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: <a href="https://github.com/AlejandroVegaUR/Proyecto-Ingeniera">https://github.com/AlejandroVegaUR/Proyecto-Ingeniera</a>

## 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.

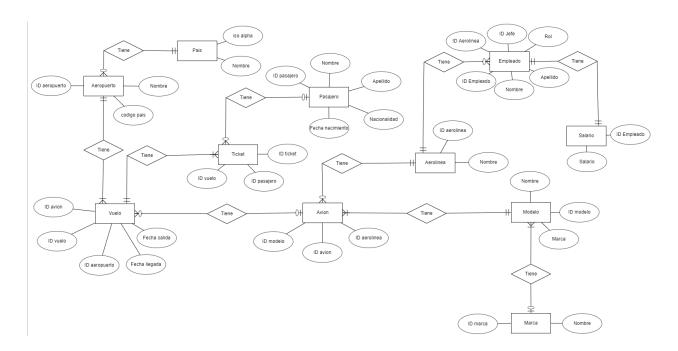
También está la tabla empleados, con las columnas id\_empleado de tipo integer y es la primary key, nombre, apellido y rol de tipo varchar, y, por último, id\_jefe e id\_aerolinea de tipo smallint, además, la variable id\_aerolinea es un foreign key que referencia a la tabla aerolíneas y la columna id\_aerolinea.

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

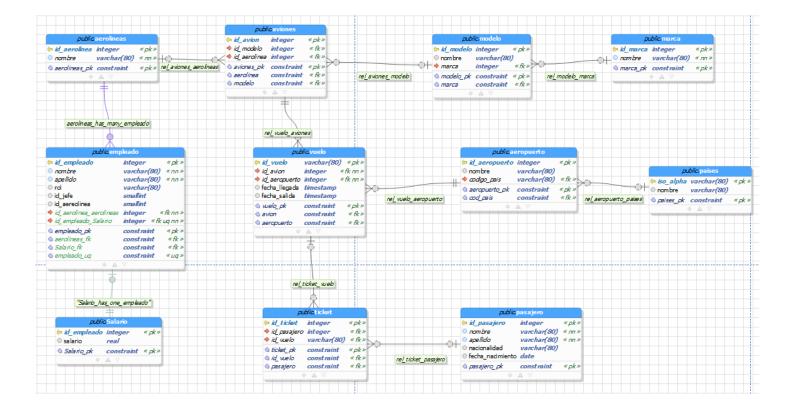
# Modelo Entidad/Relación:



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 relación 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:

```
REATE TABLE marcas(
                                                                     REATE TABLE modelos(
                                                                        id_modelo int PRIMARY KEY,
       id marca int PRIMARY KEY,
                                                                         nombre varchar(80) NOT NULL,
       nombre varchar(80) NOT NULL
                                                                         marca INT,
                                                                         FOREIGN KEY (marca) REFERENCES marcas(id_marca)
 ;
                                                                    CREATE TABLE aerolineas(
CREATE TABLE pasajeros(
                                                                         ID_aerolinea INT PRIMARY KEY,
       id pasajero int PRIMARY KEY,
                                                                        nombre VARCHAR(80) NOT NULL
       nombre varchar(80) NOT NULL,
       apellido varchar(80) NOT NULL, CREATE TABLE aeropuertos(
                                                                         ID aeropuerto INT PRIMARY KEY,
       nacionalidad varchar(80),
                                                                        nombre varchar(80) NOT NULL,
                                                                       codigo_pais VARCHAR(80) NOT NULL,
       fecha nacimiento date NOT NULL
                                                                        FOREIGN KEY(codigo_pais) REFERENCES paises(iso_alpha)
                                                                           CREATE TABLE empleados(
    ID_avion INT PRIMARY KEY,
ID_modelo int,
                                                                               id_empleado INT PRIMARY KEY,
                                                                               nombre VARCHAR(80) NOT NULL,
    ID_aerolinea INT,
FOREIGN KEY(ID_aerolinea) REFERENCES aerolineas(ID_aerolinea),
FOREIGN KEY(ID_modelo) REFERENCES modelos(id_modelo)
                                                                               apellido VARCHAR(80) NOT NULL,
                                                                               rol VARCHAR(80) NOT NULL,
                                                                               id_jefe smallint,
 REATE TABLE vuelo(
id_vuelo VARCHAR(80) PRIMARY KEY,
                                                                               id aerolinea smallint REFERENCES aerolineas(id aerolinea)
    id_avion INT,
   Tageropherto III,

fecha_llegada timestamp NOT NULL,

fecha_salida TIMESTAMP NOT NULL,

FOREIGN KEY (id_avion) REFERENCES aviones(ID_avion),

FOREIGN KEY (id_aeropuerto) REFERENCES aeropuertos(ID_aeropuerto)
                                                                            CREATE TABLE salarios(
                                                                               id_empleado INT PRIMARY KEY REFERENCES empleados(id_empleado)
 REATE TABLE tickets(
    vie inale tikeets
id_ticket int PRIMARY KEY,
id_pasajero INT,
id_vuelo VARCHAR(80),
FOREIGN KEY (id_pasajero) REFERENCES pasajeros(id_pasajero),
```

