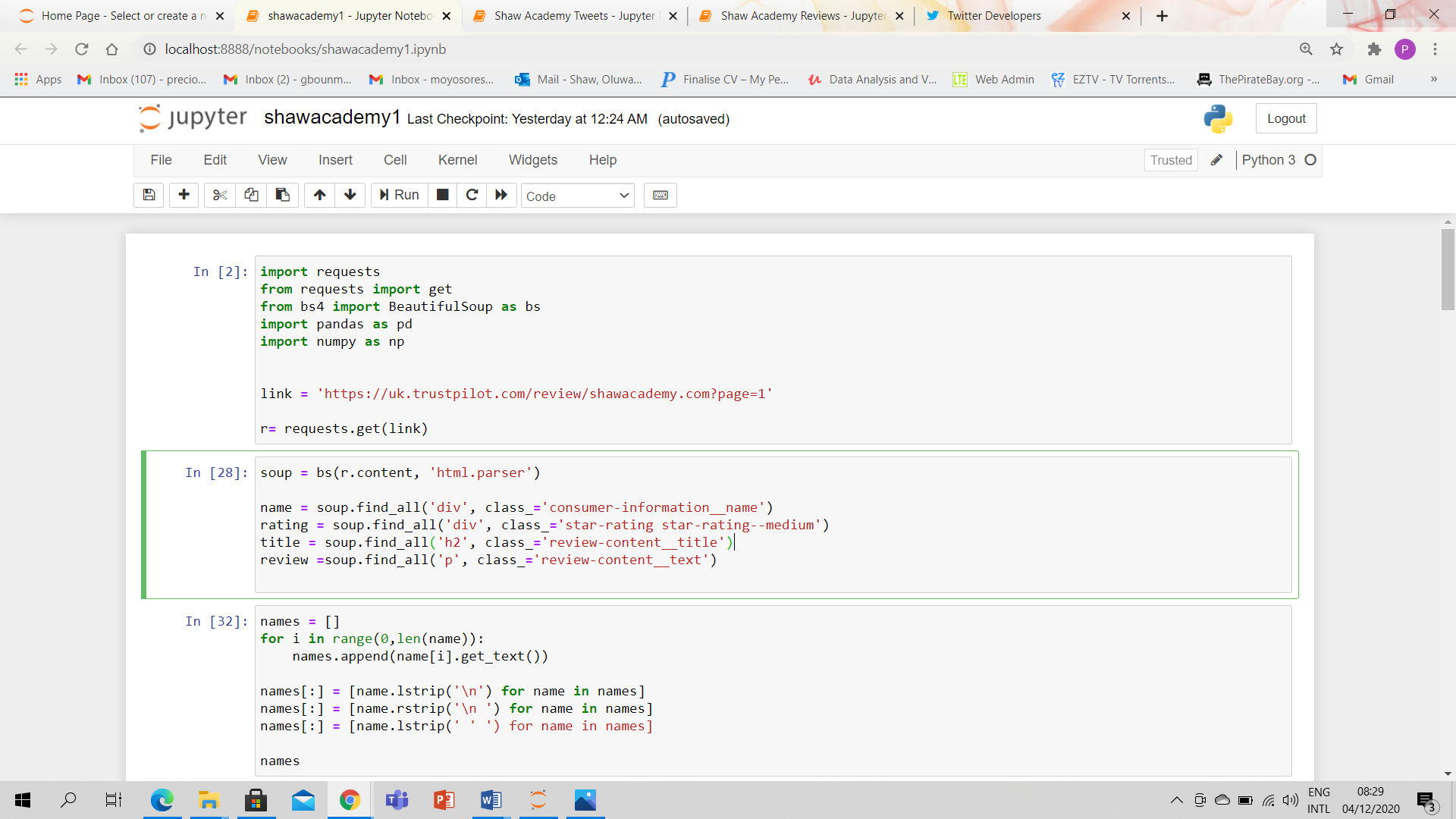


This picture is an illustration of right clicking on the name of the individual that gave the review. We can see that the information is contained in a div class and has a class of ‘customer-information\_\_name’. The same process is repeated for all the elements we want to scrap.

