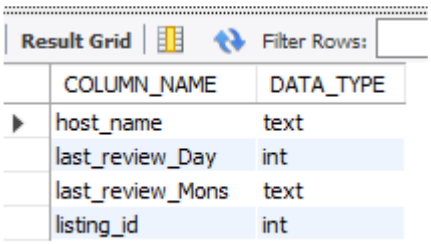
Airbnb exploratory analysis

Q1 what is the Data type of each column in all the 3 table.

#column - airbnb_last_review :

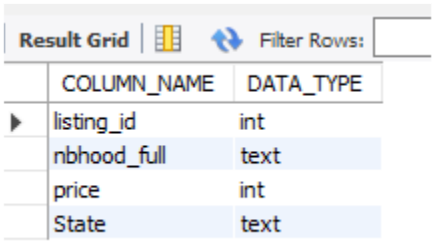
```
select column_name ,data_type
from INFORMATION_SCHEMA.COLUMNS
where table_schema = 'airbnb'
and table_name ='airbnb_last_review' ;
```



COLUMN_NAME	DATA_TYPE
host_name	text
last_review_Day	int
last_review_Mons	text
listing_id	int

column - airbnb_price:

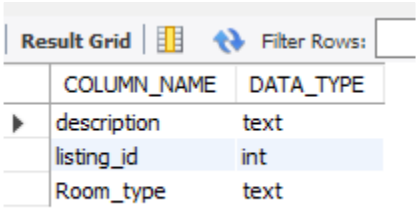
```
select column_name ,data_type
from INFORMATION_SCHEMA.COLUMNS
where table_schema = 'airbnb'
and table_name ='airbnb_price' ;
```



COLUMN_NAME	DATA_TYPE
listing_id	int
nbhood_full	text
price	int
State	text

column - airbnb_room_type:

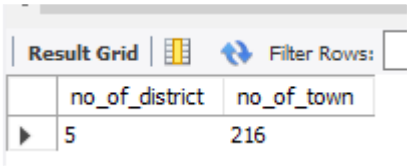
```
select column_name ,data_type
from INFORMATION_SCHEMA.COLUMNS
where table_schema = 'airbnb'
and table_name ='airbnb_room_type' ;
/*
```



COLUMN_NAME	DATA_TYPE
description	text
listing_id	int
Room_type	text

Q. count the number of unique airbnb present in different district and town of new york. */

```
select count( distinct(nbhood_full)) as no_of_district ,
    count(distinct(State)) as no_of_town
from airbnb_price ;
```

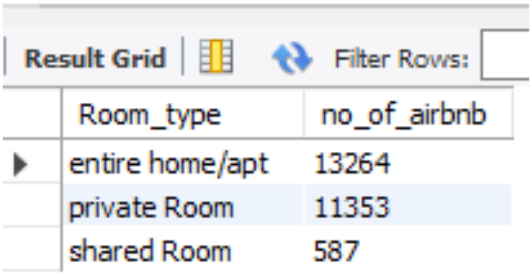


no_of_district	no_of_town
5	216

Insight: In 5 district we have 216 town huge opportunity to expand business here .

Q. Based on room type how many airbnb are present in whole new york.

```
select Room_type ,
 count(listing_id) as no_of_airbnb
from airbnb_room_type
group by Room_type
having count(listing_id) > 1
```

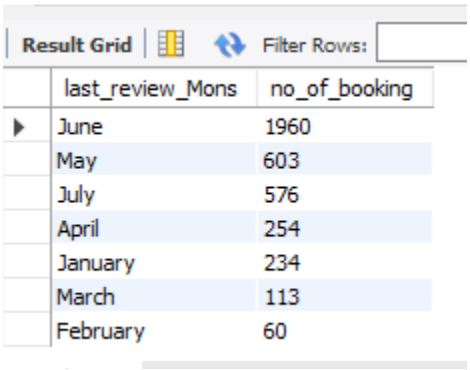


Room_type	no_of_airbnb
entire home/apt	13264
private Room	11353
shared Room	587

Insight: we can see that most of the people refer entire home or private room , so we can assume that people are more comfortable in private space. By decreasing the shared room we can adjust the revenue.

/*
Q. Monthly seasonality in terms of no of booking is done. (so it will give at what time the booking are in peak) */

```
select last_review_Mons ,
    count(listing_id) as no_of_booking
from airbnb_last_review
group  by last_review_Mons
order by no_of_booking desc ;
```



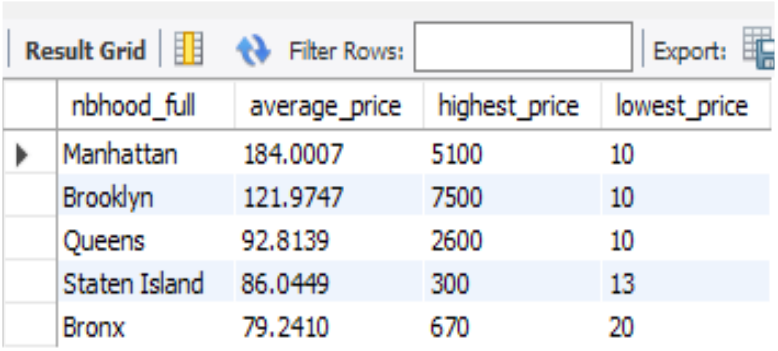
last_review_Mons	no_of_booking
June	1960
May	603
July	576
April	254
January	234
March	113
February	60

Insight: The booking in the month of June is more compare to other month , maybe tourist are prefer some particular season to travel. By increasing brokerage fee and introducing some good offer we can make some good profit here.

