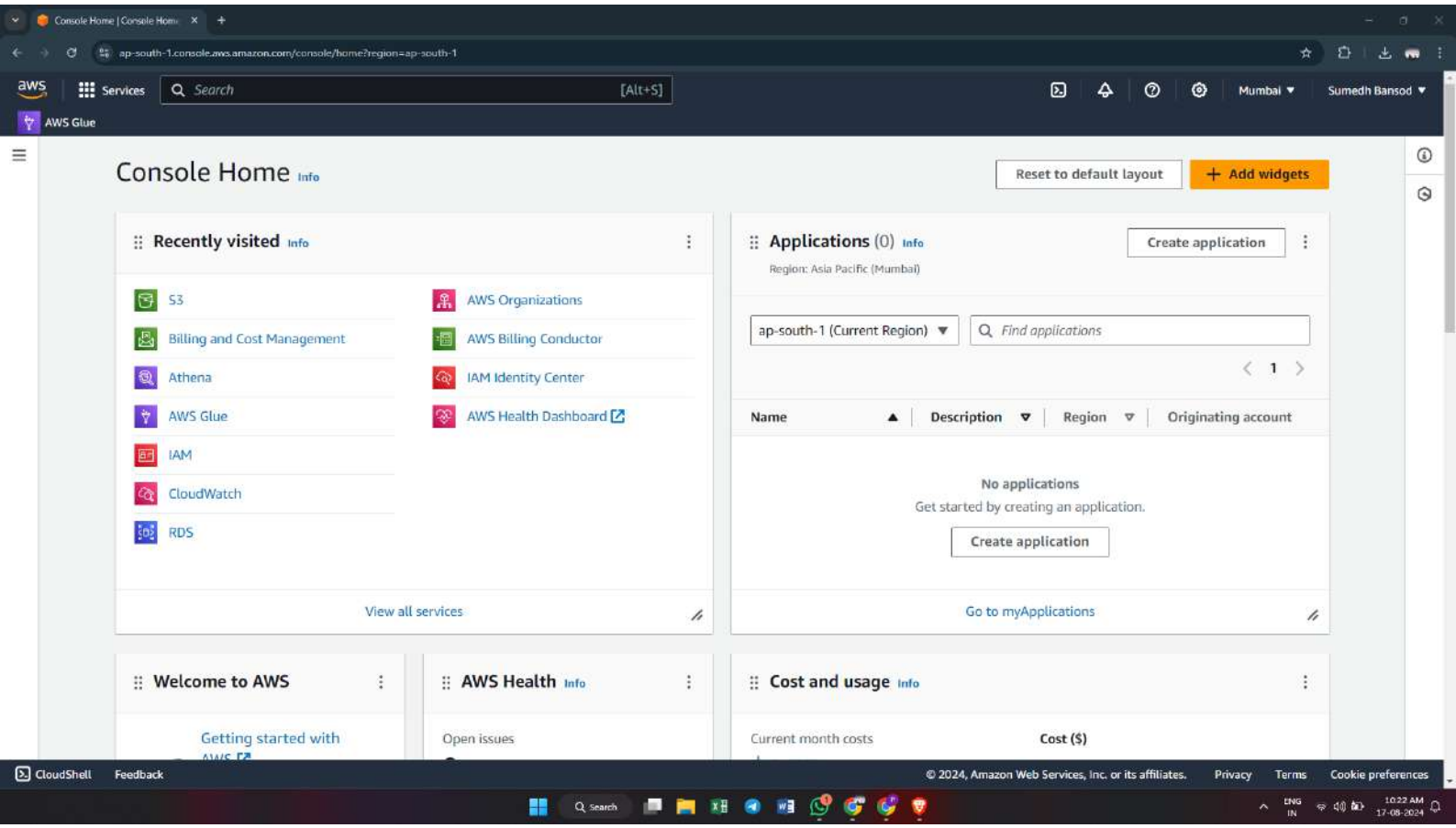
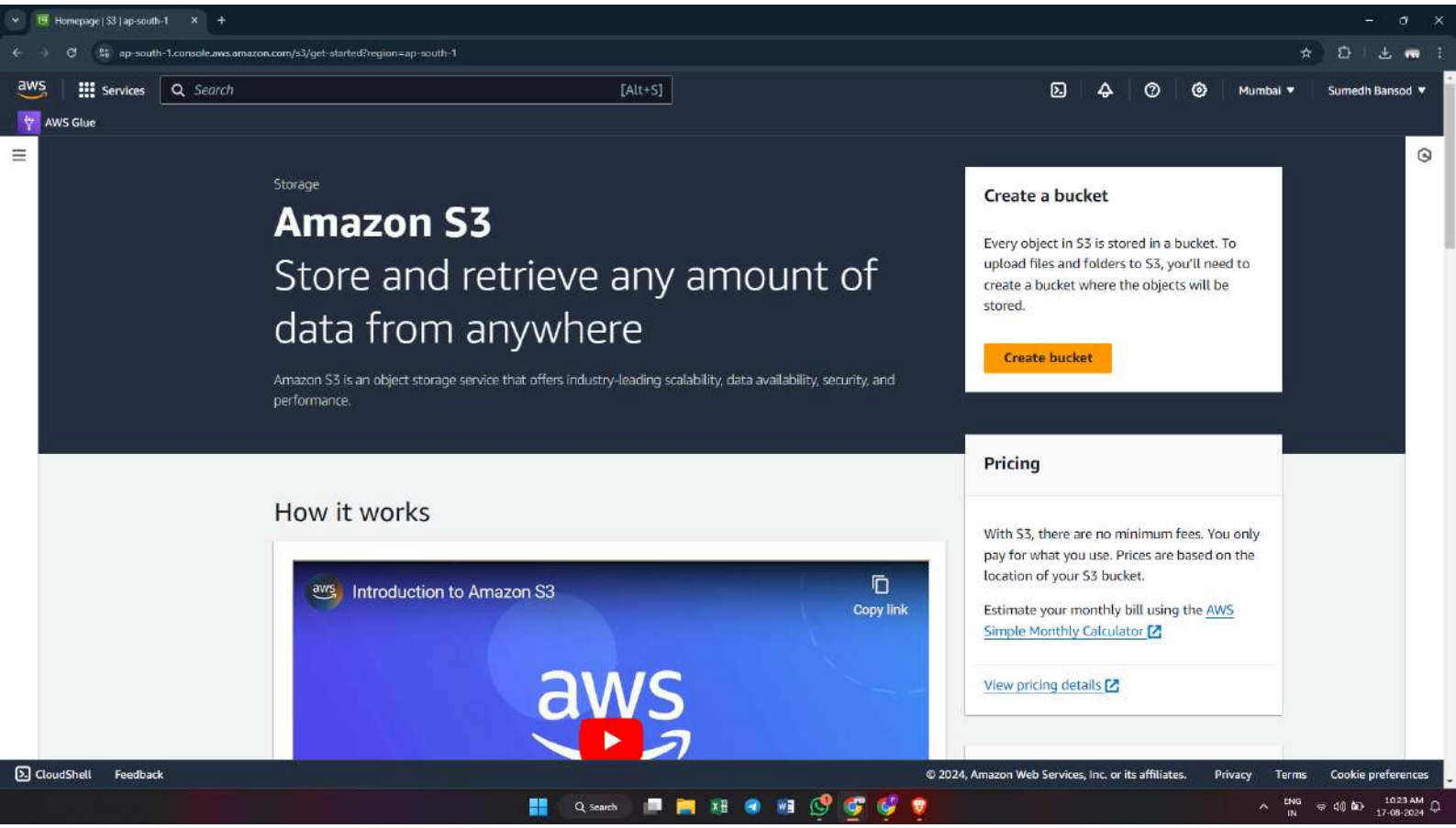
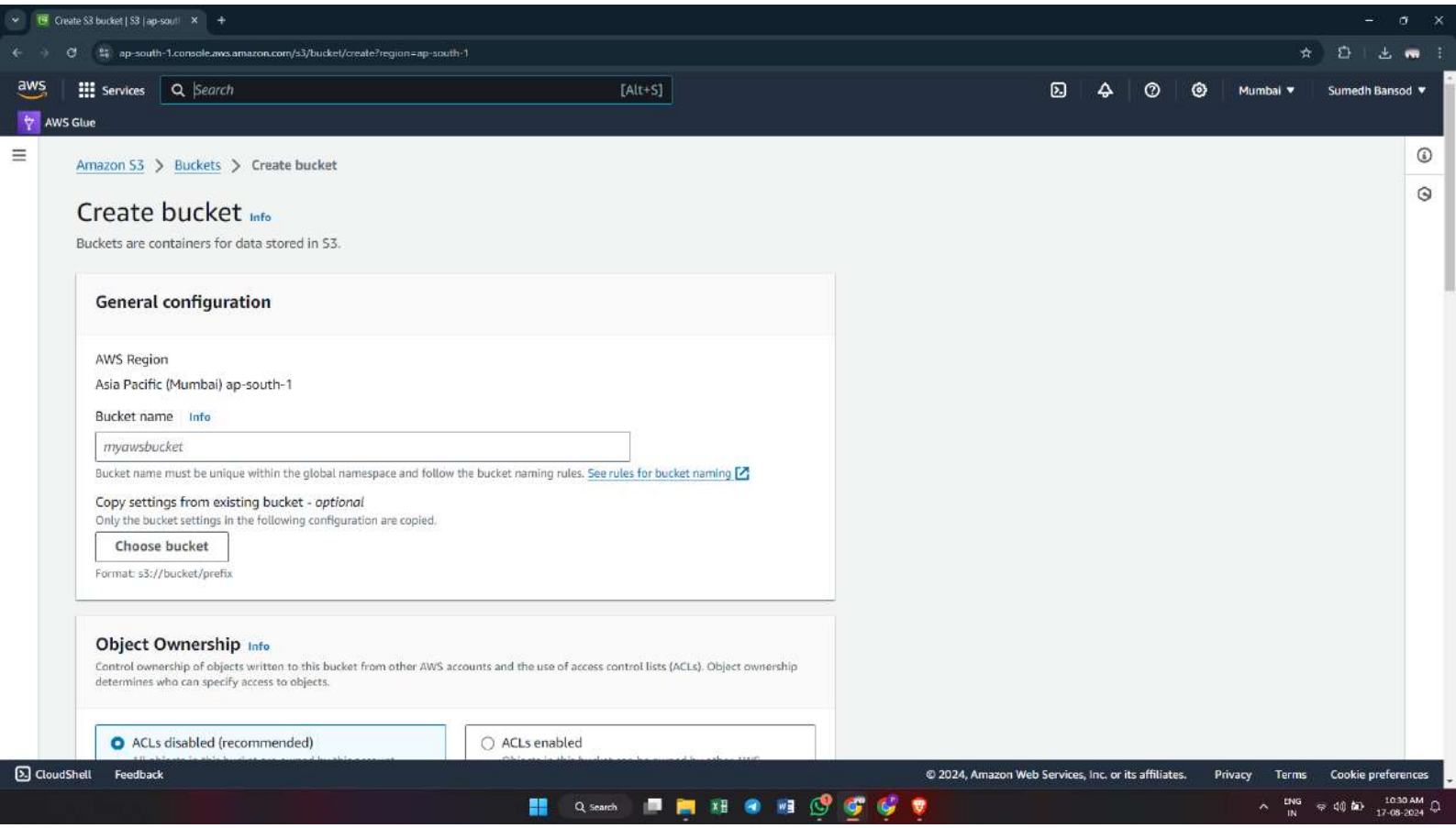
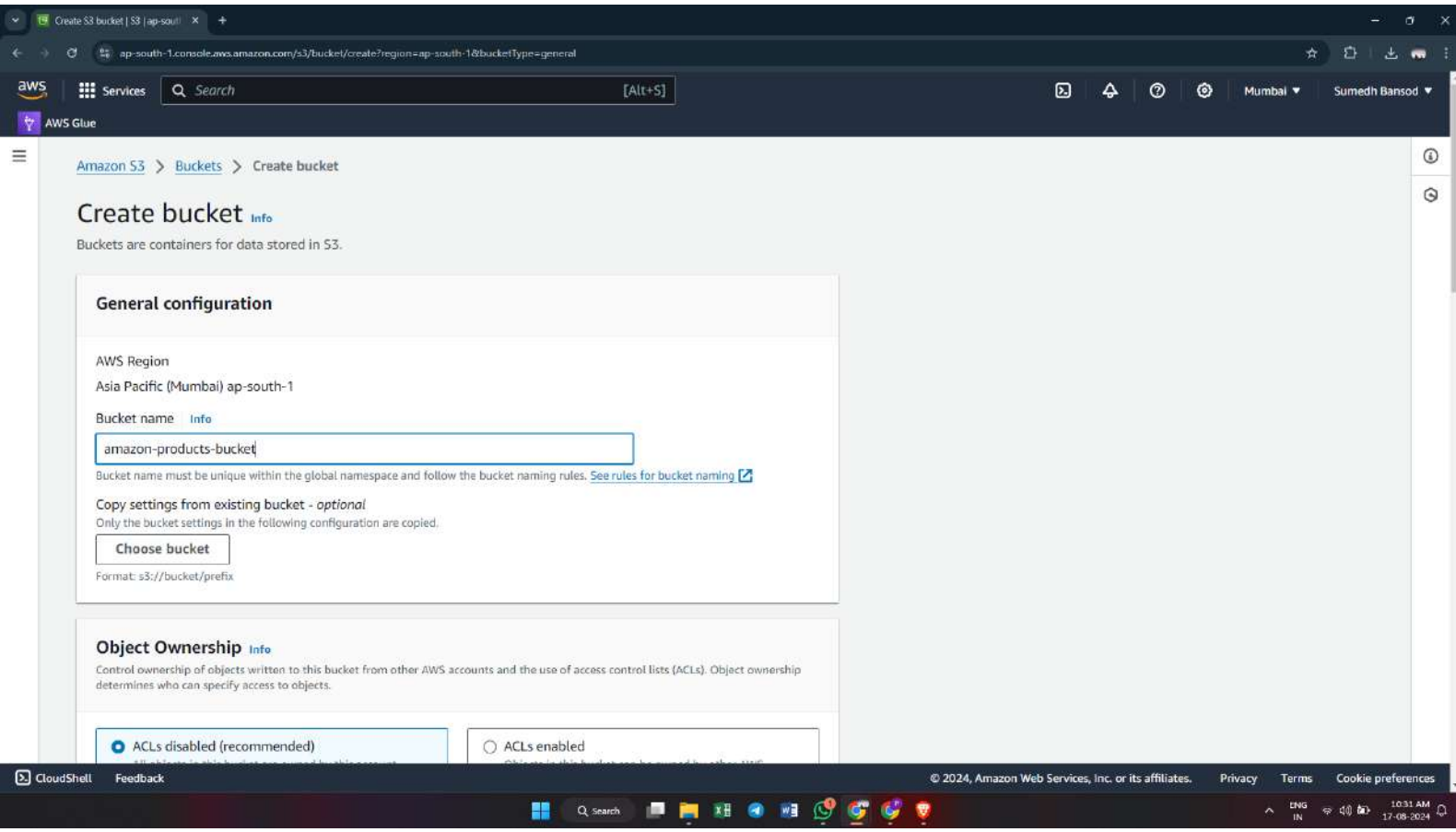
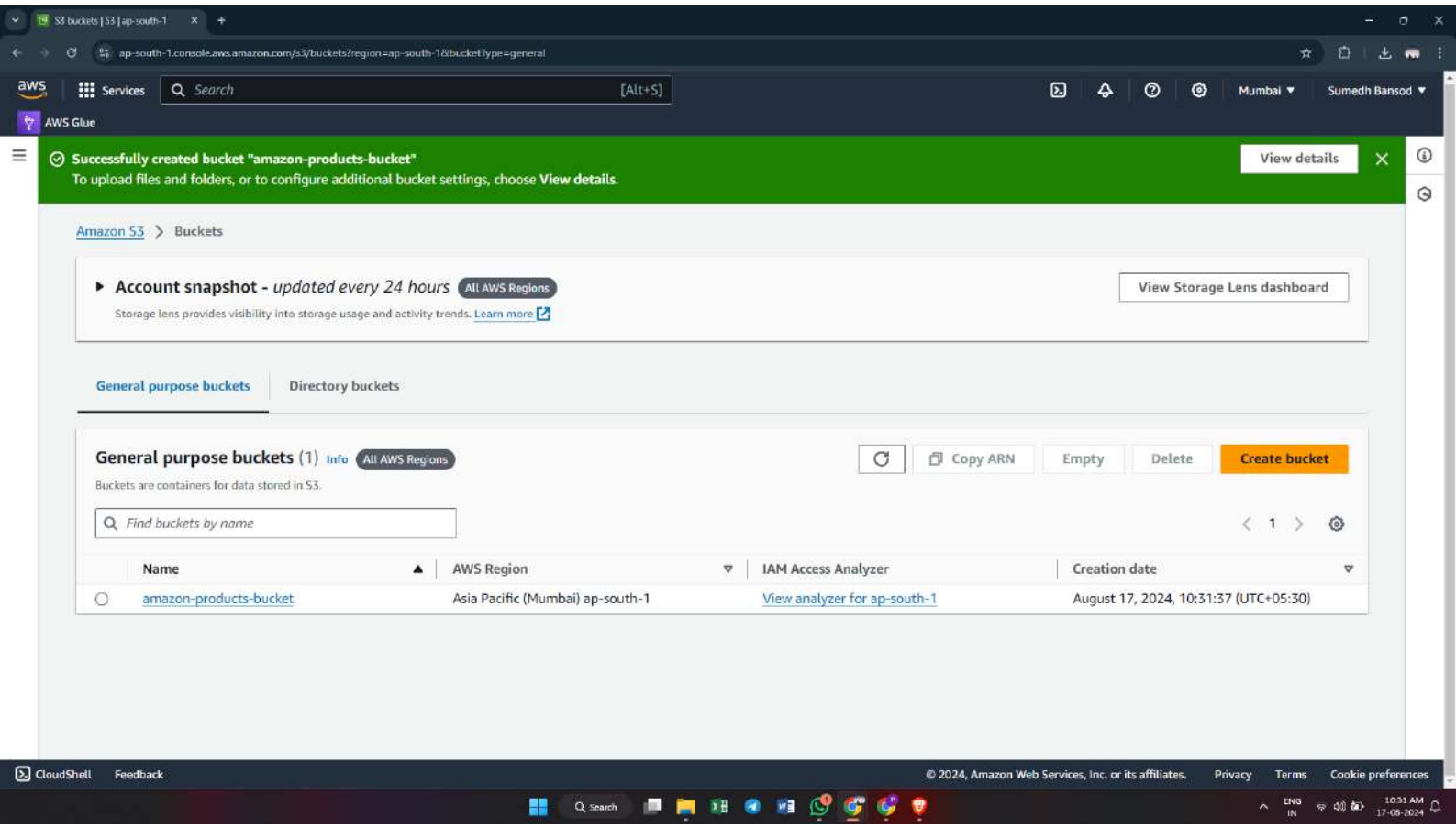
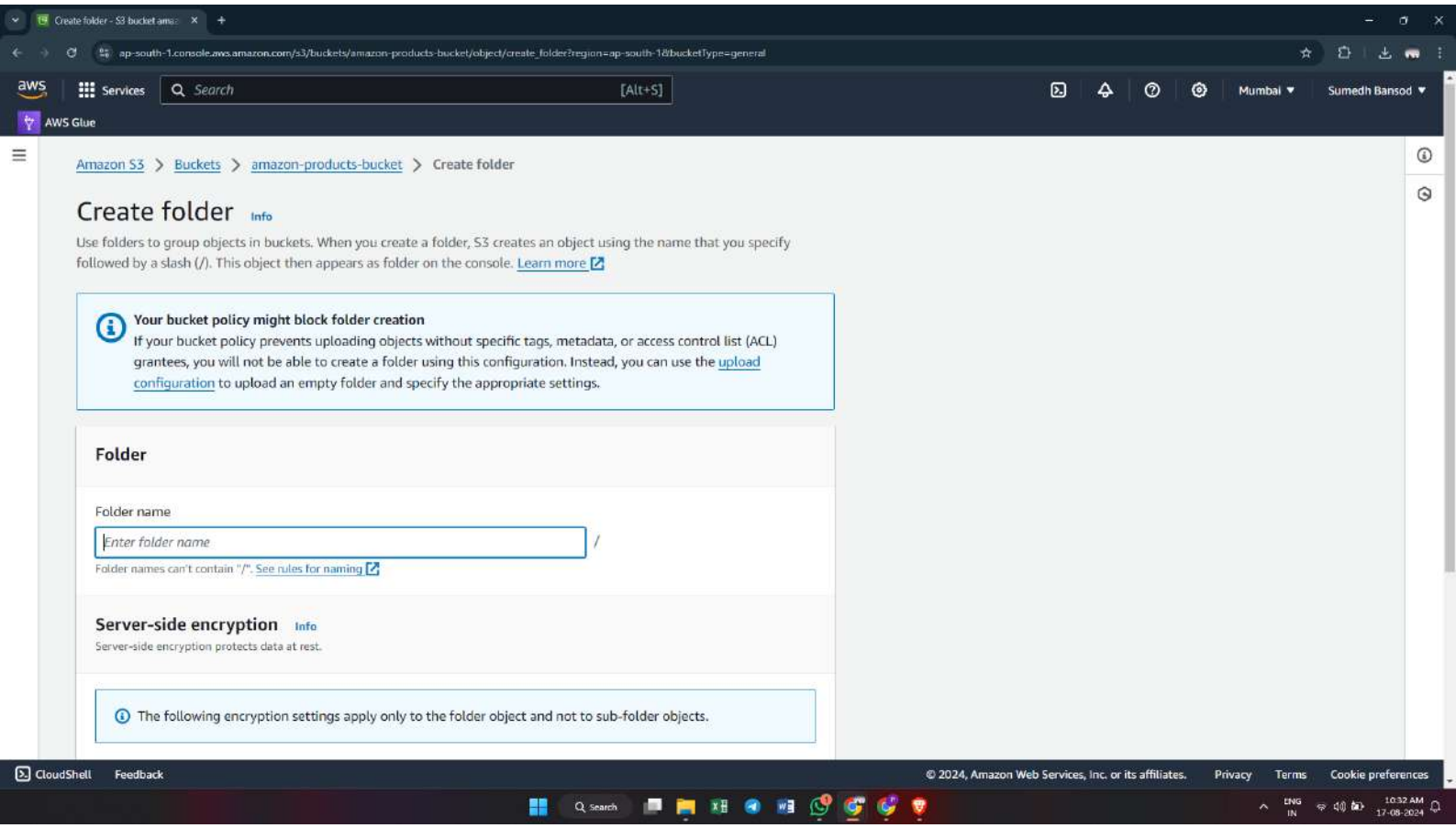

Console Home | Console Hom
ap-south-1.console.aws.amazon.com/console/home?region=ap-south-1
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Console Home Info
Reset to default layout
+ Add widgets
Recently visited Info
S3
Billing and Cost Management
Athena
AWS Glue
IAM
CloudWatch
RDS
AWS Organizations
AWS Billing Conductor
IAM Identity Center
AWS Health Dashboard
View all services
Applications (0) Info
Region: Asia Pacific (Mumbai)
Create application
ap-south-1 (Current Region)
Find applications
1
Name
Description
Region
Originating account
No applications
Get started by creating an application.
Create application
Go to myApplications
Welcome to AWS
Getting started with
AWS
AWS Health Info
Open issues
Cost and usage Info
Current month costs
Cost ($)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:22 AM
17-08-2024

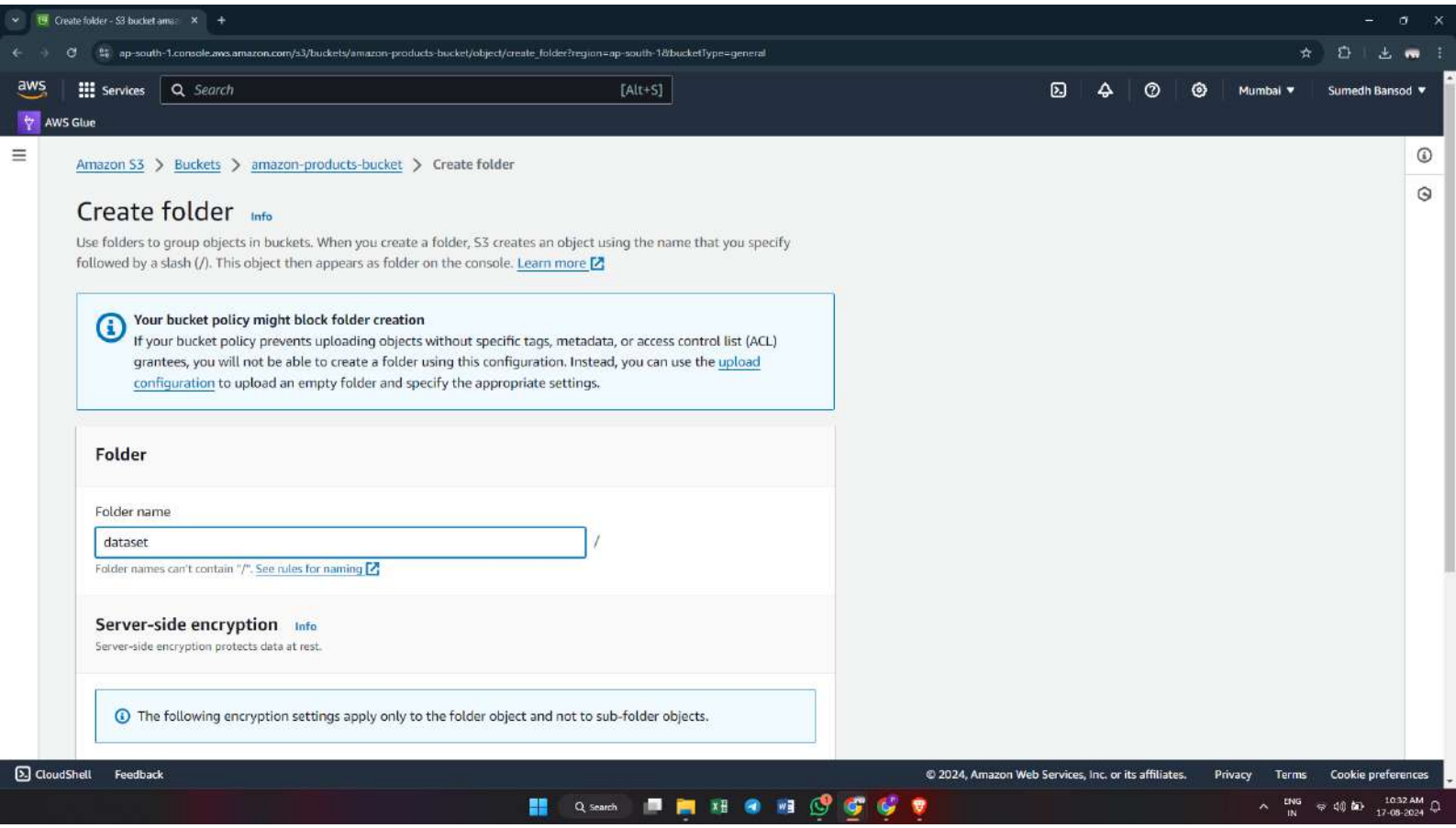

Homepage | S3 | ap-south-1
ap-south-1.console.aws.amazon.com/s3/get-started?region=ap-south-1
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Storage
Amazon S3
Store and retrieve any amount of data from anywhere
Amazon S3 is an object storage service that offers industry-leading scalability, data availability, security, and performance.
Create a bucket
Every object in S3 is stored in a bucket. To upload files and folders to S3, you'll need to create a bucket where the objects will be stored.
Create bucket
Pricing
With S3, there are no minimum fees. You only pay for what you use. Prices are based on the location of your S3 bucket.
Estimate your monthly bill using the AWS Simple Monthly Calculator
View pricing details
How it works
Introduction to Amazon S3
Copy link
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

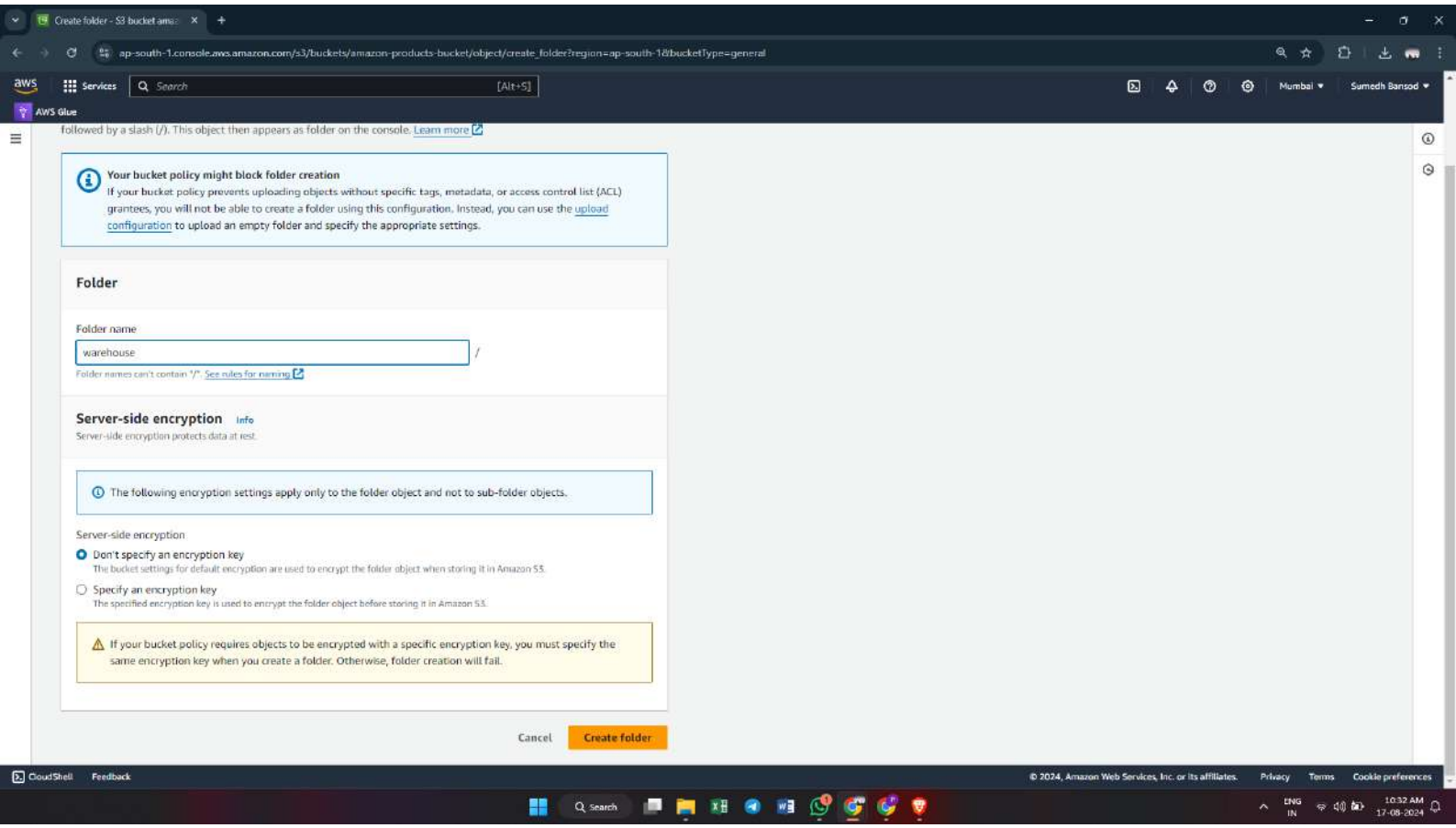

Create S3 bucket | S3 | ap-south-1
ap-south-1.console.aws.amazon.com/s3/bucket/create?region=ap-south-1
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > Create bucket
Create bucket Info
Buckets are containers for data stored in S3.
General configuration
AWS Region
Asia Pacific (Mumbai) ap-south-1
Bucket name Info
myawsbucket
Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming
Copy settings from existing bucket - optional
Only the bucket settings in the following configuration are copied.
Choose bucket
Format: s3://bucket/prefix
Object Ownership Info
Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.
ACLs disabled (recommended)
ACLs enabled
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:30 AM
17-08-2024

