
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

AIM 320

Boost ML development productivity with managed SageMaker notebooks

Sumit Thakur (he/him/his)

Principal Product Manager
Amazon Web Services (AWS)

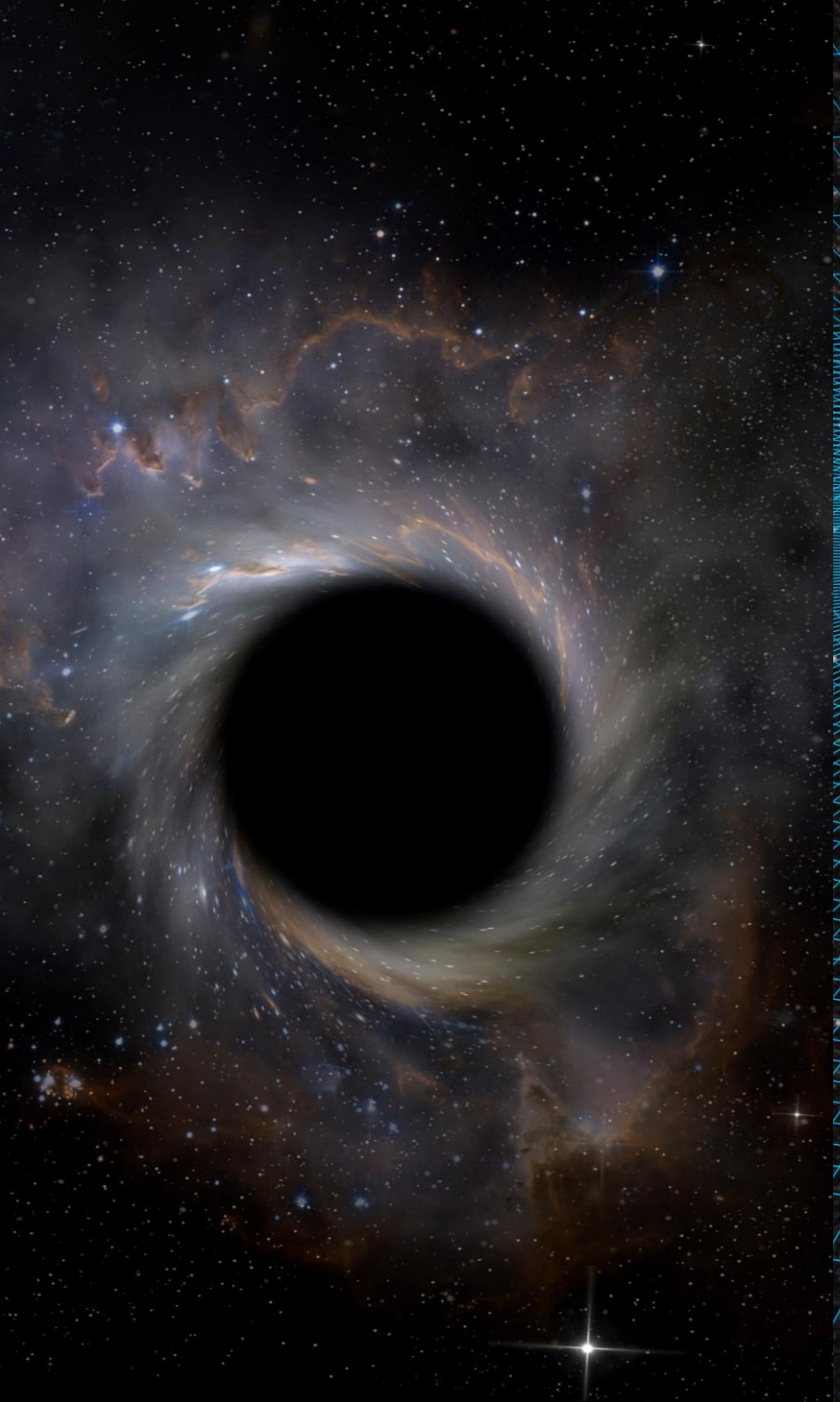
Sean Morgan (he/him/his)

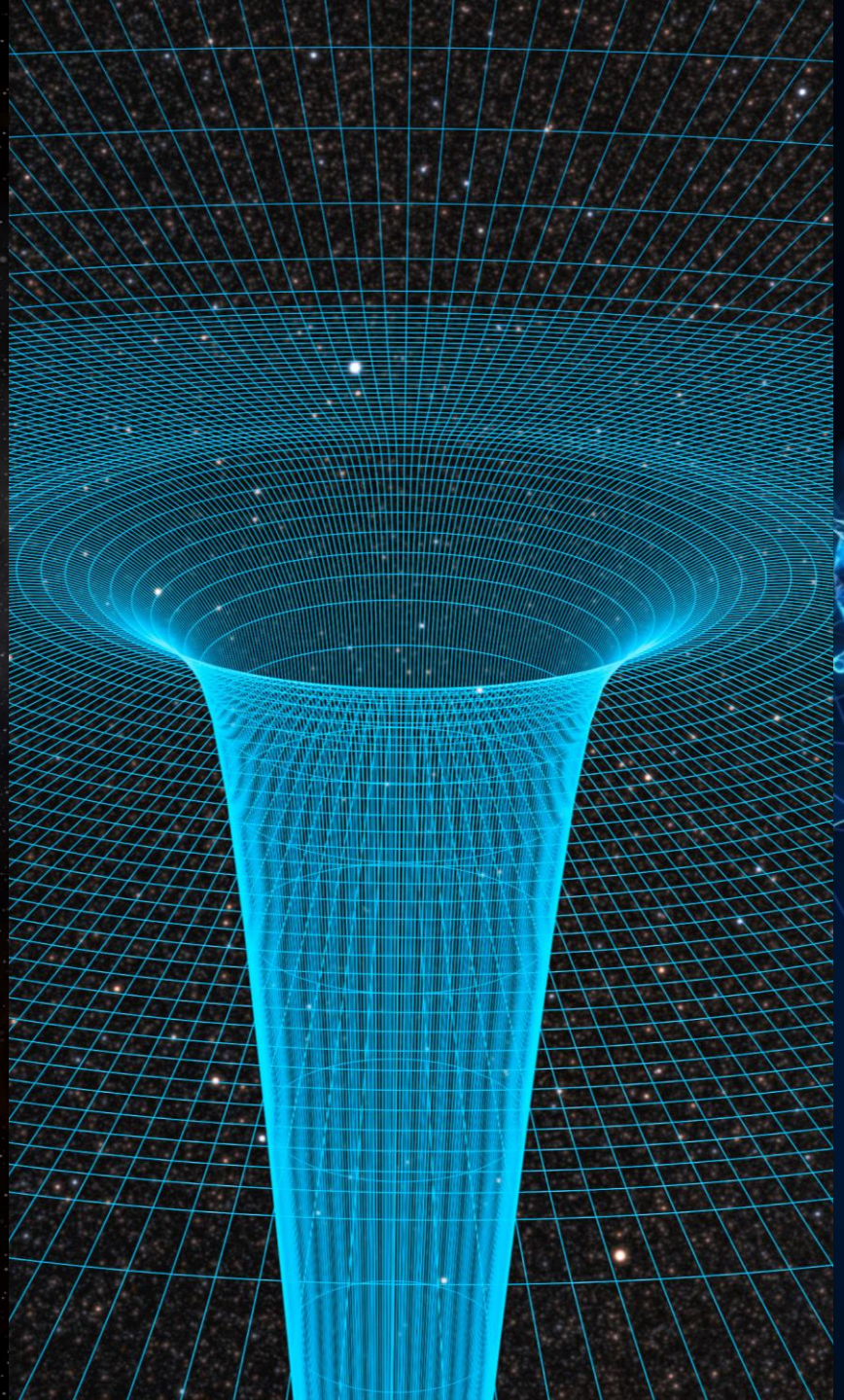
Sr. ML Specialist SA
Amazon Web Services (AWS)

