Getting started with g9 Anonymizer

g9 Anonymizer

Version 1.2.0

August 12, 2020

Table of Contents

1 How does it work?

In this getting started section, you will use the 'Anonymizer Hotel Example' project to explore several of the g9 Anonymizer capabilities. This example comes with an embedded Derby database 'hotelsystem' that is automatically created in the wizard which runs various SQL scripts to create the database, indexes and insert sample data. This example project does not represent a real-life example as it has a simplified data structure with a small data set, but it contains various tasks already defined which contains data anonymizations, data deletions and data creations.
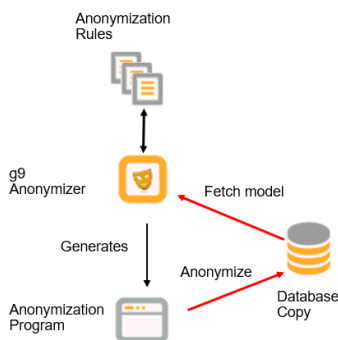
The figure shows how g9 Anonymizer works:

- **Database Copy**: Create a copy of your production database.
- **Fetch model**: Import the data structure from a running database (**Database Copy**) into g9 Anonymizer.
- **g9 Anonymizer**: The data structure displayed in an editor with database information and rule definitions.
- **Anonymization Rules**: Definition of your tasks, be it anonymization rules, inspection reports and/or deletions.
- **Generates**: The Java **Anonymization Program** application will be automatically created. You can override code.
- **Anonymize**: Run the program against the database that you are going to anonymize. You will get a database that can be used for testing and development purposes.

2 Installation and License

2.1 Requirements

Before working with the g9 Anonymizer:

The g9 Anonymizer tool-set consists of code that runs on Windows and Linux operating systems. The tool set runs as Eclipse plug-ins and requires 2018-09 or newer versions of Eclipse.Start with the *Eclipse IDE for Java and DSL Developers* package. Otherwise the Xtext and Xcore plugins must be installed prior to the g9 Anonymizer installation.

Some Eclipse packages (**Eclipse IDE for Java and DSL Developers**) already have Xcore and Xtext installed. If you choose **Eclipse IDE for Java Developers** or **Eclipse IDE for Enterprise Java Developers**, it is not. To install Xcore/Xtext, start Eclipse and select *Help > Install New Software …*, which opens the *Available Software* dialog. Select the eclipse version update site (2019-12) in the *Work with:* combo box and type *Xcore* in the filter. Select the **EMF - Eclipse Modeling Framework Xcore SDK** version shown (1.12) and install it. Do the same for Xtext, type *Xtext* in the filter and select the **Xtext Complete SDK** (2.20) and install it.

- Install Eclipse IDE
- The g9 Anonymizer comes with pre-installed JDBC drivers settings using Maven GAV settings for a set of pre-selected database systems. The Derby database is pre-installed. The other drivers need installation, use the [Install] button or use the [Edit] button to configure supported drivers or to add new (See Maintain JDBC Drivers).
- The Maven build process requires a JDK version 8 or later. Make sure to set the default Java execution environment to a valid JDK installation. (Window → Preferences → Java → Installed JREs).
- Install g9 Anonymizer trial or full license.

2.1.1 Xcore/Xtext versions

For Xcore installation, the following plugin versions have been tested:

Eclipse version	Xcore	Xtext
2020-06-R	1.14.0	2.22.0
2020-03-R	1.13.0	2.21.0
2019-12-R	1.12.0	2.20.0
2019-09-R	1.11.0	2.19.0
2019-06-R	1.10.0	2.18.0
2019-03-R	1.9.0	2.17.0
2018-12-R	1.8.0	2.16.0

2.2 Installation from update site

The g9 Anonymizer is available from the update site: **http://www.esito.no/updatesite/anonymizer**.

1. Start Eclipse and select from the main menu *Help > Install New Software…*

2. In the *Install* dialog, click the *Add…* button located at the top right of the dialog to add an install site. This brings up the *Add Repository* dialog
3. You can install the most recent version from the update site.
    a. In the *Name* field, type **g9 Anonymizer**
    b. In the *Location* field, type **http://www.esito.no/updatesite/anonymizer**
    c. Click OK to return to the Install dialog.
4. This should automatically select **g9 Anonymizer** in the *Work with* field.
5. Once *Esito Modeling tools* is displayed, select the checkbox and then select the *Next* button.
6. Click *Next* again to confirm installation.
7. Read and accept the license agreement. To continue installing, select the checkbox "*I accept the terms of the license agreement*" and click *Finish*.
8. When prompted to restart Eclipse, click *Restart Now* to restart.
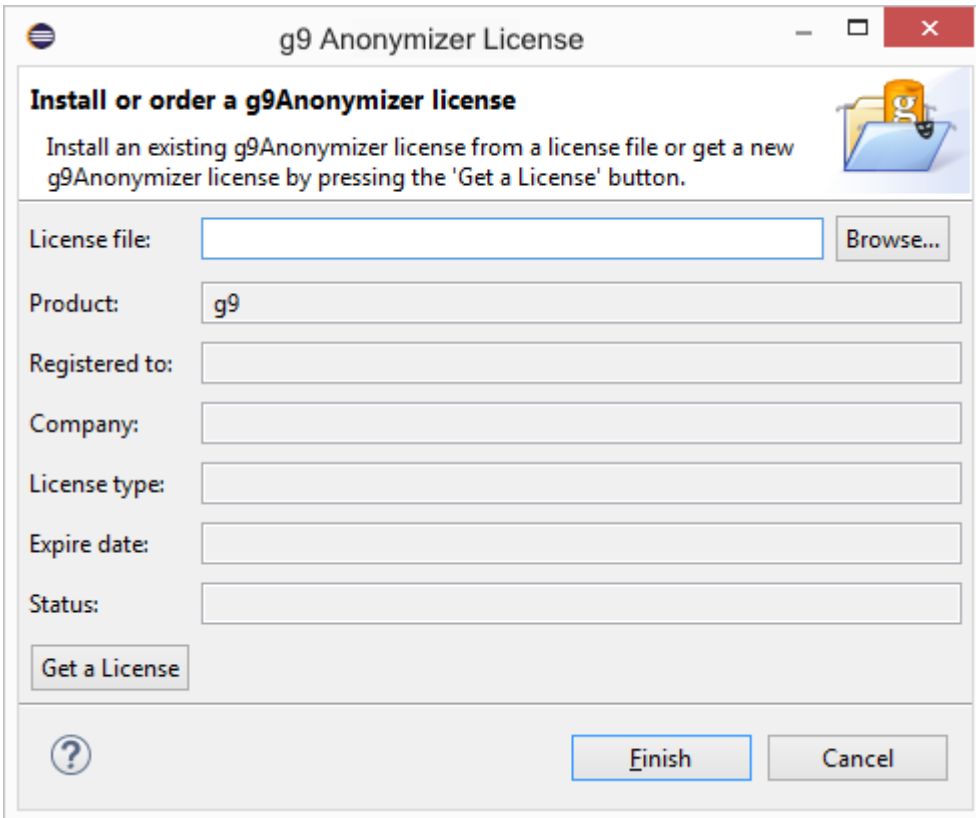
2.3 License

