

# SIRS\_eligibility\_report

## Context

In view of the COVID-19 situation, Singapore government introduced SELF-EMPLOYED PERSON INCOME RELIEF SCHEME (SIRS) to help self-employed Singaporeans. More context: <https://www.income.com.sg/blog/sirs-covid-19-support>

In order for one to be qualified for the scheme, one has to fulfill a list of requirements. full eligibility = <https://www.ntuc.com.sg/sirs/eligibility/>

I was asked by a friend regarding one of the eligibility requirement regarding property annual value. The criteria of interest is: I live in a property with an annual value of \$21,000 or less

Since annual value is computed by rental per month X 12 months, we are looking at flat with monthly rental  $\leq$  \$1750 I went around to look for Singapore's rental data. Calculation for Annual Value: <https://www.iras.gov.sg/IRASHome/Property/Property-owners/Learning-the-basics/About-Annual-Value/>

I used the median rent by town and flat type dataset in Singapore government's database. Data source: <https://data.gov.sg/dataset/median-rent-by-town-and-flat-type>

I have uploaded my codes and dataset here: [https://github.com/lingjie00/SIRS\\_HDB\\_eligibility](https://github.com/lingjie00/SIRS_HDB_eligibility)

## Data Overview

The dataset contains 4 columns:

1. Singapore rental date
2. Town of the rental flat
3. The kind of flat
4. The median rent

```
## 'data.frame':   9405 obs. of  4 variables:
## $ quarter      : chr  "2005-Q2" "2005-Q2" "2005-Q2" "2005-Q2" ...
## $ town         : chr  "ANG MO KIO" "ANG MO KIO" "ANG MO KIO" "ANG MO KIO" ...
## $ flat_type    : chr  "1-RM" "2-RM" "3-RM" "4-RM" ...
## $ median_rent : num  NA NA 800 950 NA NA NA NA 800 900 ...
```

The rental date ranged from 2005Q2 to 2020Q1

```
unique(rental$quarter)
```

```
## [1] "2005-Q2" "2005-Q3" "2005-Q4" "2006-Q1" "2006-Q2" "2006-Q3" "2006-Q4"
## [8] "2007-Q1" "2007-Q2" "2007-Q3" "2007-Q4" "2008-Q1" "2008-Q2" "2008-Q3"
## [15] "2008-Q4" "2009-Q1" "2009-Q2" "2009-Q3" "2009-Q4" "2010-Q1" "2010-Q2"
## [22] "2010-Q3" "2010-Q4" "2011-Q1" "2011-Q2" "2011-Q3" "2011-Q4" "2012-Q1"
## [29] "2012-Q2" "2012-Q3" "2012-Q4" "2013-Q1" "2013-Q2" "2013-Q3" "2013-Q4"
## [36] "2014-Q1" "2014-Q2" "2014-Q3" "2014-Q4" "2015-Q1" "2015-Q2" "2015-Q3"
## [43] "2015-Q4" "2016-Q1" "2016-Q2" "2016-Q3" "2016-Q4" "2017-Q1" "2017-Q2"
## [50] "2017-Q3" "2017-Q4" "2018-Q1" "2018-Q2" "2018-Q3" "2018-Q4" "2019-Q1"
## [57] "2019-Q2" "2019-Q3" "2019-Q4" "2020-Q1"
```

The towns includes

```
unique(rental$town)
```

```
## [1] "ANG MO KIO"      "BEDOK"           "BISHAN"          "BUKIT BATOK"
## [5] "BUKIT MERAH"    "BUKIT PANJANG"   "BUKIT TIMAH"     "CENTRAL"
## [9] "CHOA CHU KANG"   "CLEMENTI"        "GEYLANG"         "HOUGANG"
## [13] "JURONG EAST"     "JURONG WEST"     "KALLANG/WHAMPOA" "LIM CHU KANG"
## [17] "MARINE PARADE"    "PASIR RIS"       "PUNGGOL"         "QUEENSTOWN"
## [21] "SEMBAWANG"       "SENGKANG"        "SERANGOON"       "TAMPINES"
## [25] "TOA PAYOH"       "WOODLANDS"       "YISHUN"
```

The kind of flats are

```
unique(rental$flat_type)
```

```
## [1] "1-RM" "2-RM" "3-RM" "4-RM" "5-RM" "EXEC"
```

