

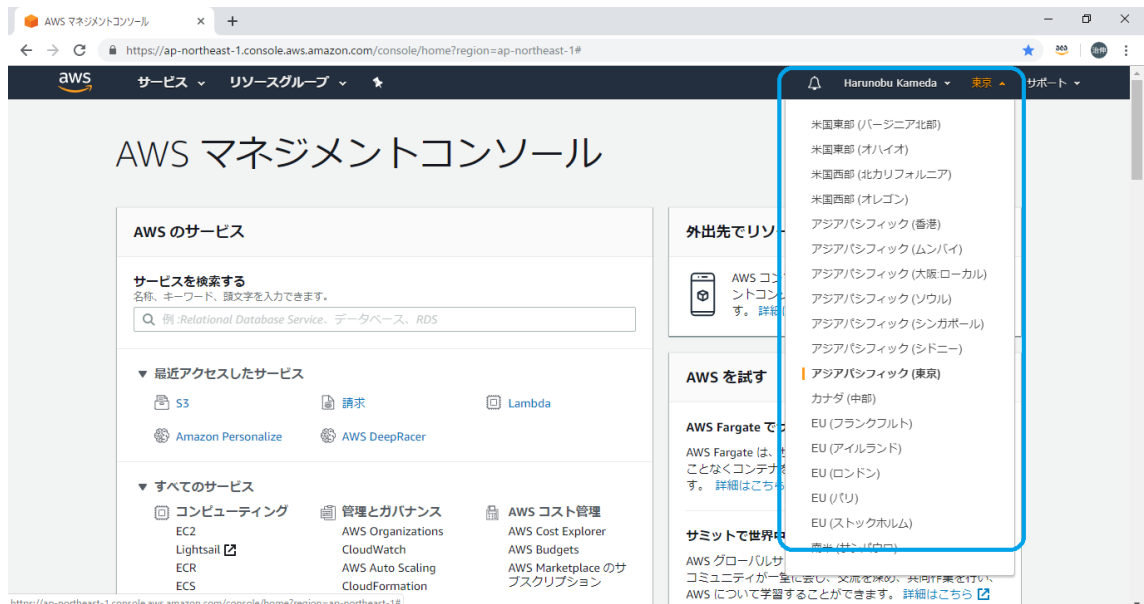
Amazon Personalize ハンズオン (No.Codes)

2022/08/15

シニアエバンジェリスト 亀田

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
ドが 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/*"

    ]

}

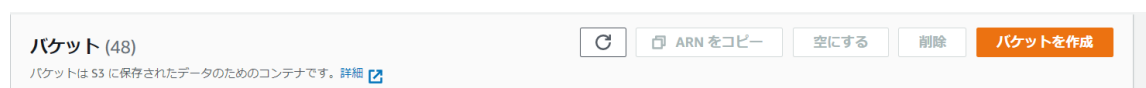
]

}
```

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    }
20  ]
21 }
```

作業が完了したら【アップロード】ボタンをおします

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

オブジェクト (0)

オブジェクトは、Amazon S3 に保存された基本的なエンティティです。Amazon S3 インベントリを使用して、バケット内のすべてのオブジェクトのリストを取得できます。他のユーザーが自分のオブジェクトにアクセスできるように、明示的にアクセス権限を付与する必要があります。[詳細はこちら](#)

