Universidad del Rosario

Escuela de Ingeniería, Ciencia y Tecnología

Matemáticas Aplicadas y Ciencias de la Computación

Ingeniería de datos

Nelson Sánchez Sánchez

Alejandro Vega, Juan Diego Martínez, Laura Hernández

Semestre 1, 2022

Repositorio en GitHub: <https://github.com/AlejandroVegaUR/Proyecto-Ingeniera>

Introducción

Muchas veces es complicado conseguir información relacionada con vuelos aéreos específicos, como hallar los datos del avión o de la aerolínea, de tal manera, se decidió crear el proyecto, el cual sirve para conseguir toda la información relacionada con la aerolínea, el avión, el aeropuerto, el vuelo, entre otros; en un solo lugar y de forma sencilla.

Fase de análisis

En esta fase se nombrará a detalle cada requerimiento de información que el sistema debe cumplir para ser funcional y lograr el objetivo del proyecto.

El sistema debe ser capaz de dar información la siguiente información, con relación a los aviones, se debe poder hallar la marca del avión, el modelo y la aerolínea a la que pertenece, por otra parte, en relación a los pasajeros, debe proporcionar el nombre, apellido, la nacionalidad y la fecha de nacimiento de cada uno, además, también debe ser posible obtener el número de vuelo y del ticket, la fecha de salida y de llegada y el aeropuerto del vuelo de los pasajeros.

Fase de diseño

En esta fase se nombrará las tablas de la base de datos, sus columnas, los tamaños, los tipos de datos y las restricciones de dominio y de integridad.

Antes de iniciar con las tablas, cabe destacar que para las columnas de tipo varchar, todas tienen un límite de caracteres de 80, por lo tanto, cuando se mencione este tipo de dato, este será su límite. Como primera tabla tenemos marcas, la cual tiene como columnas id\_marca que es de tipo integer y es el primary key, y nombre que es de tipo varchar.

Ahora esta la tabla países, con sus columnas las cuales son, iso\_alpha la cual es la primary key, y nombre, ambas son de tipo varchar.

Luego esta la tabla pasajeros, que tiene como columnas id\_pasajero de tipo integer y es la primary key, nombre, apellido y nacionalidad que todas son de tipo varchar, y fecha\_nacimiento que es de tipo date.

Sigue la tabla modelos, que tiene a las columnas id\_modelo de tipo integer y es la primary key, nombre de tipo varchar, y marca de tipo integer, además, marca es un foreign key que referencia la tabla marcas y la columna id\_marca.

Ahora esta la tabla aerolíneas, con las columnas id\_aerolinea de tipo integer y es la primary key, y nombre de tipo varchar.

Luego se tiene la tabla aeropuertos que tiene como columnas id\_aeropuerto de tipo integer y es la primary key, y nombre, ciudad y codigo\_pais de tipo varchar, además, código\_pais es un foreign key que referencia la tabla países y la columna iso\_alpha.

Sigue la tabla aviones que tiene las siguientes columnas, id\_avion de tipo integer y es la primary key, id\_modelo e id\_aerolinea ambas de tipo integer, además, las dos son foreign keys, id\_aerolinea referencia a la tabla aerolínea y la columna id\_aerolinea, e id\_modelo referencia la tabla modelos y la columna id\_modelo.

Ahora viene la tabla vuelo, con las columnas id\_vuelo de tipo varchar y es la primary key, id\_avion e id\_aeropuerto ambas de tipo integer, y fecha\_llegada y fecha\_salida ambas de tipo timestamp, además, id\_avion y id\_aeropuerto son foreign keys, id\_avion referencia a la tabla aviones y la columna id\_avion, e id\_aeropuerto referencia la tabla aeropuertos y la Columba id\_aeropuerto.

