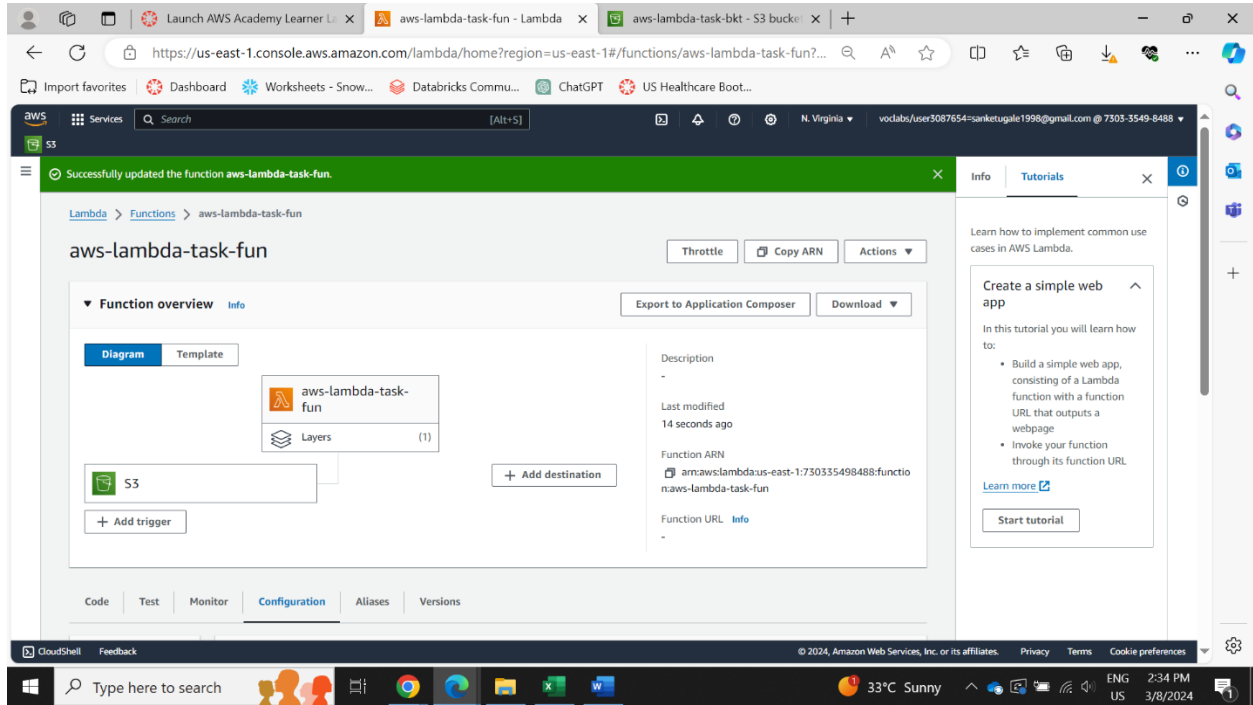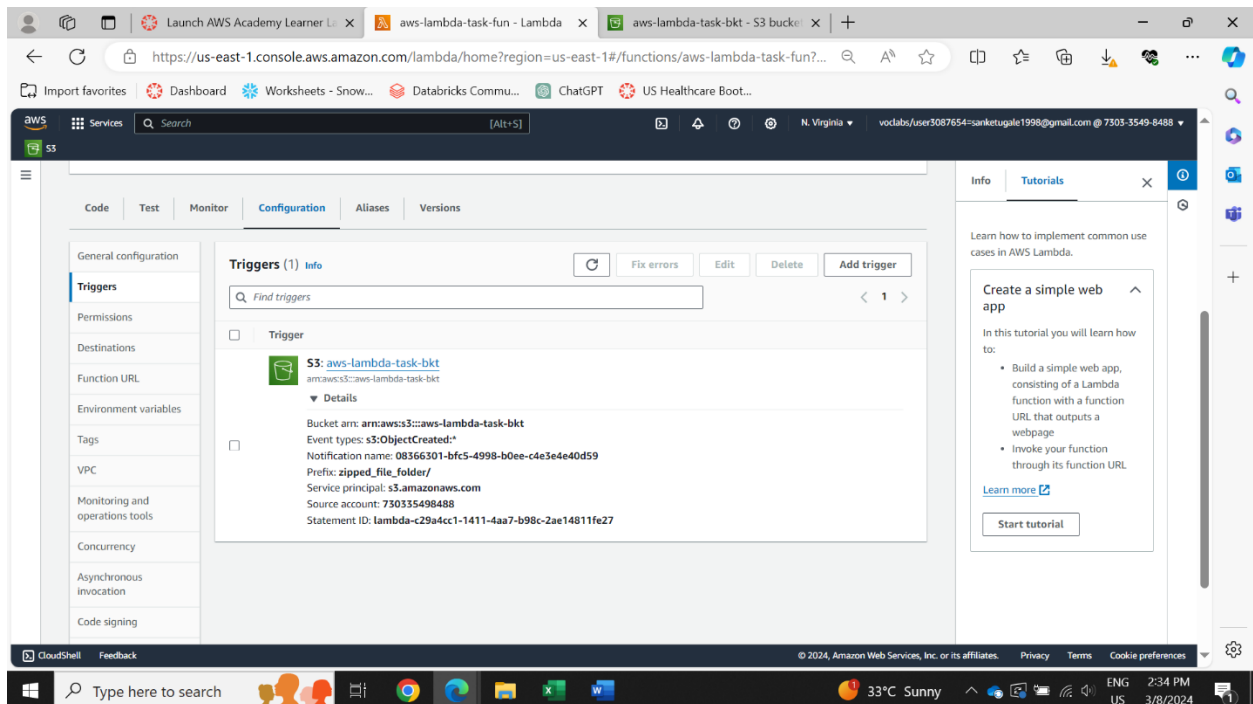# AWS Lambda Task

