

# **SNOWFLAKE INTEGRATION WITH AWS S3 STORAGE**

**Group ID – 04: Varun Randive, Aneri Vasani, Shivam Singh**

**UMB Email:** [v.randive001@umb.edu](mailto:v.randive001@umb.edu), [anerijiten.vasani001@umb.edu](mailto:anerijiten.vasani001@umb.edu), [s.singh003@umb.edu](mailto:s.singh003@umb.edu)

**Note: Please Run every code written in snowflake by selecting the code and then Running it**

## **Introduction**

In today's data-driven world, efficient data management and analysis are crucial for businesses to stay competitive. Snowflake, a cloud-based data warehousing platform, offers powerful capabilities for managing and analyzing data at scale. One key aspect of Snowflake's flexibility is its ability to integrate with various data sources, including Amazon S3 (Simple Storage Service).

Amazon S3 is a highly scalable, secure, and durable object storage service provided by Amazon Web Services (AWS). By integrating Snowflake with S3, organizations can leverage the storage capabilities of S3 to store and access large volumes of data while utilizing Snowflake's analytical prowess for querying and analysis.

The integration of Snowflake with AWS S3 is particularly useful in scenarios where organizations need to store and analyze large datasets efficiently. For instance, consider a retail company that collects vast amounts of transactional data daily. By storing this data in an S3 bucket, the company can take advantage of S3's scalability and cost-effectiveness. Meanwhile, integrating Snowflake with S3 enables the company to perform complex analytics, generate insights, and make data-driven decisions.

Compared to traditional data warehousing solutions, such as on-premises databases or legacy data warehouses, the Snowflake-S3 integration offers several advantages. Traditional solutions often require significant upfront investment in hardware and infrastructure, and they may struggle to handle the volume and variety of data generated in today's digital landscape. On the other hand, the Snowflake-S3 integration provides a scalable and flexible solution that can adapt to changing business needs without the hassle of managing hardware or worrying about capacity.

## **Advantages:**

- **Scalability:** Both Snowflake and S3 are highly scalable services, allowing organizations to store and analyze petabytes of data without worrying about infrastructure constraints.
- **Cost-Effectiveness:** AWS S3 follows a pay-as-you-go pricing model, ensuring that organizations only pay for the storage and resources they use. Snowflake also offers a similar pricing model, making it cost-effective for organizations of all sizes.
- **Flexibility:** The integration allows users to store data in various formats, such as CSV, JSON, or Parquet, in S3 buckets, giving flexibility in data storage. Snowflake can seamlessly query and analyze these data formats, providing flexibility in data analysis.
- **Security:** AWS IAM (Identity and Access Management) ensures secure access to S3 buckets, and Snowflake provides robust security features, including encryption and access controls, ensuring data security and compliance.

## **Disadvantages:**

- **Learning Curve:** Setting up and configuring the integration between Snowflake and S3 may require some expertise in AWS services and SQL queries, which could pose a challenge for organizations new to cloud data warehousing.
- **Complexity:** Managing permissions, policies, and configurations for both Snowflake and AWS S3 can become complex, especially in large-scale deployments, requiring careful planning and administration.

## **Section 2 – Requirements**

### **Requirements**

- S3 Bucket, AWS IAM Services, Integration, External Stage object
- S3 Bucket (Estimated Time: 5 minutes)
  - i) Estimated Cost: Minimal (depends on storage usage)
  - ii) Upload a file (e.g., CSV, Excel) to your S3 bucket.
  - iii) Note down the file path for later use.
- AWS IAM Services (Estimated Time: 10 minutes)
  - i) Estimated Cost: Minimal (depends on IAM usage)
  - ii) Create IAM policies for permissions.
  - iii) Upload a predefined policy file granting necessary permissions (e.g., PutObject, GetObject).
  - iv) Specify the bucket's file path as a resource.
  - v) Review and create the policy.\
- Create an AWS role:
  - i) Select a trusted entity (AWS account).
  - ii) Allow access to third-party services (Snowflake).
  - iii) Generate a unique ID.
  - iv) Attach the previously created policy.
  - v) Review and create the role.
- Copy the ARN path for communication with Snowflake.
- Integration (Estimated Time: 15 minutes)
- Estimated Cost: Minimal (depends on Snowflake usage)
- Perform Snowflake commands for integration:
  - i) Set storage provider as S3, type as external stage.
  - ii) Specify storage\_aws\_role\_arn and storage\_allowed\_locations as important parameters.
  - iii) After running the query, go to roles and edit trust relationships.
  - iv) Click on trust policy, provide ARN and external ID.
  - v) Successfully authenticate.

### External Stage Object (Estimated Time: 5 minutes)

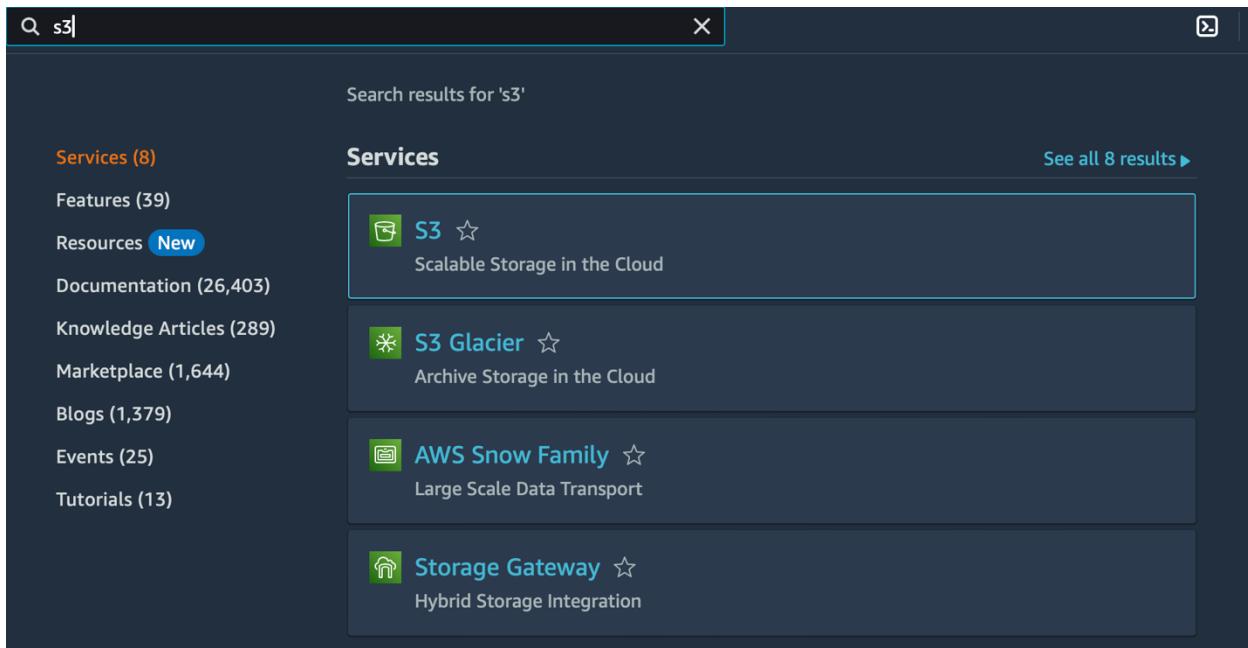
- Estimated Cost: Minimal (depends on Snowflake usage)
- Create a file format to read the file.
- Create a stage ensuring proper integration.
- Describe the details of the storage integration as aws\_s3\_integration.
- Specify the default CSV file extension.
- Provide the URL of the bucket you'll access.

