aws

Amazon SageMaker Studio

ハンズオン

Amazon Web Services
2020/04

ハンズオン事前準備と環境の確認

- ハンズオンの事前準備
  - AWSアカウントの作成
- 実行環境の前提
  - Web Browser: Chrome or Firefox を利用
  - リージョン: Ohioリージョン (us-east-2)

ハンズオンのリソース

今日のコンテンツは全て以下のYoutubeで公開されています。

Amazon SageMaker Studio - AutoML with Amazon SageMaker AutoPilot (part 1) - YouTube

- https://www.youtube.com/watch?v=qMEtqJPhqpA

Amazon SageMaker Studio - AutoML with Amazon SageMaker AutoPilot (part 2) - YouTube

- https://www.youtube.com/watch?v=WsfRAeGzgm8

Amazon SageMaker Studio - AutoML with Amazon SageMaker AutoPilot (part 3) - YouTube

- https://www.youtube.com/watch?v=KZSTsWrDGXs

Amazon SageMaker Studio - AutoML with Amazon SageMaker AutoPilot (part 4) - YouTube

- https://www.youtube.com/watch?v=vRHyX3kDstI

Amazon SageMaker Studio - Deploying a model to an endpoint - YouTube

- https://www.youtube.com/watch?v=xf46PdAjbY8

dataset.ipynb · master · Julien Simon / amazon-studio-demos · GitLab

- https://gitlab.com/juliensimon/amazon-studio-demos/blob/master/dataset.ipynb

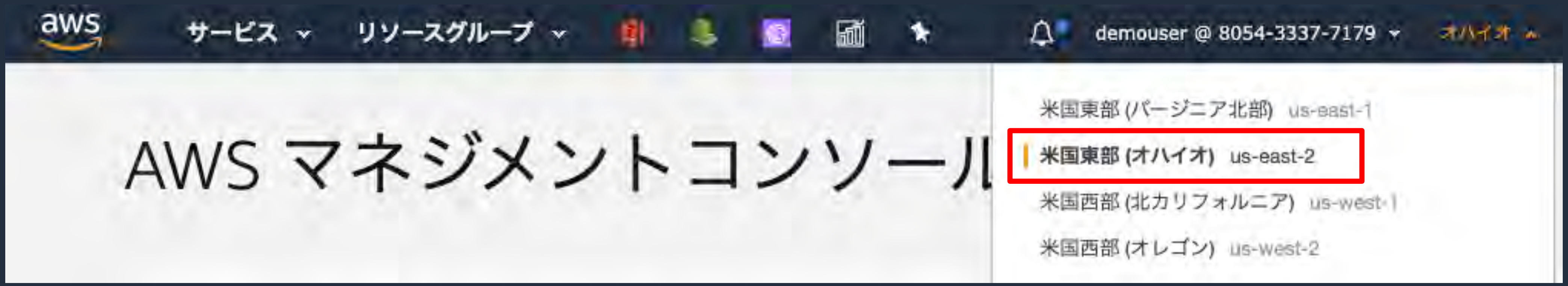
End to end demo with AutoML on Amazon SageMaker - YouTube

- https://www.youtube.com/watch?v=DRjOOaR2prQ

ハンズオンの流れ

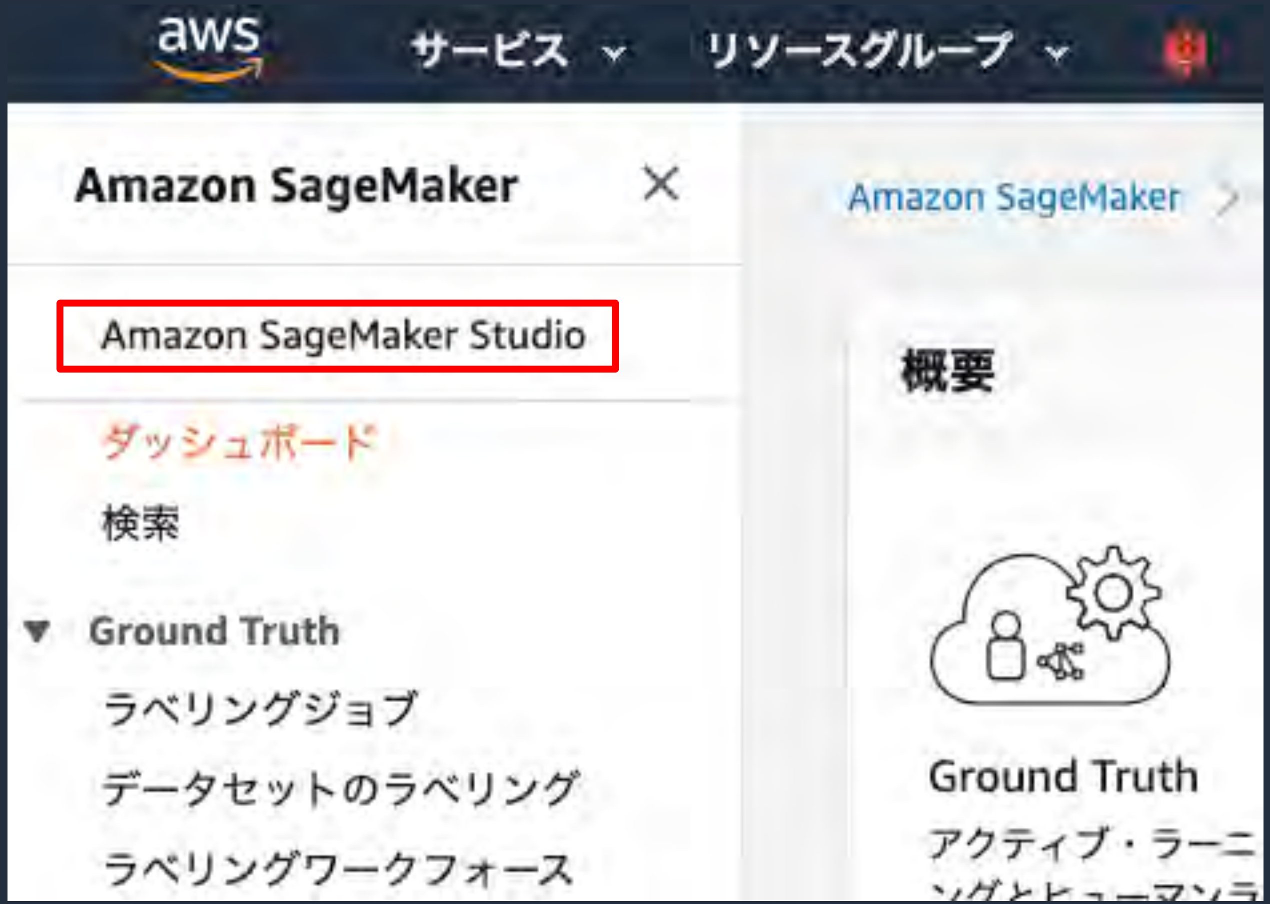
- SageMaker Studioドメイン作成（SSO方式）
- SSOユーザー登録
- SageMaker Studioの起動
- Datasetの準備
- Experiment(Autopilot)の実行
- Training結果の確認・評価
- モデルのDeploy
- Endpointへのリクエスト
- 後片付け

リージョン選択 & SageMakerコンソールへ移動

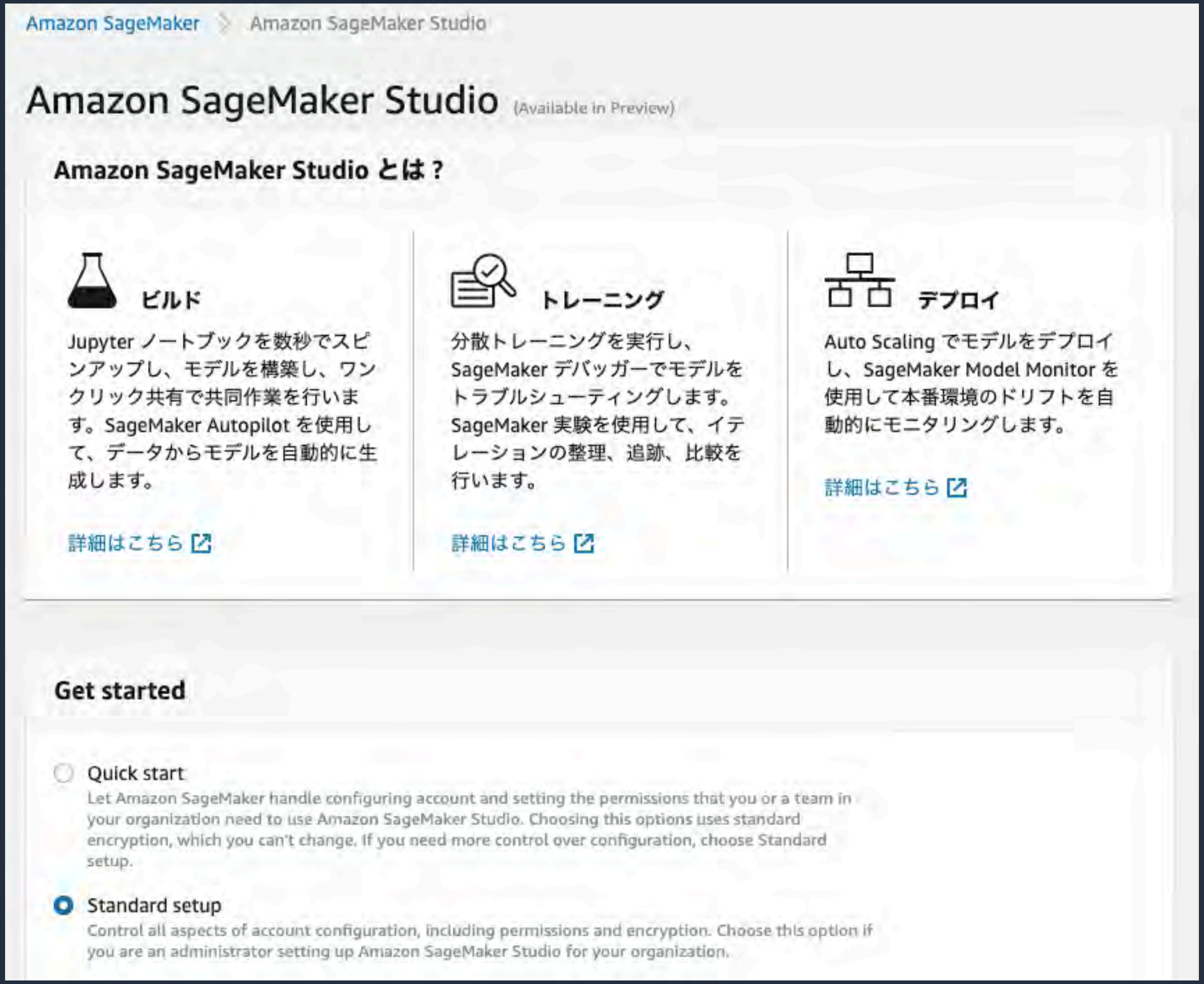


オハイオ us-east-2
を選択

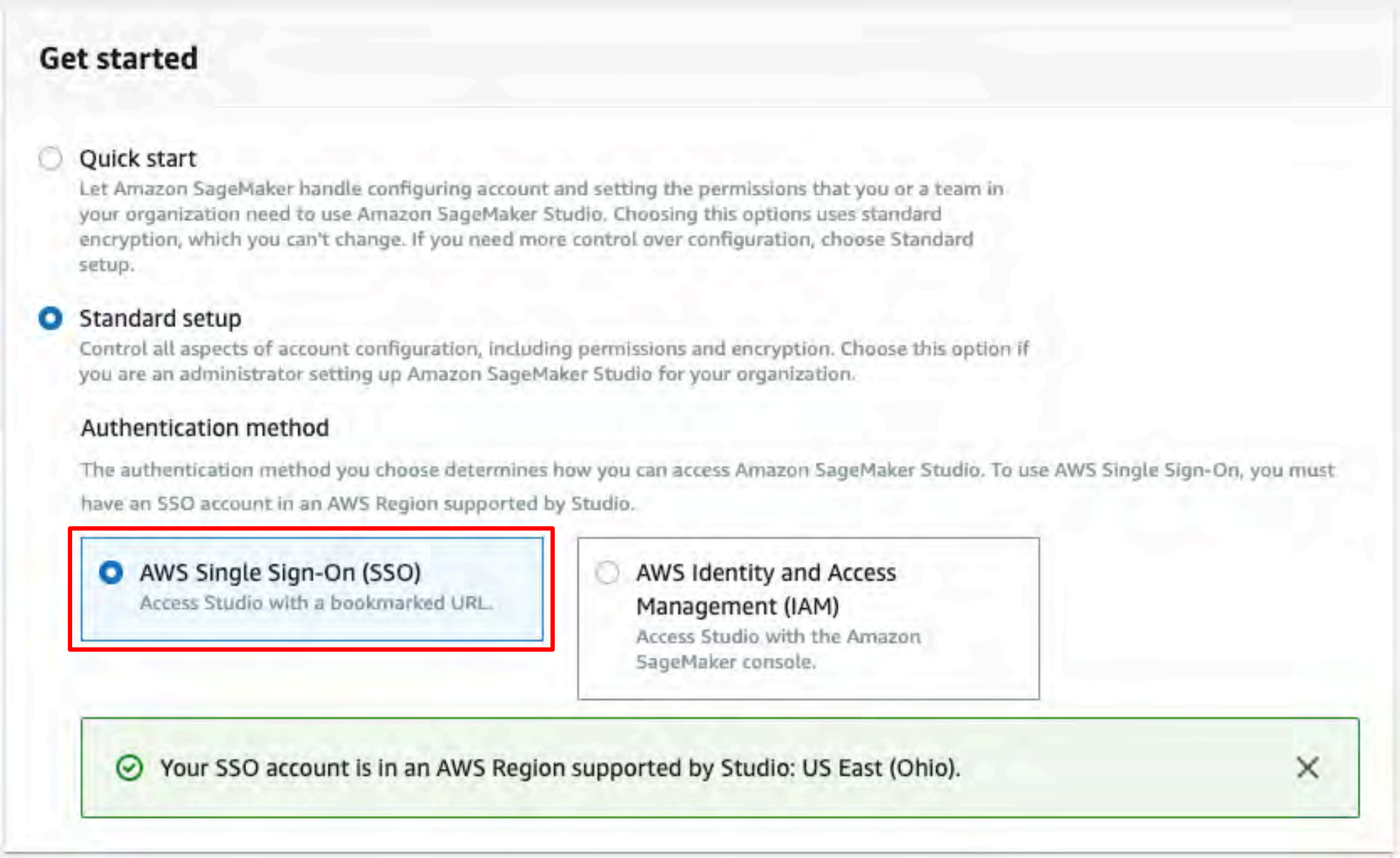
SageMaker Studioコンソールに移動



コンソール画面確認


Amazon SageMaker
Amazon SageMaker Studio
Amazon SageMaker Studio (Available in Preview)
Amazon SageMaker Studio とは？
ビルド
Jupyter ノートブックを数秒でスピンアップし、モデルを構築し、ワンクリック共有で共同作業を行います。SageMaker Autopilot を使用して、データからモデルを自動的に生成します。
詳細はこちら
トレーニング
分散トレーニングを実行し、SageMaker デバッガーでモデルをトラブルシューティングします。SageMaker 実験を使用して、イテレーションの整理、追跡、比較を行います。
詳細はこちら
デプロイ
Auto Scaling でモデルをデプロイし、SageMaker Model Monitor を使用して本番環境のドリフトを自動的にモニタリングします。
詳細はこちら
Get started
Quick start
Let Amazon SageMaker handle configuring account and setting the permissions that you or a team in your organization need to use Amazon SageMaker Studio. Choosing this options uses standard encryption, which you can't change. If you need more control over configuration, choose Standard setup.
Standard setup
Control all aspects of account configuration, including permissions and encryption. Choose this option if you are an administrator setting up Amazon SageMaker Studio for your organization.

SSO方式を選択



Get started
Quick start
Let Amazon SageMaker handle configuring account and setting the permissions that you or a team in your organization need to use Amazon SageMaker Studio. Choosing this options uses standard encryption, which you can't change. If you need more control over configuration, choose Standard setup.
Standard setup
Control all aspects of account configuration, including permissions and encryption. Choose this option if you are an administrator setting up Amazon SageMaker Studio for your organization.
Authentication method
The authentication method you choose determines how you can access Amazon SageMaker Studio. To use AWS Single Sign-On, you must have an SSO account in an AWS Region supported by Studio.
AWS Single Sign-On (SSO)
Access Studio with a bookmarked URL.
AWS Identity and Access Management (IAM)
Access Studio with the Amazon SageMaker console.
Your SSO account is in an AWS Region supported by Studio: US East (Ohio).

IAM ロール作成

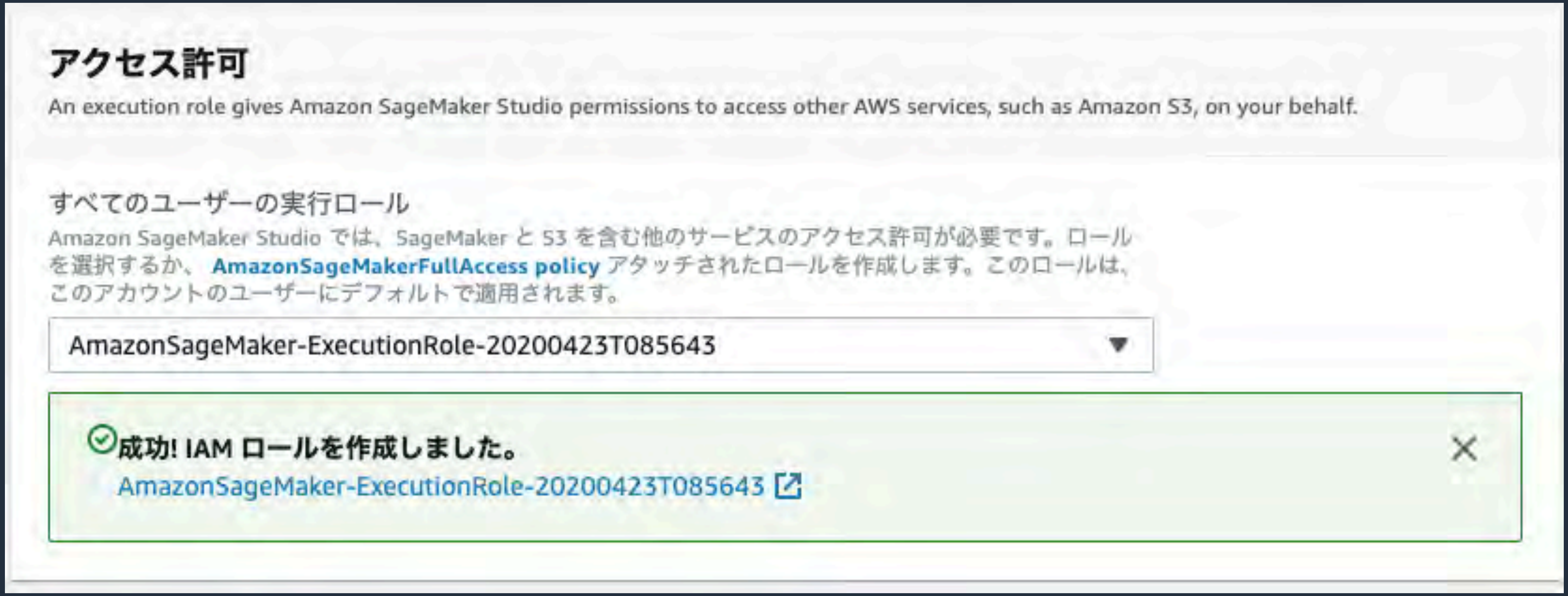

アクセス許可
An execution role gives Amazon SageMaker Studio permissions to access other AWS services, such as Amazon S3, on your behalf.
すべてのユーザーの実行ロール
Amazon SageMaker Studio では、SageMaker と S3 を含む他のサービスのアクセス許可が必要です。ロールを選択するか、 AmazonSageMakerFullAccess policy アタッチされたロールを作成します。このロールは、このアカウントのユーザーにデフォルトで適用されます。
IAM ロールを選択する


新しいロールの作成
カスタム IAM ロールの ARN の入力
既存のロールの使用

IAM ロール作成

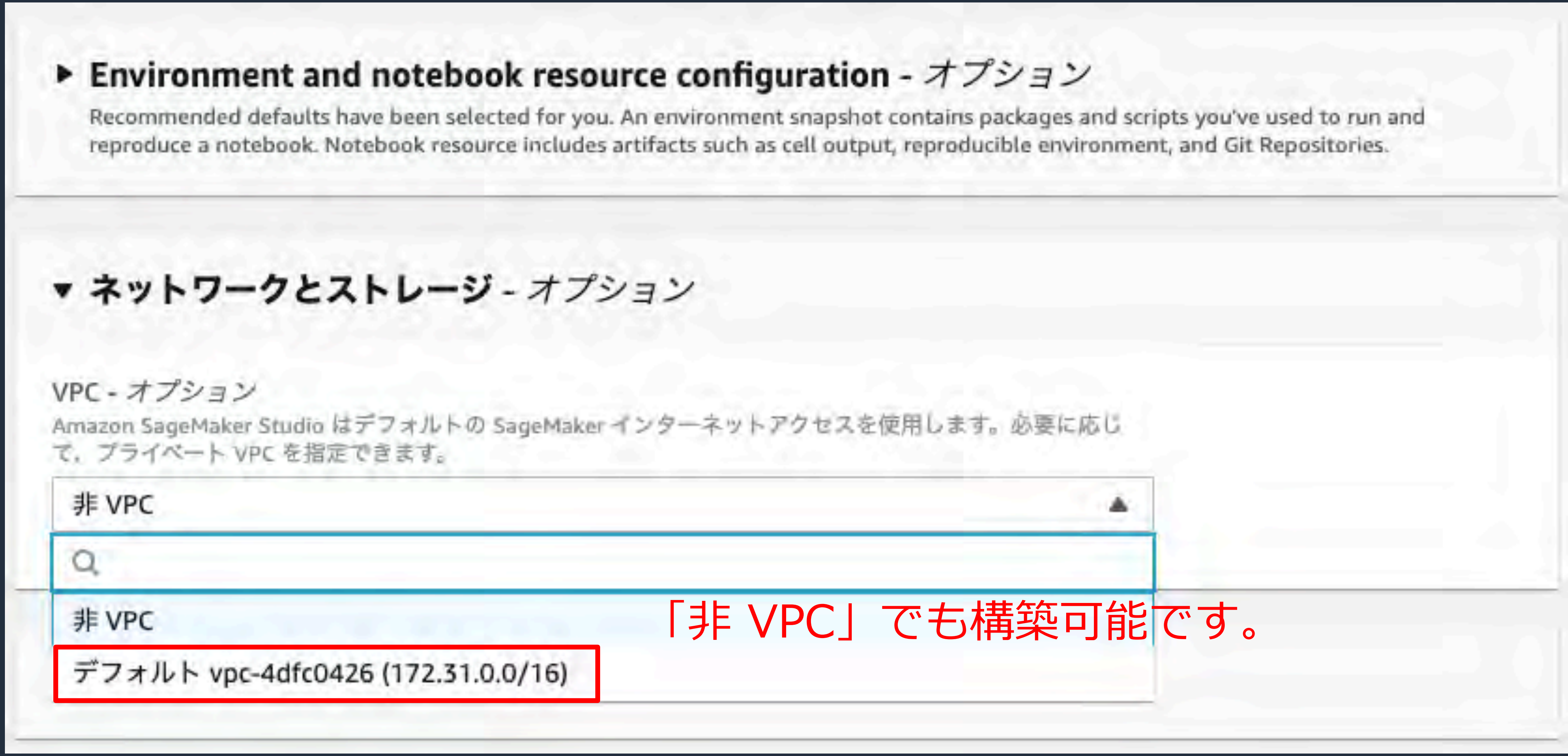

IAM ロールを作成する
IAM ロールを渡すと、ユーザーに代わって他の AWS のサービスでアクションを実行するアクセス許可が Amazon SageMaker に与えられます。ここでロールを作成すると、 AmazonSageMakerFullAccess 作成する IAM ポリシーで記述されたアクセス許可が付与されます。
作成する IAM ロールにより、以下へのアクセスが提供されます。
指定する S3 バケット - オプション
任意の S3 バケット
自分のノートブックインスタンスにアクセスできるユーザーが、アカウント内の任意のバケットとそのコンテンツにアクセスすることを許可します。
特定の S3 バケット
例: bucket-name-1、bucket-name-2、
カンマ区切り。ARN、「*」、および「/」はサポートされません。
なし
名前に「sagemaker」が含まれる任意の S3 バケット
名前に「sagemaker」が含まれる任意の S3 オブジェクト
タグ「sagemaker」と値「true」が含まれる任意の S3 オブジェクト
オブジェクトのタグ付けの表示
SageMaker へのアクセスを許可するバケットポリシーを持つ S3 バケット
S3 バケットポリシーの表示
キャンセル
ロールの作成

IAM ロール作成確認

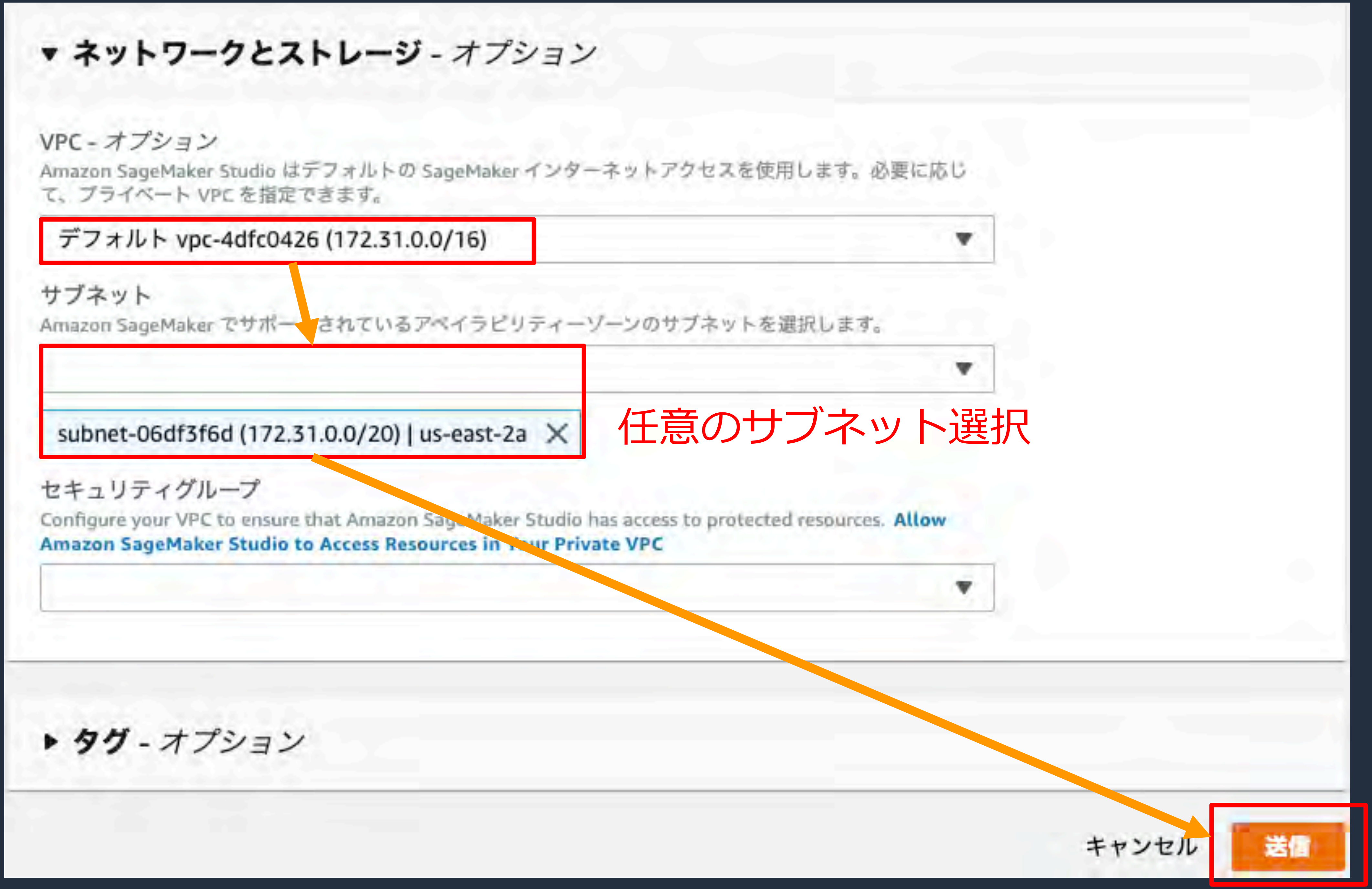


アクセス許可
An execution role gives Amazon SageMaker Studio permissions to access other AWS services, such as Amazon S3, on your behalf.
すべてのユーザーの実行ロール
Amazon SageMaker Studio では、SageMaker と S3 を含む他のサービスのアクセス許可が必要です。ロールを選択するか、 AmazonSageMakerFullAccess policy アタッチされたロールを作成します。このロールは、このアカウントのユーザーにデフォルトで適用されます。
AmazonSageMaker-ExecutionRole-20200423T085643
成功! IAM ロールを作成しました。
AmazonSageMaker-ExecutionRole-20200423T085643

