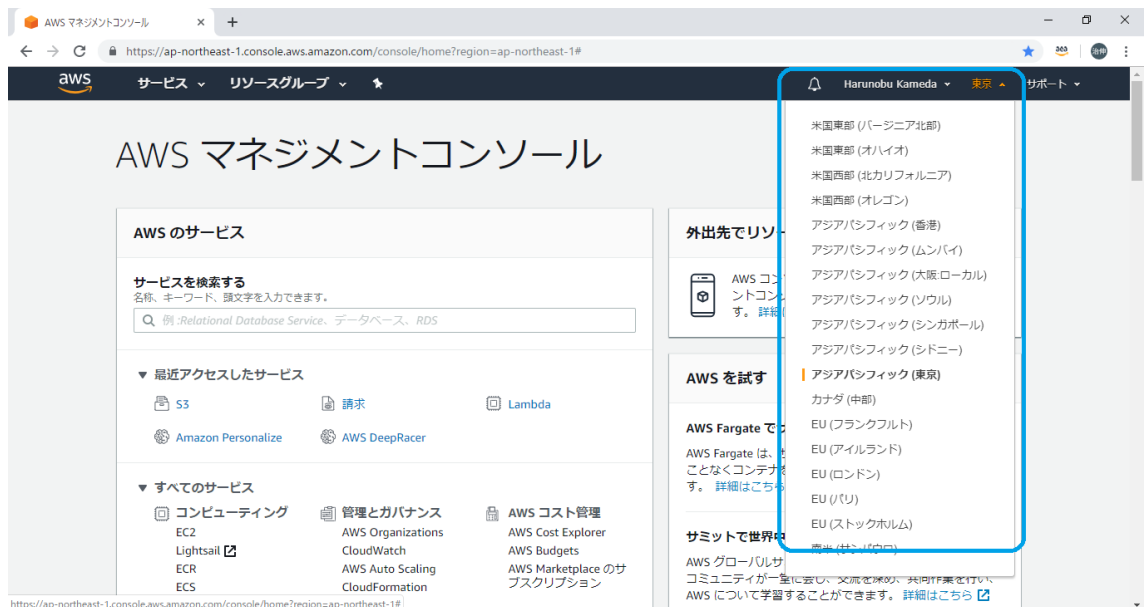


Amazon Personalize ハンズオン (No.Codes)

はじめに：Amazon Personalize は AWS が提供している過去の購買履歴などから個人化レコメンデーションを行う機械学習サービスです。データの投入や学習などで待ち時間が発生するため、Amazon Forecast ハンズオンとの同時並行をお勧めします。

1. ハンズオンを行うリージョンを画面右上から選択。リージョンは講師の指示に従ってください。



2. 映画の視聴履歴ファイルをダウンロード

<https://github.com/harunobukameda/Amazon-Personalize->

Handson/blob/master/ratings.csv

ユーザー610 人による 9700 本の映画視聴履歴とその評価が 10 万件以上ファイルに含まれています。

USER_ID : ユーザーの識別 ID (商用環境では別途ユーザーマスタと連結してください)

ITEM_ID : 映画の ID (商用環境では別途アイテムマスタと連結してください)

TIMESTAMP : 映画批評を入力された時間です。(このハンズオンでは用いませんが、Amazon Personalize の実行では必須なカラムです)

Amazon Personalize はこれらのデータをもとに、批評の相関図を作成しレコメンドを行います。従来これらのデータ分析は、Amazon EMR などを用いたヒューリスティック分析が必要でしたが、Amazon Personalize ではこれらが自動化されることが特徴です。

例) User (1) が 映画 A、B、C を視聴した場合、User(2)が B を視聴した場合、A と C がレコメンドされる。レコメンド以外にも、データの相関関係を見出す用途で様々な用途が今後想定されます。また、本シナリオでは使用しませんが、CSV ファイルには Rating (視聴済映画に対するユーザーからの評価) も準備されており、ユーザー評価に基づいたさらに細かいレ

コメントが Amazon Personalize では可能です。

3. Amazon S3 バケットに先ほど DL した ratings.csv をアップロードします。S3 バケットは既存のものを作成してもよいですし、新規で作成しても問題ありません。S3 バケットは以下のバケットポリシーを設定してください。すでにバケットポリシーが設定されている既存 S3 バケットでの作業はお勧めしていません。

```
{  
  
  "Version": "2012-10-17",  
  
  "Id": "PersonalizeS3BucketAccessPolicy",  
  
  "Statement": [  
  
    {  
  
      "Sid": "PersonalizeS3BucketAccessPolicy",  
  
      "Effect": "Allow",  
  
      "Principal": {  
  
        "Service": "personalize.amazonaws.com"  
  
      },  
  
      "Action": [  
  
        "s3:GetObject",
```

```
        "s3:ListBucket"

    ],

    "Resource": [

        "arn:aws:s3:::bucket-name",

        "arn:aws:s3:::bucket-name/*"

    ]

}

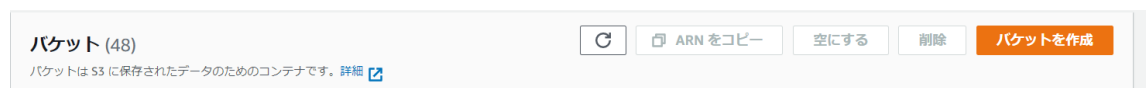
]

}
```

bucket-name はご自身で作業するバケットの名前を入れてください。

S3 バケットを新規に作成する場合は、全てデフォルトで作成してください。

【バケットを作成する】のボタンを押します。



バケット名は全 AWS ユーザーの全リージョンで一意である必要がありますので、ある程度長い名称にします。リージョンは、講師から指示されたりリージョンを指定してください。

