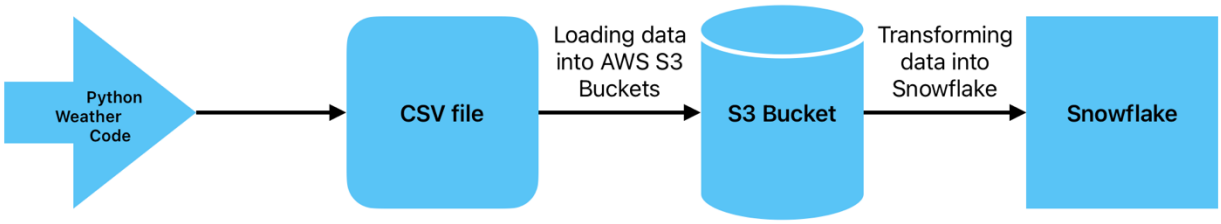
Weather data forecast (Loading data into Snowflake)

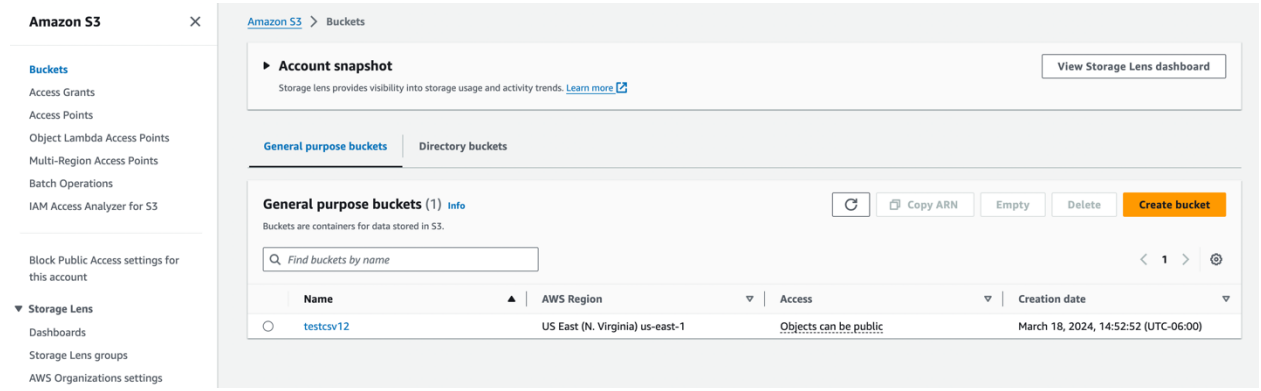
This document outlines the detailed process of transferring a CSV file stored in an Amazon Web Services (AWS) S3 bucket to Snowflake, a cloud-based data warehousing platform. The process involves setting up AWS S3 and IAM roles for access control, configuring Snowflake components such as schemas and stages, and executing the data transfer using Snowflake's capabilities. Once the data is loaded into Snowflake, it can be queried, analyzed, and integrated with other data sources for further analysis and reporting.



1. Creation of S3 Bucket:

→ The process begins with the creation of an S3 bucket on Amazon Web Services (AWS). This bucket serves as the storage location.

