aws INNOVATE
DATA EDITION

23 August, 2022

Go beyond insights to predictive analytics
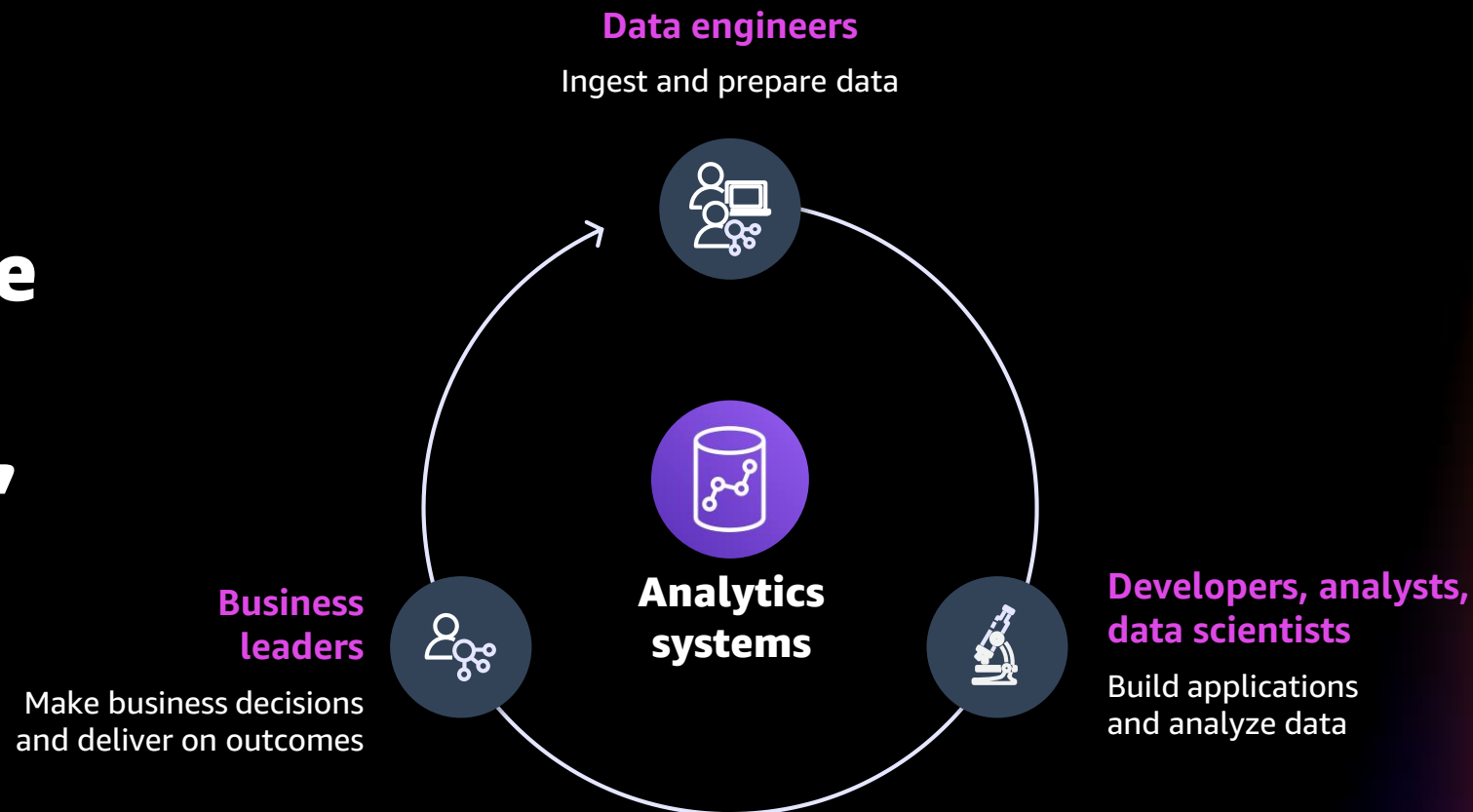# with Amazon Redshift ML and Amazon SageMaker Canvas

Mary Law

Senior Manager
Analytics Acceleration Lab
Amazon Web Services

aws

Agenda

- Amazon Redshift overview

- Benefits of machine learning

- Introduction to Amazon Redshift Machine Learning

- Introduction to Amazon SageMaker Canvas

- Demo: Build ML model with Amazon Redshift ML and BYOM

- Summary and additional resources

Data can work more effectively for diverse users through easy-to-deploy, self-serve, and auto-scaling analytics systems