🔄

📄 S3 URI をコピー

📄 URL をコピー

📄 ダウンロード

🔍 開く

🗑️ 削除

アクション ▼

📁 フォルダの作成

📄 アップロード

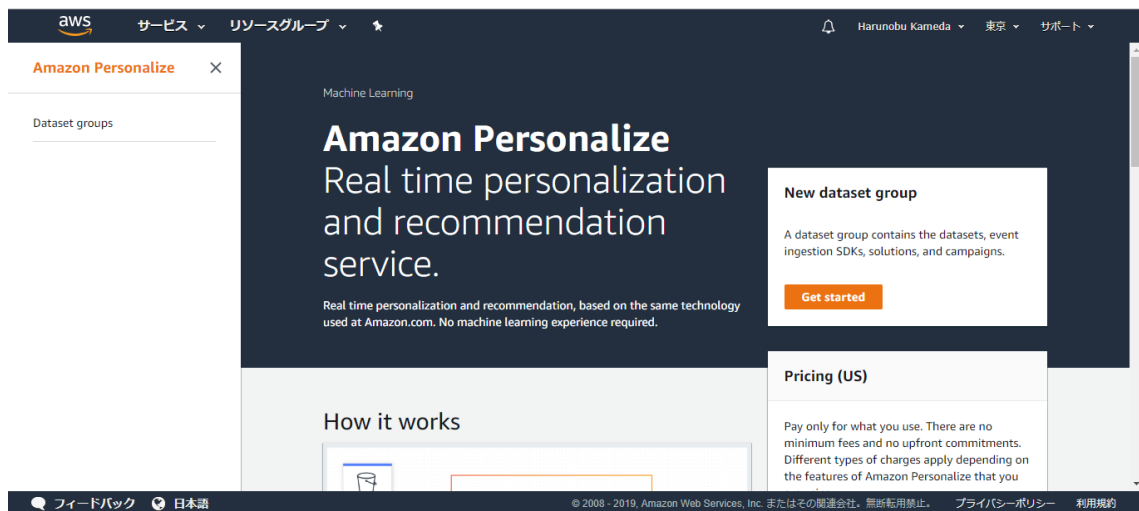
【ファイルの追加】をおします。rating.csv を追加して、【アップロード】

をおします

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 [Create 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、【Custom】を選んでください。

Dataset group details

Name

The name you enter here distinguishes this dataset group from others.

my-personalize-dataset-group

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

Domain

Choose a domain for your use cases.

☐ **E-commerce**

Grow your business by recommending the right products at the right time.

☐ **Video on demand**

Increase engagement by recommending relevant content to your users.

☐ **Custom**

Create and manage custom resources for your use cases.

7. 【Create dataset group and continue】をおします
8. 【Dataset name】と【Scheme name】に適切な名前をいれます

Dataset details

Dataset name

The name you enter here can help you distinguish this dataset import job from others.

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

Dataset schema

The schema you provide allows Amazon Personalize to understand and import your data.

☒ Create new schema

☐ Use an existing schema

Schema name

The name you enter here can help you distinguish this schema from others.


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

9. 【Create dataset and continue】をおします
10. 【Dataset import job name】に適切な名前をつけます
11. 【Data location】に先程作成した s3 バケット名およびファイル名を以下
のルールに従い指定します

s3://<bucket-name>/ratings.csv

<bucket-name>は先程皆さんが作成した名前に置き換えます

Data import source

 **Additional S3 bucket policy required**
In addition to the IAM service role defined above, Amazon Personalize also requires you to add a bucket policy to the S3 bucket containing your data files so that it can process them. Follow the instructions [described here](#) to add the required bucket policy to your S3 bucket.

Data location [Info](#)
Choose the S3 location of your data.

s3://vairignia20220815personalize/ratings.csv

Your file needs to be in a CSV format and reflect the schema.

12. 【IAM Role】では [Create a new role]を選びます。表示されるウィンドウで[Any S3 bucket]を選び[Create role]をおします

Create an IAM role ×

Passing an IAM role gives Amazon Personalize permission to perform actions in other AWS services on your behalf. Creating a role here will grant permissions described by the [AmazonPersonalizeFullAccess](#) IAM policy to the role you create.