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:

```
COPY paises FROM 'pegue aqui su direccion de memoria\pais.csv' HEADER CSV DELIMITER ';';

COPY aeropuertos FROM 'pegue aqui su direccion de memoria\Aeropuerto.csv' HEADER CSV DELIMITER ';';

COPY marcas FROM 'pegue aqui su direccion de memoria\marcas.csv' HEADER CSV DELIMITER ';';

COPY modelos FROM 'pegue aqui su direccion de memoria\modelos.csv' HEADER CSV DELIMITER ';';

COPY aerolineas FROM 'pegue aqui su direccion de memoria\aerolineas.csv' HEADER CSV DELIMITER ';';

COPY pasajeros FROM 'pegue aqui su direccion de memoria\pasajeros.csv' HEADER CSV DELIMITER ';';

COPY aviones FROM 'pegue aqui su direccion de memoria\aviones.csv' HEADER CSV DELIMITER ';';

COPY vuelo FROM 'pegue aqui su direccion de memoria\vuelos.csv' HEADER CSV DELIMITER ';';

COPY tickets FROM 'pegue aqui su direccion de memoria\tickets.csv' HEADER CSV DELIMITER ';';

COPY Empleados FROM 'pegue aqui su direccion de memoria\empleado.csv' HEADER CSV DELIMITER ';';
```

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

```
CREATE ROLE Tecnico_administrativo with login;
GRANT SELECT,INSERT,DELETE,UPDATE ON ALL TABLES IN SCHEMA PUBLIC TO Tecnico_administrativo with GRANT OPTION;

CREATE ROLE Tecnico_Operaciones_aeroportuarias with login;
GRANT SELECT,INSERT,DELETE,UPDATE on public.tickets,public.vuelo,public.pasajeros to
Tecnico_Operaciones_aeroportuarias;

CREATE ROLE Auxiliar_tierra with login;
GRANT SELECT ON public.tickets,public.vuelo,public.pasajeros to Auxiliar_tierra;

Roles:
```

# Usuarios:

```
CREATE USER TOA1 with login PASSWORD 'TOA1' IN ROLE Tecnico_Operaciones_aeroportuarias;

CREATE USER Tecni_admi1 PASSWORD 'tecadmi' IN ROLE Tecnico_administrativo;

CREATE USER aux1 PASSWORD 'aux1' IN ROLE Auxiliar_tierra;
```

## Consultas:

## Todos los modelos de los aviones:

```
CREATE VIEW Modelos_aviones AS

SELECT F.id_avion,X.nombre marca, F.referencia, F.id_modelo,F.aerolinea

FROM

(SELECT M.id_avion,A.nombre aerolinea,M.nombre referencia,M.marca id_marca, M.id_modelo

FROM

(SELECT *

FROM modelos M

NATURAL JOIN aviones A) M

INNER JOIN

aerolineas A

on A.id_aerolinea = M.id_aerolinea) F

INNER JOIN marcas X

ON F.id marca = X.id marca;
```

