Project 2 Report:

Selenium-Based Amazon Web Scraper

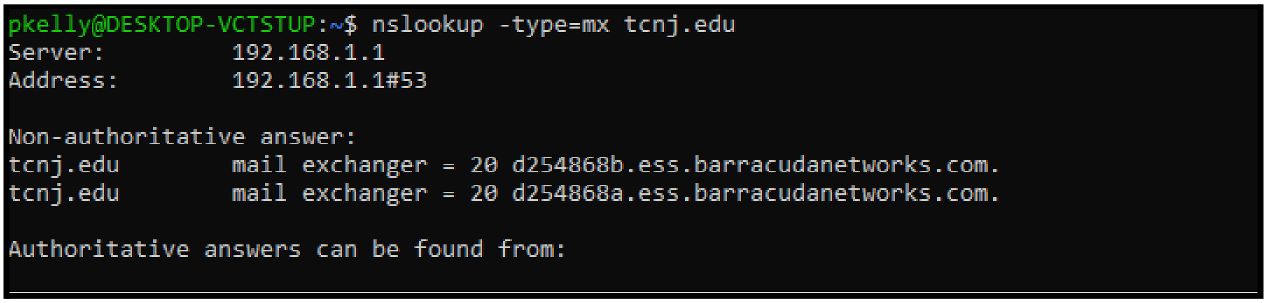
Our Process and Results

Payton Shaltis, Sterly Deracy, Peter Kelly

CSC 360-02: Computer Networking

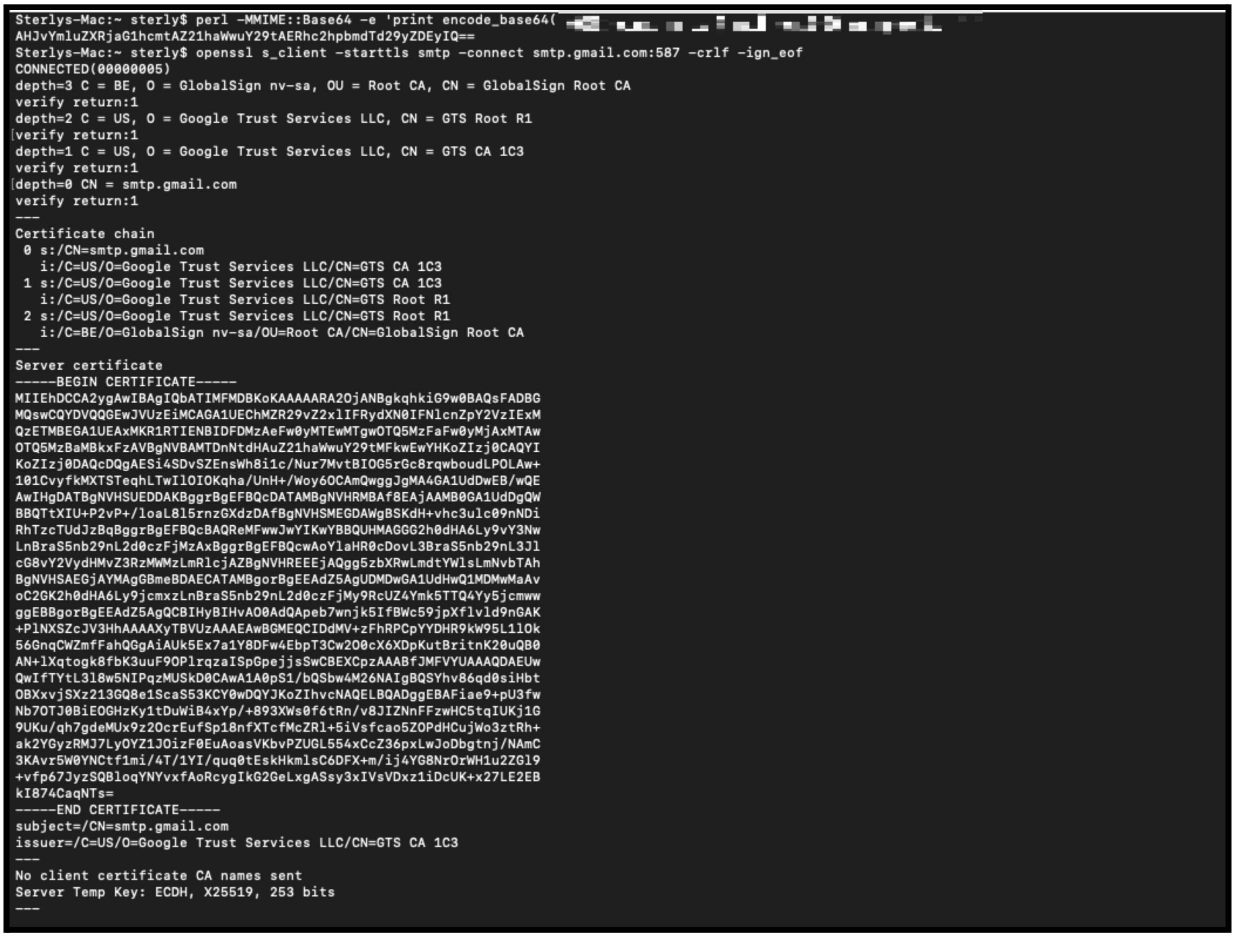
November 16, 2021

TCNJ and Gmail Mail Server Screenshots



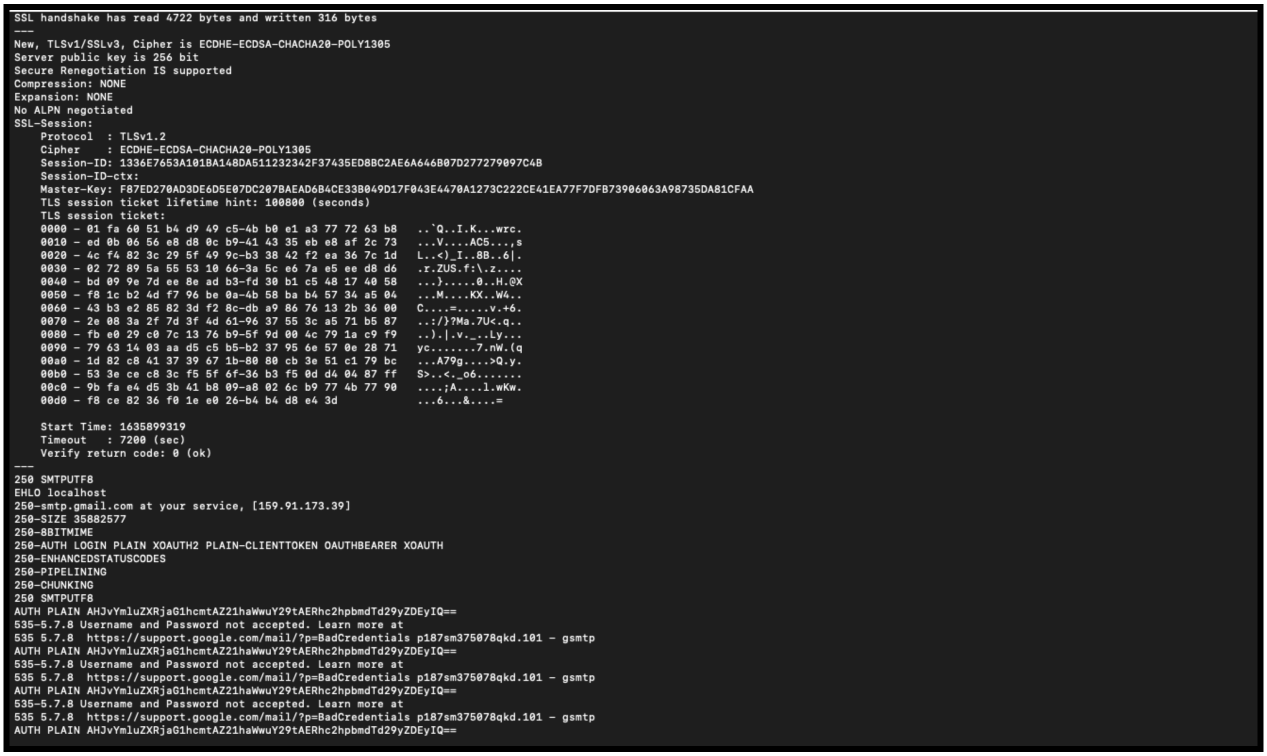
```
pkelly@DESKTOP-VCTSTUP:~$ nslookup -type=mx tcnj.edu
Server:         192.168.1.1
Address:        192.168.1.1#53

Non-authoritative answer:
tcnj.edu        mail exchanger = 20 d254868b.ess.barracudanetworks.com.
tcnj.edu        mail exchanger = 20 d254868a.ess.barracudanetworks.com.

Authoritative answers can be found from:
```

Figure 1. *Looking up the mail server hostnames using the nslookup tool from a terminal. Two results are given for the mail exchangers (which are of record type MX), having the domain of barracudanetworks.com.*



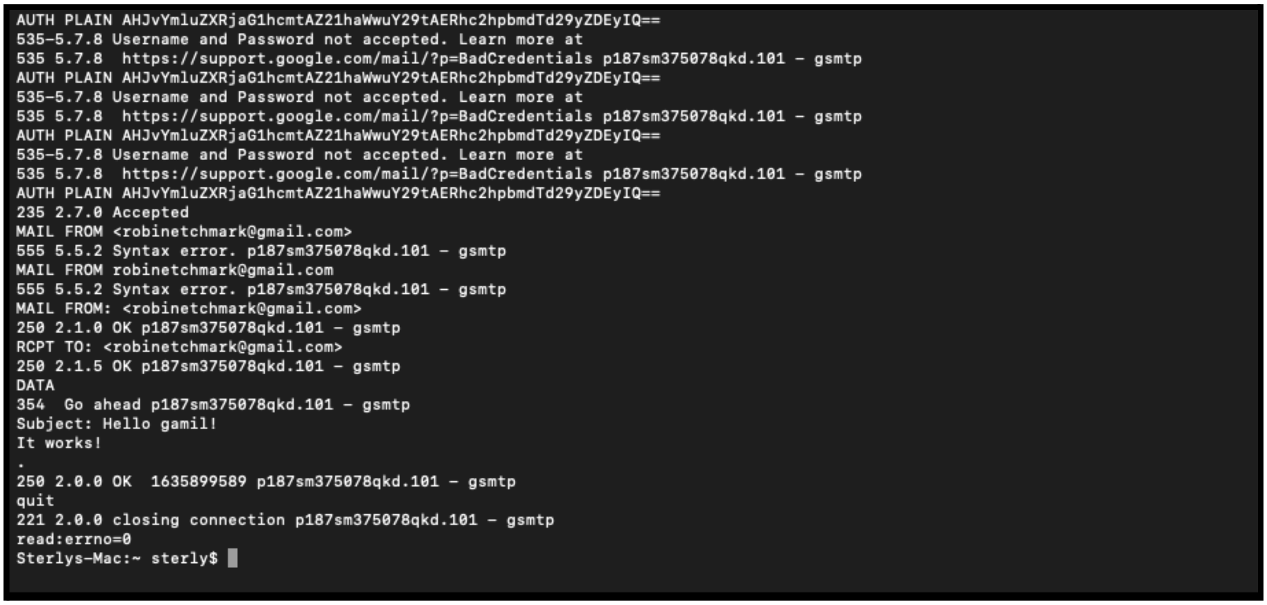
```
Sterlys-Mac:~ sterly$ perl -MMIME::Base64 -e 'print encode_base64( [redacted]
AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
Sterlys-Mac:~ sterly$ openssl s_client -starttls smtp -connect smtp.gmail.com:587 -crlf -ign_eof
CONNECTED(00000005)