g9 Anonymizer requires a valid license, either a trial or a full license. If a license is not installed, a trial license will automatically be installed the first time Eclipse starts and a g9Anonymizer project is created or refreshed. A Message telling about the installation is displayed.
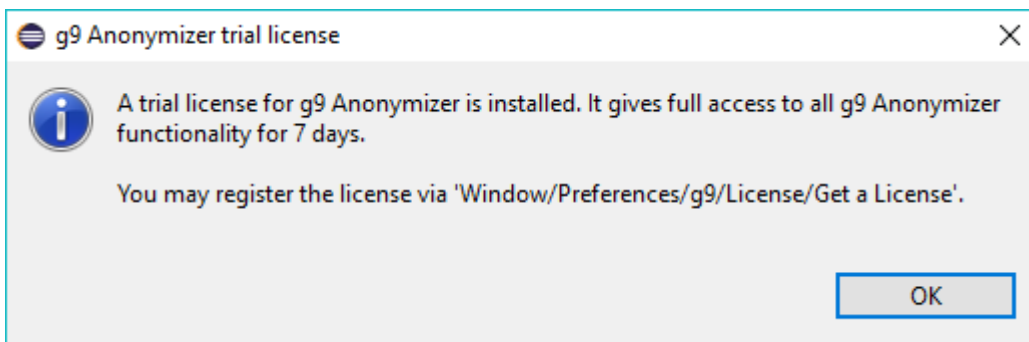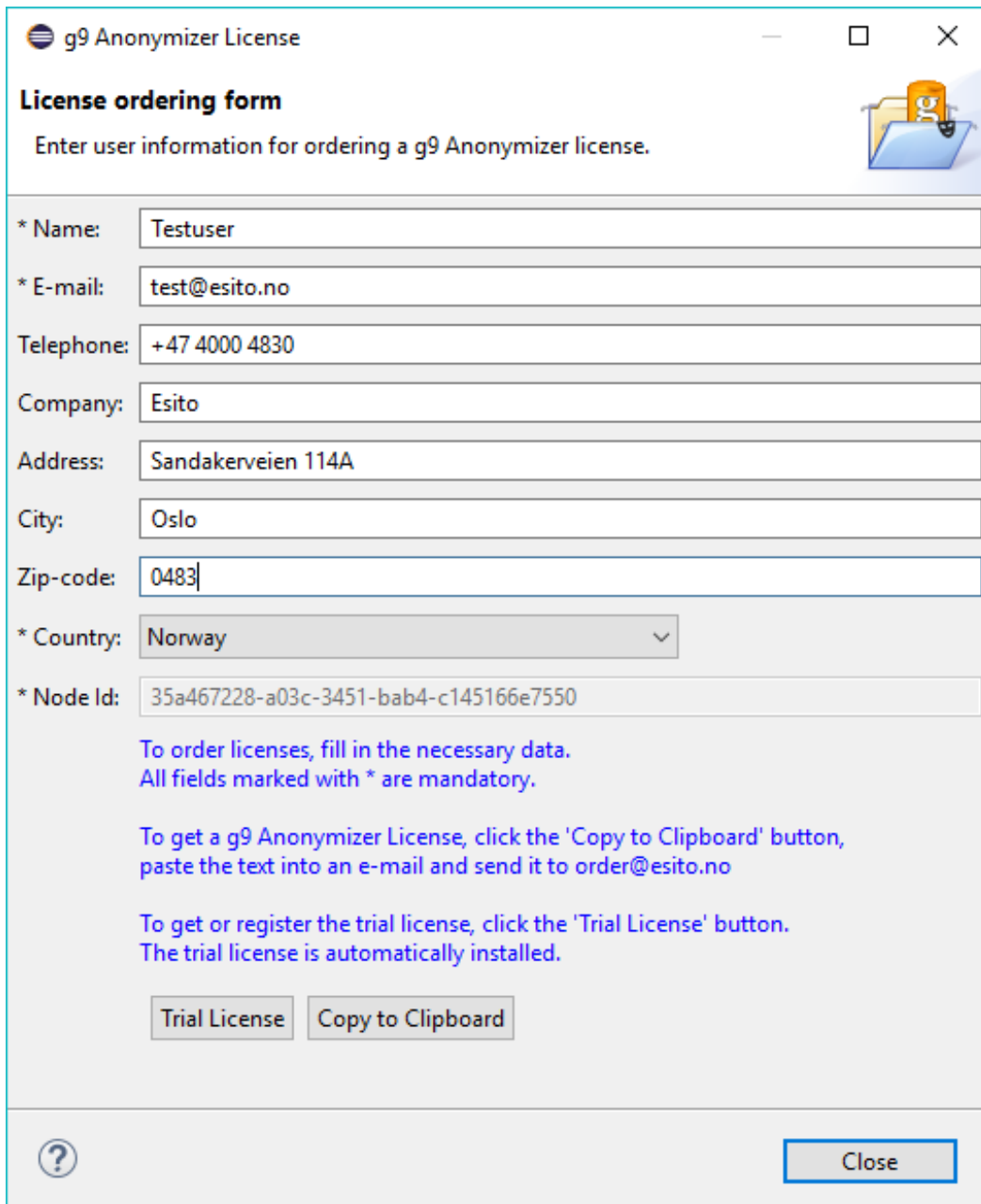