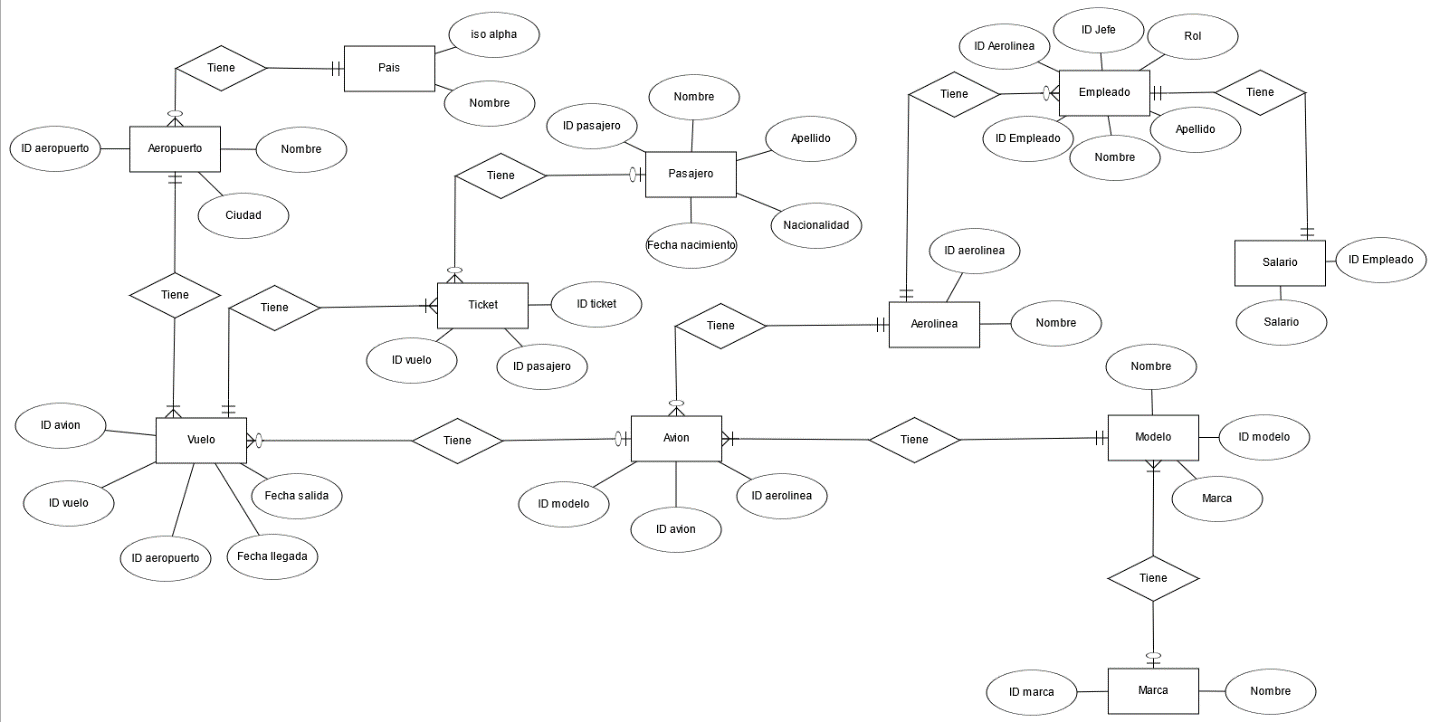
También está la tabla empleados, con las columnas id\_empleado de tipo integer y es la primary key, nombre, apellido y rol de tipo varchar, y, por último, id\_jefe e id\_aerolinea de tipo smallint, además, la variable id\_aerolinea es un foreign key que referencia a la tabla aerolíneas y la columna id\_aerolinea.

Luego se tiene la tabla salario, con las columnas id\_empleado de tipo integer y es la primary key, y salario de tipo real, además la columna id\_empleado es un foreign key que referencia la tabla empleados y la columna id\_empleado.

Por último, está la tabla tickets, con las columnas id\_ticket de tipo integer y es la primary key, id\_pasajero de tipo integer, e id\_vuelo de tipo varchar, además, id\_avion e id\_aeropuerto son foreign keys, id\_avion referencia la tabla aviones y la columna id\_avion, e id\_aeropuerto referencia la tabla aeropuertos y la columna id\_aeropuerto.

