**Fairoz Ahmed**

**22138042**

**CSE AI-DS**

**ONLINE REVIEW ANALYTICS**

**PROGRAM :**

1. import requests

2. from bs4 import BeautifulSoup

3. from textblob import TextBlob

4. from wordcloud import WordCloud

5. import matplotlib.pyplot as plt

6. from pprint import pprint

7. import re

1. def clean\_text(text):

2.     # Remove HTML tags

3.     text = re.sub(r'<[^>]\*>', '', text)

4.     # Remove special characters and numbers

5.     text = re.sub(r'[^A-Za-z\s]', '', text)

6.     # Convert to lowercase

7.     text = text.lower()

8.     # Remove extra whitespace

9.     text = ' '.join(text.split())

10.     return text

1. def get\_reviews\_with\_oxylabs(username, password, url):

2.     api\_url = 'https://realtime.oxylabs.io/v1/queries'

3.     payload = {

4.         'source': 'universal',

5.         'url': url

6.     }

7.     response = requests.post(api\_url, auth=(username, password), json=payload)

8.     if response.status\_code == 200:

9.         data = response.json()

10.         html\_content = data.get('results', [])[0].get('content', '')

11.         soup = BeautifulSoup(html\_content, 'html.parser')

12.         review\_elements = soup.find\_all('div', itemprop='reviewBody')

13.         reviews = []

14.         for review\_element in review\_elements:

15.             likes\_section = review\_element.find('h3', class\_='input-fields sub-heading', string='Likes')

16.             dislikes\_section = review\_element.find('h3', class\_='input-fields sub-heading', string='Dislikes')

17.             likes = likes\_section.find\_next\_sibling('p').get\_text(strip=True) if likes\_section else ""

18.             dislikes = dislikes\_section.find\_next\_sibling('p').get\_text(strip=True) if dislikes\_section else ""

19.             likes = clean\_text(likes)

20.             dislikes = clean\_text(dislikes)

21.             reviews.append((likes, dislikes))

22.         return reviews

23.     else:

24.         print("Failed to retrieve reviews from Oxylabs API.")

25.         return []

26.

1. def perform\_sentiment\_analysis(reviews):

2.     positive\_reviews = []

3.     negative\_reviews = []

4.     for likes, dislikes in reviews:

5.         likes\_analysis = TextBlob(likes)

6.         dislikes\_analysis = TextBlob(dislikes)

7.         if likes\_analysis.sentiment.polarity >= 0:

8.             positive\_reviews.append(likes)

9.         if dislikes\_analysis.sentiment.polarity < 0:

10.             negative\_reviews.append(dislikes

11.     return positive\_reviews, negative\_reviews

1. def create\_word\_cloud(reviews, title):

2.     text = ' '.join(reviews)

3.     wordcloud = WordCloud(width=800, height=400, background\_color='white').generate(text)

4.     plt.figure(figsize=(10, 5))

5.     plt.imshow(wordcloud, interpolation='bilinear')

6.     plt.title(title)

7.     plt.axis('off')

8.     plt.show()

9.

1. def main()

2.     url = input("Enter the URL to scrape reviews: ")

3.     username = 'XXXXXXYYYYYYYZZZZZ'

4.     password = 'XXXXXXYYYYYYYZZZZZ'

5.     reviews = get\_reviews\_with\_oxylabs(username, password, url)

6.     if not reviews:

7.         return

8.     positive\_reviews, negative\_reviews = perform\_sentiment\_analysis(reviews

9.     create\_word\_cloud(positive\_reviews, 'Positive Reviews Word Cloud')

10.     create\_word\_cloud(negative\_reviews, 'Negative Reviews Word Cloud'

11. if \_\_name\_\_ == "\_\_main\_\_":

12.     main()

13.

**# OUTPUT**

**C:\Users\91904\OneDrive\Documents\Online\_Review\_Scrapping>** python "Digital Marketing Assignment-3.py"

**Enter the URL to scrape reviews:** <https://www.ambitionbox.com/reviews/coca-cola-company-reviews>

****

**Positive Review Negative Review**

**REQUISITES :**

**Product Chosen :** Coco-Cola

**Web Scrapping API Tool Used :** Oxylabs API

**Extracted Website :** <https://www.ambitionbox.com/reviews/coca-cola-company-reviews>

**INFERENCE :**

* **Positive Inference :**

**Pay Scale :   
 🡪** Most of the Employees stated that they have a good pay scale in the beginning.

**Learning and Benefits :**

**🡪** Several reviews from new trainee stated that the training in company is decent and appreciated, where they have perks to encourage them.

* **Negative Inference :**

**Poor Work Culture :**

**🡪** Almost everyone has reviewed that the work culture is not good and hierarchy is maintained.

**Work and Life Balance :**

**🡪** Balancing the stress of work and the personal life is difficult.

* **Funny Review :**

**🡪** A Supervisor in Coco-Cola stated that he liked coco-cola brand as positive review and same time he mentioned negative review as he doesn’t like coco-cola which made me like Before joining It was his dream company after joining it was his nightmare haha!.  
  
Before Joining : Secret of coco-cola is Energetic.

After Joining : Secret of coco-cola is Chaotic.

* **Informatic Review:**

**🡪** Poor work culture started after the Franchise of **Coco-Cola** taken over by **Moon Beverages Ltd**, where the resources are not utilised well in a proper manner and being wasted as granted along with the management being biased hierarchally.

* Key findings :

🡪 The Coco-cola brand may be great and awesome as a product but the behind the scene are not as same.

🡪 The Work culture is poor and hierarchy is maintained with more work pressure and no proper work life balance.

🡪 But good for trainees and beginner as they encourage them by giving perks.

🡪 They have a good pay scale but the salary hike and increment is not that appreciated.

🡪 The shift rotation is not maintained properly and monopoly culture.