11.18.2021

PARTNER DEV DAY

aws | SingleStore | looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

09.2021

Operational Analytics

Sugandan Barathy
Partner Solutions Architect
SingleStore

Saurabh Shanbhag
Partner Solutions Architect
AWS

Sebastian Fabri
Solutions Architect
Looker

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Today’s Agenda (1 Hour)

- **Introduction**
- **Workshop Preparation**
- **Building The Operational Analytics Solution**
  - Step 1 - Setup Database and Load Data
  - Step 2 - Deploy a Real-Time Analytics Stored Procedure
  - Step 3 - AWS Glue ETL leveraging SingleStore Spark Connecter
  - Step 4 - Looker Visualization
- **Cleanup AWS Resources**

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Remote Workshop Teaching Format

Note (1) : Each student should have both the Workshop Presentation and Student Guide PDF downloaded to your local machine.

Note (2) : Each Step will taught in the following format

- Lecture
- Demonstration

Listening to the Lecture and watching the demonstration will increase your success rate.

Note (3) : You will always have a week time to finish after the session.

We will share with you which steps have dependencies.

You will gain valuable knowledge and we can help you later if you cannot complete a section.

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Student / Classroom Interaction with Instructors

We will not be using GoToMeeting **“Questions”** Panel option during this workshop.

Students, please enter your questions in **“Chat”.** They will be responded to in “Chat” by our instructors.

Instructors will also monitor and respond to **Student Questions in “Chat”**

- Instructors will be monitoring this Classroom iNotepad as well and providing answers.
- Note that we will keep Common Student Questions listed at the top of the Q&A section

A classroom **iSpreadsheet** will be used to distribute students their AWS credits

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Introduction

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HANDS ON WORKSHOP

Today's Agenda

- **Introduction**
- **Workshop Preparation**
- **Building The Operational Analytics Solution**
  - Step 1 - Setup Database and Load Data
  - Step 2 - Deploy a Real-Time Analytics Stored Procedure
  - Step 3 - AWS Glue ETL leveraging SingleStore Spark Connecter
  - Step 4 - Looker Visualization
- **Cleanup AWS Resources**

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why Operational ML?

Predictive Maintenance Through Anomaly Detection
Examples: Proactive Airline Maintenance; Reduced Manufacturing Downtime

Personalization
Example: Improve Custom Website Experience

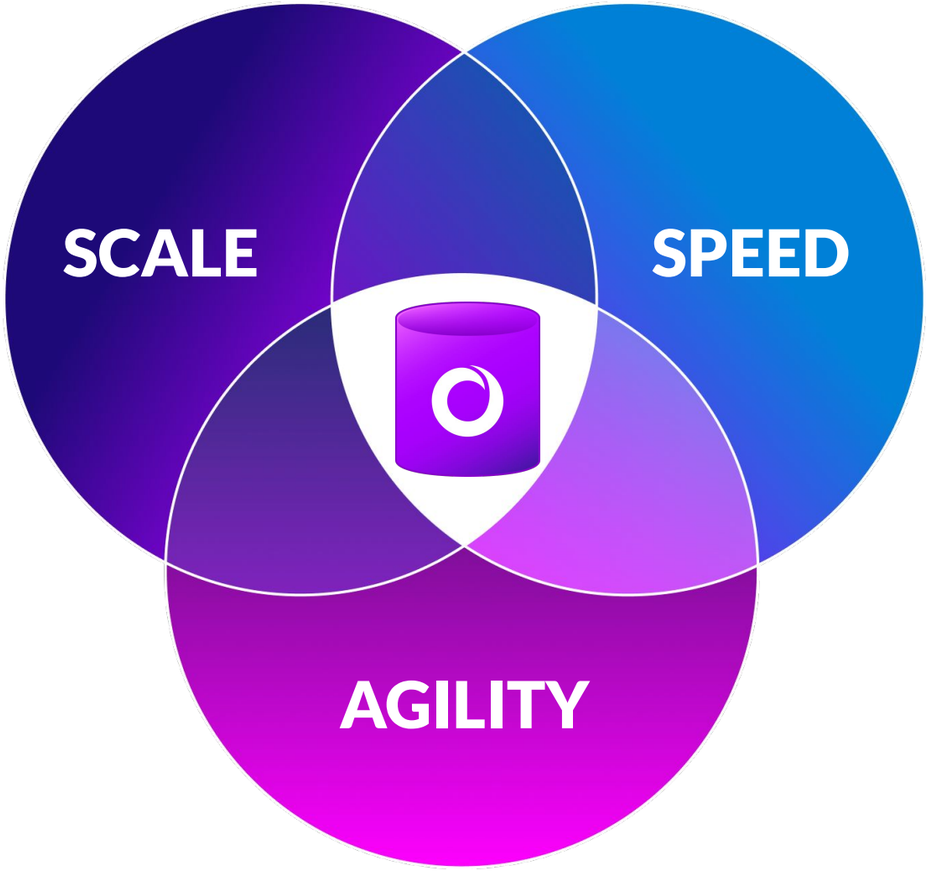
Cross Sell, Upsell
Example: Increase Sales for Retail eCommerce

Improved Customer Service
Example: Increase Customer Satisfaction with Time-to-Solution

aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The Unified Database For Fast Analytics



Modern Hyper-scale Applications

Tier-1 applications in a business where complex decisions must be made immediately

> **"Can't Wait" Analytics**
> Business facing applications and dashboards with low latency, high concurrency analytics needs

> **"Won't Wait" Customer Experiences**
> Customer-facing applications where speed and personalization drives everything

Predictive Analytics (ML & AI)

SingleStore supports the creation and operation of models at scale across streaming & historical data

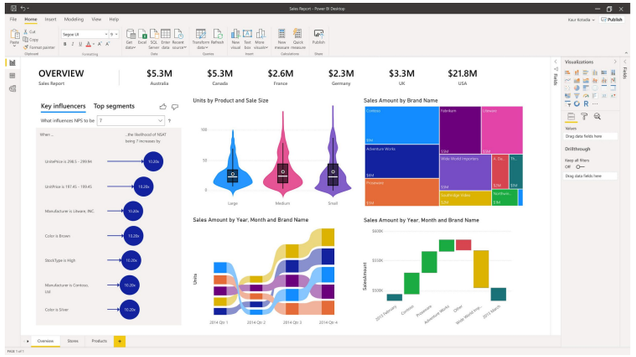
IoT Analytics

Stream real-time device and sensor data and make immediate decisions as data is streamed

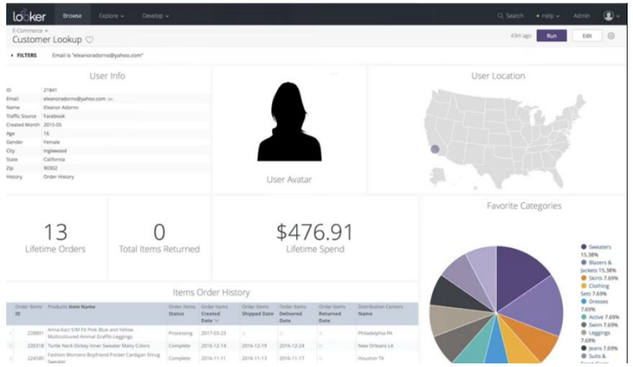
SingleStore

SingleStore for Fastboards

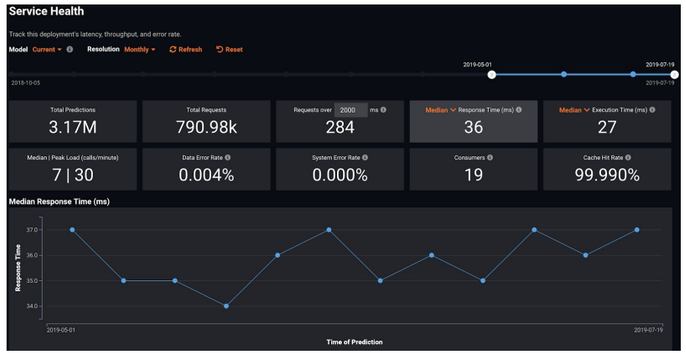
Real-Time Analytics



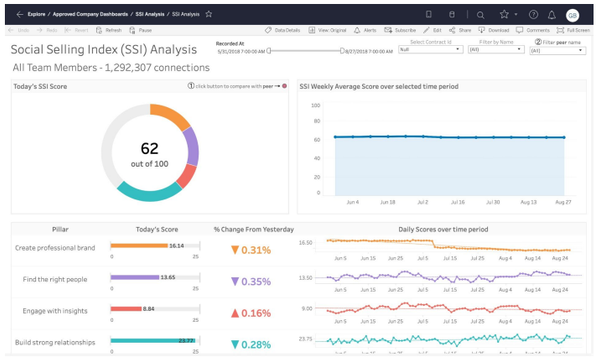
Operational BI



Ad Hoc Data Discovery



ML-Driven
Analytics



Historical
Reporting

SingleStore

How SingleStore Compares for Operational Data Workloads

	SingleStore	MongoDB	Oracle Exadata	PostgreSQL
Modern Distributed highly available architecture, legacy and modern data sources, structured and semi-structured	●	◑	◑	◔
Ingest and Transactional Performance Non-blocking, stream ingest, high concurrency, transaction management	●	◔	◕	◑
Analytical Query Performance ANSI SQL, modern query execution	●	○	◑	◔
Cloud and deployment flexibility Public or hybrid cloud, on-prem via standard hardware, containers or VMs	●	●	◑	◔
AI/ML integrations Real-time scoring, fast exploration and training, predictive analytics	◕	◑	◕	◑
Manageability Devops-ready with high availability	◕	◕	◑	◕
Cost Software and on-demand licensing	●	◕	○	◕

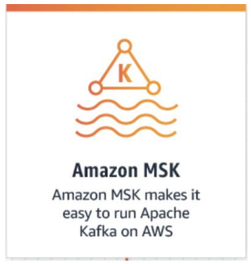
● Excellent ◕ Very Good ◑ Good ◔ Fair ○ Poor

SingleStore

SingleStore Compliments AWS DB Strategy and drives Cloud Consumption

Streaming Services

Kafka, Spark, CDC



Amazon EC2

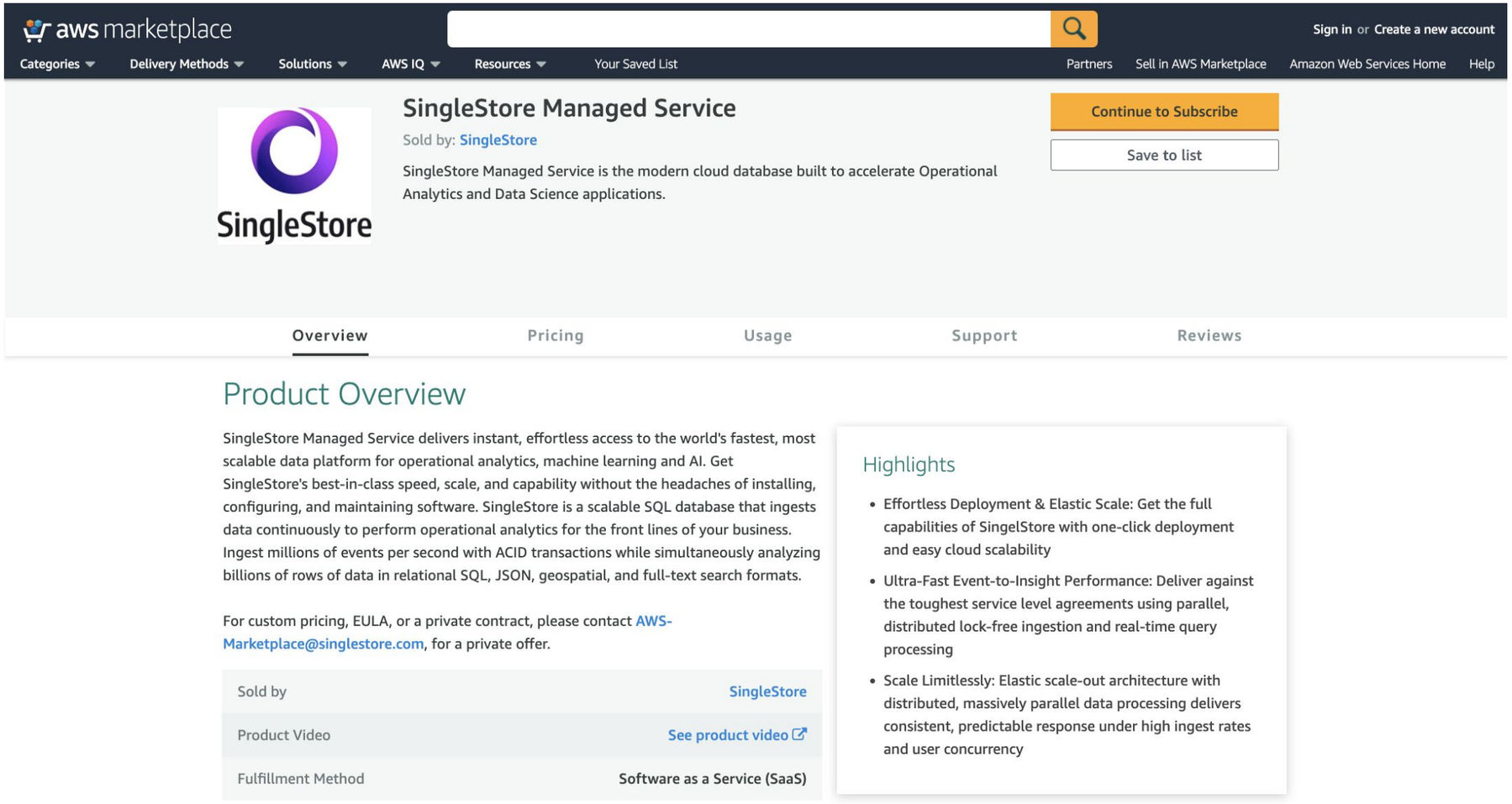
Amazon S3

BI, Analytics, Applications and Solutions

SingleStore

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

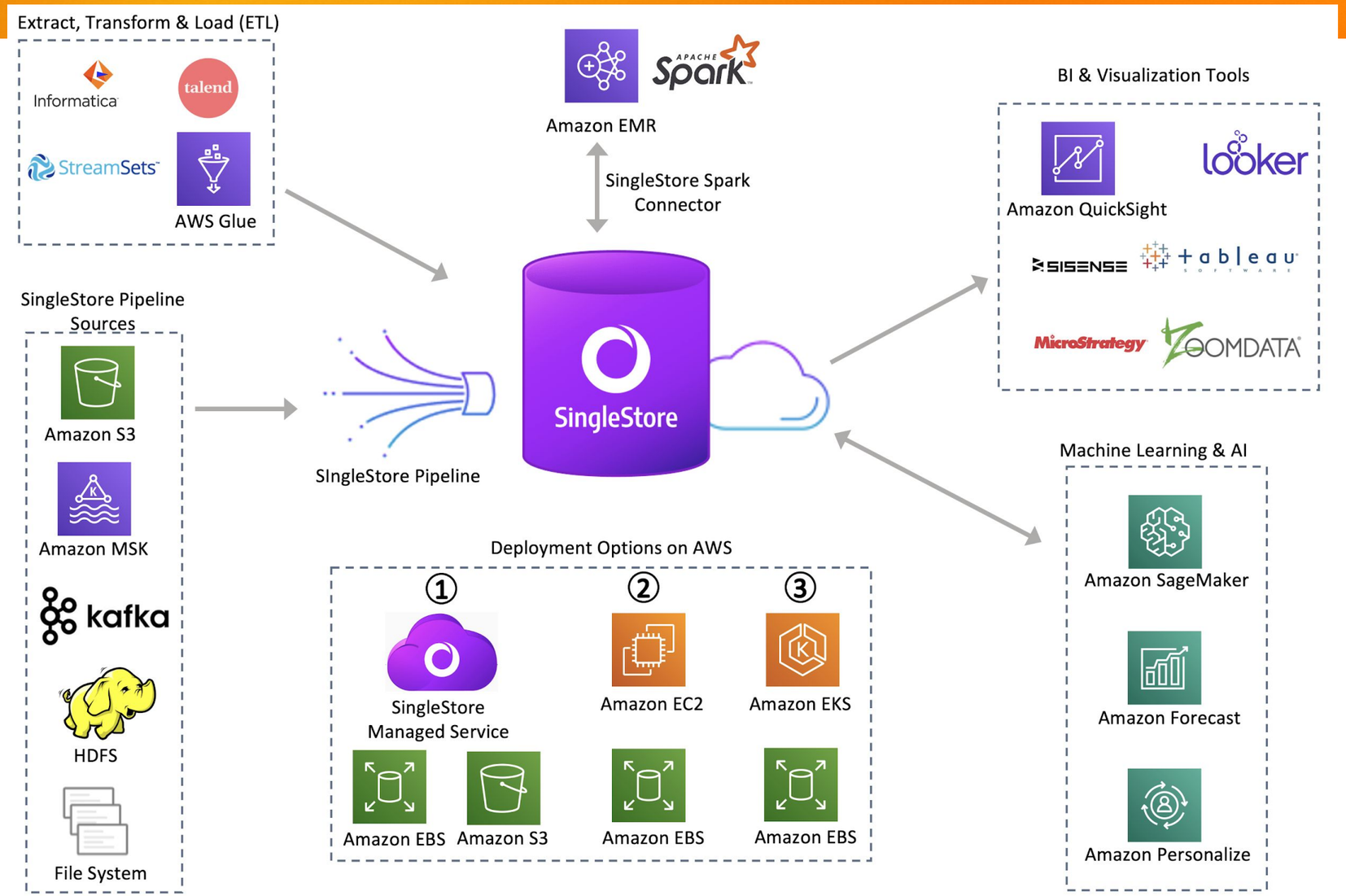
SingleStore Managed Service



Cloud Native Database as a Service

- **Effortless Deployment and Elastic Scale**
- **Ultra-Fast Event-to-Insight Performance**
- **Scale Limitlessly**
  - Ingestion Rates
  - Concurrency (Users, Jobs)
  - Workloads

SingleStore Compliments AWS DB Strategy and drives Cloud Consumption



SingleStore:

1. Managed Service
2. Self Managed DB on Instances
3. Self Managed DB on Containers

Driving Storage and AWS Services!!!!!!!

SingleStore Bottomless on S3 GA since July 2021 !!!