STEP : Create a S3 Bucket and a Lambda Function
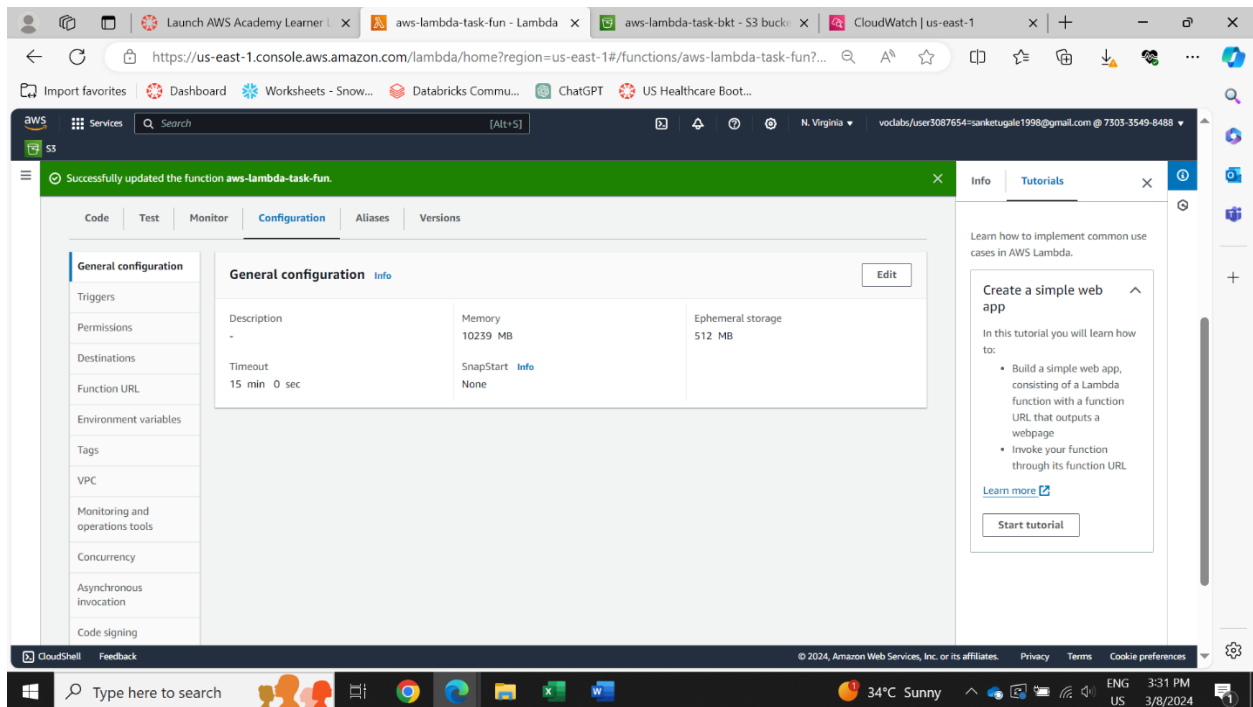
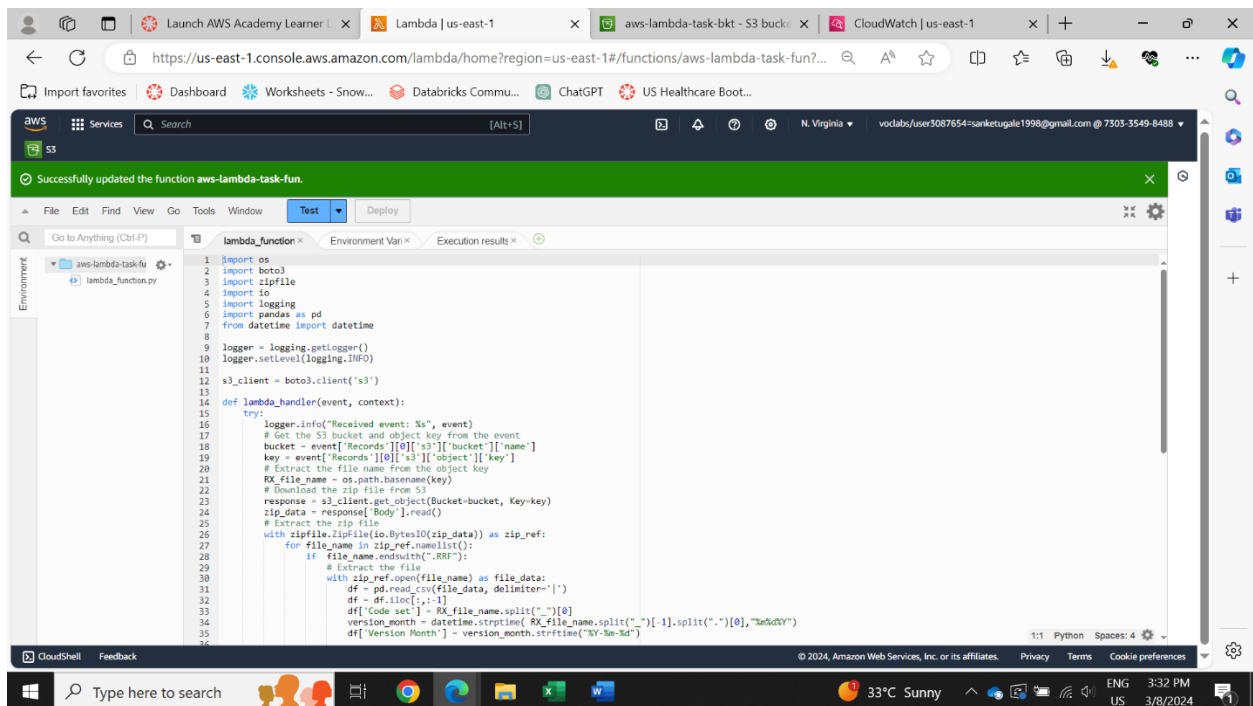1) In