depth=3 C = BE, O = GlobalSign nv-sa, OU = Root CA, CN = GlobalSign Root CA
verify return:1
depth=2 C = US, O = Google Trust Services LLC, CN = GTS Root R1
verify return:1
depth=1 C = US, O = Google Trust Services LLC, CN = GTS CA 1C3
verify return:1
depth=0 CN = smtp.gmail.com
verify return:1
---
Certificate chain
 0 s:/CN=smtp.gmail.com
   i:/C=US/O=Google Trust Services LLC/CN=GTS CA 1C3
 1 s:/C=US/O=Google Trust Services LLC/CN=GTS CA 1C3
   i:/C=US/O=Google Trust Services LLC/CN=GTS Root R1
 2 s:/C=US/O=Google Trust Services LLC/CN=GTS Root R1
   i:/C=BE/O=GlobalSign nv-sa/OU=Root CA/CN=GlobalSign Root CA
---
Server certificate
-----BEGIN CERTIFICATE-----
MIIEhDCCA2ygAwIBAgIQbATIMFMDBKoKAAAAARA2OjANBgkqhkiG9w0BAQsFADBG
MQswCQYDVQQGEwJVUzEiMCAGA1UEChMZR29vZ2xlIFRydXN0IFNlcnZpY2VzIExM
QzETMBEGA1UEAxMKR1RTIENBIDFDMzAeFw0yMTEwMTgwOTQ5MzFaFw0yMjAxMTAw
OTQ5MzBaMBkxFzAVBgNVBAMTDnNtdHAuZ21haWwuY29tMFkwEwYHKoZIzj0CAQYI
KoZIzj0DAQcDQgAESi4SDvSZEnsWh8i1c/Nur7MvtBIOG5rGc8rqwboudLPOLAw+
101CvyfkMXTSTeqhLTwIlOIOKqha/UnH+/Woy6OCAmQwggJgMA4GA1UdDwEB/wQE
AwIHgDATBgNVHSUEDDAKBggrBgEFBQcDATAMBgNVHRMBAf8EAjAAMB0GA1UdDgQW
BBQTtXIU+P2vP+/loaL8l5rnzGXdzDAfBgNVHSMEGDAWgBSKdH+vhc3ulc09nNDi
RhTzcTUdJzBqBggrBgEFBQcBAQReMFwwJwYIKwYBBQUHMAGGG2h0dHA6Ly9vY3Nw
