GOM4DW to Star Schema Conversion

To run this tool, please execute **WelcomeScreen.java** if you want to work on a new project. If you want to launch an already existing project then execute **LoginScreen.java** and provide the project name as **P4.** The java files are located in Java Tool **\src\com\FinalInfo**

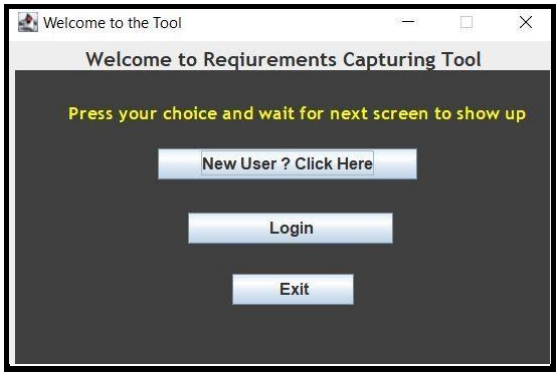
The database P4 is exported to an excel file named **P4.xlsx** and this is included in the project folder.

To import this excel file as a database, please refer to the following link.

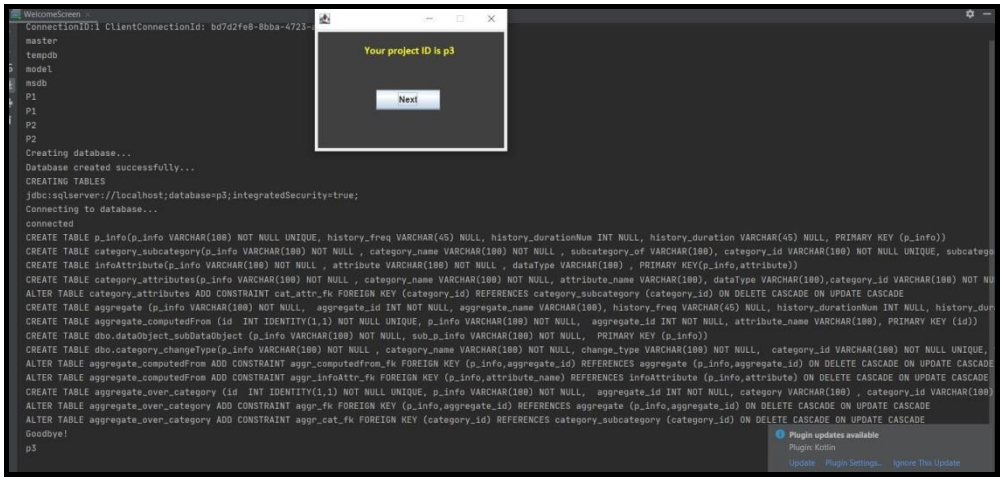
https://www.sqlserverlogexplorer.com/import-and-export-database/

How was the tool used to capture the GOM4DW data and category objects?

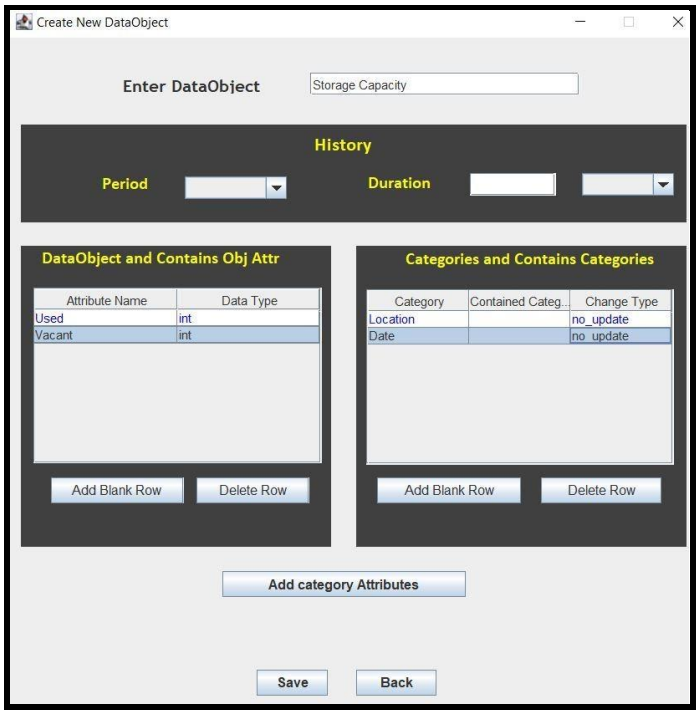
Step 1: When the welcome screen was launched, the following GUI showed up.