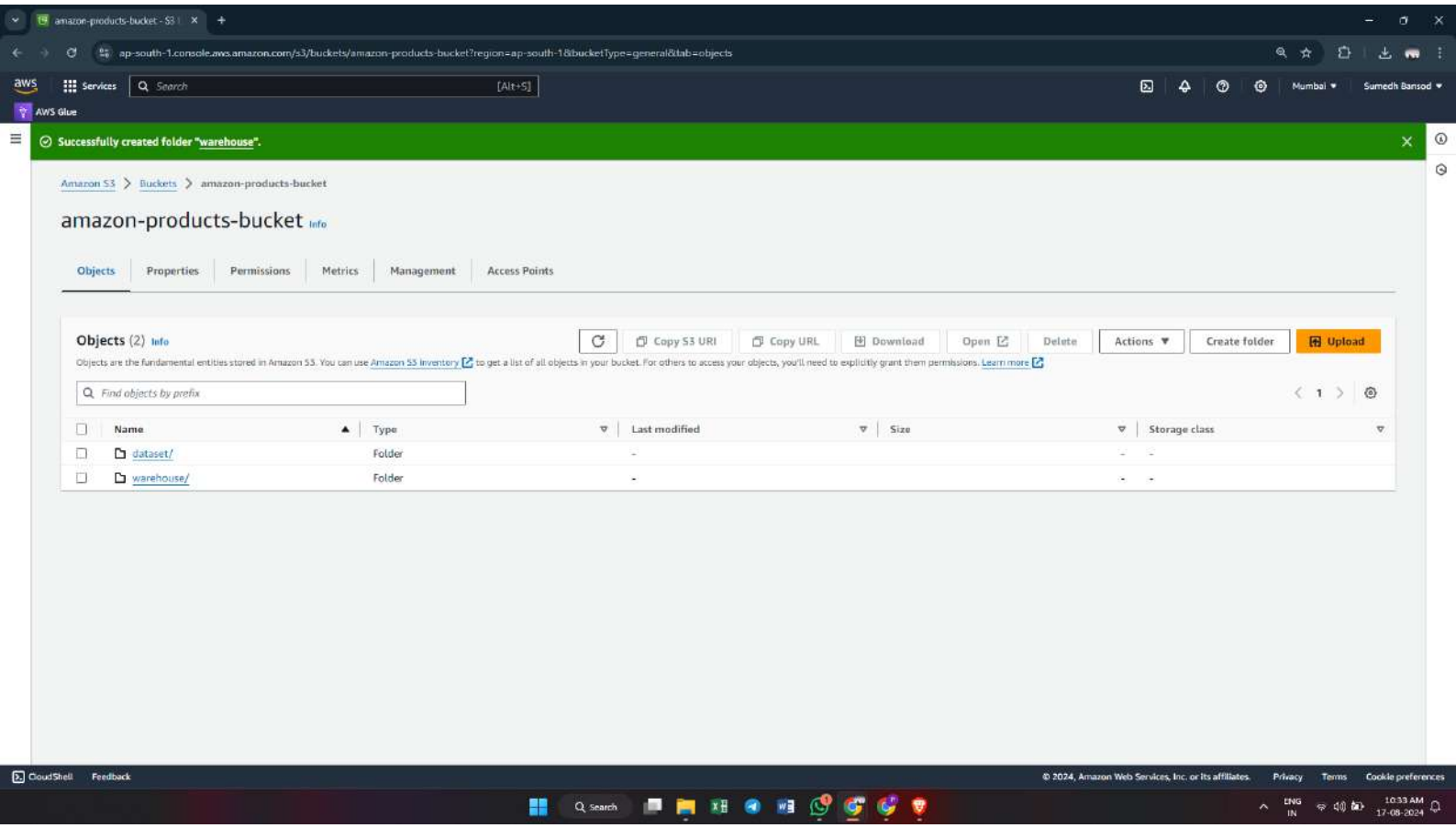

Create S3 bucket | S3 | ap-south
ap-south-1.console.aws.amazon.com/s3/bucket/create?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > Create bucket
Create bucket Info
Buckets are containers for data stored in S3.
General configuration
AWS Region
Asia Pacific (Mumbai) ap-south-1
Bucket name Info
amazon-products-bucket
Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming
Copy settings from existing bucket - optional
Only the bucket settings in the following configuration are copied.
Choose bucket
Format: s3://bucket/prefix
Object Ownership Info
Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.
ACLs disabled (recommended)
ACLs enabled
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:31 AM
17-08-2024

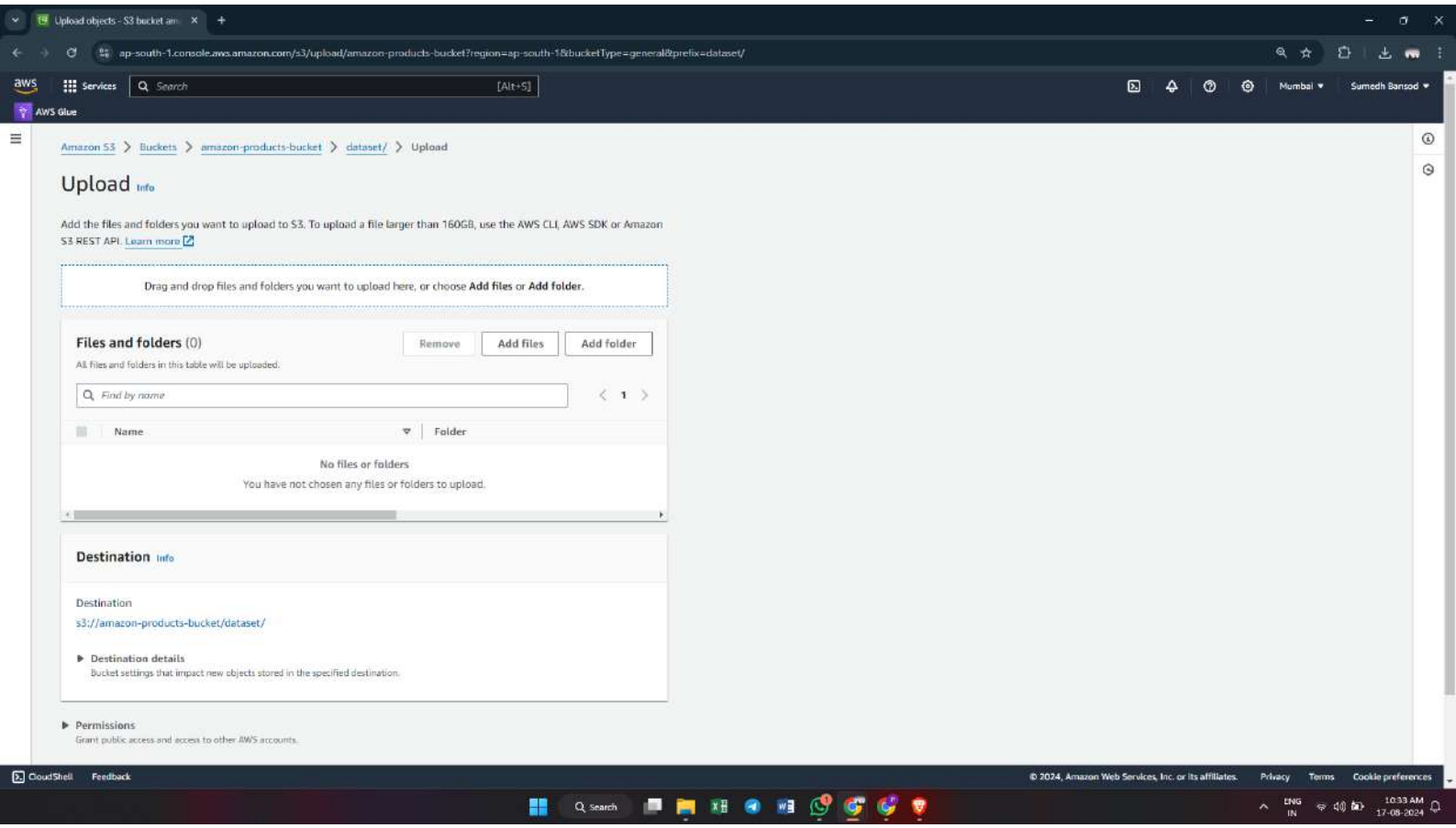

S3 buckets | S3 | ap-south-1
ap-south-1.console.aws.amazon.com/s3/buckets?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Successfully created bucket "amazon-products-bucket"
To upload files and folders, or to configure additional bucket settings, choose View details.
View details
Amazon S3 > Buckets
Account snapshot - updated every 24 hours
All AWS Regions
Storage lens provides visibility into storage usage and activity trends. Learn more
View Storage Lens dashboard
General purpose buckets
Directory buckets
General purpose buckets (1) Info
All AWS Regions
Buckets are containers for data stored in S3.
Copy ARN
Empty
Delete
Create bucket
Find buckets by name
1
Name
AWS Region
IAM Access Analyzer
Creation date
amazon-products-bucket
Asia Pacific (Mumbai) ap-south-1
View analyzer for ap-south-1
August 17, 2024, 10:31:37 (UTC+05:30)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

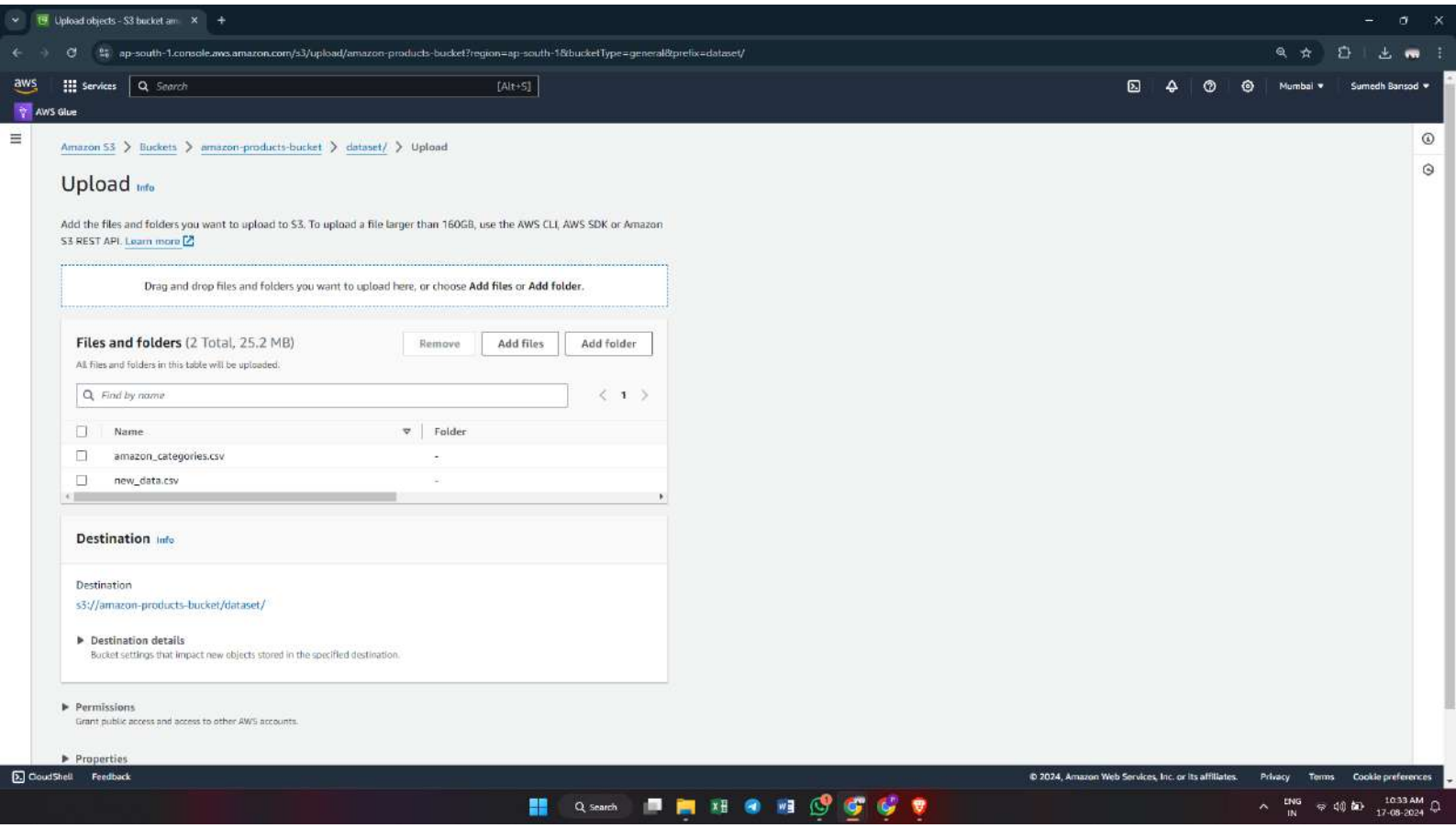

Create folder - S3 bucket ams
ap-south-1.console.aws.amazon.com/s3/buckets/amazon-products-bucket/object/create_folder?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > amazon-products-bucket > Create folder
Create folder Info
Use folders to group objects in buckets. When you create a folder, S3 creates an object using the name that you specify followed by a slash (/). This object then appears as folder on the console. Learn more
Your bucket policy might block folder creation
If your bucket policy prevents uploading objects without specific tags, metadata, or access control list (ACL) grantees, you will not be able to create a folder using this configuration. Instead, you can use the upload configuration to upload an empty folder and specify the appropriate settings.
Folder
Folder name
Enter folder name
/
Folder names can't contain "/". See rules for naming
Server-side encryption Info
Server-side encryption protects data at rest.
The following encryption settings apply only to the folder object and not to sub-folder objects.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:32 AM
17-08-2024

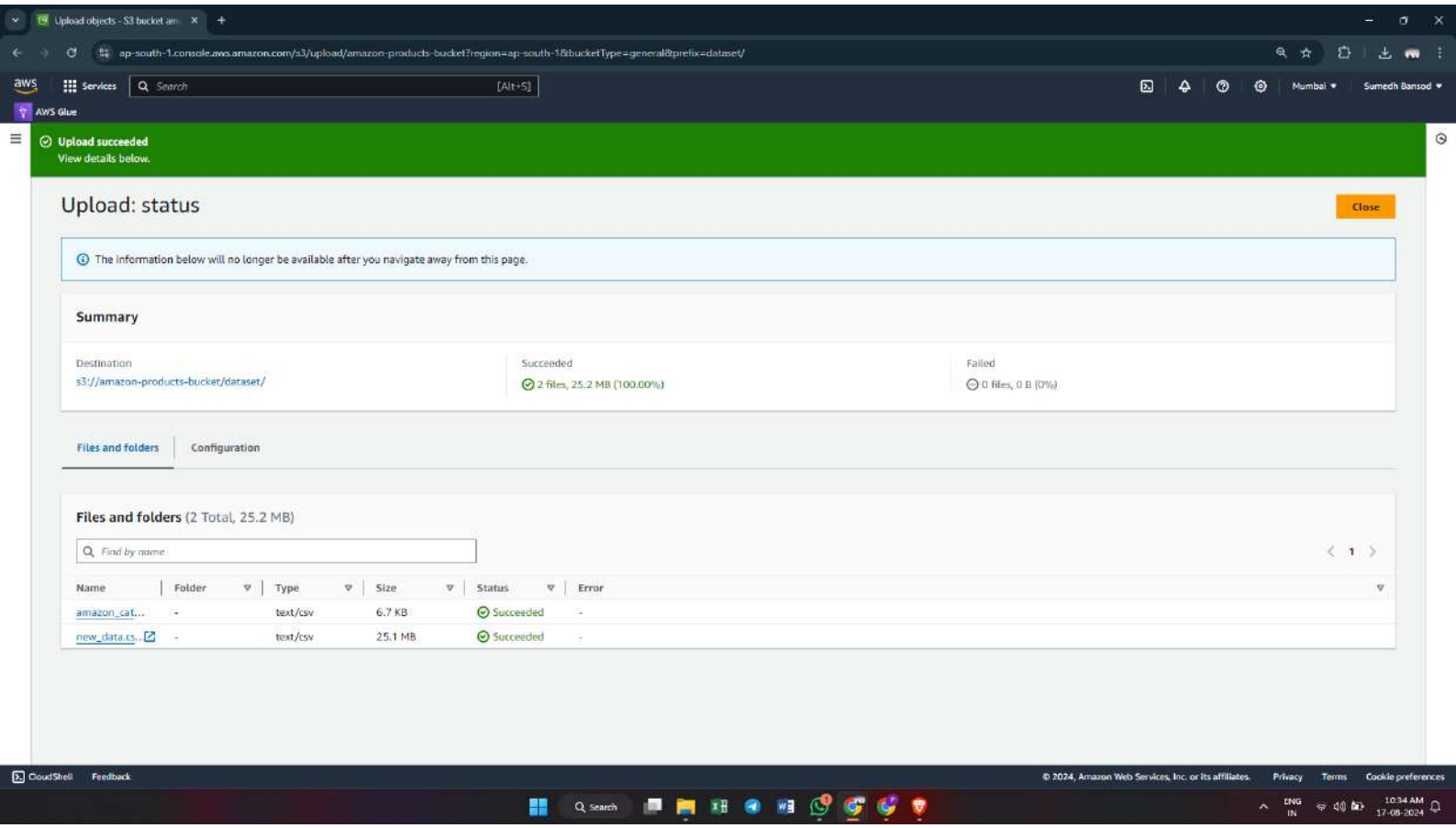

Create folder - S3 bucket ams
ap-south-1.console.aws.amazon.com/s3/buckets/amazon-products-bucket/object/create_folder?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > amazon-products-bucket > Create folder
Create folder Info
Use folders to group objects in buckets. When you create a folder, S3 creates an object using the name that you specify followed by a slash (/). This object then appears as folder on the console. Learn more
Your bucket policy might block folder creation
If your bucket policy prevents uploading objects without specific tags, metadata, or access control list (ACL) grantees, you will not be able to create a folder using this configuration. Instead, you can use the upload configuration to upload an empty folder and specify the appropriate settings.
Folder
Folder name
dataset
/
Folder names can't contain "/". See rules for naming
Server-side encryption Info
Server-side encryption protects data at rest.
The following encryption settings apply only to the folder object and not to sub-folder objects.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:32 AM
17-08-2024

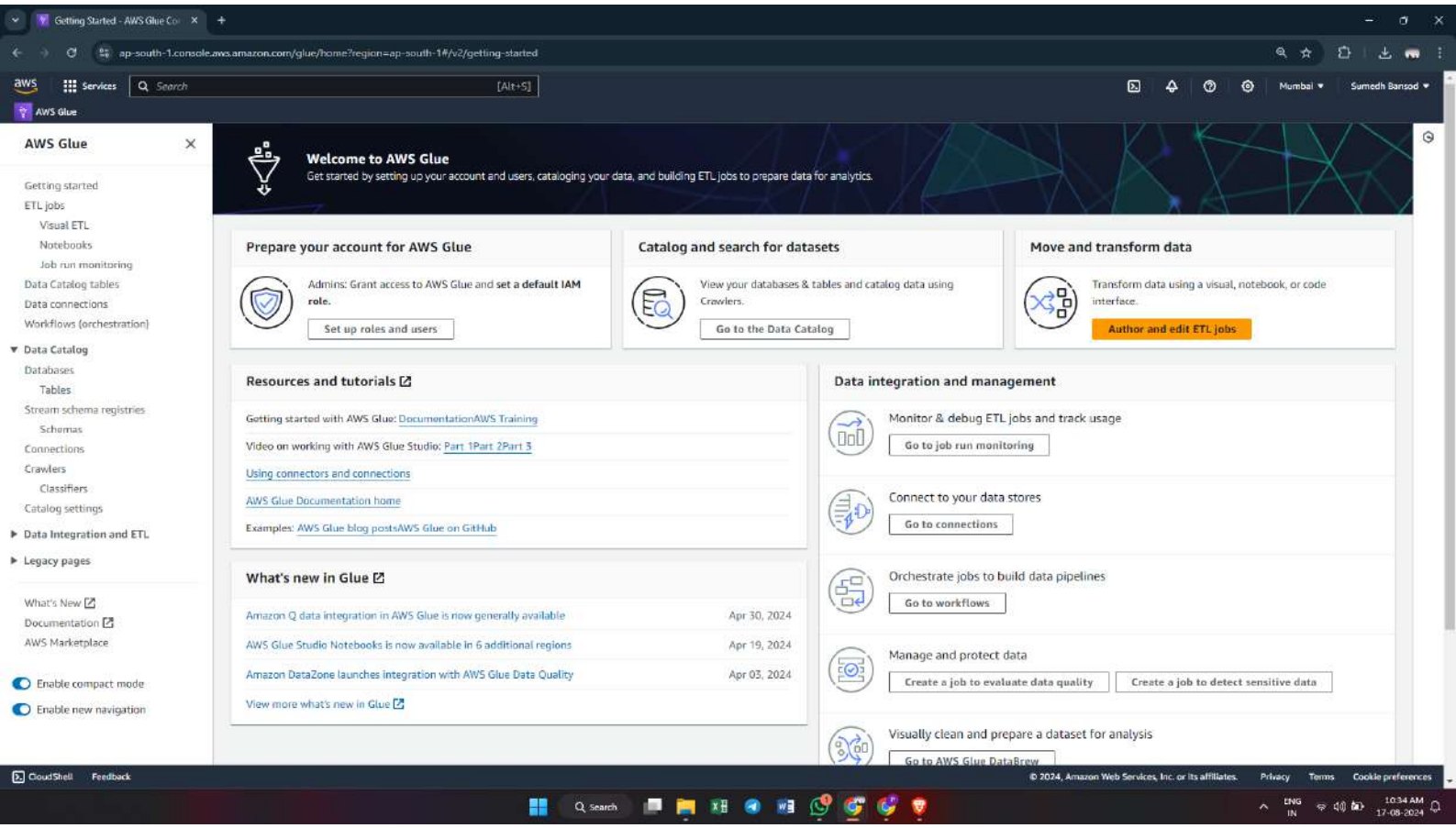

Create folder - S3 bucket ams
ap-south-1.console.aws.amazon.com/s3/buckets/amazon-products-bucket/object/create_folder?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
followed by a slash (/). This object then appears as folder on the console. Learn more
Your bucket policy might block folder creation
If your bucket policy prevents uploading objects without specific tags, metadata, or access control list (ACL) grantees, you will not be able to create a folder using this configuration. Instead, you can use the upload configuration to upload an empty folder and specify the appropriate settings.
Folder
Folder name
warehouse
/
Folder names can't contain "/". See rules for naming
Server-side encryption
Info
Server-side encryption protects data at rest.
The following encryption settings apply only to the folder object and not to sub-folder objects.
Server-side encryption
Don't specify an encryption key
The bucket settings for default encryption are used to encrypt the folder object when storing it in Amazon S3.
Specify an encryption key
The specified encryption key is used to encrypt the folder object before storing it in Amazon S3.
If your bucket policy requires objects to be encrypted with a specific encryption key, you must specify the same encryption key when you create a folder. Otherwise, folder creation will fail.
Cancel
Create folder
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:32 AM
17-08-2024

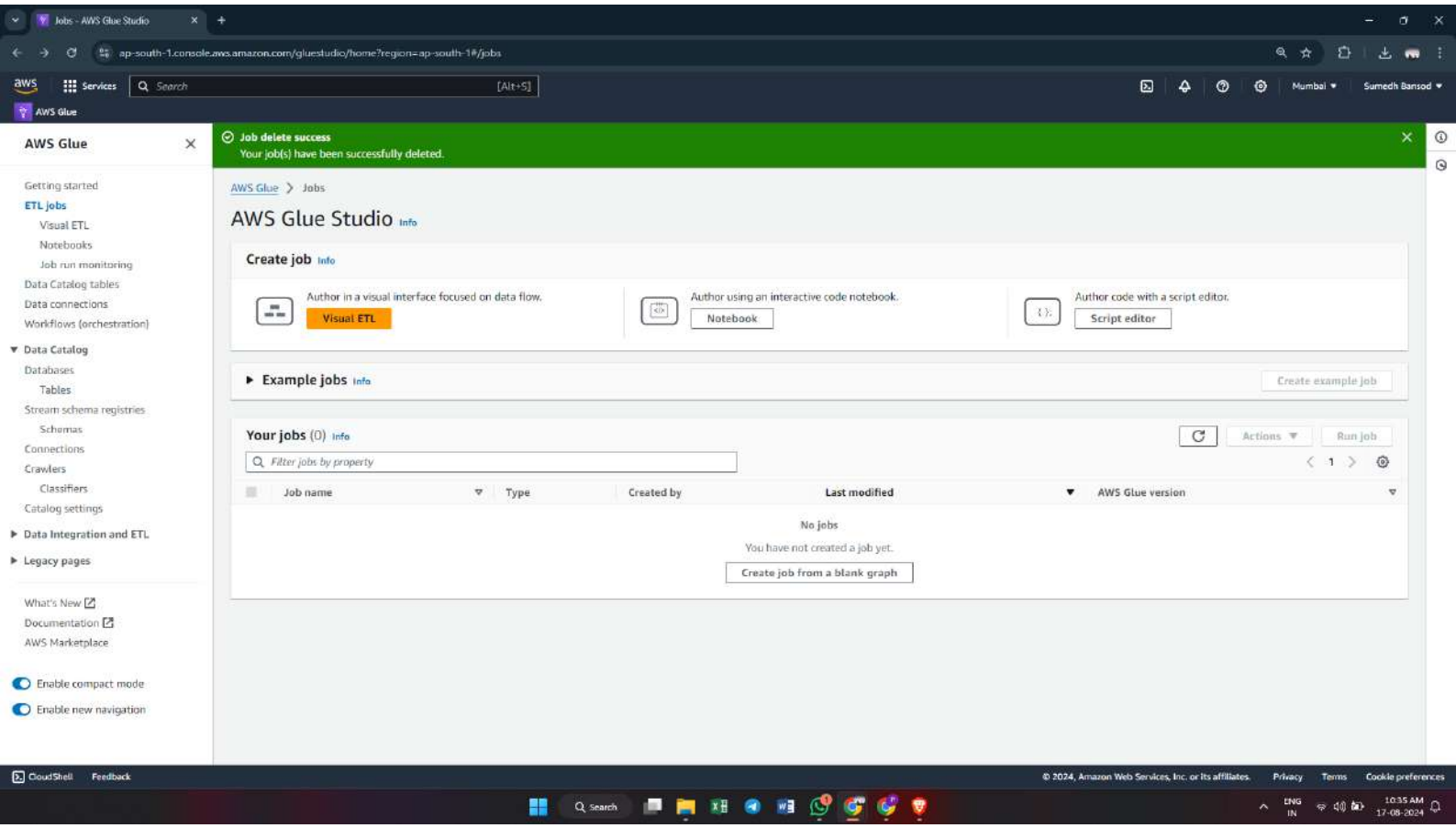

amazon-products-bucket - S3
ap-south-1.console.aws.amazon.com/s3/buckets/amazon-products-bucket?region=ap-south-1&bucketType=general&tab=objects
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Successfully created folder "warehouse".
Amazon S3 > Buckets > amazon-products-bucket
amazon-products-bucket Info
Objects
Properties
Permissions
Metrics
Management
Access Points
Objects (2) Info
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
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
Find objects by prefix
Name
Type
Last modified
Size
Storage class
dataset/
Folder
warehouse/
Folder
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:33 AM
17-08-2024

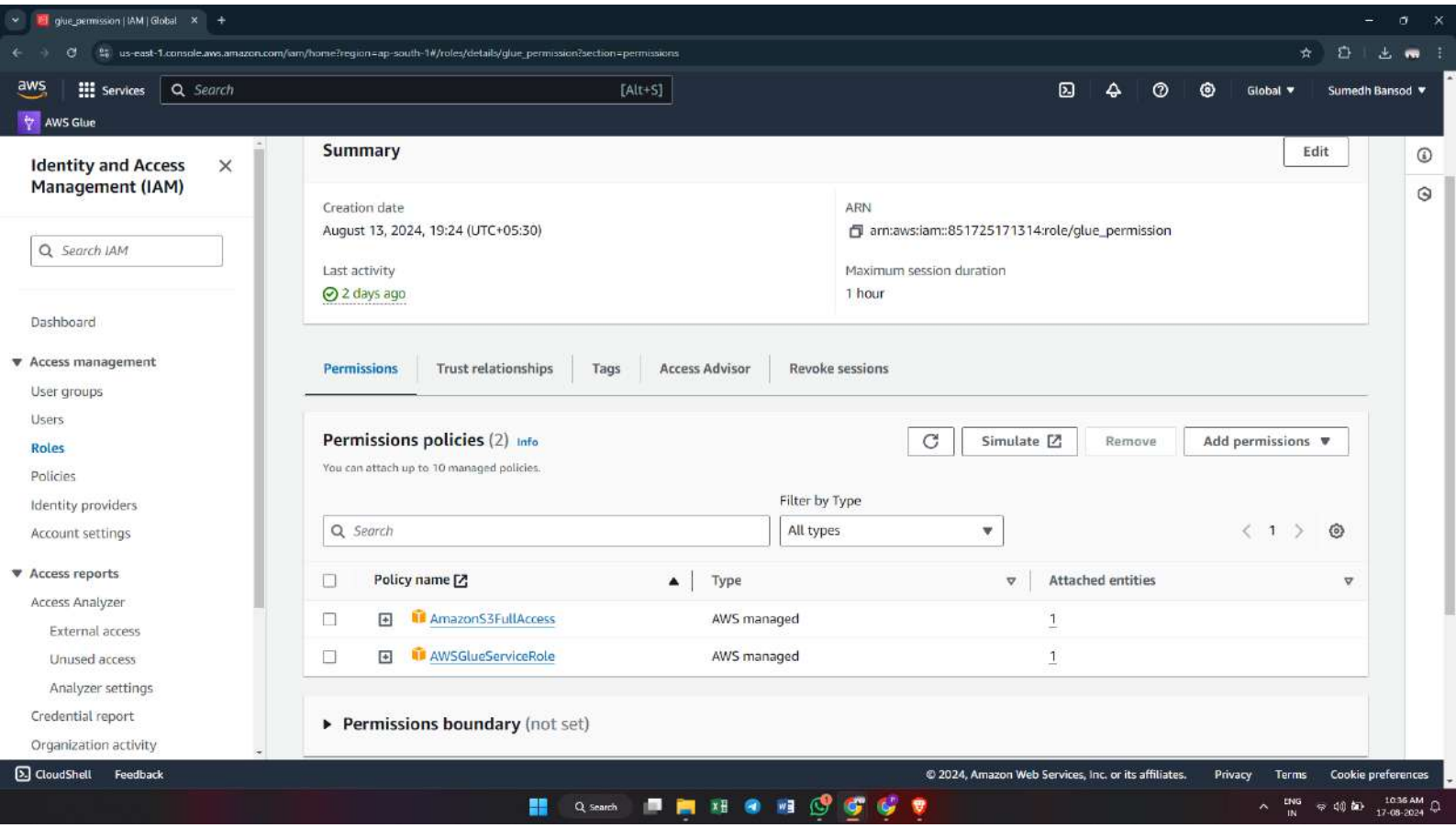

Upload objects - S3 bucket am
ap-south-1.console.aws.amazon.com/s3/upload/amazon-products-bucket?region=ap-south-1&bucketType=general&prefix=dataset/
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > amazon-products-bucket > dataset/ > Upload
Upload Info
Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. Learn more
Drag and drop files and folders you want to upload here, or choose Add files or Add folder.
Files and folders (0)
Remove
Add files
Add folder
All files and folders in this table will be uploaded.
Find by name
1
Name
Folder
No files or folders
You have not chosen any files or folders to upload.
Destination Info
Destination
s3://amazon-products-bucket/dataset/
Destination details
Bucket settings that impact new objects stored in the specified destination.
Permissions
Grant public access and access to other AWS accounts.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:33 AM
17-08-2024

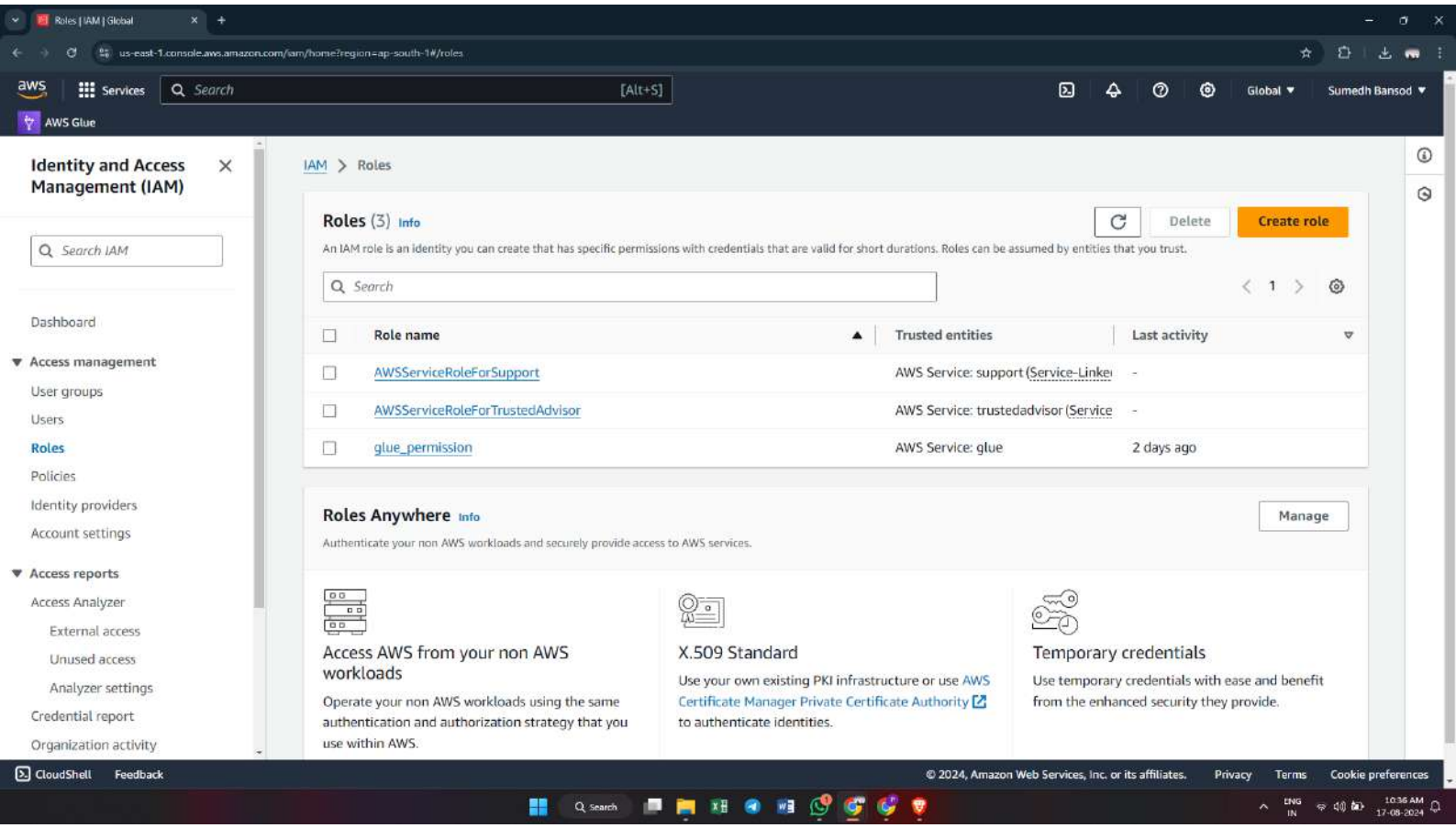

Upload objects - S3 bucket am
ap-south-1.console.aws.amazon.com/s3/upload/amazon-products-bucket?region=ap-south-1&bucketType=general&prefix=dataset/
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3 > Buckets > amazon-products-bucket > dataset/ > Upload
Upload Info
Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. Learn more
Drag and drop files and folders you want to upload here, or choose Add files or Add folder.
Files and folders (2 Total, 25.2 MB)
Remove
Add files
Add folder
All files and folders in this table will be uploaded.
Find by name
1
Name
Folder
amazon_categories.csv
-
new_data.csv
-
Destination Info
Destination
s3://amazon-products-bucket/dataset/
Destination details
Bucket settings that impact new objects stored in the specified destination.
Permissions
Grant public access and access to other AWS accounts.
Properties
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:33 AM
17-08-2024

