

Amazon Forecast ハンズオン (No.Codes)

はじめに：Amazon Forecast は AWS が提供している時系列データをもとに予測を行う機械学習サービスです。

1. リージョンの選択

本ハンズオンで作業を行うリージョンは講師に確認をしてください。画面右上の地名を選択すると変更できます。



2. 学習用データのダウンロード

<https://bit.ly/2Oof8qe>

以下のようなクライアントごとの消費電力データが含まれています。

2014/1/1 1:00	2.538071066	client_0
2014/1/1 1:00	23.64864865	client_1
2014/1/1 1:00	0	client_2
2014/1/1 1:00	144.8170732	client_3
2014/1/1 1:00	75	client_4
2014/1/1 1:00	266.3690476	client_5
2014/1/1 1:00	6.359525155	client_6
2014/1/1 1:00	246.6329966	client_7

このハンズオンでは、この時系列データをもとに将来の電力消費予測を行います。Forecast では、時系列データであればなんでも処理が可能ですので、売り上げデータ、在庫消費データなどを用いることもできます。

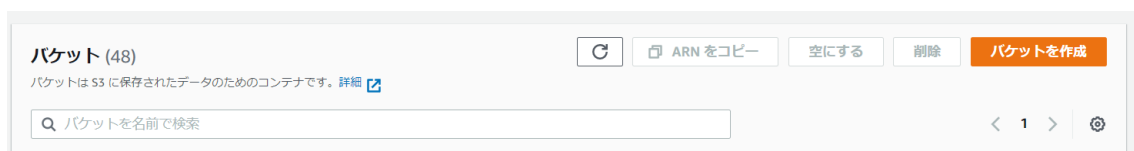
3. S3 バケットの作成

Forecast は学習データを任意の S3 バケットよりインポートして利用するため、先ほどダウンロードした CSV ファイルを格納する S3 バケットを作成します。

3-1. 以下のように[s3]と入力して S3 のマネージメントコンソールにアクセスします。



3-2. 【バケットを作成】を押します。



3-3. 【バケット名】に適切な名前を入れます。

S3 バケットは AWS 全ユーザーの全バケットで一意である必要があるため、ある程度長い名前が必要です。

リージョンは Forecast の作業を行うリージョンと同じものを指定してください。

バケットを作成

バケットは S3 に保存されたデータのためのコンテナです。[詳細](#)

一般的な設定

バケット名

20210214forecast

バケット名は一意的である必要があり、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

3-4. すべてデフォルトで S3 バケットを作りますので画面右下の【バケットを作成】を押してください。

3-5. 作成されたバケットをクリックして開いてください。

20210214forecast

[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。他のユーザーを自分のオブジェクトにアクセスさせるには、明示的にアクセス許可を付与する必要があります。[詳細](#)



削除

アクション ▼

フォルダの作成

アップロード

Find objects by prefix

< 1 > ⚙

名前	タイプ	最終更新日時	サイズ	ストレージクラス
----	-----	--------	-----	----------

オブジェクトがありません

このバケットにはオブジェクトがありません。

アップロード

3-6. 【アップロード】を押します。

S3 にアップロードするファイルとフォルダを追加します。160 GB を超えるファイルをアップロードするには、AWS CLI、AWS SDK、または Amazon S3 REST API を使用します。詳細 [🔗](#)

ここにアップロードするファイルとフォルダをドラッグアンドドロップするか、[Add files] または [Add folders] を選択します。

ファイルとフォルダ (0)

このテーブル内のすべてのファイルとフォルダがアップロードされます。

削除

ファイルを追加

フォルダの追加

🔍 名前を検索

< 1 >

<input type="checkbox"/>	名前	▲	フォルダ	▼	タイプ	▼	サイズ	▼
--------------------------	----	---	------	---	-----	---	-----	---

ファイルまたはフォルダがありません

アップロードするファイルまたはフォルダを選択していません。

3-7. [ファイルの追加]を選んで先ほどダウンロードした

[electricityusedata.csv]を選択します。

ファイルとフォルダ (1 合計, 3.9 MB)

このテーブル内のすべてのファイルとフォルダがアップロードされます。

削除

ファイルを追加

フォルダの追加

🔍 名前を検索

< 1 >

<input type="checkbox"/>	名前	▲	フォルダ	▼	タイプ	▼	サイズ	▼
<input type="checkbox"/>	electricityusedata.csv		-		application/vnd.ms-excel		3.9 MB	