The average median rent breakdown by town are (exclude NA)

```
rental_table_town = rental %>% group_by(town) %>%
  summarize(avg_median_rent = round(mean(median_rent, na.rm=TRUE), digits=2))
kable(rental_table_town)
```

town	avg_median_rent
ANG MO KIO	1959.16
BEDOK	1961.40
BISHAN	2120.84
BUKIT BATOK	1923.38
BUKIT MERAH	2255.14
BUKIT PANJANG	1850.85
BUKIT TIMAH	1816.67
CENTRAL	2400.69
CHOA CHU KANG	1903.02
CLEMENTI	2114.29
GEYLANG	2006.06
HOUGANG	1960.65
JURONG EAST	2031.27
JURONG WEST	1954.42
KALLANG/WHAMPOA	2156.46
LIM CHU KANG	NaN
MARINE PARADE	1994.30
PASIR RIS	2020.88
PUNGGOL	1994.31
QUEENSTOWN	2126.43
SEMBAWANG	1982.77
SENGKANG	2033.33
SERANGOON	2072.96
TAMPINES	2006.44
TOA PAYOH	2109.69
WOODLANDS	1774.99
YISHUN	1813.76

The average median rent breakdown by flat type are (exclude NA)

```
rental_table_flat = rental %>% group_by(flat_type) %>%
  summarize(avg_median_rent = round(mean(median_rent, na.rm=TRUE), digits=2))
kable(rental_table_flat)
```

flat_type	avg_median_rent
1-RM	NaN
2-RM	1465.92
3-RM	1656.14
4-RM	2011.50
5-RM	2219.49
EXEC	2310.71

## Missing Data

I have selected only 2020Q1 data to reflect the latest rent price.

A first look at the data discover that 75 rows have missing data with ~66% in the smallest flat types and 20% in the largest flat type. I suspect that since not all town has 1, 2 room flats or executive flats, the missing values could be due to the non-existence of flat type. But this require a further analysis.

```
missing_entries = rental_2020 %>% filter(is.na(rental_2020$median_rent))
missing_entries %>% group_by(flat_type) %>%
  summarize(n = n(), percent = round(n()/nrow(missing_entries)*100, digits=2))
```

```
## # A tibble: 6 x 3
##   flat_type      n percent
##   <chr>      <int>   <dbl>
## 1 1-RM         26   34.7
## 2 2-RM         24   32
## 3 3-RM          6    8
## 4 4-RM          1   1.33
## 5 5-RM          3    4
## 6 EXEC        15   20
```

Here is a detail breakdown of the 75 missing entries in 2020Q1

```
kable(missing_price)
```

town	flat_type
ANG MO KIO	1-RM
ANG MO KIO	2-RM
ANG MO KIO	EXEC
BEDOK	1-RM
BEDOK	2-RM
BISHAN	1-RM
BISHAN	2-RM
BISHAN	EXEC
BUKIT BATOK	1-RM
BUKIT BATOK	2-RM
BUKIT MERAH	1-RM
BUKIT MERAH	EXEC
BUKIT PANJANG	1-RM
BUKIT PANJANG	2-RM
BUKIT PANJANG	EXEC

town	flat_type
BUKIT TIMAH	1-RM
BUKIT TIMAH	2-RM
BUKIT TIMAH	3-RM
BUKIT TIMAH	4-RM
BUKIT TIMAH	5-RM
BUKIT TIMAH	EXEC
CENTRAL	1-RM
CENTRAL	2-RM
CENTRAL	5-RM
CENTRAL	EXEC
CHOA CHU KANG	1-RM
CHOA CHU KANG	2-RM
CHOA CHU KANG	3-RM
CLEMENTI	1-RM
CLEMENTI	2-RM
CLEMENTI	EXEC
GEYLANG	1-RM
GEYLANG	2-RM
GEYLANG	EXEC
HOUGANG	1-RM
HOUGANG	2-RM
JURONG EAST	1-RM
JURONG EAST	2-RM
JURONG EAST	EXEC
JURONG WEST	1-RM
JURONG WEST	2-RM
KALLANG/WHAMPOA	1-RM
KALLANG/WHAMPOA	2-RM
KALLANG/WHAMPOA	EXEC
MARINE PARADE	1-RM
MARINE PARADE	2-RM
MARINE PARADE	5-RM
MARINE PARADE	EXEC
PASIR RIS	1-RM
PASIR RIS	2-RM
PASIR RIS	3-RM
PUNGGOL	1-RM
PUNGGOL	2-RM
PUNGGOL	3-RM
PUNGGOL	EXEC
QUEENSTOWN	1-RM
QUEENSTOWN	EXEC
SEMBAWANG	1-RM
SEMBAWANG	2-RM
SEMBAWANG	3-RM
SENGKANG	1-RM
SENGKANG	2-RM
SENGKANG	3-RM
SERANGOON	1-RM
SERANGOON	2-RM
SERANGOON	EXEC
TAMPINES	1-RM

town	flat_type
TAMPINES	2-RM
TOA PAYOH	1-RM
TOA PAYOH	2-RM
TOA PAYOH	EXEC
WOODLANDS	1-RM
WOODLANDS	2-RM
YISHUN	1-RM
YISHUN	2-RM

