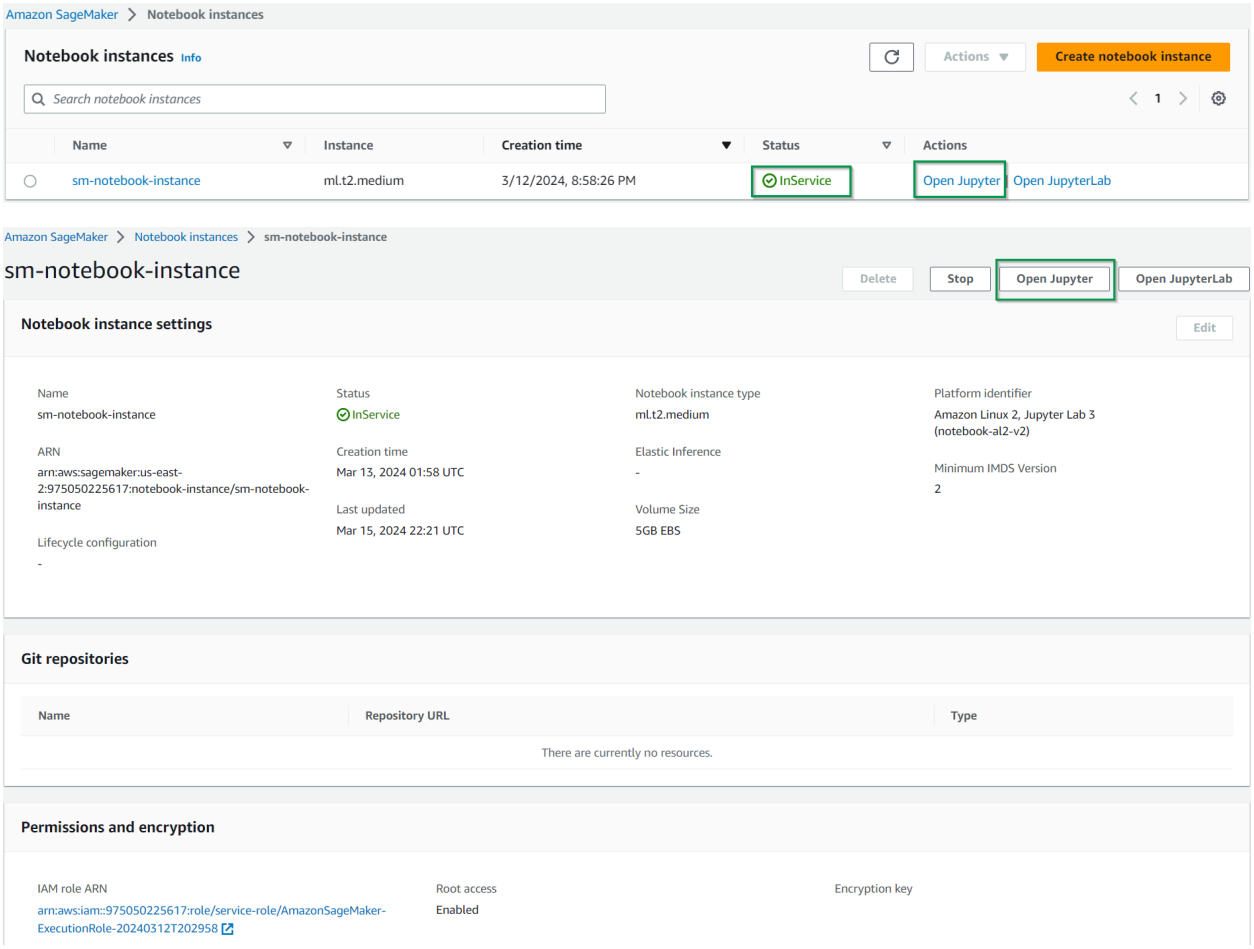
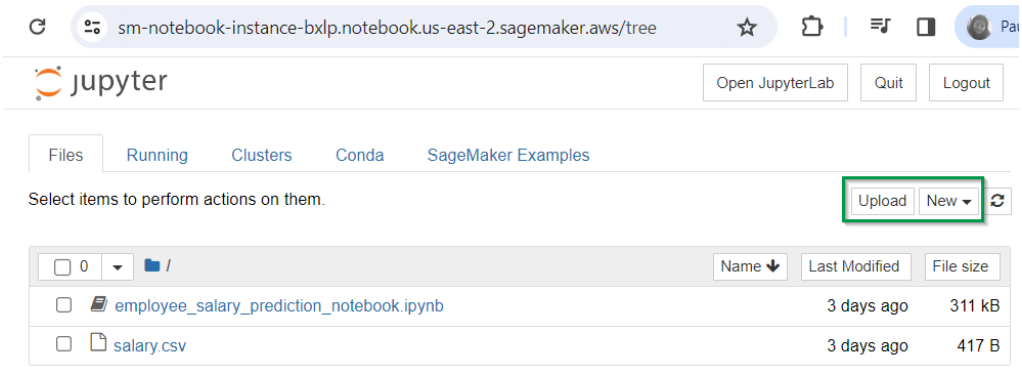
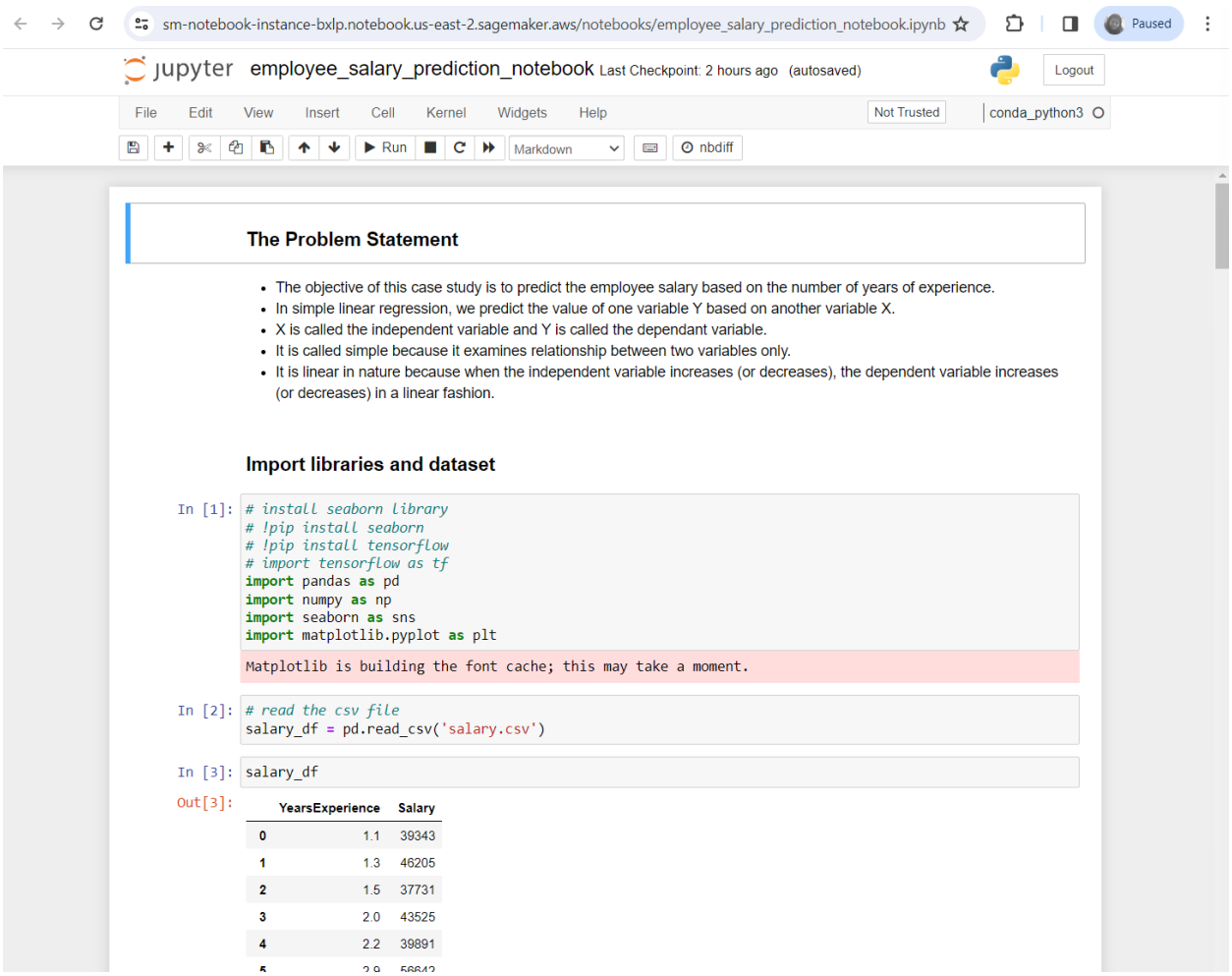
Utilizing AWS SageMaker for Advanced Data Analysis and Machine Learning Deployment