3-8.画面左下の【アップロード】ボタンをおすとアップロード作業が開始されます。



3-9. アップロードが完了すると以下の画面となります。

概要

送信先

s3://20210214forecast

成功しました

✔ 1 ファイル, 3.9 MB (100.00%)

失敗

⊖ 0 個のファイル, 0 B (0%)

ファイルとフォルダ

設定

ファイルとフォルダ (1 合計, 3.9 MB)

🔍 名前を検索

< 1 >

名前	▲	フォルダ	▼	タイプ	▼	サイズ	▼	ステータス	▼	エラー	▼
electricityusagedata.csv		-		application/vnd.ms-excel		3.9 MB		✔ 成功しました		-	

4. IAM ロールの作成

Forecast が S3 へアクセスするために必要な権限を作成します。ここで作成された権限は、後ほど Forecast にアタッチされ、S3 へアクセスできるようになります。

4-1. S3 と同じ要領で IAM のマネージメントコンソールへアクセスします。

aws

サービス ▾ リソースグループ ▾

Identity and Access Management (IAM)

▼ AWS Account (294963776963)

ダッシュボード

グループ

ユーザー

ロール

ポリシー

ID プロバイダー

アカウント設定

認証情報レポート

▼ AWS Organizations

Organization activity

Service control policies (SCPs)

Identity and Access Management へようこそ

IAM ユーザーのサインインリンク:
<https://294963776963.signin.aws.amazon.com/console>
[カスタマイズ](#)

