**Documentacion del proceso de la CARGA MASIVA DE DATOS**

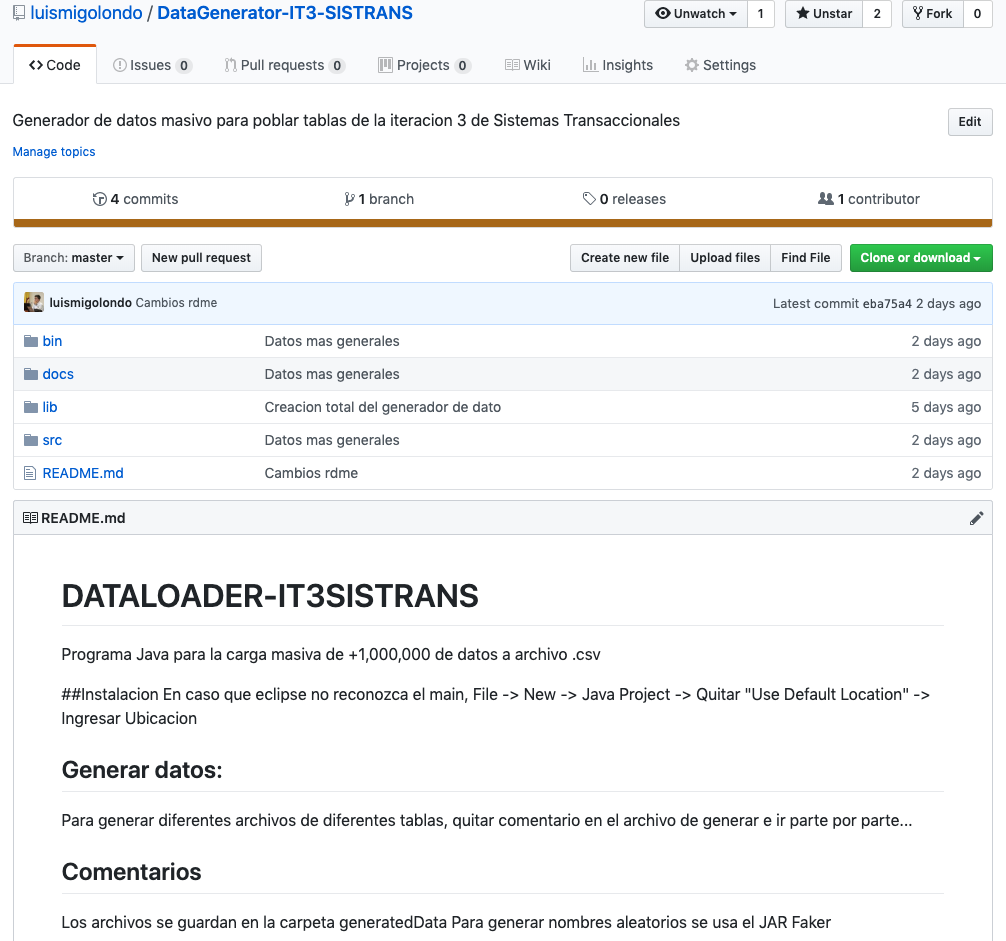
**Sistemas Transaccionales 2019-01**

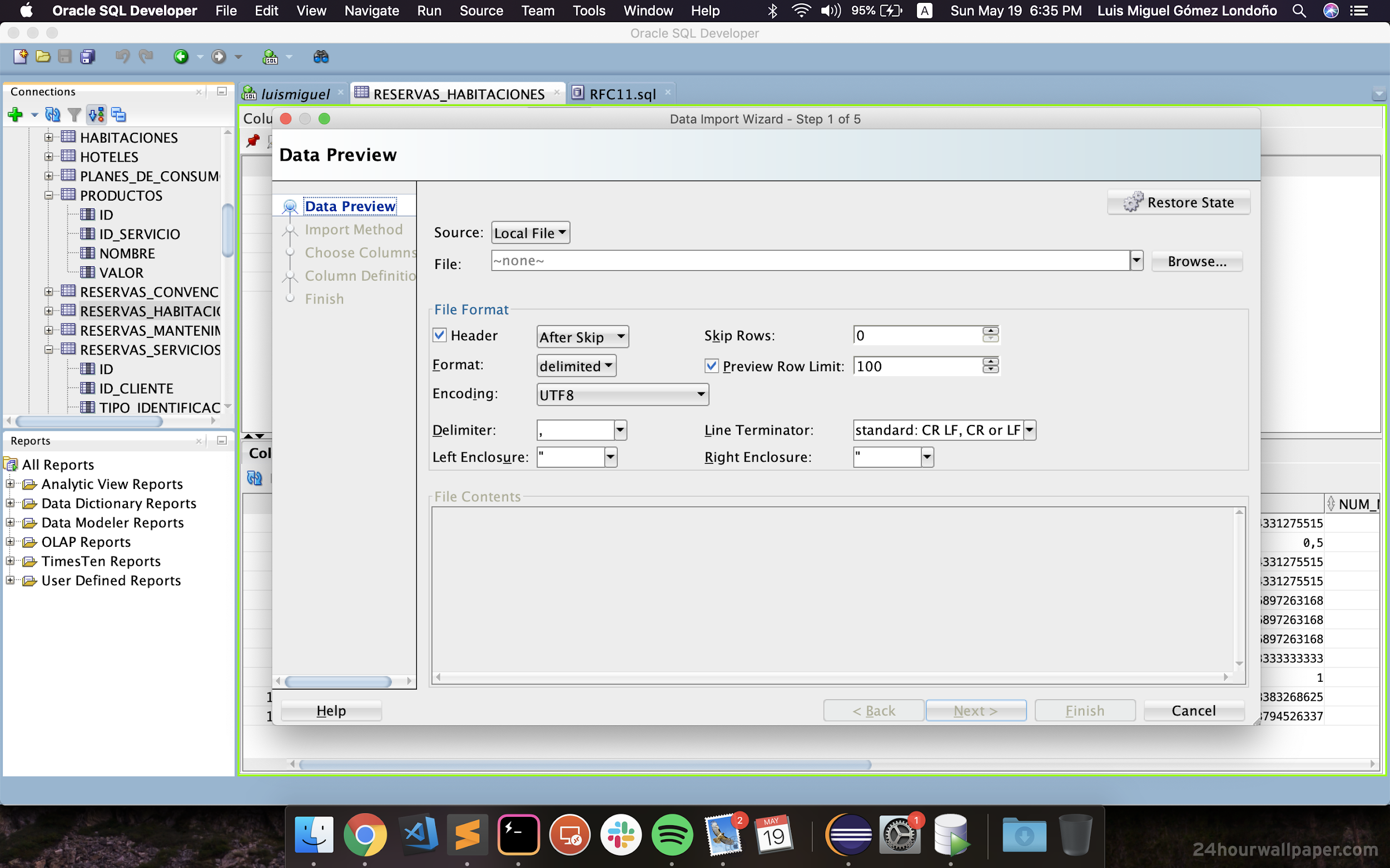
**Iteracion 3**

**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 quisieramos y ademas con mas felxibilidad y rapidez. La mayoria 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 SMBD. 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 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**](https://github.com/luismigolondo/DataGenerator-IT3-SISTRANS)