

JAPAN | JUNE 20, 2024

aws SUMMIT



DOL-11

時系列データ活用の課題を一掃! AWS による時系列データ解析

Shoji Takano

Solutions Architect
Amazon Web Service Japan G.K.

Hiroki Tsuji

Solutions Architect
Amazon Web Service Japan G.K.



自己紹介

高野 翔史 (たかの しょうじ)

アマゾンウェブサービスジャパン

ソリューションアーキテクト

製造業のお客様を担当

好きなサービス：

Amazon Timestream



自己紹介

辻 浩季 (つじ ひろき)
アマゾンウェブサービスジャパン
ソリューションアーキテクト
製造業のお客様を担当

好きなサービス：
Amazon SageMaker



Agenda

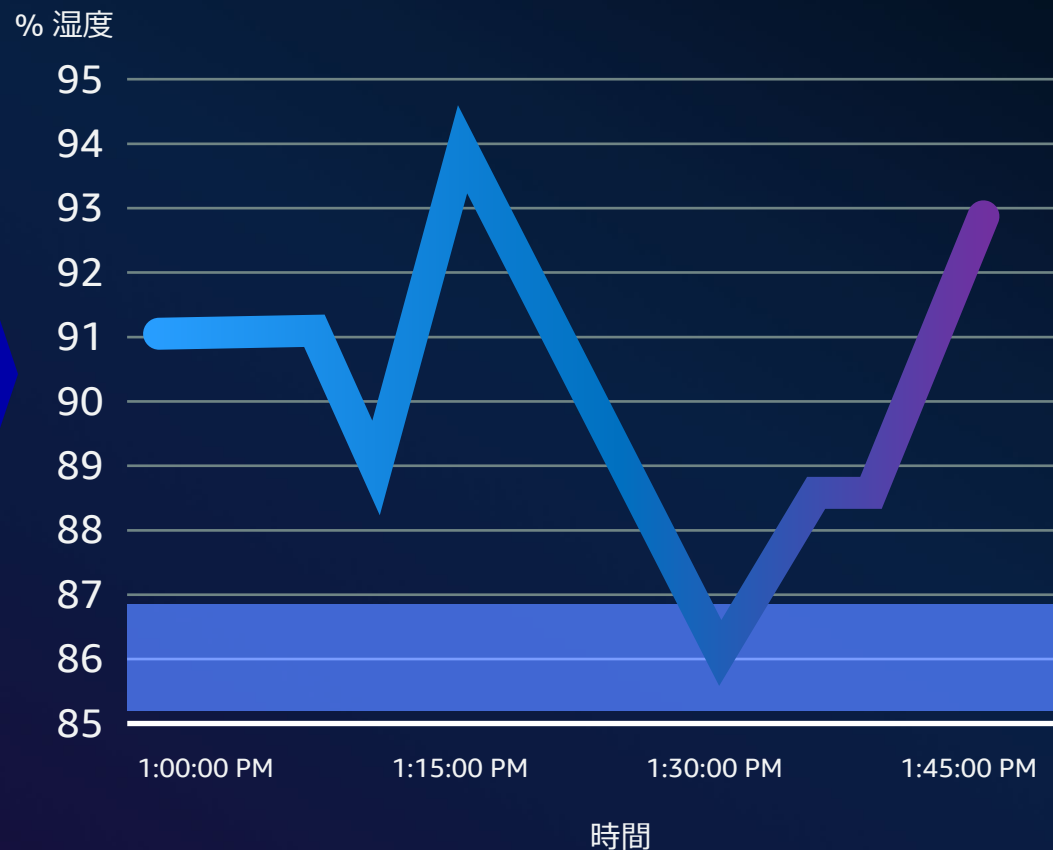
- 時系列データのユースケース
- 時系列データを扱う上でのよくあるご要望
- 時系列データを扱う AWS サービスの紹介
- まとめ

時系列データのユースケース

時系列データとは

時間経過とともに変化する事象を測定するために、
タイムスタンプとともに記録されたシーケンスデータ

timestamp, 測定情報, 属性, ...
2023-12-08 13:00:00, 91, Tokyo ...
2023-12-08 13:15:00, 94, Tokyo ...
2023-12-08 13:30:00, 86, Tokyo ...
⋮



湿度
94.0
%

時系列データのユースケース

現状の傾向を知る

これから先起きる未来を予測する



IoT applications



センサー
ネットワーク監視



在庫計画



DevOps analysis



クリックストリーム
分析



商品需要予測



App analysis



物理システムの
モニタリング



アプリケーションの
モニタリング

時系列データを扱う上での よくあるご要望

時系列データを扱う上でのよくあるご要望

時系列データを
収集する

- ✓ 大量のデータ書き込みをしたい
- ✓ データスキーマ変更に対応したい

時系列データを
保管する

- ✓ ストレージコストをリーズナブルにしたい

時系列データを
分析・可視化する

- ✓ 時系列専用の分析機能を利用したい
- ✓ 時系列予測の知見がなくても予測したい

時系列データを扱う AWS サービスの紹介



時系列データを効率的に扱う AWS サービス

Amazon Timestream for LiveAnalytics



時系列データを専用に扱う
サーバレスデータベースサービス

Amazon SageMaker Autopilot



データに基づいて最適なモデルを
自動的に構築 (AutoML)