IAM リソース
ユーザー: 6 ロール: 63
グループ: 2 ID プロバイダ: 0
カスタマー管理ポリシー: 43

セキュリティステータス

5 項目中 4 項目が完了しています。

✔ ルートアクセスキーの削除

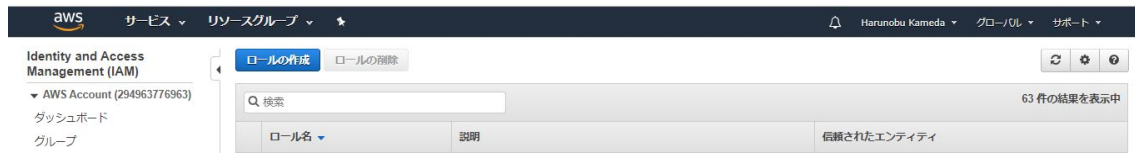
⚠ ルートアカウントの MFA を有効化

✔ 個々の IAM ユーザーの作成

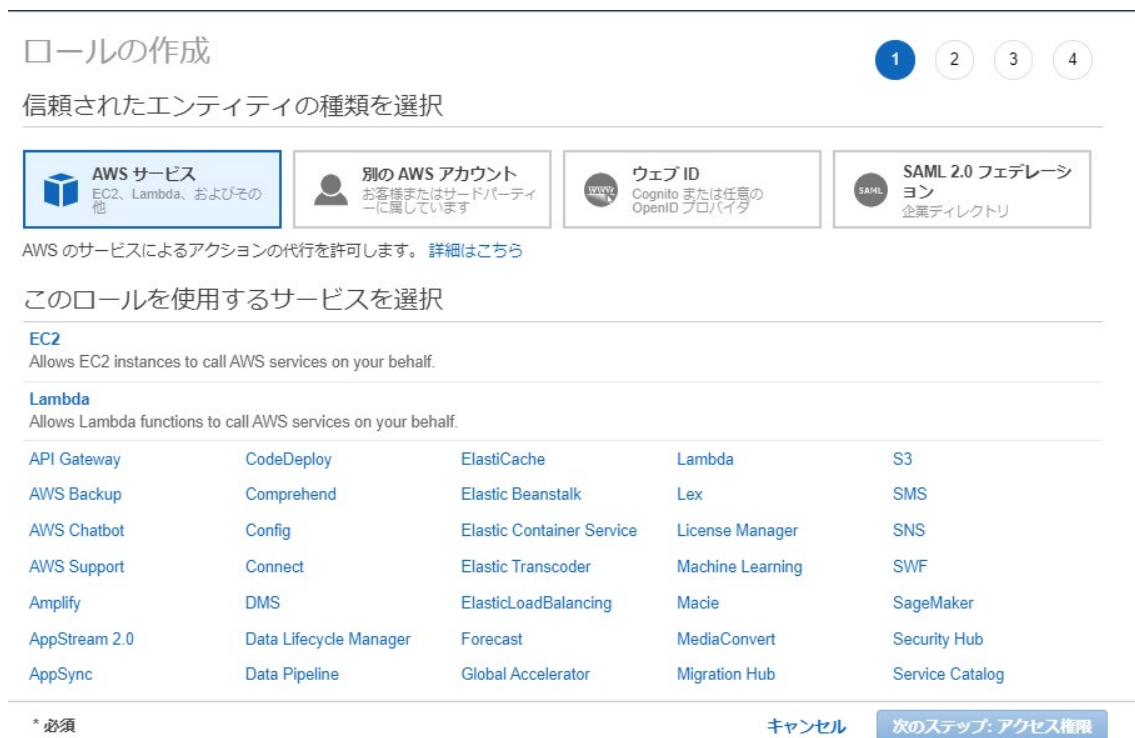
✔ グループを使用したアクセス許可の割り当て

✔ IAM パスワードポリシーの適用

4-2. 画面左のナビゲーションペインでロールを選択します。



4-3. 【ロールの作成】をおします。



4-4. 【このロールを使用するサービスを選択】 から【Forecast】 選び、【次のステップ】 を押します。



4-5. 画面右下の【次のステップ】を押します。その次に表示されるタグの画面はそのまま何も入力せずもう一度【次のステップ】を押します。

ロールの作成 1 2 3 4

確認

以下に必要な情報を指定してこのロールを見直してから、作成してください。

ロール名*

英数字と「+','=','_」を使用します。最大 64 文字。

ロールの説明

最大 1000 文字。英数字と「+','=','_」を使用します。

信頼されたエンティティ AWS のサービス: forecast.amazonaws.com

ポリシー  AmazonS3FullAccess [🔗](#)

アクセス権限の境界 アクセス権限の境界が設定されていません

追加されたタグはありません。

4-6. 【ロール名】に適当な名前を入力し【ロールの作成】を押します。

✔ ロール 20191005forecastsenario が作成されました。 ✕

ロールの作成 ロールの削除 🔄 ⚙️ ⓘ

4-7. 作成が完了しました。作成されたロールをクリックして詳細画面に行きます。後ほどこのロールを Forecase に付与して、Forecast に対する S3 へのアクセス権限を付与しますので【ロール ARN】をメモっておいてください。

ロール > 20191005forecastsenario

概要

ロールの削除

ロール ARN `arn:aws:iam::294963776963:role/20191005forecastsenario` [🔗](#)
ロールの説明 `Allows Amazon Forecast to access data in S3.` [編集](#)
インスタンスプロファイル ARN [🔗](#)
パス `/`
作成時刻 2019-10-05 15:01 UTC+0900
最大 CLI/API セッション期間 1 時間 [編集](#)

アクセス権限 信頼関係 タグ アクセスアドバイザー セッションの無効化

▼ Permissions policies (1 適用済みポリシー)

ポリシーをアタッチします

[+ インラインポリシーの追加](#)

ポリシー名 ▼	ポリシータイプ ▼	
▶ AmazonS3FullAccess	AWS 管理ポリシー	✕

▶ Permissions boundary (not set)

5. S3,IAM と同じように Forecast の画面にいきます。

Machine Learning

Amazon Forecast

Easily build accurate forecasting models

Amazon Forecast is a fully managed service that uses machine learning to deliver highly accurate forecasts. Based on the same machine learning forecasting technology used by Amazon.com.

Create new dataset group

Dataset groups are containers for your datasets, predictors, and forecasts. To create a dataset group, choose Create dataset group.

[Create dataset group](#)

Pricing (US)

Pay only for what you use. There are no minimum fees and no upfront commitments. Different types of charges apply depending on the features of Amazon Forecast that you are using.

[View pricing details](#)

How it works

Historical Data: Sales and traffic, inventory, weather, calendar, etc.

Upload your historical and related data to Amazon Forecast.

Amazon Forecast: Amazon Forecast automatically trains models on the data you upload.

Customized Forecasting: Amazon Forecast generates forecasts for your data.

Forecast can be exported to Amazon S3, Amazon Athena, and Amazon Redshift.

5-1. 【Create dataset group】を押します。

Create dataset group [Info](#)

Dataset groups are containers for all your datasets.

