



# Case Study

on

# Snowflake

Case study by,

Mouli S,  
Batch – 02,  
[moulisankar2002@outlook.com](mailto:moulisankar2002@outlook.com).



# Contents

1	Understandings of Case Study .....	3
2	Case Study Questions .....	4-14
2.1	Question 1 .....	4
2.2	Question 2 .....	7
2.3	Question 3 .....	11



## 1. Understandings of Case Study

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-COV-2 virus.

Most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, some will become seriously ill and require medical attention. Older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illness. Anyone can get sick with COVID-19 and become seriously ill or die at any age.

European Centre for Disease Prevention and Control (also known as ECDC) - an agency of the European Union tracks COVID-19 cases and vaccination status worldwide.

Reference Website - <https://www.ecdc.europa.eu/en/>



## 2. Case Study Questions

### 2.1 Question 1

Participants will create materialized view to copy data from \$3 bucket for customer data

([https://s3.ap-southeast-2.amazonaws.com/snowflake-essentials/ingesting data/new\\_customer/2019-09](https://s3.ap-southeast-2.amazonaws.com/snowflake-essentials/ingesting data/new_customer/2019-09)

24/generated\_customer\_data.csv) and display out having customers whose age is greater than 30 years and less than 50 years with following columns:

- a. CustomerName
- b. CustomerAge (as on today, Integer)
- c. CustomerCity

SQL Query:

```
CREATE DATABASE Covid_CaseStudy;
-- Question 1
-- Table creation
CREATE TABLE generated_customer_CSV (
  Customer_ID INTEGER,
  Customer_Name VARCHAR(100),
  Customer_Email VARCHAR(100),
  Customer_City VARCHAR(100),
  Customer_State VARCHAR(100),
  Customer_DOB DATE
)
CLUSTER BY (Customer_ID);
```



```
-- Stage creation
CREATE STAGE my_s3_stage_CS url='s3://snowflake-essentials/'
COPY INTO generated_customer_CSV
FROM s3://snowflake-essentials/ingesting_data/new_customer/2019-09-
24/generated_customer_data.csv
FILE_FORMAT = (TYPE = CSV FIELD_DELIMITER = '|' SKIP_HEADER = 1)
-- Age column added
ALTER TABLE generated_customer_CSV
ADD COLUMN age INT AS (DATE_PART('year', CURRENT_DATE()) - DATE_PART('year',
CUSTOMER_DOB));
SELECT * FROM generated_customer_CSV;
-- Conditions
-- Materialized view creation with conditions
CREATE MATERIALIZED VIEW generated_customer_mv AS
SELECT CUSTOMER_NAME AS CustomerName,
AGE AS CustomerAge,
CUSTOMER_CITY AS CustomerCity
FROM generated_customer_CSV
WHERE AGE > 30 AND AGE < 50;
SELECT * FROM generated_customer_mv;
SHOW TABLES;
ALTER ACCOUNT SET AUTO_CLUSTERING = ON;
```



## Output:

Worksheets CaseStudy Question\_1

ACCOUNTADMIN SNOW\_TUTO\_WH Share Updated 1 minute ago

Worksheets Databases

Search + ...

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1
- ass-1-demo
- batch\_02\_day1

COVID\_CASESTUDY.PUBLIC

```
23  
24 COPY INTO generated_customer_CSV  
25 FROM s3://snowflake-essentials/ingesting_data/new_customer/2019-09-24/generated_customer_data.csv  
26 FILE_FORMAT = (TYPE = CSV FIELD_DELIMITER = '|' SKIP_HEADER = 1)  
27  
28 -- Age column added
```

Objects Editor Results Chart

	CUSTOMERNAME	...	CUSTOMERAGE	CUSTOMERCITY
1	Jennifer Frederick		41	Canberra
2	Julian Hammond		43	Palmerston
3	David Nicholson		39	Redcliffe
4	Martena Perez		43	Cessnock
5	Tad Riddle		42	Lithgow
6	Dexter Stone		37	Brisbane
7	Maile Gallagher		40	Belgrave
8	Alisa Keith		34	Maryborough
9	Gannon Glenn		37	Cessnock
10	Kuame Casey		47	Geraldton-Greenough
11	Caldwell Dennis		38	Tamworth
12	Devine Duke		25	Bedford