Todos los nombres de las aerolíneas:

```
CREATE VIEW nom_aerolineas
AS
SELECT nombre FROM Aerolineas;
```

# Datos del pasajero, tickets de vuelo y aeropuerto:

```
CREATE VIEW Pasajeros_vuelos AS
```

```
SELECT p.id_pasajero,p.id_ticket, P.nombre nombre_pasajero, P.apellido apellido_pasajero, P.id_vuelo,
V.id_aeropuerto, V.nombre aeropuerto,V.fecha_salida, V.fecha_llegada --V.nombre,V.
FROM
(SELECT *
FROM vuelo
NATURAL JOIN aeropuertos) AS V
INNER JOIN
(SELECT *
FROM pasajeros
NATURAL JOIN tickets) AS P
ON V.id_vuelo = P.id_vuelo;
```

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

```
CREATE VIEW nacionalidad_pasajeros AS

SELECT p.id_pasajero,p.id_ticket, P.nombre nombre_pasajero, P.apellido apellido_pasajero, P.nacionalidad

FROM

(SELECT *

FROM vuelo

NATURAL JOIN aeropuertos) AS V

INNER JOIN

(SELECT *

FROM pasajeros

NATURAL JOIN tickets) AS P

ON V.id_vuelo = P.id_vuelo;
```

## La aerolínea con más vuelos:

```
SELECT P.nombre, Count(*), dense_rank() over(order by count(nombre)

desc) as s_rank

FROM

(SELECT *

FROM aerolineas A

INNER JOIN aviones V

ON A.id_aerolinea = V.id_aerolinea) AS P

INNER JOIN vuelo F

on P.id_avion = F.id_avion

GROUP BY P.nombre;
```

# Pasajeros que más han viajado:

```
SELECT id_pasajero,nombre_pasajero, count(*) num_tickets, DENSE_RANK() OVER (ORDER BY count(nombre_pasajero) DESC) as rank
FROM Pasajeros_vuelos
GROUP BY id_pasajero,nombre_pasajero
```

# Jerarquía del número de pasajeros por nacionalidad:

```
SELECT nacionalidad, count(*), DENSE_RANK() OVER (ORDER BY count(nacionalidad) DESC) as rank FROM nacionalidad_pasajeros
GROUP BY 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:
drop table if exists tr_1;
create table tr_1(
   id_ticket int PRIMARY KEY,
   id_pasajero INT,
   id_vuelo_anterior VARCHAR(80),
   id vuelo nuevo VARCHAR(80),
   fecha_cambio timestamp
);
drop function if exists f tr 1();
CREATE FUNCTION f_tr_1() returns Trigger
$$
begin
insert into tr_1(id_ticket,id_pasajero, id_vuelo_anterior,id_vuelo_nuevo,fecha_cambio)
   values(old.id ticket, old.id pasajero, old.id vuelo, new.id vuelo, now());
return new;
End
$$
Language plpgsql;
drop trigger if exists Tr_1 on tickets;
create trigger Tr_1 before update on tickets
for each row
execute procedure f_tr_1();
select *
from tickets;
update tickets set id vuelo='WN4239' where id pasajero=2;
SELECT *
FROM tr_1
   Edad de los pasajeros:
CREATE OR REPLACE FUNCTION edad_viajero (id_viajero integer)
RETURNS integer
AS $$
    DECLARE edad INT;
BEGIN
edad:=(SELECT EXTRACT (YEAR FROM AGE(fecha_nacimiento) )
FROM pasajeros
WHERE id_pasajero = id_viajero);
return edad;
end;
$$
language 'plpgsql';
SELECT * FROM edad_viajero(1);
```