Ritesh Shah (he/him/his)

Chief Architect - AI/ML
Vanguard

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.





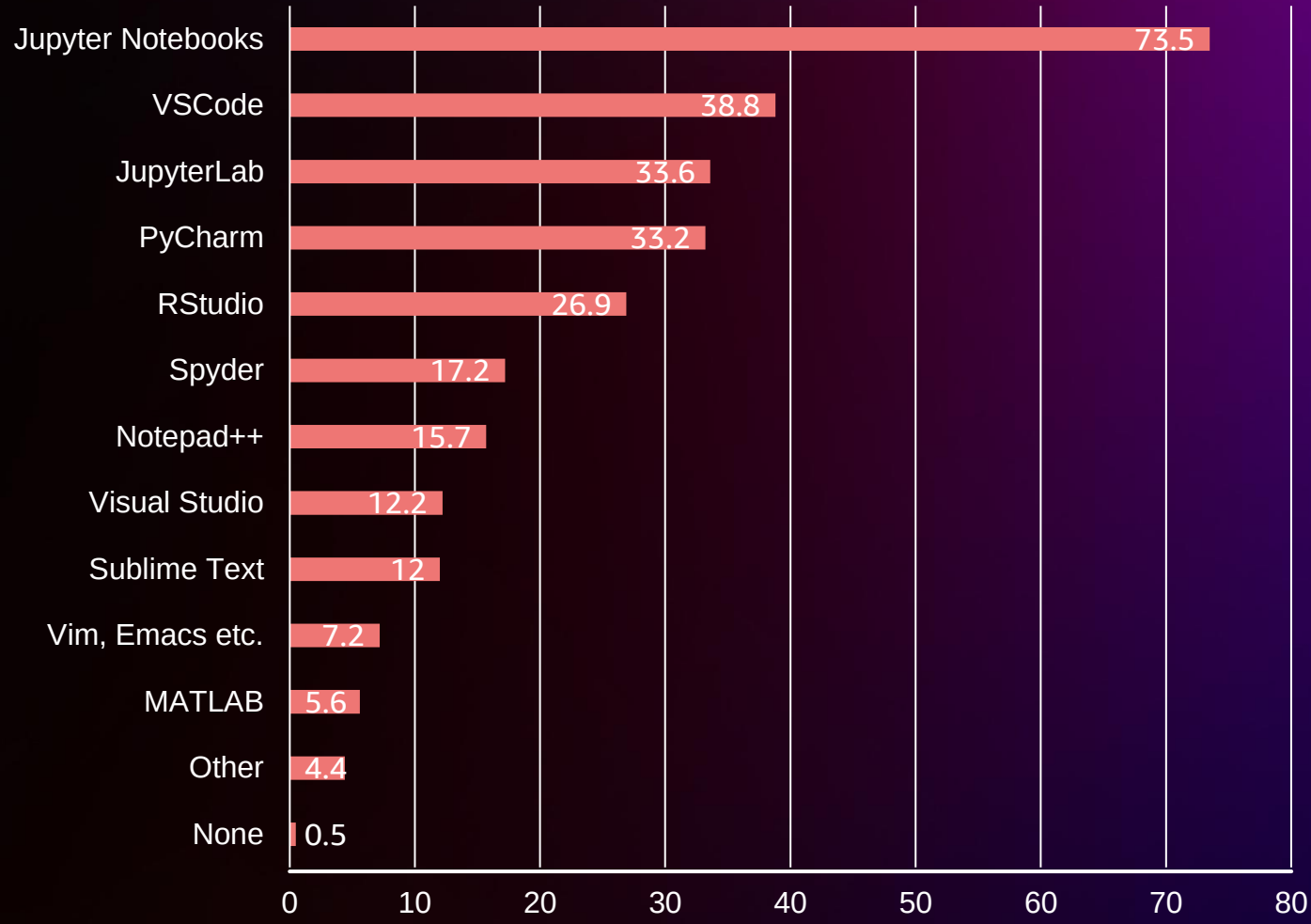


Growing adoption of Jupyter Notebooks



2015 2022

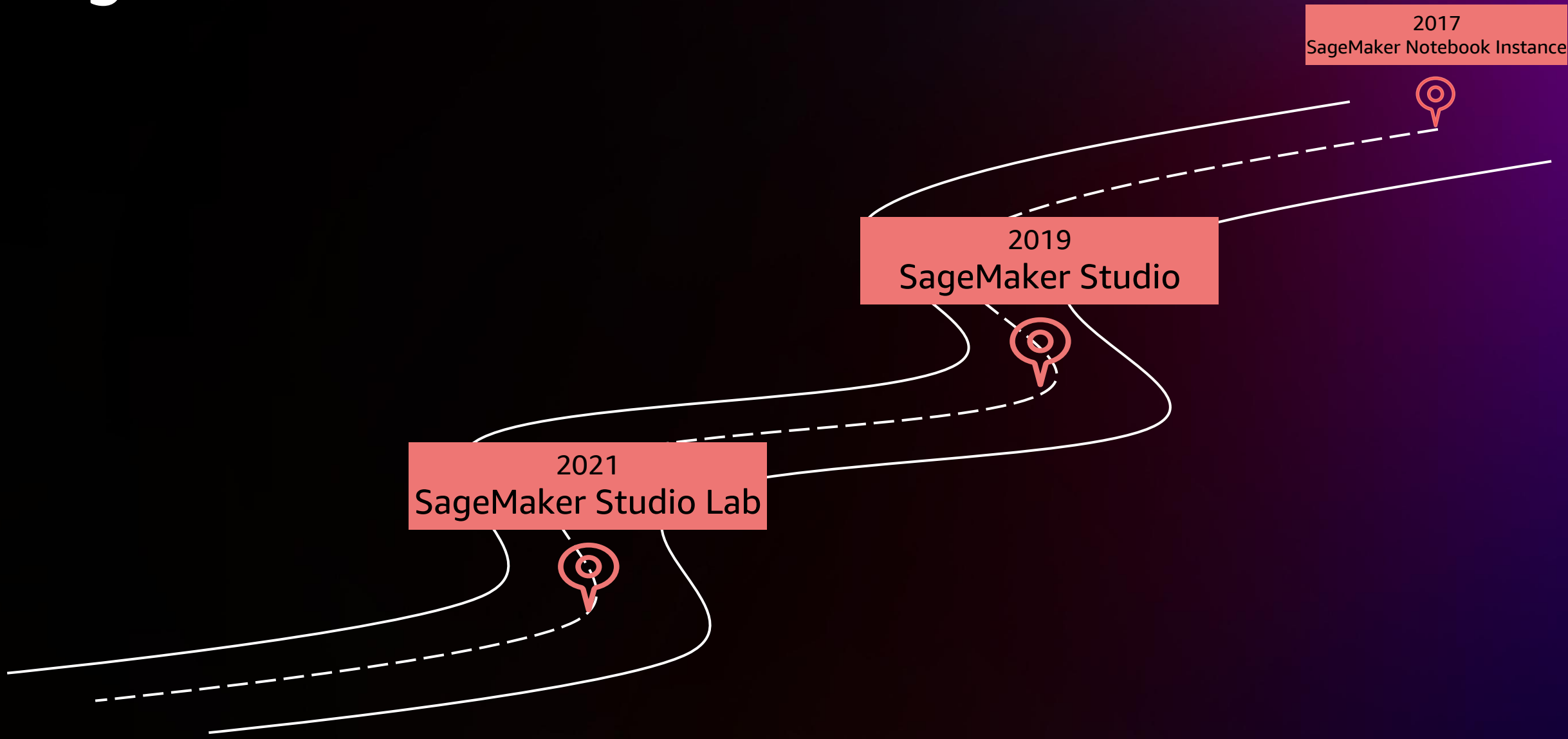
Number of notebooks on GitHub



State of Machine Learning and Data Science -2021 by Kaggle

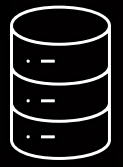
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Notebooks innovations



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML builder workflow



Data preparation

Explore, clean, and transform data



Experimentation

Build, share, and compare models



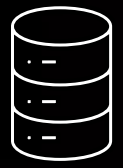
Productionization

Deploy, scale, and automate workflow

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML builder challenges



Data preparation

- Boilerplate code for exploring data
- Overheads in provisioning and managing clusters



Experimentation

- Fragmented, asynchronous collaboration