一般的な設定

バケット名

myawsbucket

バケット名は一意である必要があり、スペース、または大文字を含めることはできません。[バケットの命名規則をご参照ください](#)

リージョン

アジアパシフィック (東京) ap-northeast-1

既存のバケットから設定をコピー - オプション
次の設定のバケット設定のみがコピーされます。

バケットを選択する

バケットが作成されたらバケット名をダブルクリックして詳細画面へ遷移し、【アクセス許可】タブを押します。

Amazon S3 > 20210214personalize

20210214personalize

オブジェクト

プロパティ

アクセス許可

メトリクス

管理

アクセスポイント

アクセス許可の概要

アクセス

非公開のバケットとオブジェクト

バケットポリシーの【編集する】をボタンを押し、上記のポリシー内容をコピーし、【変更の保存】を押します。

バケットポリシーを編集

バケットポリシー

JSON で記述されたアクセスポイントポリシーは、バケットに保存されたオブジェクトへのアクセスを提供します。バケットポリシーは、他のアカウントが所有するオブジェクトには適用されません。[詳細](#)

[ポリシーの例](#)

[ポリシージェネレータ](#)

バケット ARN

 arn:aws:s3:::20210214personalize

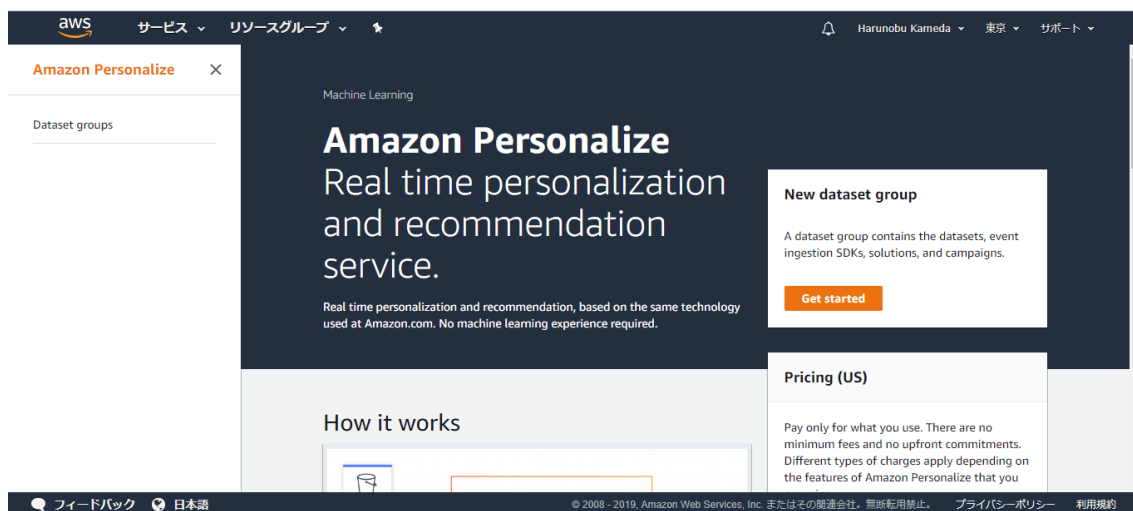
ポリシー

```
1 {  
2   "Version": "2012-10-17",  
3   "Id": "PersonalizeS3BucketAccessPolicy",  
4   "Statement": [  
5     {  
6       "Sid": "PersonalizeS3BucketAccessPolicy",  
7       "Effect": "Allow",  
8       "Principal": {  
9         "Service": "personalize.amazonaws.com"  
10      },  
11      "Action": [  
12        "s3:GetObject",  
13        "s3:ListBucket"  
14      ],  
15      "Resource": [  
16        "arn:aws:s3:::20210214personalize",  
17        "arn:aws:s3:::20210214personalize/*"  
18      ]  
19    }  
  ]  
}
```

