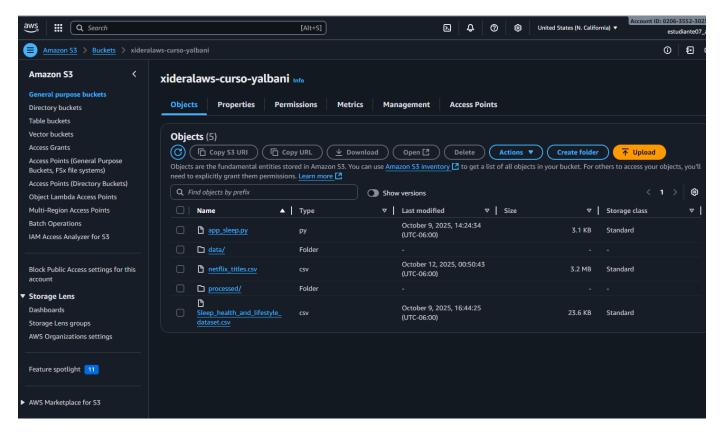
## TAREA 5

## LAMBDA A STRAMLIT

Contenido S3



• Funcion LAMBDA. Guarda procesamiento en carpeta processed

```
∠ xideralaws-yalbani

₱ lambda_function.py ×
                                                                                                                                                                                                                        □ ...
=
       ZIDERΔΙ ΔWS-VΔΙ RΔNΙ
                                   lambda_function.py
Ф
        lambda_function.py
                                          import json
Q
                                          import pandas as pd
                                          import datetime
$ \
                                          def lambda handler(event, context):
8
                                                 --- Configuracion de S3 --
                                              bucket = "xideralaws-curso-yalbani"
key = "netflix_titles.csv"
OUTPUT_PREFIX = "processed/"
                                    10
\mathbb{A}
                                   12
                                    14
                                              # Generar nuevo nombre
                                              timestamp = datetime.datetime.now().strftime("%V %m %d %H %N %S")

OUTPUT_KEY = f"{OUTPUT_PREFIX}netflix_directores_count_{timestamp}.csv"
                                   16
17
18
                                               s3 = boto3.client("s3")
                                    19
20
                                                   # LECTURA
                                                   obj = s3.get_object(Bucket=bucket, Key=key)
body = obj["Body"].read()
df = pd.read_csv(io.BytesIO(body))
                                   21
22
      ∨ DEPLOY
                                   23
24
25
26
27
                                                   # PROCESAMIENTO EJEMPLO
                                                   # ESCRITURA
                                                   # Convertir el DataFrame procesado a CSV en memoria (StringIO)
                                                   csv_buffer = io.StringIO()
      V TEST EVENTS [SELECTED: T...
                                  PROBLEMS OUTPUT CODE REFERENCE LOG TERMINAL
                                                                                                                                                                                 Execution Results
                                                                                                                                                                                                       + Create new test...
                                  Status: Succeeded

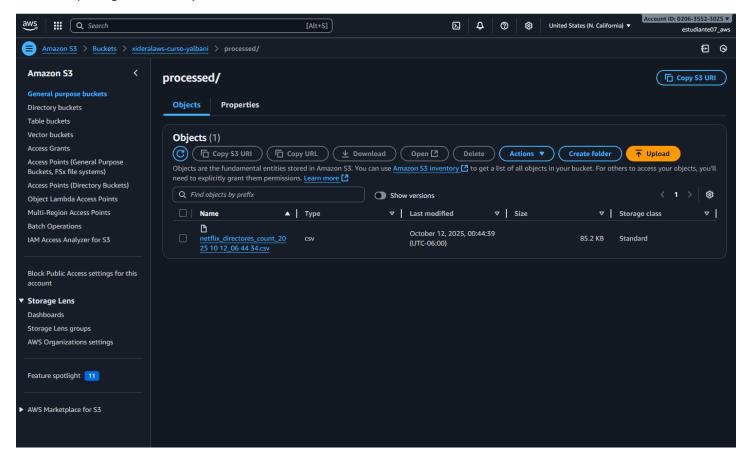
∨ A Private saved ev...

                                  Test Event Name: test
           test
                                  Response:
                                    "body": '"status'": \"success\", \"message\": \"Datos de Netflix procesados y guardados en S3.\", \"output_file\": \"s3://xideralaws-curso-yalbani/processed/netflix_directores_count_2025 10 12_06 44 34.csv\"}"
                                                                                                                                                           1 No changes to deploy
      > ENVIRONMENT VARIABLES
```

• Con el csv guardado, se activa el streamlit

```
Jupyter app_lambda.py Last Checkpoint: 6 minutes ago
 File Edit View Settings Help
 1 import pandas as pd
    import streamlit as st
 3 import boto3
 5 import requests
 7 import plotly.express as px
 9 # --- Carga de Datos Procesados desde S3 ---
10 @st.cache_data
11 def cargar_datos_procesados():
12     s3 = boto3.client("s3")
        bucket = "xideralaws-curso-yalbani"
        OUTPUT_PREFIX = "processed/"
15
16
17
18
19
20
21
22
            # Listar todos los archivos en la carpeta procesada
response = s3.list_objects_v2(Bucket=bucket, Prefix=OUTPUT_PREFIX)
             if "Contents" not in response:
                  \mathsf{st.warning}(\texttt{"Lambda no ha ejecutado el procesamiento. No se encontraron archivos procesados."})
                 return pd.DataFrame()
23
24
             # Encontrar el archivo mas reciente (timestamp mas alto)
             all_files = response["Contents"]
25
26
27
28
             #Filtrar solo archivos CSV si fuera necesario, y ordenar por la ultima modificacion all_files = sorted(all_files, key=lambda x: x['LastModified'], reverse=True)
29
30
31
             latest_key = all_files[0]['Key']
             # Leer el archivo mas reciente
32
33
34
35
             st.info(f"Cargando el ultimo archivo procesado: {latest_key.split('/')[-1]}")
             obj = s3.get_object(Bucket=bucket, Key=latest_key)
             body = obj["Body"].read()
36
37
             df_final = pd.read_csv(io.BytesIO(body))
38
39
             return df final
40
        except Exception as e:
41
             st.error(f"Error al cargar el CSV procesado desde S3: {e}")
42
43
             return pd.DataFrame()
44 # --- Streamlit ---
45
```

Aquí se guarda el csv procesado



## Resultado Streamlit