LnBraS5nb29nL2d0czFjMzAxBggrBgEFBQcwAoYlaHR0cDovL3BraS5nb29nL3Jl
cG8vY2VydHMvZ3RzMWMzLmRlcjAZBgNVHREEEjAQgg5zbXRwLmdtYWlsLmNvbTAh
BgNVHSAEGjAYMAgGBmeBDAECATAMBgorBgEEAdZ5AgUDMDwGA1UdHwQ1MDMwMaAv
oC2GK2h0dHA6Ly9jcmxzLnBraS5nb29nL2d0czFjMy9RcUZ4Ymk5TTQ4Yy5jcmww
ggEBBgorBgEEAdZ5AgQCBIHyBIHvAO0AdQApeb7wnjk5IfBWc59jpXflvld9nGAK
+PlNXSZcJV3HhAAAAXyTBVUzAAAEAwBGMEQCIDdMV+zFhRPCpYYDHR9kW95L1lOk
56GnqCWZmfFahQGgAiAUk5Ex7a1Y8DFw4EbpT3Cw2O0cX6XDpKutBritnK20uQB0
AN+lXqtogk8fbK3uuF9OPlrqzaISpGpejjsSwCBEXCpzAAABfJMFVYUAAAQDAEUw
QwIfTYtL3l8w5NIPqzMUSkD0CAwA1A0pS1/bQSbw4M26NAIgBQSYhv86qd0siHbt
OBXxvjSXz213GQ8e1ScaS53KCY0wDQYJKoZIhvcNAQELBQADggEBAFiae9+pU3fw
Nb7OTJ0BiEOGHzKy1tDuWiB4xYp/+893XWs0f6tRn/v8JIZNnFFzwHC5tqIUKj1G
9UKu/qh7gdeMUx9z2OcrEufSp18nfXTcfMcZRl+5iVsfcao5ZOPdHCujWo3ztRh+
ak2YGyzRMJ7LyOYZ1JOizF0EuAoasVKbvPZUGL554xCcZ36pxLwJoDbgtnj/NAmC
3KAvr5W0YNCtf1mi/4T/1YI/quq0tEskHkmlsC6DFX+m/ij4YG8NrOrWH1u2ZGl9
+vfp67JyzSQBloqYNYvxfAoRcygIkG2GeLxgASsy3xIVsVDxz1iDcUK+x27LE2EB
kI874CaqNTs=
-----END CERTIFICATE-----
subject=/CN=smtp.gmail.com
issuer=/C=US/O=Google Trust Services LLC/CN=GTS CA 1C3
---
No client certificate CA names sent
Server Temp Key: ECDH, X25519, 253 bits
---
```