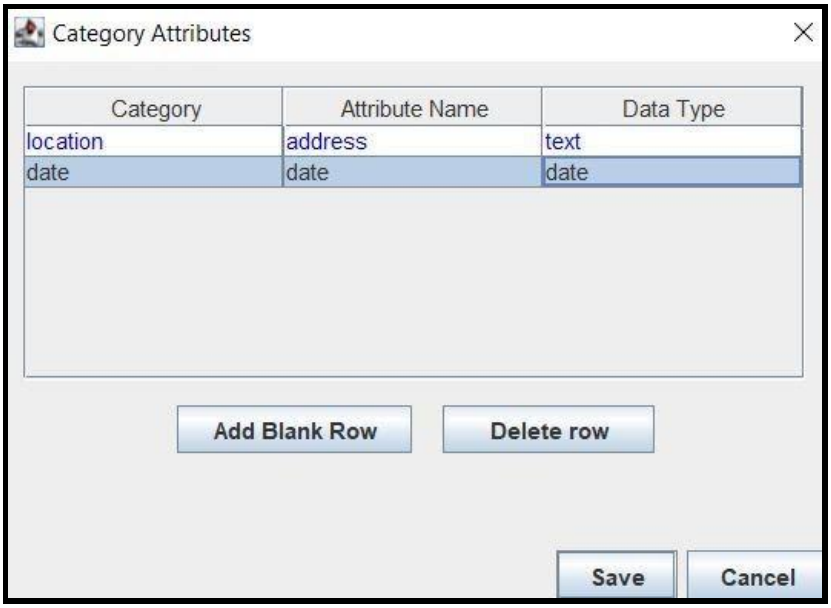
Step 2: Once we click on "New User? Click Here", the tool generates all the tables required to capture the data and category objects.



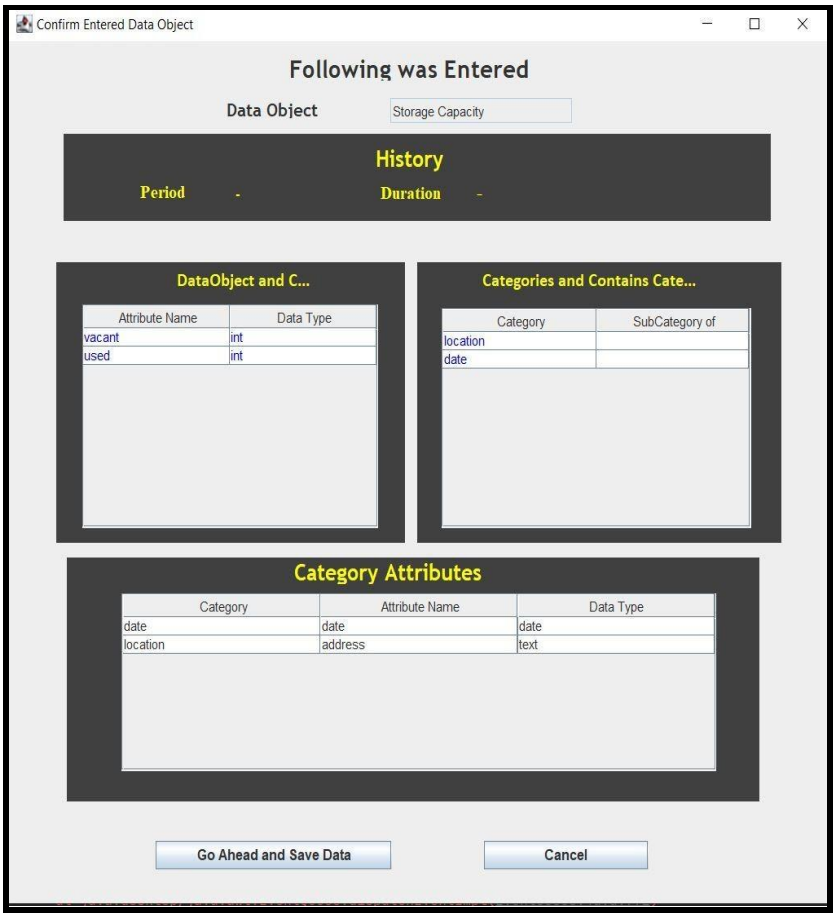
Step 3: Next. click on “Create new object” and make entries for all data objects, their attributes, categories, sub-categories (if any) and their change types.