Data engineers
Ingest and prepare data

Analytics systems

Business leaders
Make business decisions and deliver on outcomes

Developers, analysts, data scientists
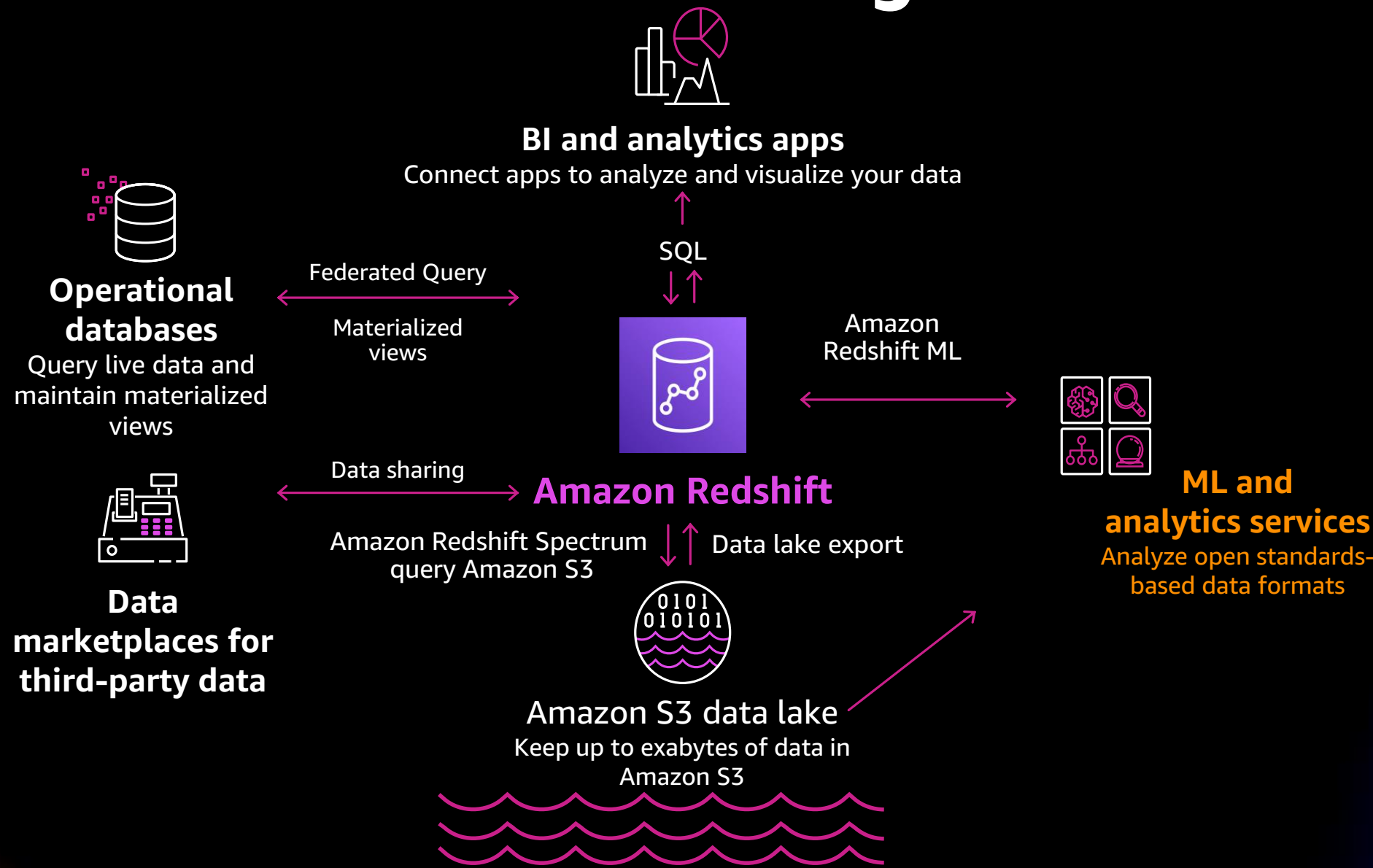Build applications and analyze data

Improve operational efficiency; make more informed decisions; accelerate innovation

*By making 10% more data accessible, a typical Fortune 1000 company will see a **$65 million increase in net income**[1]*

[1] Dykes, *"The Four Key Pillars to Fostering a Data-Driven Culture"*
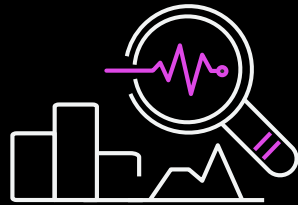
Amazon Redshift break through data silos

