

Imagine that you work at a company that is planning to use Amazon Simple Storage Service (Amazon S3) as the storage layer for their data lakes solution. Initially, the data that will be ingested into the data lake will come from three locations:

- Internet of Things (IoT) sensors that send real-time data
- A database with historical records
- Supplemental data from third-party entities for enriching internally generated data

The company has tasked you with designing solutions for ingesting this data into their data lake, and each location (IoT sensors, database, and third party) will need its own ingestion solution. From there, you will need to also design a solution for how to clean or transform the data so that it can be analyzed. The company currently uses Apache Hadoop-based software. When possible, they prefer to use similar technologies in the cloud so that they don't need to retrain

Pre-Requisites

- **AWS Account**
- **1 or more IoT devices sending messages to IoT Core**

© 2021, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

- In the "Execution role" section, create a new role or choose an existing role with appropriate permissions to read from the

Kinesis stream.

- Click "Create function."

Step 3: Configure the Lambda Function Trigger

- In the Lambda function's configuration, scroll down to the "Add triggers" section.
- Add a trigger and select "Kinesis" as the trigger source.
- Choose the Kinesis Data Stream that you want to consume data from.

- Set the batch size, which specifies how many records Lambda will process in a single invocation.
- Configure the starting position (e.g., latest, trim horizon) based on your use case.

Step 4: Write Lambda Function Code

- In the Lambda function editor, write code to process the data from the Kinesis stream. This code should be specific to your application's logic.
- You can access the Kinesis records as event data within your Lambda function code.

Step 5: Testing

- Test your Lambda function with sample data or by sending data to your Kinesis Data Stream.
- Ensure that your Lambda function processes the data as expected.

Step 6: Deployment

- Once you've tested your Lambda function and it works correctly, you can deploy it for production use.

Step 7: Monitoring and Scaling

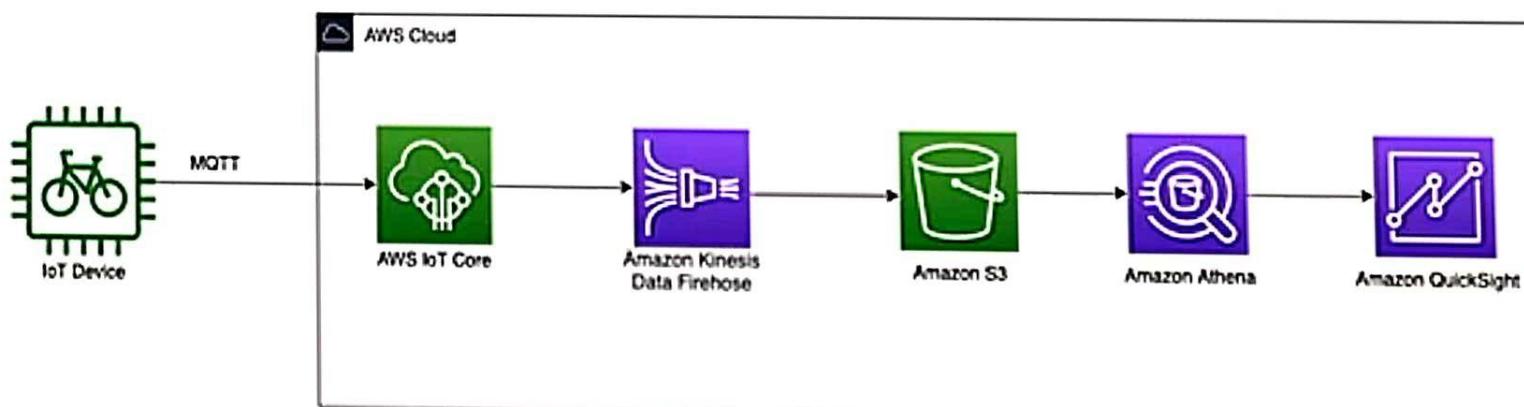
- Monitor the performance of your Lambda function and configure appropriate error handling.
- Consider auto-scaling options to handle increased data volume.

Step 8: Continuous Maintenance

- Regularly monitor your Lambda function's performance and adjust configurations if needed.
- Keep an eye on your Kinesis Data Stream to ensure data retention settings meet your requirements.

By following these steps, you can use AWS Lambda functions to consume and process data from an Amazon Kinesis Data Stream. This architecture allows you to build real-time data processing and analytics applications with ease.

Pipeline to build



The screenshot shows the AWS Management Console homepage. At the top, there's a navigation bar with links for Services, Search, and Support. The main title "AWS Management Console" is prominently displayed. Below the title, there's a section titled "AWS services" with a "Recently visited services" list. This list includes icons for Lambda, RDS, IAM, CloudFormation, API Gateway, Amazon Translate, EC2, Amazon EventBridge, Support, and VPC. There are also links for "SS" and "DynamoDB". A "All services" link is located at the bottom of this list. To the right, there's a sidebar with sections for "Stay connected to your AWS resources on-the-go" (mentioning the AWS Mobile App), "Explore AWS" (with links to AWS Certification and Free AWS Training), and "Get Started with Amazon S3" (with a "Get started now" button). At the bottom, there are sections for "Build a solution" (with icons for launching a virtual machine, building a web app, using virtual servers, registering a domain, connecting an IoT device, migrating to AWS, starting a development project, and deploying a serverless microservice) and a "Have feedback?" section.

https://eu-west-1.console.aws.amazon.com/iot/home?region=eu-west-1&tab=home

AWS Services ▾ AWS IoT [Option+1] Ireland Support ▾

AWS IoT

X Introducing the new AWS IoT console experience. We're updating the console experience for you. Learn more Try the new experiences and let us know what you think. You can turn off the new experience from the navigation menu.

Monitor Activity Onboard Manage Fleet Hub Greengrass Wireless connectivity Secure Defend Act Test Device Advisor MQTT test client

Software Settings Learn Feature spotlight Documentation ▾

New console experience Tell us what you think



AWS IoT

AWS IoT is a managed cloud platform that lets connected devices - cars, light bulbs, sensor grids, and more - easily and securely interact with cloud applications and other devices.


Connect and manage your devices
Connect devices to the cloud using the protocol that best fits your requirements - HTTP, MQTT, or WebSocket. Devices can communicate with each other even if they are using different protocols.


Process and act upon device data
Filter, transform, and act upon data from devices on the fly, based on business rules. AWS IoT can be easily integrated with AWS services like Amazon DynamoDB, Amazon Kinesis, Amazon Machine Learning, and AWS Lambda.


Read and set device state at any time
AWS IoT stores the latest state of a device so that it can be read or set anytime, even when the device is offline.

[Learn more](#) [Learn more](#) [Learn more](#)

Introducing the new AWS IoT console experience. We're updating the console experience for you. [Learn more](#). Try the new experiences and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > MQTT test client

MQTT test client

You can use the MQTT test client to monitor the MQTT messages being passed in your AWS account. Devices publish MQTT messages that are identified by topics to communicate their state to AWS IoT. AWS IoT also publishes MQTT messages to inform devices and apps of changes and events. You can subscribe to MQTT message topics and publish MQTT messages to topics by using the MQTT test client.

[Subscribe to a topic](#) [Publish to a topic](#)

Topic filter: `#`
The Topic Filter describes the topic(s) to which you want to subscribe. The Topic Filter can include MQTT wildcards characters.

`iot-tank/tester#`

[Additional configuration](#)

[Subscribe](#)

Subscriptions	Topic
You have no topic subscriptions.	Subscribe or select a topic to view incoming messages.

New console experience
[Tell us what you think](#)

Introducing the new AWS IoT console experience

We're updating the console experience for you. Learn more Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > MQTT test client

MQTT test client

You can use the MQTT test client to monitor the MQTT messages being passed in your AWS account. Devices publish MQTT messages that are identified by topics to communicate their state to AWS IoT. AWS IoT also publishes MQTT messages to inform devices and apps of changes and events. You can subscribe to MQTT message topics and publish MQTT messages to topics by using the MQTT test client.

Subscribe to a topic Publish to a topic

Topic Filter info

The Topic Filter describes the topic(s) to which you want to subscribe. The Topic Filter can include MQTT wildcards characters.

octank/escoter

Additional configuration

Subscribe

Subscriptions

octank/escoter

octank/escoter

[{"deviceID": "escoter28", "timestamp": 1629189252, "battery": 28, "speed": 70, "location": {"lat": 48.21438112333178, "long": 11.59591183643197} }]

August 17, 2021, 11:54:12 (UTC+0200)

Pause Clear Export Edit

Introducing the new AWS IoT console experience. We're updating the console experience for you. Learn more Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT Services Search for services, features, marketplace products, and docs [Option+I] https://eu-west-1.console.aws.amazon.com/iot/home?region=eu-west-1&test

AWS IoT MQTT test client

You can use the MQTT test client to monitor the MQTT messages being passed in your AWS account. Devices publish MQTT messages that are identified by topics to communicate their state to AWS IoT. AWS IoT also publishes MQTT messages to inform devices and apps of changes and events. You can subscribe to MQTT message topics and publish MQTT messages to topics by using the MQTT test client.

Subscribe to a topic Publish to a topic

Topic Filter info The Topic Filter describes the topicId to which you want to subscribe. The Topic Filter can include MQTT wildcards characters.

octank/escoter

Additional configuration

Subscribe

Subscriptions octank/escoter

octank/escoter Pause Clear Export Edit

octank/escoter

[{"deviceId": "escoter3", "timestamp": 1629389272, "battery": 42, "speed": 8, "location": {"lat": 48.1372494445474, "long": 11.58849885584583}}

August 17, 2021, 11:54:42 (UTC+0200)

New console experience Tell us what you think

- Onboard
 - Manage
 - Fleet Hub
 - Greengrass
 - Wireless connectivity
 - Secure
 - Defend
- ▼ Act
- Rules**
 - Destinations
- ▼ Test
- Device Advisor
 - MQTT test client



You don't have any rules yet

Rules give your things the ability to interact with AWS and other web services. Rules are analyzed and actions are performed based on the messages sent by your things.

[Learn more](#)

[Create a rule](#)

Introducing the new AWS IoT console experience. We're updating the console experience for you. Learn more. Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

Rule query statement
Indicate the source of the messages you want to process with this rule.
Using SQL version
2016-05-25

Rule query statement
SELECT <Attribute> FROM <Topic Filter> WHERE <Condition>. For example: SELECT temperature FROM 'iot/topic' WHERE temperature > 50. To learn more, see AWS IoT SQL Reference.

```
1 SELECT * FROM 'iot/topic'
```

Set one or more actions
Select one or more actions to happen when the above rule is matched by an inbound message. Actions define additional activities that occur when messages arrive, like storing them in a database, invoking cloud functions, or sending notifications. (* required)

Add action

Error action
Optionally set an action that will be executed when something goes wrong with processing your rule.

Add action

Tags

AWS IoT

Monitor Activity Onboard Manage Fleet Hub Greengrass Wireless connectivity Secure Device Defend Art Roles Destinations Test Device Advisor MQTT test client Software Settings Learn Feature spotlight Documentation [?]

New console experience Tell us what you think

Search for services, features, marketplace products, and docs [Option+I]

Introducing the new AWS IoT console experience. We're updating the console experience for you. Learn more Try the new experiences and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > Rules >

Select an action

Select an action:

-  Insert a message into a DynamoDB table DYNAMODB
-  Split message into multiple columns of a DynamoDB table (DynamoDBv2) DYNAMODBV2
-  Send a message to a Lambda function LAMBDA
-  Send a message as an SNS push notification SNS
-  Send a message to an SQS queue SQS
-  Send a message to an Apache Kafka cluster APACHE KAFKA
-  Send a message to an Amazon Kinesis Stream AMAZON KINESIS
-  Republish a message to an AWS IoT topic AWS IoT Republisher

Monitor

Activity

Onboard

Manage

Fleet Hub

Greengrass

Wireless connectivity

Secure

Defend

Alerts

Roles

Destinations

Test

Device Advisor

MQTT test client

Software

Settings

Learn

Feature spotlight

Documentation

New console experience
Tell us what you think

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream AMAZON-KINESIS-FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name Choose a resource

Separator (?) No separator

Batch mode (?)

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

Cancel

-  Send a message to the Amazon Elasticsearch Service
AMAZON ELASTICSEARCH
-  Send a message to a Salesforce IoT Input Stream
SALESFORCE IoT
-  Send a message to IoT Analytics
IOT ANALYTICS
-  Send a message to an IoT Events Input
IOT EVENTS
-  Send message data to asset properties in AWS IoT SiteWise
IOT SITESWISE
-  Start a Step Functions state machine execution
STEP FUNCTIONS
-  Send a message to a downstream HTTPS endpoint
LAMBDA
-  Write a message into a Timestream table
TIMESTREAM

Cancel

Configure action 

AWS IoT

Introducing the new AWS IoT console experience
We're updating the console experience for you. Learn more Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name: Choose a resource Create a new resource

Separator: No separator

Batch mode:

Choose or create a role to grant AWS IoT access to perform this action.