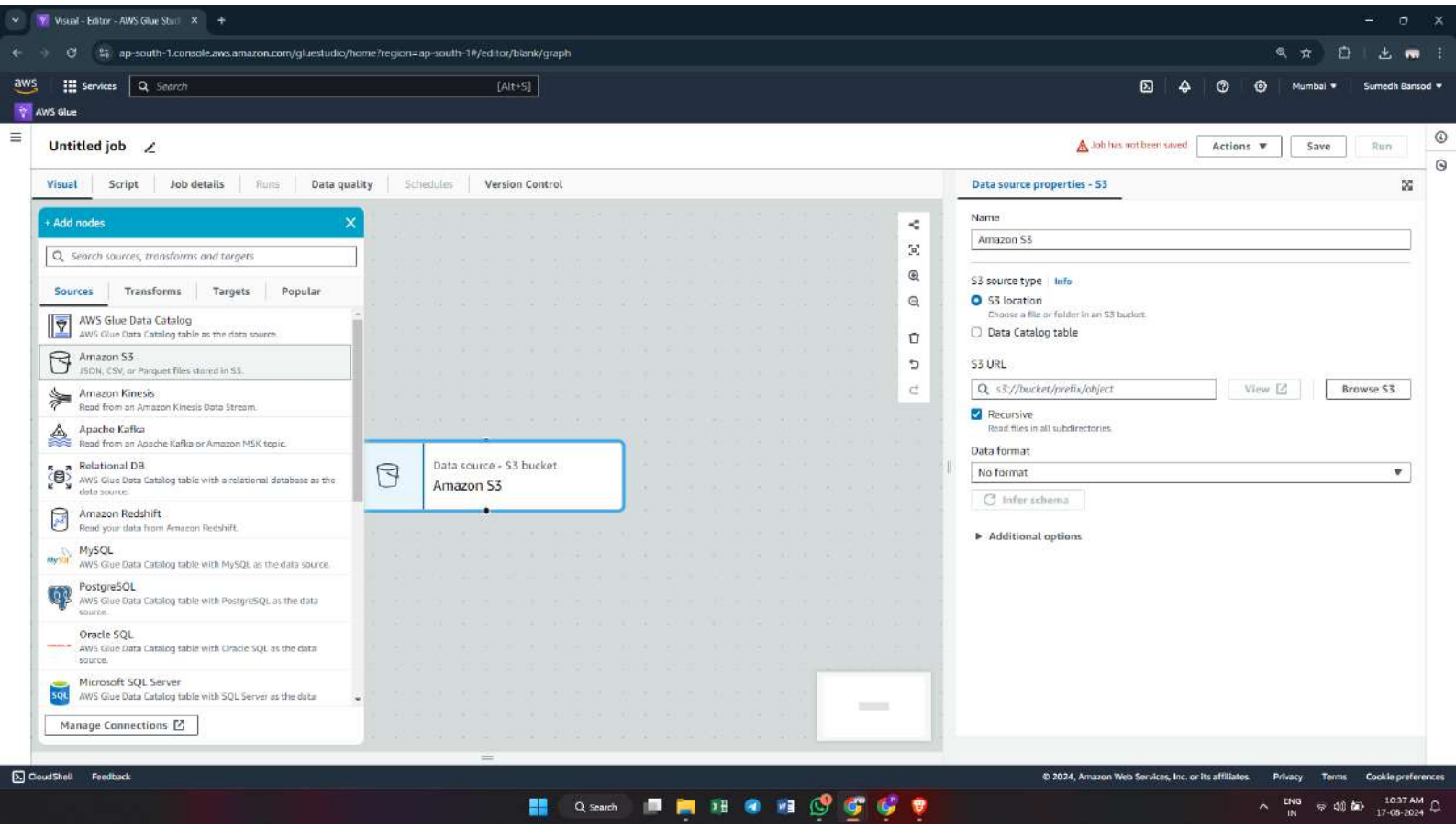

Upload objects - S3 bucket am
ap-south-1.console.aws.amazon.com/s3/upload/amazon-products-bucket?region=ap-south-1&bucketType=general&prefix=dataset/
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Upload succeeded
View details below.
Upload: status
Close
The information below will no longer be available after you navigate away from this page.
Summary
Destination
s3://amazon-products-bucket/dataset/
Succeeded
2 files, 25.2 MB (100.00%)
Failed
0 files, 0 B (0%)
Files and folders
Configuration
Files and folders (2 Total, 25.2 MB)
Find by name
1
Name
Folder
Type
Size
Status
Error
amazon_cat...
-
text/csv
6.7 KB
Succeeded
-
new_data.cs...
-
text/csv
25.1 MB
Succeeded
-
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:34 AM
17-08-2024

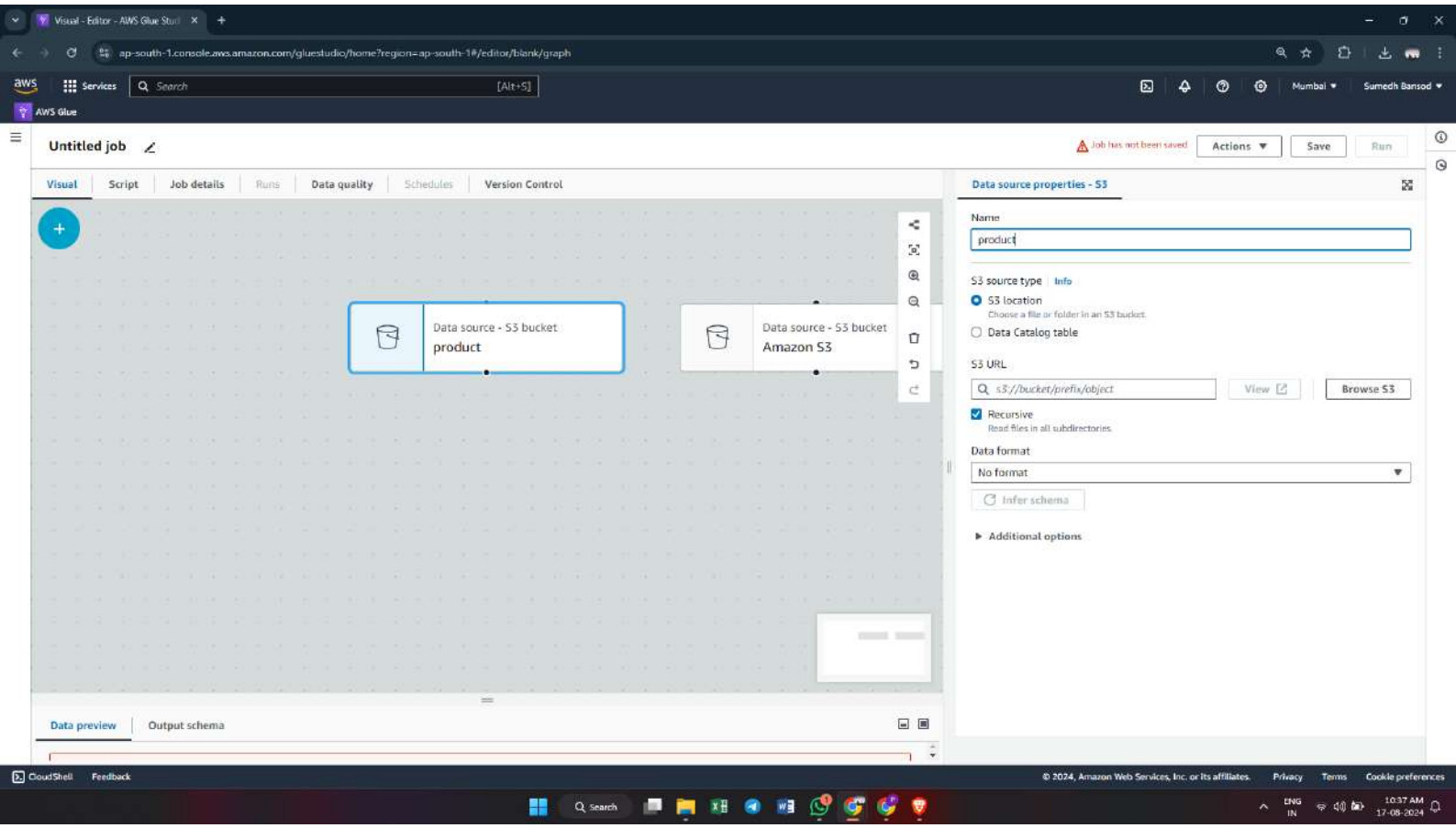

Getting Started - AWS Glue Co
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/getting-started
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Welcome to AWS Glue
Get started by setting up your account and users, cataloging your data, and building ETL jobs to prepare data for analytics.
Prepare your account for AWS Glue
Admins: Grant access to AWS Glue and set a default IAM role.
Set up roles and users
Catalog and search for datasets
View your databases & tables and catalog data using Crawlers.
Go to the Data Catalog
Move and transform data
Transform data using a visual, notebook, or code interface.
Author and edit ETL jobs
Resources and tutorials
Getting started with AWS Glue: DocumentationAWS Training
Video on working with AWS Glue Studio: Part 1Part 2Part 3
Using connectors and connections
AWS Glue Documentation home
Examples: AWS Glue blog postsAWS Glue on GitHub
What's new in Glue
Amazon Q data integration in AWS Glue is now generally available
Apr 30, 2024
AWS Glue Studio Notebooks is now available in 6 additional regions
Apr 19, 2024
Amazon DataZone launches integration with AWS Glue Data Quality
Apr 03, 2024
View more what's new in Glue
Data integration and management
Monitor & debug ETL jobs and track usage
Go to job run monitoring
Connect to your data stores
Go to connections
Orchestrate jobs to build data pipelines
Go to workflows
Manage and protect data
Create a job to evaluate data quality
Create a job to detect sensitive data
Visually clean and prepare a dataset for analysis
Go to AWS Glue DataBrew
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:34 AM
17-08-2024

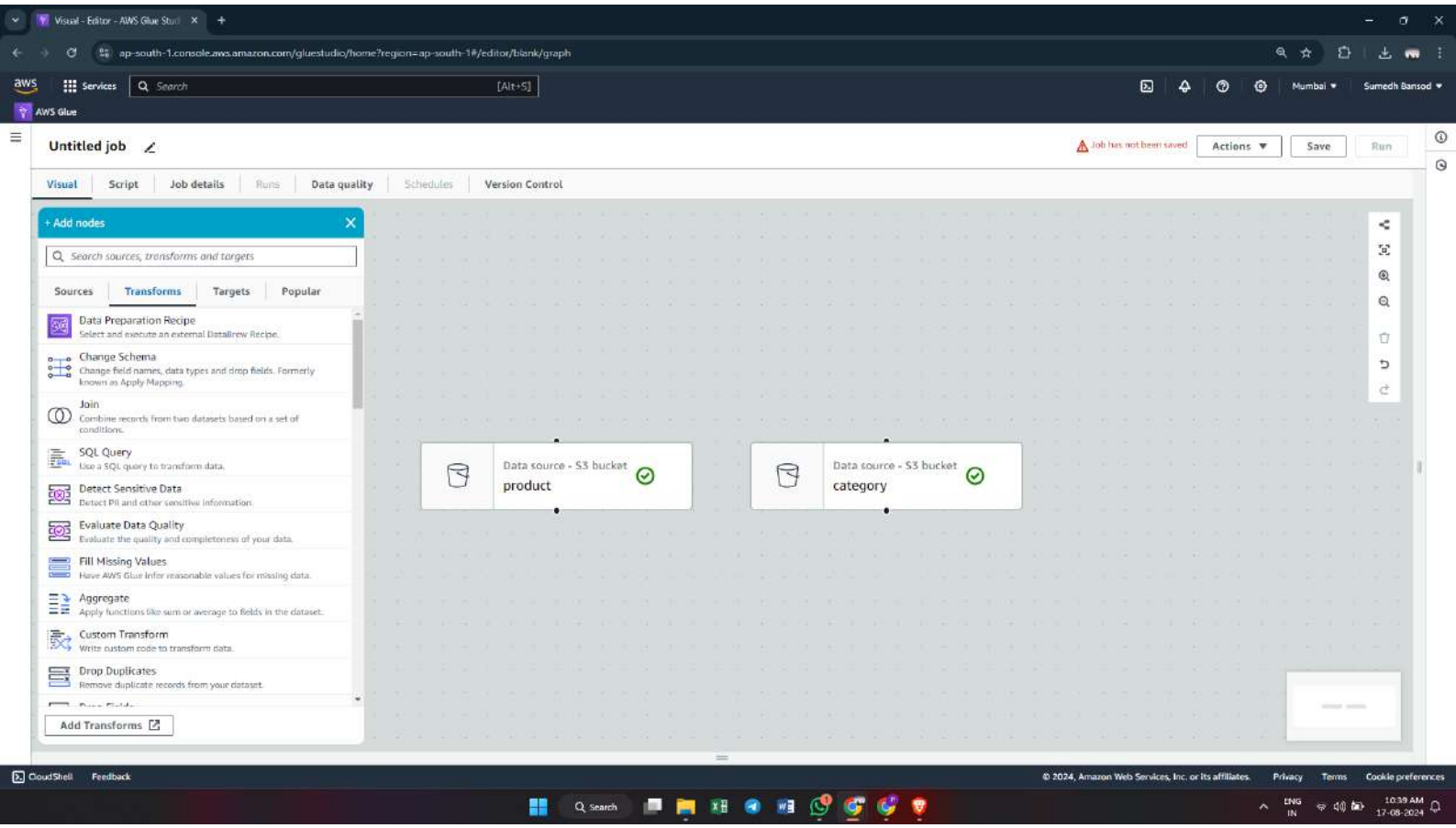

Jobs - AWS Glue Studio
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/jobs
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Job delete success
Your job(s) have been successfully deleted.
AWS Glue > Jobs
AWS Glue Studio Info
Create job Info
Author in a visual interface focused on data flow.
Visual ETL
Author using an interactive code notebook.
Notebook
Author code with a script editor.
Script editor
Example jobs Info
Create example job
Your jobs (0) Info
Actions
Run job
Filter jobs by property
Job name
Type
Created by
Last modified
AWS Glue version
No jobs
You have not created a job yet.
Create job from a blank graph
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:35 AM
17-08-2024

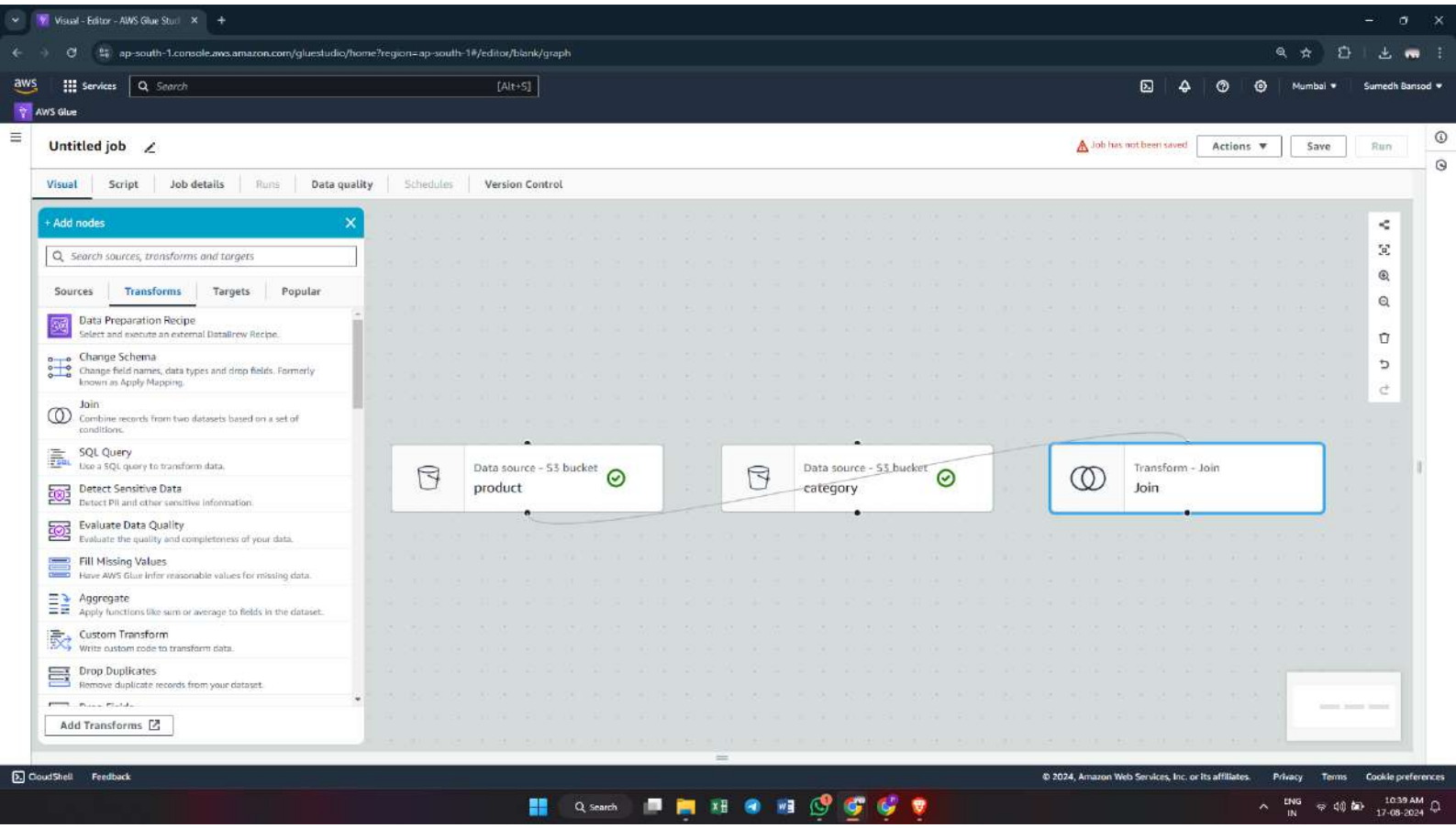

glue_permission | IAM | Global
us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/details/glue_permission?section=permissions
aws
Services
Search
[Alt+S]
Global
Sumedh Bansod
AWS Glue
Identity and Access Management (IAM)
Search IAM
Dashboard
Access management
User groups
Users
Roles
Policies
Identity providers
Account settings
Access reports
Access Analyzer
External access
Unused access
Analyzer settings
Credential report
Organization activity
Summary
Edit
Creation date
August 13, 2024, 19:24 (UTC+05:30)
ARN
arn:aws:iam::851725171314:role/glue_permission
Last activity
2 days ago
Maximum session duration
1 hour
Permissions
Trust relationships
Tags
Access Advisor
Revoke sessions
Permissions policies (2) Info
You can attach up to 10 managed policies.
Simulate
Remove
Add permissions
Filter by Type
Search
All types
1
Policy name
Type
Attached entities
AmazonS3FullAccess
AWS managed
1
AWSGlueServiceRole
AWS managed
1
Permissions boundary (not set)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:36 AM
17-08-2024

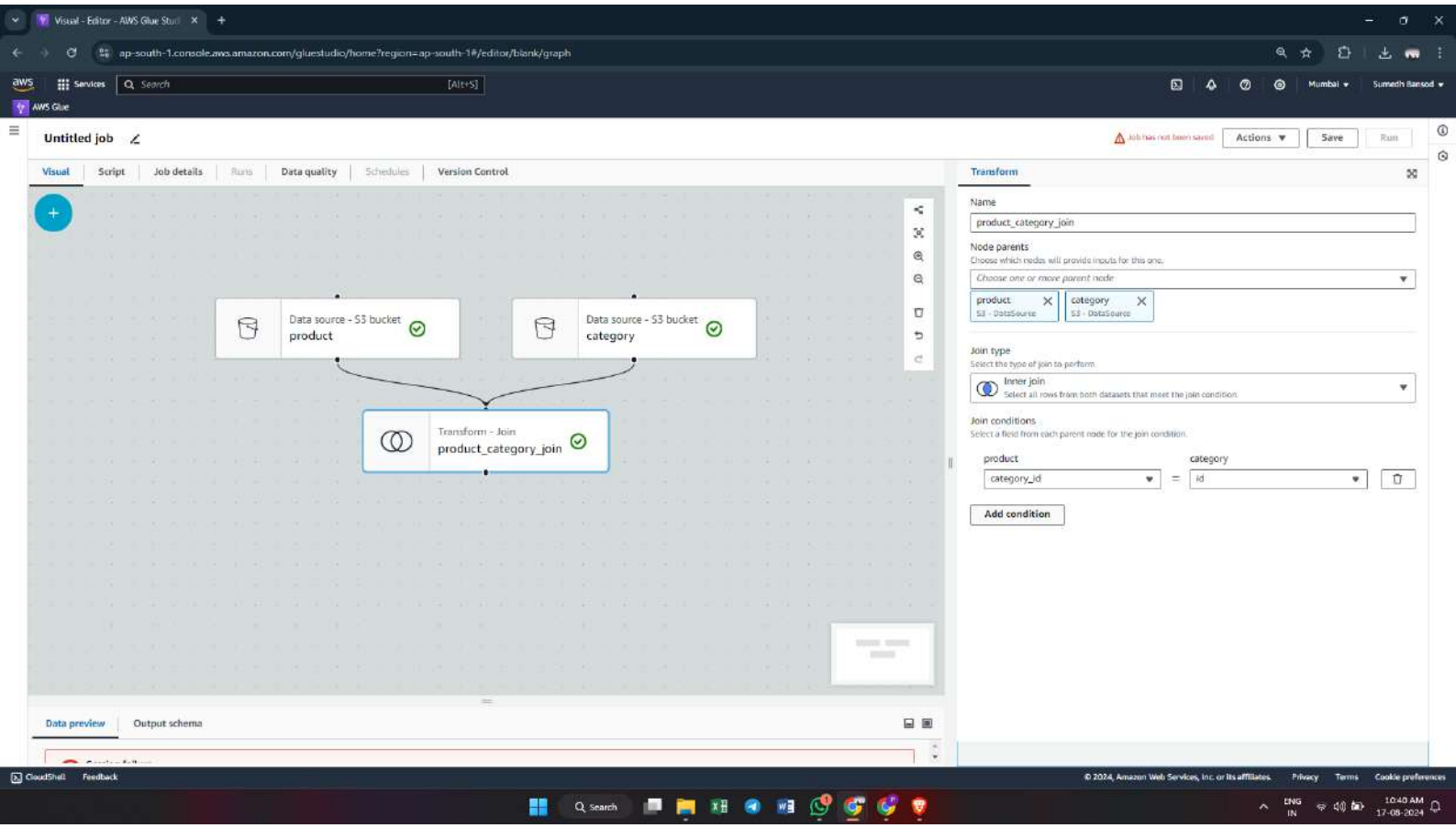

Roles | IAM | Global
us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles
Services
Search
[Alt+S]
Global
Sumedh Bansod
AWS Glue
Identity and Access Management (IAM)
Search IAM
Dashboard
Access management
User groups
Users
Roles
Policies
Identity providers
Account settings
Access reports
Access Analyzer
External access
Unused access
Analyzer settings
Credential report
Organization activity
IAM > Roles
Roles (3) Info
An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.
Delete
Create role
Search
1
Role name
Trusted entities
Last activity
AWSServiceRoleForSupport
AWS Service: support (Service-Linke
-
AWSServiceRoleForTrustedAdvisor
AWS Service: trustedadvisor (Service
-
glue_permission
AWS Service: glue
2 days ago
Roles Anywhere Info
Manage
Authenticate your non AWS workloads and securely provide access to AWS services.
Access AWS from your non AWS workloads
Operate your non AWS workloads using the same authentication and authorization strategy that you use within AWS.
X.509 Standard
Use your own existing PKI infrastructure or use AWS Certificate Manager Private Certificate Authority to authenticate identities.
Temporary credentials
Use temporary credentials with ease and benefit from the enhanced security they provide.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

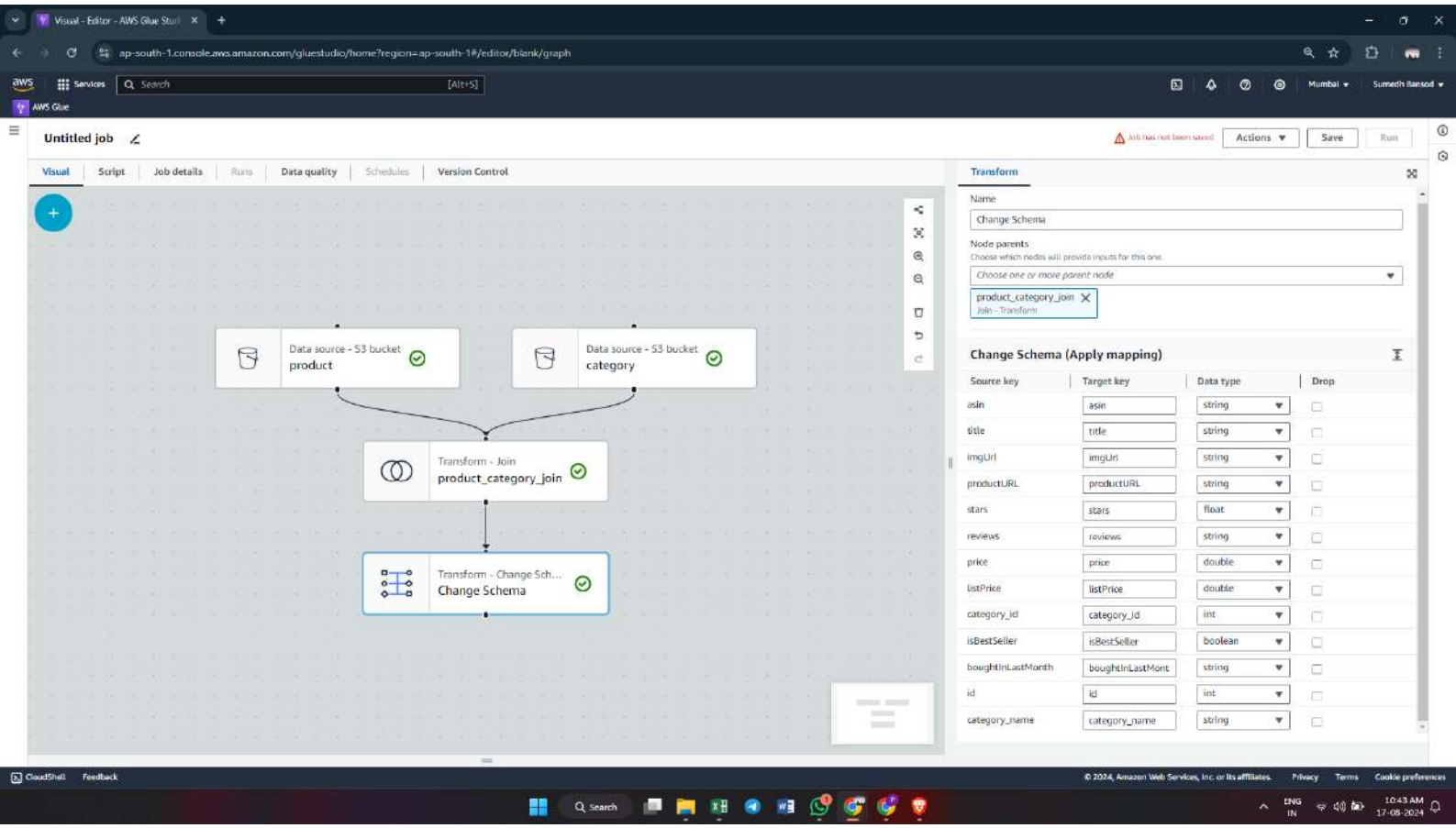

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
+ Add nodes
Search sources, transforms and targets
Sources
Transforms
Targets
Popular
AWS Glue Data Catalog
AWS Glue Data Catalog table as the data source.
Amazon S3
JSON, CSV, or Parquet files stored in S3.
Amazon Kinesis
Read from an Amazon Kinesis Data Stream.
Apache Kafka
Read from an Apache Kafka or Amazon MSK topic.
Relational DB
AWS Glue Data Catalog table with a relational database as the data source.
Amazon Redshift
Read your data from Amazon Redshift.
MySQL
AWS Glue Data Catalog table with MySQL as the data source.
PostgreSQL
AWS Glue Data Catalog table with PostgreSQL as the data source.
Oracle SQL
AWS Glue Data Catalog table with Oracle SQL as the data source.
Microsoft SQL Server
AWS Glue Data Catalog table with SQL Server as the data
Manage Connections
Data source - S3 bucket
Amazon S3
Data source properties - S3
Name
Amazon S3
S3 source type
Info
S3 location
Choose a file or folder in an S3 bucket.
Data Catalog table
S3 URL
s3://bucket/prefix/object
View
Browse S3
Recursive
Read files in all subdirectories.
Data format
No format
Infer schema
Additional options
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:37 AM
17-08-2024

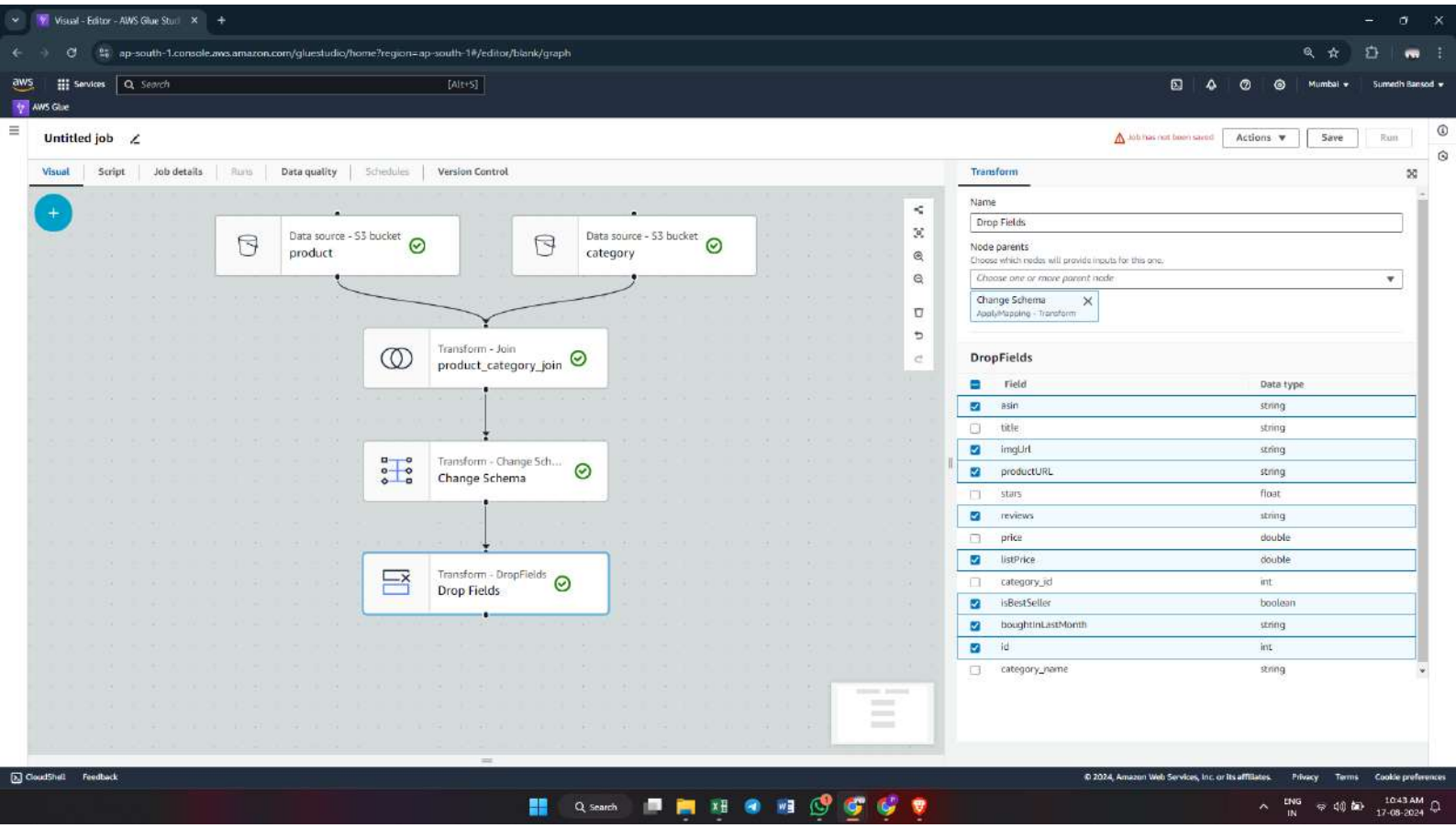

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - S3 bucket
product
Data source - S3 bucket
Amazon S3
Data source properties - S3
Name
product
S3 source type
Info
S3 location
Choose a file or folder in an S3 bucket.
Data Catalog table
S3 URL
s3://bucket/prefix/object
View
Browse S3
Recursive
Read files in all subdirectories.
Data format
No format
Infer schema
Additional options
Data preview
Output schema
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:37 AM
17-08-2024

