



72.37 - Base de Datos I

TPE

INTEGRANTES:

(n° 62161) Bruno Enzo Baumgart - Encargado de la investigación

(n° 62169) Santiago Jose Hirsch - Encargado del trigger

(n° 62556) Paz Aramburu - Encargada del funcionamiento general del proyecto

Introducción

El siguiente informe realiza un análisis del código de base de datos proporcionado, el cual tiene como objetivo generar un informe consolidado de nacimientos en base al archivo `us_births_2016_2021.csv` proporcionado por la cátedra. El código utiliza el lenguaje de programación PL/pgSQL y PostgreSQL como el sistema de gestión de bases de datos relacional.

Roles

Para el desarrollo del trabajo práctico trabajamos los tres juntos. Sin embargo, podemos definir el rol principal que cumplió cada uno mientras íbamos trabajando.

Bruno Enzo Baumgart - Encargado de la investigación.

Santiago Jose Hirsch - Encargado del trigger y funciones.

Paz Aramburu - Encargada del funcionamiento general del proyecto.

Investigación

Para el debido funcionamiento del proyecto, principalmente acudimos al material otorgado por la cátedra, revisando clases y haciendo lecturas profundas en los pdfs proporcionados. Por otra parte también se indagó en la documentación de PGADMIN para poder correr la función en la terminal y se tuvo que buscar información en internet sobre algunas cosas del código, especialmente cuando se encontraba un error o una traba.

Dificultades

A la hora del desarrollo las dificultades encontradas no fueron muchas, sin embargo la principal problemática la tuvimos a la hora de poder correr el programa desde la terminal, pero con la investigación nombrada anteriormente pudimos arreglarlo. Además tuvimos pequeñas problemáticas de sintaxis o armado de funciones/trigger que fueron rápidamente solucionadas con los pdfs de la cátedra.

Desarrollo

El código se divide en varias secciones, cada una de las cuales se explicará a continuación:

Definición de tablas

En esta sección se definen las tablas necesarias para almacenar la información relacionada con los nacimientos. Las tablas creadas son:

- ESTADO: Almacena los nombres y códigos de los estados.
- ANIO: Almacena los años y si son bisiestos o no.

- NIVEL_EDUCACION: Almacena las descripciones y niveles de educación.
- BIRTHS_TEMP: Tabla temporal utilizada para cargar los datos de los nacimientos desde el archivo CSV.
- BIRTHS_DEF: Tabla final que almacena los datos consolidados de los nacimientos.

Definición de funciones y triggers

Las funciones y triggers realizan las operaciones necesarias durante la carga de datos y generación del informe consolidado. Solamente contamos con un trigger y con cuatro funciones. A continuación se detalla una lista con cada uno de ellos:

- *distribute*: Trigger que ejecuta la función *distribute()* antes de insertar o actualizar tuplas en la tabla BIRTHS_TEMP.
- *distribute()*: Función que se ejecuta con el trigger *distribute*. Verifica si los datos relacionados con el año, estado y nivel de educación ya existen en las tablas correspondientes (ANIO, ESTADO, NIVEL_EDUCACION) y, en caso contrario, los inserta.
- *ReporteConsolidado_nivel_ed(_anio INTEGER, niv_ed INTEGER)*: Función que genera el informe consolidado para un nivel de educación específico en un año determinado. Calcula el total de nacimientos, la edad promedio de las madres, el mínimo y máximo de la edad de las madres, el peso promedio de los recién nacidos, y el mínimo y máximo del peso de los recién nacidos.
- *ReporteConsolidado_genero(_anio INTEGER, genero CHAR(1))*: Función que genera el informe consolidado para un género específico en un año determinado. Calcula el total de nacimientos, la edad promedio de las madres, el mínimo y máximo de la edad de las madres, el peso promedio de los recién nacidos, y el mínimo y máximo del peso de los recién nacidos.
- *ReporteConsolidado(num_anios INTEGER)*: Función principal que genera el informe consolidado para un número determinado de años. Recorre los años y estados, y llama a las funciones *ReporteConsolidado_genero* y *ReporteConsolidado_nivel_ed* para generar el informe consolidado respectivo.

Carga de datos

Para cargar los datos de los nacimientos desde un archivo CSV a la tabla BIRTHS_TEMP se utilizó la instrucción COPY, vista en la teoría de la cátedra.

Es importante aclarar que en el código proporcionado el PATH utilizado es el que nos permitió correr correctamente el código desde PGADMIN, por lo que para poder correr correctamente el programa, debería modificarse el PATH dentro del COPY por alguno donde su PGADMIN tenga permisos.

Generación del informe consolidad

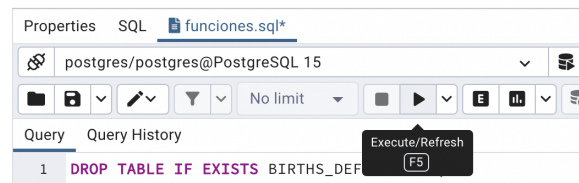
Se invoca la función *ReporteConsolidado* pasando como parámetro el número de años para generar el informe consolidado correspondiente.

Se deja como ejemplo comentado un llamado a la función ReporteConsolidado(7) dentro del código funciones.sql por si se quiere descomentar y probarlo directamente desde PGADMIN o corriendo el código.

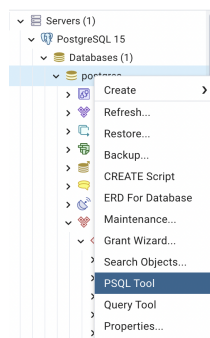
Ejecución

Para la ejecución del Trabajo Práctico son necesarios los siguientes pasos (las imágenes se obtuvieron utilizando PGADMIN):

1. Ejecutar funciones.sql



2. Abrir una terminal de psql



3. Ejecutar el comando “select reporteconsolidado(n);”, siendo n la cantidad de años que se quieren visualizar en el reporte

```
Properties  SQL  funciones.sql*  postgres/postgres@PostgreSQL 15
psql (15.2)
Type "help" for help.

postgres=# select reporteconsolidado(n);
```

Conclusiones

El archivo funciones.sql es el encargado de generar reportes consolidados sobre nacimientos en estados unidos desde 2016 hasta 2021. Esto es posible gracias a la importación de los datos a la tabla temporal, realizando de antemano el trigger para distribuir los datos en las tablas ANIO, ESTADO Y NIVEL_EDUCATIVO, y luego el manejo de los mismos con las funciones mencionadas en el desarrollo.