VPC 選択

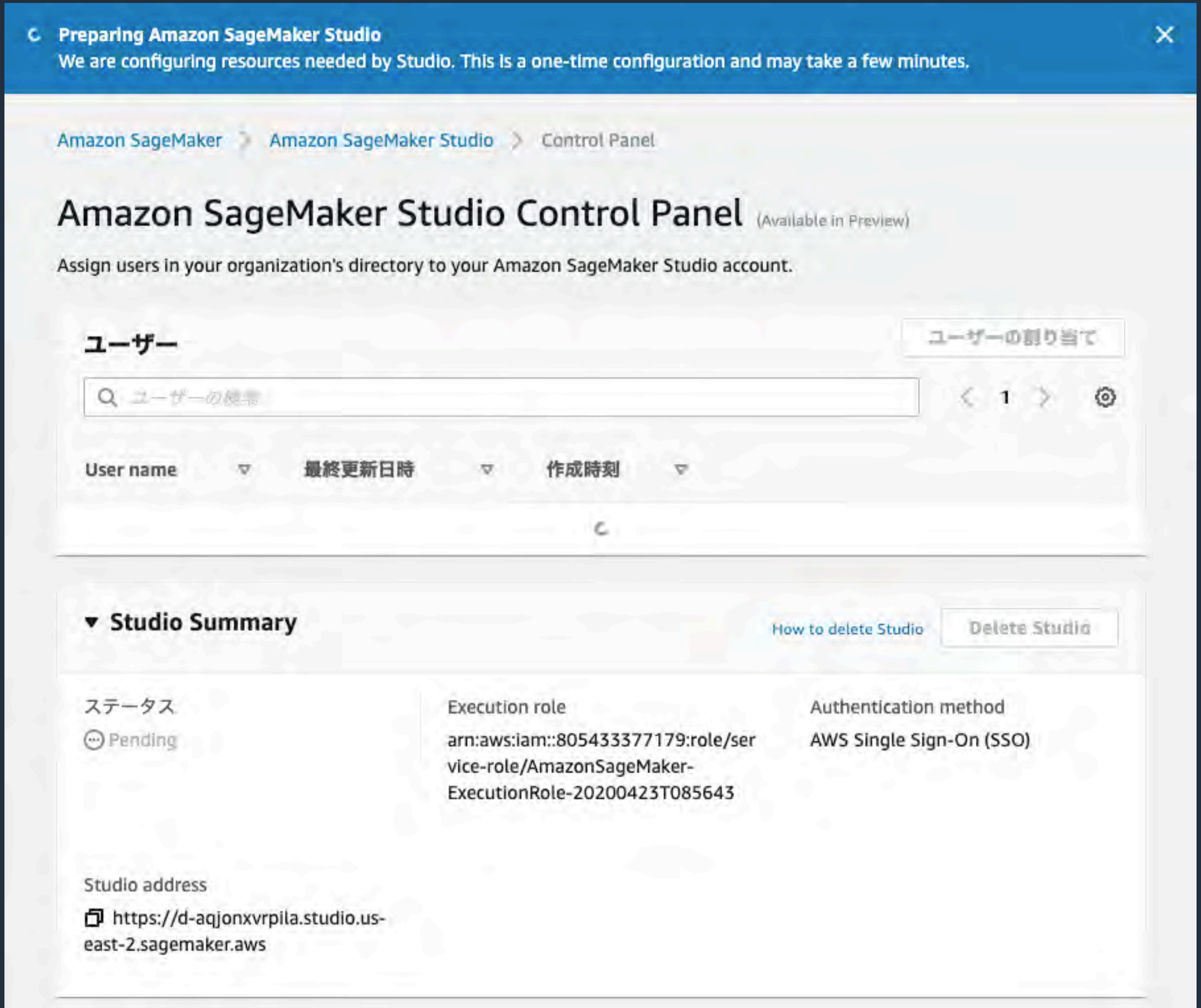

▶ Environment and notebook resource configuration - オプション
Recommended defaults have been selected for you. An environment snapshot contains packages and scripts you've used to run and reproduce a notebook. Notebook resource includes artifacts such as cell output, reproducible environment, and Git Repositories.
▼ ネットワークとストレージ - オプション
VPC - オプション
Amazon SageMaker Studio はデフォルトの SageMaker インターネットアクセスを使用します。必要に応じて、プライベート VPC を指定できます。
非 VPC
非 VPC
デフォルト vpc-4dfc0426 (172.31.0.0/16)
「非 VPC」でも構築可能です。

サブネット選択

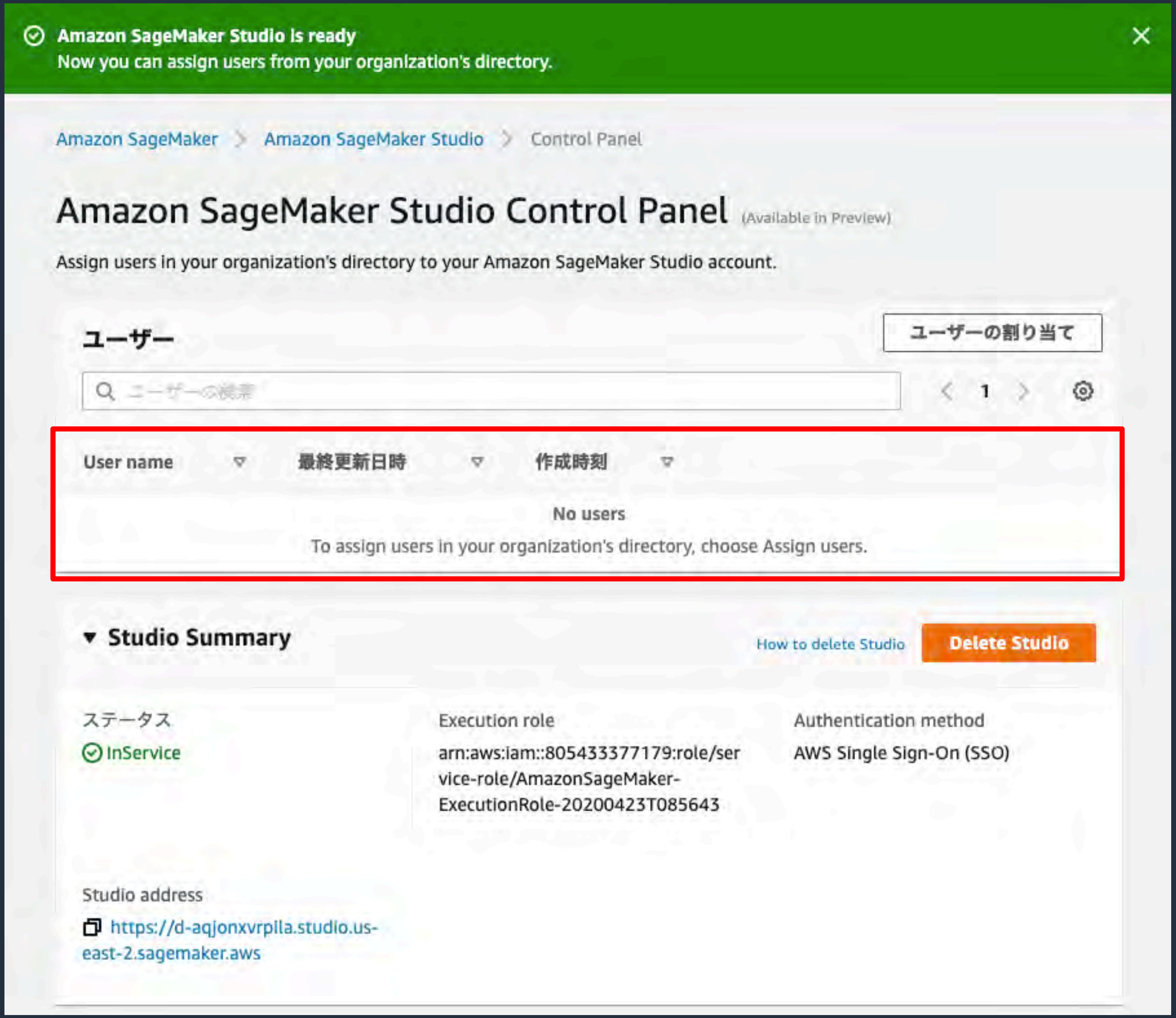


▼ ネットワークとストレージ - オプション
VPC - オプション
Amazon SageMaker Studio はデフォルトの SageMaker インターネットアクセスを使用します。必要に応じて、プライベート VPC を指定できます。
デフォルト vpc-4dfc0426 (172.31.0.0/16)
サブネット
Amazon SageMaker でサポートされているアベイラビリティーゾーンのサブネットを選択します。
subnet-06df3f6d (172.31.0.0/20) | us-east-2a
任意のサブネット選択
セキュリティグループ
Configure your VPC to ensure that Amazon SageMaker Studio has access to protected resources. Allow Amazon SageMaker Studio to Access Resources in Your Private VPC
▶ タグ - オプション
キャンセル
送信

ドメイン作成中


Preparing Amazon SageMaker Studio
We are configuring resources needed by Studio. This is a one-time configuration and may take a few minutes.
Amazon SageMaker
Amazon SageMaker Studio
Control Panel
Amazon SageMaker Studio Control Panel (Available in Preview)
Assign users in your organization's directory to your Amazon SageMaker Studio account.
ユーザー
ユーザーの割り当て
ユーザーの検索
1
User name
最終更新日時
作成時刻
Studio Summary
How to delete Studio
Delete Studio
ステータス
Pending
Execution role
arn:aws:iam::805433377179:role/service-role/AmazonSageMaker-ExecutionRole-20200423T085643
Authentication method
AWS Single Sign-On (SSO)
Studio address
https://d-aqjonxvrpila.studio.us-east-2.sagemaker.aws

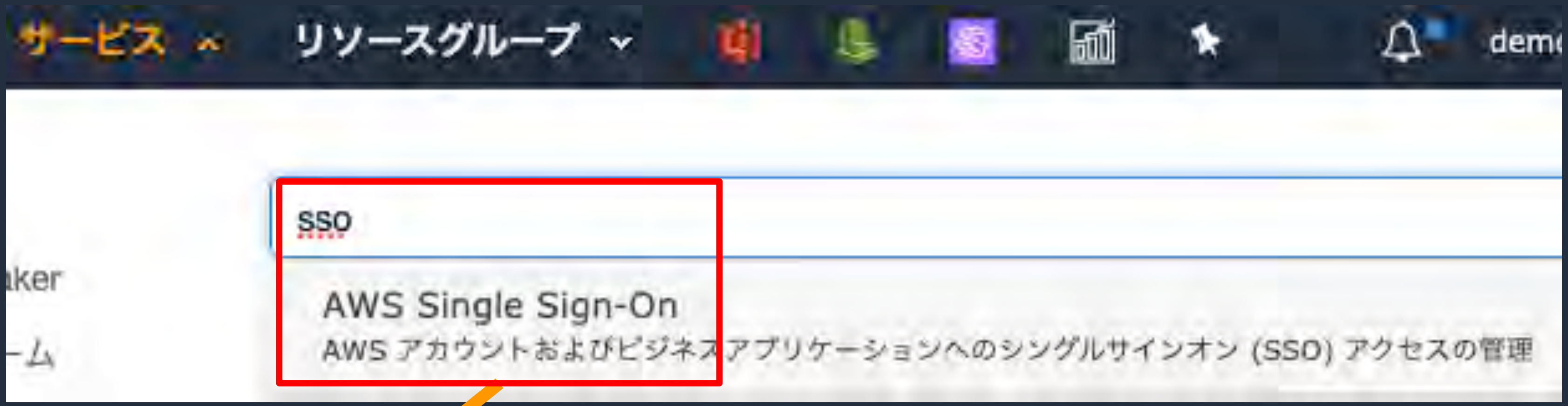
ドメイン作成完了



SSOユーザーを追加
していく

SSOユーザー登録

AWS Single Sign—On 画面に遷移



ユーザー詳細の入力


ユーザーの追加
1
詳細
2
グループ
ユーザー詳細
ユーザー名*
demouser
このユーザー名は、ユーザーポータルにサインインするために必要になります。後で変更することはできません。
パスワード
パスワードの設定手順が記載された E メールをユーザーに送信します。詳細はこちら
ユーザーと共有することができる一時的なパスワードを生成します。詳細はこちら
E メールアドレス*
yshiy@amazon.co.jp
E メールアドレスの確認*
yshiy@amazon.co.jp
名*
Yoshiyuki
姓*
Ito
表示名*
Yoshiyuki Ito
連絡方法 (オプション)
職務関連情報 (オプション)
住所 (オプション)
キャンセル
次: グループ

グループ設定

ユーザーの追加
1
2
詳細
グループ
ユーザーをグループに追加
グループに追加するユーザーは、グループに割り当てられた AWS アカウント、ロール、アプリケーションへのアクセスを継承します。
グループの作成
グループ名で検索
名前
説明
項目が見つかりません
そのまま次へ
キャンセル
戻る
ユーザーの追加

ユーザー作成完了確認

ダッシュボード
AWS アカウント
アプリケーション
ユーザー
グループ
設定
AWS SSO > ユーザー
ユーザー
ここにリストされているユーザーは、ユーザーポータルにサインインして、割り当てられた AWS アカウントまたはアプリケーションにアクセスできます。詳細はこちら
ユーザー「Yoshiyuki Ito」は正常に追加されました。
このユーザーには、パスワードを設定するリンクと、ユーザーポータルに接続する手順を示した E メールが届きます。このリンクは最長 7 日間まで有効です。これで、このユーザーにアカウントまたはアプリケーションへのアクセス許可を付与できるようになりました。これにより、ユーザーは、ユーザーポータルにサインインすると、割り当てられたアカウントとアプリケーションにアクセスできます。詳細はこちら
ユーザーを追加
ユーザーを削除
表示名
検索条件
表示名
ユーザー名
ステータス
Yoshiyuki Ito
demouser
有効

【ユーザー側】EメールにてAccept

件名　Invitation to join AWS Single Sign-On

From　no-reply@login.awsapps.com <no-reply@login.awsapps.com>

Hello Yoshiyuki Ito,

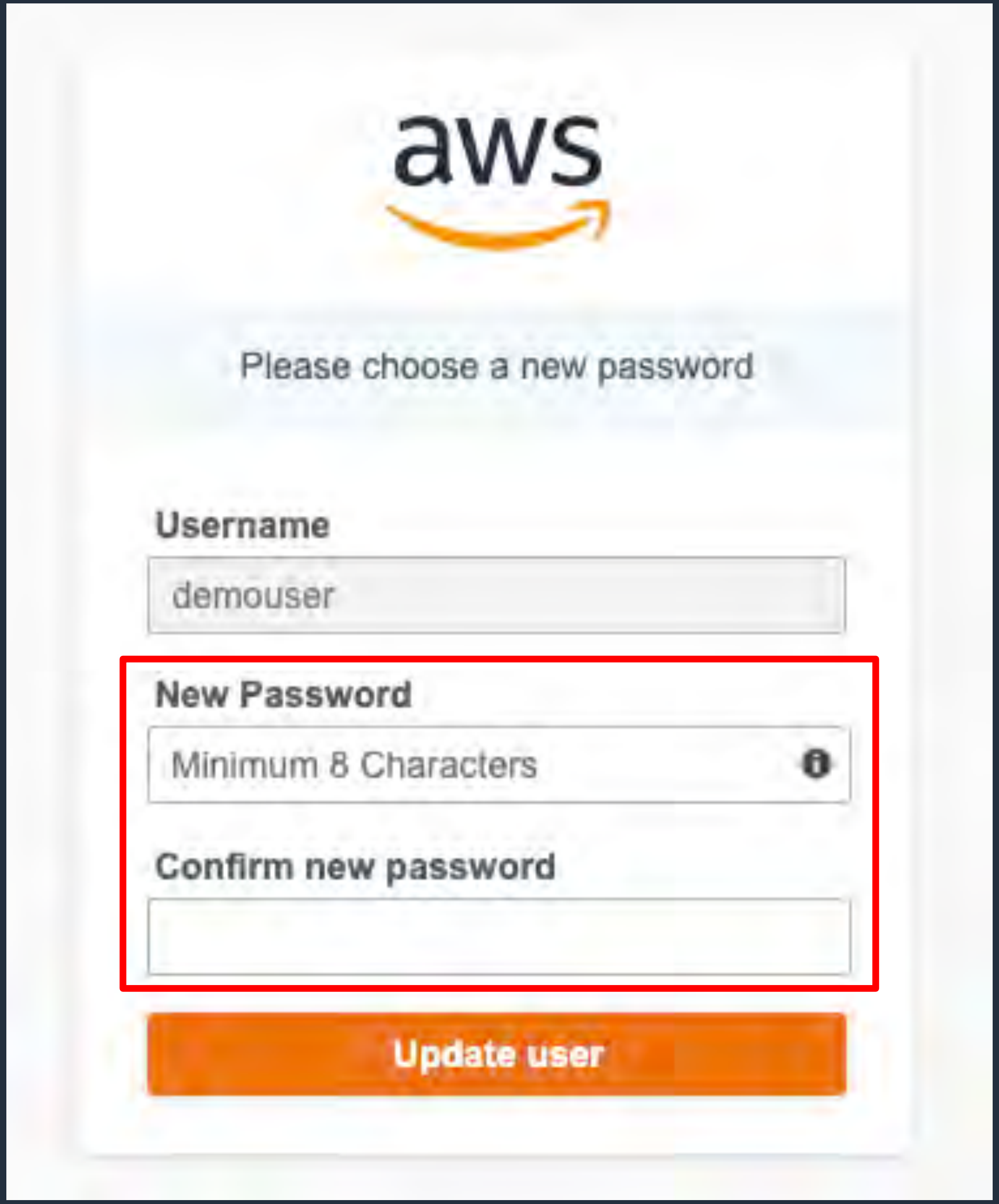
Your AWS Organization (AWS Account #805433377179) uses AWS Single Sign-On (SSO) to provide access to AWS accounts and business applications.

Your administrator has invited you to access the AWS Single Sign-On (SSO) user portal. Accepting this invitation activates your AWS SSO user account so that you can access assigned AWS accounts and applications. Click on the link below to accept this invitation.

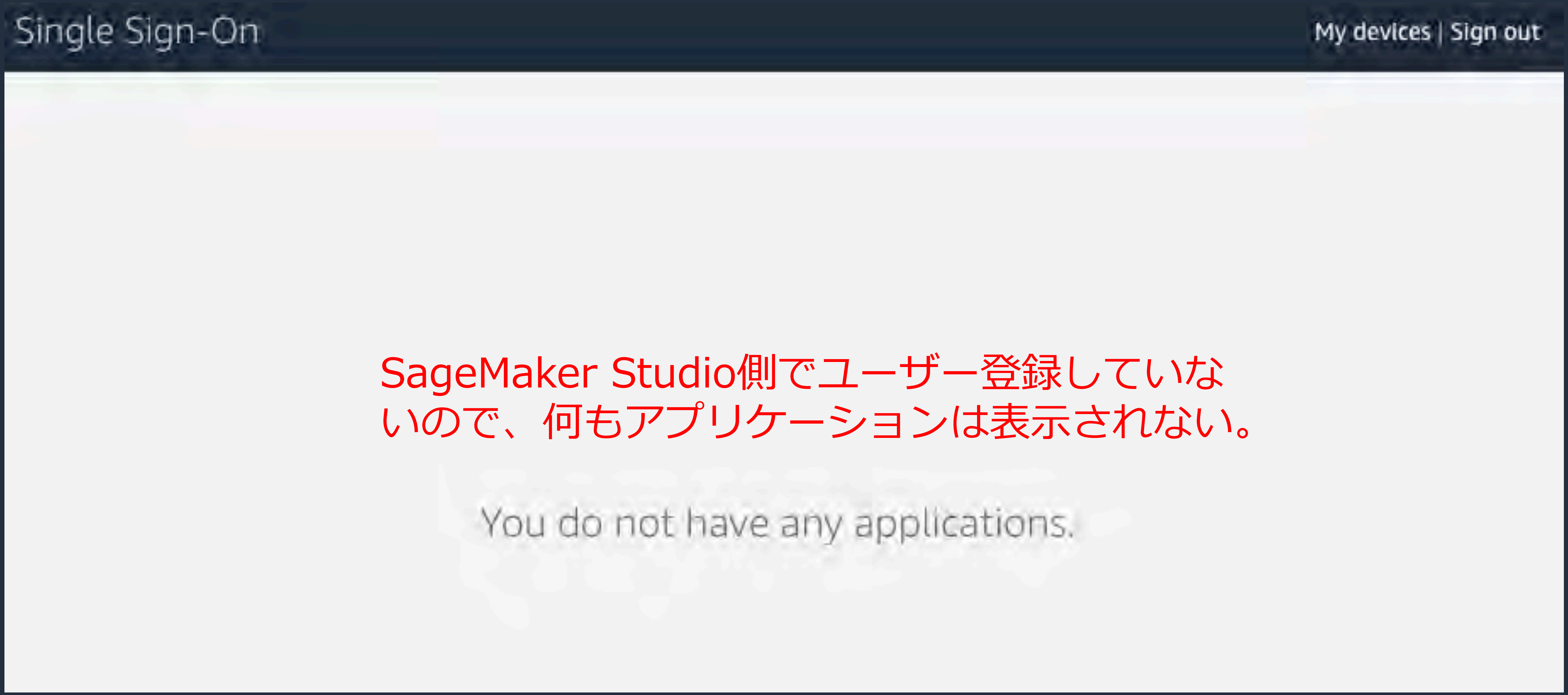
Accept invitation

This invitation will expire in 7 days.

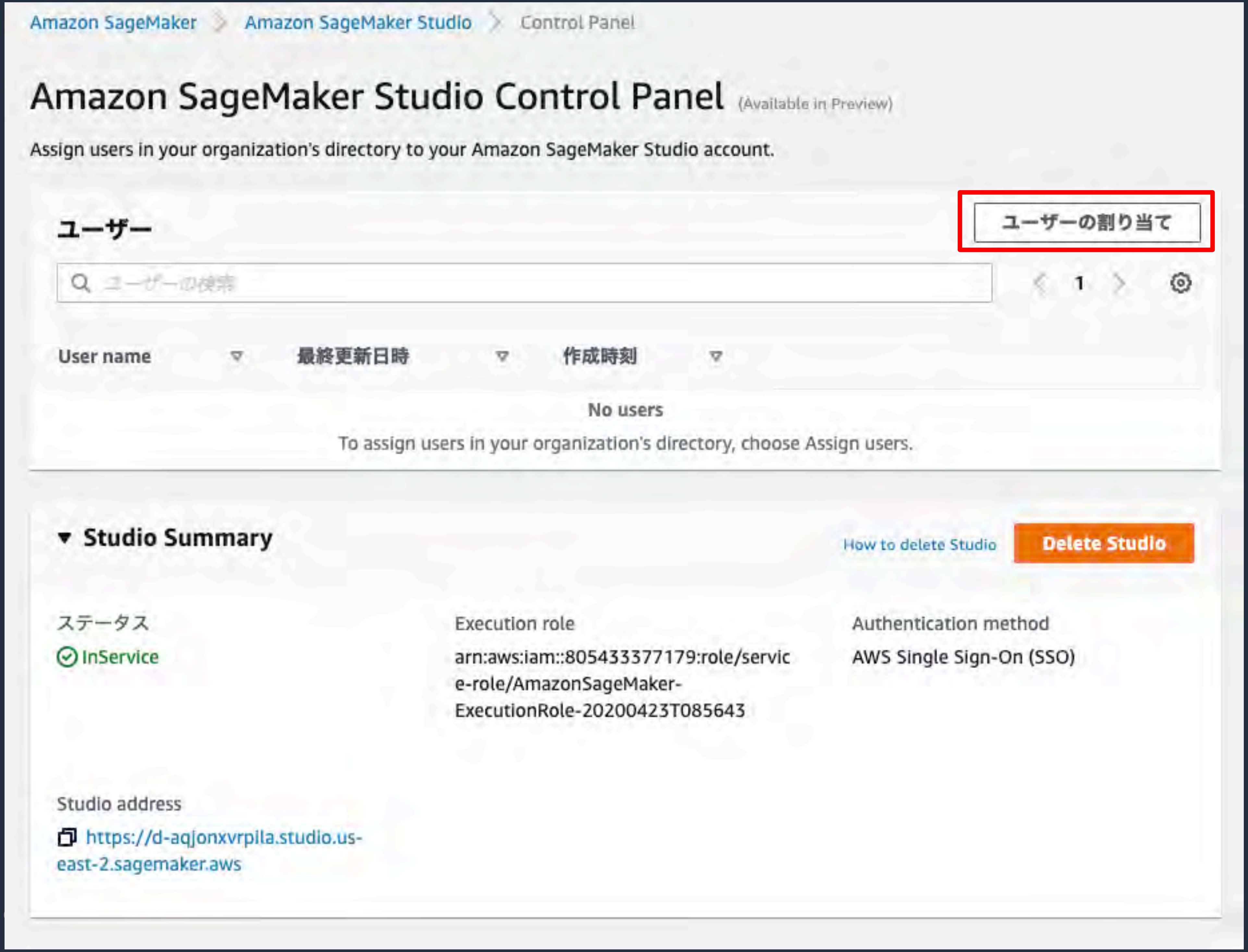
SSOパスワード設定


aws
Please choose a new password
Username
demouser
New Password
Minimum 8 Characters
Confirm new password
Update user