#### **Cost Estimate:**

The cost associated with this tutorial is minimal, primarily dependent on usage. AWS charges for S3 storage and IAM usage, while Snowflake charges for storage and compute usage. However, for the purposes of this tutorial, the cost is likely to be negligible as it involves setting up basic configurations and may fall within free tier usage for AWS services.

## Section 3 – Steps to follow.

### Step 0: Open your AWS account and search for S3.



The screenshot shows the AWS search interface with the query 's3' entered in the search bar. The results are filtered under the 'Services' category, which contains 8 items. The top result is 'S3' (Scalable Storage in the Cloud), followed by 'S3 Glacier' (Archive Storage in the Cloud), 'AWS Snow Family' (Large Scale Data Transport), and 'Storage Gateway' (Hybrid Storage Integration). A link to 'See all 8 results' is visible at the top right of the search results.

Search results for 's3'

Services (8)

Features (39)

Resources New

Documentation (26,403)

Knowledge Articles (289)

Marketplace (1,644)

Blogs (1,379)

Events (25)

Tutorials (13)

Services

S3 Scalable Storage in the Cloud

S3 Glacier Archive Storage in the Cloud

AWS Snow Family Large Scale Data Transport

Storage Gateway Hybrid Storage Integration

See all 8 results ►

## Step 1: Create a bucket, keep the other things as default and save.

The screenshot shows the AWS S3 console under the 'General purpose buckets' tab. It lists three buckets:

Name	AWS Region	IAM Access Analyzer	Creation date
aws-logs-885168717860-us-east-2	US East (Ohio) us-east-2	<a href="#">View analyzer for us-east-2</a>	February 18, 2024, 14:44:14 (UTC-05:00)
msis685-aneri	US East (Ohio) us-east-2	<a href="#">View analyzer for us-east-2</a>	February 18, 2024, 09:10:54 (UTC-05:00)
tutorial1233	US East (Ohio) us-east-2	<a href="#">View analyzer for us-east-2</a>	April 19, 2024, 15:18:06 (UTC-04:00)

The screenshot shows the 'Create bucket' configuration page. The 'General configuration' section is active. The 'AWS Region' is set to 'US East (Ohio) us-east-2'. The 'Bucket name' field contains 'bigdatatutorial'. A note below the field states: 'Bucket name must be unique within the global namespace and follow the bucket naming rules. [See rules for bucket naming](#)'.

The 'Copy settings from existing bucket - optional' section indicates that only the bucket settings in the following configuration are copied. A 'Choose bucket' button is available. The format for the prefix is shown as 'Format: s3://bucket/prefix'.

## Step 2: Add your CSV file to the bucket you created.

The screenshot shows the AWS S3 console interface for the bucket 'bigdatatutorial1'. The 'Objects' tab is selected. At the top, there are buttons for 'Copy S3 URI', 'Copy URL', 'Download', 'Open', 'Delete', 'Actions', 'Create folder', and a prominent orange 'Upload' button. Below these buttons is a dashed blue box with the text 'Drag and drop files and folders you want to upload here, or choose Add files or Add folder.' A table titled 'Files and folders (1 Total, 1.1 MB)' lists one file: 'people-10000.csv' (text/csv). There are 'Remove', 'Add files', and 'Add folder' buttons above the table. A search bar with placeholder 'Find by name' and a page navigation bar with '1' are also visible. In the 'Destination' section, the URL 's3://bigdatatutorial1' is specified. Under 'Destination details', it says 'Bucket settings that impact new objects stored in the specified destination.' Below that are sections for 'Permissions' (Grant public access and access to other AWS accounts) and 'Properties' (Specify storage class, encryption settings, tags, and more). At the bottom right are 'Cancel' and 'Upload' buttons, with 'Upload' being orange.

bigdatatutorial1 [Info](#)

Objects Properties Permissions Metrics Management Access Points

Objects (0) [Info](#)     [Open](#) [Delete](#) Actions Create folder

Drag and drop files and folders you want to upload here, or choose Add files or Add folder.

Files and folders (1 Total, 1.1 MB)

All files and folders in this table will be uploaded.

Remove Add files Add folder

Find by name < 1 >

<input type="checkbox"/>	Name	Folder	Type
<input type="checkbox"/>	people-10000.csv	-	text/csv

Destination [Info](#)

Destination  
<s3://bigdatatutorial1>

▶ Destination details  
Bucket settings that impact new objects stored in the specified destination.

▶ Permissions  
Grant public access and access to other AWS accounts.

▶ Properties  
Specify storage class, encryption settings, tags, and more.

Cancel

### **Step 3: Search IAM in the search bar.**

Search results for 'iam'

Services (11)      Services      See all 11 results ▶

Features (24)

Resources New

Documentation (58,143)

 IAM ☆ Manage access to AWS resources

**Step 4: On the left hand side, search for policies and select create policy.**

Identity and Access Management (IAM) X

Search IAM

IAM > Policies

**Policies (1198)** [Info](#)

A policy is an object in AWS that defines permissions.