Query Details

Query duration 504ms

Rows 33

CUSTOMERNAME Aa 100% filled

CUSTOMERAGE 123

CUSTOMERCITY Aa

Cessnock 2



## 2.2 Question 2

**Design Job (using ADF) to read the file (data\_20230330.csv) from blob storage and load into SQL Server.**

### SQL Query:

```
SELECT continentExp AS continents,
       countriesAndTerritories AS countries,
       CAST(year AS VARCHAR(255)) || '-' || CAST(month AS VARCHAR(255)) AS Year_Month,
       SUM(CAST(cases AS INT)) AS cases,
       SUM(CAST(deaths AS INT)) AS deaths
FROM COVID19_DATA
WHERE year = '2020'
GROUP BY continentExp, countriesAndTerritories, CAST(year AS VARCHAR(255)) || '-' ||
CAST(month AS VARCHAR(255));
-- Performing the integration process
CREATE NOTIFICATION INTEGRATION SNOWPIPE_INT_EVENT_CS
ENABLED = TRUE
TYPE = QUEUE
NOTIFICATION_PROVIDER = azure_storage_queue
azure_storage_queue_primary_uri =
'https://mousnowpipeaccount.queue.core.windows.net/snowpipe-noti-queue'
AZURE_TENANT_ID = '94124bcb-dea0-47d3-bfba-57b969f441dd'
SHOW INTEGRATIONS;
DESC notification integration SNOWPIPE_INT_EVENT_CS;
CREATE OR REPLACE STAGE "COVID_CASESTUDY"."PUBLIC"."AZURE_STAGE_CS"
url = 'azure://mousnowpipeaccount.blob.core.windows.net/snowflake-blob-src/'
credentials = (azure_sas_token=
'?sv=2021-12-02&ss=bfqt&srt=sco&sp=rwdlacupiytfx&se=2023-05-
09T19:43:45Z&st=2023-04-
07T11:43:45Z&spr=https&sig=Cxk0WRu%2BRYAec3CzDXEFtSAbkqYaAw4TEITCL%2F4lds
```



```
k%3D'
);
Is @"COVID_CASESTUDY"."PUBLIC"."AZURE_STAGE_CS";
-- Creating the pipe
CREATE OR REPLACE TABLE "COVID_CASESTUDY"."PUBLIC"."COVID19_DATA" (
  dateRep STRING,
    day STRING,
    month STRING,
    year STRING,
    cases STRING,
    deaths STRING,
    countriesAndTerritories STRING ,
    geold STRING,
    countryterritoryCode STRING,
    popData2019 STRING,
    continentExp STRING,
    "Cumulative_number_for_14_days_of_COVID-19_cases_per_100000" STRING
);
-- Creating the PIPE Between the Blob and Sowlake
CREATE OR REPLACE PIPE "COVID_CASESTUDY"."PUBLIC"."AZURE_SNOWFLAKE_PIPE"
  auto_ingest = true
  integration = 'SNOWPIPE_INT_EVENT_CS'
AS
COPY INTO "COVID_CASESTUDY"."PUBLIC"."COVID19_DATA"
FROM @"COVID_CASESTUDY"."PUBLIC"."AZURE_STAGE_CS"
file_format = (type = 'CSV' FIELD_DELIMITER = ',');
SHOW PIPES;
-- Load the files to the database
ALTER PIPE "COVID_CASESTUDY"."PUBLIC"."AZURE_SNOWFLAKE_PIPE" REFRESH;
SELECT * FROM "COVID_CASESTUDY"."PUBLIC"."COVID19_DATA";
```





## Azure Active Directory:

Microsoft Azure Upgrade Search resources, services, and docs (G+)

Home > Default Directory | Enterprise applications > Enterprise applications

### Enterprise applications | All applications

Default Directory - Azure Active Directory

+ New application Refresh Download (Export) Preview info Columns Preview features Got feedback?

Overview

View, filter, and search applications in your organization that are set up to use your Azure AD tenant as their identity Provider.

The list of applications that are maintained by your organization are in [application registrations](#).

Search by application name or object ID Application type == Enterprise Applications Application ID starts with Add filters

1 application found

Name	Object ID	Application ID	Homepage URL	Created on	Certificate Expiry Status
miicfsnowflakepacint	62d369d6-9ff4-4cc8-9c6b-913353844937	7db56686-4203-44ca-919e-dcd9ae2f2379	https://snowflake.net	4/7/2023	-

Manage

- All applications
- Application proxy
- User settings
- App launchers
- Custom authentication extensions (Preview)

## Role Assignment for Snowflake:

Microsoft Azure Upgrade Search resources, services, and docs (G+)

Home > mousnowpipeaccount | Access Control (IAM)

### mousnowpipeaccount | Access Control (IAM)

Storage account

Search Add Download role assignments Edit columns Refresh Remove Feedback

Check access Role assignments Roles Deny assignments Classic administrators

Number of role assignments for this subscription

2 4000

Search by name or email Type: All Role: All Scope: All scopes Group by: Role

1 items (1 Service Principals)

Name	Type	Role	Scope	Condition
miicfsnowflakepacint	App	Storage Queue Data Contributor	This resource	Add