/*

Q. what is the average, expensive and cheap cost airbnb present in new york city based on district. */

```
select nbhood_full ,
       avg(price) as average_price ,
       max(price) as highest_price ,
       min(price) as lowest_price
from airbnb_price
where price != 0
group by nbhood_full ;
```



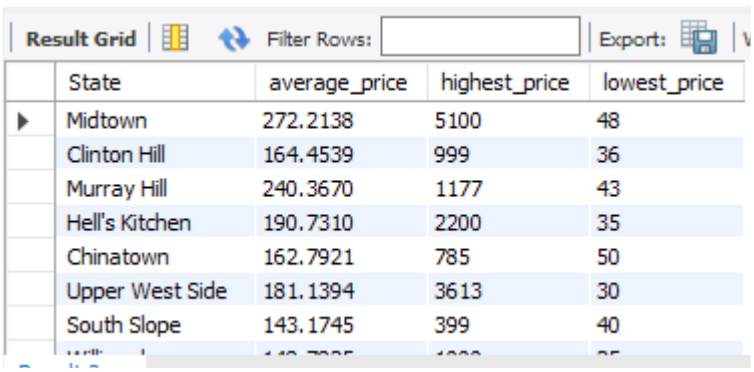
Result Grid | Filter Rows: | Export:

nbhood_full	average_price	highest_price	lowest_price
Manhattan	184.0007	5100	10
Brooklyn	121.9747	7500	10
Queens	92.8139	2600	10
Staten Island	86.0449	300	13
Bronx	79.2410	670	20

Insight: A middle class person can easily afford Airbnb present in new York city . It indicate that people prefer Airbnb more compare to other company because it is budget friendly.

Q . How the airbnb host are distributed all across the new york city.

```
select distinct nbhood_full ,
       count(listing_id) over (partition by nbhood_full) as no_of_host
from airbnb_price ;
```



Result Grid | Filter Rows: | Export:

State	average_price	highest_price	lowest_price
Midtown	272.2138	5100	48
Clinton Hill	164.4539	999	36
Murray Hill	240.3670	1177	43
Hell's Kitchen	190.7310	2200	35
Chinatown	162.7921	785	50
Upper West Side	181.1394	3613	30
South Slope	143.1745	399	40

#Q. what are the top 5 expensive airbnb and top 5 cheap airbnb based on state.

```
( select nbhood_full ,
      State ,
    listing_id ,
    'Expensive' as expensive_or_cheap,
    max(price) as expensive_airbnb
from airbnb_price
group by nbhood_full , State ,listing_id
order by expensive_airbnb desc
limit 5  )

union

(select nbhood_full ,
     State ,
     listing_id ,
     'Cheap' as expensive_or_cheap,
     min(price) as expensive_airbnb
from airbnb_price
where price != 0
group by nbhood_full , State ,listing_id
order by expensive_airbnb asc
limit 5 )
```

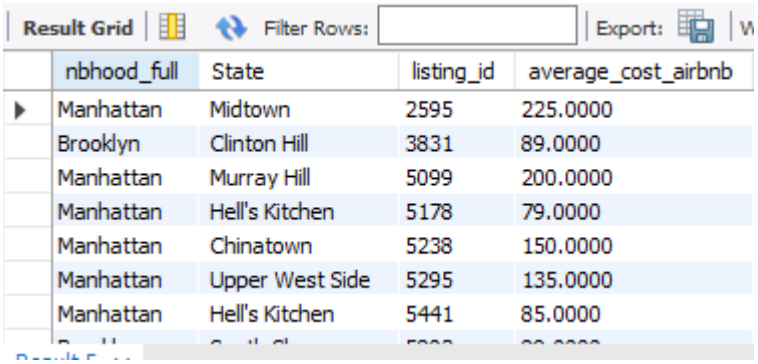
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

nbhood_full	State	listing_id	expensive_or_cheap	expensive_airbnb
Brooklyn	East Flatbush	34895693	Expensive	7500
Manhattan	Midtown	33397385	Expensive	5100
Manhattan	Harlem	30035166	Expensive	5000
Manhattan	Upper West Side	33029434	Expensive	3613
Manhattan	SoHo	22263855	Expensive	3000
Brooklyn	Bushwick	35642891	Cheap	10
Brooklyn	Greenpoint	21869057	Cheap	10

Insight:
According to above data we can assume that people prefer Brooklyn more compare to other because it have most expensive Airbnb to most cheapest Airbnb.

Q. what is the average cost of airbnb in different district of new york

```
select nbhood_full,
    State ,
    listing_id ,
    avg(price) as average_cost_airbnb
from airbnb_price
group by nbhood_full , State ,listing_id
```



Result Grid | Filter Rows: | Export:

nbhood_full	State	listing_id	average_cost_airbnb
Manhattan	Midtown	2595	225.0000
Brooklyn	Clinton Hill	3831	89.0000
Manhattan	Murray Hill	5099	200.0000
Manhattan	Hell's Kitchen	5178	79.0000
Manhattan	Chinatown	5238	150.0000
Manhattan	Upper West Side	5295	135.0000
Manhattan	Hell's Kitchen	5441	85.0000

#Q. based on room type what is the average , highest and lowest price airbnb available in new York.

```
select Room_type , avg(price) as average_price , max(price) as highest_price , min(price) as lowest_price
from airbnb_price a
join airbnb_room_type  b
on a.listing_id = b.listing_id
where price != 0
group by Room_type
limit 3 ;
```