

ASSIGNMENT 4: PART C

TASK A: Build an event-driven serverless application using AWS Comprehend.

LINK TO GITLAB:

https://git.cs.dal.ca/rtyagi/csci5410_b00872269_ruhirajnish_tyagi/-/tree/main/assignment-4

Step 1: Create **twitterdatab00872269** bucket in S3 [1] and programmatically upload files to it.

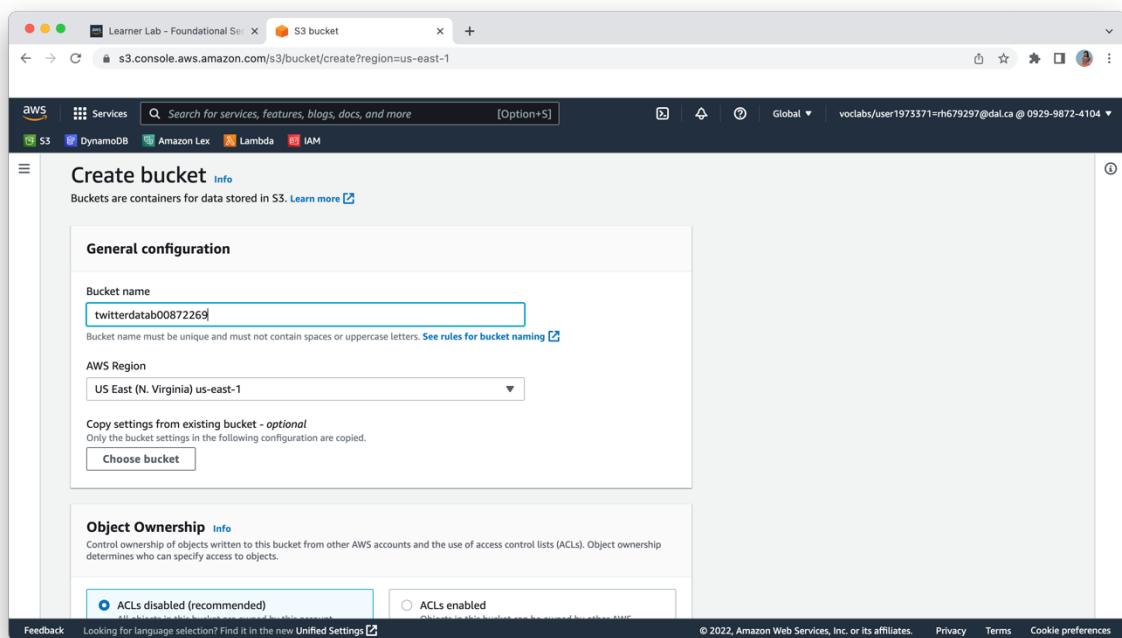


Figure 1. Screenshot of creating bucket **twitterdatab00872269** in S3.

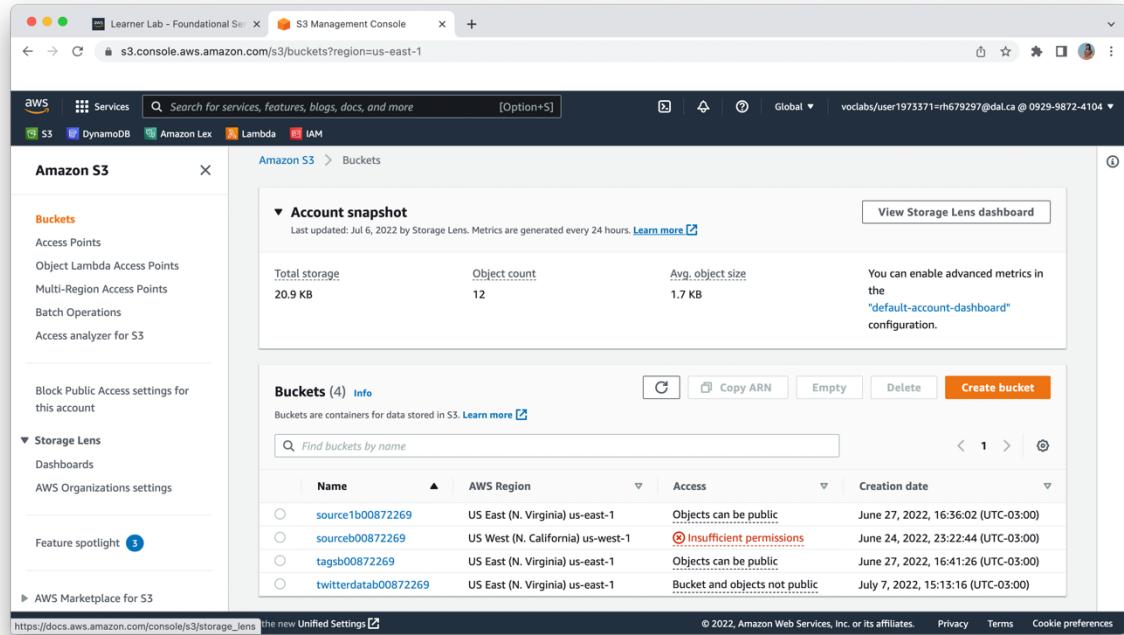


Figure 2. Screenshot of created bucket **twitterdatab00872269** in S3.