If a license is installed, but is illegal for some reason, an "Install g9 Anonymizer license" message will be displayed the first time Eclipse starts and a g9Anonymizer project is refreshed.
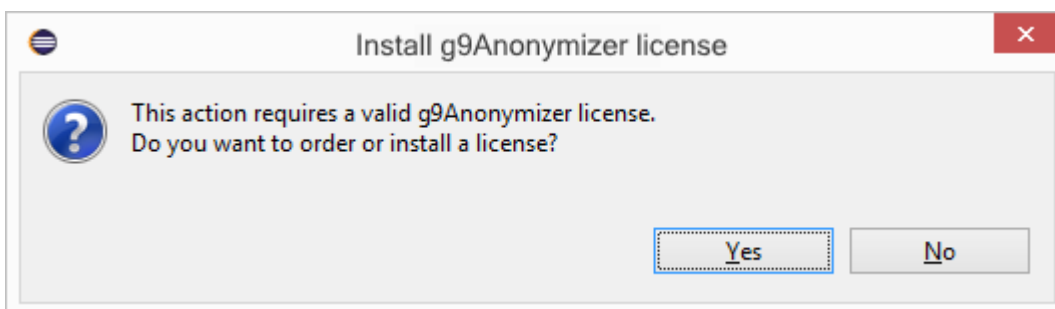
Click *Yes* to order either a full or trial license.

If you have received a license file (l4j extension), click *Browse* and select the file in the dialog and click Open. The license fields are displayed including the Expiry date. Click the *Finish* button to save and close the license dialog.

If you do not have a License or if you want to register your automatically installed trial license, click *Get a License.*

g9 Anonymizer License

License ordering form

Enter user information for ordering a g9 Anonymizer license.

* Name:	Testuser
* E-mail:	test@esito.no
Telephone:	+47 4000 4830
Company:	Esito
Address:	Sandakerveien 114A
City:	Oslo
Zip-code:	0483
* Country:	Norway
* Node Id:	35a467228-a03c-3451-bab4-c145166e7550

To order licenses, fill in the necessary data.
All fields marked with * are mandatory.

To get a g9 Anonymizer License, click the 'Copy to Clipboard' button, paste the text into an e-mail and send it to order@esito.no

To get or register the trial license, click the 'Trial License' button.
The trial license is automatically installed.

[Trial License] [Copy to Clipboard]

[Close]

Fill in the fields. (The fields beginning with an asterisk are mandatory) and then click the *Trial License* button.

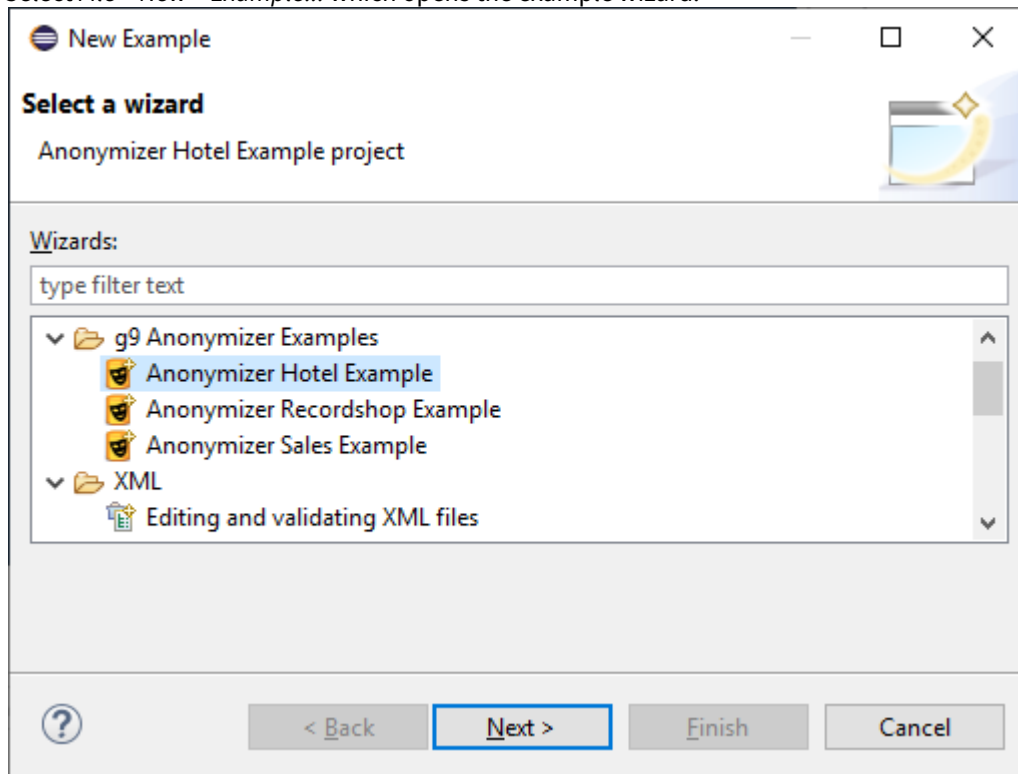**The trial license provides full functionality for 7 days.**

3 Getting started

You should start investigating g9 Anonymizer example project in order to understand the functionality of the tools.

3.1 Open Anonymizer Hotel Example project

To open the Anonymizer Hotel Example:

1. Select *File > New > Example...* which opens the example wizard:

2. Choose **Anonymizer Hotel Example**. Click *Next* and *Finish* . The Anonymizer project **AnonymizerHotel** is displayed as a node in the Project Explorer view.

3.2 Suggested perspective and views

You may organize the eclipse perspective to suit your preferences. Necessary Views are:

- Package Explorer or Project Explorer
- Properties
- Console
- Problems

To organize your perspective:

1. Select the Open Perspective button on the toolbar and select the Java perspective.
2. Select the menu *Window > Show View .. Other.* In the Show View dialog, expand the General node and select the *Properties, Console and Problems* views. (You can hold down the Ctrl button to select multiple views). Press the *Open* button to select the views.
3. Select the menu *Window > Perspective >Customize Perspective ...*, go to the *Shortcuts* tab, select the *Anonymizer Project* menu and click *Apply and Close*.
4. Select the menu *Window > Perspective >Save Perspective As*. In the dialog enter the name *Anonymizer* and click the Save button.

