
aws
INNOVATE
DATA AND AI/ML EDITION
22 February 2023

aws INNOVATE
DATA AND AI/ML EDITION

From accuracy to business case: Scoping a successful demand forecasting PoC

Julia Ang
Solutions Architect
Amazon Web Services

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

1. Building the business case for forecasting
2. Creating machine learning based forecasts on AWS
3. Scoping a successful PoC
4. Demo
5. Next steps

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Building the business case for forecasting

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

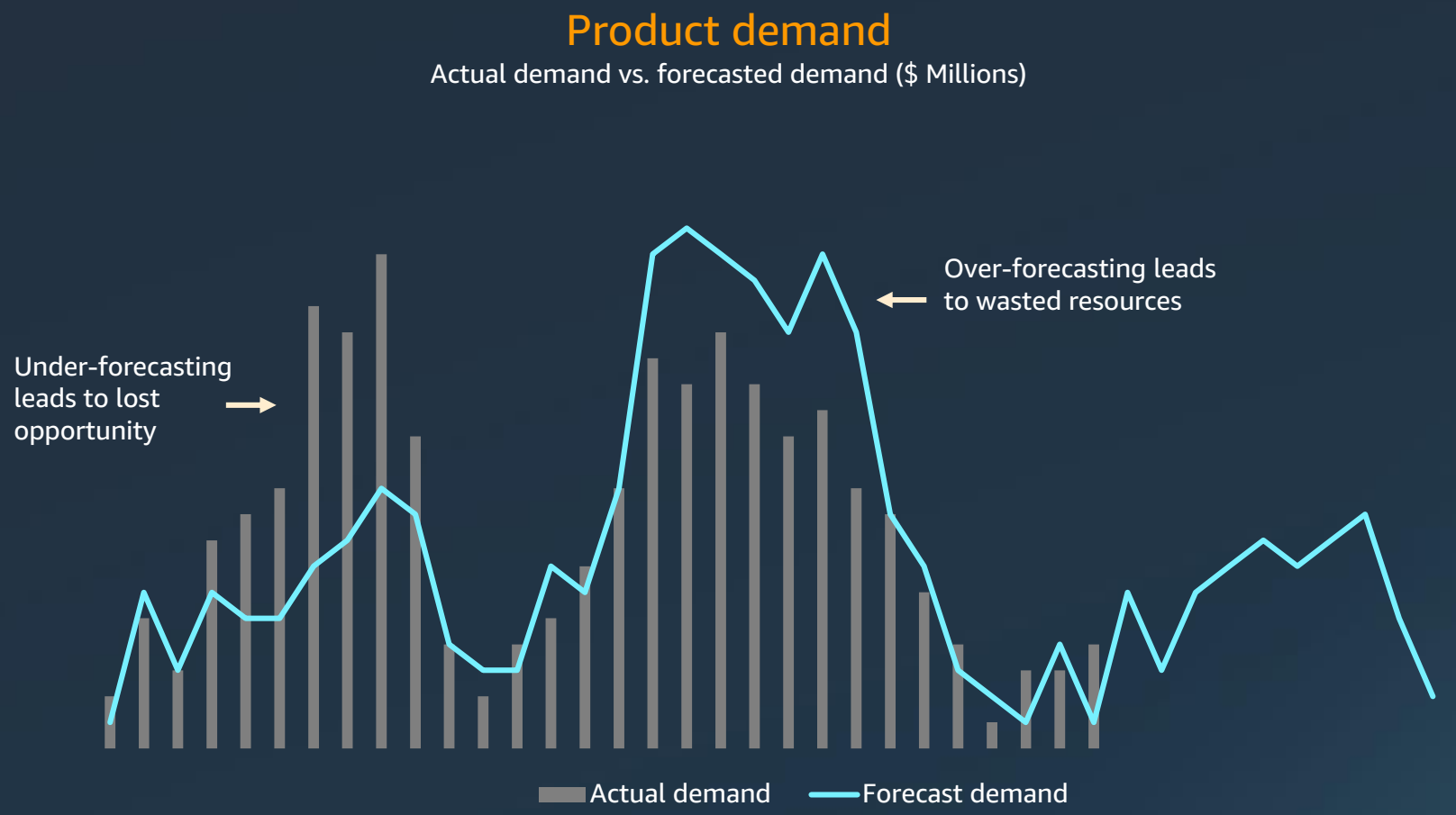
Cost

Revenue / Margin

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The case for forecasting

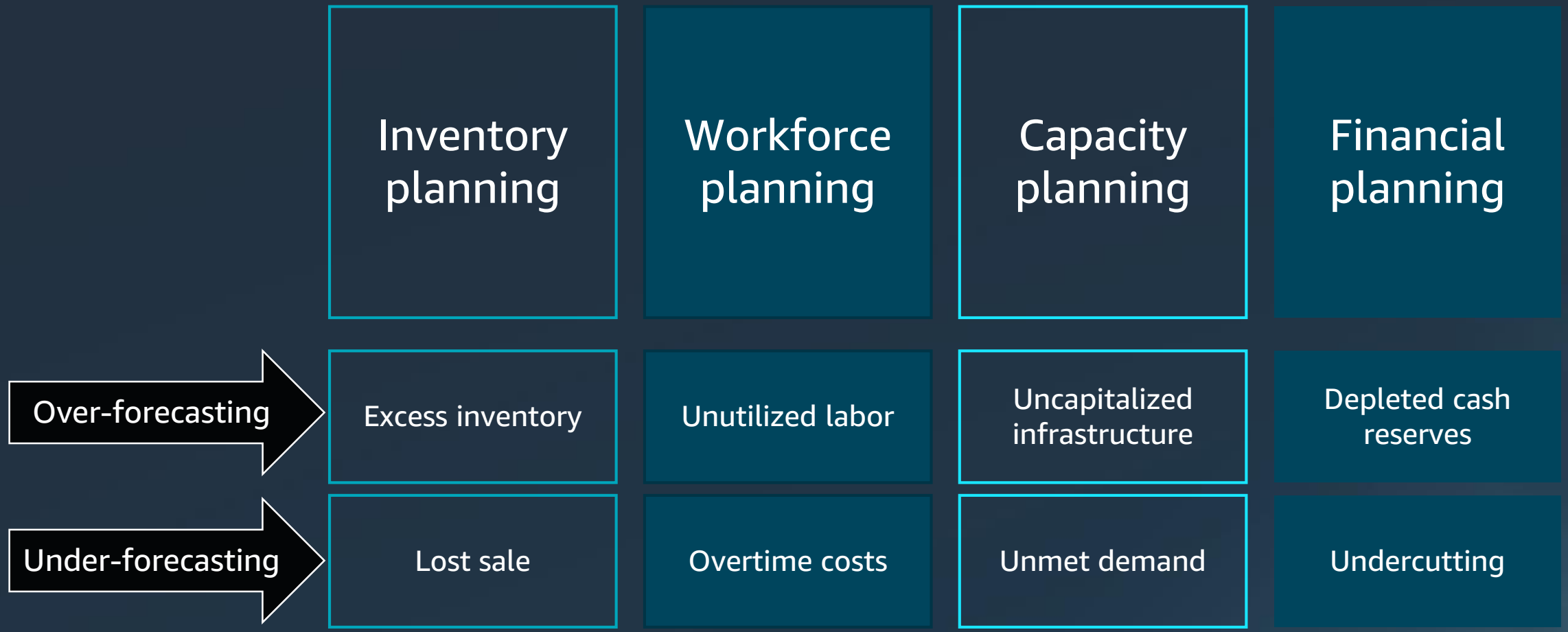
Forecasting is the science of predicting the future



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The case for forecasting

Impact of under and over forecasting



aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How customers built a case for ML-based forecasting

Heroleads

more QUALITY 1st

FOXCONN PCE TECHNOLOGY DE JUÁREZ, S.A. DE C.V.

