DQLab

Live Class

airbnb

# FINAL PROJECT

By : Kenneth Wahyudi, S.Si.

This project was created as the final project of the DQLab Bootcamp Data Analyst with Python & SQL (Batch No. 2) that I participated in.

I especially want to thank DQLab, the mentors, and the assistant mentors who have taught and guided me all the way up to this final project.

# Background

Airbnb, Inc. is an American San Francisco-based company operating an online marketplace for short-term homestays and experiences, and has listings in over 220 countries and regions. For this project, the datasets provided was for listings in Singapore. So that's what we're going to look at. Specifically, listings from the year 2018 until 2022.

All analysis and visualization were done using Python and its packages, namely Pandas and Plotly. Link to the datasets and Python file are provided below.

https://drive.google.com/drive/folders/1GzXK1KXCZ2nfTYqWUuPoC6XTk6wXGO7l?usp=share_link

DQLab

Live Class

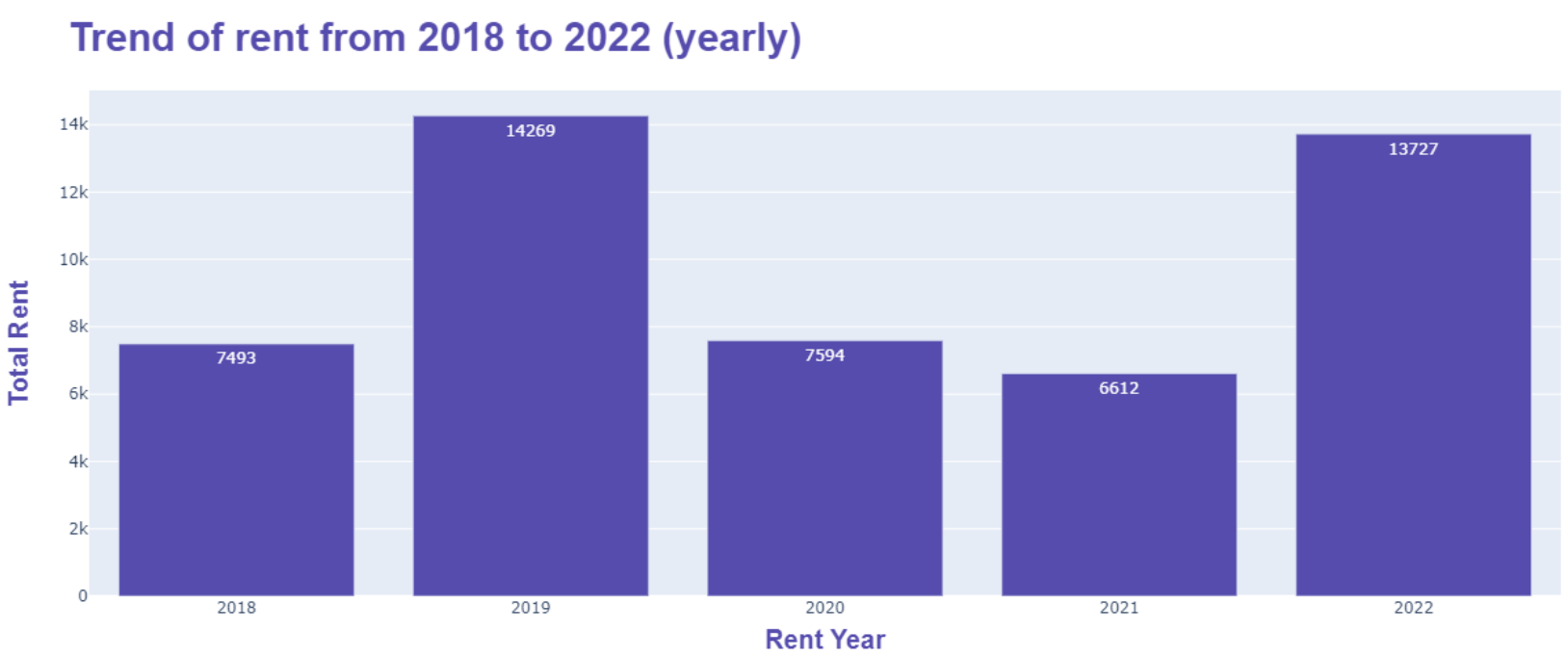
# The things we're going to look at :

- The listings. Specifically the trend of rents (monthly and yearly), and comparisons between active listings before and after early 2020.
- The prices. Specifically the distribution, mean and median, and average prices per neighborhood group
- The reviews. Specifically the top reviewed listings, reviews per room type, and reviews percentage growth by period year.