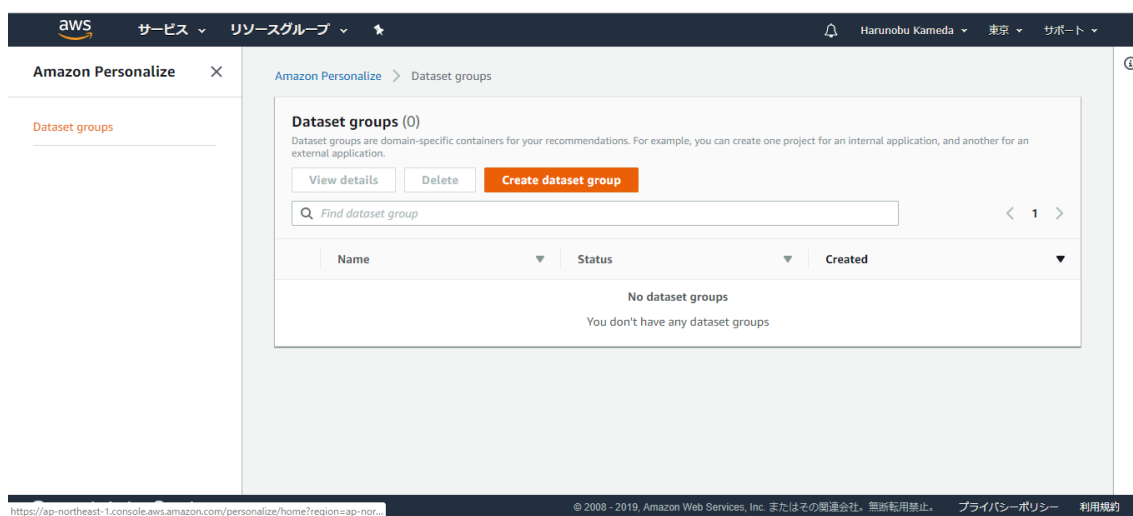
4. Amazon Personalize のマネージメントコンソール

(<https://console.aws.amazon.com/personalize/>) にアクセスします。

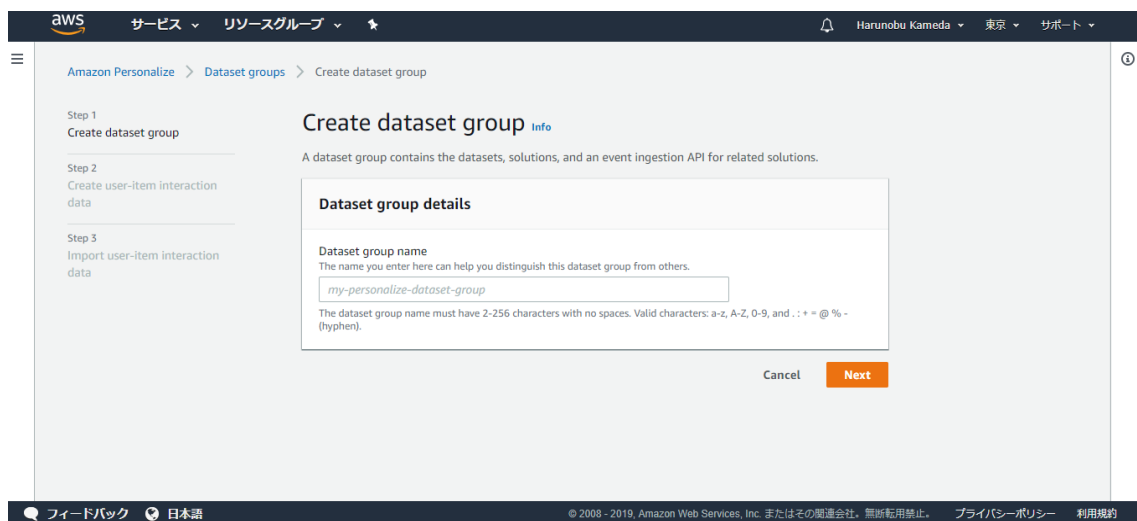
先ほど指定したリージョンになっていることを改めて確認します。



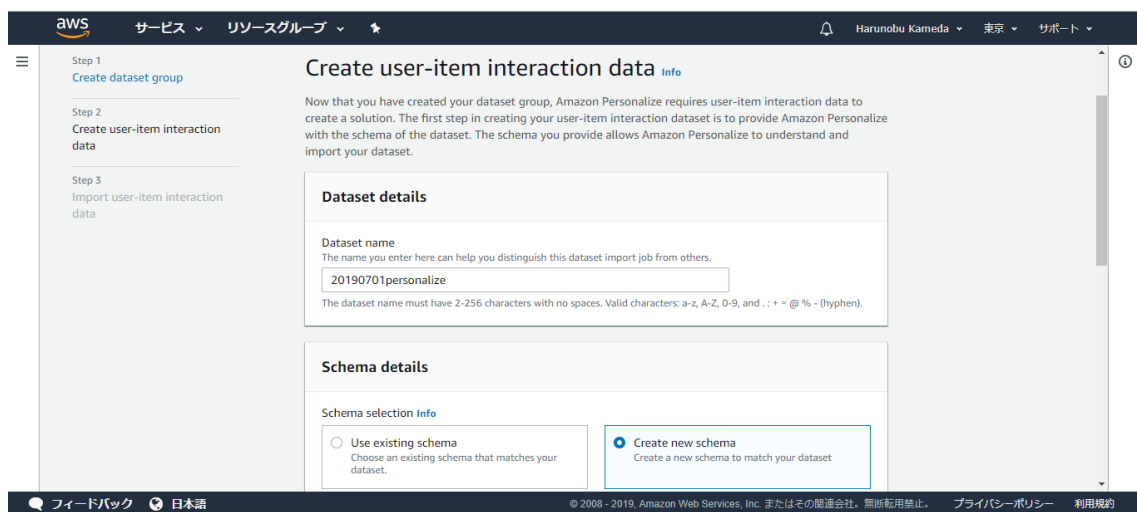
5. 画面左の [Dataset groups]を選択し、[Cerate dataset group]を押してください。



6. 作成する Dataset Group に任意の名前を付けてください。



7. 作成する Dataset に名前を付けます



8. [Create new schema]を選択し、Schema に任意の名前を設定します。

Schema selection [Info](#)

☐ Use existing schema
Choose an existing schema that matches your dataset.