S3 bucket create 3 folders source,Header, Target

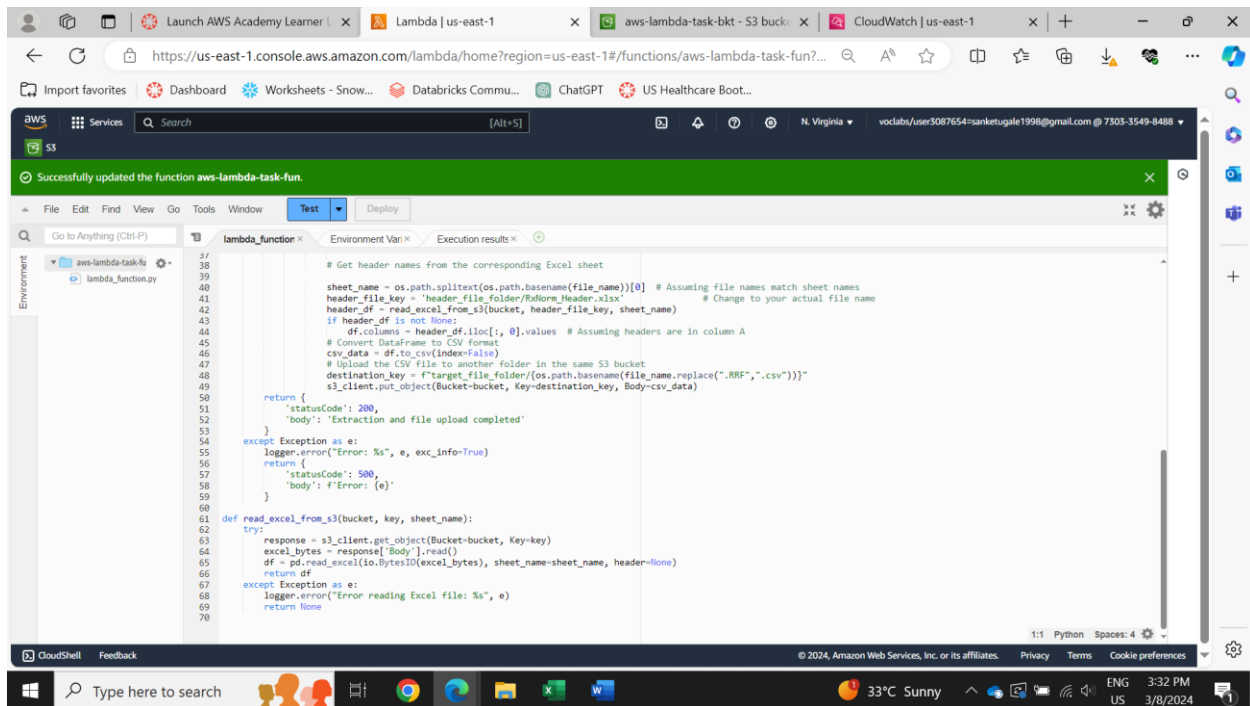STEP 2 : Create a trigger which will invoke when the (.rrf) file is inserted into source table in the S3 bucket.

