```
CREATE TABLE house_details(
	id varchar(50),
	home_type varchar(50),
	bed	int ,
	bath int,
	play_ground varchar(10),
	swming_tank varchar(10),
	acre_lot numeric ,
	city varchar(50),
	state varchar(50),
	zip_code int ,
	house_size int
);

select * from house_details;

CREATE TABLE dates_price(
	id varchar(50),
	listing_date date ,
	listing_price int,
	sold_date date ,
	actual_sold_price int ,
	status varchar(20)
);

select * from dates_price;

CREATE TABLE zip(
	city varchar(50),
	state varchar(50),
	zip_code int ,
	zip_code_population int
);

select * from zip;
```

Note : /* here in zip file zip_code_population column is not interger type so I convert it into integer in csv file only */

```
/* Delete 1 row from zip table where zip code has 2 values */
DELETE FROM zip AS z1
WHERE EXISTS (
```

```
    SELECT 1
    FROM zip AS z2
    WHERE z1.zip_code = z2.zip_code
    AND z1.zip_code_population > z2.zip_code_population
);
```

/* top 5 cities with the highest number of houses sold */

```
SELECT z.city, COUNT(1) AS houses_sold_count
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
JOIN zip z ON hd.zip_code = z.zip_code
WHERE dp.status = 'Sold'
GROUP BY z.city
ORDER BY houses_sold_count DESC
LIMIT 5;
```

/* the months in which the most houses are being sold */

```
SELECT TO_CHAR(dp.sold_date, 'Month') AS sold_month, COUNT(1) AS houses_sold_count
FROM dates_price dp
WHERE dp.status = 'Sold'
GROUP BY sold_month
ORDER BY houses_sold_count DESC;
```

/*the average sold price for each home type */

```
SELECT hd.home_type, ROUND(AVG(dp.actual_sold_price),0)AS avg_sold_price
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
GROUP BY hd.home_type;
```

/*details of houses sold where difference between listing_price ,actual_sold_price is 100000.*/

```
SELECT hd.id, hd.city, hd.house_size, dp.listing_price, dp.actual_sold_price
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE  dp.status = 'Sold' and (dp.listing_price - dp.actual_sold_price) > 100000 ;
```

/* cities with the highest average sold price */

```
SELECT z.city, ROUND(AVG(dp.actual_sold_price),0) AS avg_sold_price
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
JOIN zip z ON hd.zip_code = z.zip_code
WHERE  dp.status = 'Sold'
```

```
GROUP BY z.city
ORDER BY avg_sold_price DESC;
```

/* details of houses sold with a swimming pool and a play ground */

```
SELECT hd.id, hd.city, hd.house_size, dp.listing_price, dp.actual_sold_price
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE hd.swming_tank = 'Y' AND hd.play_ground = 'Y' and dp.status = 'Sold';
```

/* details of houses that don't have bed or bathroom */

```
SELECT hd.*, dp.listing_price, dp.actual_sold_price
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE hd.bed = 0 or hd.bath = 0;
```

/* the average difference between listing price and actual sold price*/

```
SELECT ROUND(AVG(actual_sold_price - listing_price  ),0) AS avg_price_difference
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE dp.status = 'Sold' AND dp.listing_price  < dp.actual_sold_price;
```

/* the number of houses sold in each state for which the sold price is above the average sold price */

```
SELECT z.state, COUNT(*) AS houses_sold_above_avg
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
JOIN zip z ON hd.zip_code = z.zip_code
WHERE dp.actual_sold_price > (
   SELECT AVG(actual_sold_price)
   FROM dates_price)
GROUP BY z.state;
```

/* the listing date of houses that occurred later than their sold date * i.e. house are sold before listing */

```
SELECT hd.id, hd.home_type, hd.city, hd.state, hd.zip_code,  dp.listing_date, dp.sold_date
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE dp.listing_date > dp.sold_date AND dp.status = 'Sold';
```

/* the top customer based on actual sold price and determine the city associated with that top customer */

```
SELECT hd.city, dp.id, SUM(dp.actual_sold_price) AS total_purchase_amount
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
GROUP BY hd.city, dp.id
ORDER BY total_purchase_amount DESC
LIMIT 1;
```

/* total number of houses sold in each city */

```
SELECT hd.city, COUNT(*) AS total_houses_sold
FROM house_details hd
JOIN dates_price dp ON hd.id = dp.id
WHERE dp.status = 'Sold'
GROUP BY hd.city
ORDER BY total_houses_sold DESC;
```

/* the customer who has spent the most on houses for each city */

```
WITH CustomerWithCity AS (
   SELECT hd.city, dp.id, SUM(dp.actual_sold_price) AS total_purchase_amount,
      ROW_NUMBER() OVER(PARTITION BY hd.city ORDER BY SUM(dp.actual_sold_price) DESC) AS
RowNo
   FROM house_details hd
   JOIN dates_price dp ON hd.id = dp.id
         WHERE dp.status = 'Sold'
   GROUP BY hd.city, dp.id
)
SELECT cc.city, cc.id, cc.total_purchase_amount
FROM CustomerWithCity cc
WHERE cc.RowNo = 1;
```