Date

DIFFERENT Ways to use copy command OPTIONS

# Goals

Copy command options will come handy for different scenarios. This assignment is about how to leverage one of the copy command option for our advantage.

**Business case :**

Historical files have data required for two tables merged in only one file.

Data volume is , 1TB.

**Challenges,**

Because the data volume is very huge, you can’t use any ETL tool to split the files and load data.

You can’t use Unix commands as it’s consumes lot of time to split the file based on some criteria.

# **Peparation**

Create tables,

**Table 1 :**

CREATE TABLE HR\_EMP\_DATA

(

Emp\_ID VARCHAR,

Name\_Prefix VARCHAR,

First\_Name VARCHAR,

Middle\_Initial VARCHAR,

Last\_Name VARCHAR,

Gender VARCHAR,

E\_Mail VARCHAR,

Father\_Name VARCHAR,

Mother\_Name VARCHAR,

Mother\_Maiden\_Name VARCHAR,

Date\_of\_Birth VARCHAR,

Time\_of\_Birth VARCHAR,

Age\_in\_Yrs VARCHAR,

Weight\_in\_Kgs VARCHAR,

Date\_of\_Joining VARCHAR,

Quarter\_of\_Joining VARCHAR,

Half\_of\_Joining VARCHAR,

Year\_of\_Joining VARCHAR,

Month\_of\_Joining VARCHAR,

Month\_Name\_of\_Joining VARCHAR,

Short\_Month VARCHAR,

Day\_of\_Joining VARCHAR,

DOW\_of\_Joining VARCHAR,

Short\_DOW VARCHAR,

Age\_in\_Company VARCHAR,

Salary VARCHAR,

Last\_Hike VARCHAR,

SSN VARCHAR,

Phone VARCHAR,

Place\_Name VARCHAR,

County VARCHAR,

City VARCHAR,

State VARCHAR,

Zip VARCHAR,

Region VARCHAR,

User\_Name VARCHAR,

Password VARCHAR

)

**Table 2 :**

CREATE TABLE CREDIT\_CARD

(

Date VARCHAR,

Description VARCHAR,

Deposits VARCHAR,

Withdrawls VARCHAR,

Balance VARCHAR

)

Download data from, below location,

<https://drive.google.com/file/d/1g_Q_sKINfhvG45BvI1wB35-ZHc2dlcQO/view?usp=sharing>

Data for both tables has been saved in only one file.

Task is to copy data from this single file to both the tables.

You should not be using unix command to split this file.

Let’s solve this problem,

# **Solution**

**STEP 1 : Create named staging area.**

CREATE STAGE HR\_CREDIT\_STAGE

A picture containing circuit, sitting, holding, person

Description automatically generated

**STEP 2 : Use put command and upload file to staging area.**

put file:///data-vol/HR\_CREDIT.csv @DEMO\_DB.PUBLIC.HR\_CREDIT\_STAGE;

A picture containing graphical user interface

Description automatically generated

**STEP 3 : Copy data to Employee table.**

Create file format,

create or replace file format csv\_format

type = csv field\_delimiter = ',' skip\_header = 1 null\_if = ('NULL', 'null') empty\_field\_as\_null = true compression = gzip;

Text

Description automatically generated

Check how many records you have in staging area.

Graphical user interface, text

Description automatically generated

Execute copy command,

COPY INTO HR\_EMP\_DATA

from @DEMO\_DB.PUBLIC.HR\_CREDIT\_STAGE

file\_format = (FORMAT\_NAME =’csv\_format’ ERROR\_ON\_COLUMN\_COUNT\_MISMATCH= TRUE field\_optionally\_enclosed\_by='"')

ON\_ERROR = ‘CONTINUE’

Graphical user interface

Description automatically generated

Query and check records in HR\_EMP\_DATA,

Table

Description automatically generated

**STEP 4 : Check rejected records,**

select \* from table(validate(HR\_EMP\_DATA, job\_id=>'<query\_id>'));

**STEP 5 : Execute copy command to load data to CREDIT\_CARD table,**

COPY INTO CREDIT\_CARD

from @DEMO\_DB.PUBLIC.HR\_CREDIT\_STAGE

file\_format = (FORMAT\_NAME =’csv\_format’ ERROR\_ON\_COLUMN\_COUNT\_MISMATCH= TRUE field\_optionally\_enclosed\_by='"')

ON\_ERROR = ‘CONTINUE’

Graphical user interface

Description automatically generated

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Try to improvise the copy command. You have to enable few more options to maintain data integrity.

#### Alter file format object , **csv\_format** above to enable **ERROR\_ON\_COLUMN\_COUNT\_MISMATCH** property.

#### Write you command below, Also re write your copy command.

#### You have to satisfy one more requirement,

###### *If column values are more than the specified length, then , record should be rejected.*

###### *Ex: if you have defined varchar(10) and you receive string of length 15 then record should get rejected.*

###### *Which property you will enable ?*

###### *Write your command below,*

Hint : Refer to documentation below,

<https://docs.snowflake.com/en/sql-reference/sql/copy-into-table.html>

Go to this section and figure out !!!!

![Graphical user interface, text, application, email

Description automatically generated]()