

Amazon Personalize ハンズオン

Amazon Web Services Japan

今回実施するハンズオンの内容

- AWS 公式の Getting started に従って実施
- 以下のURLから、Amazon Personalize の Getting started のページに移動 (URLのメモ, bookmarkしておくと便利)

https://docs.aws.amazon.com/ja_jp/personalize/latest/dg/getting-started.html





ハンズオンの手順

- 1. データの準備
- 2. データセットのインポート
- 3. ソリューションの学習
- 4. キャンペーンを作成してリコメンデーションを行う
- 5. リソースの削除



1. データの準備



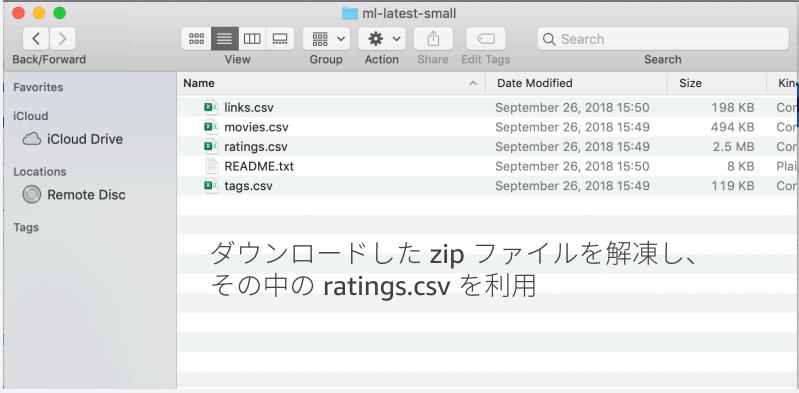
1-1. データの準備

- 9700 本の映画に対する 600 ユーザーの評価履歴にも とづいて、映画のリコメンデーションを行う
- 学習データとなる評価履歴のデータを以下のURLから ダウンロード・解凍

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



1-2. ダウンロードしたデータ

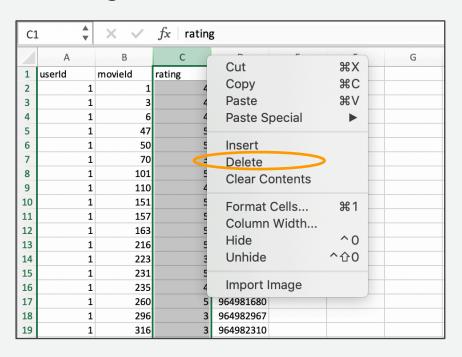




1-3. データの加工

csv を編集できる適当なソフトで加工して保存する

(1) ratings の列を削除する



(2) ヘッダを Personalize にあわせて USER_ID,ITEM_ID,TIMESTAMP にする

C1	C1 ★ X ✓ fx TIMESTAMP				
	Α	В	С	D	Е
1	USER_ID	ITEM_ID	TIMESTAMP		
2	i	1	964982703		
3	1	3	964981247		
4	1	6	964982224		
5	1	47	964983815		
6	1	50	964982931		
7	1	70	964982400		
8	1	101	964980868		
9	1	110	964982176		
10	1	151	964984041		
11	1	157	964984100		
12	1	160	06/1002650		



1-4. Amazon S3 にアクセス

データをS3にアップロードするためS3にアクセス

コンソールにサインイン



マネジメントコンソールから **S3**を検索





Amazon S3	Amazon S3 へようこそ。新しいパケットを作成するか、既存のパケットを選択してプロパティを表示、設定します。
パケット	ドキュメント
バッチオペレーショ ン	S3 パケット ロコンソールのご紹介
ブロックパブリック	Q パケット検索 すべてのアクセスタイプ v
アクセス (アカウン ト設定)	+ パケットを作成する パブリックアクセス設定を編集する 空にする 削除
注目機能 2	62 パケット 4 リージョン
ALD WARE	「パケット名 ▼ アクセス ① ▼ リージョン 作成日 ▼
	「バケットとオブジェ 米国西部 (オレ 7月 30, 2019 1:48:25 午前 クトは非公開 ゴン) GMT+0900
	大ブジェクトは公開 米国東部 6月 7, 2019 (パージニア 10:35:26 午前 10:45:25 で能 北部
	オブジェクトは公開 米国東部 6月 6, 2019 (パージニア 2:34:15 午後 北部