時系列データを扱う上でのよくあるご要望に対する Timestream と SageMaker の対応範囲

時系列データを
収集する

時系列データを
保管する

時系列データを
分析・可視化する



- ✓ 大量のデータ書き込みをしたい
- ✓ データスキーマ変更に対応したい



- ✓ ストレージコストをリーズナブルにしたい



- ✓ 時系列専用の分析機能を利用したい

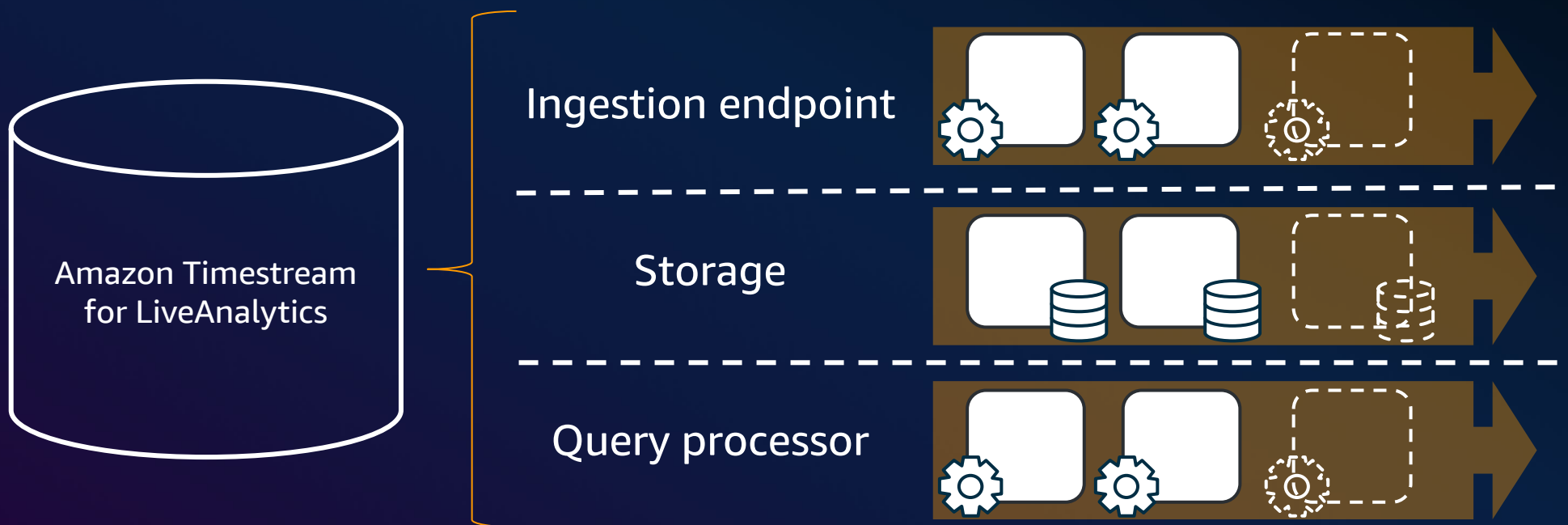


- ✓ 時系列予測の知見がなくても予測したい

Amazon Timestream for LiveAnalytics アーキテクチャ

大量のデータ
書き込みをしたい

データ挿入・ストレージ・クエリプロセッシングが分離されたアーキテクチャ構成。
それぞれ独立して拡張され、1日あたり数兆ものイベントに対応可能。



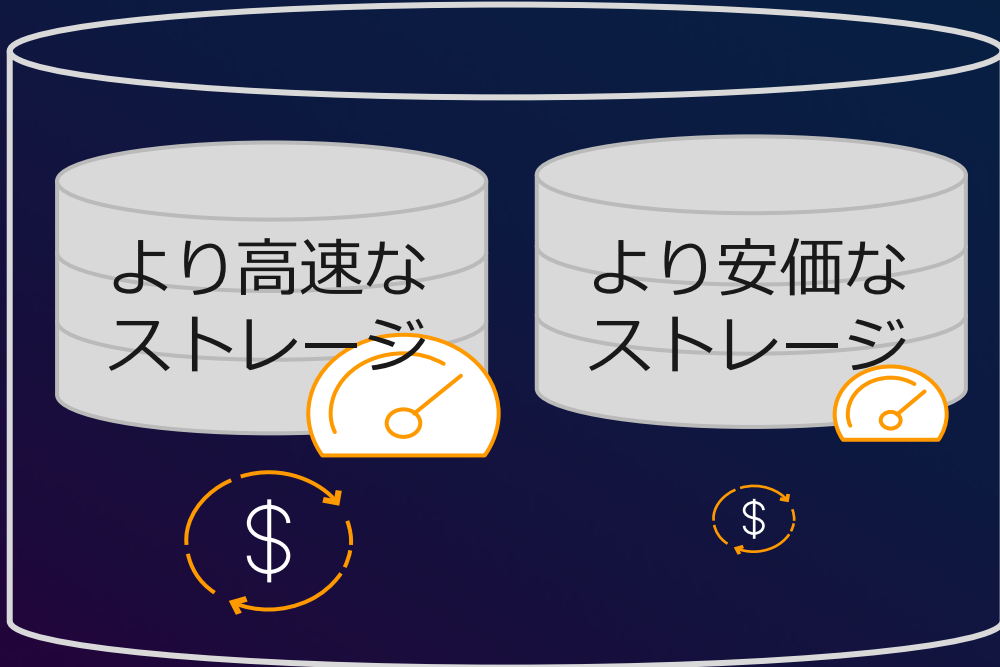
容量とパフォーマンス状況に合わせ、個別に自動でスケールリング

コスト最適化

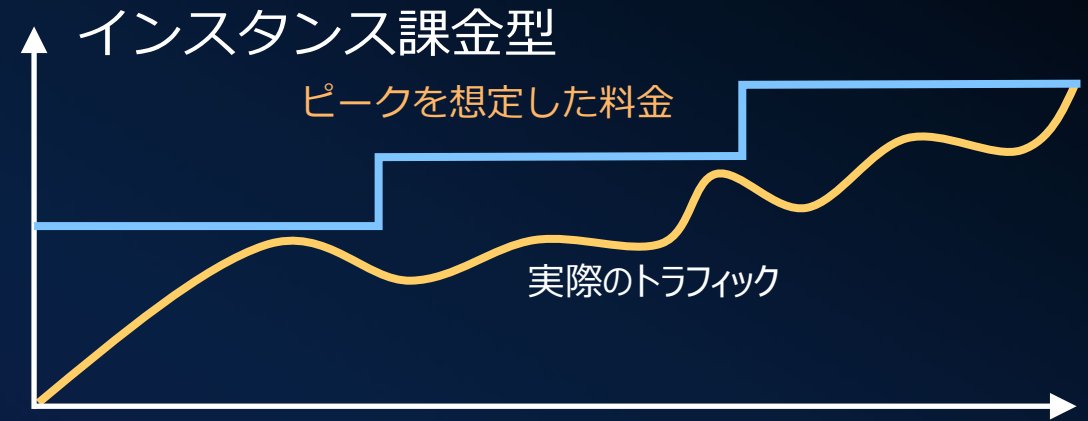
ストレージコストを
リーズナブルにしたい

パフォーマンス要求に応じてコストが最適化された異なるストレージを提供

Amazon Timestream for LiveAnalytics



取り込んだデータ量、保存したデータ量、クエリする処理時間に応じた利用料金



時系列専用のデータモデルと関数

データスキーマ変更に対応したい

時系列専用の分析機能を利用したい

時系列データ指向で、事前にスキーマ定義不要の柔軟なデータモデルで、平滑化/近似/補間などの時系列関数を提供