Forenamics

FABULYST

Heroleads

99%

increase in forecast accuracy to track marketing campaign KPIs

More Retail

30%

reduced food waste by 30% while improving in-stock availability

Foxconn

$553K

projected annual savings with an increase in forecasting accuracy of 8%

Forenamics

35%

reduction in overproduction

25%

reduction in underproduction

Fabulyst

10%

boost in incremental revenue for retailers by forecasting monthly trends

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Creating machine learning based forecasts on AWS

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS low/no code options to incorporate ML in forecasts

Amazon SageMaker Canvas

- No ML knowledge required
- No code, visual point-and-click interface
- Automatically identifies the problem type, using ML techniques such as linear regression, logistic regression, deep learning, time-series forecasting, and gradient boosting
- Share your Amazon SageMaker Canvas models with data scientists who use Amazon SageMaker Studio

Amazon Forecast

- No ML knowledge required
- API integration
- Provides the option to include related time-series data and item metadata; can incorporate holiday and weather data
- Deploy recurring workflows with no-code using AWS CloudFormation and AWS Step Functions

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Canvas

Build ML models and generate accurate predictions — no code required

Quickly access and prepare data for machine learning

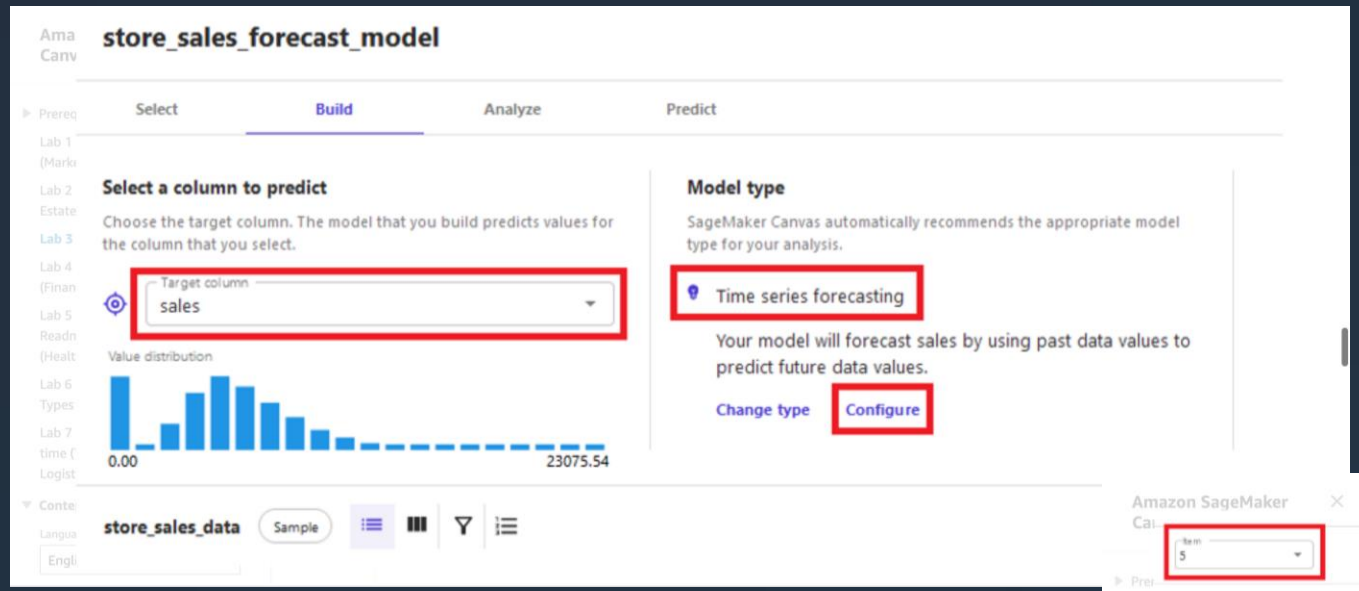
Built-in AutoML to build models and generate accurate predictions

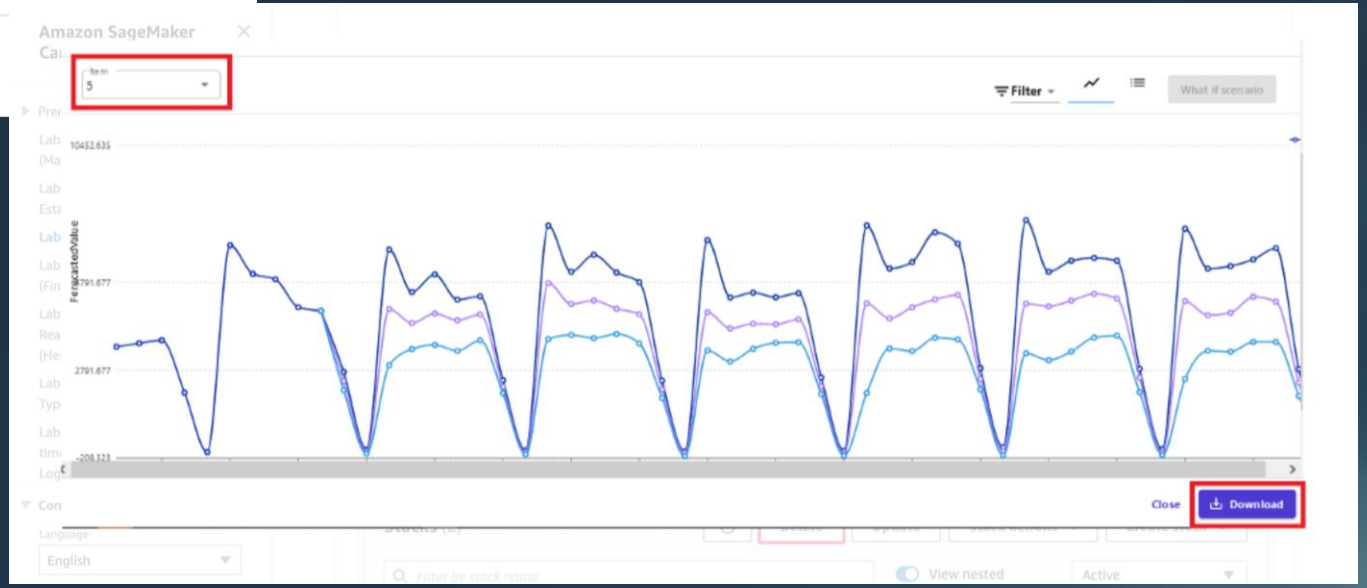
Share ML models and collaborate with data science teams

Usage-based pricing to avoid licensing fees and reduce total cost of ownership

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Canvas





© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Forecast

Advanced machine learning time-series forecasting

Create highly accurate models without deep learning expertise

Get started easily with data in Amazon S3 using the console or API

Automatically train and select the best model using AutoML

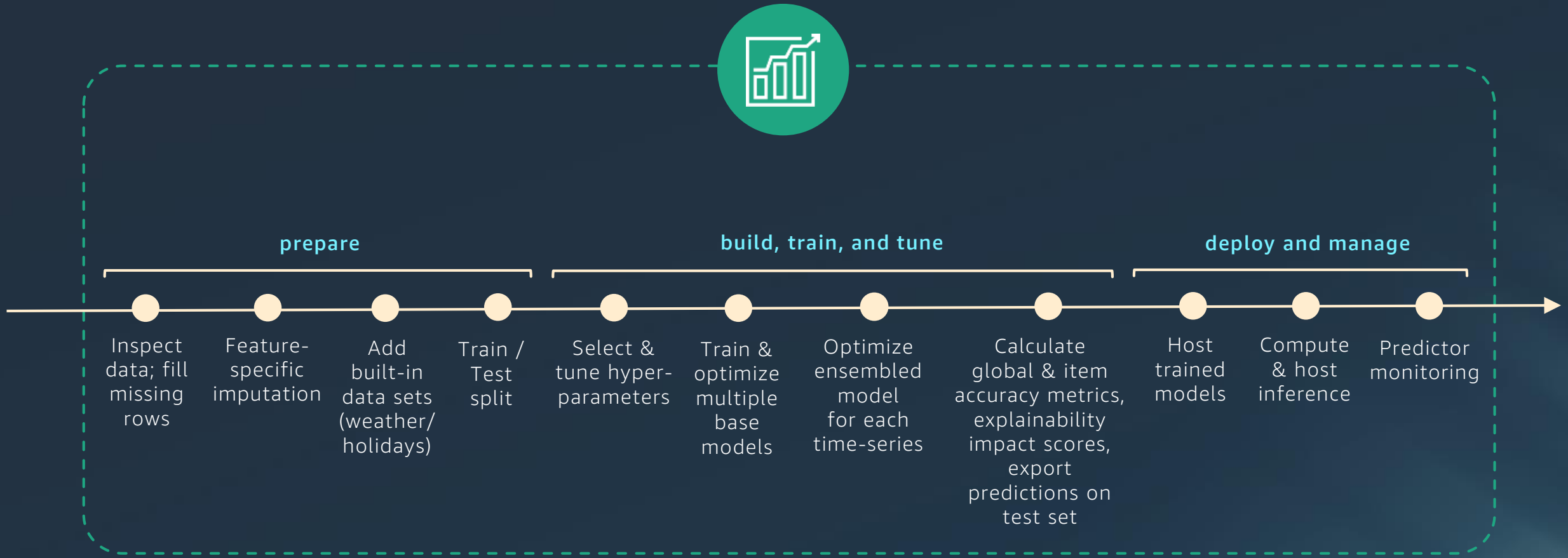
Incorporate item metadata, related time-series, holiday and weather information

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Forecast is a fully managed ML service

PREPARE, BUILD, TRAIN, TUNE, DEPLOY, AND MANAGE



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Examples of related time-series and item metadata

Use Case	Related time-series	Item metadata
Quick service restaurants	Store hours	Menu categorization (breakfast, lunch, dinner)
Grocery retail	Price, store hours, holidays, promotions	Department, store size, shelf life, type of stores, affluent cluster, competitive store, stores throughput
Manufacturing	Economic indicators from Bloomberg, backlog data	Marketing family, department
Hyper local on-demand services	Holidays, weather	City characteristic