SSOユーザー作成完了確認

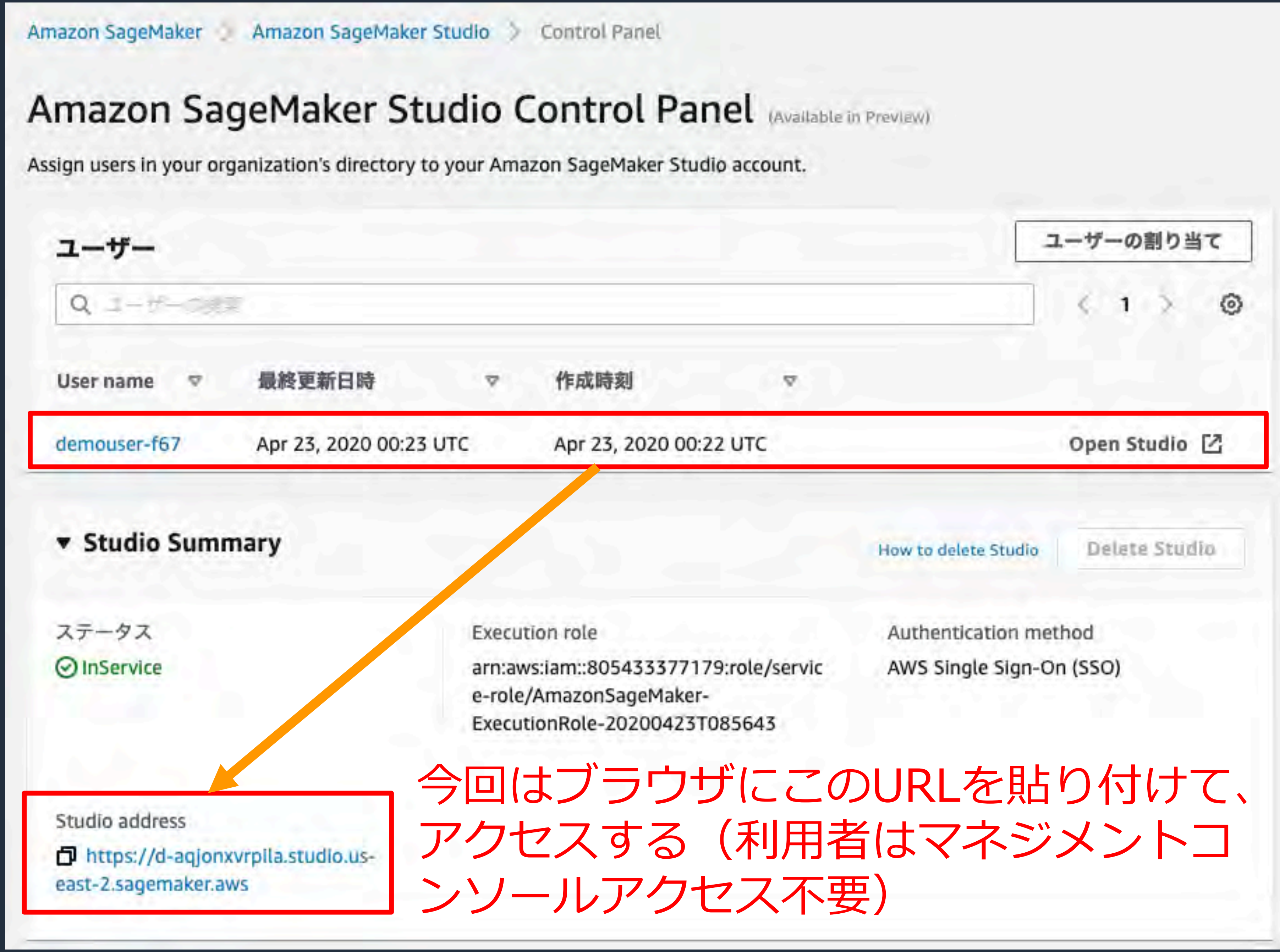


SageMaker Studioにユーザー割り当て



作成したSSOユーザーを選択

ユーザー割り当て完了確認

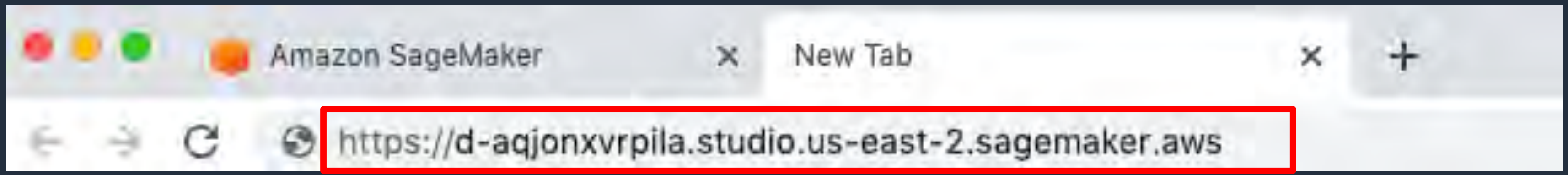


確認。Open Studioから開くことも可能

今回はブラウザにこのURLを貼り付けて、アクセスする（利用者はマネジメントコンソールアクセス不要）

SageMaker Studio の起動

ブラウザにドメインURLを入力

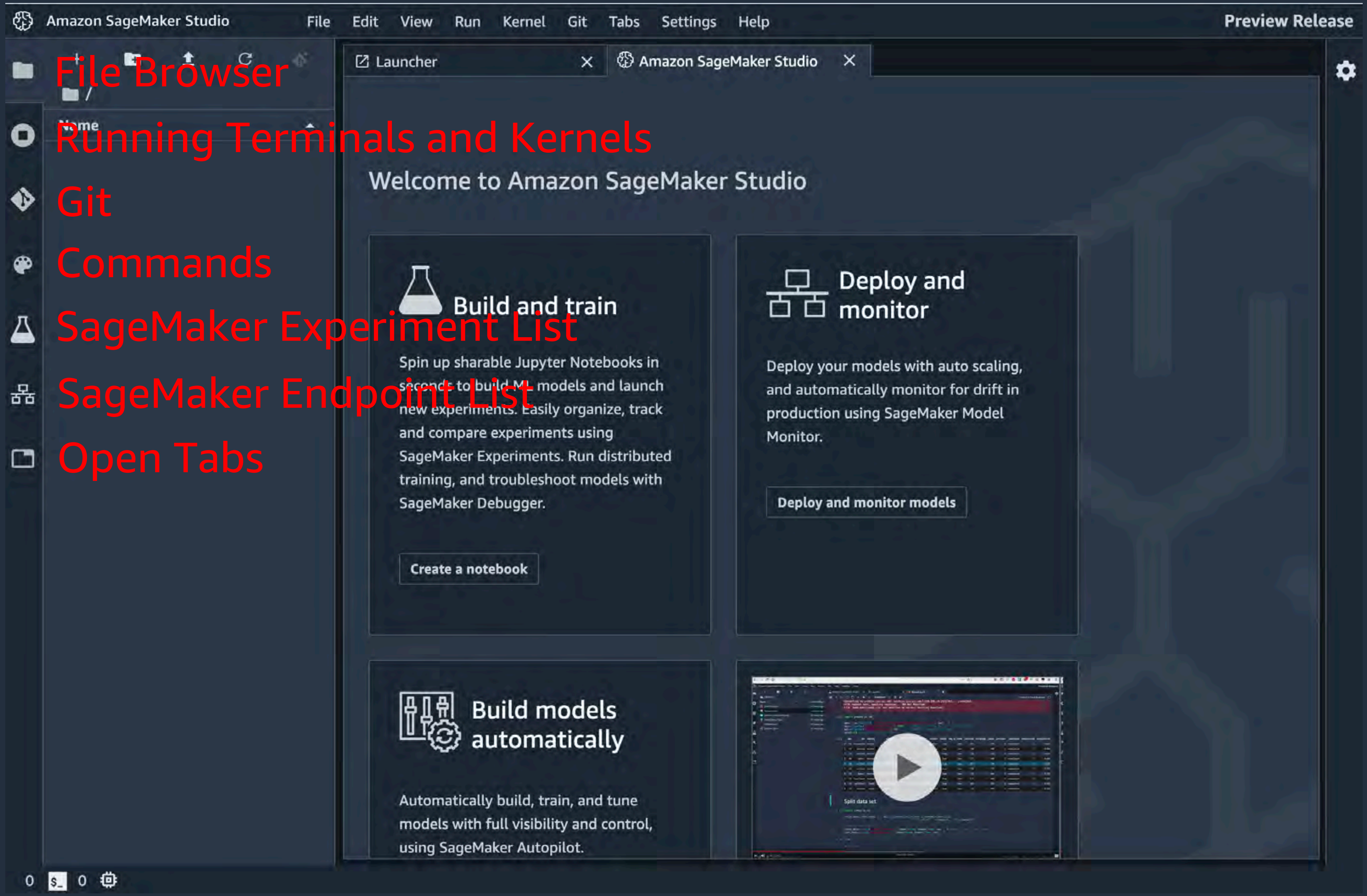


SSOを実行していない場合、ログイン画面が表示される
（今回はすでに認証しているため、表示されない）

SageMaker Studio の起動

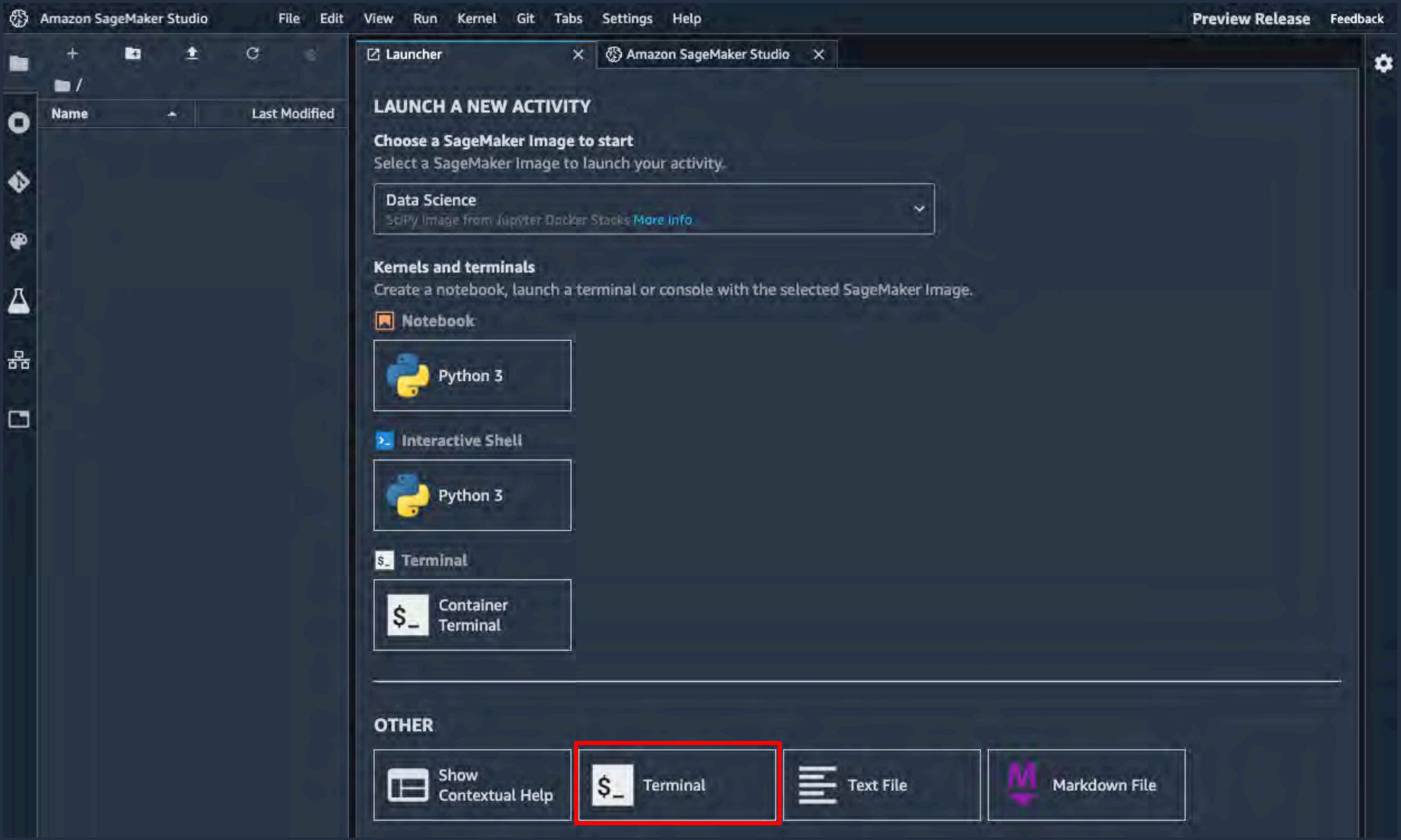
5分程度待つ

SageMaker Studioの初期画面



Datasetの準備

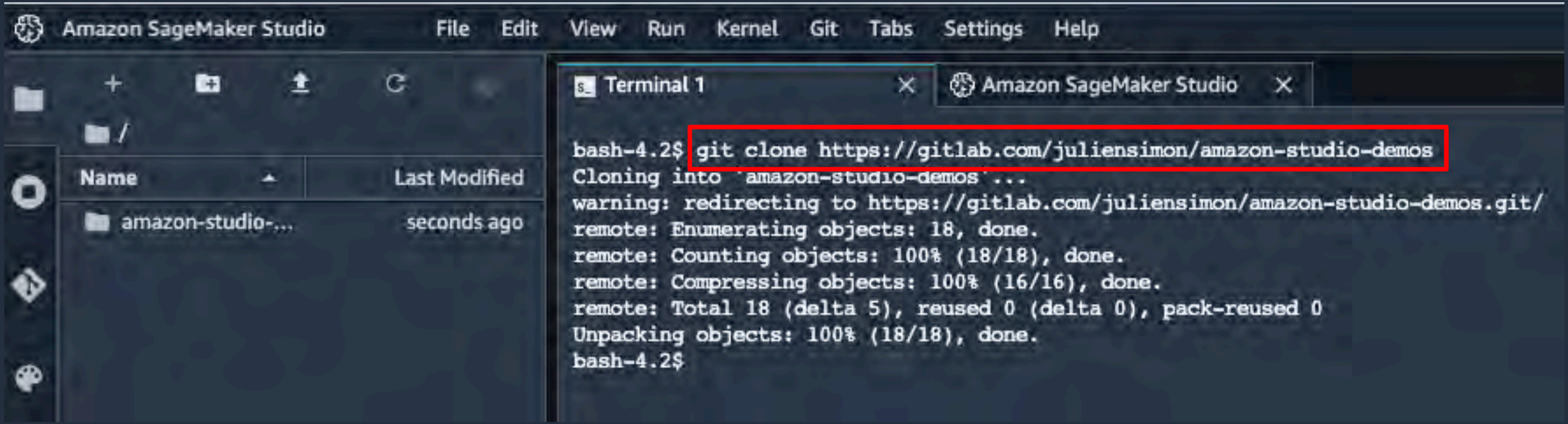
Terminalの起動



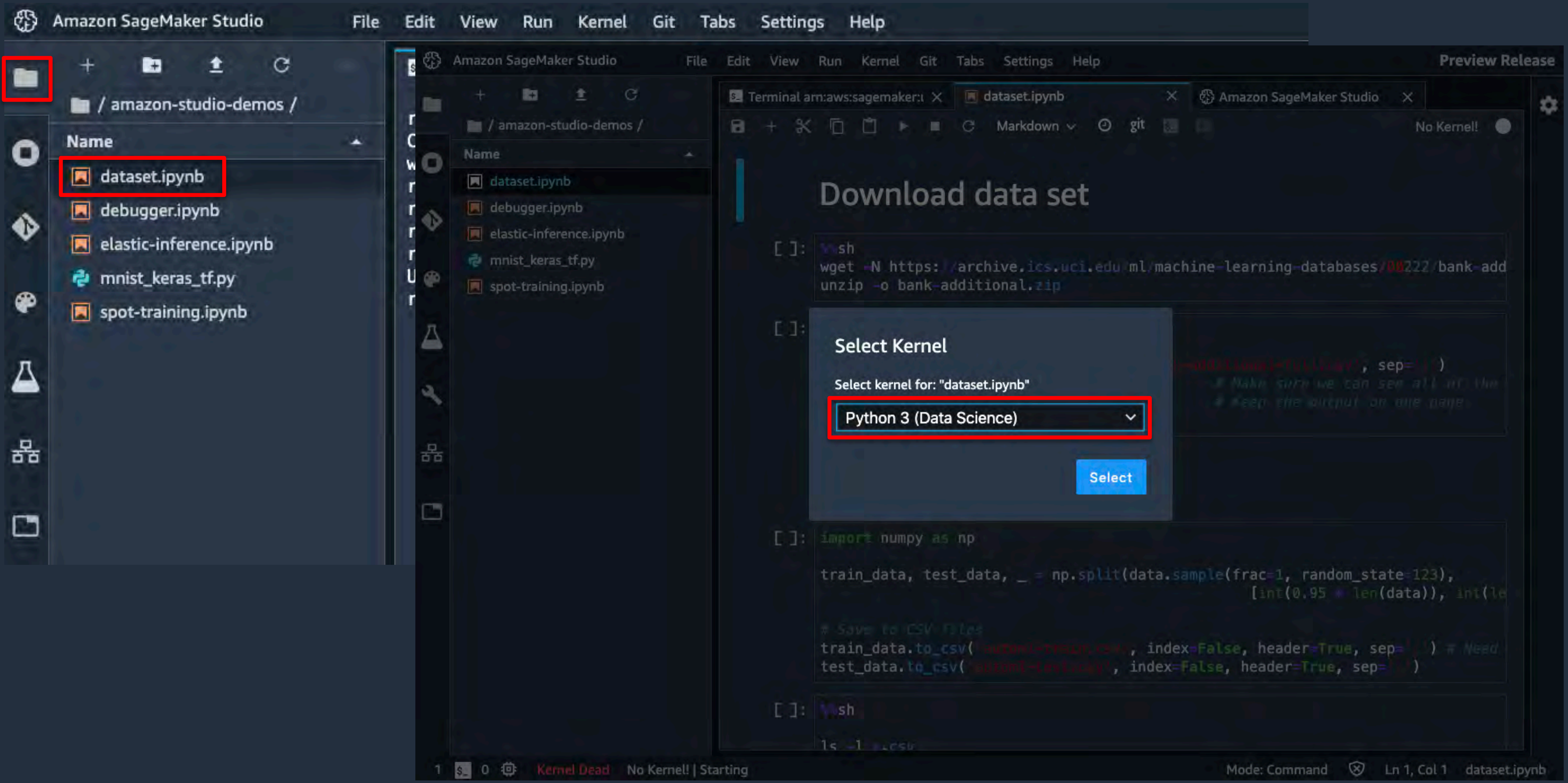
Assetのダウンロード

以下のコマンドでGitLabからAssetをダウンロード

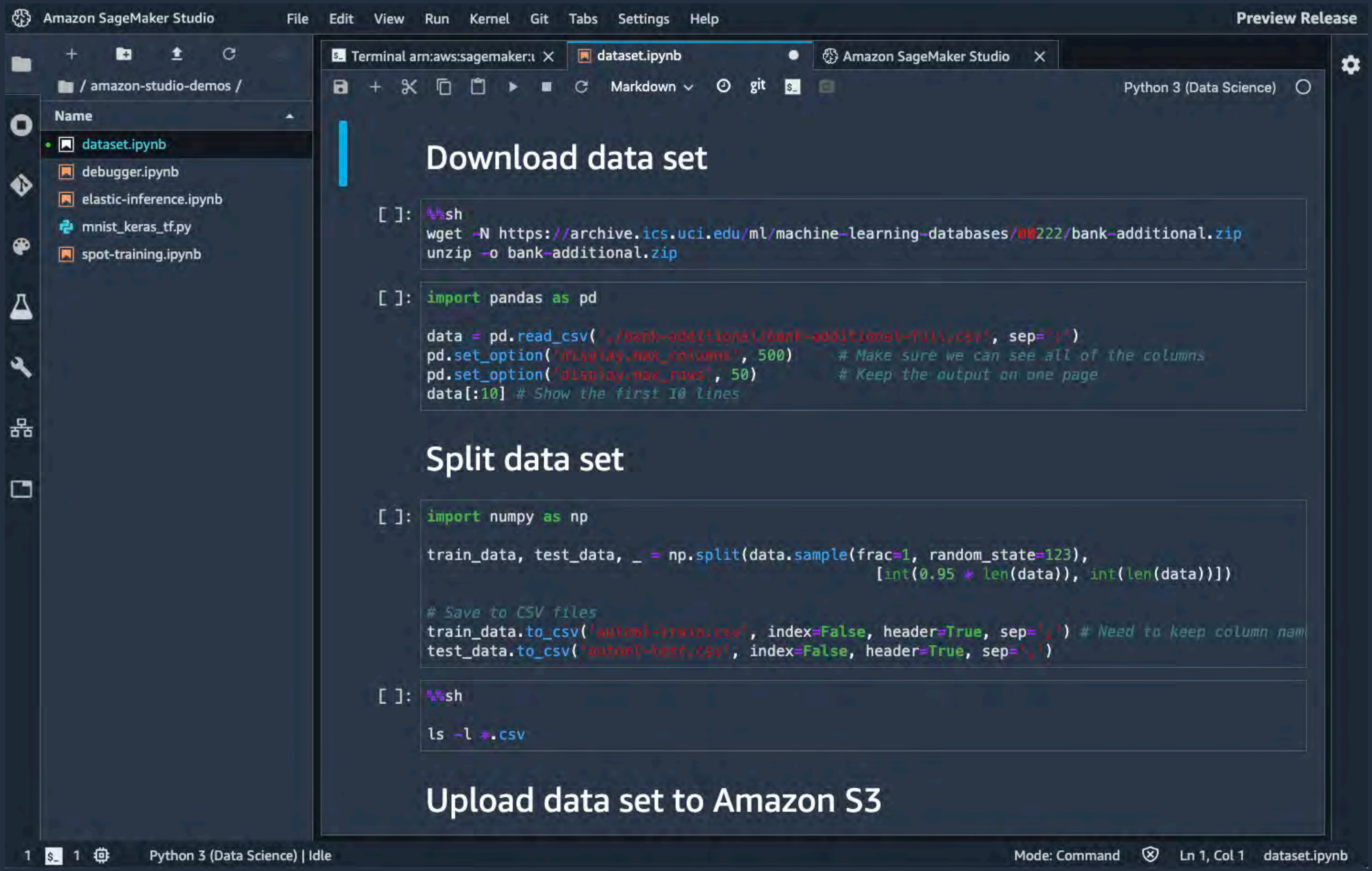
```
git clone https://gitlab.com/juliensimon/amazon-studio-demos
```



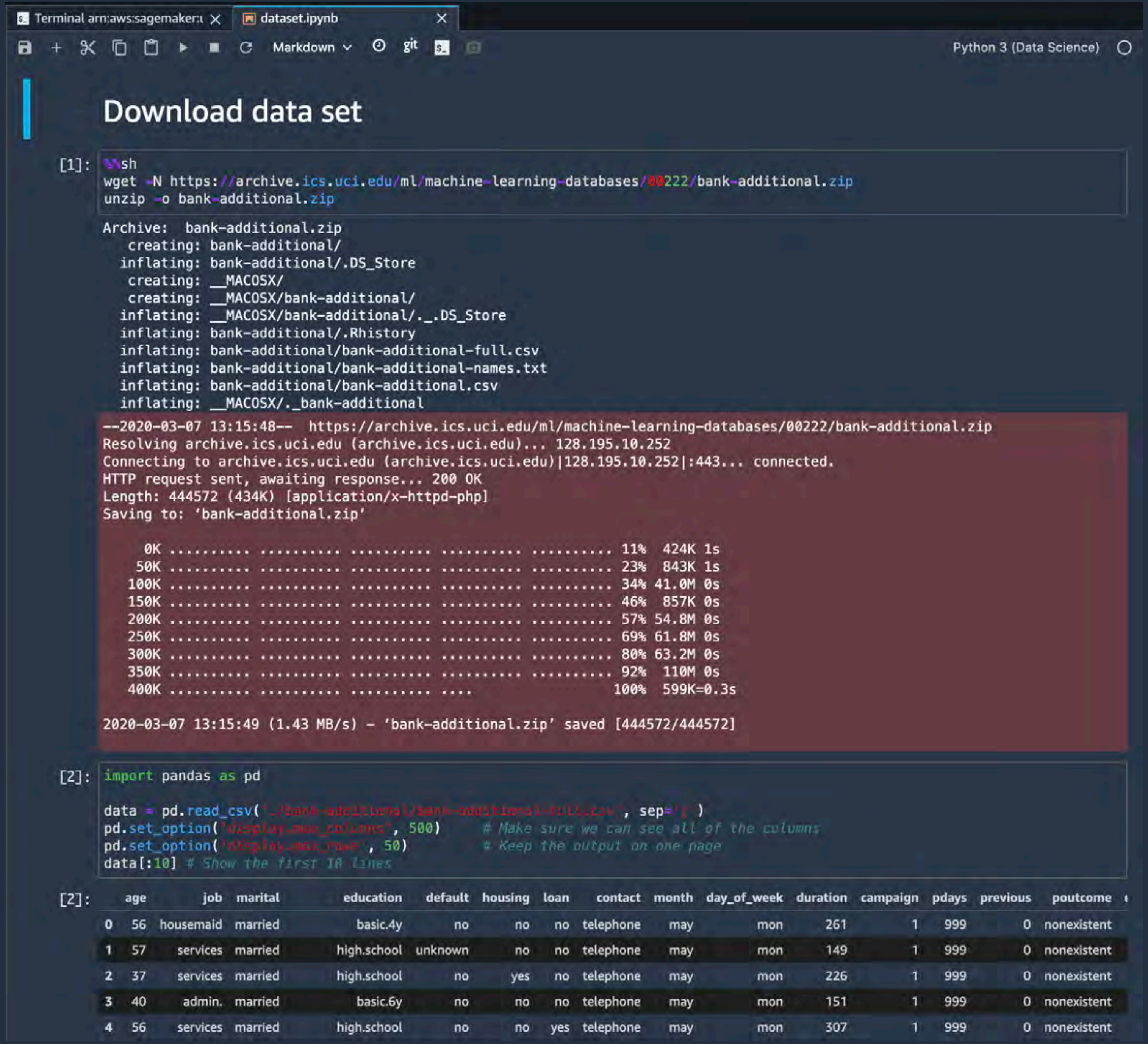
Notebookを開く



Notebookの表示



Notebookの実行



Notebookの実行
Datasetのサンプル

```
[2]: import pandas as pd

     data = pd.read_csv('./bank-additional/bank-additional-full.csv', sep=';')
     pd.set_option('display.max_columns', 500)     # Make sure we can see all of the columns
     pd.set_option('display.max_rows', 50)         # Keep the output on one page
     data[:10] # Show the first 10 lines
```

[2]:

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaigr
0	56	housemaid	married	basic.4y	no	no	no	telephone	may	mon	261	
1	57	services	married	high.school	unknown	no	no	telephone	may	mon	149	
2	37	services	married	high.school	no	yes	no	telephone	may	mon	226	
3	40	admin.	married	basic.6y	no	no	no	telephone	may	mon	151	
4	56	services	married	high.school	no	no	yes	telephone	may	mon	307	
5	45	services	married	basic.9y	unknown	no	no	telephone	may	mon	198	
6	59	admin.	married	professional.course	no	no	no	telephone	may	mon	139	
7	41	blue-collar	married	unknown	unknown	no	no	telephone	may	mon	217	
8	24	technician	single	professional.course	no	yes	no	telephone	may	mon	380	
9	25	services	single	high.school	no	yes	no	telephone	may	mon	50	

Notebookの実行
Train/Testデータに分割

Split data set

[3]:
```
import numpy as np

train_data, test_data, _ = np.split(data.sample(frac=1, random_state=123),
                                    [int(0.95 * len(data)), int(len(data))])

# Save to CSV files
train_data.to_csv('automl-train.csv', index=False, header=True, sep=',') # Need to keep column nam
test_data.to_csv('automl-test.csv', index=False, header=True, sep=',')
```

[4]:
```
%%sh

ls -l *.csv
```

```
-rw-r--r-- 1 root root  257339 Mar  7 12:43 automl-test.csv
-rw-r--r-- 1 root root 4889516 Mar  7 12:43 automl-train.csv
```

Notebookの実行
S3 Bucketにアップロード

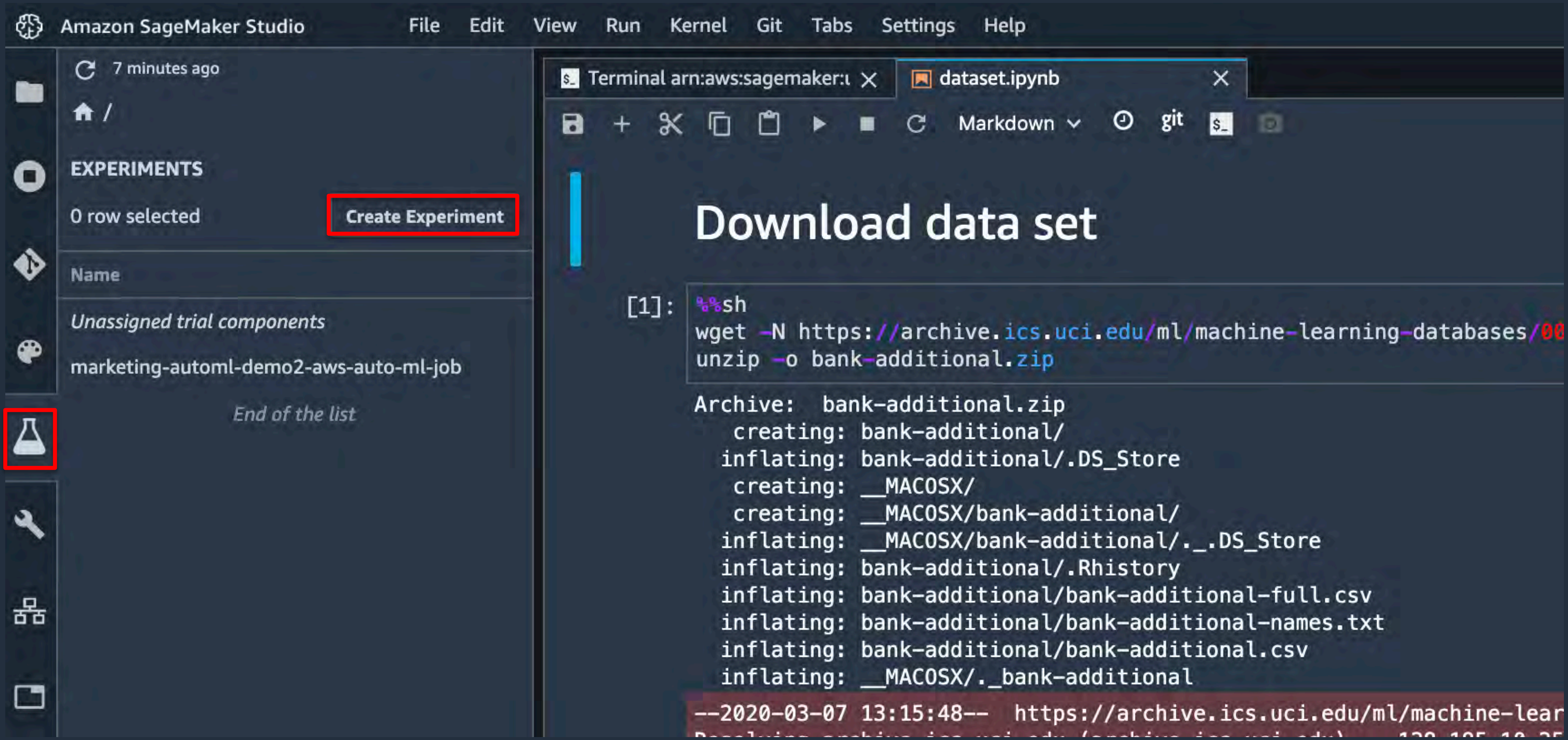
Upload data set to Amazon S3

```
[5]: import sagemaker

     prefix = 'sagemaker/DEMO-automl-dm/input'
     sess   = sagemaker.Session()

     uri = sess.upload_data(path='automl-train.csv', key_prefix=prefix)
     print(uri)

s3://sagemaker-us-east-2-226851534715/sagemaker/DEMO-automl-dm/input/automl-train.csv
```