アプリケーション側で複雑な計算ロジックの対応が不要に

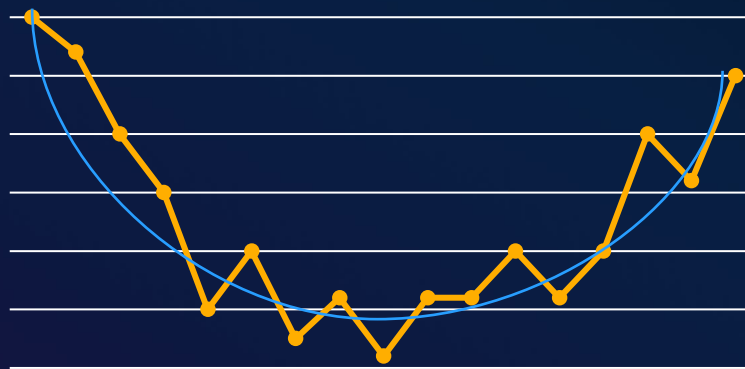
時系列データモデル



- ✓ タイムスタンプをキーに、属性情報の付加が柔軟に可能なデータモデル
- ✓ パーティションや索引も自動管理

時系列専用関数

平滑化関数



- ✓ ノイズやその他の微細なポイント、急激な現象を排除しながらデータの重要なパターンを算出し、結果データを返す関数

補間関数



- ✓ 欠落している値を補間したデータを返す関数

時系列データを扱う上でのよくあるご要望に対する Timestream と SageMaker の対応範囲

時系列データを
収集する

時系列データを
保管する

時系列データを
分析・可視化する



- ✓ 大量のデータ書き込みをしたい
- ✓ データスキーマ変更に対応したい



- ✓ ストレージコストをリーズナブルにしたい



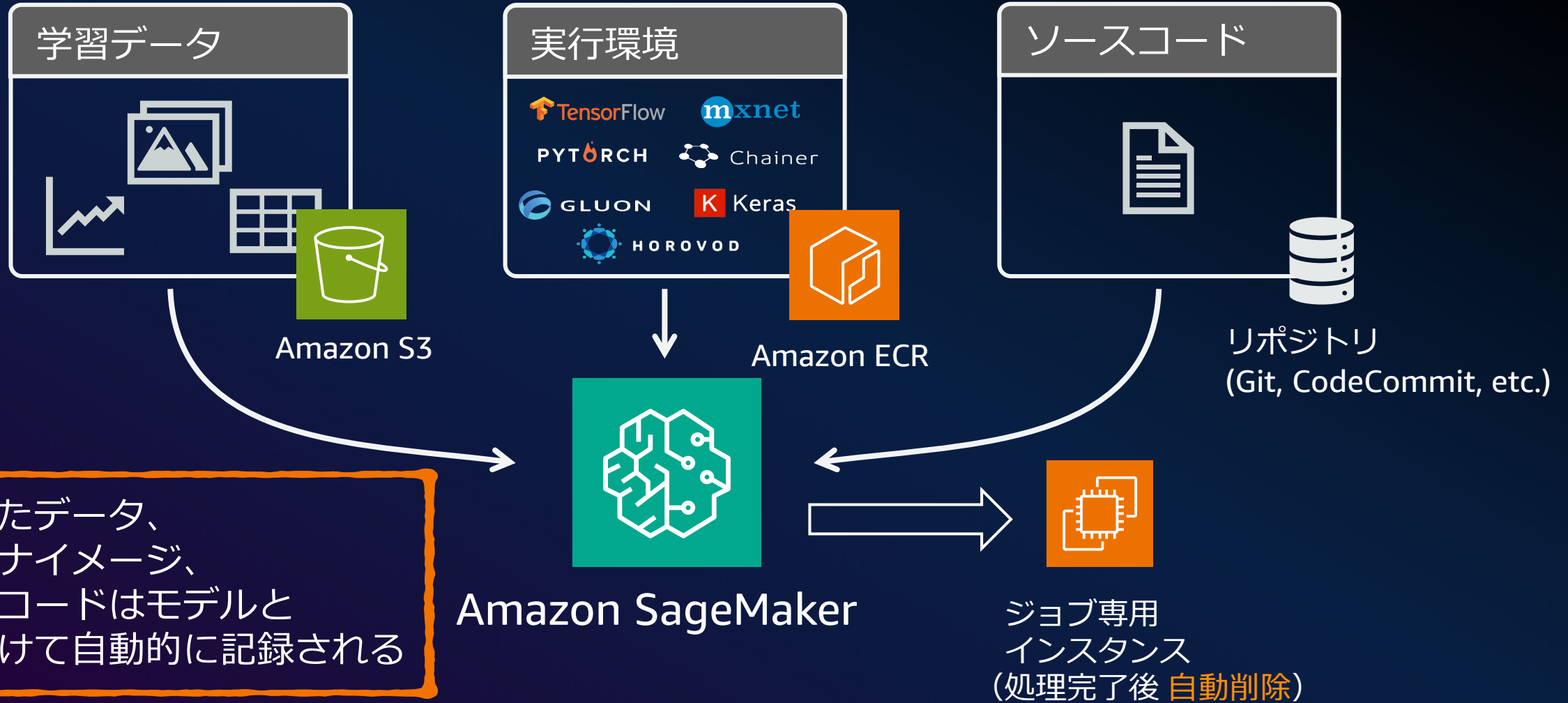
- ✓ 時系列専用の分析機能を利用したい



- ✓ 時系列予測の知見がなくても予測したい

Amazon SageMaker

必要なリソースを自由に組み合わせて学習・推論を実行





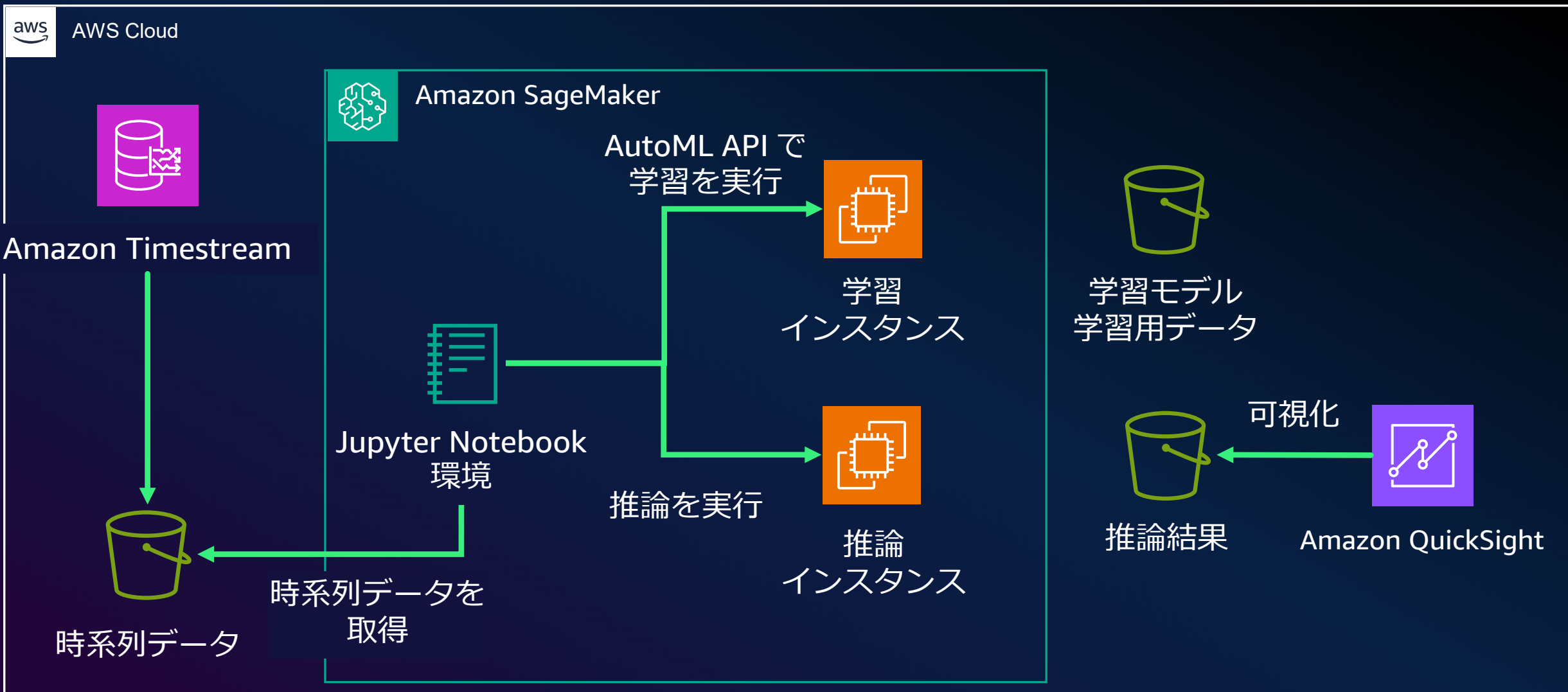
Amazon SageMaker Autopilot を用いた時系列予測

- **AutoML REST API v2 を呼び出すだけ**で、時系列予測モデルを作成できる
 - ✓ **多変量時系列予測**も可能
- API 呼び出し時のリクエストパラメータに予測頻度、予測分位数などの設定を書く