Note:

- Most of the forecasting power comes from the historical demand data
- Adding additional related data does not always increase forecasting accuracy
- First create a model with only the target time-series and assess the accuracy gain, before adding additional related data
- Gaining benefit from adding related data is dependent on feature engineering

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Ensemble model algorithms to improve accuracy

Ensemble modeling selects the best combination of algorithms for each product

Neural Networks **• Designed for larger data sets** **• External data**		**Statistical Algorithms** **• Higher level forecasting** **• Fast & accurate**			
CNN-QR	**DeepAR+**	**Prophet**	**NPTS**	**ARIMA**	**ETS**
Uses causal convolutional neural networks (CNNs)	Global neural model	Additive model with Gaussian likelihood	Non-parametric time-series	Auto-regressive integrated moving average	Statistical algorithm that uses exponential smoothing
Works best with large datasets containing hundreds of time-series; accepts related time-series data without future values	Uses related time-series and attributes to train a model	Can find trend, seasonality, cyclical, and holiday effects	Performs well for intermittent spikes	Works well with a small number of time-series; classical approach to model autocorrelations	Works well with a small number of time-series Finds trends, seasonality, and residual

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Ensemble model algorithms to improve accuracy

Ensemble modeling selects the best combination of algorithms for each product

Neural Networks **• Designed for larger data sets** **• External data**		**Statistical Algorithms** **• Higher level forecasting** **• Fast & accurate**			
CNN-QR	**DeepAR+**	**Prophet**	**NPTS**	**ARIMA**	**ETS**
Uses causal convolutional neural networks (CNNs)	Global neural model	Additive model with Gaussian likelihood	Non-parametric time-series	Auto-regressive integrated moving average	Statistical algorithm that uses exponential smoothing
Works best with large datasets containing hundreds of time-series; accepts related time-series data without future values	Uses related time-series and attributes to train a model	Can find trend, seasonality, cyclical, and holiday effects	Performs well for intermittent spikes	Works well with a small number of time-series; classical approach to model autocorrelations	Works well with a small number of time-series Finds trends, seasonality, and residual
AutoPredictor					

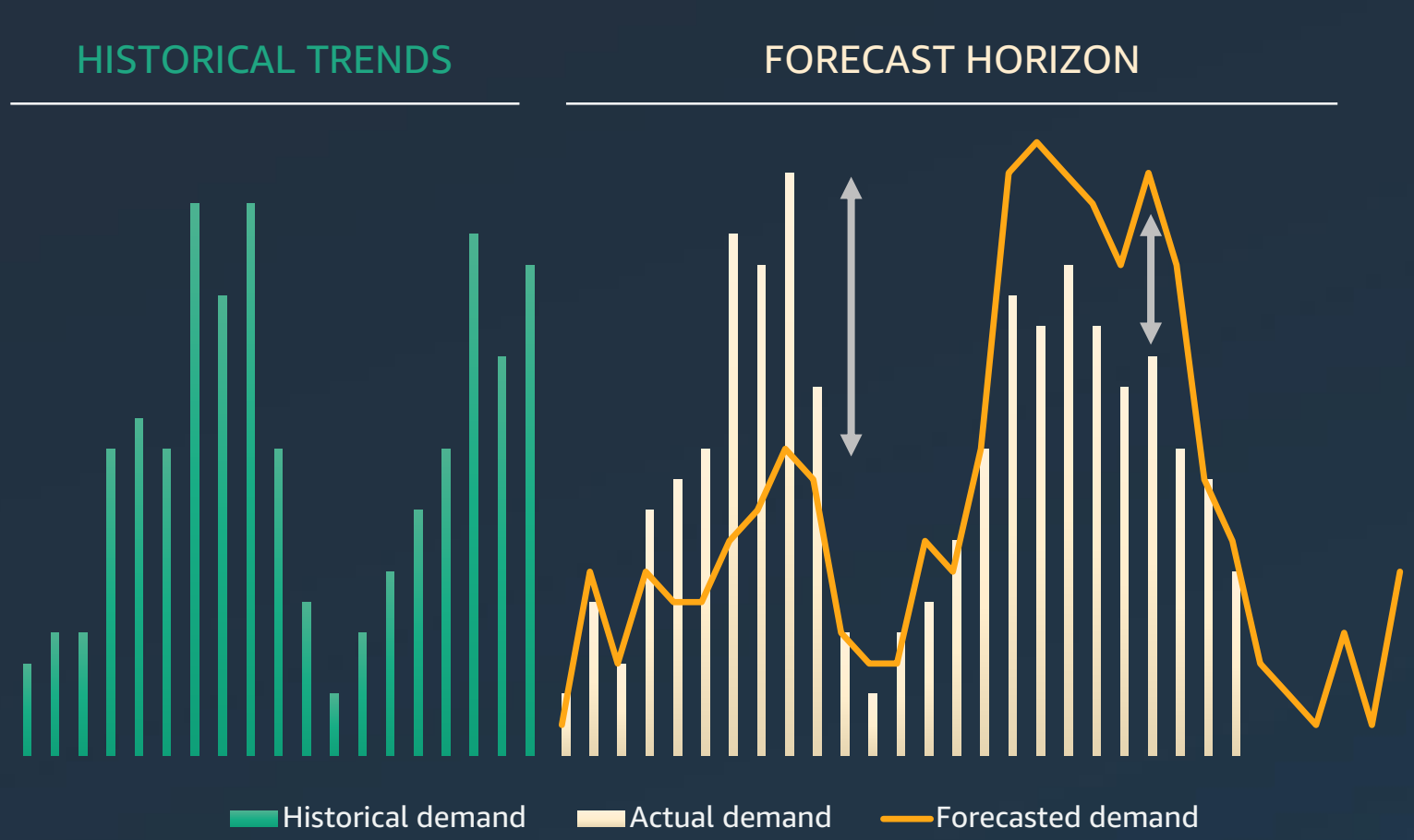
aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scoping a successful PoC

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

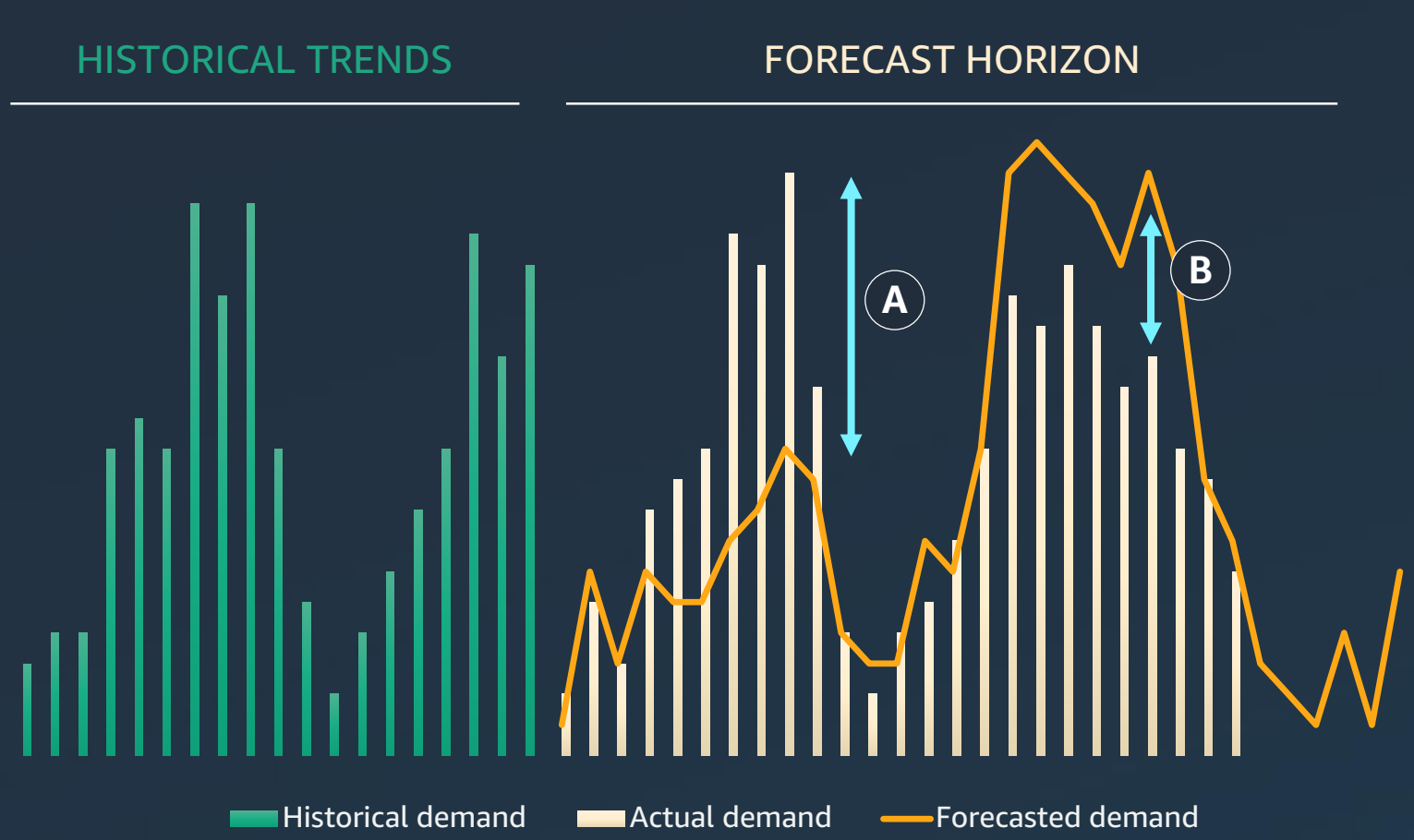
Traditional metrics for evaluating your forecast