We start with setting up a Notebook Instance in Amazon SageMaker where Jupyter notebook can be opened to build, train and test machine learning models.

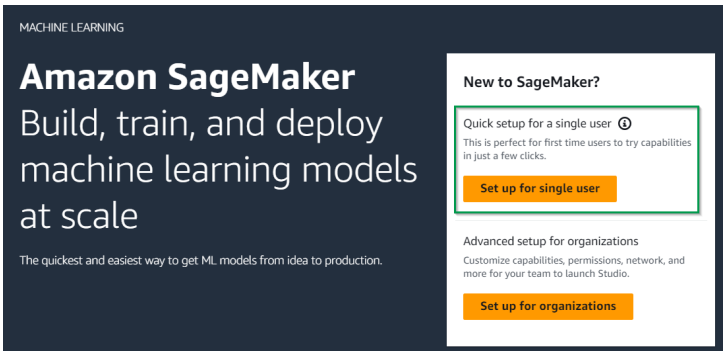


Here, data file and any pre-existing Jupyter notebook can be upload for further processing. It also has an option to create a new notebook if preferred.



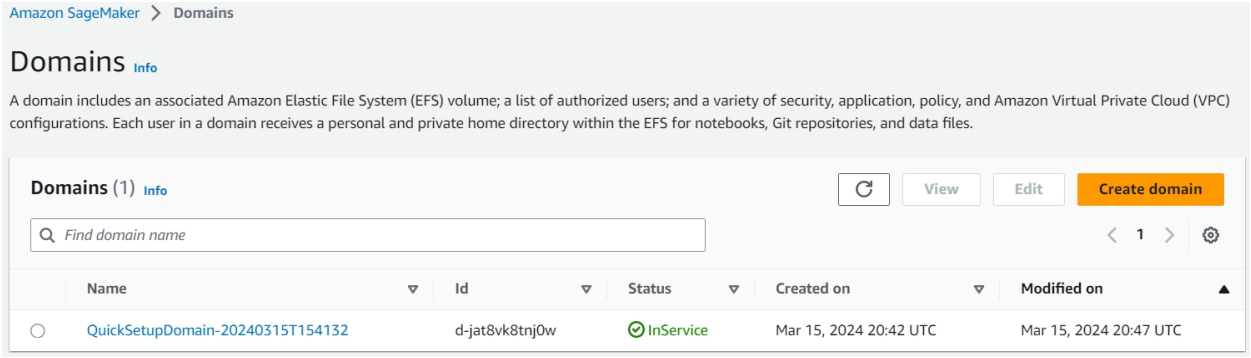


SageMaker Studio is set up with **Set up for single user** option which quickly onboards single users. It is also *recommended* by Amazon for first-time users.



Using the **Set up for single user** option, a **domain** and **user profile** are automatically created for us with default settings. Along with those, it also creates a Domain execution role, User profile execution role, Shared space execution role, SageMaker Canvas time series forecasting role, **Amazon S3 bucket** and selects a VPC.

Domain:



QuickSetupDomain-20240315T154132

Domain details

Configure and manage the domain.

User profiles | Space management | Environment | Domain settings

General settings Info

Edit

Name	Status	Domain ID
QuickSetupDomain-20240315T154132	Ready	d-jat8vk8tnj0w
Created	**Last modified**	**VPC**
Fri Mar 15 2024 15:42:02 GMT-0500 (Central Daylight Time)	Fri Mar 15 2024 15:47:41 GMT-0500 (Central Daylight Time)	vpc-0a8d9ed8d33096940
Authentication method	**Execution role**	
AWS Identity and Access Management (IAM)	arn:aws:iam::975050225617:role/service-role/AmazonSageMaker-ExecutionRole-20240315T154133	

Storage configurations

Edit

Encryption key

No custom encryption

Default space size	Maximum space size
5	100

Apps

Configure apps to manage your ML workflow using SageMaker.

