Analysis of London Airbnb Data

Data source

- http://insideairbnb.com/

Data processing

- Study the file

  - Neighourhood

  - Room type

  - Price

  - Minimum nights

  - Number of reviews

  - Last reviews

	A	B	C	D	Re
	Field	**Type**	**Calculated**	**Description**	
	id	integer		Airbnb's unique identifier for the listing	
	name	string			
	host_id	integer			
	host_name	string			
	neighbourhood_group	text	y	The neighbourhood group as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.	
	neighbourhood	text	y	The neighbourhood as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.	
	latitude	numeric		Uses the World Geodetic System (WGS84) projection for latitude and longitude.	
	longitude			Uses the World Geodetic System (WGS84) projection for latitude and longitude.	
	room_type	string			
	price	currency		daily price in local currency. Note, $ sign may be used despite locale	
	minimum_nights	integer		minimum number of night stay for the listing (calendar rules may be different)	
	number_of_reviews	integer		The number of reviews the listing has	
	last_review	date	y	The date of the last/newest review	
	calculated_host_listings_count	integer	y	The number of listings the host has in the current scrape, in the city/region geography.	
	availability_365	integer	y	avaliability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may be available because it has been booked by a guest or blocked by the host.	
	number_of_reviews_ltm	integer	y	The number of reviews the listing has (in the last 12 months)	
	license	string			

Data processing

- Put the data to correct field

- Remove data which may be difficult to be cleaned

- Add news columns and create grouping (number of review, number of minimum nights, Inner/ Outer London) to the data

- Only use data with last review in 2021 and 2022, and price larger than 0
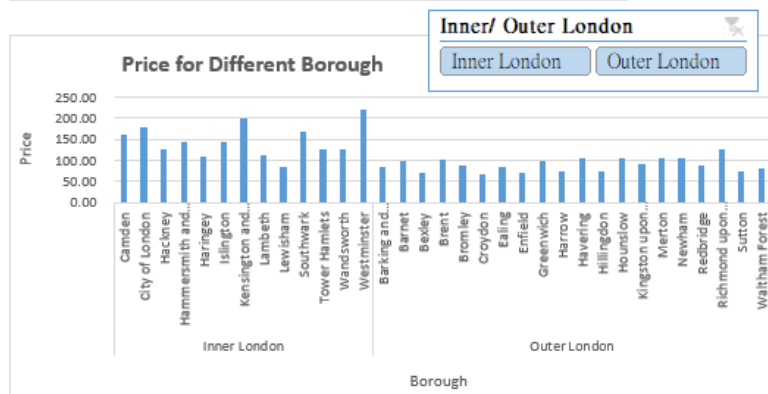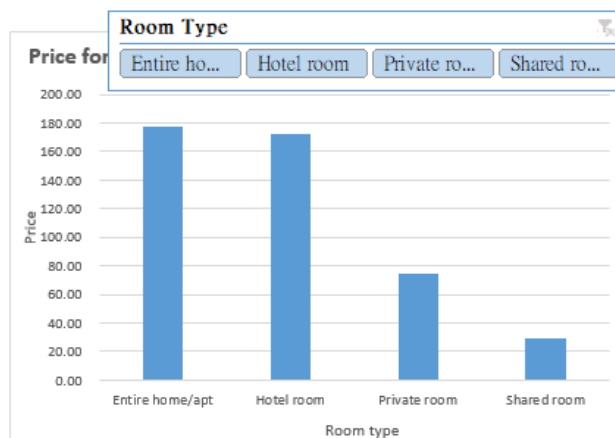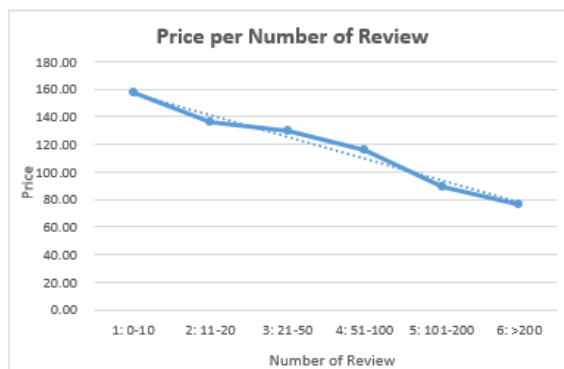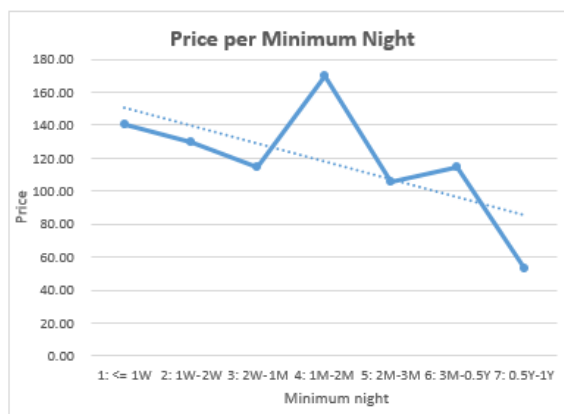
Finding

- Rental will be lower if minimum nights is larger.

- Rental in Inner London is larger than Outer London

- Rental is lower for owner with more review (Note: the relation may be caused by other factor, say number of review is correlated to whether the property was placed in Airbnb recently or long time ago (which may related to when the property was bought), such that the rental would be higher for property which is bought more recently.)