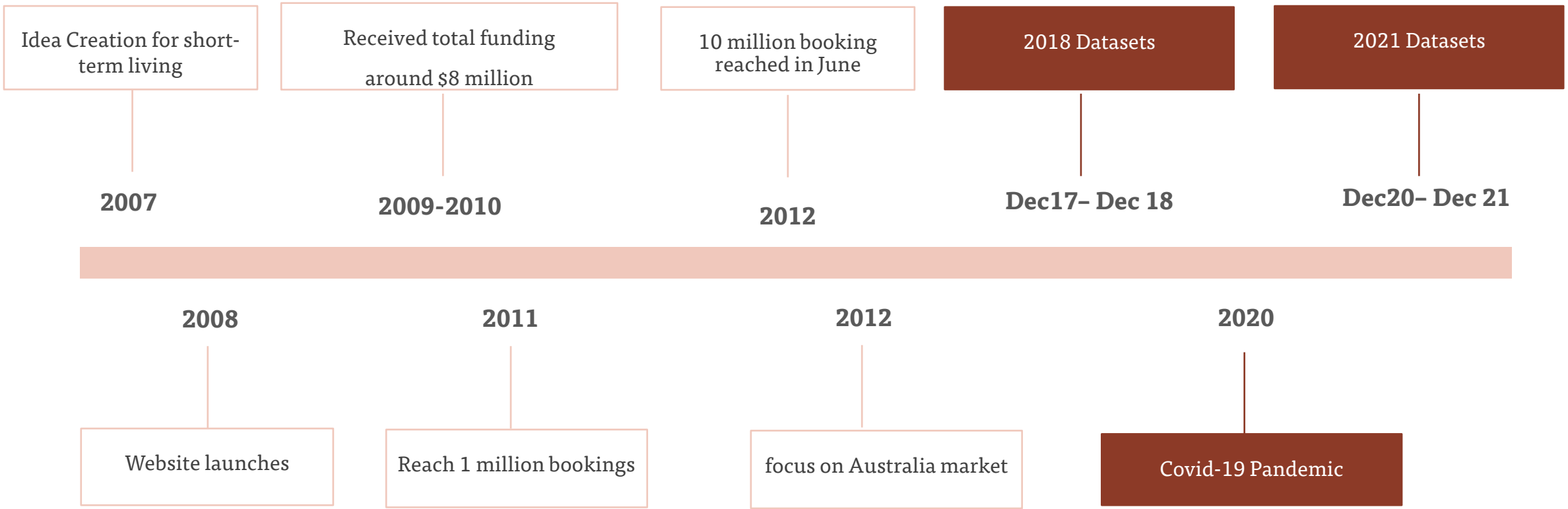
Comparing Airbnb listing price between 2018 and 2021