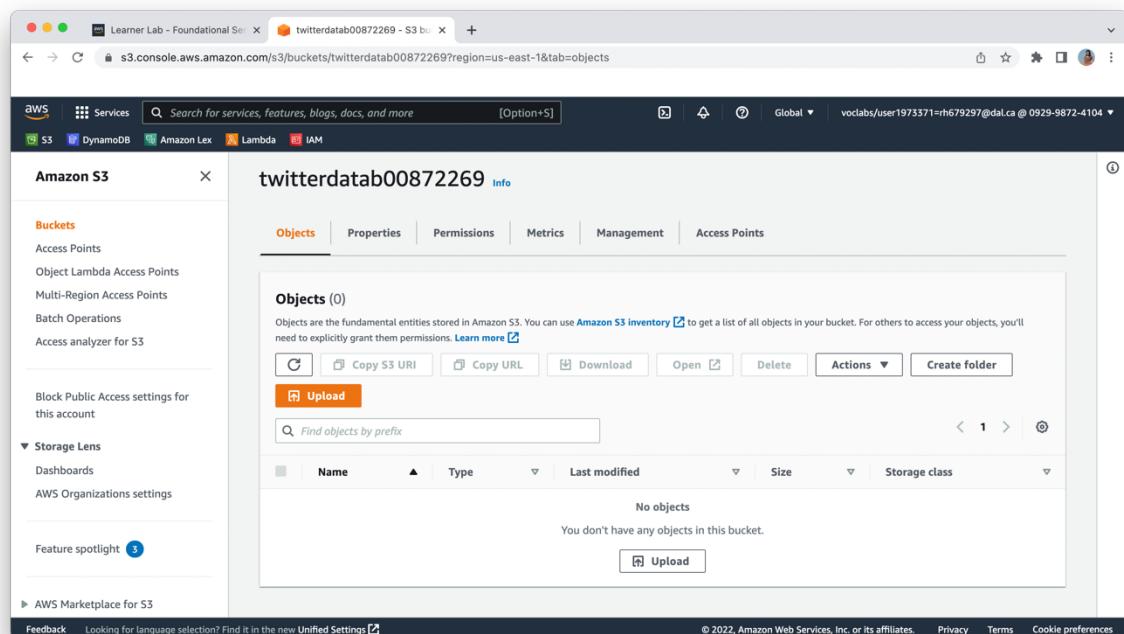


Figure 3. Screenshot of EMPTY bucket **twitterdatab00872269** in S3 .

The screenshot shows the PyCharm IDE interface. The project structure on the left includes files like main.py, sentiment_analysis.py, uploadFile.py, and file_mongo_tweets.txt. The main editor window displays the uploadFile.py script. The run output at the bottom shows the command being run and a success message: "Process finished with exit code 0".

```

assignment-4 - main.py
assignment-4 ~/Documents/Assignments_5410
  venv
    > bin
    > lib
      > ignore
      > Python
        pyvenv.cfg
      file_mongo_tweets.txt
    main.py
    sentiment_analysis.py
    uploadFile.py
  External Libraries
  Scratches and Consoles

Run: main
  /Users/ruhityagi/Documents/Assignments_5410_SDP/Coding/assignment-4/venv/bin/python /Users/ruhityagi/Documents/Assignments_5410_SDP/Coding/assignment-4/main.py
  File is pushed to twitterdataab00872269

Process finished with exit code 0

```

Figure 4. Screenshot of uploading `file_mongo_tweets.txt` to bucket `twitterdataab00872269` in S3 programmatically.

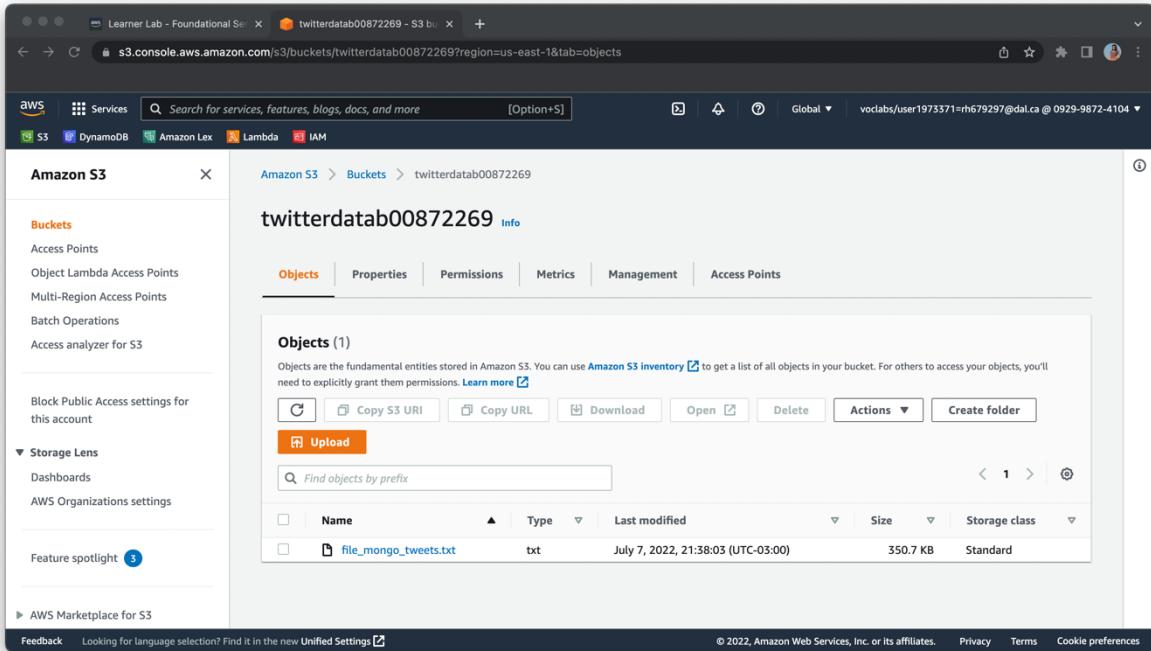


Figure 5. Screenshot of SUCCESSFUL upload of `file_mongo_tweets.txt` to bucket `twitterdataab00872269` in S3 programmatically.

Step 2: Create **sentimentAnalysis** lambda function and perform sentiment analysis employing AWS Comprehend[2] service.