- Weighted Quantile Loss (wQL)
- Weighted Absolute Percentage Error (WAPE)
- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- Mean Absolute Scaled Error (MASE)

aws
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Traditional metrics for evaluating your forecast



- Weighted Quantile Loss (wQL)
- Weighted Absolute Percentage Error (WAPE)
- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- Mean Absolute Scaled Error (MASE)

Making sense of these errors:

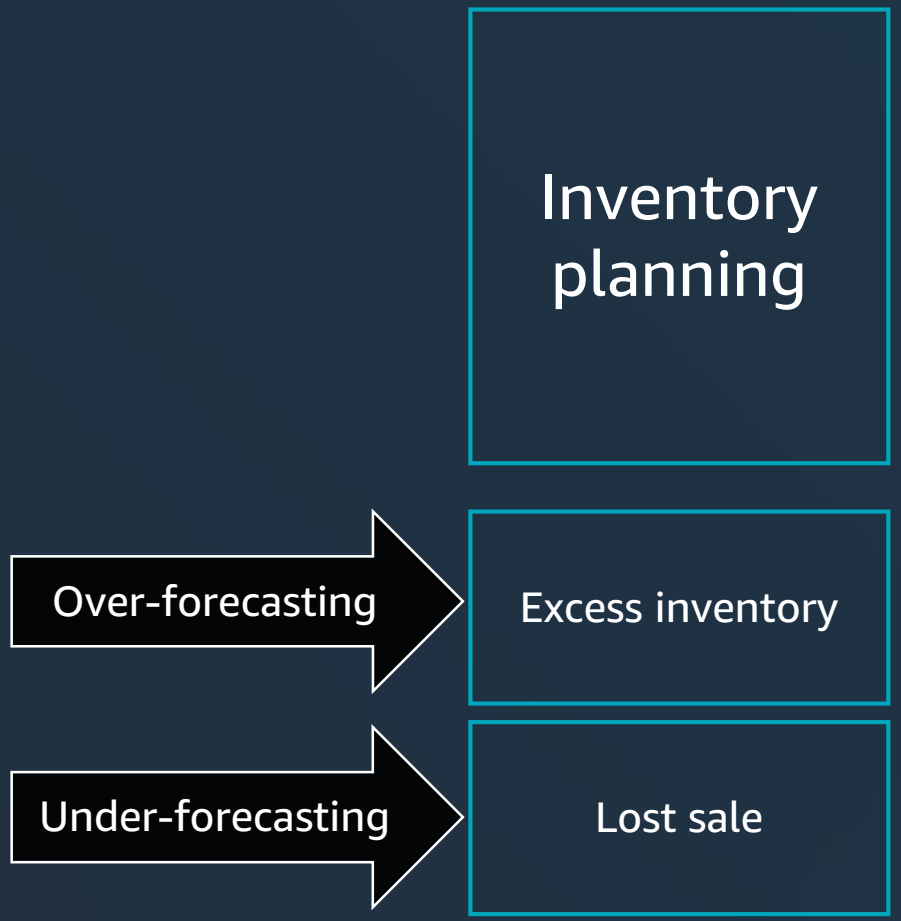
A - Stockouts

B - Excess inventory

aws

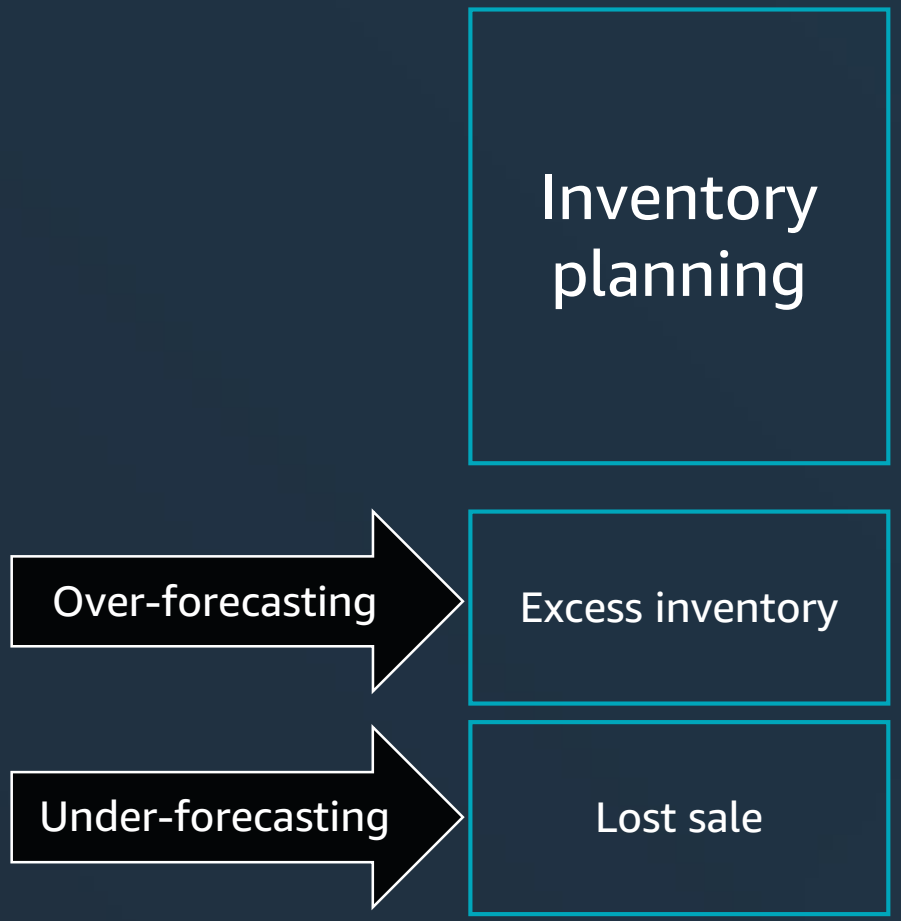
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tying back to business outcomes



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tying back to business outcomes



The cost of excess inventory is affected by:

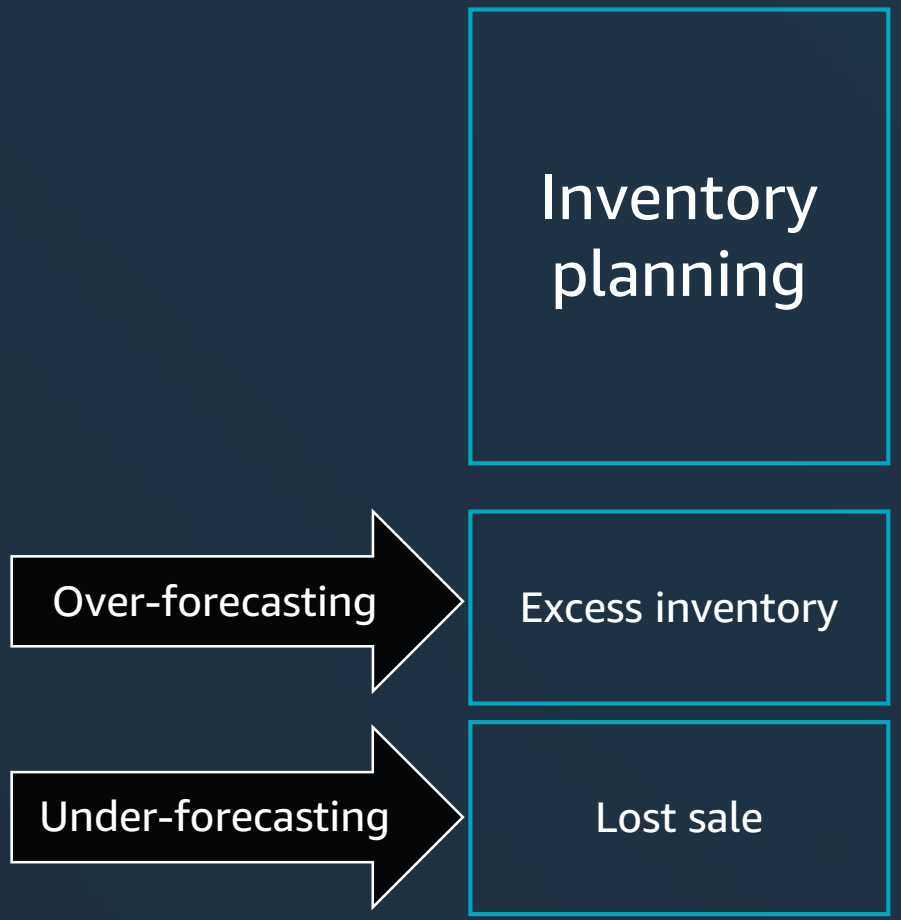
Storage space

Shelf life

Exceeding capacity

aws
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tying back to business outcomes