Clivia Kong

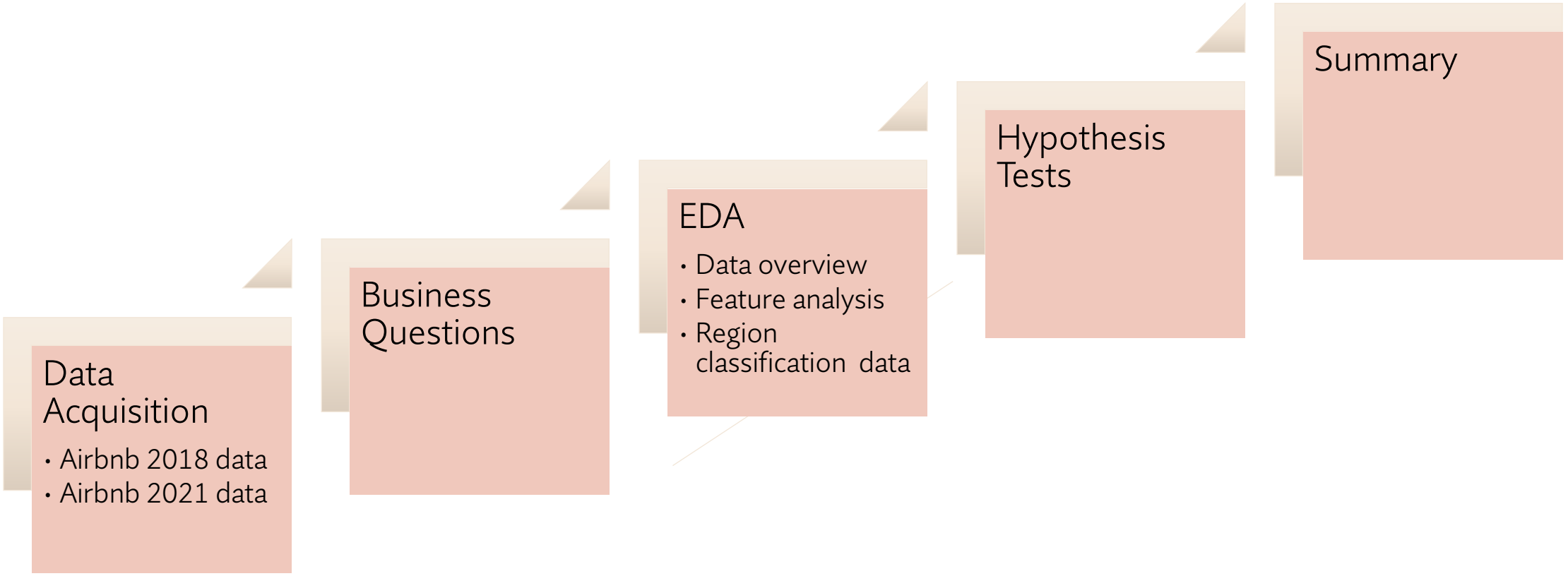
Timeline of Airbnb



Business Question

Is there any difference in the price between
“Period from Dec17 to Dec18” and “Period from Dec20 to Dec21(Covid-19 pandemic)” ?

Data Pipeline



2018 Data Set and 2021 Data Set of Sydney Airbnb

Key Features

- Id : Airbnb product ID (unique)
- Neighbourhood : suburb
- Room type: Entire home/apt, Private Room, Shared room, Hotel room
- Price: Daily price of listing

	id	name	host_id	host_name	neighbourhood	latitude	longitude	room_type	price
0	11156	An Oasis in the City	40855	Colleen	Sydney	-33.86767	151.22497	Private room	65
1	14250	Manly Harbour House	55948	Heidi	Manly	-33.80084	151.26378	Entire home/apt	470