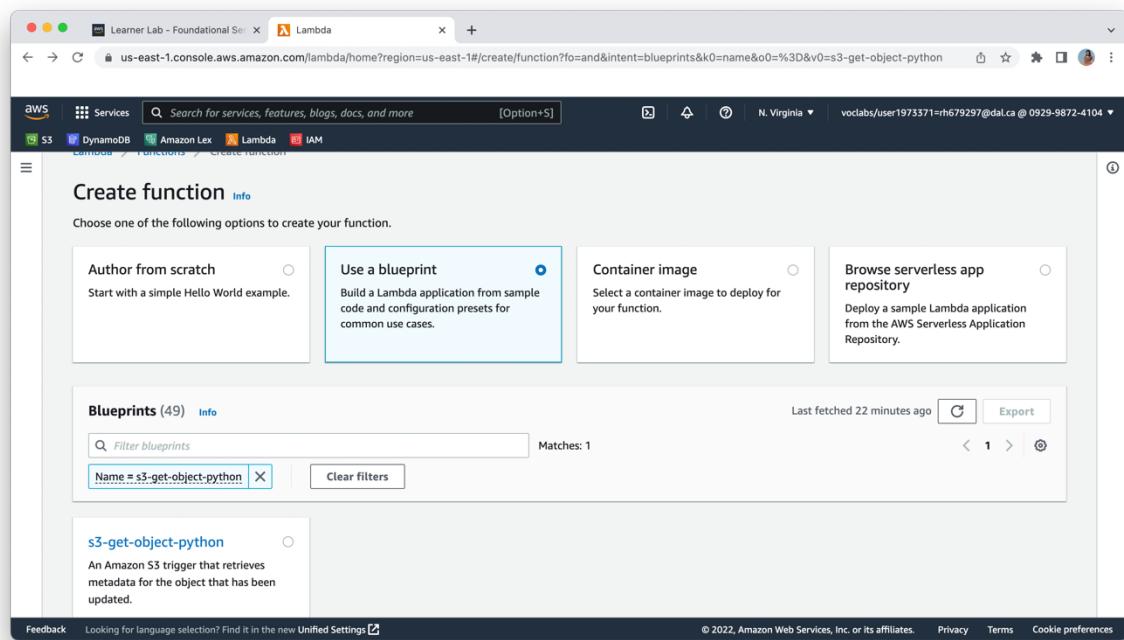


Figure 6. Screenshot of creating python based lambda function using blueprint.

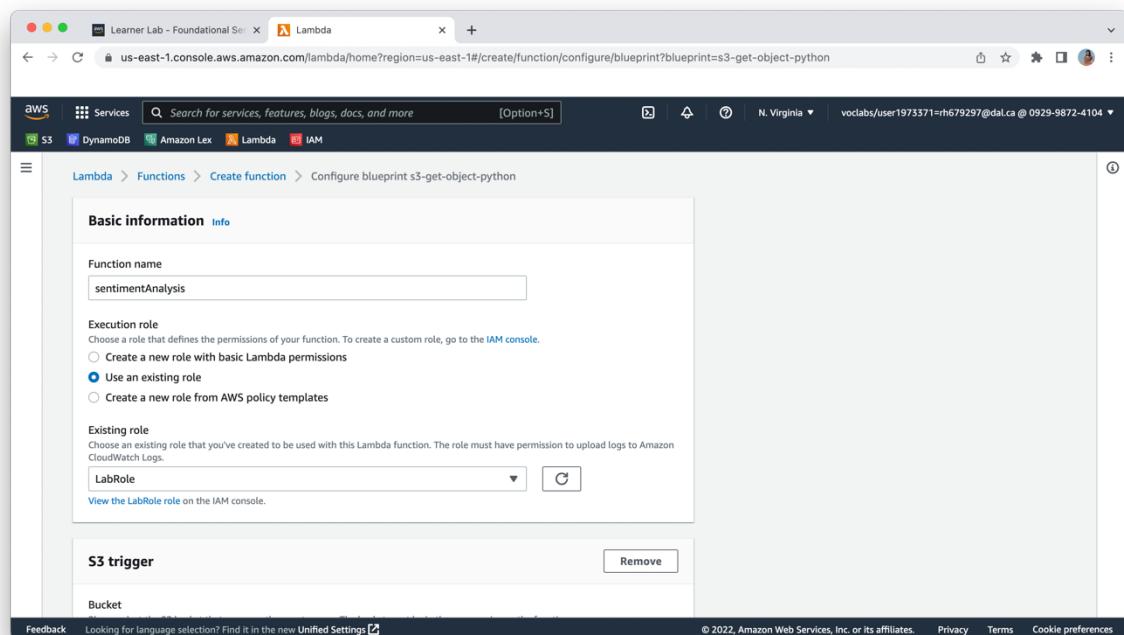


Figure 7. Screenshot of creating function **sentimentAnalysis**.