High level approximation for each item per month:

$$\frac{Total\ storage\ costs}{Capacity\ for\ incremental\ storage\ costs} \times factor\ for\ the\ cost\ of\ an\ unsold\ item$$

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Decide on your evaluation strategy

Backtesting

- Create a baseline using a selected timeframe of historical data
- Generate forecasts with the new algorithm using historical time-series data to compare against the baseline
- Calculate statistical and/or business metrics and compare between the algorithms
- Amazon Forecast uses backtesting to tune parameters and produce accuracy metrics

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Decide on your evaluation strategy

Backtesting

- Create a baseline using a selected timeframe of historical data
- Generate forecasts with the new algorithm using historical time-series data to compare against the baseline
- Calculate statistical and/or business metrics and compare between the algorithms
- Amazon Forecast uses backtesting to tune parameters and produce accuracy metrics

Forward-looking A/B testing

- Choose items to include in the test group
- Generate forecasts with the new algorithm using historical time-series data for this group, while generating forecasts as per usual using any existing algorithm
- Implement a purchasing policy with input from the generated forecasts
- Calculate statistical and/or business metrics and compare between the algorithms

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A/B testing considerations

Metric(s) being measured should be assignable to dollar values

Results from offline model metrics and/or past tests should justify the risk of testing

Factor in the difference in the price or margin of items among groups

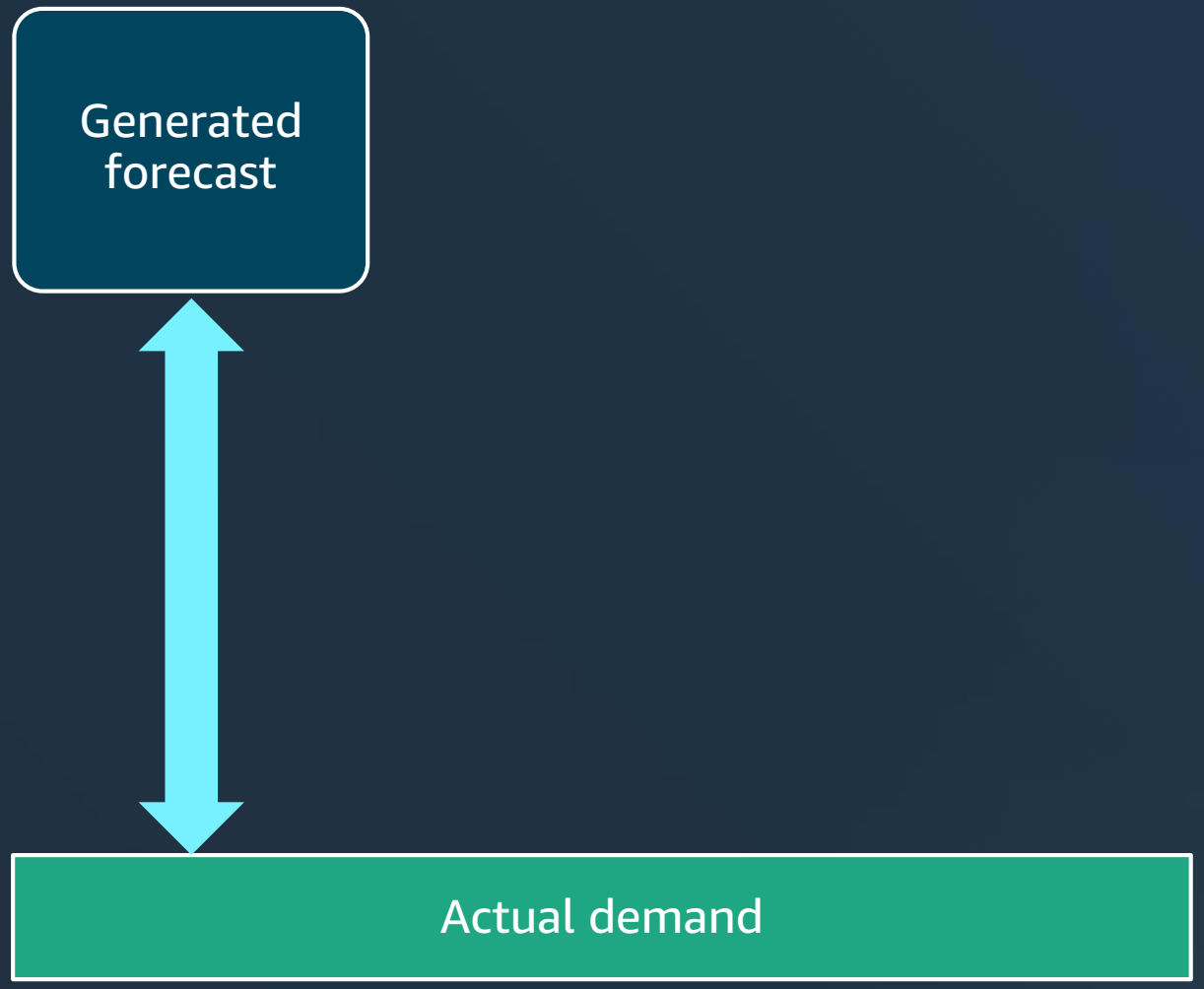
Take note of difference in the sales volumes and variability of items among groups

Consider the spread of items in A and B among different category managers

aws

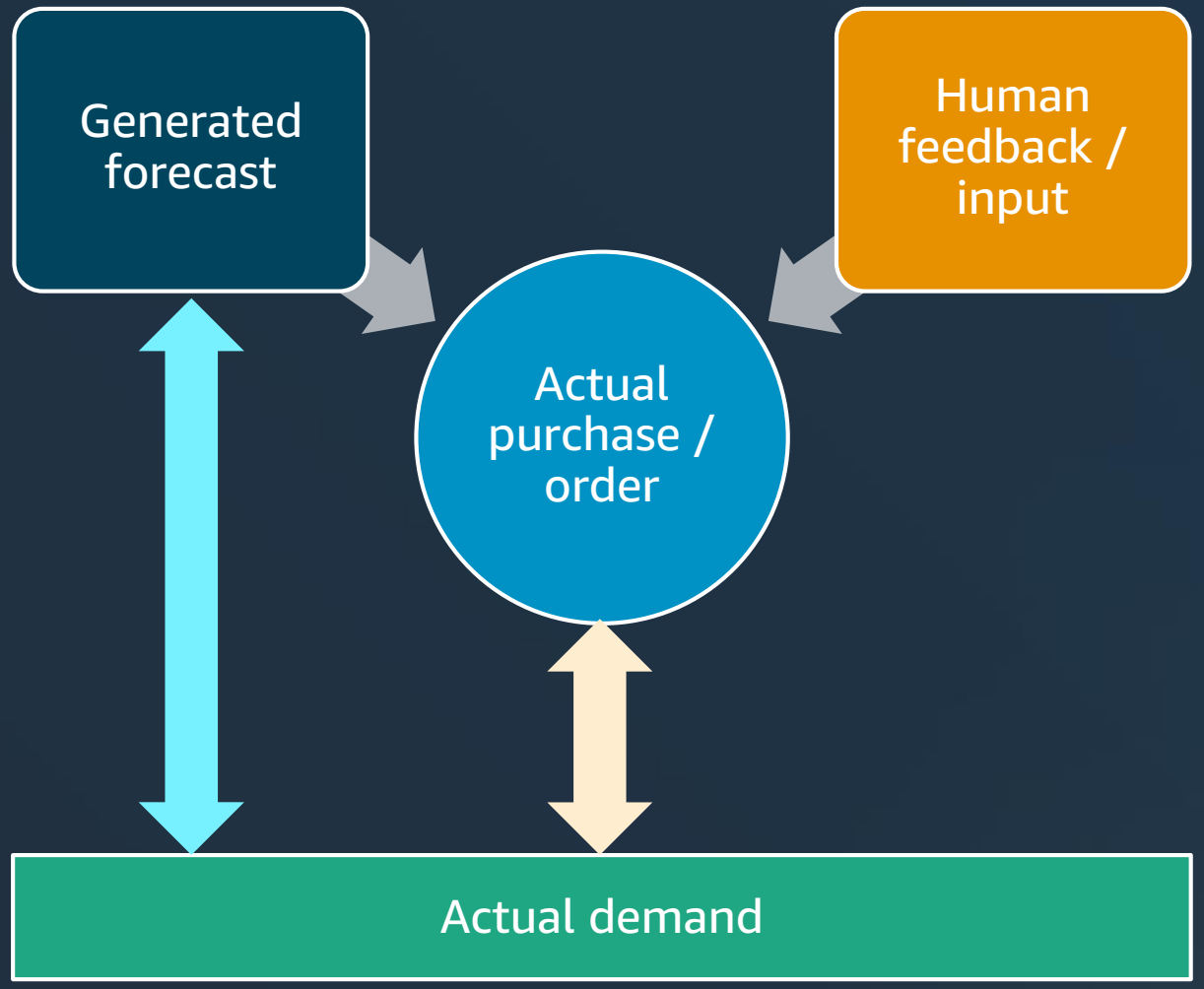
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Factoring in human feedback