Dataset group details

Dataset group name
The name can help you distinguish this dataset group from other dataset groups on the dataset groups dashboard.

The dataset group name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Forecasting domain [Info](#)
A forecasting domain defines a forecasting use case. You can choose a predefined domain, or you can create your own domain.

Choose a forecasting domain ▼

Cancel **Next**

5-2. 適当な名前を【Dataset group name】に入力し、【Forecasting domain】を【Custom】に設定し【Next】を押します。（数字から始まる名前はエラーとなりますので気を付けてください）

Create target time series dataset [Info](#)

Dataset details

Dataset name
The name can help you distinguish this dataset from other datasets on your Datasets dashboard.

The dataset name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Frequency of your data
This is the frequency at which entries are registered into your data file.

Your data entries have a time interval of 1 ▼

day ▼

Data schema [Info](#)

5-3. 【Dataset name】に適切な名前をいれます。【Frequency of your data】は【hour】を選びます。（この値は学習させる元データの時系列間隔と合わせる必要があります。）

5-4. [Data Schema]で JSNO を選び、【Data Schema】の値を学習させる元データの並び順に合わせるため以下のように入れ替えます。

Data schema [Info](#)

Use the data schema section to specify the attribute types for each column in your dataset. You can specify the schema in two ways:

☐ Schema builder

Specify your Attribute Name, Attribute Type, and attribute order in the text boxes provided.

☒ JSON schema

Specify AttributeName and AttributeType in the JSON format.

```
1 {
2   "Attributes": [
3     {
4       "AttributeName": "timestamp",
5       "AttributeType": "timestamp"
6     },
7     {
8       "AttributeName": "target_value",
9       "AttributeType": "float"
10    },
11    {
12      "AttributeName": "item_id",
13      "AttributeType": "string"
14    }
15  ]
16 }
```

5-5. [Dataset Import details]の欄に同様に適切な名前を入れます。

Dataset import details

Dataset import name
The name can help you distinguish this dataset import from other imports on your dataset detail page.

The dataset import name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Select time zone [Info](#)
Select a time zone for your dataset.

Do not use time zone ▼

Data location [Info](#)
The location is the path to the file in your S3 bucket that contains your data.

View

Browse S3

Your files must be in CSV format.

IAM role [Info](#)
Dataset groups require permissions from IAM to read your dataset files in S3. Choose or create a role using this control.

Enter a custom IAM role ARN ▼

Custom IAM role ARN

5-6. [Data location]で先ほど作成した S3 バケットとアップロード下ファイルを以下のように設定します。

Data location [Info](#)
The location is the path to the file in your S3 bucket that contains your data.

X

View

Browse S3

Your files must be in CSV format.

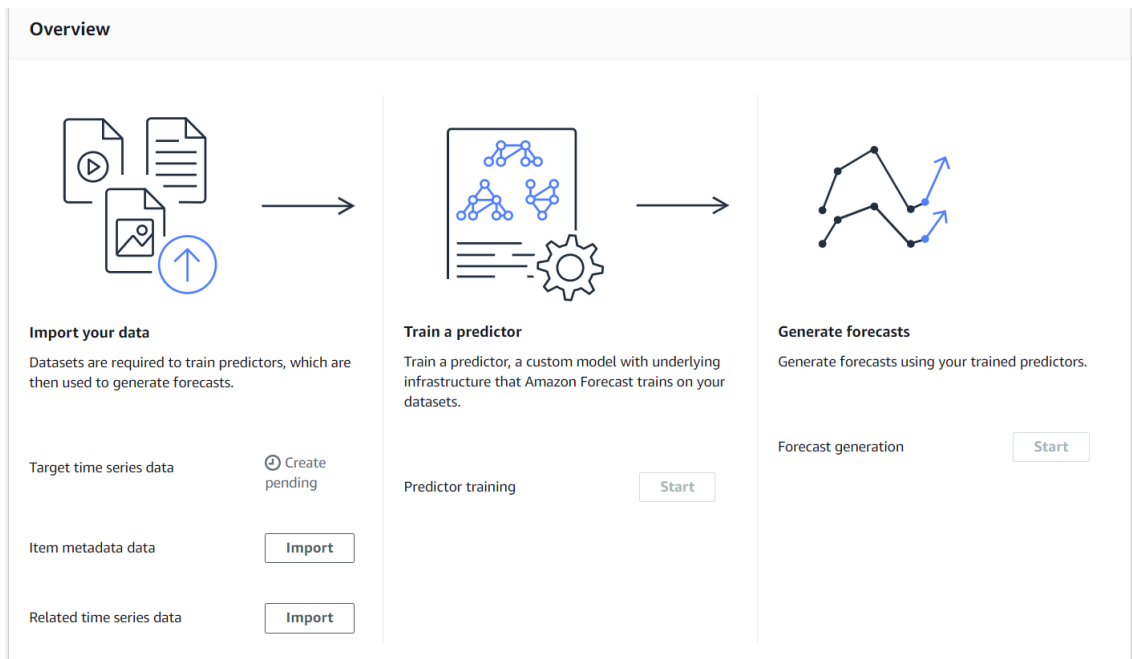
5-7.[Custom IAM role ARN]に作ほどコピーした新しく作成した IAM Role の ARN を貼り付けます。