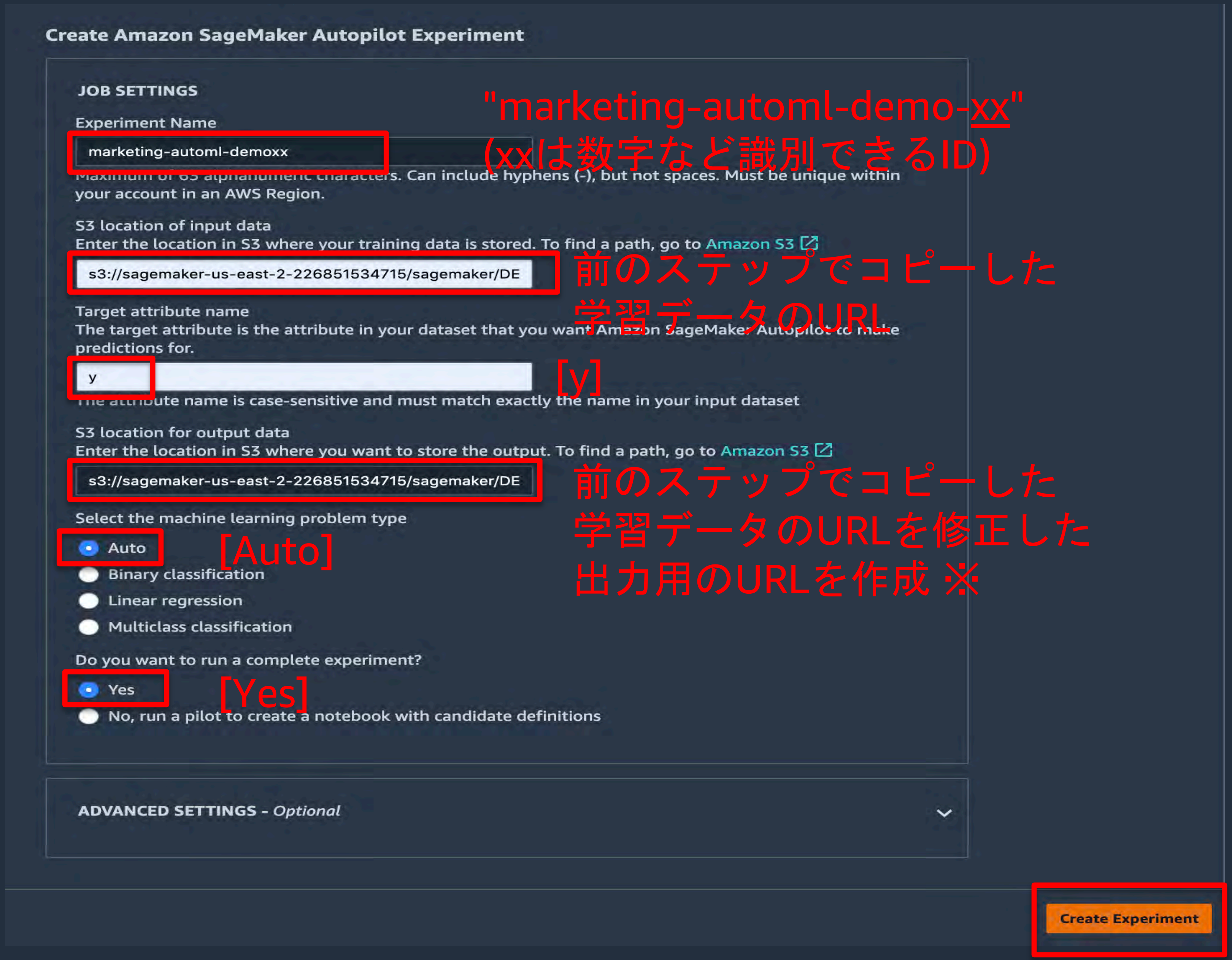
このファイルが学習データのデータソースとなる
後でExperimentを作成するために、URLをコピーしておく

Experiment(Autopilot)の実行

Experimentの作成



Experimentの作成



"marketing-automl-demo-xx"
(xxは数字など識別できるID)

前のステップでコピーした
学習データのURL

[y]

前のステップでコピーした
学習データのURLを修正した
出力用のURLを作成 ※

[Auto]

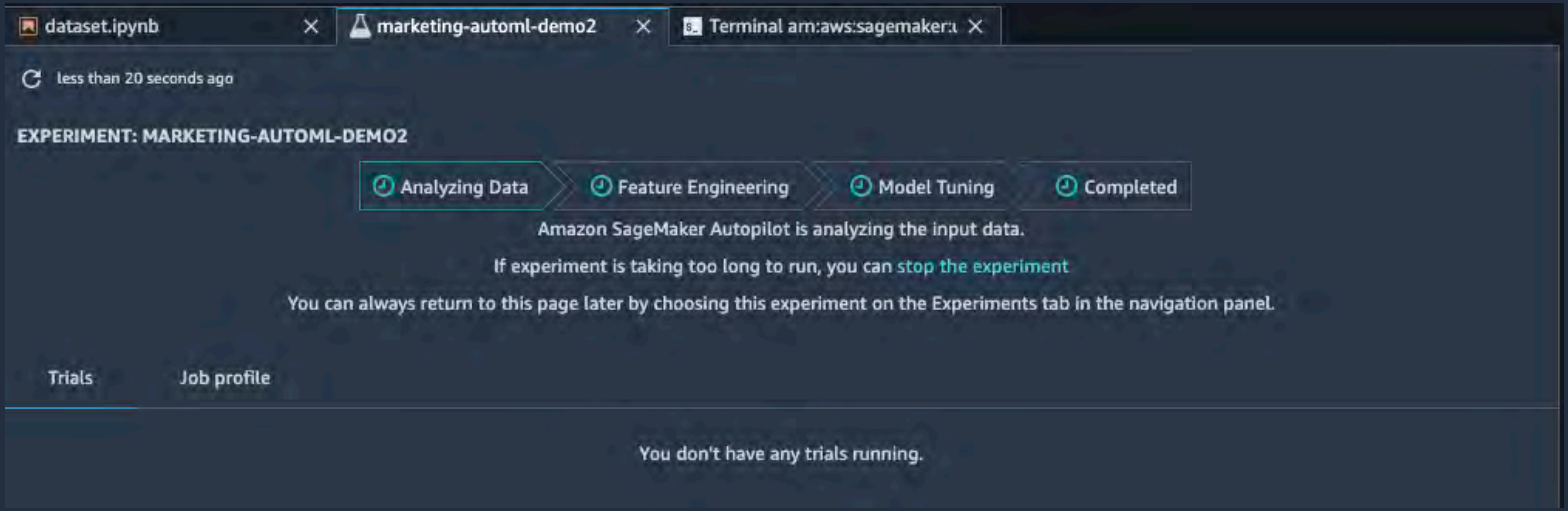
[Yes]

※例
(S3 Bucket名は、自身のもの
に変更)

データソース
s3://sagemaker-us-east-2-226851534715/sagemaker/DEMO-automl-dm/input/automl-train.csv

出力用URL
s3://sagemaker-us-east-2-226851534715/sagemaker/DEMO-automl-dm/output/

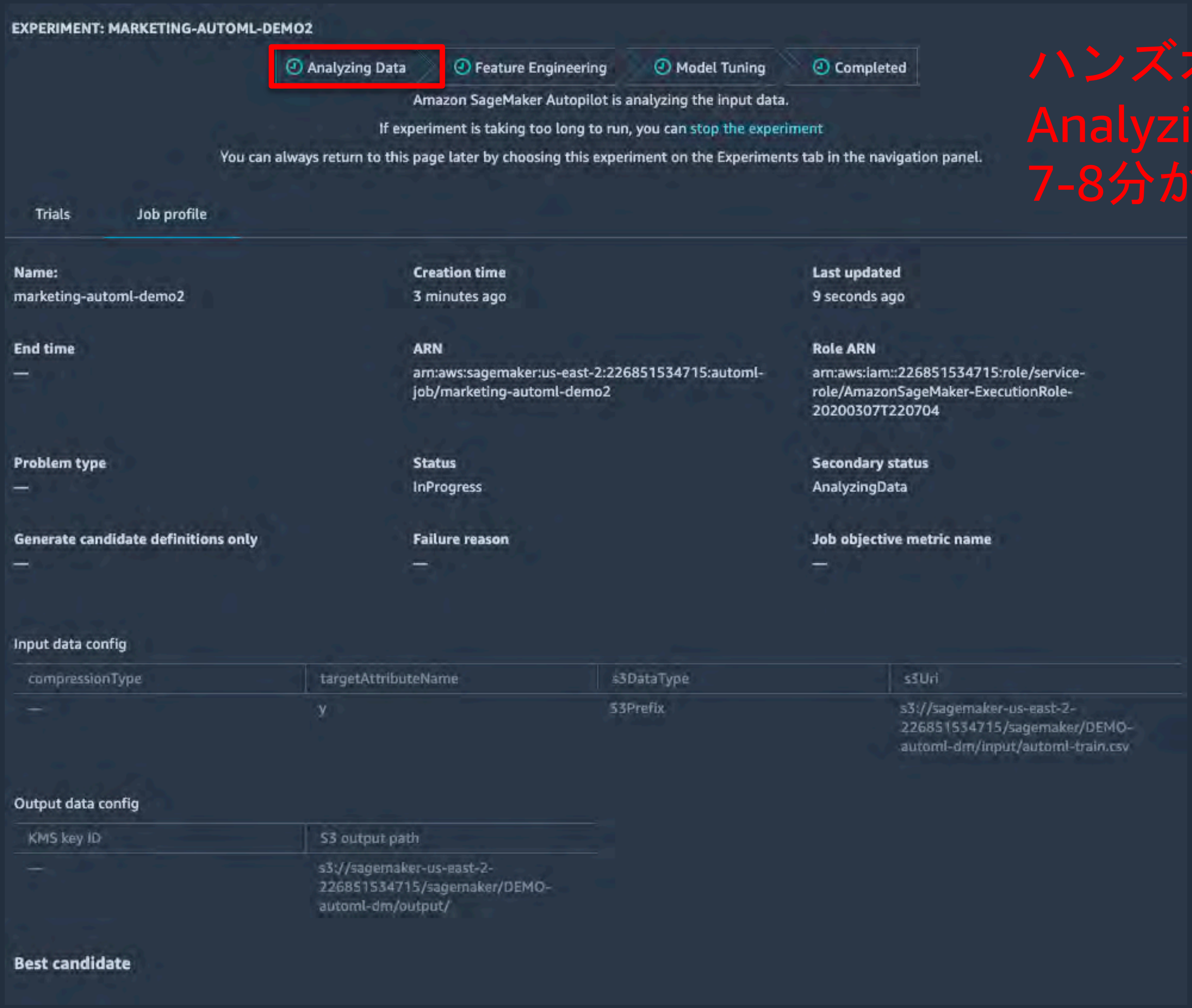
Experimentの開始



Experimentは4つのStageで実行される

- Analyzing Data
  - AutoPilotがDatasetを参照し、統計情報を分析
  - 自動的に問題タイプを判別し複数アルゴリズムを選択、Preprocessing用のScriptを生成
- Feature Engineering
  - Preprocessing scriptを実行し、候補となる学習アルゴリズムに合わせてデータ分割、データ変換等の処理を実行
- Model Tuning
  - 複数の候補アルゴリズムに対してHPO(Hyper-Parameter Optimization)を実行
  - 選択されアルゴリズムとHPOに従い、TrainingとValidationを実行し、Training結果を出力
- Completed

Analyzing Data
Job Profileの確認



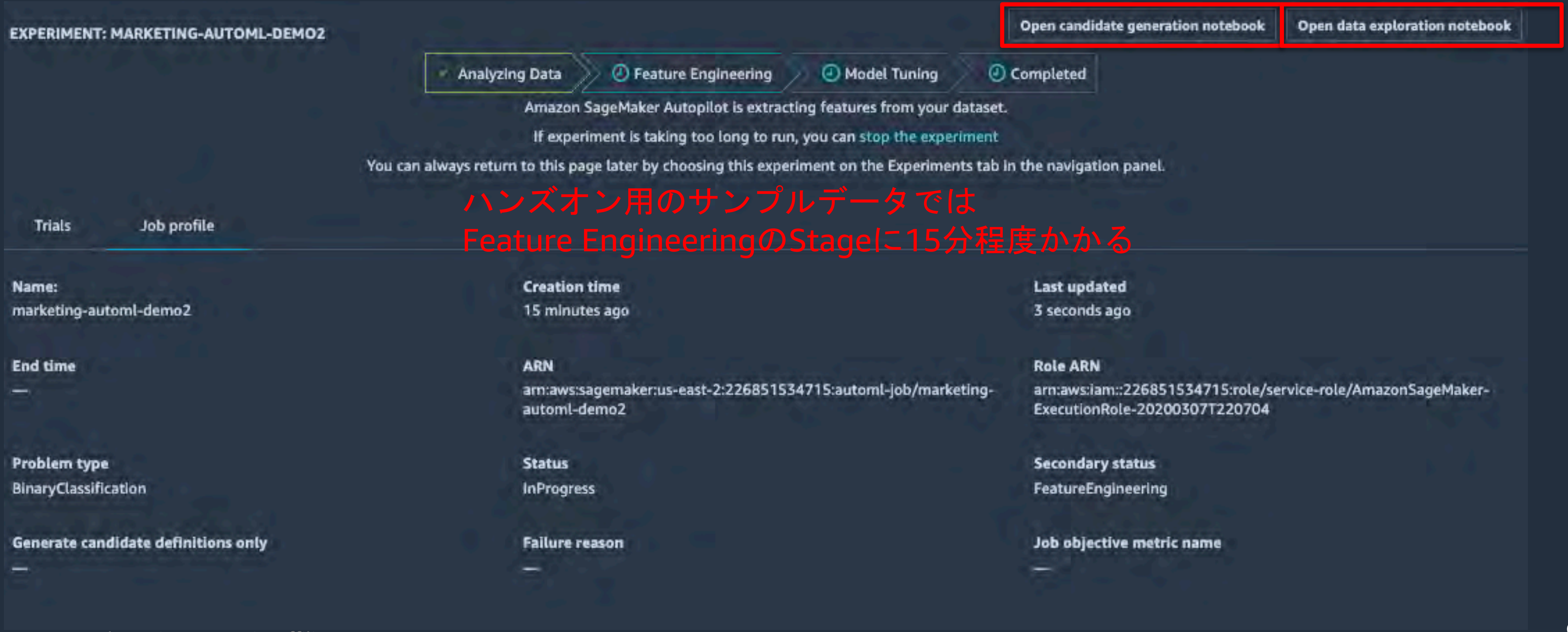
ハンズオン用のサンプルデータでは
Analyzing DataのStageに
7-8分かかる

Feature Engineering
生成されたNotebookの確認

Analyzing Dataが完了すると、Feature Engineeringに移行
SagemakerがPreprocess, HPO, Training用に自動でNotebookを作成

Candidate Generation Notebook

Data Exploration Notebook



生成されたNotebookの確認
Data Exploration Notebook

データ探索用のNotebookを自動で生成

Datasetに関する情報を提供

- レコード数、Attribute数
- Target feature
- アルゴリズム: Binary classification
- データセットのサンプル
- 欠損値
- データセットの品質を上げるための提案
- その他データの統計情報 (max, min, avg, mean 等)

ユーザーにてカスタマイズし、個別に実行することも可能

Amazon SageMaker Autopilot Data Exploration

This report provides insights about the dataset you provided as input to the AutoML job. It was automatically generated by the AutoML training job: **marketing-automl-demo2**.

As part of the AutoML job, the input dataset was randomly split into two pieces, one for **training** and one for **validation**. The training dataset was randomly sampled, and metrics were computed for each of the columns. This notebook provides these metrics so that you can:

1. Understand how the job analyzed features to select the candidate pipelines.
2. Modify and improve the generated AutoML pipelines using knowledge that you have about the dataset.

We read `39128` rows from the training dataset. The dataset has `21` columns and the column named `y` is used as the target column. This is identified as a `BinaryClassification` problem. Here are 2 examples of labels: `['yes', 'no']`.

Suggested Action Items

- Look for sections like this for recommended actions that you can take.

Contents

1. Dataset Sample
2. Column Analysis

Dataset Sample

The following table is a random sample of **10** rows from the training dataset. For ease of presentation, we are only showing **20 of the 21** columns of the dataset.

Suggested Action Items

- Verify the input headers correctly align with the columns of the dataset sample. If they are incorrect, update the header names of your input dataset in Amazon Simple Storage Service (Amazon S3).

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	...	campaign	pdays	previous	poutcome	emp.var.rate	cons.price.id
0	52	blue-collar	married	basic.4y	unknown	yes	no	cellular	jul	fri	...	1	999	0	nonexistent	1.4	93.9179999999999
1	32	technician	single	professional.course	no	yes	no	cellular	may	wed	...	2	999	0	nonexistent	-1.8	92.8929999999999
2	47	services	married	high.school	unknown	yes	no	cellular	aug	fri	...	1	999	0	nonexistent	1.4	93.444
3	51	technician	single	university.degree	no	yes	no	cellular	jun	fri	...	1	999	0	nonexistent	-2.9	92.963
4	54	blue-collar	single	unknown	no	yes	no	cellular	jul	mon	...	3	999	0	nonexistent	1.4	93.9179999999999
5	43	housemaid	married	basic.9y	unknown	no	no	cellular	jul	wed	...	1	999	0	nonexistent	1.4	93.9179999999999
6	38	blue-collar	married	basic.6y	no	no	no	cellular	jul	wed	...	2	999	0	nonexistent	1.4	93.9179999999999
7	27	blue-collar	married	basic.9y	no	yes	no	cellular	may	wed	...	3	999	0	nonexistent	-1.8	92.8929999999999
8	29	blue-collar	single	university.degree	no	yes	no	cellular	aug	wed	...	2	999	0	nonexistent	1.4	93.444
9	32	admin.	married	university.degree	no	yes	yes	cellular	aug	mon	...	1	999	0	nonexistent	1.4	93.444

生成されたNotebookの確認
Candidate Generation Notebook

Training候補を作成するためのNotebookを自動生成

- 学習アルゴリズム
- データ変換処理
  - アルゴリズムに合わせたデータ変換処理
  - 実際に実行するスクリプトへのリンク
- HPO
  - Hyperparameter range, batch size, depth 等を定義
  - HOPは、内部でSageMakerのHyperparameter Jobを呼び出し
- Objective Metrics

ユーザーにてカスタマイズし、個別に実行することも可能

Amazon SageMaker Autopilot Candidate Definition Notebook

This notebook was automatically generated by the AutoML job **marketing-automl-demo2**. This notebook allows you to customize the candidate definitions and execute the SageMaker Autopilot workflow.

The dataset has **21** columns and the column named **y** is used as the target column. This is being treated as a **BinaryClassification** problem. The dataset also has **2** classes. This notebook will build a BinaryClassification model that **maximizes** the "**ACCURACY**" quality metric of the trained models. The "**ACCURACY**" metric provides the percentage of times the model predicted the correct class.

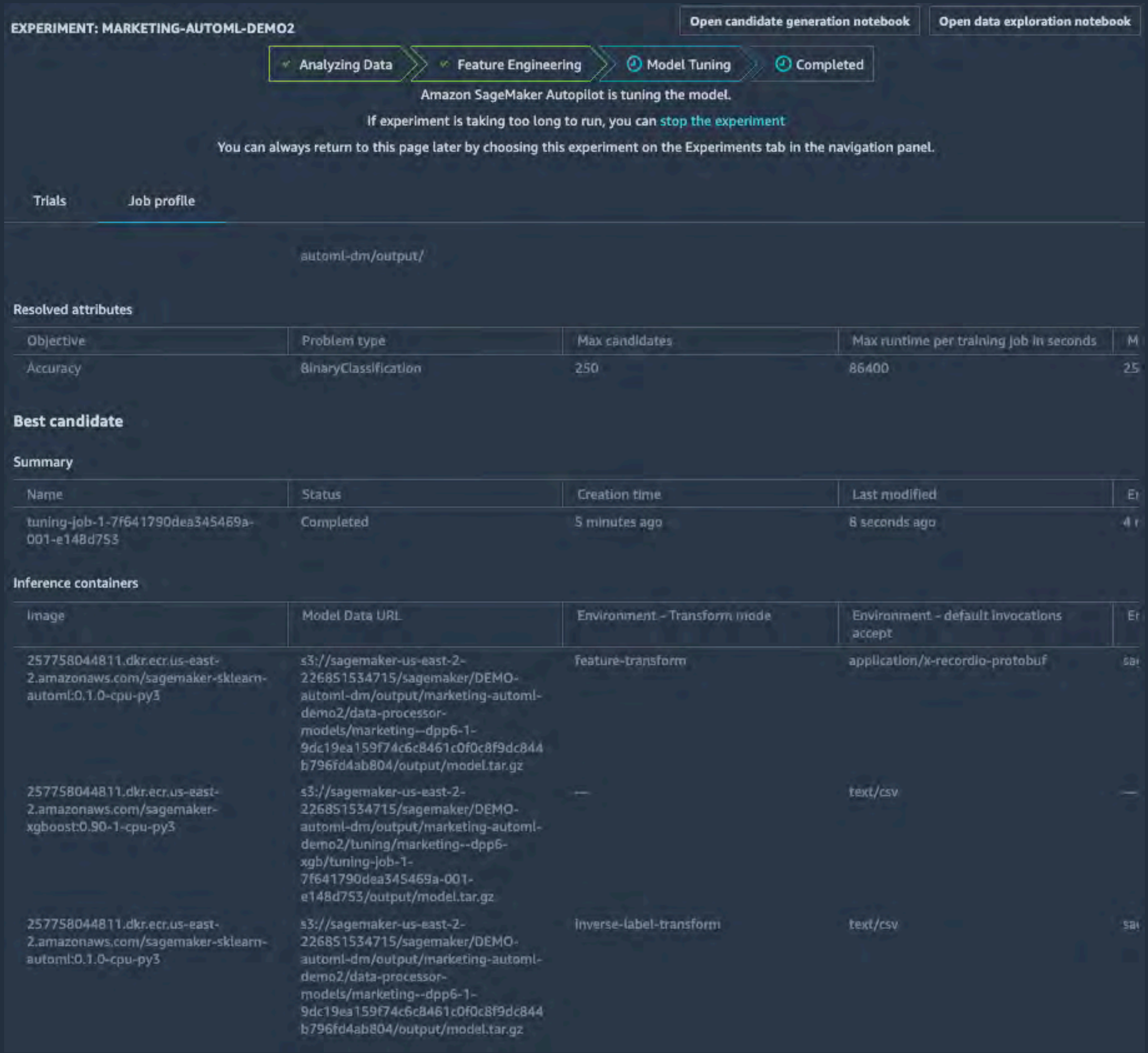
As part of the AutoML job, the input dataset has been randomly split into two pieces, one for **training** and one for **validation**. This notebook helps you inspect and modify the data transformation approaches proposed by Amazon SageMaker Autopilot. You can interactively train the data transformation models and use them to transform the data. Finally, you can execute a multiple algorithm hyperparameter optimization (multi-algo HPO) job that helps you find the best model for your dataset by jointly optimizing the data transformations and machine learning algorithms.

Available Knobs Look for sections like this for recommended settings that you can change.

Contents

Model Tuning

選択されたアルゴリズムとHPOに従い、複数のTrainingとValidationを実行し、結果を出力



Model Tuning
実行中のTraining Job

TRIAL COMPONENTS

0 row selected

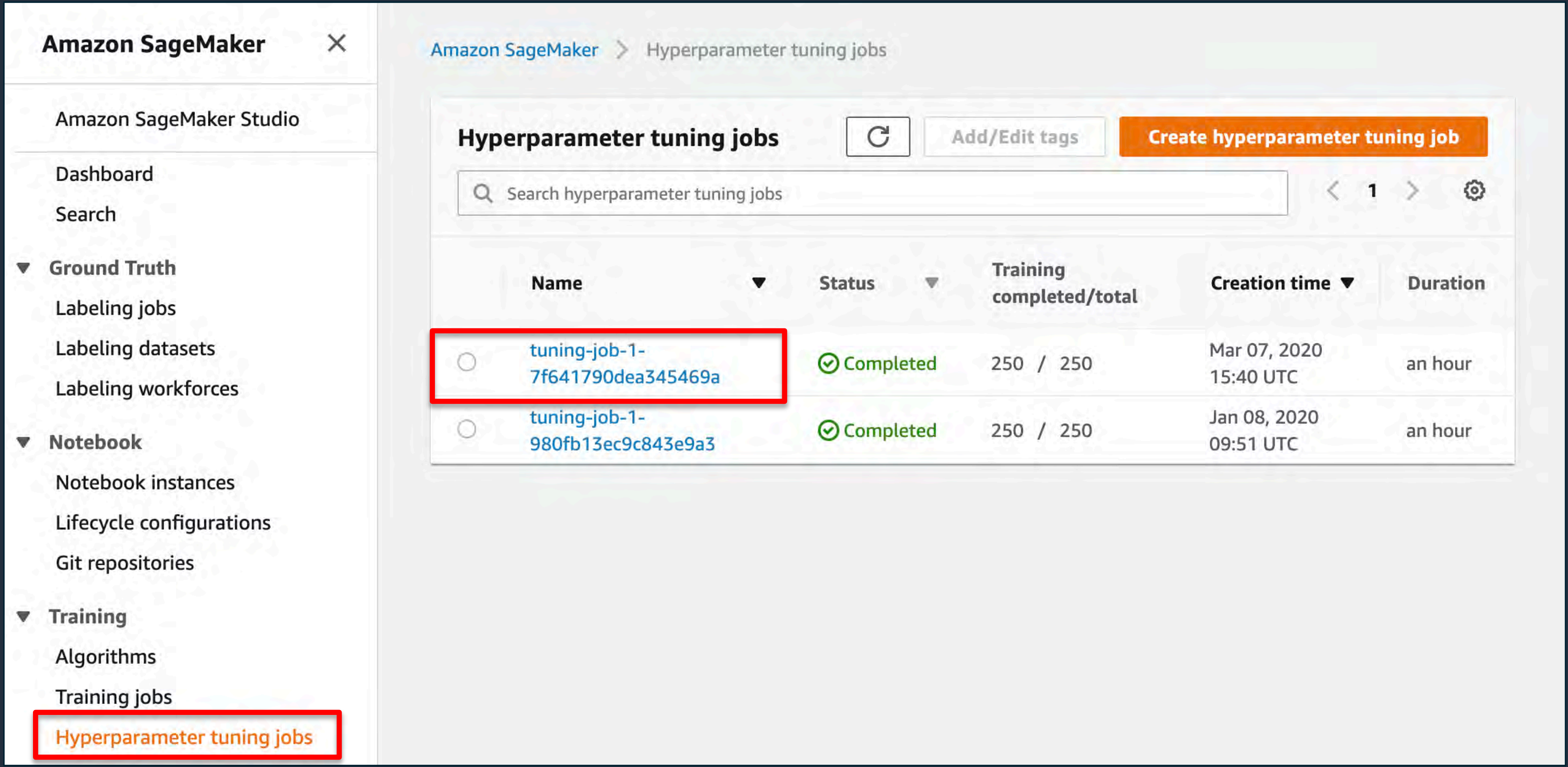
Add chart | Deploy model

Status	Experiment	Type	Trial	Trial component	Created on	Last modified	ObjectiveMetric	validation:error	validation:accuracy	train:error
In progress		Training job		tuning-job-1-7f64179...	9 seconds ago	8 seconds ago				
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	8 seconds ago	0.890351	0.109649	0.890351	0.001086
In progress		Training job		tuning-job-1-7f64179...	10 seconds ago	8 seconds ago				
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	12 seconds ago				
In progress		Training job		tuning-job-1-7f64179...	21 seconds ago	19 seconds ago				
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	27 seconds ago	0.5876038074493408			
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	29 seconds ago	0.901214057508			
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	1 minute ago	1 minute ago				
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	1 minute ago	0.112588	0.887412	0.112588	0.887519
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	4 minutes ago	1 minute ago	0.910032	0.089968	0.910032	0.071399
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	1 minute ago	1 minute ago				
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	1 minute ago	0.112588	0.887412	0.112588	0.887519
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	1 minute ago	1 minute ago				
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	1 minute ago	0.112588	0.887412	0.112588	0.887519
In progress	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	1 minute ago	1 minute ago				
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	1 minute ago	0.11258800327777863	0.8874120116233826	0.11258800327777863	0.8875190019607544
Completed	marketing-automl-de...	Training job	tuning-job-1-7f64179...	tuning-job-1-7f64179...	3 minutes ago	1 minute ago	0.112588	0.887412	0.112588	0.887519

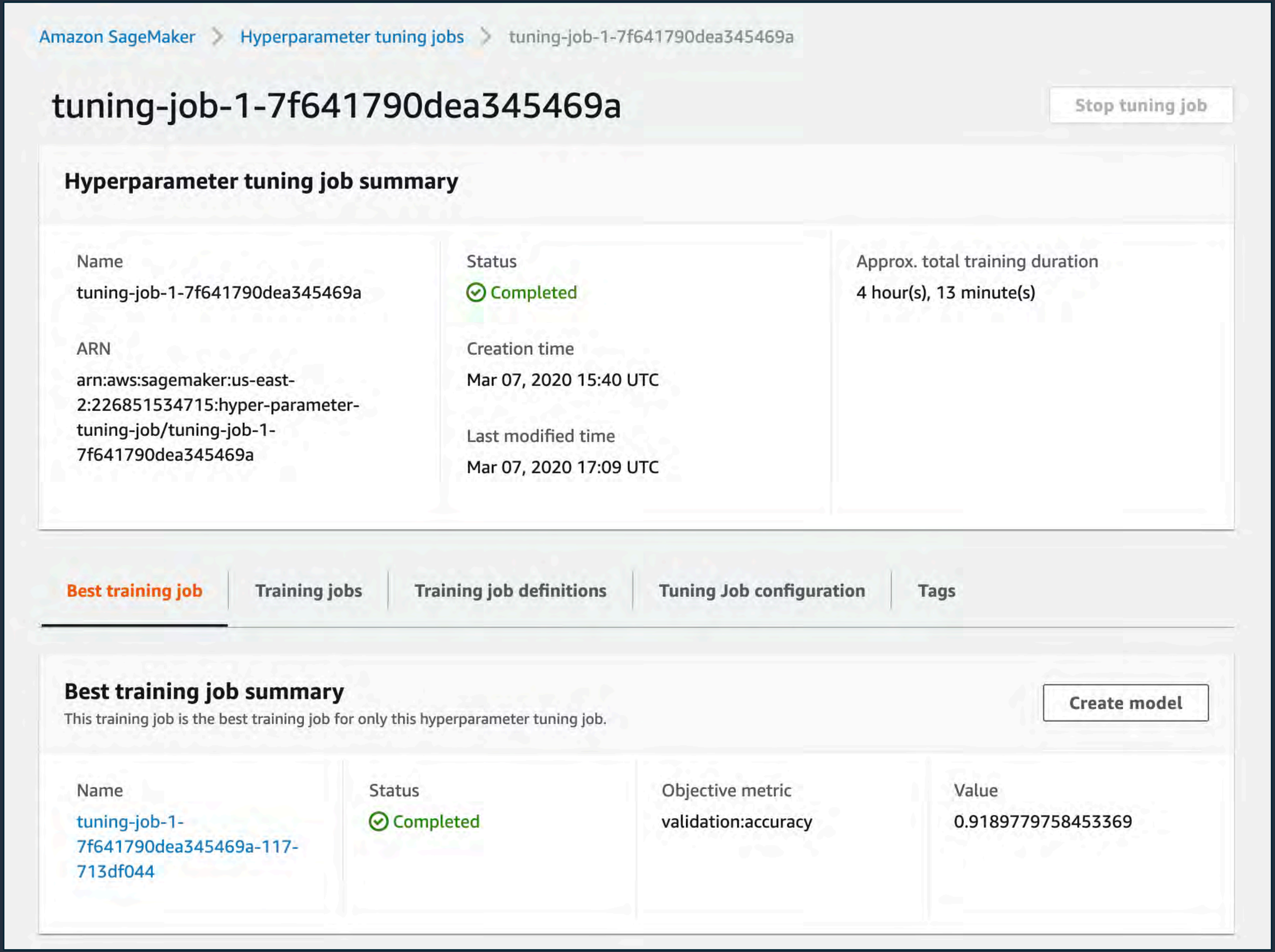
(参考)SageMaker Hyperparameter Tuning Job

- Model Tuningは、内部でSageMakerのHyperparameter Tuning Jobを呼び出し実行している
- SageMakerコンソールから確認できる