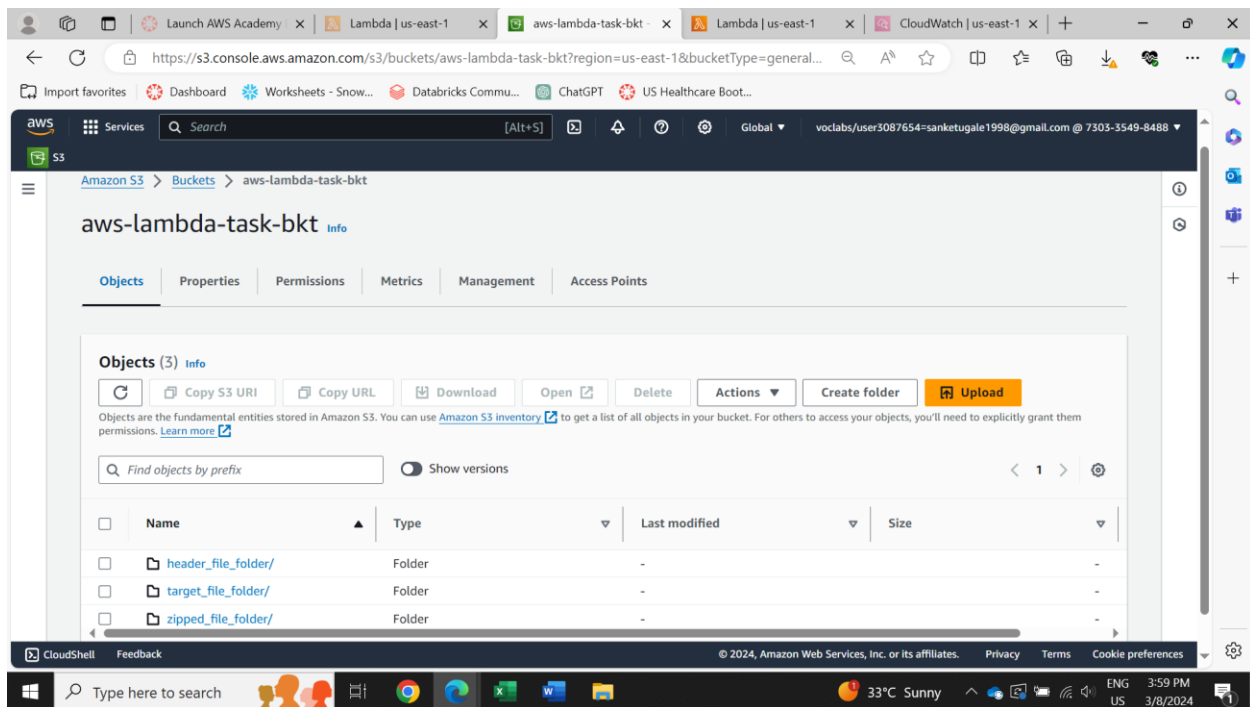## STEP : Update the configuration of the lamda Function

Lambda Handler Code


```python
import os
import boto3
import zipfile
import io
import logging
import pandas as pd
from datetime import datetime

logger = logging.getLogger()
logger.setLevel(logging.INFO)

s3_client = boto3.client('s3')

def lambda_handler(event, context):
    try:
        logger.info("Received event: %s", event)
        # Get the S3 bucket and object key from the event
        bucket = event['Records'][0]['s3']['bucket']['name']
        key = event['Records'][0]['s3']['object']['key']
        # Extract the file name from the object key
        RX_file_name = os.path.basename(key)
        # Download the zip file from S3
        response = s3_client.get_object(Bucket=bucket, Key=key)
        zip_data = response['Body'].read()
        # Extract the zip file
        with zipfile.ZipFile(io.BytesIO(zip_data)) as zip_ref:
            for file_name in zip_ref.namelist():
                if file_name.endswith(".RRF"):
                    # Extract the file
                    with zip_ref.open(file_name) as file_data:
                        df = pd.read_csv(file_data, delimiter='|')
                        df = df.iloc[:,:-1]
                        df['Code set'] = RX_file_name.split("_")[0]
                        version_month = datetime.strptime( RX_file_name.split("_")[-1].split(".")[0],"%m%d%Y")
                        df['Version Month'] = version_month.strftime("%Y-%m-%d")
```

Bucket with 3 folders

ZIP file in the source folder.

Excel File with the headers in the header folder.

Processed Files in the Target folder.