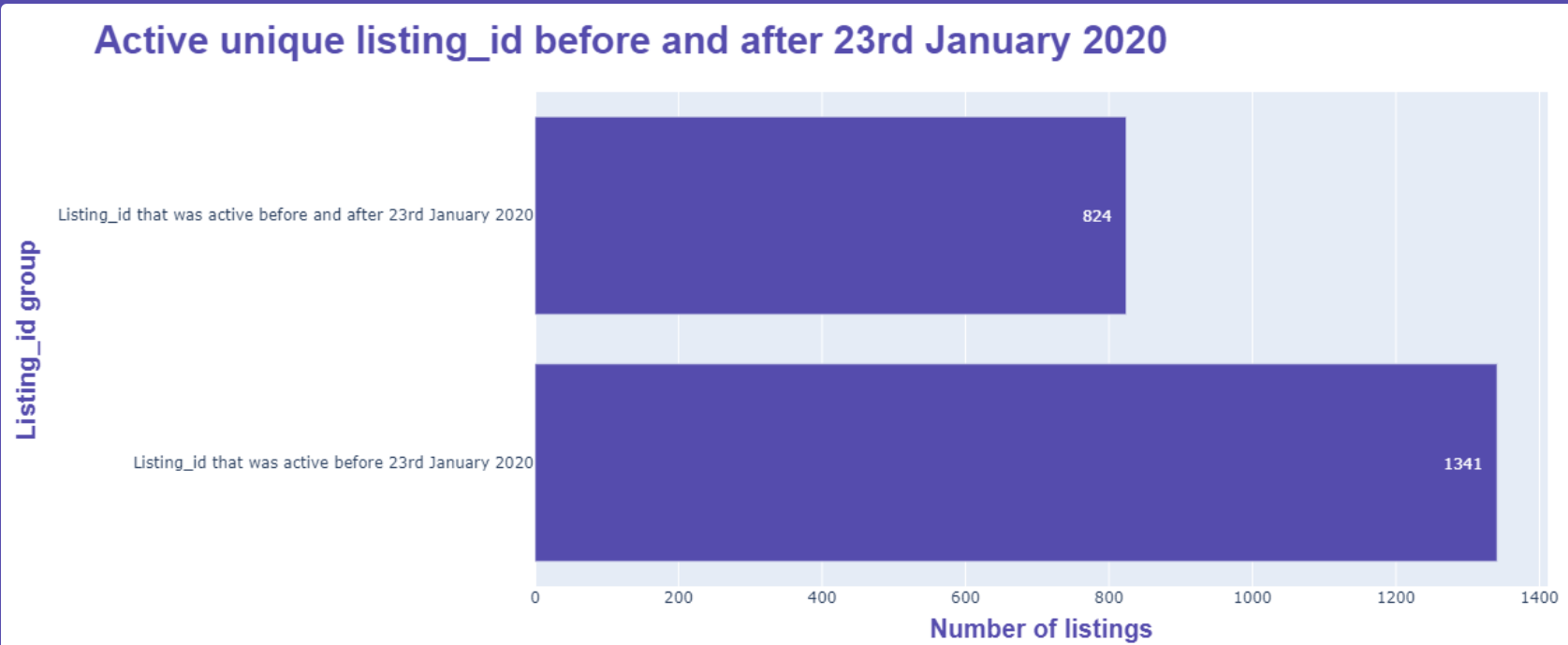
# Chart 1

Trend of rent from 2018 to 2022 (monthly)
Total Rent
2000
1500
1000
500
Jan 2018
Apr
Jul
Oct
Jan 2019
Apr
Jul
Oct
Jan 2020
Apr
Jul
Oct
Jan 2021
Apr
Jul
Oct
Jan 2022
Apr
Jul
Rent Month

