

Hands-on #12: AWS Services Lambda & Glue & Athena :

Serverless ETL Pipeline

SRIJA BANDA

801393582

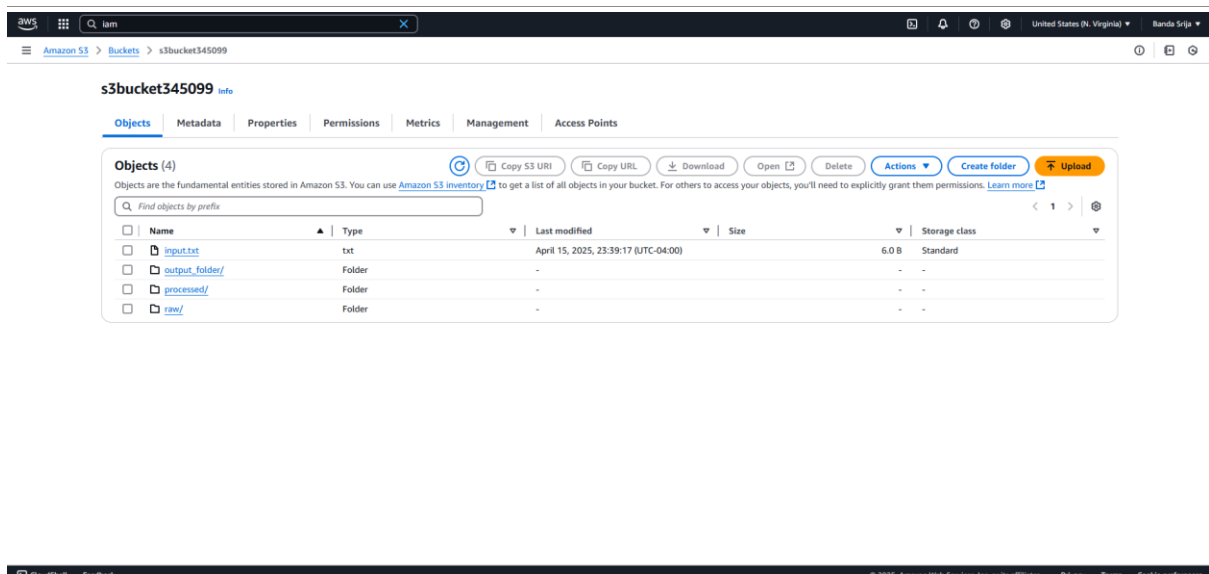
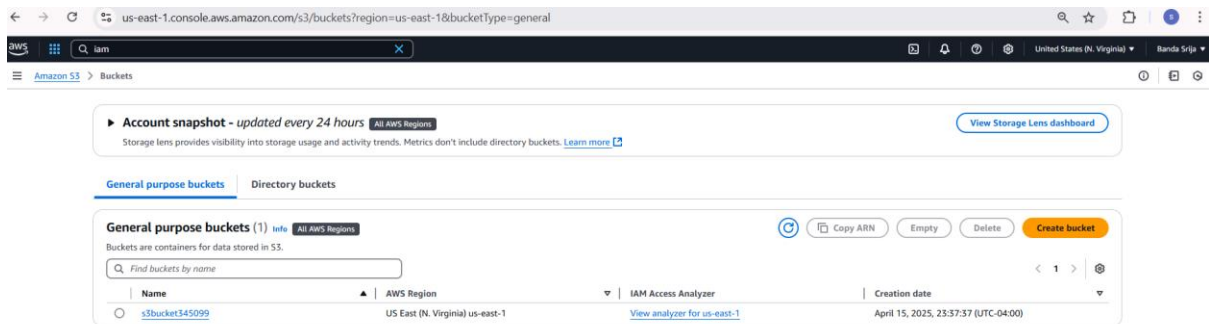
GITHUB LINK-