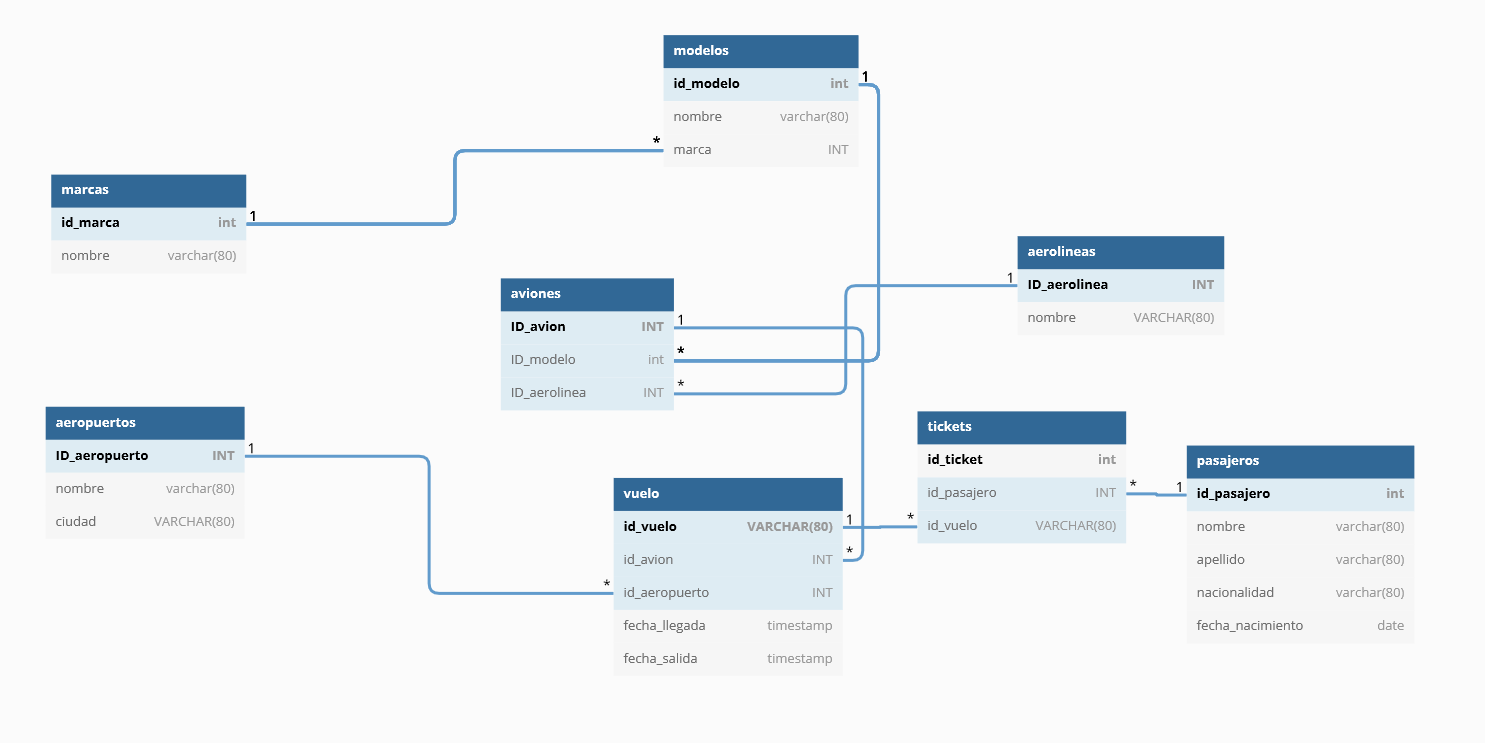
Diseño conceptual

Modelo Entidad/Relación:



En el modelo E/R anterior se puede ver las tablas descritas anteriormente y sus columnas, en este caso, los rectángulos son las tablas, los óvalos son las columnas, y los rombos son la relación que tienen entre si las tablas, también se puede ver qué tipo de relación tienen, como se presencia, la tabla vuelo y aeropuerto tienen una relación de muchos vuelos obligatoriamente a un aeropuerto, la tabla vuelo y ticket tienen una relaciona de muchos tickets obligatoriamente a un vuelo, ticket y pasajero tienen una relación de muchos tickets opcionalmente a un pasajero, vuelo y avión tienen una relación de muchos vuelos opcionalmente a un avión, avión y aerolínea tiene una relación de muchos aviones opcionalmente a una aerolínea, avión y modelo tienen una relación de muchos aviones obligatoriamente a un modelo y modelo y marca tienen una relación de muchos modelos obligatoriamente a una marca.