- Delayed feedback loops



Productionization

- Heavy lifting to package and deploy code
- Learning curve of DevOps tools and practices

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

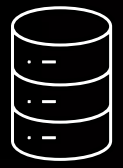
Introducing next generation SageMaker notebooks

The best notebook experience from data exploration to deployment

New capabilities include simplified data preparation, serverless notebook kernels, real time collaboration, and automatic conversion of notebook code to production-ready jobs in SageMaker Studio

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML builder challenges



Data preparation

- Boilerplate code for exploring data
- Overheads in provisioning and managing clusters



Experimentation

- Fragmented, asynchronous collaboration
- Delayed feedback loops



Productionization

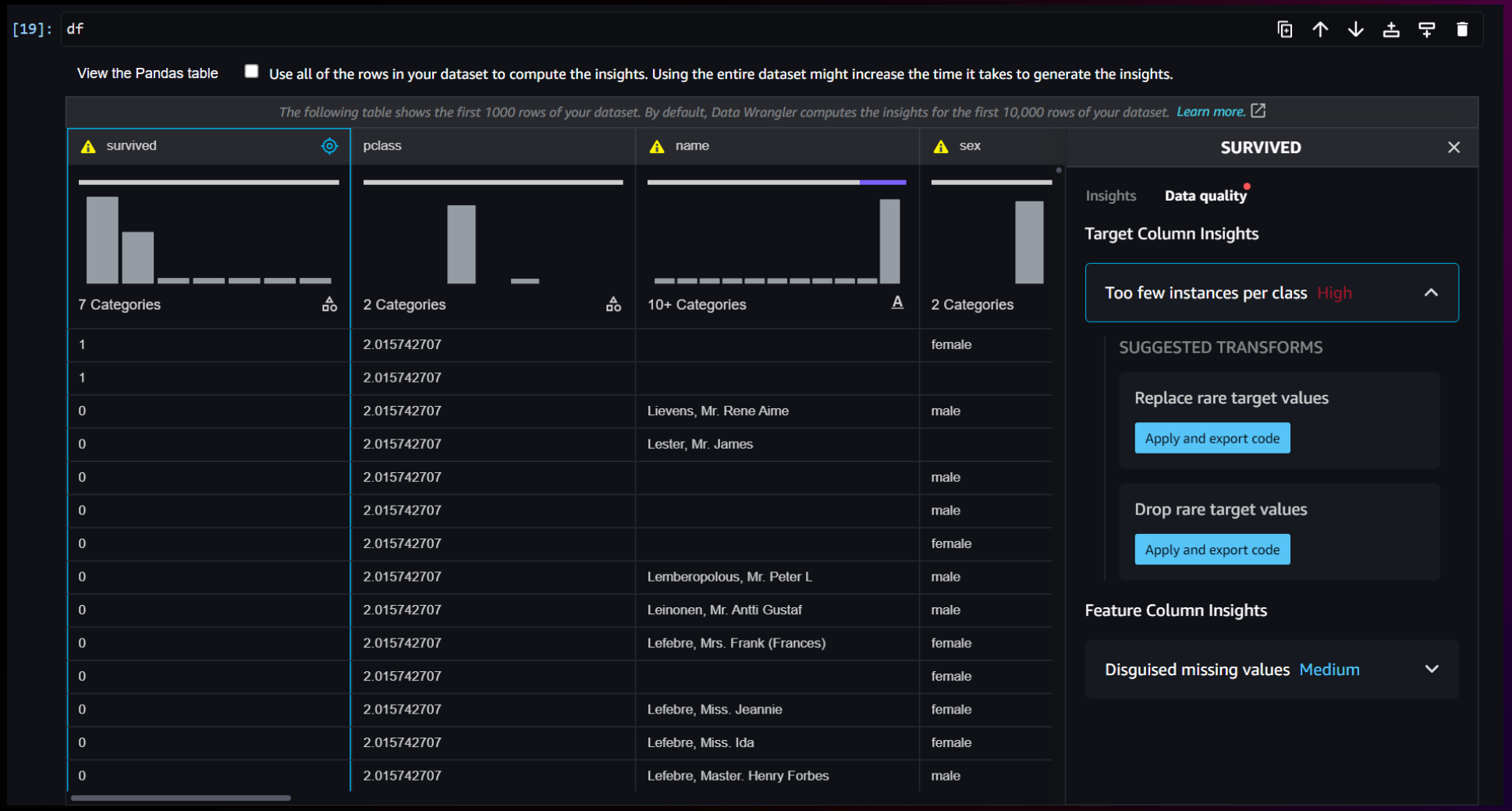
- Heavy lifting to package and deploy code
- Learning curve of DevOps tools and practices