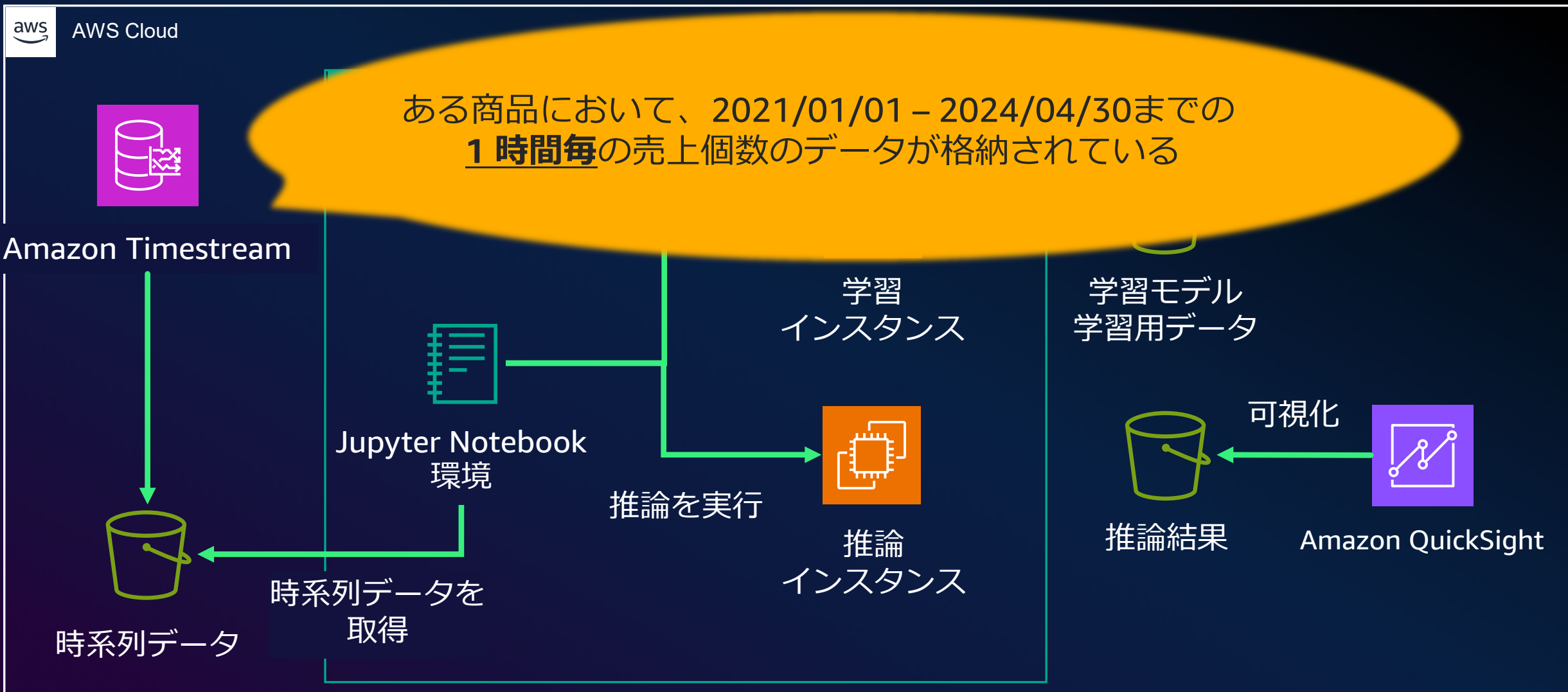
```
"AutoMLProblemTypeConfig": {  
  "ForecastFrequency": "D",  
  "ForecastHorizon": 20,  
  "ForecastQuantiles": ['p10', 'p30', 'p50', 'p70', 'p90'],  
  "TimeSeriesConfig": {  
    "TargetAttributeName": "demand",  
    "TimestampAttributeName": "timestamp",  
    "ItemIdentifierAttributeName": "item_id"  
  },  
}
```