Studio | Studio Lab | Canvas | RStudio

Studio

Amazon SageMaker Studio provides a single, web-based visual interface where you can perform all ML development steps, including building notebooks, experiment management, automatic model creation, debugging, and model and data drift detection.

Configure app | Learn more

Projects

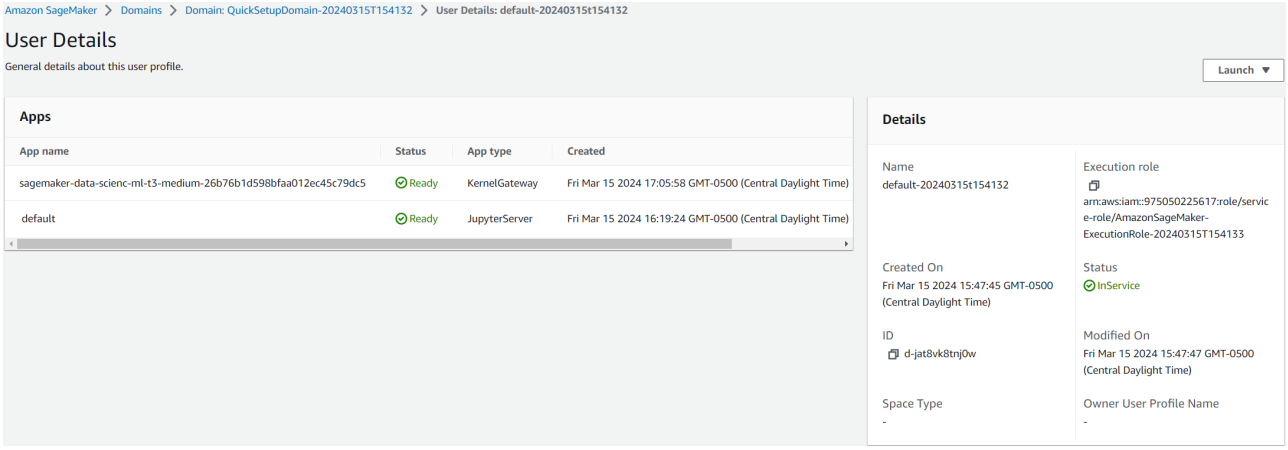
Amazon SageMaker project templates enabled for this account

- Launch constraint role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsLaunchRole
- Product use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsUseRole
- API Gateway use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsApiGatewayRole
- CloudFormation use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsCloudformationRole
- CodeBuild use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsCodeBuildRole
- CodePipeline use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsCodePipelineRole
- Events use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsEventsRole
- Firehose use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsFirehoseRole
- Glue use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsGlueRole
- Lambda use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsLambdaRole
- SageMaker use role: arn:aws:iam::975050225617:role/service-role/AmazonSageMakerServiceCatalogProductsExecutionRole

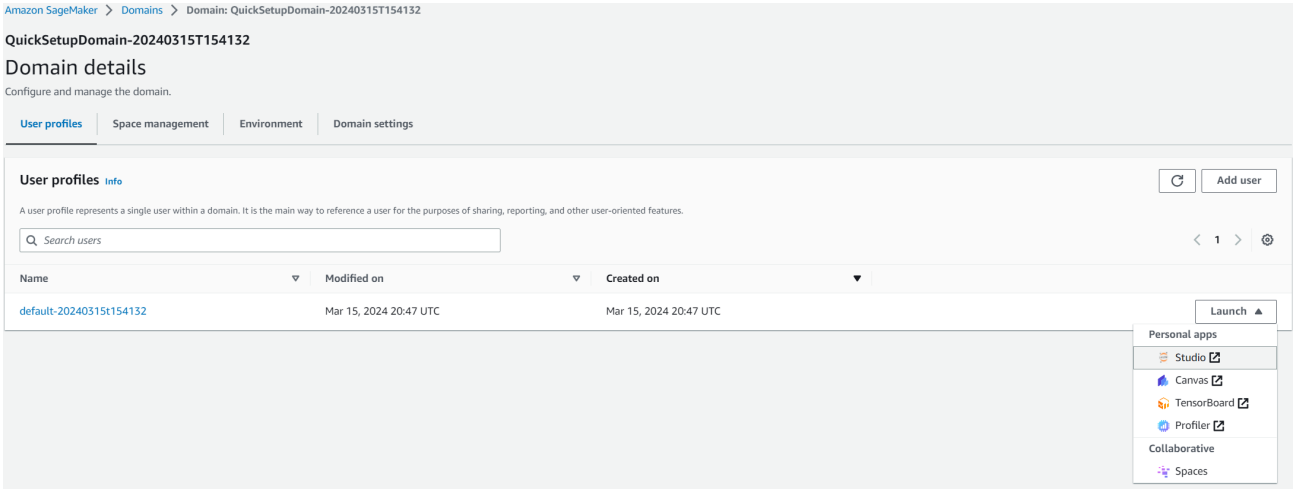
Amazon SageMaker project templates enabled for Studio users

- Execution role: arn:aws:iam::975050225617:role/service-role/AmazonSageMaker-ExecutionRole-20240315T154133

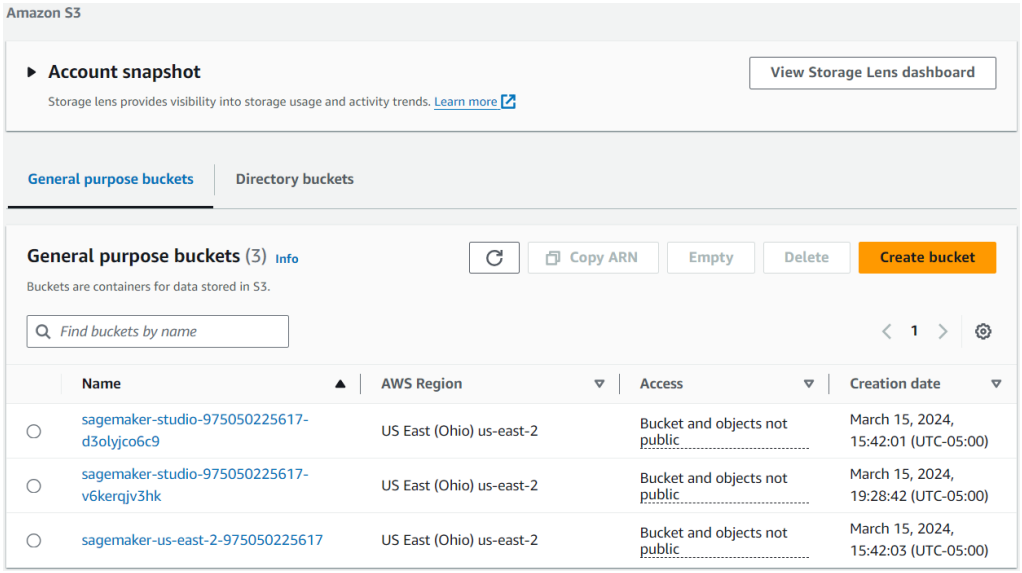
User Profile:



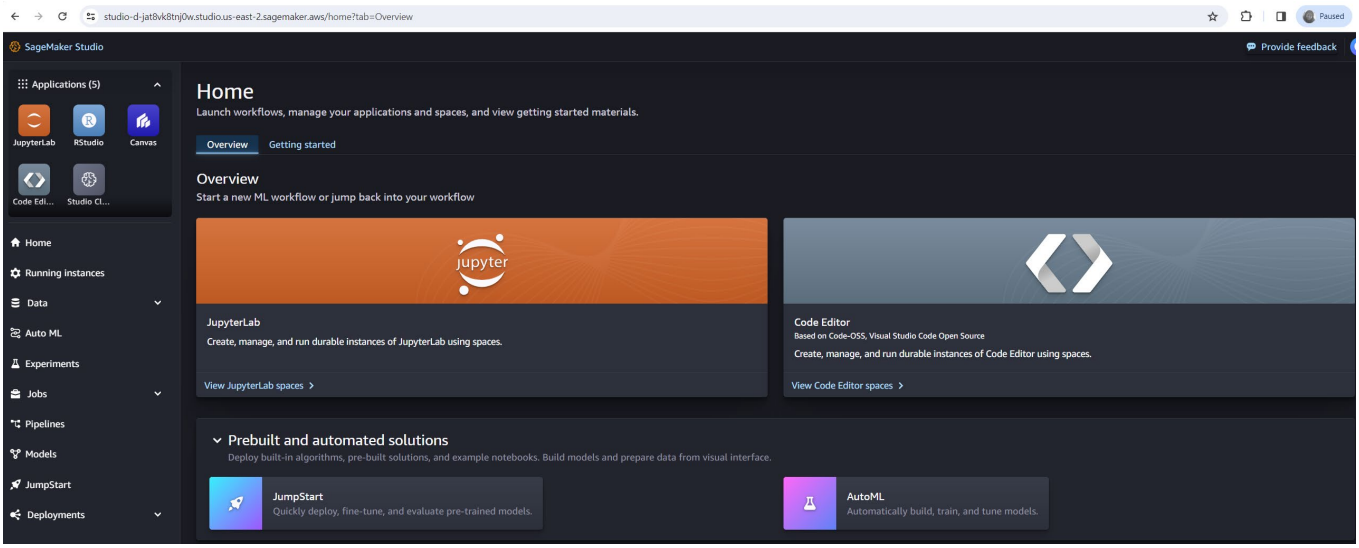
From user profile, Amazon SageMaker Studio can be launched.

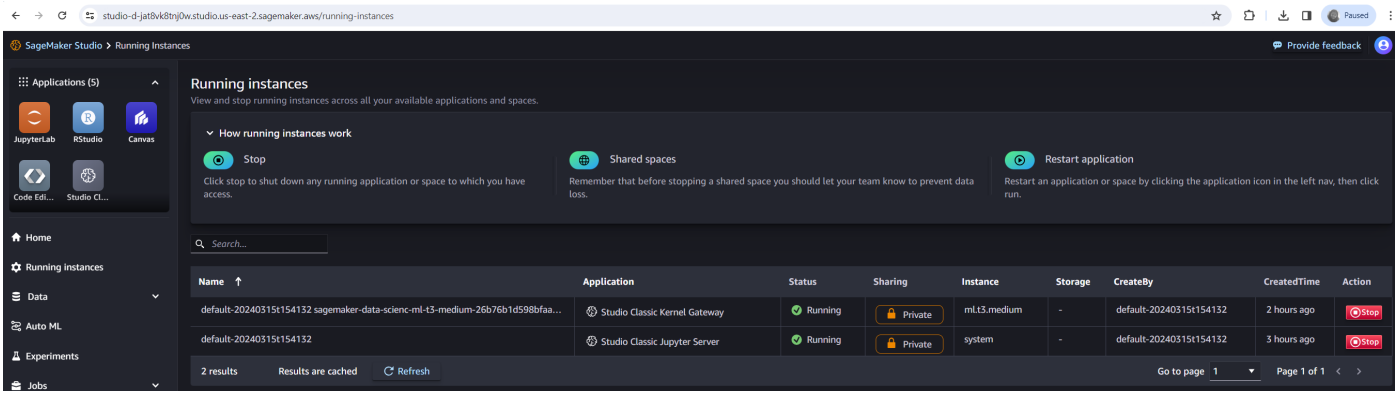


S3 Bucket:

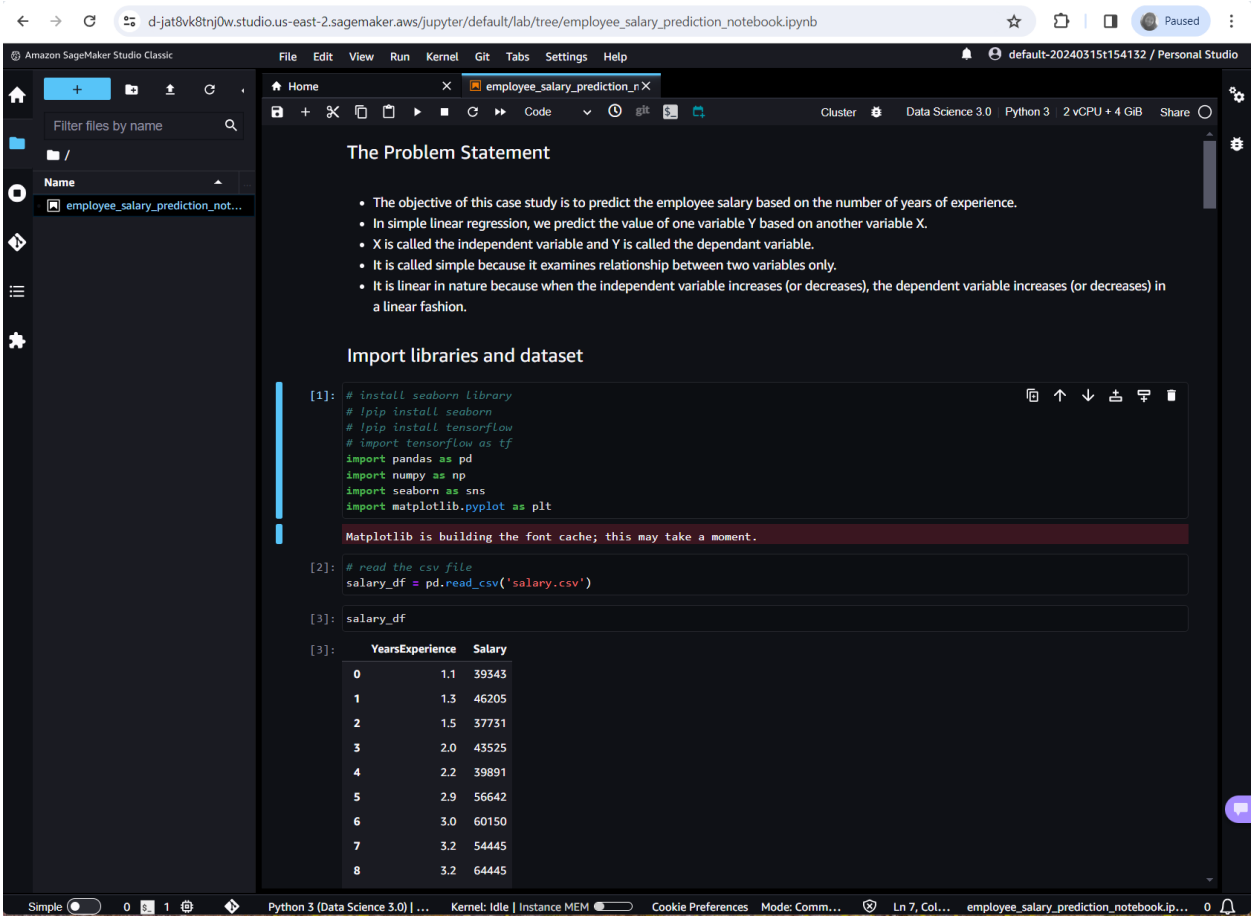


In Amazon SageMaker Studio, we have the option of various application suits based on need. For our purposes, we launch a JupyterLab application.

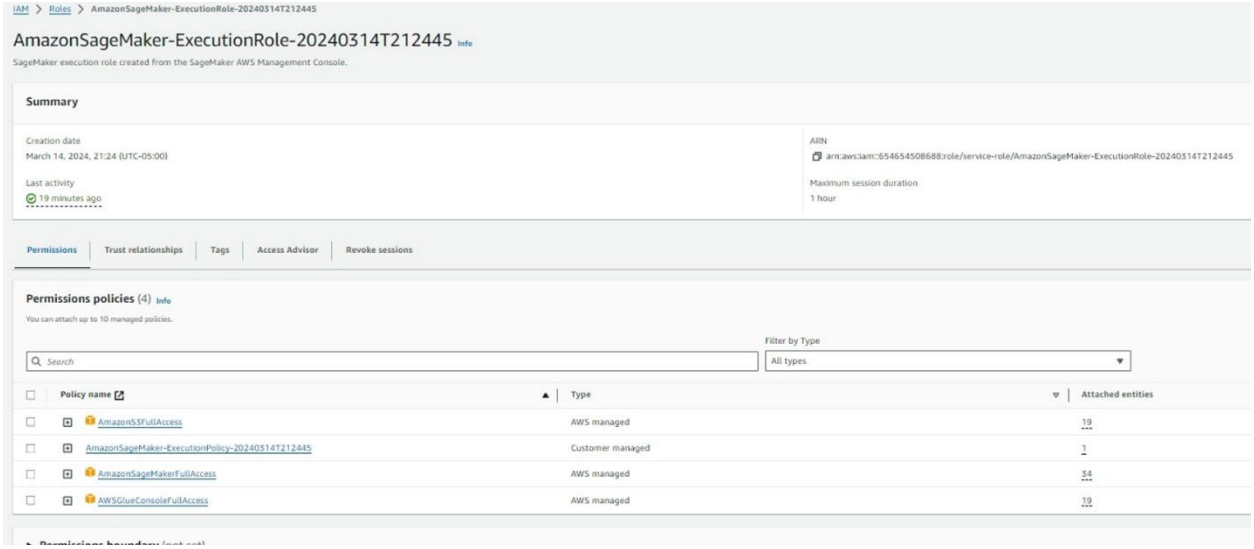




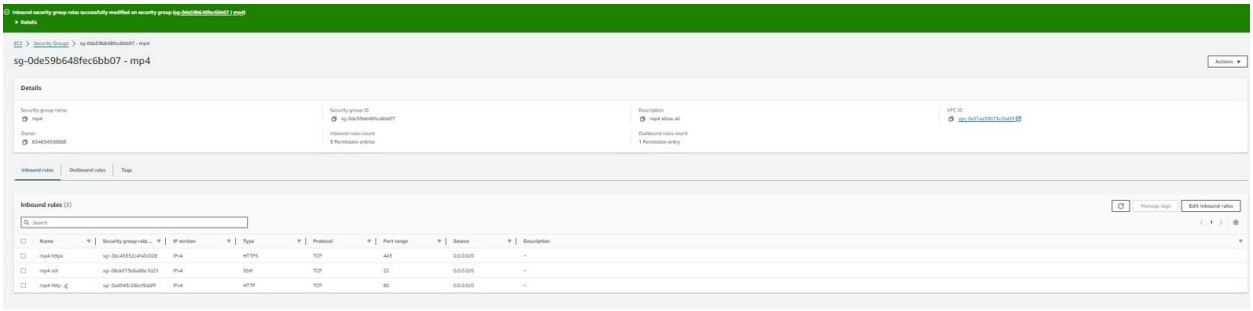
We have the option to either upload a notebook or create a new notebook to get started with machine learning tasks.