Filter by Type [Create policy](#)

Policy name	Type	Used as	Description
<a href="#">AccessAnalyzerSer...</a>	AWS managed	None	Allow Access Analyzer to analyze resou...
<a href="#">AdministratorAccess</a>	AWS managed - job function	None	Provides full access to AWS services an...
<a href="#">AdministratorAcc...</a>	AWS managed	None	Grants account administrative permis...
<a href="#">AdministratorAcc...</a>	AWS managed	None	Grants account administrative permis...
<a href="#">AlexaForBusinessD...</a>	AWS managed	None	Provide device setup access to AlexaFo...
<a href="#">AlexaForBusinessF...</a>	AWS managed	None	Grants full access to AlexaForBusiness ...
<a href="#">AlexaForBusinessG...</a>	AWS managed	None	Provide gateway execution access to A...
<a href="#">AlexaForBusinessLi...</a>	AWS managed	None	Provide access to Lifesize AVS devices
<a href="#">AlexaForBusinessN...</a>	AWS managed	None	This policy enables Alexa for Business ...
<a href="#">AlexaForBusinessP...</a>	AWS managed	None	Provide access to Poly AVS devices
<a href="#">AlexaForBusinessR...</a>	AWS managed	None	Provide read only access to AlexaForB...
<a href="#">AmazonAPIGatewa...</a>	AWS managed	None	Provides full access to create/edit/dele...

**Step 5: Go to permissions and select JSON, copy and paste the given steps for creating your own policy for the bucket.**

Specify permissions Info
Visual **JSON**
Actions ▾

```

1  {
2      "Version": "2012-10-17",
3      "Statement": [
4          {
5              "Effect": "Allow",
6              "Action": [
7                  "s3:PutObject",
8                  "s3:GetObject",
9                  "s3:GetObjectVersion",
10                 "s3:DeleteObject",
11                 "s3:DeleteObjectVersion"
12             ],
13             "Resource": "arn:aws:s3:::bigdatatutorialall/*"
14         },
15         {
16             "Effect": "Allow",
17             "Action": [
18                 "s3>ListBucket",
19                 "s3:GetBucketLocation"
20             ],
21             "Resource": "arn:aws:s3:::bigdatatutorialall",
22             "Condition": {
23                 "StringLike": {
24                     "s3:prefix": [
25                         "*"
26                     ]
27                 }
28             }
29         }
30     ]
31 }
```