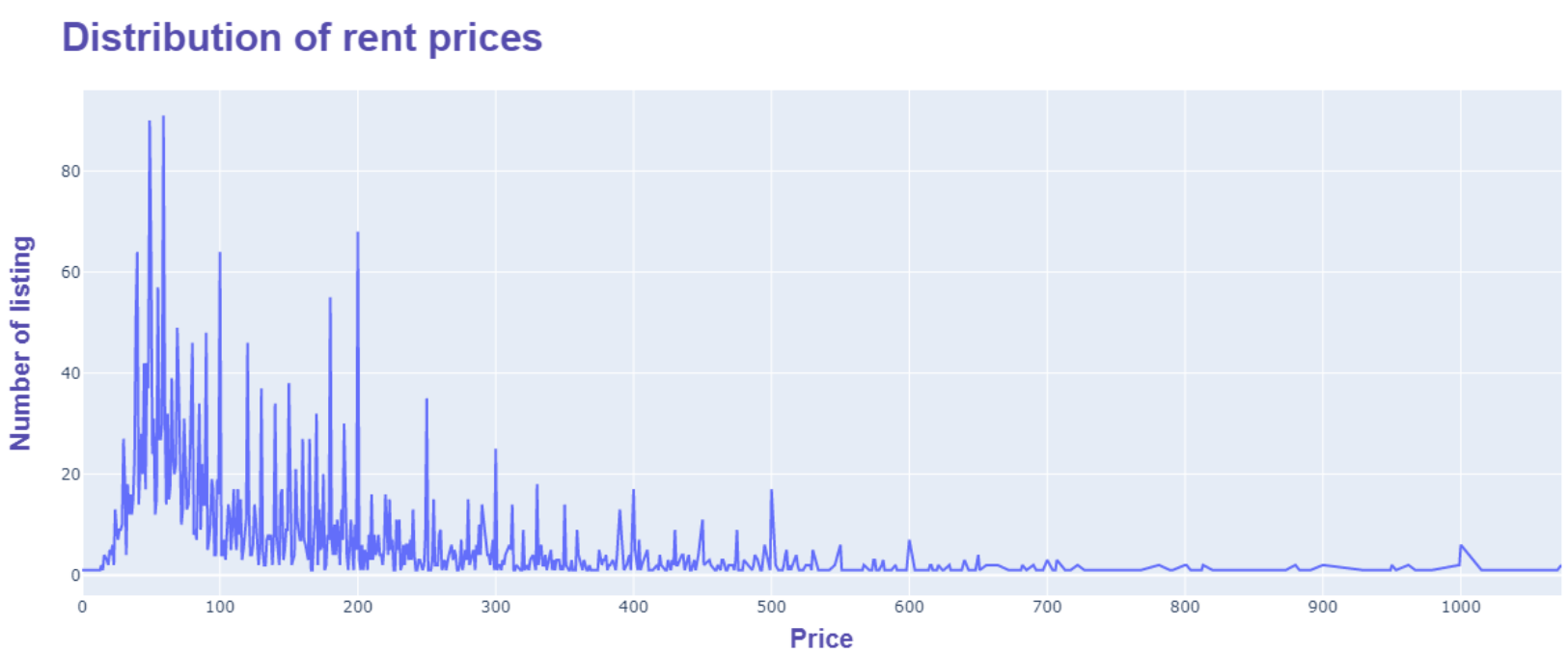
# Chart 2


Trend of rent from 2018 to 2022 (yearly)
Total Rent
14k
12k
10k
8k
6k
4k
2k
0
7493
14269
7594
6612
13727
2018
2019
2020
2021
2022
Rent Year

