create or replace database p7;

create or replace schema p7\_schema;

use database p7;

use schema p7\_schema;

create or replace storage integration a\_d

type = external\_stage

storage\_provider = s3

enabled = true

storage\_aws\_role\_arn = 'arn:aws:iam::211125294315:role/aws\_role3'

storage\_allowed\_locations=('s3://ombucket25/folderdataload/');

desc integration a\_d;

-------------------------------------------------------------------------------

CREATE or replace TABLE HR\_Ana (

EmpID VARCHAR(10),

Age INT,

AgeGroup VARCHAR(10),

Attrition VARCHAR(3),

BusinessTravel VARCHAR(20),

DailyRate INT,

Department VARCHAR(50),

DistanceFromHome INT,

Education INT,

EducationField VARCHAR(50),

EmployeeCount INT,

EmployeeNumber INT,

EnvironmentSatisfaction INT,

Gender VARCHAR(10),

HourlyRate INT,

JobInvolvement INT,

JobLevel INT,

JobRole VARCHAR(50),

JobSatisfaction INT,

MaritalStatus VARCHAR(20),

MonthlyIncome INT,

SalarySlab VARCHAR(50),

MonthlyRate INT,

NumCompaniesWorked INT,

Over18 VARCHAR(1),

OverTime VARCHAR(10),

PercentSalaryHike INT,

PerformanceRating INT,

RelationshipSatisfaction INT,

StandardHours INT,

StockOptionLevel INT,

TotalWorkingYears INT,

TrainingTimesLastYear INT,

WorkLifeBalance INT,

YearsAtCompany INT,

YearsInCurrentRole INT,

YearsSinceLastPromotion INT

);

create or replace stream data\_stream on table HR\_Ana;

select\* from data\_stream;

-------------------------------------------------------------------------------

-- comulsory add file format when data loading:

CREATE or REPLACE FILE FORMAT a\_d\_f1

TYPE = 'CSV'

COMPRESSION = 'AUTO'

FIELD\_DELIMITER =','

RECORD\_DELIMITER = '\n'

SKIP\_HEADER =1

FIELD\_OPTIONALLY\_ENCLOSED\_BY = '\042'

TRIM\_SPACE = FALSE

ERROR\_ON\_COLUMN\_COUNT\_MISMATCH = TRUE

ESCAPE = 'NONE'

ESCAPE\_UNENCLOSED\_FIELD = '\134'

DATE\_FORMAT = 'AUTO'

TIMESTAMP\_FORMAT = 'AUTO';

-- data loading stage creation:

create or replace stage s1\_stage

url ='s3://ombucket25/folderdataload/'

storage\_integration = a\_d

file\_format = a\_d\_f1;

list @s1\_stage;

-------------------------------------------------------------------------------

create or replace pipe pipe\_a auto\_ingest=true as

copy into HR\_ANA from @s1\_stage on\_error= continue;

desc pipe pipe\_a;

alter pipe pipe\_a set pipe\_execution\_paused=false;

alter pipe pipe\_a refresh;

-------------------------------------------------------------------------------

CREATE or replace TABLE HR\_Ana\_target(

EmpID VARCHAR(10),

Age INT,

AgeGroup VARCHAR(10),

Attrition VARCHAR(3),

BusinessTravel VARCHAR(20),

DailyRate INT,

Department VARCHAR(50),

DistanceFromHome INT,

Education INT,

EducationField VARCHAR(50),

EmployeeCount INT,

EmployeeNumber INT,

EnvironmentSatisfaction INT,

Gender VARCHAR(10),

HourlyRate INT,

JobInvolvement INT,

JobLevel INT,

JobRole VARCHAR(50),

JobSatisfaction INT,

MaritalStatus VARCHAR(20),

MonthlyIncome INT,

SalarySlab VARCHAR(50),

MonthlyRate INT,

NumCompaniesWorked INT,

Over18 VARCHAR(1),

OverTime VARCHAR(10),

PercentSalaryHike INT,

PerformanceRating INT,

RelationshipSatisfaction INT,

StandardHours INT,

StockOptionLevel INT,

TotalWorkingYears INT,

TrainingTimesLastYear INT,

WorkLifeBalance INT,

YearsAtCompany INT,

YearsInCurrentRole INT,

YearsSinceLastPromotion INT

);

--------------------------------------------------------------------------------

-------------------------------------------------------------------------------

-- task:

create or replace task a\_task

warehouse = compute\_wh

schedule = '1 minute'

when system$stream\_has\_data('data\_stream')

as insert into HR\_Ana\_target

select EmpID ,

Age ,

AgeGroup ,

Attrition ,

BusinessTravel ,

DailyRate ,

Department ,

DistanceFromHome ,

Education ,

EducationField ,

EmployeeCount ,

EmployeeNumber ,

EnvironmentSatisfaction ,

Gender ,

HourlyRate ,

JobInvolvement ,

JobLevel ,

JobRole ,

JobSatisfaction ,

MaritalStatus ,

MonthlyIncome ,

SalarySlab ,

MonthlyRate ,

NumCompaniesWorked ,

Over18 ,

OverTime ,

PercentSalaryHike ,

PerformanceRating ,

RelationshipSatisfaction ,

StandardHours ,

StockOptionLevel ,

TotalWorkingYears ,

TrainingTimesLastYear ,

WorkLifeBalance ,

YearsAtCompany ,

YearsInCurrentRole ,

YearsSinceLastPromotion

from data\_stream

where metadata$action='INSERT' and metadata$isupdate='FALSE';

--

show tasks;

alter task a\_task resume;

-------------------------------------------------------------------------------

select\* from HR\_ANA;

select \*from P3.P7\_SCHEMA.HR\_ANA\_TARGET;

-------------------------------------------------

-- and then this data from HR\_Ana\_target Transform to Power BI through Snowflake & this is the Combine Data Engineering & Data Analytics Project.

-- THANK YOU !