The IAM role you create will grant access to:

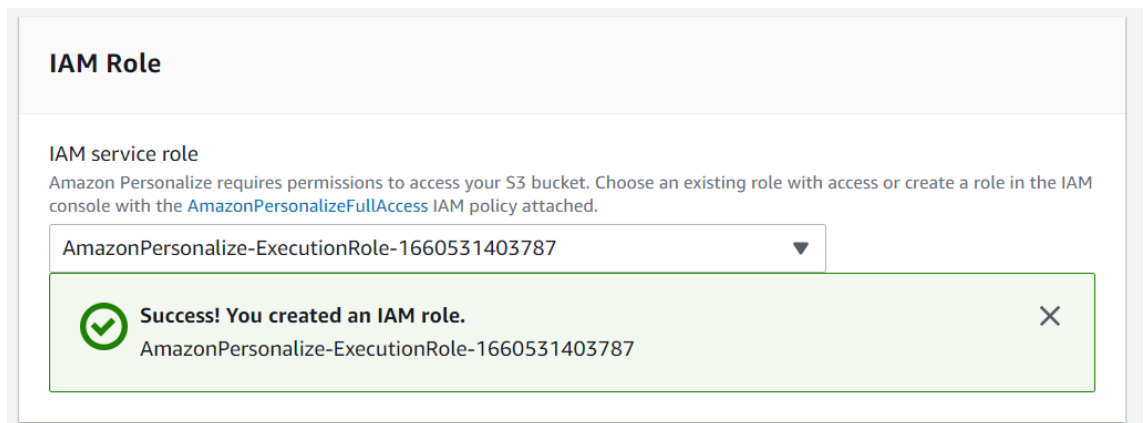
- ☒ S3 buckets you specify - *optional*
 - ☐ Specific S3 buckets

Example: bucket-name-1, bucket-name-2, bu

Comma delimited. ARNs, "*" and "/" are not supported.
 - ☒ Any S3 bucket
Allow users that have access to your notebook instance access to any bucket and its contents in your account.
 - ☐ None
- ☒ Any S3 bucket with "personalize" in the name
- ☒ Any S3 object with "Personalize" in the name
- ☒ S3 bucket with a Bucket Policy allowing access to Amazon Personalize [See S3 bucket policy](#)