	バケ	ットの作成	×
1 名前とリージョン	② オプションの設定	③ アクセス許可の設定	4 確認
名前とリージョン	唯一な名前	を入れる	
test-bucket-persona	lize		
リージョン			
米国西部 (オレゴン)			~
既存のバケットから	_{設定をコピー} Person	alizeを利用する	るリージョ
バケットを選択する	(省略可) 62 バケット		~
作成			キャンセル 次へ



Amazon S3	Amazon S3 へようこそ。新しいパケットを作成するか、既存のパケットを選択してプロパティを表示、設 定します。
バケット	ドキュメント
バッチオペレーショ ン	S3 バケット ロコンソールのご紹介
ブロックパブリック アクセス (アカウン	Q パケット検索 すべてのアクセスタイプ マ
ト設定)	+ パケットを作成する パブリックアクセス設定を編集する 空にする 削除 ご
注目機能 2	62 パケット 4 リージョン
ALDIMHO 2	□ パケット名 ▼ Pクセス ① ▼ リージョン 作成日 ▼
	□ I
	□ ■ ** ********************************
	□ ■ *** *** *** *** *** *** *** *** ***







	アップロ	ード	×
① ファイルの選択	② アクセス許可を設定する	③ プロパティを設定する	4 確認
1 ファイル サイズ: 2.0 MB	ターゲットパス: test-bucket-for-pers	onalize	
160 GB よりも大きいファイ はこちら ☑	ルをアップロードするには、AWS CLI、	AWS SDK、Amazon S3 REST API を使	用します。 詳細
十 さらにファイルを追加す			
ratings.	csv	×	
- 2.0 MB		プでアップロ	 一ド
- 2.0 MB			 一ド
- 2.0 MB			 一ド





1-6. Personalize からバケットアクセスを許可

Personalize からのアクセスを許可するようにバケットポリシーを設定



1-6. Personalize からバケットアクセスを許可

右のバケットポリシーをコピーして、S3のエディタに貼り付けて保存

bucket-name を作成したバケット名に変更



```
"Version": "2012-10-17",
"Id": "PersonalizeS3BucketAccessPolicy",
"Statement": [
    "Sid": "PersonalizeS3BucketAccessPolicy",
    "Effect": "Allow".
    "Principal": {
       "Service": "personalize.amazonaws.com"
    "Action": [
       "s3:GetObject",
       "s3:PutObject",
       "s3:ListBucket"
    "Resource": [
       "arn:aws:s3:::bucket-name",
       "arn:aws:s3:::bucket-name/*"
```

2. データセットのインポート



2-1. Personalize のコンソールに移動

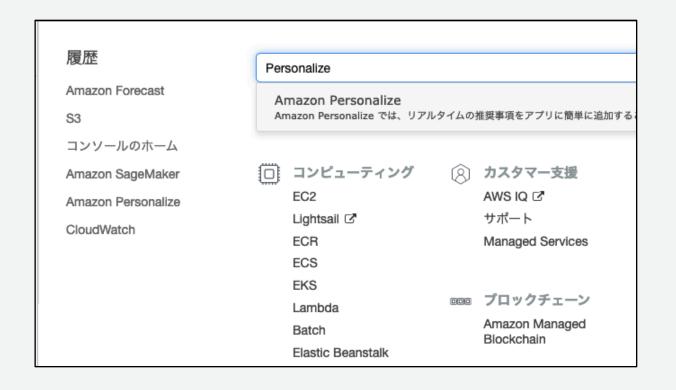
S3 から Personalize のコンソールに移動するために、 まず左上の「サービス」を選択する