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

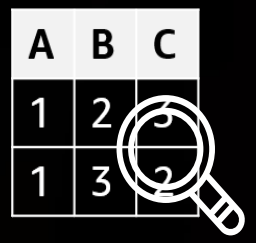
Simplified data preparation

VISUALLY INSPECT AND ADDRESS DATA-QUALITY ISSUES IN A FEW CLICKS IN SAGEMAKER NOTEBOOKS

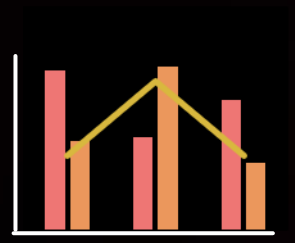


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How it works



Query and display Pandas data frame



Get automatic data distributions



Get data quality alerts with no-code remedies

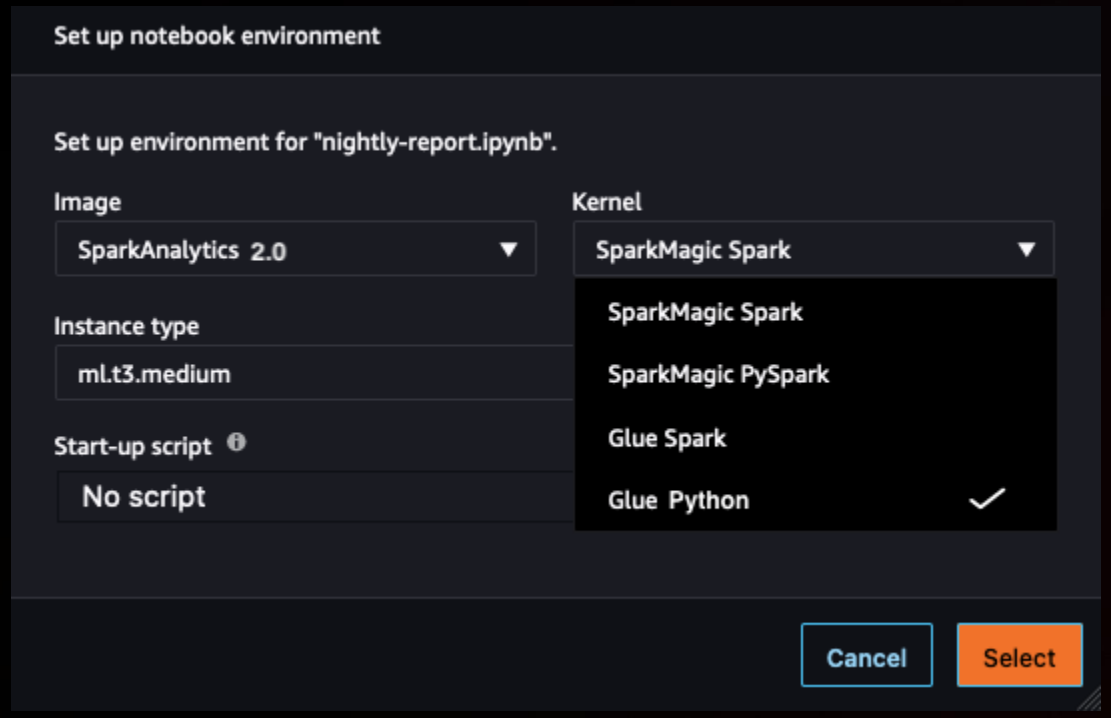


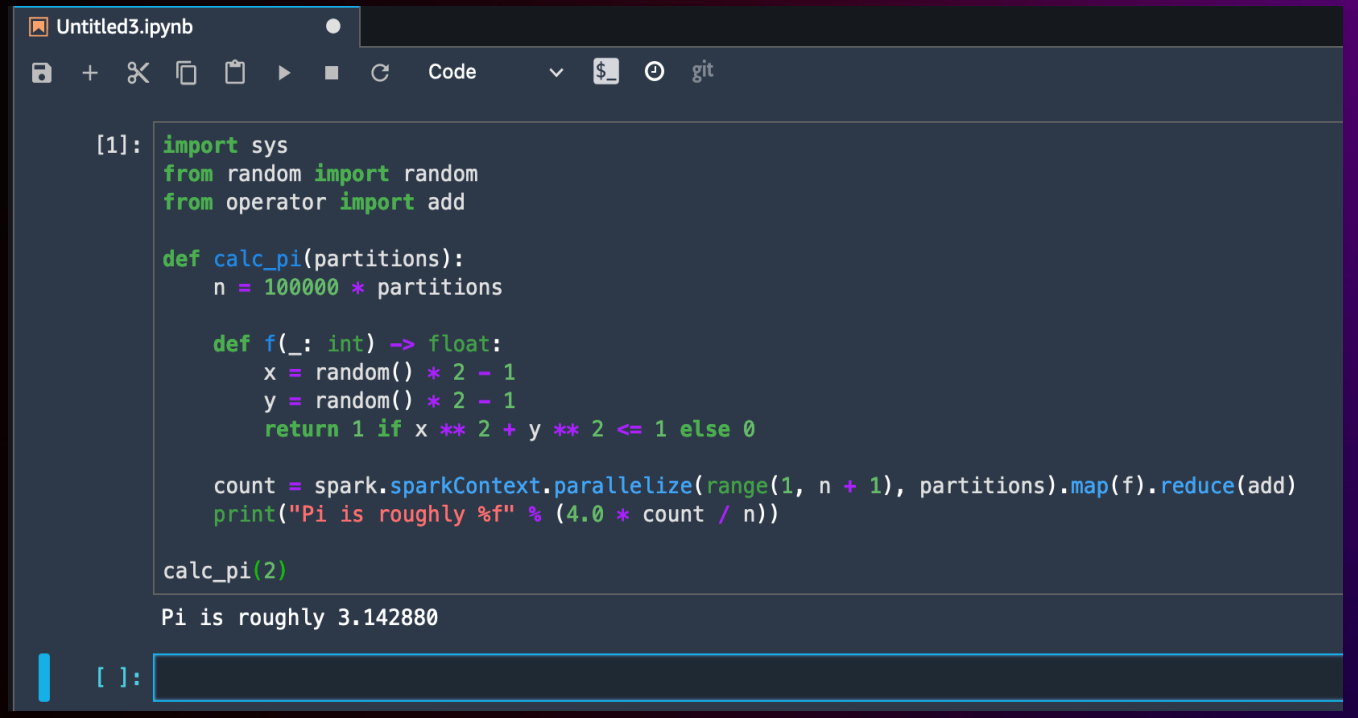
Generate code for reproducibility

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless data preparation

INTERACTIVELY PREPARE DATA AT SCALE USING SERVERLESS APACHE SPARK AND RAY IN SAGEMAKER STUDIO NOTEBOOKS



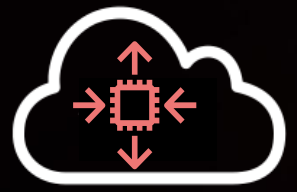


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

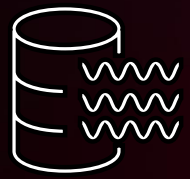
How it works



Choose the AWS Glue (Python) kernel in Studio Notebook



Start an interactive, serverless session managed by AWS Glue



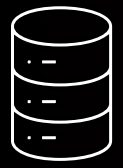
Browse data catalog
Prepare data using Apache Spark or Ray



Pay-as-you-go