Bottomless

- Separation of storage and compute for SingleStore
- Data lives on object store
- Data cached locally on SSDs
- No knobs
- Enabler for PITR
- Flexible pricing

Photo by Timo Volz on Unsplash

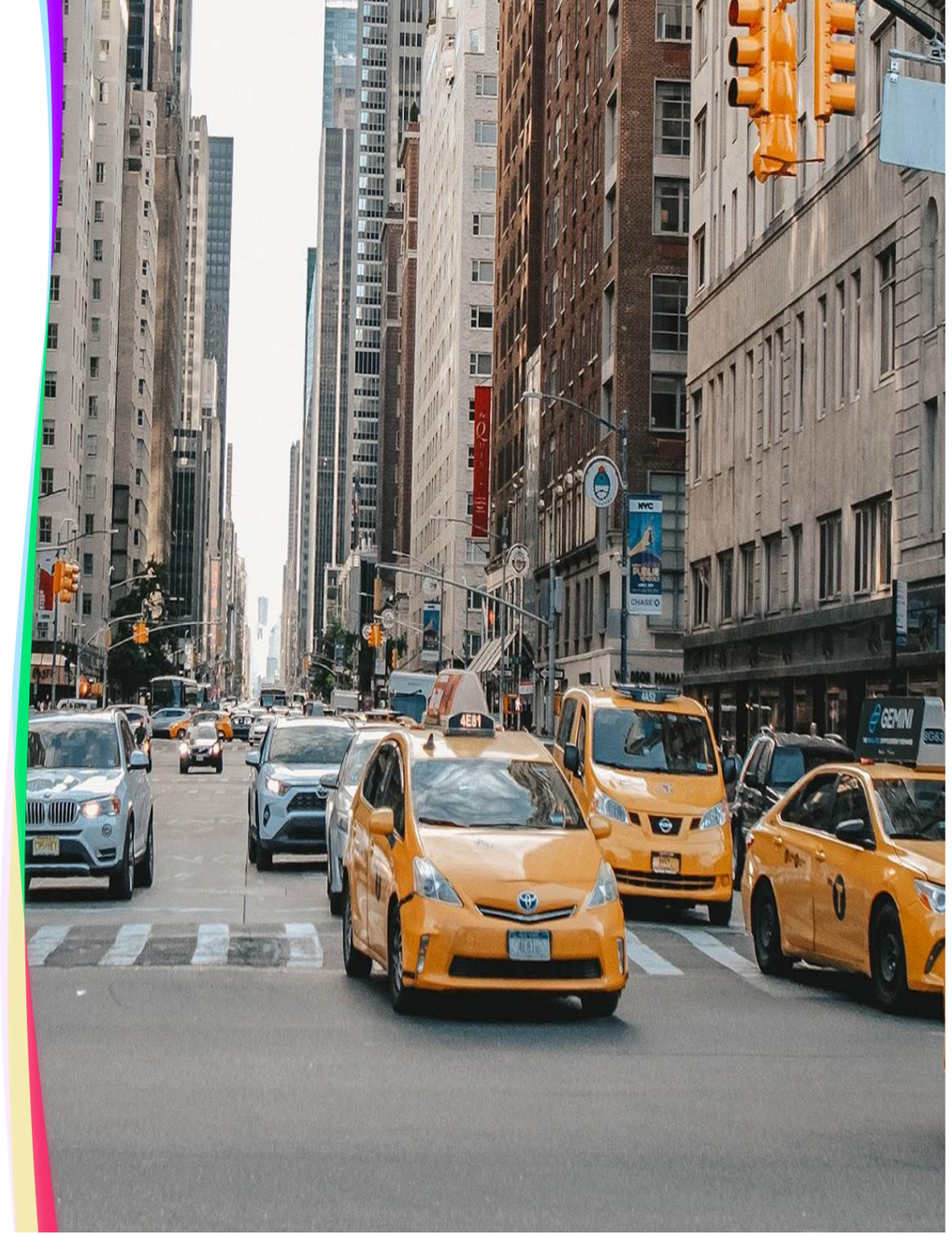
https://www.singlestore.com/blog/the-future-is-bottomless/
https://www.youtube.com/watch?v=aFpB_TSyJqE

OPERATIONAL ANALYTICS

NYC Taxicab Operational Dashboard

Demonstrate how to build an IoT-based operational “real-time” dashboard

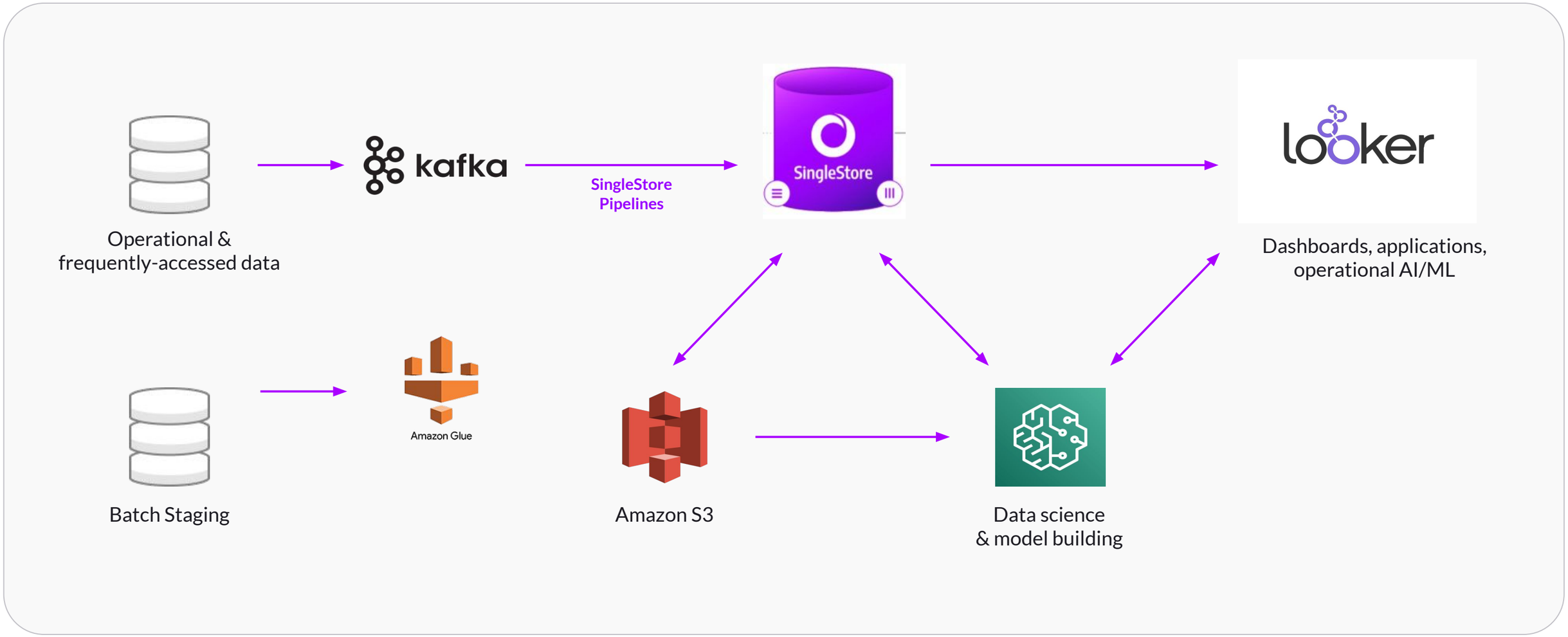
- Leveraging live NYC taxicab event streams
- Running analytics within the SingleStore database
- Scoring each streaming event using ML



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS & SingleStore

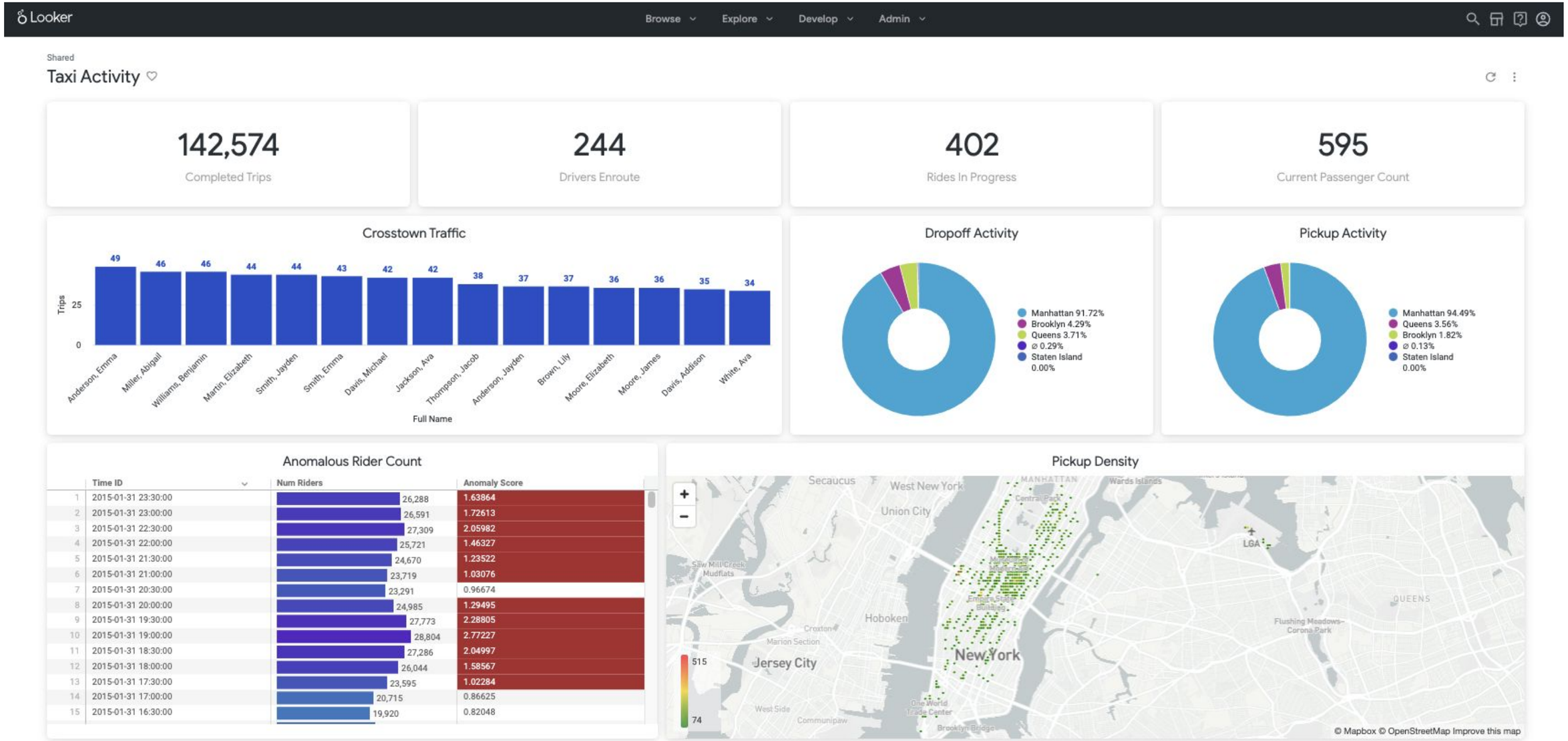
SingleStore becomes the de facto store for “always on” data and analytics “with an SLA”



aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

We'll Build This Dashboard



aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HANDS ON WORKSHOP

Today’s Agenda (3 Hrs)

- **Introduction**
- **Workshop Preparation**
- **Building The Operational Analytics Solution**
  - Step 1 - Setup Database and Load Data
  - Step 2 - Deploy a Real-Time Analytics Stored Procedure
  - Step 3 - AWS Glue ETL leveraging SingleStore Spark Connecter
  - Step 4 - Looker Visualization
- **Cleanup AWS Resources**

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Workshop Preparation

Section A - Student Guide

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Workshop Preparation

Please see detailed step by step instructions for preparation steps in the Student Guide - Section A.

By setting up the services and environment up front, we will be able to accelerate the Build Phase of this workshop (Student Guide - Section B).

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

Preparation Steps for Hands on later

1. Download Student Guide and Presentation Deck
2. Decide Which AWS Account to Use in this Workshop
3. From the AWS Console, login with your root account and perform admin tasks:
    i. Create an AWS IAM user
4. Launch your own SingleStore Fully Managed Cluster
5. AWS ETL Glue with Spark Prep
6. Launch an AWS CloudShell Service

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

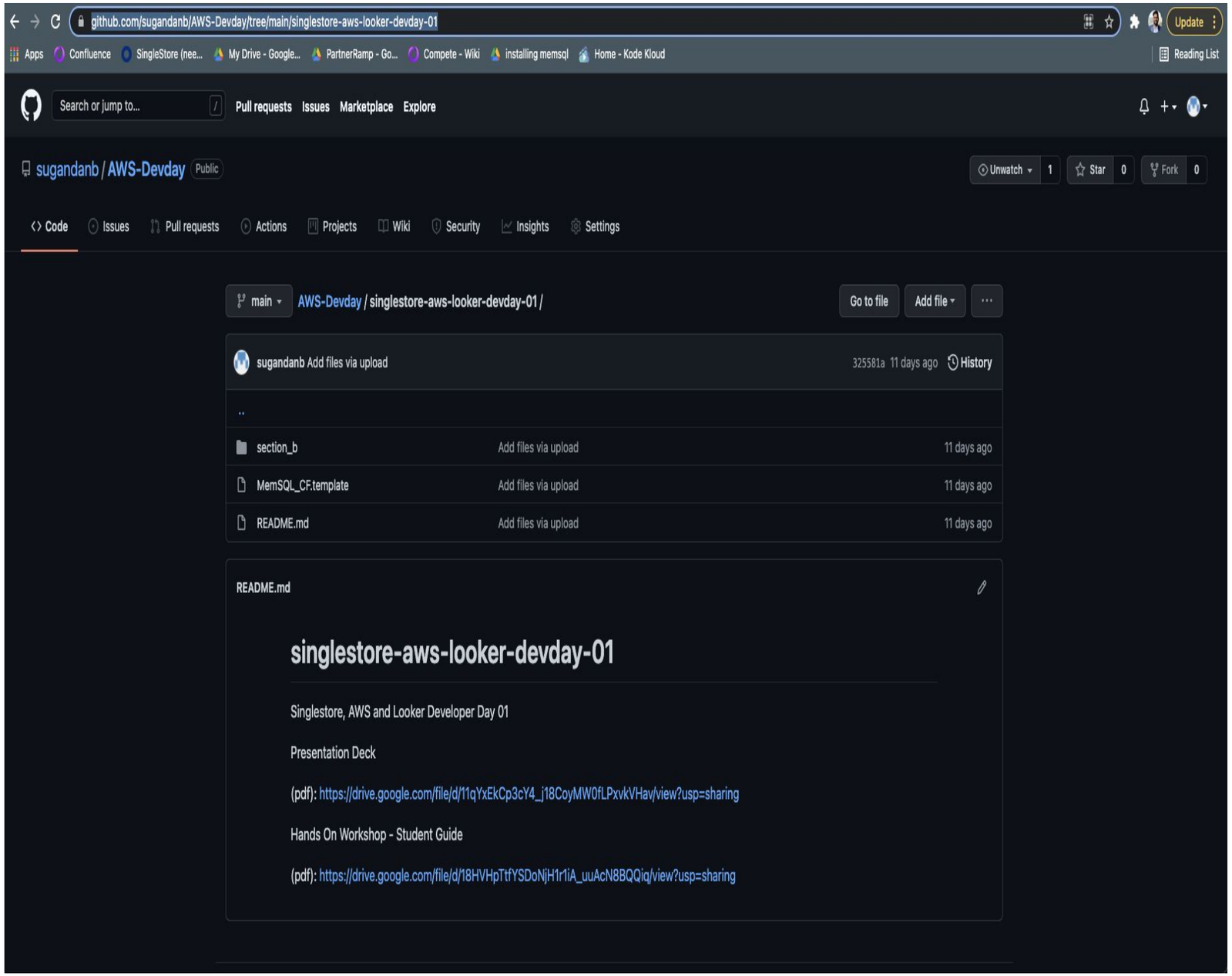
Preparation Steps for Hands on Workshop

1. **Download Student Guide and Presentation Deck**
2. Decide Which AWS Account to Use in this Workshop
3. From the AWS Console, login with your root account and perform admin tasks:
    i. Create an AWS IAM user
4. Launch your own SingleStore Fully Managed Cluster
5. AWS ETL Glue with Spark Prep
6. Launch an AWS CloudShell Service

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GitHub for Today's Workshop

[Dev Day GitHub Link]



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Step 1 : Download Student Guide and Presentation Deck

Dev Day GitHub Link

Within the Readme. Scroll down and download the following two items for today's workshop:

(1) PRESENTATION DECK

(2) STUDENT GUIDE

You can also find these two PDFs in the GoToMeeting Panel and download them there.

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

Preparation Steps for Hands on Workshop

1. Download Student Guide and Presentation Deck
2. **Decide Which AWS Account to Use in this Workshop**
3. From the AWS Console, login with your root account and perform admin tasks
    i. Create an AWS IAM user
4. Launch your own SingleStore Fully Managed Cluster
5. AWS ETL Glue with Spark Prep
6. Launch an AWS CloudShell Service

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

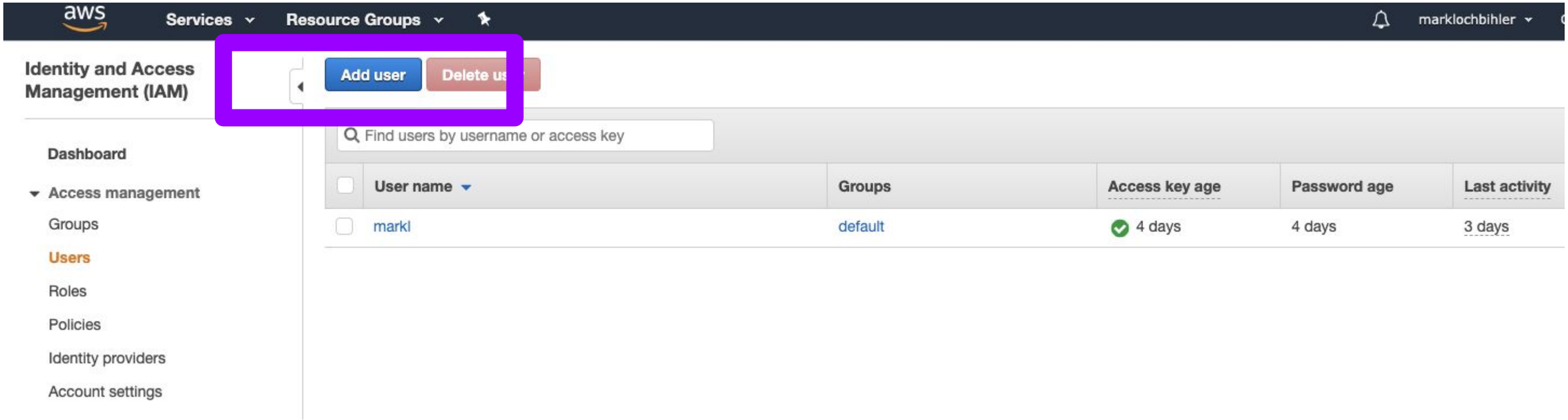
Preparation Steps for Hands on Workshop

1. Download Student Guide and Presentation Deck
2. Decide Which AWS Account to Use in this Workshop
3. **From the AWS Console, login with your root account and perform admin tasks:**
    i. **Apply an AWS Student Credit ($100) provided by the Instructor**
    ii. **Create an AWS IAM user**
4. Launch your own SingleStore Fully Managed Cluster
5. AWS ETL Glue with Spark Prep
6. Launch an AWS CloudShell Service

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Step 3.i

Create an IAM user and group



Note : The IAM account you create will be used for the remainder of this workshop. Please follow the instructure demonstration to create this IAM id and group. Make sure it has Administrative Permissions.