2-1. Personalize のコンソールに移動

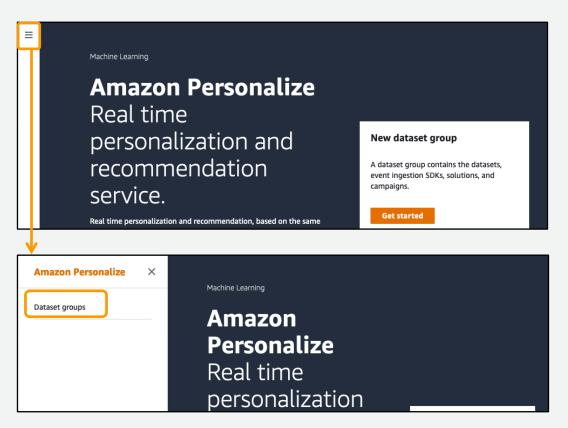
表示される画面で Personalize を検索してクリック





2-2. データセットグループを開く

- 左の量からメニューを 開く
- Dataset groups という メニューが表示される のでクリック





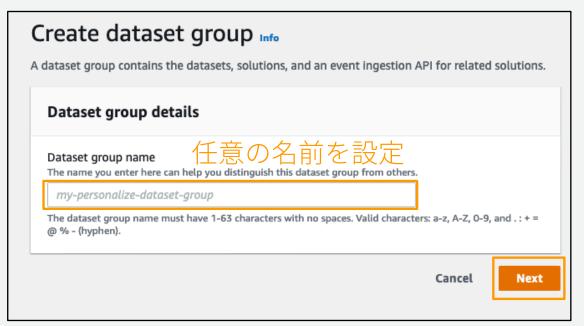
2-3. データセットグループを作成

- Create dataset group を選択してデータセットグループを作成
- Interactions (必須), Users (オプション), Items (オプション) の
 3種のデータセットをデータセットグループに登録可能
- このハンズオンでは、必須の Interactions のみ利用

Amazon Personalize $\qquad imes$	Amazon Personalize > Dataset groups
Dataset groups	Dataset groups (0) Dataset groups are domain-specific containers for your recommendations. For example, you can create one project for an internal application, and another for an external application. View details Delete Create dataset group Q Find dataset group
	\(\lambda \) \(\lambda \) \(\lambda \) \(\text{V} \) \(\text{Vected} \) \(\text{V} \) \(\text{Vected} \) \(Vect
	No dataset groups You don't have any dataset groups



2-4. データセットグループの設定



Next をクリック



2-5. Interaction データの設定

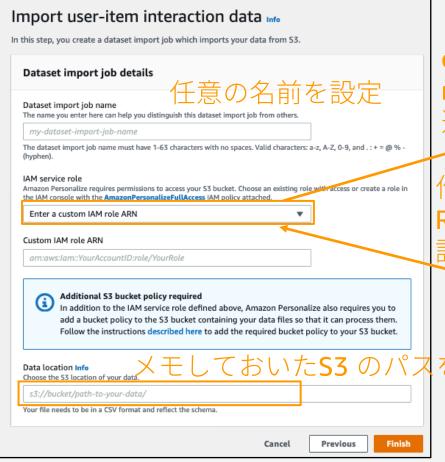


初めて schema を作るので、 Create a new schema として、 【任意の schema name をつける

schema definition はこのまま (CSV をこれにあわせたので)



2-6. データセットのインポート



create a new role

選択

作成した Role を 設定

Create an IAM role

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

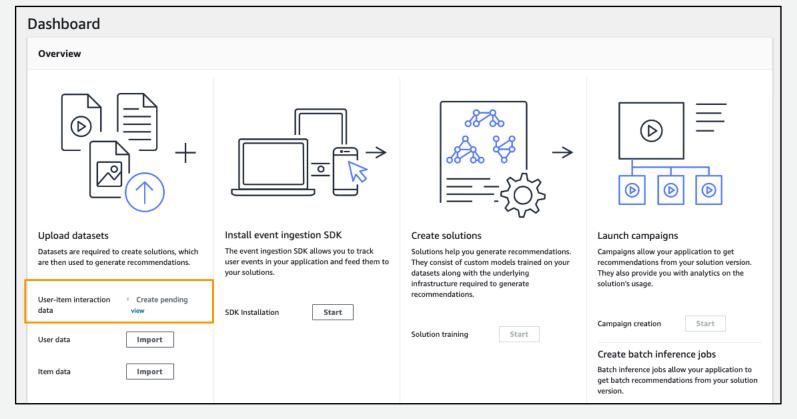
Comma delimited. ARNs, "*" and "/" are not supported.