Management consoleからSageMakerへ移動
-> [Hyperparameter tuning jobs]
または
https://us-east-2.console.aws.amazon.com/sagemaker/

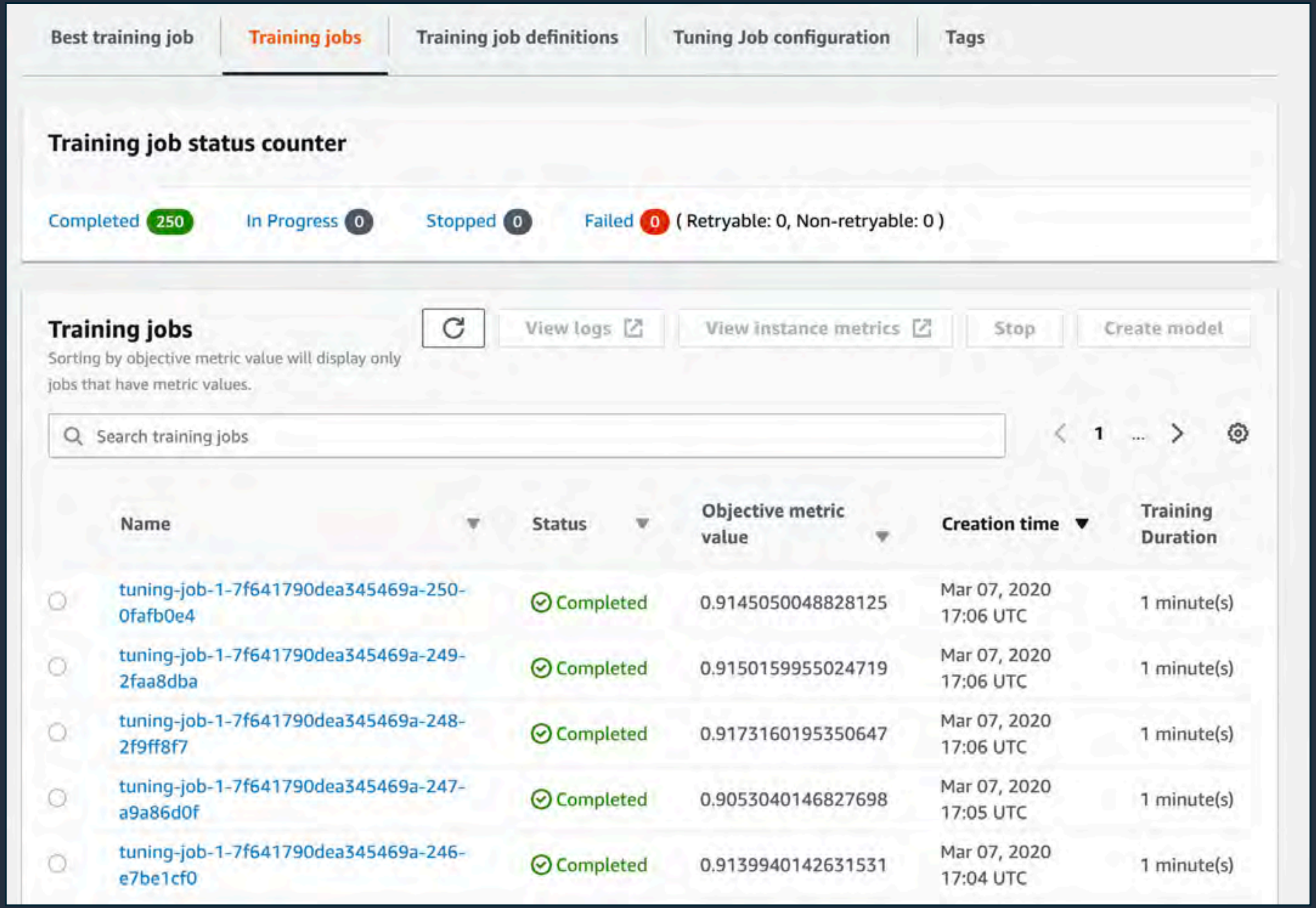


(参考)SageMaker Hyperparameter Tuning Job

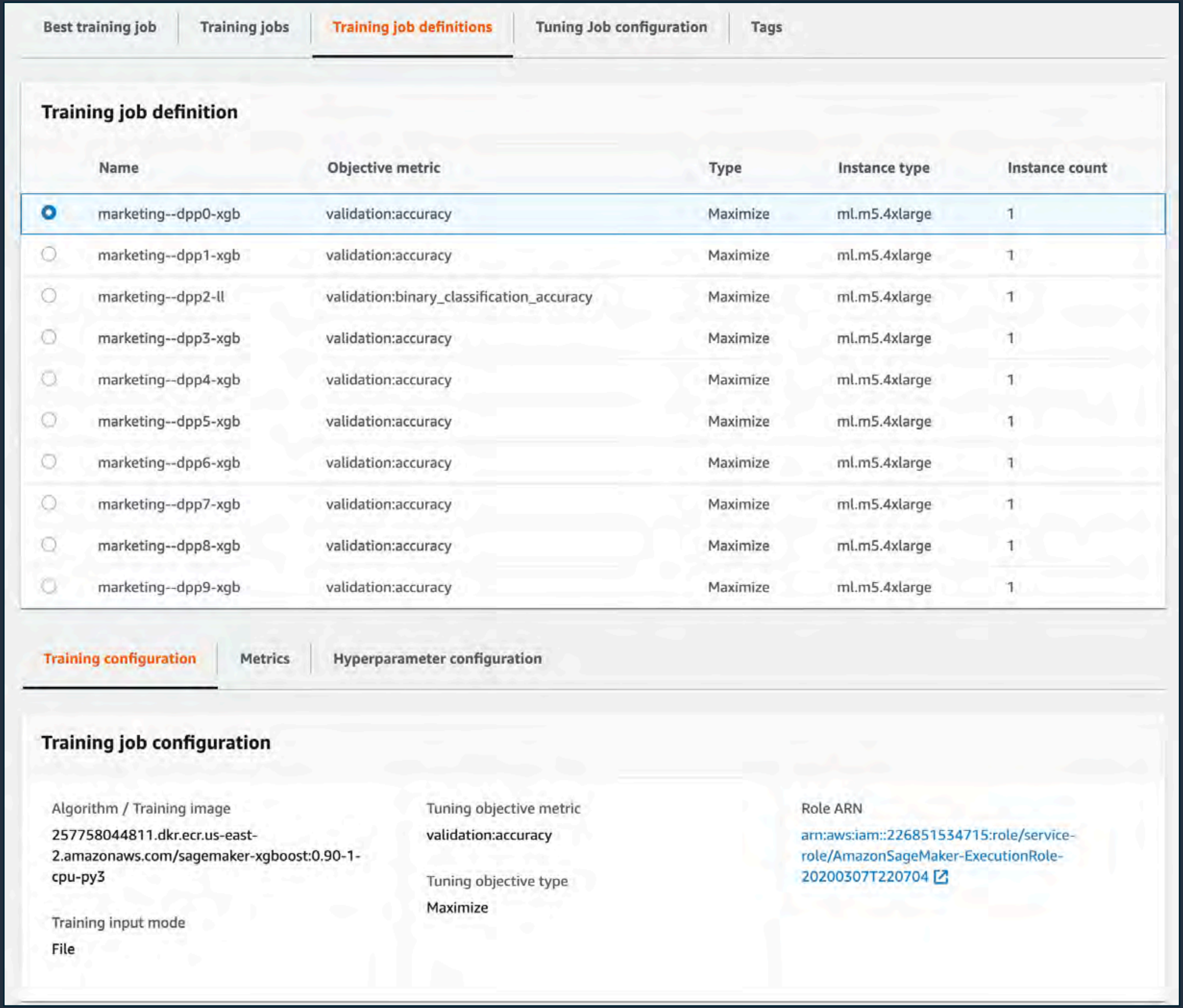


(参考)SageMaker Hyperparameter Tuning Job

Training job

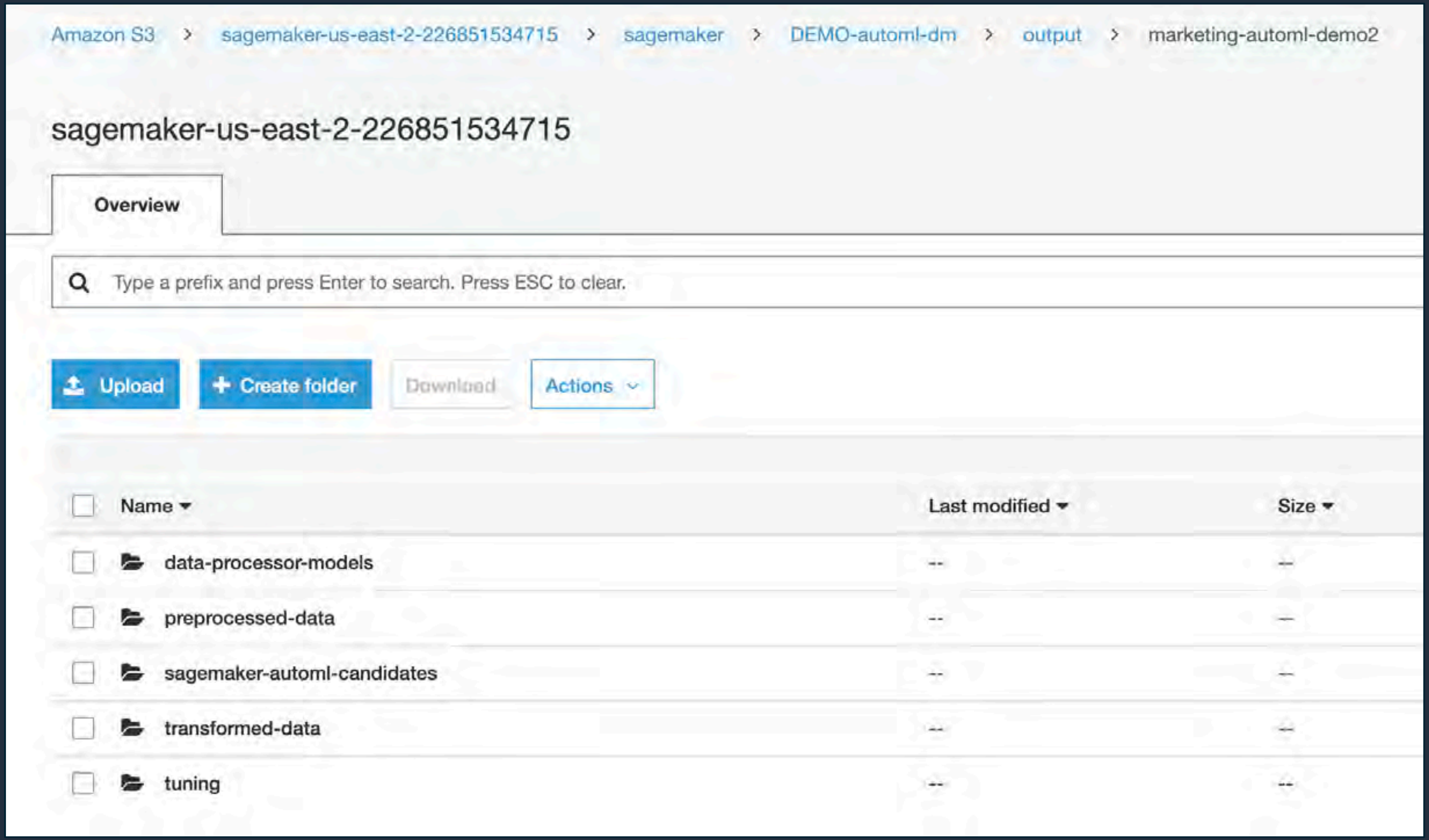


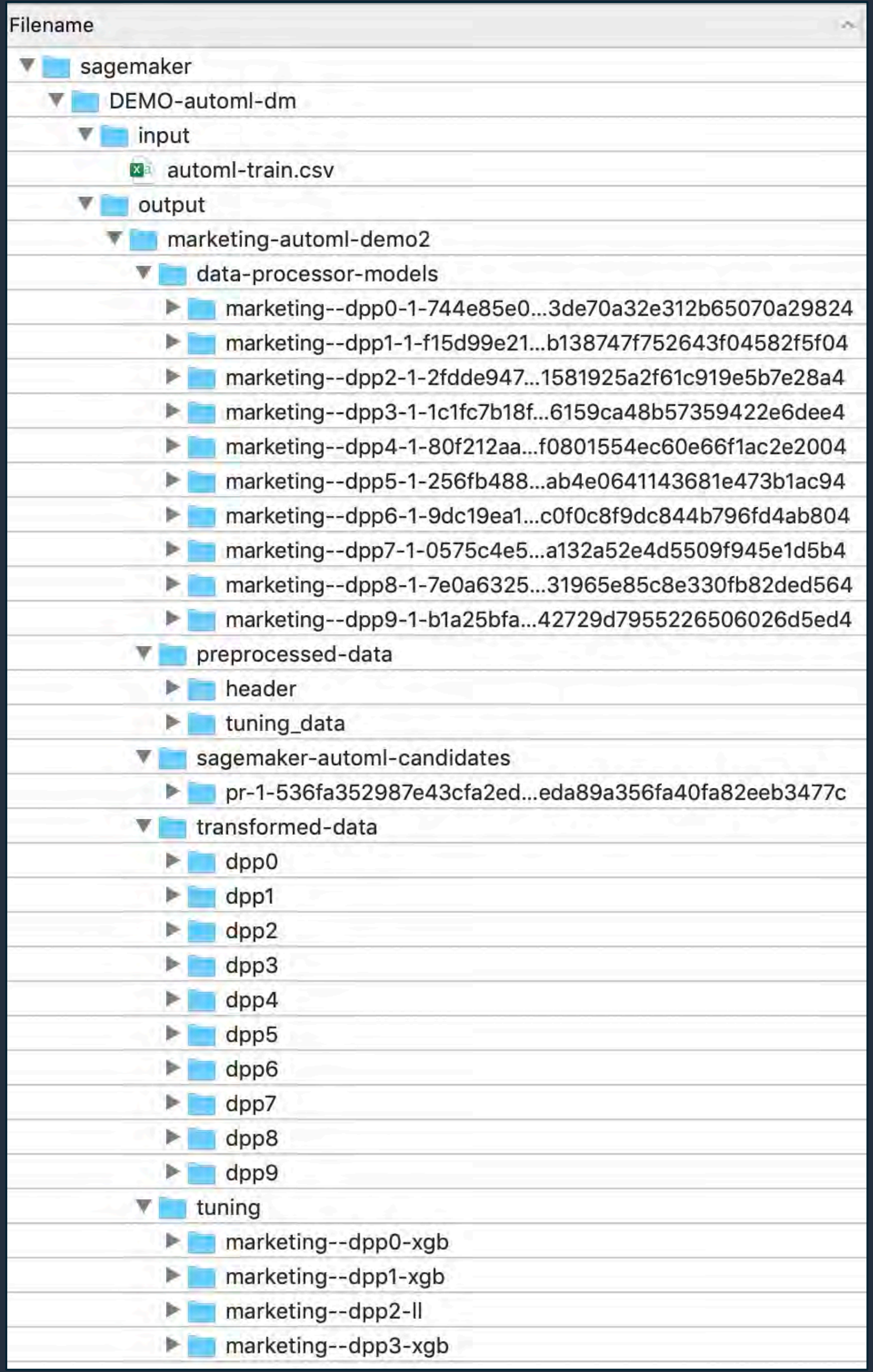
Training job definitions



(参考)Experimentで実行されるスクリプト、データ

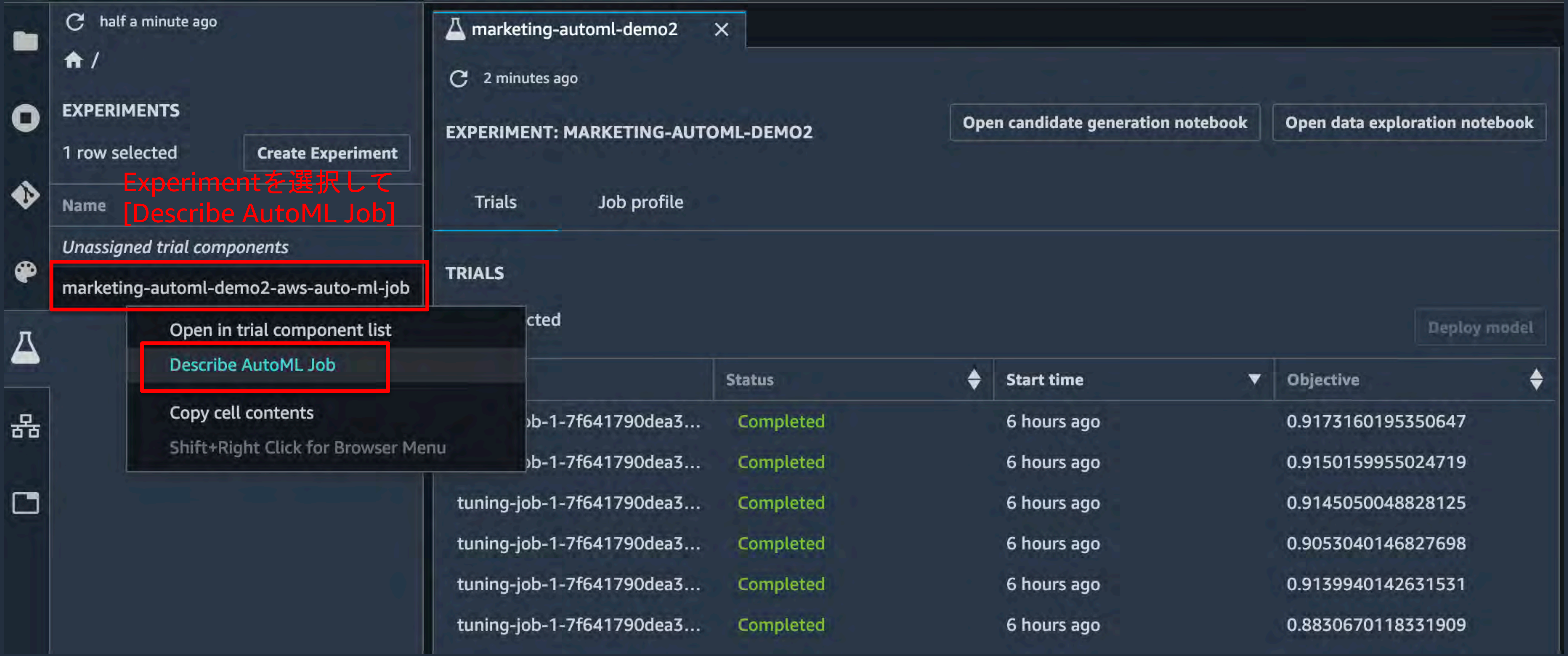
SageMaker Experimentで実行されるスクリプトや、Feature Engineeringにより生成された中間データは全てS3に出力される





Training結果の確認・評価

Training結果の確認
AutoML Jobの確認



Training結果の確認
Jobの結果一覧の表示

EXPERIMENT: MARKETING-AUTOML-DEMO2

Open candidate generation notebook | Open data exploration notebook

Trials | Job profile

TRIALS

0 row selected

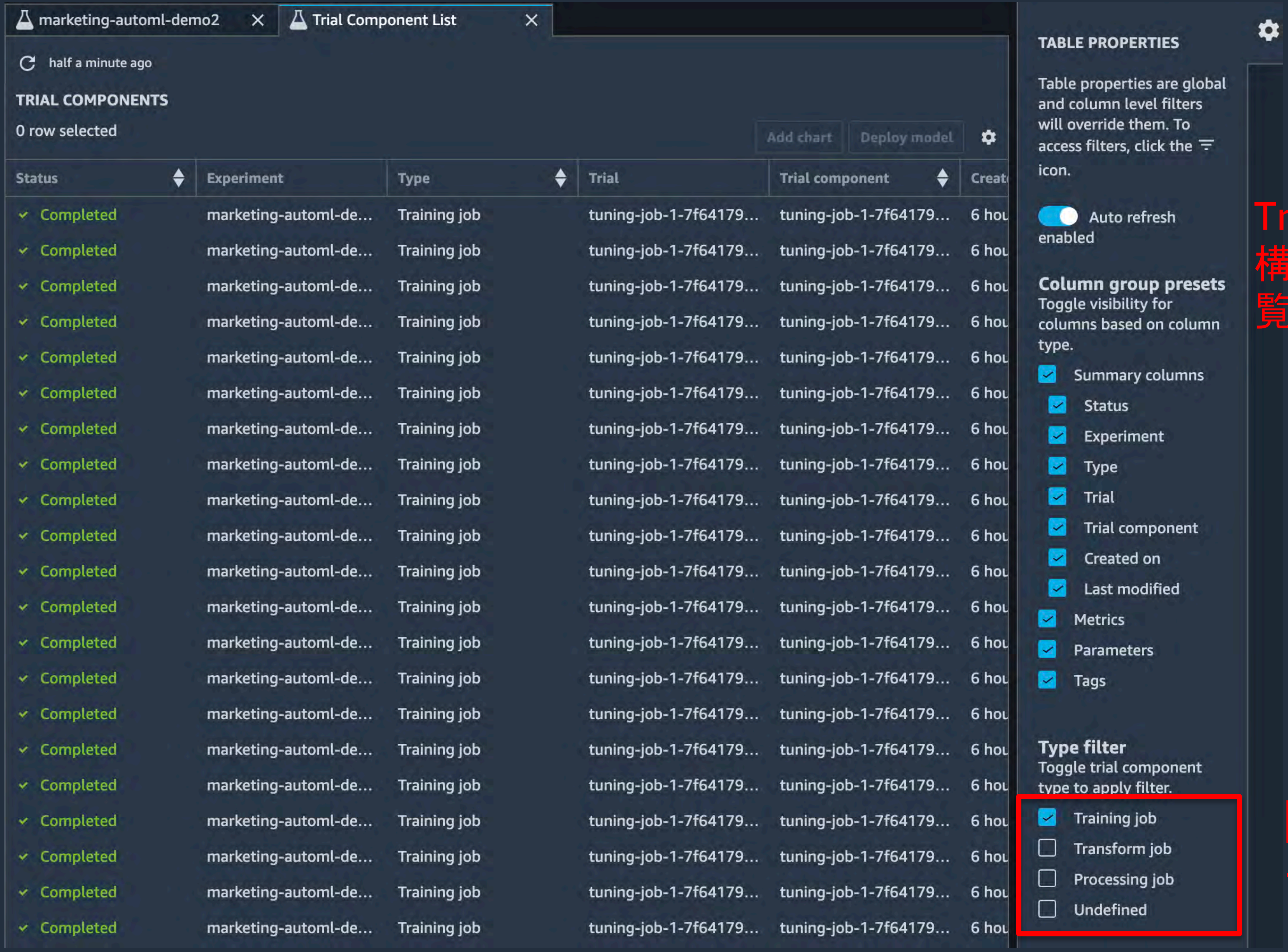
Training job等の結果が表示される

Deploy model

Trial name	Status	Start time	Objective
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9173160195350647
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9150159955024719
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9145050048828125
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9053040146827698
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9139940142631531
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.8830670118331909
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9072200059890747
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9038980007171631
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.8968690037727356
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.8828120231628418
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9101600050926208
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9096490144729614
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9171879887580872
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9079869985580444
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9179549813270569
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.912077009677887
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9077320098876953
tuning-job-1-7f641790dea345...	Completed	8 hours ago	0.9174439907073975

Training結果の確認
Trial componentの確認

Training結果の確認
Trial componentの確認



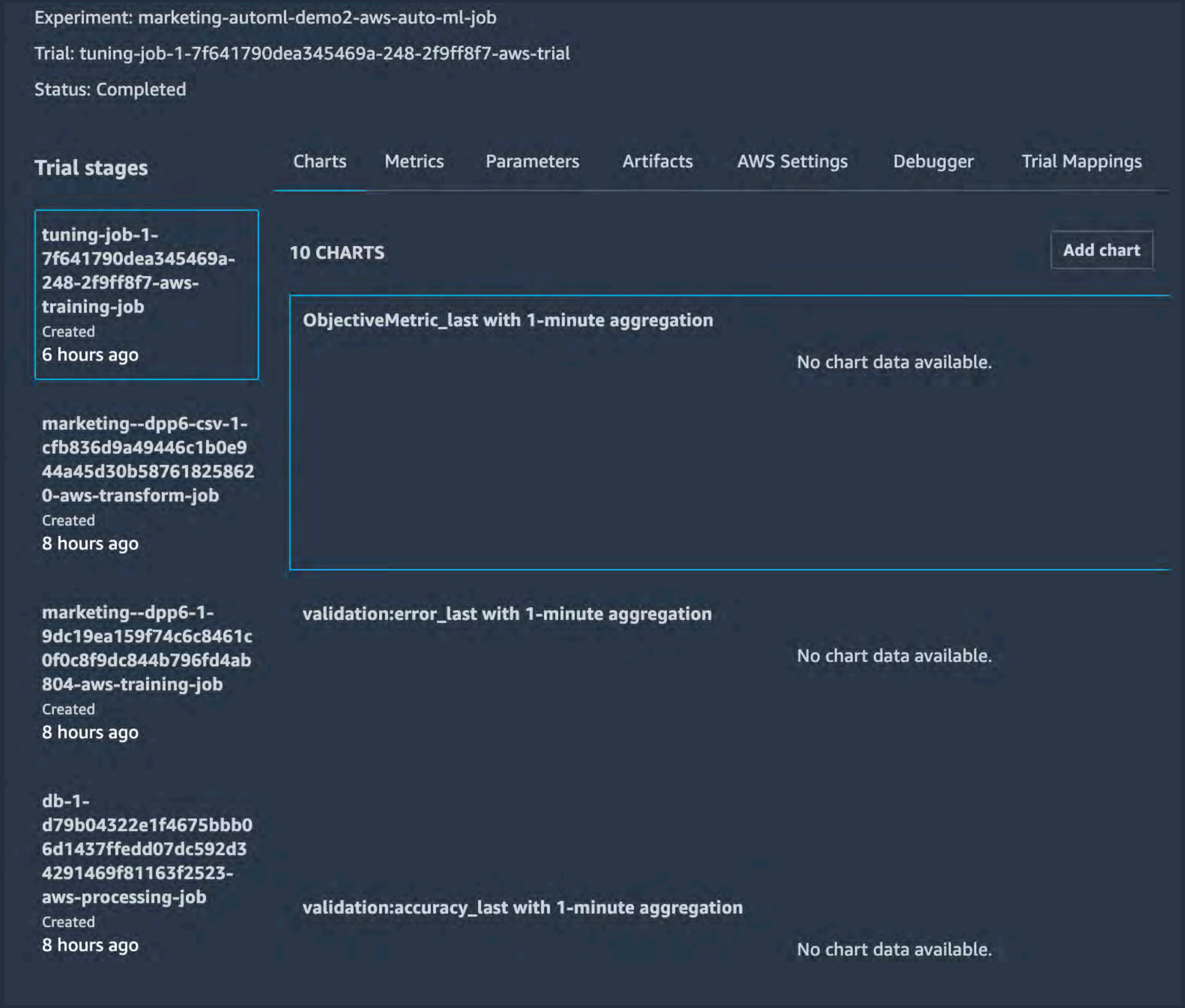
Trial components (1つのジョブを構成する実際の処理ステップ)の一覧が表示される