I decided to fix the missing values by including the past 3 years of rental data, if available. As a result, I included the following 8 past year rental data.

missing\_price\_fixed\_recent

```
##      quarter      town flat_type median_rent
## 2  2019-Q4 BUKIT PANJANG      EXEC        1980
## 7  2019-Q3 CHOA CHU KANG      3-RM        1480
## 9  2019-Q4  JURONG EAST      EXEC        2500
## 10 2019-Q3 MARINE PARADE      5-RM        2500
## 12 2019-Q3      PUNGGOL      3-RM        1700
## 13 2018-Q4      PUNGGOL      EXEC        2000
## 14 2019-Q4      SENGKANG      3-RM        1650
## 15 2019-Q4  SERANGOON      EXEC        2550
```

Now, we have around 57% of the total possible rental data in Singapore

```
non_empty_rows = rental_2020_fixed_na %>% filter(!is.na(median_rent))
cat(round(nrow(non_empty_rows)/nrow(rental_2020) * 100,digits=2),"%")
```

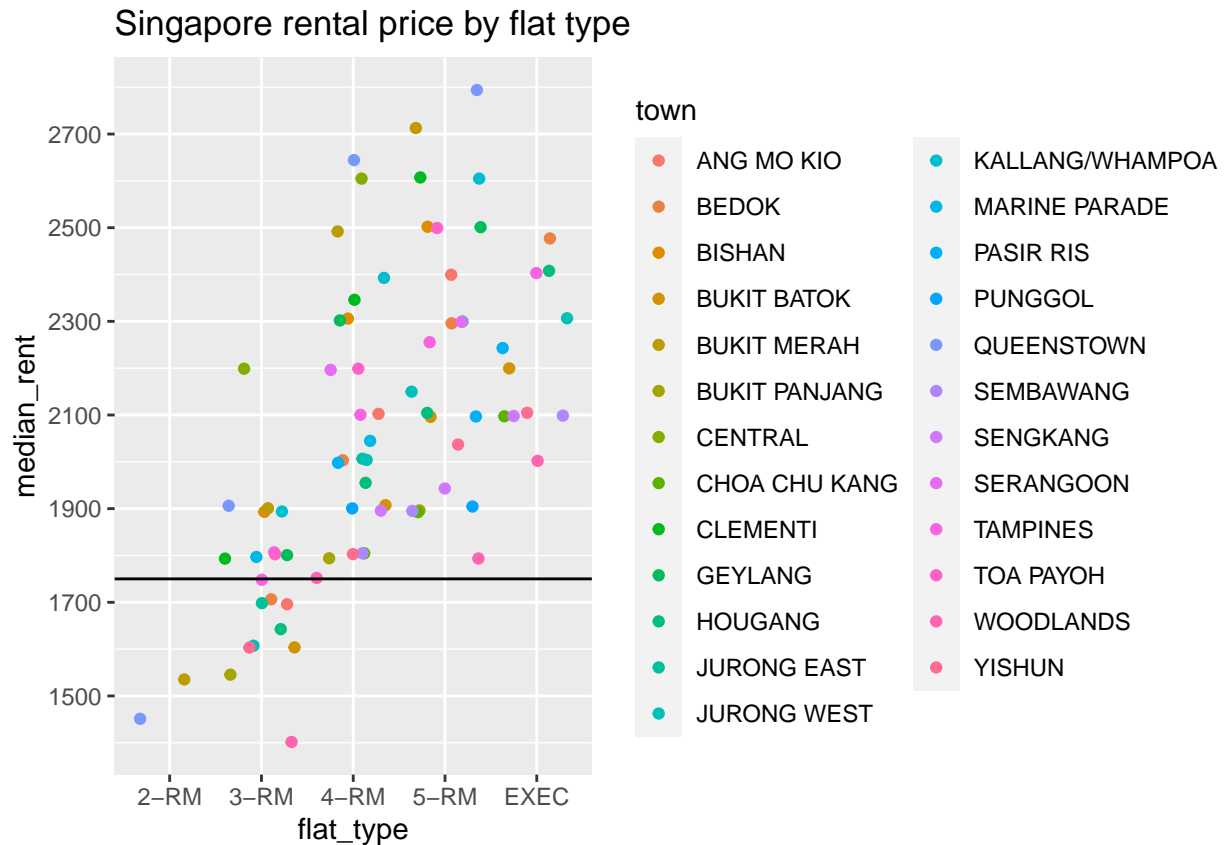
```
## 57.05 %
```

## Analysis

Since we are interested in knowing who is eligible for the SIRS, that is flat with annual value  $\leq$  \$21000, we will focus on flats with monthly rent of  $\leq$  \$1750