- Any S3 bucket Any S3 bucket Allow users that have access to your notebook instance access to any pucket and its
- None

 O Any S3 bucket with "personalize" in the name
- Any S3 object with "Personalize" in the name



2-7. インポートの完了を待つ





になる まで待つ

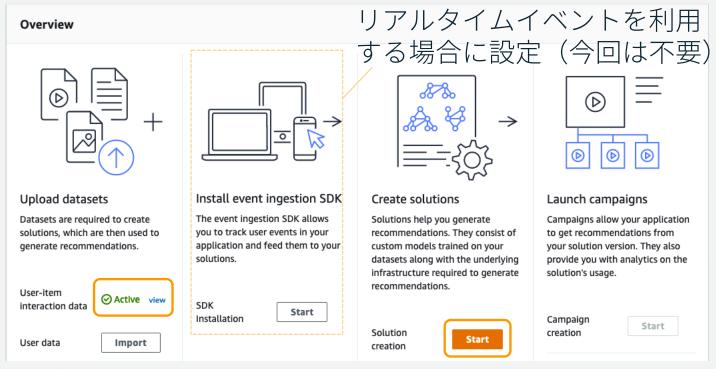


3. ソリューションを学習する



3-1. ソリューションの学習画面へ移動

データインポートが終わると、ソリューション学習を start できる



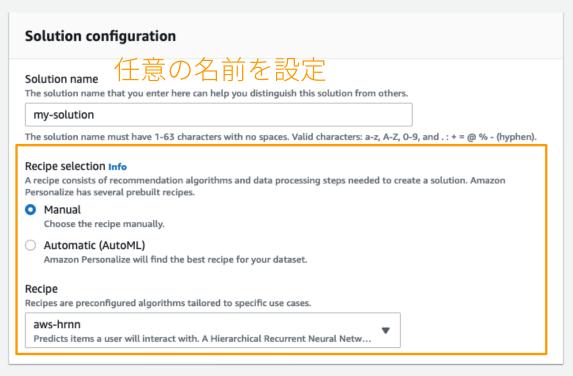
インポート完了

ここから学習を start



3-2. ソリューション学習の設定をする (1/2)

ソリューションの名前設定と学習に利用するレシピを選ぶ

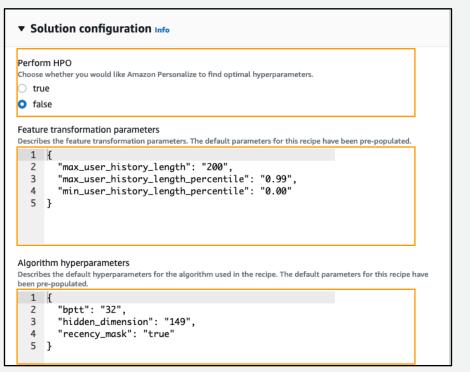


- Manual を設定
- 学習に利用する レシピに aws-hrnn を選択



3-3. ソリューション学習の設定をする (2/2)

- 履歴の長さによって、学習で考慮するユーザを選択する等の設定を行う
- デフォルトのまま、画面下の Next を実行



Perform HPO を行うと 以下を自動調整できる

履歴の長すぎる or 短すぎる ユーザを除去する

ハイパーパラメータを変更して アルゴリズムを調整する

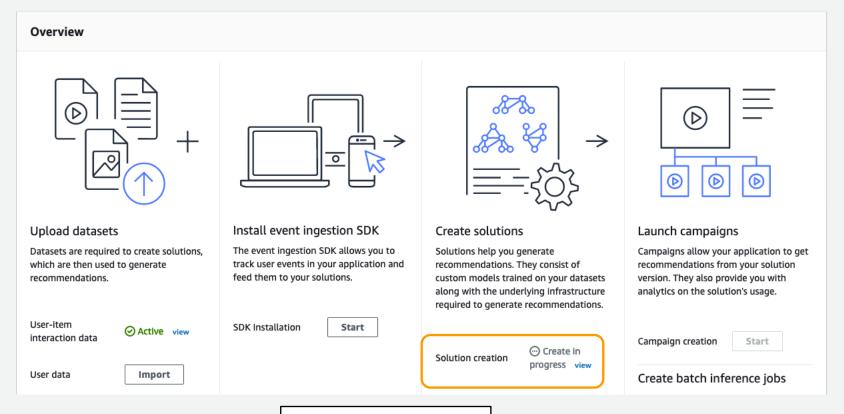
3-4. ソリューション学習の設定を確認する

問題なければ Finish を実行して学習を始める

version	
olution, click Finish button to create	a solution version (train your solution).
Perform HPO	Event type
false	
Perform AutoML	Latest solution version ARN
false	
Recipe	Created
aws-hrnn	Sun, 17 Nov 2019 00:08:35 GMT
	GMI
	Cancel Previous Finish
	Perform HPO false Perform AutoML false Recipe