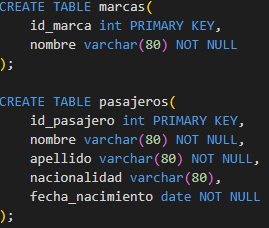
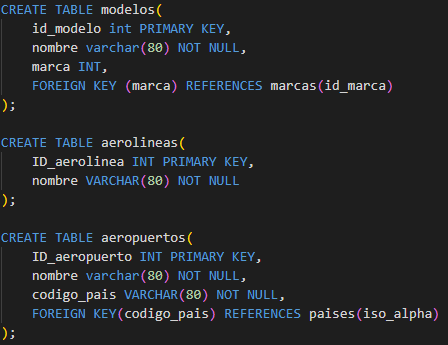
Diseño Lógico

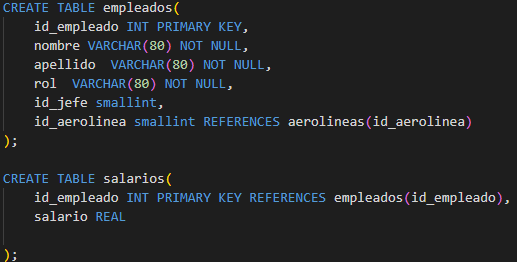
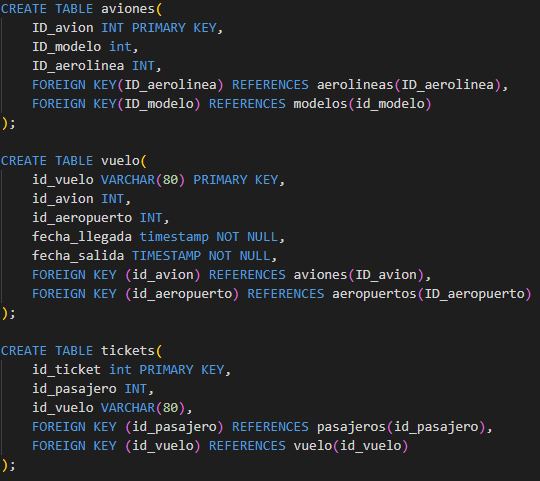
Esquema de base de datos relacional:

En este esquema se puede presenciar el diagrama E/R adaptado a una base de datos relacional, también se pueden ver las tablas con sus nombres, sus columnas con su tipo de dato, además, se puede ver que las columnas resaltadas en azul claro son las primary keys, y, por otra parte, también se pueden apreciar las relaciones entre las tablas, en este caso, el asterisco (\*) significa muchos.

Fase de construcción

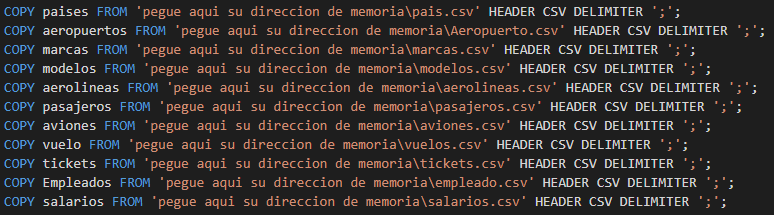
En esta fase se busca cambiar de un modelo de datos abstracto a la implementación de la base de datos incluyendo todo lo relacionado con el DML y DDL.

DDL:

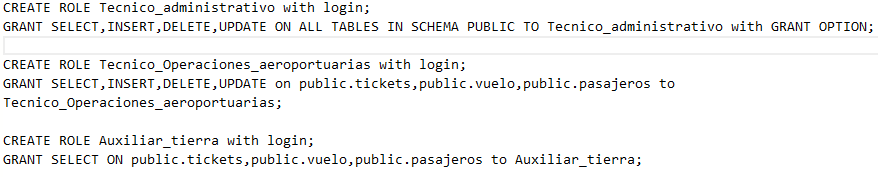


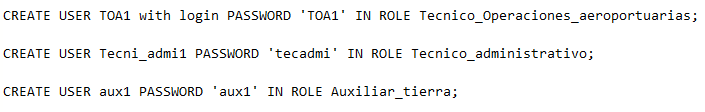
Como se puede ver, ya están creadas todas las tablas con sus columnas, los tipos de datos, y las restricciones, además, estas también están en 3NF, ya que en ninguna columna pueden haber duplicados ni múltiples valores, cada tabla solo tiene un primary key, y tampoco hay dependencia transitiva.