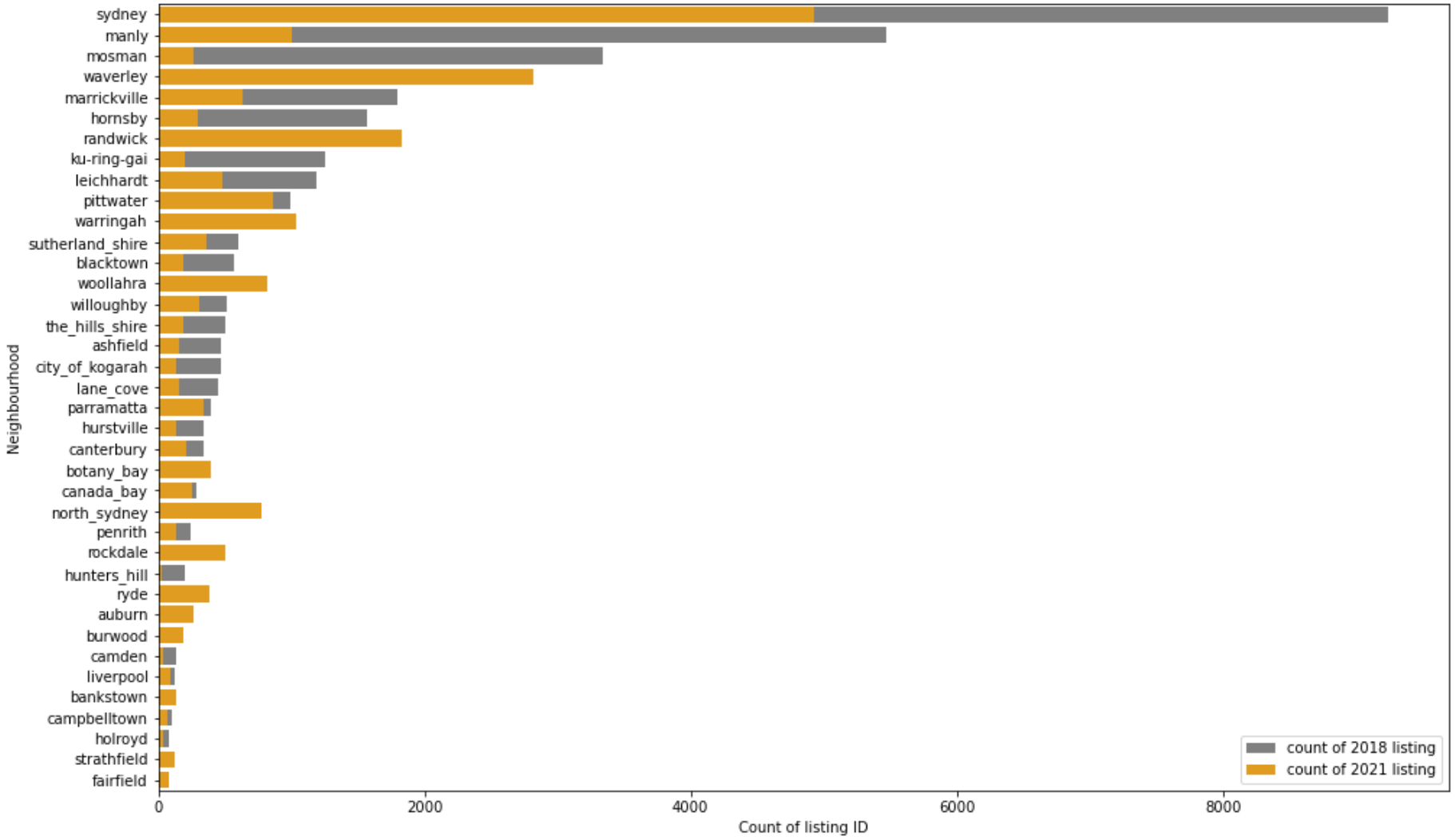
2018 Dataset:https://www.kaggle.com/datasets/tylerx/sydney-airbnb-open-data
2021 Dataset: http://insideairbnb.com/get-the-data/
Region classification table: https://sydneysuburbreviews.com/suburb-rankings/

Region classification table

- Name: suburb
- Region: region classification of suburbs

	Name	Region
0	hornsby	Upper North Shore
1	oatley	Southern Suburbs
2	dulwich_hill	Inner West
3	jannali	Sutherland Shire
4	waverton	Lower North Shore

2018 & 2021 Airbnb Data - Neighborhood



2018 Dataset: Dec17- Dec18

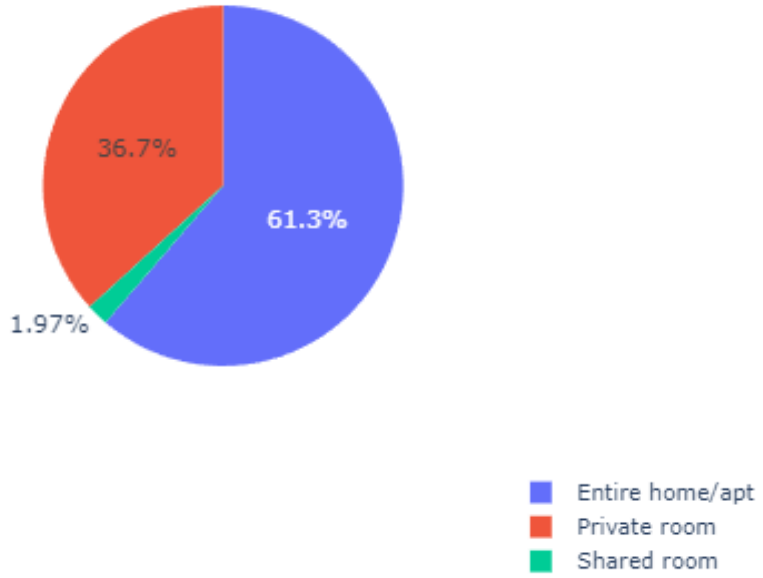
```
Number of Sydney Airbnb Listings: 36662
Number of features in dataset: 16
Avarage price is : $ 208
Number of hosts/suppliers:  27219
Number of suburbs:  38
```