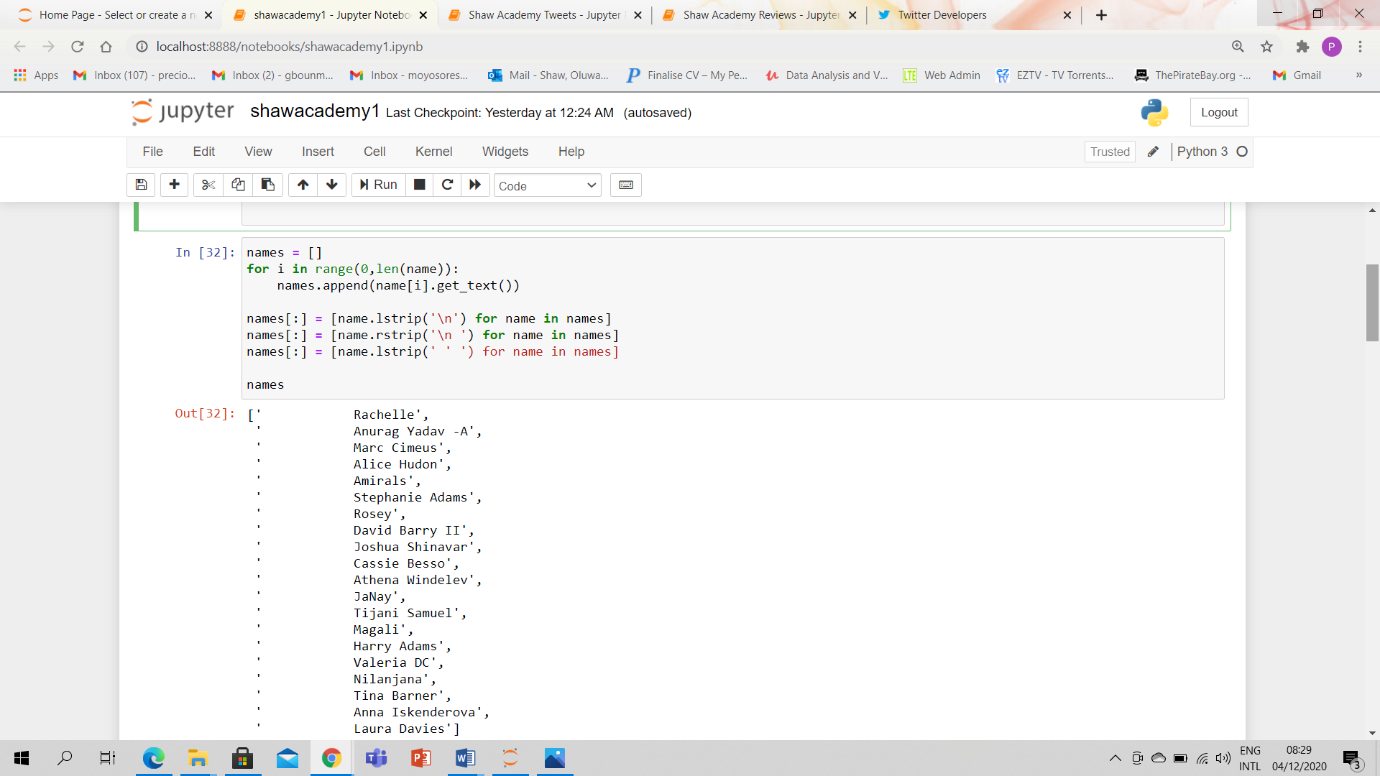
### 5.2.3 Build Scrapper With Python

After examining the website, build your scrapper using Python and BeautifulSoup.

* First, create a variable for the link then make a request to the link
* Create a variable for beautifulsoup and parse the html parser to avoid errors
* Scrape the name, rating, review and review title of the users on the first page by using the code shown in the picture below
* Strip the data of any excess spaces and \n

