In this exercise, you will use AWS SageMaker to preprocess a dataset for a machine learning model. The dataset will be cleaned, transformed, and made ready for prediction purposes. You will create a SageMaker processing job and trigger it from an AWS Glue job on a weekly schedule.

Objectives:

* Create a SageMaker processing job to clean and transform a dataset.
* Store the preprocessed dataset in an S3 bucket.
* Trigger the SageMaker processing job from an AWS Glue job on a weekly schedule.

**Step 1: DB Preparation**

* Ensure that the RDS DB is online and you have the necessary information (table name, schema, connection details) to connect to it.

**Step 2: Create a SageMaker Processing Job (Base on Glue exercise)**

* Create a new SageMaker processing job using the SageMaker Python SDK.
* In the processing job, perform the following data preprocessing steps:
  + Extract the domain name from the email column and create a new column called "domain". Remove the original email column.
  + Apply one-hot encoding to the country column and remove the original country column.
  + Create a new column called "is\_male" based on the gender column. Remove the original gender column.
  + Fill age nulls with the mean average.
  + Convert the age column into 4 different groups: 0-18, 18-30, 30-50, 50+.
  + Convert the "created\_at" column into separate columns: "created\_at\_day", "created\_at\_month", "created\_at\_year".
* Save the preprocessed dataset to an S3 bucket (create a new bucket if needed).

**Step 3: Create an AWS Glue Job to Trigger the SageMaker Processing Job:**

* Create a new AWS Glue job.
* In the Glue job:
  + Read user tables from RDS and save it to s3
  + use the AWS SDK (e.g., Boto3) to trigger the SageMaker processing job you created in Step 2 with the s3 file.
  + Set up a schedule for the Glue job to run every week at a specific time (e.g., every Sunday at 00:00).

**Step 4: Run the Glue Job and Monitor the SageMaker Processing Job**

Execute the Glue job manually to test the integration with the SageMaker processing job.

Monitor the progress of the SageMaker processing job in the SageMaker Console.

Once the processing job completes, verify that the preprocessed dataset is stored in the specified S3 bucket.