2021 Dataset: Dec 20 – Dec21

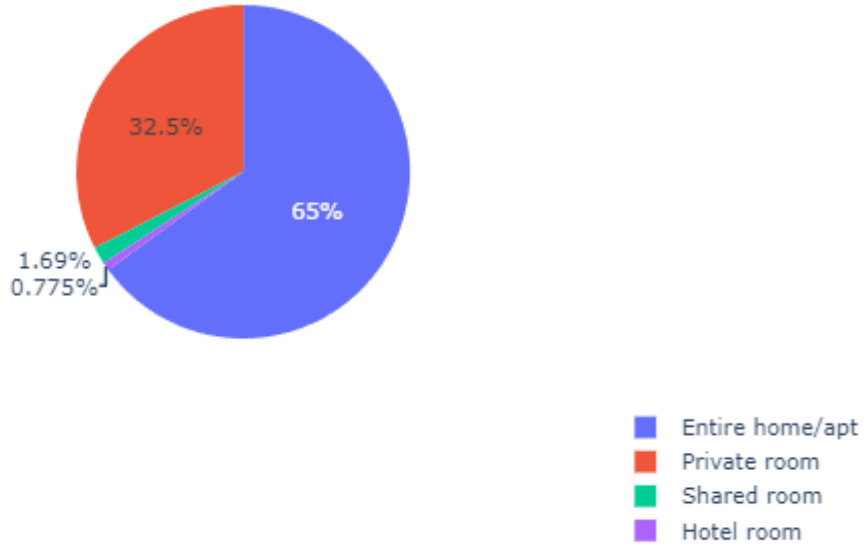
```
Number of Sydney Airbnb Listings: 20880
Number of features in dataset: 18
Avarage price is : $ 251
Number of hosts/suppliers:  15115
Number of suburbs:  38
```