<https://github.com/Srijabanda/ServerlessETLPipeline.git>

Step-by-step setup

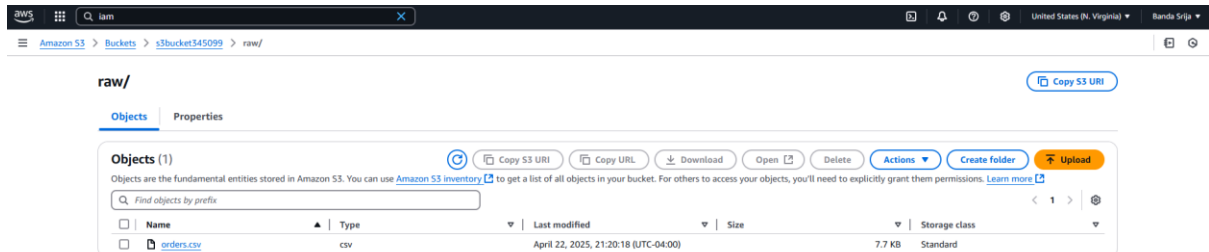
1. Create the bucket and folders

- Create an S3 bucket, for example online-orders-pipeline.
- Inside it, add two empty folders named raw/ and processed/.



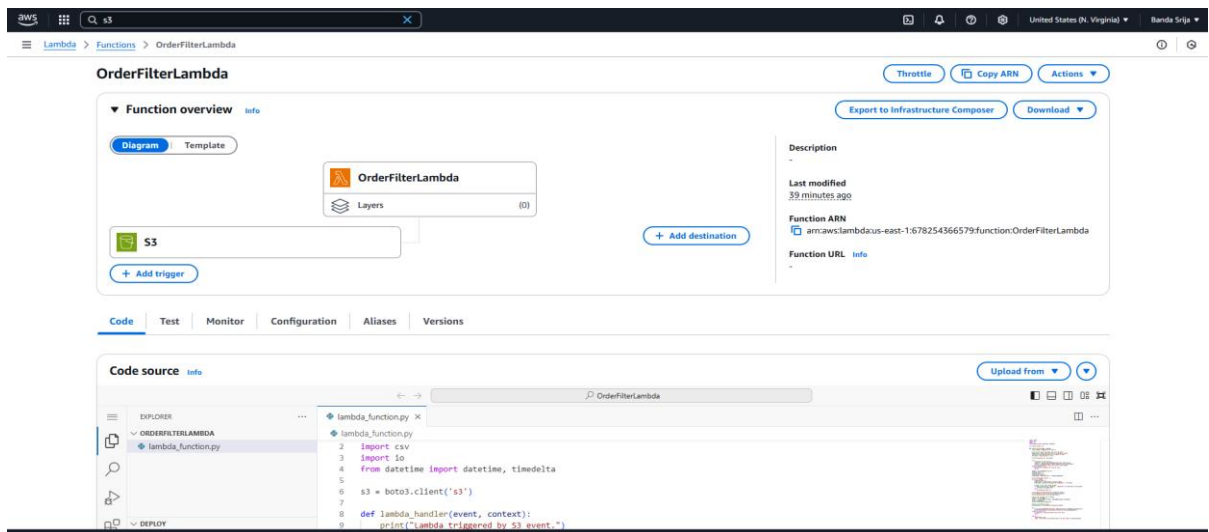
2. Generate sample data locally

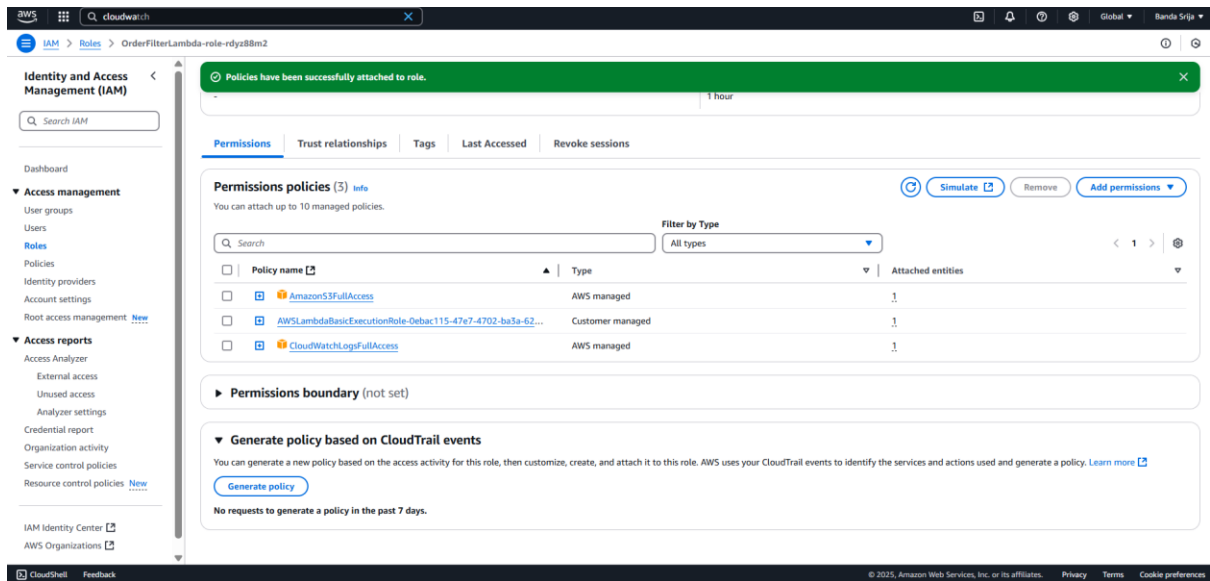
Run the following Python script; it creates orders.csv with 100 random rows.