Edit statement
Remove

Add actions

Choose a service

Included

Available

Add a resource

+ Add new statement
Add

Available
5777 of 6144 characters remaining

Security: 0 Errors: 0 Warnings: 0 Suggestions: 0
Cancel
Next

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": [  
        "s3:PutObject",  
        "s3:GetObject",  
        "s3:GetObjectVersion",  
        "s3:DeleteObject",  
        "s3:DeleteObjectVersion"  
      ],  
      "Resource": "arn:aws:s3:::bigdatatutorial/*"  
    },  
    {  
      "Effect": "Allow",  
      "Action": [  
        "s3>ListBucket",  
        "s3:GetBucketLocation"  
      ],  
      "Resource": "arn:aws:s3:::bigdatatutorial",  
      "Condition": {  
        "StringLike": {  
          "s3:prefix": [  
            "*"  
          ]  
        }  
      }  
    }  
  ]  
}
```

## Step 6: Give that policy a name and scroll down to select Create Policy.

Policy name  
Enter a meaningful name to identify this policy.

Maximum 128 characters. Use alphanumeric and '+,-,\_,@-' characters.

Description - *optional*  
Add a short explanation for this policy.

Maximum 1,000 characters. Use alphanumeric and '+,-,\_,@-' characters.

Info This policy defines some actions, resources, or conditions that do not provide permissions. To grant access, policies must have an action that has an applicable resource or condition. For details, choose [Show remaining](#). [Learn more](#)

**Permissions defined in this policy** Info [Edit](#)  
Permissions defined in this policy document specify which actions are allowed or denied. To define permissions for an IAM identity (user, user group, or role), attach a policy to it.

**Allow (1 of 408 services)**  Show remaining 407 services

Service	▲	Access level	▼	Resource	Request condition
S3		Limited: Read, List, Write		Multiple	s3:prefix  string like  All

**Add tags - *optional*** Info  
Tags are key-value pairs that you can add to AWS resources to help identify, organize, or search for resources.

## Step 7: Go back to the main IAM menu and follow the step given below to create role.

The screenshot shows the AWS IAM Roles page with 16 roles listed. The top navigation bar includes 'Roles (16)', 'Info', 'Delete', and 'Create role'. A search bar and pagination controls are also present. The 'Trusted entity type' section is highlighted, showing five options:

- AWS service: Allow AWS services like EC2, Lambda, or others to perform actions in this account.
- AWS account: Allow entities in other AWS accounts belonging to you or a 3rd party to perform actions in this account.
- Web identity: Allows users federated by the specified external web identity provider to assume this role to perform actions in this account.
- SAML 2.0 federation: Allow users federated with SAML 2.0 from a corporate directory to perform actions in this account.
- Custom trust policy: Create a custom trust policy to enable others to perform actions in this account.

Below this, a detailed configuration for an AWS account is shown:

**An AWS account**  
Allow entities in other AWS accounts belonging to you or a 3rd party to perform actions in this account.

This account (885168717860)  
 Another AWS account

**Options**

**Require external ID (Best practice when a third party will assume this role)**  
You can increase the security of your role by requiring an optional external identifier, which prevents "confused deputy" attacks. This is recommended if you do not own or have administrative access to the account that can assume this role. The external ID can include any characters that you choose. To assume this role, users must be in the trusted account and provide this exact external ID. [Learn more](#)

**External ID**  
885168717860

## Step 8: Add your predefined policy that you created in step 6.

Add permissions [Info](#)

Permissions policies (933) [Info](#)

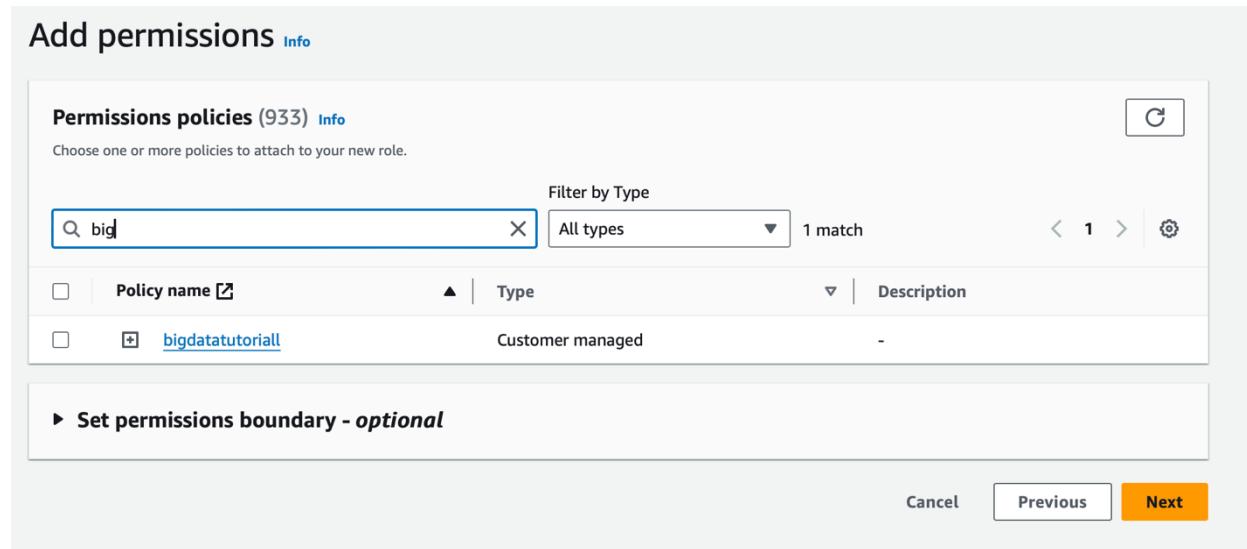
Choose one or more policies to attach to your new role.

Filter by Type

Policy name	Type	Description
<input type="checkbox"/> <a href="#">bigdatatutorial</a>	Customer managed	-

▶ Set permissions boundary - *optional*

[Cancel](#) [Previous](#) [Next](#)



## Step 9: Give a name to the role, scroll down, and select create role.

Name, review, and create

Role details

Role name

Enter a meaningful name to identify this role.

Maximum 64 characters. Use alphanumeric and '+=\_,@-' characters.

Description

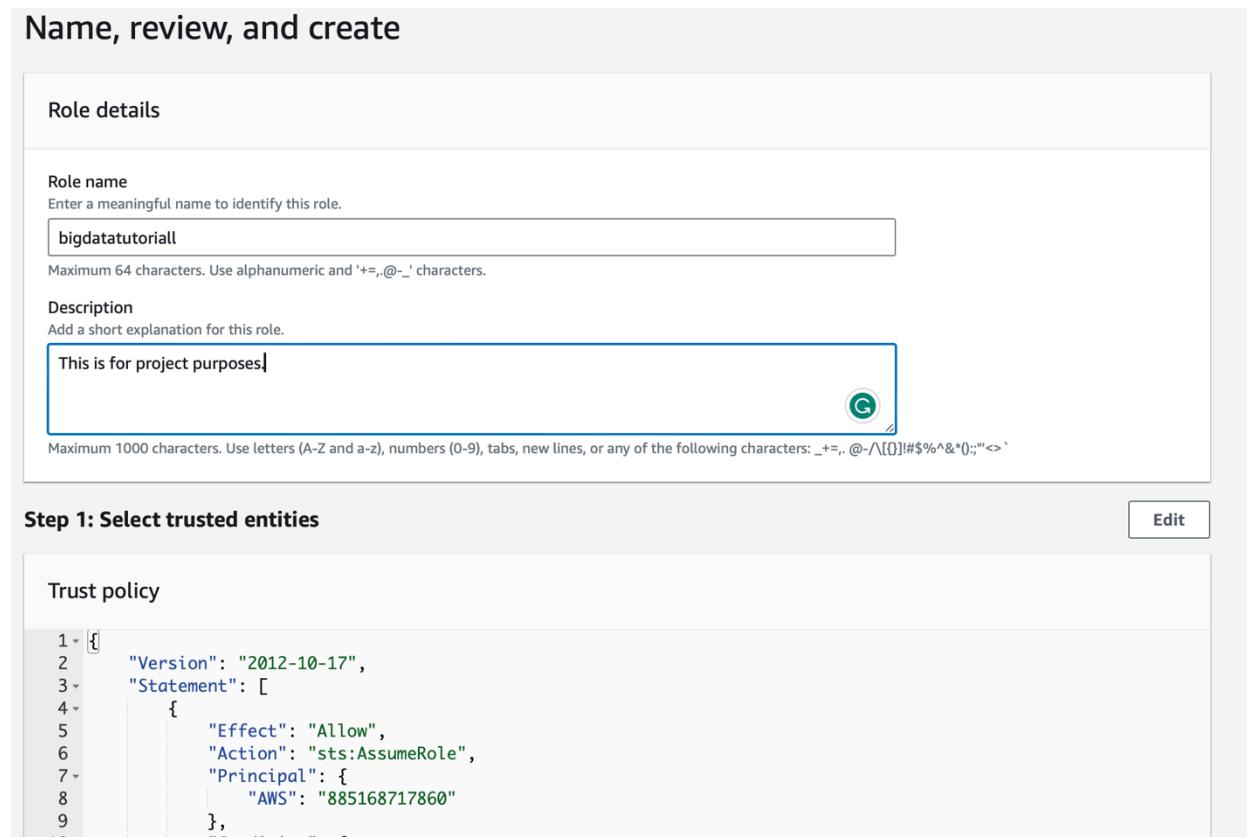
Add a short explanation for this role.

Maximum 1000 characters. Use letters (A-Z and a-z), numbers (0-9), tabs, new lines, or any of the following characters: \_+=., @-/\[\]!#\$%^&\*();:"'<>`

Step 1: Select trusted entities [Edit](#)

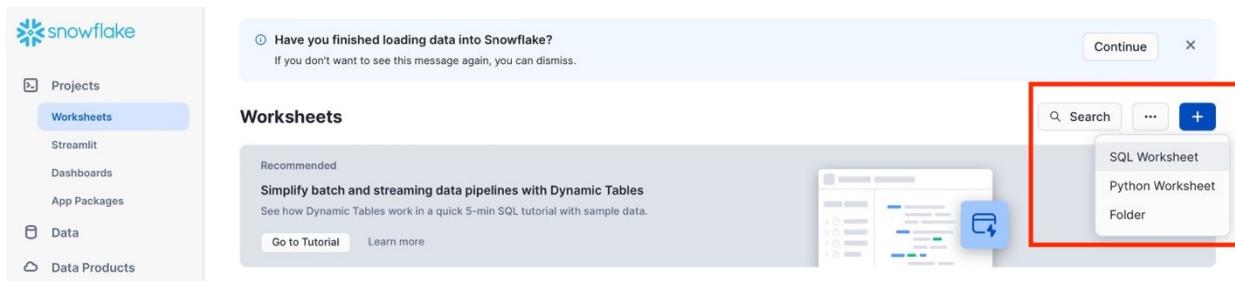
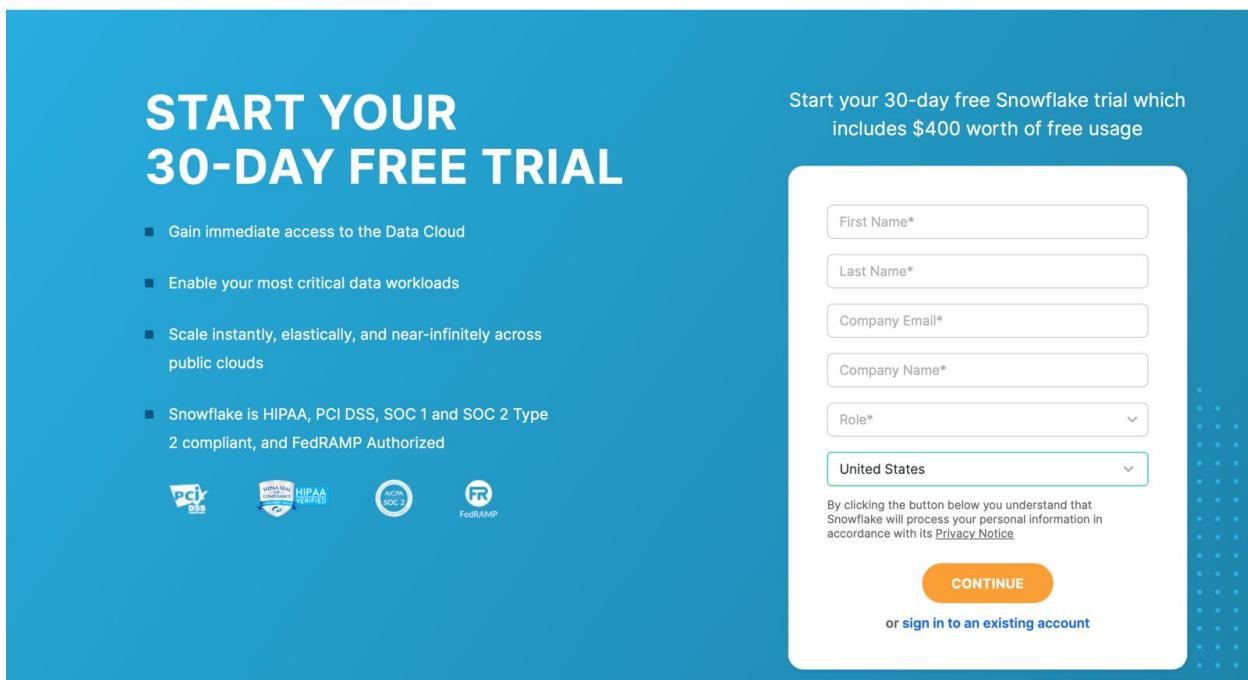
Trust policy