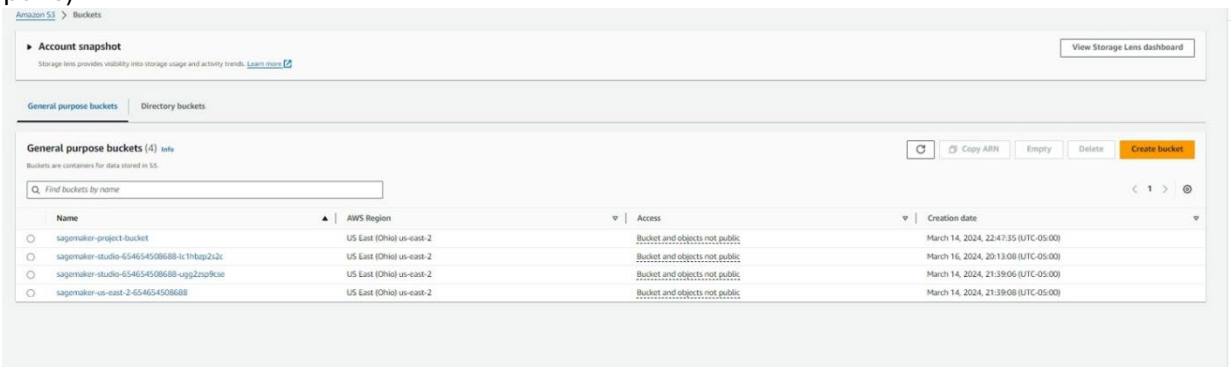
Created Execution Role Through IAM for Sagemaker Project:



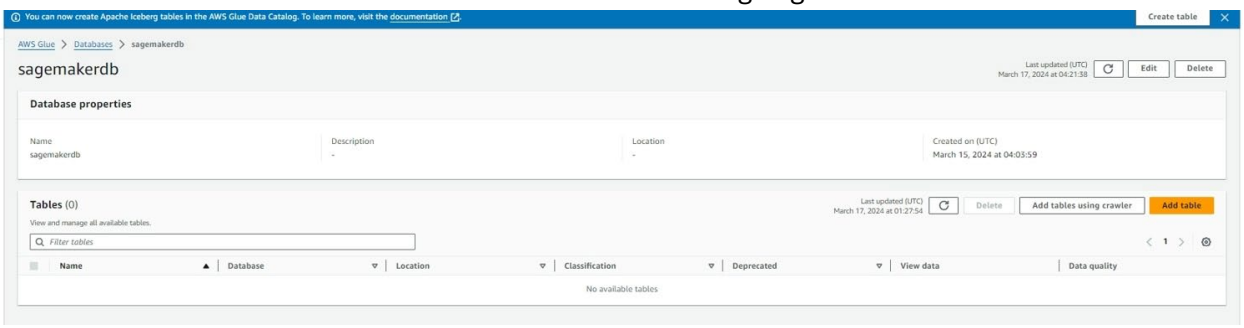
Security group for Sagemaker:



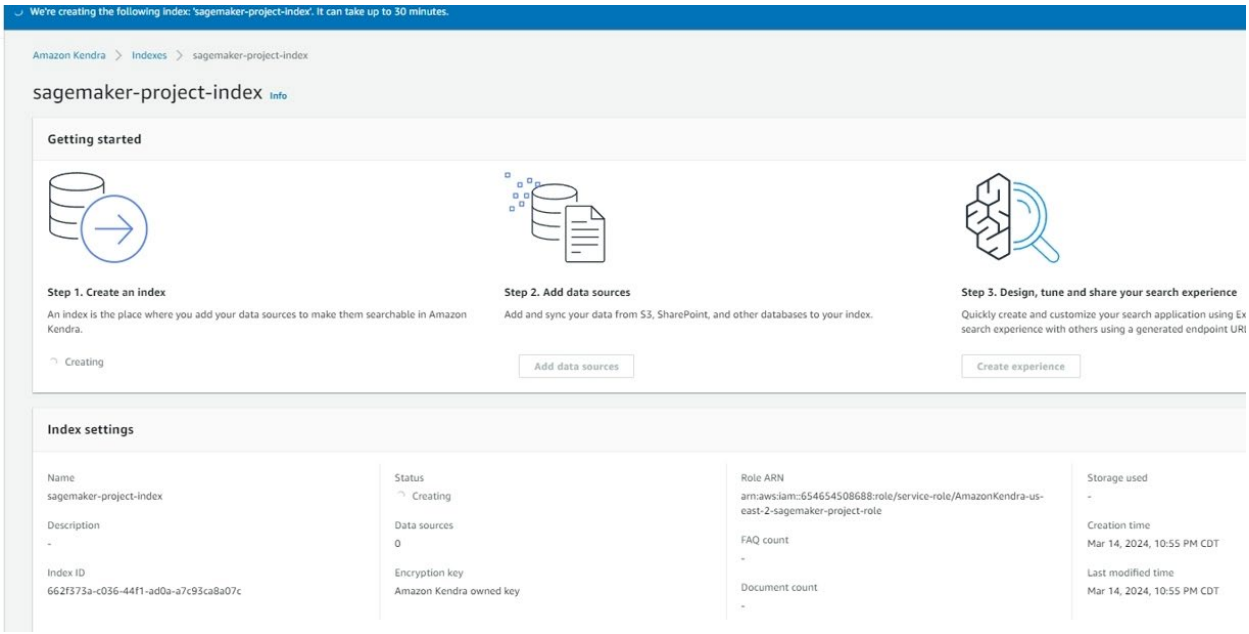
Created S3 Bucket to hold project icon and data (endpoints and model push/saves along with csv pulls):