☒ Create new schema
Create a new schema to match your dataset.

New schema name
The name you enter here appears in the Schema dashboard. It can help you distinguish this schema from others.
20190701personalize
The schema name must have 2-256 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and . : + = @ % - (hyphen).

Schema definition [Info](#)
Ensure your dataset's schema matches the following schema.

```

1 {
2   "type": "record",
3   "name": "Interactions",
4   "namespace": "com.amazonaws.personalize.schema",
5   "fields": [
6     {
7       "name": "USER_ID",
8       "type": "string"
9     },
10    {
11      "name": "ITEM_ID",
12      "type": "string"
13    },
14    {
15      "name": "TIMESTAMP",
16      "type": "long"
17    }
18  ],
19   "version": "1.0"
20 }

```

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

9. [Next]を押します

Schema definition [Info](#)
Ensure your dataset's schema matches the following schema.

```

1 {
2   "type": "record",
3   "name": "Interactions",
4   "namespace": "com.amazonaws.personalize.schema",
5   "fields": [
6     {
7       "name": "USER_ID",
8       "type": "string"
9     },
10    {
11      "name": "ITEM_ID",
12      "type": "string"
13    },
14    {
15      "name": "TIMESTAMP",
16      "type": "long"
17    }
18  ],
19   "version": "1.0"
20 }

```

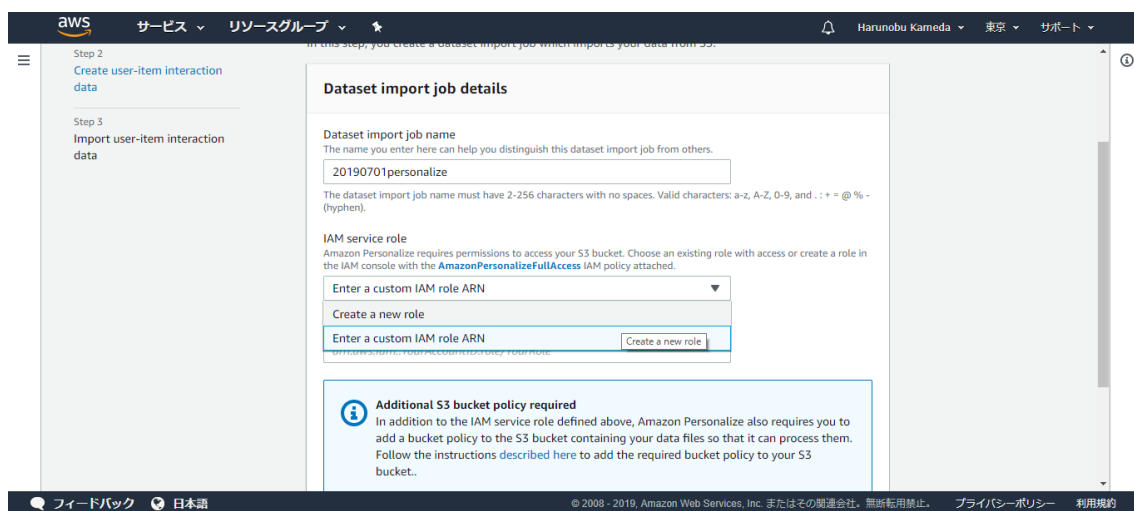
Cancel Previous **Next**

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

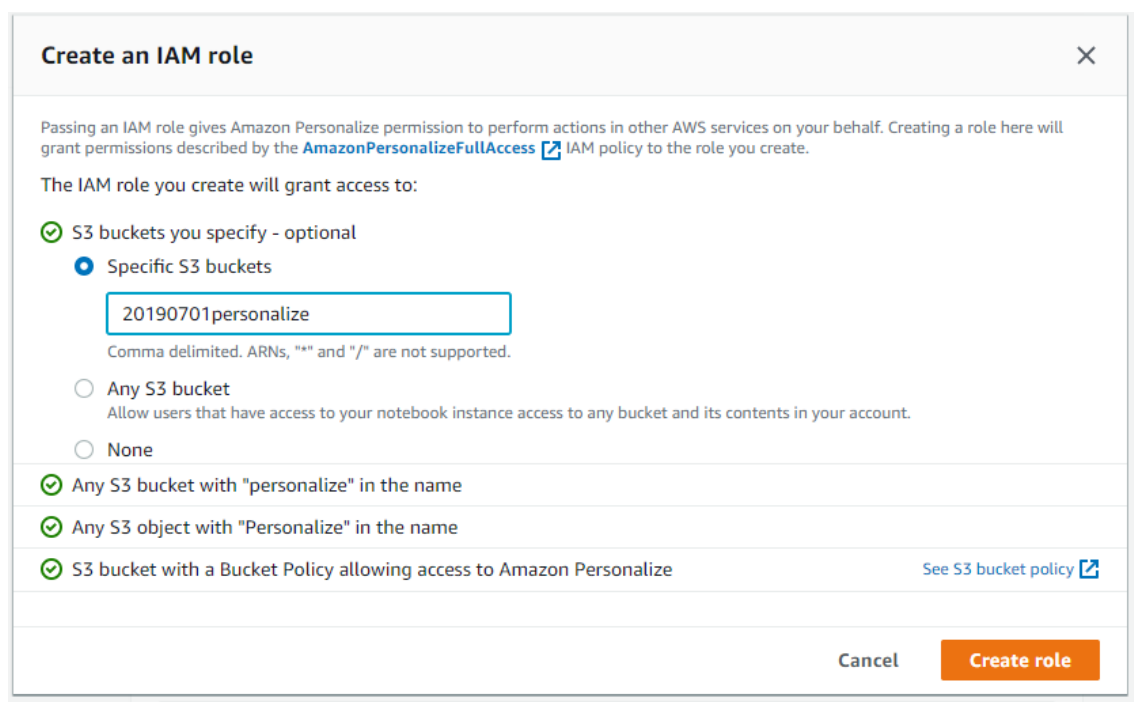
10. [user-item interaction data] (先ほどの ratings.csv)をインポート