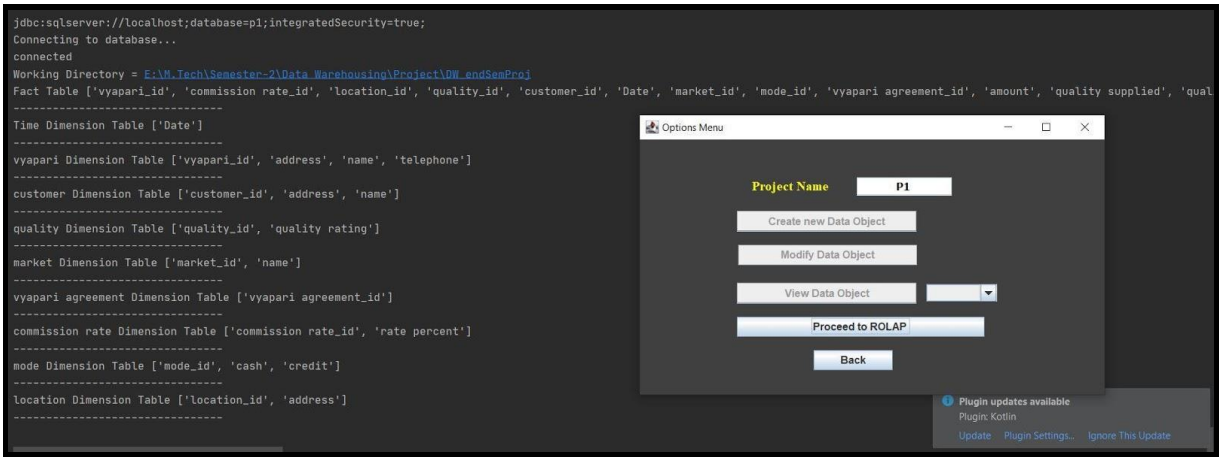
Step 4: We then click on add category attributes to add the attributes corresponding to each category.



Step 5: Once we click on the “Save” button, the GUI prompts for confirmation. On successful insertion, the GUI shows the success message “Saved successfully”.



Step 6: Proceed to ROLAP triggers button on the project home screen triggers the python script and ROLAP schema is generated as output.



How Data object within a data object was handled?

A datatype “Data Object” was added in the list of datatypes associated with attributes of a data object. This differentiates a data object from other attributes and this condition was used to check if another data object was contained in a data object.

How change type was handled?

Change type was added as a field in the “Create User Info” screen and its corresponding table was created with the schemas as follows.

dbo.category_changeType(p_info varchar, category_name varchar, attribute_name varchar, change_type varchar, category_id varchar, time_stamp datetime, primary_key(category_id,attribute_name));

“change_type” was a combo list with values “no_change”, “type-1”, “type-2”,”type-3”. The default value of “time_stamp” was set to be CURRENT_TIMESTAMP and that of “change_type” was “type-1”

An entry in the change_type table is shown below.