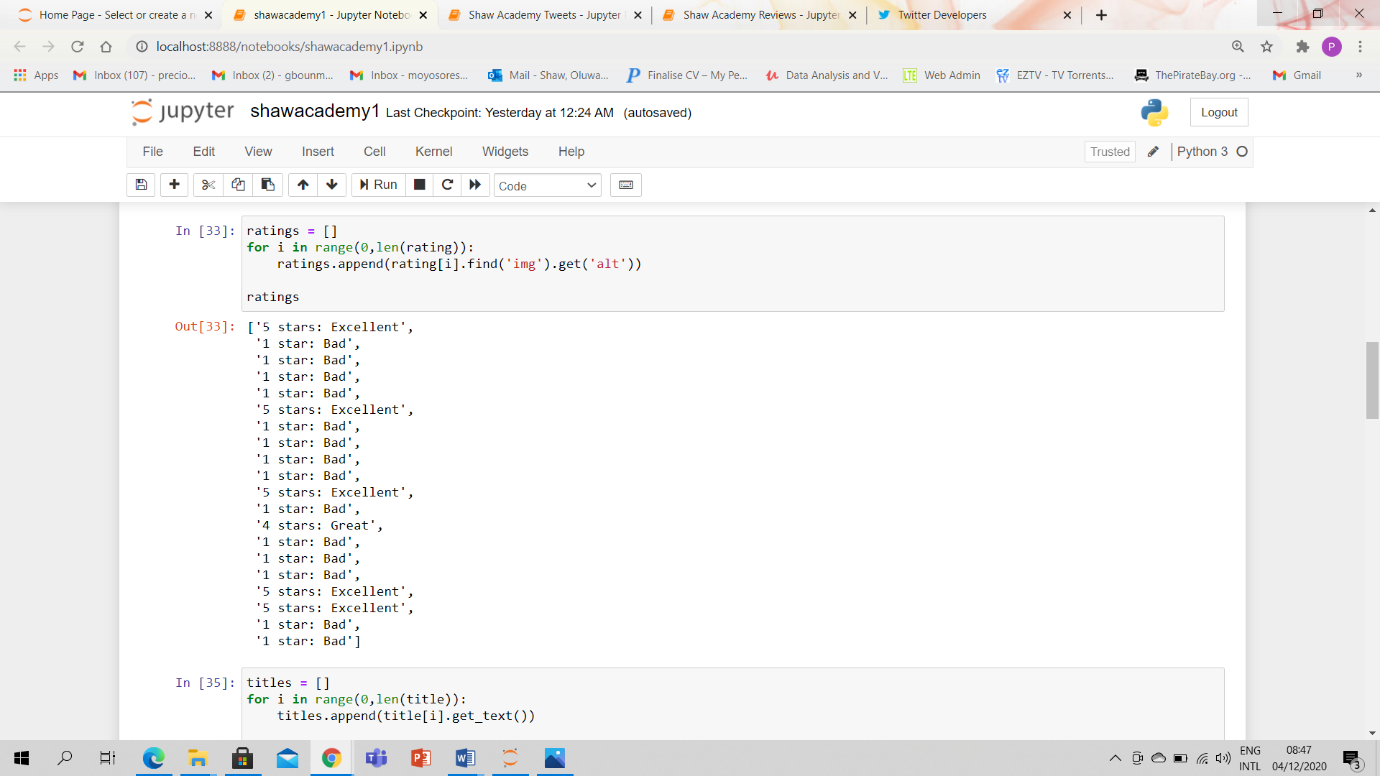


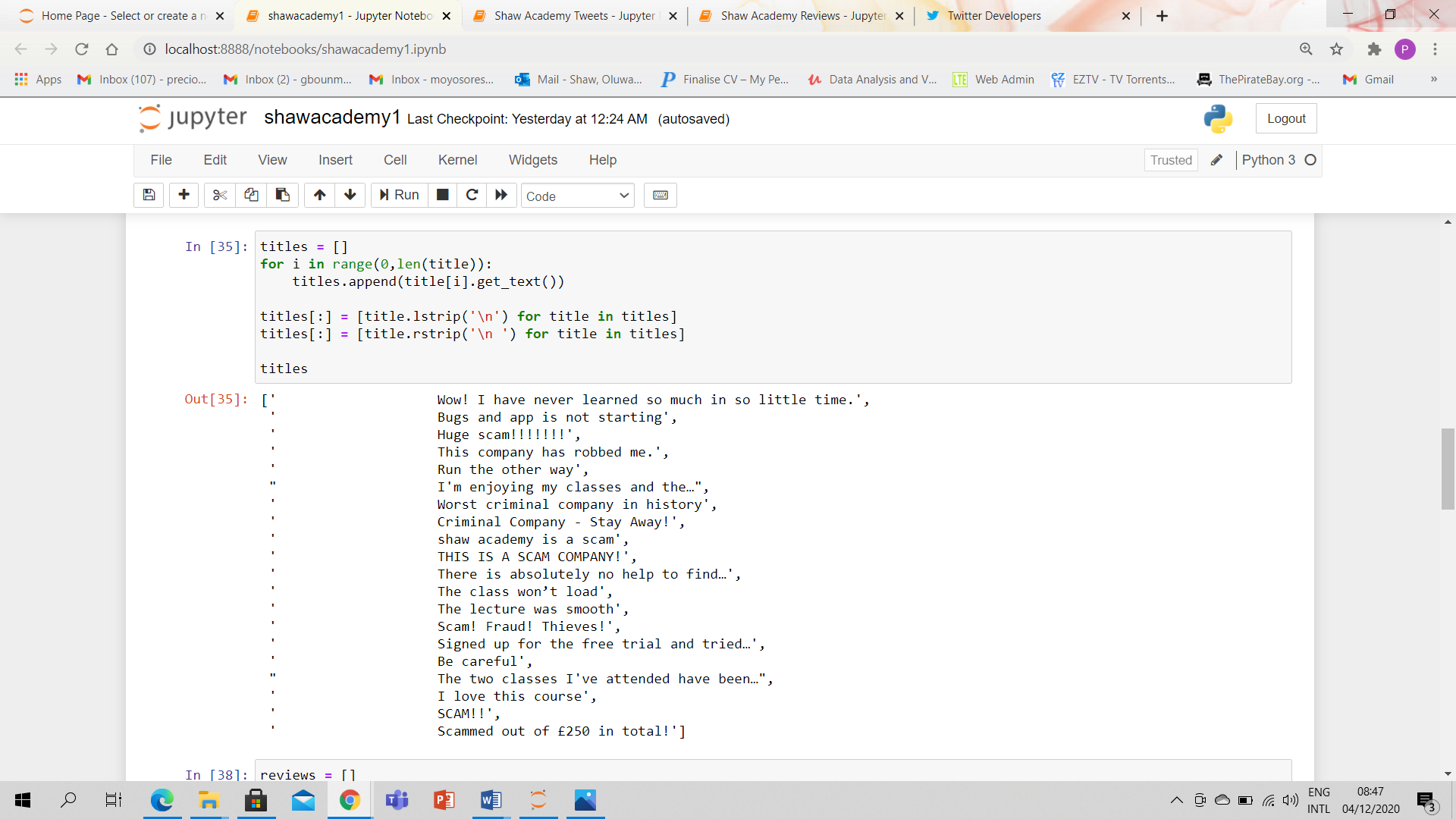
Stripping the result should give us just the names of the customers as seen below