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML builder challenges



Data preparation

- Boilerplate code for exploring data
- Overheads in provisioning and managing clusters



Experimentation

- Fragmented, asynchronous collaboration
- Delayed feedback loops



Productionization

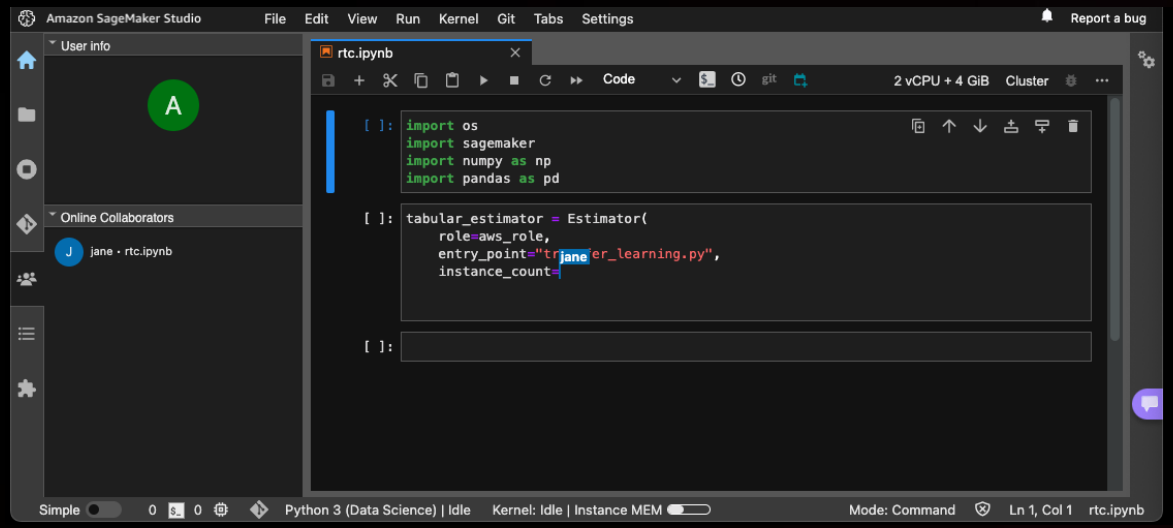
- Heavy lifting to package and deploy code
- Learning curve of DevOps tools and practices

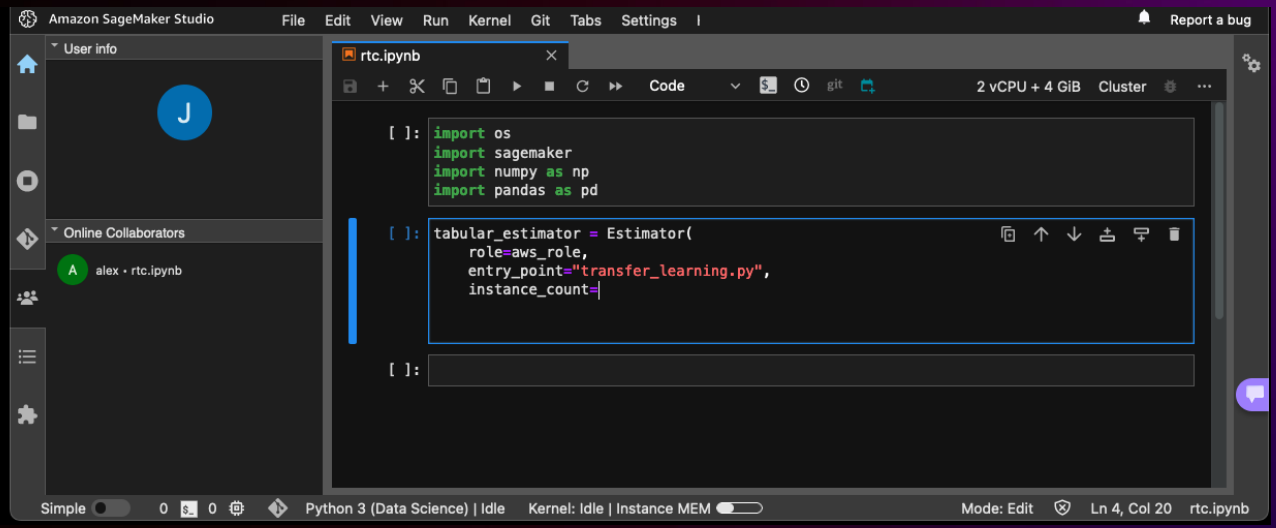
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real time collaboration

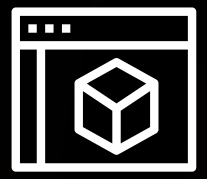
ACCELERATE COLLABORATION ACROSS DATA SCIENCE TEAMS USING SAGEMAKER NOTEBOOKS



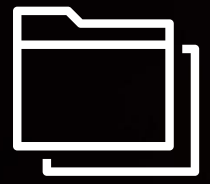


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How it works



Create a shared space



Neatly organize code, data, and experiments



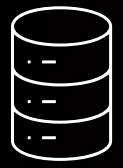
Co-edit, execute, and review notebooks together in real time



Checkpoint code in Git/Bitbucket repositories

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML builder challenges



Data preparation

- Boilerplate code for exploring data
- Overheads in provisioning and managing clusters