BI and analytics apps
Connect apps to analyze and visualize your data

Operational databases
Query live data and maintain materialized views

Federated Query

Materialized views

SQL

Amazon Redshift ML

Amazon Redshift

ML and analytics services
Analyze open standards-based data formats

Data sharing

Data marketplaces for third-party data

Amazon Redshift Spectrum query Amazon S3

Data lake export

Amazon S3 data lake
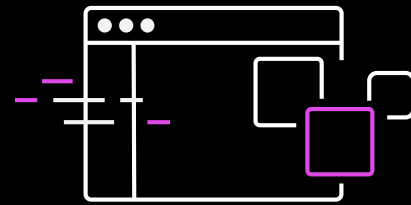Keep up to exabytes of data in Amazon S3
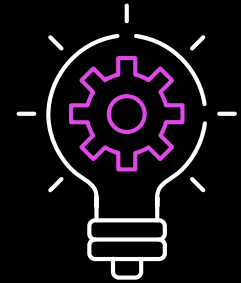
Benefits of Machine Learning

Transform
customer experience

Improve business
operations

Better and faster
decision-making
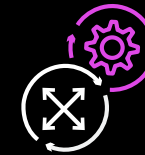
Innovate product
or service

Common ML Use Cases

Sales and Marketing

1. Sales conversion
2. Sales forecasting
3. Propensity to churn
4. Customer lifetime value prediction
5. Marketing mix modeling

Finance and Accounting

1. Credit risk scoring
2. Delayed payments prediction
3. Fraud detection
4. Portfolio optimization
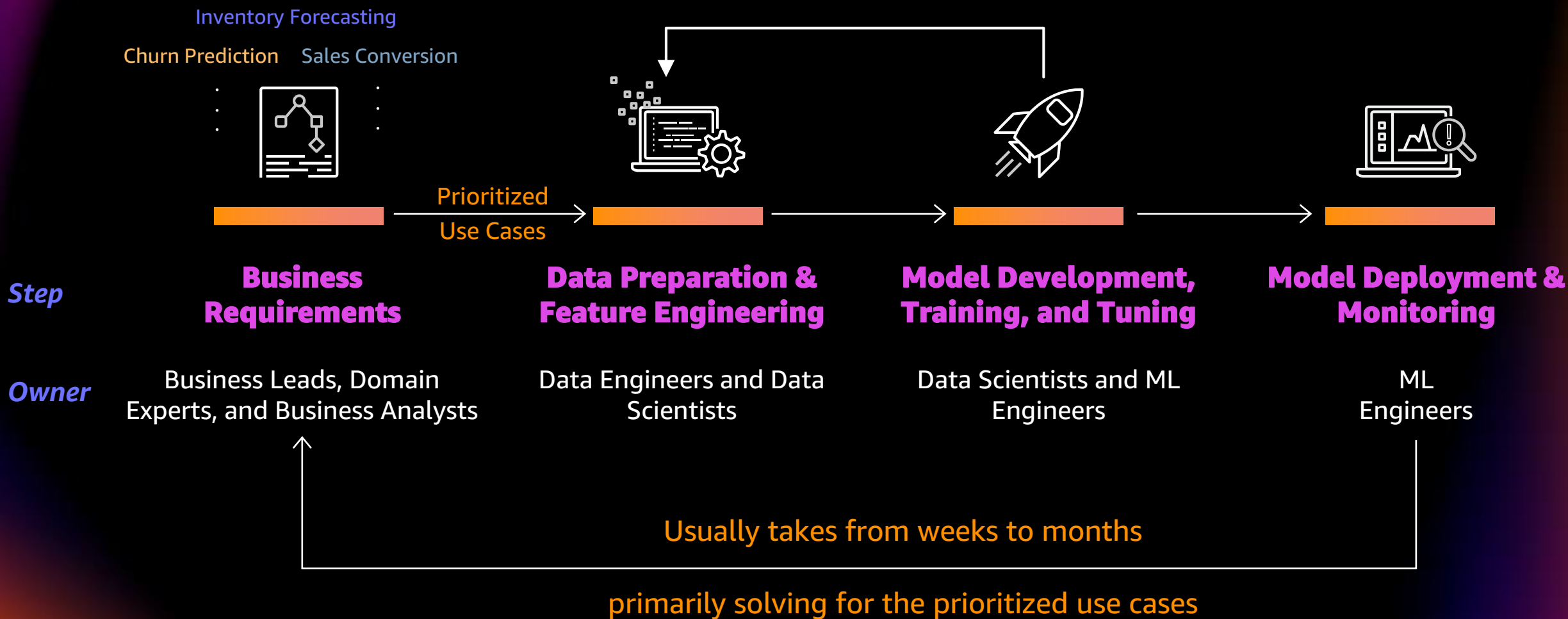5. Account payables automation