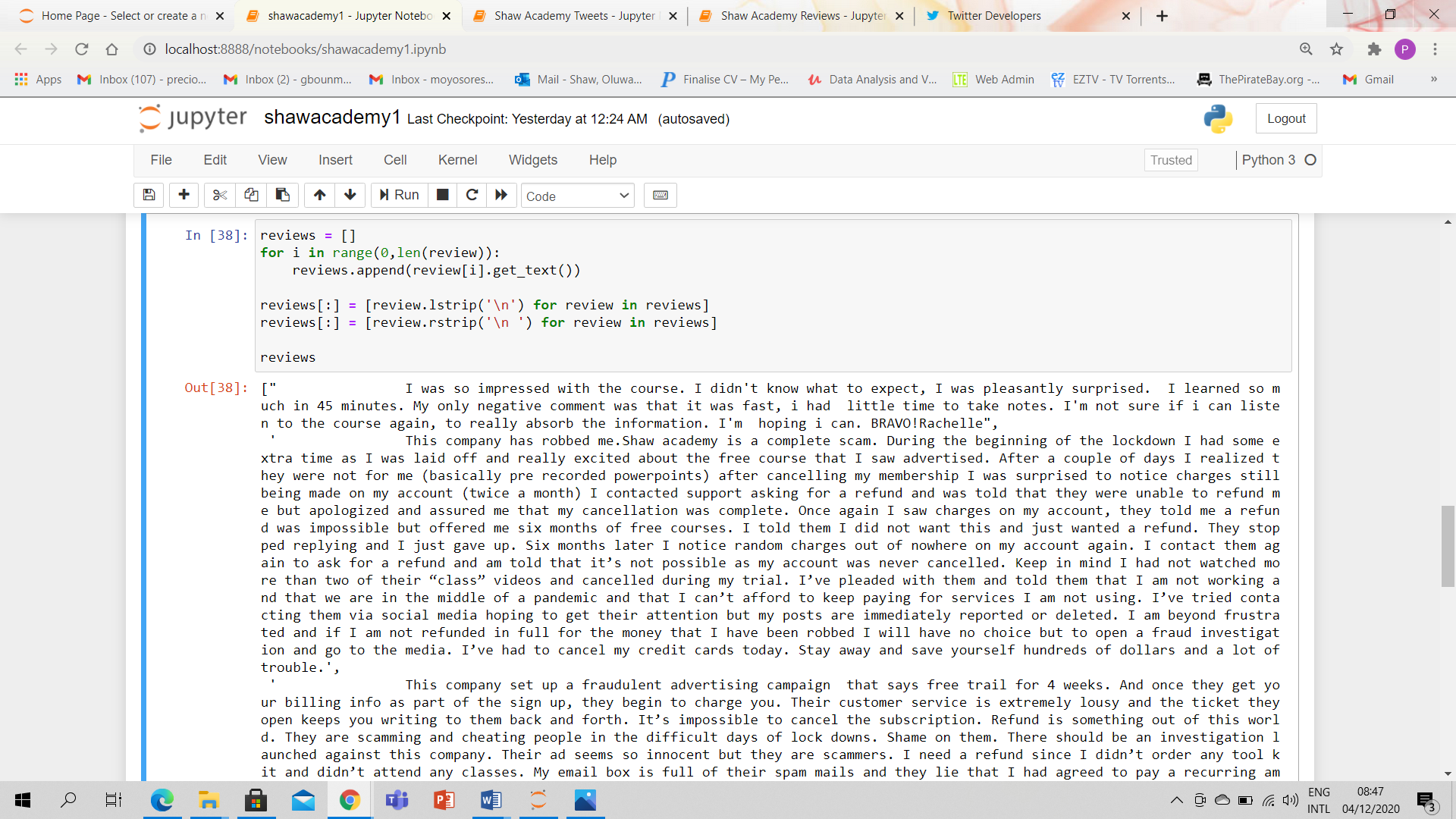


The process is repeated for all elements i.e. rating, review and review title.

### 5.2.4 Test







# 6.0 METHODOLOGY FOR SENTIMENT ANALYSIS

Sentiment analysis is the technique of establishing the emotional tone behind a text. It involves understanding the viewpoint, thoughts and feeling conveyed in text.

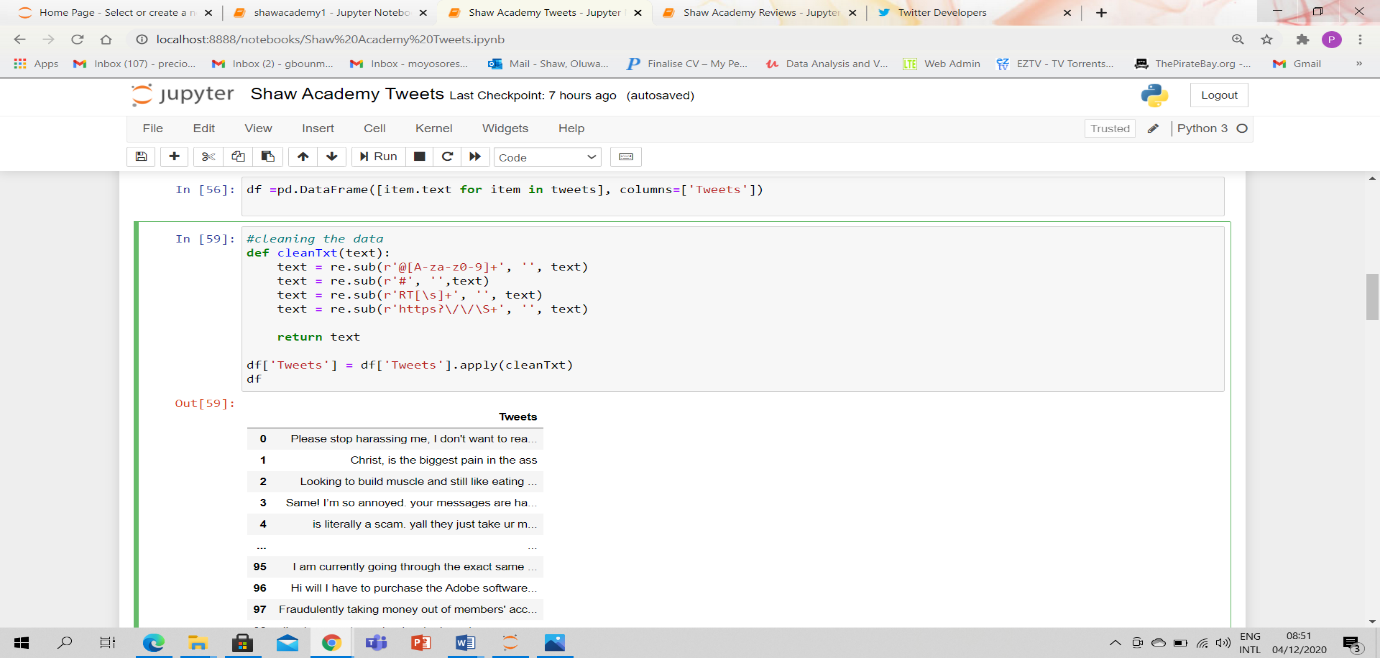
## 6.1 Python

Using Python, the level of personal opinion (subjectivity) and whether the text is positive, negative or neutral (polarity) can be determined.