No role selected Select

Add action

New console experience Tell us what you think

The screenshot displays the AWS IoT 'Configure action' configuration page. The left sidebar shows the navigation menu with 'Rules' highlighted. The main area has a blue header 'Configure action'. Below it, a section for 'Send a message to an Amazon Kinesis Firehose stream' is shown. It includes fields for 'Stream name' (with a dropdown for 'Choose a resource' and a 'Create a new resource' button), 'Separator' (set to 'No separator'), and 'Batch mode' (with a checkbox). A note says 'This will send the message to an Amazon Kinesis Firehose stream.' Below these, a section for 'Choose or create a role to grant AWS IoT access to perform this action.' is present, showing 'No role selected' and a 'Select' button. At the bottom right is an 'Add action' button.

Amazon Kinesis Data Firehose

Welcome to Amazon Kinesis Data Firehose

Amazon Kinesis Data Firehose is a fully managed, elastic service to easily deliver real-time data streams to destinations such as Amazon S3 and Amazon Redshift. You can start using Kinesis Data Firehose by:

1. Creating a delivery stream.
2. Sending your data to your delivery stream via Kinesis Agent or Kinesis Data Firehose APIs.

Data will be automatically delivered to your specified destination.

[Create Delivery Stream](#)

Kinesis Data Firehose Benefits



Easy to Use
Capture and deliver streaming data into destinations without writing any application or managing any infrastructure.



Direct to Data Stores
Batch, compress, and encrypt streaming data for delivery into your S3 bucket or Redshift cluster in as little as thirty seconds.



Zero Maintenance
Scale elastically to handle spikes in streaming data without intervention. Monitor the metrics for streaming data flowing into destinations.

Step 1: Name and source

- Step 2: Process records
- Step 3: Choose a destination
- Step 4: Configure settings
- Step 5: Review

New delivery stream

Delivery streams automatically and continuously load data to the specified destinations. Kinesis Data Firehose resources are not covered under the AWS Free Tier [?], and usage-based charges apply. For more information, see [Kinesis Data Firehose pricing](#) [?].

Delivery stream name

Acceptable characters are uppercase and lowercase letters, numbers, underscores, hyphens, and periods.

Choose a source

Choose how you would prefer to send records to the delivery stream.



Source

To learn about enabling server-side encryption (SSE), see [Data Protection in Amazon Kinesis Data Firehose](#) [?].

Direct PUT or other sources

Choose this option to send records directly to the delivery stream, or to send records from AWS IoT, CloudWatch Logs, or CloudWatch Events.

Kinesis Data Stream

Server-side encryption for source records in the delivery stream

Sales

Services ▾

Step 4: Configure settings

Step 5: Review

Delivery stream name: octank_deliverystream

Acceptable characters are uppercase and lowercase letters, numbers, underscores, hyphens, and periods.

Choose a source

Choose how you would prefer to send records to the delivery stream:

Firehose data flow overview

```
graph LR; Source[Source] --> Firehose[Firehose delivery stream]; Firehose --> Destination[Destination]
```

Source

To learn about enabling server-side encryption (SSE), see [Data Protection in Amazon Kinesis Data Firehose](#).

Direct PUT or other sources

Choose this option to send records directly to the delivery stream, or to send records from AWS IoT, CloudWatch Logs, or CloudWatch Events.

Kinesis Data Stream

Server-side encryption for source records in the delivery stream

You can use AWS Key Management Service (KMS) to create and manage customer keys (CMKs) and to control the use of encryption across a wide range of AWS services and in your applications.

Enable server-side encryption for source records in delivery stream

▶ [How to send source records to Kinesis Data Firehose](#)



Kinesis Data Firehose - Create delivery stream

Step 1: Name and source
Step 2: Process records
Step 3: Choose a destination
Step 4: Configure settings
Step 5: Review

Choose a destination

[Learn more \[F\]](#)

Destination

Amazon S3

Object storage built to store and retrieve any amount of data from anywhere.

Amazon Pinpoint

An enterprise-grade, probabilistic-aware, fully managed data warehousing service.

Amazon Elasticsearch

An open-source search and analytics engine for use cases such as log analysis, real-time application monitoring, and time-series analysis.

HTTP Endpoint

A way to extract data to your custom destinations.

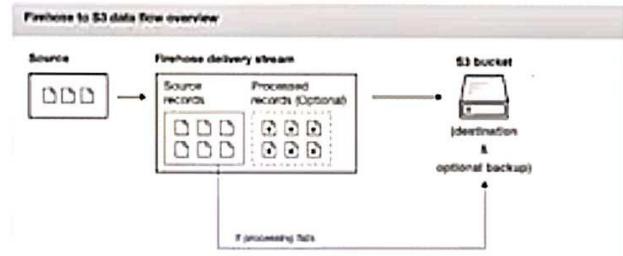
Third-party service provider

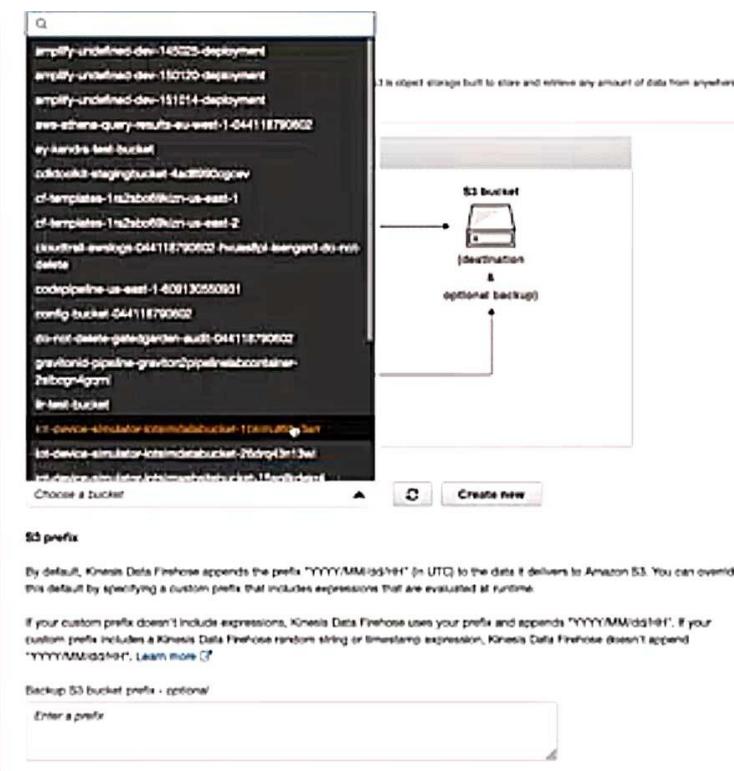
Choose from a list of third-party service providers.

S3 destination

Create a destination in Amazon S3 where your data will be stored. Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

[Learn more \[F\]](#)





S3 destination

Choose a destination in Amazon S3 where your data will be stored. Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

[Learn more](#)

Firehose to S3 data flow overview

```
graph LR; Source[Source] --> Firehose[Firehose delivery stream]; Firehose --> S3[S3 bucket]; S3 -- optional backup --> S3
```

The diagram illustrates the data flow from a source to a Firehose delivery stream, which then sends data to an S3 bucket. An optional backup path is shown from the S3 bucket back to itself.

S3 bucket

octant-database1

[View octant-database1 in S3 console](#)

S3 prefix

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

If your custom prefix doesn't include expressions, Kinesis Data Firehose uses your prefix and appends "YYYY/MM/dd/HH". If your custom prefix includes a Kinesis Data Firehose random string or timestamp expression, Kinesis Data Firehose doesn't append "YYYY/MM/dd/HH". [Learn more](#)

Backup S3 bucket prefix - optional

Enter a prefix

S3 destination

Choose a destination in Amazon S3 where your data will be stored. Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

[Learn more](#)

Firehose to S3 data flow overview

```
graph LR; Source[Source] --> Firehose[Firehose delivery stream]; Firehose --> S3[S3 bucket]; S3 -- optional backup --> S3
```

The diagram illustrates the data flow from a source to a Firehose delivery stream and finally to an S3 bucket. The Firehose delivery stream contains two sections: 'Source records' and 'Processed records (optional)'. An arrow points from the Firehose delivery stream to the S3 bucket, which is labeled '(destination & optional backup)'.

S3 bucket

octank-database1 [View octank-database1 in S3 console](#)

S3 prefix

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

If your custom prefix doesn't include expressions, Kinesis Data Firehose uses your prefix and appends "YYYY/MM/dd/HH". If your custom prefix includes a Kinesis Data Firehose random string or timestamp expression, Kinesis Data Firehose doesn't append "YYYY/MM/dd/HH". [Learn more](#)

Backup S3 bucket prefix - optional

Enter a prefix

Services ▾

Third-party service provider
Choose from a list of third-party service providers.

Amazon S3

Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

S3 destination

Create a destination in Amazon S3 where your data will be stored. Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

Learn more ↗

Firehose to S3 data flow overview

Source → Firehose delivery stream → S3 bucket

Source: CloudWatch Metrics CloudWatch Logs CloudWatch Metrics and CloudWatch Logs AWS Lambda AWS Kinesis AWS Kinesis Data Firehose AWS Lambda and AWS Kinesis Data Firehose

Firehose delivery stream:

Source records → Processed records (Optimized)

S3 bucket:

Destination

Create S3 bucket

A bucket is a container for objects stored in Amazon S3. Learn more

S3 bucket name:

Region: EU (Ireland)

S3 prefix:
octank-database1

View octank-database1 in S3 console ↗

Enter a prefix

Services ▾

Search for services, features, marketplace products, and docs [Optimize ▾]

Amazon Kinesis Data Firehose | Documentation | Support

S3 destination

Choose a destination in Amazon S3 where your data will be stored. Amazon S3 is object storage built to store and retrieve any amount of data from anywhere.

[Learn more](#) [?]

Firehose to S3 data flow overview

```
graph LR; Source[Source] --> FDS[Firehose delivery stream]; FDS --> S3[S3 bucket]; S3 --> FDS; FDS -- "If processing fails" --> S3;
```

The diagram illustrates the data flow from a source to a Firehose delivery stream and then to an S3 bucket. The Firehose delivery stream contains two sections: 'Source records' (represented by a box with four icons) and 'Processed records (Optional)' (represented by a dashed box with four icons). An arrow points from the source to the Firehose delivery stream. Another arrow points from the Firehose delivery stream to the S3 bucket. A feedback loop arrow points from the S3 bucket back to the Firehose delivery stream. A callout box labeled 'If processing fails' indicates that data can be re-submitted to the Firehose delivery stream from the S3 bucket.

S3 bucket

octank-datasink1

[View octank-datasink1 in S3 console](#) [?]

S3 prefix

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

If your custom prefix doesn't include expressions, Kinesis Data Firehose uses your prefix and appends "YYYY/MM/dd/HH". If your custom prefix includes a Kinesis Data Firehose random string or timestamp expression, Kinesis Data Firehose doesn't append "YYYY/MM/dd/HH". [Learn more](#) [?]

Backup S3 bucket prefix - optional

Enter a prefix

Processing tasks

S3 bucket: **octank-database1** View [octank-database1](#) in S3 console (?)

S3 prefix: By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

If your custom prefix doesn't include expressions, Kinesis Data Firehose uses your prefix and appends "YYYY/MM/dd/HH". If your custom prefix includes a Kinesis Data Firehose random string or timestamp expression, Kinesis Data Firehose doesn't append "YYYY/MM/dd/HH". Learn more (?)

Backup S3 bucket prefix - optional:

S3 error prefix: You can specify an S3 bucket prefix to be used in error conditions. This prefix can include expressions for Kinesis Data Firehose to evaluate at runtime. Learn more about the rules for specifying prefix expressions (?)

Backup S3 bucket error prefix - optional:



Kinesis Data Firehose - Create delivery stream

Step 1: Name and source

Step 2: Process records

Step 3: Choose a destination

Configure settings

Configure buffer, compression, logging, and IAM role settings for your delivery stream. [Learn more](#) [?]

S3 buffer conditions

Kinesis Data Firehose buffers incoming records before delivering them to your S3 bucket. Record delivery will be triggered once either of these conditions has been satisfied. [Learn more](#) [?]

Buffer size

MB

Enter a buffer size between 1 - 100 MB.

Buffer interval

seconds

Enter a buffer interval between 60 - 800 seconds.

S3 compression and encryption

Kinesis Data Firehose can compress records before delivering them to your S3 bucket. Compressed records can also be encrypted in the S3 bucket using a KMS master key. [Learn more](#) [?]

S3 compression

- Disabled
- GZIP
- Snappy
- Zstd
- Hadoop-Compatible Snappy

S3 encryption

- Disabled
- Enabled

Salesforce Services ▾

Search for services, features, marketplace products, and docs [Options+1]

Enabled

Error logging

Kinesis Data Firehose data log record delivery errors to CloudWatch Logs. If enabled, a CloudWatch log group and corresponding log streams are created on your behalf. [Learn more](#)

Enabled

Disabled

Tags - optional

You can add tags to organize your AWS resources, track costs, and control access. [Learn more](#)

Key Value (optional)

Enter key Enter value

You can add 40 more tag(s).