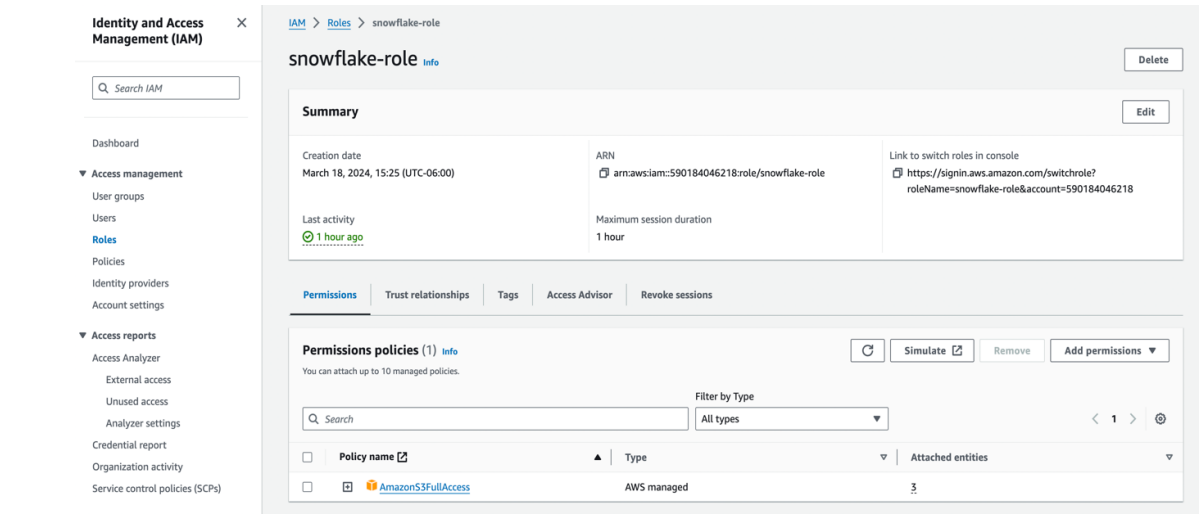
→ I created a bucket named "testcsv12" in AWS S3 to store the csv file. Then, uploaded the "weather_forecast.csv" file into the S3 bucket.

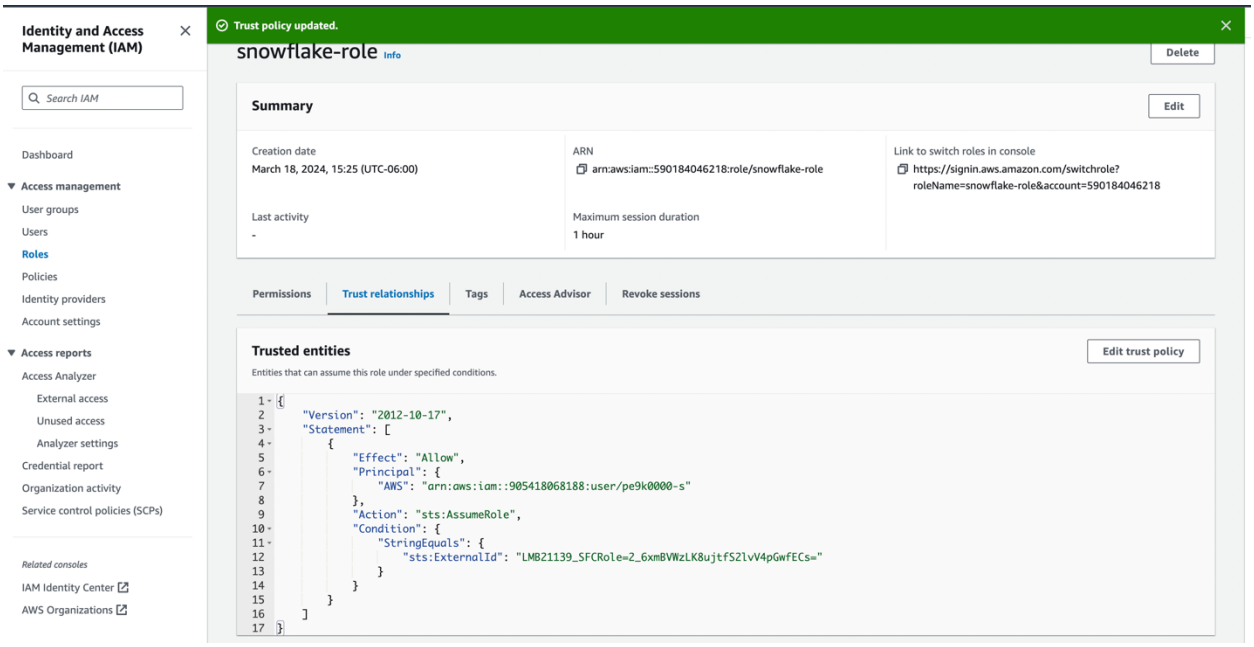


2. IAM Role Configuration: An IAM (Identity and Access Management) role is created with permissions granting access to the S3 bucket. The role is assigned the "AmazonS3FullAccess" policy to provide full access to S3 resources.