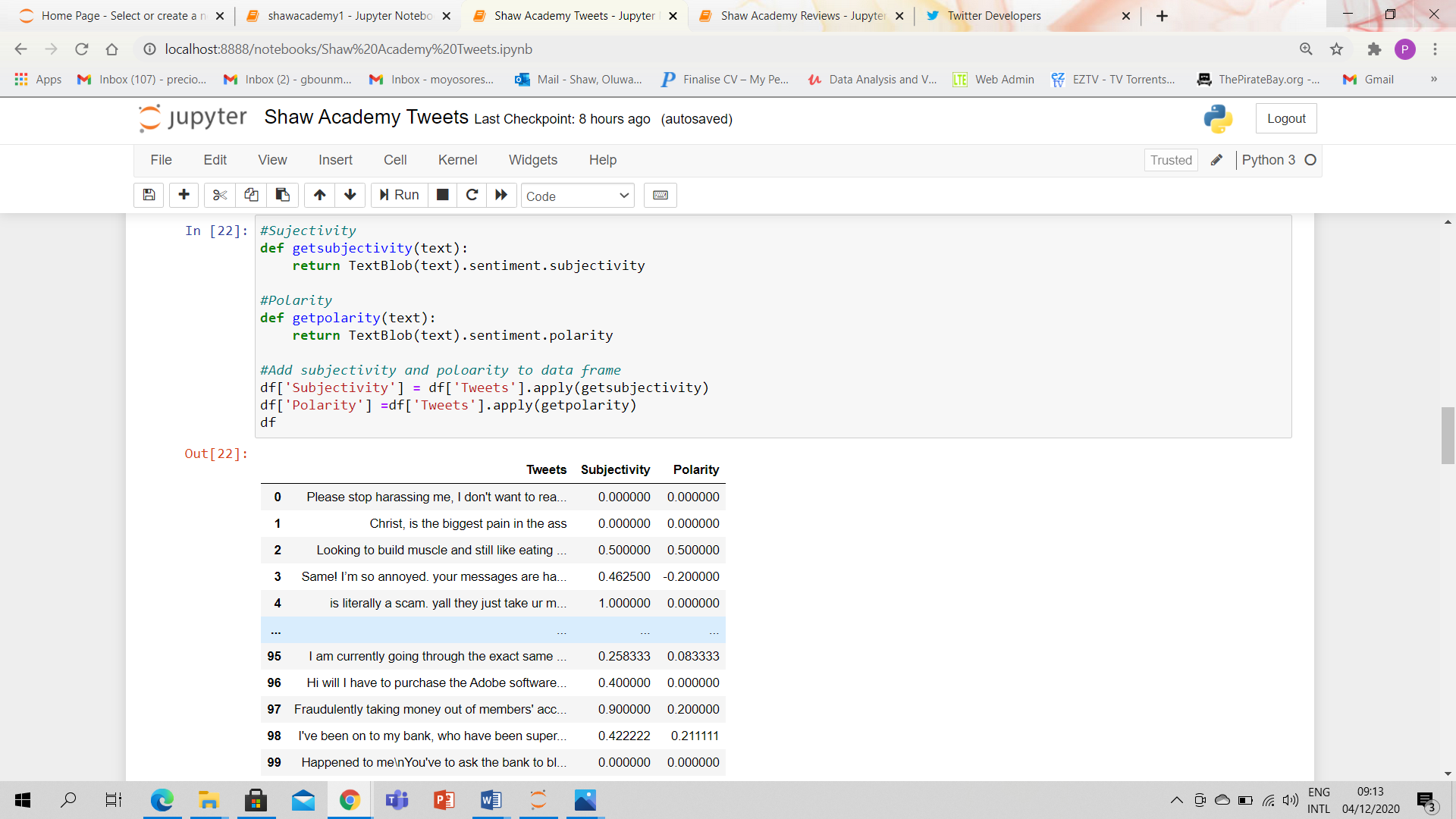
With Python it is very easy to get the subjectivity and polarity of text data with a simple code. The subjectivity of the customers is very important given their review/ tweets is as a result of their personal experience(s) with the online learning platforms. The polarity comes to play in determining if their experiences were positive, negative or neutral.

### 6.1.1 Test

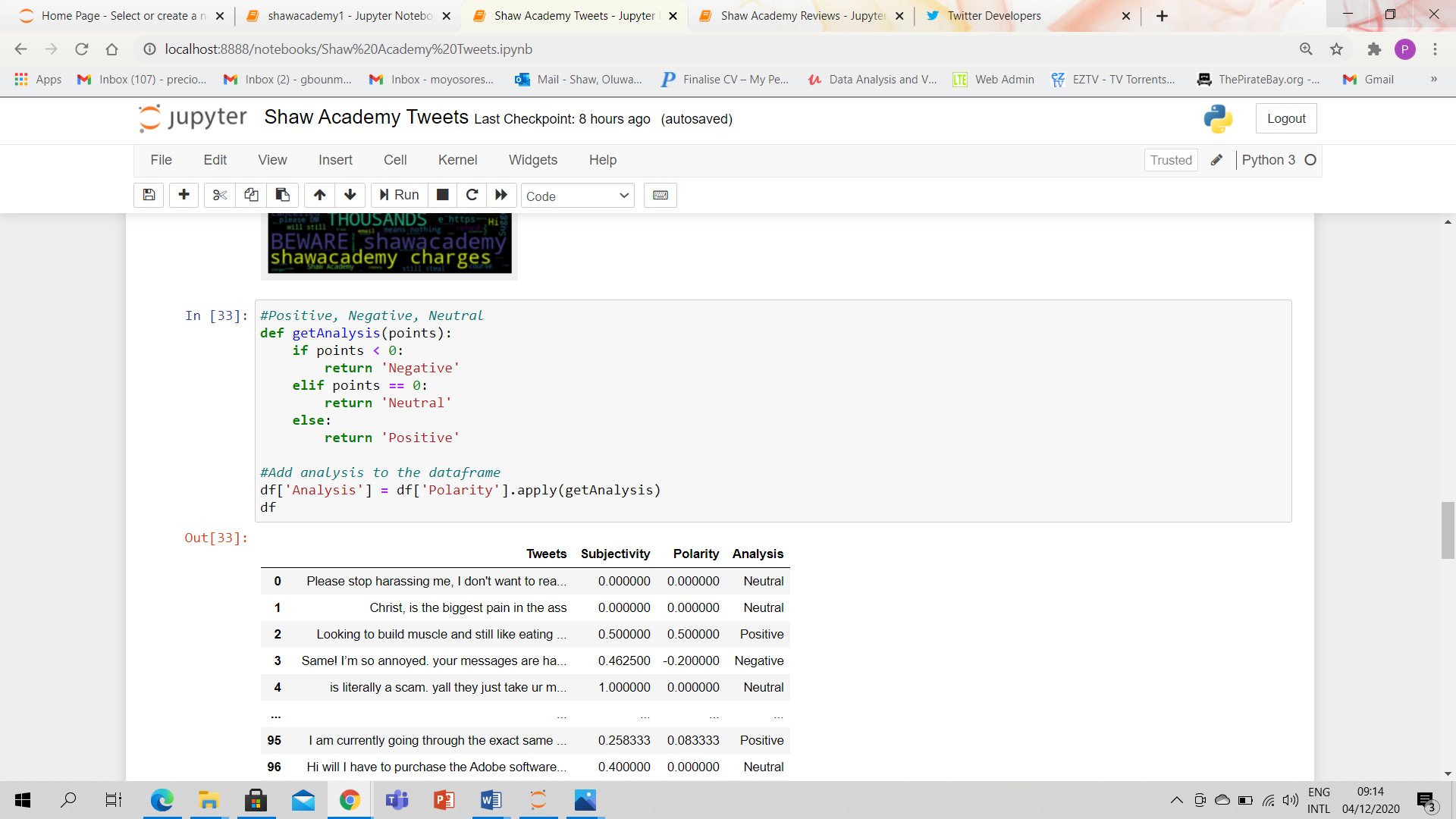
The picture below shows the cleaning of the data i.e. getting rid of bits of the data that is not needed for the sentiment analysis.