A fully working example project is imported and opened in the Package or Project Explorer. Expand the project in the Explorer and double-click the 🎭 **AnonymizerHotel.anonymizer** file. This opens the anonymizer editor with the two editor panes as shown below.
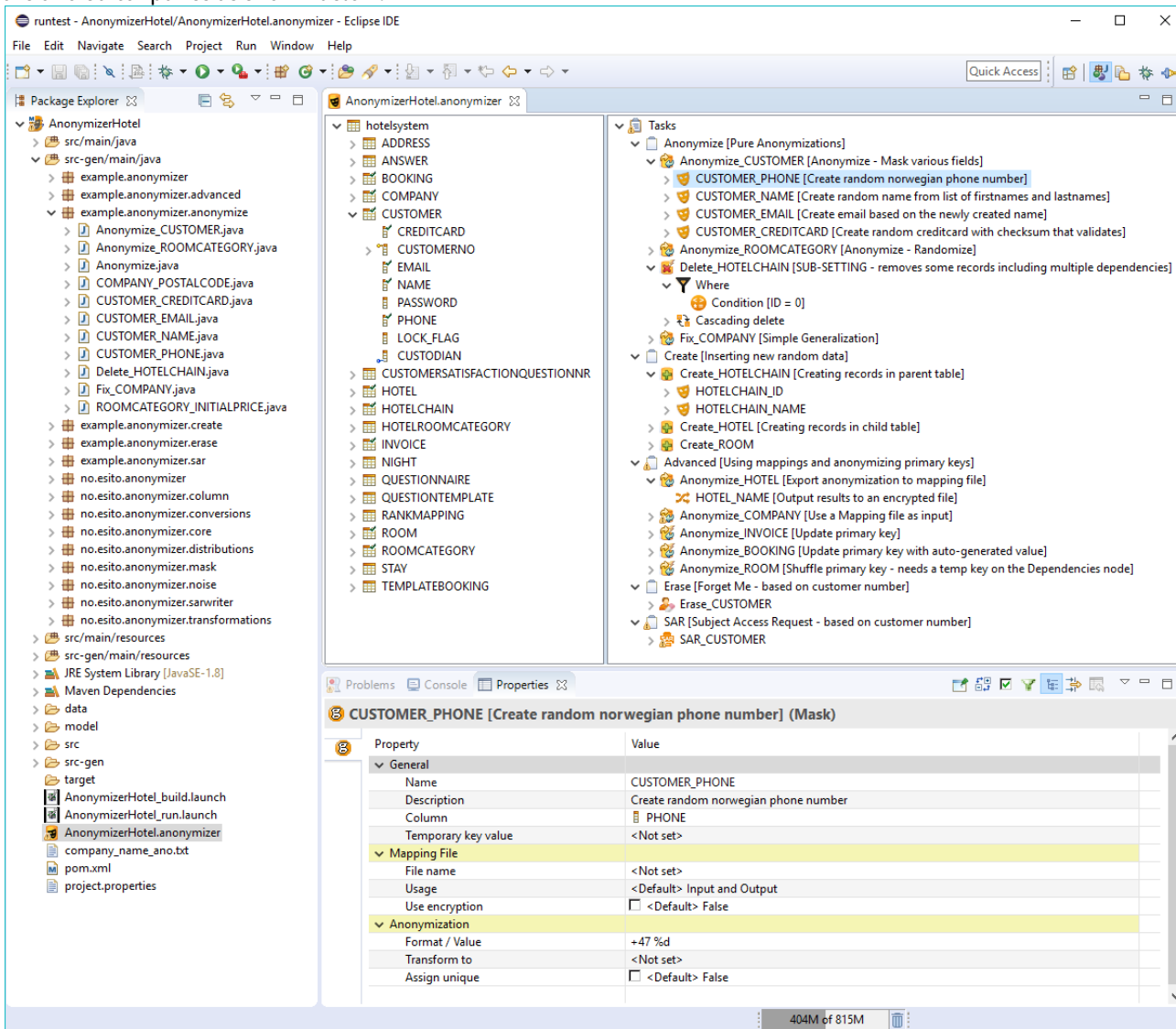
The following views are used by the g9 Anonymizer:

Package Explorer or **Project Explorer:** Shows the anonymizer project which contains the following:

- Anonymizer file which defines the anonymizing rules **(AnonymizerHotel.anonymizer)**
- Generated java source code **(src-gen/main/java)**
- Maven **pom.xml** for the project
- Launch files that can build and start generated programs (**AnonymizerHotel_build.launch** and **AnonymizerHotel_run.launch**)
- Derby example database with data and sql files **(data)**

Properties: Shows detailed information about selected elements in the editor.

Console: Shows logging of generator runs and can be used to enter command prompts.

Problems: Shows Errors and Warnings

Editor: Shows Anonymizer editor containing 2 panes. The left pane contains a tree view of the database structure with tables, columns and foreign keys (Database pane). The right pane has a tree view with tasks containing the anonymizing rules (Task pane).

3.3 Exploring the tasks

In this section, you will explore the masking rules and add a condition to an existing masking rule within a task.

1. In the task pane, expand the *Anonymize_CUSTOMER* node, expand the *Customer_PHONE* node and select it.
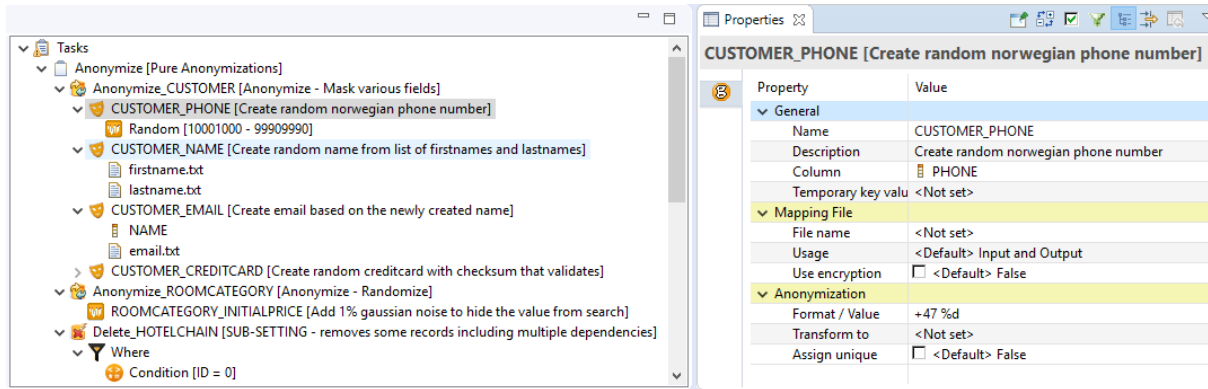

Note that the *CUSTOMER_PHONE* node has a child node *Random* and that the property value for the Format/ Value is **+47 %d**. This is a constant value (+47) followed by the substitution variable (%d). This causes the data in the PHONE column to be replaced with a random integer (between 10001000 and 99909990) prefaced with the constant +47.