Experimentation

- Fragmented, asynchronous collaboration
- Delayed feedback loops



Productionization

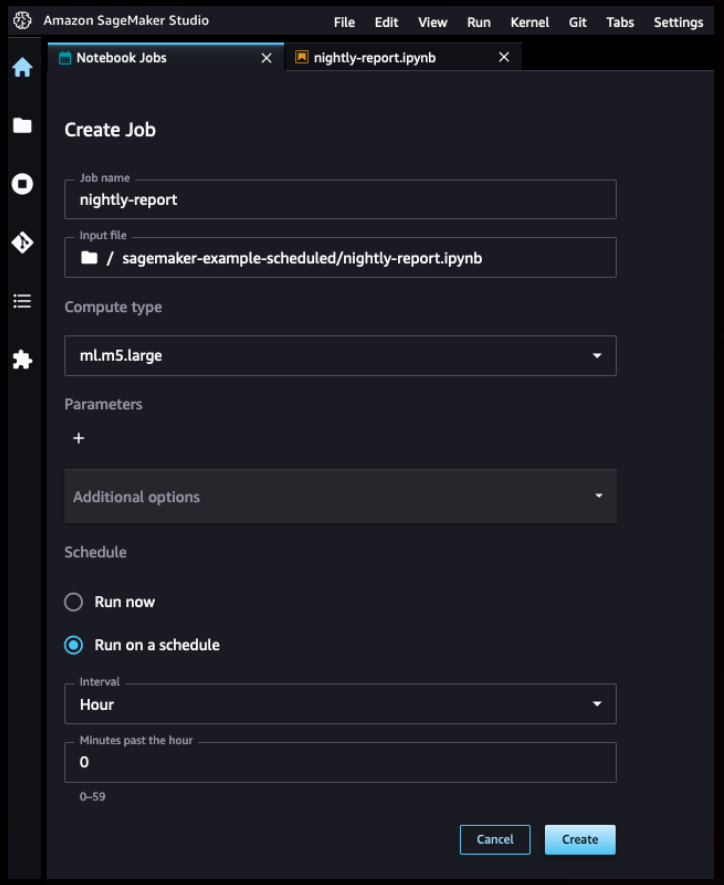
- Heavy lifting to package and deploy code
- Learning curve of DevOps tools and practices

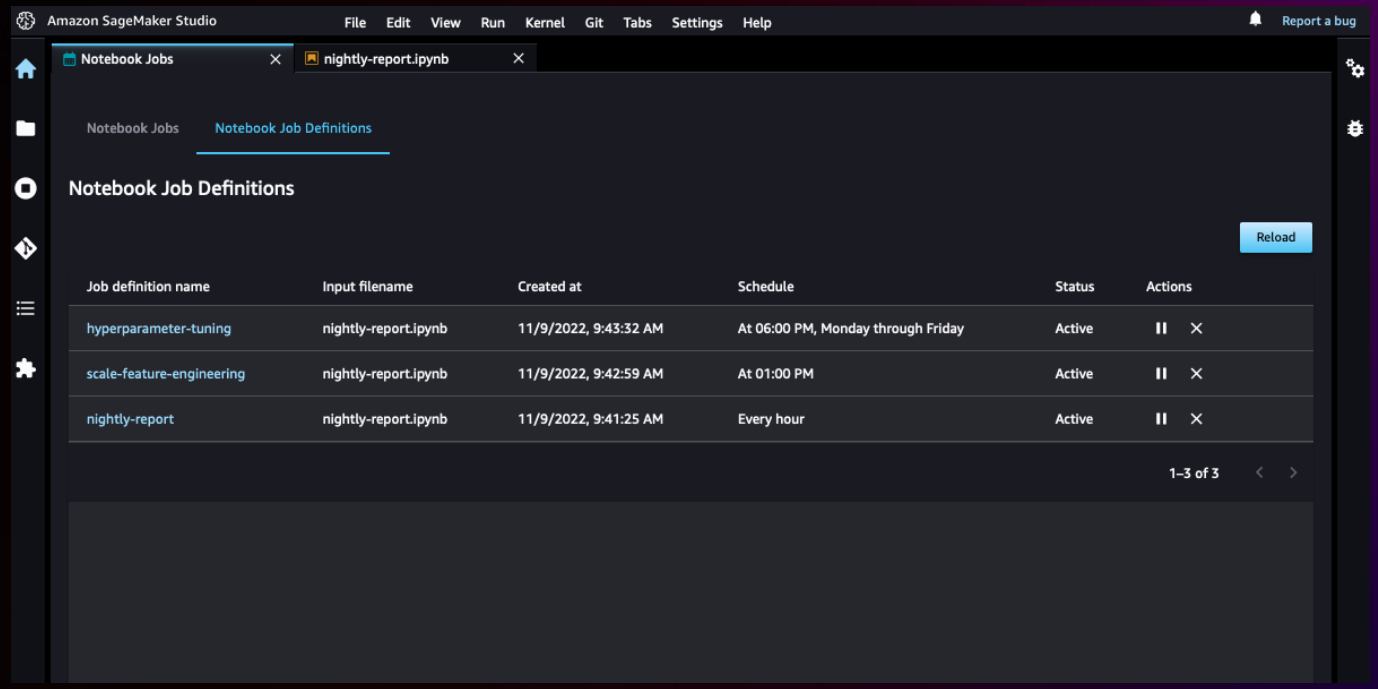
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Automatic conversion of notebook code to production-ready jobs

EXPEDITE MACHINE LEARNING DEVELOPMENT FROM EXPERIMENTATION TO PRODUCTION





© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

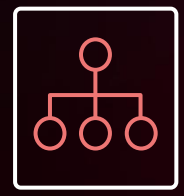
How it works



Choose a notebook file in Studio



Choose a time-based schedule



Submit notebook job
Monitor jobs in Studio



Schedule and monitor jobs from Studio Lab

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

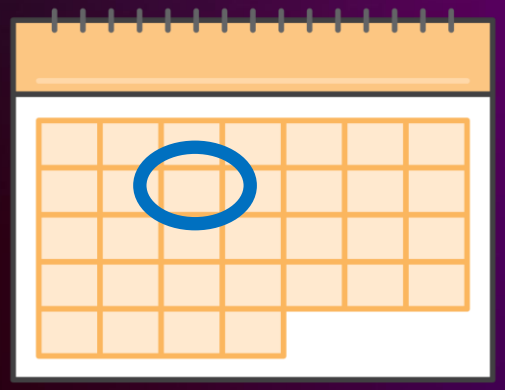
Vanguard: Decrease time to insights with unified user experience

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

About Vanguard

Core purpose: To take a stand for all investors, to treat them fairly, and to give them the best chance for investment success.