Operations and Logistics

1. Demand forecasting
2. Inventory planning and scheduling
3. Delivery time forecasting
4. Predictive maintenance
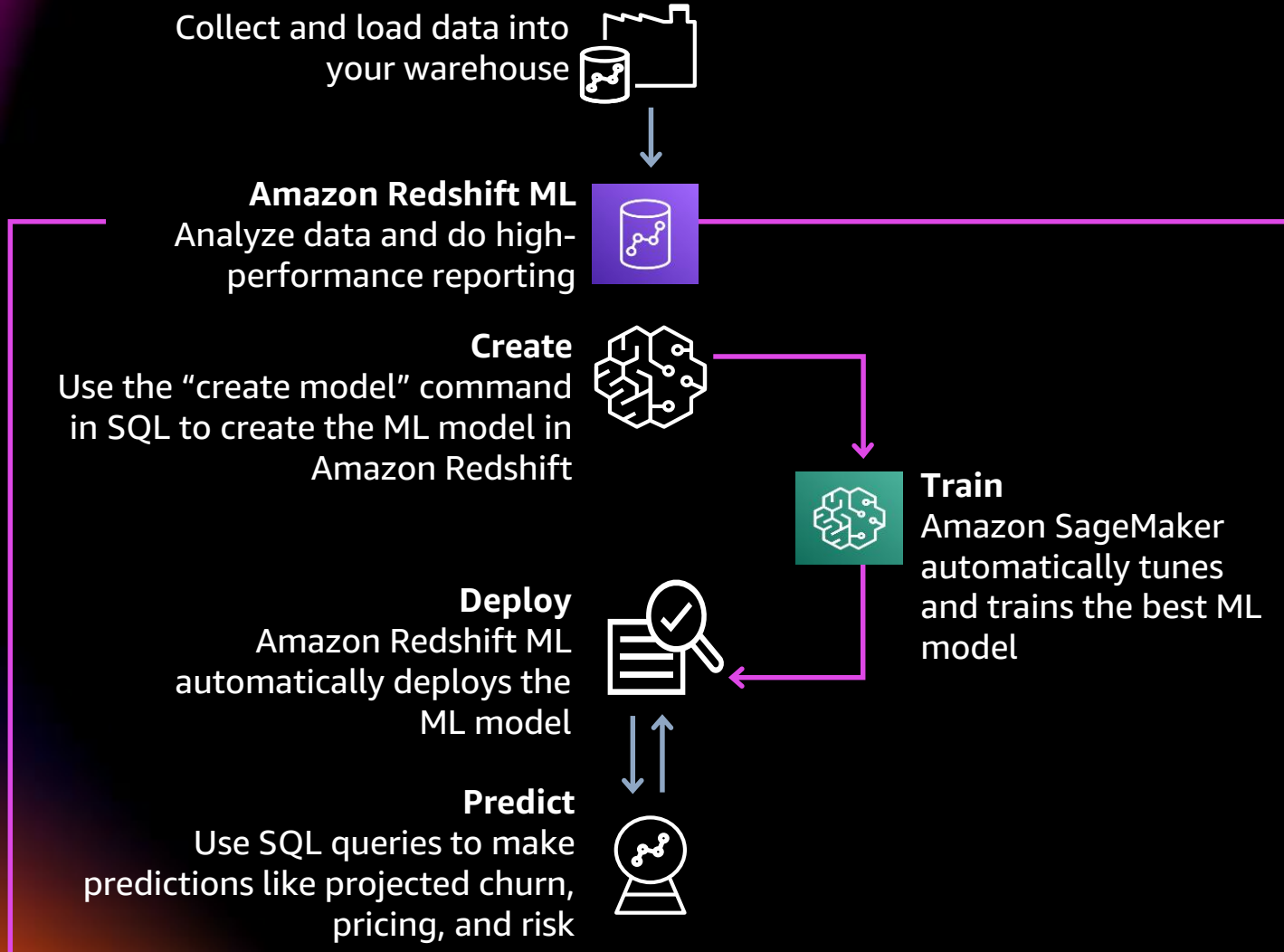
and many more…

How ML drives value creation today

Inventory Forecasting

Churn Prediction   Sales Conversion

Prioritized Use Cases

Step

Business Requirements

Data Preparation & Feature Engineering

Model Development, Training, and Tuning

Model Deployment & Monitoring

Owner

Business Leads, Domain Experts, and Business Analysts

Data Engineers and Data Scientists

Data Scientists and ML Engineers

ML Engineers

Usually takes from weeks to months

primarily solving for the prioritized use cases

aws

Amazon Redshift ML

Put the power of ML in the
hands of every data analyst, database
developer, and every data warehouse user.