→ I created an IAM role named "snowflake-role" in AWS IAM.

→ Assigned a "AmazonS3FullAccess" policy to the "snowflake-role" IAM role, ensuring sufficient permissions to access the S3 bucket.





3. Snowflake Configuration:

Schema Creation: I created a schema named “Weather” to define the structure of the data that will be loaded from the CSV file.

Storage Integration: A stage object “s3_int” is created to establish connectivity between Snowflake and AWS S3. This integration facilitates seamless data transfer between the two platforms, ensuring efficient data ingestion.

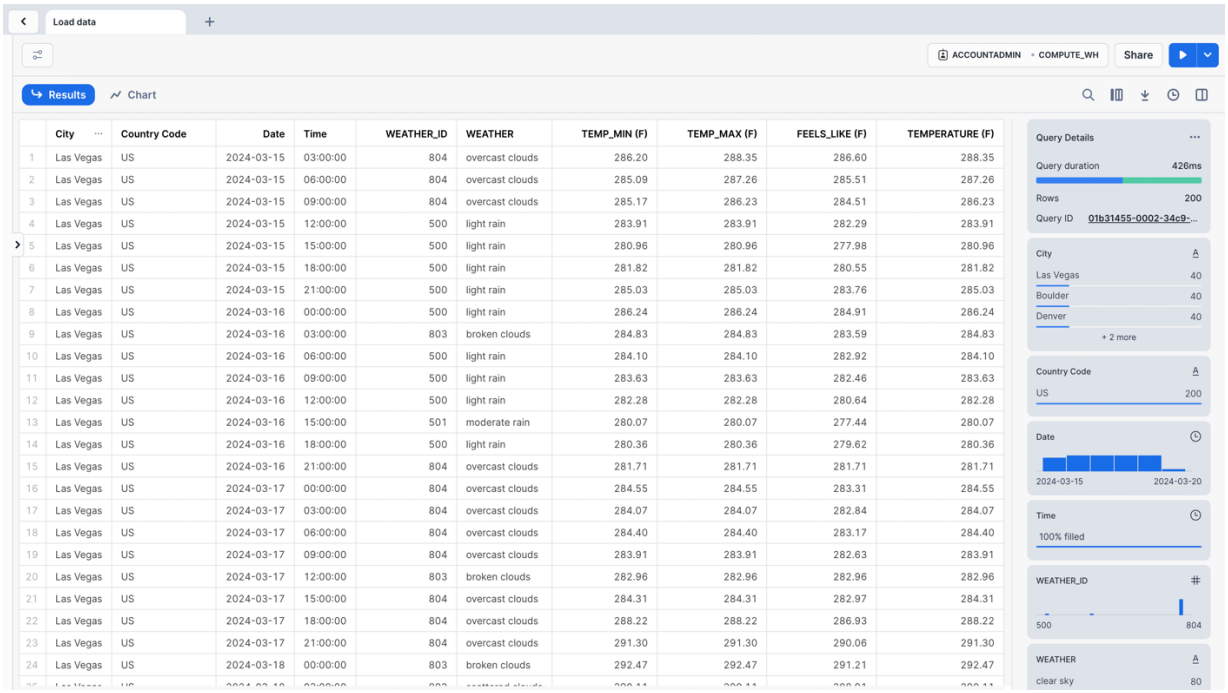
File Format Implementation: A file format object named "csv_format" is created for handling CSV files. This step ensures that the data is parsed correctly during the loading process.