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:

```
drop table if exists tr 2;
create table tr 2(
   id vuelo VARCHAR(80),
   id avion INT,
   id_aeropuerto INT,
   fecha_llegada timestamp,
   anterior_fecha_llegada timestamp,
   retraso int
);
 drop trigger if exists Tr_2 on vuelo;
drop function if exists f_tr_2();
CREATE FUNCTION f_tr_2() returns Trigger
as
$$
begin
insert into tr_2(id_vuelo,id_avion , id_aeropuerto,fecha_llegada,anterior_fecha_llegada,retraso)
   values(old.id_vuelo, old.id_avion, old.id_aeropuerto,new.fecha_llegada,old.fecha_llegada,EXTRACT('hour'
FROM new.fecha_llegada)-EXTRACT('hour' FROM old.fecha_llegada));
return new;
End
$$
Language plpgsql;
create trigger Tr_2 before update on vuelo
for each row
execute procedure f_tr_2();
update vuelo set fecha llegada='2022-04-06 23:30:00' where id vuelo='WN4239';
select * from tr_2;
   Cantidad de vuelos que llegan por hora:
SELECT EXTRACT('hour' FROM fecha_llegada ) hora_llegada, COUNT(*) total
FROM Vuelo
GROUP BY EXTRACT('hour' FROM fecha_llegada )
ORDER BY hora_llegada ASC;
   Cantidad de tickets por vuelo:
SELECT id vuelo, COUNT(*) num tickets
FROM tickets
GROUP BY id vuelo;
```

Cantidad de aviones por aerolínea:

```
SELECT nombre, COUNT(*)
FROM aviones
NATURAL JOIN aerolineas
GROUP BY nombre;
```

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

```
CREATE VIEW empleados_lufthansa AS

SELECT * FROM empleados NATURAL JOIN salarios

WHERE id_aerolinea = 9820;
```

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:

```
SELECT rol, ROUND( AVG( salario )::numeric, 2 ) avg_rental FROM empleados E
JOIN salarios S
ON E.id_empleado =S.id_empleado
GROUP BY rol
ORDER BY avg_rental DESC;
```

Transacción para aumentar el salario de un empleado:

```
-- En retribución a su gran trabajo, la aerolinea lufthansa ha decidido aumentar en un millón de pesos el salario de Keith craig
BEGIN;

UPDATE salarios
SET salario = salario +10000000
WHERE id_empleado = 4878;

--cometimos un error y aumentamos el salario en 10 millones en vez de 1 millón
ROLLBACK;

--iniciamos la transaccion nuevamente
BEGIN;
UPDATE salarios
SET salario = salario +1000000
WHERE id_empleado = 4878;

COMMIT;
```

Transacción para cambiar la hora de llegada de un vuelo:

```
--Se busca el cambio de la fecha de llegada del vuelo WN4239 debido al retraso en la salida de 2 horas
BEGIN;

UPDATE vuelo

SET fecha_salida = '2022-06-04 10:30:00'

WHERE id_vuelo = 'WN4239';

COMMIT;
BEGIN;

UPDATE vuelo

SET fecha_llegada = '2022-06-04 13:30:00'

WHERE id_vuelo = 'WN4239';

COMMIT;
```

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

```
--todos los empleados de Jeremy Hughes
with recursive subordinados
as(

select id_jefe, id_empleado, nombre, apellido
from Empleados
where id_empleado = 3866
union
select e.id_jefe, e.id_empleado, e.nombre, e.apellido
from Empleados e
inner join subordinados s on s.id_empleado = e.id_jefe
)
select * from subordinados;
```