How Amazon Redshift ML works

SQL users can create models and make predictions on data in their data warehouse

Collect and load data into your warehouse

Amazon Redshift ML
Analyze data and do high-performance reporting

Create
Use the "create model" command in SQL to create the ML model in Amazon Redshift

Train
Amazon SageMaker automatically tunes and trains the best ML model

Deploy
Amazon Redshift ML automatically deploys the ML model

Predict
Use SQL queries to make predictions like projected churn, pricing, and risk

Automatically create and train ML models using machine learning algorithms supported by Autopilot

Apply ML models on Amazon Redshift data using standard SQL

Embed predictions like fraud detection and risk scoring

Bring Your Own Model - You can use a model trained outside of Redshift with Amazon SageMaker for in-database inference local in Amazon Redshift

aws

Amazon Redshift ML: Optimized for cost

Typically predictions drive cost in production

You only pay for training while **prediction comes at no extra cost** when you use Amazon Redshift ML
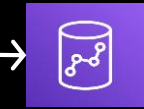
Training
10%

Cost

Prediction (inference)
90%
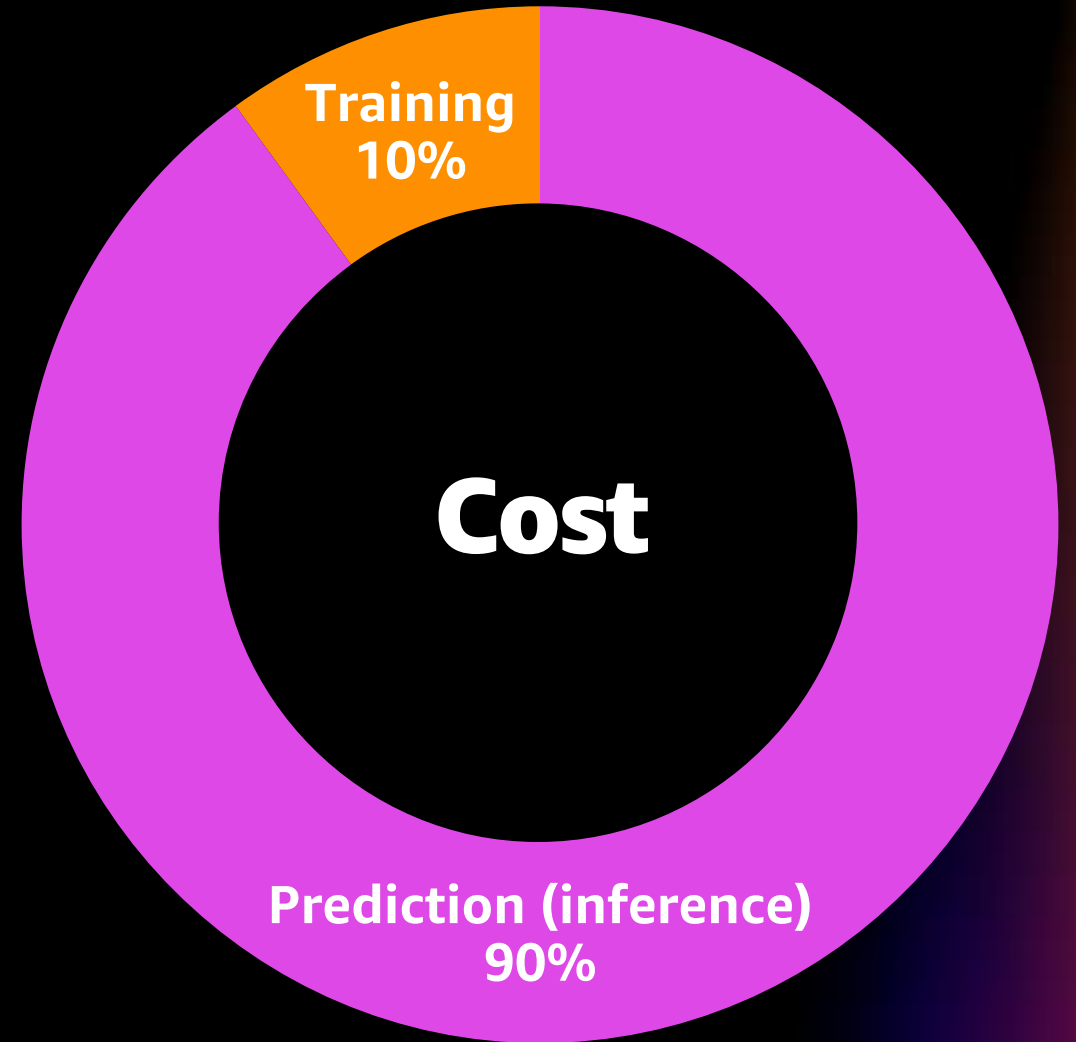
How Amazon Redshift ML works

CREATE & TRAIN

```
CREATE MODEL
cust_act_churn_auto_acc
FROM (SELECT state,account_length,
area_code,average_daily_spend,
average_daily_cases,churn
FROM customer_act  WHERE
ACCOUNT_LENGTH > 120)
TARGET churn
FUNCTION
ml_fn_cust_act_churn_auto_acc
```