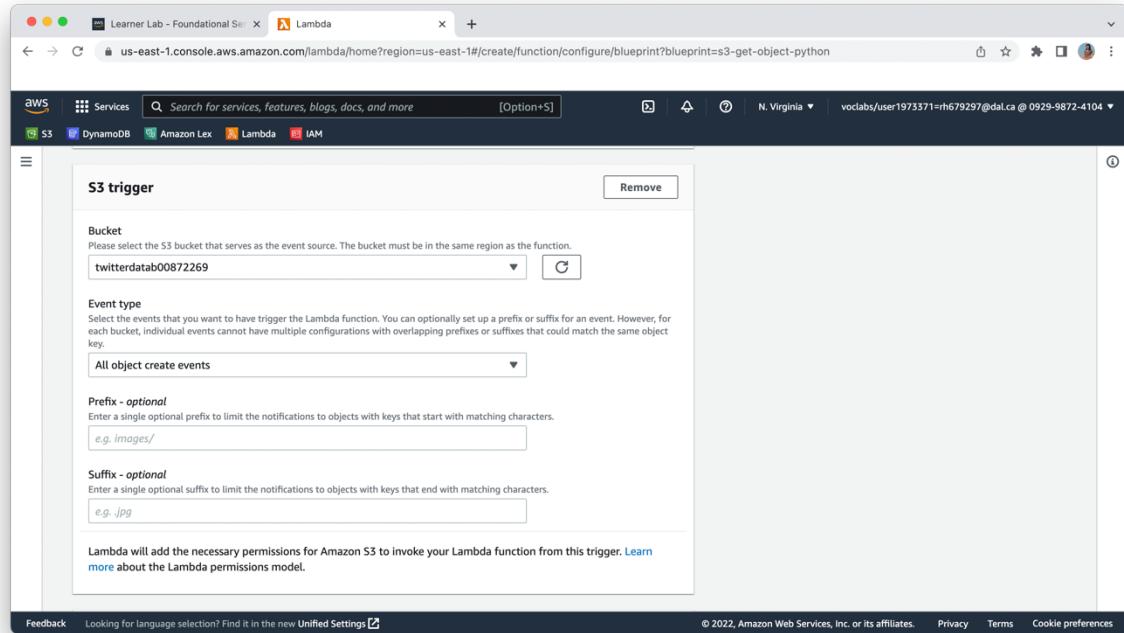


Figure 8. Screenshot of setting trigger as S3 bucket **twitterdatab00872269**.

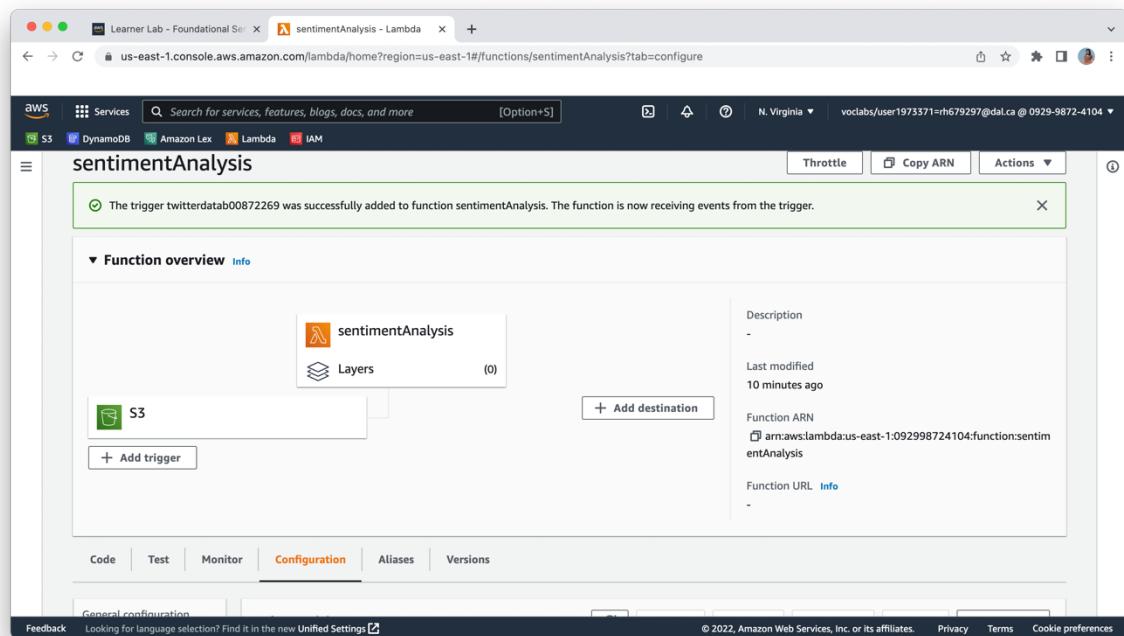


Figure 9(a). Screenshot of **SUCCESSFUL** creation of function **sentimentAnalysis** with S3 trigger as **twitterdatab00872269**.

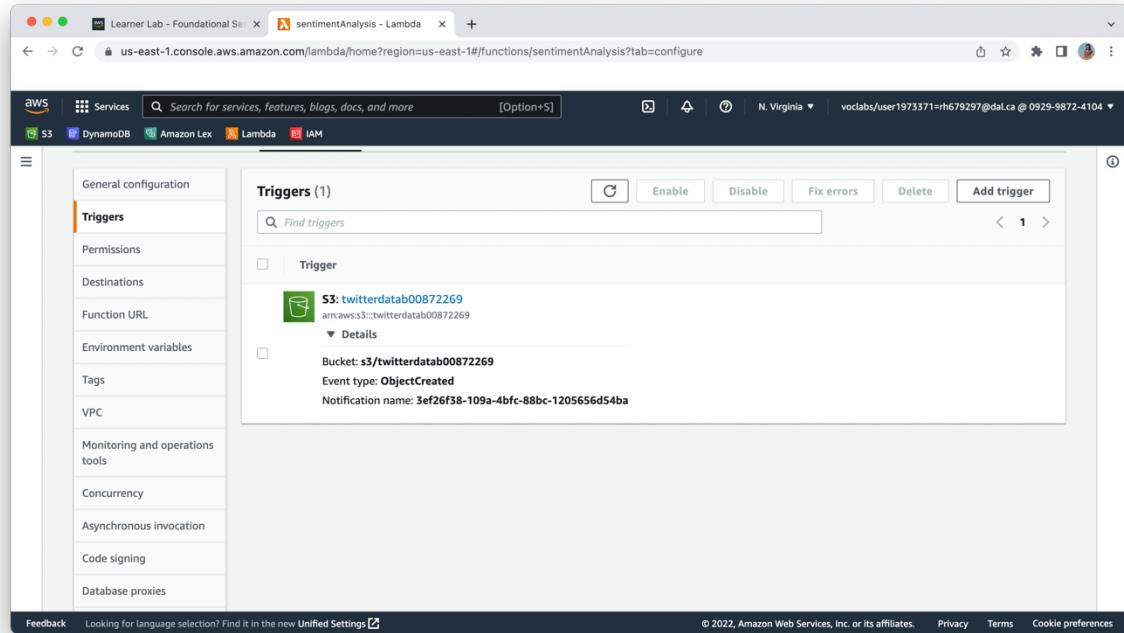


Figure 9(b). Screenshot of *SUCCESSFUL* creation of function **sentimentAnalysis** with S3 trigger as **twitterdata00872269**.

