The Complete Technical Guide

On Analyzing Google Reviews

By Eldon Lim Kai Jie

Table of Contents

Section 1: How to Use Python to Scrape Google Reviews and Convert them to Excel

1.0 Prerequisites

Before you start, make sure you have the following installed on your machine:

1. Python 3.x
2. Visual Studio Code
3. Google Chrome browser
4. ChromeDriver (compatible with your Chrome version)

You will also need to install some Python packages by running this line of code in your terminal.

```
pip install selenium numpy pandas openpyxl
```

1.1 Step-by-Step Guide

1. Set up Chrome Driver

Download the ChromeDriver from the official site and place it in a directory of your choice. Note the path to the ‘chromedriver.exe’ file.

2. Import Necessary Libraries

Start by importing the required libraries for web scraping and data manipulation in your Visual Studio Code:

```
from selenium import webdriver
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.common.by import By
import time
import numpy as np
from selenium.webdriver.common.action_chains import ActionChains
from selenium.webdriver.common.actions.wheel_input import ScrollOrigin
import pandas as pd
from openpyxl import Workbook
```

3. Configure WebDriver

Specify the path to your ChromeDriver and configure the WebDriver options:

```
PATH = "C:\\Path\\To\\Your\\chromedriver.exe"

cService = webdriver.chrome.service.Service(executable_path=PATH)  # instantiate
options = webdriver.ChromeOptions()  # instantiate
options.add_experimental_option("detach", True)
driver = webdriver.Chrome(service=cService, options=options)
```

4. Navigate to the Target Page

Use the WebDriver to open the Google Maps page with the reviews you want to scrape:

```
driver.get("https://www.google.com/maps/place/Adidas+Outlet+-+DVOM/@5.244397,100.29339,12z/data=!4m12!1m2!2m1!1sadidas!3m8!1s0x304ab9dfc5f8f8c9:0xbe30133fcec8059c!8m2!3d5.244397!4d100.4375856!9m1!1b1!15sCgZhZGlkYXMiA4gBAZIBEHNwb3J0c3dlYXJfc3RvcmXgAQA!16s%2Fg%2F11c2kzbw8q?entry=ttu")
```

5. Define the Review Mining Function

Create a function to scrape reviews from the page:

```
def MineReviews():
    action = ActionChains(driver)
    allReviews =
wait.until(EC.element_to_be_clickable((By.CSS_SELECTOR,
'[class="hh2c6 G7m0Af"]')))
    allReviews = driver.find_element(By.CSS_SELECTOR, '[class="hh2c6
G7m0Af"]')
    allReviews.click()

    comments =
wait.until(EC.element_to_be_clickable((By.CSS_SELECTOR, '[class =
"wiI7pd"]')))
    comments = driver.find_elements(By.CSS_SELECTOR, '[class =
"wiI7pd"]')

    while len(comments) < 1000:
        var = len(comments)
        scroll_origin =
ScrollOrigin.from_element(comments[len(comments)-1])
        action.scroll_from_origin(scroll_origin, 0, 1000).perform()
        time.sleep(2)
        comments = driver.find_elements(By.CSS_SELECTOR, '[class =
"wiI7pd"]')

        if len(comments) == var:
            le += 1
            if le > 20:
                break
        else:
            le = 0

    for line in comments:
        print(line.text)
        comment.append(line.text)

    np.savetxt('review.csv', comment, delimiter=';' ,fmt='%s',
encoding="utf-8")
    df = pd.read_csv("review.csv", on_bad_lines='skip')
    df.to_excel('review.xlsx')
```

6. Main Function to Execute the Script

Create the main function to initialize the WebDriverWait and call the MineReviews function:

```
def main():
    global wait
    global comment
    comment = []
    wait = WebDriverWait(driver, 10)

    MineReviews()

main()
```

7. Running the Script

Save the script as ‘scrape_reviews.py’ and run it:

This script will scrape the reviews from the specified Google Maps page, save them to a CSV file, and then convert the CSV file to an Excel file.

1.2 Troubleshooting

- Ensure that your ChromeDriver version matches your installed Chrome browser version.
- If the script fails to find elements, the HTML structure of the page might have changed. Inspect the page to get the updated selectors.

With this guide, you should be able to scrape Google reviews and convert them to an Excel file using Python.

Section 2: How to Perform Sentiment Analysis in Excel using Azure Machine Learning Add-On

2.0 Prerequisites

Ensure you have the Excel file ‘review.xlsx’ created from Section 1.