2018 & 2021 Airbnb Data – Room Type of listing

2018 Dataset: Dec17- Dec18

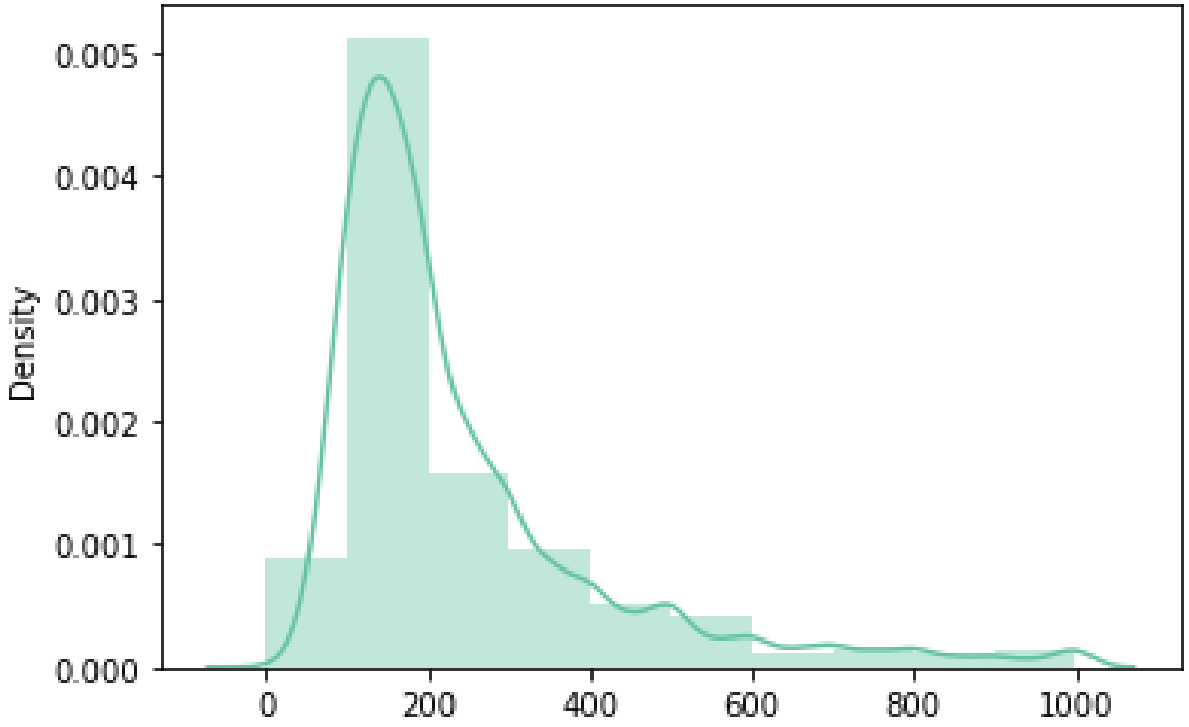


2021 Dataset: Dec 20 – Dec21

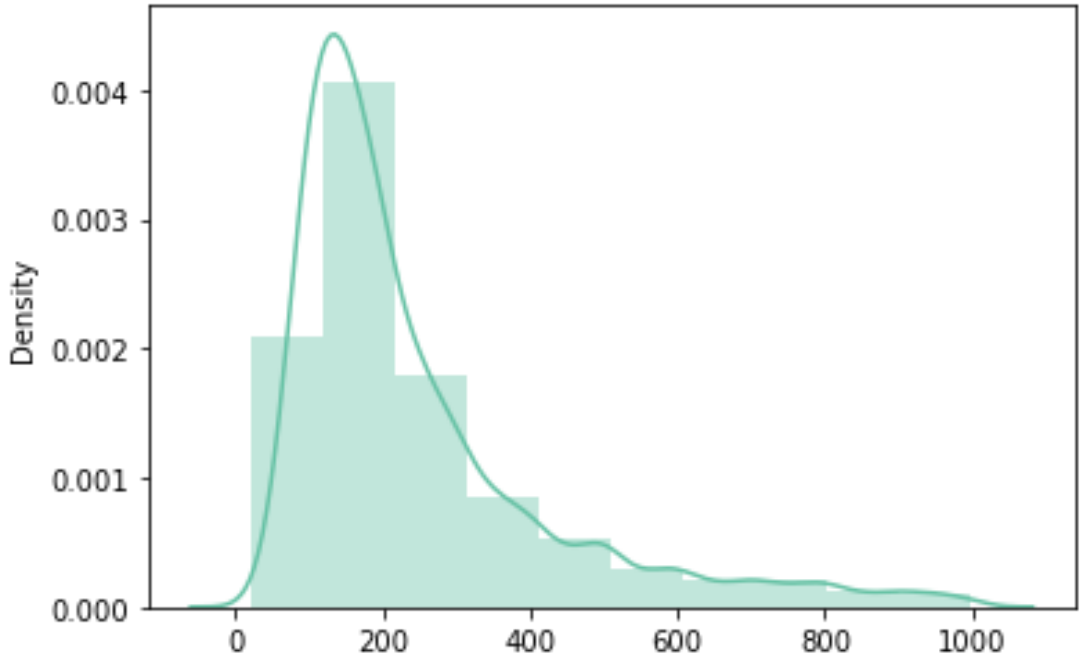


2018 & 2021 Airbnb Data – Price (Entire home/apt)