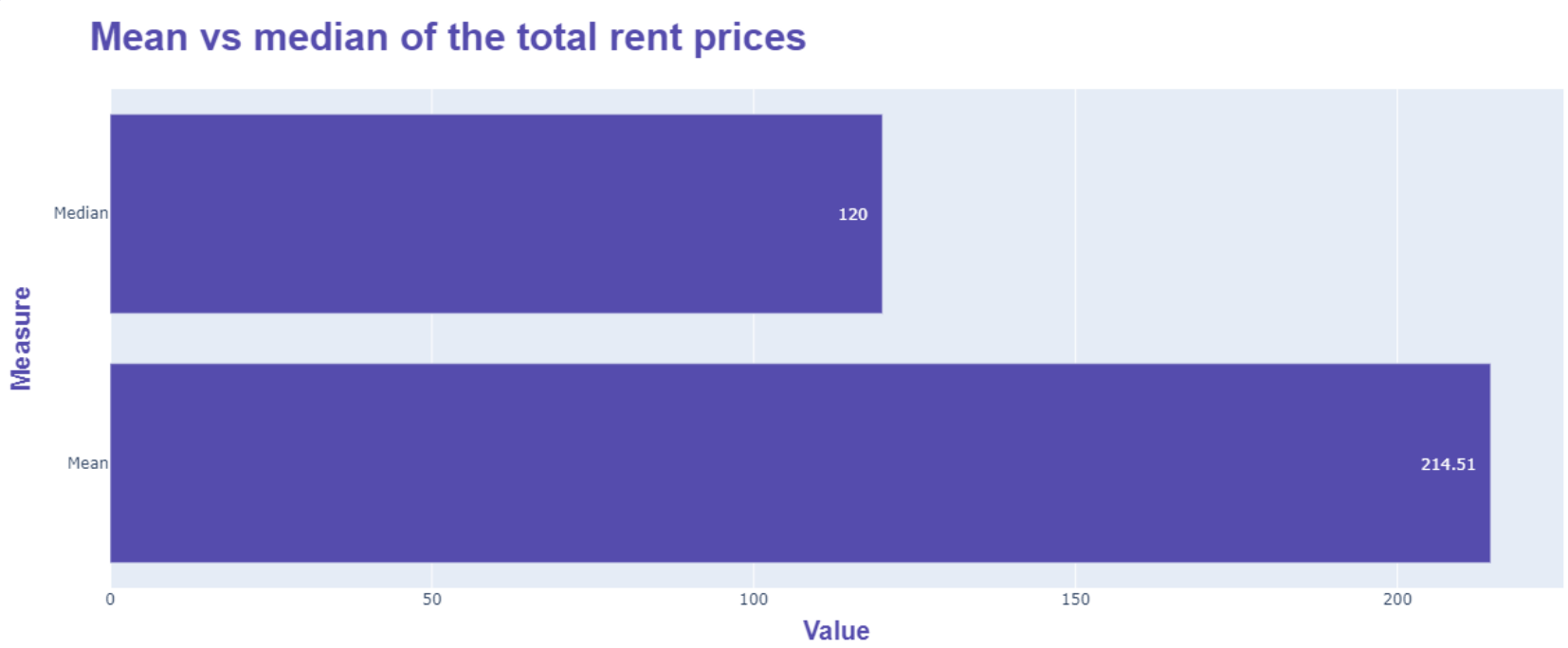
# Chart 3


Active unique listing_id before and after 23rd January 2020
Listing_id group
Listing_id that was active before and after 23rd January 2020
824
Listing_id that was active before 23rd January 2020
1341
0
200
400
600
800
1000
1200
1400
Number of listings

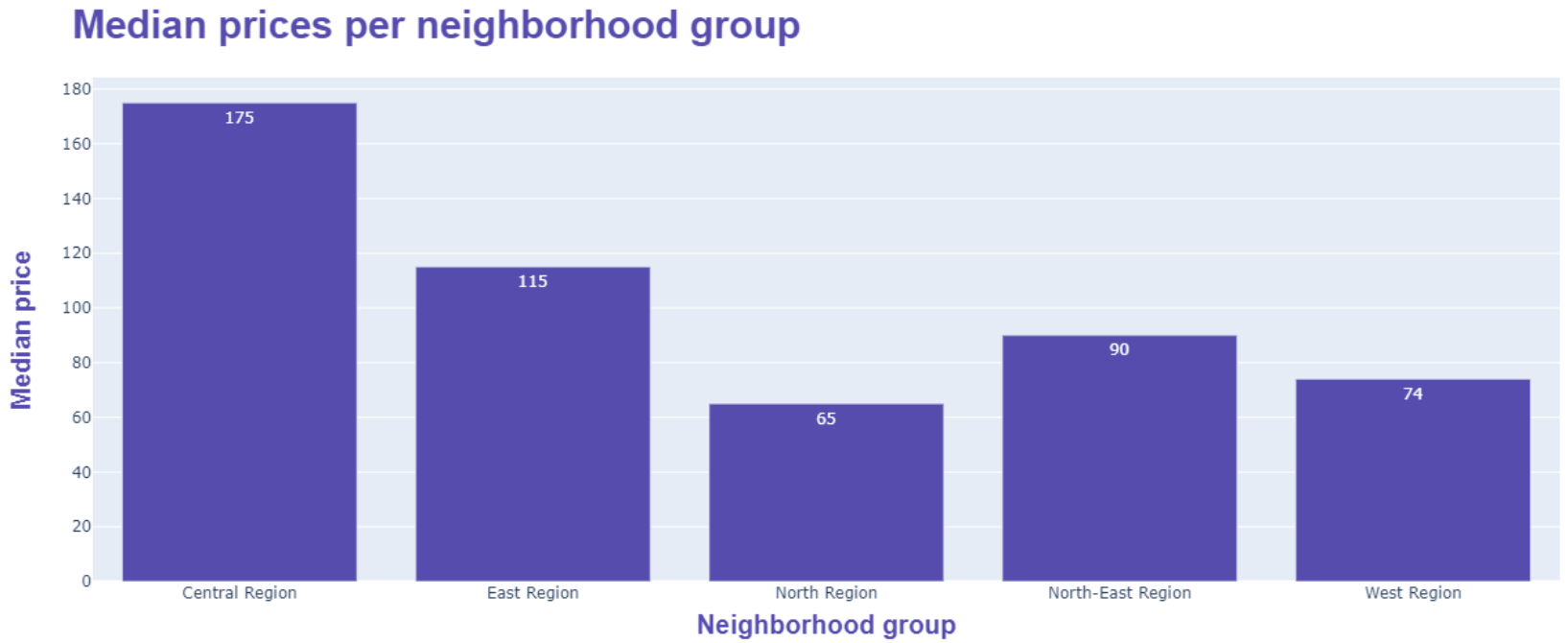
# Chart 4


Distribution of rent prices
Number of listing
0
20
40
60
80
0
100
200
300
400
500
600
700
800
900
1000
Price