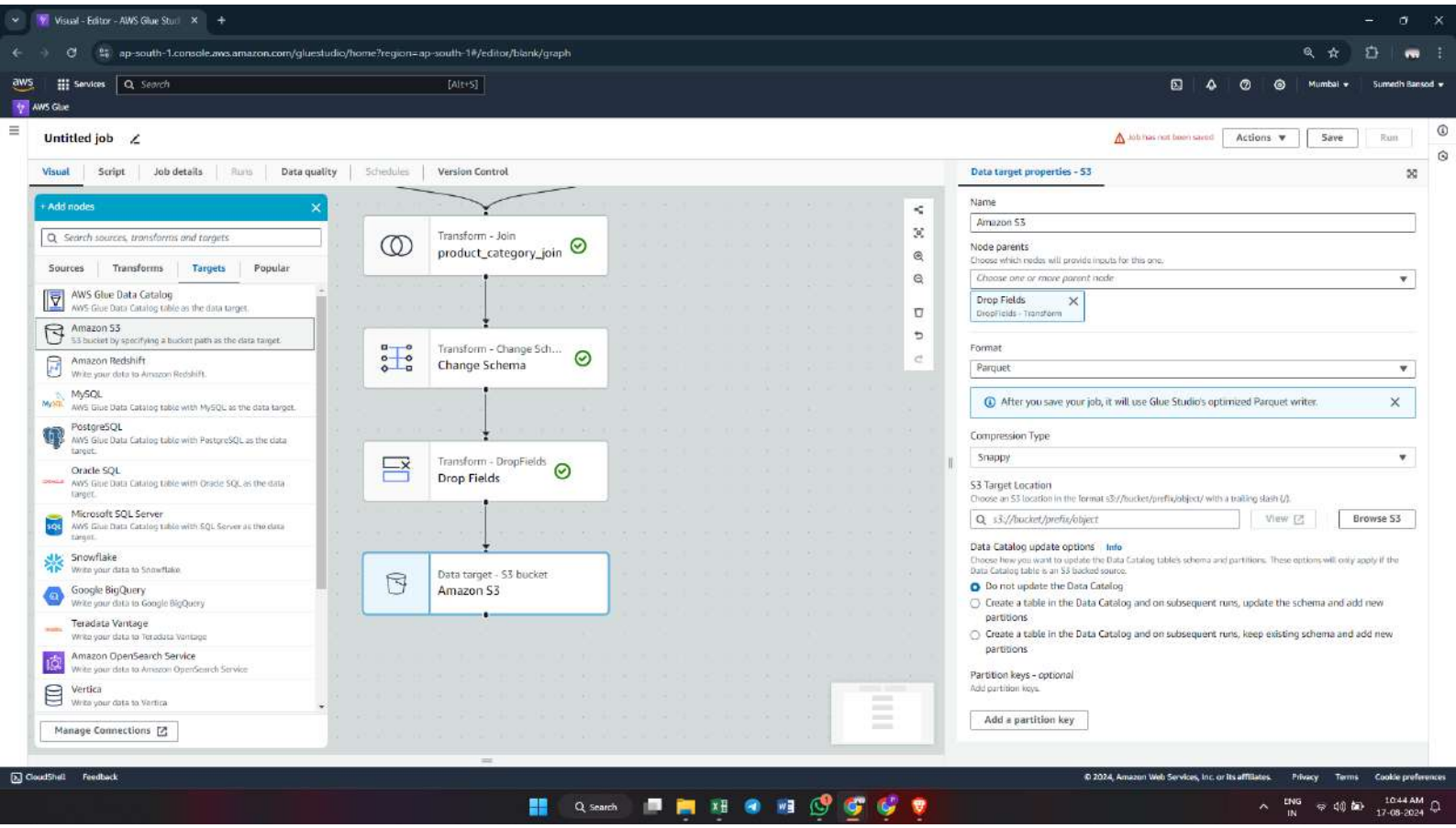

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
+ Add nodes
Search sources, transforms and targets
Sources
Transforms
Targets
Popular
Data Preparation Recipe
Select and execute an external DataBrew Recipe.
Change Schema
Change field names, data types and drop fields. Formerly known as Apply Mapping.
Join
Combine records from two datasets based on a set of conditions.
SQL Query
Use a SQL query to transform data.
Detect Sensitive Data
Detect PII and other sensitive information.
Evaluate Data Quality
Evaluate the quality and completeness of your data.
Fill Missing Values
Have AWS Glue infer reasonable values for missing data.
Aggregate
Apply functions like sum or average to fields in the dataset.
Custom Transform
Write custom code to transform data.
Drop Duplicates
Remove duplicate records from your dataset.
Add Transforms
Data source - S3 bucket
product
Data source - S3 bucket
category
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:39 AM
17-08-2024

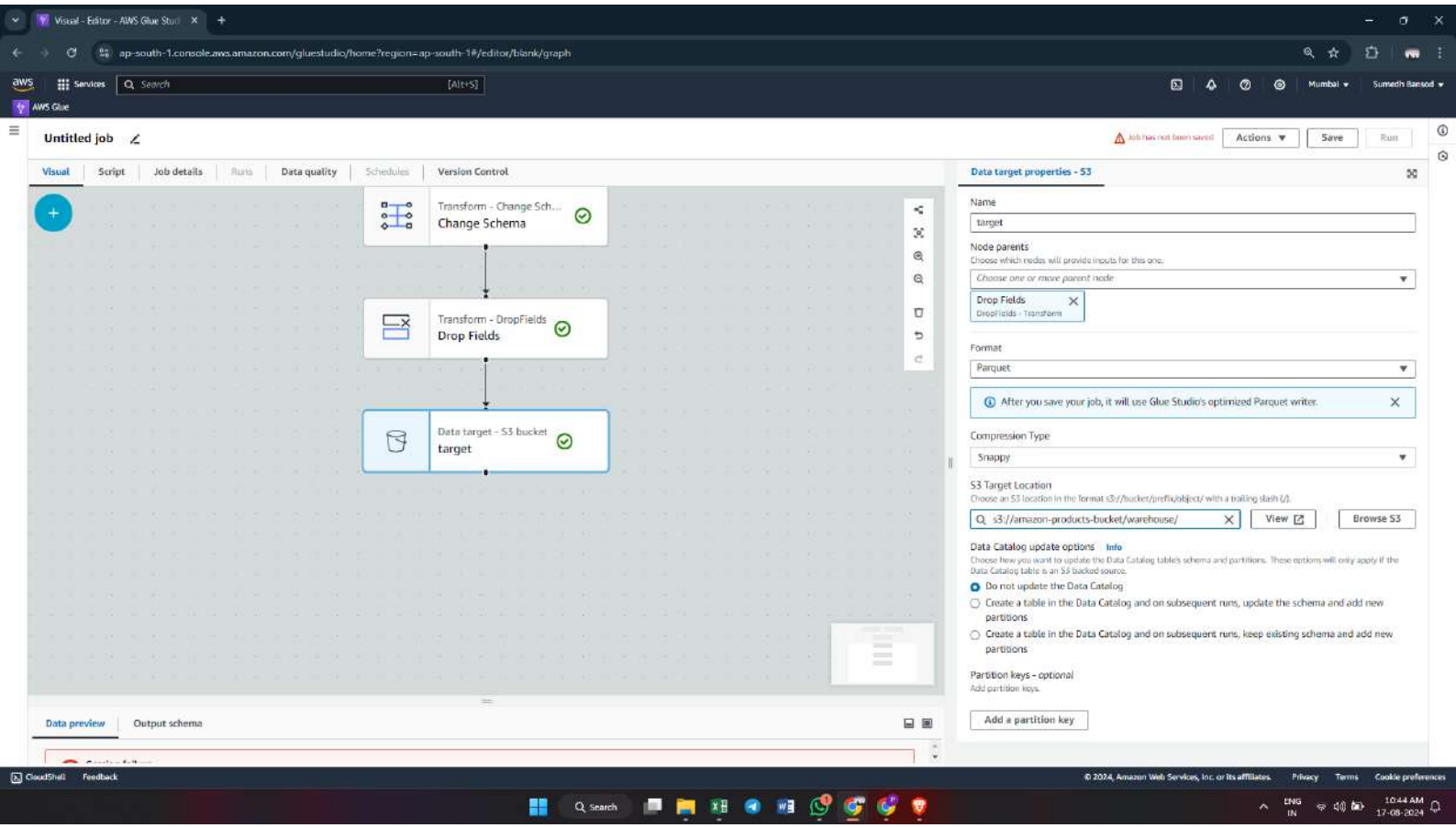

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
+ Add nodes
Search sources, transforms and targets
Sources
Transforms
Targets
Popular
Data Preparation Recipe
Select and execute an external DataBrew Recipe.
Change Schema
Change field names, data types and drop fields. Formerly known as Apply Mapping.
Join
Combine records from two datasets based on a set of conditions.
SQL Query
Use a SQL query to transform data.
Detect Sensitive Data
Detect PII and other sensitive information.
Evaluate Data Quality
Evaluate the quality and completeness of your data.
Fill Missing Values
Have AWS Glue infer reasonable values for missing data.
Aggregate
Apply functions like sum or average to fields in the dataset.
Custom Transform
Write custom code to transform data.
Drop Duplicates
Remove duplicate records from your dataset.
Add Transforms
Data source - S3 bucket
product
Data source - S3 bucket
category
Transform - Join
Join
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:39 AM
17-08-2024

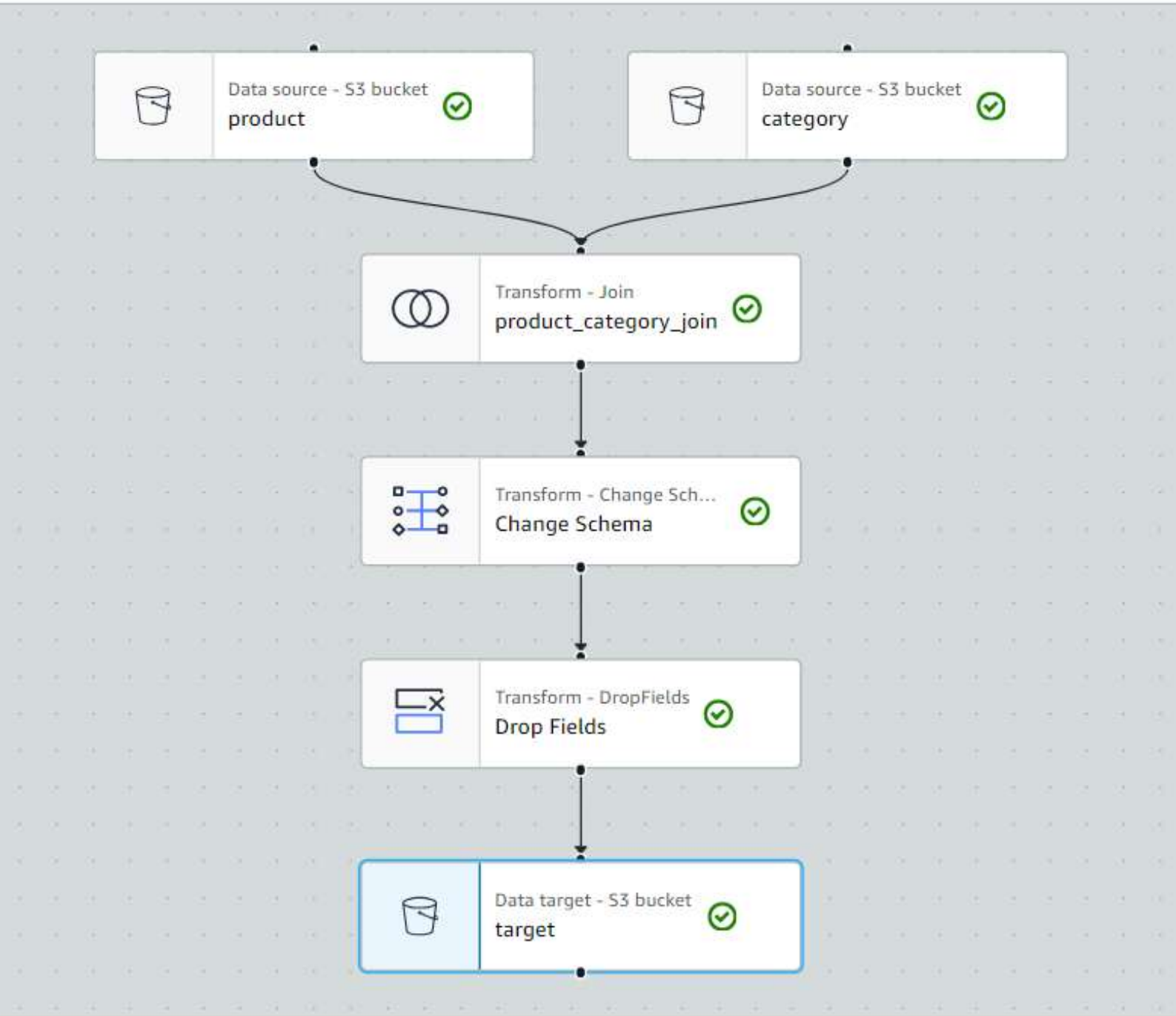

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - S3 bucket
product
Data source - S3 bucket
category
Transform - Join
product_category_join
Transform
Name
product_category_join
Node parents
Choose which nodes will provide inputs for this one.
Choose one or more parent node
product
S3 - DataSource
category
S3 - DataSource
Join type
Select the type of join to perform.
Inner join
Select all rows from both datasets that meet the join condition.
Join conditions
Select a field from each parent node for the join condition.
product
category
category_id
=
id
Add condition
Data preview
Output schema
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:40 AM
17-08-2024

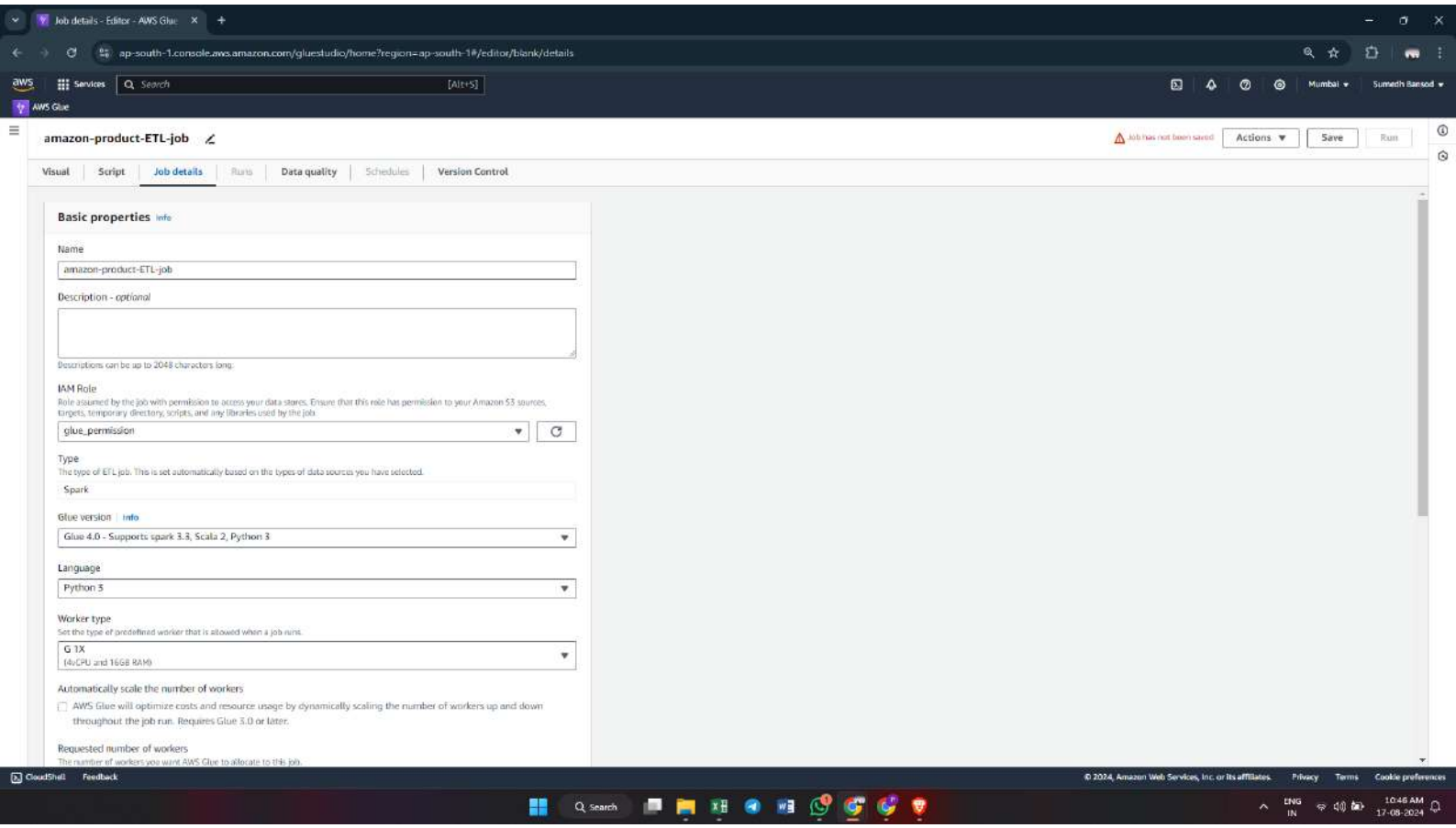

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - S3 bucket
product
Data source - S3 bucket
category
Transform - Join
product_category_join
Transform - Change Sch...
Change Schema
Transform
Name
Change Schema
Node parents
Choose which nodes will provide inputs for this one.
Choose one or more parent node
product_category_join
Join - Transform
Change Schema (Apply mapping)
Source key
Target key
Data type
Drop
asin
asin
string
title
title
string
imgUrl
imgUrl
string
productURL
productURL
string
stars
stars
float
reviews
reviews
string
price
price
double
listPrice
listPrice
double
category_id
category_id
int
isBestSeller
isBestSeller
boolean
boughtInLastMonth
boughtInLastMont
string
id
id
int
category_name
category_name
string
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:43 AM
17-08-2024

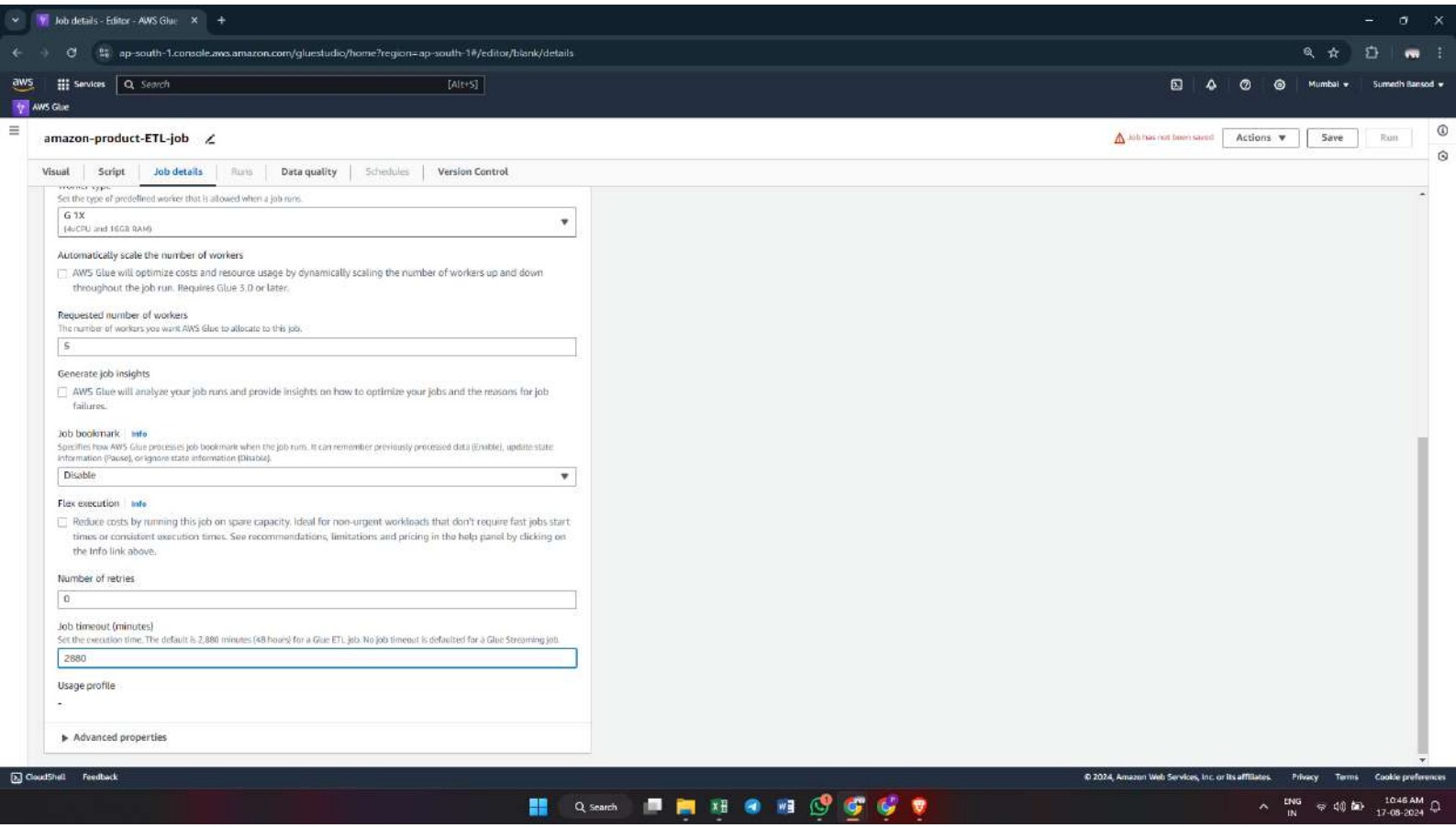

Visual - Editor - AWS Glue Studio
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - S3 bucket
product
Data source - S3 bucket
category
Transform - Join
product_category_join
Transform - Change Sch...
Change Schema
Transform - DropFields
Drop Fields
Transform
Name
Drop Fields
Node parents
Choose which nodes will provide inputs for this one.
Choose one or more parent node
Change Schema
ApplyMapping - Transform
DropFields
Field
Data type
asin string
title string
imgUrl string
productURL string
stars float
reviews string
price double
listPrice double
category_id int
isBestSeller boolean
boughtInLastMonth string
id int
category_name string
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:43 AM
17-08-2024

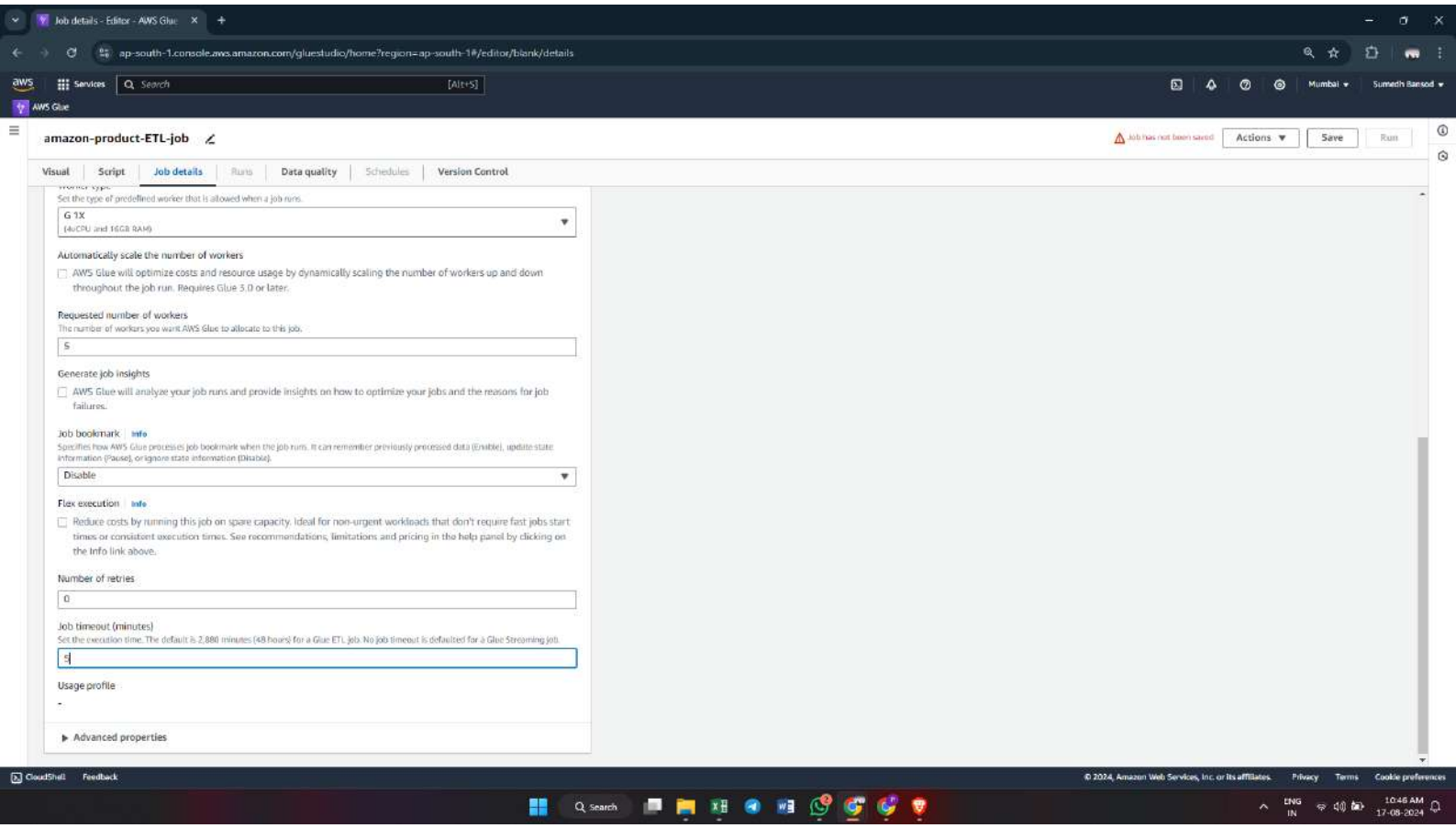

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
+ Add nodes
Search sources, transforms and targets
Sources
Transforms
Targets
Popular
AWS Glue Data Catalog
AWS Glue Data Catalog table as the data target.
Amazon S3
S3 bucket by specifying a bucket path as the data target.
Amazon Redshift
Write your data to Amazon Redshift.
MySQL
AWS Glue Data Catalog table with MySQL as the data target.
PostgreSQL
AWS Glue Data Catalog table with PostgreSQL as the data target.
Oracle SQL
AWS Glue Data Catalog table with Oracle SQL as the data target.
Microsoft SQL Server
AWS Glue Data Catalog table with SQL Server as the data target.
Snowflake
Write your data to Snowflake.
Google BigQuery
Write your data to Google BigQuery.
Teradata Vantage
Write your data to Teradata Vantage.
Amazon OpenSearch Service
Write your data to Amazon OpenSearch Service.
Vertica
Write your data to Vertica.
Manage Connections
Transform - Join
product_category_join
Transform - Change Sch...
Change Schema
Transform - DropFields
Drop Fields
Data target - S3 bucket
Amazon S3
Data target properties - S3
Name
Amazon S3
Node parents
Choose which nodes will provide inputs for this one.
Choose one or more parent node
Drop Fields
DropFields - Transform
Format
Parquet
After you save your job, it will use Glue Studio's optimized Parquet writer.
Compression Type
Snappy
S3 Target Location
Choose an S3 location in the format s3://bucket/prefix/object/ with a trailing slash (/).
s3://bucket/prefix/object
View
Browse S3
Data Catalog update options
Info
Choose how you want to update the Data Catalog table's schema and partitions. These options will only apply if the Data Catalog table is an S3 backed source.
Do not update the Data Catalog
Create a table in the Data Catalog and on subsequent runs, update the schema and add new partitions
Create a table in the Data Catalog and on subsequent runs, keep existing schema and add new partitions
Partition keys - optional
Add partition keys.
Add a partition key
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:44 AM
17-08-2024

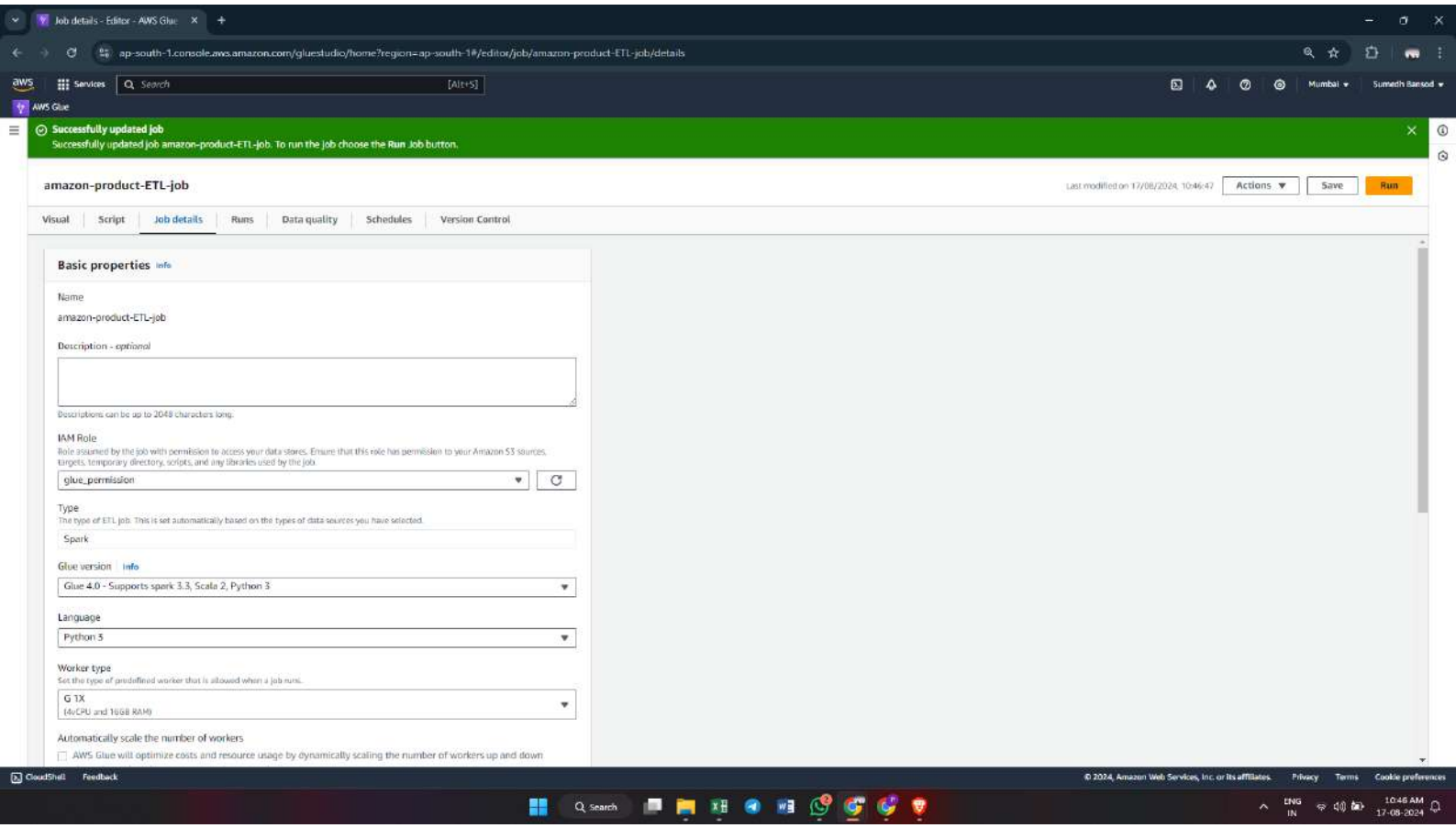

Visual - Editor - AWS Glue Stud
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/graph
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
Untitled job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Transform - Change Sch...
Change Schema
Transform - DropFields
Drop Fields
Data target - S3 bucket
target
Data target properties - S3
Name
target
Node parents
Choose which nodes will provide inputs for this one.
Choose one or more parent node
Drop Fields
DropFields - Transform
Format
Parquet
After you save your job, it will use Glue Studio's optimized Parquet writer.
Compression Type
Snappy
S3 Target Location
Choose an S3 location in the format s3://bucket/prefix/object/ with a trailing slash (/).
s3://amazon-products-bucket/warehouse/
View
Browse S3
Data Catalog update options
Info
Choose how you want to update the Data Catalog table's schema and partitions. These options will only apply if the Data Catalog table is an S3 backed source.
Do not update the Data Catalog
Create a table in the Data Catalog and on subsequent runs, update the schema and add new partitions
Create a table in the Data Catalog and on subsequent runs, keep existing schema and add new partitions
Partition keys - optional
Add partition keys.
Add a partition key
Data preview
Output schema
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:44 AM
17-08-2024