Permissions

IAM role

Kinesis Data Firehose uses this IAM role for all the permissions that the delivery stream needs. To specify different roles for the different permissions, use the API or the CLI. [Learn more](#)

Create or update IAM role **KinesisFirehoseServiceRole-octant_delivery-eo-west-1-1629182097571**

This creates the role or updates it if it already exists, adds the required policies to it, and enables Kinesis Data Firehose to assume it.

Choose existing IAM role

The role you choose must have policies that include the permissions that Kinesis Data Firehose needs.

Kinesis Data Firehose - Create delivery stream

Step 1: Name and source Step 2: Process records Step 3: Choose a destination Step 4: Configure settings Step 5: Review

Review

Review your configuration details before creating your delivery stream.

Name and source

Delivery stream name: octank_deliverystream Edit

Source: Direct PUT or other sources Edit

Server-side encryption for source records: Disabled

Process records

Source record transformation: Disabled Edit

Record format conversion: Disabled Edit

Destination

Destination: Amazon S3 Edit

S3 bucket: octank-datasink1 Edit

S3 bucket Prefix:

Sales

Services ▾

Search for services, features, marketplace products, and docs [Options+1]

Amazon CloudWatch Metrics > Metrics > Metrics Stream > Create delivery stream

Ireland Support

Destination

Destination: Amazon S3

S3 bucket: octane-datasink-1

S3 bucket Prefix: -

S3 bucket error prefix: -

Settings

S3 buffer conditions: 5 MB or 300 seconds

S3 compression: Disabled

S3 encryption: Disabled

Error logging: Enabled

Tags: no tags specified

IAM role: KinesisInboxServiceRole-octane_delivery-eu-west-1-1629193087571

Cancel Previous **Create delivery stream**

Amazon Kinesis

Kinesis Data Firehose delivery streams continuously collect, transform, and load streaming data into the destinations that you specify.

Creating delivery stream octank_deliverystream
It can take up to a minute before the status is updated.

Find delivery streams

Create delivery stream

Name	Status	Creation time	Source	Data transformation	Destination
octank_deliverystream	Creating	2021-08-17T11:39:02Z	Direct PUT and other sources	Disabled	Amazon S3 octank-database1 [7]

Services ▾

Search for services, features, monitoring, products, and docs [Option+Shift]

Kinesis Data Firehose delivery streams continuously collect, transform, and load streaming data into the destinations that you specify.

Amazon Kinesis

Dashboard

Data Streams

Data Firehose

Data Analytics

Video Streams

External resources

What's new

Kinesis Data Firehose delivery streams

Creating delivery stream octane_deliverystream. It can take up to a minute before the status is updated.

Find delivery streams

Create delivery stream

Name	Status	Creation Time	Source	Data Transformation	Destination
Loading delivery streams					

Amazon Kinesis

Kinesis Data Firehose delivery streams continuously collect, transform, and load streaming data into the destinations that you specify.

Creating delivery stream octank_deliverystream.
It can take up to a minute before the status is updated.

Find delivery streams

Name Status Creation Time Source Data transformation Destination

Name	Status	Creation Time	Source	Data transformation	Destination
octank_deliverystream	Creating	2021-08-17T11:39:02Z	Direct PUT and other sources	Disabled	Amazon S3 octank-database1 [7]

Amazon Kinesis

Kinesis Data Firehose delivery streams continuously collect, transform, and load streaming data into the destinations that you specify.

Success! You created delivery stream octane_deliverystream.

Next, send records directly to the delivery stream using the Amazon Kinesis Agent or the Kinesis Data Firehose API using the AWS SDK, or send records from AWS IoT, CloudWatch Logs, or CloudWatch Events. Learn more.

Find delivery streams

Name Status Creation time Source Data transformation Destination

Name	Status	Creation time	Source	Data transformation	Destination
octane_deliverystream	Active	2021-08-17T11:29:02Z	Direct PUT and other sources	Disabled	Amazon S3 octane-database1 [7]

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream AMAZON KINESIS FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name

Separator

Batch mode

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

Select

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream AMAZON-KINESIS-FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name:

  Create a new resource

Separator 



Batch mode: 

Choose or create a role to grant AWS IoT access to perform this action.

 Select

 Cancel

Add action 

AWS IoT

Search for services, features, marketplace products, and docs [Option+I]

Introducing the new AWS IoT console experience
We're updating the console experience for you. Learn more [?] Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream
AMAZON KINESIS FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name: octank_drivenstream

Separator:

Batch mode:

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

New console experience
Tell us what you think

AWS IoT

Introducing the new AWS IoT console experience
We're updating the console experience for you. Learn more Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > Rules >

Configure action

 Send a message to an Amazon Kinesis Firehose stream
AMAZON KINESIS FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name: octank_deliverystream

Separator: No separator

Batch mode:

Choose or create a role to grant AWS IoT access to perform this action.

All role selected

New console experience Tell us what you think

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name

Separator 

Batch mode 

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

AWS IoT

Introducing the new AWS IoT console experience. We're updating the console experience for AWS IoT. [Learn more](#)

Monitor

Activity

Onboard

Manage

Fleet Hub

Greengrass

Wireless connectivity

Secure

Defend

Art

Roles

Destinations

Test

Device Advisor

MQTT test client

Software

Settings

Learn

Feature spotlight

Documentation

New console experience
Tell us what you think

Create a new role

Configure action

Send a message to an Amazon Kinesis Firehose

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name: octank_deliverystream

Separator: No separator

Batch mode:

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

Cancel

The screenshot shows the AWS IoT console interface. A modal window titled "Create a new role" is open over the main "Configure action" screen. The main screen shows a step to "Send a message to an Amazon Kinesis Firehose". The configuration details include a stream name "octank_deliverystream" and a separator setting. Below this, there's a section to choose or create a role to grant AWS IoT access. The "No role selected" button is highlighted. At the bottom right of the modal, there are "Create Role" and "Select" buttons, and at the bottom center is a large "Create Role" button.

AWS IoT

Introducing the new AWS IoT console experience
We're updating the console experience for you. [Leave more](#) [?] try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream

AMAZON KINESIS FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name: octank_deliverystream

Separator:

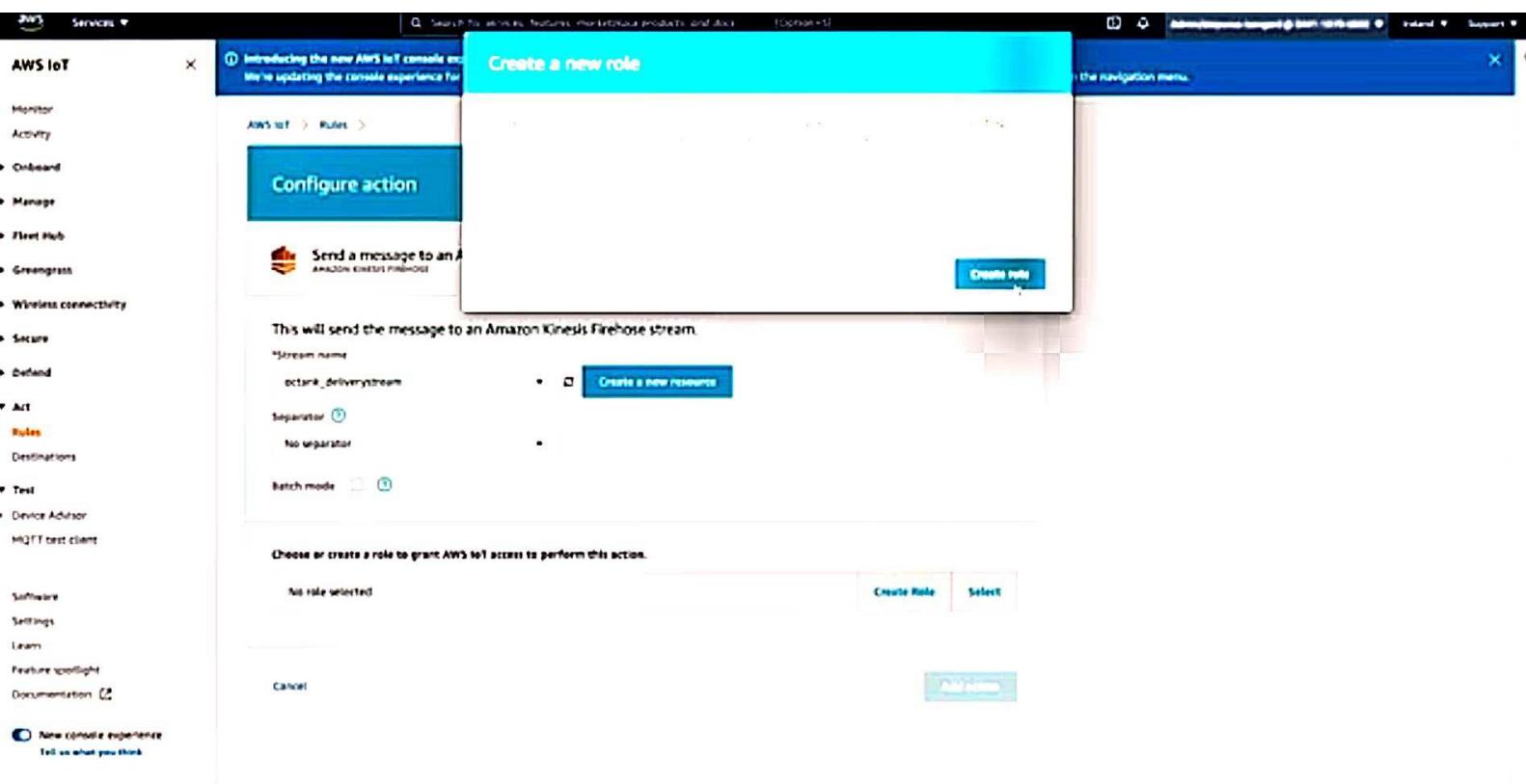
Batch mode:

Choose or create a role to grant AWS IoT access to perform this action.

No role selected

Cancel

New console experience [Tell us what you think](#)



AWS IoT > Rules >

Configure action

Send a message to an Amazon Kinesis Firehose stream AMAZON KINESIS FIREHOSE

This will send the message to an Amazon Kinesis Firehose stream.

*Stream name

octank_deliverystream

[Create a new resource](#)

Separator

No separator

Batch mode

Choose or create a role to grant AWS IoT access to perform this action.

[awsiotfirehose-role](#) Policy Attached

[Create Role](#)

[Select](#)

[Cancel](#)

[Next step](#)

Set one or more actions

Select one or more actions to happen when the above rule is matched by an inbound message. Actions define additional activities that occur when messages arrive, like storing them in a database, invoking cloud functions, or sending notifications. (* required)



Send a message to an Amazon Kinesis Firehose stream
aws_kinesis_delivery

Remove Edit

[Add action](#)

Error action

Optionally set an action that will be executed when something goes wrong with processing your rule.

[Add action](#)

Tags

Apply tags to your resources to help organize and identify them. A tag consists of a case-sensitive key-value pair. [Learn more](#) about tagging your AWS resources.

Tag name

Value

[Clear](#)

[Add another](#)

[Cancel](#)

[Create rule](#)

AWS IoT Services Search for services, features, marketplace products, and docs [Option+I]

Introducing the new AWS IoT console experience. We're updating the console experience for you. Learn more [?] Try the new experience and let us know what you think. You can turn off the new experience from the navigation menu.

Success Successfully created rule.

AWS IoT Rules Create

Rules

Name	Status
OctankAnalyticsRule	ENABLED

Search rules

New console experience Tell us what you think

AWS IoT Services Monitor Activity Onboard Manage Fleet Hub Greengrass Wireless connectivity Secure Defend Roles Destinations Test Device Advisor MQTT test client Software Settings Learn Feature spotlight Documentation

Screenshot of the AWS IoT Rule Hub console search results for "S3".

The search results are categorized into three sections:

- Services (7)**:
 - S3 (Scalable Storage in the Cloud)
 - S3 Glacier (Archive Storage in the Cloud)
 - Athena (Query Data in S3 using SQL)
 - AWS Snow Family (Large Scale Data Transport)
- Features (10)**:
 - Amazon S3 File Gateway (Storage Gateway feature)
 - Datasets (IoT Analytics feature)
 - Batch Operations (S3 feature)
 - Buckets (S3 feature)
- Documentation (401,673)**: See all 401,673 results in Documentation.

The interface includes a sidebar with navigation links like Monitor, Activity, Onboard, Manage, Fleet Hub, Greengrass, Wireless connectivity, Secure, Behind, Rules, Destinations, Test, Device Advisor, MQTT test client, Software, Settings, Learn, Feature spotlight, Documentation, and a feedback link for the new console experience.

Amazon S3

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, please provide feedback.

Provide Feedback

Amazon S3

Amazon S3

Account snapshot

Storage lens provides visibility into storage usage and activity trends. Learn more

View Storage Lens dashboard

Buckets

Access Points

Object Lambda Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

Storage Lens

Dashboards

Alerts, Organizations settings

Feature spotlight

AWS Marketplace for S3

Buckets

Buckets are containers for data stored in S3. Learn more

Search buckets by name

Name	AWS Region	Access	Creation date
Leading buckets			

Amazon S3

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, click [Provide Feedback](#).