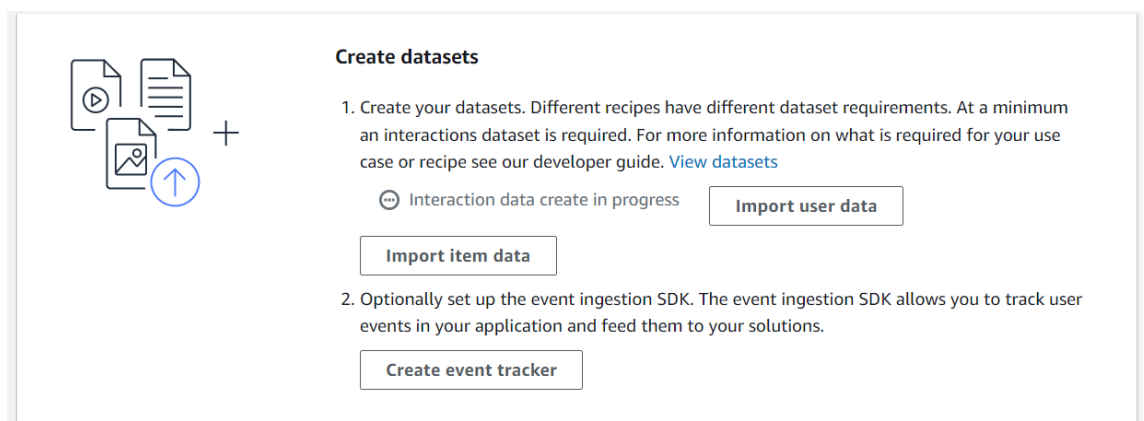
Cancel Create role

13. サイド、ドロップダウンから先程作成された IAM ロールを指定します

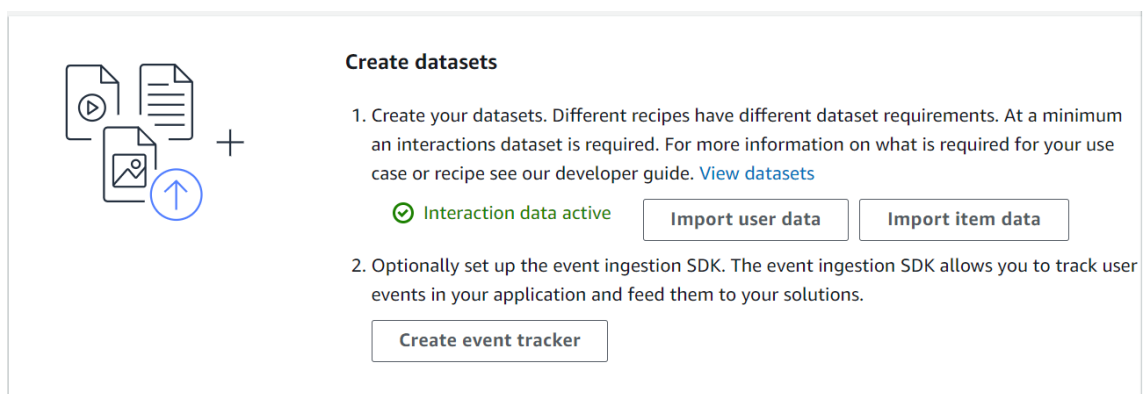


14. 【Finish】をおします

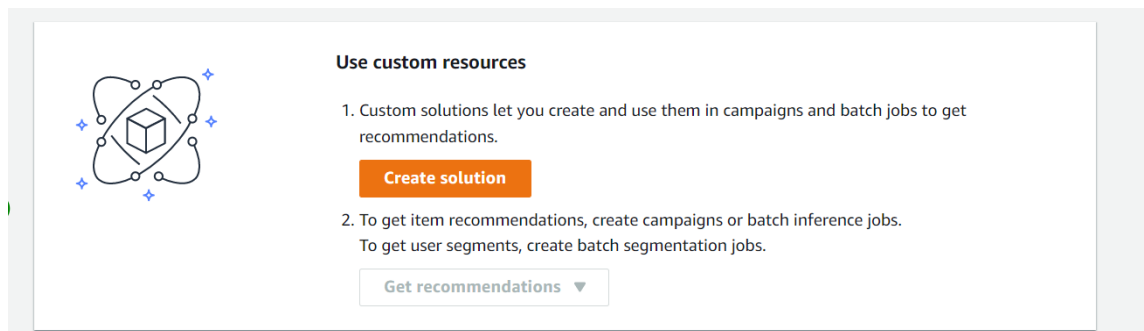
15. データ読み込み中になりますので、しばらく待ちます。（この間講師によるサービス説明があります）



16. 以下のように[Active]と表示されれば完了です



17. 続いてメインの学習にはいります。【Create Solution】をおします



18. 適当な名前を【Solution-name】に指定し、[Item recommendation] を選択します

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).

Solution type [Info](#)
Choose the type of solution you want to create. The type determines what recipes are available for solution creation.

☒ **Item recommendation**
Create a solution that generates item recommendations.

☐ **User segmentation - new**
Create a solution that predicts groups of users based on item input data.

19. 【Recipe】では[aws-similar-items]を選びます。この Recipe は学習に用いるデータの性質毎に最適なものを選ぶ必要がありますがこのハンズオンでは、似ているアイテムをユーザーの視聴履歴から特定し、似ているアイテムを良く視聴している似ているユーザーを洗い出します。