3. Create the Lambda function

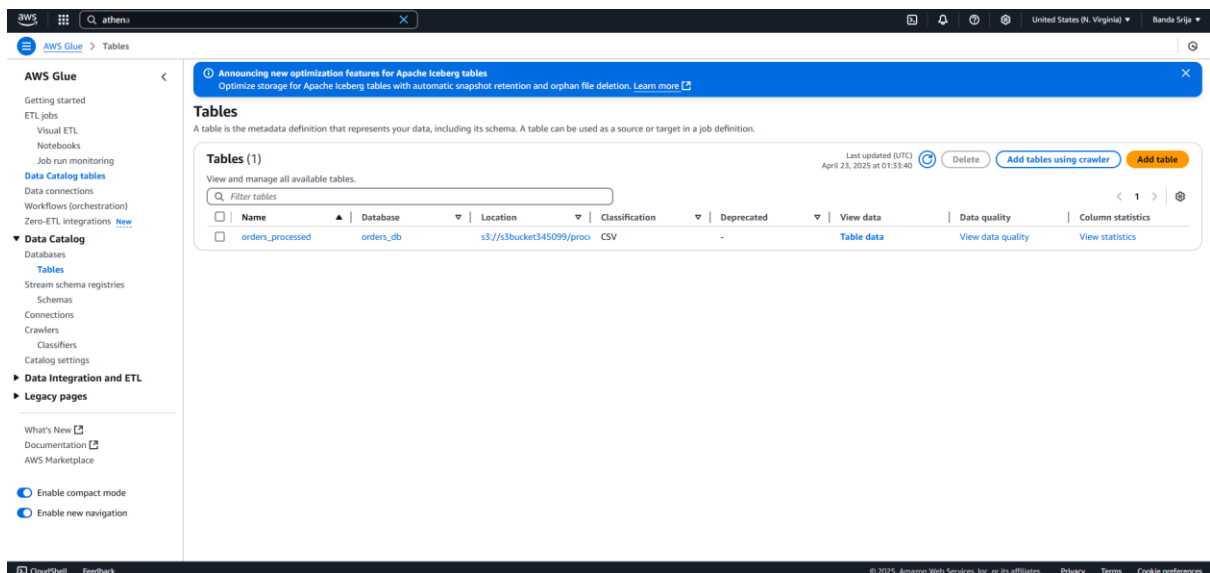
Handler (Python 3.9 or newer)





4. Set up a Glue crawler

1. AWS Glue → Crawlers → Create.
2. Source: s3://<bucket>/processed/
3. Target database: create (or reuse) one, e.g. orders_db.
4. Table name: orders_processed.
5. Run the crawler after Lambda finishes.