2. Expand the *CUSTOMER_NAME* node and you will see two child nodes: **firstname.txt** and **lastname.txt**. Select the firstname.txt node and press the F3 key and it will open the file in an editor. These files contain lists of first names and last names respectively. Click on the *CUSTOMER_NAME* node and you will see the *Format /Value* of **%s %s**. This instructs the system to replace the data in the NAME column with random first and last names extracted from the specified text files.

3. Expand the other nodes to view the rules. The *CUSTOMER_EMAIL* will use the masked value of the NAME column and append it with a random value obtained from the email.txt file.

4. We will now look at the data in the CUSTOMER table. Right-click the CUSTOMER table in the Database pane and select the menu option '*Show Table'.* This opens a 'Show Table' view displaying all columns and values.

5. You will now add a Where condition for the masking of the Customer columns. In the Task pane, right-click on the *Anonymize_CUSTOMER* node and select *Add > Where.* This creates a *Where* node. Right-click on the *Where* node and select *Add > Condition.*
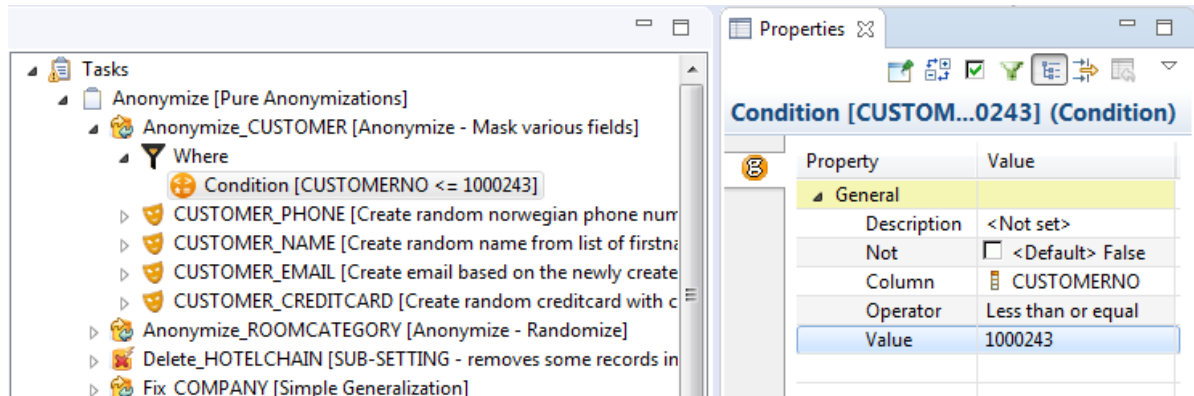
6. Click on the Condition node and enter the following in the properties page:

Column	CUSTOMERNO
Operator	Less than or equal
Value	1000243

7. Save the editor by pressing Ctrl + S. You have now added a where condition that limits the customer records that will be anonymized.

3.4 Building the project and running a task

Each time changes are saved to the anonymizer file, the generator starts and generates a complete program reflecting all your defined rules. (If the Project > Build Automatically is selected) The Console View displays the output of the generation process.

To generate the anonymization java files and run a task, do the following:

1. Select the AnonymizerHotel project
2. Select *Clean…* from the *Project* menu and press the *Clean* button.
3. The application code is regenerated and you can inspect it in the Project Explorer. Expand the `src-gen/ main/java` directory and open the `example.anonymizer.anonymize` package. If you would open the Anonymize_CUSTOMER.java file, you would notice the `getWhere()` method that defines the Where condition that you added.
4. A pre-defined launch script runs the generated program, simply right-click the **`AnonymizerHotel_run.launch`** file and select ***Run As > AnonymizerHotel_run***.
5. The program is started and starts the Anonymizer Console in the Console View.
6. Place the cursor at the > and write the command **trace** and press the Enter key.
7. Then enter **run anonymize** (the task name) and press the Enter key. This command performs all the anonymizations specified under the task *anonymize* and the program displays the sql commands performed on the database.

```
Console ␦                                     ■ ✕ ✕ | ␦ ␦ ␦ | ␦ ␦ | ␦ ␦ ▾ ␦ ▾ ▬ ␦
AnonymizerHotel_run [Java Application] C:\Program Files\Java\jre1.8.0_151\bin\javaw.exe (Feb 22, 2018, 1:40:36 PM)
>trace
>run anonymize
>>> START 13:40:53(Task commit)
Running example.anonymizer.anonymize.Anonymize
>>>    Anonymize
>>>       Anonymize_CUSTOMER
Reading internal config.properties
SELECT CUSTOMERNO, PHONE, NAME, EMAIL, CREDITCARD FROM APP.CUSTOMER WHERE CUSTOMERNO <= 1000243
UPDATE APP.CUSTOMER SET PHONE = ?, NAME = ?, EMAIL = ?, CREDITCARD = ? WHERE CUSTOMERNO = ?
['+47 51289513','Evie Ellis','evie.ellis@gmail.com','4142834051119496',1000234]
['+47 87236119','Evelyn Johnston','evelyn.johnston@inmeta.no','4142834087046101',1000235]
['+47 77004074','Hannah Lewis','hannah.lewis@yahoo.com','4142834076954067',1000236]
['+47 53391573','Landon Hodges','landon.hodges@crayon.com','4142834053251552',1000237]
['+47 30209650','Dangelo Vega','dangelo.vega@crayon.com','4142834030059631',1000238]
['+47 44319116','Martha Hopkins','martha.hopkins@crayon.com','4142834044249103',1000239]
['+47 23238608','Sarah Harvey','sarah.harvey@crayon.com','4142834023108593',1000240]
['+47 81452779','Lucas Gonzales','lucas.gonzales@esito.no','4142834081432778',1000241]
['+47 34493845','Isabella Preston','isabella.preston@esito.no','4142834034353832',1000242]
['+47 69972016','Lilliana Saunders','lilliana.saunders@msn.com','4142834069791997',1000243]
<<<       Anonymize_CUSTOMER (252ms)
>>>       Anonymize_ROOMCATEGORY
SELECT ID, INITIALPRICE FROM APP.ROOMCATEGORY
UPDATE APP.ROOMCATEGORY SET INITIALPRICE = ? WHERE ID = ?
[603.81,1]
[771.98,2]
[591.05,3]
[502.81,4]
[5251.20,5]
[668.55,6]
[1779.51,7]
[2693.10,8]
[2689.47,9]
[3394.02,10]
[999.52,11]
[10253.18,12]
<<<       Anonymize_ROOMCATEGORY (55ms)
>>>       Delete_HOTELCHAIN
SELECT ID FROM APP.HOTELCHAIN WHERE ID = 0
SELECT ID FROM APP.HOTEL WHERE CHAIN_ID = 0
SELECT ID FROM APP.BOOKING WHERE HOTEL_ID = 16
UPDATE APP.CUSTOMERSATISFACTIONQUESTIONNR SET VISITEDHOTELS_ID = NULL WHERE VISITEDHOTELS_ID = 16
DELETE FROM APP.HOTELROOMCATEGORY WHERE HOTEL_ID = 16
SELECT ID FROM APP.ROOM WHERE HOTEL_ID = 16
SELECT SERIALNO, CUSTOMER_CUSTOMERNO FROM APP.STAY WHERE LOCATION_ID = 16
DELETE FROM APP.TEMPLATEBOOKING WHERE HOTEL_ID = 16
DELETE FROM APP.HOTEL WHERE CHAIN_ID = 0
DELETE FROM APP.HOTELCHAIN WHERE ID = 0
<<<       Delete_HOTELCHAIN (317ms)
>>>       Fix_COMPANY
SELECT ORGANIZATIONID, POSTALCODE FROM APP.COMPANY
UPDATE APP.COMPANY SET POSTALCODE = ? WHERE ORGANIZATIONID = ?
[1000,123456789]
[3000,234567890]
[3000,345678901]
[5000,456789012]
[9000,567890123]
[7000,678901234]
[1000,789012345]
[9000,846295628]
[3000,984567126]
<<<       Fix_COMPANY (34ms)
<<<    Anonymize (665ms)
<<< FINISHED 804ms
```