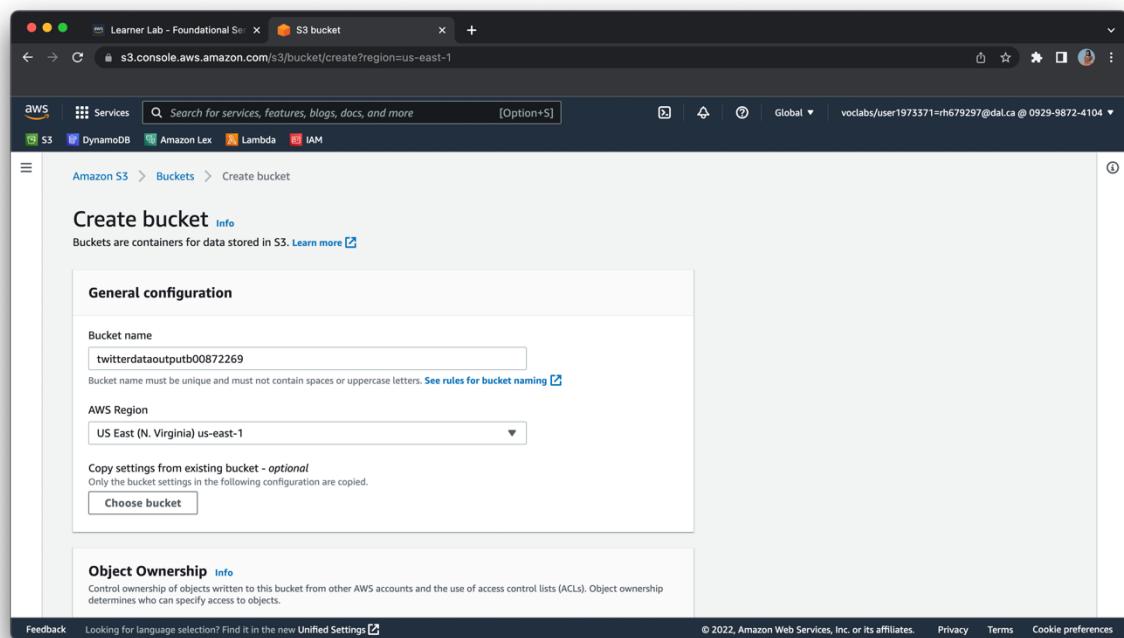


Figure 10. Screenshot of creating of S3 bucket **twitterdataoutputb00872269** to save the **output.json** generated by AWS comprehend.

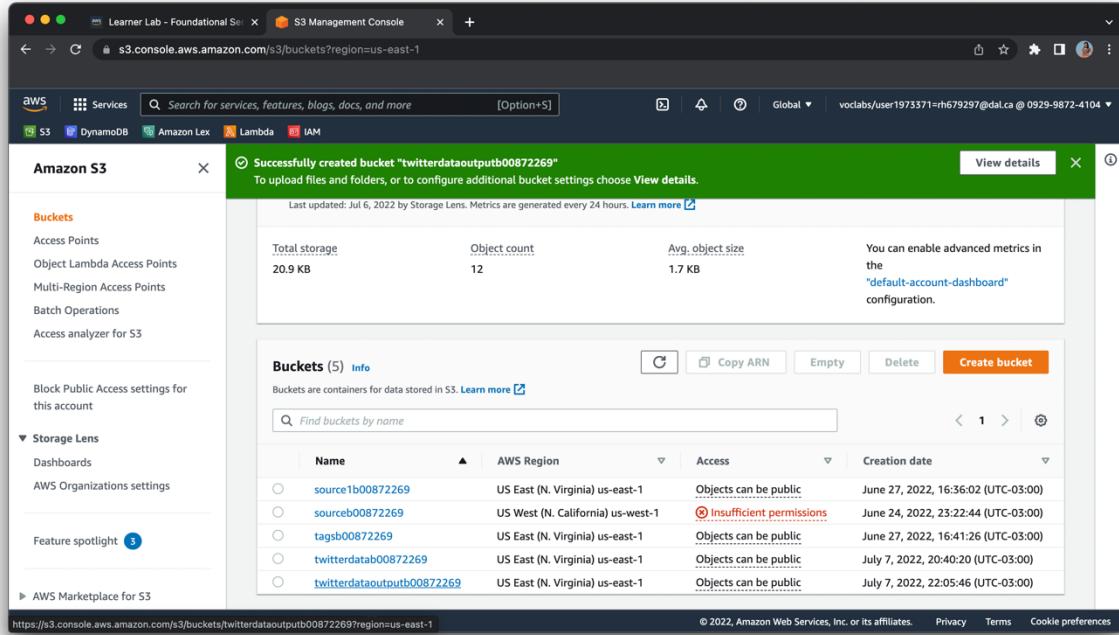


Figure 11. Screenshot of *SUCCESSFUL* creation of S3 bucket **twitterdataoutputb00872269**.