- 6 種類のアプローチにより計算されたモデルからスタッキングアンサンブル法を用いてモデル候補を組み合わせ、**最適なモデルを自動で作成**

デモ：ある商品の日次の売上個数予測



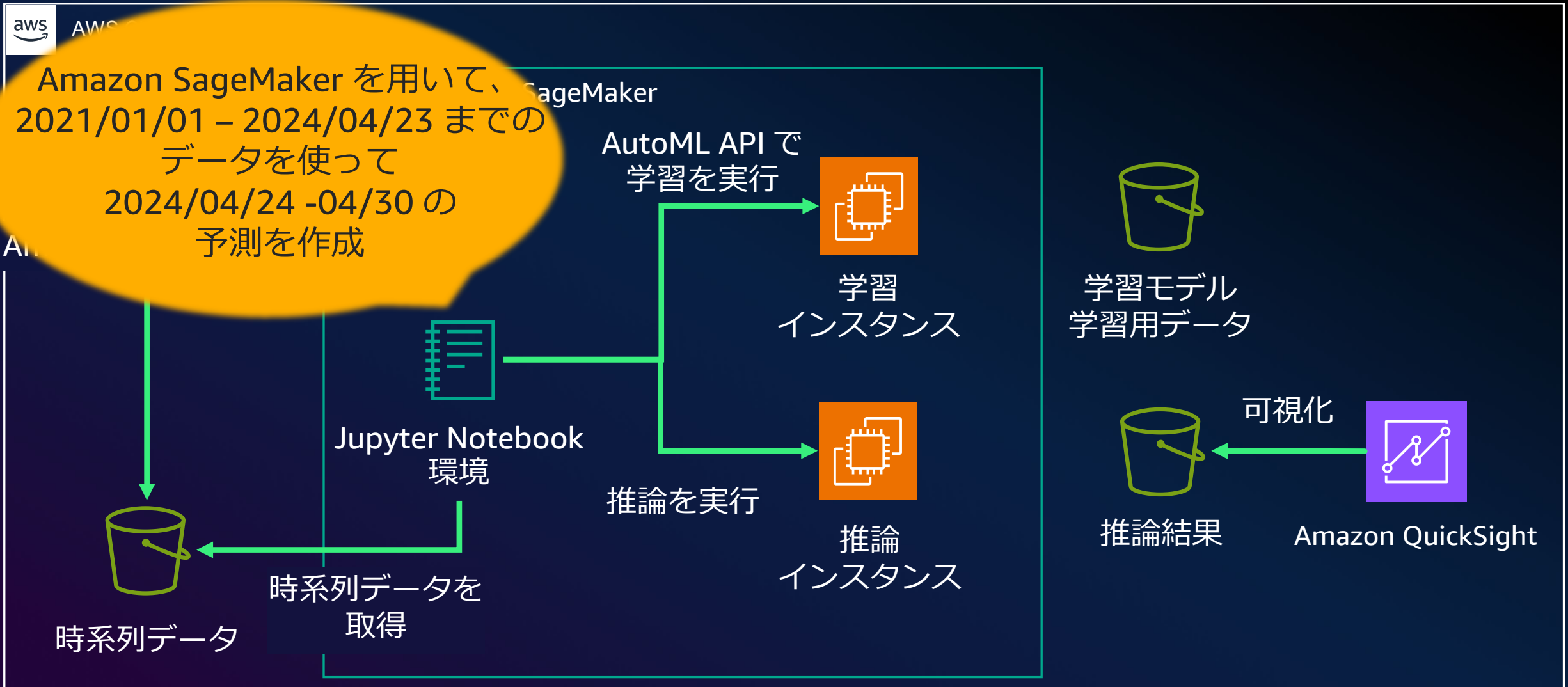
デモ：ある商品の日次の売上個数予測



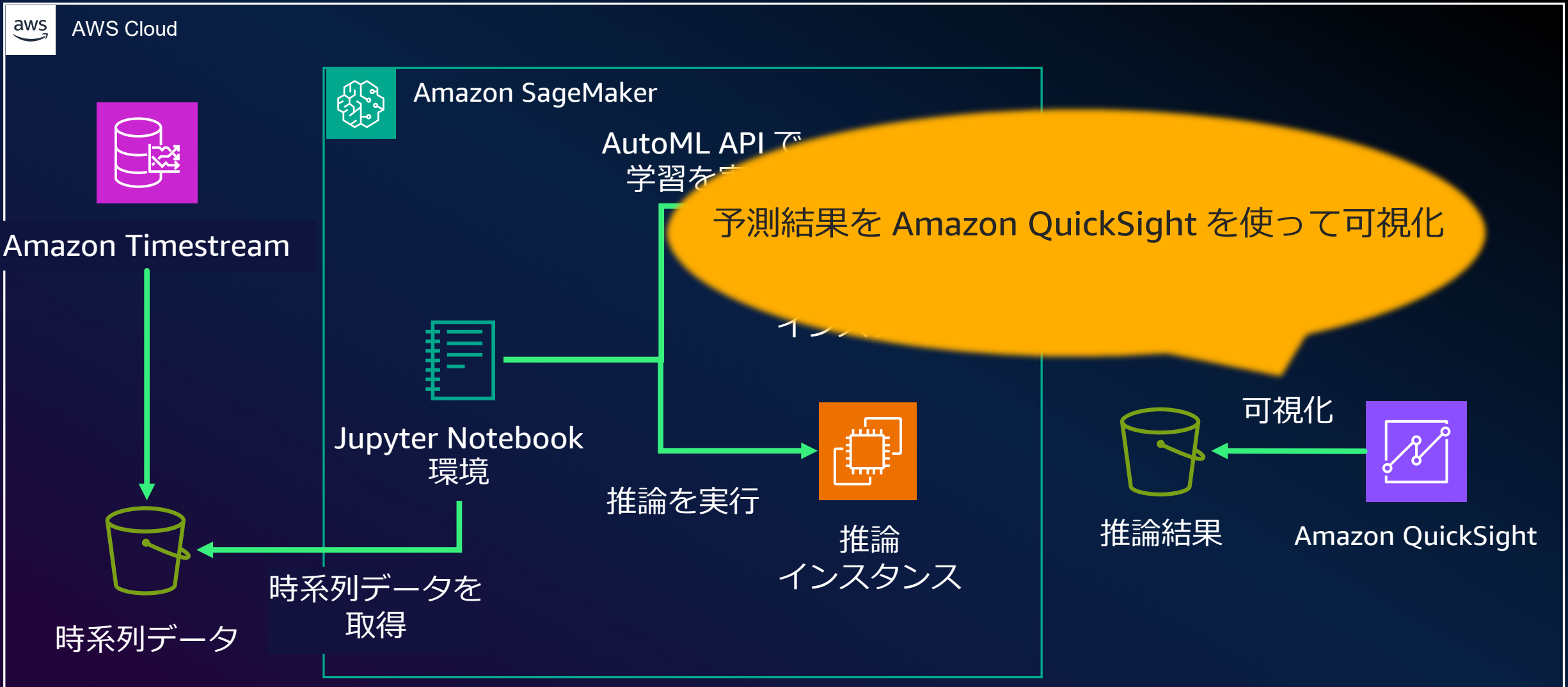
デモ：ある商品の日次の売上個数予測



デモ：ある商品の日次の売上個数予測



デモ：ある商品の日次の売上個数予測



デモ：ある商品の日次の売上個数予測

デモの注目ポイント

- Amazon Timestream でデータをクエリする際、
時系列関数を使って SQL で簡単にデータを集計できる！
- Amazon SageMaker Autopilot の AutoML API を用いて、
API 呼び出しのコードを書くだけで簡単に時系列予測を実現できる！



bin 関数

bin(timestamp, interval)

→ timestamp (時刻) を、指定した interval (例 : 1秒、5分、1時間) で切り下げを行う。

bin(TIMESTAMP '2024-06-20 10:15:20', 1m) → 2024-06-20 10:15:00.000000000
bin(TIMESTAMP '2024-06-20 10:15:20', 1d) → 2024-06-20 00:00:00.000000000

bin 関数の使い方例

```
SELECT BIN(time, 1d) AS DAY,  
       measure_name,  
       sum(sales) AS SUM  
FROM "timeseries_demo_db"."timeseries_demo_table"  
WHERE measure_name = 'product_a' AND region = 'location_a'  
GROUP BY BIN(time, 1d), measure_name  
ORDER BY DAY
```

DAY	measure_name	SUM
2024-06-20 10:00:00	product_a	10
2024-06-20 11:00:00	product_a	12
⋮	⋮	⋮
2024-06-20 23:00:00	product_a	14
2024-06-21 00:00:00	product_a	15
⋮	⋮	⋮

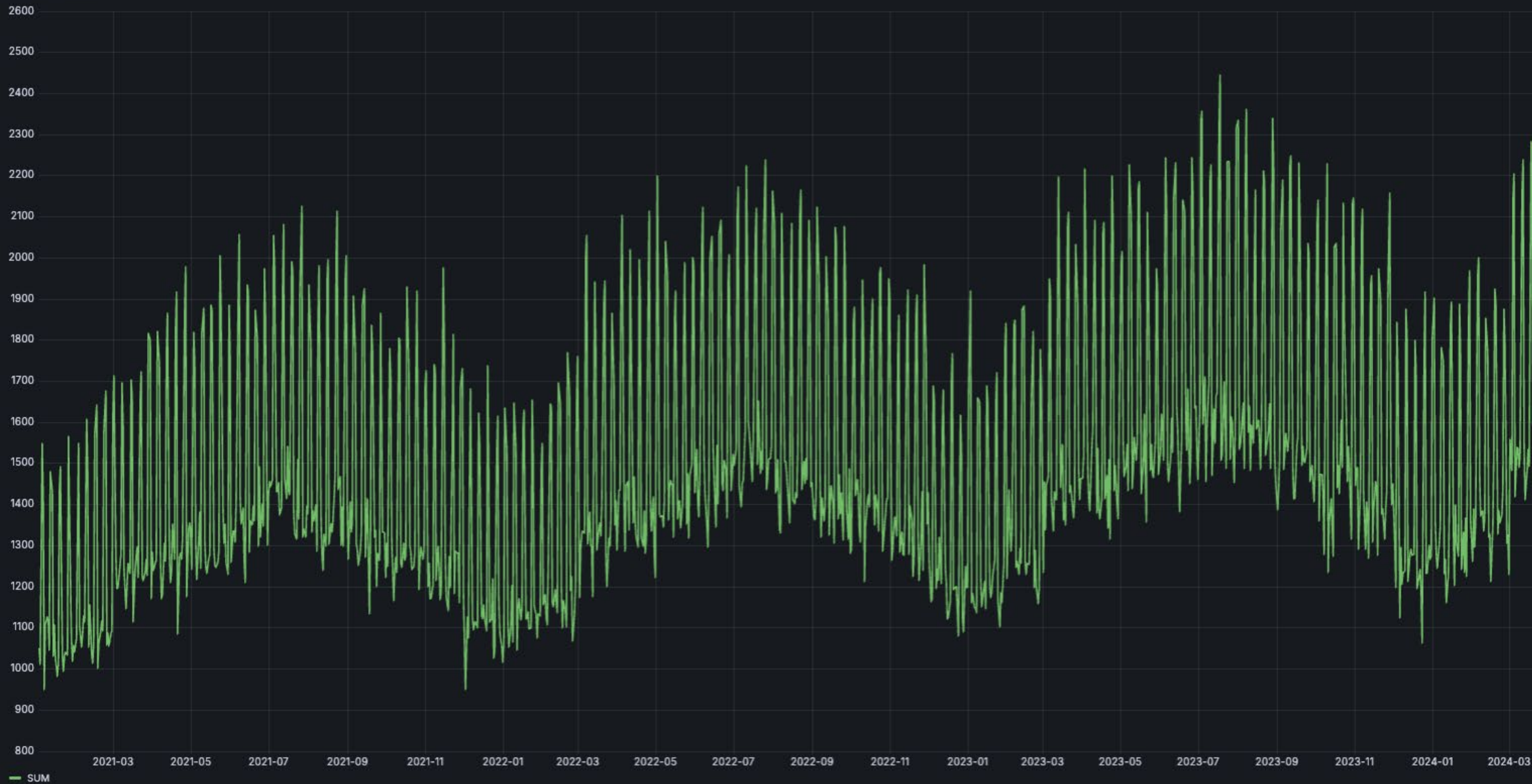
DAY	measure_name	SUM
2024-06-20 00:00:00	product_a	10
2024-06-20 00:00:00	product_a	12
⋮	⋮	⋮
2024-06-20 00:00:00	product_a	14
2024-06-21 00:00:00	product_a	15
⋮	⋮	⋮