3-5. 学習完了まで待つ



View Solutions

が表示されるまで待つ

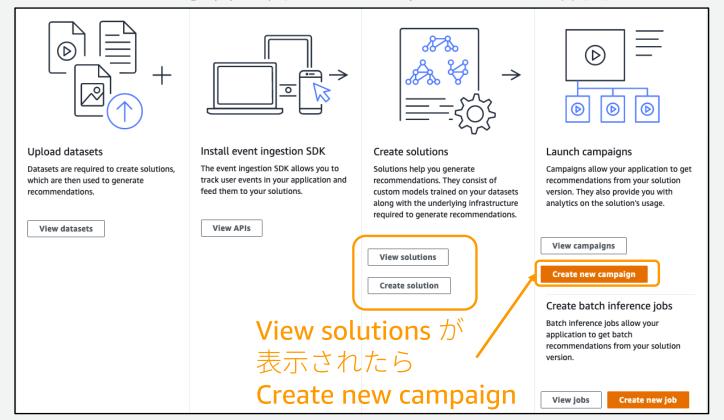


4. キャンペーンを作成して リコメンデーションを行う



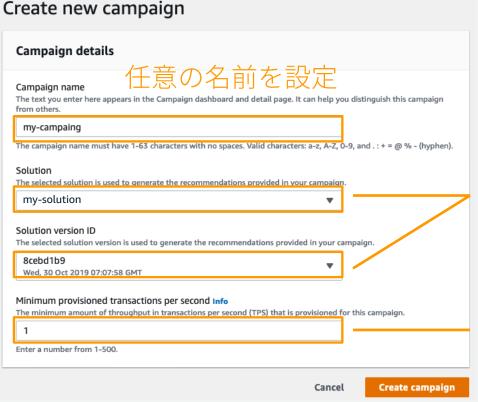
4-1. キャンペーン作成画面へ移動

ソリューションの学習が終わるとキャンペーンを作成できる





4-2. キャンペーンの作成

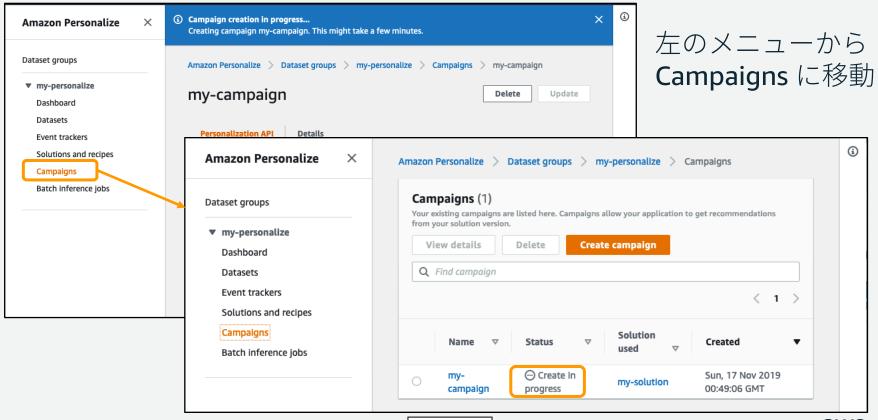


作成したソリューションと ソリューションバージョンを 選択

1秒間にリアルタイムで 応答するリクエスト数 (数に応じた料金がかかる)