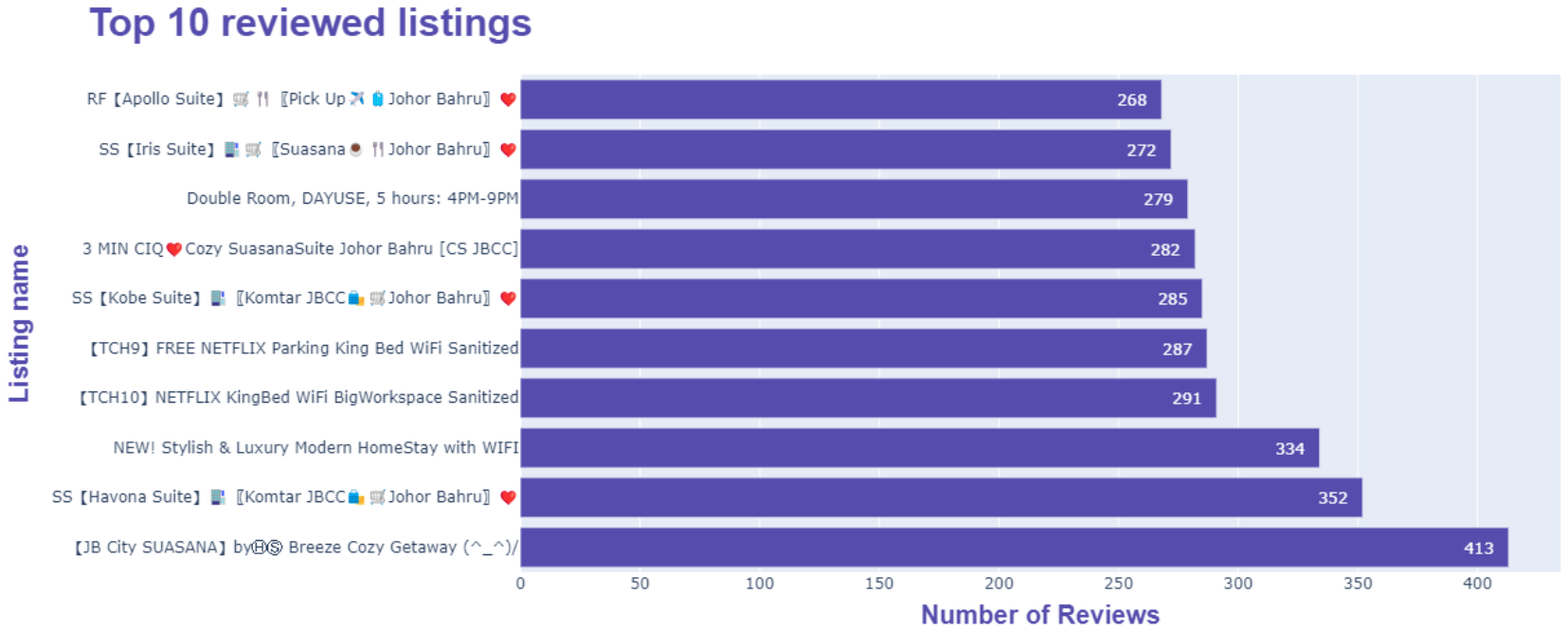
# Chart 5


Mean vs median of the total rent prices
Measure
Median
Mean
120
214.51
0
50
100
150
200
Value