*Runs
Amazon SageMaker Autopilot
and returns model*

Amazon
Redshift

Amazon
SageMaker

Create, train, and deploy model
with a simple SQL command

Auto-selection of model,
pre-processing, and training
using SageMaker Autopilot

Trained model gets compiled by
SageMaker Neo in Amazon Redshift
data warehouse so that you can
make predictions using SQL

PREDICT

```
SELECT phone,
ml_fn_cust_act_churn_auto_acc(
state,account_length,area_code,
average_daily_spend,average_daily_
cases)AS active
FROM customer_act
WHERE record_date >'2020-01-01';
```

Amazon
Redshift

*Uses previously built model to predict in-place
(inference executed entirely in Amazon Redshift)*

Demo overview:

1 – Amazon Redshift ML : AUTO On

2 – BYOM and intro to Amazon SageMaker Canvas

aws

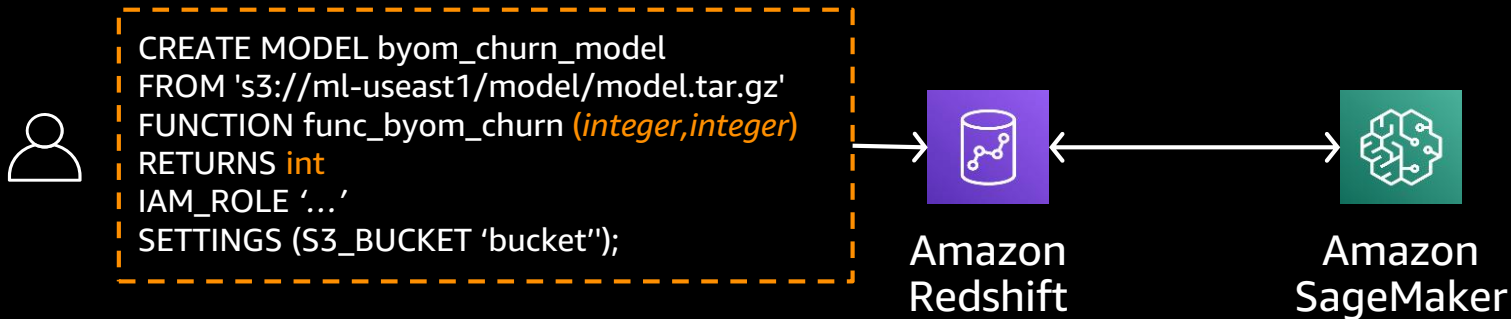# Demo 1 – Amazon Redshift ML : AUTO On

Demo 2 –
Build Your Own Model with Amazon SageMaker Canvas and
Bring Your Own Model back into Amazon Redshift

Bring your model to Amazon Redshift ML

IMPORT SAGEMAKER TRAINED MODELS FOR LOCAL INFERENCE

CREATE MODEL

```
CREATE MODEL byom_churn_model
FROM 's3://ml-useast1/model/model.tar.gz'
FUNCTION func_byom_churn (integer,integer)
RETURNS int
IAM_ROLE '...'
SETTINGS (S3_BUCKET 'bucket");
```

Amazon
Redshift

Amazon
SageMaker

Create, train, and deploy model
in Amazon SageMaker. Make
available in Amazon Redshift via
SageMaker External Function

Invoke your model from
Amazon Redshift

Provides you full flexibility
and algorithms of
Amazon SageMaker

PREDICT

```
SELECT phone,
func_byom_churn (state,
account_length, area_code,…)
As active
FROM customer_activity
```

Amazon
Redshift

Amazon
SageMaker

We will demo the new Amazon SageMaker Canvas

aws

Amazon SageMaker Canvas

Build ML models and generate accurate predictions—no code required

Can quickly **connect to and access data** from disparate sources (local disk, Amazon S3, Amazon Redshift)

Leverages powerful AutoML technology to automatically **train and build models** based on your dataset
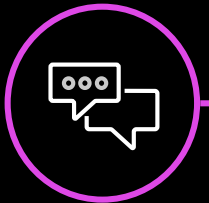
Integrates with SageMaker Studio, making it easy to **share models** with data scientists

