

## IMT577 Summer 2022

### Assignment 5.12-Develop ELT and Staging in Snowflake

Lizzy Schmitz

Steps I took to complete assignment:

#### 1.Created Database and Virtual Data Warehouse in Snowflake

Databases				
<div><div><div><div><div></div><div>Create...</div></div><div><div></div><div>Clone...</div></div><div><div></div><div>Drop...</div></div><div><div></div><div>Transfer Ownership</div></div></div></div></div>				
<div><div><div>Search Databases</div></div><div>4 databases</div></div>				
Database	Origin	↓ Creation Time	Owner	Comment
IMT577_DW_LIZZY_SCHMITZ_...		12:58 PM	ACCOUNTADMIN	

Warehouses										
Manage your warehouses from this page. To operate on your data, you need to create one or more warehouses.										
<div><div><div><div><div></div><div>Create...</div></div><div><div></div><div>Configure...</div></div><div><div></div><div>Suspend...</div></div><div><div></div><div>Resume...</div></div><div><div></div><div>Drop...</div></div><div><div></div><div>Transfer Ownership</div></div></div></div></div>										
Status	Warehouse Name	Size	Runn...	Que...	Auto Suspe...	Auto Resume	Created On ▼	Resumed On	Owner	Comment
Started	IMT577_DW_LIZZY_SCHMITZ_...	X-Small	0	0	10 minutes	Yes	1:01:34 PM	3:59:49 PM	ACCOUNTADMIN	

#### 2.Created Stages within datawarehouse (match one for one on data extracts listed in assignment)

Databases > IMT577\_DW\_LIZZY\_SCHMITZ\_M5

TablesViewsSchemasStagesFile FormatsSequencesPipes

Create...

Clone...

Edit...

Drop...

Transfer Ownership

Stage	Schema	Location	Creation Time ▼	Owner	Comment
STAGE_TARGETDATAPRODUCT	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:15:22 PM	ACCOUNTADMIN	
STAGE_TARGETDATACHANNE...	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:14:54 PM	ACCOUNTADMIN	
STAGE_STORE	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:13:56 PM	ACCOUNTADMIN	
STAGE_SALESHEADER	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:13:13 PM	ACCOUNTADMIN	
STAGE_SALESDetail	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:12:45 PM	ACCOUNTADMIN	
STAGE_RESELLER	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:12:22 PM	ACCOUNTADMIN	
STAGE_PRODUCTTYPE	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:11:59 PM	ACCOUNTADMIN	
STAGE_PRODUCTCATEGORY	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:11:36 PM	ACCOUNTADMIN	
STAGE_PRODUCT	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:11:04 PM	ACCOUNTADMIN	
STAGE_CUSTOMER	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:10:37 PM	ACCOUNTADMIN	
STAGE_CHANNELCATEGORY	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:10:10 PM	ACCOUNTADMIN	
STAGE_CHANNEL	PUBLIC	azure://imt577sourcedata.blob.core.windows.ne...	1:08:57 PM	ACCOUNTADMIN	

### 3.Created tables and loaded the CSV files into (matching table to corresponding CSV)

Databases > IMT577\_DW\_LIZZY\_SCHMITZ\_M5

Tables Views Schemas Stages File Formats Sequences Pipes

+ Create... + Create Like... Clone... Load Data... Drop... Transfer Ownership

Table Name	Schema	Creation Time ▼	Owner	Rows	Size	Comment
STAGE_RESELLER	PUBLIC	3:59:15 PM	ACCOUNTADMIN	4	5KB	
STAGE_TARGETDATAPRODUCT	PUBLIC	2:57:54 PM	ACCOUNTADMIN	48	2.5KB	
STAGE_TARGETDATACHANNE...	PUBLIC	2:54:37 PM	ACCOUNTADMIN	22	2KB	
STAGE_STORE	PUBLIC	2:45:09 PM	ACCOUNTADMIN	6	5KB	
STAGE_SALESHEADER	PUBLIC	2:41:36 PM	ACCOUNTADMIN	187.3K	397KB	
STAGE_SALESDetail	PUBLIC	2:37:03 PM	ACCOUNTADMIN	187.3K	822.5KB	
STAGE_PRODUCTTYPE	PUBLIC	2:29:09 PM	ACCOUNTADMIN	10	2.5KB	
STAGE_PRODUCTCATEGORY	PUBLIC	2:24:29 PM	ACCOUNTADMIN	4	2KB	
STAGE_PRODUCT	PUBLIC	2:22:28 PM	ACCOUNTADMIN	24	5.5KB	
STAGE_CUSTOMER	PUBLIC	2:14:32 PM	ACCOUNTADMIN	3	5.5KB	
STAGE_CHANNELCATEGORY	PUBLIC	2:09:08 PM	ACCOUNTADMIN	2	2KB	
STAGE_CHANNEL	PUBLIC	2:05:37 PM	ACCOUNTADMIN	5	2.5KB	

### 4. Confirmed all data was loaded into tables

< New Worksheet +

Find database objects Starting with...