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

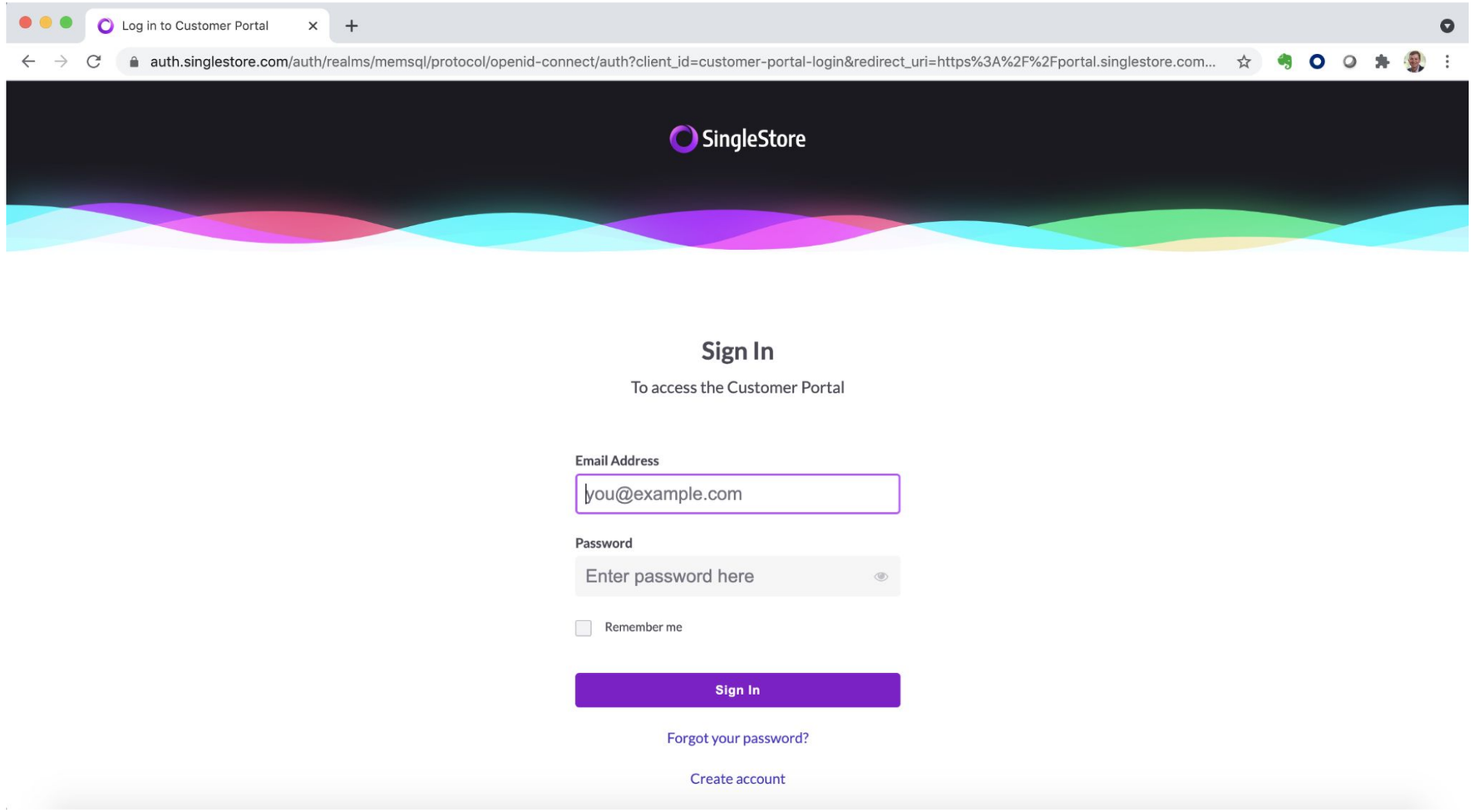
Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

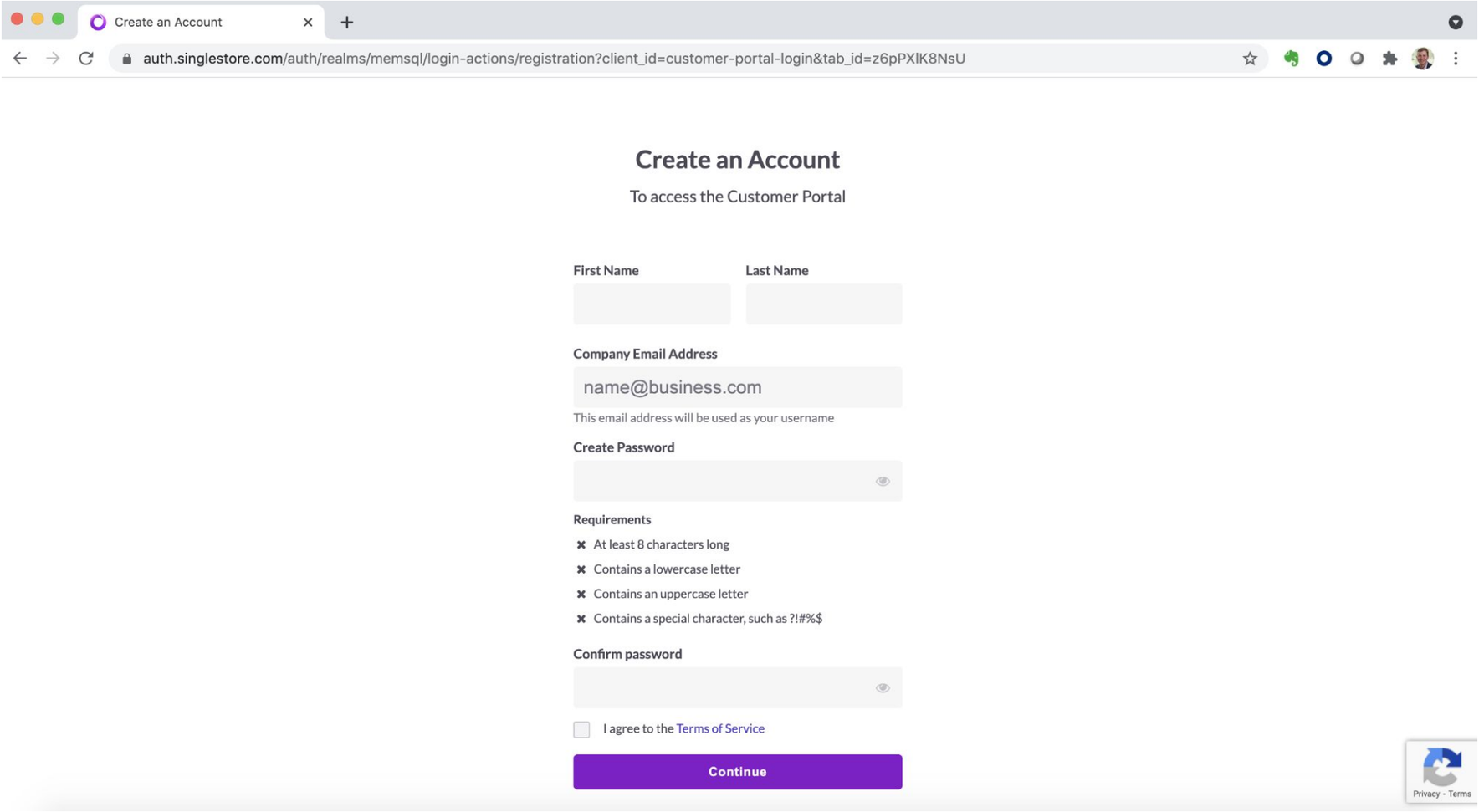
Preparation Steps for Hands on Workshop

1. Download Student Guide and Presentation Deck
2. Decide Which AWS Account to Use in this Workshop
3. From the AWS Console, login with your root account and perform admin tasks:
    i. Create an AWS IAM user
4. **Launch your own SingleStore Fully Managed Cluster**
5. AWS ETL Glue with Spark Prep
6. Launch an AWS CloudShell Service

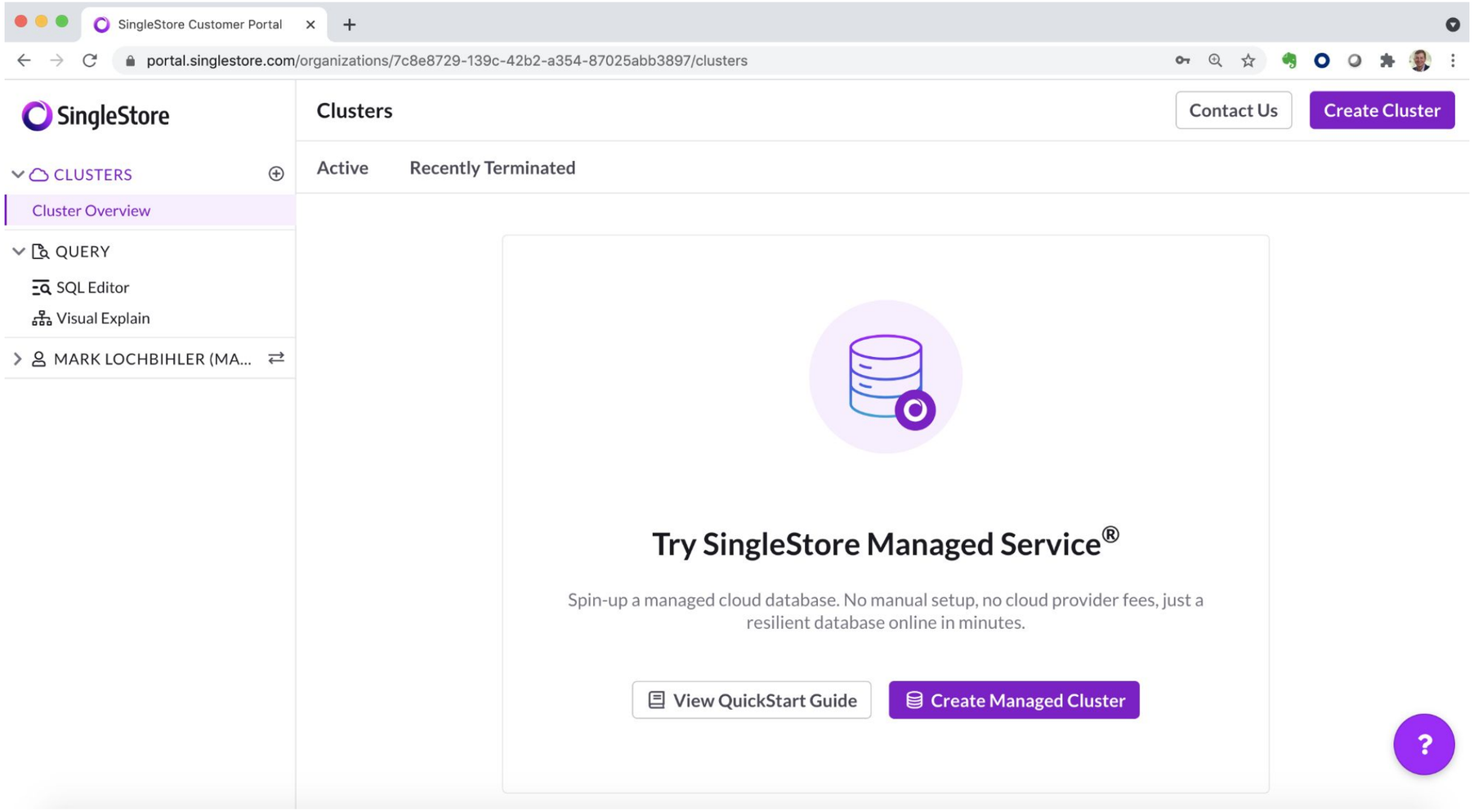
© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

