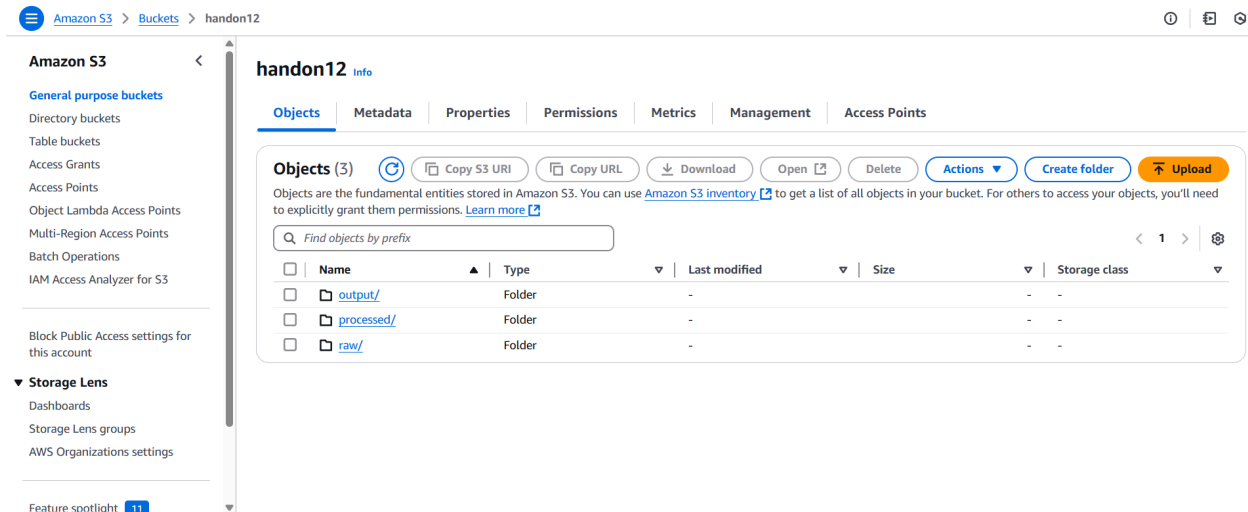


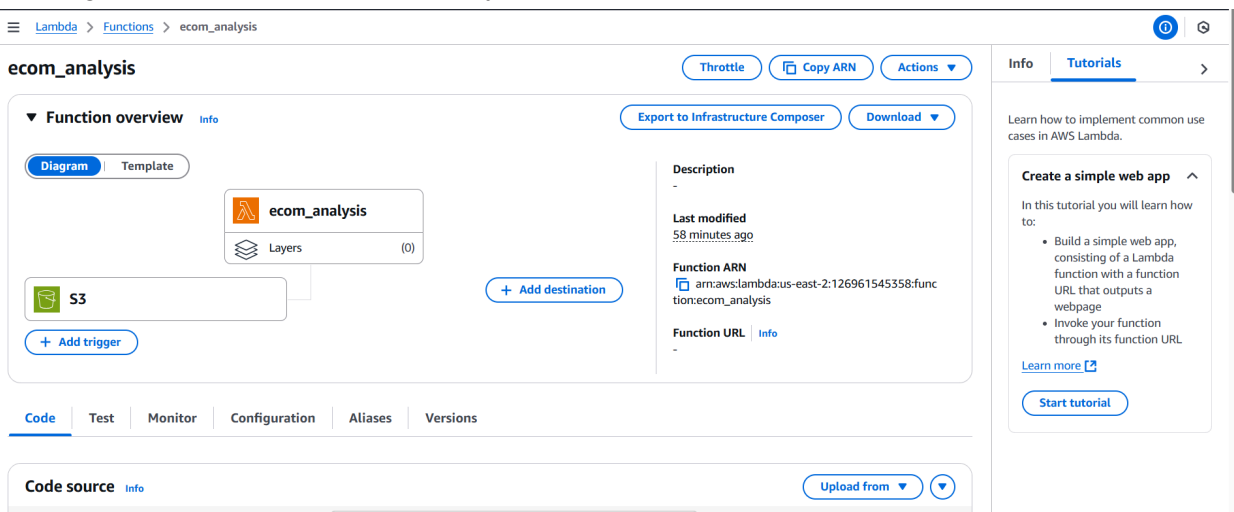
S3 BUCKET

Creating Bucket



LAMBDE

Creating Lamdde Function ecom_analysis.