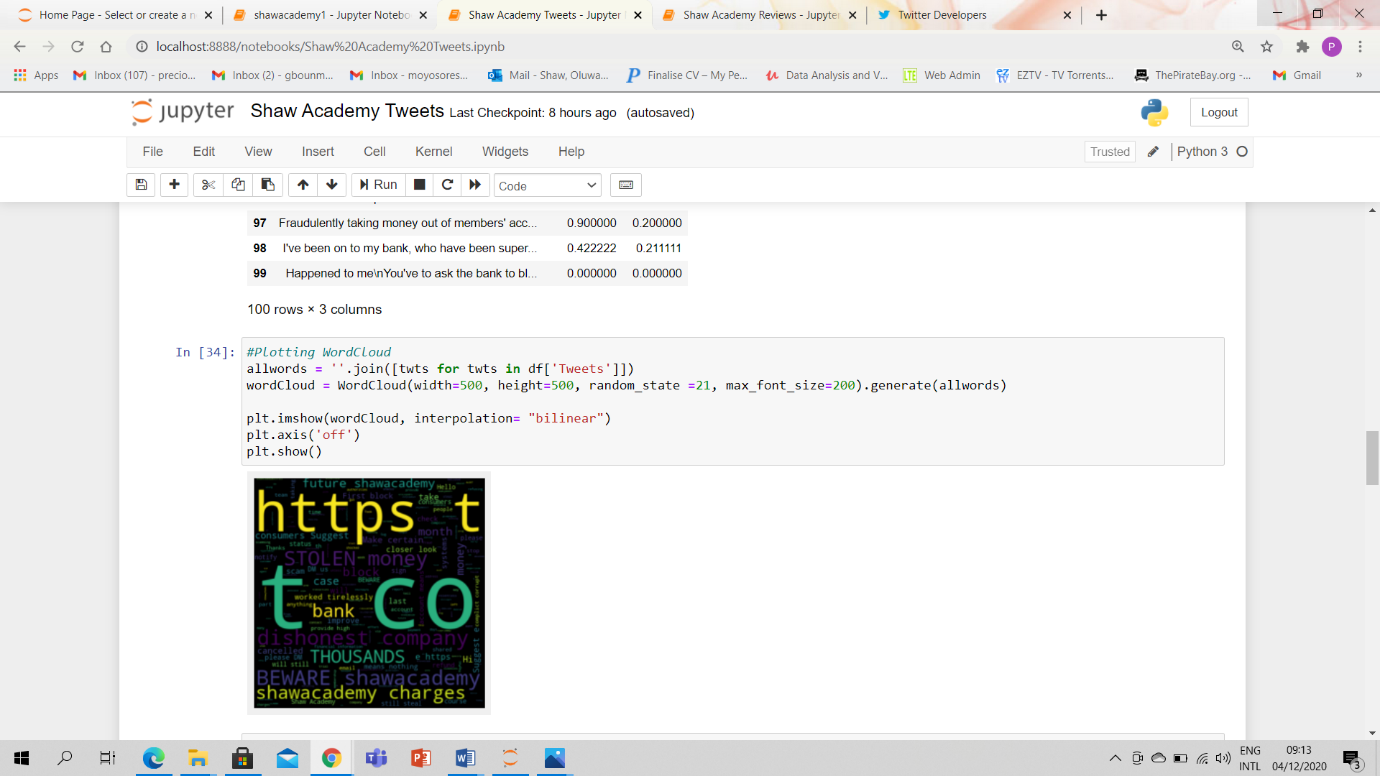
This picture shows the creation of the getsubjectivity and getpolarity functions



The picture below depicts defining an analysis function that prints the polarity of the text as positive, negative or neutral.



This picture shows the word cloud of the tweets collected.



# 7.0 RESULTS

## 7.1 Tweets

### 7.1.1 Shaw Academy Tweets

From the word cloud for Shaw Academy tweets, it is evident that a lot of its users complaints about charges and stolen money. The polarity shows 65% of the user tweets are neutral, 19% of user tweets are positive and 16% of user tweets are negative. The subjectivity shows that the experiences are unique to each user.

### 7.1.2 Udemy Tweets

The word cloud for Udemy tweets shows that most of the users its enrolled for computing courses as we can see words like coding, data science, advanced nlp etc. The polarity shows that 48% of the user tweets are neutral, 51% of the user tweets are positive and 1% of the user tweets is negative. The subjectivity here also shows that user experiences are unique to each user.

### 7.1.3 EdxOnline Tweets

The word cloud for EdxOnline tweets is not definitive on what its users are tweeting about as there are so many different words with no correlation. On the other hand, the polarity shows that 71% of the user tweets are neutral, 27% of the user tweets are positive and 2% of the user tweets are negative. The subjectivity shows that each tweet is a function of the user’s experience.