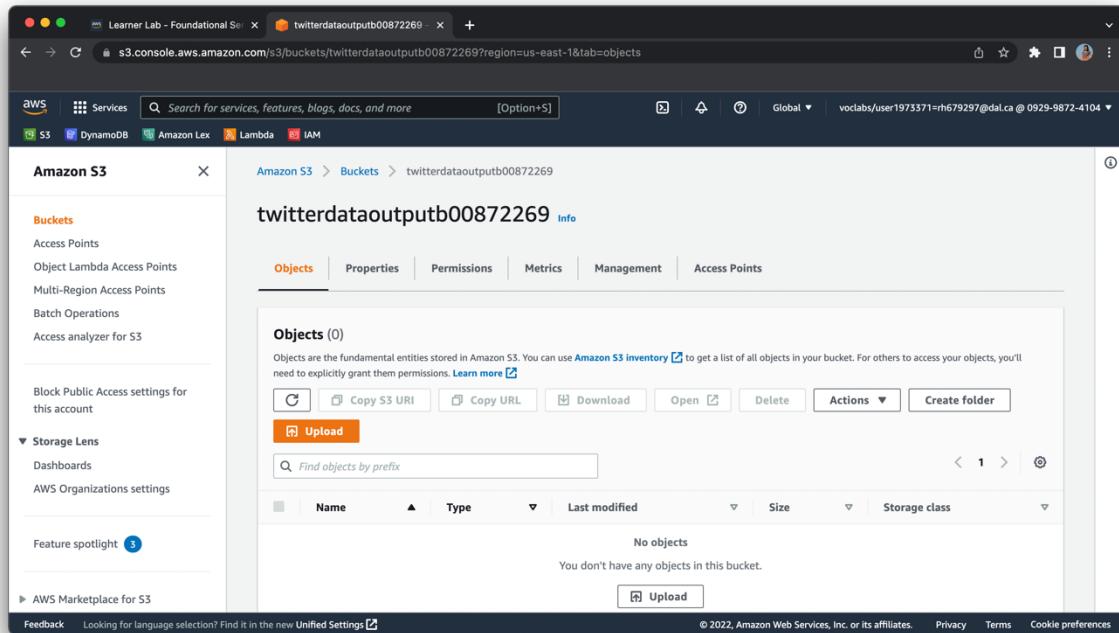


Figure 12. Screenshot of *EMPTY* bucket **twitterdataoutputb00872269** in S3 .

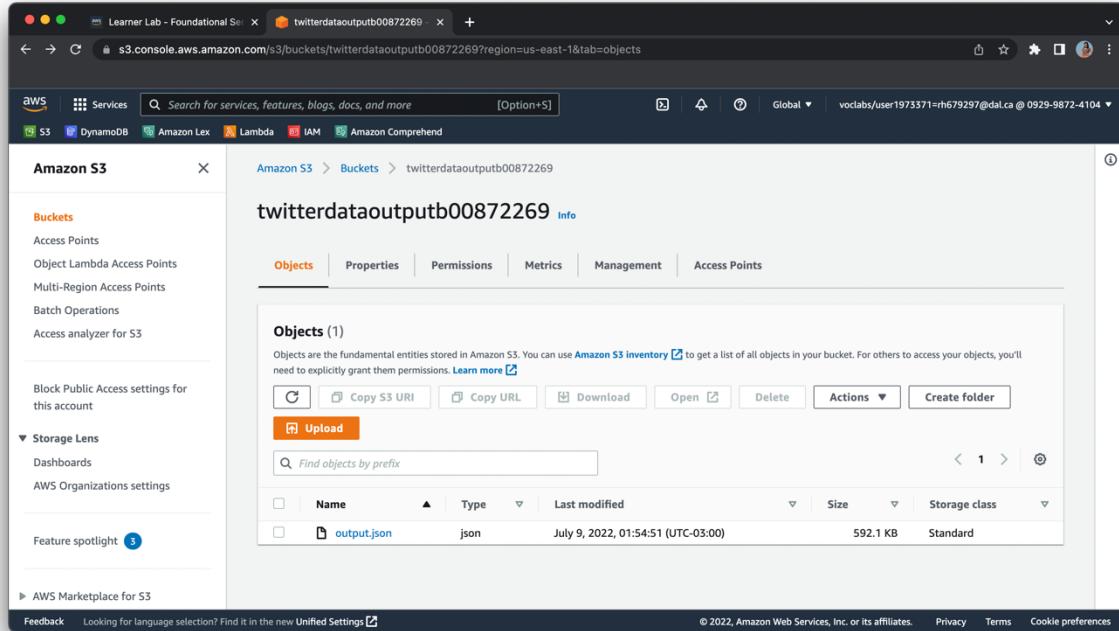


Figure 13. Screenshot of **output.json** saved in bucket **twitterdataoutputb00872269** in S3 .