Amazon S3

Amazon S3

Account snapshot

Buckets [24]

Name	AWS Region	Access	Creation date
amplify-undefined-dev-245211-deployment	US East (N. Virginia) us-east-1		December 31, 2020, 14:50:33 UTC+01:00
amplify-undefined-dev-150120-deployment	US East (N. Virginia) us-east-1		December 31, 2020, 15:01:28 UTC+01:00
amplify-undefined-dev-151014-deployment	US East (N. Virginia) us-east-1		December 31, 2020, 15:10:21 UTC+01:00
aws.athena-query-results-eu-west-1-044118790602	EU (Ireland) eu-west-1		May 5, 2021, 18:24:32 UTC+02:00
ay-kendra-test-bucket	EU (Ireland) eu-west-1		June 28, 2021, 16:27:58 UTC+02:00
cdkstack8-stagingbucket-4ad899094eav	EU (Ireland) eu-west-1		July 5, 2021, 12:42:11 UTC+02:00
cf-templates-1fa2bb67942m-us-east-1	US East (N. Virginia) us-east-1		November 11, 2020, 11:52:10 UTC+01:00
cf-templates-1fa2bb67942m-us-east-2	US East (Ohio) us-east-2		December 5, 2020, 14:22:19 UTC+01:00
cloudtrail-auditlog-044118790602-fvwsdpl-bergard-do-not-delete	US East (N. Virginia) us-east-1		October 20, 2020, 19:01:47 UTC+02:00
compliance-cis-east-1-609130550931	US East (N. Virginia) us-east-1		December 15, 2020, 14:47:55 UTC+01:00
config-bucket-044118790602	US East (N. Virginia) us-east-1		January 25, 2021, 17:02:01 UTC+01:00
do-not-delete-patentedgarden-audit-044118790602	US West (Oregon) us-west-2		October 27, 2020, 01:29:06 UTC+01:00
gravitonid-pipeline-graviton2/pipelineLabContainer-2zbogn4gpm	EU (Ireland) eu-west-1		July 5, 2021, 12:51:51 UTC+02:00
fr-test-bucket	EU (Ireland) eu-west-1		August 2, 2021, 17:46:38 UTC+02:00
iot-device-simulator-lonemudatabucket-151005605am	US East (N. Virginia) us-east-1		February 2, 2021, 17:13:28 UTC+01:00
iot-device-simulator-lonemudatabucket-26drq4tn15r	EU (Ireland) eu-west-1		February 2, 2021, 17:21:15 UTC+01:00
iot-device-simulator-lonewebdatabucket-1fba0j4kx4nt	US East (N. Virginia) us-east-1		February 2, 2021, 17:13:27 UTC+01:00

Amazon S3

Services ▾

Amazon S3 > octank-datalake1

octank-datalake1

Objects Properties Permissions Metrics Management Access Points

Objects [1]

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permission. [Learn more](#)

Action Copy URL Copy URL Download Open Delete Actions Create Folder Upload

Find objects by prefix:

Name	Type	Last modified	Size	Storage class
2021/	Folder	-	-	-

Actions ▾

Provide Feedback

Block Public Access settings for this account

Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight

AWS Marketplace for S3

The screenshot shows the AWS S3 console interface. On the left, there's a sidebar with various navigation links like 'Buckets', 'Access Points', 'Object Lambda Access Points', etc. The main area shows a single object named '2021/' in the 'octank-datalake1' bucket. The object is listed as a folder. At the top, there's a blue bar with a message about improving the console and a 'Provide Feedback' button. Below the bar, there are tabs for 'Objects', 'Properties', 'Permissions', 'Metrics', 'Management', and 'Access Points'. The 'Objects' tab is selected. The main content area shows a table with columns for Name, Type, Last modified, Size, and Storage class. The single object '2021/' is listed under the 'Name' column.

Amazon S3

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose Provide feedback.

Provide Feedback X

Bucket

Access Points

Object Lambda Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

▼ Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight 3

▶ AWS Marketplace for S3

Amazon S3 > octank-datalake1

octank-datalake1

Objects Properties Permissions Metrics Management Access Points

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Actions Create folder

Name	Type	Last modified	Size	Storage class
2021/	Folder			

Amazon S3

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose [Provide feedback](#).

Provide feedback

Amazon S3 > octank-datalake1 > 2021/

2021/

Objects Properties

Objects (2)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Copy S3 URL Copy URL Download Open Delete Actions Create folder Upload

Find objects by prefix

Name	Type	Last modified	Size	Storage class
o1/	Folder	-	-	-
o2/	Folder	-	-	-

https://s3.console.aws.amazon.com/buckets/octank-datalake?region=eu-west-1&prefix=2021/08/17/09&listType=detail

Amazon S3 Services ▾

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose Provide feedback.

Provide feedback

Buckets

Amazon S3 > octank-datalake1 > 2021/ > 08/ > 17/ > 09/

09/

Copy S3 URL

Objects Properties

Objects (4)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 Inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

Actions Create folder Upload

Find objects by prefix

Name	Type	Last modified	Size	Storage class
octank_deliverystream-1-2021-08-17-09-14-62ec7053-a7e2-4f4e-8390-f269ef755471	Object	August 17, 2021, 11:18:17 (UTC-02:00)	8.6 KB	Standard
octank_deliverystream-1-2021-08-17-09-14-90fb4ad-0276-4f98-a7e2-c140c671268	Object	August 17, 2021, 11:21:17 (UTC-02:00)	8.7 KB	Standard
octank_deliverystream-1-2021-08-17-09-23-40-c7572a51-4f46-4e84-9c79-977b91af2044	Object	August 17, 2021, 11:28:42 (UTC-02:00)	8.6 KB	Standard
octank_deliverystream-1-2021-08-17-09-28-41-4b416bd-8184-457d-4734-4b0141758409	Object	August 17, 2021, 11:29:42 (UTC-02:00)	1.4 KB	Standard

Storage Lens

Dashboard

AWS Organizations settings

Feature spotlight

AWS Marketplace for S3

Feedback [Provide feedback](#)

[Copy 53 links](#)

objects; you'll need to explicitly grant them permission. Learn more [\[?\]](#)

[Folder](#) [Upload](#)

< 1 > [\[?\]](#)

Last modified	Size	Storage class
August 17, 2021, 11:18:37 (UTC+02:00)	8.6 KB	Standard
August 17, 2021, 11:21:17 (UTC+02:00)	8.7 KB	Standard
August 17, 2021, 11:28:42 (UTC+02:00)	8.6 KB	Standard
August 17, 2021, 11:29:42 (UTC+02:00)	1.4 KB	Standard

https://s3.console.aws.amazon.com/buckets/sectorbank-datalake?region=eu-west-1&prefix=Athena/

Q. Athena

Feedback. Provide feedback

Amazon S3 Services

Services (1) Documentation (10,971) Knowledge Articles (30) Marketplace (84)

Athena Query data in S3 using SQL

Search results for Athena

Services

Athena

Documentation (10,971) See all 10,971 results in Documentation

Identity and Access Management In Athena - Amazon Athena User Guide

Resilience in Athena - Amazon Athena User Guide

Troubleshooting in Athena - Amazon Athena User Guide

Monitoring Athena Queries with CloudWatch Events - Amazon Athena User Guide

Documentation (10,971) See all 10,971 results in Documentation

Knowledge Articles

Resilience in Athena - Amazon Athena User Guide

What Is Amazon Athena? - Amazon Athena User Guide

When should I use Athena? - Amazon Athena User Guide

Copy S3 URI

Upload

Last modified	Size	Storage class
August 17, 2021, 11:18:37 (UTC+02:00)	8.6 kB	Standard
August 17, 2021, 11:21:37 (UTC+02:00)	8.7 kB	Standard
August 17, 2021, 11:28:42 (UTC+02:00)	8.6 kB	Standard
August 17, 2021, 11:29:42 (UTC+02:00)	1.4 kB	Standard

Athena [Query editor](#) [Saved queries](#) [History](#) [Data sources](#) [Workgroup primary](#)

Data source [Connect data source](#)

Amazon Data Catalog

Database

Filter tables and views...

New query 1

Run query Save as Create

Ctrl + Enter to run query, Ctrl + Space to autocomplete

Format query Clear

Athena engine version 2 Release version: 1.0

Results

Amazon S3 Services

Feedback

Provide feedback

Copy S3 URL

Search results for 'Athena'

Services (1)

Athena

Query data in S3 using SQL

Documentation (30,971)

Identity and Access Management in Athena - Amazon Athena

User Guide

Resilience in Athena - Amazon Athena

User Guide

Troubleshooting in Athena - Amazon Athena

User Guide

Monitoring Athena Queries with CloudWatch Events - Amazon Athena

User Guide

Knowledge Articles (30)

Resilience in Athena - Amazon Athena

Amazon Athena

What is Amazon Athena? - Amazon Athena

Amazon Athena

When should I use Athena? - Amazon Athena

Amazon Athena

Last modified

Size

Storage class

Upload

Last modified	Size	Storage class
August 17, 2021, 11:18:37 (UTC-02:00)	8.6 kB	Standard
August 17, 2021, 11:21:17 (UTC-02:00)	8.7 kB	Standard
August 17, 2021, 11:28:42 (UTC-02:00)	8.6 kB	Standard
August 17, 2021, 11:29:42 (UTC-02:00)	14 kB	Standard

Athena Query editor Saved queries History Data sources Workgroup: primary

New query 1

Run query Save as Create -

Use Ctrl + Enter to run query, Ctrl + Space to autocomplete

Athena engine version 2 Release versions (7)

Data source: AvnDataCatalog Connect data source

Database: Default

Tables (0) Create table

The selected database has no tables.

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

Results

The screenshot shows the AWS Athena Query Editor. On the left, there's a sidebar with 'Data source' set to 'AvnDataCatalog' and 'Database' set to 'Default'. Below this are sections for 'Tables (0)' and 'Views (0)', both of which are currently empty. The main workspace is titled 'New query 1' and contains a single line of code: '1'. At the bottom of the workspace, there are buttons for 'Run query', 'Save as', and 'Create -'. A note below the workspace says 'Use Ctrl + Enter to run query, Ctrl + Space to autocomplete'. In the bottom right corner, it says 'Athena engine version 2' and 'Release versions (7)'. The top of the screen has a standard browser-style header with tabs like 'File', 'Edit', 'View', etc., and a status bar at the bottom right with 'Settings', 'Tutorial', 'Help', and 'What's new'.

Athena Query editor Saved queries History Data sources Workgroup: primary [Options+I]

Data source: AmazonDataCatalog Connect data source

Database: default

Filter tables and views

Tables (0) Create table

The selected database has no tables.

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1 +

Run query Save as Create -

Use Ctrl + Enter to run query, Ctrl + Space to autocomplete

Format query Clear

Athena Engine version 2 - Release version: 1.0

Results

This screenshot shows the AWS Athena Query Editor. On the left, there's a sidebar with 'Data source' set to 'AmazonDataCatalog', 'Database' set to 'default', and a note that the selected database has no tables or views. The main area is titled 'New query 1' and contains three buttons: 'Run query', 'Save as', and 'Create -'. Below these buttons is a note: 'Use Ctrl + Enter to run query, Ctrl + Space to autocomplete'. In the bottom right corner, it says 'Athena Engine version 2 - Release version: 1.0'.

Connect data source

Step 1: Choose a data source

Choose where your data is located

Athena queries data where it is. Data is not loaded or moved. Learn more [?]

 Query data in Amazon S3

Choose an external data catalog

 Query a data source

Configure a connector for common data sources.



Choose a metadata catalog

The catalog contains the schema for the source data such as column names, data types and table names. Learn more [?]

AWS Glue Data Catalog

Apache Hive metastore

Cancel

Next

Connect data source**Step 1: Choose a data source**

Choose where your data is located

Athena queries data where it is. Data is not loaded or moved. [Learn more](#) ⓘ **Query data in Amazon S3**

Choose an external data catalog

 **Query a data source**

Configure a connector for common data sources.



Choose a metadata catalog

The catalog contains the schema for the source data such as column names, data types and table names. [Learn more](#) ⓘ AWS Glue Data Catalog Apache Hive metastore[Cancel](#)[Next](#) 

Connect data source

Step 1: Choose a data source

Step 2: Connection details

Connection details: choose an AWS Glue Data Catalog

Choose an AWS Glue Data Catalog in your account or in another account.

 AWS Glue Data Catalog in this account

Create a table in AWS Glue Data Catalog

 AWS Glue Data Catalog in another account

Register an external AWS Glue Data Catalog for cross account access.