AWS Glue > **Tables** > **orders_processed**

Announcing new optimization features for Apache Iceberg tables
Optimize storage for Apache Iceberg tables with automatic snapshot retention and orphan file deletion. [Learn more](#)

Last updated (UTC) April 23, 2025 at 01:51:47 **Version 0 (Current version)** **Actions**

orders_processed

Table overview | Data quality - new

Table details

Name orders_processed	Classification CSV	Deprecated -
Database orders_db	Location s3://s3bucket345099/processed/	Column statistics No statistics
Description -	Connection -	
Last updated April 23, 2025 at 01:32:51		

Advanced properties

Schema | Partitions | Indexes | Column statistics - new

Schema (5)
View and manage the table schema. [Edit schema as JSON](#) [Edit schema](#)

#	Column name	Data type	Partition key	Comment
1	orderid	string	-	-
2	customer	string	-	-
3	amount	double	-	-
4	status	string	-	-

AWS Glue > **Crawlers** > **orders-processed-crawler**

Last updated (UTC) April 23, 2025 at 01:52:17 **Run crawler** **Edit** **Delete**

orders-processed-crawler

Crawler properties

Name orders-processed-crawler	IAM role AWSGlueServiceRole-orders	Database orders_db	State READY
Description -	Security configuration -	Lake Formation configuration -	Table prefix orders_
Maximum table threshold -			

Advanced settings

Crawler runs | Schedule | Data sources | Classifiers | Tags

Crawler runs (1)
The list of crawler runs for this crawler. [Stop run](#) [View CloudWatch logs](#) [View run details](#)

Start time (UTC)	End time (UTC)	Current/last duration	Status	DPU hours	Table changes
April 23, 2025 at 01:31:33	April 23, 2025 at 01:32:51	01 min 18 s	Completed	0.040	1 table change, 0 partition changes

5. Query the data with Athena

Make sure Athena's query-result location is set, then run:

Amazon Athena > **Query editor**

Editor | Recent queries | Saved queries | Settings

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. [Edit preferences](#)

Data

Data source
AwsDataCatalog

Catalog
None

Database
orders_db

Tables and views
[Create](#)

Tables (1)
orders_processed

Views (0)