IAM role [Info](#)
Dataset groups require permissions from IAM to read your dataset files in S3. Choose or create a role using this control.

Enter a custom IAM role ARN ▼

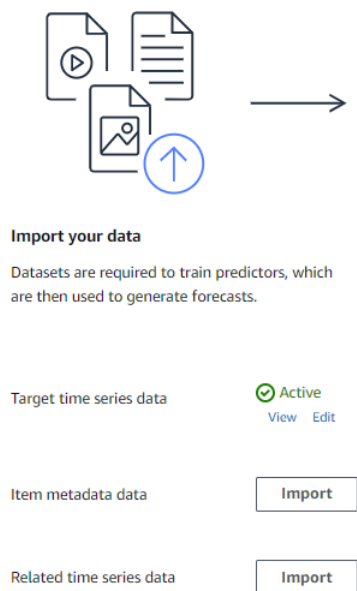
Custom IAM role ARN

5-9. [Start Import]を押すと、以下の画面に遷移します。



待ち時間の間に Personalize のハンズオンを行います。

5-10. 以下のように【Active】と緑で表示されればインポート完了です。



6. 学習

データのインポートが完了したので、学習を行います。

6-1. 【Train a predictor】の【Start】を押します。

aws サービス リソースグループ

Amazon Forecast > Dataset groups > forecastsenario > Train predictor

Train predictor info

Amazon Forecast uses an algorithm to train a predictor on the data in your dataset group.

Predictor details

Predictor name
The name can help you distinguish this predictor from your other predictors.

The predictor name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Forecast horizon info
This number tells Amazon Forecast how far into the future to predict your data at the specified forecast frequency.

Forecast frequency
This is the frequency at which your forecasts are generated.
Your forecast frequency is

6-2. 【Predictor name】に適当な名前を入れます。【Forecast horizon】は予測を行う時間間隔をしています。このハンズオンでは 36（36 時間分）の予測を作成するため、【36】と入力します。

Predictor settings

Predictor name
The name can help you distinguish this predictor from your other predictors.

The predictor name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Forecast horizon info
This number tells Amazon Forecast how far into the future to predict your data at the specified forecast frequency.

Forecast frequency
This is the frequency at which your forecasts are generated.
Your forecast frequency is

6-3. 【Forecast frequency】は元データと同じである必要があるため【hour】を選びます。【Algorithm selection】は【Manual】とし【Algorithm】は【ETS】を選びます。

Algorithm selection **Info**

An algorithm is used to train your predictor.

☐ Automatic (AutoML)
Let Amazon Forecast choose the right algorithm for your dataset.

☒ Manual
Explore the algorithms and choose one.

Algorithm

The algorithm that you want Amazon Forecast to use to train your predictor.

☐ CNN-QR
An Amazon proprietary convolutional neural network algorithm. Useful for 100+ time series and flexible with all forms of related time series and item metadata.

☐ DeepAR+
An Amazon proprietary recurrent neural network algorithm. Useful for 100+ time series with forward-looking related time series and item metadata.

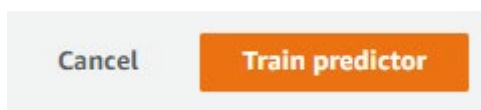
☐ Prophet
A Bayesian structural time series model. Useful for time series with strong seasonal effects and several seasons of historical data.

☐ NPTS
Non-Parametric Time Series. An Amazon proprietary algorithm useful for sparse and intermittent time series.