Figure 2. *Connecting to smtp.gmail.com manually using the ‘perl’ command. The username and password used are obviously blurred out in order to preserve privacy. The ‘openssl’ command is then used in order to connect to the mail server, and the CONNECTED message above clearly shows that it was a success. The certificate for the server was then printed to the terminal as well, and we were ready to begin sending SMTP messages.*



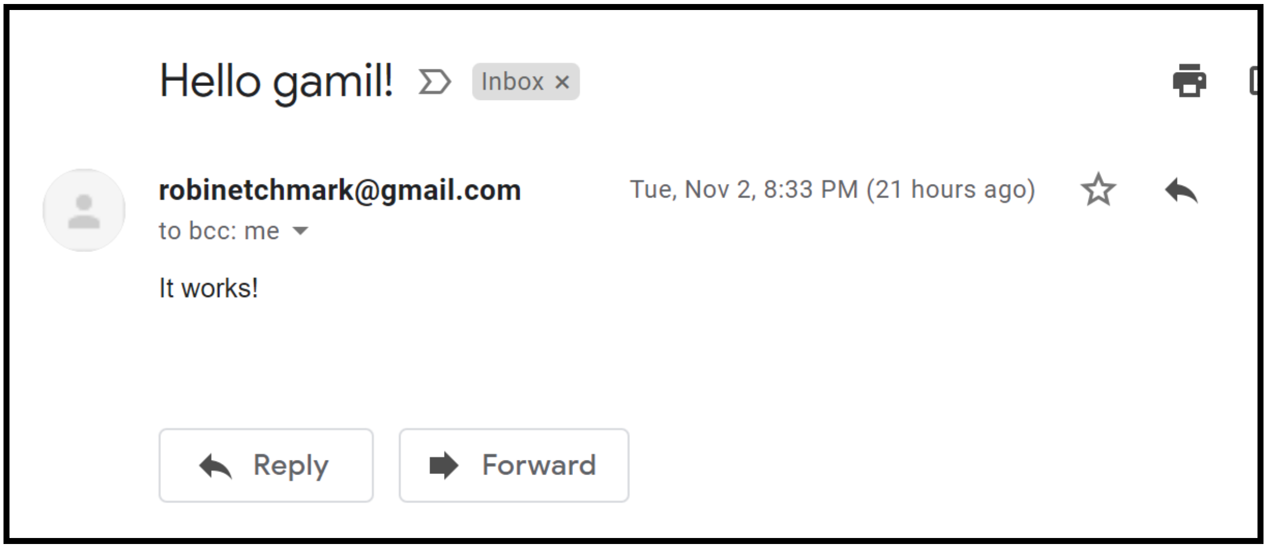
```
SSL handshake has read 4722 bytes and written 316 bytes
---
New, TLSv1/SSLv3, Cipher is ECDHE-ECDSA-CHACHA20-POLY1305
Server public key is 256 bit
Secure Renegotiation IS supported
Compression: NONE
Expansion: NONE
No ALPN negotiated
SSL-Session:
    Protocol  : TLSv1.2
    Cipher    : ECDHE-ECDSA-CHACHA20-POLY1305
    Session-ID: 1336E7653A101BA148DA511232342F37435ED8BC2AE6A646B07D277279097C4B
    Session-ID-ctx:
    Master-Key: F87ED270AD3DE6D5E07DC207BAEAD6B4CE33B049D17F043E4470A1273C222CE41EA77F7DFB73906063A98735DA81CFAA
    TLS session ticket lifetime hint: 100800 (seconds)
    TLS session ticket:
    0000 - 01 fa 60 51 b4 d9 49 c5-4b b0 e1 a3 77 72 63 b8   ..`Q..I.K...wrc.
    0010 - ed 0b 06 56 e8 d8 0c b9-41 43 35 eb e8 af 2c 73   ...V....AC5...,s
    0020 - 4c f4 82 3c 29 5f 49 9c-b3 38 42 f2 ea 36 7c 1d   L..<)_I..8B..6|.
    0030 - 02 72 89 5a 55 53 10 66-3a 5c e6 7a e5 ee d8 d6   .r.ZUS.f:\.z....
    0040 - bd 09 9e 7d ee 8e ad b3-fd 30 b1 c5 48 17 40 58   ...}.....0..H.@X
    0050 - f8 1c b2 4d f7 96 be 0a-4b 58 ba b4 57 34 a5 04   ...M....KX..W4..
    0060 - 43 b3 e2 85 82 3d f2 8c-db a9 86 76 13 2b 36 00   C....=.....v.+6.
    0070 - 2e 08 3a 2f 7d 3f 4d 61-96 37 55 3c a5 71 b5 87   ..:/}?Ma.7U<.q..
    0080 - fb e0 29 c0 7c 13 76 b9-5f 9d 00 4c 79 1a c9 f9   ..).|.v._..Ly...
    0090 - 79 63 14 03 aa d5 c5 b5-b2 37 95 6e 57 0e 28 71   yc.......7.nW.(q
    00a0 - 1d 82 c8 41 37 39 67 1b-80 80 cb 3e 51 c1 79 bc   ...A79g....>Q.y.
    00b0 - 53 3e ce c8 3c f5 5f 6f-36 b3 f5 0d d4 04 87 ff   S>..<._o6.......
    00c0 - 9b fa e4 d5 3b 41 b8 09-a8 02 6c b9 77 4b 77 90   ....;A....l.wKw.
    00d0 - f8 ce 82 36 f0 1e e0 26-b4 b4 d8 e4 3d            ...6...&....=

    Start Time: 1635899319
    Timeout   : 7200 (sec)
    Verify return code: 0 (ok)
---
250 SMTPUTF8
EHLO localhost
250-smtp.gmail.com at your service, [159.91.173.39]
250-SIZE 35882577
250-8BITMIME
250-AUTH LOGIN PLAIN XOAUTH2 PLAIN-CLIENTTOKEN OAUTHBEARER XOAUTH
250-ENHANCEDSTATUSCODES
250-PIPELINING
250-CHUNKING
250 SMTPUTF8
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
```

Figure 3. *The first message sent to the server was the 'EHLO localhost' message. The response from the server is shown below this command above.*



```
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
535-5.7.8 Username and Password not accepted. Learn more at
535 5.7.8  https://support.google.com/mail/?p=BadCredentials p187sm375078qkd.101 - gsmtp
AUTH PLAIN AHJvYmluZXRjaG1hcmtAZ21haWwuY29tAERhc2hpbmdTd29yZDEyIQ==
235 2.7.0 Accepted
MAIL FROM <robinetchmark@gmail.com>
555 5.5.2 Syntax error. p187sm375078qkd.101 - gsmtp
MAIL FROM robinetchmark@gmail.com
555 5.5.2 Syntax error. p187sm375078qkd.101 - gsmtp
MAIL FROM: <robinetchmark@gmail.com>
250 2.1.0 OK p187sm375078qkd.101 - gsmtp
RCPT TO: <robinetchmark@gmail.com>
250 2.1.5 OK p187sm375078qkd.101 - gsmtp
DATA
354  Go ahead p187sm375078qkd.101 - gsmtp
Subject: Hello gamil!
It works!
.
250 2.0.0 OK  1635899589 p187sm375078qkd.101 - gsmtp
quit
221 2.0.0 closing connection p187sm375078qkd.101 - gsmtp
read:errno=0
Sterlys-Mac:~ sterly$ █
```

Figure 4. *This is a screenshot of the complete transaction of messages between the client and the server. After the initial 'EHLO' message from Figure 3, we were able to send and receive messages to and from the mail server, and concluded by closing the connection. The final step was to check our email and determine if the message was actually sent.*



Hello gamil!

robinetchmark@gmail.com

Tue, Nov 2, 8:33 PM (21 hours ago)

to bcc: me

It works!

Figure 5. *As you can see from this figure as well as Figure 4, the communication with the Gmail SMTP mail server worked! The mail subject (with typo), sender name, and body all match what was sent in the transactions shown above.*

Selenium-Based Amazon Web Scraper Report

The first thing the program does is utilize the ChromeDriver to open up Google Chrome and navigate to the product link. From there, it navigates to the "See All Reviews" section. This allows the program to see all the reviews that would have otherwise been hidden by Amazon. The program has a special parameter to limit the number of pages to scrape if the user does not want to process all review pages. Regardless, it will scrape pages, analyzing each user's profile and storing their profile link, their product rating, and their product review into corresponding arrays "profile_links", "reviews", and "ratings". In order to avoid running into any issues with scraping too quickly, this is where we use a randomly generated sleep timer. Another benefit of this randomly generated timer is that it ensures a much greater chance of keeping the user's ip in good graces with Amazon's server, and not receiving a ban. Along with storing data into the three arrays, we also gather data using variables "vp_badges" which is an array and "num_reviews". The purpose of these will be discussed later. Everything up to this point deals with scraping the pages for user reviews. Figure 6 shows the scraping process for hundreds of reviews of a Bose Wireless Earphone product.