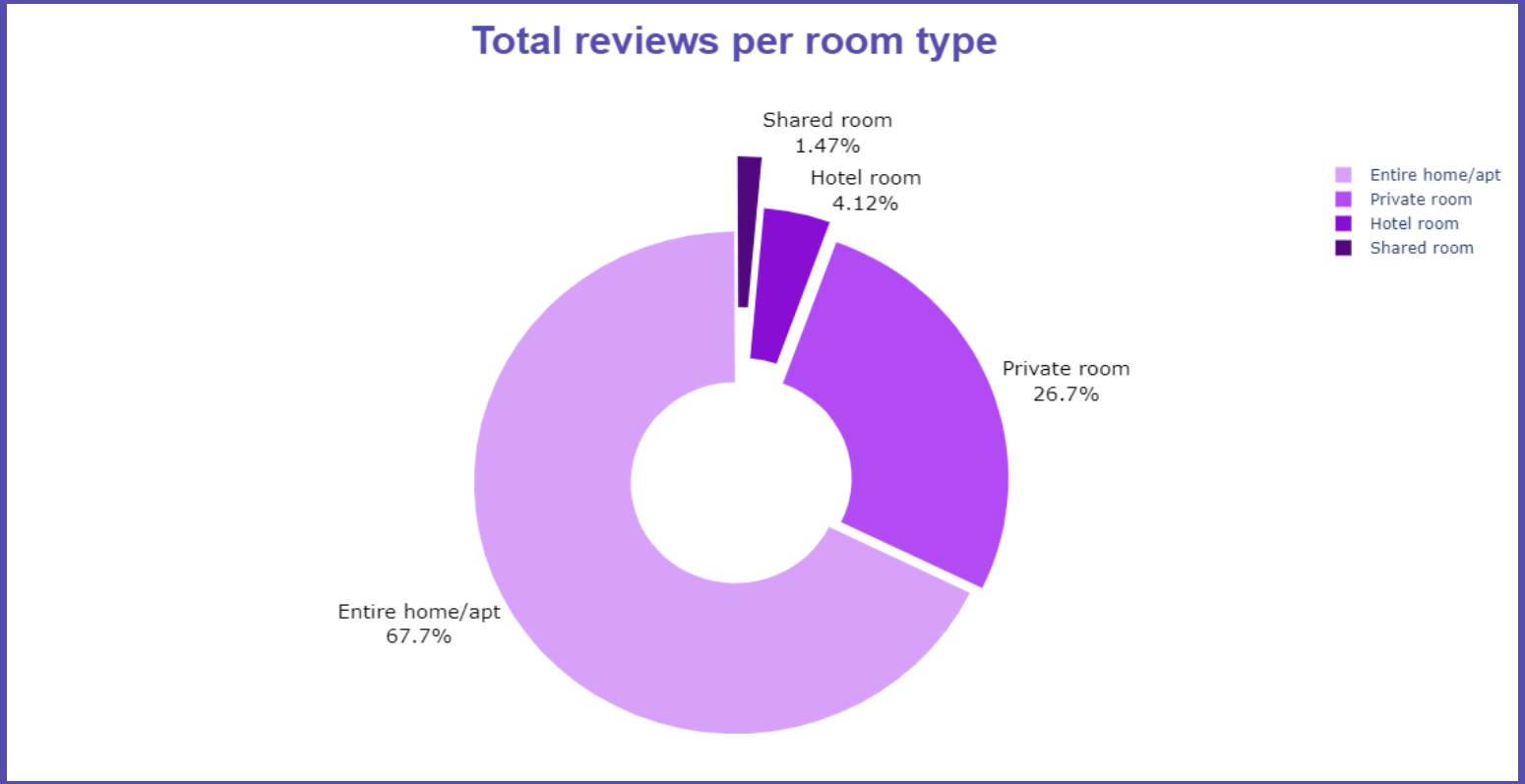
# Chart 6


Median prices per neighborhood group
Median price
180
160
140
120
100
80
60
40
20
0
175
115
65
90
74
Central Region
East Region
North Region
North-East Region
West Region
Neighborhood group