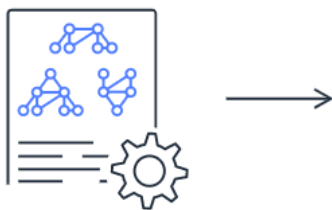
☐ ARIMA
Autoregressive Integrated Moving Average. A statistical algorithm useful for datasets with fewer time series.

☒ ETS
Exponential Smoothing. A statistical algorithm useful for datasets with fewer time series and seasonality.

6-4. 残りの値は全てそのまま【Train predictor】を押します。



6-5. 以下が表示されたら学習が開始されています。



Train a predictor

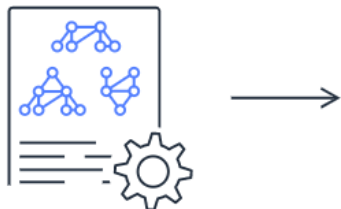
Train a predictor, a custom model with underlying infrastructure that Amazon Forecast trains on your datasets.

Predictor training

⋮ Create in progress...
[View](#)

待ち時間の間に Personalize のハンズオンを行います。

6-6. 以下のように【Active】と表示されれば学習が完了です。



Train a predictor

Train a predictor, a custom model with underlying infrastructure that Amazon Forecast trains on your datasets.

Predictor training

✔ Active
View

7. 予測エンドポイント作成

学習が開始したので予測を行うエンドポイントを作成します。

7-1. 画面右の【Generate forecasts】のところで【Start】ボタンを押します。

Create a forecast [Info](#)

Use a predictor to create forecasts based on your datasets.

Forecast details

Forecast name

The name can help you distinguish this forecast from your other forecasts.

The forecast name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Predictor [Info](#)

This is the predictor that you want to use to create forecasts.

Cancel

Create a forecast

- 7-2. 【Forecast name】に適当な名前を入れます。【Predictor】のドロップダウンから先ほど学習環境につけた名前を選びます。
- # 何も表示されない場合は、cancel を押して再度やり直すと表示されます。

Create a forecast [Info](#)

Use a predictor to create forecasts based on your datasets.

Forecast details

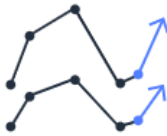
Forecast name
The name can help you distinguish this forecast from your other forecasts.

The forecast name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Predictor [Info](#)
This is the predictor that you want to use to create forecasts.

[Cancel](#) [Create a forecast](#)

7-3. 【Create new forecast】のボタンを押します。



Generate forecasts

Generate forecasts using your trained predictors.

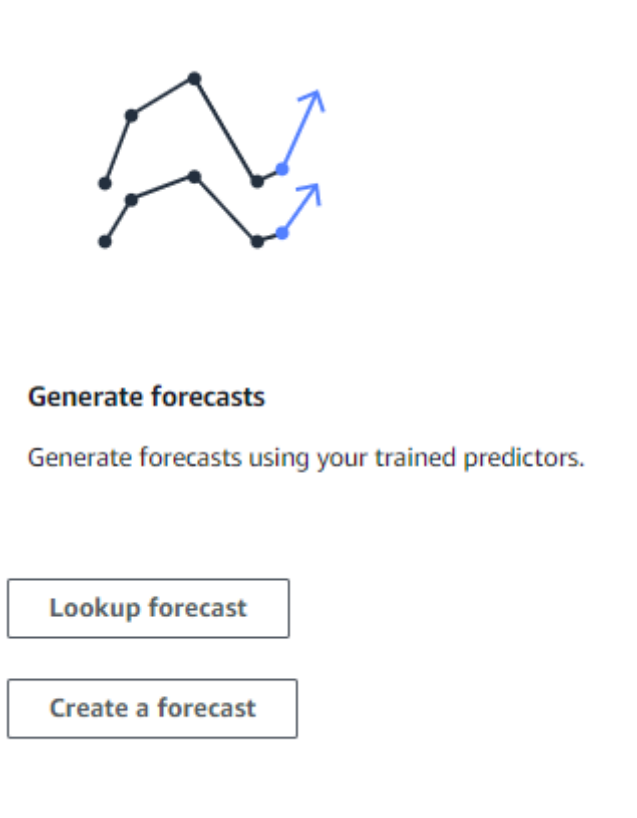
Forecast generation

🕒 Create
pending

7-4. 予測エンドポイントの生成が開始されるので少し待ちます。

待ち時間の間に Personalize のハンズオンを行います。

以下の表示になれば完了です。予測ができるようになりました。



8. 予測を行う

以上の手順で学習済データから予測を行うことができます。

【Look up Forecast】を押します。

Amazon Forecast > Dataset groups > forecastsenario > Forecast lookup

Forecast lookup [Info](#)

When you create a forecast, Amazon Forecast generates forecasts for each unique item in your target time-series dataset. Use the forecast lookup to find your forecasts.