Storage Queue Data Contributor

miicfsnowflakepacint

Storage Queue Data Contributor

This resource

Add

Overview

- Activity log
- Tags
- Diagnose and solve problems
- Access Control (IAM)
- Data migration
- Events
- Storage browser
- Data storage
- Containers
- File shares
- Queues
- Tables
- Security + networking



## Pipe Status:

Worksheets CaseStudy Question\_2

ACCOUNTADMIN SNOW\_TUTO\_WH Share Updated 5 seconds ago

Worksheets Databases

Search

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1
- ass-1-demo
- batch\_02\_day1

COVID\_CASESTUDY.PUBLIC

```
-- Load the files to the database
67 AS
68 COPY INTO "COVID_CASESTUDY"."PUBLIC"."COVID19_DATA"
69 FROM @"COVID_CASESTUDY"."PUBLIC"."AZURE_STAGE_CS"
70 file_format = (type = 'CSV' FIELD_DELIMITER = ',' );
71
72 SHOW PIPES;
73
74
75
```

Objects Editor Results Chart

File	Status
data_20230330.csv	SENT

Query Details

Query duration 1.1s

Rows 1

File 100% filled

Status 100% filled

## Output:

Worksheets CaseStudy Question\_2

ACCOUNTADMIN SNOW\_TUTO\_WH Share Updated 1 minute ago

Worksheets Databases

Search

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1
- ass-1-demo
- batch\_02\_day1

COVID\_CASESTUDY.PUBLIC

```
72 SHOW PIPES;
73
74 -- Load the files to the database
75
76 ALTER PIPE "COVID_CASESTUDY"."PUBLIC"."AZURE_SNOWFLAKE_PIPE" REFRESH;
77
```

Objects Editor Results Chart

	DATEREP	DAY	MONTH	YEAR	CASES	DEATHS	COUNTRIESA
1	14/12/2020	14	12	2020	746	6	Afghanistan
2	13/12/2020	13	12	2020	298	9	Afghanistan
3	12/12/2020	12	12	2020	113	11	Afghanistan
4	11/12/2020	11	12	2020	63	10	Afghanistan
5	10/12/2020	10	12	2020	202	16	Afghanistan
6	09/12/2020	9	12	2020	135	13	Afghanistan
7	08/12/2020	8	12	2020	200	6	Afghanistan
8	07/12/2020	7	12	2020	210	26	Afghanistan
9	06/12/2020	6	12	2020	234	10	Afghanistan
10	05/12/2020	5	12	2020	235	18	Afghanistan
11	04/12/2020	4	12	2020	119	5	Afghanistan

Query Details

Query duration 36ms

Rows 13

DATEREP 100% filled

DAY 100% filled

MONTH 100% filled



## 2.3 Question 3

Participants will create tasks for the snow pipe they have created in Question No. 2. Output table and file will consist of following columns for Year 2020:

- a. Continent (e.g... Asia)
- b. Countries (e.g... Afghanistan)
- c. Year Month (e.g., 2020-01)
- d. Cases (e.g., 0)
- e. Deaths (e.g., 0)

SQL Query:

```
-- Create a task that references the stored procedure and specifies the schedule for the task to run
```

```
CREATE TASK COVID19_DATA_task2
WAREHOUSE = SNOW_TUTO_WH
SCHEDULE = 'USING CRON 0 1 * * * America/Los_Angeles'
AS
CALL COVID19_DATA_task1_pro();
CREATE OR REPLACE TASK my_task
  WAREHOUSE = COMPUTE_WH
  SCHEDULE = '1 MINUTE'
AS
```

```
-- Create a target table to store the result
```

```
CREATE TABLE IF NOT EXISTS output_table (
  continents VARCHAR(255),
  countries VARCHAR(255),
  Year_Month VARCHAR(255),
  cases INT,
```



```
    deaths INT
);
-- Insert the result of the query into the target table
INSERT INTO output_table (continents, countries, Year_Month, cases, deaths)
SELECT continentExp AS continents,
    countriesAndTerritories AS countries,
    CAST(year AS VARCHAR(255)) || '-' || CAST(month AS VARCHAR(255)) AS Year_Month,
    SUM(CAST(cases AS INT)) AS cases,
    SUM(CAST(deaths AS INT)) AS deaths
FROM COVID19_DATA
WHERE year = '2020'
GROUP BY continentExp, countriesAndTerritories, CAST(year AS VARCHAR(255)) || '-' ||
CAST(month AS VARCHAR(255));
-- Task check
SHOW TASKS;
SHOW TABLES
-- Resume task
ALTER TASK MY_TASK RESUME;
--CHECK TASK HISTORY
SELECT * FROM TABLE(INFORMATION_SCHEMA.TASK_HISTORY()) WHERE NAME =
'MY_TASK';
--DISPLAY
SELECT * FROM OUTPUT_TABLE;
```



## Created Task:

Worksheets CaseStudy Question\_3

ACCOUNTADMIN SNOW\_TUTO\_WH Share Updated 8 seconds ago

Worksheets Databases

Search + ...