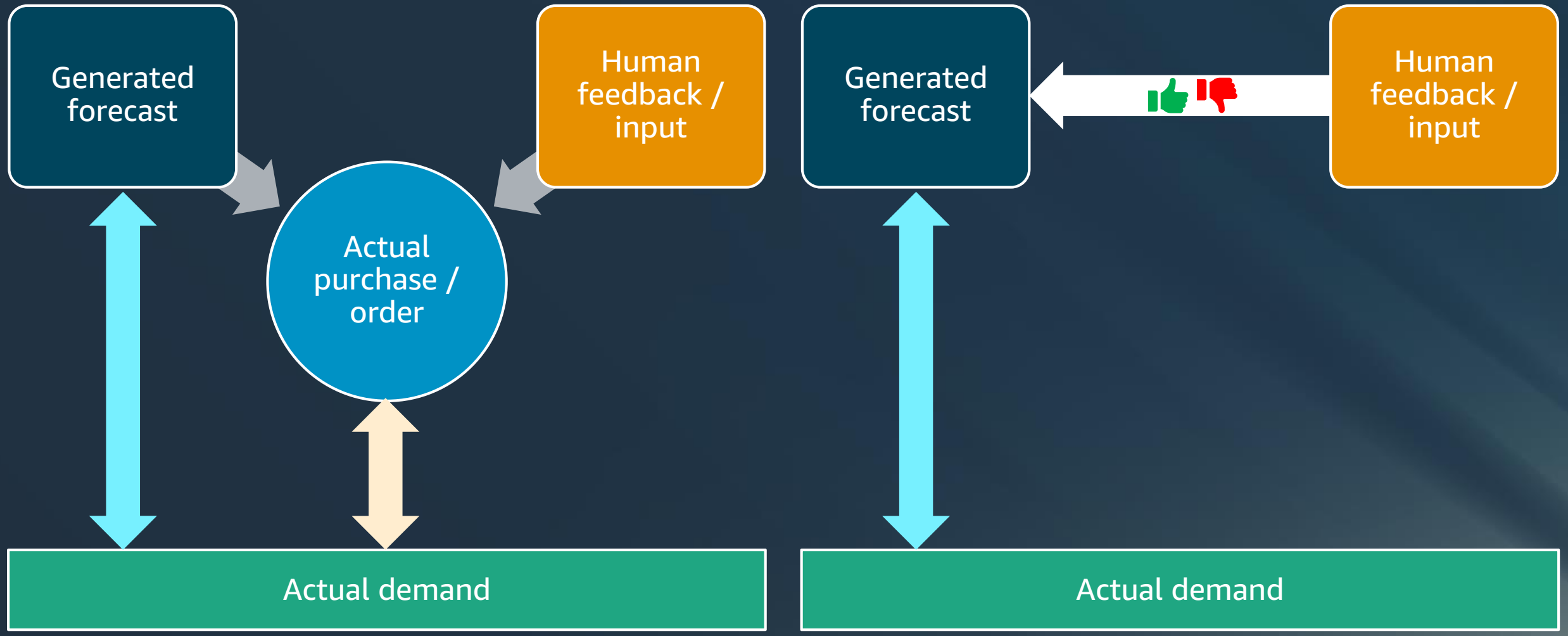
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Factoring in human feedback



aws
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Factoring in human feedback

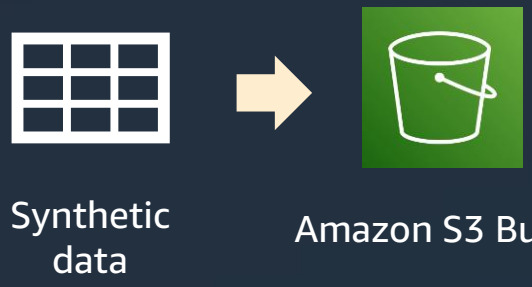


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo

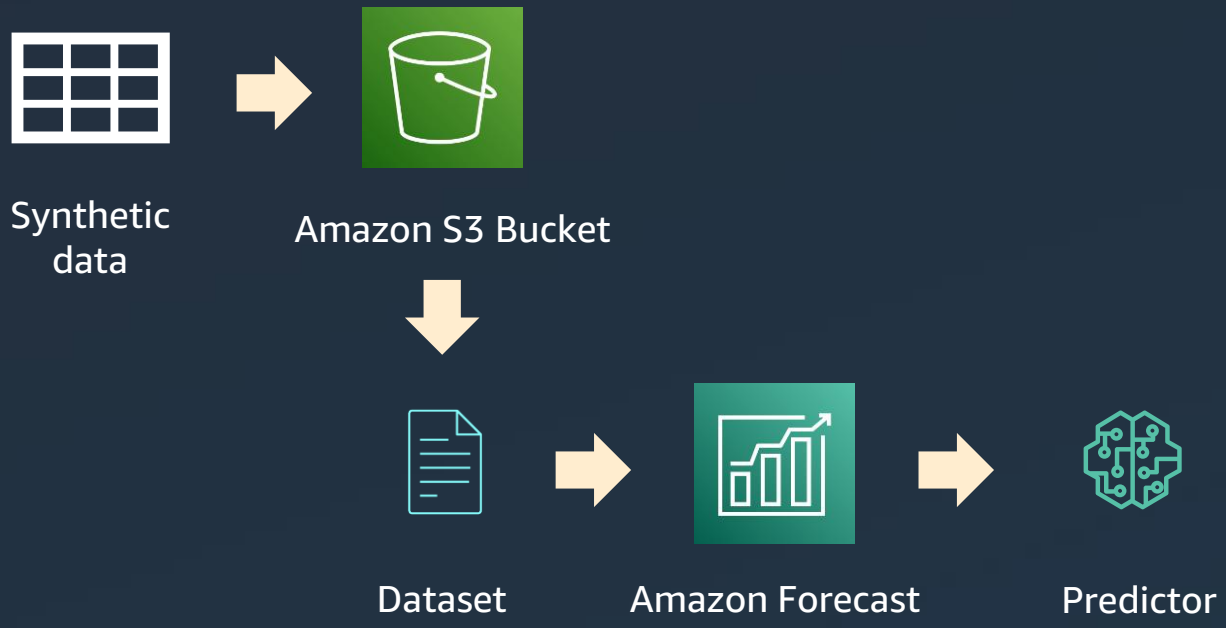
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture components



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture components

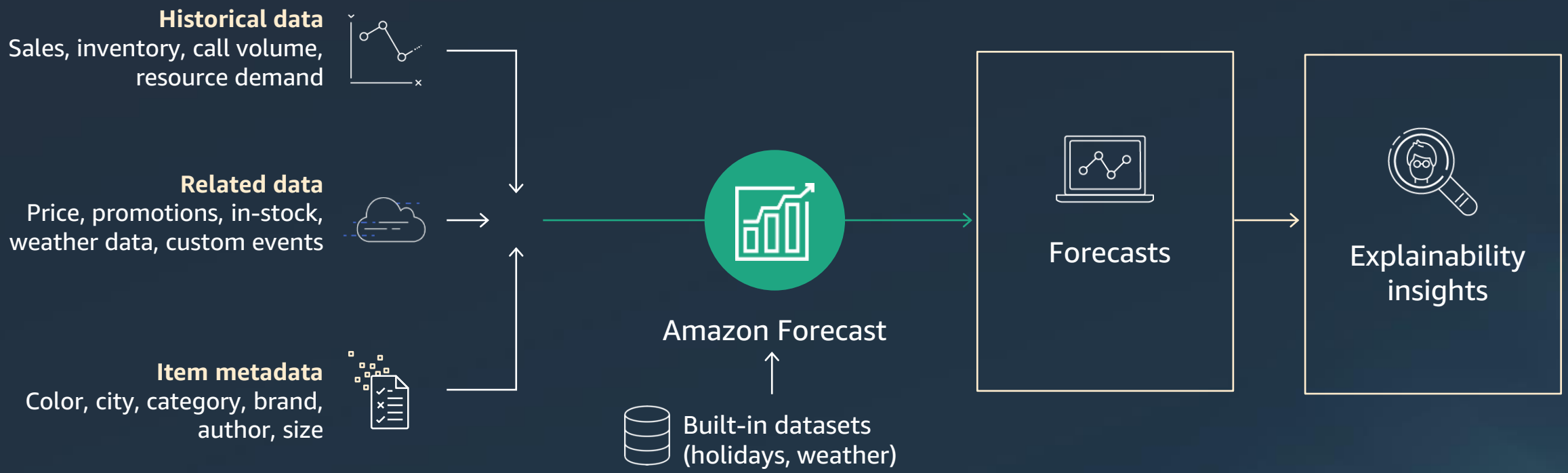


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture components

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Forecast – key inputs and outputs