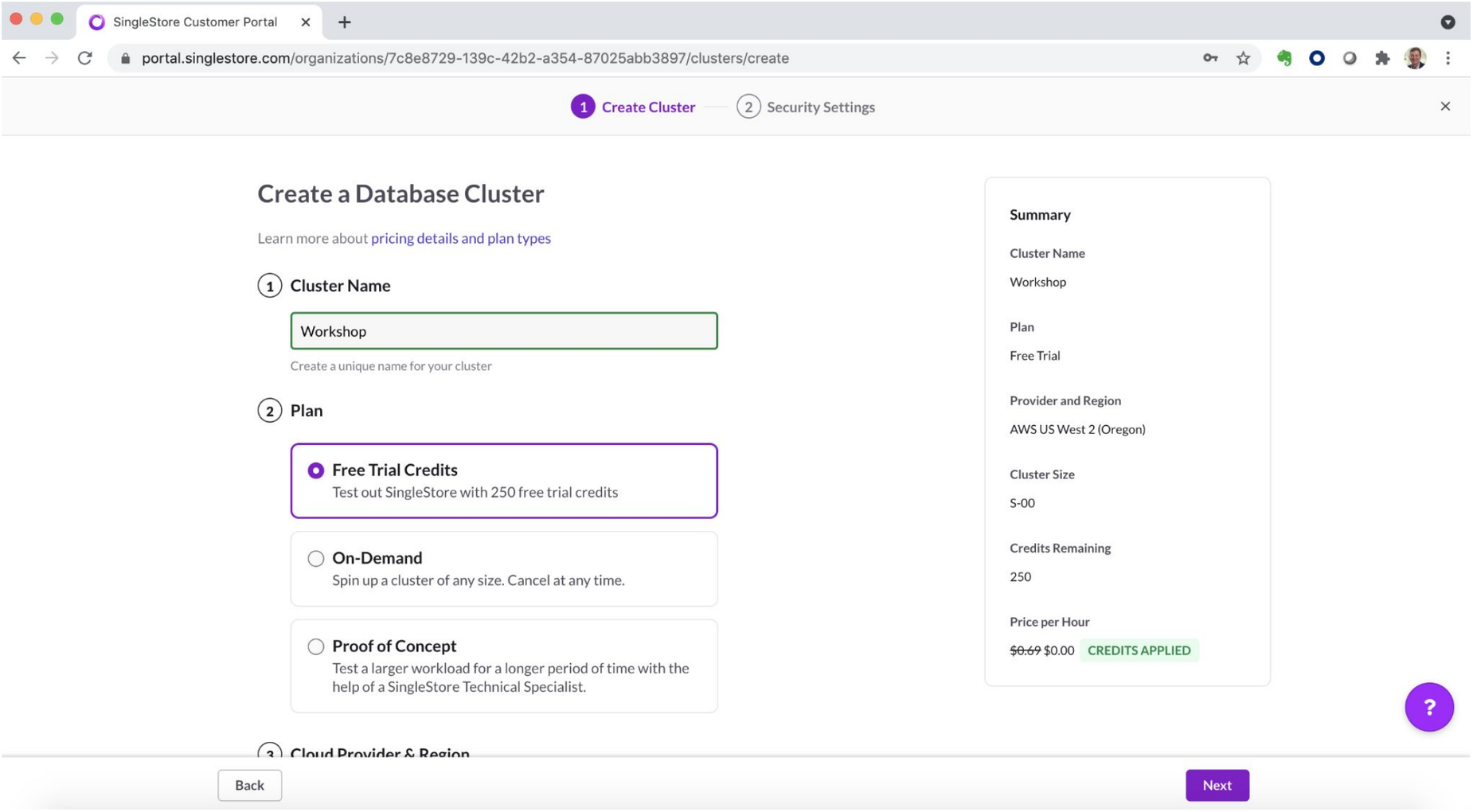


© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

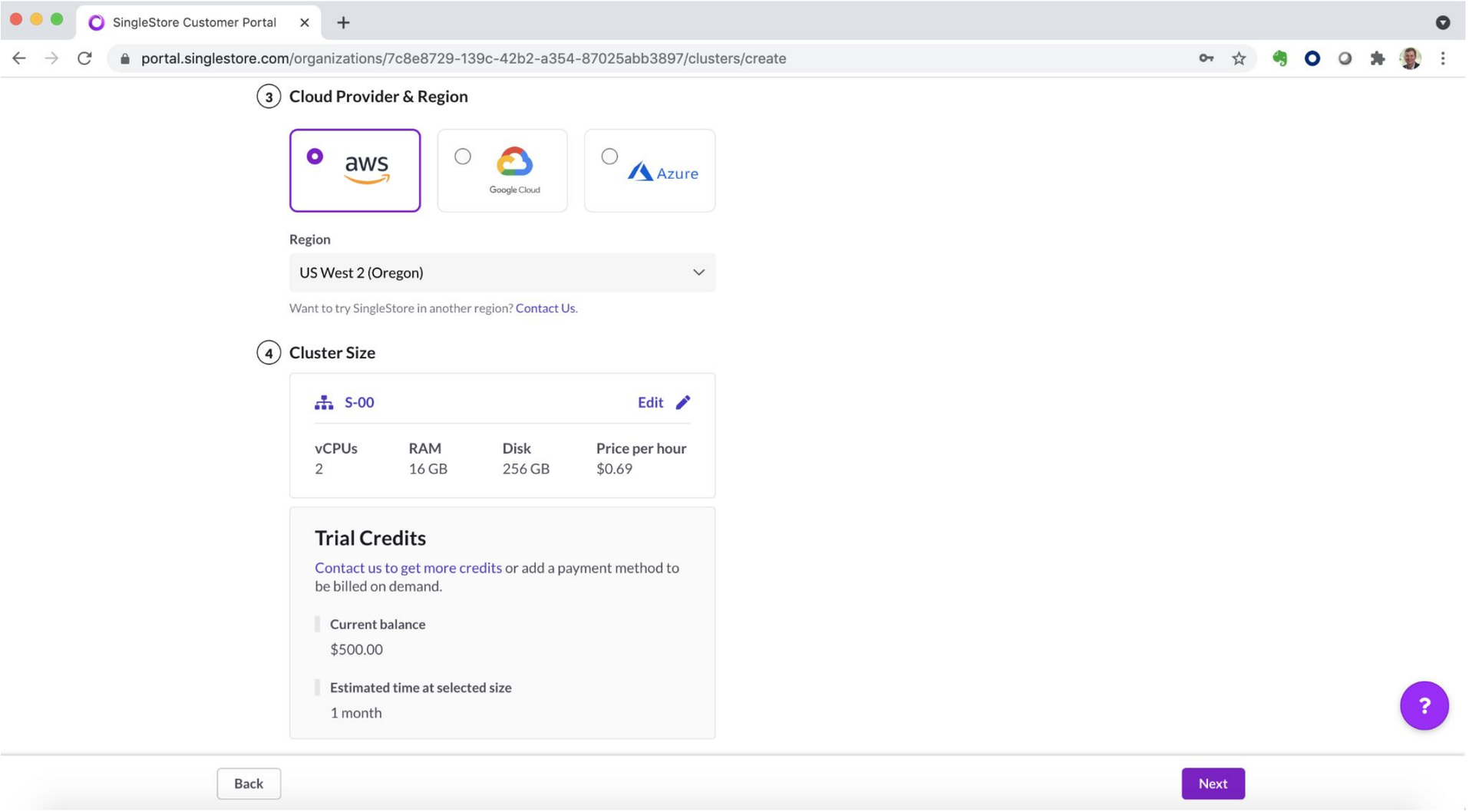


aws | SingleStore Looker

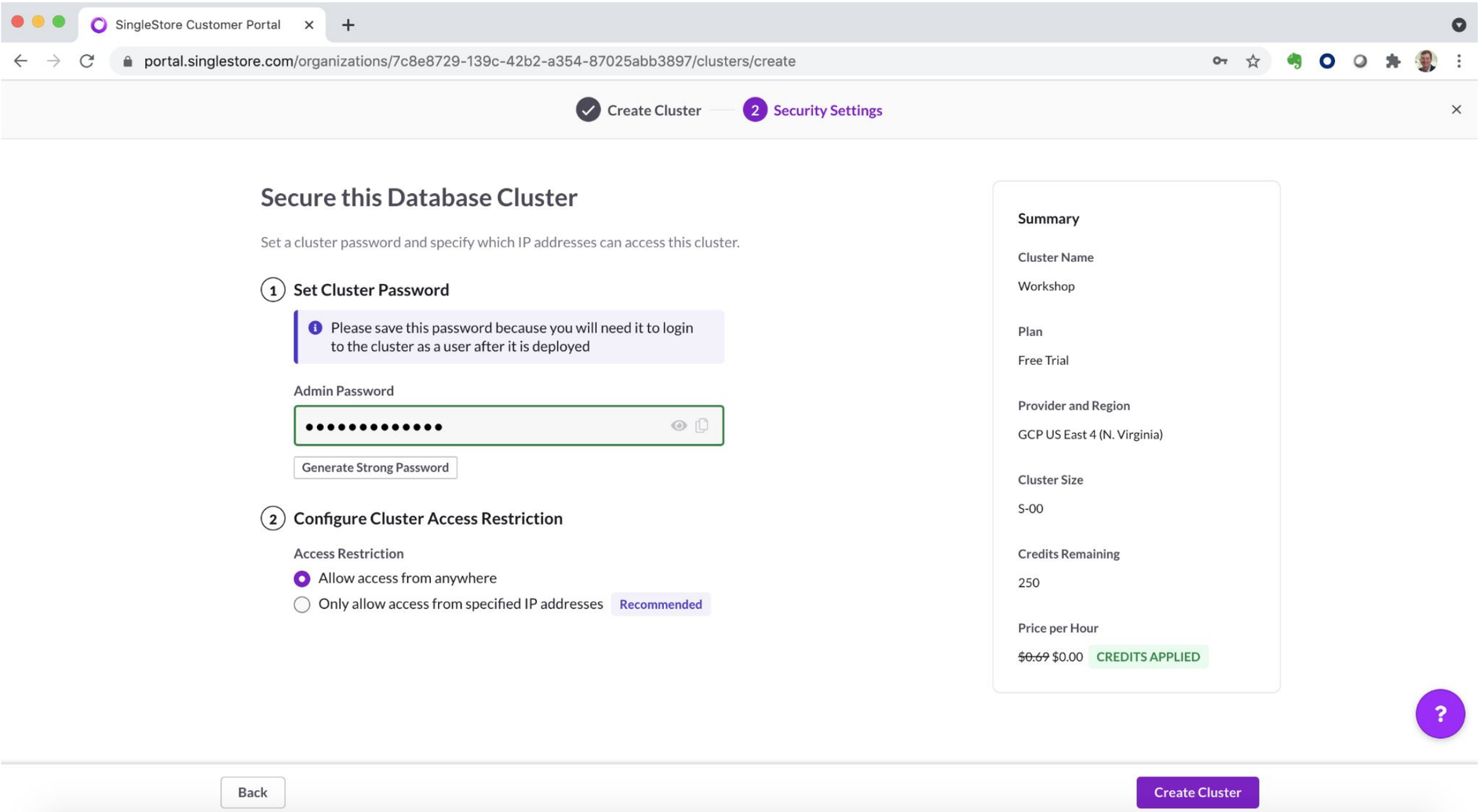
© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SingleStore Customer Portal

portal.singlestore.com/organizations/21827726-255e-4fdb-b47d-1f8b4b63e180/clusters/d5fa3be3-2443-4e2c-bc13-93cedca1abdf

SingleStore

CLUSTERS

Cluster Overview

Workshop

QUERY

SQL Editor

Visual Explain

MARK LOCHBIHLER (MARK LOC...

Workshop +

15% / Deploying Nodes

Clusters > Workshop

Last updated: 9:26:19 AM

Load Data

Connect

Overview | Firewall | Security | Updates | Monitoring

Cluster Properties

Cluster Type
Production Edit

Size
S-00
2 vCPUs | 16 GB Memory | 256 GB Storage
Edit

Region
AWS US East 1 (N. Virginia)

Created
5/27/21 at 9:25:48 AM

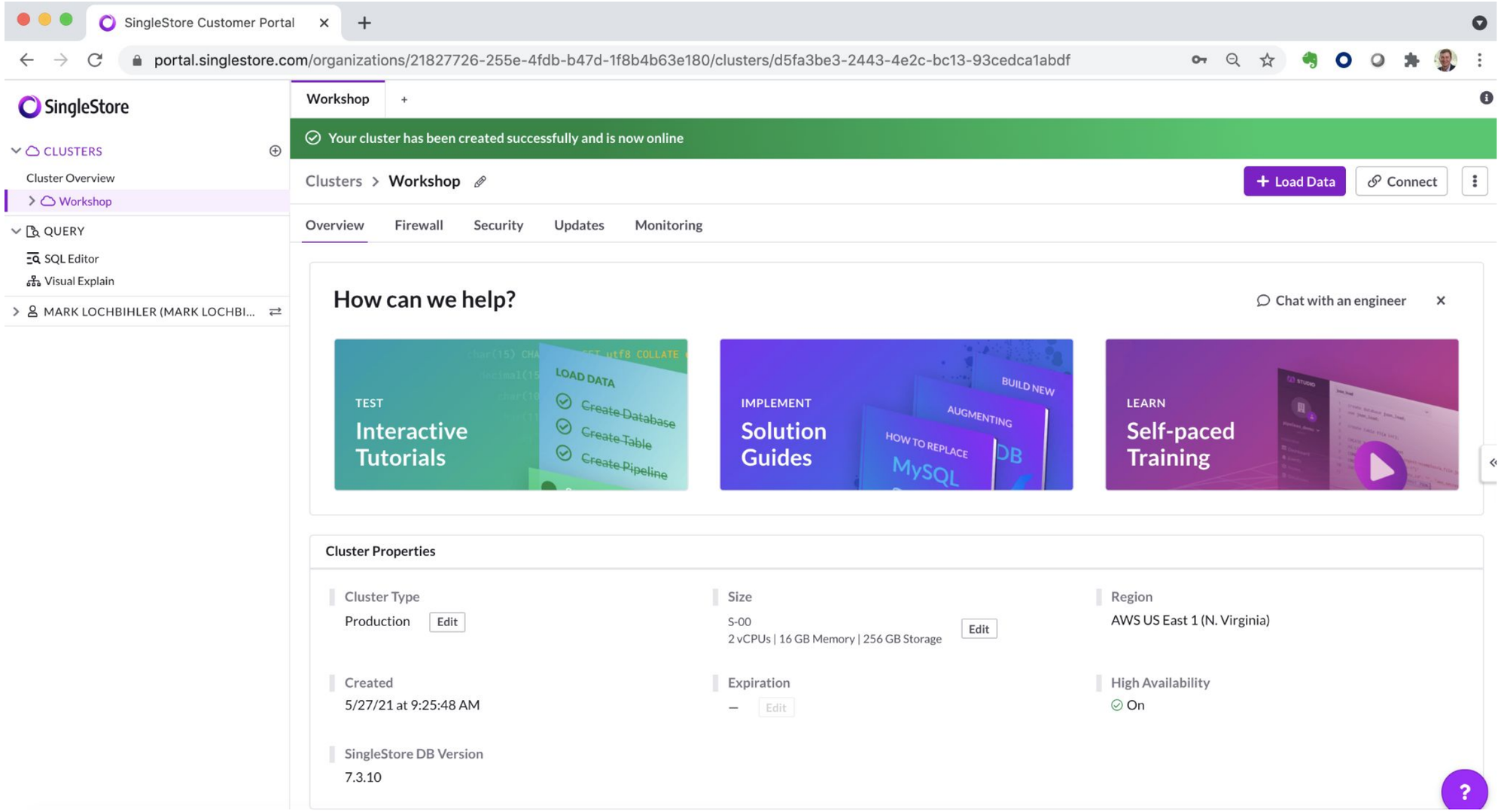
Expiration
— Edit

High Availability
On

SingleStore DB Version
7.3.10

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Connect to Your Cluster

CLI Client | SQL IDE | Your App | BI Tools

MySQL Client

MySQL Command

```
mysql -u admin -h svc-d5fa3be3-2443-4e2c-bc13-93cedca1abdf-ddl.aws-virginia-1.svc.singlestore.com -P 3306 --default-auth=mysql_native_password -p
```

To connect to a database instance using the MySQL client, type the command at a command prompt. Enter the user password when prompted.

TLS/SSL Certification

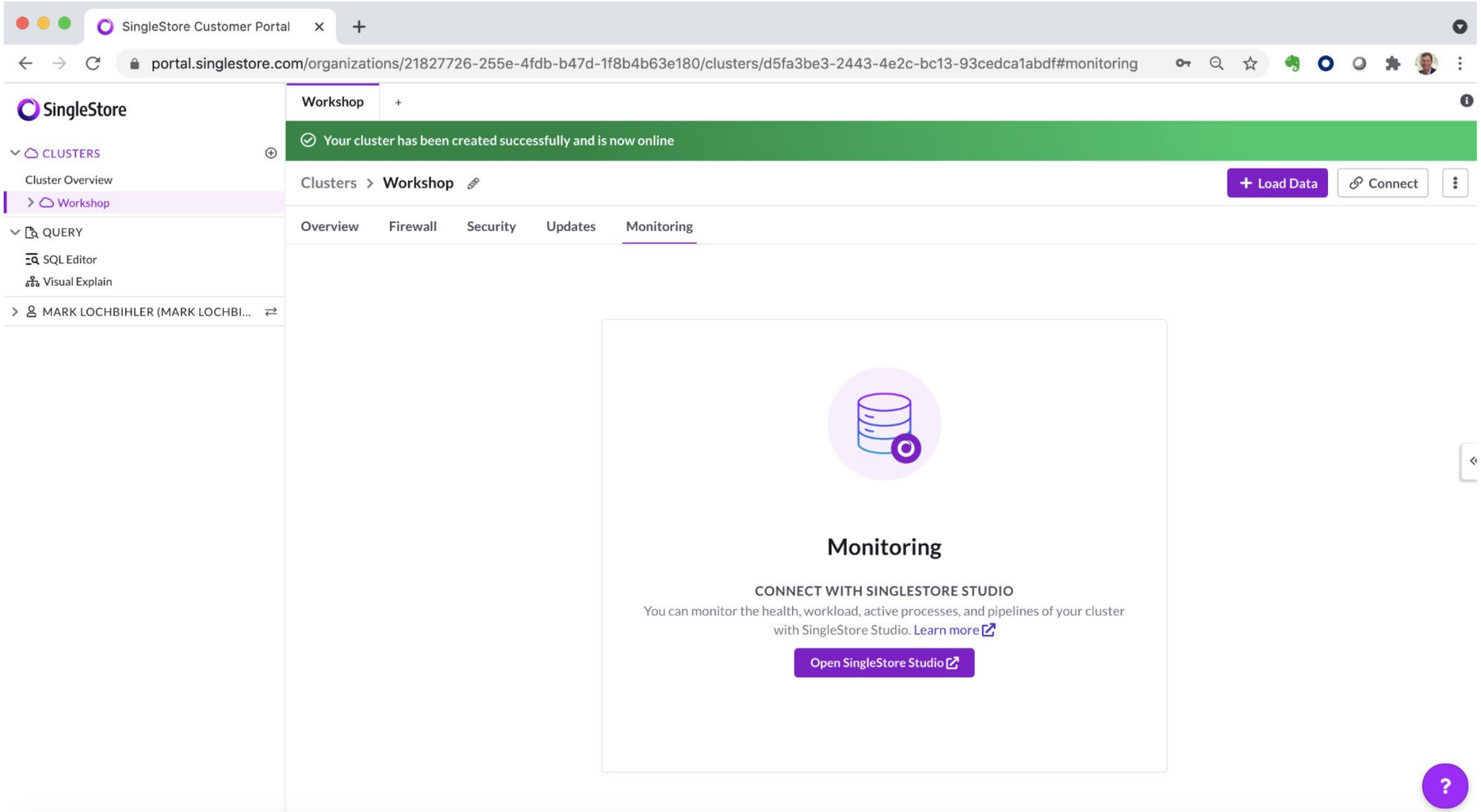
Download

Having trouble connecting?

- Make sure the IP from which you will connect to the cluster has been added to the inbound connection list.
- You can reset your password in your cluster's User Credentials card in case you forgot.
- Don't see the connection you're looking for? Tell us about it.

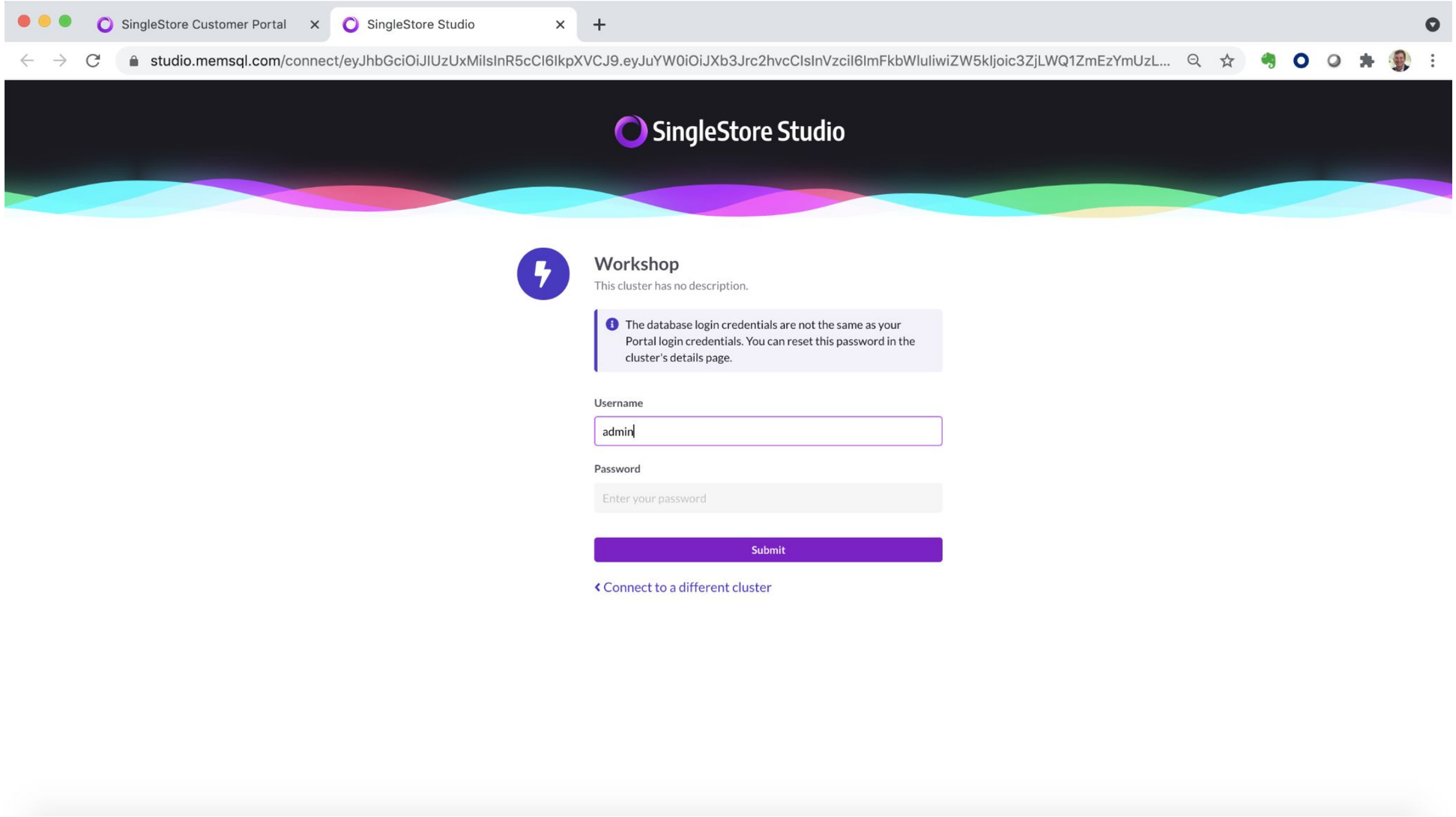
aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