## 7.2 Reviews

### 7.2.1 Shaw Academy Reviews

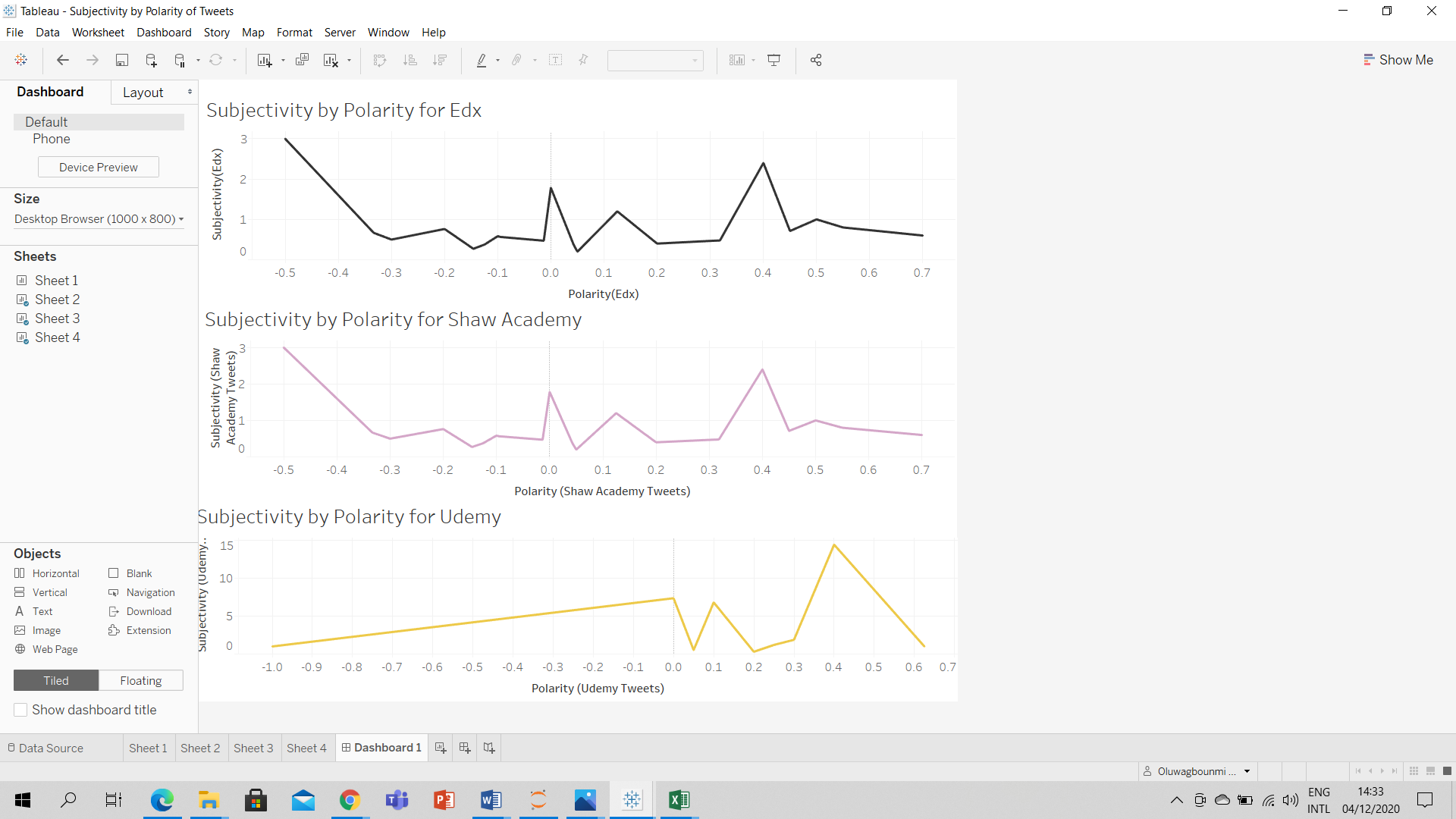
The word cloud for Shaw Academy reviews fromTrustPilot explicitly displays scam and fraud. The polarity shows that 43.33% of the reviews are neutral, 36.67% of the reviews are positive and 20% of the reviews are negative.

### 7.2.2 EdxOnline Reviews

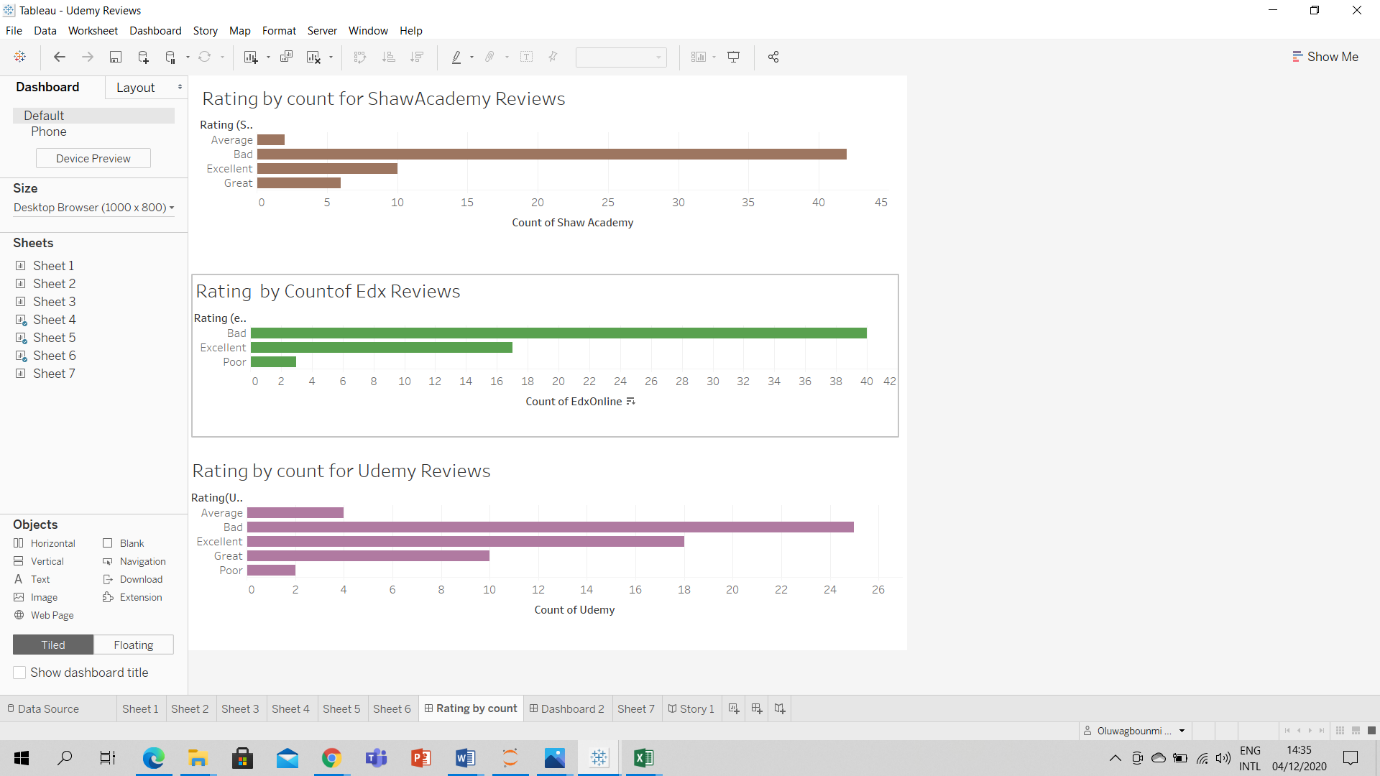
The word cloud for EdxOnline reviews is not so definitive but there is correlation between some words like terrible, error, deception which shows that there are complaints from Edx users. The polarity shows that 46.67% of the reviews are neutral, 30% of the reviews are positive and 23.33% of the reviews are negative.