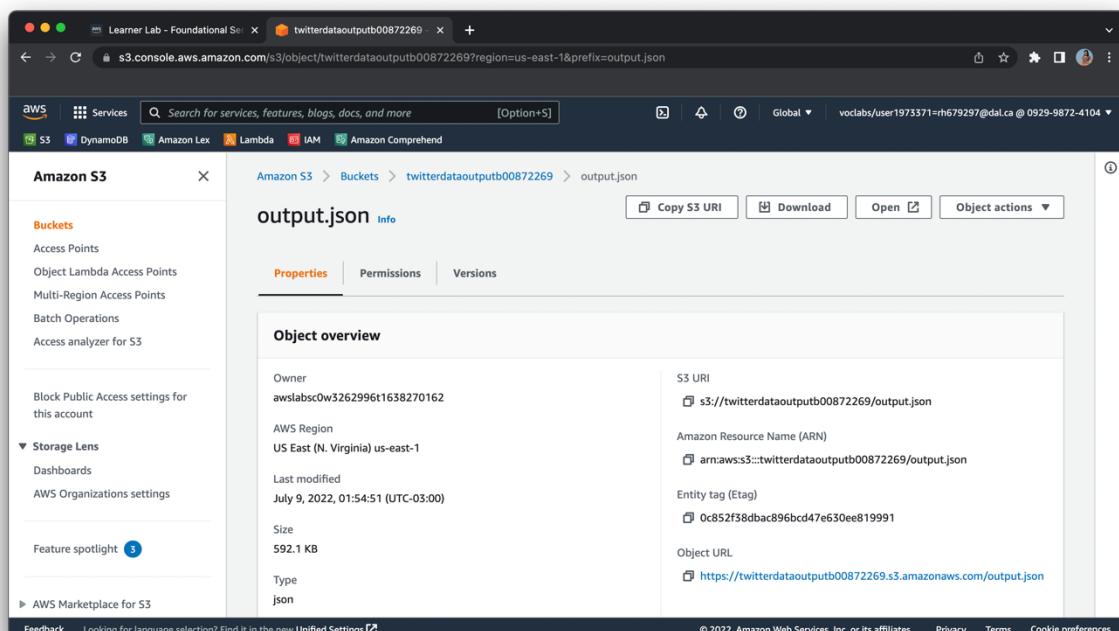


Figure 14. Screenshot of details of **output.json** saved in bucket **twitterdataoutputb00872269** in S3 .

```

[{"Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.7472493052482605, "Negative": 0.010755833238363266, "Neutral": 0.24056804180145264, "Mixed": 0, "RequestId": "0dead776-9858-4d6e-a4a2-967cf8e5237", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "0dead776-9858-4d6e-a4a2-967cf8e5237"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "165", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\ufeffRT @UnionSuite: Port-au-Prince fans takes to the streets to celebrate Haiti\u2019s Quarterfinal Gold Cup win against Canada! \u2026 Canada, thank you for a successful #StrokeMonth!\u2026 Thanks to FAST, more people in Canada are recognizing the signs of\u2026 \u2026 Authorities say they plan to return the artifacts in the near future and are working to determine who is criminally\u2026 \u2026 Many families travel to a cottage or lake house to enjoy the Canada Day long weekend! If you have a First Aid Kit \u2026 \u2026 Arnprior is setting fireworks off, as they do each year, from the island below the bridge at Hydro Park.", "Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.9827692672576904, "Negative": 0.001431510590912938, "Neutral": 0.09539277851581573, "Mixed": 0.00041544236592017114}, "RequestId": "378f7c5c-1ffe-43f5-9094-57fbccf79656", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "378f7c5c-1ffe-43f5-9094-57fbccf79656"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "166", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\u2026 Canada Day \u2019 Open, Closed, \u2026 Fireworks\u2026 Barrie #Huskoka #SimcoeCounty\u2026 \u2026 Jason Chatfield Canada sucks\u2026 Looking for something to do this Canada Day long weekend? Learn more about our planned celebrations, events and act\u2026 \u2026 We hope everyone is enjoying the long weekend. Happy Canada Day weekend!", "Sentiment": "NEUTRAL", "SentimentScore": {"Positive": 0.4551204741001129, "Negative": 0.0026463456451892853, "Neutral": 0.542102989288978e-05}, "RequestId": "84145770-4bcf-4be8-81da-023089e93b559", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "84145770-4bcf-4be8-81da-023089e93b559"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "163", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\u2026 #longweekend celebration \u2019The Ambassador, \u2026 Bridge, spanning the Detroit River between Detroit and Windsor, Ontario, Canada, under construction\u2026 \u2019Arguably the biggest star to come from #Canada and play in, \u2026 the #BBA is Steve Nash. Nash is an eight-time All-Star \u2026 Setting off fireworks to celebrate Canada\u2019s birthday! Make sure you stay safe and always abide by, \u2026 your municipal by\u2026 \u2019The museum will be closed on Monday, July 1 for the statutory holiday. Happy Canada Day to all! \u2026 Le mus\u00e9e sera ferm\u00e9 le lundi 1er juillet en raison du jour f\u00f4t\u00e9 du Canada \u2019ou\u00e9 ou\u00e9! \u2026 Fireworks kick off in Pembroke and Petawawa around 10:00pm on Monday night.", "Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.001132692443269663, "Negative": 0.0191855302178669, "Neutral": 0.22326034924481e-06}, "RequestId": "fff8b5a5-49f8-4376-9375-e77925805ecc", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "fff8b5a5-49f8-4376-9375-e77925805ecc"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "164", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\u2026 \u2019Nuit\u2019s Red Party Day! Why celebrate Canada Day anywhere else than at Toronto\u2019s BIGGEST pool party. ", "Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.7709631323814392, "Negative": 0.001188945141628934, "Neutral": 0.2278228104114525}, "RequestId": "2a6b5365-62e6-4009-8f06-041cbef0d87b", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "2a6b5365-62e6-4009-8f06-041cbef0d87b"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "165", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\u2026 \u2019Free before 3pm\u2026 \u2019Where in Canada would you like to travel to the most? #LongWeekendVibes \u2019Soyez un bon voisin. Respectez les autres lorsque vous \u00e7u\u00e9rez la fl\u00e9e du \u2019O\u00e9ate du Canada! ", "Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.0195770630836487, "Negative": 0.002611982403320543, "Neutral": 0.377801265716553, "Mixed": 3, "RequestId": "1270a7db-1667-43d3-9b5a-fbaa6702ca5b", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "1270a7db-1667-43d3-9b5a-fbaa6702ca5b"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "164", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}, "Tweet": "\u2026 \u2019Program director with Renfrew\u2019s Parks and Recreations, Jo-anne Caldwell, says the town will host a spectacular pyro\u2026 \u2019QUINTE WEST CANADA DAY", "Sentiment": "POSITIVE", "SentimentScore": {"Positive": 0.930101811885837, "Negative": 0.0003170739219658395, "Neutral": 0.06951446086168289, "Mixed": 6, "RequestId": "3bf14a84-72f7-4311-9ab2-cea3d9e678c2", "HTTPStatusCode": 200, "HTTPHeaders": {"x-amzn-requestid": "3bf14a84-72f7-4311-9ab2-cea3d9e678c2"}, "Content-Type": "application/x-amz-json-1.1", "Content-Length": "164", "Date": "Sat, 09 Jul 2022 04:54:27 GMT", "RetryAttempts": 0}
]

```

Figure 15. Screenshot of contents of `output.json`.

Step 3: Code samples of **uploadFile.py** [3] and **sentiment_analysis.py**.