Forecast details

Forecast Choose the forecast whose forecasts you want to view. <div>Choose a forecast name ▼</div>	Start date This is the start date for the forecast that you want to view. The date must be later than the latest entry for your item. <div>YYYY/MM/DD</div> <div>00:00:00</div> <div>Use 24-hour format.</div>	End date This is the end date for the forecast that you want to view. The date must be later than the start date and earlier than the start date plus the forecast horizon. <div>YYYY/MM/DD</div> <div>00:00:00</div> <div>Use 24-hour format.</div>
---	--	--

Get Forecast

8-1. 【Forecast】に先ほど作成したエンドポイントの名前をドロップダウン

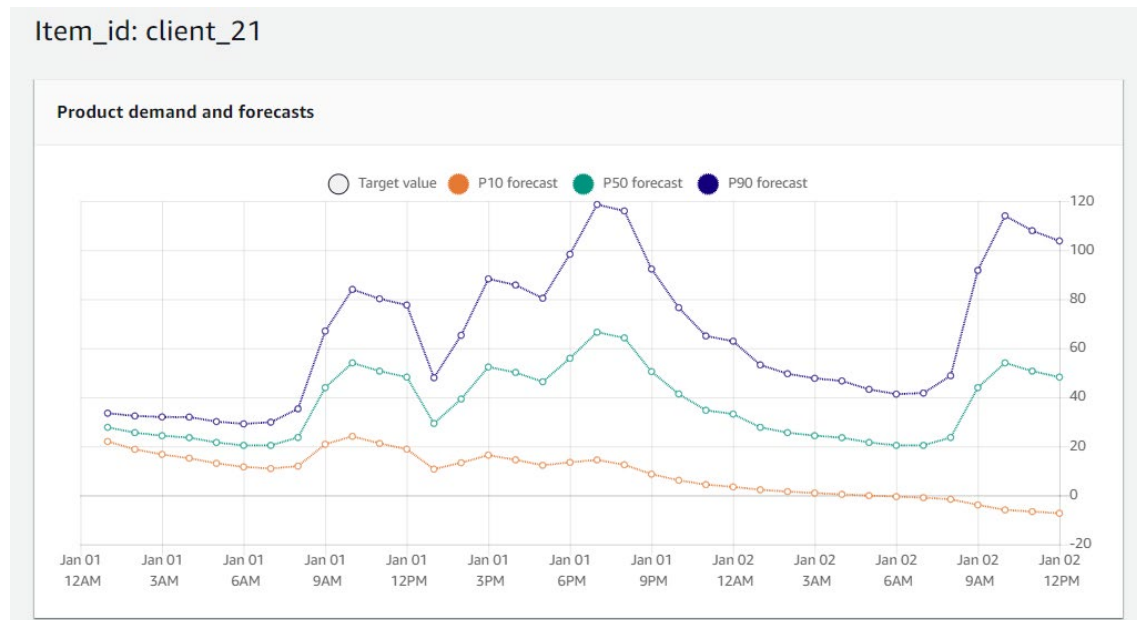
から選びます。

- **【開始日】** – **2015/01/01** と入力します。デフォルト時間である 00:00:00 のままにします。
- **【終了日】** – **2015/01/02** と入力します。時間を 12:00:00 に変更します。

【Value】には学習元データに入っているクライアントの ID を入力します。

(例 : client_2)

8-2. **【Get Forecast】** を押すと以下のように予測がグラフで出力されます。



P10、P50、P90 の数値は、それぞれ 10%、50%、90% の確率で、実際の需要量が含まれる値になります。90%の確率でその範囲内 (値より下) に入

る、という意味になっています。P××の数字が大きくなれば大きくなるほど予測がその値に収まる確率は高まるものの、ブレ幅が大きくなっていきますので、まずは P50 の値を用いることをお勧めしています。

9. お疲れ様でした！

削除は環境の構築順番と逆でおこなっていきます。

Forecasts→Predictors→Datasets→Dataset groups

S3 バケット