# 8.0 VISUALIZATION

For visualization of the results, Python and Tableau were used.



Above the representation of subjectivity against polarity graph for tweets obtained.



Above id the rating by count graph for the reviews of the three platforms.

# 9.0 BUSINESS STORY TELLING

From the data gotten from both tweets and reviews, users are given free trial for a period of time. Before the free trial can be activated, the debit/ credit card details of the users are requested which is not to be billed if the user cancels their subscription. In addition, to subscribe, the email address of the potential users is requested. The complaints given by the users is that their emails are spammed by Shaw Academy and that their cards were billed even after cancelling subscription, some even complained of being billed more than twice a month.

From the analysed, it can be deduced Shaw Academy has the highest bad ratings compared to Edx and Udemy. Shaw Academy is the only platform of three being considered with card fraud allegations against them. We can also tell from the results that even though some users give Shaw Academy a bad rating, they were somewhat satisfied with the services rendered, at least to the barest minimum. Majority of the users labelled the platform a scam, fraud and deceptive which is not good for the reputation of the business.

The data collected on Udemy shows that there is a dispersion between the tweets and data collected. While the tweets show that the users are majorly enrolled for computing courses and enjoying their experience, the reviews have labelled a scam, cheat and also deceptive. It may seem as if the users felt they weren’t getting the value of their money in the training they enrolled for. Udemy seems to be a great platform for computing and science related courses but not for great for business and management courses.

Analysis of Udemy tweets and review show that they are have an average standing with their users unlike Shaw Academy.

Edx data depicts that among the three platforms being considered in this study, it is the best. Edx has the highest amount of positive rating. The users seem to be satisfied with their courses but have issues with communicating the with customer support team.

Overall, the data analysed shows that most users across the three platforms have a neutral opinion about the platforms.

# 10.0 CONCLUSION

## 10.1 Review Aim and Objectives

**Aim:** for Shaw Academy and EdxOnline users, there was a lot similarities in their tweets and reviews while for Udemy users, the tweets seem to be more positive than the reviews.

**Objectives:** most of users feel neutral about their experiences and are not afraid to give the platform a bad rating. For the purpose of this study only, Udemy is a better online learning platform than Shaw Academy and EdxOnline.

## 10.2 Recommendations

From this research, I recommend Udemy as the best online learning platform of the considered group. I also recommend that potential users do a thorough research on any platform before subscribing to it.

# REFERENCES

Bannister, K., 2020. *Sentiment Analysis: How Does It Work? Why Should We Use It?*. [online] Brandwatch. Available at: <https://www.brandwatch.com/blog/understanding-sentiment-analysis/> [Accessed 4 December 2020].

Bbc.co.uk. 2020. *BBC One - Watchdog - Trustpilot*. [online] Available at: <https://www.bbc.co.uk/programmes/articles/3SkLXs4RK24Lzx0sPYXVRpg/trustpilot> [Accessed 4 December 2020].

Bucher, T., 2020. *Objects Of Intense Feeling: The Case Of The Twitter API*. [online] Computationalculture.net. Available at: <http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/> [Accessed 4 December 2020].

*Cbtnuggets*, 2020. Why Data Scientists Love Python. Available at: <https://www.cbtnuggets.com/blog/technology/data/why-data-scientists-love-python> [Accessed 4 December 2020].

Earth Data Science - Earth Lab. 2020. *Automate Getting Twitter Data In Python Using Tweepy And API Access*. [online] Available at: <https://www.earthdatascience.org/courses/use-data-open-source-python/intro-to-apis/twitter-data-in-python/> [Accessed 4 December 2020].

Fontanella, C., 2020. *How To Get, Use, & Benefit From Twitter's API*. [online] Blog.hubspot.com. Available at: <https://blog.hubspot.com/website/how-to-use-twitter-api> [Accessed 4 December 2020].

Forsey, C., 2020. *What Is Twitter And How Does It Work?*. [online] Blog.hubspot.com. Available at: <https://blog.hubspot.com/marketing/what-is-twitter> [Accessed 4 December 2020].

# APPENDIXES

## Appendix A - Twitter API Test

## Appendix B – Web Scrapping Using Python Test

## Appendix C – Sentiment Analysis Test

## Appendix D – Tableau Data Visualisation Test

## Appendix E – Result CSV Files