```
1 [ {  
2   "Version": "2012-10-17",  
3   "Statement": [  
4     {  
5       "Effect": "Allow",  
6       "Action": "sts:AssumeRole",  
7       "Principal": {  
8         "AWS": "885168717860"  
9       },  
10      "Condition": {}  
11    }  
12  ]  
13}  
14
```



**Step 10: Login to your snowflake account, create a new sheet and copy and paste the SQL code to create database and schema.**

<https://www.snowflake.com/en/>



[CREATE DATABASE PROJECT;](#)  
[CREATE SCHEMA PROJECT1233;](#)

[CREATE DATABASE PROJECT;](#)  
[CREATE SCHEMA PROJECT1233;](#)

**Step 11: Go back to AWS and click on IAM -> Roles -> ‘your role-name’-> Summary-> Copy the ARN**

IAM > Roles

**Roles (17) 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.

Search: big

Role name Trusted entities Last activity

Role name	Trusted entities	Last activity
<a href="#">bigdatatutorial1</a>	Account: 891377318427	42 minutes ago

**bigdatatutorial1** [Info](#)

This is for project purposes.

**Summary** [Edit](#)

Creation date	ARN	Link to switch roles in console
April 21, 2024, 18:26 (UTC-04:00)	<a href="#">arn:aws:iam::885168717860:role/bigdatatutorial1</a>	<a href="https://signin.aws.amazon.com/switchrole?roleName=bigdatatutorial1&amp;account=885168717860">https://signin.aws.amazon.com/switchrole?roleName=bigdatatutorial1&amp;account=885168717860</a>
Last activity	Maximum session duration	
<a href="#">42 minutes ago</a>	1 hour	

**Step 11.1: For the location for bucket, go to S3-> Buckets-> ‘your bucket name’ -> ‘your uploaded csv file’ -> and copy the S3 URI.**

Properties Permissions Versions

**Object overview**

Owner	S3 URI
dfcb3f6b216c674d516120e7944d34ca3c8fcbead9d1d98bbf1405f643976561	<a href="#">s3://bigdatatutorial1/people-10000.csv</a>
AWS Region	Amazon Resource Name (ARN)
US East (Ohio) us-east-2	<a href="#">arn:aws:s3:::bigdatatutorial1/people-10000.csv</a>
Last modified	Entity tag (Etag)
April 21, 2024, 20:53:37 (UTC-04:00)	<a href="#">e6b23a66152c0d50958bcf17259fd6aa</a>
Size	Object URL
1.1 MB	<a href="https://bigdatatutorial1.s3.us-east-2.amazonaws.com/people-10000.csv">https://bigdatatutorial1.s3.us-east-2.amazonaws.com/people-10000.csv</a>
Type	
csv	
Key	
<a href="#">people-10000.csv</a>	

**Step 12: Copy and paste the following code and in the space for storage\_aws\_role\_arn paste the ARN and S3 URI in storage\_allowed\_location without the “/filename.csv” you copied in Step 11 and Step 11.1.**

```
4 CREATE OR REPLACE STORAGE INTEGRATION aws_s3_integration
5   type = external_stage
6   storage_provider='S3'
7   enabled= true
8   storage_aws_role_arn='arn:aws:iam::885168717860:role/bigdatatutorial1'
9   storage_allowed_locations = ('s3://bigdatatutorial1');
```

status
1 Integration AWS_S3_INTEGRATION successfully created.

Query Details

Query duration

Rows

Query ID 91b3d38c-0001-

