create or storage integration s3\_int

type = external\_stage

storage\_provider = 'S3'

enabled = true

storage\_aws\_role\_arn = 'arn:aws:iam::627741535532:role/kajalpatelrole'

storage\_allowed\_locations = ('s3://kajalpatelbucket/heart\_2020.csv');

DESC INTEGRATION s3\_int;

--grant create or replace stage on schema public to role kajalpatelrole;

--grant usage on integration s3\_int to role kajalpatelrole;

-- Create External Stage --

create or replace DATABASE DBPROJECT2;

create or replace table "DBPROJECT2"."PUBLIC"."Table\_Heart"

(HeartDisease varchar,BMI float,Smoking varchar,AlcoholDrinking varchar,

Stroke varchar,PhysicalHealth int,MentalHealth int,DiffWalking varchar,Sex varchar,

AgeCategory string,Race varchar,Diabetic varchar,PhysicalActivity varchar,GenHealth varchar,

SleepTime int,Asthma varchar,KidneyDisease varchar,SkinCancer varchar,id int);

create or replace FILE FORMAT "DBPROJECT2"."PUBLIC".KajalpatelFileformate\_CSV TYPE = 'CSV' COMPRESSION = 'AUTO' FIELD\_DELIMITER = ','

RECORD\_DELIMITER = '\n' SKIP\_HEADER = 1 FIELD\_OPTIONALLY\_ENCLOSED\_BY = 'NONE'

TRIM\_SPACE = FALSE ERROR\_ON\_COLUMN\_COUNT\_MISMATCH = TRUE ESCAPE = 'NONE' ESCAPE\_UNENCLOSED\_FIELD = '\134'

DATE\_FORMAT = 'AUTO' TIMESTAMP\_FORMAT = 'AUTO' NULL\_IF = ('\\N');

CREATE or replace STAGE "DBPROJECT2"."PUBLIC".kajalpatelstage URL = 's3://kajalpatelbucket/heart\_2020.csv'

CREDENTIALS = (AWS\_KEY\_ID = 'AKIAZEKCPVUWPYB4MA7J' AWS\_SECRET\_KEY = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*');

create or replace STAGE "DBPROJECT2"."PUBLIC"."Kajalpatelstage" URL = 's3://kajalpatelbucket/heart\_2020.csv'

STORAGE\_INTEGRATION = s3\_int;

COPY INTO "DBPROJECT2"."PUBLIC"."Table\_Heart" FROM '@"DBPROJECT2"."PUBLIC"."KAJALPATELSTAGE1"'

FILE\_FORMAT = '"DBPROJECT2"."PUBLIC"."KajalpatelFileformate\_CSV"' ON\_ERROR = 'CONTINUE' ;

copy into "DBPROJECT2"."PUBLIC"."Table\_Heart"

from s3://kajalpatelbucket/heart\_2020.csv

storage\_integration = s3\_int FILE\_FORMAT = '"DBPROJECT2"."PUBLIC"."KajalpatelFileformate\_CSV"';

create or replace TASK mytask\_minute1 WAREHOUSE = COMPUTE\_WH, SCHEDULE = 'USING CRON 0 0 \* \* \*/THU Asia/Kolkata' AS

copy into "DBPROJECT2"."PUBLIC"."Table\_Heart"

from (select t.$1, t.$2, t.$3, t.$4, t.$5,t.$6,t.$7, t.$8, t.$9, t.$10, t.$11, t.$12, t.$13, t.$14, t.$15, t.$16, t.$17, t.$18, t.$19

from "DBPROJECT2"."PUBLIC"."Table\_Heart") on\_error = 'skip\_file' file\_format = my\_csv\_format;

alter task mytask\_minute1 resume;

alter task mytask\_minute suspend;

SHOW TASKS;

/\*\*\*creating snowpipe\*\*\*\*\*/

create or replace table "DBPROJECT2"."PUBLIC"."Table\_Heart" (HeartDisease varchar,BMI float,Smoking varchar,AlcoholDrinking varchar,Stroke varchar,

PhysicalHealth int,MentalHealth int,DiffWalking varchar,Sex varchar,AgeCategory string,

Race varchar,Diabetic varchar,PhysicalActivity varchar,GenHealth varchar,SleepTime int,

Asthma varchar,KidneyDisease varchar,SkinCancer varchar,id int);

create or replace pipe "DBPROJECT2"."PUBLIC"."mypipe\_s3" auto\_ingest = true as

copy into "DBPROJECT2"."PUBLIC"."Table\_Heart"

from (select t.$1 , t.$2, t.$3, t.$4, t.$5 ,t.$6, t.$7, t.$8, t.$9, t.$10, t.$11, t.$12, t.$13, t.$14, t.$15, t.$16, t.$17, t.$18, t.$19

from "DBPROJECT2"."PUBLIC"."Kajalpatelstage" );

list "DBPROJECT2"."PUBLIC"."Kajalpatelstage";

select \* from stadium;