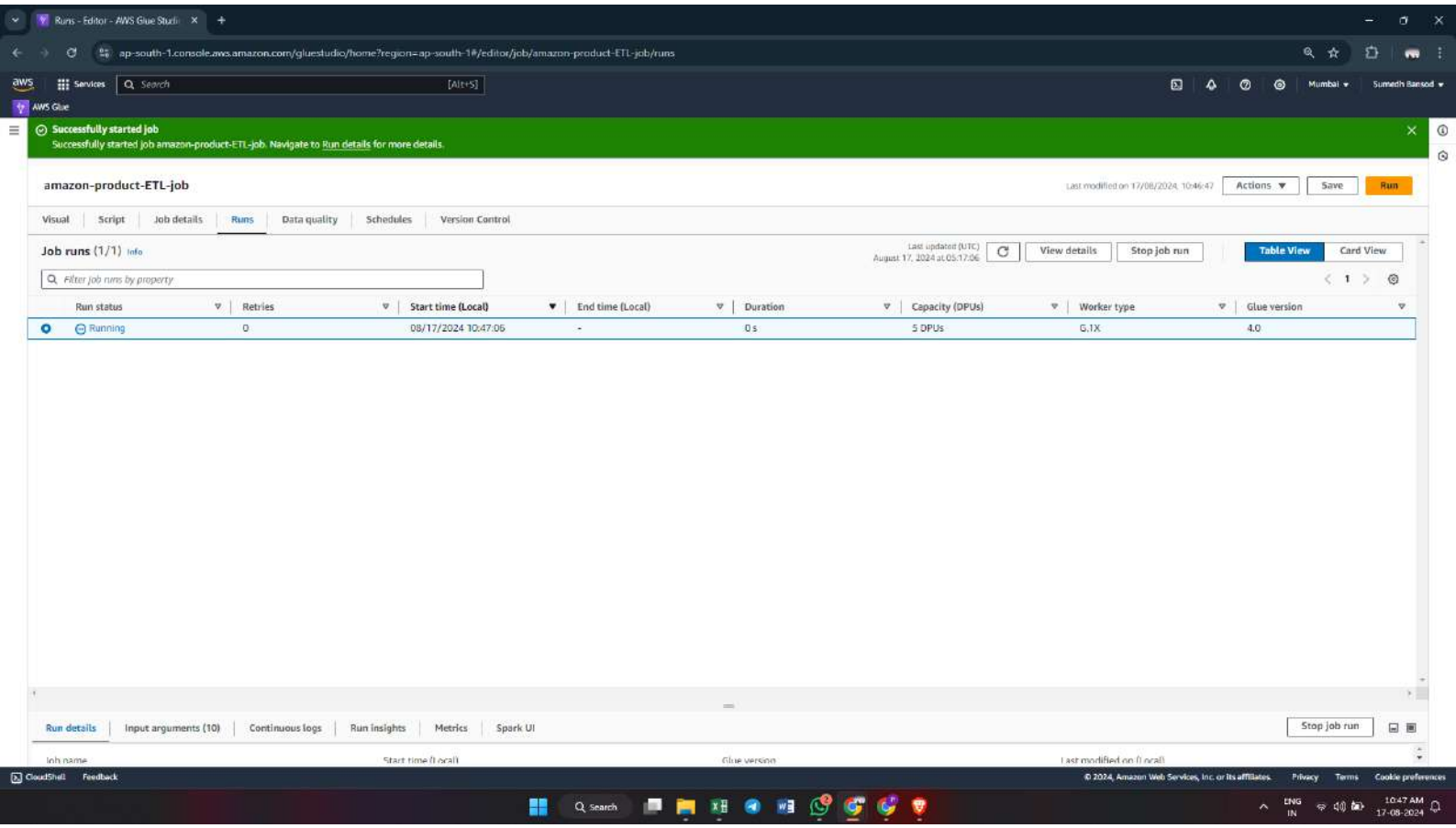

Data source - S3 bucket
product
Data source - S3 bucket
category
Transform - Join
product_category_join
Transform - Change Sch...
Change Schema
Transform - DropFields
Drop Fields
Data target - S3 bucket
target

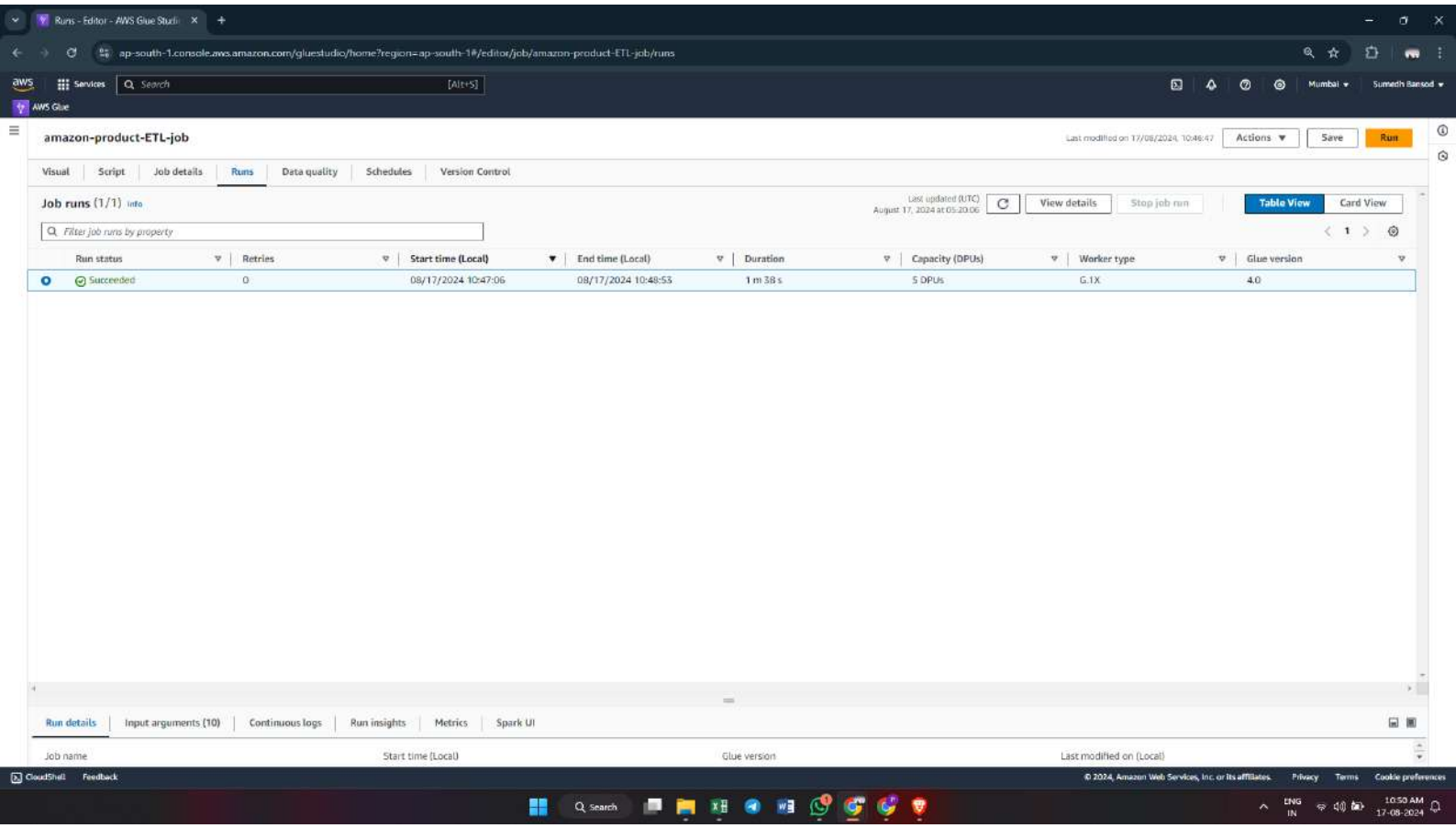

Job details - Editor - AWS Glue
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/details
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
amazon-product-ETL-job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Basic properties info
Name
amazon-product-ETL-job
Description - optional
Descriptions can be up to 2048 characters long.
IAM Role
Role assumed by the job with permission to access your data stores. Ensure that this role has permission to your Amazon S3 sources, targets, temporary directory, scripts, and any libraries used by the job.
glue_permission
Type
The type of ETL job. This is set automatically based on the types of data sources you have selected.
Spark
Glue version info
Glue 4.0 - Supports spark 3.3, Scala 2, Python 3
Language
Python 3
Worker type
Set the type of predefined worker that is allowed when a job runs.
G 1X
(4vCPU and 16GB RAM)
Automatically scale the number of workers
AWS Glue will optimize costs and resource usage by dynamically scaling the number of workers up and down throughout the job run. Requires Glue 3.0 or later.
Requested number of workers
The number of workers you want AWS Glue to allocate to this job.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:46 AM
17-08-2024

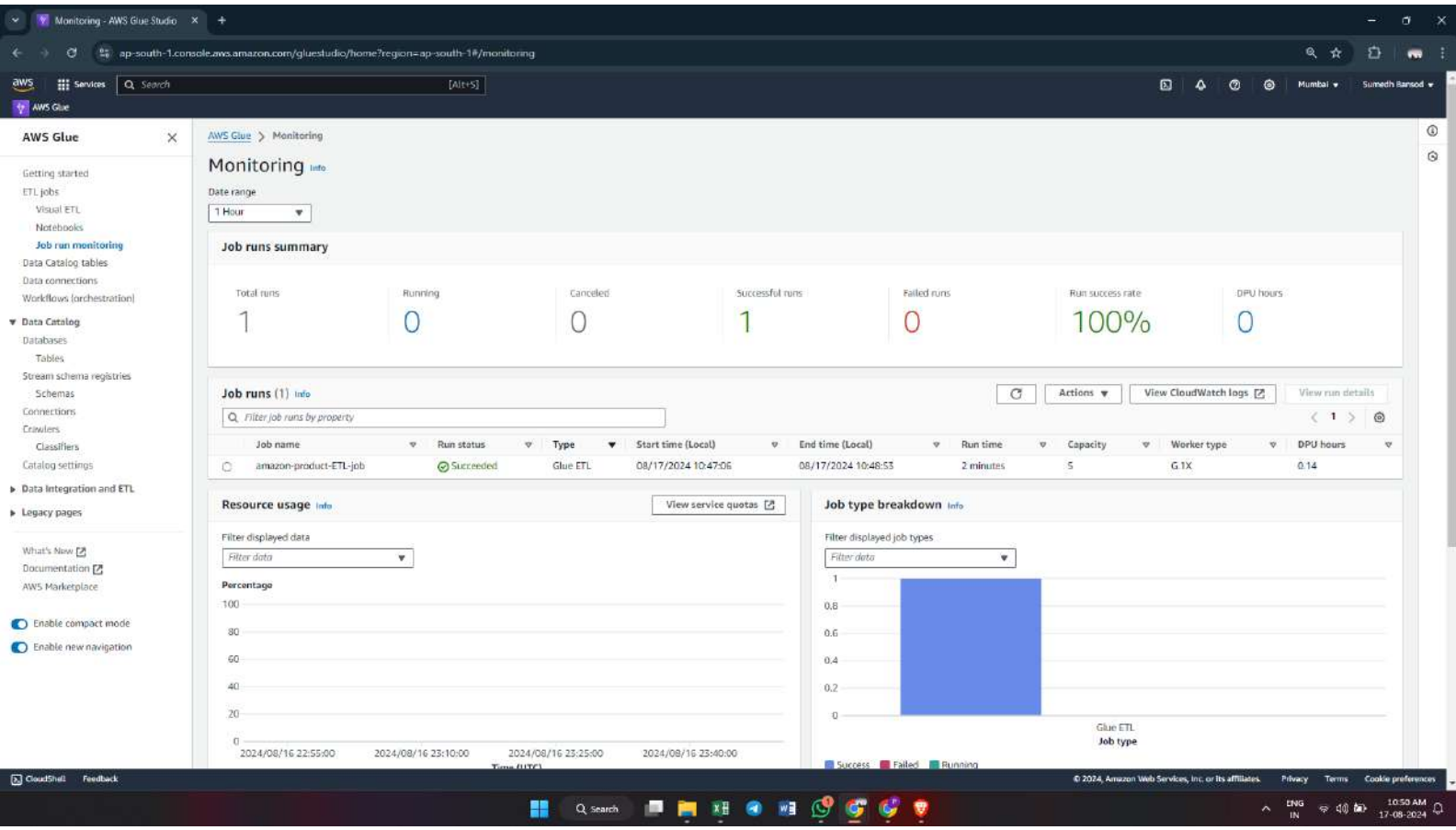

Job details - Editor - AWS Glue
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/details
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
amazon-product-ETL-job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Set the type of predefined worker that is allowed when a job runs.
G 1X
(4vCPU and 16GB RAM)
Automatically scale the number of workers
AWS Glue will optimize costs and resource usage by dynamically scaling the number of workers up and down throughout the job run. Requires Glue 3.0 or later.
Requested number of workers
The number of workers you want AWS Glue to allocate to this job.
5
Generate job insights
AWS Glue will analyze your job runs and provide insights on how to optimize your jobs and the reasons for job failures.
Job bookmark
Info
Specifies how AWS Glue processes job bookmark when the job runs. It can remember previously processed data (Enable), update state information (Pause), or ignore state information (Disable).
Disable
Flex execution
Info
Reduce costs by running this job on spare capacity. Ideal for non-urgent workloads that don't require fast jobs start times or consistent execution times. See recommendations, limitations and pricing in the help panel by clicking on the Info link above.
Number of retries
0
Job timeout (minutes)
Set the execution time. The default is 2,880 minutes (48 hours) for a Glue ETL job. No job timeout is defaulted for a Glue Streaming job.
2880
Usage profile
-
Advanced properties
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:46 AM
17-08-2024

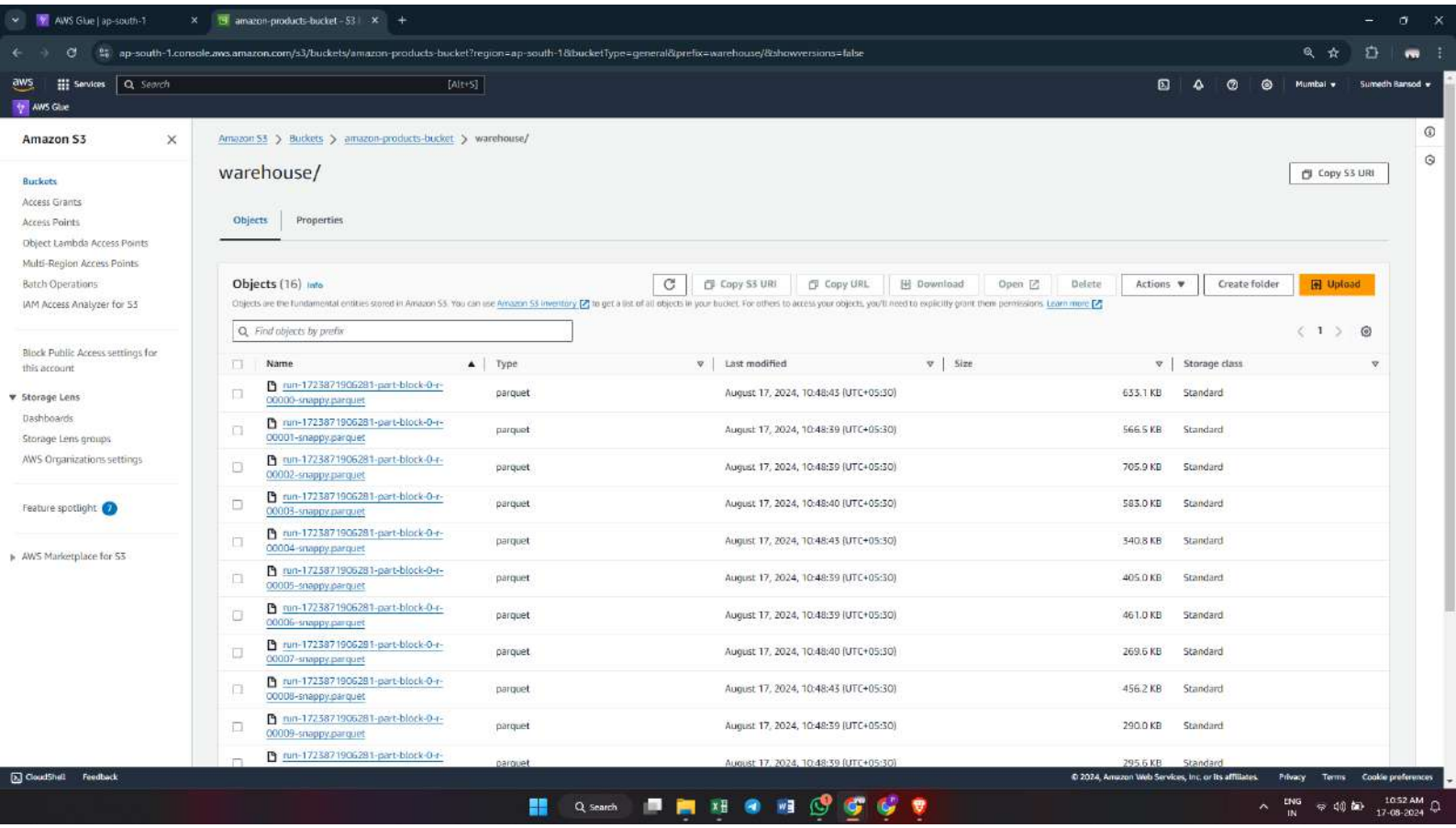

Job details - Editor - AWS Glue
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/blank/details
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
amazon-product-ETL-job
Job has not been saved
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Set the type of predefined worker that is allowed when a job runs.
G 1X
(4vCPU and 16GB RAM)
Automatically scale the number of workers
AWS Glue will optimize costs and resource usage by dynamically scaling the number of workers up and down throughout the job run. Requires Glue 3.0 or later.
Requested number of workers
The number of workers you want AWS Glue to allocate to this job.
5
Generate job insights
AWS Glue will analyze your job runs and provide insights on how to optimize your jobs and the reasons for job failures.
Job bookmark
Info
Specifies how AWS Glue processes job bookmark when the job runs. It can remember previously processed data (Enable), update state information (Pause), or ignore state information (Disable).
Disable
Flex execution
Info
Reduce costs by running this job on spare capacity. Ideal for non-urgent workloads that don't require fast jobs start times or consistent execution times. See recommendations, limitations and pricing in the help panel by clicking on the Info link above.
Number of retries
0
Job timeout (minutes)
Set the execution time. The default is 2,880 minutes (48 hours) for a Glue ETL job. No job timeout is defaulted for a Glue Streaming job.
5
Usage profile
-
Advanced properties
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:46 AM
17-08-2024

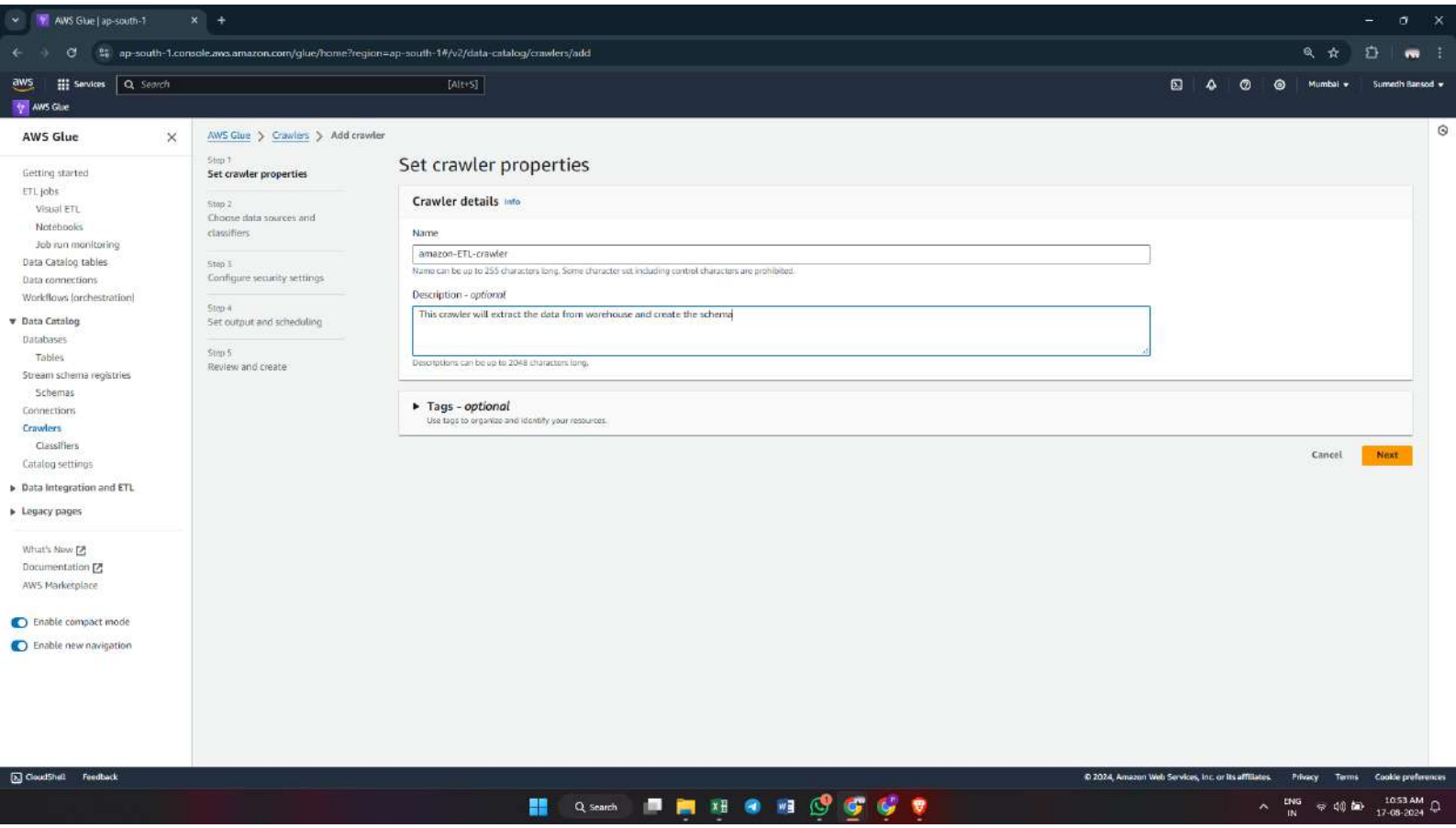

Job details - Editor - AWS Glue
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/job/amazon-product-ETL-job/details
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
Successfully updated job
Successfully updated job amazon-product-ETL-job. To run the job choose the Run Job button.
amazon-product-ETL-job
Last modified on 17/08/2024, 10:46:47
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Basic properties
Info
Name
amazon-product-ETL-job
Description - optional
Descriptions can be up to 2048 characters long.
IAM Role
Role assumed by the job with permission to access your data stores. Ensure that this role has permission to your Amazon S3 sources, targets, temporary directory, scripts, and any libraries used by the job.
glue_permission
Type
The type of ETL job. This is set automatically based on the types of data sources you have selected.
Spark
Glue version
Info
Glue 4.0 - Supports spark 3.3, Scala 2, Python 3
Language
Python 3
Worker type
Set the type of predefined worker that is allowed when a job runs.
G 1X
(4vCPU and 16GB RAM)
Automatically scale the number of workers
AWS Glue will optimize costs and resource usage by dynamically scaling the number of workers up and down
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:46 AM
17-08-2024

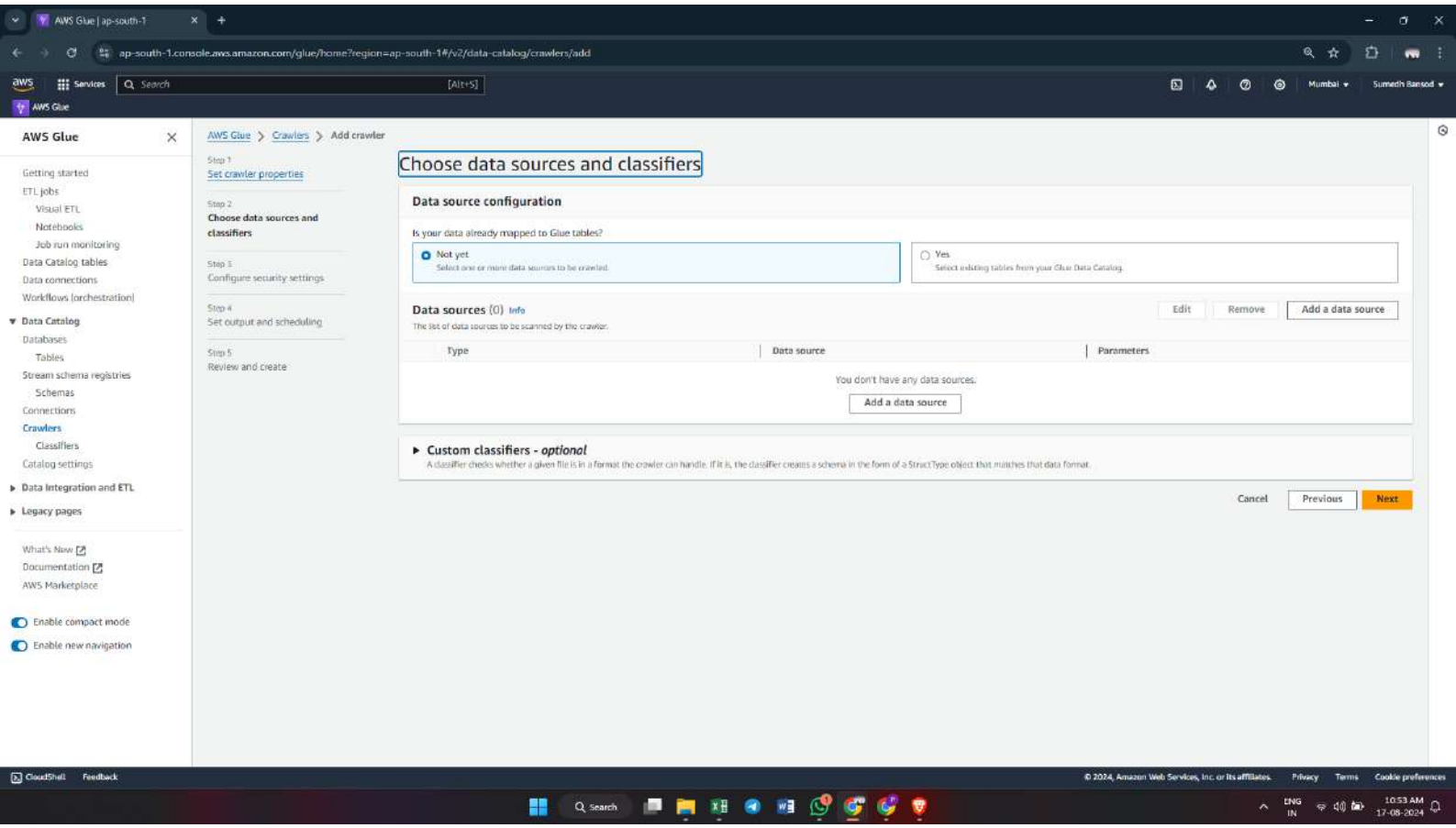

Runs - Editor - AWS Glue Studio
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/job/amazon-product-ETL-job/runs
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Successfully started job
Successfully started job amazon-product-ETL-job. Navigate to Run details for more details.
amazon-product-ETL-job
Last modified on 17/08/2024, 10:46:47
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/1) Info
Last updated (UTC)
August 17, 2024 at 05:17:06
View details
Stop job run
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Running
0
08/17/2024 10:47:06
-
0 s
5 DPUs
G.1X
4.0
Run details
Input arguments (10)
Continuous logs
Run insights
Metrics
Spark UI
Stop job run
Job name
Start time (Local)
Glue version
Last modified on (Local)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:47 AM
17-08-2024

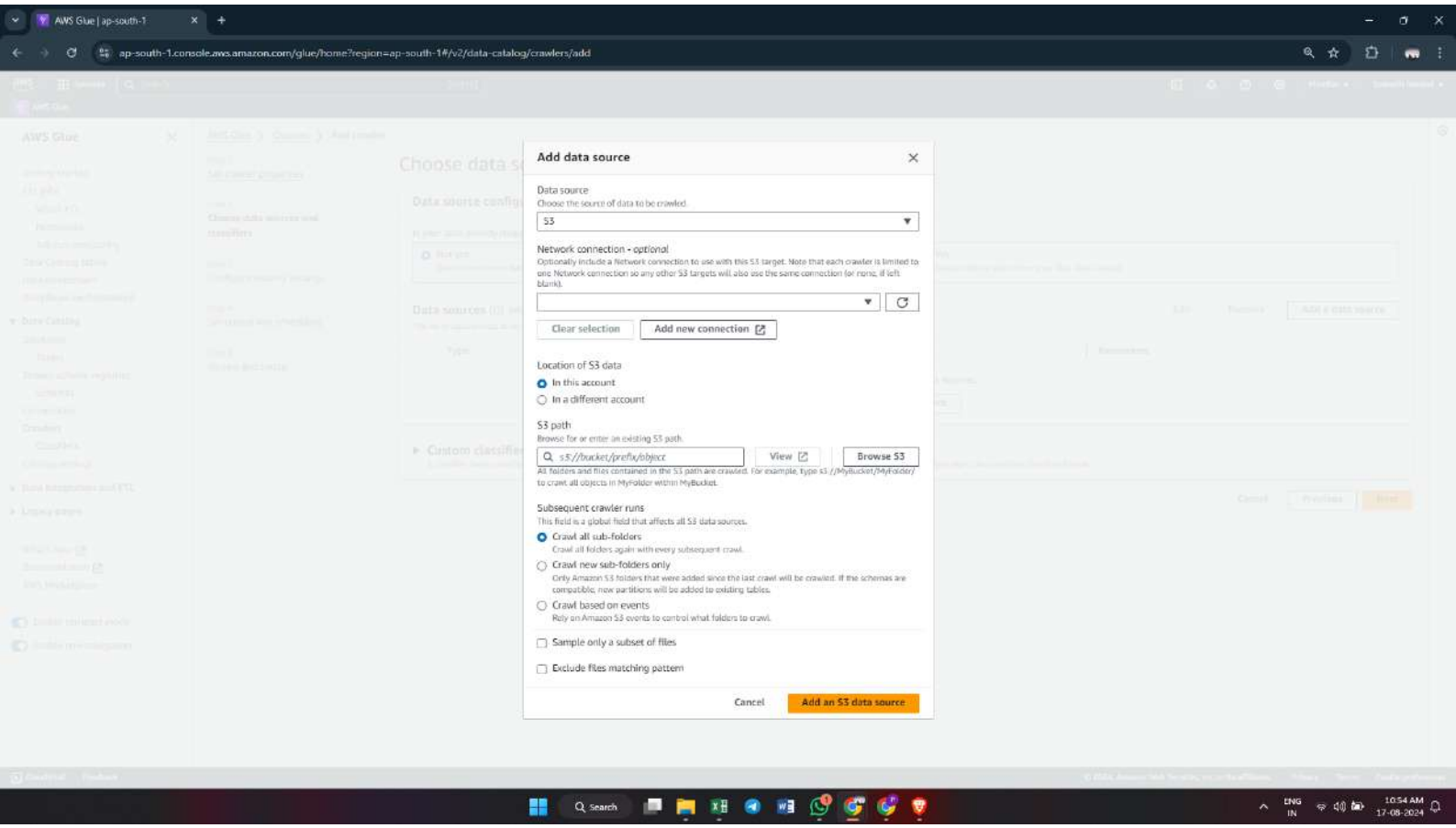

Runs - Editor - AWS Glue Studio
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/editor/job/amazon-product-ETL-job/runs
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
amazon-product-ETL-job
Last modified on 17/08/2024, 10:46:47
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/1) info
Last updated (UTC)
August 17, 2024 at 05:20:06
View details
Stop job run
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
08/17/2024 10:47:06
08/17/2024 10:48:53
1 m 38 s
5 DPUs
G.1X
4.0
Run details
Input arguments (10)
Continuous logs
Run insights
Metrics
Spark UI
Job name
Start time (Local)
Glue version
Last modified on (Local)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:50 AM
17-08-2024

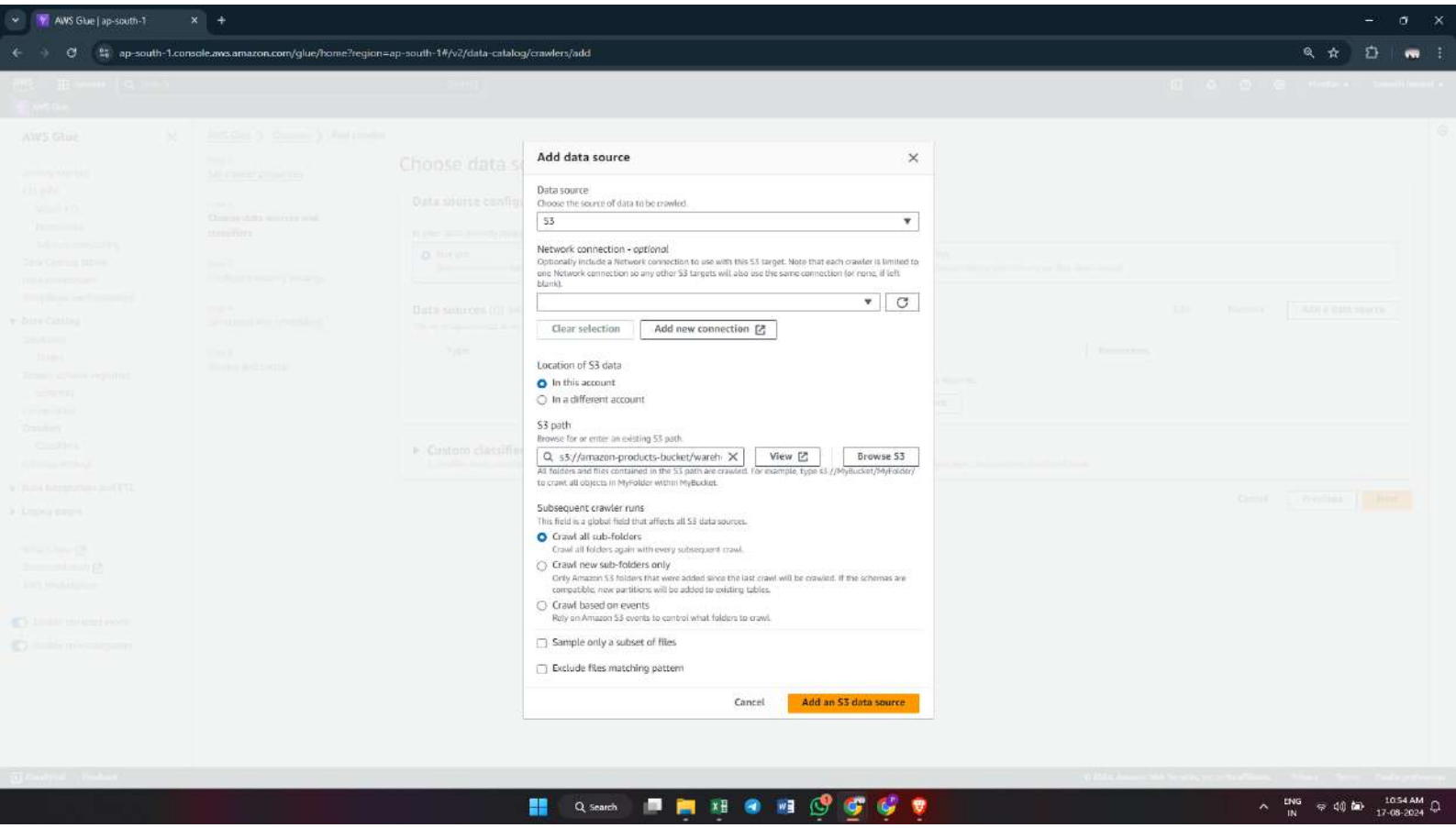

Monitoring - AWS Glue Studio
ap-south-1.console.aws.amazon.com/gluestudio/home?region=ap-south-1#/monitoring
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Monitoring
Monitoring Info
Date range
1 Hour
Job runs summary
Total runs
1
Running
0
Canceled
0
Successful runs
1
Failed runs
0
Run success rate
100%
DPU hours
0
Job runs (1) Info
Actions
View CloudWatch logs
View run details
Filter job runs by property
Job name
Run status
Type
Start time (Local)
End time (Local)
Run time
Capacity
Worker type
DPU hours
amazon-product-ETL-job
Succeeded
Glue ETL
08/17/2024 10:47:06
08/17/2024 10:48:53
2 minutes
5
G.1X
0.14
Resource usage Info
View service quotas
Filter displayed data
Filter data
Percentage
100
80
60
40
20
0
2024/08/16 22:55:00
2024/08/16 23:10:00
2024/08/16 23:25:00
2024/08/16 23:40:00
Job type breakdown Info
Filter displayed job types
Filter data
1
0.8
0.6
0.4
0.2
0
Glue ETL
Job type
Success
Failed
Running
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:50 AM
17-08-2024