する job の起動を行います。[Dataset import job name]に名前を入力します。

[IAM service role]で[Create a new role]を選択します。



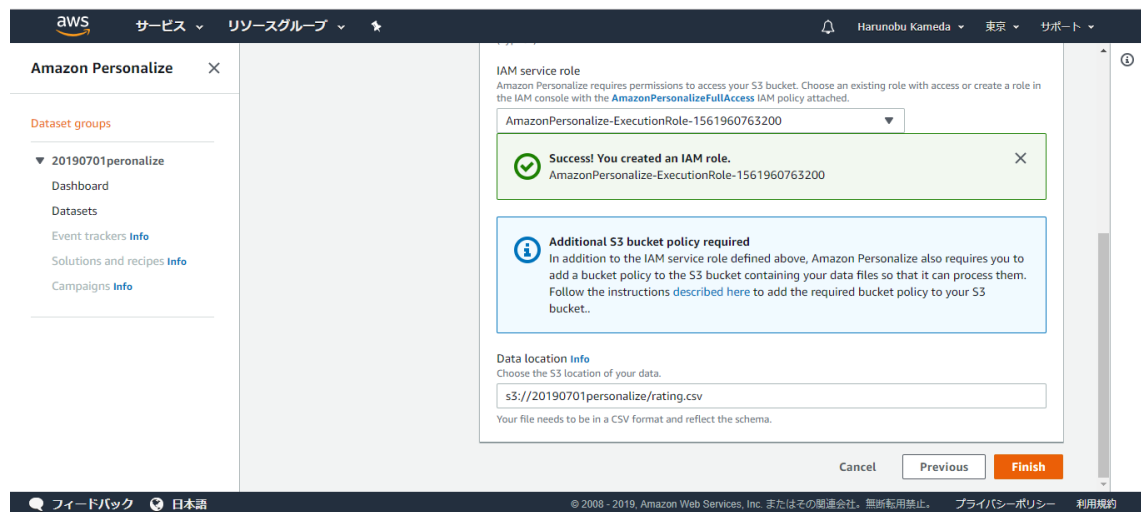
1 1. 先ほど作成した S3 バケットの名前を[Specify S3 buckets]に入力して[Create role]を押します。



1 2. [Success! You created an IAM role.]が表示されていることを確認します。

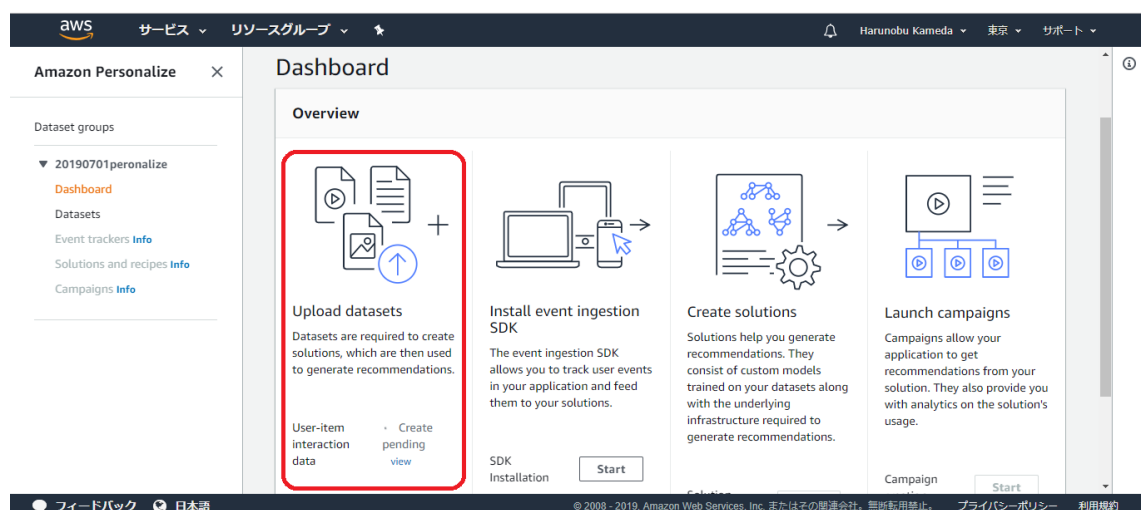
[Data location]に s3://<bucketname>/ratings.csv を入力します。作業