[Training Job]のみチェック

Training結果の確認
Training jobの詳細確認

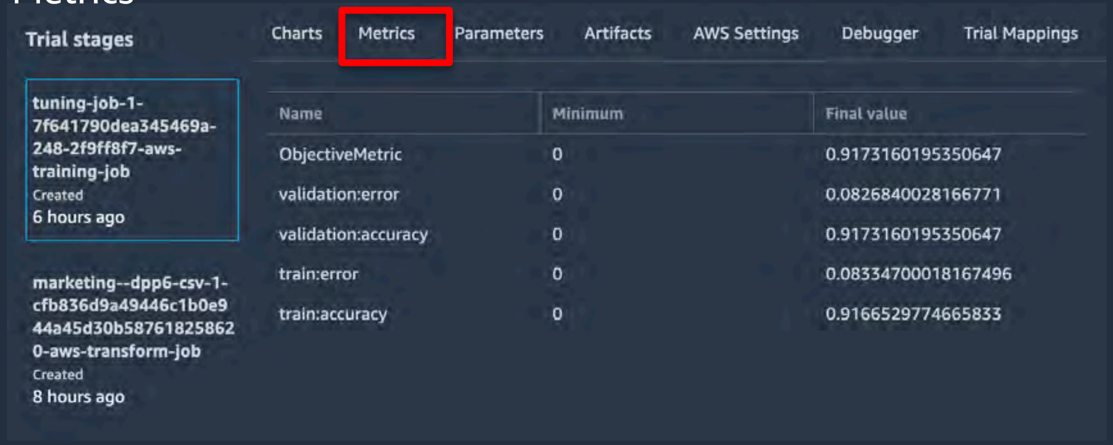
Training結果の確認
Training jobの詳細確認

Trialを構成する
複数のステップが表示される



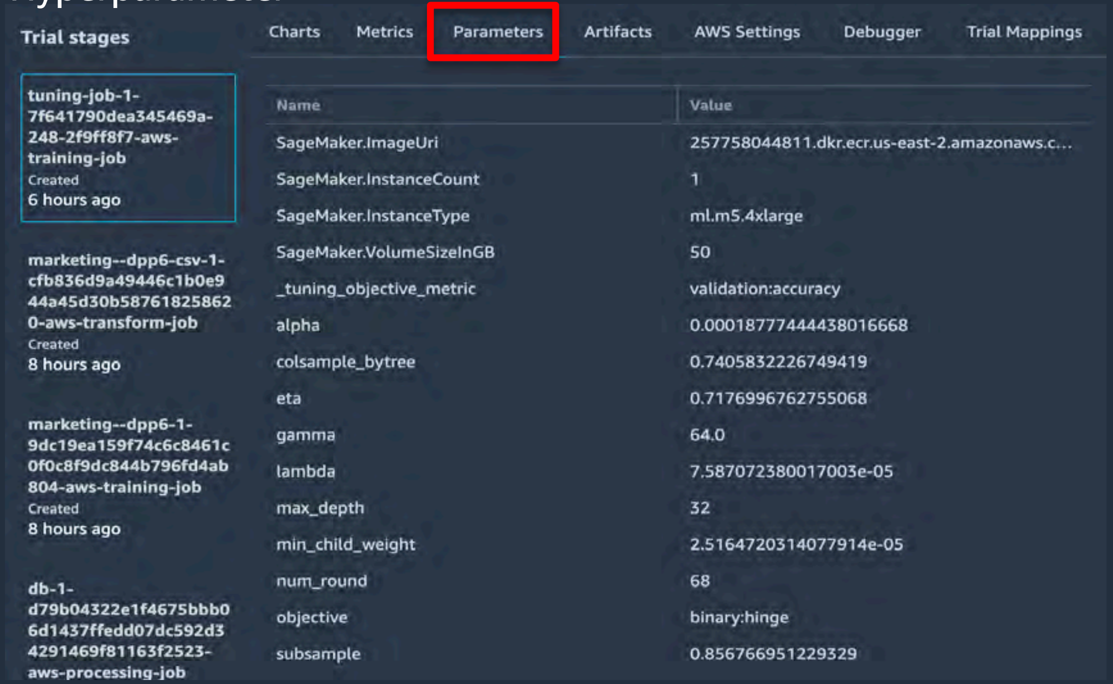
Training結果の確認
Training jobの詳細確認

Metrics



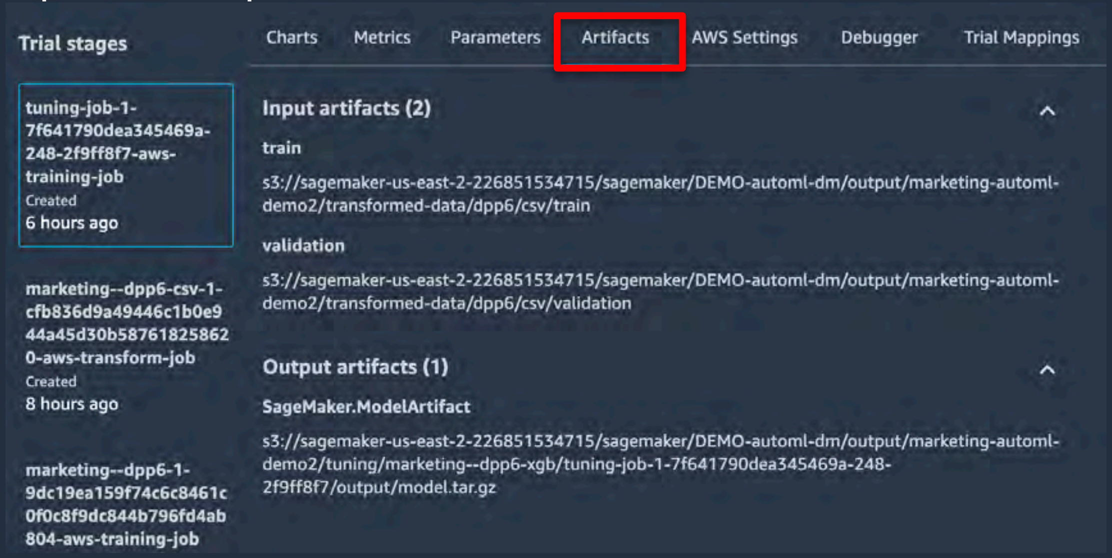
Name	Minimum	Final value
ObjectiveMetric	0	0.9173160195350647
validation:error	0	0.0826840028166771
validation:accuracy	0	0.9173160195350647
train:error	0	0.08334700018167496
train:accuracy	0	0.9166529774665833

Hyperparameter



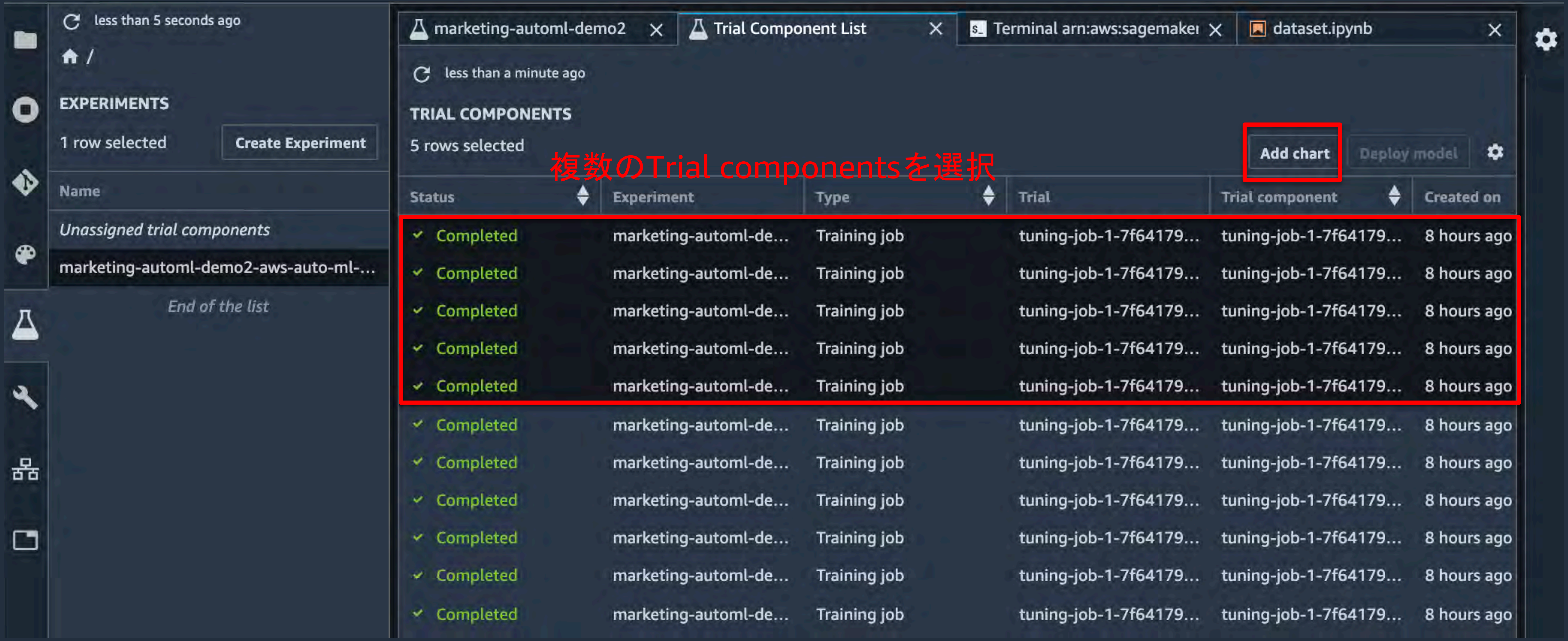
Name	Value
SageMaker.ImageUri	257758044811.dkr.ecr.us-east-2.amazonaws.c...
SageMaker.InstanceCount	1
SageMaker.InstanceType	ml.m5.4xlarge
SageMaker.VolumeSizeInGB	50
_tuning_objective_metric	validation:accuracy
alpha	0.00018777444438016668
colsample_bytree	0.7405832226749419
eta	0.7176996762755068
gamma	64.0
lambda	7.587072380017003e-05
max_depth	32
min_child_weight	2.5164720314077914e-05
num_round	68
objective	binary:hinge
subsample	0.856766951229329

Input and Output Artifacts

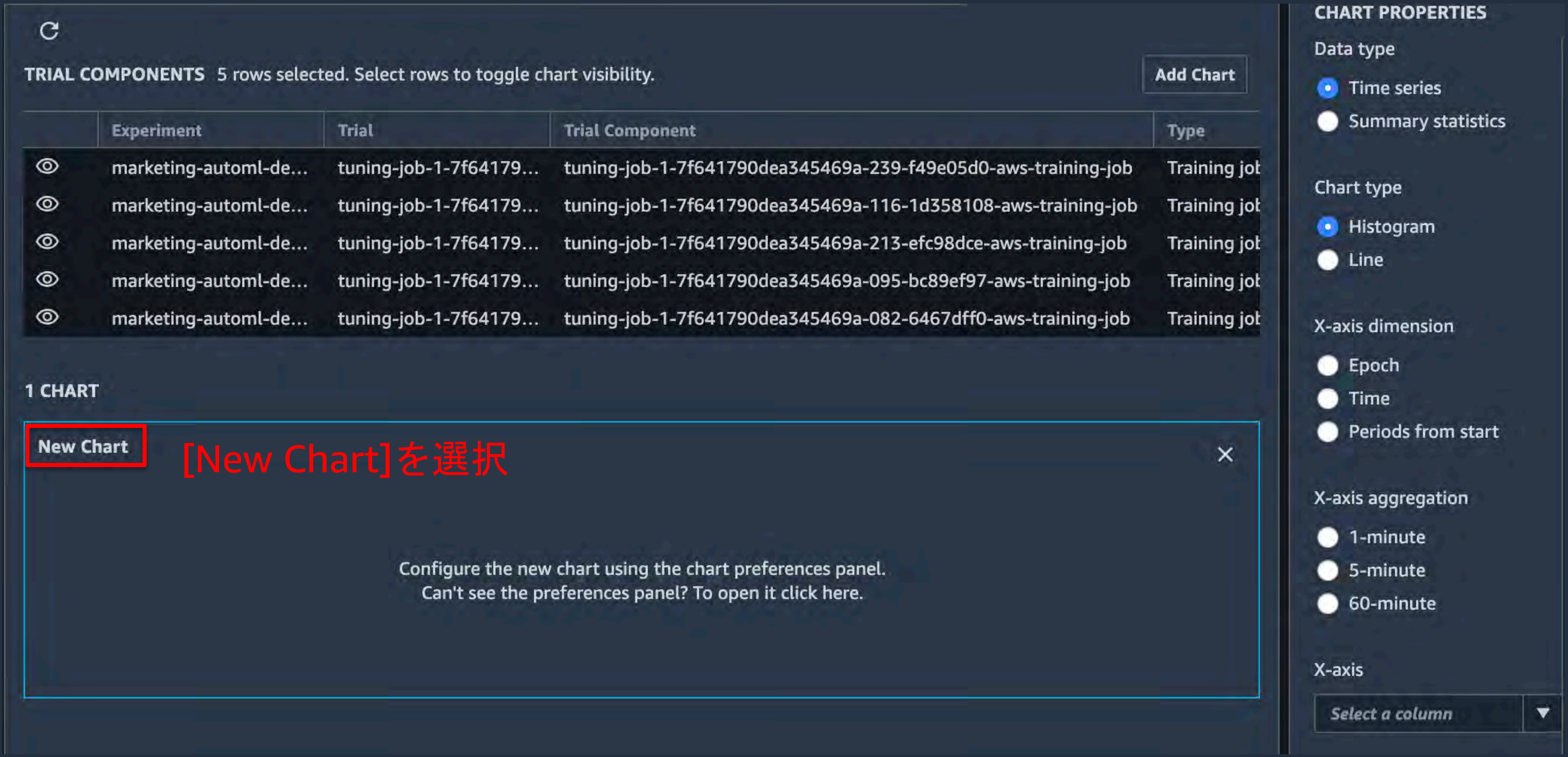


Training結果の比較
Chartの作成

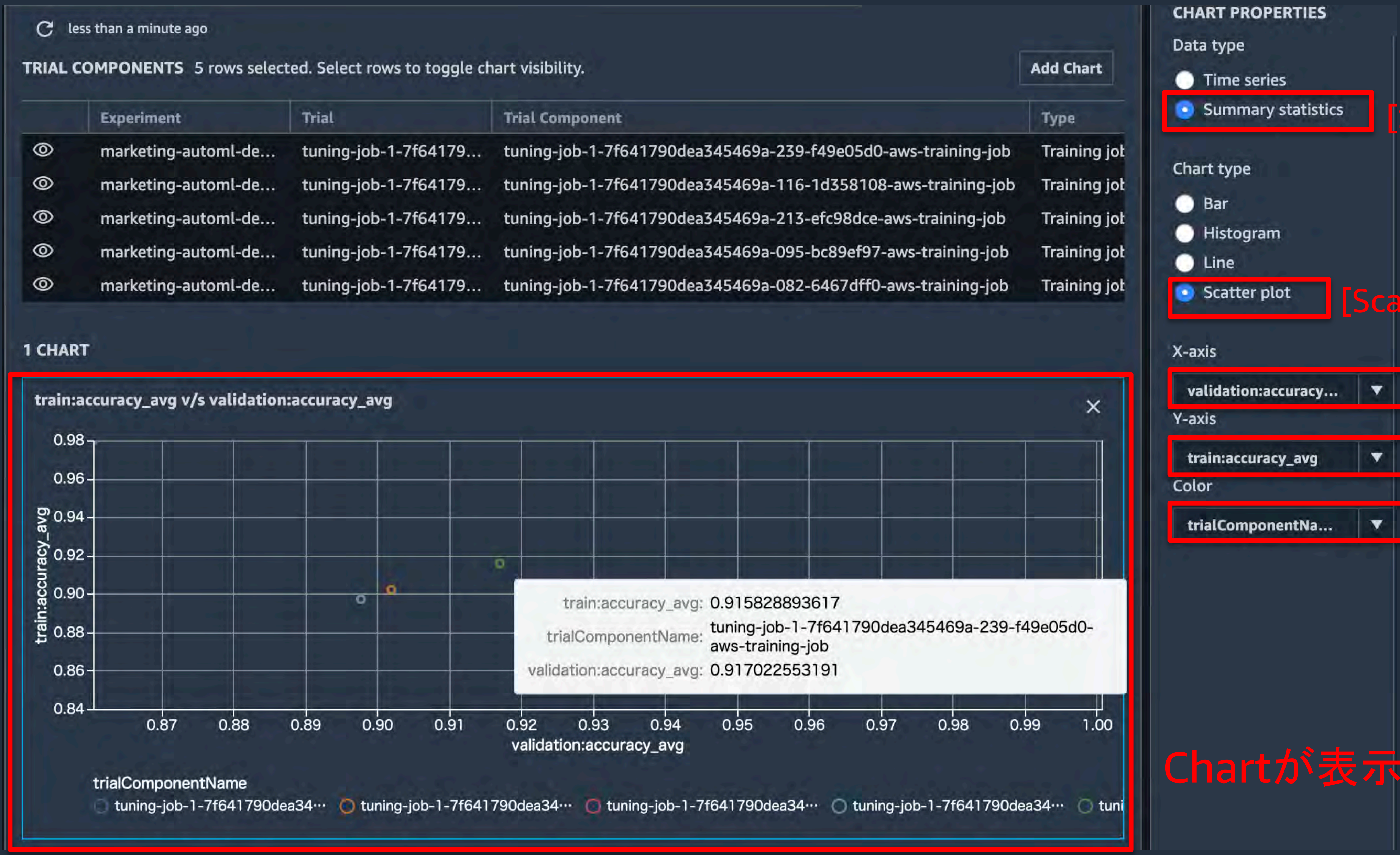
複数のTrialをChartで比較し、最も"良い"モデルを選択する



Training結果の比較
Chartの作成



Training結果の比較
Chartの作成



[Summary statistics]

[Scatter plot]

[validation:accuracy_avg]

[train:accuracy_avg]

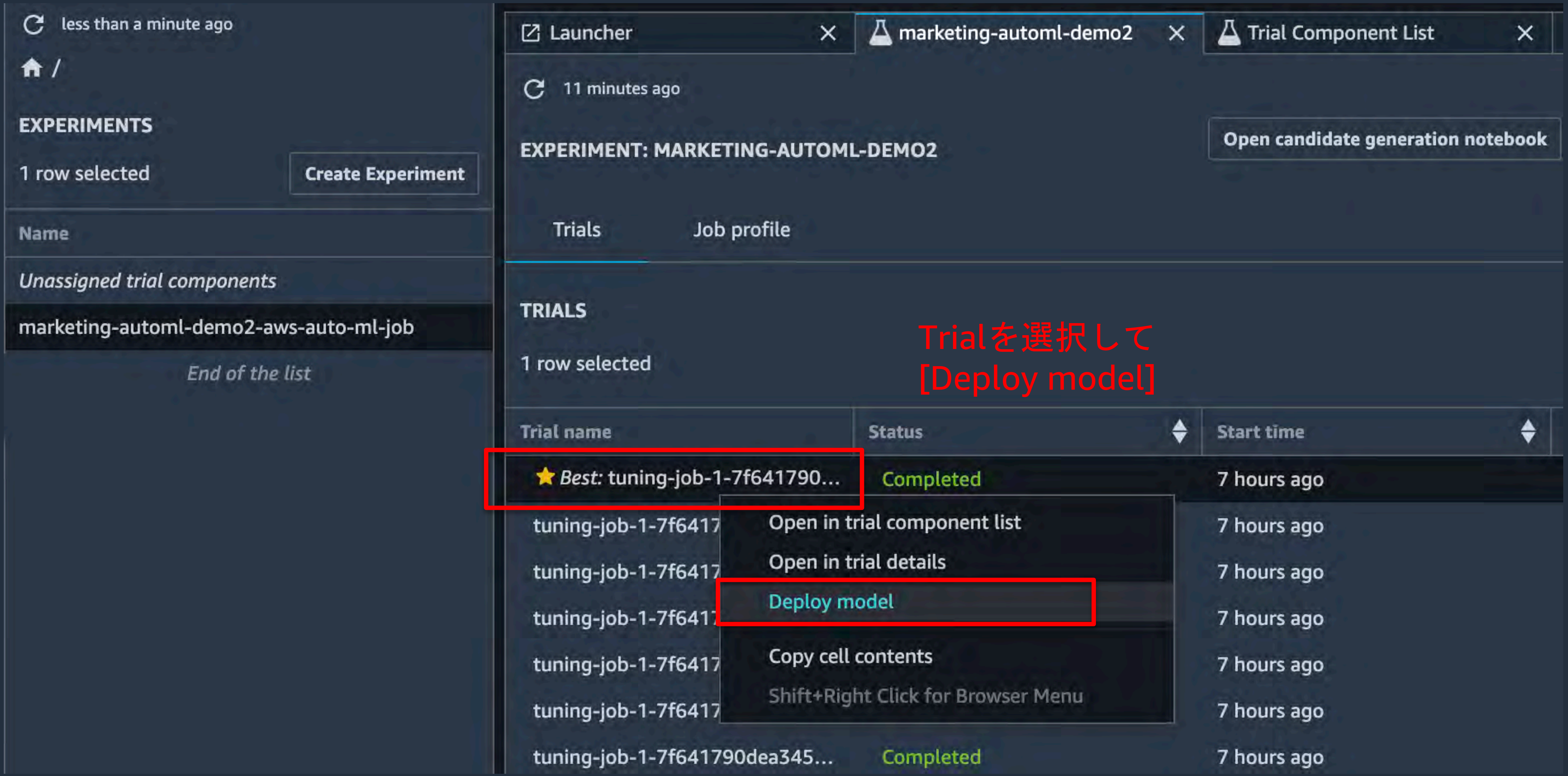
[trialComponentName]

Chartが表示される

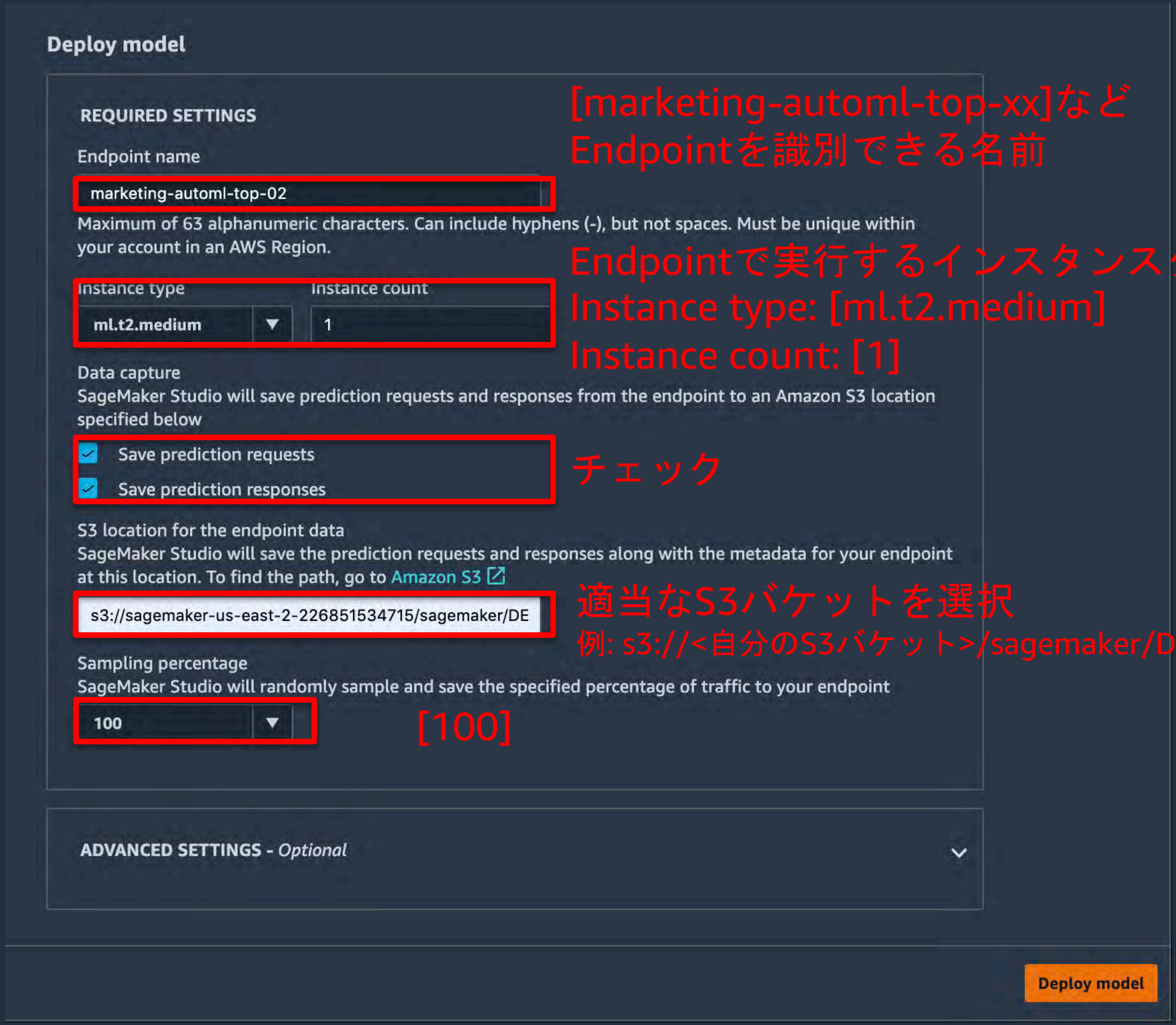
モデルのDeploy

モデルのDeploy

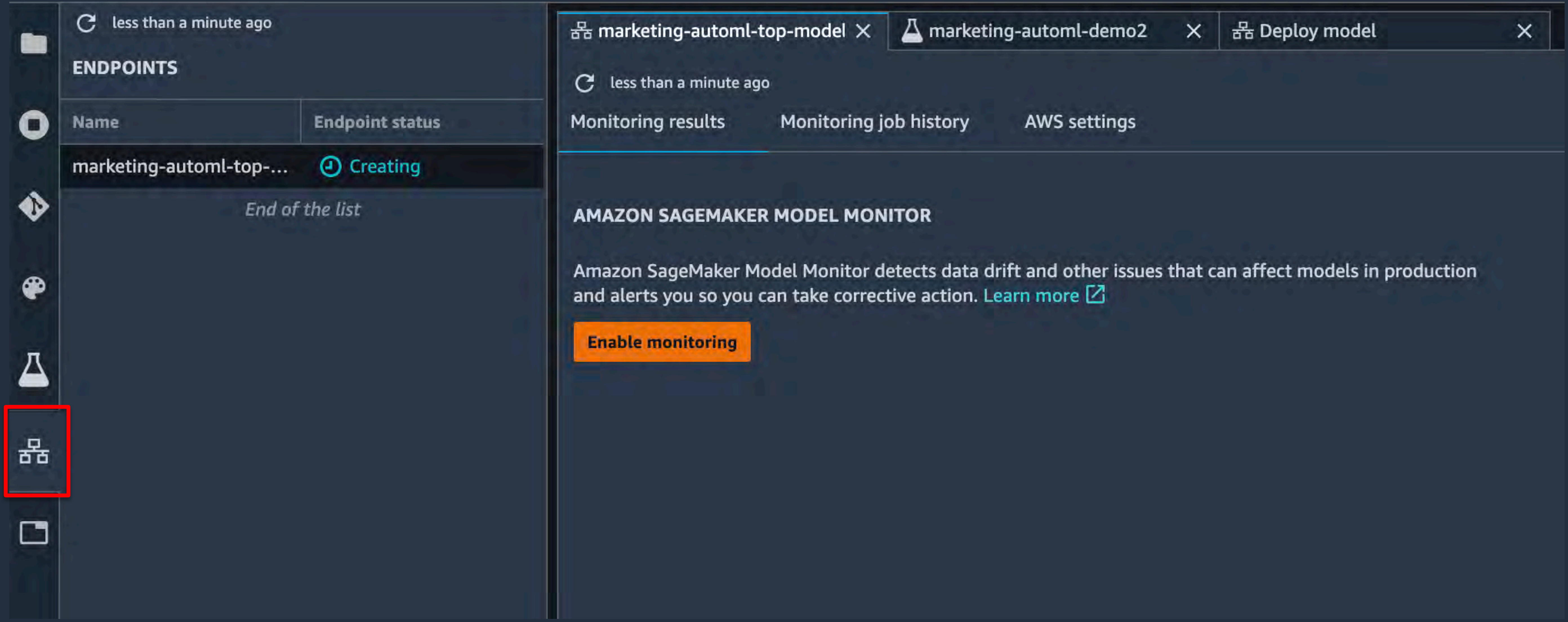
作成した推論モデルをDeployし、APIから利用できるEndpointを作成する