Once this is done, we move to analyzing user profiles. We use a custom Bias Determination algorithm, which looks at a key number of factors. We look at the size of a user's review, with any review over 100 characters considered "long" and any review under that character limit "short". We also consider whether the user has a "verified purchase" badge, which indicates that Amazon recognizes them as a legitimate buyer of the product under consideration. If a user has a long review and they have been verified as a purchaser of this item by Amazon, it is more likely that they are unbiased; it demonstrates the fact that the user put some amount of effort into submitting their review. We also consider whether a user is reviewing from a country outside of the United States, since Amazon's display of those pages differ from those of the United States. If a user has a long review and they're verified, we consider this review in our final evaluation regardless of what country they're from. If the user is from the United States and they have either an unverified review or a short review, we flag the review as one that "Needs Further Investigation." If a user from outside the United States has similar parameters, we drop their score completely. Because Amazon does not allow users to look at the profiles of foreign users, these reviews cannot be marked as "Needing Further Investigation," so we play it safe and drop these reviews. Now we're left with two groups of reviews: one group whose ratings *will* contribute to the final score, and another group whose profiles will need to be investigated, and whose ratings *may* contribute to the final score.

For the group that needs more consideration, we analyze the user's last ten reviews and determine whether they have given either a majority of very positive reviews, a majority of negative reviews, or a good balance of the two. We do this by comparing the percentage of 5-star reviews and the percentage of 1-star reviews with a constant percentage that is currently set to 85%. We chose 85% in order to be more fair towards users that don't have 10 reviews; for example, a user with only 5 reviews in which 4 of them are 1-star won't get flagged, since this user only has 80% completely negative reviews. Any user profiles who seem to overwhelmingly give 5-star or 1-star ratings are automatically dropped from consideration. The idea here is that they are more biased and more likely to leave an unfair rating over a fair and balanced one. At this point, we have carefully constructed a list of reviews that are unbiased. Figures 7 and 8 give good samples of the program scraping the profiles of users that we deemed necessary to investigate further, using the Bias Determination algorithm described above.

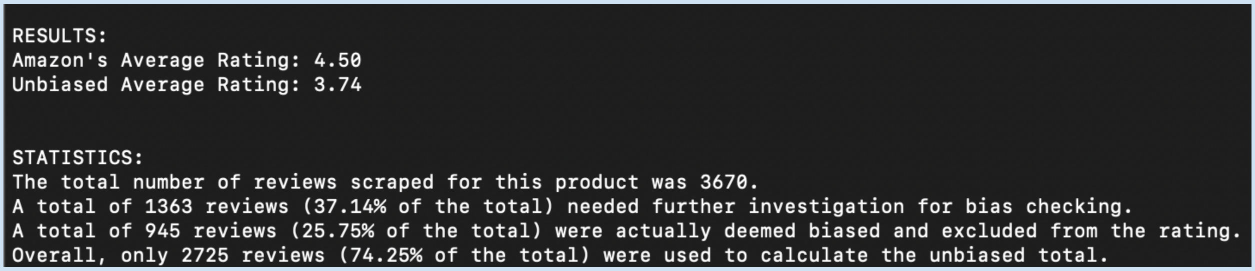
The final part of our program uses these unbiased ratings in order to calculate a new score for the product, as well as displaying other helpful statistics to the user. Before we display the statistics, we close the Chrome browser since the program has already gathered all the data it needs from the product. As you can see in our snippets, the program displays Amazon's average rating and the program's calculated unbiased average rating. We also discuss how this rating came to be by telling the user how many reviews were scraped, how many needed further investigation, how many were dropped from consideration, and how many were actually used to calculate the unbiased average. It should be noted that the program has some additional helper methods, such as get_int_rating and investigate_profile. The method get_int_rating returns the actual number from a rating, since the value of the element's attribute is actually a string formatted like '3-star-review.' investigate_profile parses through an individual user profile.

Because some of the functions in Selenium have the possibility of throwing exceptions, any time we call a function such as 'find_elements()' or 'click()', it is done in a 'try/except' block. NoSuchElementException, StaleElementReferenceException, and ElementClickInterceptedException were thrown when an element was not found on the page, the page refreshed after getting a reference to the element, or when the element was not clickable, respectively. Most of the time, this was due to our program using Selenium functions before a page was completely loaded. The workaround was trivial: anytime an exception was caught, our program would wait some amount of time (usually only a second) and retry the ‘find_elements()’ or ‘click()’ method call.

The screenshots placed above this explanation give all the outputs that the user should see while the program is scraping, verifying, etc., and the code has further comments contained within to further clarify exactly what is being added to or changed from each array. Once the user receives both our unbiased average rating and the Amazon rating, the choice to buy the product is up to the user, who would most certainly be appreciative of the carefully curated and more accurate user score. Below are some links to Amazon product links, as well as the output that was generated by our program for each. Note that the total number of reviews may not align with the current number of reviews; each day, new reviews are added, so there will likely be more when the links are accessed:

Bose Wireless Earbuds

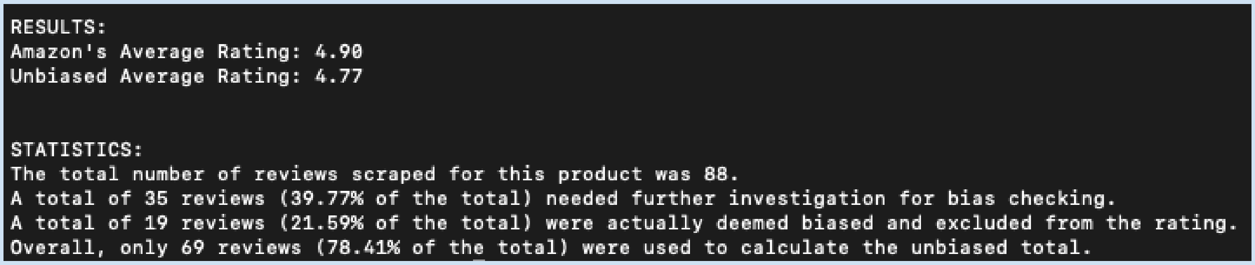
https://www.amazon.com/Bose-QuietComfort-Noise-Cancelling-Earbuds/dp/B08C4KWM9T/ref=sr_1_2?dchild=1&qid=1635902902&refinements=p_89%3ABose%2Cp_n_feature_four_browse-bin%3A12097501011&s=aht&sr=1-2



```
RESULTS:
Amazon's Average Rating: 4.50
Unbiased Average Rating: 3.74


STATISTICS:
The total number of reviews scraped for this product was 3670.
A total of 1363 reviews (37.14% of the total) needed further investigation for bias checking.
A total of 945 reviews (25.75% of the total) were actually deemed biased and excluded from the rating.
Overall, only 2725 reviews (74.25% of the total) were used to calculate the unbiased total.
```