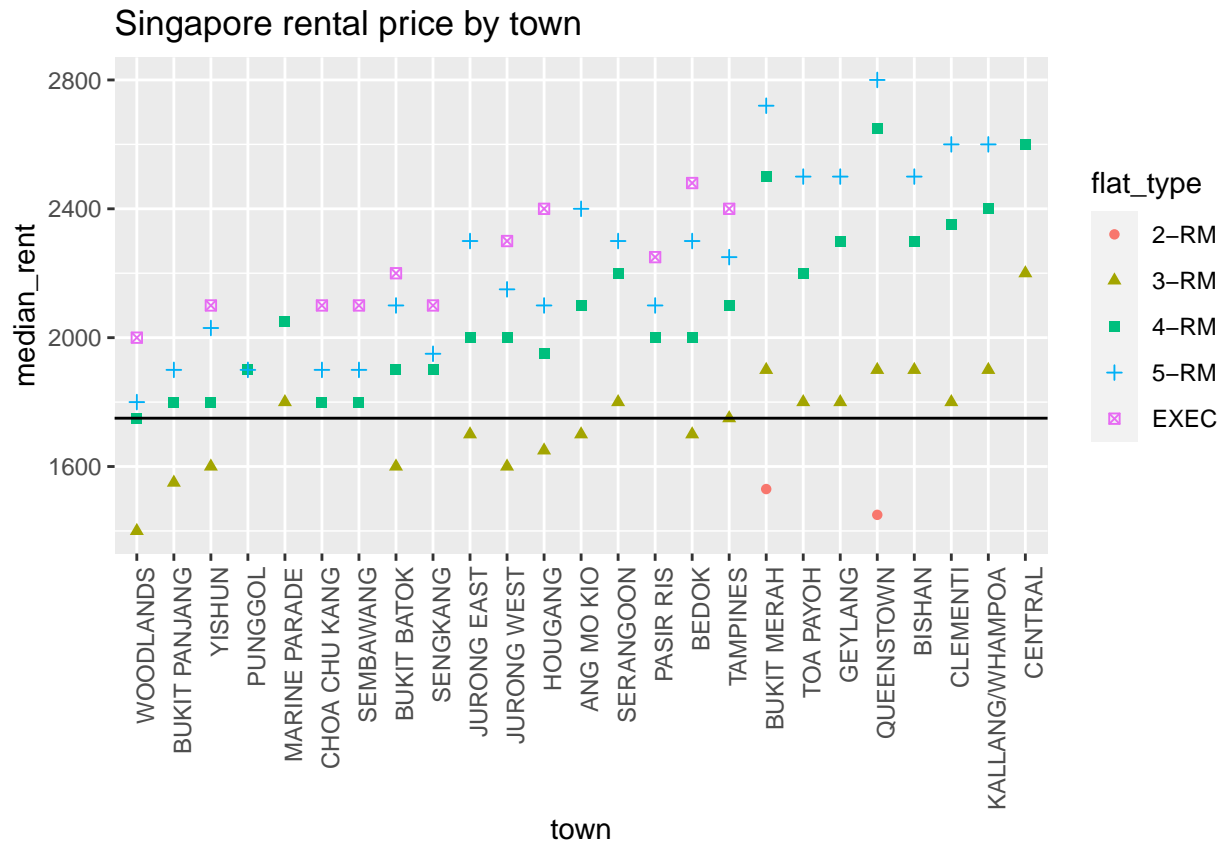
We first take a look at the difference in median rent across flat type. It is clear from the graph that most of the eligible flat types are 3 room and 2 room flats. The anomalies include one 4 room flat and a percentage of 3 room flats.