Founded on May 1, 1975, Vanguard is one of the world's largest investment companies, with around $7.1 trillion AUM, 412 funds offered worldwide, 0.09% asset-weighted average U.S. fund expenses, and approximately 18,800 crew members worldwide

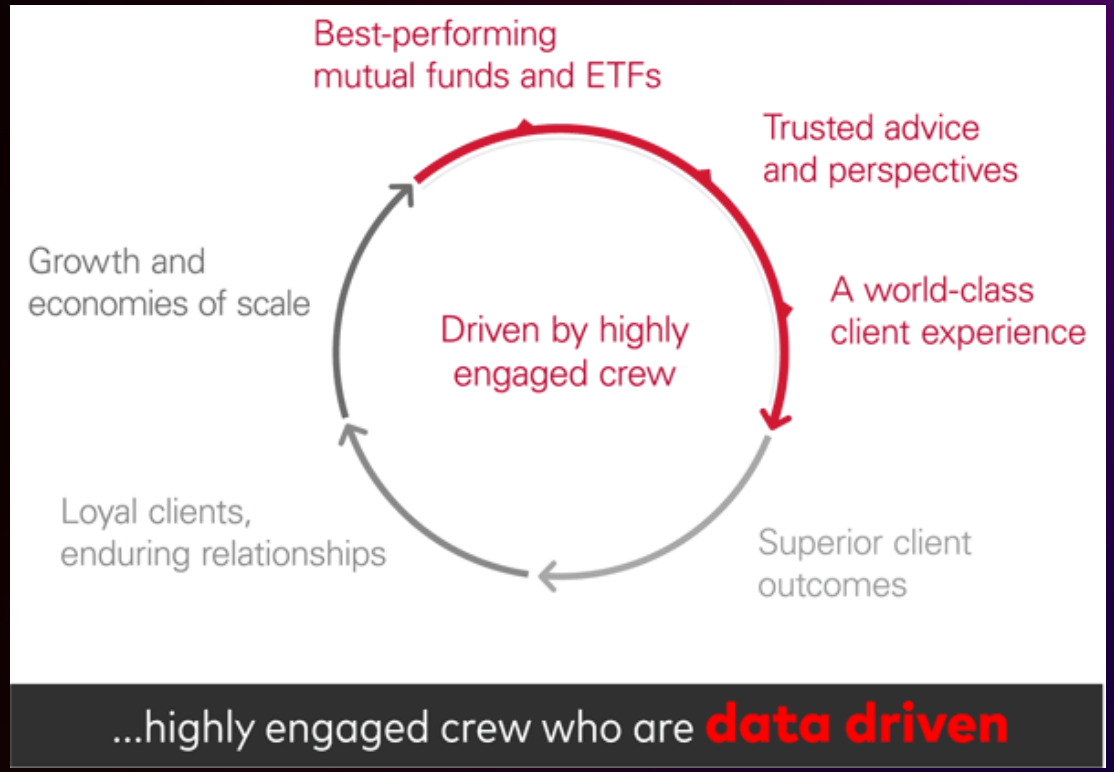
* Oct 31st, 2022

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CDAO - Vision

Vanguard operates in a highly regulated industry; therefore we need strong **data defense** as well as strong **data offense** to differentiate ourselves in a very competitive dynamic environment

Our two-sided data strategy will transform data into an asset and drive Vanguard's fly wheel



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CDAO – Strategic imperatives and key results

We will enable Vanguard to be a data-driven organization by building a data culture and ecosystem that drives business outcomes using data, analytics, and AI/ML at scale while adhering to our privacy obligations to properly use our client's information

STRATEGIC IMPERATIVES



EXECUTION LEVERS

- Data talent strategy
- Data literacy
- Collaboration with privacy and protection teams

- Data driven architecture (lake 2.0)
- Simple, effective platforms
- Elevated governance

- Standardized data products
- Expand global data management

- Ethical AI/ ML at scale
- Integrate and strengthen D&A organizations

KEY RESULTS

WIG

Decreased time to insights
from weeks to days

Decreased time to fill position
Increase retention & decrease cost of recruiting

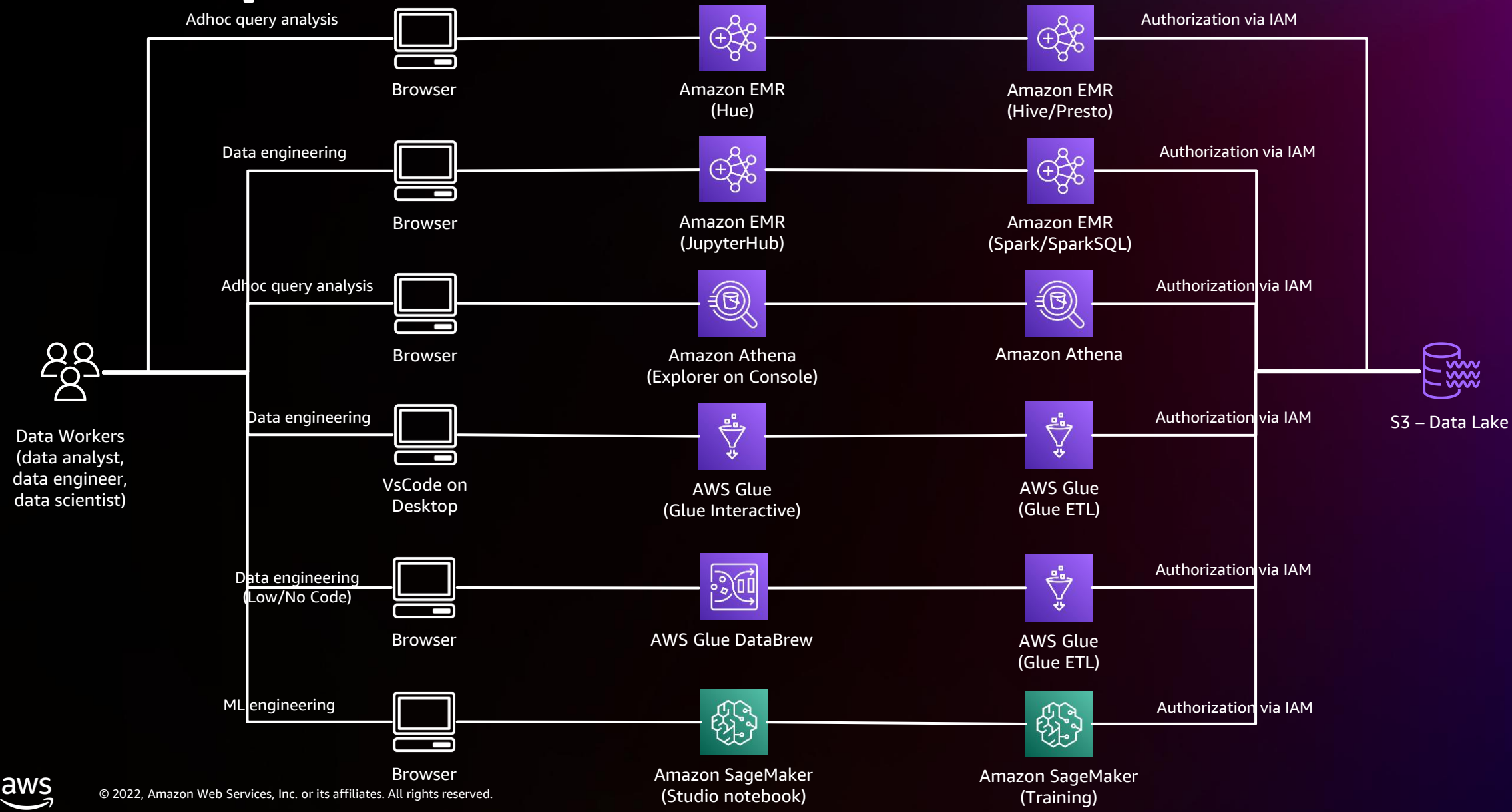
Increased NPV
Maintain and scale CAI's ROI of 3.5x cost