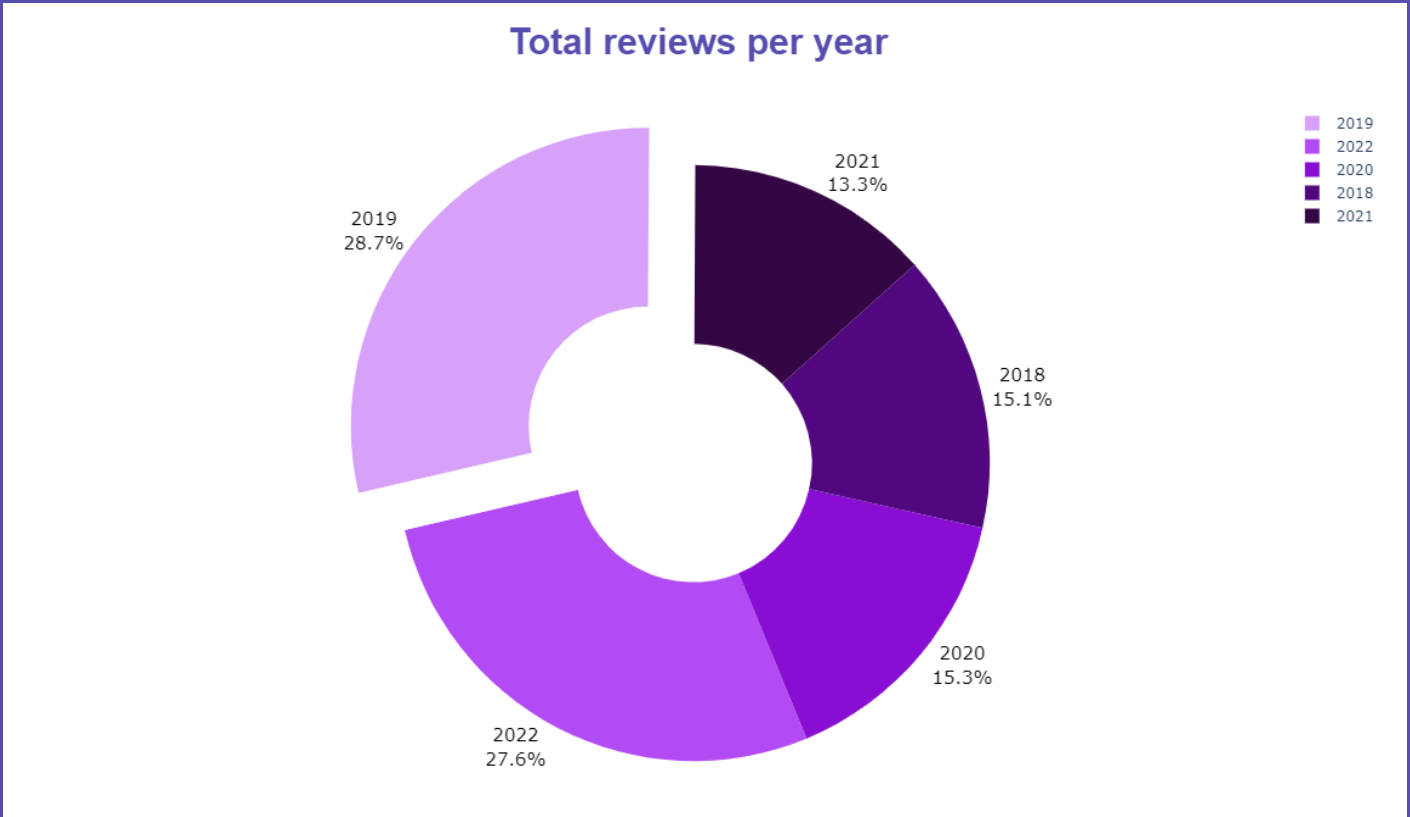
# Chart 7


Top 10 reviewed listings
Listing name
RF【Apollo Suite】🛒🍴〖Pick Up✈️🧳Johor Bahru〗❤️
SS【Iris Suite】🛒〖Suasana🤎🍴Johor Bahru〗❤️
Double Room, DAYUSE, 5 hours: 4PM-9PM
3 MIN CIQ❤️Cozy SuasanaSuite Johor Bahru [CS JBCC]
SS【Kobe Suite】〖Komtar JBCC🛒Johor Bahru〗❤️
【TCH9】FREE NETFLIX Parking King Bed WiFi Sanitized
【TCH10】NETFLIX KingBed WiFi BigWorkspace Sanitized
NEW! Stylish & Luxury Modern HomeStay with WIFI
SS【Havona Suite】〖Komtar JBCC🛒Johor Bahru〗❤️
【JB City SUASANA】by Breeze Cozy Getaway (^_^)/
268
272
279
282
285
287
291
334
352
413
0
50
100
150
200
250
300
350
400
Number of Reviews

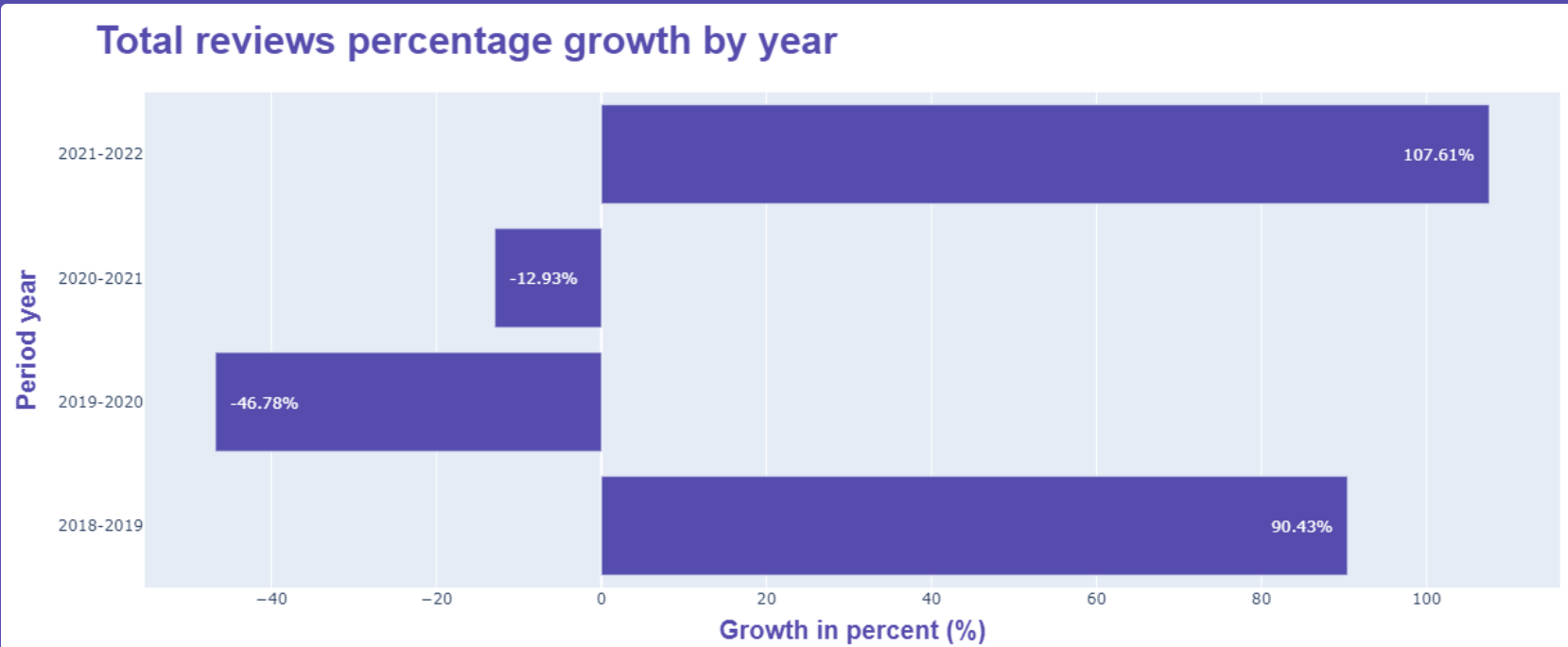
# Chart 8


Total reviews per room type
Shared room
1.47%
Hotel room
4.12%
Private room
26.7%
Entire home/apt
67.7%
Entire home/apt
Private room
Hotel room
Shared room