Stage Object Configuration: Another stage object, “ext_stage” is created, locates the CSV files in the AWS S3 bucket. This stage is configured to utilize the previously defined storage integration (“s3_int”) and file format (“csv_format”).

4. Data Transfer Process:

Initiating Data Loading: The “COPY INTO” command is executed to initiate the data loading process. This command specifies the “Weather” table and the source stage “ext_stage”, ensuring that the data is copied from the correct location in AWS S3.

Results:



	City	Country Code	Date	Time	WEATHER_ID	WEATHER	TEMP_MIN (F)	TEMP_MAX (F)	FEELS_LIKE (F)	TEMPERATURE (F)
1	Las Vegas	US	2024-03-15	03:00:00	804	overcast clouds	286.20	288.35	286.60	288.35
2	Las Vegas	US	2024-03-15	06:00:00	804	overcast clouds	285.09	287.26	285.51	287.26
3	Las Vegas	US	2024-03-15	09:00:00	804	overcast clouds	285.17	286.23	284.51	286.23
4	Las Vegas	US	2024-03-15	12:00:00	500	light rain	283.91	283.91	282.29	283.91
5	Las Vegas	US	2024-03-15	15:00:00	500	light rain	280.96	280.96	277.98	280.96
6	Las Vegas	US	2024-03-15	18:00:00	500	light rain	281.82	281.82	280.55	281.82
7	Las Vegas	US	2024-03-15	21:00:00	500	light rain	285.03	285.03	283.76	285.03
8	Las Vegas	US	2024-03-16	00:00:00	500	light rain	286.24	286.24	284.91	286.24
9	Las Vegas	US	2024-03-16	03:00:00	803	broken clouds	284.83	284.83	283.59	284.83
10	Las Vegas	US	2024-03-16	06:00:00	500	light rain	284.10	284.10	282.92	284.10
11	Las Vegas	US	2024-03-16	09:00:00	500	light rain	283.63	283.63	282.46	283.63
12	Las Vegas	US	2024-03-16	12:00:00	500	light rain	282.28	282.28	280.64	282.28
13	Las Vegas	US	2024-03-16	15:00:00	501	moderate rain	280.07	280.07	277.44	280.07
14	Las Vegas	US	2024-03-16	18:00:00	500	light rain	280.36	280.36	279.62	280.36
15	Las Vegas	US	2024-03-16	21:00:00	804	overcast clouds	281.71	281.71	281.71	281.71
16	Las Vegas	US	2024-03-17	00:00:00	804	overcast clouds	284.55	284.55	283.31	284.55
17	Las Vegas	US	2024-03-17	03:00:00	804	overcast clouds	284.07	284.07	282.84	284.07
18	Las Vegas	US	2024-03-17	06:00:00	804	overcast clouds	284.40	284.40	283.17	284.40
19	Las Vegas	US	2024-03-17	09:00:00	804	overcast clouds	283.91	283.91	282.63	283.91
20	Las Vegas	US	2024-03-17	12:00:00	803	broken clouds	282.96	282.96	282.96	282.96
21	Las Vegas	US	2024-03-17	15:00:00	804	overcast clouds	284.31	284.31	282.97	284.31
22	Las Vegas	US	2024-03-17	18:00:00	804	overcast clouds	288.22	288.22	286.93	288.22
23	Las Vegas	US	2024-03-17	21:00:00	804	overcast clouds	291.30	291.30	290.06	291.30
24	Las Vegas	US	2024-03-18	00:00:00	803	broken clouds	292.47	292.47	291.21	292.47

Overall, the process involves the setup and configuration of both AWS S3 and Snowflake components to enable seamless data transfer from an S3 bucket to Snowflake. This approach ensures efficient, secure, and automated data loading from AWS S3 to Snowflake.