1. uploadFile.py:

```
import logging
import boto3
from botocore.exceptions import ClientError

def upload_file(bucket):
    # Upload the file
    s3_client = boto3.client('s3', region_name='us-east-1',
                            aws_access_key_id='ASIARLJZH5IEEO7WKHK5',
                            aws_secret_access_key='EXxOXA2v5IyzQa40ZbH1RJC0UOhj6aPrDCwVS6k4',
                            aws_session_token='FwoGZXIvYXdzEMn//////////wEaDOQmyY9KmAiUH+OsbCLAAZYFUqOo
8OdZAymm9rrD6C3fFfe9nOxDuqqoyUL6aWMbyGVluRz0pMWr3SOMF2JjhxqCqcwe7sIqb/LUWap
YSjrucy+ar0zvGS0dG80e6z4Fa7qmZvUTskAWN2PZLDFnej/dGH/mYLsThCCPSBFFPSB5sQdYne7
6UI+10j10K5iXQcmyb+eSLj1R72s3fxdstJbf0Oz+9OG8zf9SsTGB5dcnw20B4zU+ByMtuPlP5a
55R8PdPsEK7ggEMtgpvkZbsCiic2p2WBjItWBMMt5YXD2cA+URIZ5Mv9mcfuNOC4zQ1vKPQ2dVK
15zK6mTMO3CtiG6boXkC'
    )

    try:

        s3_client.upload_file('/Users/ruhityagi/Documents/Assignments_5410_SDP/Codi
ng/assignment-4/file_mongo_tweets.txt', bucket, 'file_mongo_tweets.txt')
        print('File is pushed to {}'.format(bucket))
    except ClientError as e:
        logging.error(e)
        return False
    return True
```

2. sentiment_analysis.py:

```

import json
import urllib.parse
import re

import boto3

print('Loading function')

s3_client = boto3.client('s3', region_name='us-east-1',
                        aws_access_key_id='ASIAIRLJZH5IEBTCUQXWB',
                        aws_secret_access_key='tjxOEK2AvtiUL1/e4RjMYBO91ovfkdpLf6ehV/2t',
                        aws_session_token='FwoGZXIvYXdzEOP//////////wEaDKohCECUKwMBDQq29iLAAX3REKh9
fublmxPgG9Gcu6/da07jVgMBcR1Y6J7AqcA/enRbvnaR8LHL4iMfeJz/lY7ihs0JjrCIM34d1kg
KeRGPP6E10J83mHCDwzxjiN1rQALNuH49h5b+jPQiJR+S2kAUwkx3aEppN/KdwkP1o1A4Iyoq9M
xFOMxCld8k1JXj4lfZ2vAzYXy4k1miwBCwcyAnBmXy93JW2zUvR1SXdFmxzZsJPppVqYlaCZz1
hPRB5B3nySbyrVBuwUwWyTtQyjAvKOWBjIthN18KPasyDrGIssSNV1oq2yjVBSzHwsm+4LrHjehi
W4wcZfRK8f8JN2dsyk+A'
                    )

comprehend = boto3.client('comprehend', region_name='us-east-1',
                        aws_access_key_id='ASIAIRLJZH5IEBTCUQXWB',
                        aws_secret_access_key='tjxOEK2AvtiUL1/e4RjMYBO91ovfkdpLf6ehV/2t',
                        aws_session_token='FwoGZXIvYXdzEOP//////////wEaDKohCECUKwMBDQq29iLAAX3REKh9
fublmxPgG9Gcu6/da07jVgMBcR1Y6J7AqcA/enRbvnaR8LHL4iMfeJz/lY7ihs0JjrCIM34d1kg
KeRGPP6E10J83mHCDwzxjiN1rQALNuH49h5b+jPQiJR+S2kAUwkx3aEppN/KdwkP1o1A4Iyoq9M
xFOMxCld8k1JXj4lfZ2vAzYXy4k1miwBCwcyAnBmXy93JW2zUvR1SXdFmxzZsJPppVqYlaCZz1
hPRB5B3nySbyrVBuwUwWyTtQyjAvKOWBjIthN18KPasyDrGIssSNV1oq2yjVBSzHwsm+4LrHjehi
W4wcZfRK8f8JN2dsyk+A'
                    )

def lambda_handler(event, context):
    # print("Received event: " + json.dumps(event, indent=2))
    try:
        bucket = 'twitterdataab00872269'
        key = 'file_mongo_tweets.txt'
        response = s3_client.get_object(Bucket=bucket, Key=key)
        paragraph = response["Body"].read().decode('utf-8')
        paragraph = paragraph.split('\n\n')
        print(len(paragraph))
        result = []
        for i in range(len(paragraph)):
            try:
                if len(paragraph[i]) and len(paragraph[i].encode('utf-8')) <= 5000:
                    sentiment =
comprehend.detect_sentiment(Text=paragraph[i], LanguageCode="en")
                    sentiment["tweet"] = paragraph[i]
                    result.append(sentiment)
            except Exception as e:
                print(e)
        s3_client.put_object(Bucket='twitterdataoutputb00872269',
Key='output.json', Body=json.dumps(result))

    return response['ContentType']

```

```
except Exception as e:  
    print(e)  
    print(  
        'Error getting object {} from bucket {}. Make sure they exist  
and your bucket is in the same region as this function.'.format(  
            key, bucket))  
    raise e
```

REFERENCES:

- [1] "Amazon S3 buckets — Boto3 Docs 1.24.17 documentation", Boto3.amazonaws.com, 2022. [Online]. Available: <https://boto3.amazonaws.com/v1/documentation/api/latest/guide/s3-examplecreating-buckets.html>. [Accessed: 9- Jul- 2022]
- [2] "Amazon Comprehend", Amazon.com, 2022. [Online]. Available: <https://docs.aws.amazon.com/comprehend/index.html>. [Accessed: 09- Jul- 2022]
- [3] S. Dey, Assignment3. Halifax: Saurabh Dey, 2022, p. Part A.