2018 Dataset: Dec17- Dec18, price < 1000



2021 Dataset: Dec 20 – Dec21, price < 1000

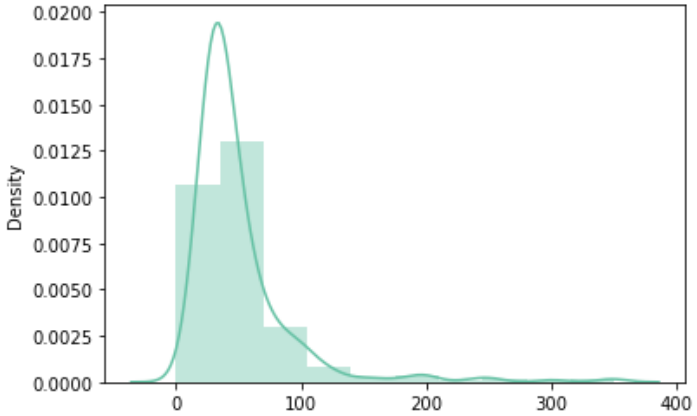


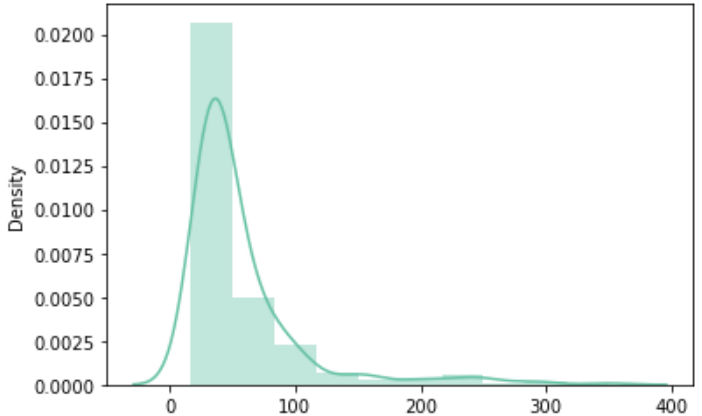
2018 & 2021 Airbnb Data – Price (Shared/Private)

2018 Dataset: Dec17- Dec18, price < 1000

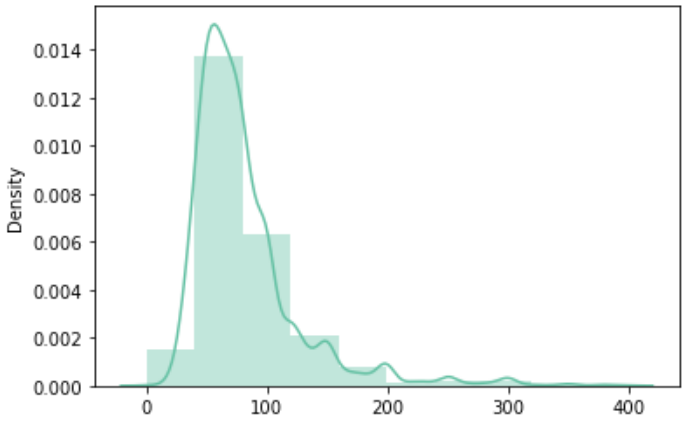
2021 Dataset: Dec 20 – Dec21, price < 1000

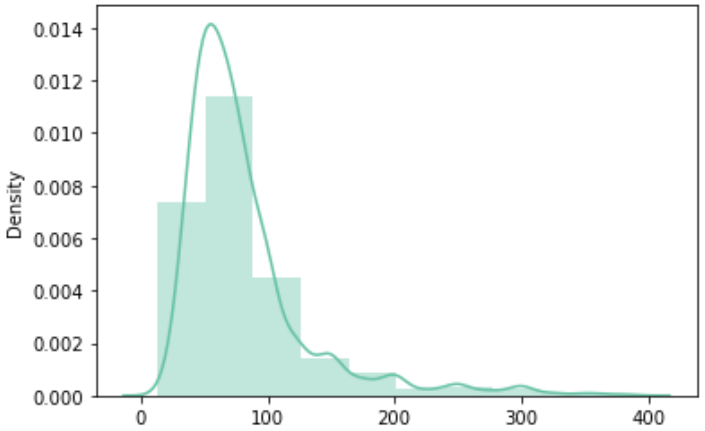
Shared room





Private room

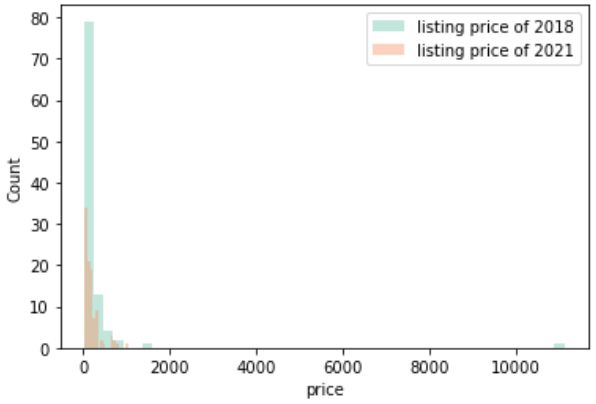




1. Is there any difference in the price between
"Period from Dec17 to Dec18" and "Period from Dec20 to Dec21"

- H0: There is no difference in the price between 2018 listings and 2021 listings
- H1: There is a difference in the price between 2018 listings and 2021 listings
- Confidence level: 95%
- Sample size: 100

```
t = -0.9143260108128004
p = 0.3616576539888706
We fail to reject our null hypothesis, accept H0 statement
There is no difference in the price between 2018 listings and 2021 listings
```