8. Enter quit to stop the console program.
9. Inspect the data in the database. You can use the Refresh button if you already have a CUSTOMER tab open, otherwise right-click the CUSTOMER table and select the *Show Table* menu. Note all of the changes are performed on PHONE, NAME, EMAIL and CREDITCARD for your specified range of customers.

🖥 Console	🐞 Problems	🏛 CUSTOMER ✕				
CREDITCARD	CUSTOMERNO	EMAIL	NAME	PASSWORD	PHONE	CUSTODIAN
4142834051119...	1000234	evie.ellis@gm...	Evie Ellis	<null>	+47 51289513	<null>
4142834087046...	1000235	evelyn.johnsto...	Evelyn Johnston	<null>	+47 87236119	<null>
4142834076954...	1000236	hannah.lewis...	Hannah Lewis	<null>	+47 77004074	<null>
4142834053251...	1000237	landon.hodge...	Landon Hodges	<null>	+47 53391573	<null>
4142834030059...	1000238	dangelo.vega...	Dangelo Vega	<null>	+47 30209650	<null>
4142834044249...	1000239	martha.hopki...	Martha Hopkins	<null>	+47 44319116	<null>
4142834023108...	1000240	sarah.harvey@...	Sarah Harvey	<null>	+47 23238608	<null>
4142834081432...	1000241	lucas.gonzales...	Lucas Gonzales	<null>	+47 81452779	1000234
4142834034353...	1000242	isabella.presto...	Isabella Preston	<null>	+47 34493845	<null>
4142834069791...	1000243	lilliana.saunde...	Lilliana Saund...	<null>	+47 69972016	<null>
4142834096149...	1000244	tristan.becker...	Tristan Becker	<null>	+47 96219140	1000240
4142834013319...	1000245	desiree.becker...	Desiree Becker	<null>	+47 13379731	<null>
4142834027917...	1000246	kaya..garner@...	Kaya Garner	<null>	+47 28007554	<null>
4142834067619...	1000247	estella.morriso...	Estella Morrison	<null>	+47 67759388	1000234
4142834099651...	1000248	isabella.joyce...	Isabella Joyce	<null>	+47 99851264	<null>

ⓘ **Derby database server**

The AnonymizerHotel example project uses a Derby database. The database is embedded in the project at `data/hotelsystem` and not running after eclipse start/restart. The first database access will start the database. Be aware that the Derby database server will keep the database locked until Eclipse is restarted. That means you cannot delete the example project if you have already looked at the database in the project. To stop the server without an eclipse restart, right-click the AnonymizerHotel project and select *g9 Anonymizer > Shutdown Derby* menu. Note that the console program must be exited.

The anonymizer example program changes the database content.

If you want to start with a fresh database, select Shutdown Derby, delete the database (directory) `hotelsystem` and create it with *g9 Anonymizer > Create Derby.*

3.5 Adding your own anonymization task

In this section you will add a new anonymization task to generate new names for existing hotels.


ⓘ **g9 Preference**

These steps assume that the preference 'Prompt for name on added objects' is selected. Select the menu Window > Preferences and click the g9 > Design node. Check the preference and then click the Apply and Close button.

---

1. To see the hotel names prior to the anonymization, right-click the HOTEL node in the database pane and select 'Show Table'. A new 'Hotel' tab is created

ID	LOCATION	LOGO	NAME	CHAIN_ID	LOCK_FLAG
1	Waterside meact	Logo1	Antique Canyon Hotel	1	<null>
2	West kainnict	Logo2	Ivory Refuge Hotel & Spa	1	<null>
3	South tasiob	Logo3	Historic Heritage Resort & Spa	1	<null>
4	Andebu	Logo4	Sunny Jungle Resort & Spa	1	<null>
5	Waterside gramict	Logo5	Sunny Ribbon Resort	1	<null>
6	Bergen	Logo6	DeLuxor	1	<null>
7	Tronheim	Logo7	Da Hood	1	<null>
8	Nuword Garden	Logo8	Royal Ocean Hotel	1	<null>
9	Cunnic Center	Logo9	Grand Hotel	1	<null>
10	Lillehammer	Logo10	Prism Hotel	1	<null>
11	Sandvika	Logo11	The House of Dilbert	1	<null>

2. You will now add the Anonymization mask for the hotel NAME attribute:
   a. Select the "Anonymize" task in the task pane. This ensures the added anonymization is added to that task.
   b. Expand the HOTEL node in the database pane and right-click the NAME node.
   c. Choose Anonymize Column > Mask
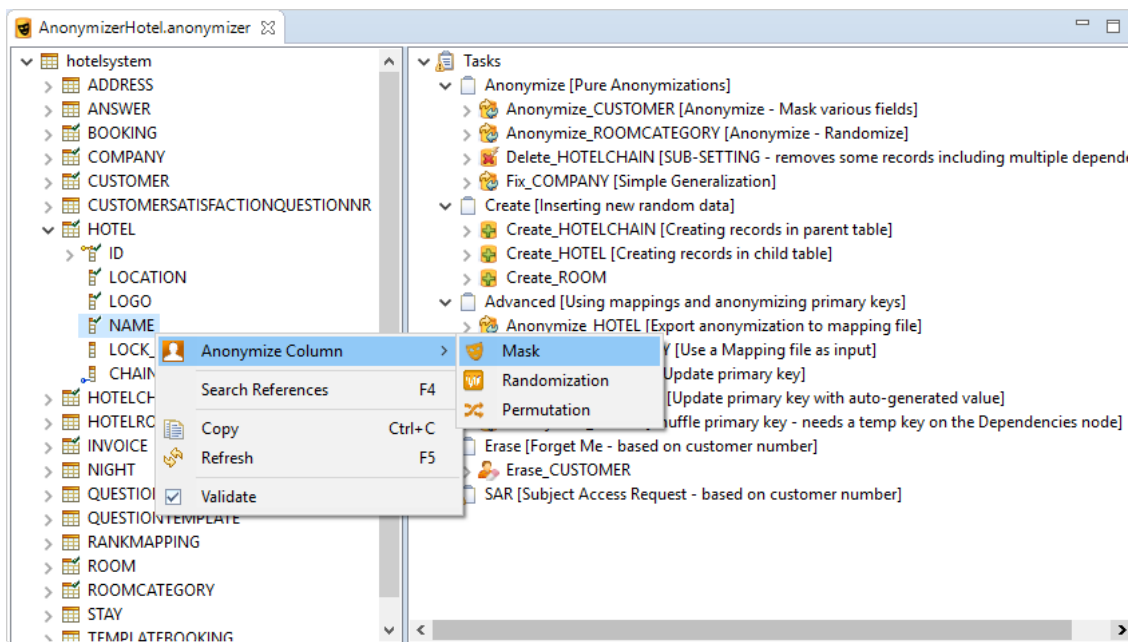

   d. Enter "FixHotelName" and press OK.
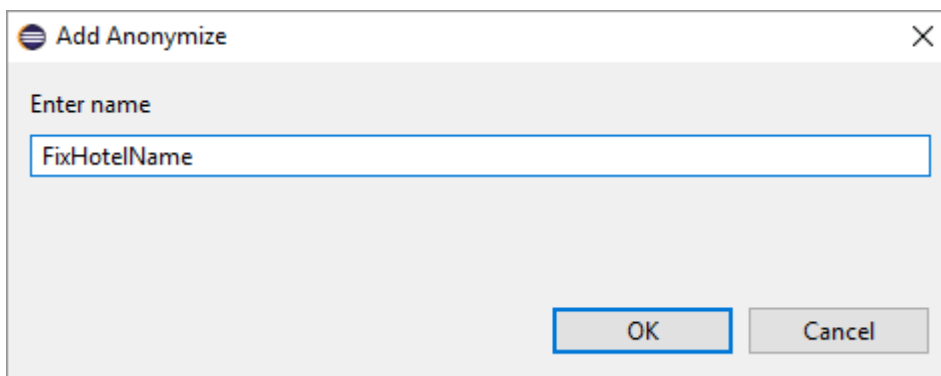
  e. Accept the default HOTEL_NAME name and press OK.

3. Next you will define the Hotel naming rule. This rule will replace the name with a random name from a text file and add the suffix "International Hotel":
  a. Click the HOTEL_NAME node in your new task.
  b. In the Property view enter the following

Property	Value	Description
Format / Value	%s International Hotel	The %s variable is replaced with the random text entry followed by the constant text 'International Hotel'
Assign unique	Checked	Checking this option ensures that all entries are unique. (Unless the number of records exceeds the number of text entries in the file)

  c. Right-click the HOTEL_NAME node and choose Add > File
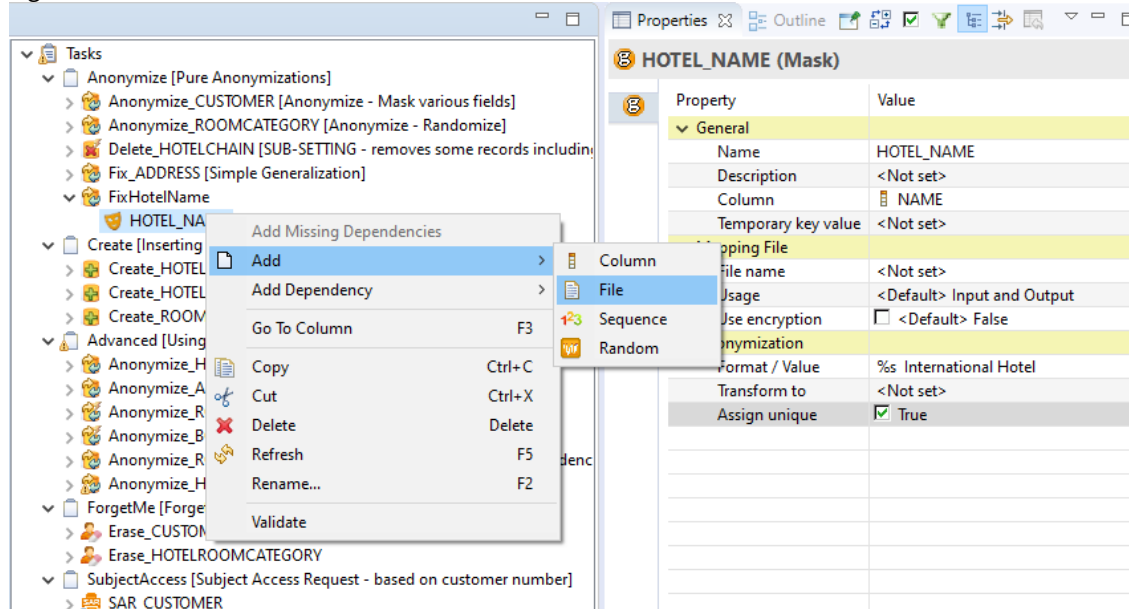
4. Next you will specify the file "lastname.txt" containing multiple text entries of names.
  a. Click the new <File> node
  b. In the Property view, click the File property and click the browse button at the far right. This opens up a Select File dialog.
  c. Select the file *src/main/resources/lastname.txt* and press OK.
  d. Save your changes to rebuild the application. (This assumes the option Project > Build Automatically is checked. If not, select Project > Clean and then click the Clean button)
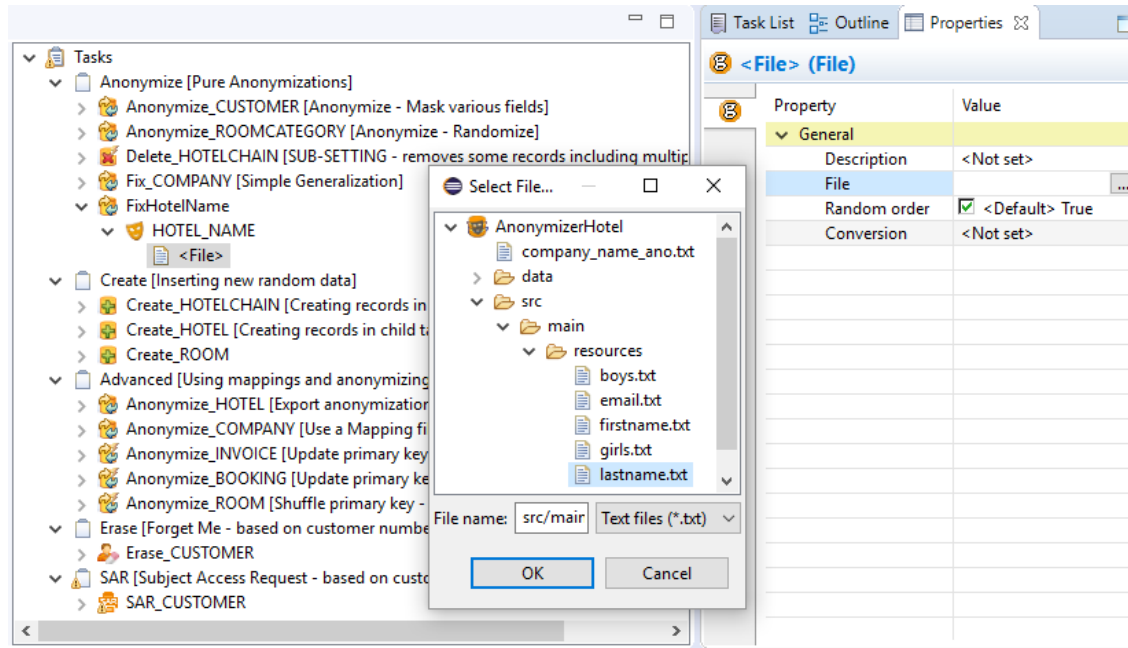
5. You will now run the new anonymization task within the Console view.
   a. Right-click on the "AnonymizerHotel_run.launch" in the project explorer
   b. Choose Run As > AnonymizerHotel_run. The Console view is displayed.
   c. Click in the Console view to position the cursor in front of the > prompt and type **run FixHotelName**
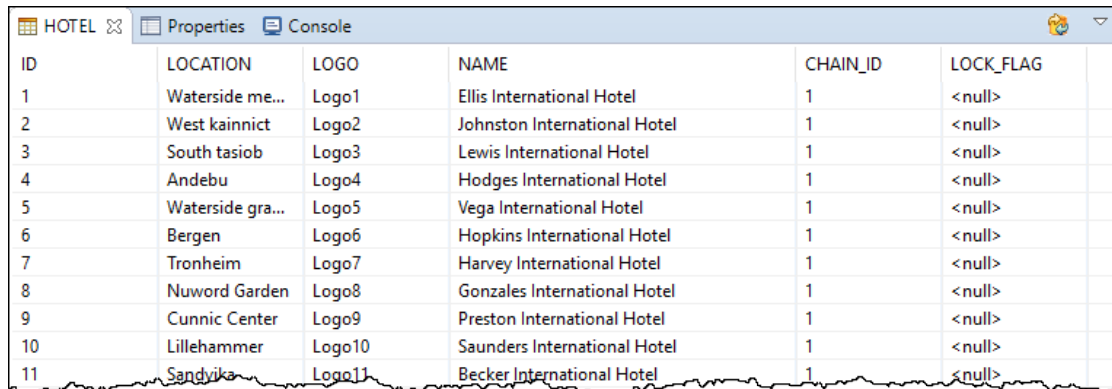   d. Hit the enter key.

6. To verify the results of running your new anonymization task:
   a. If the HOTEL tab is already displayed, click the Refresh icon located at the top right of the view

b. If the HOTEL tab is closed, right-click HOTEL in the database pane and then select Show Table.

HOTEL	Properties	Console			
ID	LOCATION	LOGO	NAME	CHAIN_ID	LOCK_FLAG
1	Waterside me...	Logo1	Ellis International Hotel	1	<null>
2	West kainnict	Logo2	Johnston International Hotel	1	<null>
3	South tasiob	Logo3	Lewis International Hotel	1	<null>
4	Andebu	Logo4	Hodges International Hotel	1	<null>
5	Waterside gra...	Logo5	Vega International Hotel	1	<null>
6	Bergen	Logo6	Hopkins International Hotel	1	<null>
7	Tronheim	Logo7	Harvey International Hotel	1	<null>
8	Nuword Garden	Logo8	Gonzales International Hotel	1	<null>
9	Cunnic Center	Logo9	Preston International Hotel	1	<null>
10	Lillehammer	Logo10	Saunders International Hotel	1	<null>
11	Sandvika	Logo11	Becker International Hotel	1	<null>