モデルのDeploy
Endpoint設定とモデルのデプロイ

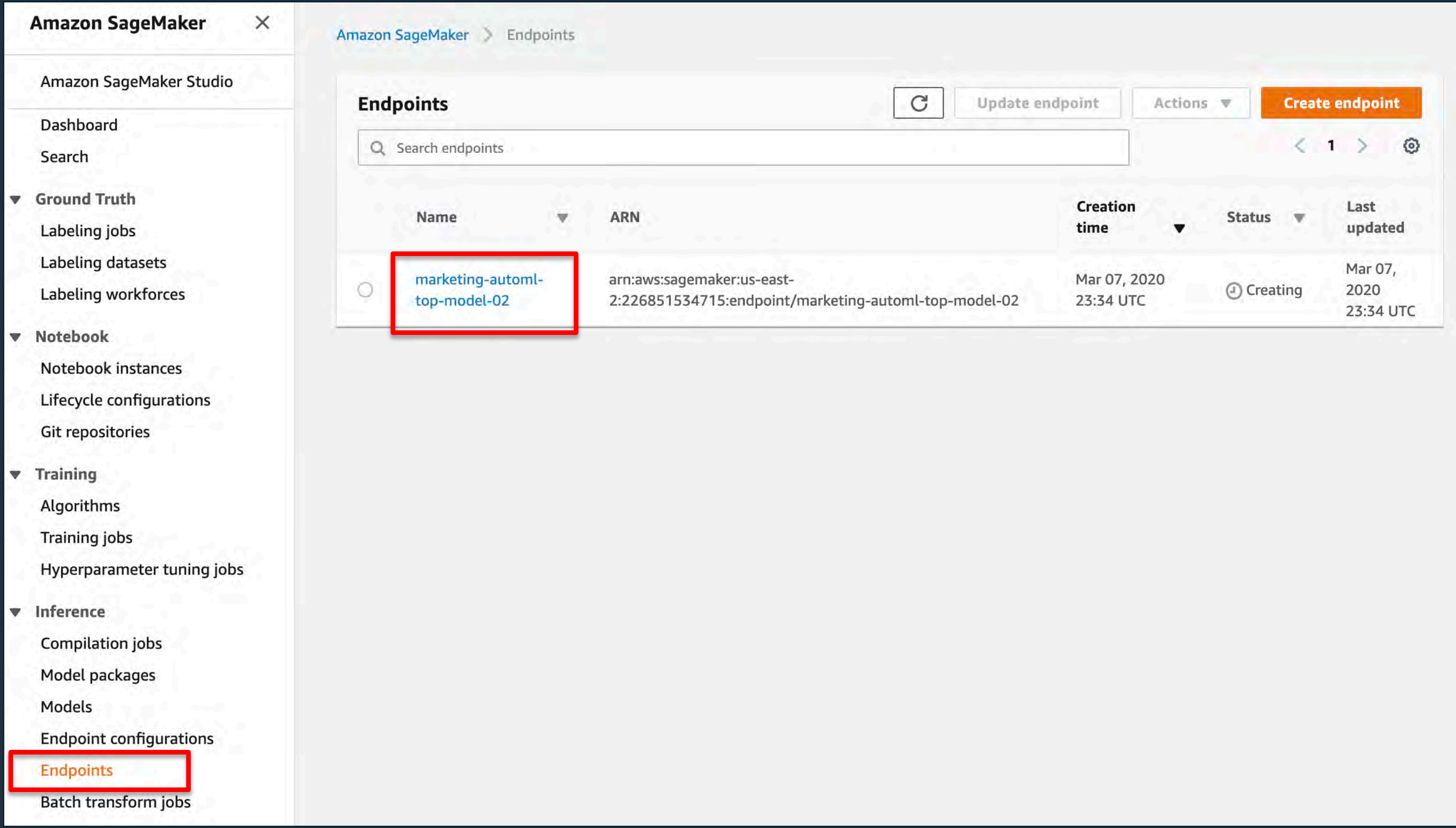


モデルのDeploy
Endpointの確認



(参考)SageMakerコンソール Endpointの確認

- Endpoint作成は、内部でSageMakerのEndpoint作成を呼び出し実行している
- SageMakerコンソールからも確認できる



(参考)SageMakerコンソール Endpointの確認

Amazon SageMaker > Endpoints > marketing-automl-top-model-02

marketing-automl-top-model-02

Delete

Endpoint settings

Name
marketing-automl-top-model-02

ARN
arn:aws:sagemaker:us-east-2:226851534715:endpoint/marketing-automl-top-model-02

Status
Creating

Creation time
Sun Mar 08 2020 08:34:28 GMT+0900 (Japan Standard Time)

Last updated
Sun Mar 08 2020 08:34:28 GMT+0900 (Japan Standard Time)

URL
https://runtime.sagemaker.us-east-2.amazonaws.com/endpoints/marketing-automl-top-model-02/invocations
Learn more about the API

Data capture settings

Enable data capture
Yes

Data capture status
Started

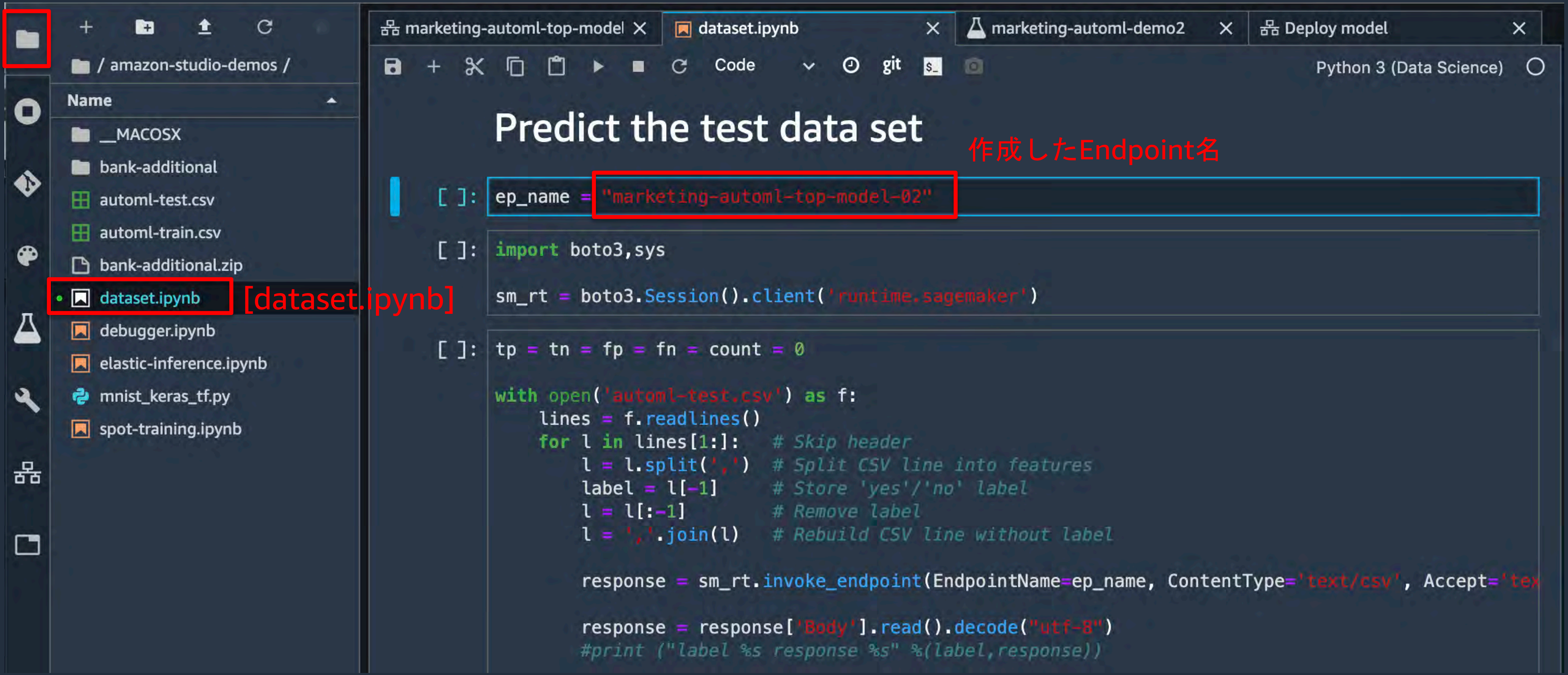
Current sampling percentage (%)
10

S3 location to store data collected
s3://sagemaker-us-east-2-226851534715/sagemaker/DEMO-automl-dm/monitor/

Endpointへのリクエスト

Endpointへのリクエスト

作成したEndpointにリクエストを実行し、結果を確認する



Endpointへのリクエスト
リクエストの実行

```
[9]: tp = tn = fp = fn = count = 0

     with open('automl-test.csv') as f:
         lines = f.readlines()
         for l in lines[1:]:   # Skip header
             l = l.split(',')  # Split CSV line into features
             label = l[-1]     # Store 'yes'/'no' label
             l = l[:-1]        # Remove label
             l = ','.join(l)   # Rebuild CSV line without label

             response = sm_rt.invoke_endpoint(EndpointName=ep_name, ContentType='text/csv', Accept='text/csv', Body=l)

             response = response['Body'].read().decode("utf-8")
             #print ("label %s response %s" %(label,response))

             if 'yes' in label:
                 # Sample is positive
                 if 'yes' in response:
                     # True positive
                     tp=tp+1
                 else:
                     # False negative
                     fn=fn+1
             else:
                 # Sample is negative
                 if 'no' in response:
                     # True negative
                     tn=tn+1
                 else:
                     # False positive
                     fp=fp+1
             count = count+1
             if (count % 100 == 0):
                 sys.stdout.write(str(count)+' ')

     print ("Done")

     100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000 Done

[10]: print ("%d %d" % (tn, fp))
      print ("%d %d" % (fn, tp))

      accuracy  = (tp+tn)/(tp+tn+fp+fn)
      precision = tp/(tp+fp)
      recall    = tn/(tn+fn)
      f1        = (2*precision*recall)/(precision+recall)

      print ("%.4f %.4f %.4f %.4f" % (accuracy, precision, recall, f1))

      1747 75
      91 147
      0.9194 0.6622 0.9505 0.7806
```

Endpointにtext/csvで
リクエストを送信

True Positive, False Negativeを集計

True Negative, False Positiveを集計

処理レコード数を出力

Accuracy, Precision, Recall, F1値を出力

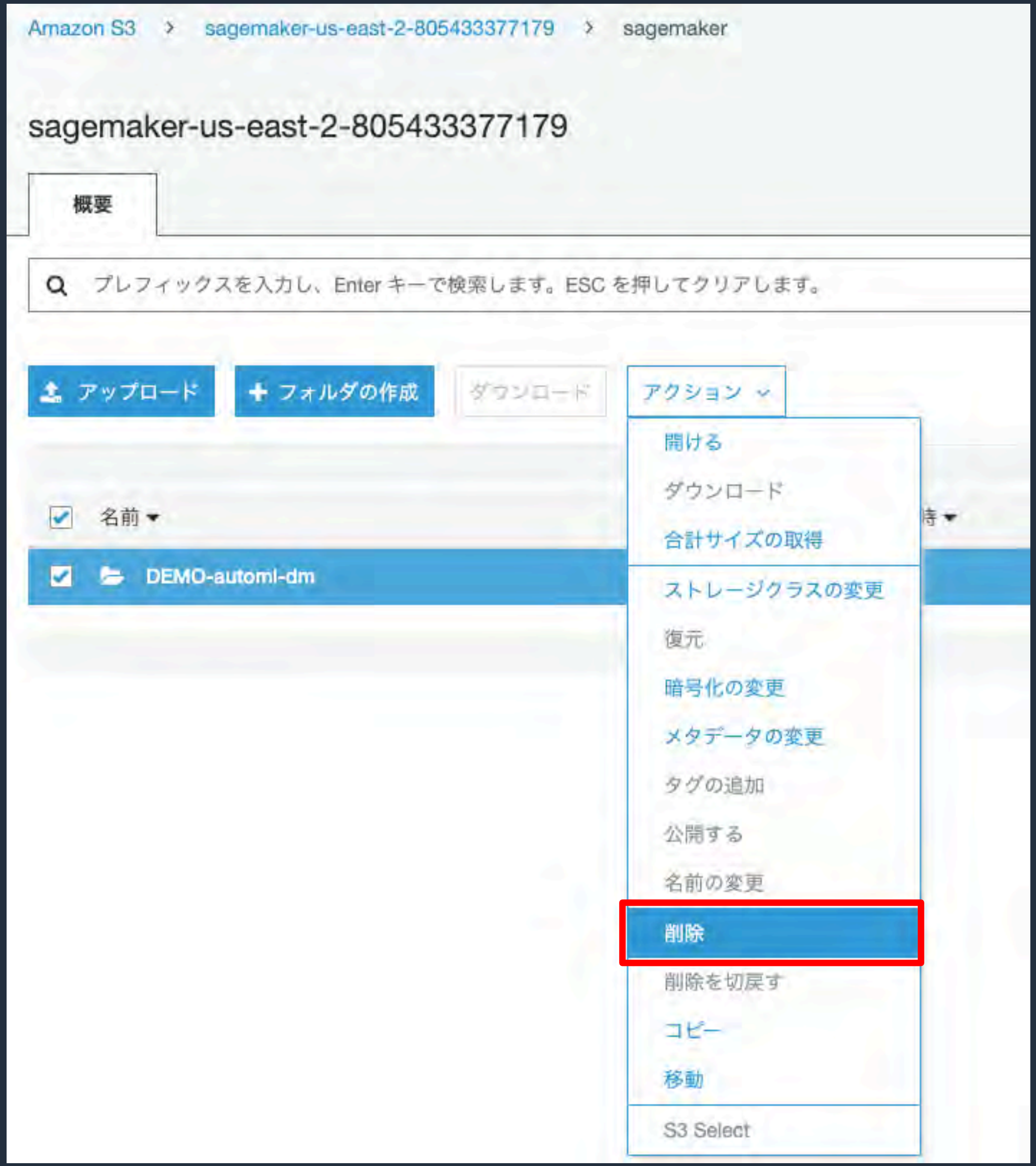
後片付け

削除：S3バケット(Experiments)

Upload data set to Amazon S3

```
[4]: import sagemaker

     prefix = 'sagemaker/DEMO-automl-dm/input'
     sess   = sagemaker.Session()

     uri = sess.upload_data(path="automl-train.csv", key_prefix=prefix)
     print(uri)

     s3://sagemaker-us-east-2-805433377179/sagemaker/DEMO-automl-dm/input/automl-train.csv
```

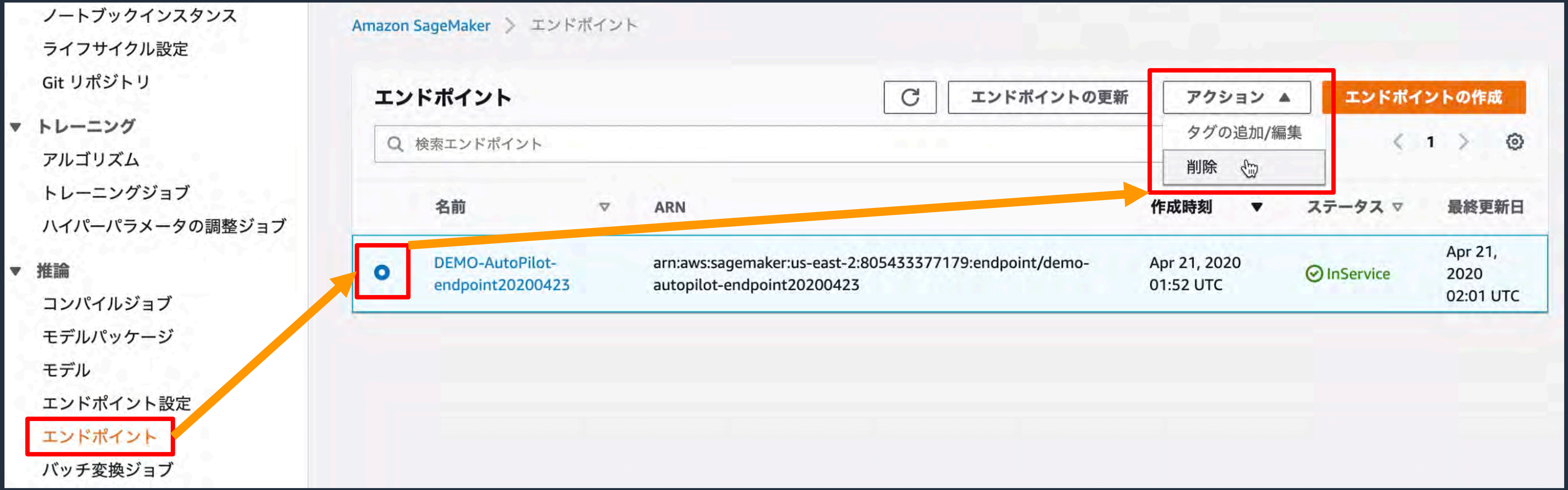
Experimentのinput/outputを格納ディレクトリ



削除：S3バケット(エンドポイント)

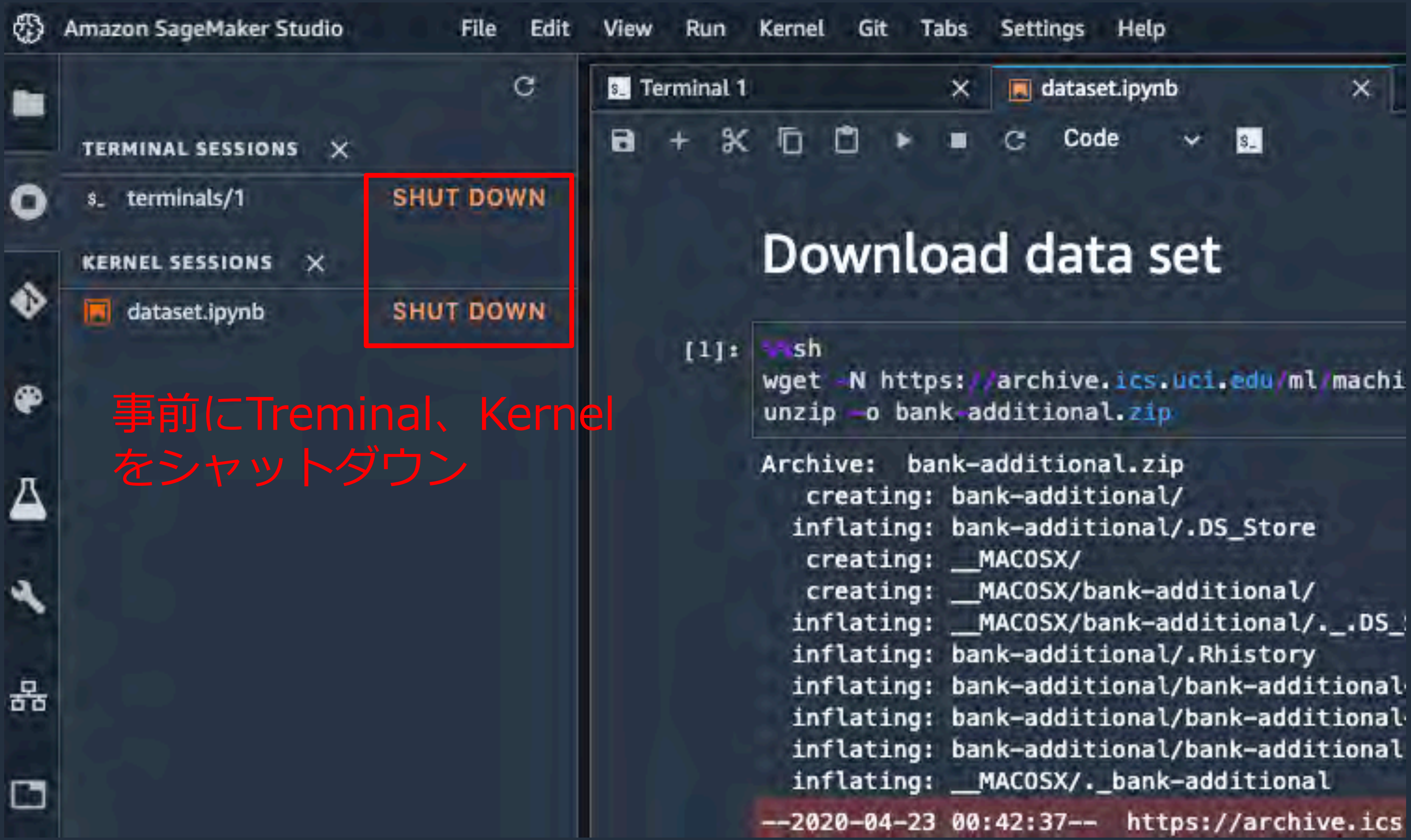
バケットを空にした後、バケットの削除

削除：エンドポント

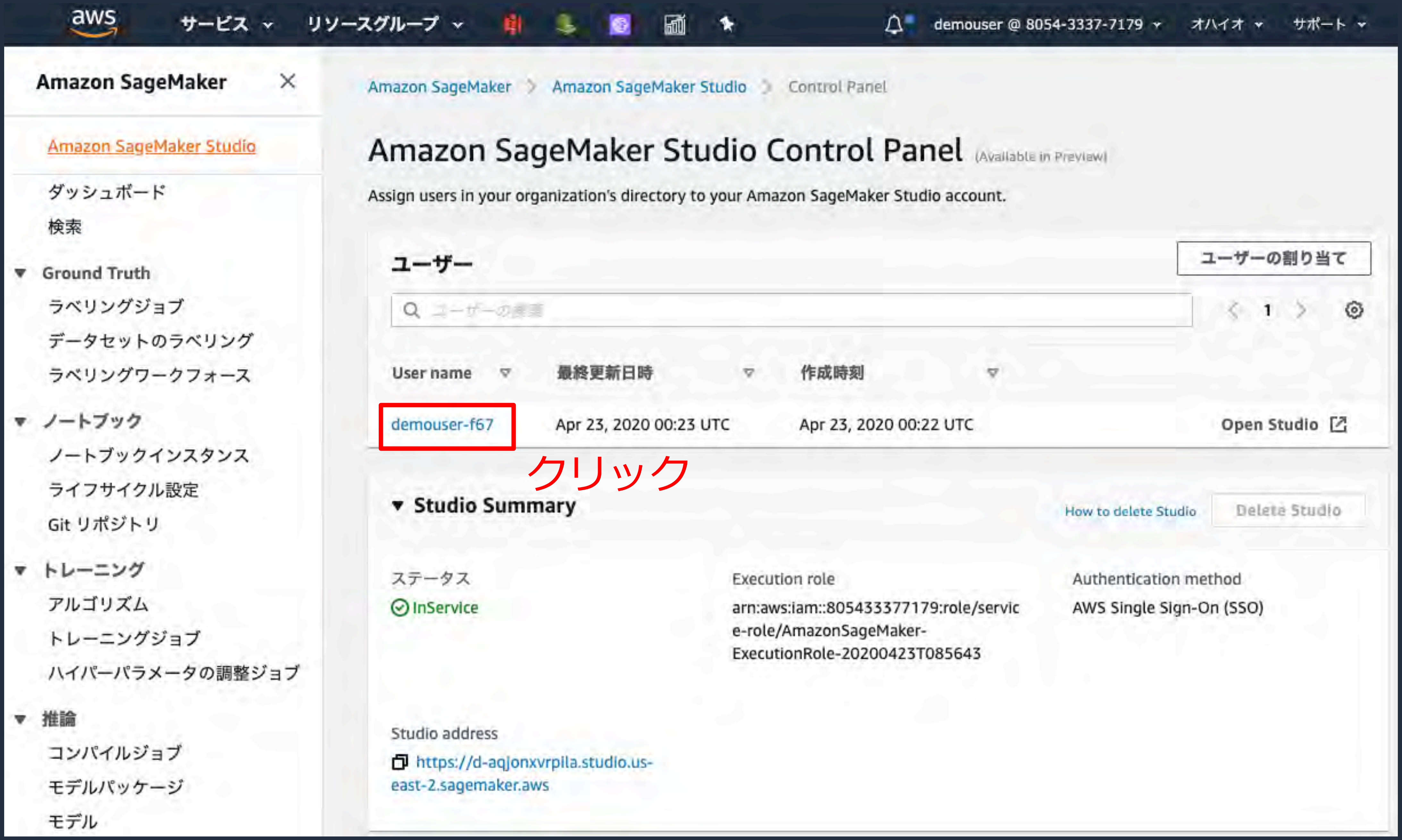


ノートブックインスタンス
ライフサイクル設定
Git リポジトリ
トレーニング
アルゴリズム
トレーニングジョブ
ハイパーパラメータの調整ジョブ
推論
コンパイルジョブ
モデルパッケージ
モデル
エンドポイント設定
エンドポイント
バッチ変換ジョブ
Amazon SageMaker > エンドポイント
エンドポイント
エンドポイントの更新
アクション
エンドポイントの作成
タグの追加/編集
削除
検索エンドポイント
1
名前
ARN
作成時刻
ステータス
最終更新日
DEMO-AutoPilot-endpoint20200423
arn:aws:sagemaker:us-east-2:805433377179:endpoint/demo-autopilot-endpoint20200423
Apr 21, 2020 01:52 UTC
InService
Apr 21, 2020 02:01 UTC