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1
- ass-1-demo
- batch\_02\_day1

COVID\_CASESTUDY.PUBLIC

```
54 1 MINUTE year = 2020
55 GROUP BY continentExp, countriesAndTerritories, CAST(year AS VARCHAR(255)) || '-' || CAST(month AS
56 VARCHAR(255));
57 -- Task check
58
59 SHOW TASKS;
```

Objects Editor Results Chart

	created_on	name	id
1	2023-04-07 05:52:12.460 -0700	COVID19_DATA_TASK	01ab77c4-9af7-c045-0000-000000000000
2	2023-04-07 07:05:31.986 -0700	COVID19_DATA_TASK1	01ab780d-5c94-1dfe-0000-000000000000
3	2023-04-07 07:07:02.914 -0700	COVID19_DATA_TASK2	01ab780f-9a2c-0ef8-0000-000000000000
4	2023-04-07 07:31:48.360 -0700	MY_TASK	01ab7827-2ae3-385b-0000-000000000000

Query Details ...

Query duration 45ms

Rows 4

created\_on 100% filled

name Aa 100% filled

id Aa

## Resumed Task:

Worksheets CaseStudy Question\_3

ACCOUNTADMIN SNOW\_TUTO\_WH Share Updated 24 seconds ago

Worksheets Databases

Search + ...

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2

COVID\_CASESTUDY.PUBLIC

```
58
59 SHOW TASKS;
60 SHOW TABLES;
61 -- Resume task
62
63 ALTER TASK MY_TASK RESUME;
64
65 SHOW METADATA
```

Objects Editor Results Chart

	warehouse	schedule	predecessors	state
1	SNOW_TUTO_WH	USING CRON 0 1 * * * America/Los_Angeles	[]	started
2	SNOW_TUTO_WH	USING CRON 0 1 * * * America/Los_Angeles	[]	started
3	SNOW_TUTO_WH	USING CRON 0 1 * * * America/Los_Angeles	[]	started
4	COMPUTE_WH	1 MINUTE	[]	started

Query Details ...

Query duration 44ms

Rows 4

created\_on 100% filled



## Task History:

Worksheets CaseStudy Question\_3

ACCOUNTADMIN SNOW\_TUTO\_WH Share

Updated 17 seconds ago Draft

Worksheets Databases

Search + ...

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1
- ass-1-demo
- batch\_02\_day1

COVID\_CASESTUDY.PUBLIC

```
--CHECK TASK HISTORY
SELECT * FROM TABLE(INFORMATION_SCHEMA.TASK_HISTORY()) WHERE NAME = 'MY_TASK';
```

Objects Editor Results Chart

	SCHEMA_NAME	QUERY_TEXT	CONDITION_TEXT	...	STATE
1	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SCHEDULED
2	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
3	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
4	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
5	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
6	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
7	PUBLIC	CREATE TABLE IF NOT EXISTS output_table ( continents VARCHAR(255),	null		SUCCEEDED
8	PUBLIC	SELECT continentExp AS continents, countriesAndTerritories AS countries,	null		SUCCEEDED
9	PUBLIC	SELECT continentExp AS continents, countriesAndTerritories AS countries,	null		SUCCEEDED
10	PUBLIC	SELECT continentExp AS continents, countriesAndTerritories AS countries,	null		SUCCEEDED

## Output Table named as 'output\_table:

Worksheets CaseStudy Question\_3

ACCOUNTADMIN SNOW\_TUTO\_WH Share

Updated 1 minute ago

Worksheets Databases

Search + ...

CaseStudy

- Question\_3
- Question\_2
- Question\_1

learn\_from\_vinit

- Pipe
- vinit\_28
- VINIT\_27-03
- tuto
- pipe\_tuto
- json\_file
- day2
- snowflake\_ass\_csv\_2
- snowflake\_ass\_csv\_1

COVID\_CASESTUDY.PUBLIC

```
-- Task check
SHOW TASKS;
SHOW TABLES;
-- Resume task
ALTER TASK MY_TASK RESUME;
SHOW METADATA
SELECT * FROM OUTPUT_TABLE;
--CHECK TASK HISTORY
SELECT * FROM TABLE(INFORMATION_SCHEMA.TASK_HISTORY()) WHERE NAME = 'MY_TASK';
```

Objects Editor Results Chart

	CONTINENTS	COUNTRIES	YEAR_MONTH	...	CASES	DEATHS
1	Asia	Afghanistan	2020-12		3,157	197