```
CREATE OR REPLACE STORAGE INTEGRATION aws_s3_integration
type = external_stage
storage_provider='S3'
enabled= true
storage_aws_role_arn='arn:aws:iam::885168717860:role/bigdatatutorial1'
storage_allowed_locations = ('s3://bigdatatutorial1');
```

**bigdatatutorial1** Info

This is for project purposes.

**Summary** Edit

Creation date April 21, 2024, 18:26 (UTC-04:00)	ARN <a href="#">arn:aws:iam::885168717860:role/bigdatatutorial1</a>
Last activity <a href="#">42 minutes ago</a>	Maximum session duration 1 hour

[Link to switch roles in console](#)  
<https://signin.aws.amazon.com/switchrole?roleName=bigdatatutorial1&account=885168717860>

**Properties** Permissions Versions

**Object overview**

Owner dfcb3f6b216c674d516120e7944d34ca3c8fcbead9d1d98bbf1405f643976561	S3 URI <a href="#">s3://bigdatatutorial1/people-10000.csv</a>
AWS Region US East (Ohio) us-east-2	Amazon Resource Name (ARN) <a href="#">arn:aws:s3:::bigdatatutorial1/people-10000.csv</a>
Last modified April 21, 2024, 20:53:57 (UTC-04:00)	Entity tag (Etag) <a href="#">e6b23a66152c0d50958bcf17259fd6aa</a>
Size 1.1 MB	Object URL <a href="https://bigdatatutorial1.s3.us-east-2.amazonaws.com/people-10000.csv">https://bigdatatutorial1.s3.us-east-2.amazonaws.com/people-10000.csv</a>
Type csv	
Key <a href="#">people-10000.csv</a>	

## Step 13: Copy and paste the following steps to show and desc integration.

PROJECT.PROJECT1233 ▾ Settings ▾

Code Versions 

```
1 CREATE DATABASE PROJECT;
2 CREATE SCHEMA PROJECT1233;
3
4 CREATE OR REPLACE STORAGE INTEGRATION aws_s3_integration
5   type = external_stage
6   storage_provider='S3'
7   enabled= true
8   storage_aws_role_arn='arn:aws:iam::885168717860:role/bigdatatutorialall'
9   storage_allowed_locations = ('s3://bigdatatutorialall');
10
11 SHOW INTEGRATIONS;
12
13 DESC integration aws_s3_integration;
```

Results  PREVIEW     

	name	type	category	enabled	comment
1	AWS_S3_INTEGRATION	EXTERNAL_STAGE	STORAGE	true	null
2	SNOWSERVICES_INGRESS_OAUTH	OAUTH - SNOWSERVICES_INGRESS	SECURITY	true	null

Query Details   
Query duration 52ms  
Rows 2  
Query ID [01b3d3d2-0001-cc1d-0...](#)

PROJECT.PROJECT1233 ▾    Settings ▾    ACCOUNTADMIN . COMPUTE\_WH Share ▶ ▾

Code Versions    Q

```

1 CREATE DATABASE PROJECT;
2 CREATE SCHEMA PROJECT1233;
3
4 CREATE OR REPLACE STORAGE INTEGRATION aws_s3_integration
5   type = external_stage
6   storage_provider='S3'
7   enabled= true
8   storage_aws_role_arn='arn:aws:iam::885168717860:role/bigdatatutorialall'
9   storage_allowed_locations = ('s3://bigdatatutorialall');
10
11 SHOW INTEGRATIONS;
12
13 DESC integration aws_s3_integration;
```

👉 🌙

**Results**    **Chart**

property	property_type	property_value	proper
1 ENABLED	Boolean	true	false
2 STORAGE_PROVIDER	String	S3	
3 STORAGE_ALLOWED_LOCATIONS	List	s3://bigdatatutorialall	□
4 STORAGE_BLOCKED_LOCATIONS	List		□
5 STORAGE_AWS_IAM_USER_ARN	String	arn:aws:iam::891377318427:user/lr4m0000-s	
6 STORAGE_AWS_ROLE_ARN	String	arn:aws:iam::885168717860:role/bigdatatutorialall	
7 STORAGE_AWS_EXTERNAL_ID	String	LQ97468_SFCRole=2_HxR1udsjgwYKs+iTOVlNoO-	
8 COMMENT	String		

**Query Details** ...
   
Query duration 47ms
   
Rows 8
   
Query ID 01b3d3d3-0001-cc3f-0...
   
  
**property** A
   
100% filled
   
  
**property\_type** A
   
String 5

SHOW INTEGRATIONS;

DESC integration aws\_s3\_integration;

**Step 14: Copy the storage\_aws\_iam\_user\_arn and storage\_aws\_external\_id and paste it somewhere, as we will use it while editing trust policies.**

The screenshot shows the AWS Database Migration Service (DMS) console. At the top, there are tabs for 'ACCOUNTADMIN' and 'COMPUTE\_WH', and buttons for 'Share', 'Code Versions', and a search icon. Below the tabs, the project name 'PROJECT.PROJECT1233' and a 'Settings' dropdown are visible. On the right, there are icons for a yellow hand, a green checkmark, and a blue circular arrow.

The main area displays a SQL query result:

```
1 CREATE DATABASE PROJECT;
2 CREATE SCHEMA PROJECT1233;
3
4 CREATE OR REPLACE STORAGE INTEGRATION aws_s3_integration
5   type = external_stage
6   storage_provider='S3'
7   enabled= true
8   storage_aws_role_arn='arn:aws:iam::885168717860:role/bigdatatutorialall'
9   storage_allowed_locations = ('s3://bigdatatutorialall');
10
11 SHOW INTEGRATIONS;
12
13 DESC integration aws_s3_integration;
```

The line 'storage\_aws\_role\_arn='arn:aws:iam::885168717860:role/bigdatatutorialall'' is highlighted with a red box.

Below the query result, there are two tabs: 'Results' (selected) and 'Chart'. The 'Results' tab shows a table of properties for the integration:

property	property_type	property_value	proper...
1 ENABLED	Boolean	true	false
2 STORAGE_PROVIDER	String	S3	
3 STORAGE_ALLOWED_LOCATIONS	List	s3://bigdatatutorialall	□
4 STORAGE_BLOCKED_LOCATIONS	List		□
5 STORAGE_AWS_IAM_USER_ARN	String	arn:aws:iam::891377318427:user/lr4m0000-s	
6 STORAGE_AWS_ROLE_ARN	String	arn:aws:iam::885168717860:role/bigdatatutorialall	
7 STORAGE_AWS_EXTERNAL_ID	String	LQ97468_SFCRole=2_HxR1udsjgwYKs+iTOVINoO-	
8 COMMENT	String		

On the right side of the results table, there is a 'Query Details' panel with the following information:

- Query duration: 47ms
- Rows: 8
- Query ID: 01b3d3d3-0001-cc3f-0...

Below the table, there are two more panels:

- property: 100% filled
- property\_type: String

**Step 15: Go to IAM-> Roles-> ‘your role name’ -> scroll down and search Trust Relationships-> Edit trust policy -> Paste storage\_aws\_iam\_user\_arn and storage\_aws\_external\_id that you copied in step 14 in AWS and sts.ExternalId.**

The screenshot shows the IAM Roles page with the role 'bigdatatutorial1' selected. The 'Summary' tab is active, displaying basic information like creation date (April 21, 2024), ARN, and last activity (55 minutes ago). Below the summary, there are tabs for 'Permissions', 'Trust relationships', 'Tags', 'Access Advisor', and 'Revoke sessions'. The 'Permissions' tab is selected, showing one managed policy named 'bigdatatutorial1'. The policy details are shown in a modal window, including its ARN and a link to switch roles in the console.

**bigdatatutorial1** Info

This is for project purposes.

**Summary**

Creation date  
April 21, 2024, 18:26 (UTC-04:00)

ARN  
arn:aws:iam::885168717860:role/bigdatatutorial1

Last activity  
55 minutes ago

Maximum session duration  
1 hour

**Permissions** **Trust relationships** **Tags** **Access Advisor** **Revoke sessions**

**Permissions policies (1) Info**

You can attach up to 10 managed policies.

Filter by Type

Policy name	Type	Attached entities
<a href="#">bigdatatutorial1</a>	Customer managed	1

Creation date	ARN	Link to switch roles in console
April 21, 2024, 18:26 (UTC-04:00)	<a href="#">arn:aws:iam::885168717860:role/bigdatatutorialall</a>	<a href="https://signin.aws.amazon.com/switchrole?roleName=bigdatatutorial&amp;account=885168717860">https://signin.aws.amazon.com/switchrole?roleName=bigdatatutorial&amp;account=885168717860</a>
Last activity	Maximum session duration	
-	1 hour	

Permissions    **Trust relationships**    Tags    Access Advisor    Revoke sessions

**Trusted entities**

Entities that can assume this role under specified conditions.

[Edit trust policy](#)

```
1- [
2   "Version": "2012-10-17",
3-   "Statement": [
4-     {
5       "Effect": "Allow",
6       "Principal": {
7         "AWS": "arn:aws:iam::885168717860:root"
8       },
9       "Action": "sts:AssumeRole",
10      "Condition": {
11        "StringEquals": {
12          "sts:ExternalId": "885168717860"
13        }
14      }
15    }
16  ]
17 ]
```

## Edit trust policy

```
1▼ {
2   "Version": "2012-10-17",
3▼   "Statement": [
4▼     {
5       "Effect": "Allow",
6       "Principal": {
7         "AWS": "arn:aws:iam::891377318427:user/1r4m0000-s"
8       },
9       "Action": "sts:AssumeRole",
10      "Condition": {
11        "StringEquals": {
12          "sts:ExternalId": "LQ97468_SFCRole=2_HxR1udsjgwYKs+iTOVINO0+EnIE="
13        }
14      }
15    }
16  ]
17 }
```

**Step 16: Write the code to grant aws to role accountadmin and click on Run.**

```
15 | GRANT usage on integration aws_s3_integration to role accountadmin;
```

↳ Results    ↗ Chart

status	
1	Statement executed successfully.

GRANT usage on integration aws\_s3\_integration to role accountadmin;

**Step 17: Create or replace file format and stage codes, use the S3 url that you copied in step 11.1, copy the code below and then change as mentioned.**

```
17 | create or replace file format demo_format
18 | type='CSV'
19 | field_delimiter='|'
20 | skip_header=1;
21 |
22 | CREATE OR REPLACE stage demo_aws_stage
23 |   storage integration = aws_s3_integration
24 |   url = 's3://bigdatatutorial1/';
```

↳ Results    ↗ Chart

status	
1	Stage area DEMO_AWS_STAGE successfully created.

create or replace file format demo\_format

type='CSV'

field\_delimiter='|'

skip\_header=1;

CREATE OR REPLACE stage demo\_aws\_stage

storage\_integration = aws\_s3\_integration

url = 's3://bigdatatutorial1/';

**Step 18: Copy and paste codes to list and to remove the csv file as shown below.**

```
26 | List @demo_aws_stage;  
27 | remove @demo_aws_stage/people-10000.csv;
```

	name	size	md5	last_modified
1	s3://bigdatatutorial1/people-10000.csv	1117138	e6b23a66152c0d50958bcf17259fd6aa	Sun, 21 Apr 2024

```
27 | remove @demo_aws_stage/people-10000.csv;
```

	name	result
1	s3://bigdatatutorial1/people-10000.csv	removed

List @demo\_aws\_stage;

remove @demo\_aws\_stage/people-10000.csv;

**Step 19:** After following step 18, you will see that the csv file you uploaded in your bucket is no more, so you will have to upload it again manually.

The screenshot shows the AWS S3 console interface for a bucket named "bigdatatutorial1". The top navigation bar includes tabs for Objects, Properties, Permissions, Metrics, Management, and Access Points. The "Objects" tab is selected, showing 0 objects. Below the tabs is a toolbar with actions: Copy S3 URI, Copy URL, Download, Open, Delete, Actions, and Create folder. A prominent orange "Upload" button is highlighted. A message below the toolbar states: "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](#)". A search bar labeled "Find objects by prefix" is present. A table header for object listing includes columns for Name, Type, Last modified, Size, and Storage class. The main content area displays the message "No objects" and "You don't have any objects in this bucket.", with another "Upload" button at the bottom.

**Step 20: After uploading copy and paste the following codes to create a table and list it.**

```
29  create or replace temporary table demo_customer_info (
30    user_id string,
31    first_name string,
32    last_name string
33  );
34
```

↳ Results    ↵ Chart    PREVIEW         

status	
1	Table DEMO_CUSTOMER_INFO successfully created

Query Details

Query duration

Rows