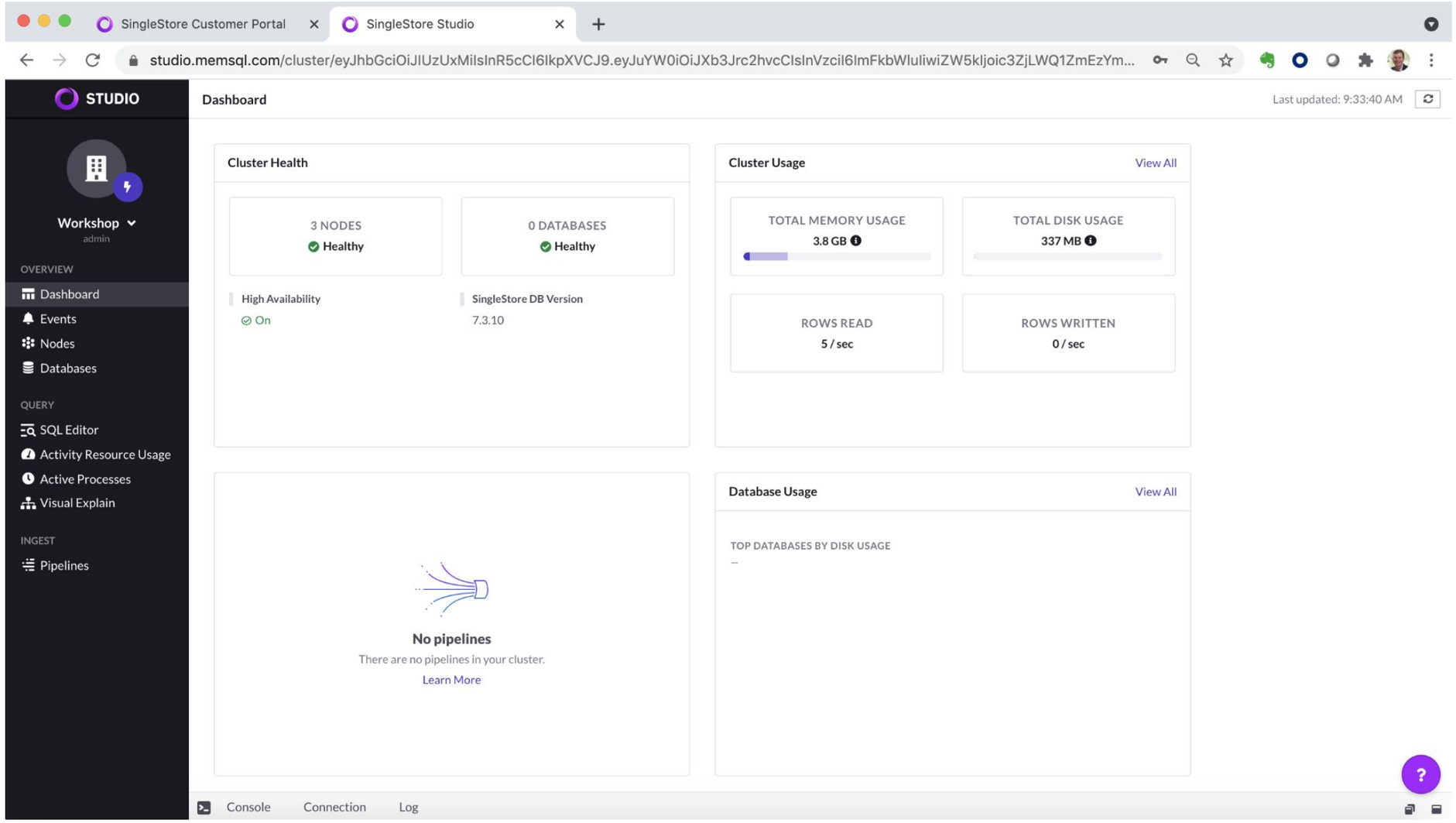


aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

Preparation Steps for Hands on Workshop

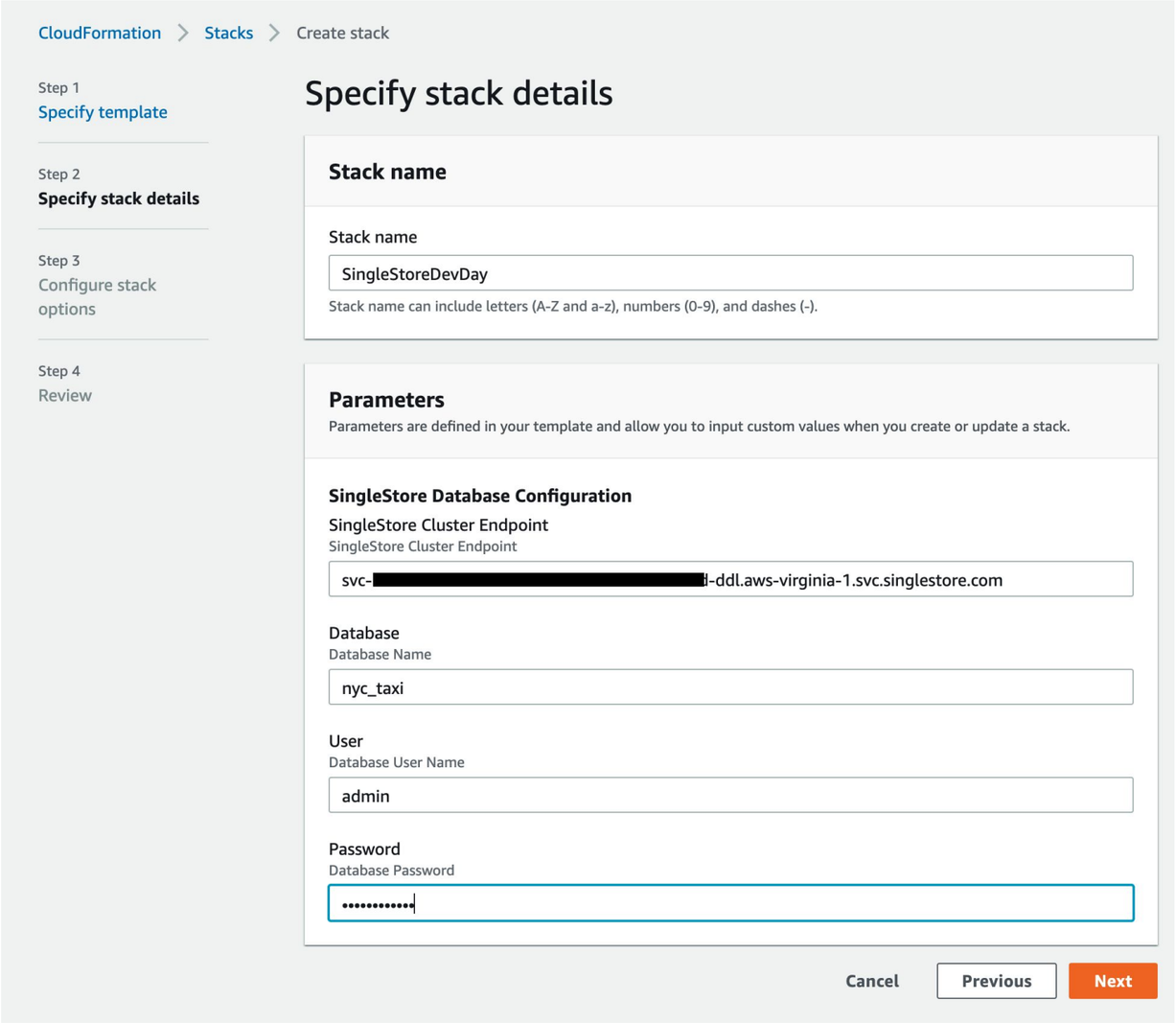
1. Download Student Guide and Presentation Deck
2. Decide Which AWS Account to Use in this Workshop
3. From the AWS Console, login with your root account and perform admin tasks:
    i. Create an AWS IAM user
4. Launch your own SingleStore Fully Managed Cluster
5. **AWS ETL Glue with Spark Prep**
6. Launch an AWS CloudShell Service

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Glue Prep

AWS CloudFormation template

- Creates IAM Role
- Stores credentials in AWS Secrets Manager.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

WORKSHOP PREPARATION

Your AWS account will be access the following Services on the cloud:

AWS S3, Sagemaker, Glue ETL, and CloudShell

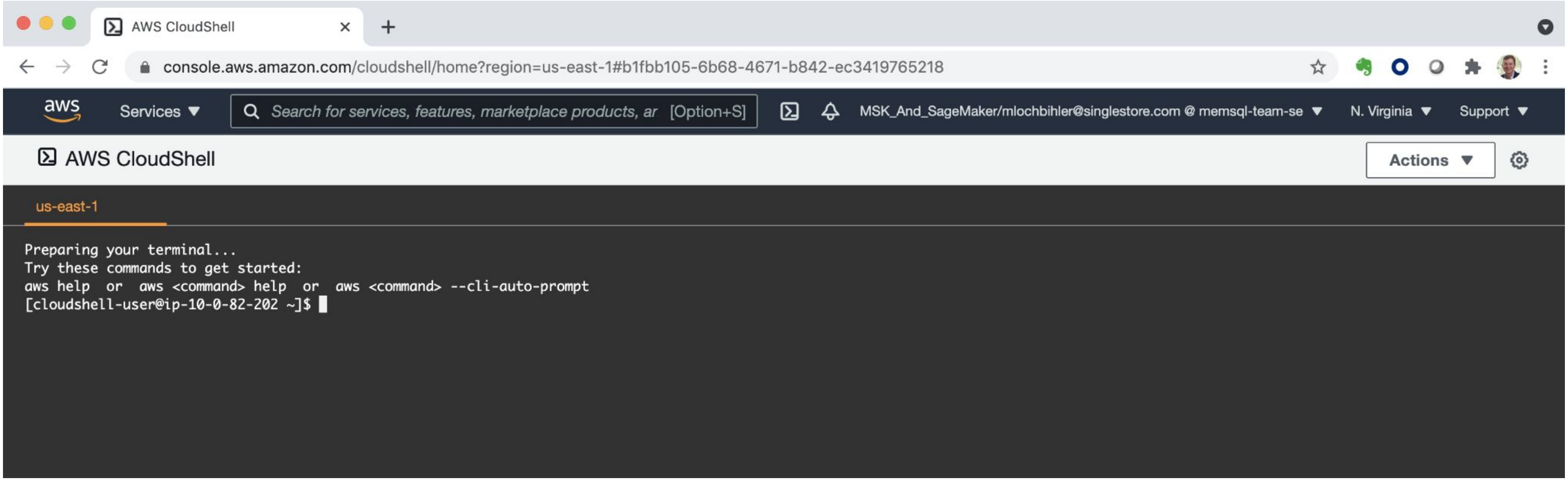
Preparation Steps for Hands on Workshop

1. Download Student Guide and Presentation Deck
2. Decide Which AWS Account to Use in this Workshop
3. From the AWS Console, login with your root account and perform admin tasks:
    i. Create an AWS IAM user with your root AWS account
4. Launch your own SingleStore Fully Managed Cluster
5. AWS ETL Glue with Spark Prep
6. **Launch an AWS CloudShell Service**

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Step 7 : Launch a CloudShell Service

Search for CloudShell, within your AWS Console, then launch the environment by taking the defaults and start the service.



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Today's Agenda (3 Hrs)

- Introduction
- Workshop Preparation
- **Building The Operational Analytics Solution**
  - Step 1 - Setup Database and Load Data
  - Step 2 - Deploy a Real-Time Analytics Stored Procedure
  - Step 3 - AWS Glue ETL leveraging SingleStore Spark Connecter
  - Step 4 - Looker Visualization
- Cleanup AWS Resources

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Building the Operational Analytics Solution

Section B - Student Guide

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS, SingleStore and Looker - Operational Analytics

Walking through the Build

Describe Development Process to Create a Real Time Operational Dashboard

1) Setup Database and Load Data
2) Deploy a Real Time Analytics Stored Procedure
3) Leverage AWS Glue ETL with SingleStore Spark Connector
4) Bring it all together with an Operational Dashboard

Where can you get all the details?

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Section 2 - Student Guide
BUILD PHASE

1) Setup Database and Load Data

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New York City Taxi and Limousine Commission (NYC TLC) yellow taxi trip dataset [1]

Trip Event Sample:

30,requested,POINT(40.754100 -73.980000),POINT(40.755300 -73.962600),1483984668,1,0,0,0,1,9.299954999999471,-1

Driver Event Sample:

21,Ella,Harris,driving,POINT(40.73281182 -73.99684899),POINT(40.73950010 -74.00639998),NULL,1783796

Sample Event Neighborhood:

281 Jamaica Queens POLYGON((40.70537600 -73.77582000, 40.70429051 -73.77592224, 40.70272444 -73.77547673, 40.69983099 -73.77395775, 40.69822334 -73.77734617, 40.69976700 -73.77854600, 40.69739000 -73.78446600, 40.69599500 -73.78277100, 40.69523700 -73.78211000, 40.68836400 -73.77662000, 40.68528000 -73.77240000, 40.68377800 -73.76974000, 40.68432922 -73.76963696, 40.68644322 -73.76116940, 40.68337600 -73.76042400, 40.68284600 -73.76073100, 40.67905300 -73.76183200, 40.67750200 -73.76289000, 40.67606842 -73.76346750, 40.67600300 -73.76349400, 40.67364000 -73.76383000, 40.67185300 -73.76435300, 40.66703500 -73.76696000, 40.66744600 -73.77113700, 40.66743600 -73.77295100, 40.66673000 -73.78367300, 40.66678400 -73.78511800, 40.66632163 -73.78941799, 40.66631400 -73.78948900, 40.66718129 -73.78930714, 40.67301000 -73.78808500, 40.67339100 -73.79047400, 40.67434400 -73.80121600, 40.67768100 -73.80323000, 40.68291367 -73.80583205, 40.68617474 -73.80745368, 40.68857916 -73.80870357, 40.69793100 -73.81414400, 40.70404100 -73.81702000, 40.70440200 -73.81443900, 40.70771600 -73.80278700, 40.70844600 -73.80022200, 40.70978288 -73.79553437, 40.71050718 -73.79353164, 40.71071100 -73.79304000, 40.71203000 -73.78932100, 40.71186500 -73.78771800, 40.71291943 -73.78271314, 40.71293800 -73.78262500, 40.70705100 -73.77997500, 40.70782300 -73.77678100, 40.70537600 -73.77582000))

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Create and Start SingleStore Pipelines

Driver Event Sample:

public-kafka.memcompute.com/drivers

Trip Event Sample:

public-kafka.memcompute.com/trips

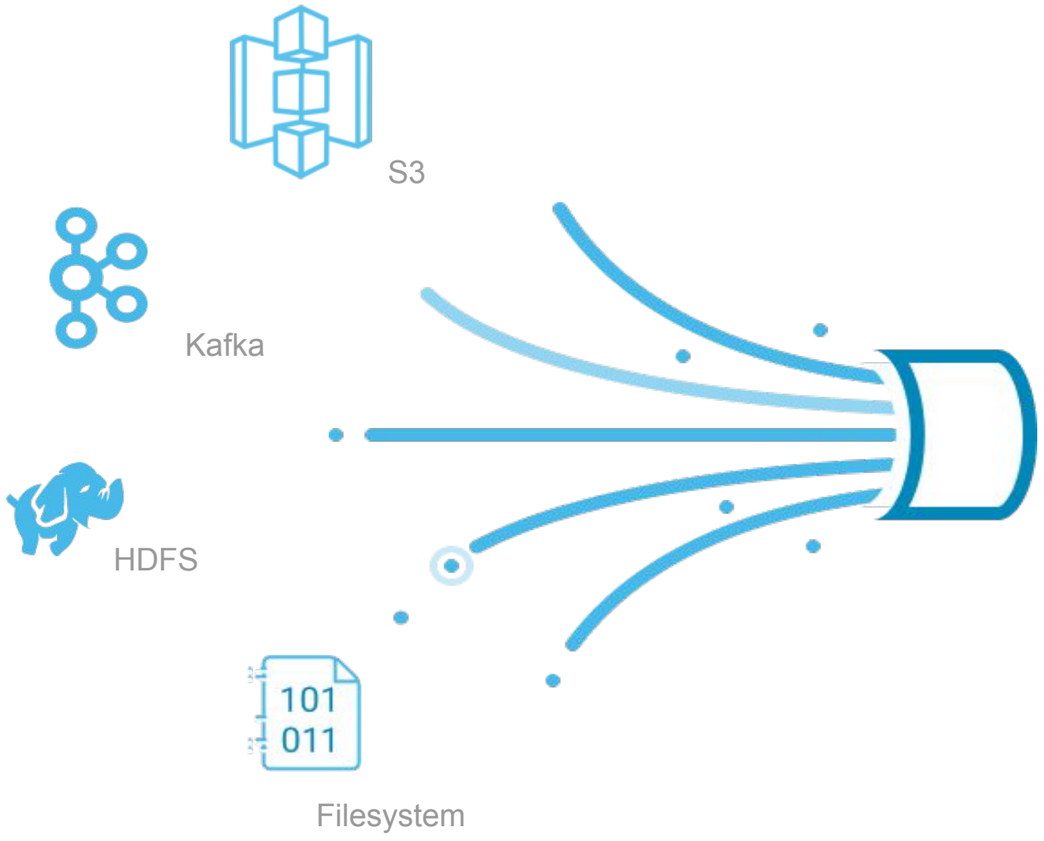
Sample Event Neighborhood:

memsql-datafeeds/nyc-taxi-data/

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use SingleStore Pipelines for Scoring on Load

- SingleStore Pipelines
  - Automatic streaming loader
- Transform types
  - Python
  - Executable
- Use
  - Compute new "score" column from other columns during load
- Benefits
  - Use with any external scoring code or libraries



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

portal.singlestore.com/organizations/21827726-255e-4fdb-b47d-1f8b4b63e180/editor?connectionID=750788d6-cb22-443d-864f-e3f91a6acad3

SingleStore

CLUSTERS

Cluster Overview

AWS Dev Day

QUERY

SQL Editor

Visual Explain

MARK LOCHBIHLER (MARK LOCH...

SQL Editor +

AWS Dev Day • Select Database

+ Ingest UI + Load Data Run ⌘↵

```
CREATE PIPELINE neighborhoods
AS LOAD DATA S3 'memsql-datafeeds/nyc-taxi-data/'
config '{"region":"us-east-1"}'
REPLACE
INTO TABLE neighborhoods
FIELDS TERMINATED BY '\t' ENCLOSED BY '' ESCAPED BY '\\'
LINES TERMINATED BY '\n' STARTING BY '';

USE nyc_taxi;

ALTER PIPELINE drivers SET OFFSETS EARLIEST;
```

Message Logs

Select one or more queries and hit ⌘ + Return to run them.

Results are limited to 300 rows.

aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Section 2 - Student Guide
BUILD PHASE

2) Deploy a Real-Time Analytics Stored Procedure

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Leveraging In Database Scoring With Extensibility

Support for:

- User-defined functions (UDFs)
- Stored procedures (SPs)
- User-defined aggregates (UDAFs)
- Arrays
- Records
- Control structures

Implement PMML or other models for in-DB scoring

Benefits

- Combine with other SQL (joins, filters, etc.)
- Use with SQL-compatible tools
- Compiled code and scale-out performance

aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

portal.singlestore.com/organizations/21827726-255e-4fdb-b47d-1f8b4b63e180/editor?connectionID=750788d6-cb22-443d-864f-e3f91a6acad3

SingleStore

CLUSTERS

Cluster Overview

AWS Dev Day

QUERY

SQL Editor

Visual Explain

MARK LOCHBIHLER (MARK LOCH...