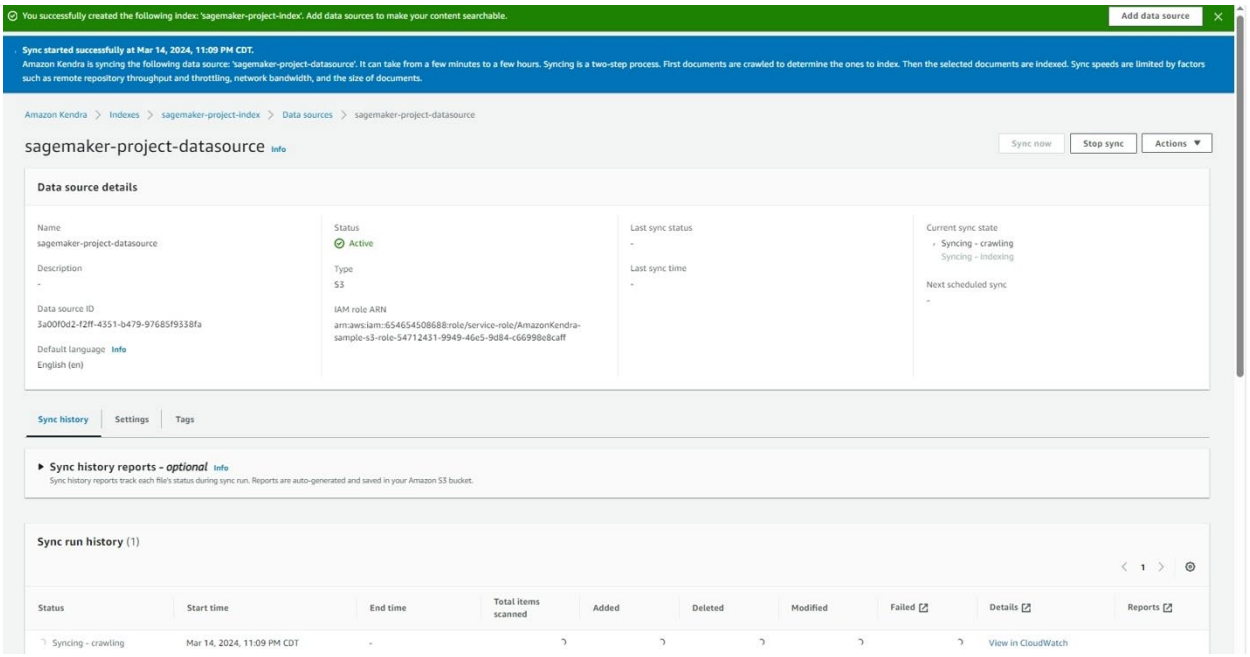
AWS Glue database created for data connection when using Sagemaker Studio Lab:



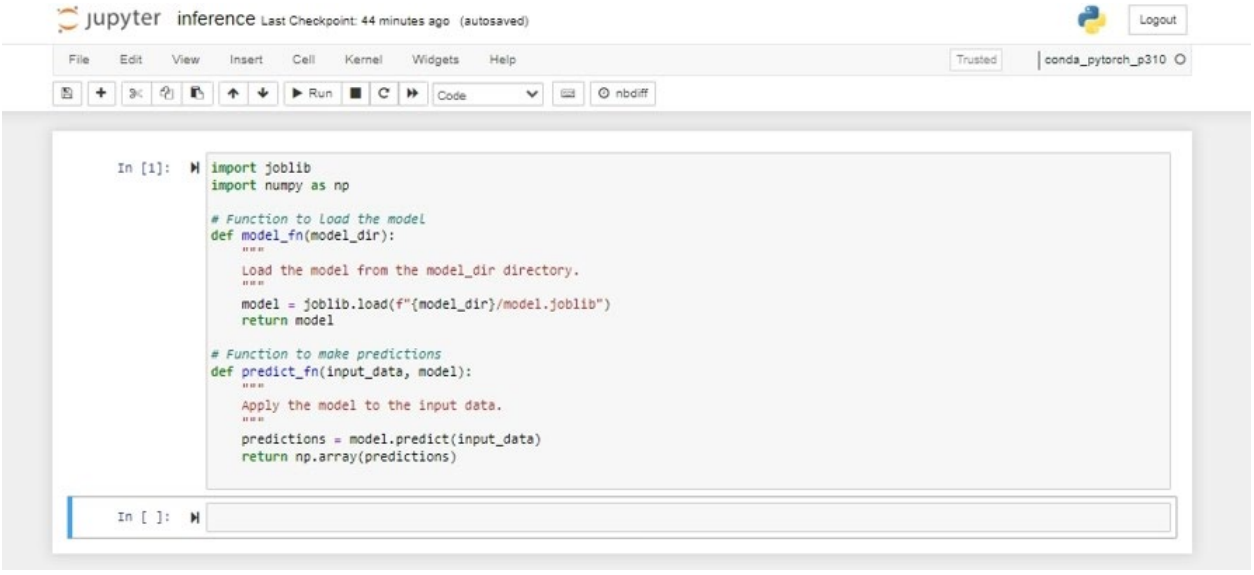
Kendra data source created and connected to sagemaker database for data connection.



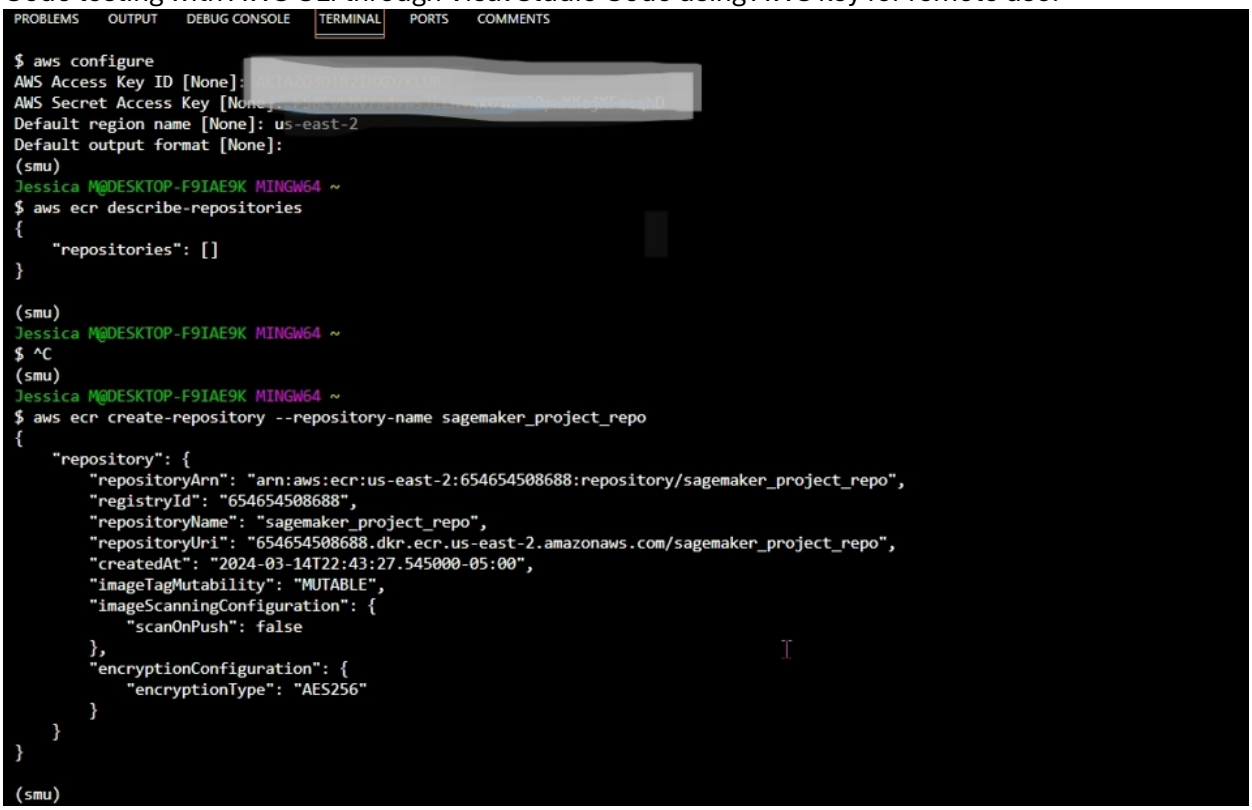
Sagemaker Role created and connected to project with IAM Sagemaker execution role privileges.