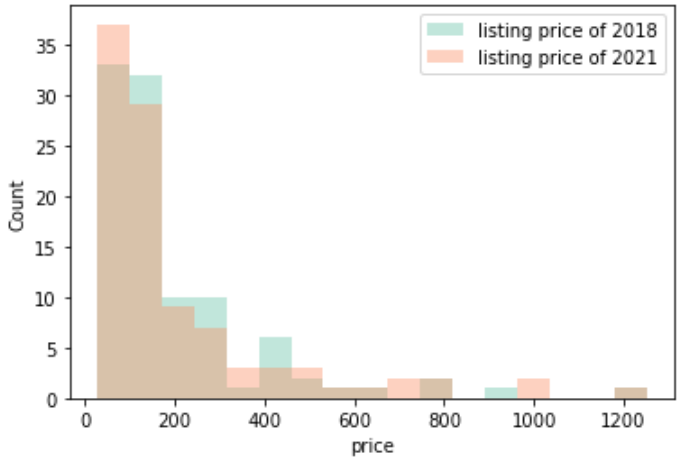
2. Is there any difference in the price between

"Period from Dec17 to Dec18" and "Period from Dec20 to Dec21"

For listings available in both periods?

- H0: There is no difference in the price between 2018 listings and 2021 listings in dataset with same listing IDs
- H1: There is a difference in the price between 2018 listings and 2021 listings in dataset with same listing IDs
- Confidence level: 95%
- Sample size: 100

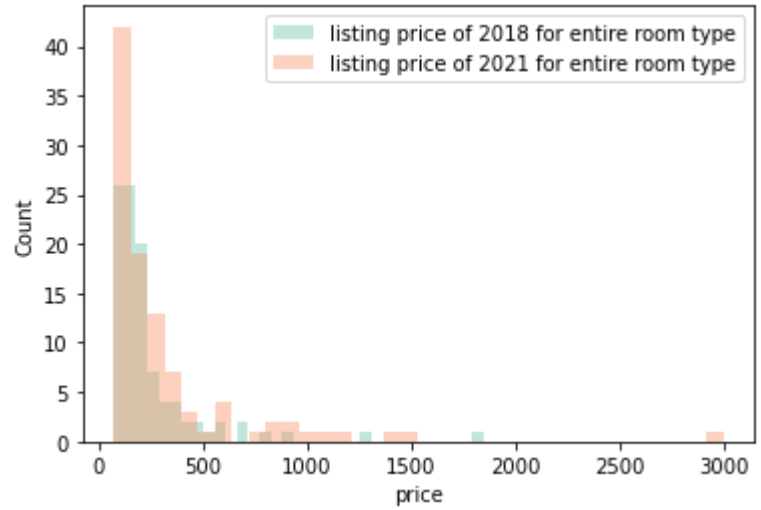
```
t = [-0.48977309]
p = [0.62483668]
We fail to reject our null hypothesis, accept H0 statement
There is no difference in the price between 2018 listings and 2021 listings
```



3. Is there any difference in the price between "Period from Dec17 to Dec18" and "Period from Dec20 to Dec21" For Entire home type group?

- H0: There is no difference in the price between 2018 listings and 2021 listings of entire home group
- H1: There is a difference in the price between 2018 listings and 2021 listings of entire home group
- Confidence level: 95%
- Sample size: 100

```
t = -1.4990692055535604
p = 0.13544837226885142
We fail to reject our null hypothesis, accept H0 statement
There is no difference in the price between 2018 listings and 2021 listings for entire home listings
```



4. Is there any difference in the price between "Period from Dec17 to Dec18" and "Period from Dec20 to Dec21" For Non-entire home type group?