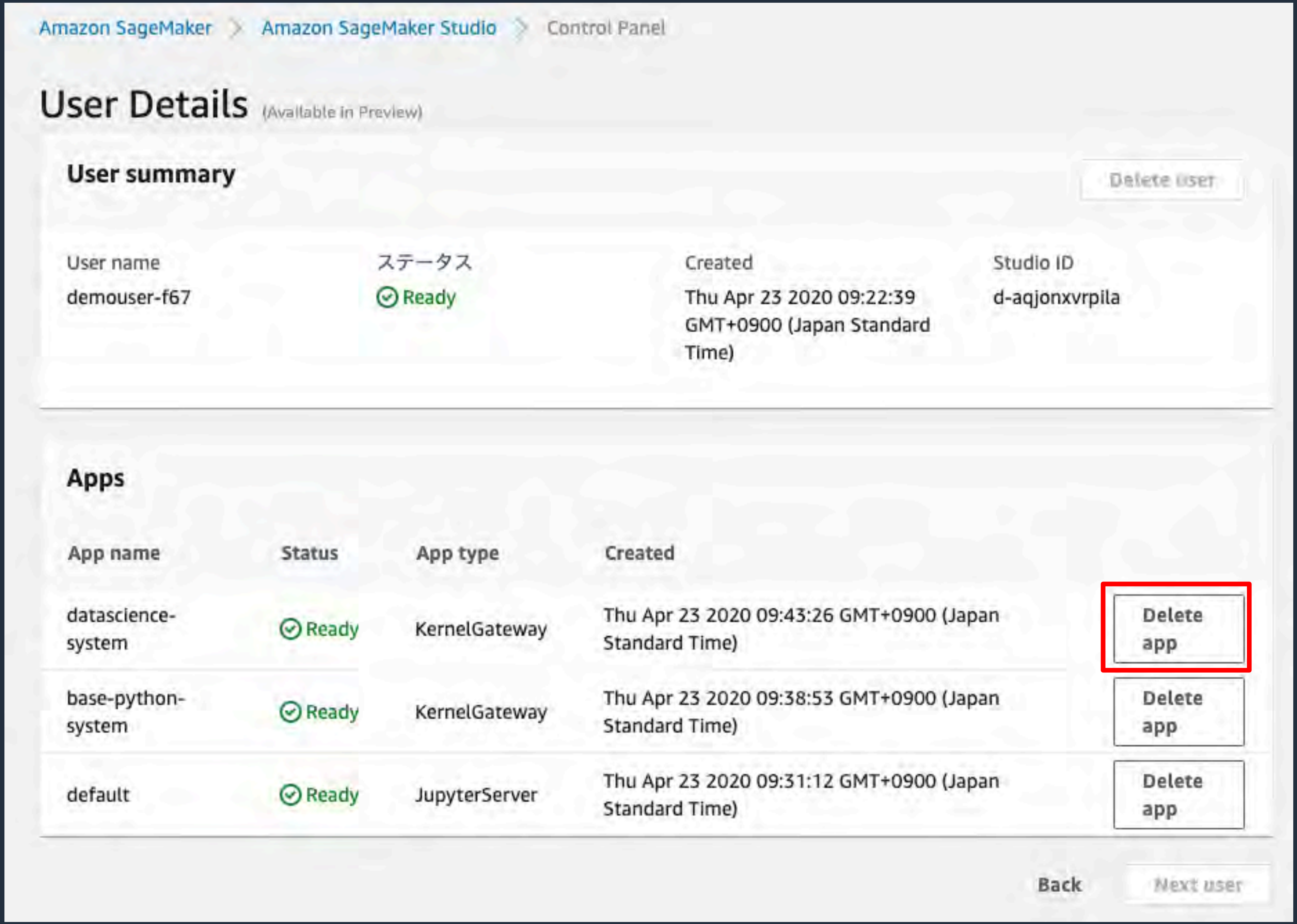
削除：SageMaker Studio ドメイン



削除：SageMaker Studio ドメイン

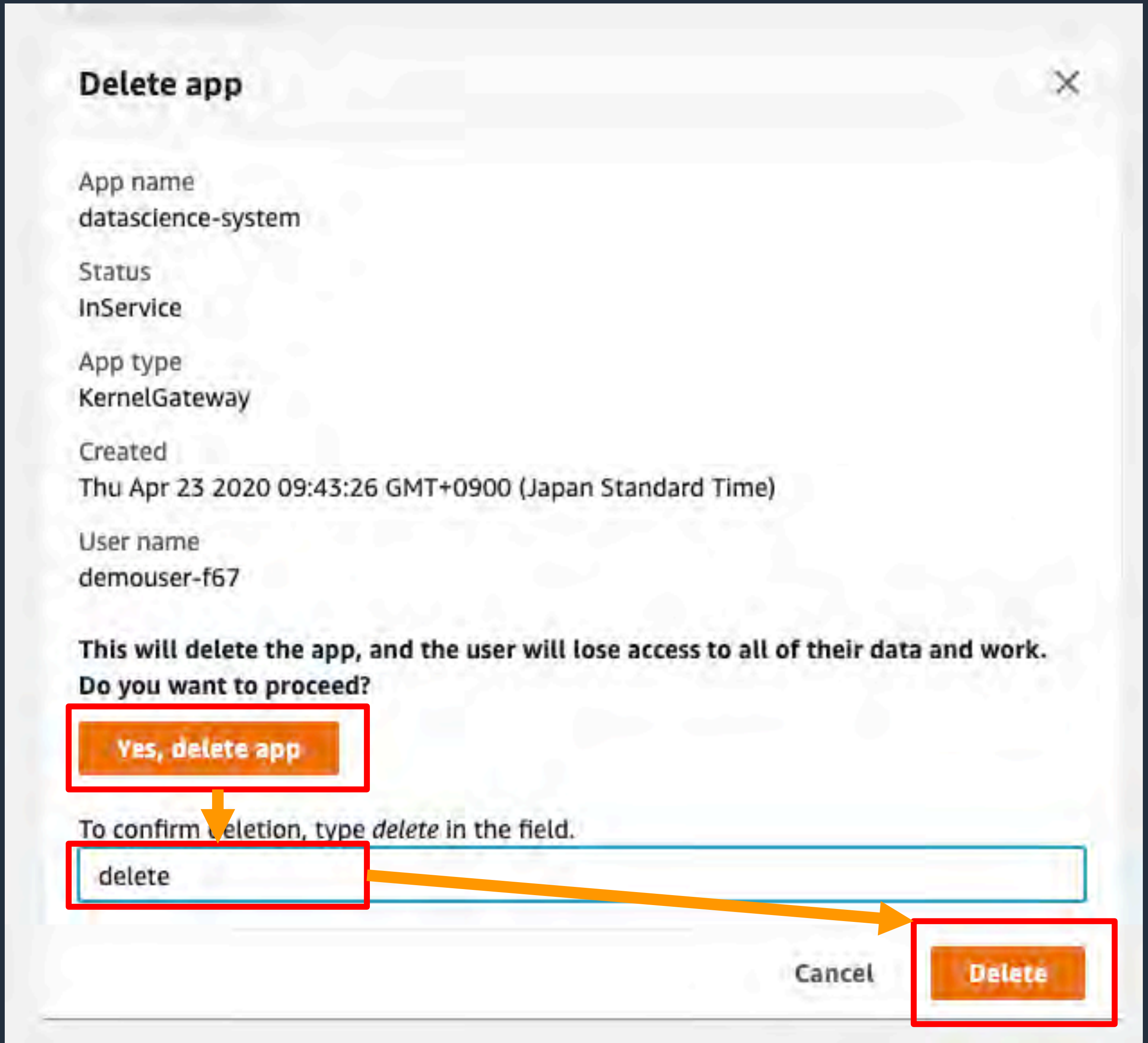


削除：SageMaker Studio ドメイン

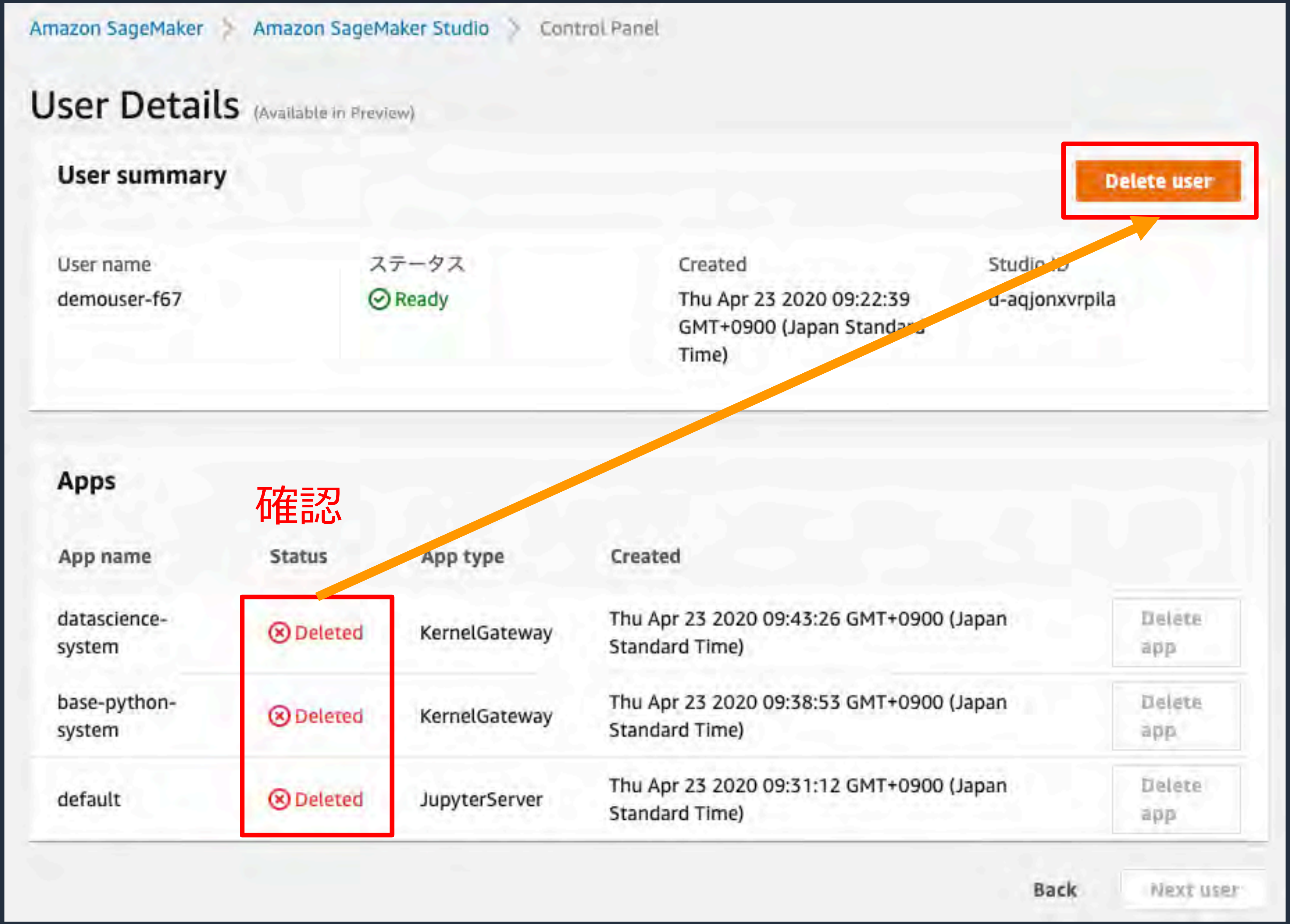


アプリケーションを
削除していく

削除：SageMaker Studio ドメイン

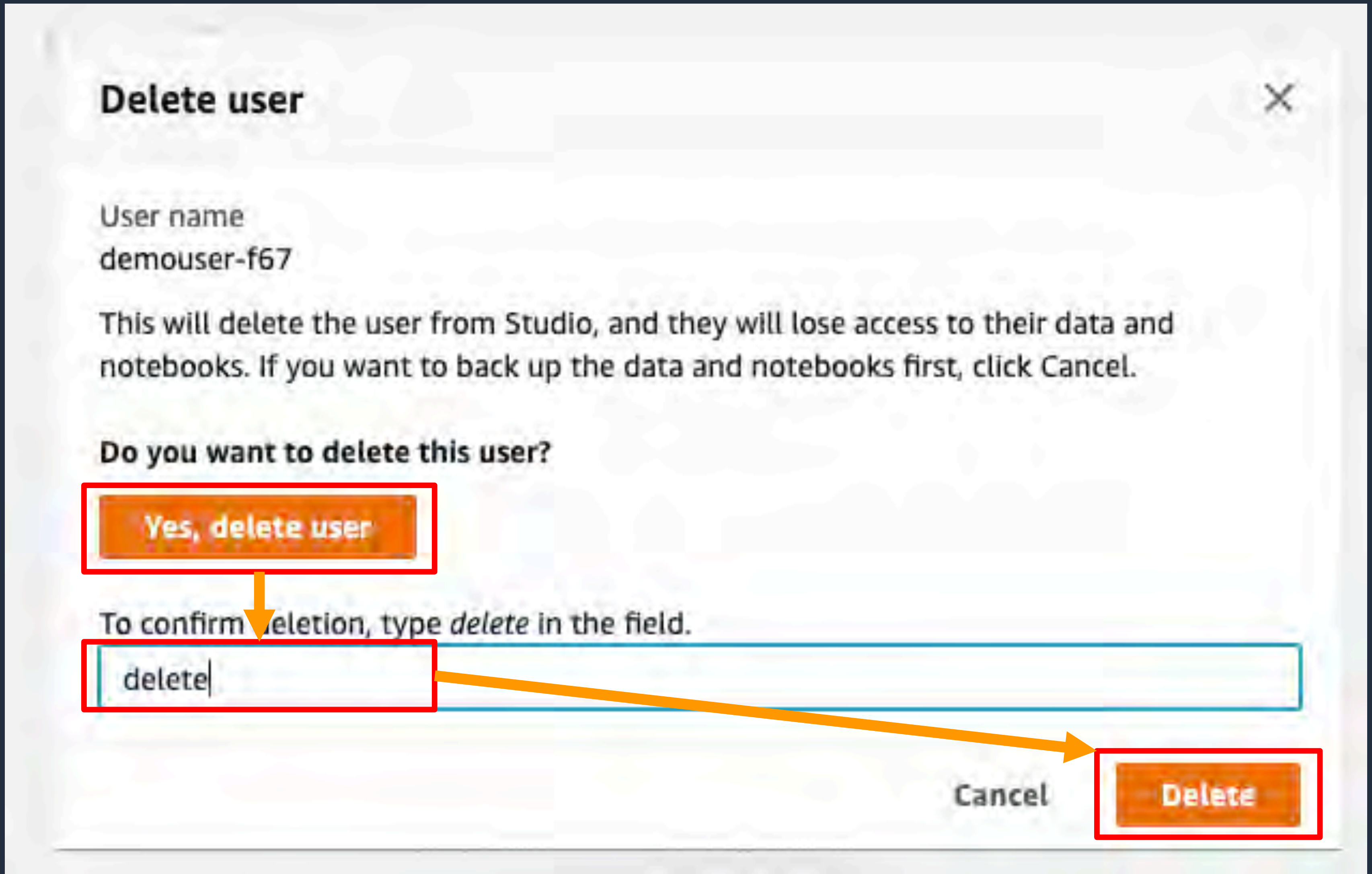


削除：SageMaker Studio ドメイン

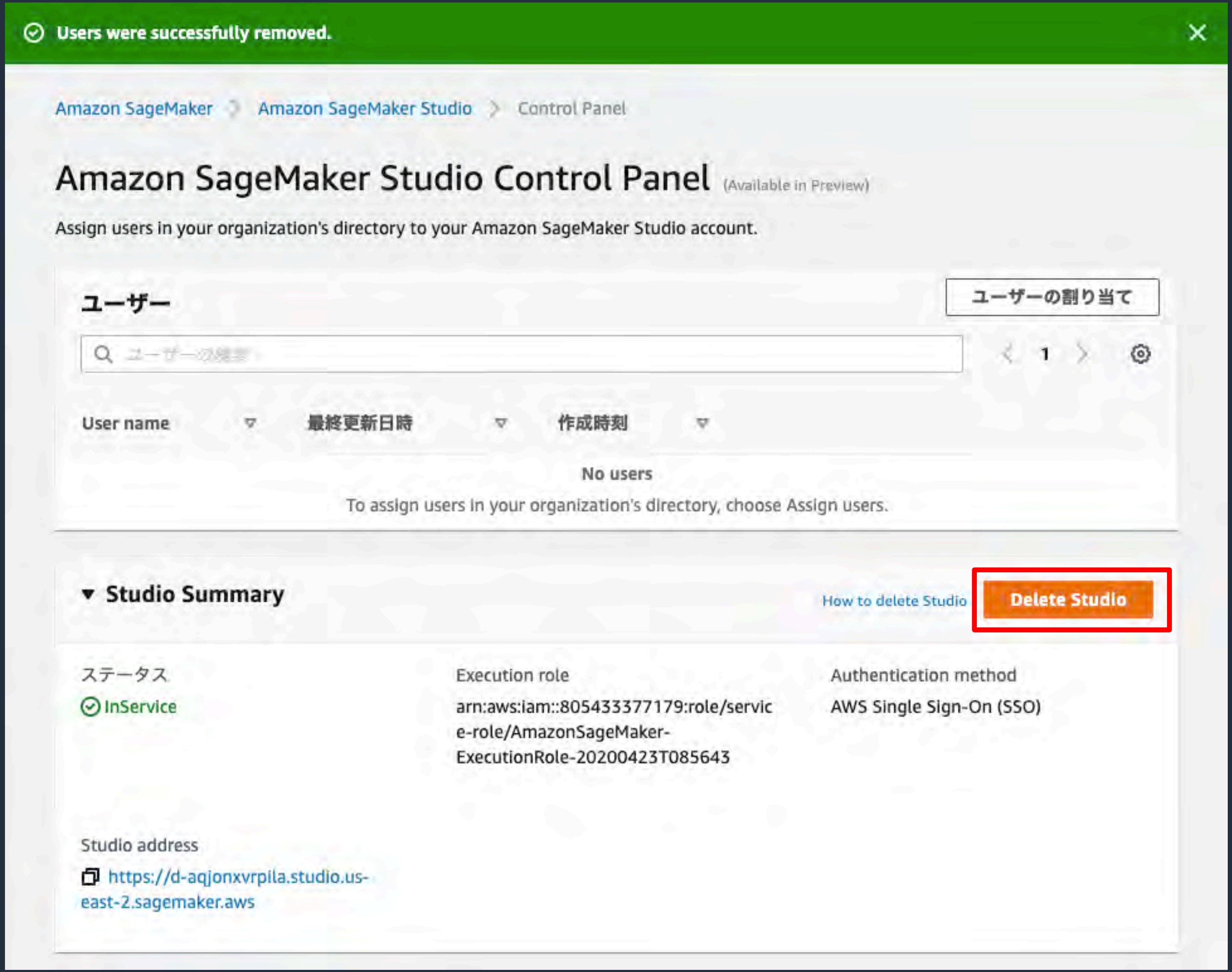


ユーザーを削除していく

削除：SageMaker Studio ドメイン

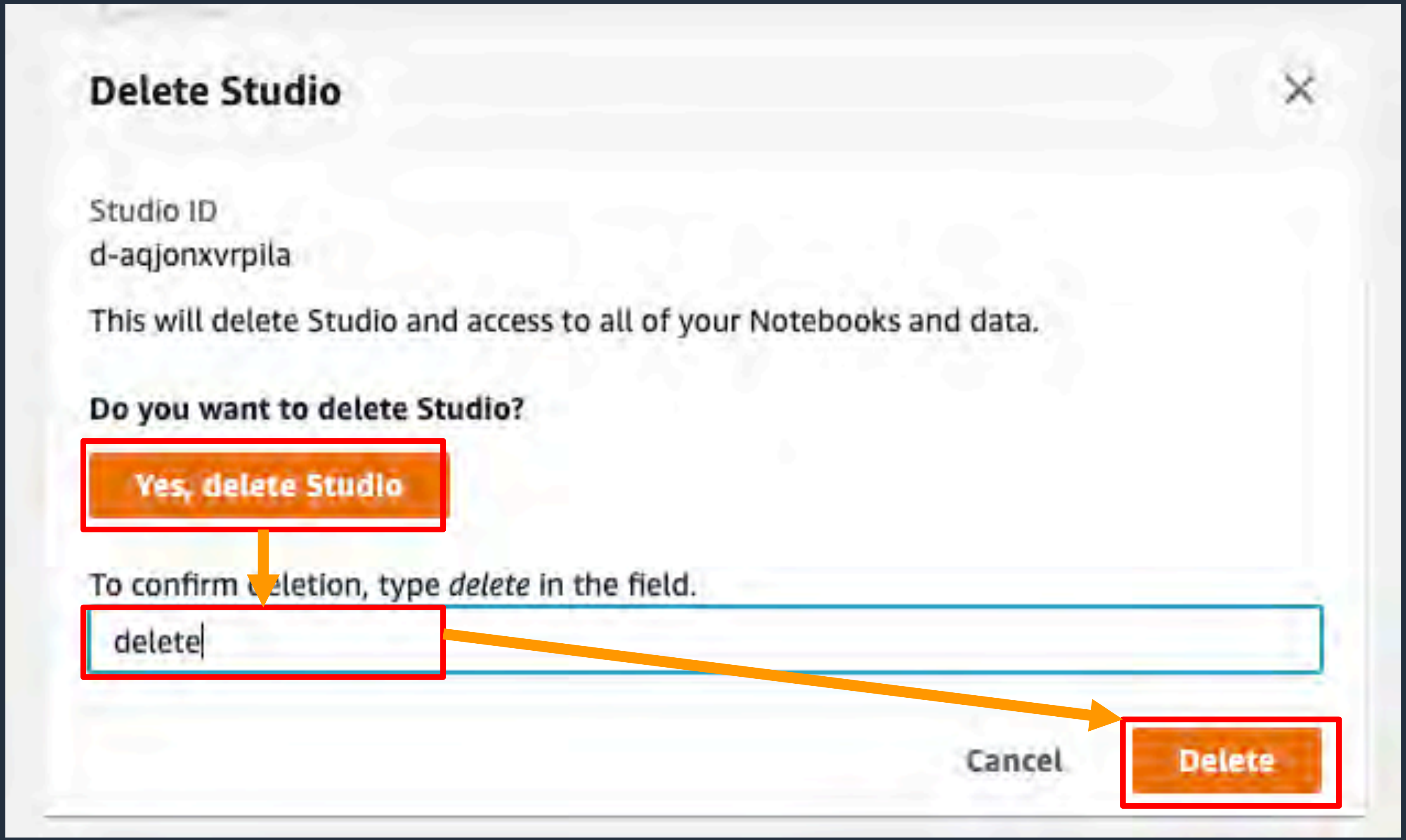


削除：SageMaker Studio ドメイン

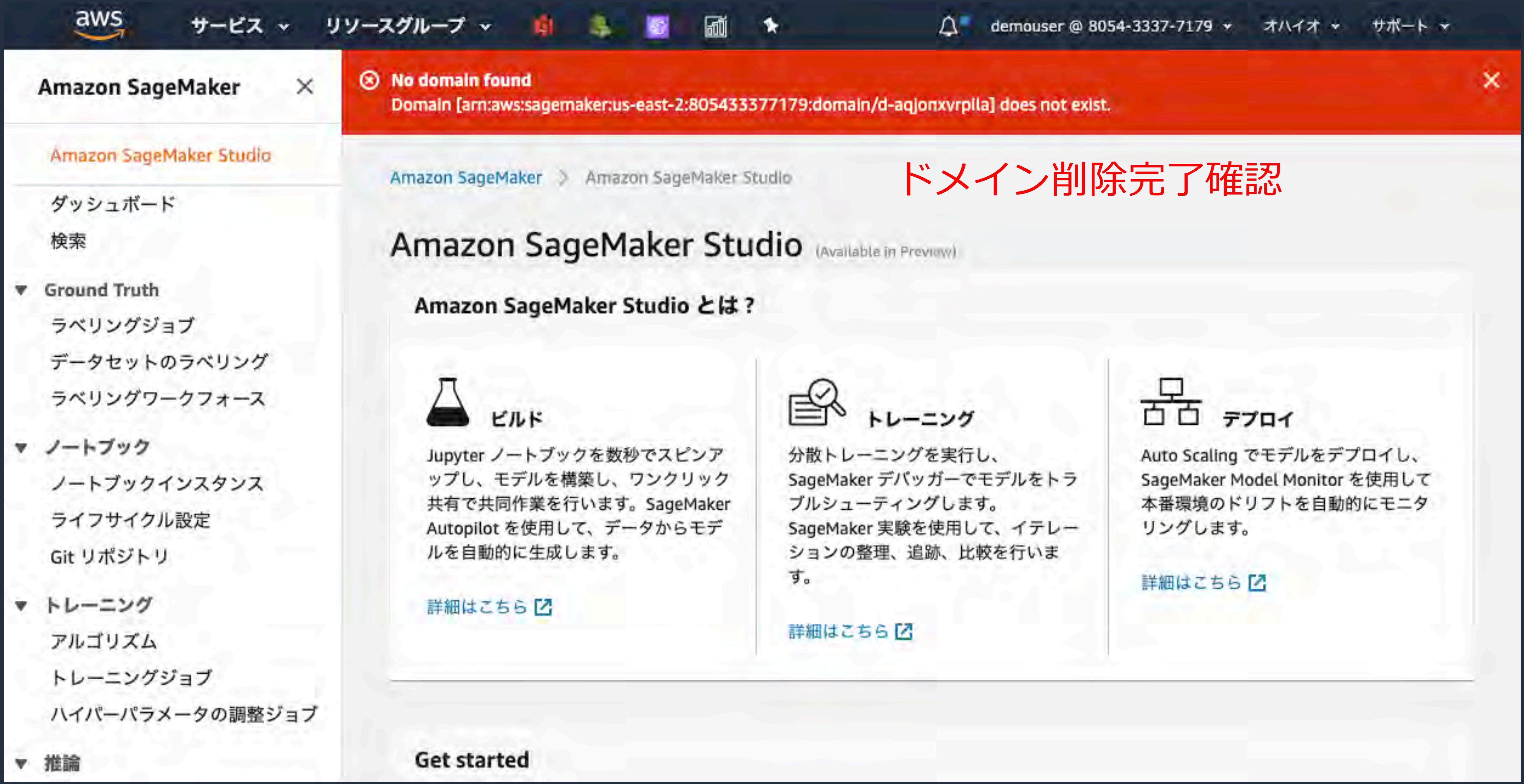


ドメインを削除していく

削除：SageMaker Studio ドメイン



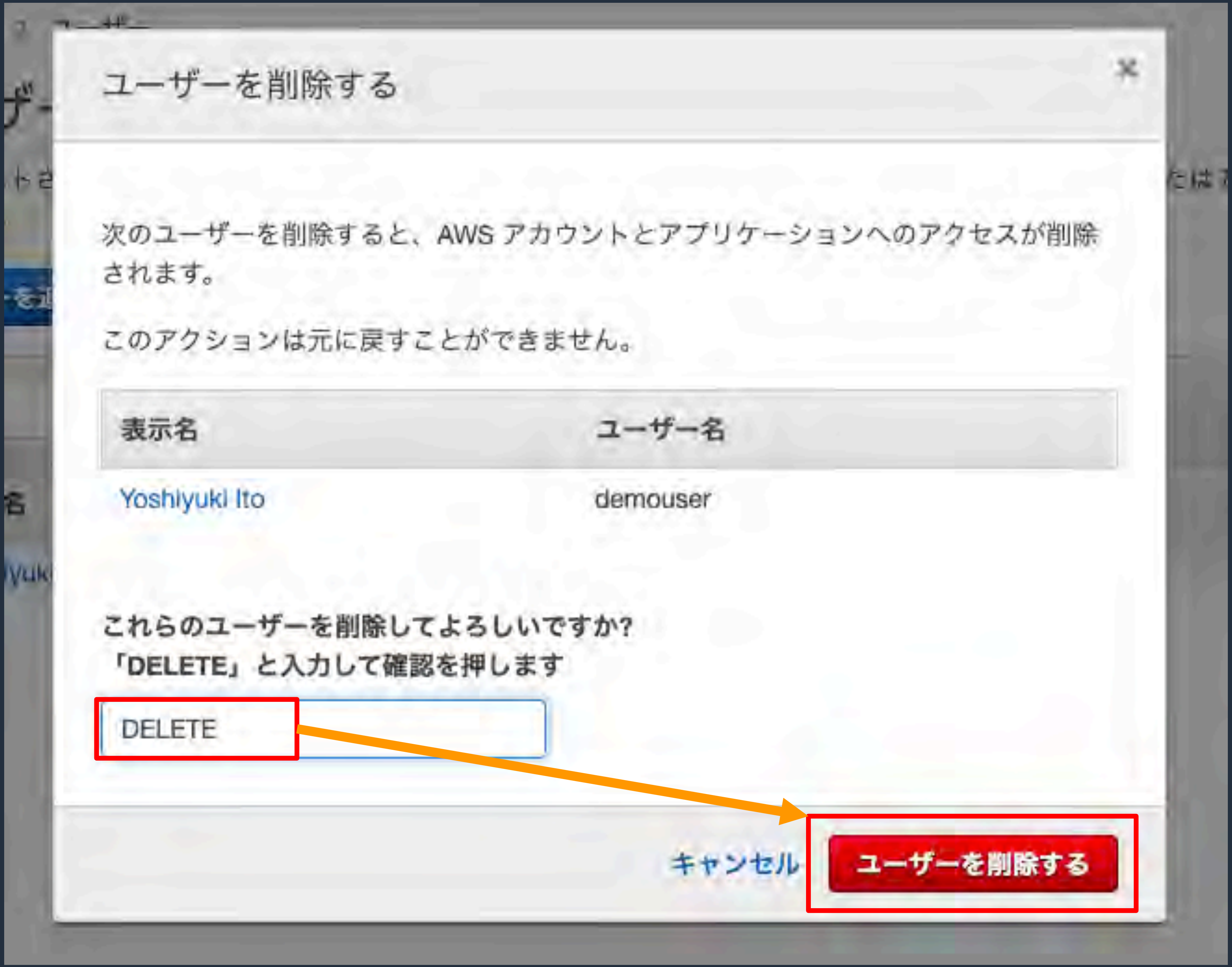
削除：SageMaker Studio ドメイン



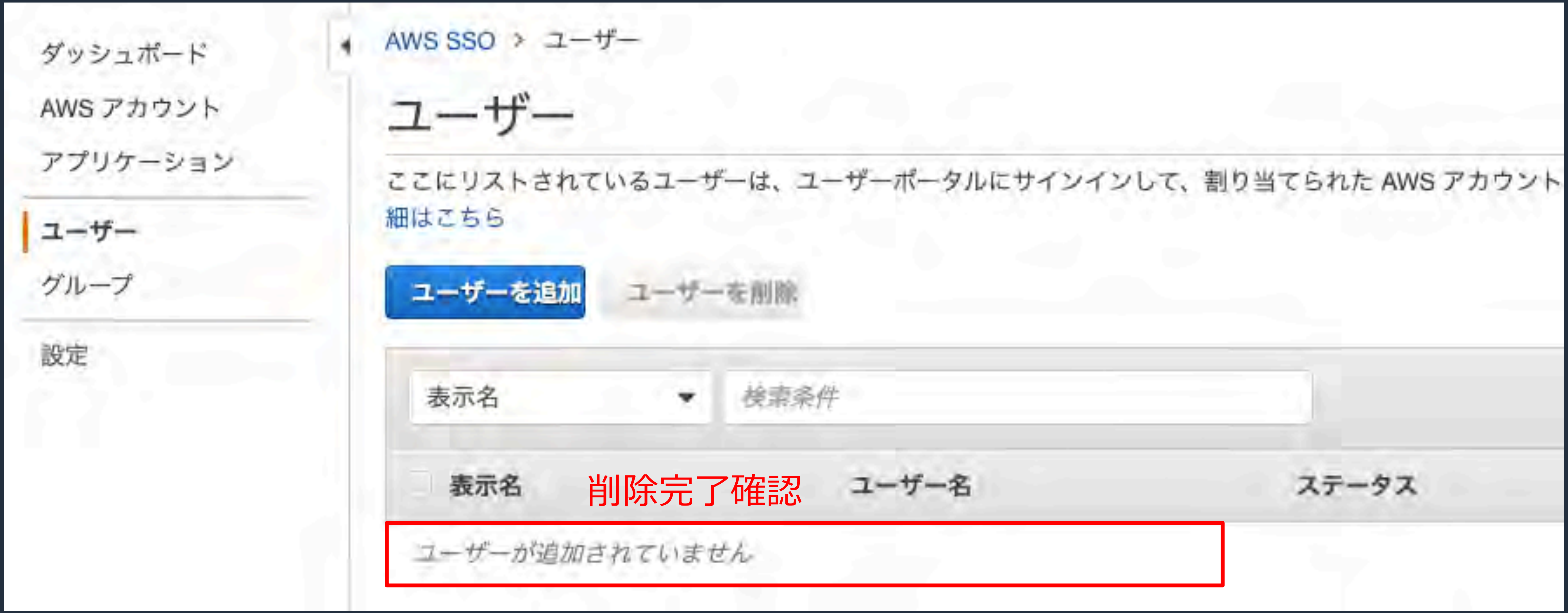
削除：SSOユーザー

ダッシュボード
AWS アカウント
アプリケーション
ユーザー
グループ
設定
AWS SSO › ユーザー
ユーザー
ここにリストされているユーザーは、ユーザーポータルにサインインして、割り当てられた AWS アカウントまたはアプリケーションにアクセスできます。 詳細はこちら
ユーザーを追加
ユーザーを削除
表示名
検索条件
表示名
ユーザー名
ステータス
Yoshiyuki Ito
demouser
有効

削除： SSOユーザー


ユーザーを削除する
次のユーザーを削除すると、AWS アカウントとアプリケーションへのアクセスが削除されます。
このアクションは元に戻すことができません。
表示名
ユーザー名
Yoshiyuki Ito
demouser
これらのユーザーを削除してよろしいですか?
「DELETE」と入力して確認を押します
DELETE
キャンセル
ユーザーを削除する

削除： SSOユーザー


ダッシュボード
AWS アカウント
アプリケーション
ユーザー
グループ
設定
AWS SSO > ユーザー
ユーザー
ここにリストされているユーザーは、ユーザーポータルにサインインして、割り当てられた AWS アカウント
細はこちら
ユーザーを追加
ユーザーを削除
表示名
検索条件
表示名
削除完了確認
ユーザー名
ステータス
ユーザーが追加されていません

Thank you!

参考

- Amazon SageMaker Studio Walkthrough
  - Experiments、Debugger、Model hosting、Model Monitorを実行できるサンプルノートブック
    - https://github.com/awslabs/amazon-sagemaker-examples/blob/master/aws_sagemaker_studio/getting_started/xgboost_customer_churn_studio.ipynb
- Amazon SageMaker 開発者ガイド
  - 上記サンプルノートブックの解説
    - https://docs.aws.amazon.com/ja_jp/sagemaker/latest/dg/gs-studio-end-to-end.html