DAY	measure_name	SUM
2024-06-20 00:00:00	product_a	160
2024-06-21 00:00:00	product_a	162
⋮	⋮	⋮

1 日のインターバルで時刻の切り下げ

GROUP BY でデータをまとめる

日次商品売上個数

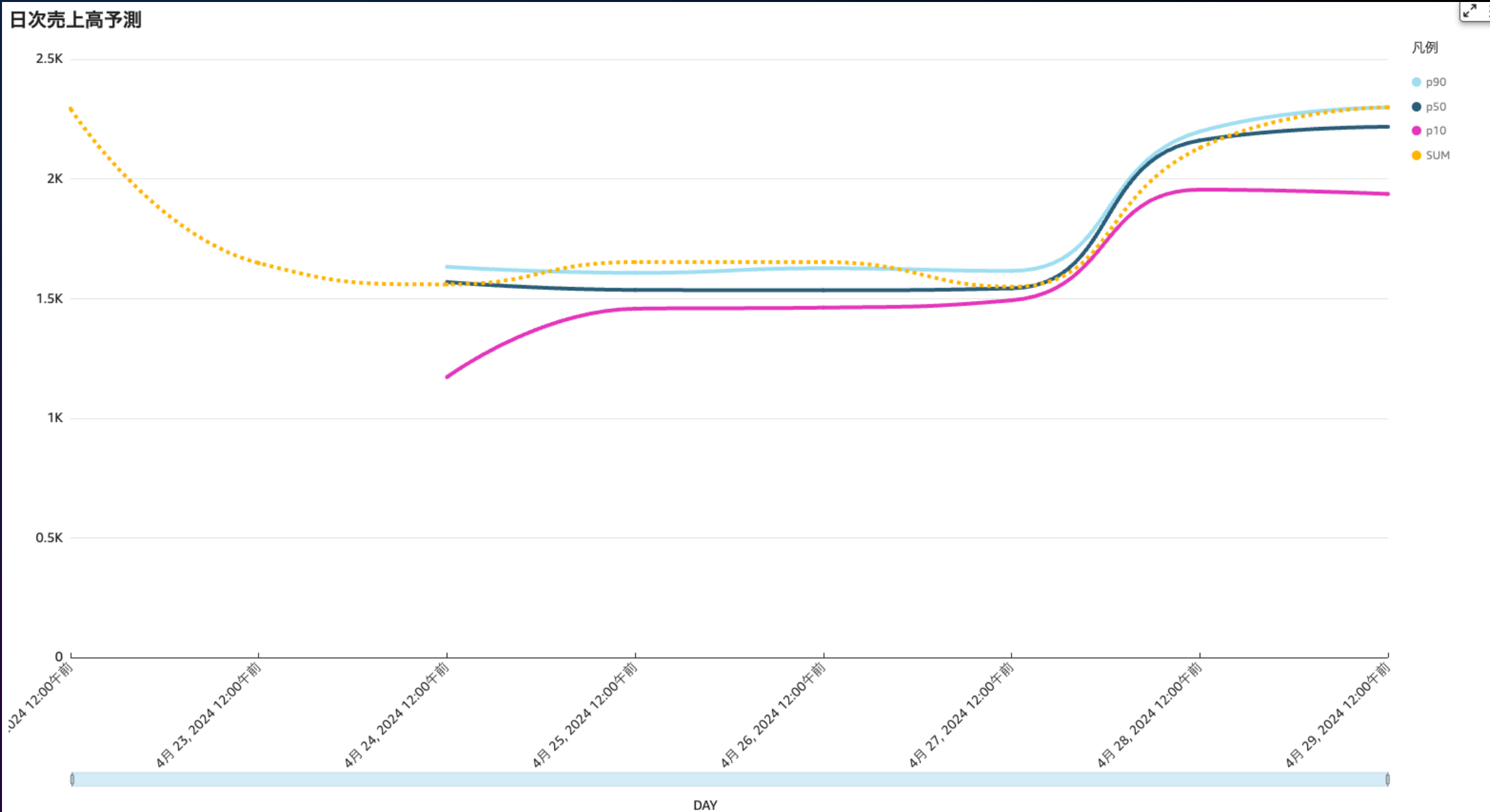


商品売上個数

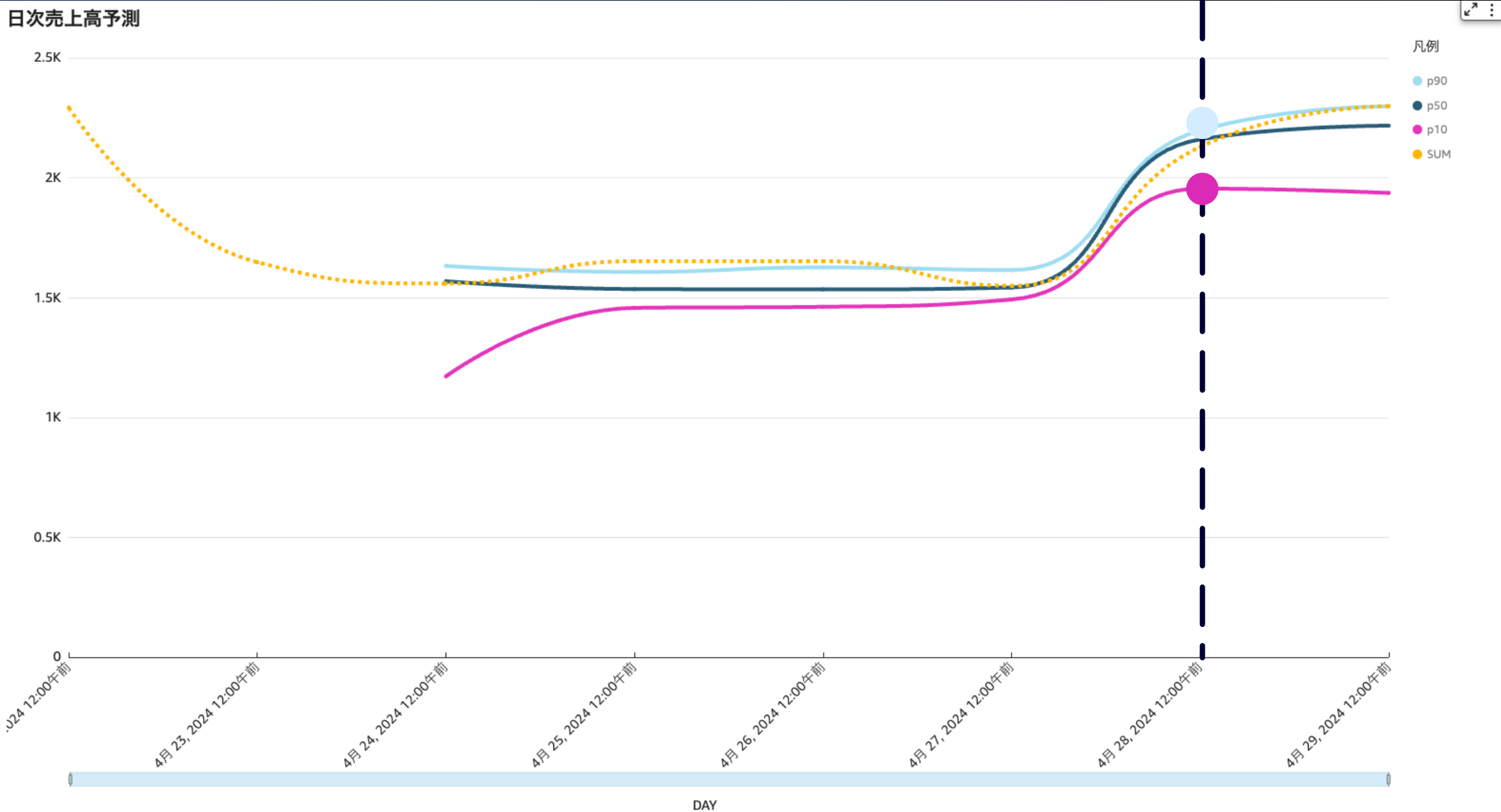


時間

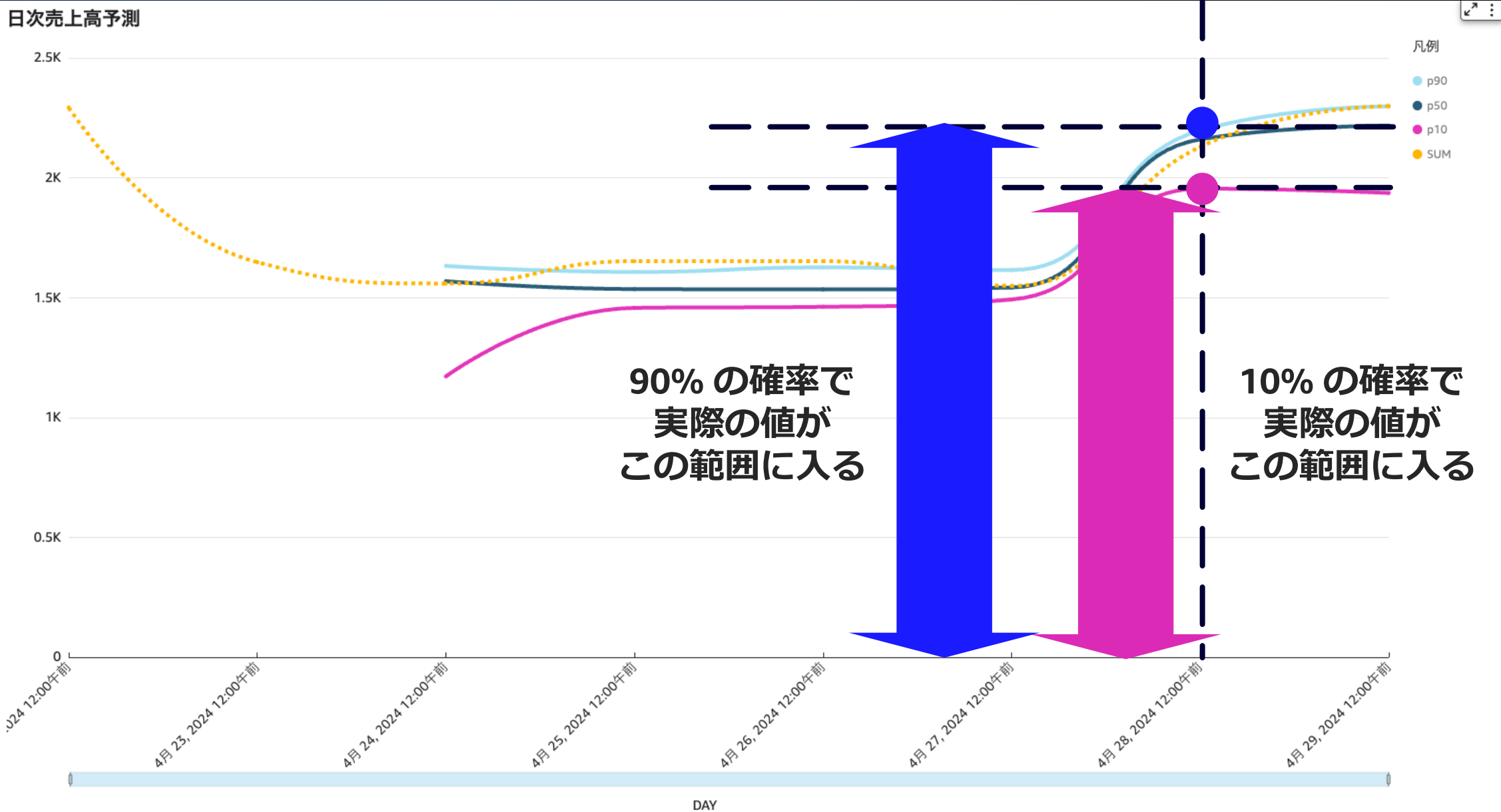
Amazon QuickSight での可視化



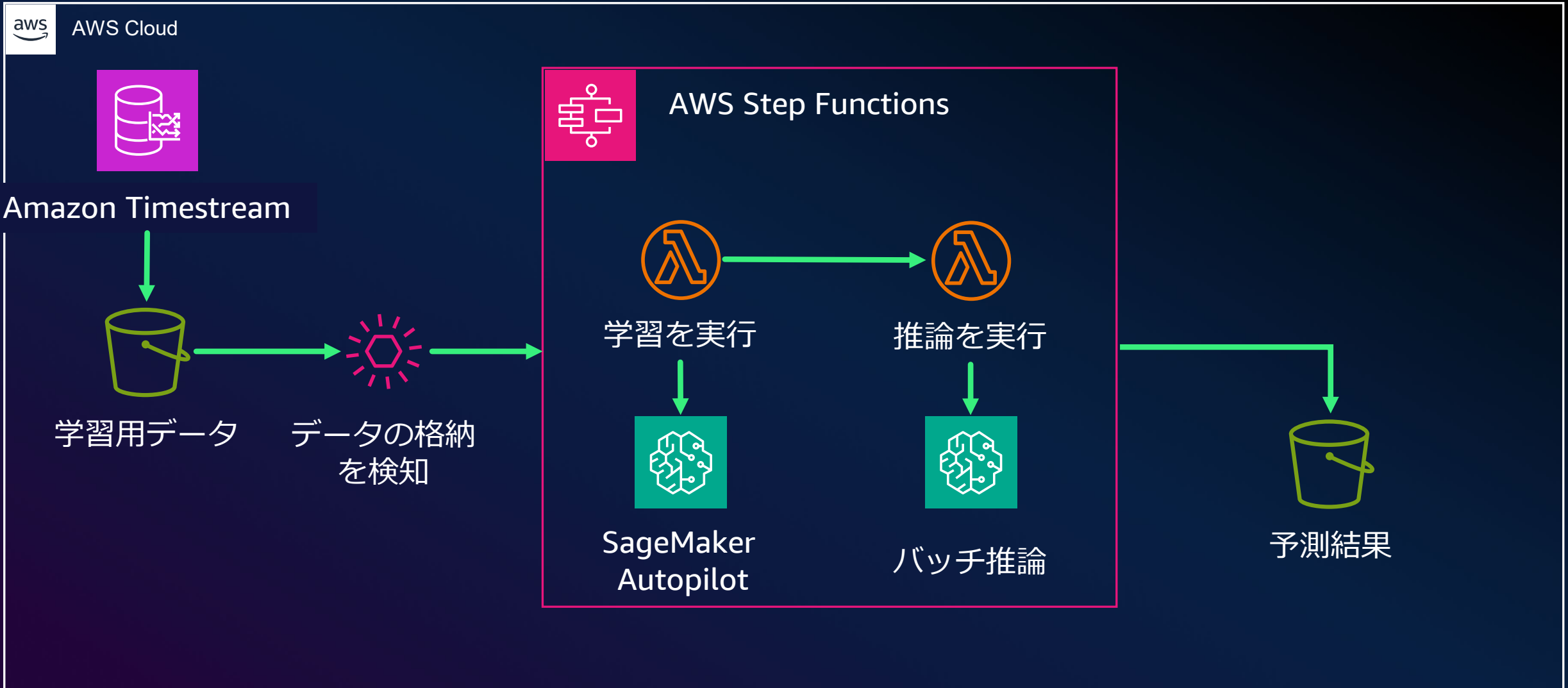
Amazon QuickSight での可視化



Amazon QuickSight での可視化



自動化することも可能

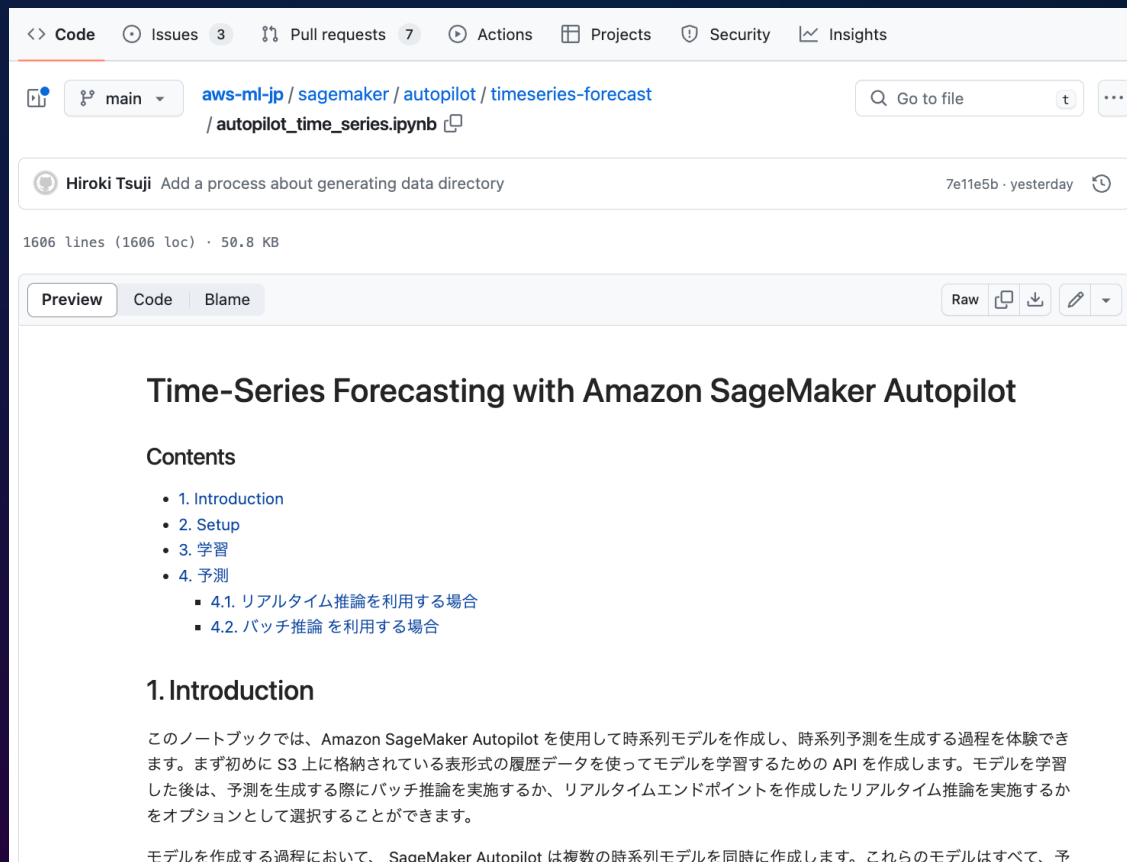


まとめ

まとめ

- AWS を利用いただくことで、時系列データを効率的に扱うことが可能です。
- Amazon Timestream for LiveAnalytics でスケーラブルにデータ収集し、コスト効率良くデータを保管でき、時系列専用の分析機能を SQL で利用することができます。