SQL Editor +

AWS Dev Day • Select Database

+ Ingest UI + Load Data Run ⌘↵

```
delimiter //
CREATE or replace PROCEDURE `derive_stats` (i int)
AS
declare x int;
begin

DELETE FROM avgcost;
DELETE FROM avgdist;
DELETE FROM avgriders;
DELETE FROM avgridetime;
DELETE FROM avgwait_driver;
DELETE FROM avgwait_passenger;

INSERT INTO avgcost VALUES (0);
INSERT INTO avgdist VALUES (0);
INSERT INTO avgriders VALUES (0);
INSERT INTO avgridetime VALUES (0);
INSERT INTO avgwait_driver VALUES (0);
INSERT INTO avgwait_passenger VALUES (0);

while True loop
```

Message Logs

aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

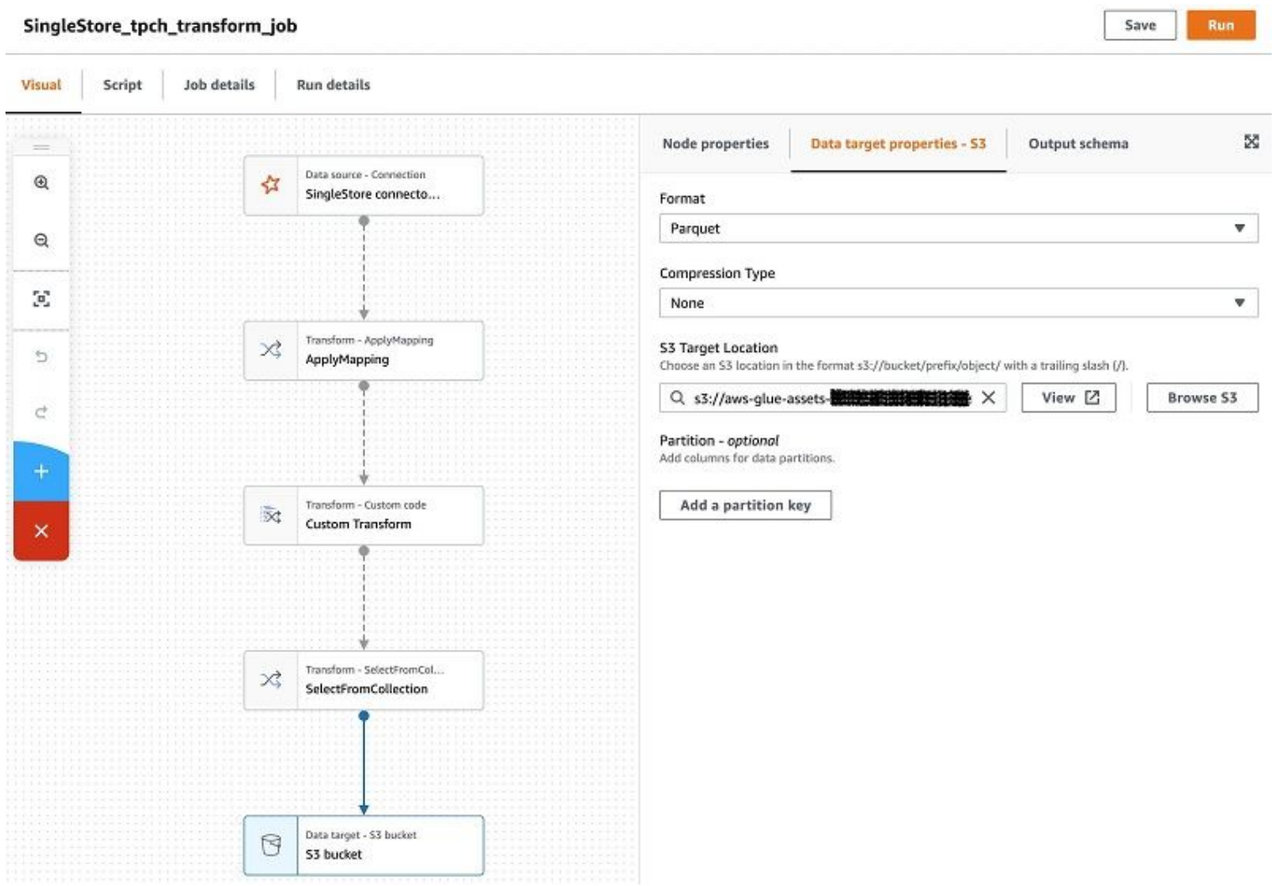
Section 2 - Student Guide
BUILD PHASE

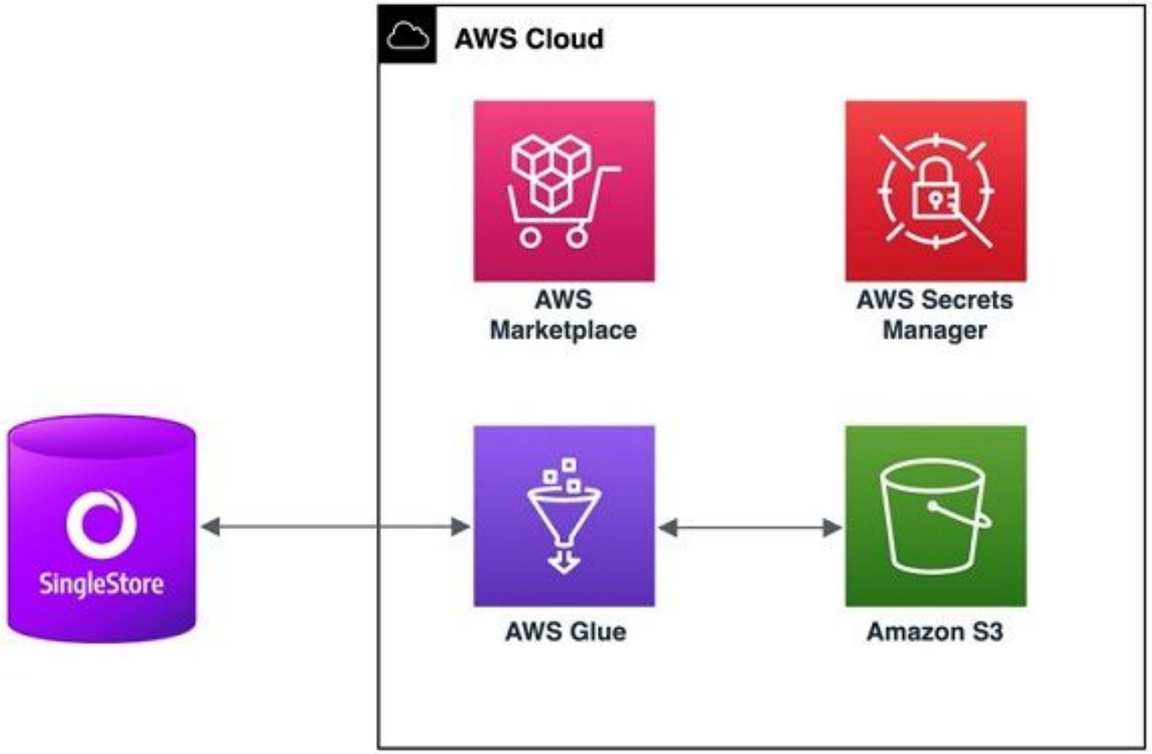
3) Leverage AWS Glue ETL with SingleStore Spark Connector

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Building Fast ETL using SingleStore and AWS Glue





https://aws.amazon.com/blogs/big-data/building-fast-etl-using-singlestore-and-aws-glue/

Sample NYC Taxi Average Ridership dataset:

```
07/01/14 0:00, 10844
07/01/14 0:30, 8127
07/01/14 1:00, 6210
07/01/14 1:30, 4656
07/01/14 2:00, 3820
07/01/14 2:30, 2873
...
...
08/01/14 10:00, 8156
```

Reference: Numenta Anomaly Benchmark (NAB) NYC Taxi dataset [2]

ML "Transform" output

```
TIMESTAMP, NUMRIDERS, ANOMALY_SCORE
07/01/14 0:00, 10844, 1.0229578266
07/01/14 0:30, 8127, 0.8665221552
07/01/14 1:00, 6210, 0.8211380899
07/01/14 1:30, 4656, 0.7894662312
07/01/14 2:00, 3820, 0.7285065354
07/01/14 2:30, 2873, 0.8211380899
...
...
08/01/14 10:00, 8156, 0.7285065354
```

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws marketplace

Hello, assumed-role/Adm

Categories | Delivery Methods | Solutions | AWS IQ | Resources | Your Saved List | Partners | Sell in AWS Marketplace | Amazon Web Services Hor

All Products > Analytic Platforms > SingleStore connector for AWS Glue

SingleStore connector for AWS Glue

By: SingleStore Latest Version: 3.0.4

The SingleStore connector for AWS Glue allows you to connect to data in AWS Glue

Linux/Unix ☆☆☆☆☆ 0 AWS reviews | 61 external reviews

Continue to Subscribe

Save to List

Overview | Pricing | Usage | Support | Reviews

Product Overview

The SingleStore connector for AWS Glue allows you to connect to data in AWS Glue and makes it easy to prepare and load data for any workload.

Version	3.0.4
By	SingleStore
Categories	Analytic Platforms Business Intelligence & Advanced Analytics
Operating System	Linux
Delivery Methods	Container

Highlights

- Simple Integration between AWS Glue and SingleStore

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws Services ▼ Search for services, features, marketplace products, and docs [Option+S]

"SingleStore connector for AWS Glue" connector is selected.
To begin using your connector in AWS Glue Studio, you must create a connection using the below form or choose "Activate connec

AWS Glue Studio > Connectors > Create connection

Create connection Info

Connection properties Info

Connector

SingleStore connector for AWS Glue
The SingleStore connector for AWS Glue allows you to connect to data in AWS Glue

Name
Enter a unique name for your connection.

singlestore_devday_connection

Names can contain letters (A-Z), numbers (0-9), hyphens (-), or underscores (_), and must be less than 256 characters long.

Description - *optional*

Descriptions can be up to 2048 characters long.

Connection access Info

AWS Secret - *optional* Info
Choose a secret from **AWS Secrets Manager.** AWS secrets eliminate hardcoding sensitive information.

aws-glue-singlestore-devday-01

Network options - *optional*

If your AWS Glue job needs to run on **Amazon Elastic Compute Cloud** (EC2) instances in a virtual private cloud (VPC) subnet, you must provide additional VPC-specific configuration information.

Cancel | Activate connector only | Create connection and activate connector

AWS Glue Studio > Connectors

Connectors Info

Marketplace connectors

Subscribe to connectors from AWS partners to expand your data sources.

Go to AWS Marketplace

Custom connectors

Provide your own connector to expand your data sources. Creating custom connectors

Create custom connector

Your connectors (1) Info

Actions ▼ | Create connection

You can manage your connectors or use them to create connections.

Find connectors

1

Name	Type	Last modified
SingleStore connector for AWS Glue	MARKETPLACE	2021-06-11 14:42:43.279000

Your connections (1) Info

Actions ▼ | Create job

You can manage your connections or use a connection in a job.

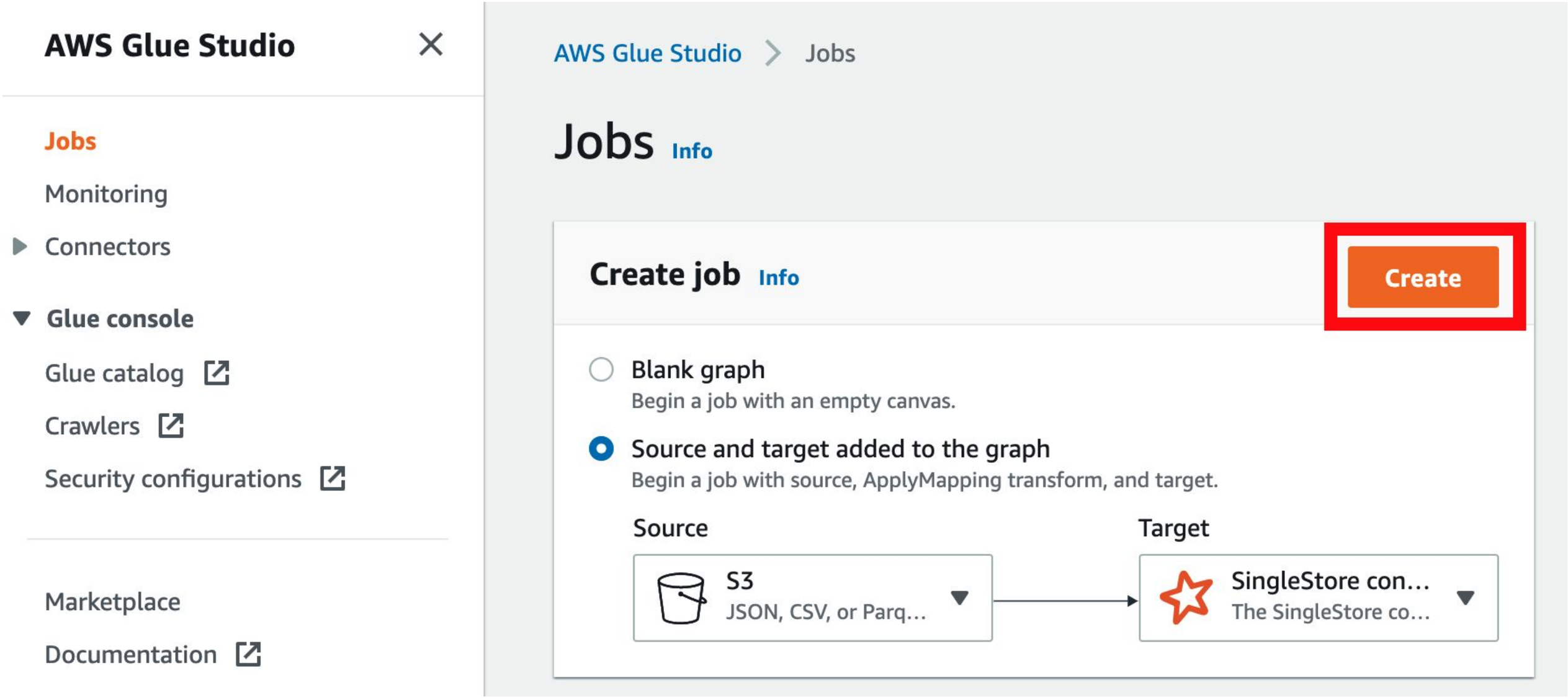
Find connections

1

Name	Type	Last modified
singlestore_devday_connection	MARKETPLACE	2021-06-11 14:42:43.453000

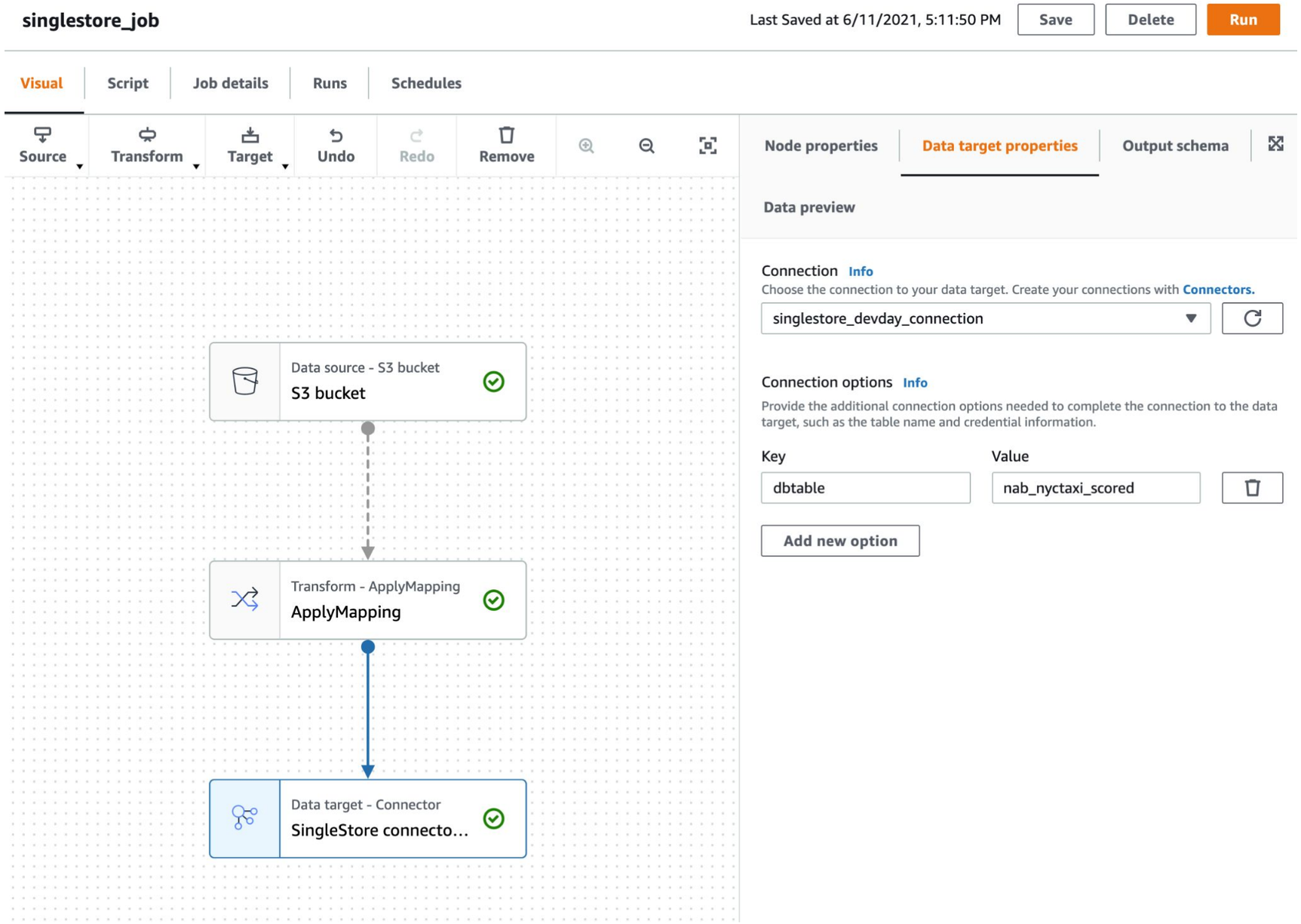
aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