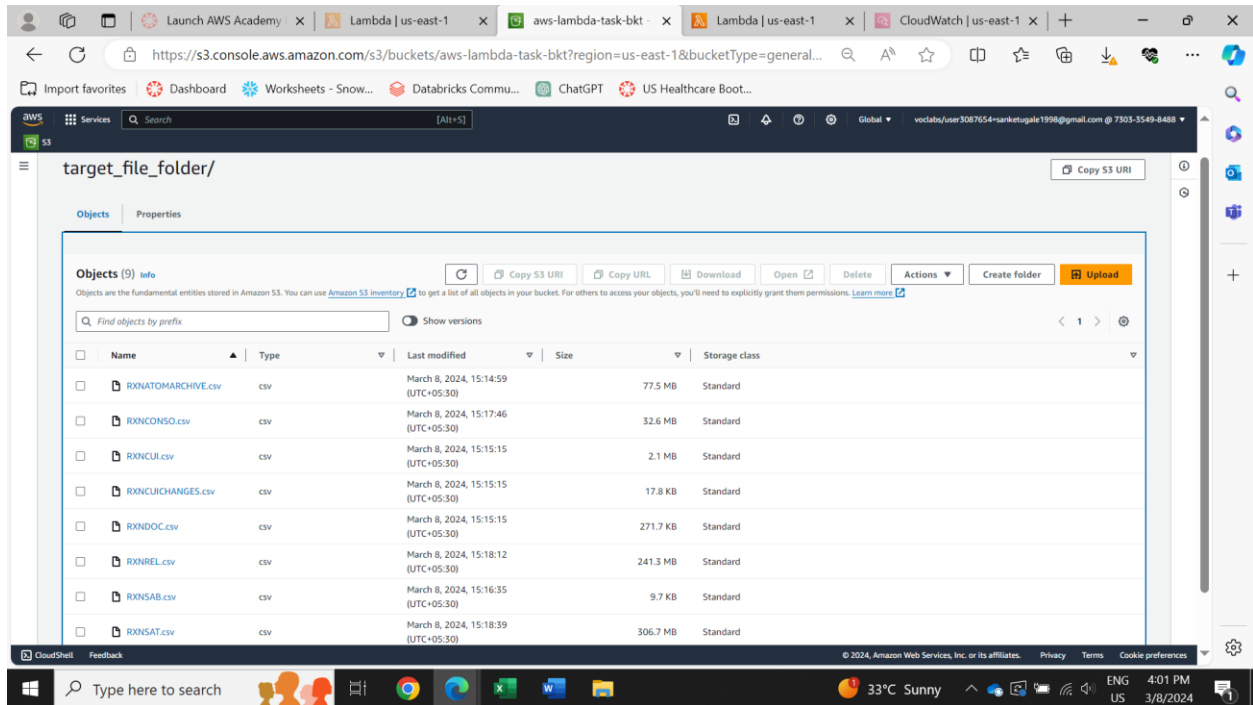
Processed File with headers and the Code set, Version month columns.