20. 【Create and train solution】をおします

21. 待っている間に、画面左ペインの[custom resources]から[Solutions

and receipes]をクリックします。

▼ 20220815dataset

Overview

Datasets

Event trackers

▼ Custom resources

Solutions and recipes

Campaigns

Batch inference jobs

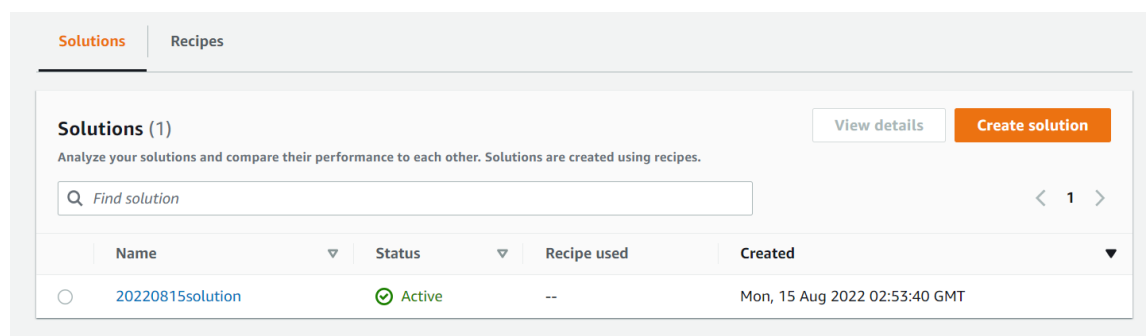
Batch segment jobs

New

Filters

22. 表示された solutions は Active となっていますが、実際は学習中で

す。ソリューション名をクリックします



Solutions (1)

Analyze your solutions and compare their performance to each other. Solutions are created using recipes.

Find solution

Name	Status	Recipe used	Created
20220815solution	Active	--	Mon, 15 Aug 2022 02:53:40 GMT

23. 学習が完了するまでしばらく待ちます

Solution versions (1)			Stop creation
<input type="text"/> Find solution version			< 1 >
Solution version ID	Solution version status	Created	
<input type="radio"/> 426b5b20	Create in progress	Mon, 15 Aug 2022 02:53:40 GMT	

この待ち時間の間に講師が追加の Personalize の説明を行います。

以下のように Active になれば学習完了です。

Solution versions (1)			Stop creation
<input type="text"/> Find solution version			< 1 >
Solution version ID	Solution version status	Created	
<input type="radio"/> 426b5b20	Active	Mon, 15 Aug 2022 02:53:40 GMT	

24. 続いて campaign を作成します。Campaign とは推論を行う API のこ

とをさします。【Create campaign】をおします

25. 適当な名前を入力し【Create campaign】をおします

26. 左ペインの[Custom resources] から [Solutions and recipes]をクリ

ックします。

27. 表示されたソリューションをクリックします。以下のように API 作成中

ですののでしばらく待ちます。

Campaign using the solution (1)			Create campaign
<input type="text"/> Find campaign			< 1 >
Name	Status	Created	
<input type="radio"/> 20220815campaign	Create in progress	Mon, 15 Aug 2022 04:06:00 GMT	

28. 以下のように Active と表示されたら完了です

Campaign using the solution (1)			Create campaign
<input type="text" value="Find campaign"/>			< 1 >
Name	Status	Created	
20220815campaign	Active	Mon, 15 Aug 2022 04:06:00 GMT	

29. 生成された campaign をクリックします

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。

30. 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。

31. 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。

以下のように JSON でレコメンドが出力されます。

```
[cloudshell-user@ip-10-0-138-90 ~]$ aws personalize-runtime get-recommendations --campaign-arn arn:aws:personalize:ap-northeast-1:294963776963:campaign/20220815campaign --item-id 1
{
  "itemList": [
    {
      "itemId": "780"
    },
    {
      "itemId": "736"
    },
    {
      "itemId": "141"
    },
    {
      "itemId": "3114"
    }
  ]
}
```

一番簡単な AWS CLI 環境の構築は CloudShell です。マネージメントコンソール

右上の、コマンドアイコンを押します



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

<arn:personalize campaign の arn> --item-id 1

のように入力し実行します。

Campaign ARN は以下の画面に表示されています

Amazon Personalize > Dataset groups > 20220815dataset > Campaigns > 20220815campaign

20220815campaign

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/20220815campaign

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"
    }
  ]
}
```