Women’s Makeup Bag

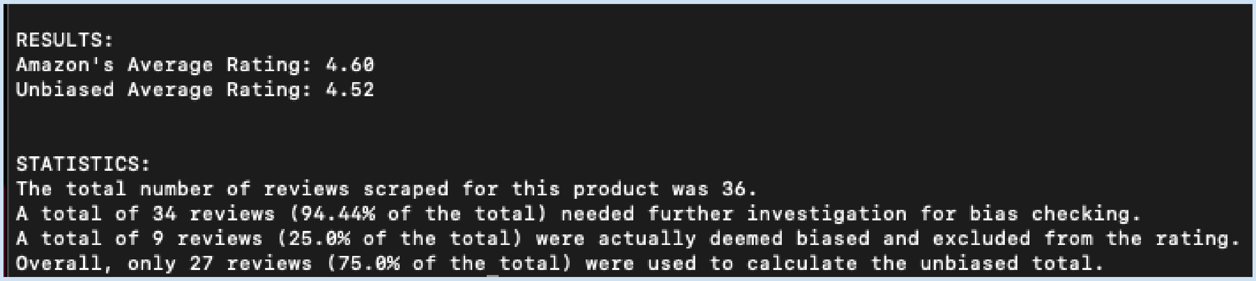
https://www.amazon.com/Lug-Womens-Trolley-Cosmetic-Bloom/dp/B087NKFQB7/?_encoding=UTF8&pd_rd_w=EorSw&pf_rd_p=8b894231-4b84-44da-9446-c27cf0e8abc2&pf_rd_r=WZNKXNMVTVQZZPV6K79G&pd_rd_r=afe07fde-25b8-44dd-97b5-bb7911ab9913&pd_rd_wg=rU5X4&ref_=pd_gw_ci_mcx_mr_hp_d



```
RESULTS:
Amazon's Average Rating: 4.90
Unbiased Average Rating: 4.77


STATISTICS:
The total number of reviews scraped for this product was 88.
A total of 35 reviews (39.77% of the total) needed further investigation for bias checking.
A total of 19 reviews (21.59% of the total) were actually deemed biased and excluded from the rating.
Overall, only 69 reviews (78.41% of the total) were used to calculate the unbiased total.
```

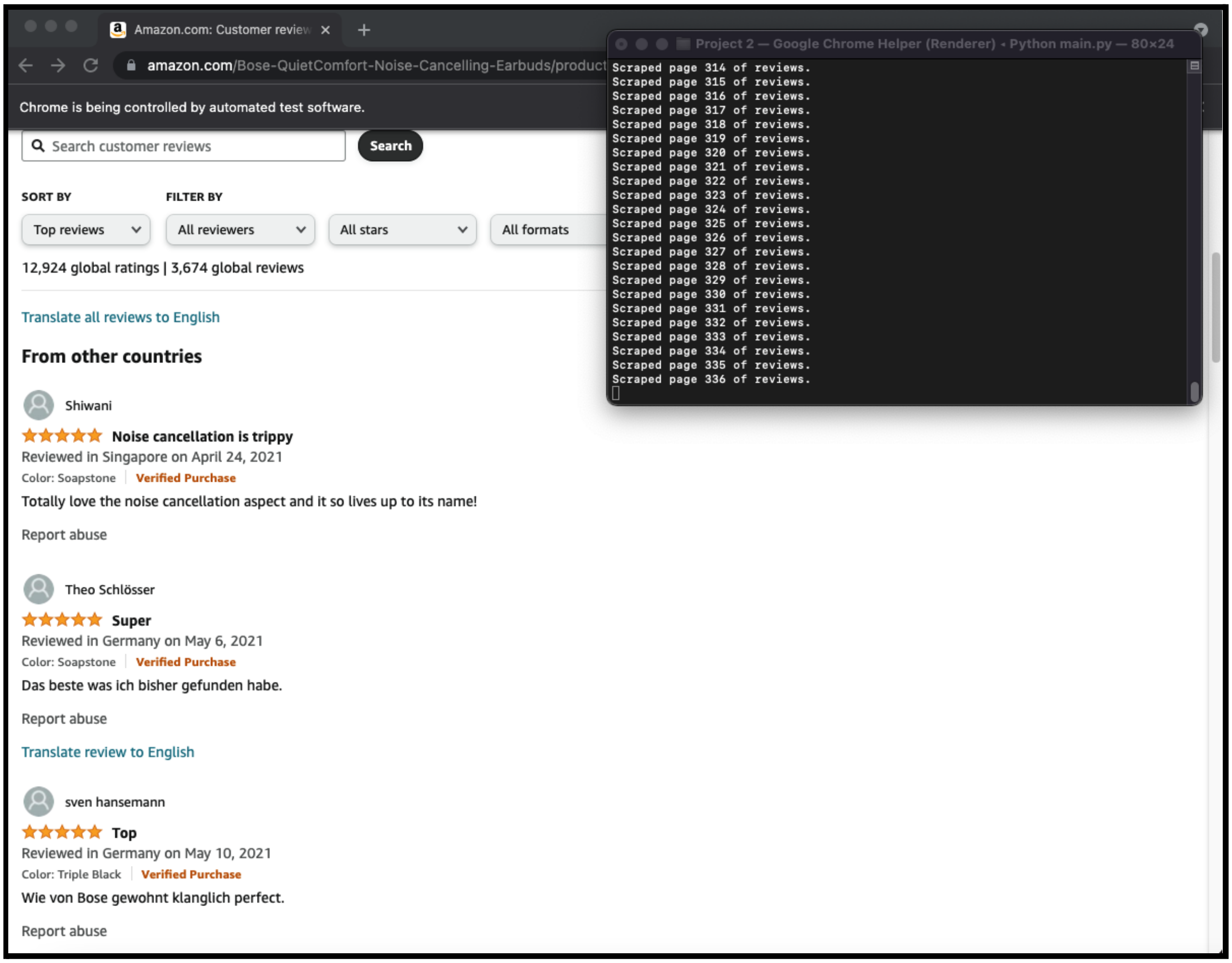
Windswept Novel

https://www.amazon.com/Windswept-Novel-WWI-Saga/dp/1736809504/ref=sr_1_4?keywords=windswept&qid=1636741044&qsid=138-8482874-4338910&s=books&sr=1-4&sres=1951142705%2C1736809504%2CB08XJRL2S4%2C0385492316%2CB07GJ6NYD2%2C1510742824%2CB07HXNW8X8%2C1529324718%2CB07JYQR7T4%2CB01LE03TDO%2CB01LM3RJ3O%2CB07K3ZZG78%2CB071V9BCD7%2CB06XJM3B6F%2C1451665555%2CB01JH39IXK



```
RESULTS:
Amazon's Average Rating: 4.60
Unbiased Average Rating: 4.52


STATISTICS:
The total number of reviews scraped for this product was 36.
A total of 34 reviews (94.44% of the total) needed further investigation for bias checking.
A total of 9 reviews (25.0% of the total) were actually deemed biased and excluded from the rating.
Overall, only 27 reviews (75.0% of the total) were used to calculate the unbiased total.
```