CUI1	VER_START	VER_END	Cardinality	CUI2	Code set	Version Month
91669	RXNORM_17AA_170905F	RXNORM_18AB_181203F	1	91669	RxNorm	2/5/2024
91714	RXNORM_10AA_100802F	RXNORM_10AA_100802F	1	1000992	RxNorm	2/5/2024
91811	RXNORM_04AC_050210F	RXNORM_09AA_090706F	1	857082	RxNorm	2/5/2024
91812	RXNORM_04AC_050210F	RXNORM_09AA_090706F	1	857087	RxNorm	2/5/2024
91919	RXNORM_04AC_050210F	RXNORM_09AA_090504F	1	849580	RxNorm	2/5/2024
91942	RXNORM_12AA_120507F	RXNORM_12AA_120507F	1	1293743	RxNorm	2/5/2024
91960	RXNORM_10AA_100503F	RXNORM_10AA_100503F	1	991063	RxNorm	2/5/2024
91961	RXNORM_10AA_100503F	RXNORM_10AA_100503F	1	991088	RxNorm	2/5/2024
92286	RXNORM_04AC_050210F	RXNORM_09AB_100104F	1	895989	RxNorm	2/5/2024
92548	RXNORM_10AB_101206F	RXNORM_10AB_101206F	1	1048284	RxNorm	2/5/2024
92549	RXNORM_10AB_101206F	RXNORM_10AB_101206F	1	1048288	RxNorm	2/5/2024
92643	RXNORM_04AC_050210F	RXNORM_09AA_090706F	1	856710	RxNorm	2/5/2024
92690	RXNORM_10AA_100607F	RXNORM_10AA_100607F	1	994091	RxNorm	2/5/2024
92691	RXNORM_10AA_100607F	RXNORM_10AA_100607F	1	994093	RxNorm	2/5/2024
92840	RXNORM_04AC_050210F	RXNORM_09AA_090803F	1	859873	RxNorm	2/5/2024
92841	RXNORM_04AC_050210F	RXNORM_09AA_090803F	1	859869	RxNorm	2/5/2024
92887	RXNORM_04AC_050210F	RXNORM_09AA_090706F	1	856955	RxNorm	2/5/2024
93008	RXNORM_04AC_050210F	RXNORM_09AA_090504F	1	793871	RxNorm	2/5/2024
93162	RXNORM_04AC_050210F	RXNORM_09AB_091102F	1	857113	RxNorm	2/5/2024

RXNCUI

Cloud Watch Logs

Timestamp	Message
	No older events at this moment. Retry
2024-03-08T15:14:47.156+05:30	INIT_START Runtime Version: python:3.12.v20 Runtime Version ARN: arn:aws:lambda:us-east-1::runtime:82aea00f37a44d68665730d559c81352fc95f2…
2024-03-08T15:14:49.843+05:30	[INFO] 2024-03-08T09:44:49.843Z Found credentials in environment variables.
2024-03-08T15:14:49.979+05:30	START RequestId: 4462c2b0-9cb3-4b04-9bdd-3ff9560ddb4a Version: $LATEST
2024-03-08T15:14:49.980+05:30	[INFO] 2024-03-08T09:44:49.980Z 4462c2b0-9cb3-4b04-9bdd-3ff9560ddb4a Received event: {'Records': [{'eventVersion': '2.1', 'eventSource': …
2024-03-08T15:15:02.488+05:30	/var/task/lambda_function.py:31: DtypeWarning: Columns (9) have mixed types. Specify dtype option on import or set low_memory=False.
2024-03-08T15:15:02.488+05:30	df = pd.read_csv(file_data, delimiter='|')
2024-03-08T15:15:24.023+05:30	/var/task/lambda_function.py:31: DtypeWarning: Columns (7) have mixed types. Specify dtype option on import or set low_memory=False.
2024-03-08T15:15:24.023+05:30	df = pd.read_csv(file_data, delimiter='|')
2024-03-08T15:16:45.453+05:30	/var/task/lambda_function.py:31: DtypeWarning: Columns (6) have mixed types. Specify dtype option on import or set low_memory=False.
2024-03-08T15:16:45.453+05:30	df = pd.read_csv(file_data, delimiter='|')
2024-03-08T15:18:17.597+05:30	/var/task/lambda_function.py:31: DtypeWarning: Columns (6) have mixed types. Specify dtype option on import or set low_memory=False.
2024-03-08T15:18:17.597+05:30	df = pd.read_csv(file_data, delimiter='|')
2024-03-08T15:18:41.085+05:30	END RequestId: 4462c2b0-9cb3-4b04-9bdd-3ff9560ddb4a
2024-03-08T15:18:41.085+05:30	REPORT RequestId: 4462c2b0-9cb3-4b04-9bdd-3ff9560ddb4a Duration: 231110.39 ms Billed Duration: 231111 ms Memory Size: 10239 MB Max Memory…
	No newer events at this moment. Auto retrying… Pause