が完了したら[Finish]ボタンをおします。



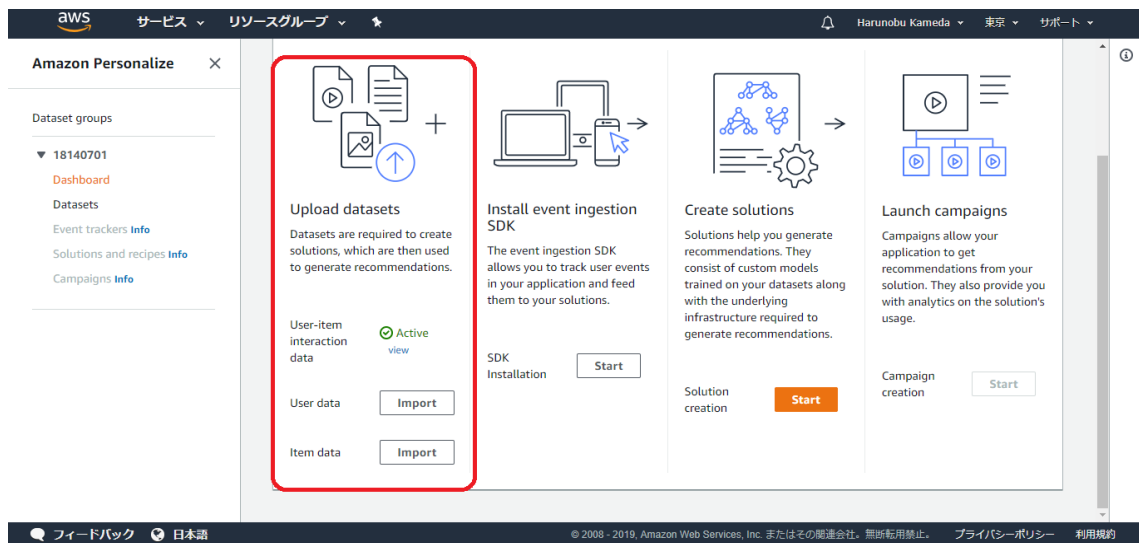
1 3. ステータスが[Create Pending] or [Create in Progress]になりますので少し待ちます。

Finish ボタンを押しても画面が遷移しない場合、ブラウザ上部の見えないところでエラーが出ていますので、右側のブラウザスクロールバーで一番上まで戻りエラーメッセージを確認してください。(ほとんどのケースにおいて、S3 のパス、もしくはファイル名のタイプミスです)

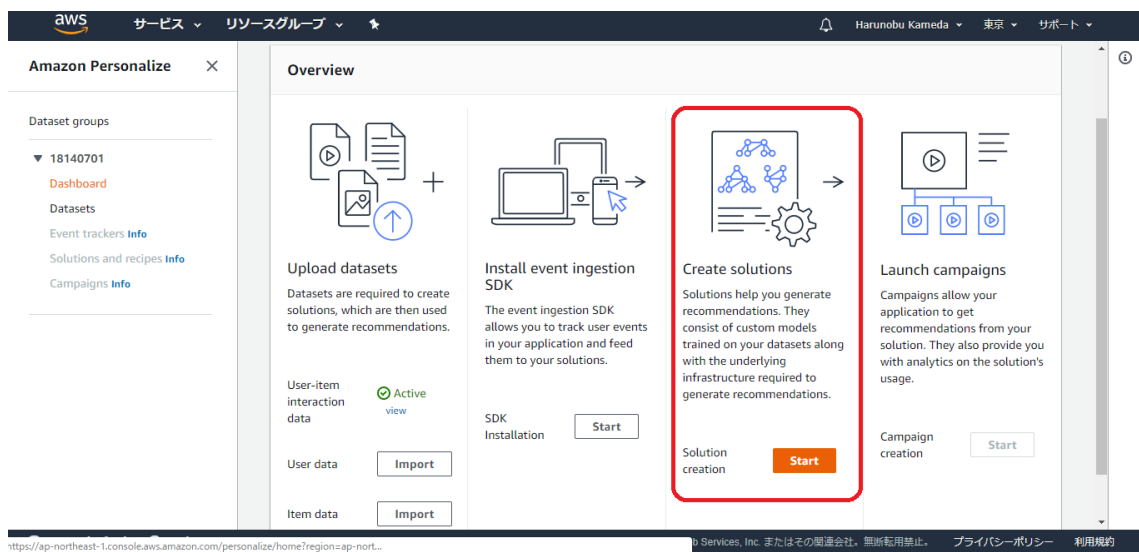


1 4. 取り込んだデータ(ratings.csv)の読み込みに少し時間がかかります。(20 分ぐらい)。待ち時間の間に、Forecast のハンズオンを続行します。

[Active]と表示されれば完了です。



1 5. [Create Solutions]の[Start]を押します。読み込んだ datasets をもとに学習を開始します。



1 6 . [Solution Name]を入力し、[Manual]を選択します。ドロップダウンリストから[aws-hrnn]を選択し以下の画面のようになれば、[Next]を押します。次に出るダイアログで[Continue with current selection]を押します。

Create solution [Info](#)

You create a solution using a recipe that is tailored to a specific use case.