Figures Referenced in Web Scraper Report



Amazon.com: Customer review
amazon.com/Bose-QuietComfort-Noise-Cancelling-Earbuds/product
Project 2 — Google Chrome Helper (Renderer) ◂ Python main.py — 80×24
Search customer reviews
Search
SORT BY
FILTER BY
Top reviews
All reviewers
All stars
All formats

Chrome is being controlled by automated test software.

```
Scraped page 314 of reviews.
Scraped page 315 of reviews.
Scraped page 316 of reviews.
Scraped page 317 of reviews.
Scraped page 318 of reviews.
Scraped page 319 of reviews.
Scraped page 320 of reviews.
Scraped page 321 of reviews.
Scraped page 322 of reviews.
Scraped page 323 of reviews.
Scraped page 324 of reviews.
Scraped page 325 of reviews.
Scraped page 326 of reviews.
Scraped page 327 of reviews.
Scraped page 328 of reviews.
Scraped page 329 of reviews.
Scraped page 330 of reviews.
Scraped page 331 of reviews.
Scraped page 332 of reviews.
Scraped page 333 of reviews.
Scraped page 334 of reviews.
Scraped page 335 of reviews.
Scraped page 336 of reviews.
```

12,924 global ratings | 3,674 global reviews

Translate all reviews to English

From other countries

Shiwani

★★★★★ **Noise cancellation is trippy**

Reviewed in Singapore on April 24, 2021

Color: Soapstone | Verified Purchase

Totally love the noise cancellation aspect and it so lives up to its name!

Report abuse

Theo Schlösser

★★★★★ **Super**

Reviewed in Germany on May 6, 2021

Color: Soapstone | Verified Purchase

Das beste was ich bisher gefunden habe.

Report abuse

Translate review to English

sven hansemann

★★★★★ **Top**

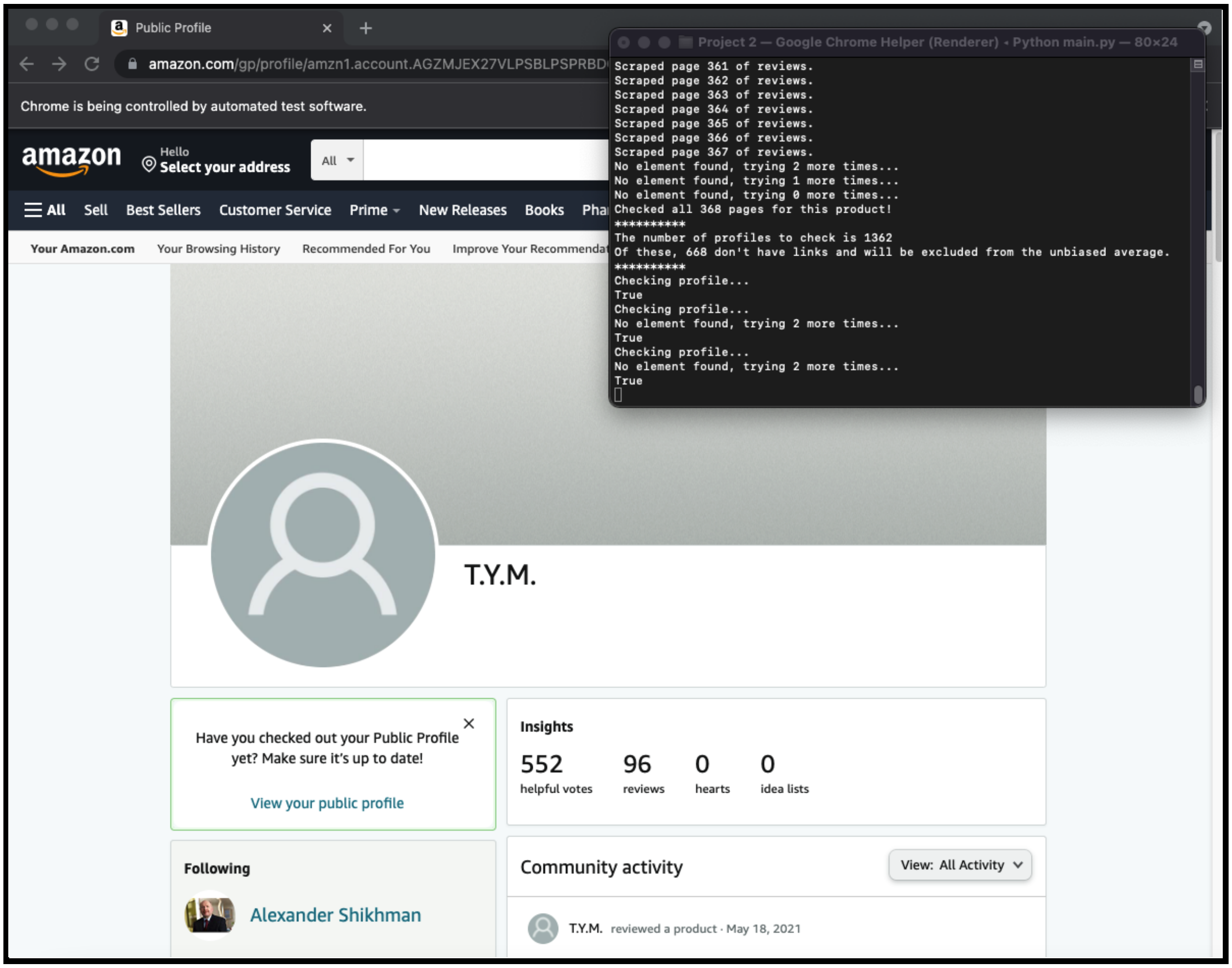
Reviewed in Germany on May 10, 2021

Color: Triple Black | Verified Purchase

Wie von Bose gewohnt klanglich perfect.