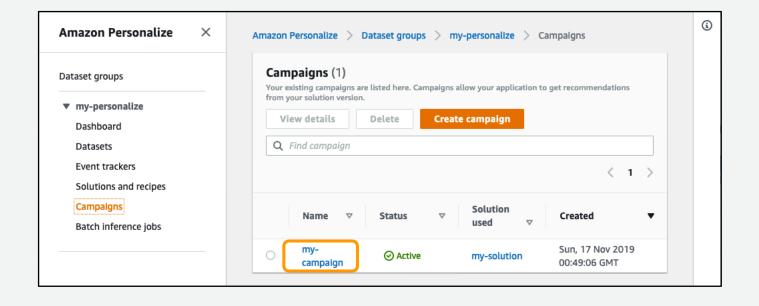


4-3. キャンペーンの作成を待つ



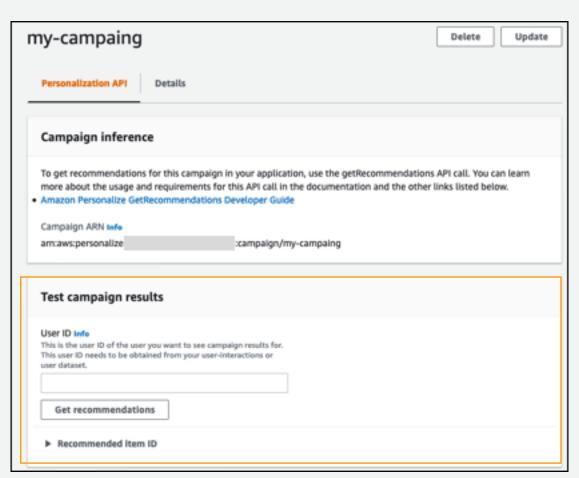
aws

4-4. 作成したキャンペーンを選ぶ





4-5. キャンペーンからリコメンデーションを行う



キャンペーンをテストする 画面が追加されている

- User ID を入力 (例えば1)
- Get recommendations を実行



4-6. 結果の確認

入力した User ID に推薦すべき Item ID のリストが出力される

Test campaign results
User ID Info This is the user ID of the user you want to see campaign results for. This user ID needs to be obtained from your user-interactions or user dataset.
1 Get recommendations
▼ Recommended item ID
104
786
2617
1035
1407
1088



4-7. バッチ推論の開始 (S3への出力)

左のメニューから Batch inference jobs を選択

Amazon Personalize $\qquad imes$	Test campaign results
Dataset groups ▼ test-dataset Dashboard Datasets Info Event trackers	User ID Info This is the user ID of the user you want to see campaign results for. This user ID needs to be obtained from your user-interactions or user dataset. 1 Get recommendations
Solutions and recipes Campaigns Batch inference jobs	▼ Recommended item ID 104 786 2617



4-8. バッチ推論ジョブの入力データ準備

Personalize ではなく手元のPCでの作業です

- リコメンデーション結果を得たいユーザのリストを json で作成
- テキストエディタで以下のファイルを input.json として保存

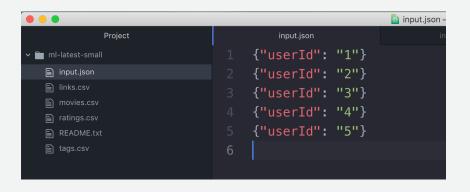
```
{"userId": "1"}

{"userId": "2"}

{"userId": "3"}

{"userId": "4"}

{"userId": "5"}
```



入力形式はアルゴリズムによって異なります https://docs.aws.amazon.com/personalize/latest/dg/gettingrecommendations.html#recommendations-batch



S3 での作業です

- Personalize の画面は開いたまま、S3 にアクセスする https://s3.console.aws.amazon.com/
- Personalize 用に作成したバケットに、input.json をアップロード



S3 での作業です

(1) 7×110) 選択 ② アクセス許可を	アップロー E設定する	ド ③ プロパティを設定	X する 4 確認
	イズ: 97.0 B ターゲットパス: test 大きいファイルをアップロードす		_	3 DEST ADI 左佶田 I .
ます。 詳細は 2 + さらにファ・	. ちら C ⁷ イルを追加する 	SICIA, AWG OLI	AWS SDA, Alliazoli S	THE STAFF ELEMO
	input.json - 97.0 B			*
アップロード				次へ



S3 での作業です

input.json のアップロードを確認

検索します。ESC を押してク	リアします。	
ダウンロード アクシ :	ョン ~	
*		,
	〈 表7	中 1 ~ 2
最終更新日時▼	サイズ▼	ストレージクラ ス ▼
11月 17, 2019 10:29:19 午前 GMT+0900	97.0 B	スタンダード
11月 17, 2019 12:52:42 午前 GMT+0900	2.0 MB	スタンダード
	ダウンロード アクシ : 米 最終更新日時 ▼ 11月 17, 2019 10:29:19 午前 GMT+0900 11月 17, 2019 12:52:42 午前	米国東部 (オー 表示 最終更新日時 ▼ サイズ ▼ 11月 17, 2019 10:29:19 午前 97.0 B GMT+0900 11月 17, 2019 12:52:42 午前 2.0 MB