Adding triggers to the lambda function

The screenshot shows the AWS Lambda console interface. The main content area is titled 'Configuration' and displays 'Triggers (1)'. A search bar is present with the text 'Find triggers'. Below it, a table lists the triggers, showing one trigger named 'S3: handon12' with the ARN 'arn:aws:s3::handon12'. The left sidebar contains a list of configuration options: General configuration, Triggers, Permissions, Destinations, Function URL, Environment variables, Tags, VPC, RDS databases, Monitoring and operations tools, and Concurrency and. The right sidebar has a 'Tutorials' section with a link to 'Create a simple web app' and a 'Start tutorial' button.

Checking Lambda logs

The screenshot shows the AWS CloudWatch console interface. The main content area is titled 'Log events' and displays a list of log events. The left sidebar contains a list of navigation options: CloudWatch, Log groups, Alarms, Logs, Metrics, X-Ray traces, Events, and Application Signals. The right sidebar has a 'Log events' section with a search bar and a 'Display' button. The log events table shows the following data:

Timestamp	Message
2025-04-22T17:24:00.764Z	INIT_START Runtime Version: python:3.13.v38 Runtime Version ARN: arn:aws:lambda:us-east-2::runtime:9ed40a0a2fed561b26d389cb8d6a2a08...
2025-04-22T17:24:01.336Z	START RequestId: f707e4bf-6e6c-4adf-a1bd-04a68e239eca Version: \$LATEST
2025-04-22T17:24:01.337Z	Lambda triggered by S3 event.
2025-04-22T17:24:01.337Z	Incoming file: raw/orders.csv
2025-04-22T17:24:01.669Z	Successfully read file from S3: orders.csv
2025-04-22T17:24:01.669Z	Processing records...
2025-04-22T17:24:01.705Z	Total records processed: 200
2025-04-22T17:24:01.705Z	Records filtered out: 61
2025-04-22T17:24:01.705Z	Records kept: 139

Crawler

The screenshot displays two AWS management console pages. The top page is the 'orders.crawler' configuration in AWS Glue. It shows a green success banner at the top stating 'Crawler successfully starting'. Below this, the crawler's properties are listed in a grid: Name (orders.crawler), IAM role (AWSGlueServiceRole-orders_data_in_s3), Database (ecommerce), State (READY), Description, Security configuration, Lake Formation configuration, and Table prefix. A section for 'Advanced settings' is partially visible. Navigation tabs include 'Crawler runs', 'Schedule', 'Data sources', 'Classifiers', and 'Tags'. The 'Crawler runs' tab shows a single run (1/1) with a filter bar and various controls like 'Stop run', 'View CloudWatch logs', and 'View run details'. The bottom page is the 'Log streams' view in CloudWatch for the same crawler. It shows a table with columns for Log stream name, Tags, Anomaly detection, Metric filters, Subscription filters, Contributor Insights, Data protection, and Field index. One log stream named 'orders.crawler' is listed with its last event time as 2025-04-22 17:35:38 (UTC).

AWS Glue

Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations [New](#)

Data Catalog

- Databases
- Tables
- Stream schema registries
- Schemas
- Connections

Crawlers

- Classifiers
- Catalog settings

Data Integration and ETL

Legacy pages

orders.crawler

Crawler successfully starting
The following crawler is now starting: "orders.crawler"

Last updated (UTC)
April 22, 2025 at 18:17:00

[Run crawler](#) [Edit](#) [Delete](#)

Crawler properties			
Name orders.crawler	IAM role AWSGlueServiceRole-orders_data_in_s3	Database ecommerce	State READY
Description	Security configuration	Lake Formation configuration	Table prefix
Maximum table threshold	-	-	-

► Advanced settings

[Crawler runs](#) | [Schedule](#) | [Data sources](#) | [Classifiers](#) | [Tags](#)

Crawler runs (1/1)

The list of crawler runs for this crawler.