Solution detail

Solution name
The solution name that you enter here can help you distinguish this solution from others.

The solution name must have 1-63 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and _ - (hyphen).

Recipe [Info](#)
Recipes are preconfigured algorithms tailored to specific use cases.

aws-hrnn (legacy)
Predicts items a user will interact with. A Hierarchical Recurrent Neural Network which models the ... ▼

► **Solution configuration - optional** [Info](#)

Choose legacy recipe? ×

You have chosen a legacy recipe

We recommend 'aws-user-personalization' recipe for this use case. Do you want to continue with current recipe choice?

Back to recipe selection

Continue with current selection

1 7 . 以下の画面で[Finish]を押すと学習が開始されます。

Create solution version

Now that you have created a solution, click Finish button to create a solution version (train your solution).

Solution overview

Solution name solution201910292	Perform HPO false	Event type -
Solution ARN arn:aws:personalize:ap-northeast-1:294963776963:solution/solution201910292	Perform AutoML false	Latest solution version ARN -
Status ✔ Active	Recipe aws-hrnn	Created Tue, 29 Oct 2019 07:51:37 GMT

► Solution config

Cancel

Previous

Finish

18. 学習は40分ぐらいかかります。

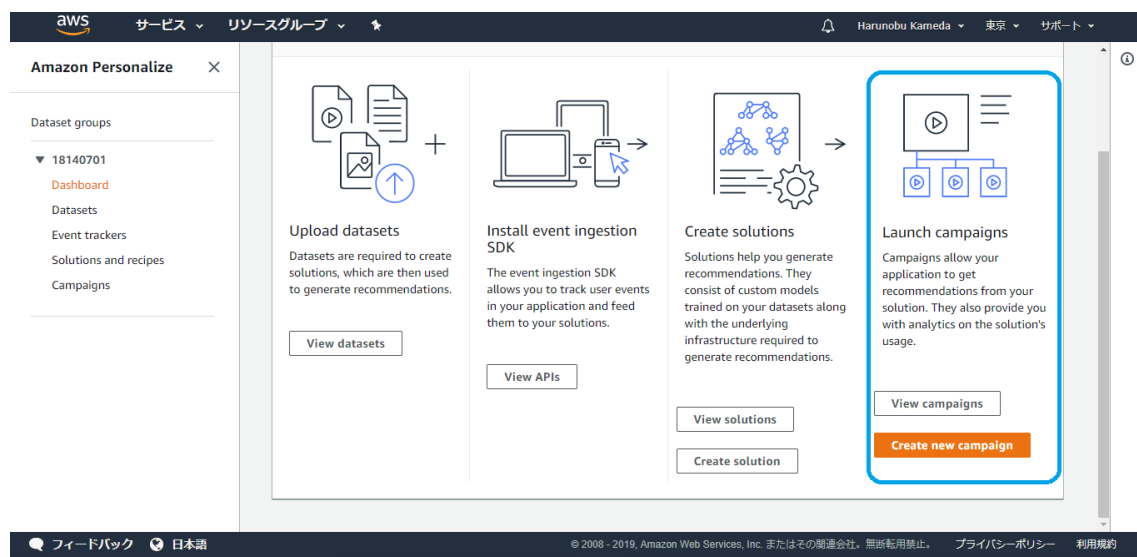
The screenshot shows the Amazon Personalize console interface. On the left is a navigation menu with 'Dataset groups' expanded, showing '18140701' and sub-items like 'Dashboard', 'Datasets', 'Event trackers', 'Solutions and recipes', and 'Campaigns'. The main area displays a four-step workflow: 1. 'Upload datasets' with 'User-item interaction data' status 'Active' and 'Import' buttons for 'User data' and 'Item data'. 2. 'Install event ingestion SDK' with a 'Start' button. 3. 'Create solutions' with a 'Create pending view' button highlighted by a blue box. 4. 'Launch campaigns' with a 'Start' button. The top navigation bar includes the AWS logo, service links, resource groups, and user information 'Harunobu Kameda' in '東京' (Tokyo). The footer contains feedback links, language settings, and copyright information for 2008-2019.

待ち時間の間に Forecast のハンズオンを続行します。

AWS CLI が使える方は待ち時間に余裕があれば巻末の手順を参考に AWS CLI

の実行環境の準備をしてみてください。

19. 学習が完了すると、[Create New Campaign]のボタンを押せるようになります。Campaign は一般的な機械学習用語でいうと、推論モデルのデプロイになります。



20. [Campaign Name]に任意の名前を入れます。[Solution]は先ほど作成した Solution の名前をドロップダウンから選択します。[Minimum provisioned transactions per second]はデフォルト状態のまま[Create Campaign]を押します。

aws サービス リソースグループ

Amazon Personalize > Dataset groups > 18140701 > Campaigns > Create new campaign

Create new campaign

Campaign details

Campaign name
The text you enter here appears in the Campaign dashboard and detail page. It can help you distinguish this campaign from others.
11310702
The campaign name must have 2-256 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and . : + = @ % - (hyphen).