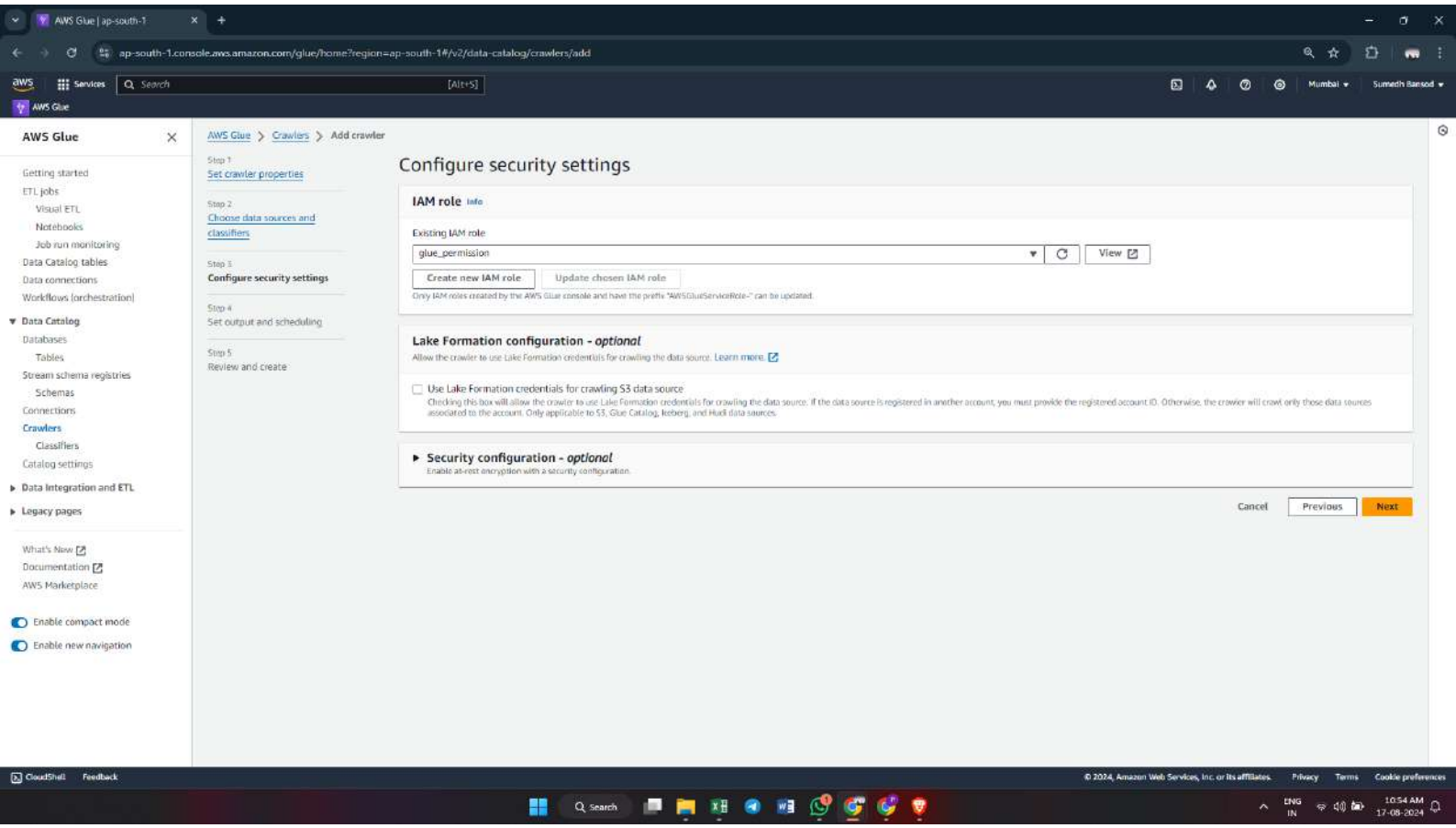

AWS Glue | ap-south-1
amazon-products-bucket - S3
ap-south-1.console.aws.amazon.com/s3/buckets/amazon-products-bucket?region=ap-south-1&bucketType=general&prefix=warehouse/&showversions=false
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon S3
Buckets
Access Grants
Access Points
Object Lambda Access Points
Multi-Region Access Points
Batch Operations
IAM Access Analyzer for S3
Block Public Access settings for this account
Storage Lens
Dashboards
Storage Lens groups
AWS Organizations settings
Feature spotlight
AWS Marketplace for S3
Amazon S3 > Buckets > amazon-products-bucket > warehouse/
warehouse/
Copy S3 URI
Objects
Properties
Objects (16) Info
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
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
Find objects by prefix
Name
Type
Last modified
Size
Storage class
run-1723871906281-part-block-0-r-00000-snappy.parquet parquet August 17, 2024, 10:48:43 (UTC+05:30) 633.1 KB Standard
run-1723871906281-part-block-0-r-00001-snappy.parquet parquet August 17, 2024, 10:48:39 (UTC+05:30) 566.5 KB Standard
run-1723871906281-part-block-0-r-00002-snappy.parquet parquet August 17, 2024, 10:48:39 (UTC+05:30) 705.9 KB Standard
run-1723871906281-part-block-0-r-00003-snappy.parquet parquet August 17, 2024, 10:48:40 (UTC+05:30) 583.0 KB Standard
run-1723871906281-part-block-0-r-00004-snappy.parquet parquet August 17, 2024, 10:48:43 (UTC+05:30) 340.8 KB Standard
run-1723871906281-part-block-0-r-00005-snappy.parquet parquet August 17, 2024, 10:48:39 (UTC+05:30) 405.0 KB Standard
run-1723871906281-part-block-0-r-00006-snappy.parquet parquet August 17, 2024, 10:48:39 (UTC+05:30) 461.0 KB Standard
run-1723871906281-part-block-0-r-00007-snappy.parquet parquet August 17, 2024, 10:48:40 (UTC+05:30) 269.6 KB Standard
run-1723871906281-part-block-0-r-00008-snappy.parquet parquet August 17, 2024, 10:48:43 (UTC+05:30) 456.2 KB Standard
run-1723871906281-part-block-0-r-00009-snappy.parquet parquet August 17, 2024, 10:48:39 (UTC+05:30) 290.0 KB Standard
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG IN
10:52 AM 17-08-2024

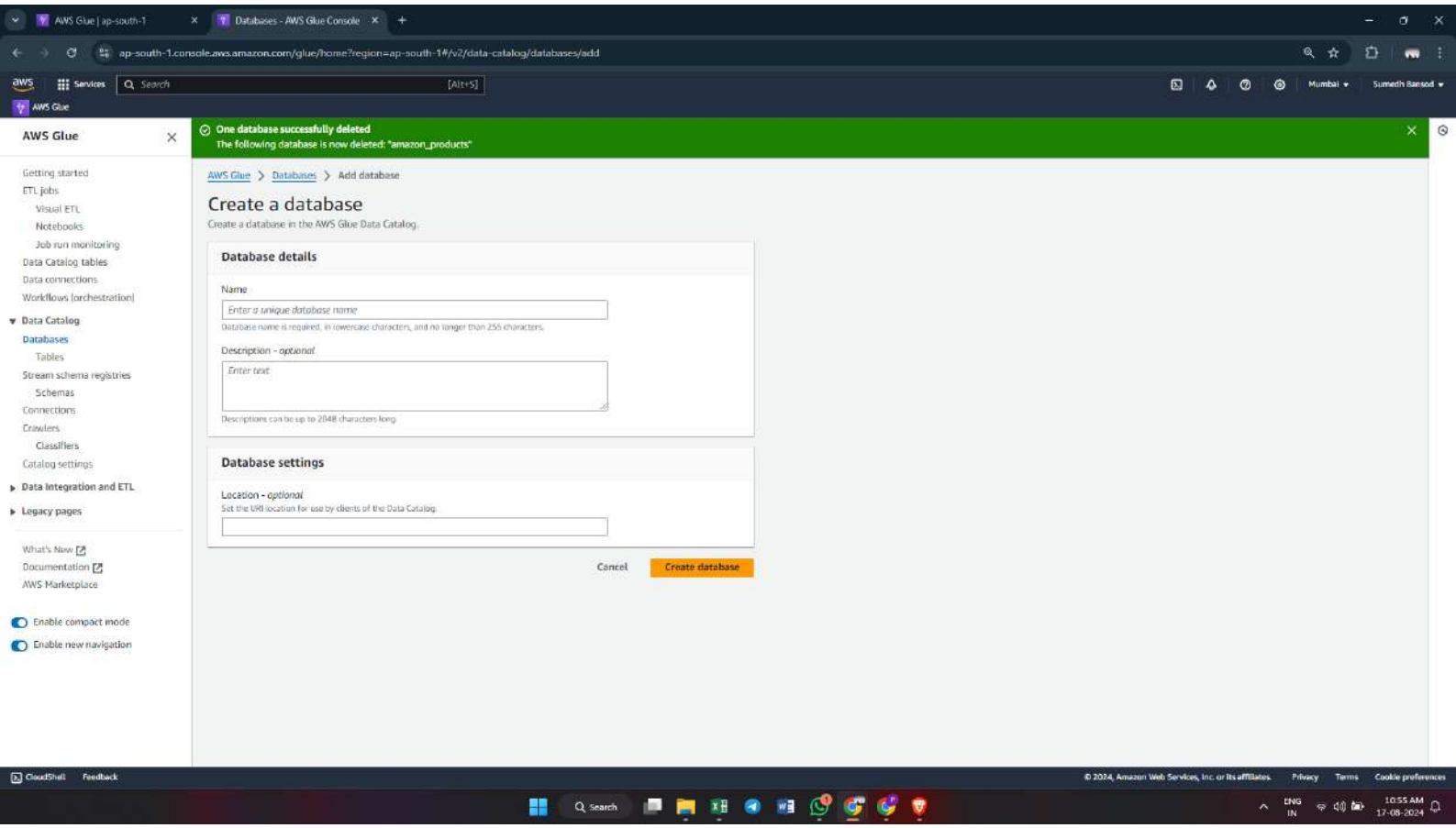

AWS Glue | ap-south-1
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Set crawler properties
Crawler details Info
Name
amazon-ETL-crawler
Name can be up to 255 characters long. Some character set including control characters are prohibited.
Description - optional
This crawler will extract the data from warehouse and create the schema
Descriptions can be up to 2048 characters long.
Tags - optional
Use tags to organize and identify your resources.
Cancel
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:53 AM
17-08-2024

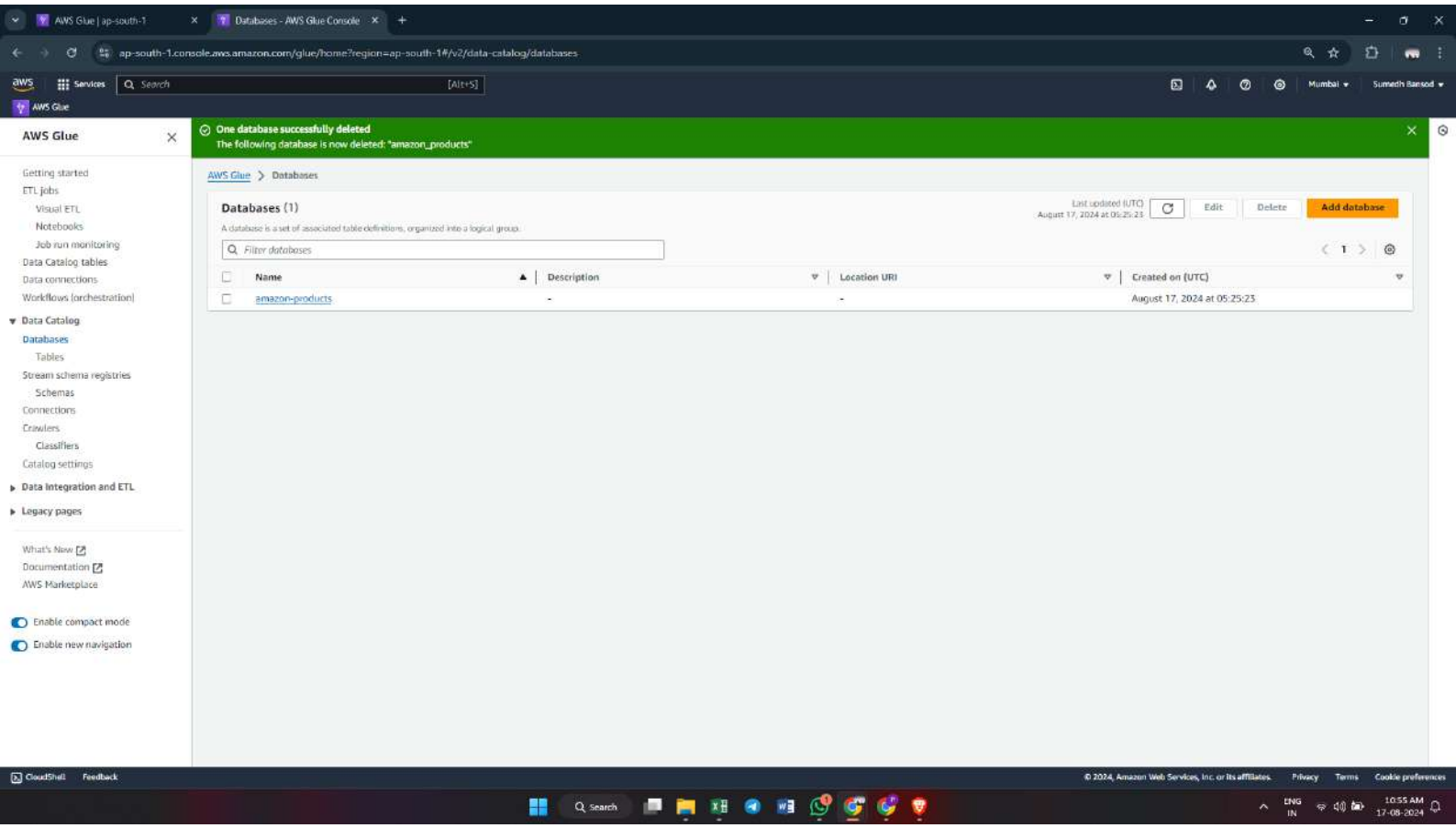

AWS Glue | ap-south-1
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Choose data sources and classifiers
Data source configuration
Is your data already mapped to Glue tables?
Not yet
Select one or more data sources to be crawled.
Yes
Select existing tables from your Glue Data Catalog.
Data sources (0) Info
The list of data sources to be scanned by the crawler.
Edit
Remove
Add a data source
Type
Data source
Parameters
You don't have any data sources.
Add a data source
Custom classifiers - optional
A classifier checks whether a given file is in a format the crawler can handle. If it is, the classifier creates a schema in the form of a StructType object that matches that data format.
Cancel
Previous
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:53 AM
17-08-2024

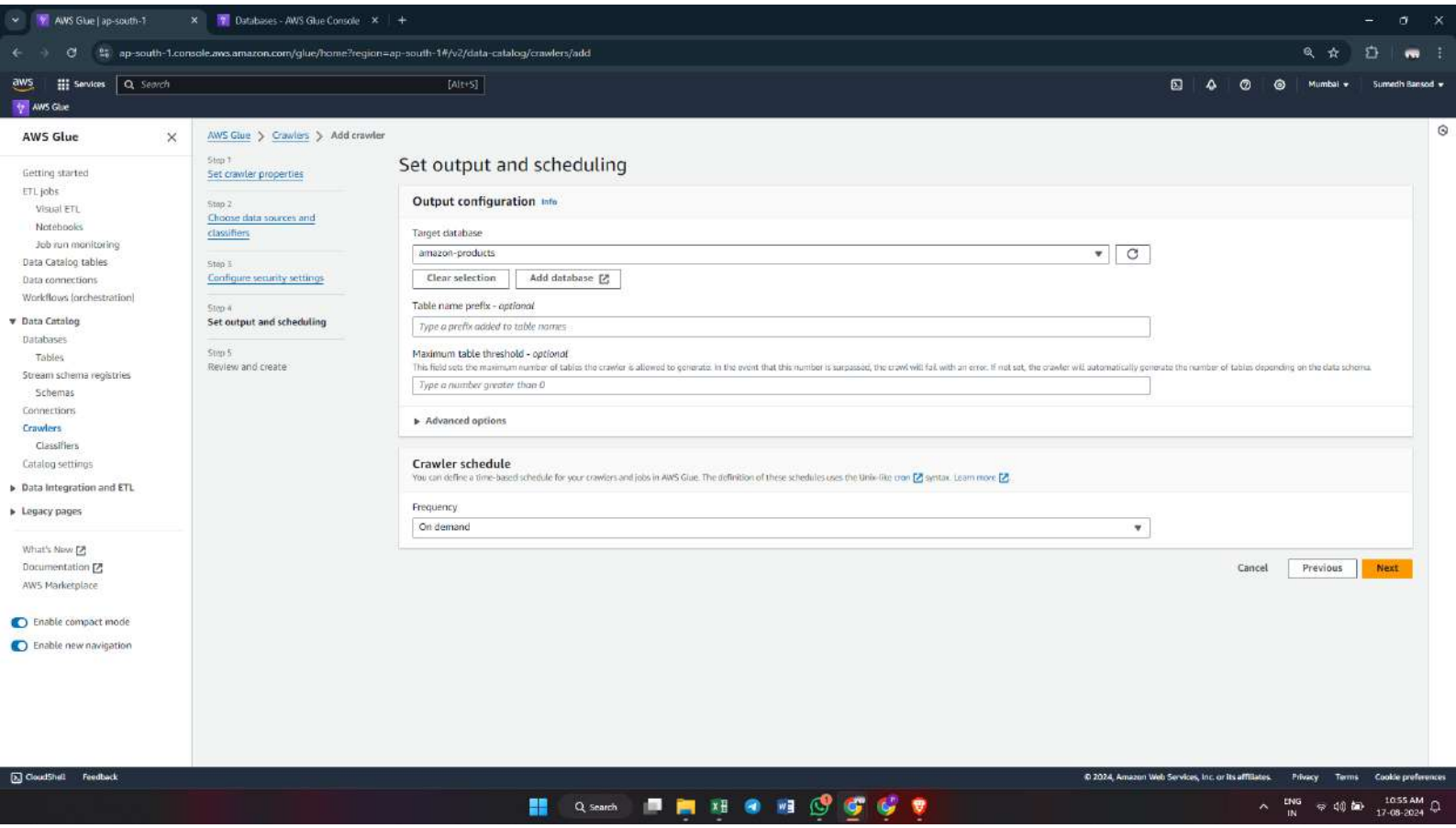

AWS Glue | ap-south-1
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Add data source
Data source
Choose the source of data to be crawled.
S3
Network connection - optional
Optionally include a Network connection to use with this S3 target. Note that each crawler is limited to one Network connection so any other S3 targets will also use the same connection (or none, if left blank).
Clear selection
Add new connection
Location of S3 data
In this account
In a different account
S3 path
Browse for or enter an existing S3 path.
s3://bucket/prefix/object
View
Browse S3
All folders and files contained in the S3 path are crawled. For example, type s3://MyBucket/MyFolder/ to crawl all objects in MyFolder within MyBucket.
Subsequent crawler runs
This field is a global field that affects all S3 data sources.
Crawl all sub-folders
Crawl all folders again with every subsequent crawl.
Crawl new sub-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.
Crawl based on events
Rely on Amazon S3 events to control what folders to crawl.
Sample only a subset of files
Exclude files matching pattern
Cancel
Add an S3 data source
Search
ENG
IN
10:54 AM
17-08-2024

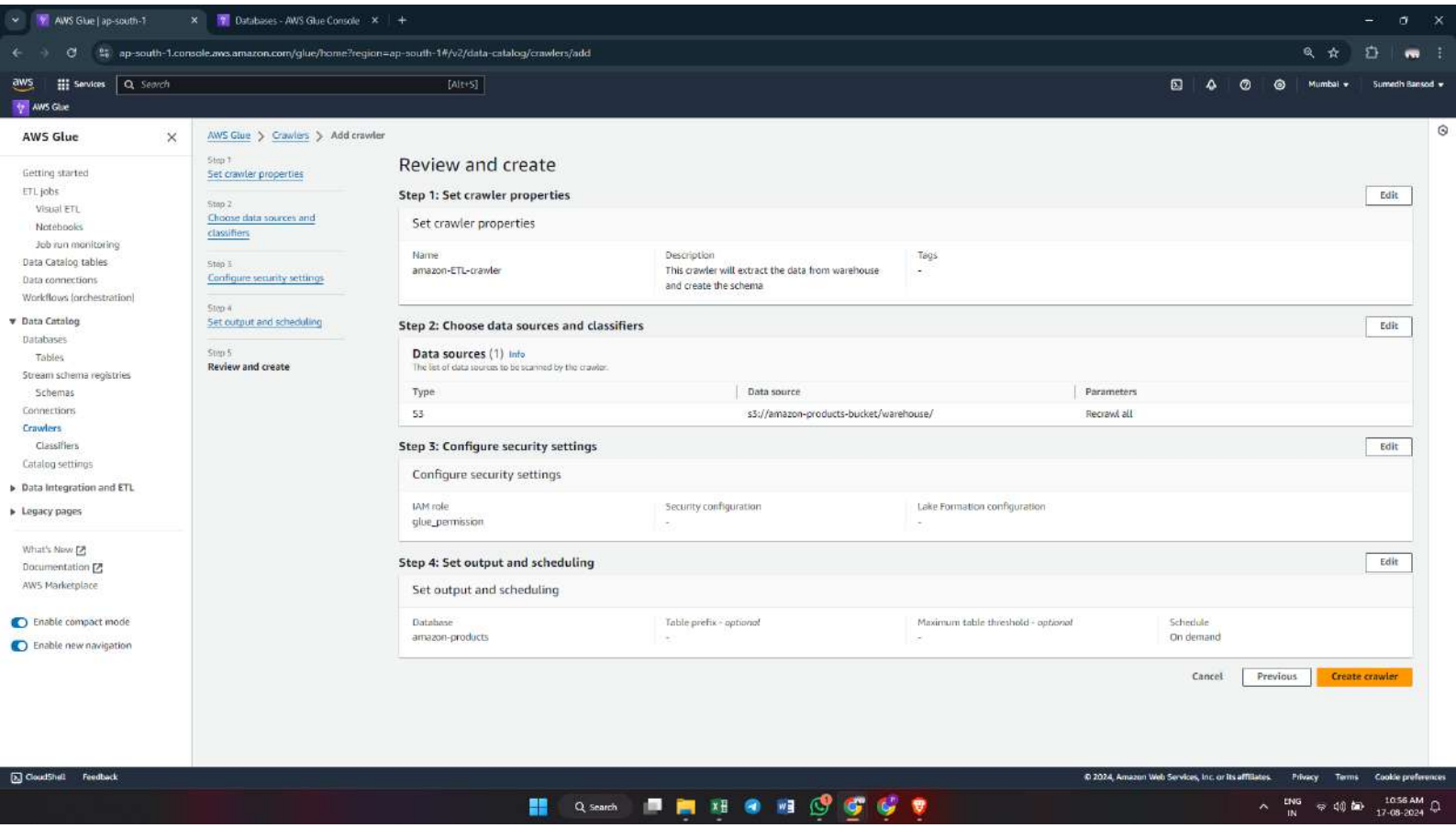

AWS Glue | ap-south-1
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Add data source
Data source
Choose the source of data to be crawled.
S3
Network connection - optional
Optionally include a Network connection to use with this S3 target. Note that each crawler is limited to one Network connection so any other S3 targets will also use the same connection (or none, if left blank).
Clear selection
Add new connection
Location of S3 data
In this account
In a different account
S3 path
Browse for or enter an existing S3 path.
s3://amazon-products-bucket/wareh
View
Browse S3
All folders and files contained in the S3 path are crawled. For example, type s3://MyBucket/MyFolder/ to crawl all objects in MyFolder within MyBucket.
Subsequent crawler runs
This field is a global field that affects all S3 data sources.
Crawl all sub-folders
Crawl all folders again with every subsequent crawl.
Crawl new sub-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.
Crawl based on events
Rely on Amazon S3 events to control what folders to crawl.
Sample only a subset of files
Exclude files matching pattern
Cancel
Add an S3 data source
Search
ENG
IN
10:54 AM
17-08-2024

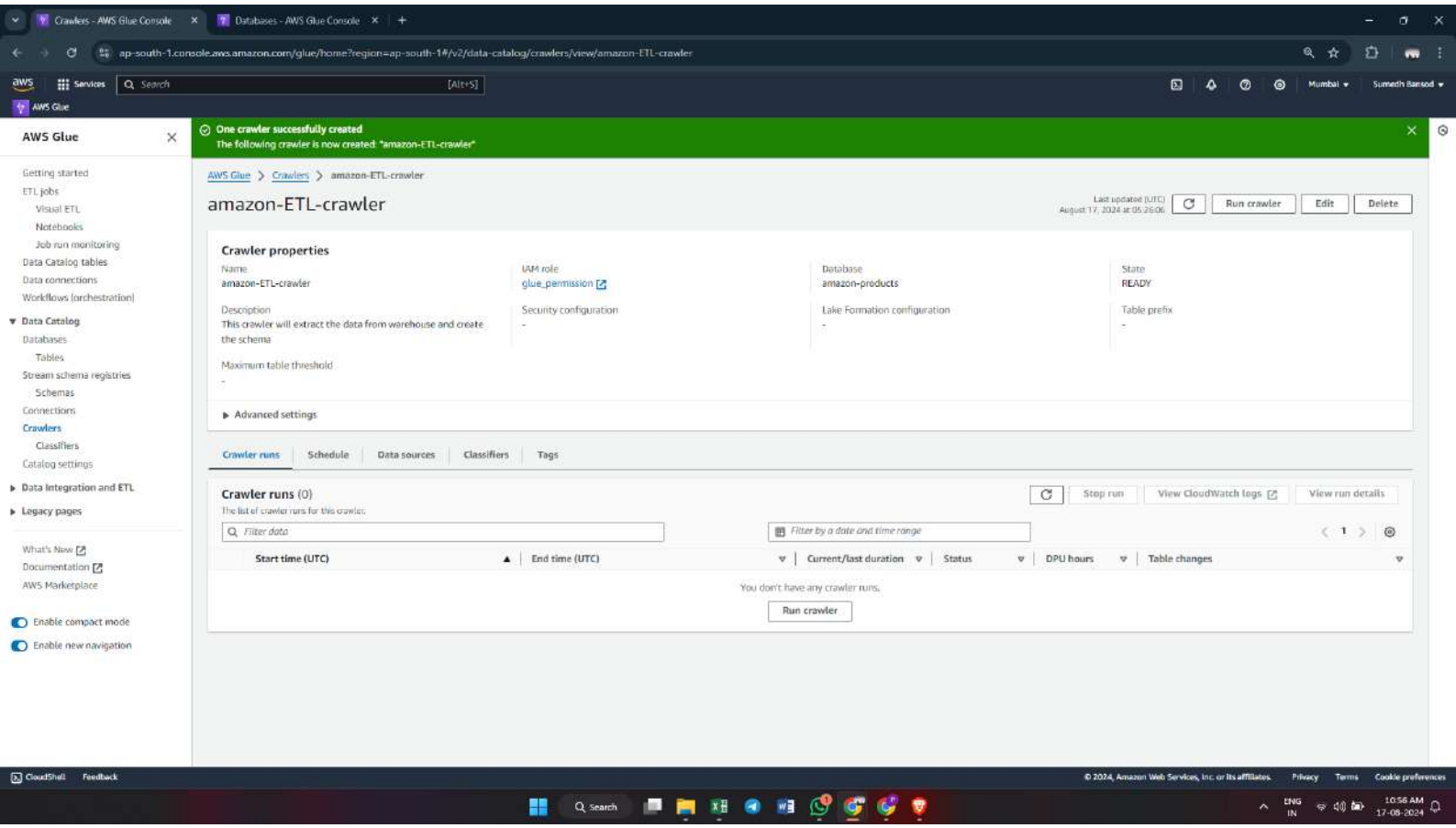

AWS Glue | ap-south-1
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Configure security settings
IAM role Info
Existing IAM role
glue_permission
View
Create new IAM role
Update chosen IAM role
Only IAM roles created by the AWS Glue console and have the prefix "AWSGlueServiceRole-" can be updated.
Lake Formation configuration - optional
Allow the crawler to use Lake Formation credentials for crawling the data source. Learn more.
Use Lake Formation credentials for crawling S3 data source
Checking this box will allow the crawler to use Lake Formation credentials for crawling the data source. If the data source is registered in another account, you must provide the registered account ID. Otherwise, the crawler will crawl only those data sources associated to the account. Only applicable to S3, Glue Catalog, Iceberg, and Hudi data sources.
Security configuration - optional
Enable at-rest encryption with a security configuration.
Cancel
Previous
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:54 AM
17-08-2024

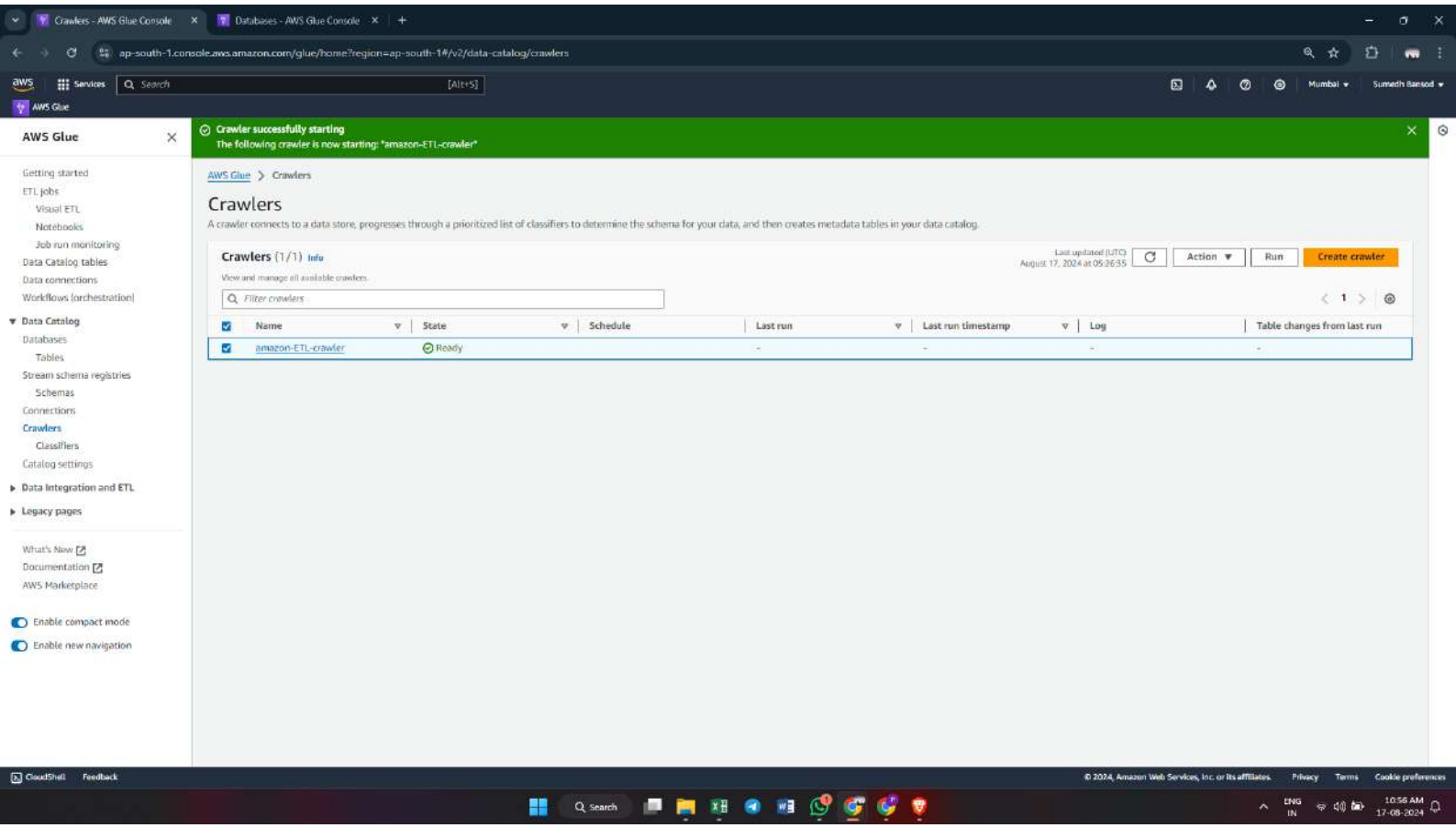

AWS Glue | ap-south-1
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/databases/add
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
One database successfully deleted
The following database is now deleted: "amazon_products"
AWS Glue > Databases > Add database
Create a database
Create a database in the AWS Glue Data Catalog.
Database details
Name
Enter a unique database name
Database name is required, in lowercase characters, and no longer than 255 characters.
Description - optional
Enter text
Descriptions can be up to 2048 characters long.
Database settings
Location - optional
Set the URI location for use by clients of the Data Catalog.
Cancel
Create database
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:55 AM
17-08-2024