Choose a way to create a table

Athena will connect to your data stored in Amazon S3 and use AWS Glue Data Catalog to store metadata, such as table and column names. Once connected, your databases, tables and views appear in Athena's query editor. [Learn more](#) Create a crawler in AWS Glue

Add a table by setting up a crawler in AWS Glue to analyse data and retrieve the schema automatically

 Create a table using the Athena table wizard

Add a table by entering schema information manually in Athena

Cancel

Previous

Connect to AWS Glue

- Crawler info
- Crawler source type
- Crawl rules
- IAM roles
- Schedule
- Output
- Preview API results

Add information about your crawler

Crawler name

Enter name

► Tags, description, security configuration, and classifiers (optional)

Next

- Crawler info
- Crawler source type
- Data store
- IAM Role
- Schedule
- Output
- Preview all steps

Add information about your crawler

Crawler name

octane-created

► Tags, description, security configuration, and classifiers (optional)

Next

- Crawler info
 - Amazon crawler
- Crawler source type
 - Data store
 - Amazon Redshift
 - Schedule
 - Output
 - Review all steps

Specify crawler source type

Choose Existing catalog tables to specify catalog tables as the crawler source. The selected tables specify the data stores to crawl. This option doesn't support JDBC data stores.

Crawler source type

- Data stores
- Existing catalog tables

Repeat crawls of S3 data stores

- Crawl all folders
 - Crawl all folders again with every subsequent crawl.
- Crawl new folders only
 - Only Amazon S3 folders that were added since the last crawl will be crawled. If the schemas are compatible, new partitions will be added to existing tables.

[Back](#)

[Next](#)

- Crawler info
- extant_crawler
- Crawler source type
- Data stores
- Data store
- DMU crawl
- Exclude
- Output
- Review all steps

Add a data store

Choose a data store

53

Connection

Select a connection

(Optional) Include a network connection to use with this S3 target. Note that each crawler is limited to one network connection so any future S3 targets will also use the same connection (or none, if left blank).

Add connection

Crawl data in

Specified path in my account
 Specified path in another account.

Include path

s3://DocumentCloud/objekt/

All folders and files contained in the include path are crawled. For example, type s3://MyBucket/MyFolder to crawl all elements in MyFolder within MyBucket.

Sample size (optional)

Enter an integer between 1 and 249

This limit sets the number of files in each and folder to be crawled. If not set, all the files are crawled.

Exclude patterns (optional)

Back Next

- Crawler info
- extdata_crawler
- Crawler source type
- Data stores
- Data store
- S3 bucket
- Redshift
- Output
- Review all steps

Add a data store

Choose a data store

S3

Connection

Select a connection

(Optional) Include a network connection to use with this S3 target. Note that each crawler is limited to one network connection so any future S3 targets will also use the same connection (or none, if left blank).

Add connection

Crawl data in

Specified path in my account
 Specified path in another account.

Include path

s3://bucketname/prefix/

All folders and files contained in the include path are crawled. For example, type s3://MyBucket/MyFolder to crawl all elements in MyFolder within MyBucket.

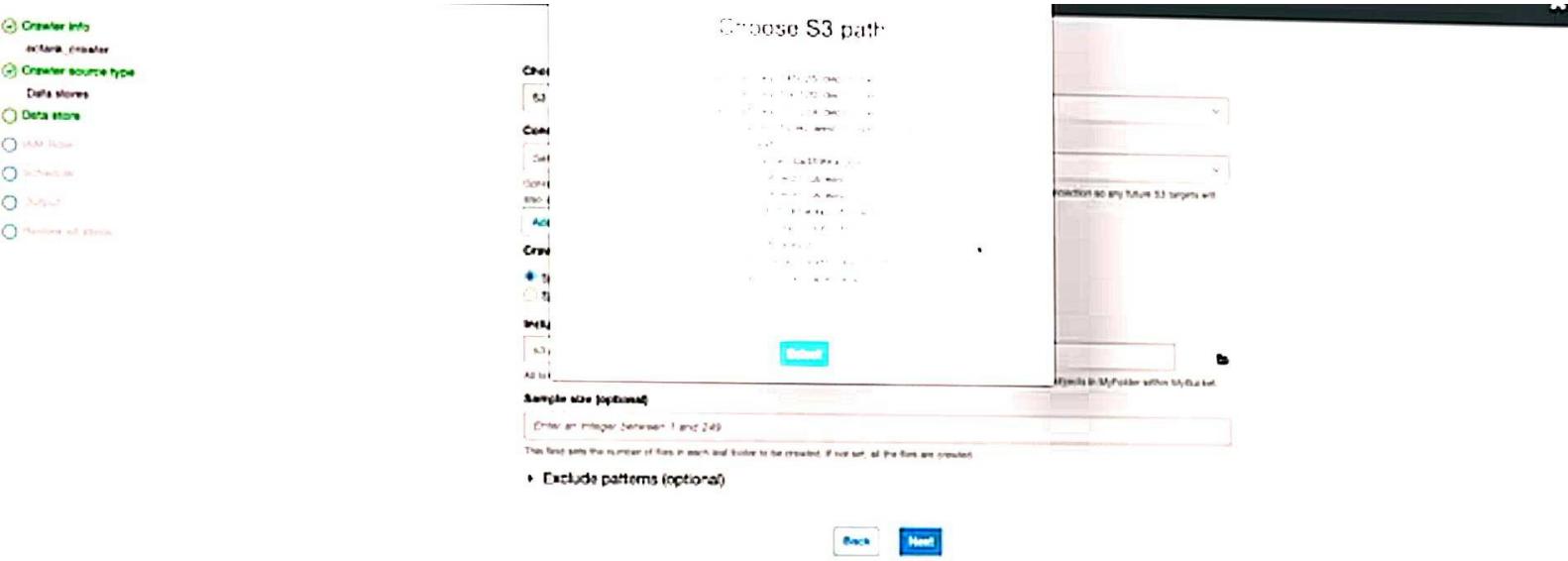
Sample size (optional)

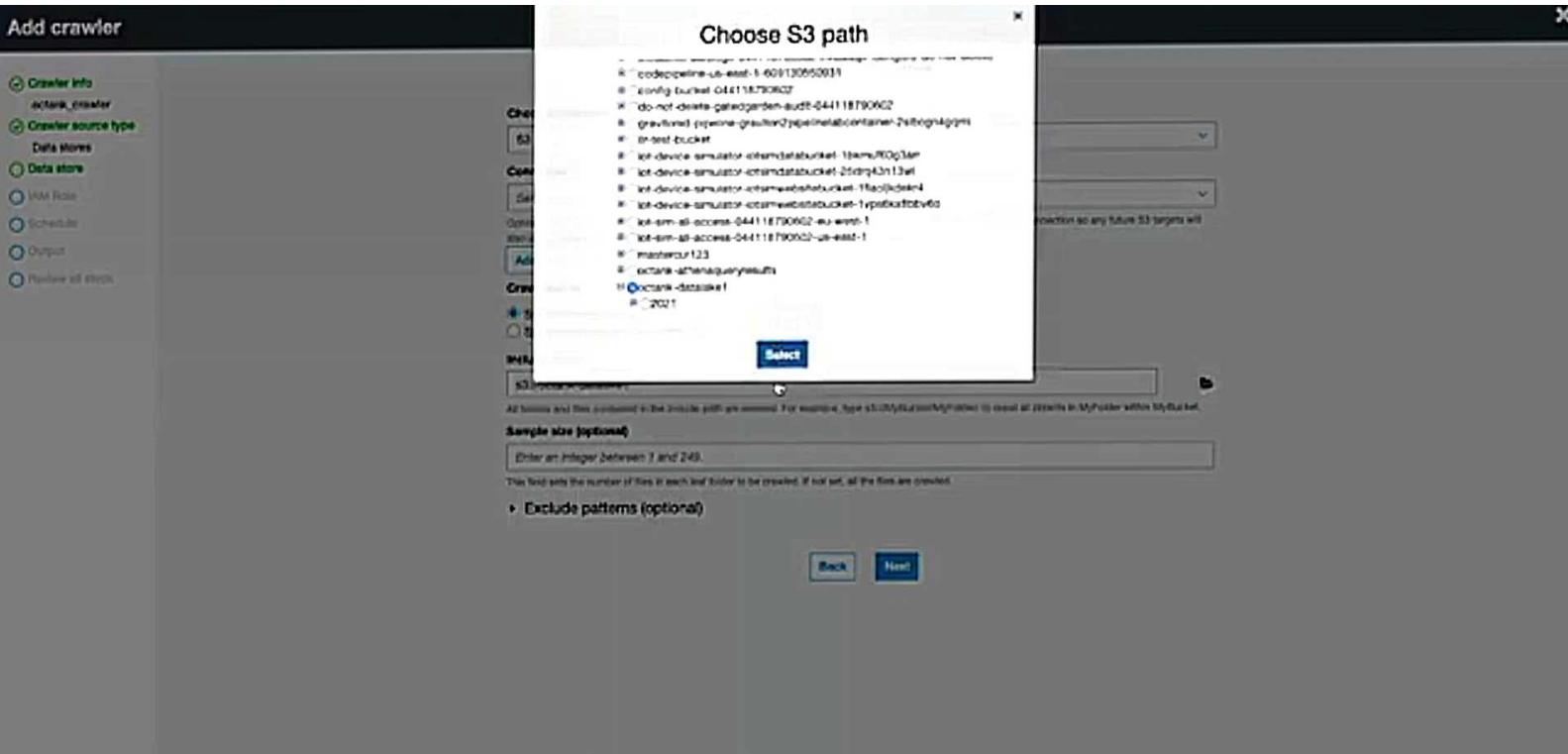
Enter an integer between 1 and 249

This field sets the number of files in each result to be returned. If not set, all the files are crawled.

Exclude patterns (optional)

Back Next





- Crawler info**
extana_crawler
- Crawler source type**
Data stores
- Data store**
S3: s3://extana-data...
Amazon Kinesis
Schedule
Output
Preview all steps

Add another data store

Yes
 No

[Back](#) [Next](#)

Chosen data stores
S3: s3://extana-data...

- Crawler info**
octank_crawler
- Crawler source type**
Data stores
- Data store**
S3: s3://octana-database1
- IAM Role**
- S3 buckets
- Output
- Review all stages

Choose an IAM role

The IAM role allows the crawler to run and access your Amazon S3 data stores. [Learn more](#)

- Update a policy in an IAM role
- Choose an existing IAM role
- Create an IAM role

IAM role

To create an IAM role, you must have `CreateRole`, `CreatePolicy`, and `AttachRolePolicy` permissions.

Create an IAM role named '`AWSGlueServiceRole-torname`' and attach the AWS managed policy `AWSGlueServiceRole`, plus an inline policy that allows read access to:

- S3:octank-database1

You can also create an IAM role on the [IAM console](#).

[Back](#)

[Next](#) 

Create a schedule for this crawler

Frequency

Run on demand

Back

Next

- Crawler info**
 - octark_crawler
- Crawler source type**
 - Data stores
- Data store**
 - S3 s3://octark-data...
...
- IAM Role**
 - arn:aws:iam::6441187:80602:role/service-role/AmazonSGlueServiceRole
Role-
octark_databricks
- Schedule**
- Output**
- Review all steps**

- Crawler info**
 - extarnl_crawler
- Crawler source type**
 - Data stores
- Data store**
 - S3 s3://Portaria-del...
- IAM Role**
 - arn:aws:iam::6441187:80602role/service-role/AmazonSGlueServiceRole
 - Role-externl_databaserole
- Schedule**
 - Run on demand
- Output**
 - (None) all_stages

Configure the crawler's output

Database 0

Choose a database to contain tables

[Add database](#)

Prefix added to tables (optional) 0

Type a prefix added to table names

► Grouping behavior for S3 data (optional)

► Configuration options (optional)

[Back](#)

[Next](#)

- Crawler info**
octank_crawler
- Crawler source type**
Data stores
- Data store**
S3: s3://octank-datalake
- IAM Role**
arn:aws:iam::0441187:
S3FullRole/service-
role-AWSGlueService
Role-
octank_datalakerole
- Schedule**
Run on demand
- Output**
- (Review all steps)**

Configure the crawler's output

Database 0
Choose a database to contain tables

Add database

Prefix added to tables (optional) 0
Type a prefix added to table names

► Grouping behavior for S3 data (optional)

► Configuration options (optional)

Back **Next**



Add database

Database name

Description and location (optional)

Resource link name

Shared database suggestions

Shared database

Shared database owner account ID





Add database

Customer name:

Octave 6

- Description and location (optional)

Resources List page

Public administration and services

Standard deviation of error term

Please e-mail me to receive

Shared database

Create an AWS account ID

https://eu-west-1.console.aws.amazon.com/glue/home?region=eu-west-1#addCrawler

Add crawler

Crawler info

Name: octank_crawler
Tags: -

Data stores

Data store: S3
Include path: s3://octank-database/
Connection: octank_database
Exclude patterns: -

IAM role

IAM role: arn:aws:iam::04411879002:role/service-role/AWSGlueServiceRole-octank_database

Schedule

Schedule: Run on demand

Output

Database: octank_db
Prefix added to tables (optional):
Create a single schema for each S3 path: false
Table level (optional):
• Configuration options

Back Next

https://eu-west-1.console.aws.amazon.com/glue/home?region=eu-west-1#AddCrawler

Add crawler

Crawler info

Name: octank_crawler
Tags: -

Data stores

Data store: S3
Include path: s3://octank-database/
Connection:
Exclude patterns: -

IAM role

IAM role: arn:aws:iam::04111879002:role/service-role/MSQLeServiceRole-octank_database

Schedule

Schedule: Run on demand

Output

Database: octank_db
Prefix added to tables (optional):
Create a single schema for each S3 path: false
Table level (optional):
• Configuration options

Back Finish

AWS Glue

Data catalog

[Databases](#)[Tables](#)[Connections](#)[Crawlers](#)[Classifiers](#)[Schema registries](#)[Schemas](#)[Settings](#)

ETL

[AWS Glue Studio](#)[Blueprints](#)[Workflows](#)[Jobs](#)[ML Transforms](#)[Triggers](#)[Dev endpoints](#)[Notebooks](#)

Security

[Security configurations](#)

Tutorials

[Add crawler](#)[Explore table](#)[Add job](#)[Resources](#)[What's new](#)

Crawlers: A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.