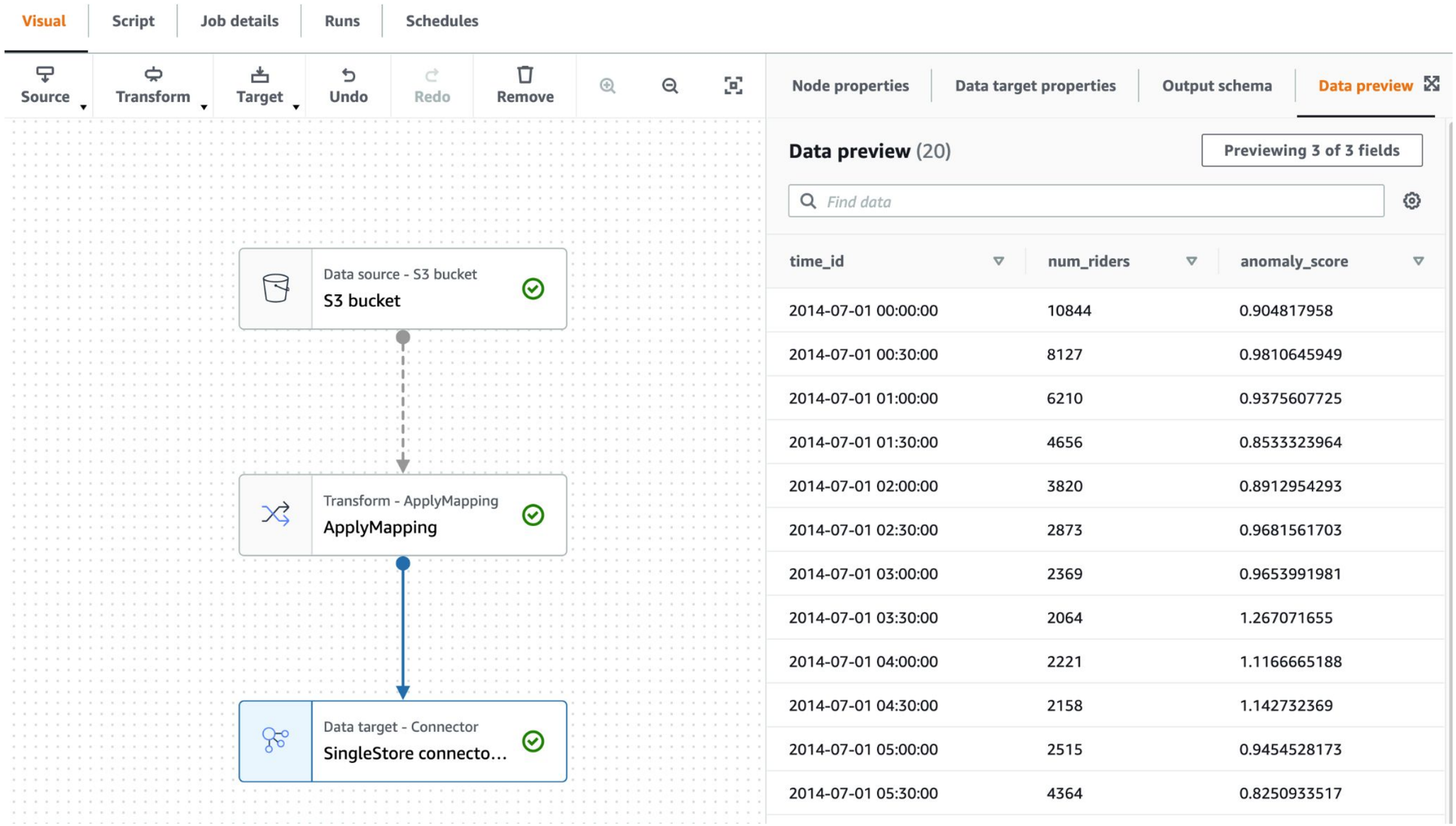
aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Section 2 - Student Guide
BUILD PHASE

4) Bring it all together with a Looker Dashboard

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Visualizing Your Data With Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Looker: The Platform for Data Experiences

Google Cloud's cloud-native Enterprise BI Platform enabling **secure access** to **near real-time data when and where** you need it.

Agility

- Respond quickly to change
- Make better decisions based on fresh data
- Build portable, repeatable data processes
- Leverage database technology, support multi-cloud

Efficiency

- Simplify and accelerate data pipelines
- Save 70% of analyst time
- Collaborate on data model development and iteration
- Reduce costs and improve performance

Governance/Control

- Unify your organization with centralized, auditable business logic
- Reduce risk, yet provide self-service access to data
- Control where and how data is used

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proprietary + Confidential

Built to be different

In-database architecture

No additional data movement or storage

Supports more than 50 database dialects

Leverage native database functionality

Semantic modeling layer

Define standard business logic across the business

Robust data permissions (row level, column level, content)

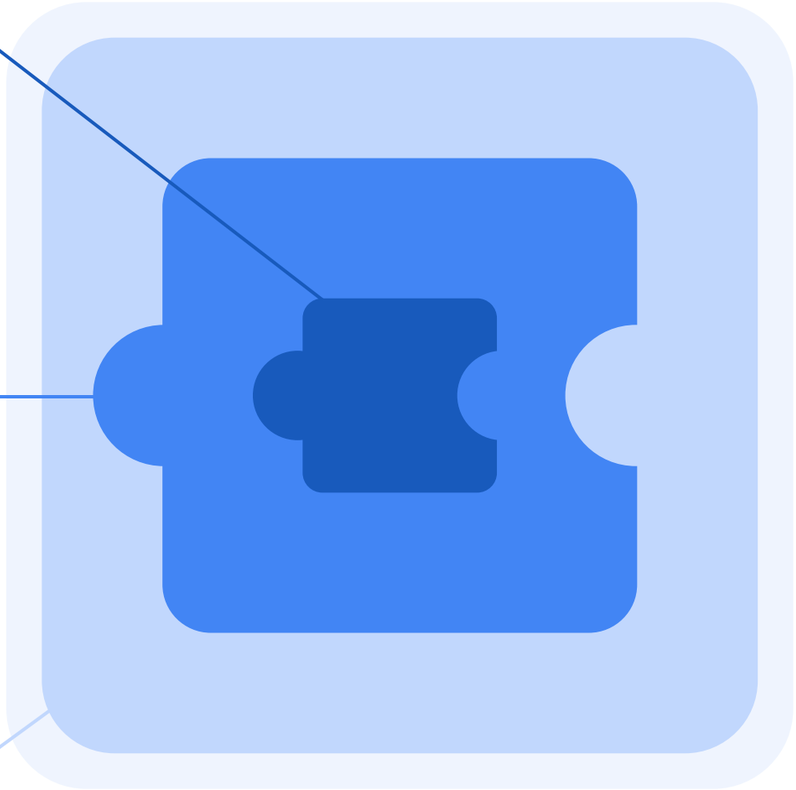
Collaborative Git-integrated IDE with version control

Cloud native and extendable

100% web-based

API extensibility

Unlimited data experiences



Built on the cloud strategy of your choice

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proprietary + Confidential

Serve up real-time, reports and dashboards that inspire more in-depth analysis.

Modern BI & Analytics

Infuse relevant data into your existing tools for an enhanced experience, and more effective results.

Integrated Insights

Super-charge operational workflows with complete, near-real time data.

Data-driven Workflows

Provide a data tool built to deliver the results you need, as needed.

Custom Applications

Looker

Governed metrics | Best-in-class APIs | In-database | Git version-control | Security | Cloud

SQL In ↓ ↑ Results Back

SingleStore

INTERCOM · eloqua · shopify · mixpanel · Adobe · servicenow

G Suite · Magento · ADP · Google Ads · HubSpot · Square · zendesk · LinkedIn · Jira Software

aws | SingleStore · Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Looker and SingleStore Integration

In-Database Architecture

The power of SingleStore is directly leveraged by Looker because all transformation is done in-database

- As real-time as data in SingleStore
- Leverages in-memory performance

Warehouse Exploration

All data in SingleStore can be made available to users in a customized and user-friendly way

- Data can be explored as real time as the data in SingleStore
- Reporting based on a single source of truth

Advanced Querying

Advanced SingleStore functionality can be directly leveraged in the LookML modeling layer including:

- Geospatial functions
- JSON Querying

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Visualizing your NYC Taxi Data with Looker

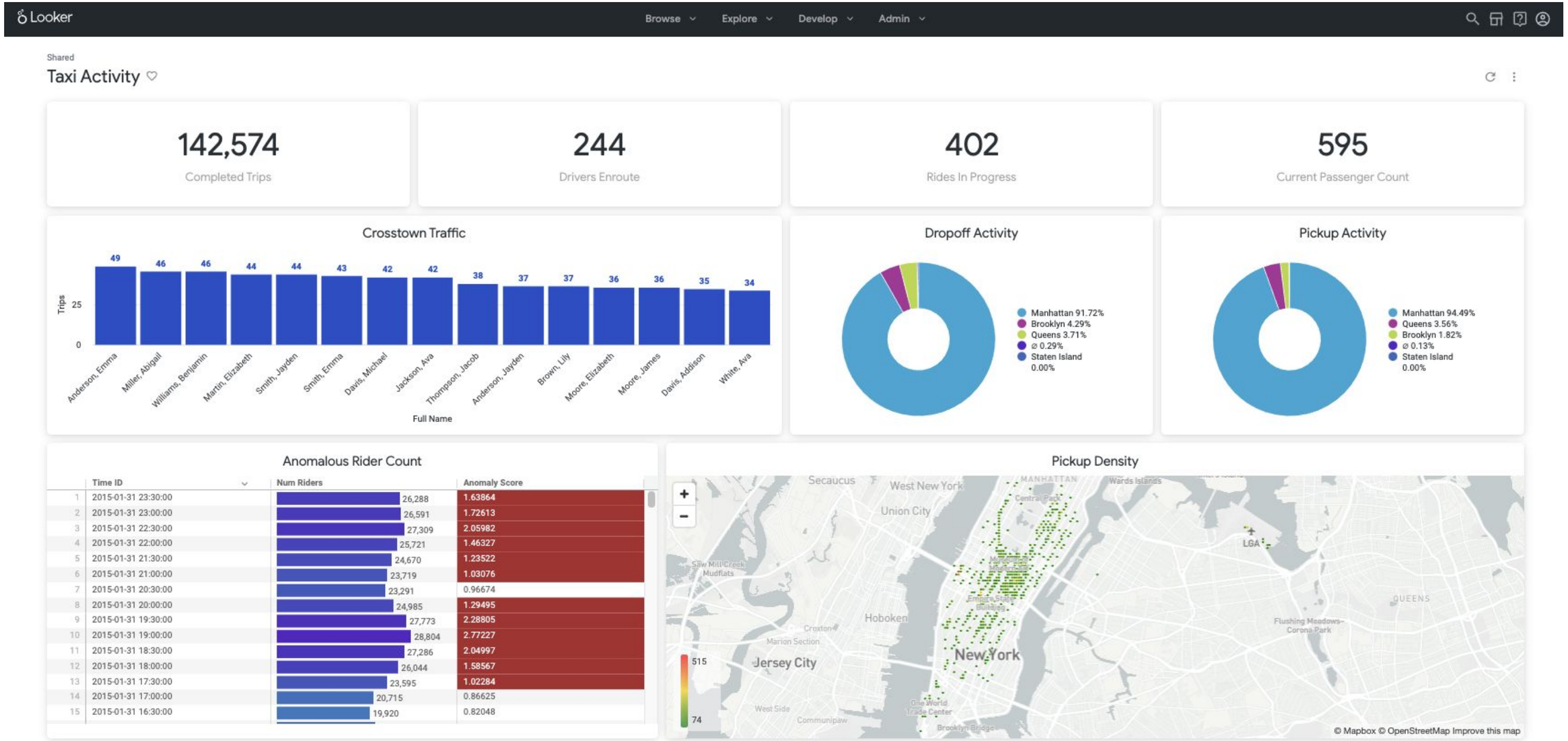
Agenda

- Anatomy of a Looker Project
- Create a connection from Looker to SingleStore
- Generate a default LookML Model
- Enhance the Model File
- Add business-specific dimensions and measures
- Create a visualization and save it to a dashboard
- Customize dashboard properties

aws | SingleStore Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

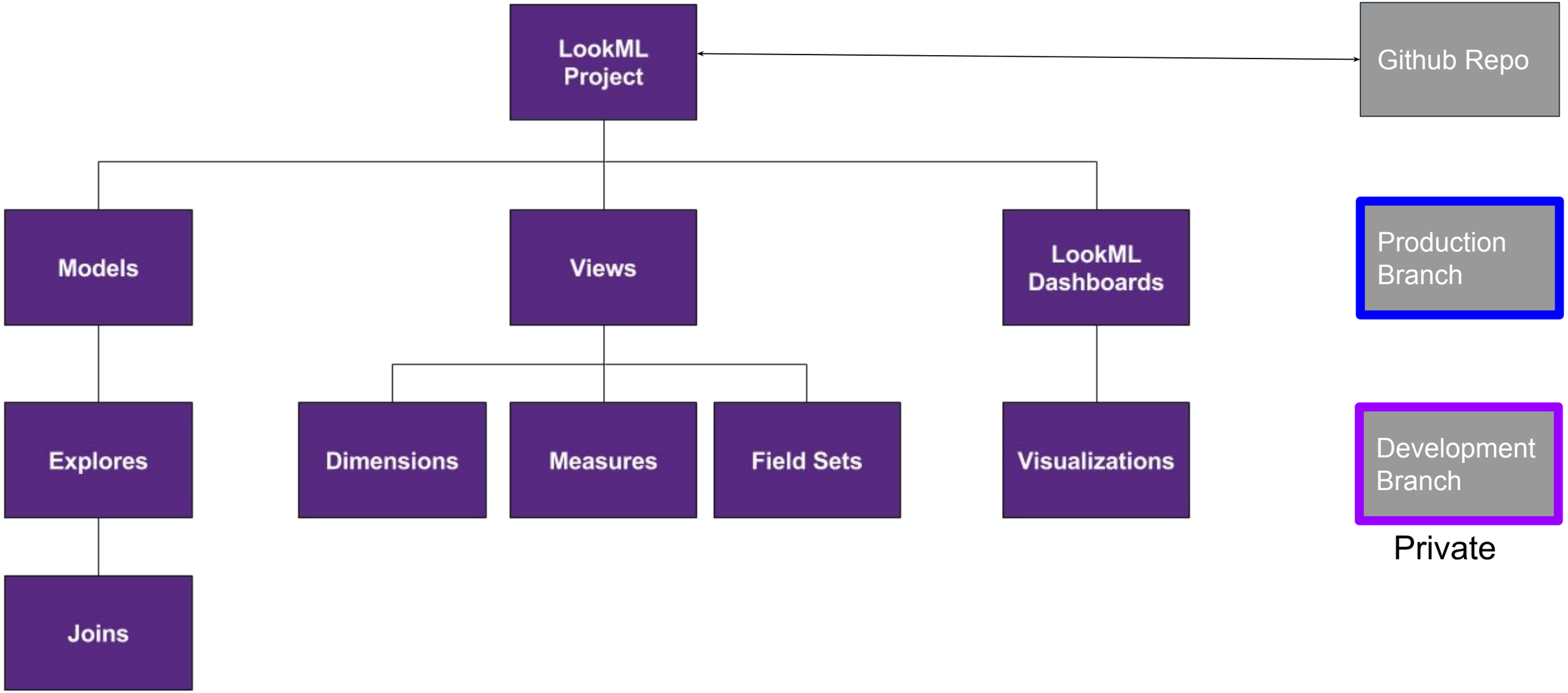
We'll Build This Dashboard



aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Anatomy of a LookML Project



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HANDS ON WORKSHOP

Today's Agenda (3 Hrs)

- **Introduction**
- **Workshop Preparation**
- **Building The Operational Analytics Solution**
  - Step 1 - Setup Database and Load Data
  - Step 2 - Deploy a Real-Time Analytics Stored Procedure
  - Step 3 - AWS Glue ETL leveraging SingleStore Spark Connecter
  - Step 4 - Looker Visualization
- **Cleanup AWS Resources**

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Recap and Cleanup AWS resources

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Hands On Session

SingleStore AWS Looker DevDay Github

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Thank you!

Sugandan Barathy
sbarathy@singlestore.com

Jonathan Dungay
dungay@google.com

Saurabh Shanbhag
saurshan@amazon.com

Sebastian Fabri
sebastienfabri@google.com

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

A Survey will be emailed to you! Please take the time to provide us with feedback!