Decreased regulatory risk
Prevent fines, preserve reputation, good will

Financial discipline target
5% efficiency realized by 2022 YE

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

User experience for data workers – Current state



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Metrics and areas of opportunities



~27K nodes used for Amazon EMR
~90 nodes used for Amazon Redshift



~5000 queries per day for Amazon Athena
~1100 workloads per day on AWS Glue



~20PB of data available in Amazon S3 based data lake



~1350 unique with ~350 active monthly users of Amazon SageMaker Studio



Cognitive burden in learning multiple user interfaces
Leads to increased start time to insight



Context switching and loss of productivity
Leads to increased time to develop insights



Operational overhead in setting up multiple environments
Leads to increased cost



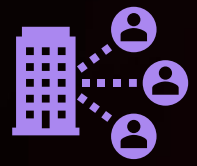
Inefficient collaboration between data workers
Leads to increased deployment time

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

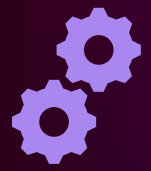
Principle tenets



Unified user experience for all data workers



Minimize handoffs and increase collaboration amongst data workers

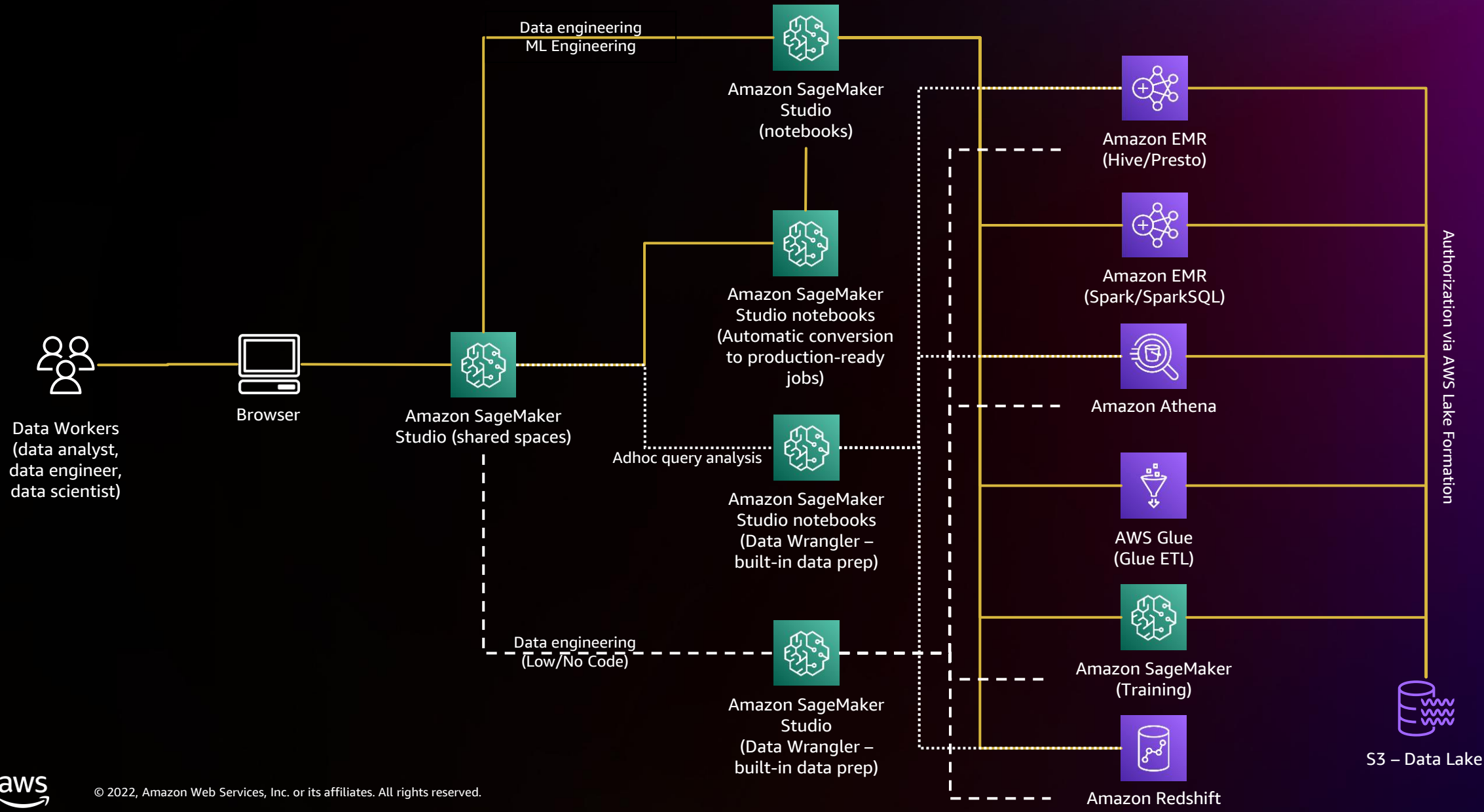


Self-service and automate setup of user experience, compute and access to data

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Unified experience for data workers - Upcoming



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways



Use Amazon SageMaker Studio integration with AWS Glue and Amazon EMR to provide a unified interface for data engineering and machine learning



Use Amazon SageMaker Studio notebooks automation to simplify deployment of pipelines and reduce operational overhead



Use the shared spaces in Amazon SageMaker to minimize handoffs and increase collaboration among data workers



Enable FinOps with continuous focus on cost transparency and cost management to reap benefits from derived insights

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

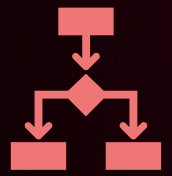
Peek into future (in partnership with AWS)



Best-in class experience for data development within Amazon SageMaker Studio (similar to model engineering)



Enhanced collaboration and integration with data product and data catalogs at AWS



Introduce DataOps capabilities as native integration within Amazon SageMaker Studio



Support higher level of resiliency through multi-domain and multi-region capabilities

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Sumit Thakur
sumthaku@amazon.com

Sean Morgan
seanmrgn@amazon.com

Ritesh Shah
ritesh_shah@vanguard.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.