2.1 Step-by-Step Guide

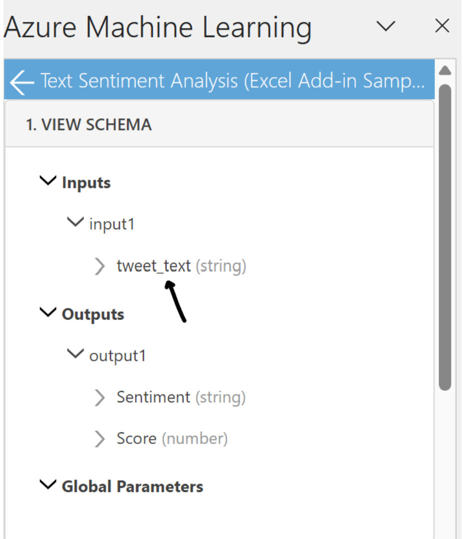
1. Open the Excel File you created in Section 1

2. Install the Azure Machine Learning Add-On

- Go to the ‘File’ tab in Excel.
- Click on ‘Get Add-ins’.
- In the Office Add-ins store, search for “Azure Machine Learning”.
- Click “Add” to install the Azure Machine Learning add-on.

3. Prepare the Data for Sentiment Analysis

Insert a new row at the top of your data to serve as header. Note that the title of the header must be the same as the one inside Azure Machine Learning for it to detect.



In this case, the title should be ‘tweet_text’

4. Perform Sentiment Analysis

- Select the Azure Machine Learning add-on from the Add-ins section.
- Choose the 'Text Sentiment Analysis' option.
- Follow the prompts to select the range of cells that contain the review texts.

5. View the Results

After the sentiment analysis is complete, two new columns will be added to your Excel file.

Sentiment: This column will indicate the sentiment of each review (Positive, Neutral, or Negative).
Score: This column will provide a numerical score for the sentiment, ranging from 0 to 1.

- Scores below 0.5 indicate negative sentiment.
- Scores of exactly 0.5 indicate neutral sentiment.
- Scores above 0.5 indicate positive sentiment.

2.2 Troubleshooting: Handling Comma-Separated Text

You may notice that reviews with commas are split across multiple columns. To concatenate these back into a single sentence, follow these steps:

1. Insert a new column next to your existing review columns.
2. Use the ‘CONCATENATE’ function to combine the split text. For example, if your text is split across columns A and B, use the following formula in a new cell:
   =CONCATENATE(A1,B1)
3. Drag the formula down to apply it to all rows.

2.3 Notes

- Sentiment analysis may not be 100% accurate and should be used as a general indicator of sentiment.
- Always review the concatenated text to ensure it is correct.

With this guide, you should be able to perform sentiment analysis on your scraped Google reviews using the Azure Machine Learning add-on in Excel.

Section 3: How to Gain Further Insights on Customer Comments Using Excel Tools

3.0 Step-by-Step Guide

1. Use the Text to Columns Tool

- Open your review.xlsx file in Excel.
- Select the column that contains the reviews.
- Go to the ‘Data’ tab and click on ‘Text to Columns’.
- Choose Delimited.
- Select space as the delimiter and click finish.

2. Construct Keyword and Count Columns

Insert two new columns next to your separated words columns. Label them as "Keyword" and "Count".

3. Identify Potential Keywords

In the "Keyword" column, list the keywords you want to track. For example:

- A2: good
- A3: bad
- A4: excellent
- A5: poor

4. Use the COUNTIF Function

In the "Count" column next to each keyword, use the COUNTIF function to count the occurrences of each keyword in the review text columns. For example, if your review words are in columns B to Z, use:

```
=COUNTIF(B:Z, A2)
```

Drag this formula down to apply it to all your keywords.

5. Create PivotCharts

- Select the range of your Keyword and Count columns.
- Go to the ‘Insert’ tab.
- Click on ‘PivotChart’ and select the type of chart you want to create (e.g. Bar Chart, Pie Chart)
- In the PivotChart Field list, drag ‘Keyword’ to the Axis (Categories) area and ‘Count’ to the Values area.
- Customize the chart as needed to better visualize the frequency of keywords in customer reviews.

6. Customize Your PivotChart

- Click on the chart to bring up the Chart Tools.
- Use the ‘Design’ and ‘Format’ tabs to customize the appearance of you chart.
- Add titles, labels, and adjust the colors to make your chart more informative and visually appealing.