Solution
The selected solution is used to generate the recommendations provided in your campaign.
19270701

Minimum provisioned transactions per second [Info](#)
The minimum amount of throughput in transactions per second (TPS) that is provisioned for this campaign.
1
Enter a number from 1-500.

Cancel Create campaign

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

2 1. 推論環境の構築が終わるまで待ちます。

aws サービス リソースグループ

Amazon Personalize × Campaign creation in progress... Creating campaign 11310702. This might take a few minutes to hours. ×

Dataset groups

▼ 18140701
Dashboard
Datasets
Event trackers
Solutions and recipes
Campaigns

Amazon Personalize > Dataset groups > 18140701 > Campaigns > 11310702

11310702

Delete Update

Personalization API Details

Campaign inference

To get recommendations for this campaign in your application, use the `getRecommendations` API call. You can learn more about the usage and requirements for this API call in the documentation and the other links listed below.

- [Amazon Personalize GetRecommendations Developer Guide](#)

Campaign ARN [Info](#)
arn:aws:personalize:ap-northeast-1:294963776963:campaign/11310702

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

待ち時間の間に Forecast のハンズオンを続行します。

2 2. 作成が完了すると[Test Campaign Result]が表示されますので、1 から 601 までの任意のユーザーID を入力し、[Get Recommendations]を押します。ratings.csv をもとに、入力されたユーザーの映画視聴履歴と相関のある別ユーザーの映画視聴履歴からレコメンドが出力されます。

例えば商用環境で使う場合、映画は ID で出力されますので以下のデータをもとに補正を行うことでユーザーに映画情報を出力することができます。

<https://grouplens.org/datasets/movielens/>

<http://files.grouplens.org/datasets/movielens/ml-latest-small.zip>

3. おつかれさまでした！

削除は作成と逆の順番で消す必要があります。

Campaigns→Solutions→Datasets→Dataset groups

S3 バケット

その他：時間に余裕があれば AWS CLI 経由での操作に挑戦してみてください。

[GetRecommendations]を用いて以下のようにレコメンドを出力できます。

```
aws personalize-runtime get-recommendations ¥
```

```
--campaign-arn arn:aws:personalize:us-west-2:acct-
```

```
id:campaign/MovieRecommendationCampaign ¥
```

```
--user-id 123
```

以下のように JSON でレコメンドが出力されます。

```
{  
  "itemList": [  
    {  
      "itemId": "14"  
    },  
    {  
      "itemId": "15"  
    },  
    {  
      "itemId": "275"  
    },  
    {  
      "itemId": "283"  
    },  
    {  
      "itemId": "273"  
    },  
    ...  
  ]  
}
```

]

}

一番簡単な AWS CLI 環境の構築は Amazon Lightsail です。

1. Amazon Lightsail の管理画面で[インスタンスの作成]を押します



2. [OS のみ]を選択します



2. Amazon Linux が選択されていることを確認し、残りは全てデフォルト値で[インスタスの作成]を押します。

3. インスタスが作成中となりますので少し待ちます



4. 実行中になったら起動が完了ですので、コンソールボタンを押します。



5. [sudo pip install -U awscli]で CLI をアップグレードします。

6. [aws configure]でデフォルトリージョンを作業中のリージョンに指定し

ます。(東京であれば[ap-northeast-1])

CLI の実行権限([AmazonPersonalizeFullAccess]と

[AmazonPersonalizeFullAccess]を合わせて設定します。)この際設定する

IAM ロールは[AmazonPersonalizeFullAccess]を保持している必要があります。

[Default output format]はデフォルトのまま[None]を設定します。

7. aws personalize-runtime get-recommendations --campaign-arn

arn:<personalize campaign の arn> --user-id 123

のように入力し実行します。

以下のようなエラーがでた場合、IAM ロールの設定ミスが発生しています。

An error occurred (AccessDeniedException) when calling the

GetRecommendations operation: User:

arn:aws:sts::320976636470:assumed-

role/AmazonLightsailInstanceRole/i-0fa8803af5e4611ee is not

authorized to access this resource

8. 成功した場合以下が表示されます。

```
Amazon_Linux-512MB-Tokyo-1 - ターミナル | Lightsail - Google Chrome
https://lightsail.aws.amazon.com/ls/remote/ap-northeast-1/instances/Amazon_Linux-512MB-Tokyo-1/terminal?protocol=ssh
[ec2-user@ip-172-26-4-242 ~]$ aws personalize-runtime get-recommendations --campaign-arn arn:aws:personalize:ap-northeast-1:294963776963:campaign/11310702 --user-id 123
{
  "itemList": [
    {
      "itemId": "38038"
    },
    {
      "itemId": "37729"
    },
    {
      "itemId": "53125"
    },
    {
      "itemId": "115149"
    },
    {
      "itemId": "106696"
    },
    {
      "itemId": "51540"
    },
    {
      "itemId": "66097"
    },
    {
      "itemId": "68791"
    },
    {
      "itemId": "63859"
    }
  ]
}
```