[User preferences](#)

<input type="checkbox"/>	Name	Schedule	Status	Logs	Last runtime	Median runtime	Tables updated	Tables added
<input type="checkbox"/>	oceanic_crawler		Ready		0 secs	0 secs	0	0

AWS Glue

Data catalog

Databases

Tables

Connections

Crawlers

Classifiers

Schema registries

Schemas

Settings

ETL

AWS Glue Studio

Blueprints

Workflows

Jobs

ML Transforms

Triggers

Dev endpoints

Notebooks

Security

Security configurations

Tutorials

Add crawler

Explore table

Add job

Resources

DF

What's new

Crawlers - A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.

User preferences

<input type="checkbox"/>	Name	Schedule	Status	Logs	Last runtime	Median runtime	Tables updated	Tables added
<input type="checkbox"/>	octane_crawler		Starting		0 secs	0 secs	0	0

https://eu-west-1.console.aws.amazon.com/athena/home?#workset=watch&tab=history

Athena Query editor Saved queries History Data sources Workgroup: primary Settings Timer Help About this page

Data source: [Amazon Athena](#) Connect data source

Database: [default](#) [default](#) [athena_05](#) [sampledb](#) [test](#)

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1 [+](#)

Run query Save as Create

Ctrl + Enter to run query, Ctrl + Space to auto-complete

Results

Athena engine version 2 · Release versions (7)

Athena Query editor Saved queries History Data sources Workgroup: primary [Options+]

Data source: AmazonDataCatalog Connect data source

Database: octane_db

Filter tables and views

Tables (1) Create table: octane_database1 (Partitioned)

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1 +

Run query Save as Create -

Use Ctrl + Enter to run query. CM + F2 to autocomplete

Format query Clear

Athena engine version 2 - Release version 5

Results

The screenshot shows the AWS Athena Query Editor. The top navigation bar includes 'Athena', 'Query editor', 'Saved queries', 'History', 'Data sources', 'Workgroup: primary', and various settings like 'Settings', 'Tutorials', 'Help', and 'What's new'. On the left, there's a sidebar for 'Data source' (AmazonDataCatalog), 'Database' (octane_db), and sections for 'Tables' (one entry: octane_database1 (Partitioned)) and 'Views' (empty). Below the sidebar is a large 'New query 1' input area with a '+' button to add another query. At the bottom of this area are buttons for 'Run query', 'Save as', and 'Create -'. A note says 'Use Ctrl + Enter to run query. CM + F2 to autocomplete'. To the right of the input area are 'Format query' and 'Clear' buttons. At the very bottom, it says 'Athena engine version 2 - Release version 5'. The main body of the page is labeled 'Results'.

Athena Query editor Saved queries History Data sources Workgroup: primary

Data source: AvnDataCatalog Connect data source

Database: extana_db

Filter: Tables and views

Tables (1): extana_database1 (Partitioned) Create table

Views (0): Create view

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1

Run query Save as Create

Use Ctrl + Enter to run query, CM + F to auto-complete

Format query Clear

Athena engine version 2 - Release version 1

Results

The screenshot shows the AWS Athena Query Editor. On the left, there's a sidebar with 'Data source' set to 'AvnDataCatalog' and 'Database' set to 'extana_db'. Below that, it says 'Tables (1)' with 'extana_database1 (Partitioned)' listed. There's also a section for 'Views (0)'. A note at the bottom of this sidebar says 'You have not created any views. To create a view, run a query and click "Create view from query".' The main area is titled 'New query 1' and contains three buttons: 'Run query', 'Save as', and 'Create'. Below these buttons is the text 'Use Ctrl + Enter to run query, CM + F to auto-complete'. At the bottom right of the main area, it says 'Format query' and 'Clear'. At the very bottom, it says 'Athena engine version 2 - Release version 1'. The top of the screen has a standard browser-style header with tabs like 'File', 'Edit', 'View', etc., and a status bar at the bottom.

Athena Query editor Saved queries History Data sources Workgroup: primary

Data source: AvnDataCatalog Connect data source

Database: octank_db

Tables (1): octank_database1 (Partitioned)

Views (0)

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1

Preview table Show properties Delete table List partitions Generate Create Table DDL

Run query Save as Create -

Use Ctrl + Enter to run query, Ctrl + Space to autocomplete

Results

Athena engine version 2 Release versions (7)

The screenshot shows the AWS Athena Query Editor. On the left, there's a sidebar with 'Data source' set to 'AvnDataCatalog' and 'Database' set to 'octank_db'. Under 'Tables', there is one entry: 'octank_database1 (Partitioned)'. Below that, under 'Views', it says '(0)'. A note at the bottom of this sidebar says, 'You have not created any views. To create a view, run a query and click "Create view from query".' The main area is titled 'New query 1'. It has a context menu open over the table 'octank_database1' with options like 'Preview table', 'Show properties', 'Delete table', 'List partitions', and 'Generate Create Table DDL'. At the bottom of the main area, there are buttons for 'Run query', 'Save as', and 'Create -'. A status bar at the bottom right indicates 'Athena engine version 2' and 'Release versions (7)'.

AWS Services ▾

Amazon Athena

Query editor

Saved queries History Data sources Workgroup: primary

Search for services, datasets, monitoring products, and docs [Options+]

New query 1 New query 2 +

```
1 SELECT * FROM "testtask_db"."testtask_detailed1" limit 10
```

Data source Connect data source

AmazonDataCatalog

Database

testtask_db

Tables (1) Create table

+ testtask_dataset1 (Partitioned)

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

Cancel Save as Create ↗

use Ctrl + Enter to run query, CM + Space to autocomplete

Results

Burning Query ...

Estimated time elapsed: 0 seconds

You can run another query by clicking on the New Query button. The current query will continue to run in the background. You can check the status of all queries in the History tab.

Format query Clear

Athena engine version 2 · Release versions [F]

https://eu-west-1.console.aws.amazon.com/athena/home?®ion=eu-west-1&queryHistoryId=47c6b3f46-414a-a000-ccf354cb6001

Athena Query editor Saved queries History Data sources Workgroup: primary Settings Tutorial Help What's new

Data source Connect data source AaaDataCatalog

Database octane_db

Tables (1) Create table octane_datatable (Partitioned)

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

New query 1 New query 2 +
SELECT * FROM "octane_db"."octane_datatable" LIMIT 100

Run query Save as Create (Run time: 0.72 seconds, Data scanned: 30.29 KB)
Use CTR + Enter to run query, CM + Space to autocomplete

Athena engine version 2 - Release version 15 Format query Clear

Results

#	deviceid	timestamp	battery	speed	location	partition_0	partition_1	partition_2	partition_3
1	encouter1	1620308524	85	42	[lat=48.19408329432654, long=-11.7000834688034]	2021	05	06	14
2	encouter25	1620308529	85	42	[lat=48.113311132384474, long=-11.6723185487349]	2021	05	06	14
3	encouter27	1620308534	76	1	[lat=48.17521815776688, long=-11.64689025927934]	2021	05	06	14
4	encouter19	1620308539	23	96	[lat=48.17628810041454, long=-11.468270210946567]	2021	05	06	14
5	encouter88	1620308544	16	20	[lat=48.147213958175835, long=-11.83638008780911]	2021	05	06	14
6	encouter12	1620308549	83	6	[lat=48.035877935800419, long=-11.534354181808029]	2021	05	06	14
7	encouter22	1620308554	95	18	[lat=48.2362973556884, long=-11.47584963384267]	2021	05	06	14
8	encouter11	1620308559	18	71	[lat=48.1216395767278, long=-11.87309173381471]	2021	05	06	14
9	encouter30	1620308564	17	67	[lat=48.0382117831533, long=-11.731347888521159]	2021	05	06	14
10	encouter10	1620308569	79	25	[lat=48.16745490290829, long=-11.4421152349161589]	2021	05	06	14

https://eu-west-1.console.aws.amazon.com/athena/home?region=eu-west-1&queryHistoryId=47ccb-3f46-414a-a080-ccf304ccb6001

Athena Query editor Saved queries History Data sources Workgroup: primary Settings Tutorial Help What's new

New query 1 New query 2 +
SELECT * FROM "octane_db"."octane_database1" LIMIT 100

Data source Connect data source
AthenaDataCatalog

Database octane_db

Tables (1) Create table
octane_database1 (Partitioned)

Views (0) Create view

You have not created any views. To create a view, run a query and click "Create view from query".

Run query Save as Create (Run time: 0.72 seconds, Data scanned: 30.29 KB)
Use Ctrl + Enter to run query, Ctrl + Space to autocomplete

Athena engine version 2 - Release version 1.0

Results

#	deviceid	Timestamp	battery	speed	location	partition_0	partition_1	partition_2	partition_3
1	escooter1	1620308524	85	42	[lat=48.19408529432654, long=-11.7000935498034]	2021	05	06	14
2	escooter25	1620308529	85	42	[lat=48.11331130238474, long=-11.67421854873845]	2021	05	06	14
3	escooter27	1620308534	78	1	[lat=48.17521915776668, long=-11.66589925927934]	2021	05	06	14
4	escooter19	1620308539	23	96	[lat=48.17808510041454, long=-11.489270216986587]	2021	05	06	14
5	escooter68	1620308544	16	20	[lat=48.147213581773835, long=-11.826360097880911]	2021	05	06	14
6	escooter12	1620308549	83	6	[lat=48.03597783580415, long=-11.534354181808229]	2021	05	06	14
7	escooter22	1620308554	95	18	[lat=48.2362973556684, long=-11.43584953598267]	2021	05	06	14
8	escooter11	1620308559	14	71	[lat=48.121043995767278, long=-11.573059133581473]	2021	05	06	14
9	escooter30	1620308564	11	67	[lat=48.0382717831533, long=-11.721347888521159]	2021	05	06	14
10	escooter30	1620308569	79	25	[lat=48.10745490290829, long=-11.44715239161589]	2021	05	06	14

https://eu-west-1.console.aws.amazon.com/athena/home?#®ion=eu-west-1

Amazon Athena Manager 2.0

Search results for 'Quick'

Services [6]

See all 6 results ▾

- QuickSight**
Fast, easy-to-use business analytics
- AWS Glue DataBrew**
Visual data preparation tool to clean and normalize data for analytics and machine learning
- Amazon Machine Learning**
Build Smart Applications Quickly and Easily
- CodeStar**
Quickly develop, build, and deploy applications

Features [4]

See all 6 results ▾

- Quick Setup**
Systems Manager feature
- Quick start**
IoT Analytics feature
- Quick Connect workflow**
FreeRTOS Native
- Dashboard**
EC2 feature

Documentation [20,108]

See all 20,108 results in Documentation ▾

Format query Clear

Athena engine version 2 · Release versions: 1.7

	partition_0	partition_1	partition_2	partition_3
140404	2021	05	06	14
173445	2021	05	06	14
219346	2021	05	06	14
468877	2021	05	06	14
540111	2021	05	06	14
606220	2021	05	06	14
642677	2021	05	06	14
644138	2021	05	06	14
711598	2021	05	06	14
815899	2021	05	06	14

https://eu-west-1.console.aws.amazon.com/athena/home?®ion=eu-west-1&queryHistoryID=47cbdb-3f46-414a-a080-ccf3b4cb0d81

Amazon Athena - eu-west-1 - https://eu-west-1.console.aws.amazon.com/athena/home?®ion=eu-west-1

Search results for 'quick'

Services (6)

Features (4)

Documentation (20,108)

Knowledge Articles (38)

Marketplace (3,491)

Services

See all 6 results ▾

- Quicksight**
Fast, easy-to-use business analytics
- AWS Glue DataBrew**
Visual data preparation tool to clean and normalize data for analytics and machine learning
- Amazon Machine Learning**
Build smart applications quickly and easily
- CodeStar**
Quickly develop, build, and deploy applications

Features

See all 6 results ▾

- Quick Setup**
Systems Manager feature
- Quick start**
IoT Analytics feature
- Quick Connect workflow**
FreeRTOS feature
- Dashboard**
EC2 feature

Documentation

See all 20,108 results in Documentation ▾

Format query Clear

Athena engine version 2 - Release version 12

	partition_0	partition_1	partition_2	partition_3
15634	2021	05	06	14
17345	2021	05	06	14
17934	2021	05	06	14
18587	2021	05	06	14
180911	2021	05	06	14
180829	2021	05	06	14
18267	2021	05	06	14
14110	2021	05	06	14
11190	2021	05	06	14
15689	2021	05	06	14

Find analyses & more

Datasets

New dataset

★ Favorites

Recent

My folders

Shared folders

Dashboards

Analyses

Datasets

No datasets

Import or create a new dataset to start an analysis.