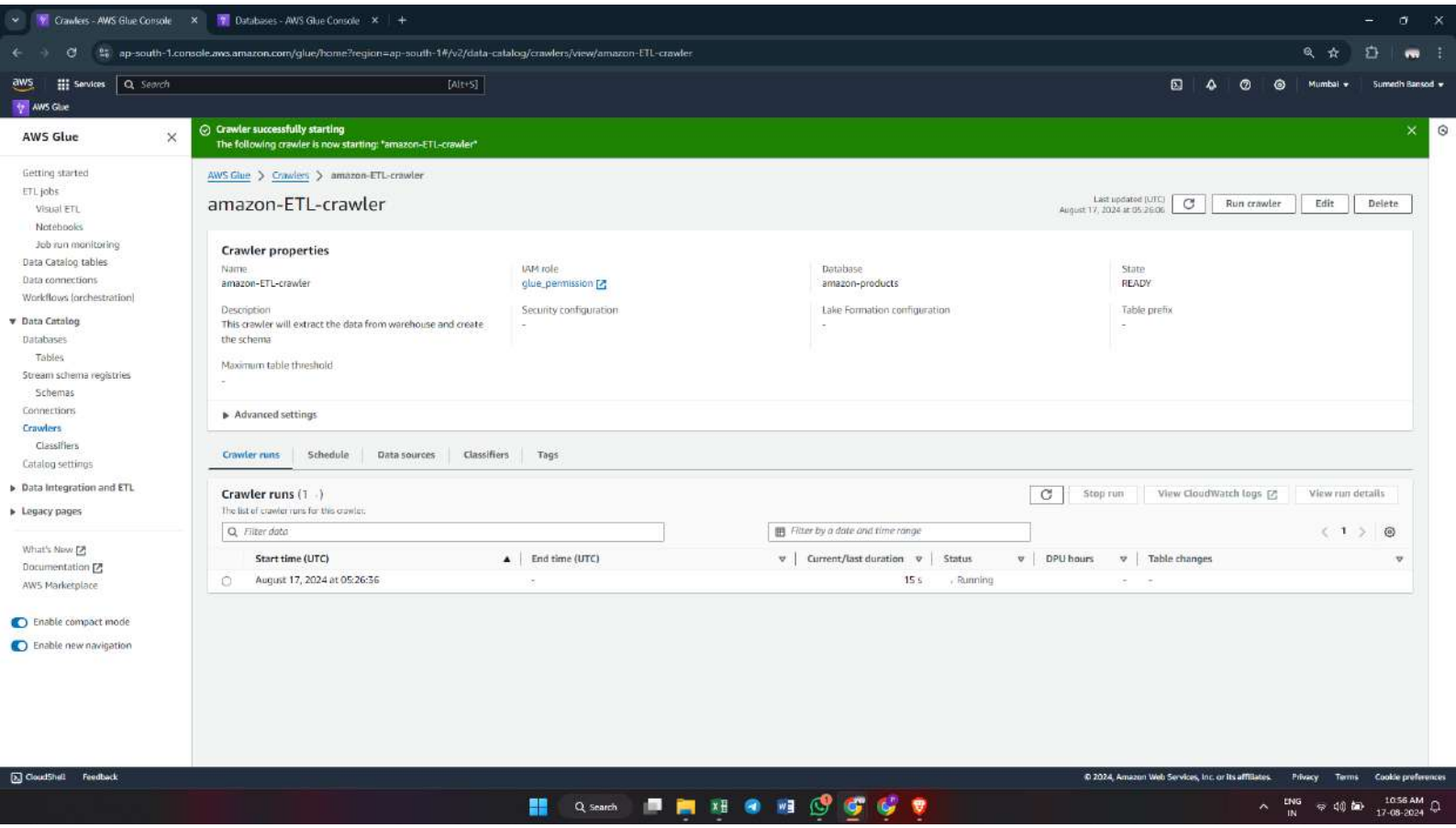

AWS Glue | ap-south-1
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/databases
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
One database successfully deleted
The following database is now deleted: "amazon_products"
AWS Glue > Databases
Databases (1)
A database is a set of associated table definitions, organized into a logical group.
Last updated (UTC)
August 17, 2024 at 05:25:23
Edit
Delete
Add database
Filter databases
Name
Description
Location URI
Created on (UTC)
amazon-products
-
-
August 17, 2024 at 05:25:23
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:55 AM
17-08-2024

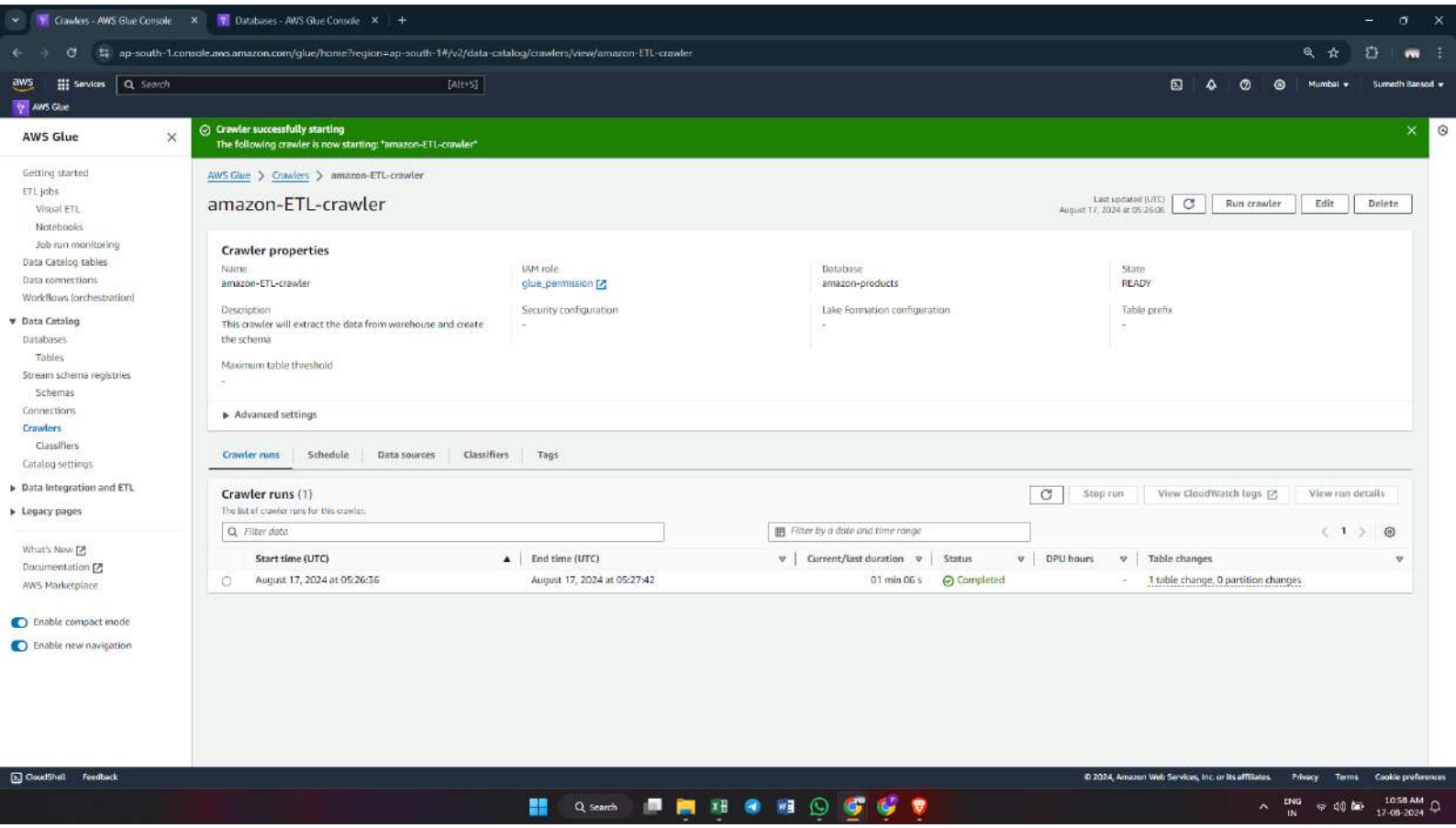

AWS Glue | ap-south-1
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Mumbai
Sumeth Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Set output and scheduling
Output configuration Info
Target database
amazon-products
Clear selection
Add database
Table name prefix - optional
Type a prefix added to table names
Maximum table threshold - optional
This field sets the maximum number of tables the crawler is allowed to generate. In the event that this number is surpassed, the crawl will fail with an error. If not set, the crawler will automatically generate the number of tables depending on the data schema.
Type a number greater than 0
Advanced options
Crawler schedule
You can define a time-based schedule for your crawlers and jobs in AWS Glue. The definition of these schedules uses the Unix-like cron syntax. Learn more
Frequency
On demand
Cancel
Previous
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:55 AM
17-08-2024

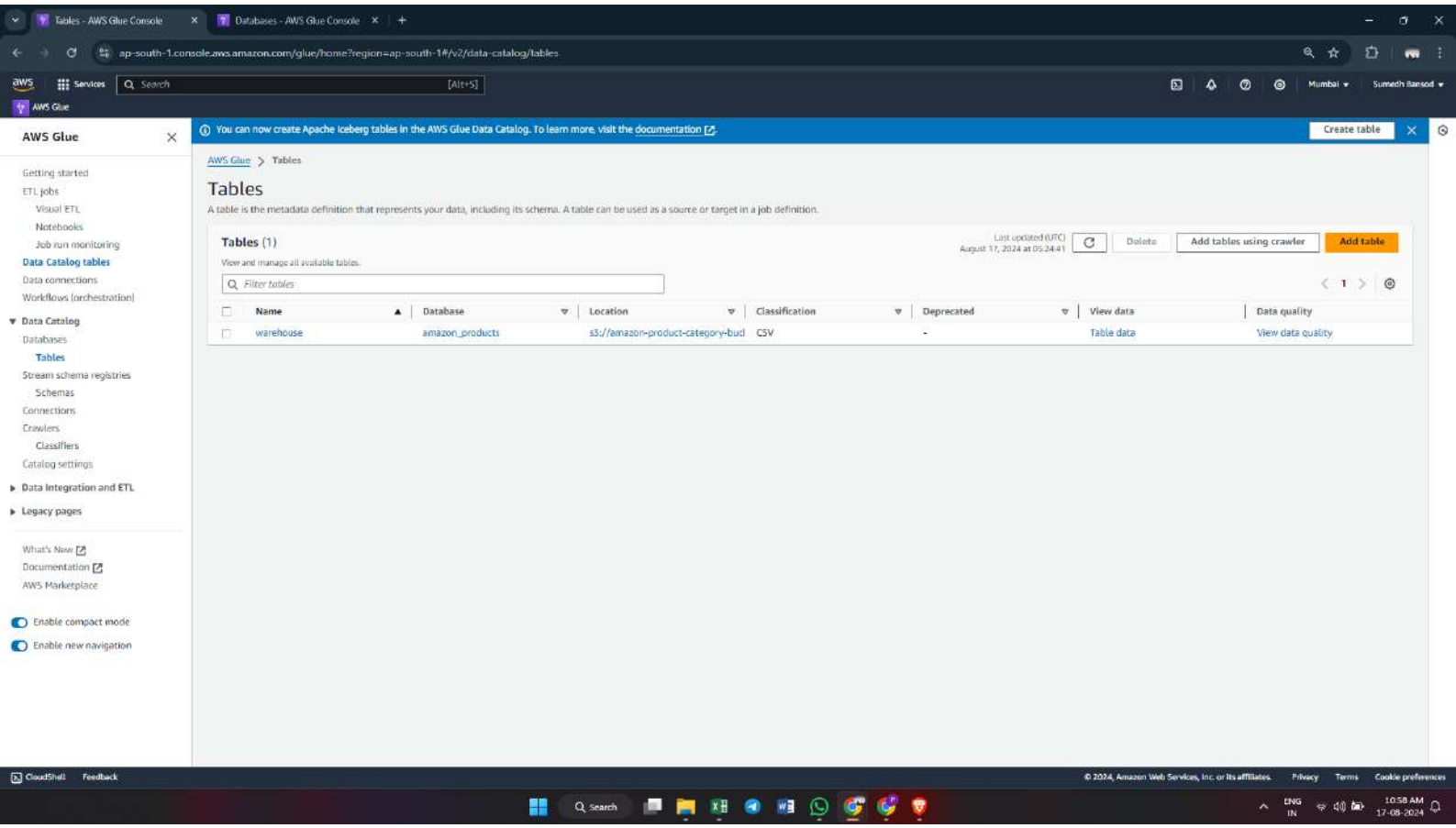

AWS Glue | ap-south-1
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Review and create
Step 1: Set crawler properties
Edit
Set crawler properties
Name
amazon-ETL-crawler
Description
This crawler will extract the data from warehouse and create the schema
Tags
-
Step 2: Choose data sources and classifiers
Edit
Data sources (1) Info
The list of data sources to be scanned by the crawler.
Type
Data source
Parameters
S3
s3://amazon-products-bucket/warehouse/
Recrawl all
Step 3: Configure security settings
Edit
Configure security settings
IAM role
glue_permission
Security configuration
-
Lake Formation configuration
-
Step 4: Set output and scheduling
Edit
Set output and scheduling
Database
amazon-products
Table prefix - optional
-
Maximum table threshold - optional
-
Schedule
On demand
Cancel
Previous
Create crawler
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:56 AM
17-08-2024


Crawlers - AWS Glue Console
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/view/amazon-ETL-crawler
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
One crawler successfully created
The following crawler is now created: "amazon-ETL-crawler"
AWS Glue > Crawlers > amazon-ETL-crawler
amazon-ETL-crawler
Last updated (UTC)
August 17, 2024 at 05:26:06
Run crawler
Edit
Delete
Crawler properties
Name
amazon-ETL-crawler
IAM role
glue_permission
Database
amazon-products
State
READY
Description
This crawler will extract the data from warehouse and create the schema
Security configuration
-
Lake Formation configuration
-
Table prefix
-
Maximum table threshold
-
Advanced settings
Crawler runs
Schedule
Data sources
Classifiers
Tags
Crawler runs (0)
The list of crawler runs for this crawler.
Stop run
View CloudWatch logs
View run details
Filter data
Filter by a date and time range
Start time (UTC)
End time (UTC)
Current/last duration
Status
DPU hours
Table changes
You don't have any crawler runs.
Run crawler
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:56 AM
17-08-2024


Crawlers - AWS Glue Console
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Crawler successfully starting
The following crawler is now starting: "amazon-ETL-crawler"
AWS Glue > Crawlers
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.
Crawlers (1/1) Info
View and manage all available crawlers.
Last updated (UTC)
August 17, 2024 at 05:26:35
Action
Run
Create crawler
Filter crawlers
1
Name
State
Schedule
Last run
Last run timestamp
Log
Table changes from last run
amazon-ETL-crawler
Ready
-
-
-
-
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:56 AM
17-08-2024


Crawlers - AWS Glue Console
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/view/amazon-ETL-crawler
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Crawler successfully starting
The following crawler is now starting: "amazon-ETL-crawler"
AWS Glue > Crawlers > amazon-ETL-crawler
amazon-ETL-crawler
Last updated (UTC)
August 17, 2024 at 05:26:06
Run crawler
Edit
Delete
Crawler properties
Name
amazon-ETL-crawler
IAM role
glue_permission
Database
amazon-products
State
READY
Description
This crawler will extract the data from warehouse and create the schema
Security configuration
-
Lake Formation configuration
-
Table prefix
-
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
Filter data
Filter by a date and time range
Start time (UTC)
End time (UTC)
Current/last duration
Status
DPU hours
Table changes
August 17, 2024 at 05:26:36
-
15 s
Running
-
-
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:56 AM
17-08-2024


Crawlers - AWS Glue Console
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/crawlers/view/amazon-ETL-crawler
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Crawler successfully starting
The following crawler is now starting: "amazon-ETL-crawler"
AWS Glue > Crawlers > amazon-ETL-crawler
amazon-ETL-crawler
Last updated (UTC)
August 17, 2024 at 05:26:06
Run crawler
Edit
Delete
Crawler properties
Name
amazon-ETL-crawler
IAM role
glue_permission
Database
amazon-products
State
READY
Description
This crawler will extract the data from warehouse and create the schema
Security configuration
-
Lake Formation configuration
-
Table prefix
-
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
Filter data
Filter by a date and time range
Start time (UTC)
End time (UTC)
Current/last duration
Status
DPU hours
Table changes
August 17, 2024 at 05:26:36
August 17, 2024 at 05:27:42
01 min 06 s
Completed
-
1 table change, 0 partition changes
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:58 AM
17-08-2024


Tables - AWS Glue Console
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/tables
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
You can now create Apache Iceberg tables in the AWS Glue Data Catalog. To learn more, visit the documentation.
Create table
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Tables
Tables
A table is the metadata definition that represents your data, including its schema. A table can be used as a source or target in a job definition.
Tables (1)
View and manage all available tables.
Last updated (UTC)
August 17, 2024 at 05:24:41
Delete
Add tables using crawler
Add table
Filter tables
1
Name
Database
Location
Classification
Deprecated
View data
Data quality
warehouse
amazon_products
s3://amazon-product-category-bucl
CSV
-
Table data
View data quality
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
10:58 AM
17-08-2024

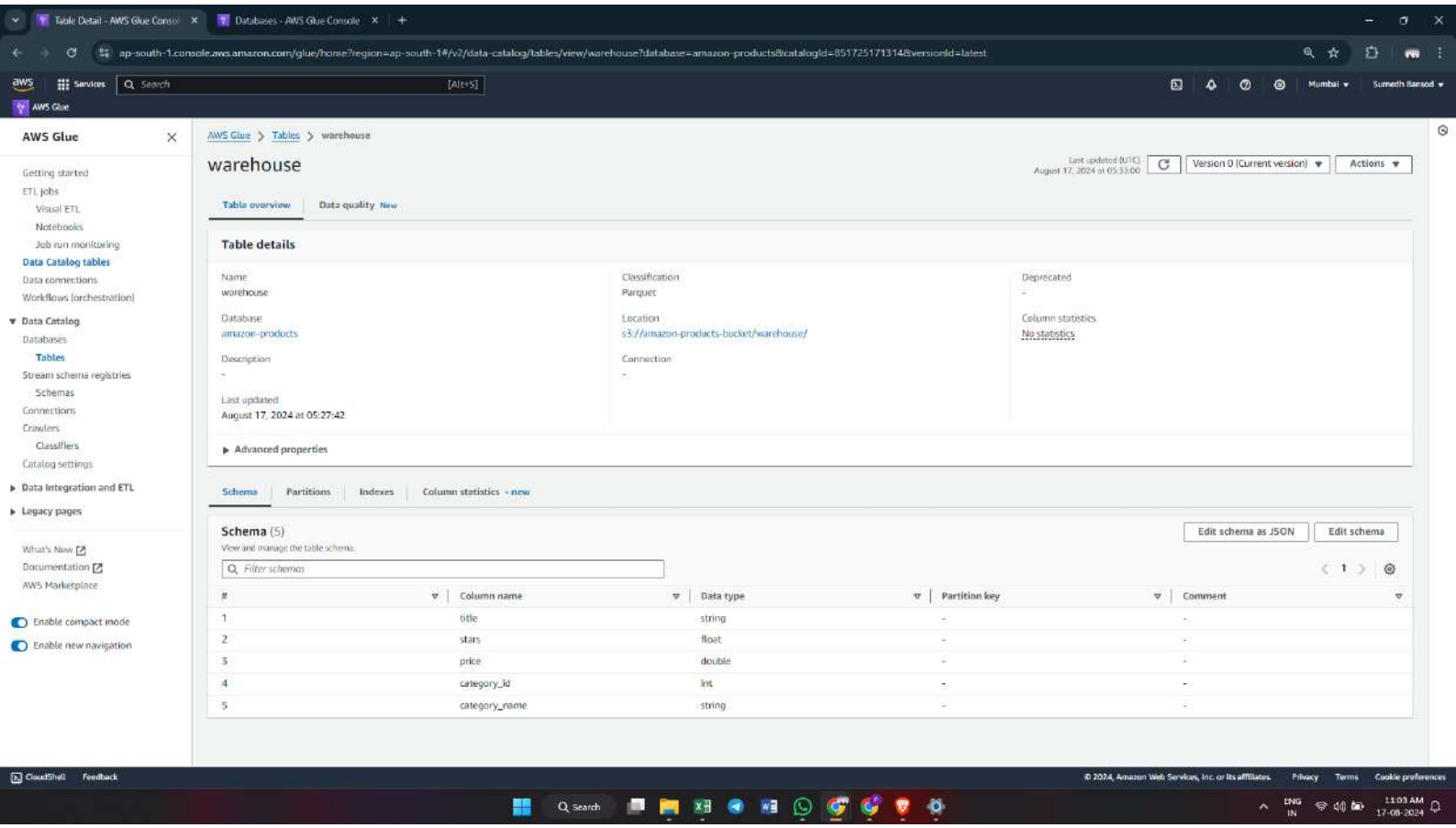


Table Detail - AWS Glue Consol
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/tables/view/warehouse?database=amazon-products&catalogId=851725171314&versionId=latest
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
AWS Glue > Tables > warehouse
warehouse
Last updated (UTC)
August 17, 2024 at 05:53:00
Version 0 (Current version)
Actions
Table overview
Data quality New
Table details
Name
warehouse
Database
amazon-products
Description
-
Last updated
August 17, 2024 at 05:27:42
Classification
Parquet
Location
s3://amazon-products-bucket/warehouse/
Connection
-
Deprecated
-
Column statistics
No statistics
Advanced properties
Schema
Partitions
Indexes
Column statistics - new
Schema (5)
View and manage the table schema.
Edit schema as JSON
Edit schema
Filter schemas
1
| Column name | Data type | Partition key | Comment
1 | title | string | - | -
2 | stars | float | - | -
3 | price | double | - | -
4 | category_id | int | - | -
5 | category_name | string | - | -
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:03 AM
17-08-2024

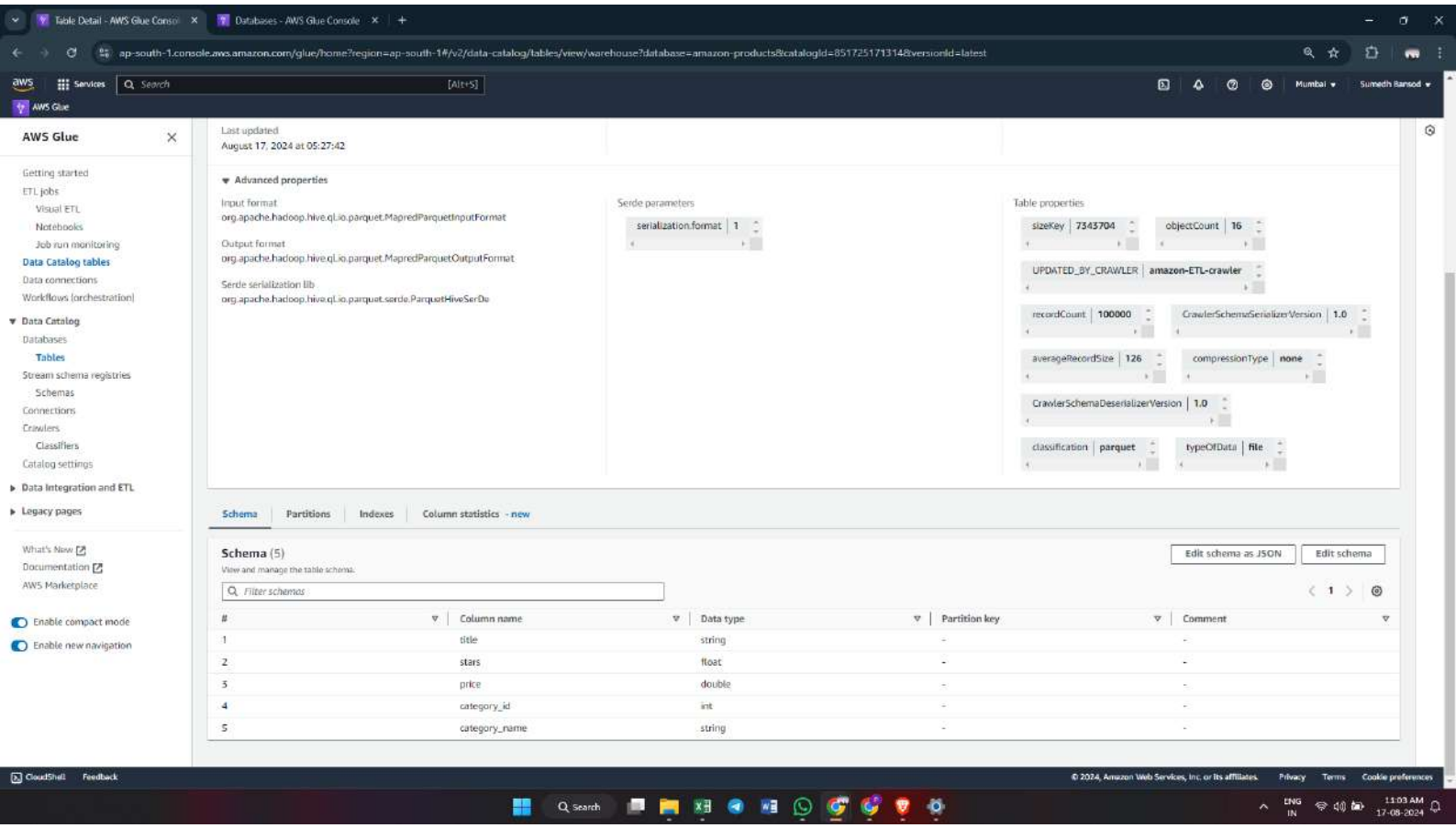
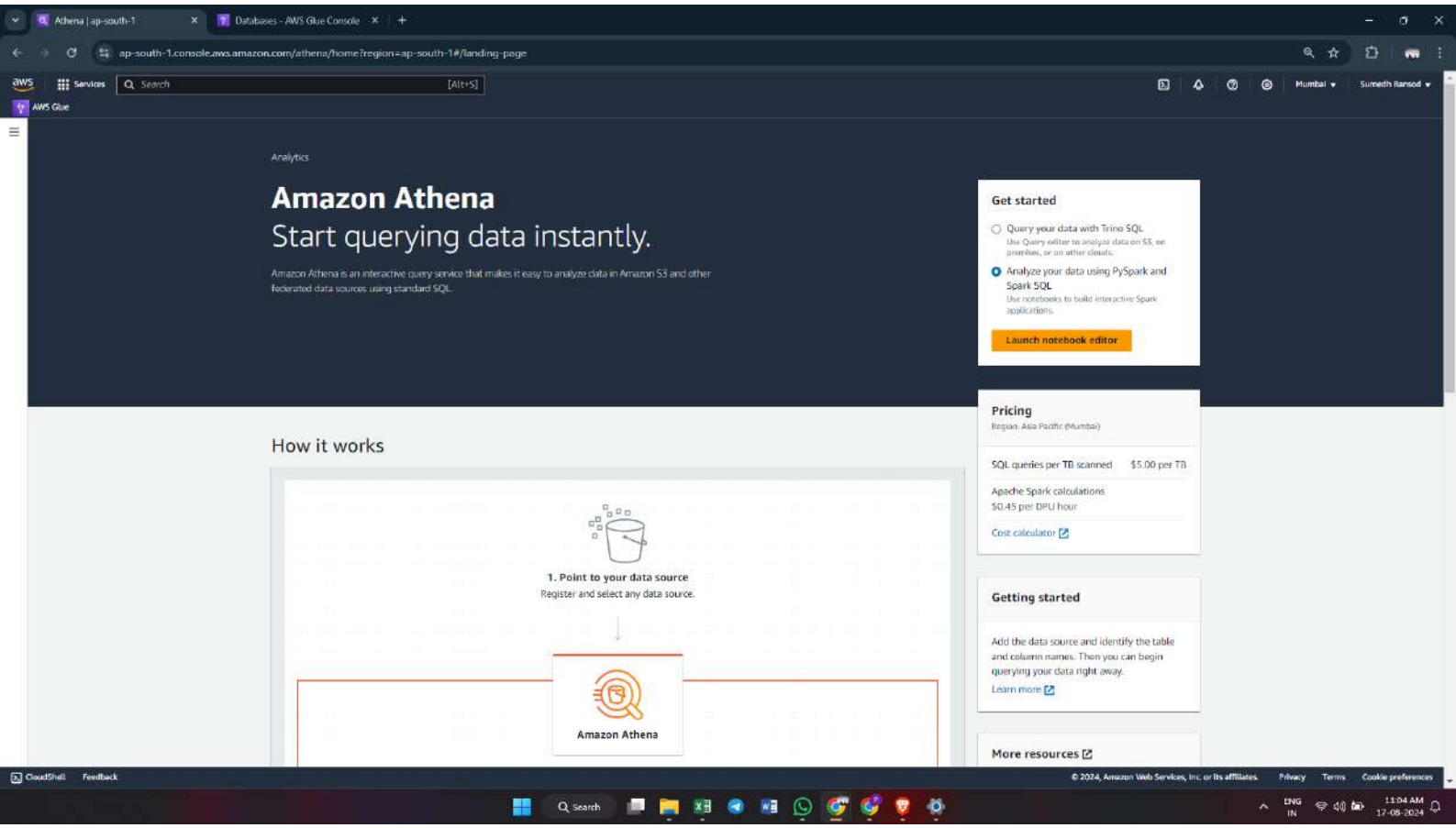
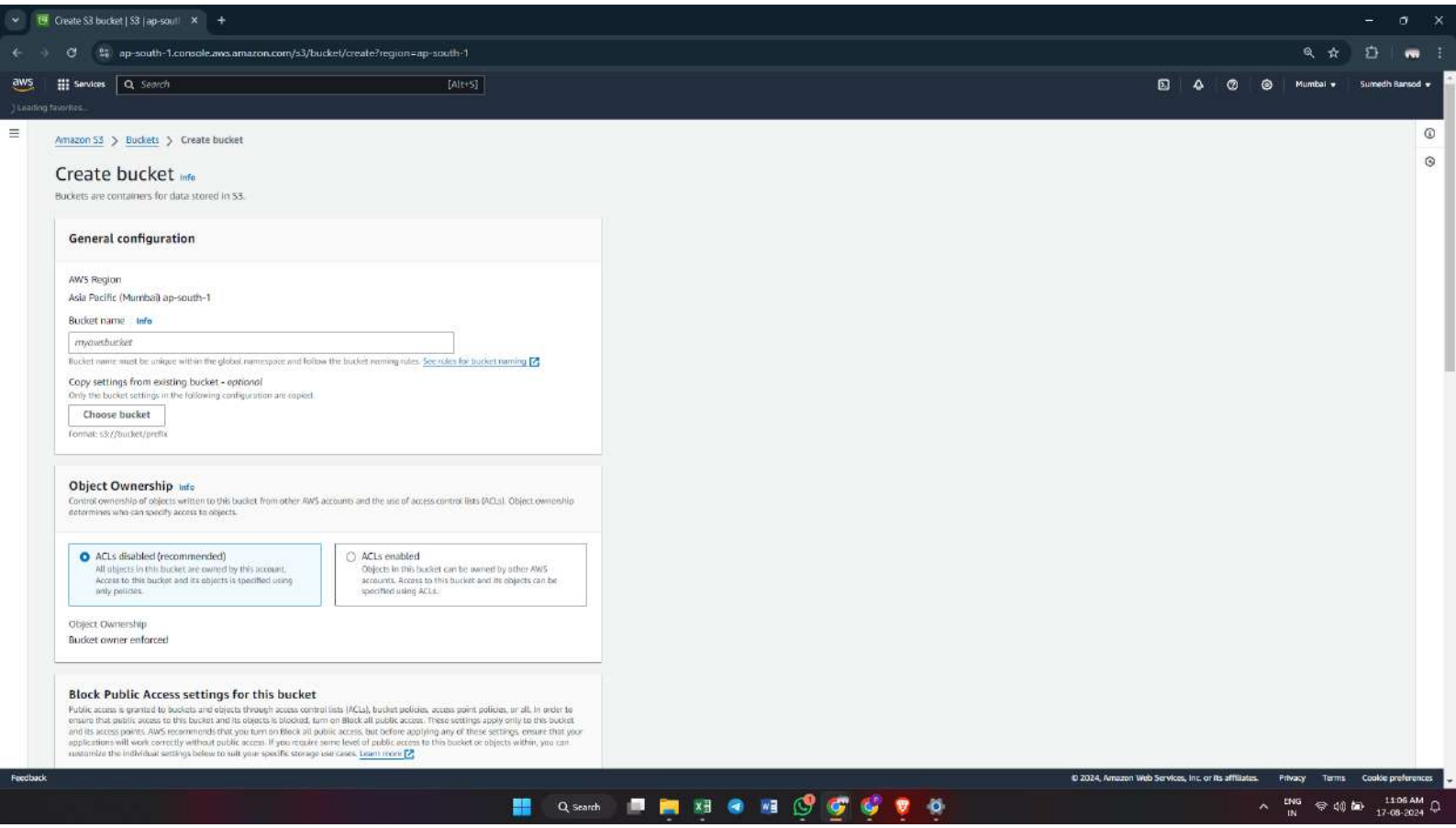
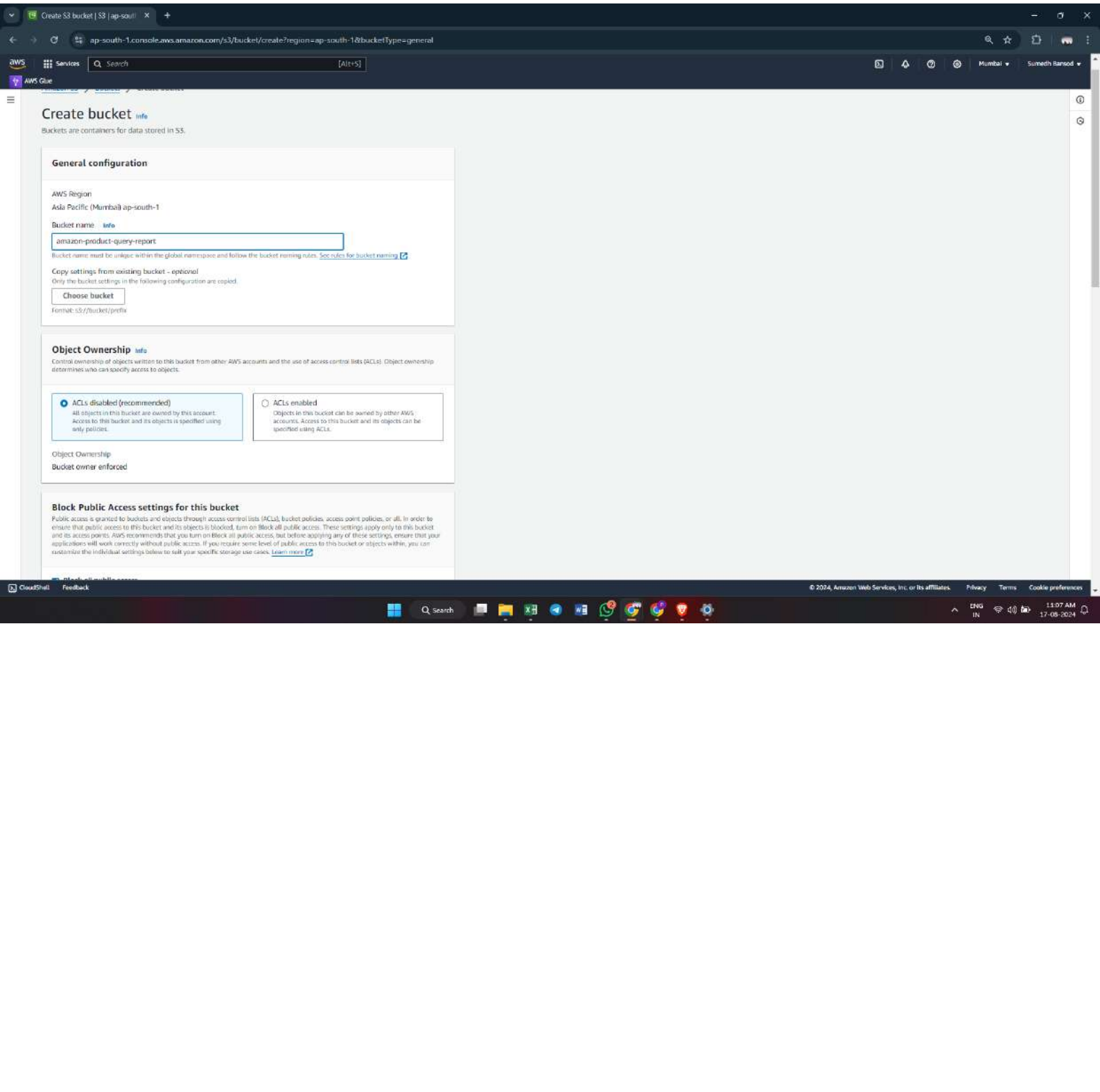
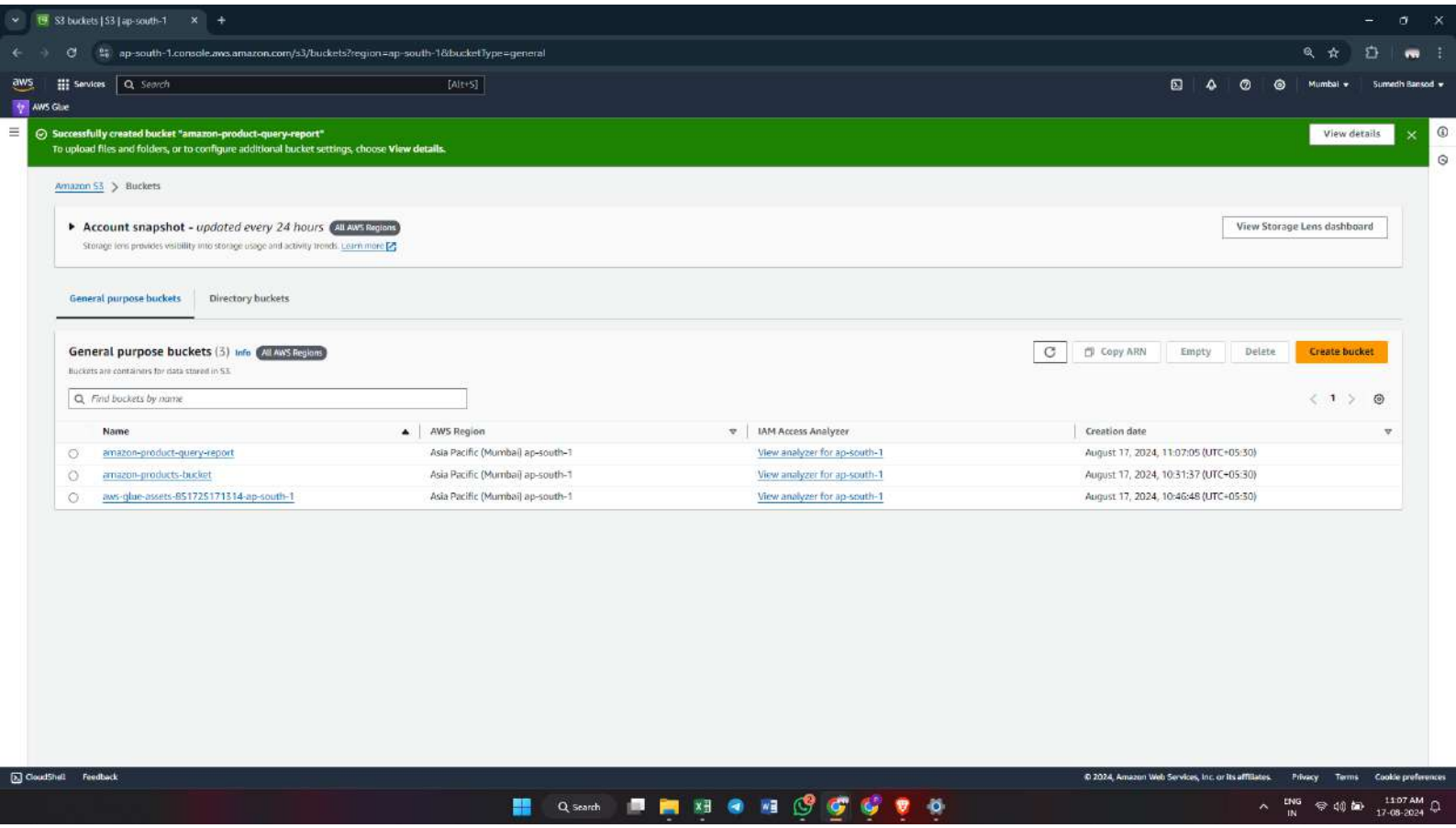