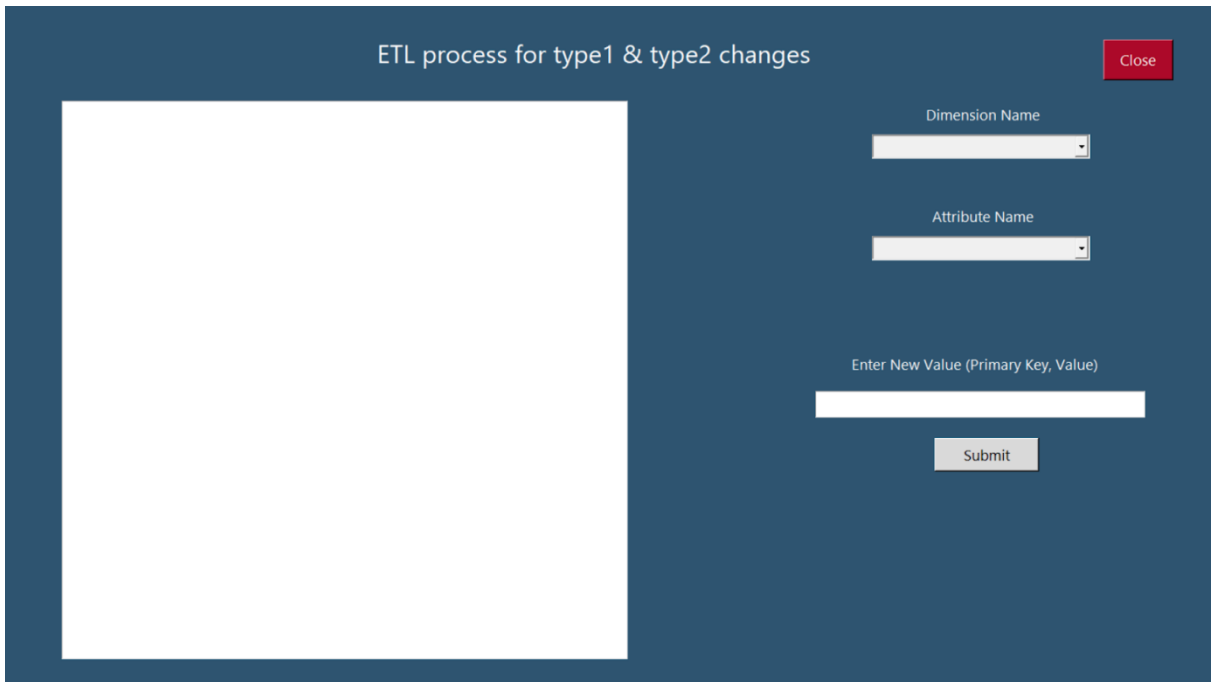
	p_info	category_name	attribute_name	change_type	category_id	time_stamp
1	Accomodate Teams	team	allocated hotel	type-1	Accomodate Teams_team	2020-05-22 21:50:23.990
2	Accomodate Teams	team	city	type-1	Accomodate Teams_team	2020-05-22 21:50:23.990
3	Accomodate Teams	team	group name	type-1	Accomodate Teams_team	2020-05-22 21:50:23.990
4	Accomodate Teams	team	head count	type-1	Accomodate Teams_team	2020-05-22 21:50:23.990
5	Accomodate Teams	team	team name	type-1	Accomodate Teams_team	2020-05-22 21:50:23.990
6	Expenditure	city	city name	type-1	Expenditure_city	2020-05-22 21:50:23.990

ETL Tool Functioning:

For the ETL (Extract, Transform, Load) process, a GUI (Graphical User Interface) is created for easy and convenient operations. It is created to handle the changes for Type-1 and Type-2 attributes that are present in schema.The GUI is implemented using Python with the help of tkinter library.

To open the ETL tool:

1) Open file ETL Tool -> etl.py file in Python console
2) Run the file using Python command
3) Following screen will appear



Brief Description:

(For detailed description and working, refer to DW_Report.pdf)

GUI includes 2 drop downs named as 'Dimension Name' which displays all the dimensions present in the data warehouse for a given problem and another as 'Attribute Name' which displays all the attributes with their type (type I or type II) for the selected dimension in the first drop down.

As soon as the attribute is selected in the second drop down, the selected dimension and attribute name is processed and the corresponding table is displayed in the white text box placed in the left part of the GUI screen.

The value which needs to be changed is entered in the "Enter New Value" field with the format of (Primary Key, New Value) where primary key is the key of the record which we want to change. After entering all the details, on clicking the submit button a message box is shown which notifies the successful or unsuccessful update procedure and the left textbox is updated to reflect the changes.

Note: For Boolean values update through ETL, please enter 1 for "True" and 0 for "False"

ASSUMPTIONS:

- “Change_type” can be set only at the time of creation of category in GUI and its addition to the database.
- “Data Object” is used as a datatype to identify data objects within a data object.
- If history is to be maintained for any data object, then the attributes **'day', 'week', 'month', 'quarter', 'year'** automatically get added to the time dimension in order to support ROLAP operations.
- All date and day related attributes are assumed as date and added to the time dimension.
- For Boolean values update through ETL, please enter 1 for “True” and 0 for “False”