```
35  Select * from demo_customer_info limit 10;
36
37  COPY INTO demo_customer_info
38  FROM @demo_aws_stage/people-10000.csv
39  FILE_FORMAT=(format_name=demo_format)
40  ON_ERROR='CONTINUE'; -- or ON_ERROR='SKIP_FILE' if you want to skip the entire file
41
42  CREATE OR REPLACE FILE FORMAT demo_format
43  TYPE = 'CSV'
44  FIELD_OPTIONALLY_ENCLOSED_BY = ''
45  SKIP_HEADER = 1
46  ERROR_ON_COLUMN_COUNT_MISMATCH = FALSE; -- Add this line to ignore the column count mismatch
```

↳ Results    ↵ Chart

file	status	rows_parsed	rows_loaded	error_limit	errors_seen	first_error
s3://bigdatatutorial1/people-10000.csv	LOAD_FAILED	10000	0	10000	10000	Number of columns in file (1) does not n

```
create or replace temporary table demo_customer_info (
  user_id string,
  first_name string,
  last_name string
);
```

```
Select * from demo_customer_info limit 10;
```

```
COPY INTO demo_customer_info
FROM @demo_aws_stage/people-10000.csv
FILE_FORMAT=(format_name=demo_format)
ON_ERROR='CONTINUE'; -- or ON_ERROR='SKIP_FILE' if you want to skip the entire file
```

```
CREATE OR REPLACE FILE FORMAT demo_format
TYPE = 'CSV'
FIELD_OPTIONALLY_ENCLOSED_BY = ""
SKIP_HEADER = 1
ERROR_ON_COLUMN_COUNT_MISMATCH = FALSE; -- Add this line to ignore the
column count mismatch
```

## Step 21: Write the following codes to copy and try different scenarios like copy and count.

```
42 CREATE OR REPLACE FILE FORMAT demo_format
43   TYPE = 'CSV'
44   FIELD_OPTIONALLY_ENCLOSED_BY = '"'
45   SKIP_HEADER = 1
46   ERROR_ON_COLUMN_COUNT_MISMATCH = FALSE; -- Add this line to ignore the column count mismatch
47
48   --Scenario :1
49
50 COPY INTO demo_customer_info
51   from @demo_aws_stage/
52   file_format=(format_name=demo_format)
53   on_error='Skip_file'; --Skip the whole file
54
```

↳ Results ↗ Chart

status	
1	File format DEMO_FORMAT successfully created.

```
50 COPY INTO demo_customer_info
51   from @demo_aws_stage/
52   file_format=(format_name=demo_format)
53   on_error='Skip_file'; --Skip the whole file
54
55 COPY INTO demo_customer_info
56   from @demo_aws_stage/
57   file_format=(format_name=demo_format)
58   on_error='Continue'; --Skip only the bad record and load the rest of the record
59
60 COPY INTO demo_customer_info
61   FROM @demo_aws_stage/
62   FILE_FORMAT=(format_name=demo_format)
```

↳ Results ↗ Chart

	file	status	rows_parsed	rows_loaded	error_limit	errors_seen	first_error	first_error_line	first_error_col
1	s3://bigdatatutorial1/people-10000.csv	LOADED	10000	10000	1	0	null	null	

```
72 List @demo_aws_stage/people-10000.csv;
73
74   SELECT count(*) from demo_customer_info;
75
76   --Scenario :3
77
78   COPY INTO demo_customer_info
79     from @demo_aws_stage/people-10000.csv
80     file_format=(format_name=demo_format)
81       force=true purge=true;
82
83
84 Validation Mode
```

↳ Results ↗ Chart

	name	size	md5	last_modified
1	s3://bigdatatutorial1/people-10000.csv	1117138	e6b23a66152c0d50958bcf17259fd6aa	Sun, 21 Apr 2024 23:57:39 GMT

```

74 | SELECT count(*) from demo_customer_info;
75 |
76 --Scenario :3
77
78 COPY INTO demo_customer_info
79   from @demo_aws_stage/people-10000.csv
80   file_format=(format_name=demo_format)
81           force=true purge=true;
82
83
84
85

```

↳ Results ⚡ Chart

	COUNT(*)
1	20000

```

78 COPY INTO demo_customer_info
79   from @demo_aws_stage/people-10000.csv
80   file_format=(format_name=demo_format)
81           force=true purge=true;
82
83
84
85

```

↳ Results ⚡ Chart

	file	status	rows_parsed	rows_loaded	error_limit	errors_seen	first_error	first_error_line	first_error_c
1	s3://bigdatatutorial1/people-10000.csv	LOADED	10000	10000	1	0	null		null

--Scenario :1

```

COPY INTO demo_customer_info
from @demo_aws_stage/
file_format=(format_name=demo_format)
on_error='Skip_file'; --Skip the whole file

```

```

COPY INTO demo_customer_info
from @demo_aws_stage/
file_format=(format_name=demo_format)
on_error='Continue'; --Skip only the bad record and load the rest of the record

```

```
COPY INTO demo_customer_info  
FROM @demo_aws_stage/  
FILE_FORMAT=(format_name=demo_format)  
ON_ERROR='ABORT_STATEMENT'; -- Change 'abort' to 'ABORT_STATEMENT'
```

--Scenario :2

```
COPY INTO demo_customer_info  
from @demo_aws_stage/people-10000.csv  
file_format=(format_name=demo_format)  
force=true;
```

```
List @demo_aws_stage/people-10000.csv;
```

```
SELECT count(*) from demo_customer_info;
```

--Scenario :3

```
COPY INTO demo_customer_info  
from @demo_aws_stage/people-10000.csv  
file_format=(format_name=demo_format)  
force=true purge=true;
```

```
Select * from demo_customer_info limit 10;
```

```
COPY INTO demo_customer_info  
FROM @demo_aws_stage/people-10000.csv
```

```
FILE FORMAT=(format_name=demo_format)
ON_ERROR='CONTINUE'; -- or ON_ERROR='SKIP_FILE' if you want to skip the entire file
```

```
CREATE OR REPLACE FILE FORMAT demo_format
```

```
TYPE = 'CSV'
```

```
FIELD_OPTIONALLY_ENCLOSED_BY = ""
```

```
SKIP_HEADER = 1
```

```
ERROR_ON_COLUMN_COUNT_MISMATCH = FALSE; -- Add this line to ignore the
column count mismatch
```