- Amazon SageMaker Autopilot で、時系列予測に関する知見がなくても、簡単に精度高い時系列予測モデルを構築できます。

時系列予測デモで使用したコンテンツ



The screenshot shows a GitHub repository page for the file `aws-ml-jp / sagemaker / autopilot / timeseries-forecast / autopilot_time_series.ipynb`. The repository is owned by Hiroki Tsuji. The file is 1606 lines (1606 loc) and 50.8 KB. The preview shows the following content:

Time-Series Forecasting with Amazon SageMaker Autopilot

Contents

- 1. Introduction
- 2. Setup
- 3. 学習
- 4. 予測
 - 4.1. リアルタイム推論を利用する場合
 - 4.2. バッチ推論を利用する場合

1. Introduction

このノートブックでは、Amazon SageMaker Autopilot を使用して時系列モデルを作成し、時系列予測を生成する過程を体験できます。まず最初に S3 上に格納されている表形式の履歴データを使ってモデルを学習するための API を作成します。モデルを学習した後は、予測を生成する際にバッチ推論を実施するか、リアルタイムエンドポイントを作成したリアルタイム推論を実施するかをオプションとして選択することができます。

モデルを作成する過程において、SageMaker Autopilot は複数の時系列モデルを同時に作成します。これらのモデルはすべて、予



SageMaker Autopilot を使った
時系列予測ハンズオンコンテンツ

<https://github.com/aws-samples/aws-ml-jp/tree/main/sagemaker/autopilot/timeseries-forecast>



Thank you!

Shoji Takano

tshoj@amazon.co.jp

Hiroki Tsuji

hitsuji@amazon.co.jp



SageMaker Autopilot を使った
時系列予測ハンズオンコンテンツ