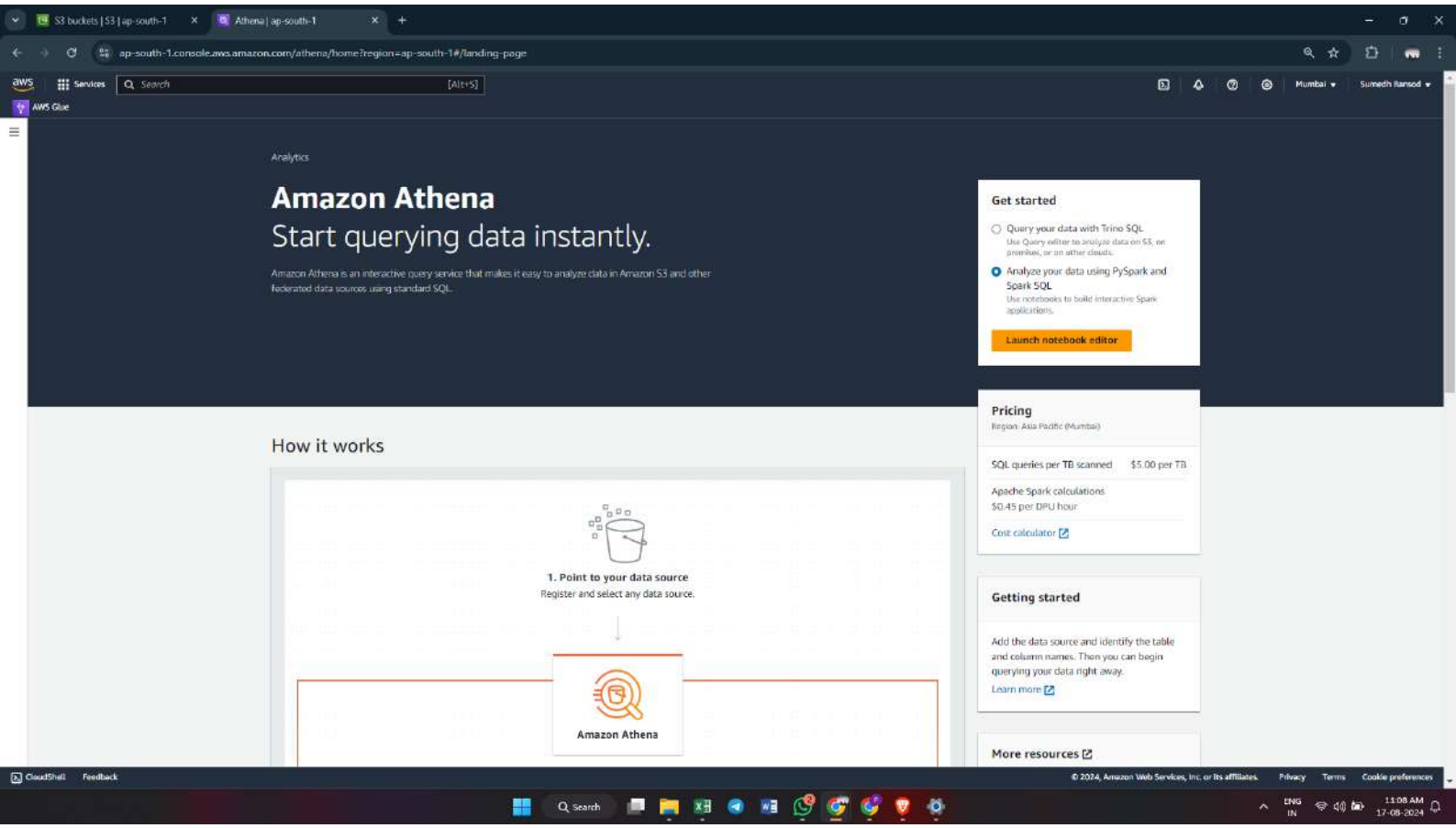
Table Detail - AWS Glue Consol
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/glue/home?region=ap-south-1#/v2/data-catalog/tables/view/warehouse?database=amazon-products&catalogId=851725171314&versionId=latest
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Last updated
August 17, 2024 at 05:27:42
Advanced properties
Input format
org.apache.hadoop.hive.ql.io.parquet.MapredParquetInputFormat
Output format
org.apache.hadoop.hive.ql.io.parquet.MapredParquetOutputFormat
Serde serialization lib
org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe
Serde parameters
serialization.format 1
Table properties
sizeKey 7343704
objectCount 16
UPDATED_BY_CRAWLER amazon-ETL-crawler
recordCount 100000
CrawlerSchemaSerializerVersion 1.0
averageRecordSize 126
compressionType none
CrawlerSchemaDeserializerVersion 1.0
classification parquet
typeOfData file
Schema
Partitions
Indexes
Column statistics - new
Schema (5)
View and manage the table schema.
Edit schema as JSON
Edit schema
Filter schemas
1
Column name Data type Partition key Comment
1 title string - -
2 stars float - -
3 price double - -
4 category_id int - -
5 category_name string - -
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:03 AM
17-08-2024

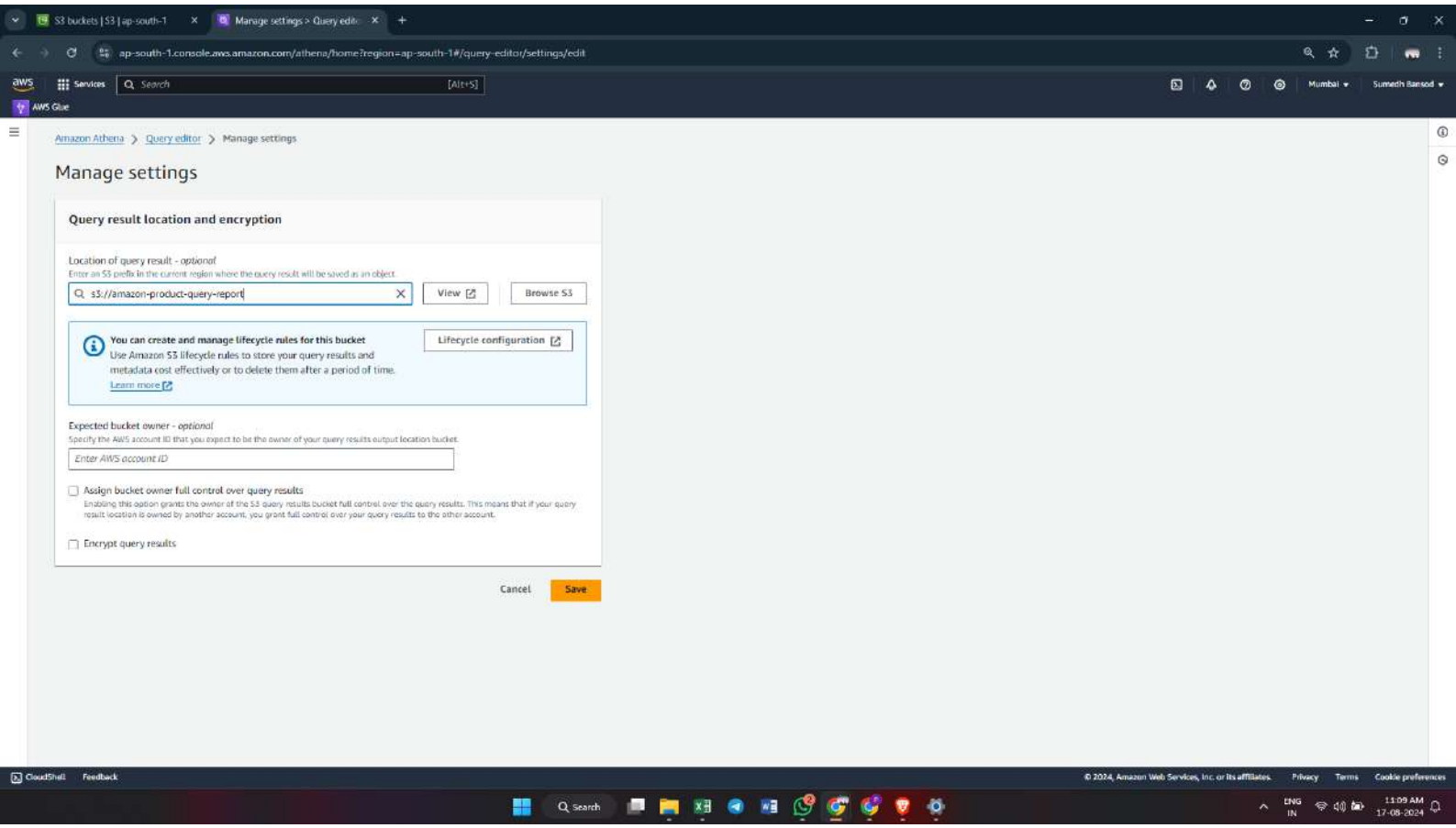

Athena | ap-south-1
Databases - AWS Glue Console
ap-south-1.console.aws.amazon.com/athena/home?region=ap-south-1#/landing-page
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Analytics
Amazon Athena
Start querying data instantly.
Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 and other federated data sources using standard SQL.
Get started
Query your data with Trino SQL
Use Query editor to analyze data on S3, on premises, or on other clouds.
Analyze your data using PySpark and Spark SQL
Use notebooks to build interactive Spark applications.
Launch notebook editor
Pricing
Region: Asia Pacific (Mumbai)
SQL queries per TB scanned
$5.00 per TB
Apache Spark calculations
$0.45 per DPU hour
Cost calculator
Getting started
Add the data source and identify the table and column names. Then you can begin querying your data right away.
Learn more
More resources
How it works
1. Point to your data source
Register and select any data source.
Amazon Athena
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:04 AM
17-08-2024

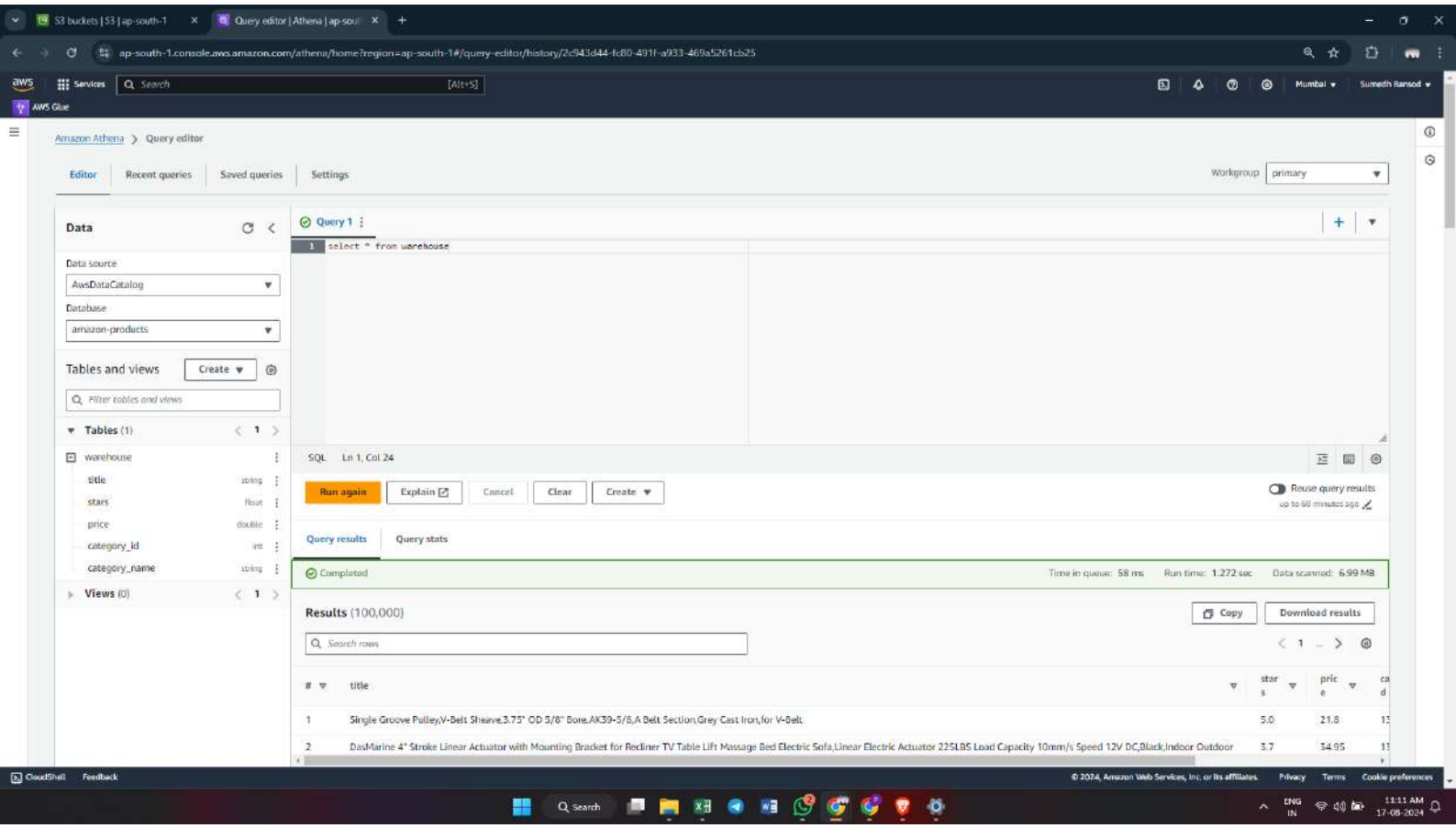

Create S3 bucket | S3 | ap-south1
ap-south-1.console.aws.amazon.com/s3/bucket/create?region=ap-south-1
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
Amazon S3 > Buckets > Create bucket
Create bucket Info
Buckets are containers for data stored in S3.
General configuration
AWS Region
Asia Pacific (Mumbai) ap-south-1
Bucket name Info
myawsbucket
Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming
Copy settings from existing bucket - optional
Only the bucket settings in the following configuration are copied.
Choose bucket
Format: s3://bucket/prefix
Object Ownership Info
Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.
ACLs disabled (recommended)
All objects in this bucket are owned by this account. Access to this bucket and its objects is specified using only policies.
ACLs enabled
Objects in this bucket can be owned by other AWS accounts. Access to this bucket and its objects can be specified using ACLs.
Object Ownership
Bucket owner enforced
Block Public Access settings for this bucket
Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to this bucket and its objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to this bucket or objects within, you can customize the individual settings below to suit your specific storage use cases. Learn more
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:06 AM
17-08-2024

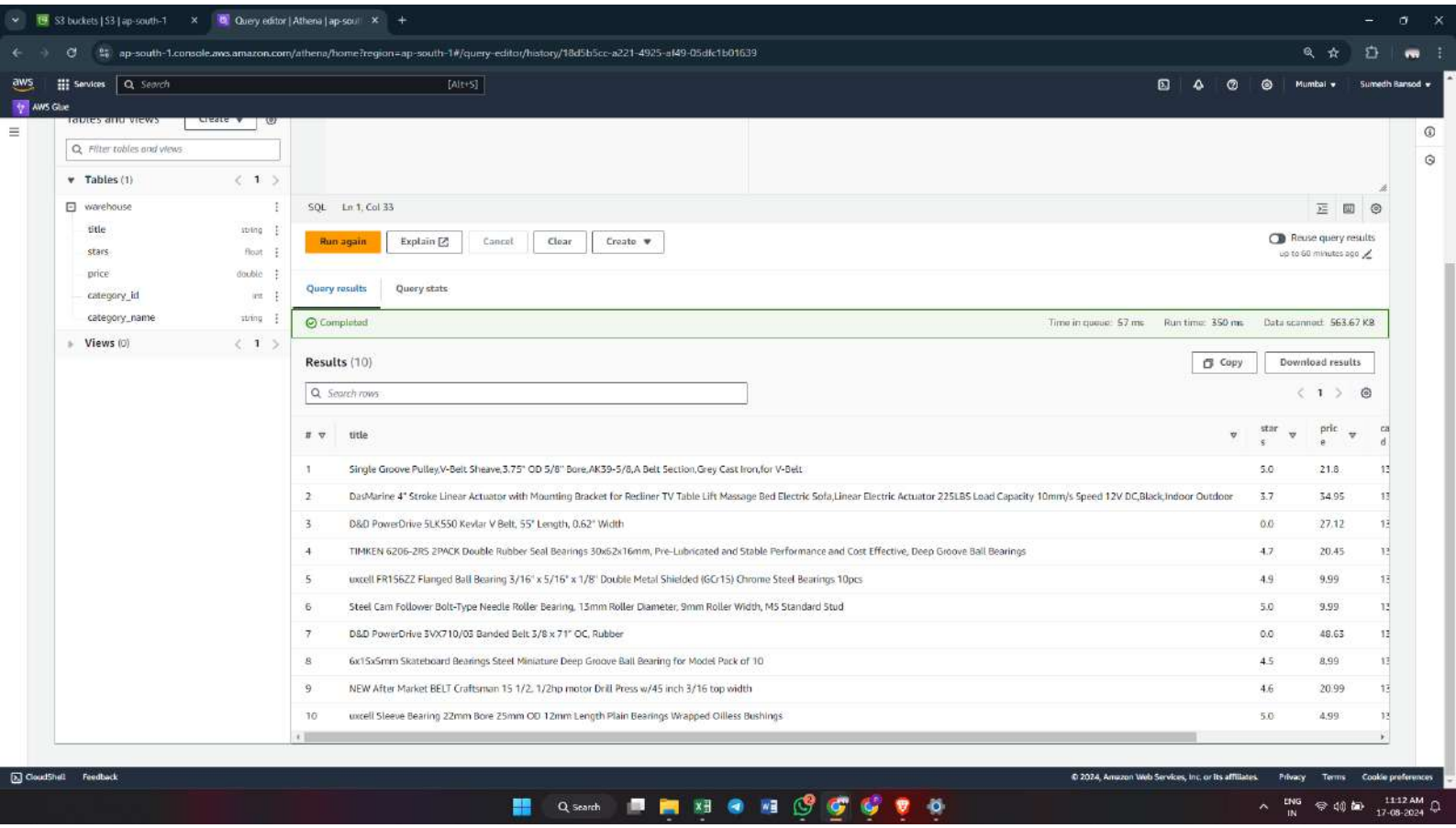

Create S3 bucket | S3 | ap-south-1
ap-south-1.console.aws.amazon.com/s3/bucket/create?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Create bucket Info
Buckets are containers for data stored in S3.
General configuration
AWS Region
Asia Pacific (Mumbai) ap-south-1
Bucket name Info
amazon-product-query-report
Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming
Copy settings from existing bucket - optional
Only the bucket settings in the following configuration are copied.
Choose bucket
Format: s3://bucket/prefix
Object Ownership Info
Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.
ACLs disabled (recommended)
All objects in this bucket are owned by this account. Access to this bucket and its objects is specified using only policies.
ACLs enabled
Objects in this bucket can be owned by other AWS accounts. Access to this bucket and its objects can be specified using ACLs.
Object Ownership
Bucket owner enforced
Block Public Access settings for this bucket
Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to this bucket and its objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to this bucket or objects within, you can customize the individual settings below to suit your specific storage use cases. Learn more
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:07 AM
17-08-2024


S3 buckets | S3 | ap-south-1
ap-south-1.console.aws.amazon.com/s3/buckets?region=ap-south-1&bucketType=general
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Successfully created bucket "amazon-product-query-report"
To upload files and folders, or to configure additional bucket settings, choose View details.
View details
Amazon S3 > Buckets
Account snapshot - updated every 24 hours
All AWS Regions
Storage lens provides visibility into storage usage and activity trends. Learn more
View Storage Lens dashboard
General purpose buckets
Directory buckets
General purpose buckets (3) Info
All AWS Regions
Buckets are containers for data stored in S3.
Copy ARN
Empty
Delete
Create bucket
Find buckets by name
Name
AWS Region
IAM Access Analyzer
Creation date
amazon-product-query-report
Asia Pacific (Mumbai) ap-south-1
View analyzer for ap-south-1
August 17, 2024, 11:07:05 (UTC+05:30)
amazon-products-bucket
Asia Pacific (Mumbai) ap-south-1
View analyzer for ap-south-1
August 17, 2024, 10:31:37 (UTC+05:30)
aws-glue-assets-851725171314-ap-south-1
Asia Pacific (Mumbai) ap-south-1
View analyzer for ap-south-1
August 17, 2024, 10:46:48 (UTC+05:30)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:07 AM
17-08-2024


S3 buckets | S3 | ap-south-1
Athena | ap-south-1
ap-south-1.console.aws.amazon.com/athena/home?region=ap-south-1#/landing-page
aws
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Analytics
Amazon Athena
Start querying data instantly.
Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 and other federated data sources using standard SQL.
Get started
Query your data with Trino SQL
Use Query editor to analyze data on S3, on premises, or on other clouds.
Analyze your data using PySpark and Spark SQL
Use notebooks to build interactive Spark applications.
Launch notebook editor
Pricing
Region: Asia Pacific (Mumbai)
SQL queries per TB scanned
$5.00 per TB
Apache Spark calculations
$0.45 per DPU hour
Cost calculator
Getting started
Add the data source and identify the table and column names. Then you can begin querying your data right away.
Learn more
More resources
How it works
1. Point to your data source
Register and select any data source.
Amazon Athena
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:08 AM
17-08-2024


S3 buckets | S3 | ap-south-1
Manage settings > Query editor
ap-south-1.console.aws.amazon.com/athena/home?region=ap-south-1#/query-editor/settings/edit
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon Athena > Query editor > Manage settings
Manage settings
Query result location and encryption
Location of query result - optional
Enter an S3 prefix in the current region where the query result will be saved as an object.
s3://amazon-product-query-report
View
Browse S3
You can create and manage lifecycle rules for this bucket
Use Amazon S3 lifecycle rules to store your query results and metadata cost effectively or to delete them after a period of time.
Learn more
Lifecycle configuration
Expected bucket owner - optional
Specify the AWS account ID that you expect to be the owner of your query results output location bucket.
Enter AWS account ID
Assign bucket owner full control over query results
Enabling this option grants the owner of the S3 query results bucket full control over the query results. This means that if your query result location is owned by another account, you grant full control over your query results to the other account.
Encrypt query results
Cancel
Save
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:09 AM
17-08-2024


S3 buckets | S3 | ap-south-1
Query editor | Athena | ap-south-1
ap-south-1.console.aws.amazon.com/athena/home?region=ap-south-1#/query-editor/history/2c943d44-fc80-491f-a933-469a5261cb25
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Amazon Athena > Query editor
Editor
Recent queries
Saved queries
Settings
Workgroup
primary
Data
Data source
AwsDataCatalog
Database
amazon-products
Tables and views
Create
Filter tables and views
Tables (1)
warehouse
title string
stars float
price double
category_id int
category_name string
Views (0)
Query 1
1 select * from warehouse
SQL Ln 1, Col 24
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 58 ms
Run time: 1.272 sec
Data scanned: 6.99 MB
Results (100,000)
Copy
Download results
Search rows
title stars price cat
1 Single Groove Pulley,V-Belt Sheave,3.75" OD 5/8" Bore,AK39-5/8,A Belt Section,Grey Cast Iron,for V-Belt 5.0 21.8
2 DasMarine 4" Stroke Linear Actuator with Mounting Bracket for Recliner TV Table Lift Massage Bed Electric Sofa,Linear Electric Actuator 225LBS Load Capacity 10mm/s Speed 12V DC,Black,Indoor Outdoor 3.7 34.95
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
IN
11:11 AM
17-08-2024


S3 buckets | S3 | ap-south-1
Query editor | Athena | ap-south-1
ap-south-1.console.aws.amazon.com/athena/home?region=ap-south-1#/query-editor/history/18d5b5cc-a221-4925-af49-05dfc1b01639
Services
Search
[Alt+S]
Mumbai
Sumedh Bansod
AWS Glue
Tables and views
Create
Filter tables and views
Tables (1)
warehouse
title string
stars float
price double
category_id int
category_name string
Views (0)
SQL Ln 1, Col 33
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 57 ms
Run time: 350 ms
Data scanned: 563.67 KB
Results (10)
Copy
Download results
Search rows
title stars price
1 Single Groove Pulley,V-Belt Sheave,3.75" OD 5/8" Bore,AK39-5/8,A Belt Section,Grey Cast Iron,for V-Belt 5.0 21.8
2 DasMarine 4" Stroke Linear Actuator with Mounting Bracket for Recliner TV Table Lift Massage Bed Electric Sofa,Linear Electric Actuator 225LBS Load Capacity 10mm/s Speed 12V DC,Black,Indoor Outdoor 3.7 34.95
3 D&D PowerDrive 5LK550 Kevlar V Belt, 55" Length, 0.62" Width 0.0 27.12
4 TIMKEN 6206-2RS 2PACK Double Rubber Seal Bearings 30x62x16mm, Pre-Lubricated and Stable Performance and Cost Effective, Deep Groove Ball Bearings 4.7 20.45
5 uxcell FR156ZZ Flanged Ball Bearing 3/16" x 5/16" x 1/8" Double Metal Shielded (GCr15) Chrome Steel Bearings 10pcs 4.9 9.99
6 Steel Cam Follower Bolt-Type Needle Roller Bearing, 13mm Roller Diameter, 9mm Roller Width, M5 Standard Stud 5.0 9.99
7 D&D PowerDrive 3VX710/03 Banded Belt 3/8 x 71" OC, Rubber 0.0 48.63
8 6x15x5mm Skateboard Bearings Steel Miniature Deep Groove Ball Bearing for Model Pack of 10 4.5 8.99
9 NEW After Market BELT Craftsman 15 1/2, 1/2hp motor Drill Press w/45 inch 3/16 top width 4.6 20.99
10 uxcell Sleeve Bearing 22mm Bore 25mm OD 12mm Length Plain Bearings Wrapped Oilless Bushings 5.0 4.99
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG IN
11:12 AM 17-08-2024