Query 1 : **Query 2** : **Query 3** : **Query 4** : **Query 5** : **Query 6** : **Query 7** : **Query 8** : **Query 9** : **Query 10** : **Query 11** : **Query 12** : **Query 13** : **Query 14** : **Query 15** : **Query 16** : **Query 17** : **Query 18** : **Query 19** : **Query 20** : **Query 21** : **Query 22** : **Query 23** : **Query 24** : **Query 25** : **Query 26** : **Query 27** : **Query 28** : **Query 29** : **Query 30** : **Query 31** : **Query 32** : **Query 33** : **Query 34** : **Query 35** : **Query 36** : **Query 37** : **Query 38** : **Query 39** : **Query 40** : **Query 41** : **Query 42** : **Query 43** : **Query 44** : **Query 45** : **Query 46** : **Query 47** : **Query 48** : **Query 49** : **Query 50** : **Query 51** : **Query 52** : **Query 53** : **Query 54** : **Query 55** : **Query 56** : **Query 57** : **Query 58** : **Query 59** : **Query 60** : **Query 61** : **Query 62** : **Query 63** : **Query 64** : **Query 65** : **Query 66** : **Query 67** : **Query 68** : **Query 69** : **Query 70** : **Query 71** : **Query 72** : **Query 73** : **Query 74** : **Query 75** : **Query 76** : **Query 77** : **Query 78** : **Query 79** : **Query 80** : **Query 81** : **Query 82** : **Query 83** : **Query 84** : **Query 85** : **Query 86** : **Query 87** : **Query 88** : **Query 89** : **Query 90** : **Query 91** : **Query 92** : **Query 93** : **Query 94** : **Query 95** : **Query 96** : **Query 97** : **Query 98** : **Query 99** : **Query 100** : **Query 101** : **Query 102** : **Query 103** : **Query 104** : **Query 105** : **Query 106** : **Query 107** : **Query 108** : **Query 109** : **Query 110** : **Query 111** : **Query 112** : **Query 113** : **Query 114** : **Query 115** : **Query 116** : **Query 117** : **Query 118** : **Query 119** : **Query 120** : **Query 121** : **Query 122** : **Query 123** : **Query 124** : **Query 125** : **Query 126** : **Query 127** : **Query 128** : **Query 129** : **Query 130** : **Query 131** : **Query 132** : **Query 133** : **Query 134** : **Query 135** : **Query 136** : **Query 137** : **Query 138** : **Query 139** : **Query 140** : **Query 141** : **Query 142** : **Query 143** : **Query 144** : **Query 145** : **Query 146** : **Query 147** : **Query 148** : **Query 149** : **Query 150** : **Query 151** : **Query 152** : **Query 153** : **Query 154** : **Query 155** : **Query 156** : **Query 157** : **Query 158** : **Query 159** : **Query 160** : **Query 161** : **Query 162** : **Query 163** : **Query 164** : **Query 165** : **Query 166** : **Query 167** : **Query 168** : **Query 169** : **Query 170** : **Query 171** : **Query 172** : **Query 173** : **Query 174** : **Query 175** : **Query 176** : **Query 177** : **Query 178** : **Query 179** : **Query 180** : **Query 181** : **Query 182** : **Query 183** : **Query 184** : **Query 185** : **Query 186** : **Query 187** : **Query 188** : **Query 189** : **Query 190** : **Query 191** : **Query 192** : **Query 193** : **Query 194** : **Query 195** : **Query 196** : **Query 197** : **Query 198** : **Query 199** : **Query 200** : **Query 201** : **Query 202** : **Query 203** : **Query 204** : **Query 205** : **Query 206** : **Query 207** : **Query 208** : **Query 209** : **Query 210** : **Query 211** : **Query 212** : **Query 213** : **Query 214** : **Query 215** : **Query 216** : **Query 217** : **Query 218** : **Query 219** : **Query 220** : **Query 221** : **Query 222** : **Query 223** : **Query 224** : **Query 225** : **Query 226** : **Query 227** : **Query 228** : **Query 229** : **Query 230** : **Query 231** : **Query 232** : **Query 233** : **Query 234** : **Query 235** : **Query 236** : **Query 237** : **Query 238** : **Query 239** : **Query 240** : **Query 241** : **Query 242** : **Query 243** : **Query 244** : **Query 245** : **Query 246** : **Query 247** : **Query 248** : **Query 249** : **Query 250** : **Query 251** : **Query 252** : **Query 253** : **Query 254** : **Query 255** : **Query 256** : **Query 257** : **Query 258** : **Query 259** : **Query 260** : **Query 261** : **Query 262** : **Query 263** : **Query 264** : **Query 265** : **Query 266** : **Query 267** : **Query 268** : **Query 269** : **Query 270** : **Query 271** : **Query 272** : **Query 273** : **Query 274** : **Query 275** : **Query 276** : **Query 277** : **Query 278** : **Query 279** : **Query 280** : **Query 281** : **Query 282** : **Query 283** : **Query 284** : **Query 285** : **Query 286** : **Query 287** : **Query 288** : **Query 289** : **Query 290** : **Query 291** : **Query 292** : **Query 293** : **Query 294** : **Query 295** : **Query 296** : **Query 297** : **Query 298** : **Query 299** : **Query 300** : **Query 301** : **Query 302** : **Query 303** : **Query 304** : **Query 