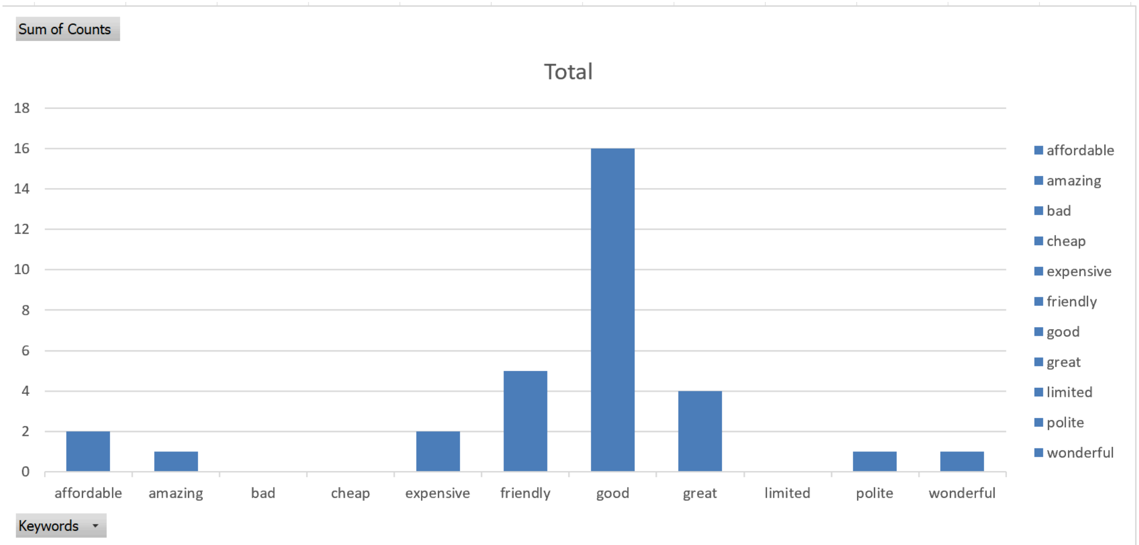
3.1 Example

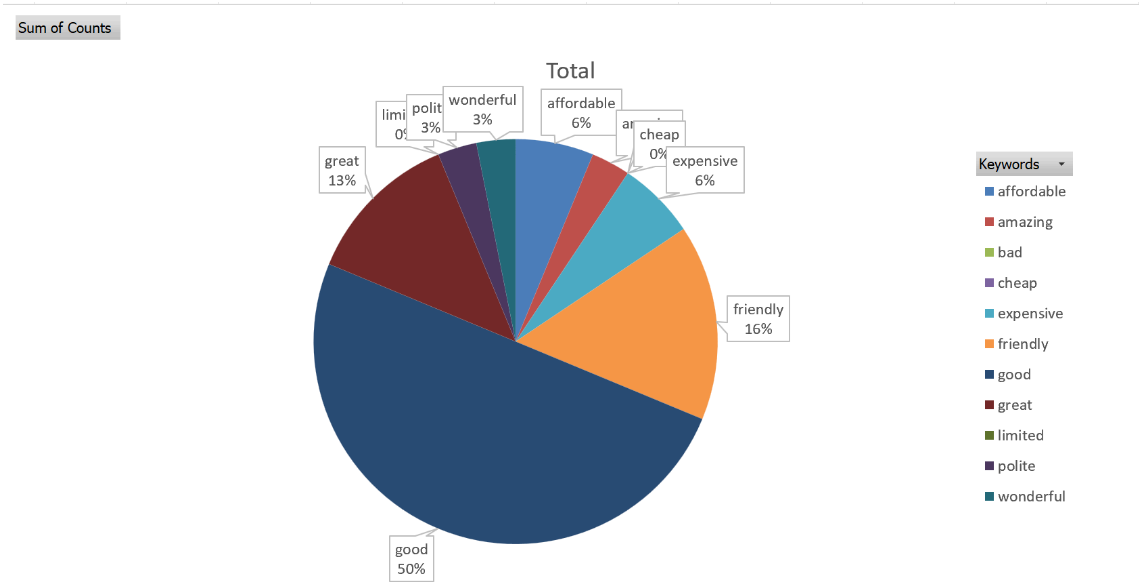
Suppose your keyword list and count columns look like this:

Keyword	Count
good	15
bad	6
excellent	10
poor	4

Suppose your keyword list and count columns look like this:

- A bar chart with keywords on the x-axis and counts on the y-axis, showing the frequency of each keyword in customer reviews.





3.2 Notes

- Choose keywords that are relevant to your analysis and customer reviews.
- The PivotChart can help you quickly identify trends and insights from customer comments.

With this guide, you should be able to gain further insights from customer reviews by identifying keyword frequencies and visualizing the data using PivotCharts in Excel.