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Target time-series (TTS)

REQUIRED FILE THAT CONTAINS WHAT YOU ARE TRYING TO PREDICT

up to 10 dimensions

referred to as **y-value** or target value

location_id	item_id	target_value	timestamp
A55	1885B	177	2016-11-08
A55	1993B	270	2016-11-08
A55	2539B	189	2016-11-08
A55	2139B	54	2016-11-08
A55	2631B	40	2016-11-08

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Related time-series (RTS)

OPTIONAL FILE THAT HELPS EXPLAIN AND INFORM THE PREDICTIONS

up to 10 dimensions

numeric features (x-value multi-variates)

location_id	item_id	checkout_price	base_price	emailer for promotion	homepage featured	timestamp
A10	1062	131.95	182.36	0	1	2019-01-29
A10	1062	157.14	157.14	0	0	2018-10-16
A10	1062	158.17	156.17	0	0	2018-10-30
A10	1062	159.08	182.39	0	0	2017-12-05
A10	1062	159.08	183.33	0	0	2017-03-21

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Item metadata (ITEM)

OPTIONAL FILE PROVIDING CATEGORICAL FEATURES

no dimensions
only key

up to 10 categorical (string)

item_id	food_category	food_cuisine
2826	Sandwich	Italian
2664	Salad	Italian
2569	Salad	Italian
1230	Beverages	Continental
1207	Beverages	Continental

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

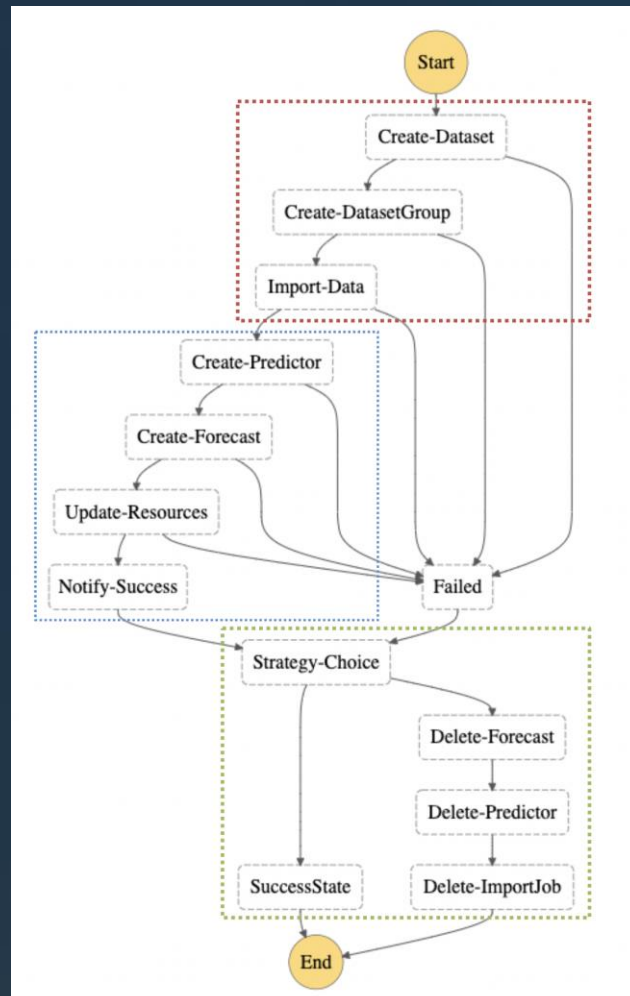
Next steps

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Next steps

Automated retraining

Create managed workflows using AWS Step Functions




```
{
  "DatasetGroup": {
    "DatasetGroupName":"…",
    "Domain": "…"
  },
  "TimestampFormat": "…",
  "Datasets": [
    {
      "DatasetName": "…",
      "Domain": "…",
      "DatasetType": "…",
      "Schema": {
          …
      }
    }],
  "Predictor": {
    "PredictorName": "…",
    "ForecastHorizon": 123,
    "AlgorithmArn": "…"
     …
  },
  "Forecast": {
    "ForecastName": "…",
    "ForecastTypes":[
      "0.10",
      "0.50",
      "0.90"
    ]
  },
  "PerformDelete": false
}
```

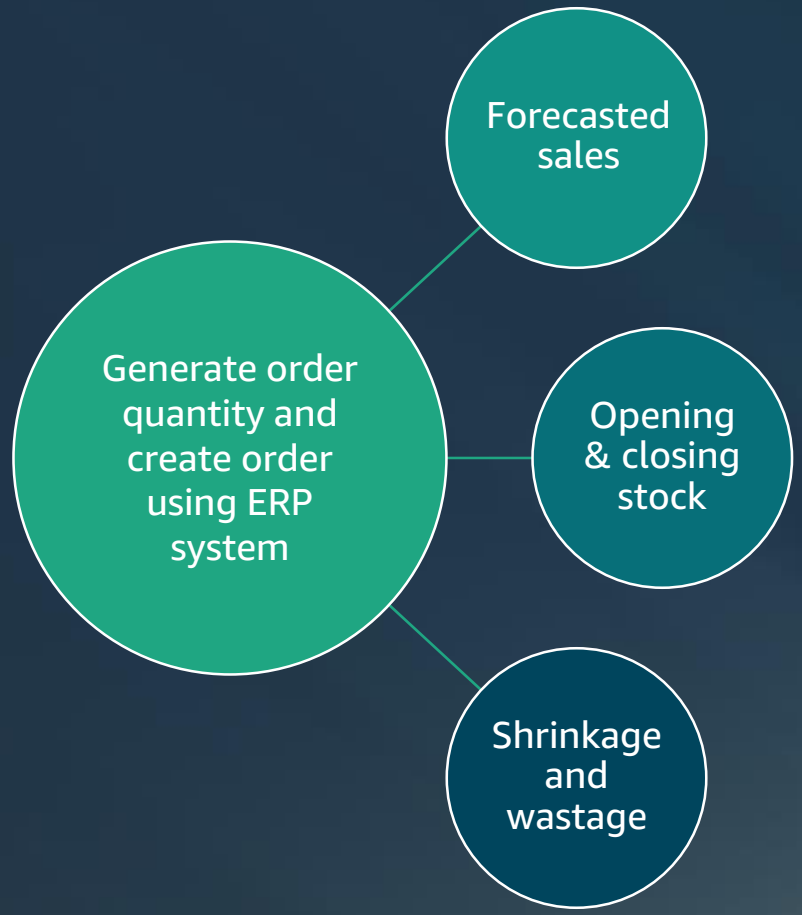
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Next steps