305** : **Query 306** : **Query 307** : **Query 308** : **Query 309** : **Query 310** : **Query 311** : **Query 312** : **Query 313** : **Query 314** : **Query 315** : **Query 316** : **Query 317** : **Query 318** : **Query 319** : **Query 320** : **Query 321** : **Query 322** : **Query 323** : **Query 324** : **Query 325** : **Query 326** : **Query 327** : **Query 328** : **Query 329** : **Query 330** : **Query 331** : **Query 332** : **Query 333** : **Query 334** : **Query 335** : **Query 336** : **Query 337** : **Query 338** : **Query 339** : **Query 340** : **Query 341** : **Query 342** : **Query 343** : **Query 344** : **Query 345** : **Query 346** : **Query 347** : **Query 348** : **Query 349** : **Query 350** : **Query 351** : **Query 352** : **Query 353** : **Query 354** : **Query 355** : **Query 356** : **Query 357** : **Query 358** : **Query 359** : **Query 360** : **Query 361** : **Query 362** : **Query 363** : **Query 364** : **Query 365** : **Query 366** : **Query 367** : **Query 368** : **Query 369** : **Query 370** : **Query 371** : **Query 372** : **Query 373** : **Query 374** : **Query 375** : **Query 376** : **Query 377** : **Query 378** : **Query 379** : **Query 380** : **Query 381** : **Query 382** : **Query 383** : **Query 384** : **Query 385** : **Query 386** : **Query 387** : **Query 388** : **Query 389** : **Query 390** : **Query 391** : **Query 392** : **Query 393** : **Query 394** : **Query 395** : **Query 396** : **Query 397** : **Query 398** : **Query 399** : **Query 400** : **Query 401** : **Query 402** : **Query 403** : **Query 404** : **Query 405** : **Query 406** : **Query 407** : **Query 408** : **Query 409** : **Query 410** : **Query 411** : **Query 412** : **Query 413** : **Query 414** : **Query 415** : **Query 416** : **Query 417** : **Query 418** : **Query 419** : **Query 420** : **Query 421** : **Query 422** : **Query 423** : **Query 424** : **Query 425** : **Query 426** : **Query 427** : **Query 428** : **Query 429** : **Query 430** : **Query 431** : **Query 432** : **Query 433** : **Query 434** : **Query 435** : **Query 436** : **Query 437** : **Query 438** : **Query 439** : **Query 440** : **Query 441** : **Query 442** : **Query 443** : **Query 444** : **Query 445** : **Query 446** : **Query 447** : **Query 448** : **Query 449** : **Query 450** : **Query 451** : **Query 452** : **Query 453** : **Query 454** : **Query 455** : **Query 456** : **Query 457** : **Query 458** : **Query 459** : **Query 460** : **Query 461** : **Query 462** : **Query 463** : **Query 464** : **Query 465** : **Query 466** : **Query 467** : **Query 468** : **Query 469** : **Query 470** : **Query 471** : **Query 472** : **Query 473** : **Query 474** : **Query 475** : **Query 476** : **Query 477** : **Query 478** : **Query 479** : **Query 480** : **Query 481** : **Query 482** : **Query 483** : **Query 484** : **Query 485** : **Query 486** : **Query 487** : **Query 488** : **Query 489** : **Query 490** : **Query 491** : **Query 492** : **Query 493** : **Query 494** : **Query 495** : **Query 496** : **Query 497** : **Query 498** : **Query 499** : **Query 500** : **Query 501** : **Query 502** : **Query 503** : **Query 504** : **Query 505** : **Query 506** : **Query 507** : **Query 508** : **Query 509** : **Query 510** : **Query 511** : **Query 512** : **Query 513** : **Query 514** : **Query 515** : **Query 516** : **Query 517** : **Query 518** : **Query 519** : **Query 520** : **Query 521** : **Query 522** : **Query 523** : **Query 524** : **Query 525** : **Query 526** : **Query 527** : **Query 528** : **Query 529** : **Query 530** : **Query 531** : **Query 532** : **Query 533** : **Query 534** : **Query 535** : **Query 536** : **Query 537** : **Query 538** : **Query 539** : **Query 540** : **Query 541** : **Query 542** : **Query 543** : **Query 544** : **Query 545** : **Query 546** : **Query 547** : **Query 548** : **Query 549** : **Query 550** : **Query 551** : **Query 552** : **Query 553** : **Query 554** : **Query 555** : **Query 556** : **Query 557** : **Query 558** : **Query 559** : **Query 560** : **Query 561** : **Query 562** : **Query 563** : **Query 564** : **Query 565** : **Query 566** : **Query 567** : **Query 568** : **Query 569** : **Query 570** : **Query 571** : **Query 572** : **Query 573** : **Query 574** : **Query 575** : **Query 576** : **Query 577** : **Query 578** : **Query 579** : **Query 580** : **Query 581** : **Query 582** : **Query 583** : **Query 584** : **Query 585** : **Query 586** : **Query 587** : **Query 588** : **Query 589** : **Query 590** : **Query 591** : **Query 592** : **Query 593** : **Query 594** : **Query 595** : **Query 596** : **Query 597** : **Query 598** : **Query 599** : **Query 600** : **Query 601** : **Query 602** : **Query 