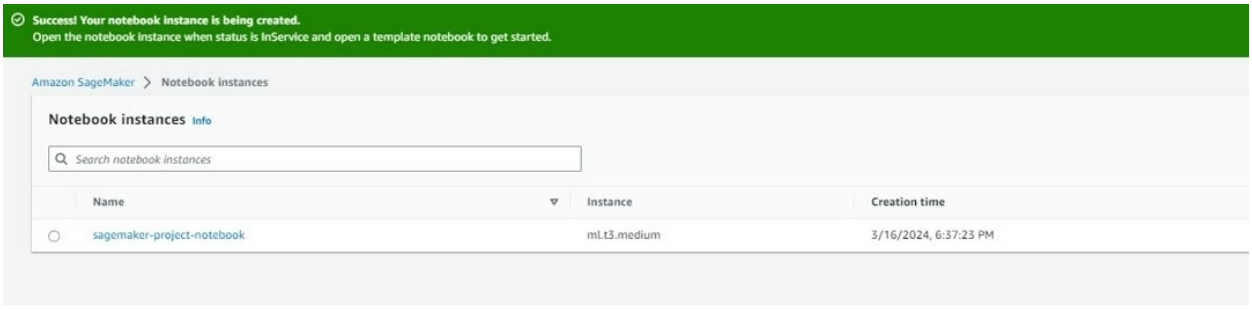
Inference.py created and saved into bucket to use with notebook and define endpoint(s):



Code testing with AWS CLI through Visal Studio Code using AWS key for remote use:

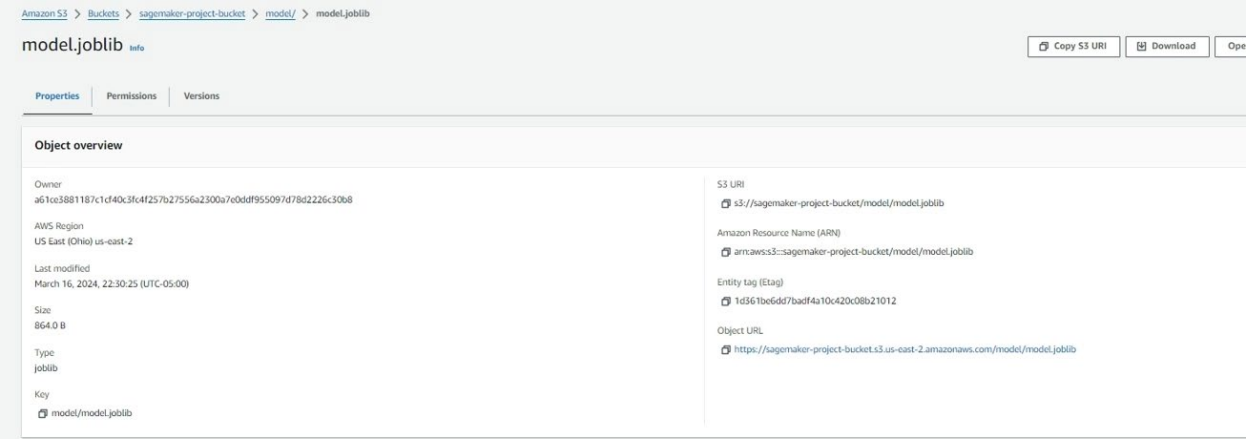


Notebook instance and files created/ saved into S3 bucket:

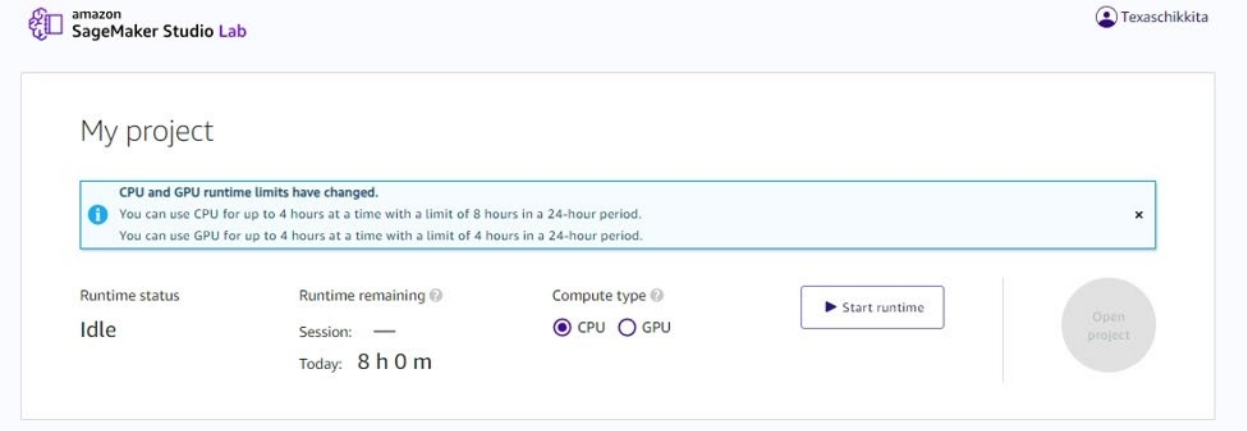


Bucket Files (viewed from notebook instance):

Model saved as model.joblib to use with Kendra and Sagemaker DB, and S3 bucket:



Sagemaker Studio Lab provision completed. Lab is now ready for use with model and resources created above:



Snapshot of instance created in sagemaker portal through aws- EDA, and saving model to bucket.