```sql
-- ------------------------------------------------
-- Visualización de datos III
-- Primera Parte - Análisis de Datos con SQL
-- ------------------------------------------------


-- ------------------------------------------------
-- Importar el archivo “kc_house_data.csv”
-- Generacion de Tabla "housing"
-- ------------------------------------------------

CREATE TABLE IF NOT EXISTS public.housing
(
    id CHARACTER VARYING(10),
    date_time timestamp,
    price BIGINT NOT NULL,
    bedrooms INT NOT NULL,
    bathrooms FLOAT NOT NULL,
    sqft_living INT NOT NULL,
    sqft_lot BIGINT NOT NULL,
    floors FLOAT NOT NULL,
    waterfront BIT(1) NOT NULL,
    view_apt INT NOT NULL,
    condition_apt INT NOT NULL,
    grade INT NOT NULL,
    sqft_above INT NOT NULL,
    sqft_basement INT NOT NULL,
    yr_built INT NOT NULL,
    yr_renovated INT NOT NULL,
    zipcode INT NOT NULL,
    lat FLOAT NOT NULL,
    long FLOAT NOT NULL,
    sqft_living15 INT NOT NULL,
    sqft_lot15 BIGINT NOT NULL
);

select * from housing;

COPY housing (id, date_time, price, bedrooms, bathrooms,
    sqft_living, sqft_lot, floors, waterfront, view_apt,
    condition_apt, grade, sqft_above, sqft_basement, yr_built,
    yr_renovated, zipcode, lat, long, sqft_living15, sqft_lot15)
FROM 'D:\WORK IN PROGRESS\Data Analytics course\parte 5 visualizacion de datos\week
45\kc_house_data.csv'
(FORMAT csv, HEADER, DELIMITER E'\t');


select distinct id,
    count(*)
from housing
group by 1
having count(*)>1;


---------------------------------------
--  Precio promedio de una casa por zipcode,
--  agrupado por año de construcción y
--  ordenado de mayor a menor (Top 50).
--  Además incluir el número promedio de habitaciones (bedrooms) y
```

```
--  baños (bathrooms) para la agrupación.
----------------------------------------

select * from housing;

SELECT
    zipcode,
    yr_built,
    ROUND(avg(price))AS PRECIO_MEDIO,
    ROUND(avg(bedrooms))AS HABITACIONES_MEDIO,
    ROUND(avg(bathrooms))AS BANOS_MEDIO
FROM housing
GROUP BY 1,2
ORDER BY 2 DESC
LIMIT 50;

----------------------------------------
--  Calcular el precio por m2 (ojo! La información viene en square feet), por zipcode.
----------------------------------------


-- divide el valor de Área entre 10.764

ALTER TABLE housing
ADD COLUMN m2_lot INT;


UPDATE housing
    SET m2_lot = ROUND(sqft_lot/10.764);

select * from housing;


SELECT
    zipcode,
    ROUND (AVG(price/m2_lot)) AS PRECIOMEDIOXM2
FROM housing
GROUP BY 1
ORDER BY 2 DESC;

----------------------------------------
-- Precio clasificado por grade confrontado con el número de viviendas por grade
----------------------------------------

SELECT
    grade,
    ROUND(AVG(price),2) AS PRECIO_MEDIO,
    COUNT(*) AS Q_VIVIENDAS
FROM housing
GROUP BY 1
ORDER BY 1;

----------------------------------------
-- Precio clasificado por número de pisos confrontado con el número de viviendas por
cantidad de pisos
----------------------------------------

SELECT
    floors,
```

```
    ROUND(AVG(price),2) AS PRECIO_MEDIO,
    COUNT(*) AS Q_VIVIENDAS
FROM housing
GROUP BY 1
ORDER BY 1;
```