

# Documentacion del proceso de la CARGA MASIVA DE DATOS

## Sistemas Transaccionales 2019-01

### Iteracion 3

Luis Miguel Gomez Londono - 201729597  
Juan David Diaz Cristancho - 201729408

En este documento se explica el proceso utilizado para la carga masiva de datos y poblacion de tablas para la realizacion completa de la iteracion 3 de SISTRANS. Primero lo primero, se requeria generar muchos datos (1,000,000+) y existian muchas alternativas para este, sin embargo algunos sin lo que requeriamos.

Los programas de generar datos que existen en internet solo nos dejaban efectuar < 1,000 datos, lo que no nos servia, a pesar de la cantidad de tipos de datos y atributos que se podian usar.

Recurrimos finalmente a la creacion de nuestro propio programa de **Java** para generar estos: la haberlo hecho nosotros nos dejaba generar los datos que quisiéramos y ademas con mas flexibilidad y rapidez. La mayoría de las tablas requerian solo valores numericos que se podian generar aleatoriamente con las librerias nativas de Java como Math.random... Para la creacion de los atributos mas personalizados como nombres, correos electronicos y demas se utilizo la libreria de **Faker**

que nos permitia generar estos datos directamente en el programa. Estos archivos se generaban en formato .csv tanto por la facilidad de crearlos como tambien de los que requeria el SMDBD. Este proceso le tomaba al programa tan solo cuestion de milisegundos.

Para la carga de los datos ya generados se utilizo la opcion de carga de datos de **SQLDeveloper**, esto, por su facilidad y rapidez en la carga de datos. El proceso era

The screenshot shows a GitHub repository page for 'DataGenerator-IT3-SISTRANS' by user 'luismigolondo'. The repository has 4 commits, 1 branch, 0 releases, and 1 contributor. The 'master' branch is selected. The file list shows: bin (Datos mas generales, 2 days ago), docs (Datos mas generales, 2 days ago), lib (Creacion total del generador de dato, 5 days ago), src (Datos mas generales, 2 days ago), and README.md (Cambios rdme, 2 days ago). The README.md file is open, showing the title 'DATALOADER-IT3SISTRANS', a description 'Programa Java para la carga masiva de +1,000,000 de datos a archivo .csv', installation instructions, a 'Generar datos:' section with a note to remove a comment, a 'Comentarios' section, and a footer note about file storage and the Faker library.

File	Description	Time
bin	Datos mas generales	2 days ago
docs	Datos mas generales	2 days ago
lib	Creacion total del generador de dato	5 days ago
src	Datos mas generales	2 days ago
README.md	Cambios rdme	2 days ago

**DATALOADER-IT3SISTRANS**

Programa Java para la carga masiva de +1,000,000 de datos a archivo .csv

##Instalacion En caso que eclipse no reconozca el main, File -> New -> Java Project -> Quitar "Use Default Location" -> Ingresar Ubicacion

**Generar datos:**

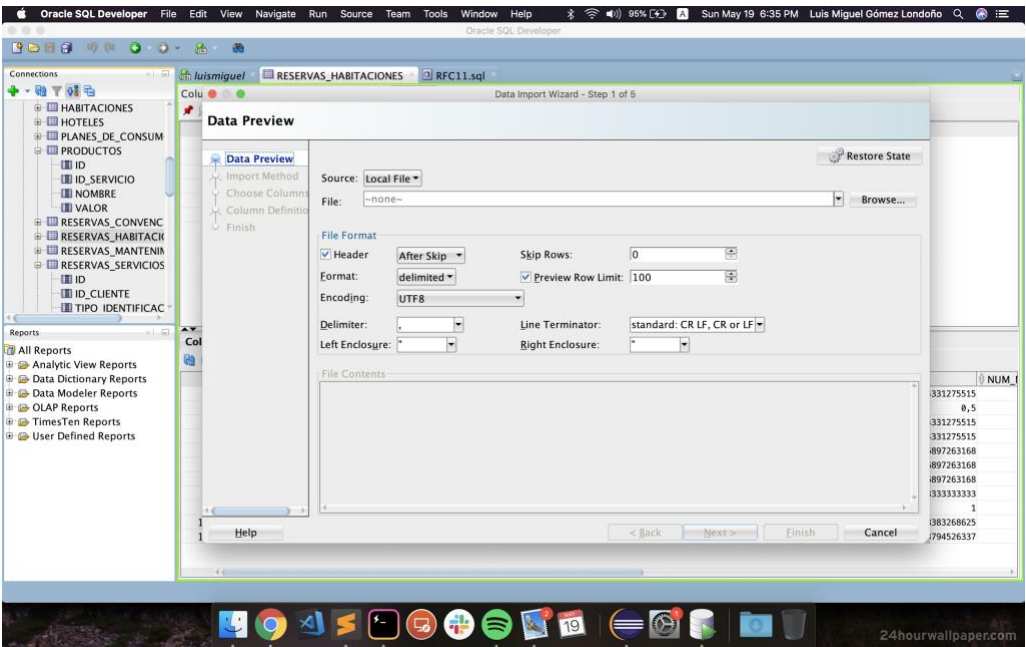
Para generar diferentes archivos de diferentes tablas, quitar comentario en el archivo de generar e ir parte por parte...

**Comentarios**

Los archivos se guardan en la carpeta generatedData Para generar nombres aleatorios se usa el JAR Faker

simple, consiste en pinchar sobre la tabla con click derecho -> Import Data y seguir con el proceso de cargar el archivo y asignar las columnas. Este proceso, al igual que el generador, tomaba tan solo cuestion de milisegundos.

Las tablas  
que se  
poblaron  
fueron  
elegidas  
segun los



requerimientos funcionales de consulta del proyecto. Estas, siguiendo la siguiente distribucion:

Tabla	#	Comentarios
RESERVAS_SERVICIOS	200145	
RESERVAS_MANTENIMIENTOS	101132	* Se divide en dos: uno para habitaciones y otro para servicios
RESERVAS_HABITACIONES	201145	
RESERVAS_CONVENCIONES	107142	*Se divide en dos: uno para habitaciones y otro para servicios (3000)
GASTOS	245599	
CLIENTES	142857	

## Links:

Carpeta OneDrive con archivos utilizados con los datos de carga: [https://uniandes-my.sharepoint.com/:f:/g/personal/lm\\_gomezl\\_uniandes\\_edu\\_co/ElbyiscFy-9FmT2U47rTNuABnkr5NcdxuwXFPLRJF3SvvA?e=7jDQf2](https://uniandes-my.sharepoint.com/:f:/g/personal/lm_gomezl_uniandes_edu_co/ElbyiscFy-9FmT2U47rTNuABnkr5NcdxuwXFPLRJF3SvvA?e=7jDQf2)

Repositorio GitHub con generador de datos para carga masiva: <https://github.com/luismigolondo/DataGenerator-IT3-SISTRANS>