/\*\*\*\*\* scd2\*\*\*\*\*\*\*/

create or replace stream Table\_Heart\_stream on table "DBPROJECT2"."PUBLIC"."Table\_Heart";

copy into "DBPROJECT2"."PUBLIC"."Table\_Heart"

from (select t.$1, t.$2, t.$3, t.$4, t.$5, t.$6, t.$7, t.$8, t.$9, t.$10, t.$11, t.$12, t.$13, t.$14, t.$15, t.$16, t.$17, t.$18, t.$19

from "DBPROJECT2"."PUBLIC"."Kajalpatelstage" t) on\_error = 'skip\_file' file\_format = my\_csv\_format;

-- Create consumer table

create or replace table Consumer\_Heart(HeartDisease varchar,BMI float,Smoking varchar,AlcoholDrinking varchar,Stroke varchar,

PhysicalHealth int,MentalHealth int,DiffWalking varchar,Sex varchar,AgeCategory string,Race varchar,

Diabetic varchar,PhysicalActivity varchar,GenHealth varchar,SleepTime int,Asthma varchar,

KidneyDisease varchar,SkinCancer varchar,id int, stream\_type string,

rec\_version number default 0,REC\_DATE TIMESTAMP\_LTZ);

-- updating task

CREATE or replace TASK tgt\_merge WAREHOUSE = compute\_wh SCHEDULE = '1 minute'

WHEN SYSTEM$STREAM\_HAS\_DATA('strm\_cyclone') AS merge into Consumer\_Heart t

using Table\_Heart\_stream s on t."ID"=s."ID" and (metadata$action='DELETE')

when matched and metadata$isupdate='FALSE' then update set rec\_version=9999, stream\_type='DELETE'

when matched and metadata$isupdate='TRUE' then update set rec\_version=rec\_version-1

when not matched then insert (HeartDisease, BMI, Smoking, AlcoholDrinking ,Stroke, PhysicalHealth, MentalHealth, DiffWalking, Sex,

AgeCategory, Race, Diabetic, PhysicalActivity, GenHealth, SleepTime, Asthma, KidneyDisease,

SkinCancer,id,stream\_type ,rec\_version,REC\_DATE) values(s.HeartDisease ,s.BMI ,s.Smoking,

s.AlcoholDrinking ,s.Stroke ,s.PhysicalHealth ,s.MentalHealth ,s.DiffWalking ,s.Sex ,s.AgeCategory,

s.Race ,s.Diabetic,s.PhysicalActivity, s.GenHealth, s.SleepTime, s.Asthma, s.KidneyDisease ,

s.SkinCancer, s.id, metadata$action,0,CURRENT\_TIMESTAMP());

ALTER TABLE Consumer\_Heart ADD ID INT;

SELECT \* from Consumer\_Heart;

SELECT \* FROM strm\_cyclone;

show tasks;

select \* from Consumer\_Heart;

DESC table target\_t;

-- Sql Query--

--1.Which gender have majority heart atteck

select count(HEARTDISEASE),sex

from "DBPROJECT2"."PUBLIC"."Kajaltable2"

where HEARTDISEASE = 'Yes'

group by sex;

-- 2.Average Age at which people genrally get heart atteck

SELECT FLOOR(AVG(SUBSTRING(AgeCategory, 1,2))) as AGE

FROM "DBPROJECT2"."PUBLIC"."Kajaltable2"

where stroke = 'Yes'

--3.Which Combination of Physical and Mental health causes stroke

SELECT HEARTDISEASE,PHYSICALHEALTH, MENTALHEALTH

from "DBPROJECT2"."PUBLIC"."Kajaltable2"

WHERE HEARTDISEASE = 'Yes' ;

--4.GenHeealth wise average BMI of People

select GENHEALTH, AVG(BMI)

from "DBPROJECT2"."PUBLIC"."Kajaltable2"

group by GENHEALTH;

--5.Average age of people having problem in walking

SELECT FLOOR(AVG(SUBSTRING(AGECATEGORY, 1,2))) as AGE

FROM "DBPROJECT2"."PUBLIC"."Kajaltable2"

where DiffWalking = 'Yes';

select FLOOR(AVG(SUBSTRING(AGECATEGORY, 1,2))) as AGE from Consumer\_Heart where DIFFWALKING = 'Yes';

--6.count the different comabination of people having heart- attack against BMI,

--Smoking, alcoholdrink strok, diabetic, kidney diseses, skin cancer.

SELECT HEARTDISEASE, SMOKING

from "DBPROJECT2"."PUBLIC"."Kajaltable2"

where HEARTDISEASE = 'Yes' AND SMOKING ='Yes' AND ALCOHOLDRINKING = 'Yes' AND STROKE = 'Yes' AND DIABETIC = 'Yes' AND KIDNEYDISEASE = 'Yes' AND SKINCANCER = 'Yes';

--7. Average BMI of Having Sking Cancer

SELECT SKINCANCER, AVG(BMI)

FROM "DBPROJECT2"."PUBLIC"."Kajaltable2"

where SKINCANCER = 'Yes'

group by SKINCANCER;

--8. possibilities/percentage of people having both heartattack and skin cancer

SELECT HEARTDISEASE,SKINCANCER, count(ID) \* 100.0 / sum(count(ID)) Over() as Percentage

FROM "DBPROJECT2"."PUBLIC"."Kajaltable2"

where HEARTDISEASE = 'Yes' and SKINCANCER = 'Yes'

GROUP BY HEARTDISEASE,SKINCANCER;