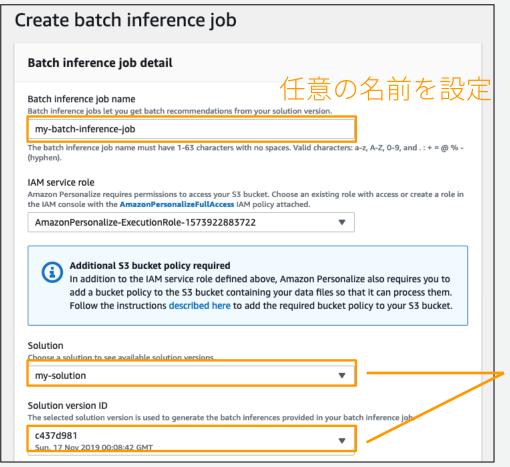
S3 での作業です

input.json のアップロードを確認・選択し、コピーパスしてメモしておく





4-10. バッチ推論ジョブの設定・実行 (1/2)



作成したソリューションと ソリューションバージョンを 選択

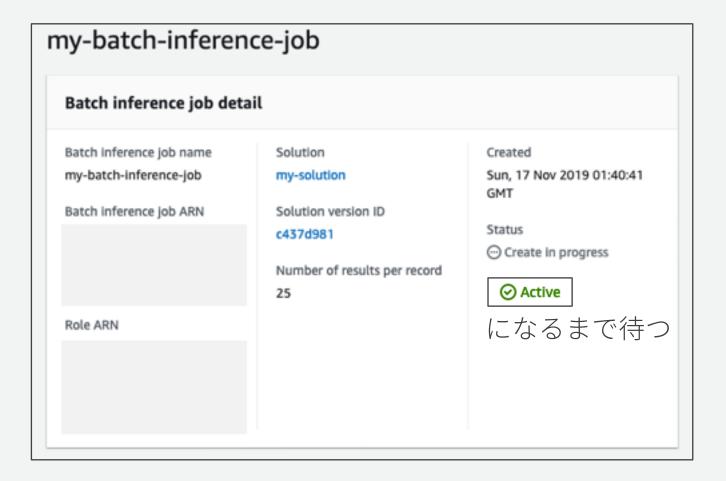


4-11. バッチ推論ジョブの設定・実行 (2/2)

S3 input path Provide the s3 prefix or file pa	th of your input		゚ーパス	した jso n	ファ	イルのバ
s3://test-bucket-person	alize/input.jso	n				
To find a path, go to Amazor	s3.			•		
► Additional configura	tion					
Output data config	uration					
Output data config S3 output path Provide the s3 prefix of your of		作成した	ニバケッ	トへのバ	パス	
S3 output path	output data.	作成した	ニバケッ	トへのバ	パス	
S3 output path Provide the s3 prefix of your o	utput data. alize	作成した	ニバケッ	トへのバ	パス	



4-12. バッチ推論ジョブの完了を待つ





4-13. バッチ推論の結果を確認する

S3 での作業です

- Personalize の画面は開いたまま、S3 にアクセスする https://s3.console.aws.amazon.com/
- Personalize 用のバケットに、出力結果 input.json.out を確認できるので選択





4-13. バッチ推論の結果を確認する

S3 での作業です

「ダウンロード」からファイルをダウンロードして、テキストエディタなどで 結果を確認する

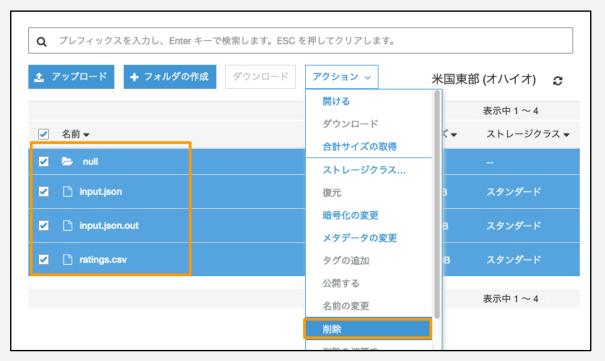


5. リソースを削除する



リソースの削除: S3

- 今回利用したデータ
- エクスポートを実施した場合は、その入出力ファイルも削除





リソースの削除: Amazon Personalize

- 左のメニューから、Campaigns, Solutions and recipes, Datasets を 順に選び、各ページでDelete を実行する
- すべて削除すると、左のメニューから Dataset groups を選んで、 Dataset groupも削除