603** : **Query 604** : **Query 605** : **Query 606** : **Query 607** : **Query 608** : **Query 609** : **Query 610** : **Query 611** : **Query 612** : **Query 613** : **Query 614** : **Query 615** : **Query 616** : **Query 617** : **Query 618** : **Query 619** : **Query 620** : **Query 621** : **Query 622** : **Query 623** : **Query 624** : **Query 625** : **Query 626** : **Query 627** : **Query 628** : **Query 629** : **Query 630** : **Query 631** : **Query 632** : **Query 633** : **Query 634** : **Query 635** : **Query 636** : **Query 637** : **Query 638** : **Query 639** : **Query 640** : **Query 641** : **Query 642** : **Query 643** : **Query 644** : **Query 645** : **Query 646** : **Query 647** : **Query 648** : **Query 649** : **Query 650** : **Query 651** : **Query 652** : **Query 653** : **Query 654** : **Query 655** : **Query 656** : **Query 657** : **Query 658** : **Query 659** : **Query 660** : **Query 661** : **Query 662** : **Query 663** : **Query 664** : **Query 665** : **Query 666** : **Query 667** : **Query 668** : **Query 669** : **Query 670** : **Query 671** : **Query 672** : **Query 673** : **Query 674** : **Query 675** : **Query 676** : **Query 677** : **Query 678** : **Query 679** : **Query 680** : **Query 681** : **Query 682** : **Query 683** : **Query 684** : **Query 685** : **Query 686** : **Query 687** : **Query 688** : **Query 689** : **Query 690** : **Query 691** : **Query 692** : **Query 693** : **Query 694** : **Query 695** : **Query 696** : **Query 697** : **Query 698** : **Query 699** : **Query 700** : **Query 701** : **Query 702** : **Query 703** : **Query 704** : **Query 705** : **Query 706** : **Query 707** : **Query 708** : **Query 709** : **Query 710** : **Query 711** : **Query 712** : **Query 713** : **Query 714** : **Query 715** : **Query 716** : **Query 717** : **Query 718** : **Query 719** : **Query 720** : **Query 721** : **Query 722** : **Query 723** : **Query 724** : **Query 725** : **Query 726** : **Query 727** : **Query 728** : **Query 729** : **Query 730** : **Query 731** : **Query 732** : **Query 733** : **Query 734** : **Query 735** : **Query 736** : **Query 737** : **Query 738** : **Query 739** : **Query 740** : **Query 741** : **Query 742** : **Query 743** : **Query 744** : **Query 745** : **Query 746** : **Query 747** : **Query 748** : **Query 749** : **Query 750** : **Query 751** : **Query 752** : **Query 753** : **Query 754** : **Query 755** : **Query 756** : **Query 757** : **Query 758** : **Query 759** : **Query 760** : **Query 761** : **Query 762** : **Query 763** : **Query 764** : **Query 765** : **Query 766** : **Query 767** : **Query 768** : **Query 769** : **Query 770** : **Query 771** : **Query 772** : **Query 773** : **Query 774** : **Query 775** : **Query 776** : **Query 777** : **Query 778** : **Query 779** : **Query 780** : **Query 781** : **Query 782** : **Query 783** : **Query 784** : **Query 785** : **Query 786** : **Query 787** : **Query 788** : **Query 789** : **Query 790** : **Query 791** : **Query 792** : **Query 793** : **Query 794** : **Query 795** : **Query 796** : **Query 797** : **Query 798** : **Query 799** : **Query 800** : **Query 801** : **Query 802** : **Query 803** : **Query 804** : **Query 805** : **Query 806** : **Query 807** : **Query 808** : **Query 809** : **Query 810** : **Query 811** : **Query 812** : **Query 813** : **Query 814** : **Query 815** : **Query 816** : **Query 817** : **Query 818** : **Query 819** : **Query 820** : **Query 821** : **Query 822** : **Query 823** : **Query 824** : **Query 825** : **Query 826** : **Query 827** : **Query 828** : **Query 829** : **Query 830** : **Query 831** : **Query 832** : **Query 833** : **Query 834** : **Query 835** : **Query 836** : **Query 837** : **Query 838** : **Query 839** : **Query 840** : **Query 841** : **Query 842** : **Query 843** : **Query 844** : **Query 845** : **Query 846** : **Query 847** : **Query 848** : **Query 849** : **Query 850** : **Query 851** : **Query 852** : **Query 853** : **Query 854** : **Query 855** : **Query 856** : **Query 857** : **Query 858** : **Query 859** : **Query 860** : **Query 861** : **Query 862** : **Query 863** : **Query 864** : **Query 865** : **Query 866** : **Query 867** : **Query 868** : **Query 869** : **Query 870** : **Query 871** : **Query 872** : **Query 873** : **Query 874** : **Query 875** : **Query 876** : **Query 877** : **Query 878** : **Query 879** : **Query 880** : **Query 881** : **Query 882** : **Query 883** : **Query 884** : **Query 885** : **Query 886** : **Query 887** : **Query 888** : **Query 889** : **Query 890** : **Query 891** : **Query 892** : **Query 893** : **Query 894** : **Query 895** : **Query 896** : **Query 897** : **Query 898** : **Query 899** : **Query 900** : **Query 90**