Usage-based pricing to avoid licensing fees and reduce TCO

aws
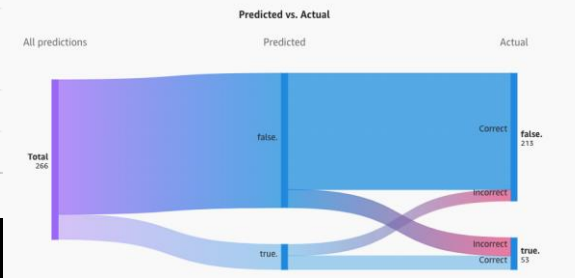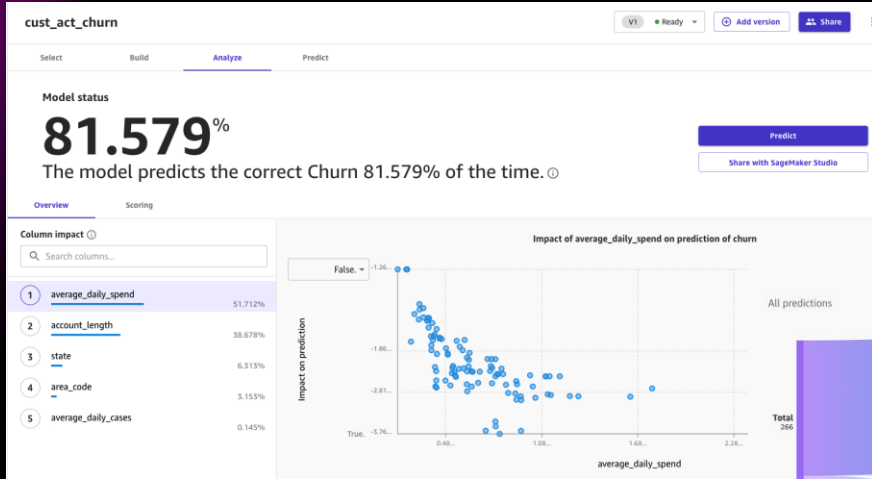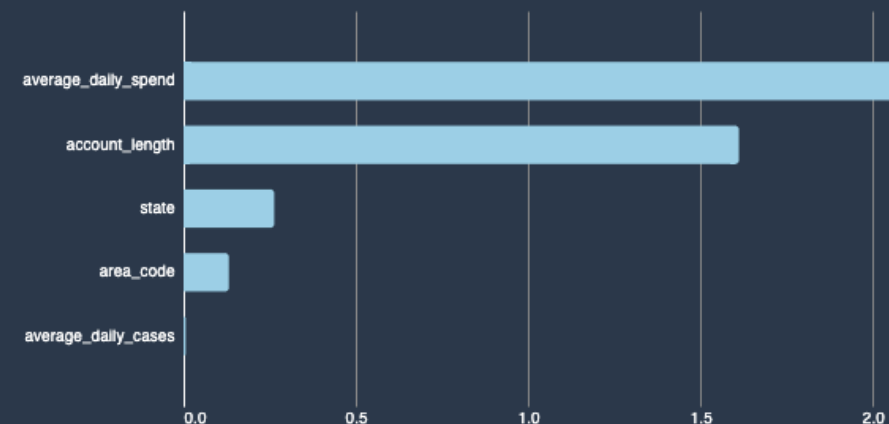
Amazon SageMaker Canvas overview
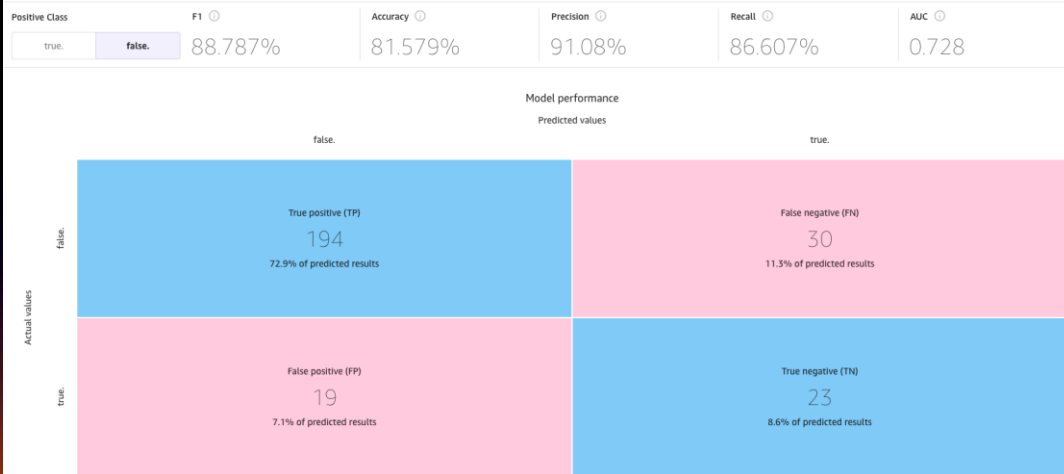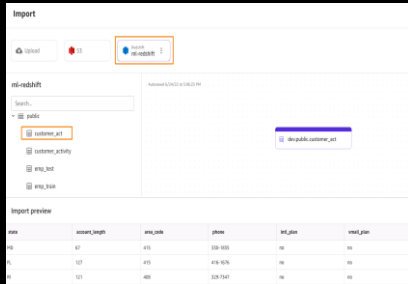
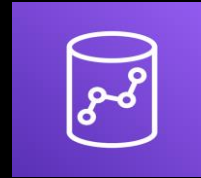# Demo 2 : High level overview