Automated retraining

Integrated replenishment

Use jobs to make forecasts & integrate with ERP system



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Next steps

Automated retraining

Integrated replenishment

Use jobs to make forecasts & integrate with ERP system

Generate order quantity and create order using ERP system

Forecasted sales

Opening & closing stock

Shrinkage and wastage

Select quantiles based on stockouts and wastage

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

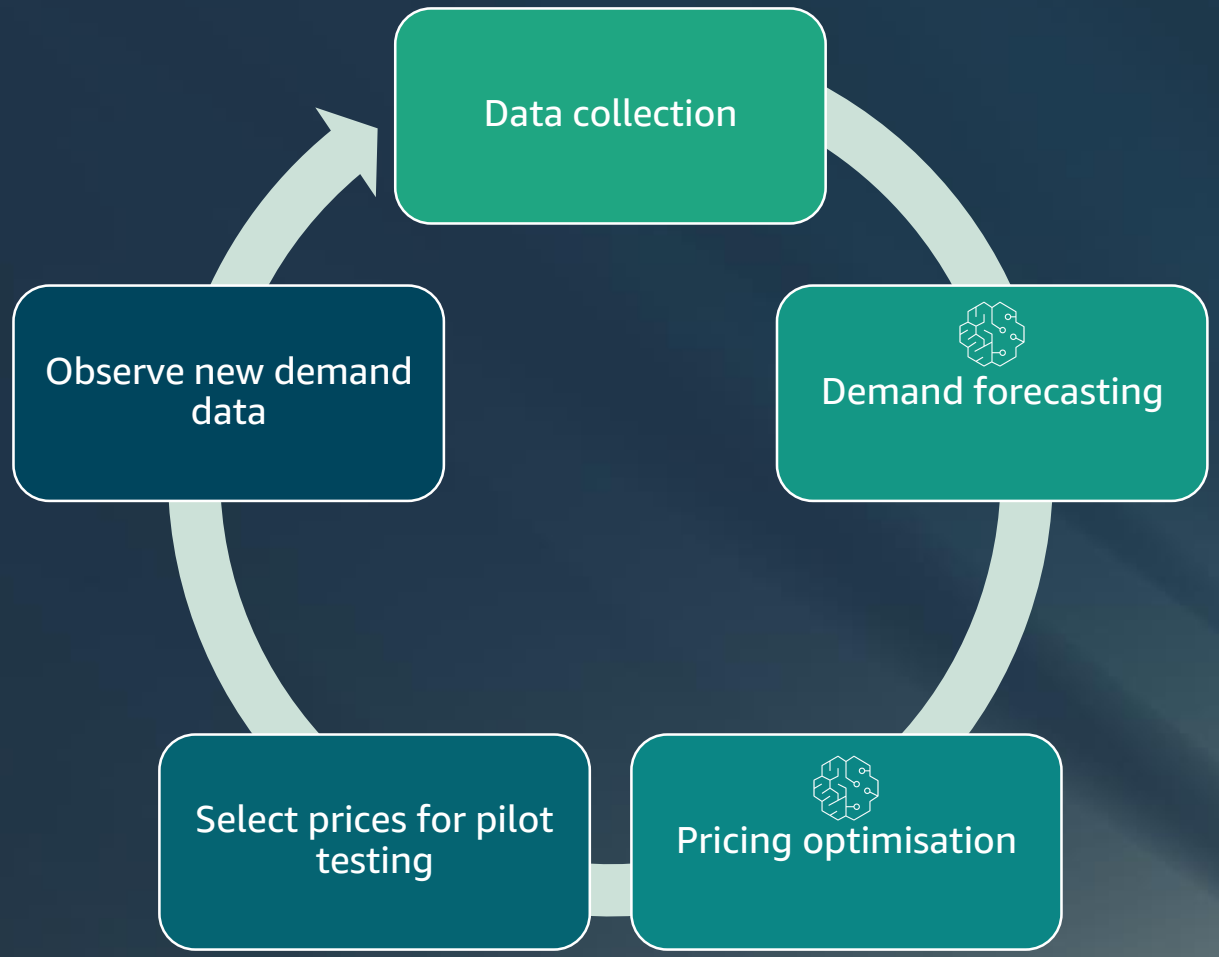
Next steps

Automated retraining

Integrated replenishment

Dynamic pricing

Optimise using iterative workflows and tests



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Recap

- Using machine learning based forecasting has brought about many tangible benefits to AWS customers in use cases such as inventory planning
- Amazon SageMaker Canvas and Amazon Forecast allow customers to perform forecasting without having a background in machine learning
- Scoping a successful PoC involves measuring business metrics such as the cost of stockouts and/or excess inventory

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Additional resources

Measuring demand forecast benefits repository
Learn how to translate forecasts to quantifiable business metrics

Implementing Amazon Forecast in the retail industry - blog post
Creating and productionizing a solution for demand forecasting in 8 weeks, saving 16 labor hours monthly

Amazon Forecast MLOps Github repository
An easily adjustable template for deploying multiple parallel workloads

AWS machine learning blog
From forecasting demand to ordering – An automated machine learning approach with Amazon Forecast to decrease stockouts, excess inventory, and costs

Amazon Forecast AWS solution
One-click deployment using an AWS CloudFormation template

Bosch blog post
Large-scale revenue forecasting at Bosch with Amazon Forecast and Amazon SageMaker custom models

aws
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Training and Certification

Access the AI & ML learning plan courses built by AWS experts on AWS Skill Builder

- Get started with digital self-paced, on-demand training and ramp-up guides to help you grow your technical skills
- Learn how to apply machine learning, artificial intelligence, and deep learning to unlock new insights and value in your role
- Take the steps today, towards validating your expertise with an AWS Certified Machine Learning – Specialty Certification

https://bit.ly/3FnxDH7

Learn your way explore.skillbuilder.aws »

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Visit the Data & AI/ML resource hub

Dive deeper into these resources, get inspired and learn how you can use AI and machine learning to accelerate your business outcomes.

- 6 steps to machine learning success e-book
- 7 leading machine learning use cases e-book
- Machine learning at scale e-book
- Achieving transformative business results with machine learning e-book
- Tackling our world's hardest problems with machine learning e-book
- Accelerating machine learning innovation through security e-book
- … and more!

https://bityl.co/FqdC

Visit resource hub

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you for attending AWS Innovate – Data & AI/ML Edition

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.


aws-apj-marketing@amazon.com


twitter.com/AWSCloud


facebook.com/AmazonWebServices

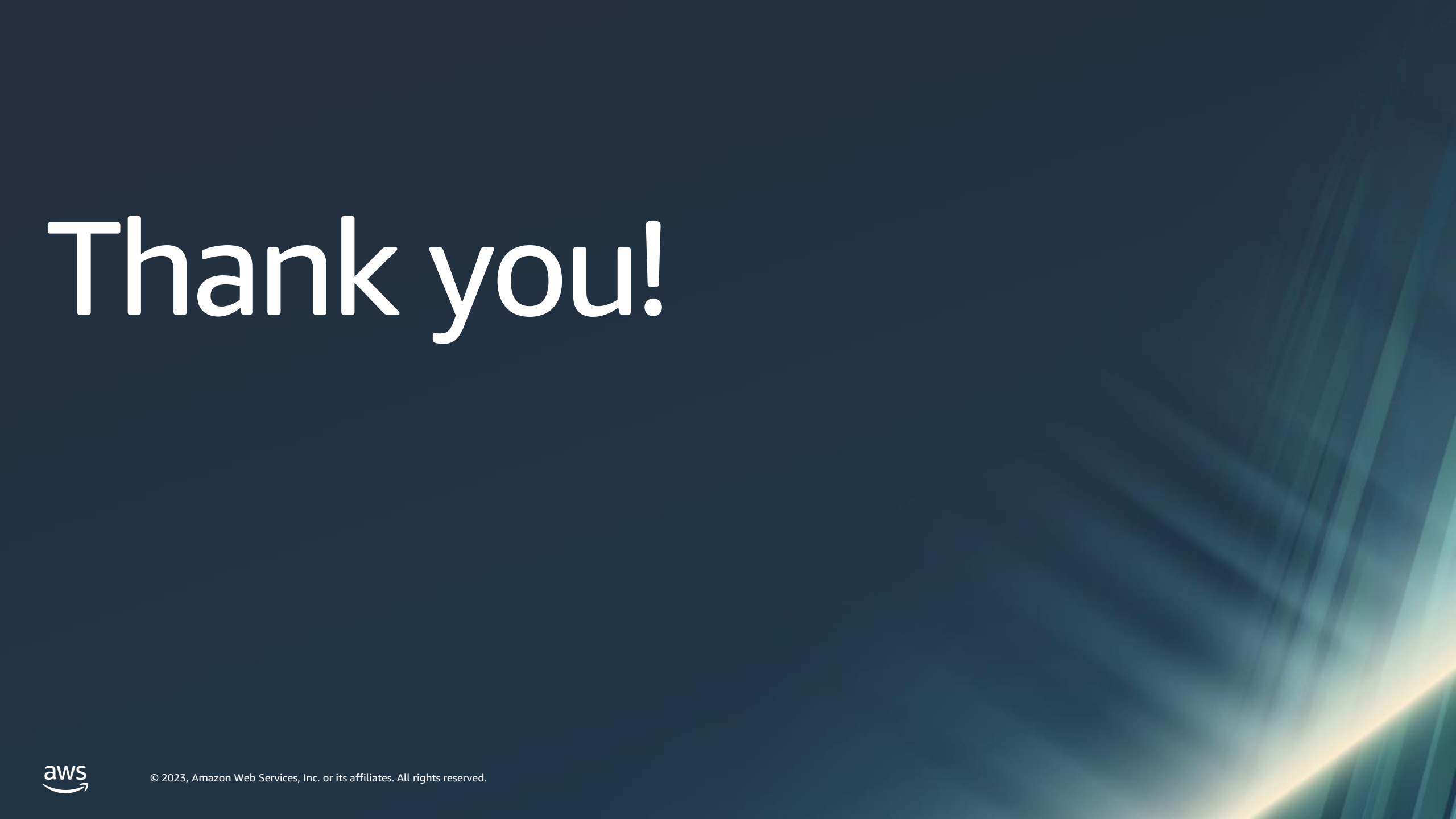

youtube.com/user/AmazonWebServices


slideshare.net/AmazonWebServices


twitch.tv/aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.