The screenshot shows the Amazon Athena Query Editor. On the left, the 'Data' sidebar shows the data source as 'AwsDataCatalog', catalog as 'None', and database as 'orders_db'. Below this, 'Tables and views' are listed, including 'orders_processed'. The main area displays a SQL query:

```
1 SELECT *
2 FROM orders_processed
3 ORDER BY orderdate DESC
4 LIMIT 10;
```

Below the query, the 'Query results' section shows a 'Completed' status with a 'Time in queue: 78 ms', 'Run time: 584 ms', and 'Data scanned: 5.35 KB'. The results are displayed in a table with 10 rows:

#	orderid	customer	amount	status	orderdate
1	O0003	Diana	230.68	confirmed	2025-04-21
2	O0063	Diana	305.64	cancelled	2025-04-21
3	O0129	Alice	206.43	shipped	2025-04-21
4	O0014	Charlie	383.52	confirmed	2025-04-21
5	O0070	Alice	162.44	shipped	2025-04-21

6. CloudWatch logs

CloudWatch Logs → Log groups → `/aws/lambda/<function-name>`

The screenshot shows the AWS CloudWatch console. The left sidebar contains navigation links for 'Dashboards', 'Alarms', 'Logs', 'Metrics', 'X-Ray traces', 'Events', 'Application Signals', 'Network Monitoring', and 'Insights'. The main area is titled 'Log groups (2)' and shows a list of log groups:

Log group	Log class	Anomaly d...	Data protection	Sensitive data co...	Retention	Metric filters	Contributor Ins...	Subscri...
/aws/lambda/OrderFilterLambda	Standard	Configure	-	-	Never expire	-	-	-
/aws-glue/crawlers	Standard	Configure	-	-	Never expire	-	-	-

CloudWatch

Log groups

/aws/lambda/OrderFilterLambda

Actions

View in Logs Insights

Start tailing

Search log group

CloudWatch

Log groups

/aws/lambda/OrderFilterLambda

Log group details

Log class

Standard

ARN

arn:aws:logs:us-east-1:678254366579:log-group:/aws/lambda/OrderFilterLambda*

Creation time

35 minutes ago

Retention

Never expire

Stored bytes

-

Metric filters

0

Subscription filters

0

Contributor insights rules

-

KMS key ID

-

Anomaly detection

Configure

Data protection

-

Sensitive data count

-

Field indexes

Configure

Transformer

Configure

Log streams

Tags

Anomaly detection

Metric filters

Subscription filters

Contributor Insights

Data protection

Field indexes - new

Transformer - new

Log streams (1)

Filter log streams or try prefix search

☐ Exact match ☐ Show expired [Info](#)

☐ Log stream

Last event time

☐ 2025/04/23/5LATE57J659d6edaa68b49ee8469424d101316eb

2025-04-23 01:20:19 (UTC)

CloudWatch

Log groups

/aws-glue/crawlers

Actions

View in Logs Insights

Start tailing

Search log group

CloudWatch

Log groups

/aws-glue/crawlers

Log group details

Log class

Standard

ARN

arn:aws:logs:us-east-1:678254366579:log-group:/aws-glue/crawlers*

Creation time

23 minutes ago

Retention

Never expire

Stored bytes

-

Metric filters

0

Subscription filters

0

Contributor insights rules

-

KMS key ID

-

Anomaly detection

Configure

Data protection

-

Sensitive data count

-

Field indexes

Configure

Transformer

Configure

Log streams

Tags

Anomaly detection

Metric filters

Subscription filters

Contributor Insights

Data protection

Field indexes - new

Transformer - new

Log streams (1)

Filter log streams or try prefix search

☐ Exact match ☐ Show expired [Info](#)

☐ Log stream

Last event time

☐ orders-processed-crawler

2025-04-23 01:31:57 (UTC)