1 Connect to Amazon Redshift to select dataset

Amazon Redshift

+

Amazon SageMaker Canvas

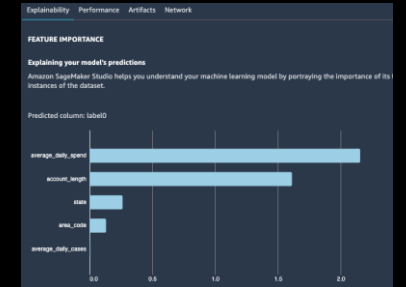**3** BYO model back into Amazon Redshift for deployment
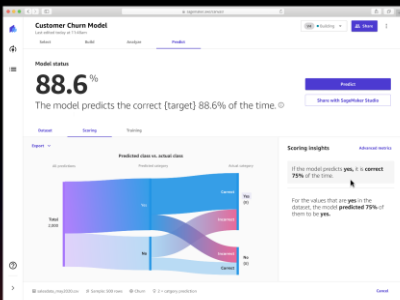
2 Build model with Amazon Sagemaker Canvas

88.6%
The model predicts the correct (target) 88.6% of the time.

4 Run local inference in Amazon Redshift
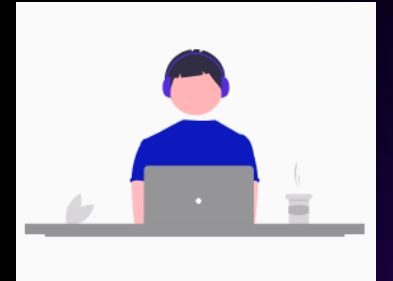
Demo 2 –
Amazon SageMaker Canvas + BYOM into Amazon Redshift

Recap

Amazon Redshift Machine Learning

1. Democratize ML to data analysts by simplifying the creation of model

2. Simplify pipeline and eliminate data movement for ML by moving models closer to data

3. Enable ML experts such as data scientists to BYOM to Amazon Redshift

Amazon SageMaker Canvas

1. No ML experience required

2. Visual point-and-click interface, allowing you to generate ML predictions on your own

Other resources

- Amazon Redshift – Overview and features

- Amazon Redshift ML Blog

- Amazon Redshift Database Developer Guide

- GitHub Repository – Getting started with Amazon Redshift ML

- Bring Your Amazon SageMaker model into Amazon Redshift for remote inference

- Amazon SageMaker Canvas – Overview and features

Visit the AWS Data resource hub

A modern data strategy can help you manage, act on, and react to your data so you can make better decisions, respond faster, and uncover new opportunities. Dive deeper with these resources today.
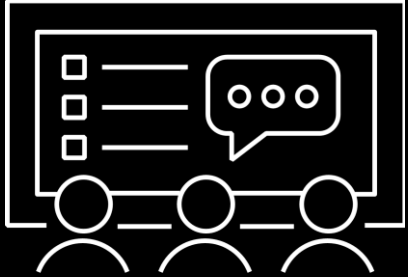
- Harness data to reinvent your organization
- In unpredictable times, a data strategy is key
- Make data a strategic asset
- Rewiring your culture to be data-driven
- Put your data to work with a modern analytics approach
- … and more!

https://tinyurl.com/data-hub-aws
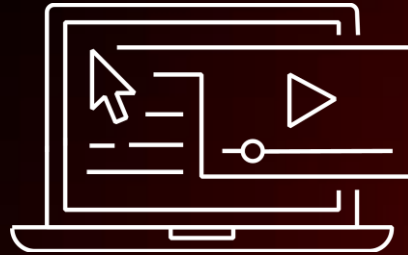
Visit resource hub

AWS Training and Certification for Data and Analytics

AWS Data & Analytics FREE Training Resources

Discover how to harness data, one of the world's most valuable resources, and innovate at scale.

https://bit.ly/3Ntlhy7

AWS Data Analytics Learning Plan

This learning plan expose you to the fastest way to get answers from all your data to all your users. It can also help prepare you for the AWS Certified Data Analytics - Specialty certification exam.

https://bit.ly/3wBVjD1

AWS Certified Data Analytics - Specialty

Earning AWS Certified Data Analytics – Specialty validates expertise in using AWS data lakes and analytics services.

https://go.aws/3lwF0RR

aws

Thank you for attending AWS Innovate – Data 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

aws

Thank you!

https://www.linkedin.com/in/awsmary/

aws