Filter data Filter by a date and time range

< 1 >

Start time (UTC) ▲ End time (UTC) ▼ Current/last duration ▼ Status ▼ DPU hours ▼ Table changes ▼

CloudWatch

/aws-glue/crawlers

Standard

ARN arn:aws:logs:us-east-2:126961545358:log-group:/aws-glue/crawlers:	Subscription filters 0	Sensitive data count 0
Creation time 35 minutes ago	Contributor Insights rules -	Field indexes Configure
Retention Never expire	KMS key ID -	Transformer Configure
Stored bytes -	Anomaly detection Configure	

< [Log streams](#) | [Tags](#) | [Anomaly detection](#) | [Metric filters](#) | [Subscription filters](#) | [Contributor Insights](#) | [Data protection](#) | [Field index](#) >

Log streams (1)

Filter log streams or try prefix search ☐ Exact match ☐ Show expired [Info](#)

< 1 >

<input type="checkbox"/>	Log stream	Last event time
<input type="checkbox"/>	orders.crawler	2025-04-22 17:35:38 (UTC)

AWS GLUE Creating Database

The screenshot shows the AWS Glue console interface. On the left is a navigation menu with categories like 'AWS Glue', 'Data Catalog', and 'Connections'. The main content area displays the 'ecommm' database properties, including its name, description, location, and creation date. Below this, a 'Tables (1)' section shows a table named 'filtered_orders_csv' with its location and classification. A blue banner at the top of the main area announces new optimization features for Apache Iceberg tables.

Database properties

Name	Description	Location	Created on (UTC)
ecommm	-	-	April 22, 2025 at 17:32:14

Tables (1)

View and manage all available tables.

Filter tables

Name	Database	Location	Classification	Deprecated	View data	Data quality	Column statis...
filtered_orders_csv	ecommm	s3://handon12/pro	CSV	-	Table data	View data quality	View statistics

Creating IAM ROLES

The screenshot shows the AWS IAM console interface for the role 'ecom_analysis-role-nkdoguyi'. The left navigation menu includes 'Identity and Access Management (IAM)', 'Access management', and 'Access reports'. The main content area shows the 'Permissions' tab, which lists two permissions policies: 'AmazonS3FullAccess' (AWS managed) and 'AWSLambdaBasicExecutionRole-8e151b25...' (Customer managed). There are buttons for 'Simulate', 'Remove', and 'Add permissions'. A section for 'Permissions boundary' is also visible, indicating it is not set.

Permissions policies (2)

You can attach up to 10 managed policies.

Filter by Type: All types

Policy name	Type	Attached entities
AmazonS3FullAccess	AWS managed	1
AWSLambdaBasicExecutionRole-8e151b25...	Customer managed	1

Permissions boundary (not set)

Generate policy based on CloudTrail events

AWS ATHENA

The screenshot shows the Amazon Athena Query Editor. At the top, there's a navigation bar with the AWS logo, a search bar, and user information. Below this, the 'Query editor' tab is active. A notification banner at the top states: 'Athena now supports typeahead code suggestions to speed up SQL query development. Typeahead suggestions are turned on by default. You can change this setting in query editor preferences.' Below the notification, the left sidebar contains 'Data' settings: 'Data source' is 'AwsDataCatalog', 'Catalog' is 'None', and 'Database' is 'ecomm'. There's a 'Tables and views' section with a 'Create' button and a search bar. The main area displays 'Query 1' with the following SQL code:

```
1 select * from filtered_orders_csv
2
3     SELECT *
4     FROM filtered_orders_csv
5     ORDER BY orderdate DESC
6     LIMIT 10;
7
8     SELECT SUM(amount) AS total_revenue
9     FROM filtered_orders_csv
10    WHERE status IN ('confirmed', 'shipped');
```

Below the query editor, there are buttons: 'Run again', 'Explain', 'Cancel', 'Clear', and 'Create'. At the bottom right, there's a toggle for 'Reuse query results up to 60 minutes ago'.

Cloudwatch logs

The screenshot shows the AWS CloudWatch 'Log groups' page. The left sidebar contains a navigation menu with 'CloudWatch' at the top, followed by 'Favorites and recents', 'Dashboards', 'AI Operations', 'Alarms', 'Logs' (selected), 'Metrics', 'X-Ray traces', 'Events', and 'Application Signals'. The 'Logs' section is expanded, showing 'Log groups', 'Log Anomalies', 'Live Tail', 'Logs Insights', and 'Contributor Insights'. The main area is titled 'Log groups (2)' and includes a search bar, a filter for 'Exact match', and a table of log groups. The table has columns: 'Log group', 'Log class', 'Anomaly d...', 'Data pr...', 'Sensitiv...', 'Retention', and 'Metric fi...'. The log groups listed are:

Log group	Log class	Anomaly d...	Data pr...	Sensitiv...	Retention	Metric fi...
/aws-glue/crawlers	Standard	Configure	-	-	Never expire	-
/aws/lambda/ecom_analysis	Standard	Configure	-	-	Never expire	-