IMT577\_DW\_LIZZY\_SCHMITZ\_M5

- INFORMATION\_SCHEMA
- PUBLIC
  - Tables
    - STAGE\_CHANNEL
    - STAGE\_CHANNELCATEGORY
    - STAGE\_CUSTOMER
    - STAGE\_PRODUCT
    - STAGE\_PRODUCTCATEGORY
    - STAGE\_PRODUCTTYPE
    - STAGE\_RESELLER
    - STAGE\_SALESDetail
    - STAGE\_SALESHEADER
    - STAGE\_STORE
    - STAGE\_TARGETDATACHANNELRESELL...
    - STAGE\_TARGETDATAPRODUCT

Run All Queries Saved 0 seconds ago

```
1 create or replace file format CSV_SKIP_HEADER
2 type = 'csv'
3 field_delimiter = ','
4 FIELD_OPTIONALLY_ENCLOSED_BY = ''
5 skip_header = 1;
6
7 --CONFIRM DATA COPIED INTO TABLE--
8 SELECT * FROM "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_CHANNEL"
9 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_CHANNELCATEGORY"
10 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_CUSTOMER"
11 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_PRODUCT"
12 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_PRODUCTCATEGORY"
13 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_PRODUCTTYPE"
14 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_RESELLER"
15 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_SALESDetail"
16 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_SALESHEADER"
17 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_STORE"
18 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_TARGETDATACHANNELRESELLER"
19 select * from "IMT577_DW_LIZZY_SCHMITZ_M5"."PUBLIC"."STAGE_TARGETDATAPRODUCT"
20
21 |
```

## Shared database

Secure Shares				
Inbound	Outbound	⊕ Create	👤 Add Consumers	✎ Edit
Search Outbound Secure Shares		1 Outbound Secure Shares		
Secure Share Name	Shared With	Database	↓ Creation Time	Owner
IMT577_DW_LIZZYSCHMITZ_M5	WE74592,NR86927	IMT577_DW_LIZZY_SCHMITZ_M5	4:38:46 PM	ACCOUNTADMIN

## Questions

- Name of your database share: **IMT577\_DW\_LIZZYSCHMITZ\_M5**
- GitHub repo link: [https://github.com/lizzyschmitz/lischmitz\\_IMT577](https://github.com/lizzyschmitz/lischmitz_IMT577)
- What went well for you? All of the steps were easy to follow after watching the videos and the UI is very intuitive. I also liked that it provided the SQL code that would be generated based on my selections in the UI (very helpful as I am newer to SQL)
- What didn't go well?
  - Loading the CSVs (Asure URL loads were not working) was straightforward, but I did encounter an error when loading the reseller table. Other files have a similar column/data, but weren't producing errors, so I knew it was an issue with the table itself. I was able to troubleshoot and figure out the columns were in a slightly different order than the CSV file. Once I corrected that issue, the load worked.

### Error

Unable to copy files into table.  
Timestamp 'company\SQLServerSalesAccount' is not recognized File  
'@STAGE\_RESELLER/ui1658701922440/Reseller.csv', line 2, character 167 Row 1, column  
"STAGE\_RESELLER"["CREATEDDATE":11] If you would like to continue loading when an error  
is encountered, use other values such as 'SKIP\_FILE' or 'CONTINUE' for the ON\_ERROR  
option. For more information on loading options, please run 'info loading\_data' in a SQL  
client.

OK

- - I had forgotten to uncheck the "Not Null" column on my first table and I was unable to figure out how to go back into that same UI to update the column characteristics. From what I was able to research, this might only be possible via the worksheet with SQL.
- Did you learn something significant to share with your peers?
  - Taking your time and double checking before you hit any "Complete, Submit, Create" buttons is very important when first setting up your database. This will save you time in the long run.
- Do you have any tips or tricks to share with the class?
  - Copy/Paste is the best way to ensure that all column headers, table names, etc. match between your source files and the staging tables.

- For me it was easier to complete similar steps all at one time, focusing on one type of task at a time vs. jumping back and forth. For example, I created all stages I needed at the same time, then created all tables, then loaded all tables. This was slightly different than the process the video used.
  - When creating tables, I created all of the column names, then went down the list and assigned the Types (this also allowed me to copy/paste a type into columns with the same types- example VARCHAR(255)), which helped to save time and ensure accuracy.
- If you ever need the entire name/path of a table, you can double click on it in the left list in worksheet view and it will populate in your worksheet.