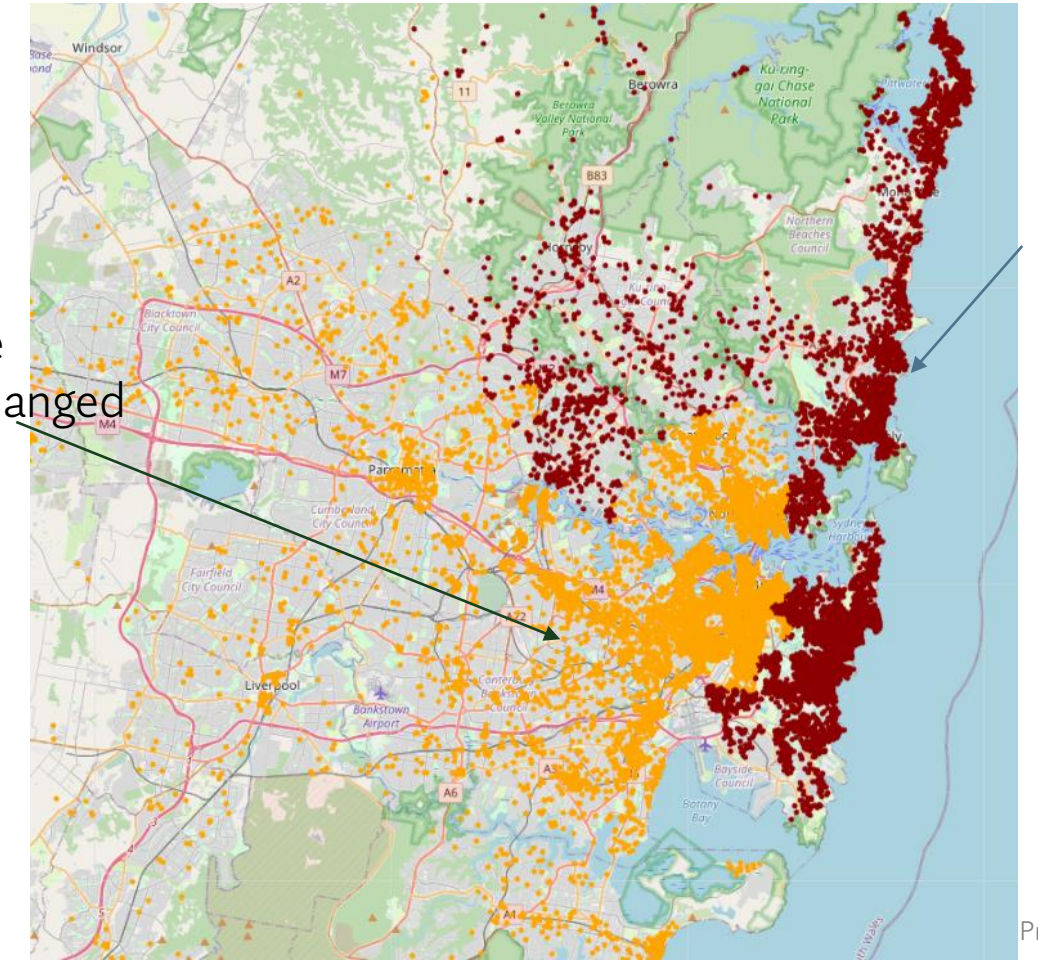
- H0: There is no difference in the price between 2018 listings and 2021 listings of Non-entire home group
- H1: There is a difference in the price between 2018 listings and 2021 listings of Non-entire home group
- Confidence level: 95%
- Sample size: 100

```
t = -0.8975420201617832
p = 0.3705198199395001
We fail to reject our null hypothesis, accept H0 statement
There is no difference in the price between 2018 listings and 2021 listings for Non-entire home listings
```

5. Is there any difference in the price between

"Period from Dec17 to Dec18" and "Period from Dec20 to Dec21"

For each region group?



Price changed

Price unchanged

Region	Result
Sydney_cbd	H0
Northern Beaches	H1
North Shore	H1
Eastern Suburbs	H1
Inner West	H0
Upper North Shore	H1
Northern Suburbs	H1
Southern Suburbs	H0
Western Suburbs	H0
Lower North Shore	H0
Northern West	H0
Northern West	H0

Summary

- The number of listing ID in 2021 is significantly lower than the listing ID in 2018.
- The listing price of 2021 is similar with listing price of 2018
- The listing price of Northern and Eastern suburbs are changed`

Next Steps:

- Further analysis the regions that have the difference by suburbs through hypothesis testing
- Identify the price change direction of the regions/suburbs which have price changed
- Analysis the reason causes the price change in these regions/suburbs
- Compare each month of transactional price between two periods to see whether price changed in specific month
- Analysis the correlation of price with other features to find out what drives the price change