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws PARTNER DEV DAY

Student Guide

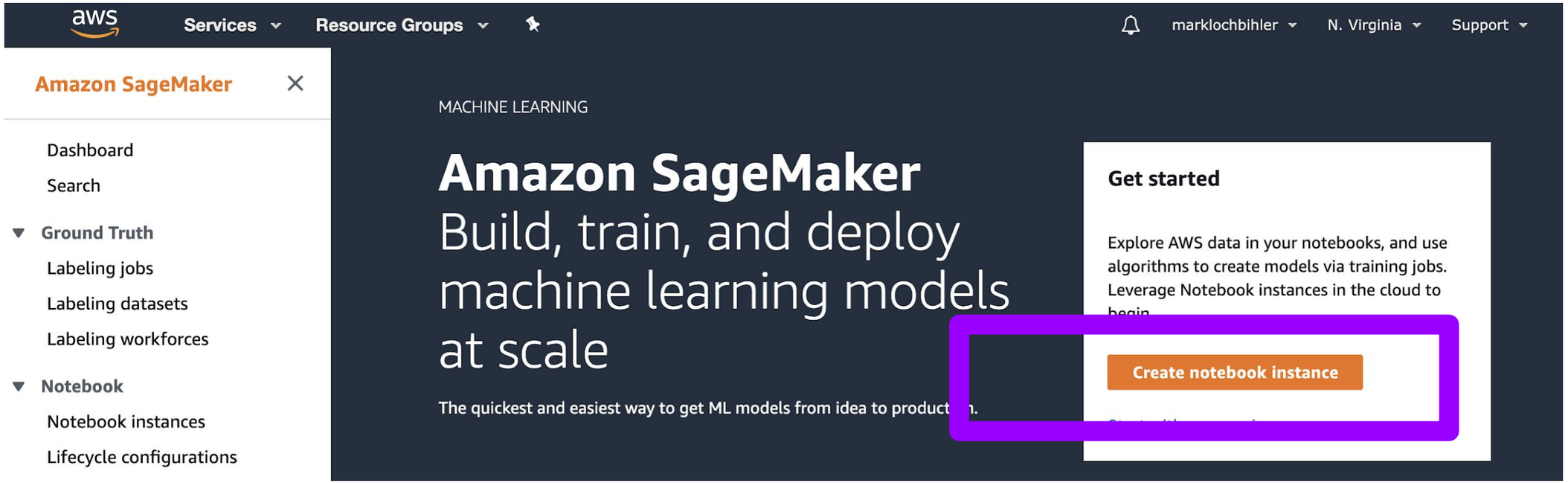
APPENDIX

aws | SingleStore looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session

Search for SageMaker, within your AWS Console and then hit “Create notebook instance”



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

Update the following template parameter

Notebook instance settings:

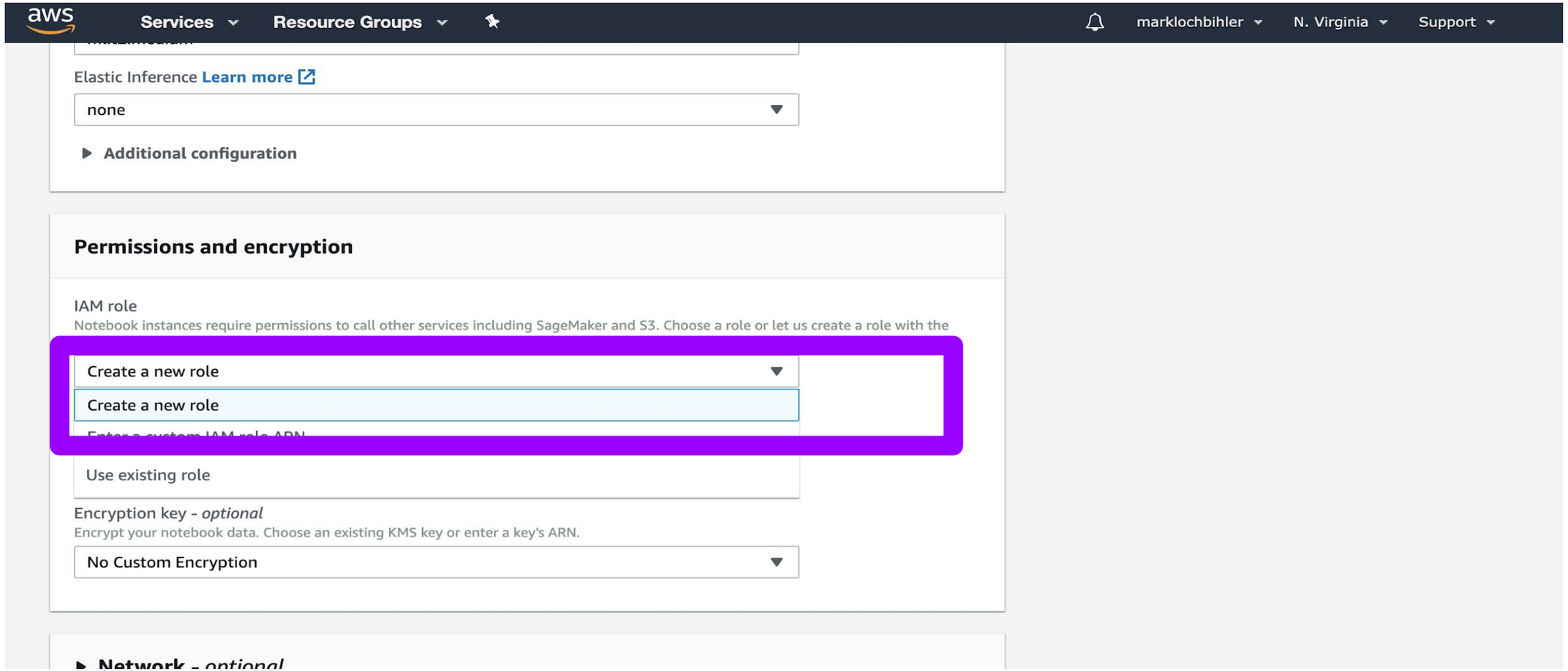
Notebook instance name = <Make up a Name for your Notebook>

Suggestion: RandomCutForestML

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

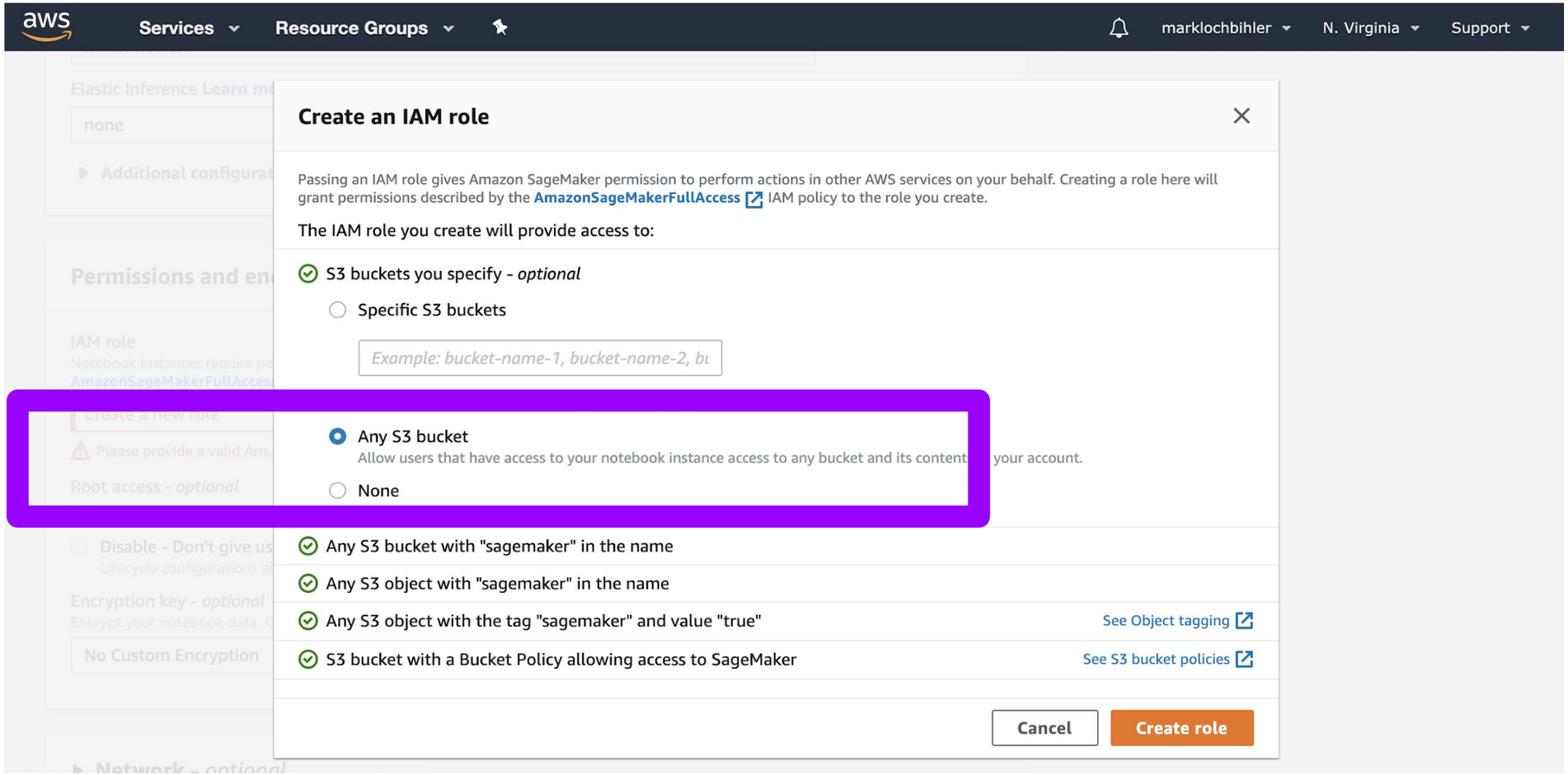
Click on "Create a new role"



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

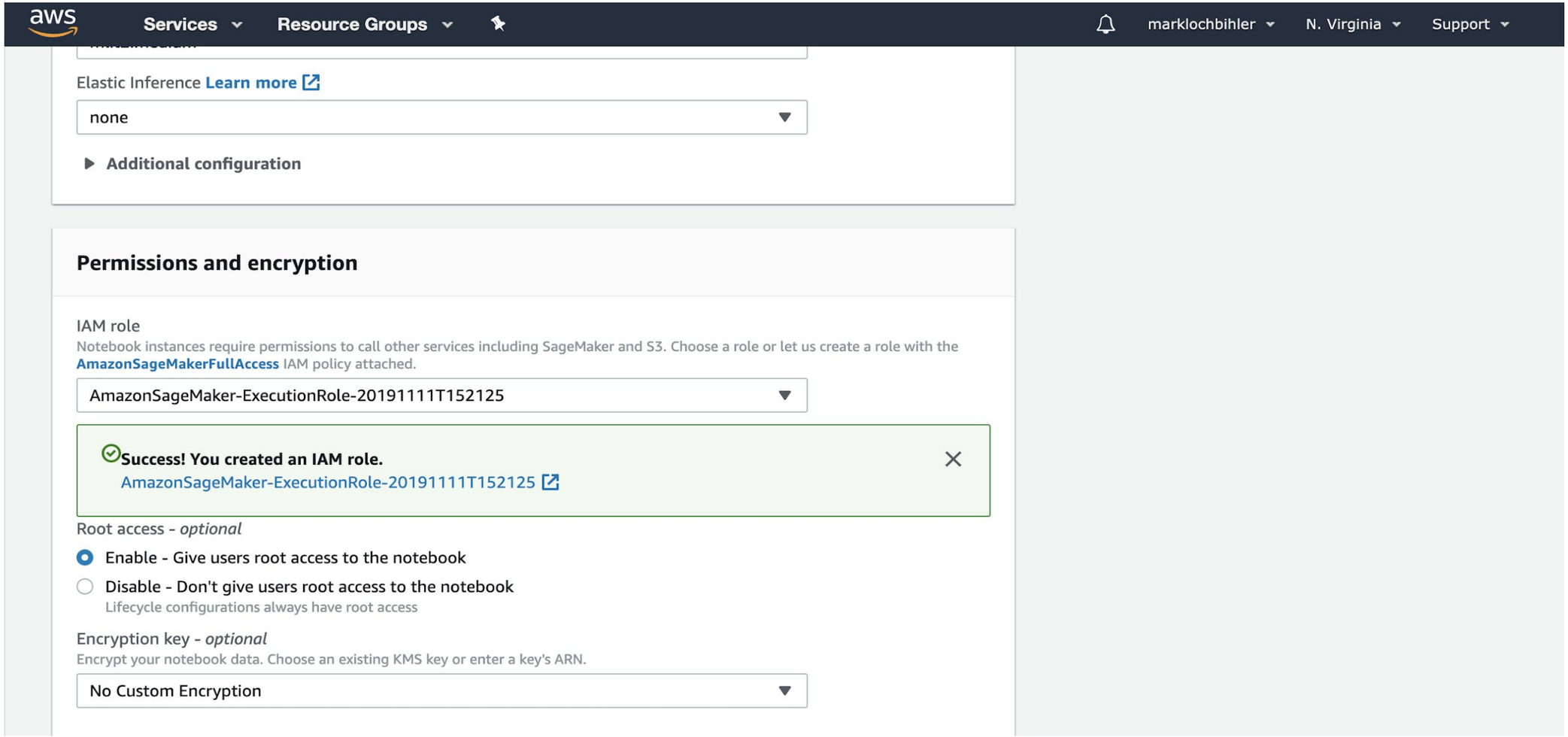
Make sure you select “Any S3 Bucket” radio button and then hit “Create Role”



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

If the role is created, you should see this screen



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

Update the rest of the template as follows

Root Access = Choose "**Enable**"

Add **Tags** to your Notebook so that you can find this instance in the EC2 Services Console.

Here are some tag suggestions:

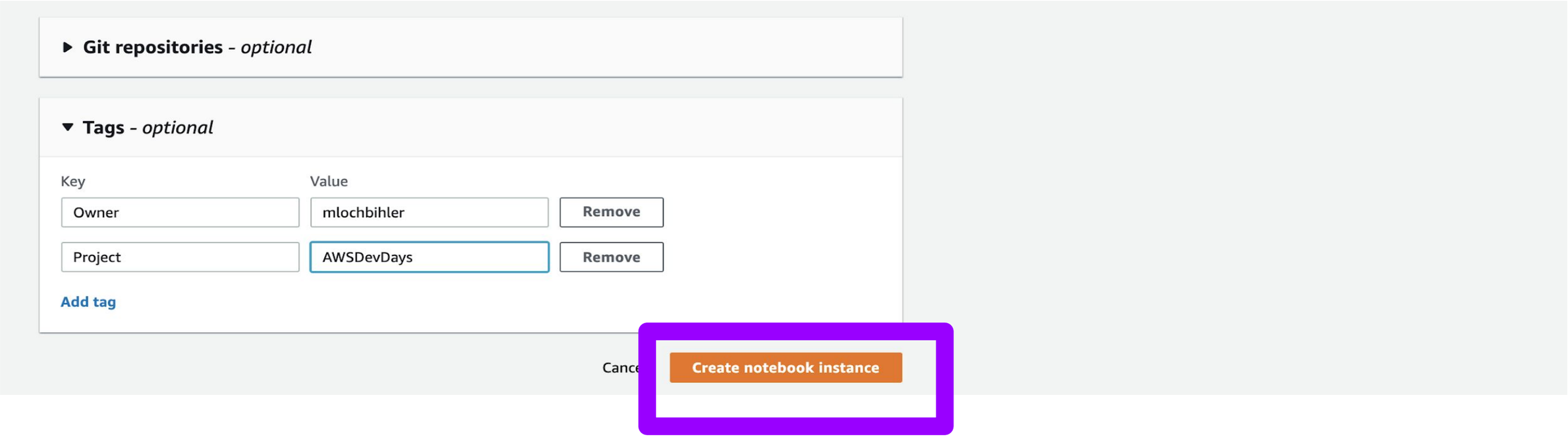
Owner = your_first_initial_last_name

Project = "AWSDevDays"

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

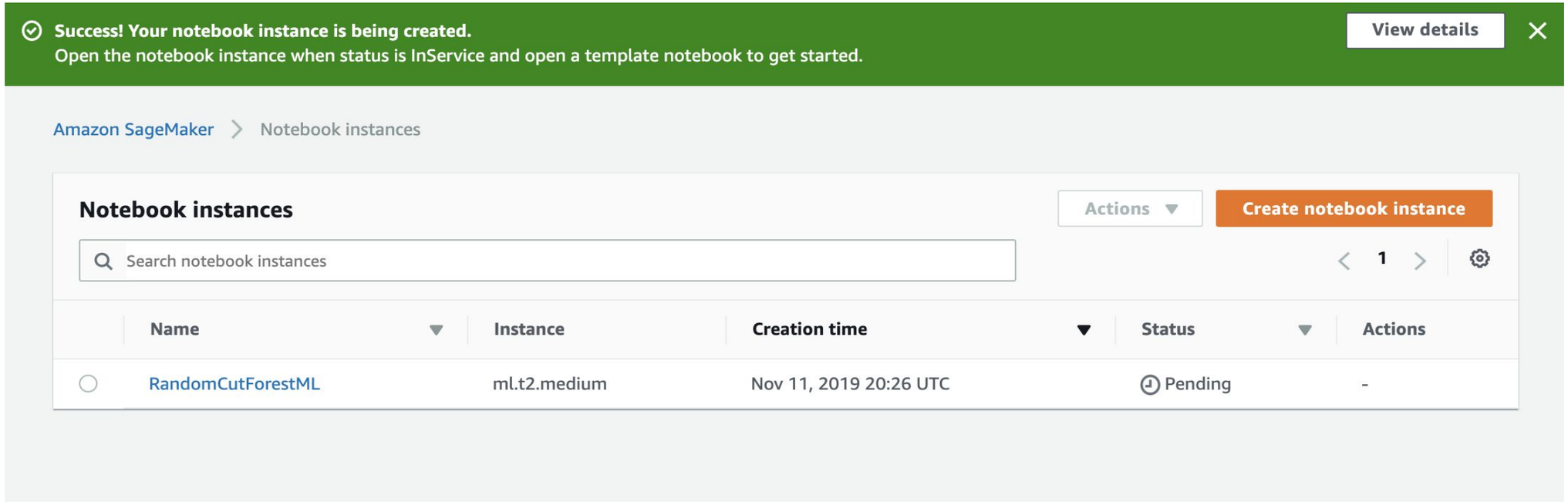
Enter Tags similar to before and then hit "Create Notebook Instance" on this screen"



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preparation - Launch an AWS SageMaker Notebook Session (continued)

You should see something similar to this the below screen. "Success!....."



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker

Machine learning for every developer and data scientist.

Get started with Amazon SageMaker

Amazon SageMaker provides every developer and data scientist with the ability to build, train, and deploy machine learning models quickly. Amazon SageMaker is a fully-managed service that covers the entire machine learning workflow to label and prepare your data, choose an algorithm, train the model, tune and optimize it for deployment, make predictions, and take action. Your models get to production faster with much less effort and lower cost.

BUILD

Collect & prepare training data

Data labeling & pre-built notebooks for common problems

Choose & optimize your ML algorithm

Built-in, high-performance algorithms and hundreds of ready to use algorithms in **AWS Marketplace**

TRAIN

Set up & manage environments for training

One-click training on the highest performing infrastructure

Train & tune model

Train once, run anywhere & model optimization

DEPLOY

Deploy model in production

One-click deployment

Scale & manage the production environment

Fully managed with auto-scaling for 75% less

Operationalizing ML with SageMaker and MemSQL

AWS SageMaker

Amazon SageMaker Random Cut Forest(RFC):

An algorithm designed to detect anomalous data points within a dataset.

Amazon S3

Sample NYC Taxi Average Ridership dataset:

lochbihler/nyctaxi/

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The only edit you need to make to this notebook is below: update “bucket” and “prefix” then run each step to train and then deploy the model to a restful endpoint. You can then download the model for use in this workshop lab.

jupyter random_cut_forest_memsql_lab (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Not Trusted conda_python3

Run Code nbdiff

Add tag

Select Amazon S3 Bucket

We first need to specify the locations where we will store our training data and trained model artifacts. ***This is the only cell of this notebook that you will need to edit.*** In particular, we need the following data:

- `bucket` - An S3 bucket accessible by this account.
- `prefix` - The location in the bucket where this notebook's input and output data will be stored. (The default value is sufficient.)

In [ ]: parameters

```
import boto3
import botocore
import sagemaker
import sys


bucket = 'lochbihler'   # <--- specify a bucket you have access to
prefix = 'sagemaker'
execution_role = sagemaker.get_execution_role()


# check if the bucket exists
try:
    boto3.Session().client('s3').head_bucket(Bucket=bucket)
except botocore.exceptions.ParamValidationError as e:
    print('Hey! You either forgot to specify your S3 bucket'
          ' or you gave your bucket an invalid name!')
except botocore.exceptions.ClientError as e:
    if e.response['Error']['Code'] == '403':
        print('Hey! You don't have permission to access the bucket, {}.'.format(bucket))
    elif e.response['Error']['Code'] == '404':
        print('Hey! Your bucket, {}, doesn't exist!'.format(bucket))
    else:
        raise
else:
    print('Training input/output will be stored in: s3://{}/{}'.format(bucket, prefix))
```

bucket = ‘lochbihler’
prefix = ‘sagemaker’

aws | SingleStore | Looker

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.