Report abuse

Figure 6. *The scraper in progress. Selenium is using the Chrome webdriver in order to display what is going on at all times in the Chrome window. Over this window is the terminal that shows output from our scraper, informing the user each time that a page is successfully scraped, keeping track of the cumulative number of pages scraped.*

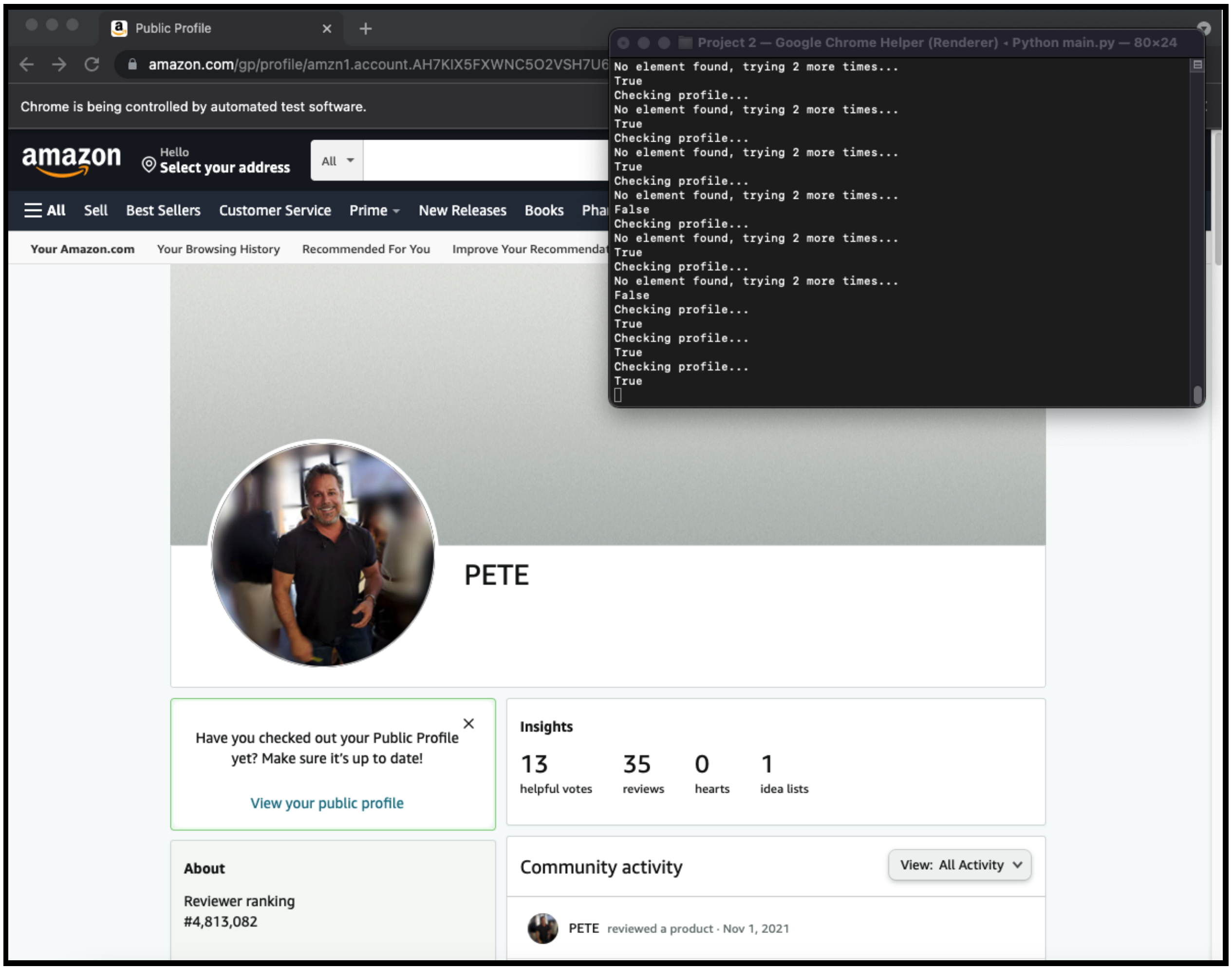


Chrome is being controlled by automated test software.

Have you checked out your Public Profile yet? Make sure it's up to date!

```
Scraped page 361 of reviews.
Scraped page 362 of reviews.
Scraped page 363 of reviews.
Scraped page 364 of reviews.
Scraped page 365 of reviews.
Scraped page 366 of reviews.
Scraped page 367 of reviews.
No element found, trying 2 more times...
No element found, trying 1 more times...
No element found, trying 0 more times...
Checked all 368 pages for this product!
**********
The number of profiles to check is 1362
Of these, 668 don't have links and will be excluded from the unbiased average.
**********
Checking profile...
True
Checking profile...
No element found, trying 2 more times...
True
Checking profile...
No element found, trying 2 more times...
True
```

*__Figure 7.__ The Bias Detection algorithm at work. The Chrome window shows the current profile that is being scraped by Selenium in our program. The terminal window displays information about the scrape. As you can see, the reviews page of the product has just concluded, indicating the number of links that should be investigated further as well as those that need to be thrown away (foreign accounts missing either a verified purchase badge or foreign accounts with a short review). The program then goes on to scrape 1362 - 668, or **694** profiles.*



Chrome is being controlled by automated test software.

```
No element found, trying 2 more times...
True
Checking profile...
No element found, trying 2 more times...
True
Checking profile...
No element found, trying 2 more times...
True
Checking profile...
No element found, trying 2 more times...
False
Checking profile...
No element found, trying 2 more times...
True
Checking profile...
No element found, trying 2 more times...
False
Checking profile...
True
Checking profile...
True
Checking profile...
True
```

PETE

Have you checked out your Public Profile yet? Make sure it's up to date!

Insights

13 helpful votes · 35 reviews · 0 hearts · 1 idea lists

About

Reviewer ranking
#4,813,082

Community activity

PETE reviewed a product · Nov 1, 2021

__Figure 8.__ Another sample of the Bias Detection algorithm at work. You can see what our program does to prevent crashing when an exception is thrown by Selenium - in the example above, the 'review' elements are not found on the page right away. We give the program 2 more tries to find the element before determining that it isn't on the page, with each try separated by a few seconds in order to allow the page enough time to load.

Notice the 'True' and 'False' that follow each 'Checking profile...' output. These indicate whether or not the profile would be considered in the final batch of unbiased ratings. False means that the user's last 10 reviews contained either at least 85% 1-star reviews or at least 85% 5-star reviews.