Screenshot of the Amazon QuickSight Data Sources interface.

The interface shows a grid of data source icons and a central modal dialog for creating a new Athena data source.

FROM NEW DATA SOURCES

- Upload a file (CSV, XLS, JSON, XML)
- Salesforce (Connect to Salesforce)
- RDS
- ORACLE
- Spark
- Timestream
- Adobe Analytics

FROM EXISTING DATA SOURCES

- Redshift (Automated connect)
- Redshift (Manual connect)
- Aurora
- SQL Server
- Teradata (Provided by Teradata)
- Github

New Athena data source

Data source name: Enter a name for the data source.

Athena workgroup: Economy 2

Validate connection: SSL is enabled

Create data source

SHCE capacity for this region: 28 SHC of 2200

Admin/DynamoDB Standard

Datasets

Create a Dataset

FROM NEW DATA SOURCES

 Upload a file <small>(CSV, Excel, XLSX, XML, JSON)</small>	 Salesforce <small>Connect to Salesforce</small>	 S3 Analytics	 S3	 Athena
 RDS	 Redshift <small>Auto-discovered</small>	 Redshift <small>Manual connect</small>	 MySQL	 PostgreSQL
 ORACLE	 SQL Server	 Aurora	 MariaDB	 Presto
 Spark	 Teradata <small>Provided by Teradata</small>	 Snowflake	 AWS IoT Analytics	 Amazon Elasticsearch Se...
 Timestream	 GitHub	 Twitter	 Jira	 ServiceNow

SPICE capacity for this region: 28.59GB of 22GB

Databases

Create a Dataset

FROM NEW DATA SOURCES

Upload a file <small>(CSV, JSON, XML, AVRO, parquet, JSONL)</small>	Salesforce <small>Connect to Salesforce</small>	S3 Analytics	S3	Athena
RDS	Redshift <small>Auto-detected</small>	Redshift <small>Manual connect</small>	MySQL	PostgreSQL
ORACLE	SQL Server	Aurora	MariaDB	Presto
Spark	Teradata <small>Powered by Teradata</small>	Snowflake	AWS IoT Analytics	Amazon Elasticsearch Se...
Timestream	GitHub	Twitter	Jira	ServiceNow
Adobe Analytics				

FROM EXISTING DATA SOURCES

QuickSight

Datasets

Create a Dataset
FROM NEW DATA SOURCES

Upload a file (CSV, XLS, XML, JSON)
Salesforce (Connect to Salesforce)

Athena

RDS Redshift Redshift MySQL PostgreSQL

ORACLE SQL Server Aurora MariaDB Presto

Spark Teradata Snowflake AWS IoT Analytics Amazon Elasticsearch Se...

Timestream GitHub Twitter Jira ServiceNow

Adobe Analytics

FROM EXISTING DATA SOURCES

New Athena data source

Data source name: octank_datasource

Athena workgroup: Devney2

Validate connection: SSL is enabled

Create data source

Datasets

Create a Dataset

FROM NEW DATA SOURCES

Upload a file (CSV, JSON, XML, XML, JSON)

Salesforce Connect to Salesforce

RDS

ORACLE

Spark

Timestream

Adobe Analytics

Catalog: contain sets of databases.

AwsDataCatalog

Database: contain sets of tables.

Select...

default
octank_db
sampledb
test

Edit/Preview Data Use custom SQL Select

SPICE capacity for this region: 28.99GB of 22GB

Athena

PostgreSQL

Presto

AWS IoT Analytics

Amazon Elasticsearch Se...

Jira

ServiceNow

Datasets

Create a Dataset
FROM NEW DATA SOURCES

Upload a file (CSV, Avro, JSON, XML, Parquet, JDBC)

Salesforce Connect to Salesforce

RDS

ORACLE

Spark

Timestream

Adobe Analytics

FROM EXISTING DATA SOURCES

Athena

PostgreSQL

Presto

Amazon Elasticsearch Service

ServiceNow

Jira

Snowflake

Twitter

GitHub

Teradata Provided by Teradata

Choose your table

ectank_datalake

Catalog: contain sets of databases.

AvlDataCatalog

Database: contain sets of tables.

ectank_0ff

Tables: contain the data you can visualize.

ectank_datalake1

View previous data Use custom SQL

SPICE capacity for this region: 28.94GB of 22GB

The screenshot shows the Amazon Quicksight interface. On the left, there's a sidebar with 'Datasets' and a main area titled 'Create a Dataset FROM NEW DATA SOURCES'. It lists various data sources like 'Upload a file', 'Salesforce', 'RDS', 'ORACLE', 'Spark', 'Timestream', and 'Adobe Analytics'. Below that is a section for 'FROM EXISTING DATA SOURCES' with icons for 'Athena', 'PostgreSQL', 'Presto', 'Amazon Elasticsearch Service...', 'ServiceNow', 'Jira', 'Snowflake', 'Twitter', 'GitHub', and 'Teradata Provided by Teradata'. A modal dialog is open in the center, titled 'Choose your table'. It shows a catalog named 'ectank_datalake' and a database named 'ectank_0ff'. Inside the database, there is one table named 'ectank_datalake1'. At the bottom of the modal are buttons for 'View previous data', 'Use custom SQL', and a blue 'Next' button.

Datasets

Create a Dataset
FROM NEW DATA SOURCES

Upload a file (CSV, XLS, XML, JSON, PDF)
Salesforce Connect to Salesforce

RDS Redshift Auto-discovered

ORACLE SQL Server

Spark Teradata Provided by Teradata

Timestream GitHub Twitter

Adobe Analytics

Choose your table
ectank_datasource

Catalog: contain sets of databases.
AmazonDataCatalog

Database: contain sets of tables.
ectank_db

Tables: contain the data you can visualize.
ectank_datasource?

Edit/Preview data Use custom SQL Select

Athena PostgreSQL Presto

AWS IoT Analytics Amazon Elasticsearch Se...

Jira ServiceNow

SPICE capacity for this region: 28.94GB of 22GB

Datasets

Create a Dataset
FROM NEW DATA SOURCES

Upload a file (CSV, XLS, XML, JSON, PDF)
Salesforce Connect to Salesforce

RDS Redshift Auto-discovered

ORACLE SQL Server

Spark Teradata Provided by Teradata

Timestream GitHub Twitter

Adobe Analytics

Choose your table
ectank_datasource

Catalog: contain sets of databases.
AmazonDataCatalog

Database: contain sets of tables.
ectank_db

Table: contain the data you can visualize.
 ectank_datasource

Edit/Preview data Use custom SQL Select

Athena PostgreSQL Presto

AWS IoT Analytics Amazon Elasticsearch Se...

Jira ServiceNow

SPICE capacity for this region: 28,948 of 2200

Finish dataset creation

Table: octank_datalake1
Data source: octank_datalake1
Schema: octank_db

Import to SPICE for quicker analytics ✓ 1208 available

Directly query your data

Email owners when a refresh fails

[Edit/Preview data](#) [Augment with SageMaker](#) [Visualize](#)

SPICE capacity for this region: 28 [details](#) of 225GB

Dataset

Create a Dataset FROM NEW DATA SOURCES

Upload a file (csv, txt, xlf, xlz, json)  Salesforce Connect to Salesforce

RDS  Redshift Auto-discovered  MySQL

ORACLE  SQL Server  Aurora  MariaDB  Presto

Spark  Teradata Provided by Teradata  AWS IoT Analytics  Amazon Elasticsearch Se...

Timestream  GitHub  Twitter  Jira  ServiceNow

Adobe Analytics

FROM EXISTING DATA SOURCES

Athena  PostgreSQL

SAPCE capacity for this region: 28.59PB of 72.00PB

Create a Dataset
FROM NEW DATA SOURCES

Upload a file (CSV, JSON, XML, AVRO, Parquet)

Salesforce Connect to Salesforce

RDS Auto-Discoverable

ORACLE

Spark

Timestream

Adobe Analytics

Redshift Auto-Discoverable

SQL Server

Teradata Provided by Teradata

Github

Twitter

Aurora

Snowflake

Jira

MySQL

MariaDB

AWS IoT Analytics

Amazon Elasticsearch Service

PostgreSQL

Presto

ServiceNow

SPICE capacity for this region: 28.59PB of 72.00PB

Finish dataset creation

Name: customer-database

Data source: customer-database

Region: US East (Ohio)

Request to AWS Lambda for user-defined ETL

0 of 1200 available

Dataset quality: good

CloudWatch Metrics refresh time: 1 minute

Table Previews (1) | Segment with LogMeister

Preview

The screenshot shows the Amazon QuickSight interface for creating visualizations. On the left, there's a sidebar with various icons for navigation and settings. The main area is titled "AutoGraph" and contains the message: "Choose 1 or more fields and let QuickSight choose the most appropriate chart". To the right, there's a status bar showing "Importing:" with details: "0 rows were imported to SPICE" and "0 rows were skipped".

Dataset: **SPICE** dataset_database1

Field edits

Sheet 1

Fields list

Search field:

- battery
- deviceid
- location
- partition_0
- partition_1
- partition_2
- partition_3
- speed
- timestamp

Importing:
0 rows were imported to SPICE
0 rows were skipped

Options

Dataset

SPICE dataset_database1 100%

Field wells

Sheet 1

Fields list

Search field:

- battery
- deviceid
- location
- partition_0
- partition_1
- partition_2
- partition_3
- speed
- timestamp

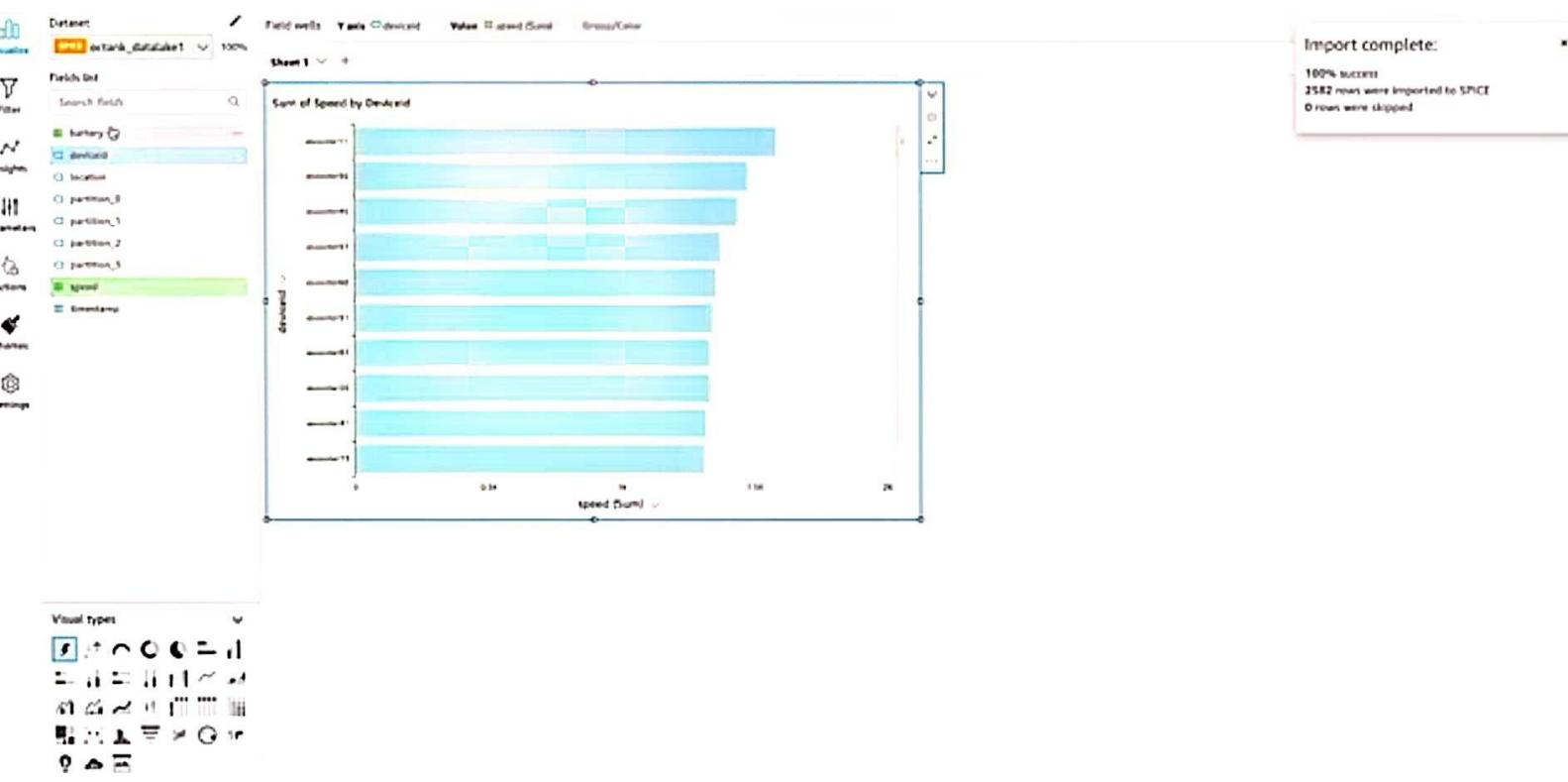
Import complete:

100% success
2542 rows were imported to SPICE
0 rows were skipped

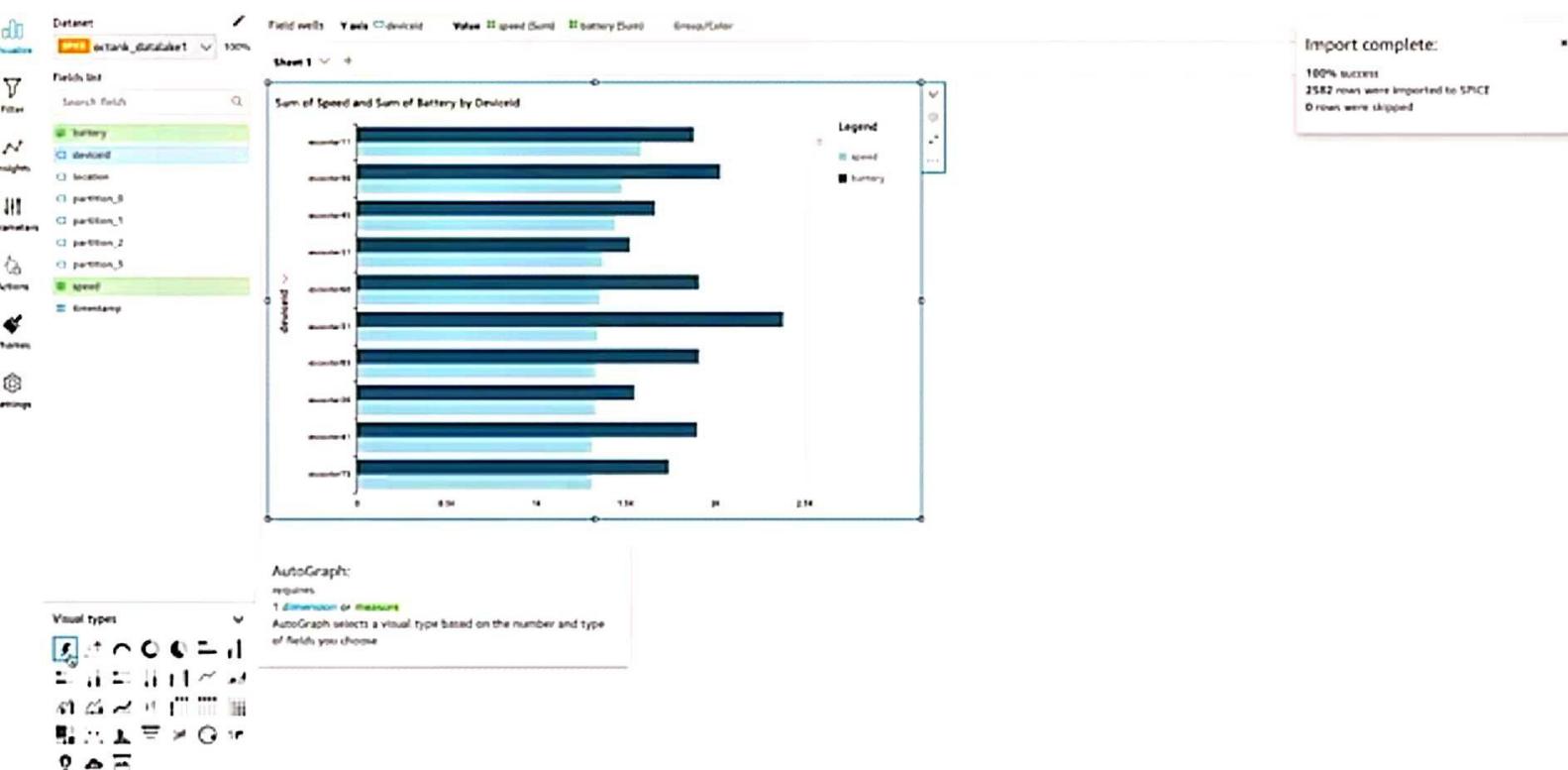
AutoGraph
Choose 1 or more fields and let QuickSight choose the most appropriate chart

The screenshot shows the Amazon QuickSight interface. On the left, there's a sidebar with various icons and a 'Dataset' section showing 'SPICE dataset_database1' and '100%'. Below it is a 'Fields list' with several fields checked: 'battery', 'speed', and 'timestamp'. In the center, a large white canvas is labeled 'AutoGraph' with the sub-instruction 'Choose 1 or more fields and let QuickSight choose the most appropriate chart'. In the top right corner, a message box indicates 'Import complete' with statistics: '100% success', '2542 rows were imported to SPICE', and '0 rows were skipped'.







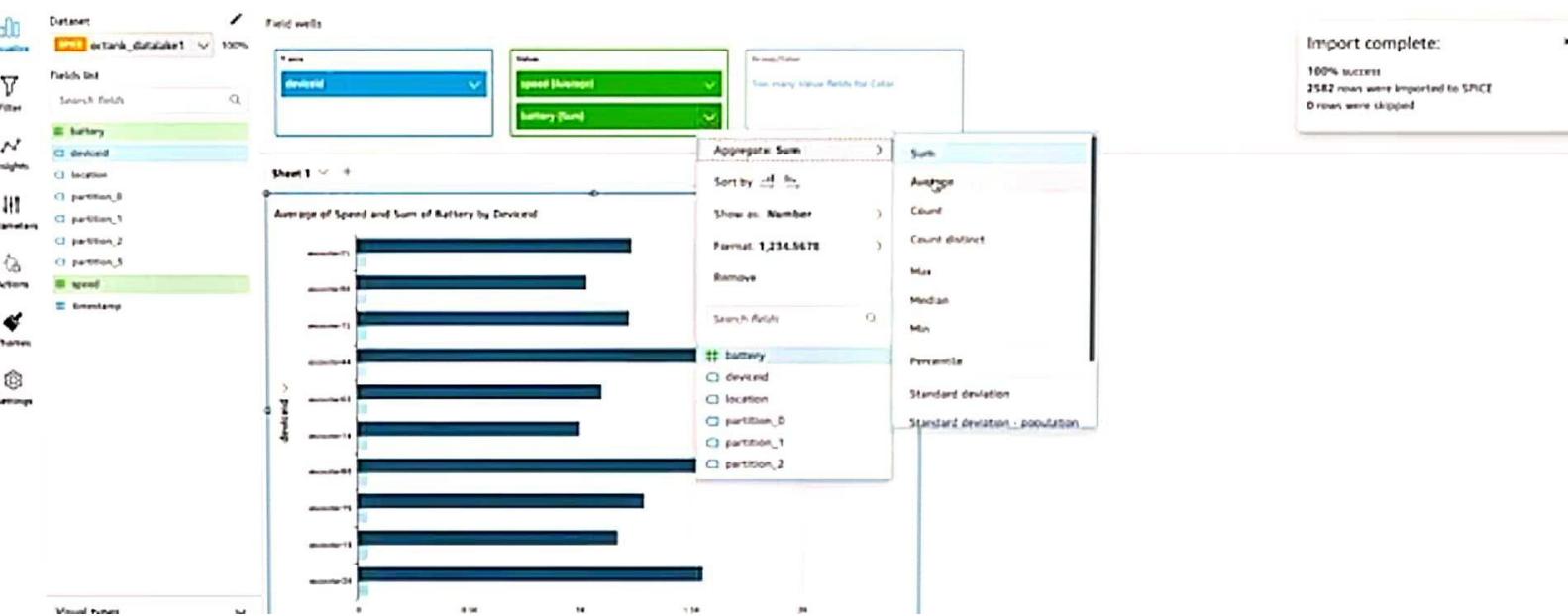








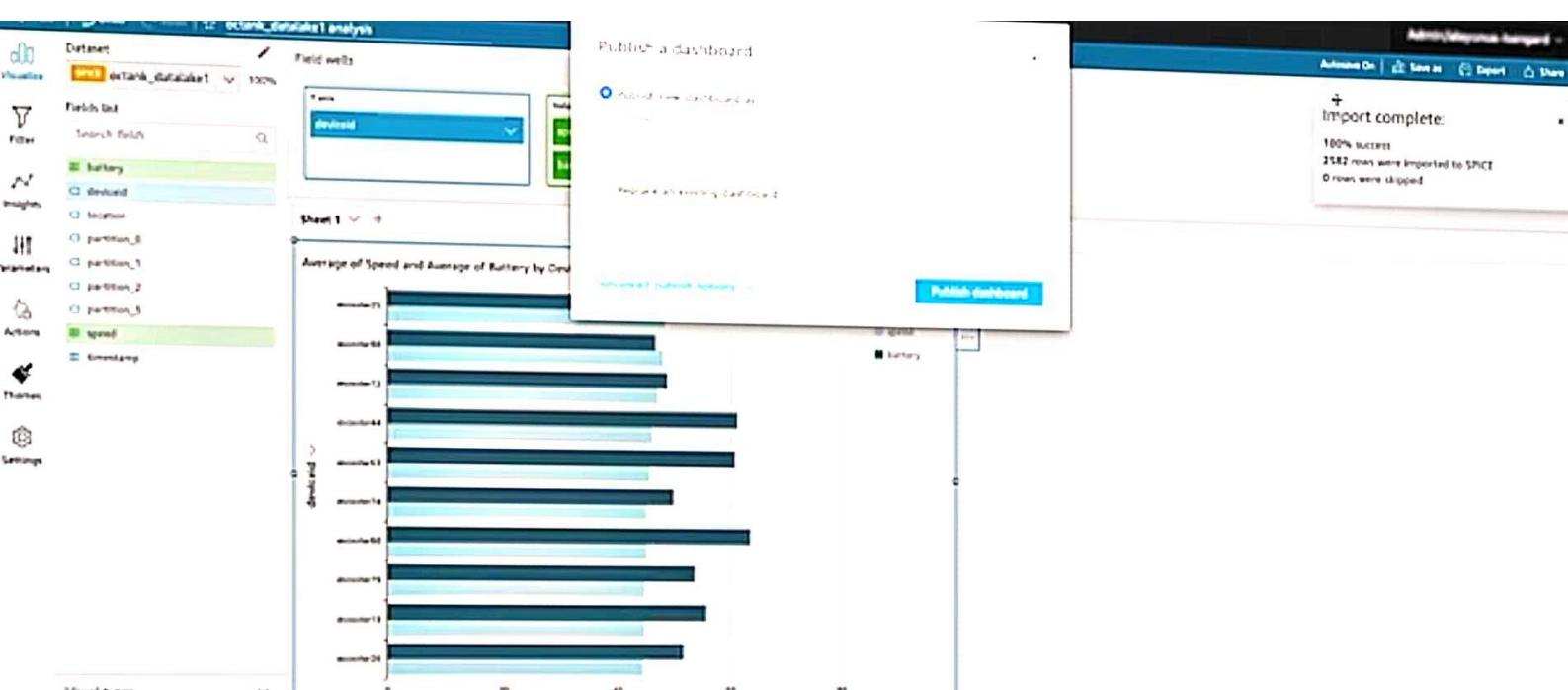










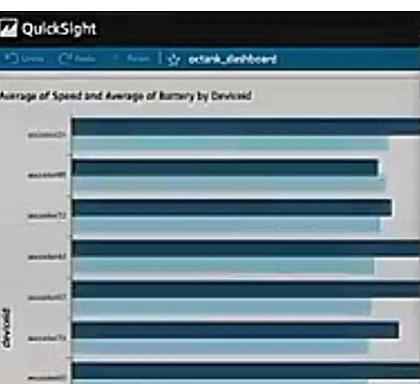


The screenshot shows a data analysis interface with a central dashboard and two overlapping windows.

Left Panel: Shows a dataset named "extank_database1" with 100% completion. The "Field wells" section lists fields: battery (selected), devicel, location, partition_0, partition_1, partition_2, partition_3, speed (selected), and timestamp. A bar chart titled "Average of Speed and Average of Battery by Devicel" displays data for 10 devices. The Y-axis is labeled "devicel" and the X-axis ranges from 0 to 60. The legend indicates blue for "battery" and teal for "speed".

Top Center Window: A "Publish a dashboard" dialog box. It has two options: "Publish new dashboard as" (selected) with the name "extank_dashboard" and "Replace an existing dashboard". Below these are "Advanced publish options" and a "Publish dashboard" button.

Right Side: A separate window showing import statistics: "Import complete: 100% success, 2582 rows were imported to SPICE, 0 rows were skipped".



Share dashboard

Find a person or group to share with.

Enter a username, group, or email address

Share with everyone in this account

Name	Email	Permissions	Role

[Manage dashboard access](#)

X Close







Pipeline to build