plot\_price\_by\_flat\_type



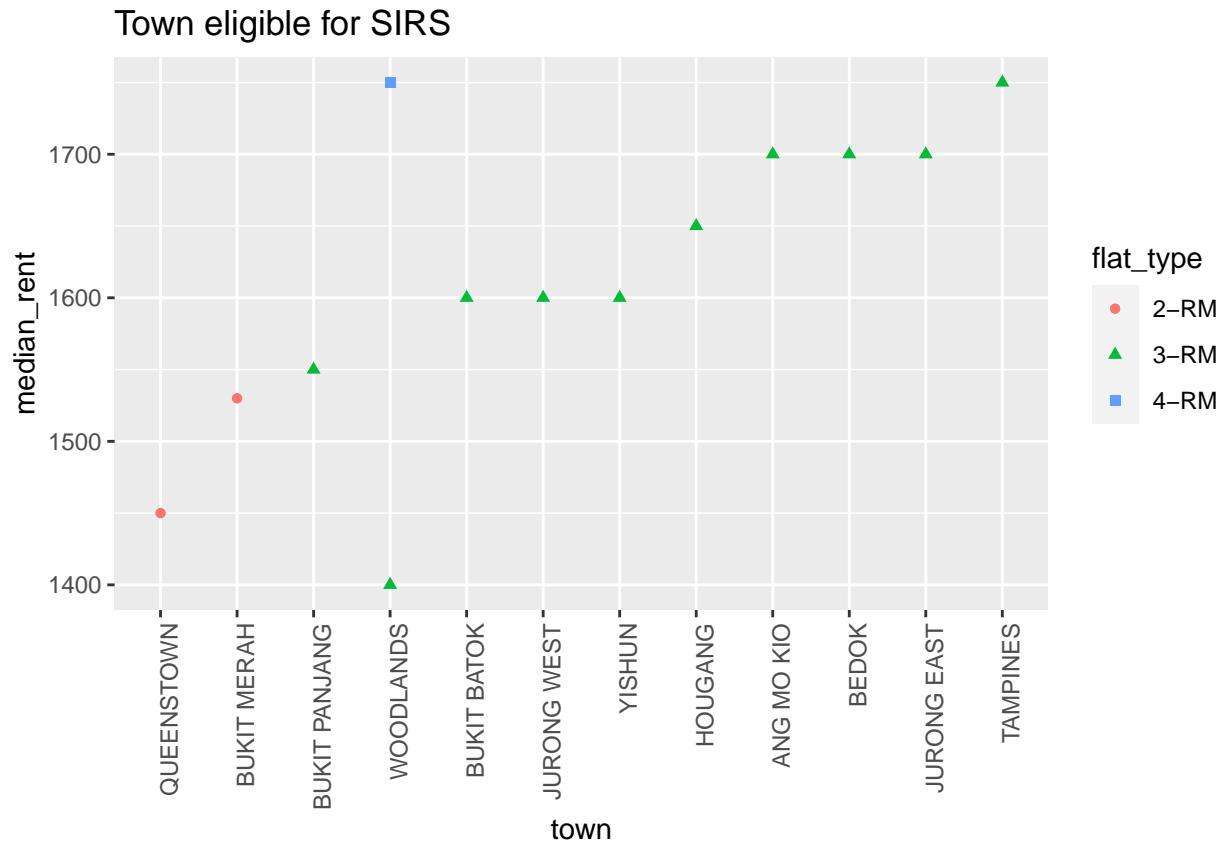
Now we will look at the median rent breakdown by town. We see that except for 4 room flat in Woodlands, all other 3 room flats are not eligible for SIRS. And about 1/3 of the 3 room flats are not eligible for SIRS as well.

```
plot_price_by_town
```



When we take a closer look at the flats eligible for SIRS, our previous observation is even clearer

```
plot_eligible_town
```



The total town and flat type combination eligible for SIRS takes up only 8% of the total possible combination

```
cat(round(nrow(rental_sirs) / nrow(rental_2020) * 100,digits=2),"%")
```

```
## 8.33 %
```

## Conclusion

There are a lot of limitations with my analysis.

Firstly, the rental prices only illustrate the median price, it does not consider the lower range of rental.

Secondly, we need to look further into the missing values. We need to investigate if the missing values is due to absence of the flat type or other reasons.

Thirdly, this datasets only captures public housing and not private housings. It might be naive to assume all private housings in Singapore has a rental value higher than \$1750/month.

Lastly, we only look at one of the criteria for SIRS: annual property value. Self-employed individuals may still apply for SIRS even if they failed one of the criteria. Thus we are not observing the actual population who are eligible for SIRS.

Nonetheless, from our current anlysis, we can tell that government has a targeted approach of targeting only around 8% of a typical self-employed individuals who lives in HDB flats.