# Chart 9


Total reviews per year
2019
28.7%
2021
13.3%
2018
15.1%
2020
15.3%
2022
27.6%
2019
2022
2020
2018
2021

# Chart 10


Total reviews percentage growth by year
Period year
2021-2022
2020-2021
2019-2020
2018-2019
107.61%
-12.93%
-46.78%
90.43%
−40
−20
0
20
40
60
80
100
Growth in percent (%)

# Key Takeaways

- The listing trend in Chart 1 from 2018-2022 has some interesting points. The trend goes upward until late 2019, when it began to free fall from 1846 listings in Dec 2019, to just 180 listings in Apr 2020. Before it finally steadily rises again up to 2331 listings in Aug 2022.
- And in the yearly listings trend in Chart 2 , we can see that the year with the lowest activity is 2021 with 6612 listings, and followed closely by 2018 and 2020, despite the fact that 2019 is the year with the highest activity with 14269 listings (followed closely by 2022).

- Finally in Chart 3 , we can see that the number of unique listings that were active before 23rd of January 2020 was 1341, whilst the number of the same unique listings that were still active after that was decreased by around 38.55% to just 824 listings. This is due to the lockdown effect of the Covid-19 Pandemic hitting places across the world, and especially Singapore. So apparently, 38.55% of listings did not survive the pandemic. And that is also the reason why the listings trend took a dive by late 2019, and also why the year 2020 and 2021 are the lowest in terms of activity when compared to the previous relatively remarkable year 2019.

- Looking at Chart 4 , we can see that the distribution of the rent prices are skewed to the left. This is because there are a significant amount of listings that has prices way above the median. Outlier handling was done to the data prior to visualization, to make it more informative. Values of greater than 1.5 times the Inter-Quartile Range (IQR) were dropped (i.e. above 1074). The distribution shows that most listings lie in the price range of 0 to 200, with significant but a minority amount being upwards of 200.

- The mean and median of listing prices are shown in Chart 5 . As one would be able to see, because of the amount of listings with the price much greater than the median, the mean value is far higher than the median value. With the median being 120, and the mean being around 214.51, which is around 78.76% higher than the median. Therefore, it is much more safe, informative, and just wise overall to use the median instead of the mean.
- That being said, the median of prices per neighborhood group are shown in Chart 6 . With Central Region being the highest at 175, and North Region being the lowest at 65.

- In Chart 7 , we can see the top 10 most reviewed listings. With the most reviewed listing being “【JB City SUASANA】byⒽⓈ Breeze Cozy Getaway (^_^)/” with the listing id of 25360027, located in the woodlands neighborhood, which is part of the North Region neighborhood group.
- Moving on to Chart 8 , we can see that the most reviewed room type is the “Entire home/apt” room type, getting 67.7% of the total reviews share. And “Shared room” being the least reviewed. And it seems that room type does have an effect towards reviews as 9 out of 10 of the most reviewed listings that we discussed earlier have the room type of “Entire home/apt”.

- As we discussed previously with the listings trend per year, the same could be said for the reviews per year as shown in Chart 9 . With 2019 being the year with the most reviews at 28.7% out of the total reviews, followed closely by year 2022 at 27.6% (2 years after the pandemic). Whilst year 2021 has the least amount of reviews at just 13.3% out of the total reviews (1 year after the pandemic), followed closely by 2018 at 15.1%, and 2020 at 15.3%. The reason why 2020 has more activity than 2021 is because that at the early part of the year 2020 (Jan - April), there were still some activity because the lockdown isn't yet in full effect. The year 2021 however, sees lockdown worldwide almost the entire year.

- And finally, if we look at Chart 10 , we can clearly see the way the pandemic moves the regulations and people's desire to travel by the year. The period year 2019-2020 sees the highest number of negative growth to the amount of reviews received at -46.78%. And apparently the following period year, regulations are tighter and people's desire to travel are even lower than the previous period year because at the period year 2020-21, the growth is still negative at -12.93% despite already being negative the previous year. And the highest positive growth in percentage are in the other two period years with 2021-2022 being the highest at +107.61%, and followed by the pre-pandemic period year of 2018-2019 being at +90.43%.

Thanks!