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:

```
rom colorama import Style
import dash
from dash import dcc
from <mark>dash import html</mark>
import plotly.express as px
import pandas as pd
 mport plotly.graph_objects as go
import proyecto as sql
app = dash.Dash(__name__)
con = Connection()
con.openConnection()
query_1 = pd.read_sql_query(sql.vuelo_tickets(), con.connection)
query_2 = pd.read_sql_query(sql.vuelo_hora(), con.connection)
query_3 = pd.read_sql_query(sql.num_aviones(), con.connection)
query_4 = pd.read_sql_query(sql.nacionalidad(), con.connection)
query_6 = pd.read_sql_query(sql.modelos(), con.connection)
query_7 = pd.read_sql_query(sql.nacionalidad_asc(), con.connection)
query_8= pd.read_sql_query(sql.mapa(), con.connection)
con.closeConnection()
aeropuertos = pd.DataFrame(query_1, columns=["id_vuelo","num_tickets"])
vuelos_hora= pd.DataFrame(query_2, columns=["hora_llegada", "total"])
num_aviones= pd.DataFrame(query_3, columns=["aerolinea", "num_aviones"])
nacionalidad= pd.DataFrame(query_4, columns=["nacionalidad", "total"])
nacionalidad_asc= pd.DataFrame(query_7, columns=["nacionalidad", _"total"])
#destino= pd.DataFrame(query_5, columns=["ciudad", "total"])
modelos= pd.DataFrame(query_6, columns=["nombre", "total"])
mapa= pd.DataFrame(query_8, columns=["codigo_pais", "total"])
print(nacionalidad)
```

```
app.layout = html.Div(children=[
html.Div('Base de datos aérea',style={ 'family': "Times New Roman", 'fontSize': 50,'text-align':"center",'color':letra}
               dcc.Graph(
               figure=fig_7,
style={'text-align':"center"}
               dcc.Graph(
               figure=fig_6,
style={'text-align':"center"}
               dcc.Graph(
               id='tiquetes',
figure=fig,
style={'text-align':"center"}
               dcc.Graph(
               id='llegadas',
               figure=fig_2,
style={'text-align':"center"}
               dcc.Graph(
               id='aviones',
               figure=fig 3,
                ),dcc.Dropdown( id = 'menu',
               options = [
    {'label':'Top 10 nacionalidades ascendentemente', 'value':'asc' },
    {'label': 'Top 10 nacionalidades descendentemente', 'value':'desc'},
               ],
style={'backgroundColor':"#161a28"}
               dcc.Graph(
```

```
@app.callback(
       dash.dependencies.Output("nacionalidad","figure"),
       dash.dependencies.Input("menu","value"),
       dash.dependencies.Input('b_ver', 'n_clicks'),
def update graph(value,n clicks):
   df=nacionalidad
           if value=="asc":
               df=nacionalidad asc
           print(df)
   fig_5=px.bar(df, x='nacionalidad',y="total",title="Flujo de turistas por nacionalidad")
   fig_5.update_layout(paper_bgcolor = "#161a28", font = {'color': letra, 'family': "Times New Roman"})
   fig_5.update_traces(marker_color='rgb(199, 0, 57)', marker_line_color='rgb(144, 12, 63)',
                 marker_line_width=1.5, opacity=0.6)
   return fig_5
  __name__ == "__main__":
   app.run_server(debug=True)
```

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

```
def vuelo_hora():
               SELECT EXTRACT('hour' FROM fecha_llegada ) hora_llegada, COUNT(*) total
               FROM Vuelo
                GROUP BY EXTRACT('hour' FROM fecha llegada )
               ORDER BY hora_llegada ASC;""
def vuelo_tickets():
             FROM tickets
def num_aviones():
   return """ SELECT nombre as aerolinea, COUNT(*) num_aviones
FROM aviones
               GROUP BY nombre;"
def nacionalidad_asc():
               FROM pasajeros
def nacionalidad():
              FROM pasajeros
               NATURAL JOIN tickets GROUP BY nacionalidad ORDER BY total desc LIMIT 20"""
def modelos():
           SELECT nombre, COUNT(*) total
           FROM aviones natural JOIN modelos group by nombre;"""
def mapa():
  select codigo_pais, count(*) as total
 rom vuelo natural join aeropuertos group by (codigo_pais);
```