DML:

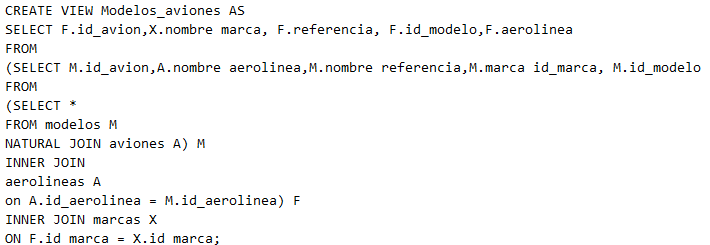
Para la inserción de información se llevo a cabo mediante archivos csv de la siguiente manera:

Utilizando estas inserciones se realizaron las siguientes roles, usuarios y consultas:

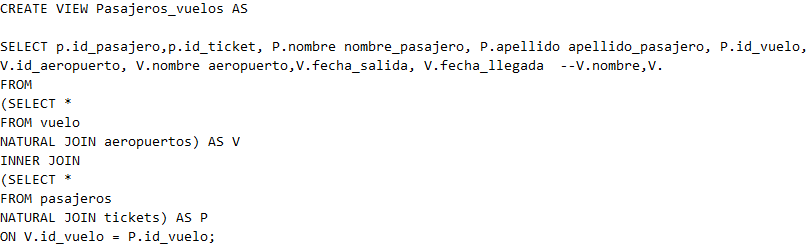
Roles:

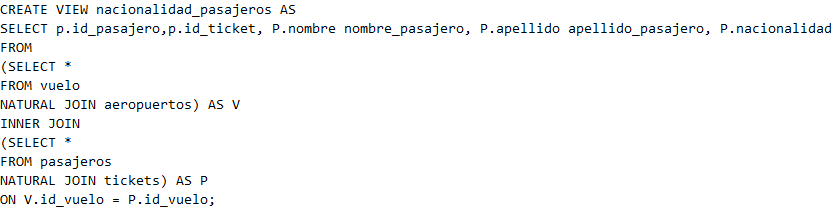
Usuarios:

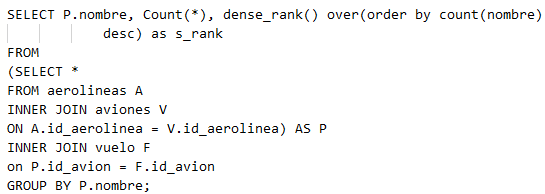
Consultas:

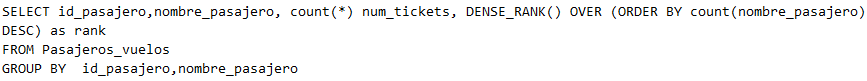
Todos los modelos de los aviones:

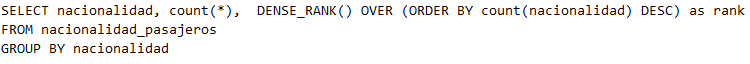
Todos los nombres de las aerolíneas:

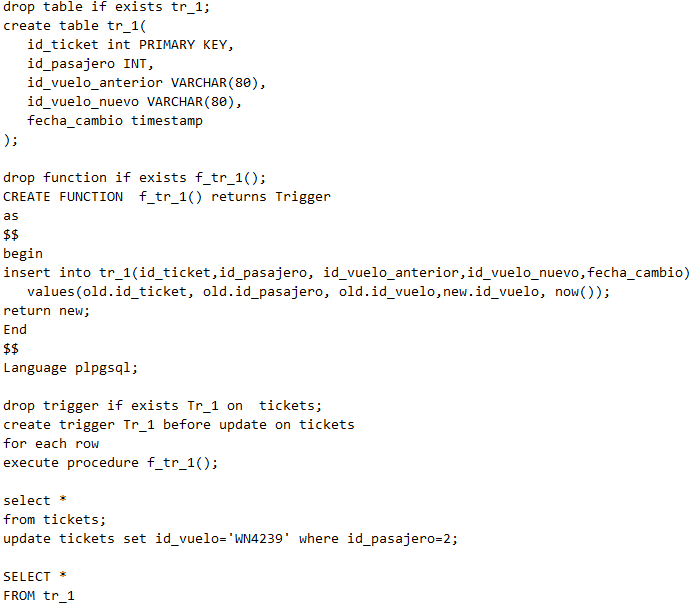
Datos del pasajero, tickets de vuelo y aeropuerto:

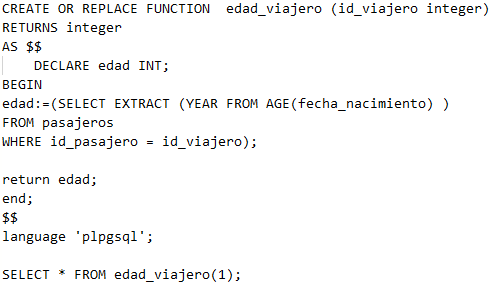
Información de vuelo de los pasajeros y el id, nombre, apellido y nacionalidad:

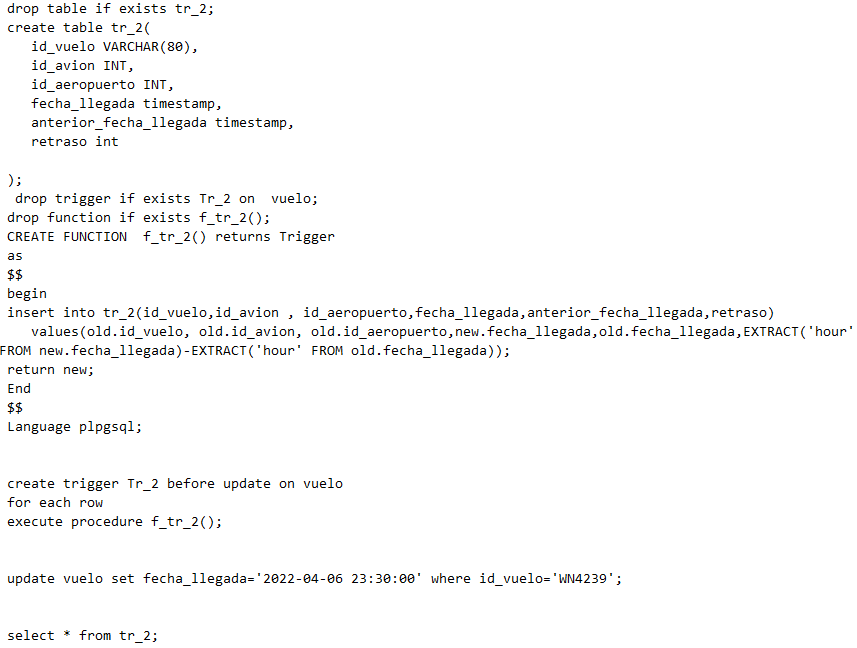
La aerolínea con más vuelos:

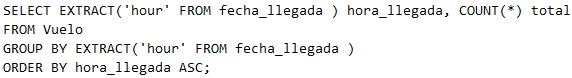
Pasajeros que más han viajado:

Jerarquía del número de pasajeros por nacionalidad:

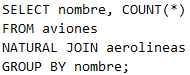
Almacenar la información de los cambios efectuados al numero de vuelo de un pasajero y la fecha en la que se realizó el cambio:

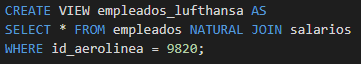
Edad de los pasajeros:

Almacenar la información de los cambios efectuados a la hora de llegada de los vuelos y evidenciar el retraso desde la última fecha en horas:

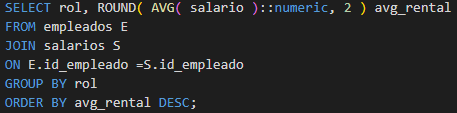
Cantidad de vuelos que llegan por hora:

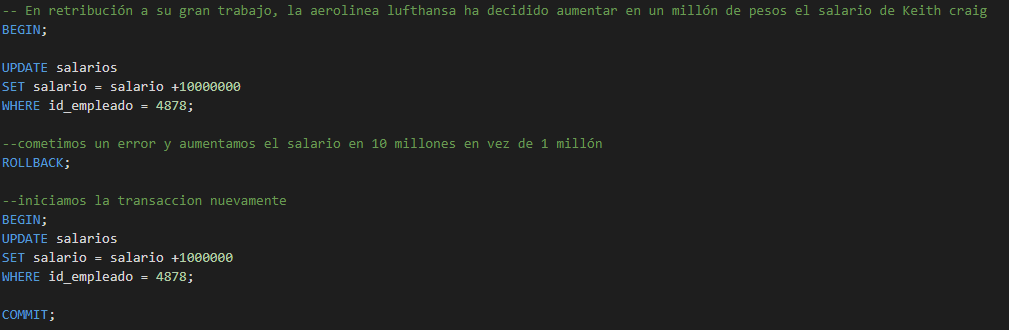
Cantidad de tickets por vuelo:

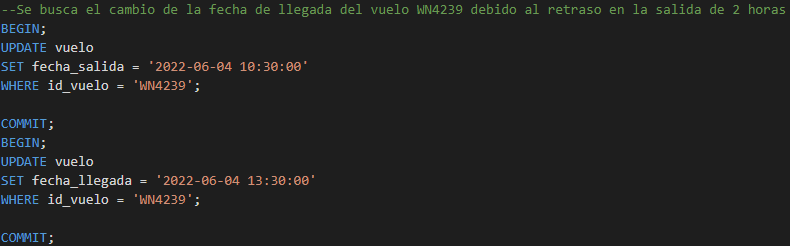
 Cantidad de aviones por aerolínea:

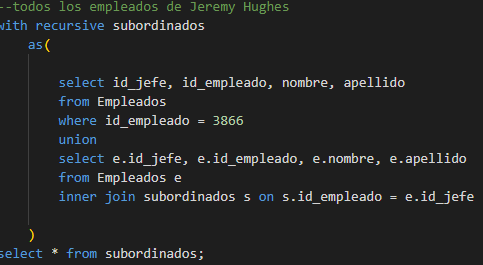
También se hicieron consultas sobre los empleados para conocer todos los empleados de las empresas:

En este caso son todos los empleados de la aerolínea con id 9820, la cual es la aerolínea Lufthansa, también hay consultas para todas las otras consultas, pero son iguales, solo cambia el id de la aerolínea.

Promedio de salario por cargo:

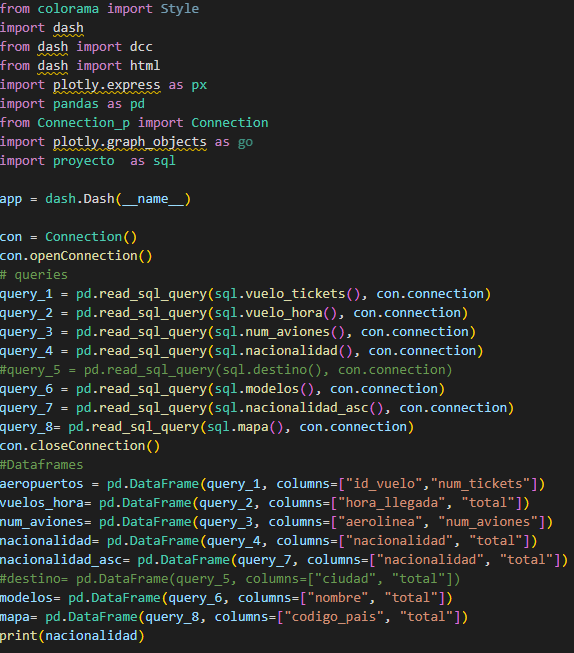
Transacción para aumentar el salario de un empleado:

Transacción para cambiar la hora de llegada de un vuelo:

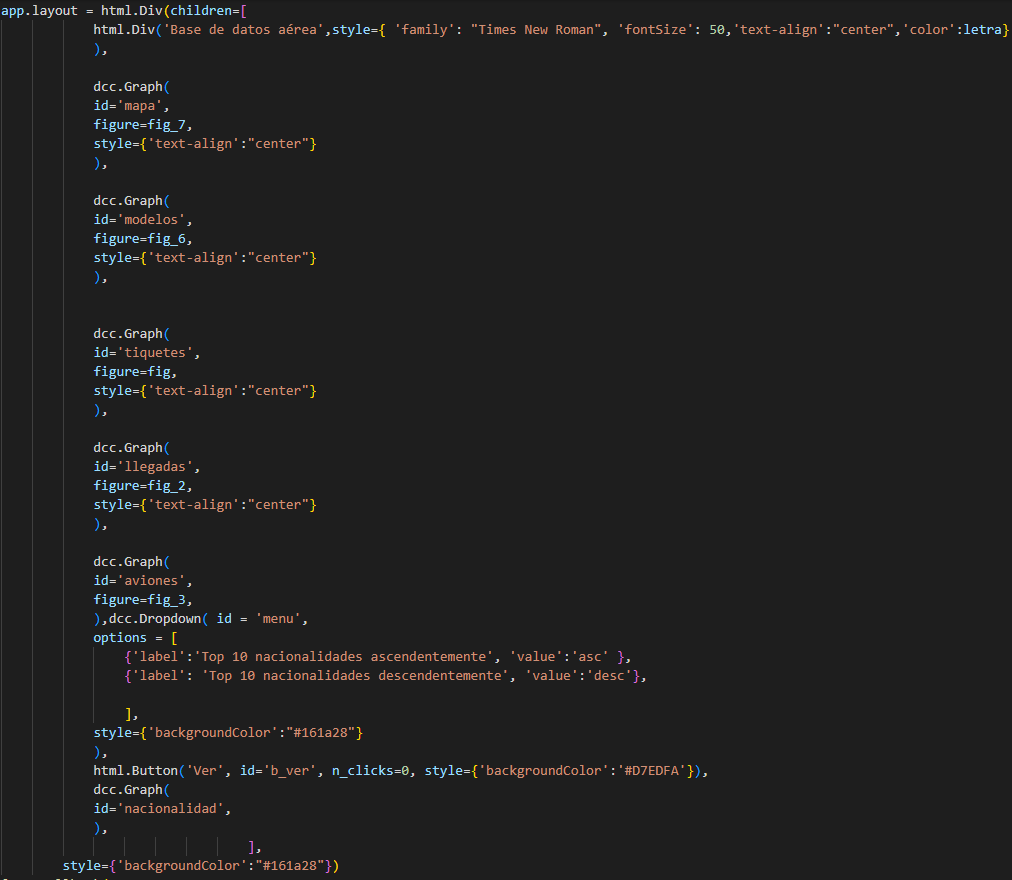
Por otra parte, se realizaron consultas recursivas para hallar todos los empleados de los jefes:

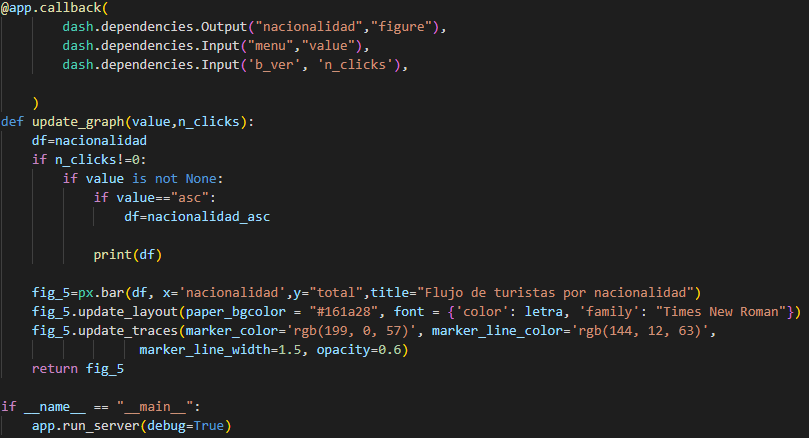
En este caso, se obtienen todos los empleados del jefe Jeremy Hughes, pero también se hicieron consultas para obtener los empleados de todos los otros jefes, el código es igual, solo cambia el id del empleado.

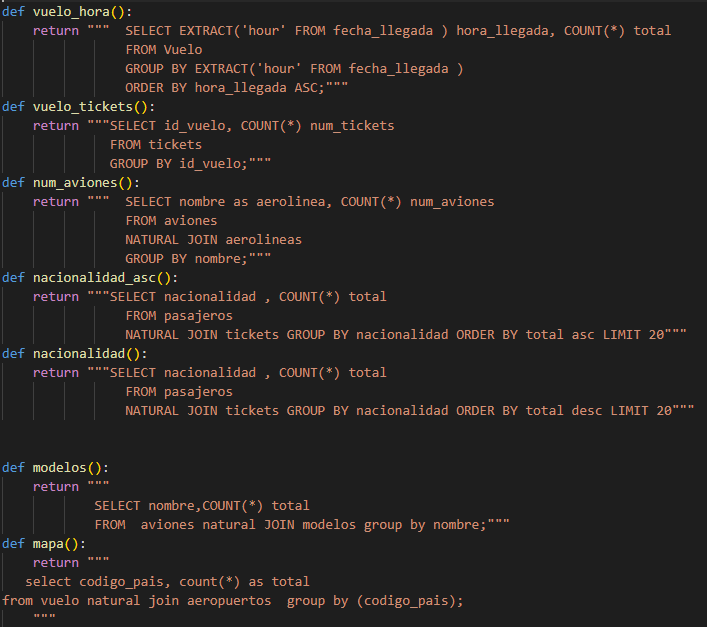
Además, también se realizó la conexión con Python de la siguiente manera:

Con el objetivo de utilizar dash de la siguiente forma:







Para graficar la siguiente consulta la cual nos da toda la información relacionada a los aeropuertos: