

# Behaviour of residential construction and real estate markets in W.A. during Covid-19 Crisis - report

Team: 'Kalbarri Sun'

## Team members

- Lauren Prins
- Daniel Sobral
- Warren Villarosa
- Sylvia Broadbent

+Behaviour of residential construction and real estate markets in W.A. during Covid-19 Crisis - slides  
Github link: <https://github.com/Supasyl/Data-Analytics-Project-1.git>

---

## Summary

### Hypothesis

Considering the information we keep receiving from the government and media, it seems that the whole world has been affected by COVID-19. We are looking at various factors of the residential real estate and construction market in WA to see if it has really been affected.

### Questions

We asked ourselves what areas we can look into to get a diversified look at the residential real estate market as that is an area of shared interest in our team. We have narrowed it down to property sales, building permits issued and their value and the first home owner grant in relation to dwellings, townhouses and apartments.

We then looked further into the behaviour of the stock value of the 10 largest property groups in WA.

### Available timeframe

We based our timeframe for the research on the dataset with the most limited data. This turned out to be the real estate information as most of that information requires payment. We started from the last quarter of the year in 2018 to the second quarter this year.

### Covid-19 period

In WA the regional travel ban started on 31 March so we have used the second quarter of 2020 as the 'COVID-19' quarter to compare against.

---

# Stock Market - Daniel Sobral

**Question:** How was the Share Market impacted from Q4 2018 to Q2 2020?

How does the behavior of the market relates to other macro economic factors in the Real Estate Sector?

To do so we will look at the top 10 companies with highest market capitalization in the property sector.

**Data:** Extracted share prices and traded volume data from the last 5 years using Bloomberg API available on RapidAPI.

National Real Estate Sales Data - extracted from Kaggle

## Cleanup & exploration:

Libraries and modules used:

- requests
- pandas
- numpy
- datetime
- json
- matplotlib.pyplot
- matplotlib.dates
- scipy.stats
- csv

Save extracted Data into json format

Use Pandas module to treat our data creating tables to analyse changes in volume traded and share price change. Individual tables per share were created, holding the price value and traded volume value.

Convert our tables into quarterly % changes to analyse possible sudden changes as Australia dealt with the Covid-19 Crisis.

Convert Share Market Price change into monthly % changes, increasing our number of data points in order to look for correlation with changes in national changes.

Table 1 - Share Price % change

	GMG	SCG	DXS	SGP	MGR	LLC	GPT	VXS	CHC	GOZ	Avg (%) Change
<b>Quarter</b>											
2018Q4	2.33%	-6.94%	3.16%	-4.64%	-4.88%	-11.34%	3.93%	-0.72%	5.30%	-2.59%	-1.64%
2019Q1	22.06%	-1.20%	14.88%	-5.60%	17.52%	-24.08%	11.89%	-3.99%	36.03%	9.21%	7.67%
2019Q2	12.62%	-3.65%	7.06%	15.72%	16.73%	0.92%	0.97%	1.13%	12.17%	4.82%	6.85%
2019Q3	6.56%	4.29%	1.91%	4.68%	3.74%	24.79%	3.51%	0.37%	9.81%	2.76%	6.24%
2019Q4	-7.78%	-3.63%	-12.01%	7.45%	1.20%	9.57%	-2.62%	0.00%	-6.40%	-1.57%	-1.58%
2020Q1	10.93%	1.26%	8.18%	5.35%	3.26%	-1.53%	-1.27%	-4.83%	17.15%	0.91%	3.94%
2020Q2	-6.51%	-37.72%	-24.04%	-25.56%	-28.45%	-32.31%	-28.37%	-32.03%	-27.18%	-22.97%	-26.51%
2020Q3	19.78%	-9.56%	-5.87%	-1.52%	-8.03%	-2.68%	-1.34%	-16.67%	23.86%	0.58%	-0.14%
2020Q4	3.80%	1.32%	-2.33%	5.13%	-1.31%	-0.63%	-4.99%	-1.38%	7.35%	1.16%	0.81%

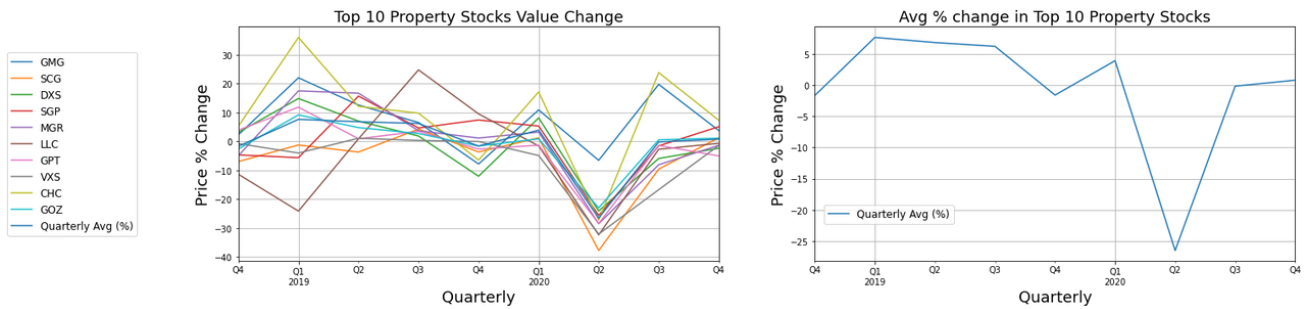
Table 2 - Trade Volume % Change

	GMG Vol	SCG Vol	DXS Vol	SGP Vol	MGR Vol	LLC Vol	GPT Vol	VXS Vol	CHC Vol	GOZ Vol	Avg Vol (%) Change
<b>Quarter</b>											
2018Q4	12.94%	15.40%	30.89%	32.31%	40.78%	405.66%	18.34%	25.13%	65.31%	-9.43%	63.73%
2019Q1	74.95%	2.50%	13.50%	12.77%	-20.03%	-43.93%	-20.04%	37.30%	-13.79%	28.82%	7.21%
2019Q2	-54.91%	45.26%	20.81%	52.15%	49.15%	-36.48%	85.35%	-9.20%	18.19%	-2.08%	16.82%
2019Q3	379.43%	-11.37%	-34.90%	-28.95%	-34.03%	66.04%	-28.70%	-34.84%	43.78%	33.84%	35.03%
2019Q4	-74.88%	133.76%	4.35%	-23.63%	-4.92%	-47.50%	-14.90%	-9.85%	-36.93%	12.91%	-6.16%
2020Q1	125.26%	9.23%	123.99%	173.58%	153.32%	40.26%	101.78%	201.02%	127.95%	79.87%	113.63%
2020Q2	-46.30%	-28.57%	-16.06%	-24.25%	-29.95%	37.62%	-5.50%	52.47%	-38.81%	-27.11%	-12.64%
2020Q3	-15.24%	-13.10%	-40.55%	-37.36%	-23.71%	-54.17%	-45.22%	-49.28%	-40.67%	-31.30%	-35.06%
2020Q4	-37.56%	-50.23%	-33.40%	-24.22%	-24.56%	-25.70%	-24.27%	-33.12%	-32.08%	-48.46%	-33.36%

### Analysis:

To visualise our data we decided to plot line charts as they express strong changes during the Covid-19 Outbreak. The analysis is split into 2 approaches, outlook on share valuation (Figure 1) and market liquidity (Figure 2)

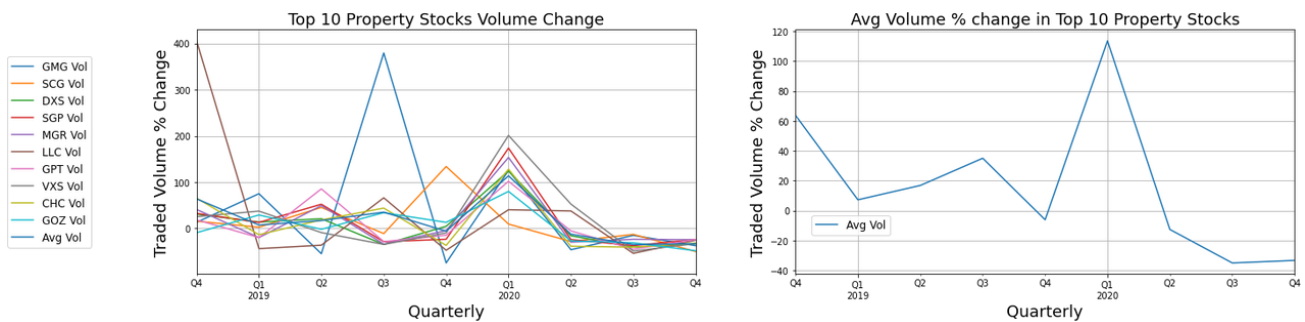
Figure 1 - % Change in Share Value



### Share Price Analysis

- From start of 2020 the Top 10 Property stocks display closer price movements,
- Leading to end of Q2 2020 (covid outbreak and subsequent lockdown) measures all shares dropped.
- Followed by a V-shape recovery until end of Q3 of 2020.
- GMG stands has the biggest realestate company with 3 times the market cap of the second largest.
- The Average Price Movement will allow us to compare the stock market behaviour with other macroeconomic variables.

Figure 2 - % change in Traded Value



Volume as a Liquidity indicator:

- GMG displays a more independent behaviour with an individual spike of liquidity leading to 2020 Q3.
- Spike in volume traded in the beginning of 2020(pre-covid) - Followed of a massive drop heading to end of Q2 2020(post-covid).

### Overall Share Market Conclusion Conclusion:

- Shares Value and Volume traded dropped after March.
- Although the % change in prices saw a V-shape recovery, the % change in traded volume kept negative.
- Despite the recovery in the value these stocks, they now offer significantly less liquidity than they did pre-covid.

**Correlation:**

After analysing the Share Market behaviour in regards to real estate, we set out to find if the market behaviour has any correlation with changes in National Real Estate Sales.

Due to limitations on National Sales Data we limit our table to beginning of April 2020.

Using pandas module I transformed absolute sales value into monthly percentage change and compare it with the share market percentage change. This change allowed me to increase the number of data points to 18 (Table 3) and create a scatter (Figure 3) plot to compare the 2 variables.

Table 3 - Change in Property Sales Vs Share Value Change

	Year	Month	Sales	Monthly Avg (%)
1	2018	11	1.58%	-1.23%
2	2018	12	-19.42%	0.72%
3	2019	1	-27.19%	2.06%
4	2019	2	40.16%	6.66%
5	2019	3	20.53%	2.20%
6	2019	4	-10.62%	0.73%
7	2019	5	11.05%	3.92%
8	2019	6	-4.94%	1.74%
9	2019	7	4.80%	4.50%
10	2019	8	5.47%	1.19%
11	2019	9	-1.53%	-3.24%
12	2019	10	17.47%	0.73%
13	2019	11	8.06%	1.92%
14	2019	12	-22.54%	-1.19%
15	2020	1	-34.51%	2.40%
16	2020	2	63.99%	2.28%
17	2020	3	0.17%	-9.46%
18	2020	4	-35.84%	-28.31%

Figure 3 - Change in Property Sales vs Share Value Change

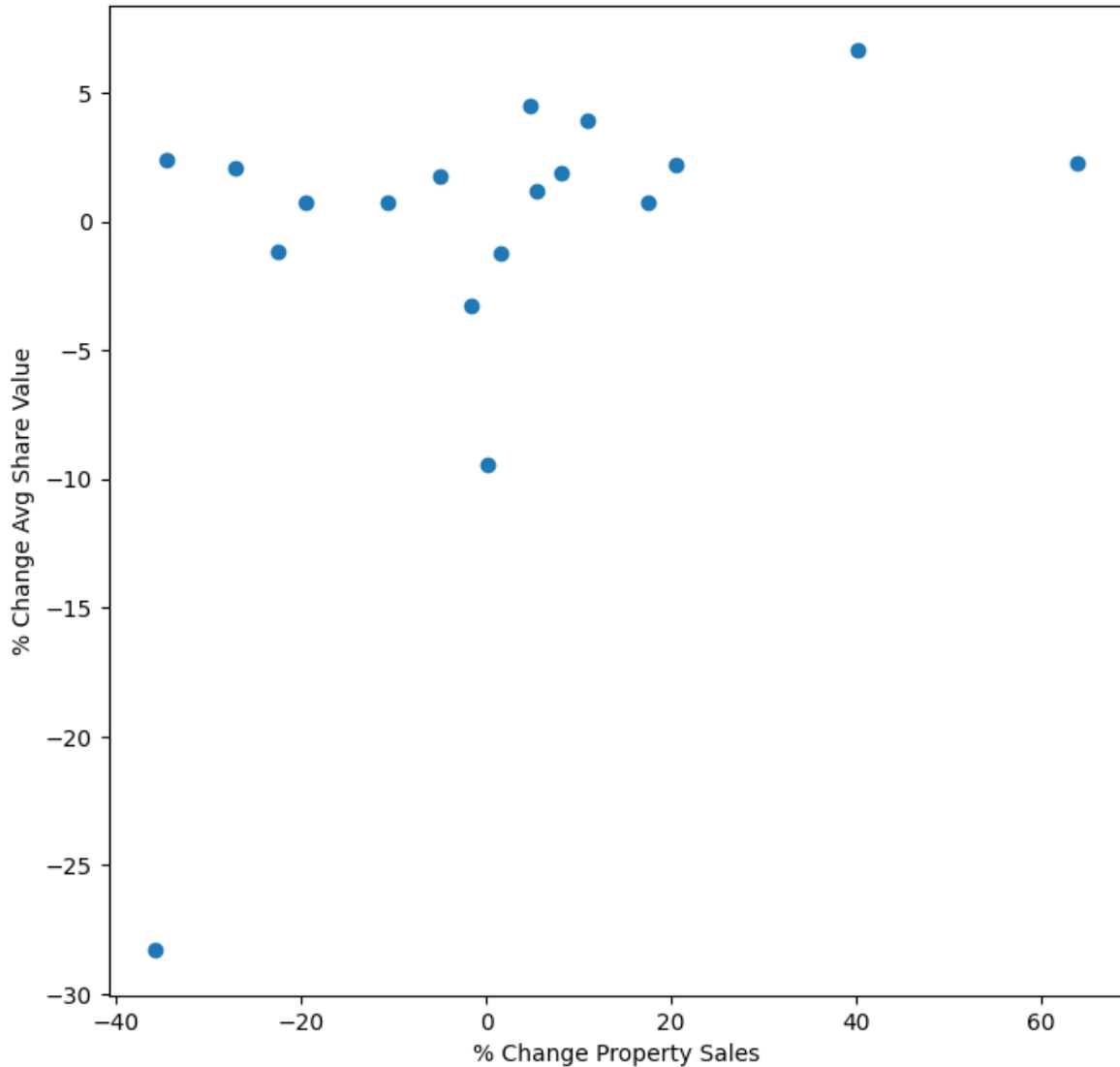


Figure 3 shows that share price value variation presented some stability when compared with the variation in Property Sales.

We can notice a clear exception when both variables revealed a record drop.

As we've seen previously on Figure 1 this drop relates to the beginning of Q2 2022, when the covid-19 outbreak gave origin to lockdown measures Australia wide.

Normal Distribution Hypothesis:

Using - scipy.stats module

Both Variables exceeded the necessary p-value to be considered normal distribution so we cannot run a linear regression and evaluate a degree of correlation.

# Real estate - Warren Villarosa

## **Introduction:**

The following data was sourced from Kaggle as a CSV file and included real estate sales nation wide. The dataset did not include commercial real estate sales. Each row was a confirmed sale and included date, price, suburb, property type and the coordinates for the suburb

## **Methodology:**

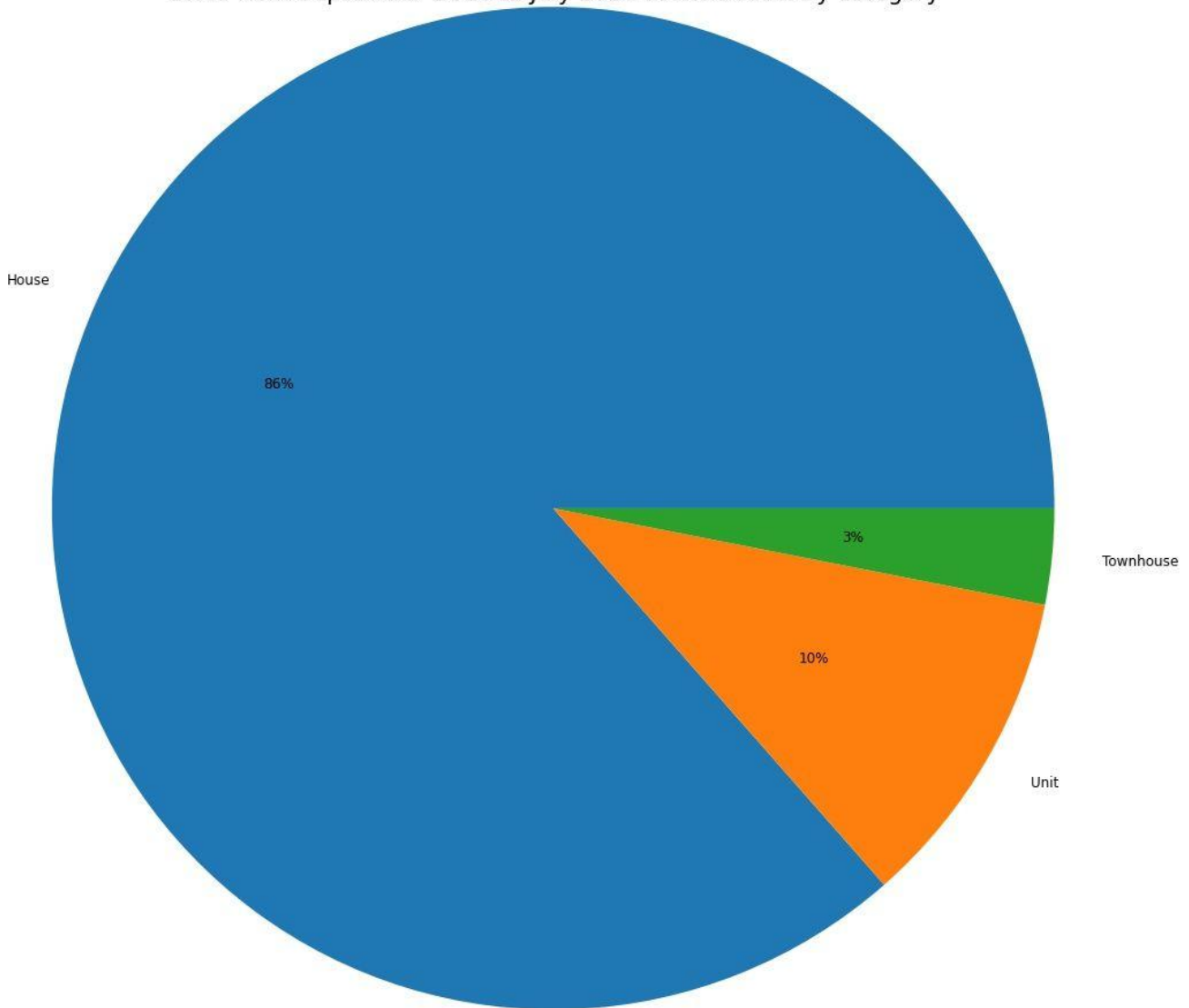
The Market of interest for the analysis was Western Australia so I cleaned the dataset and limited it to sales within Western Australia. I wanted to make sure that I used as much of the information available. Some time was spent changing the strings in the date column into Datetime format and breaking the dates into year, month and quarter columns to make grouping a little simpler.

The main libraries and modules used for the processing of this data were:

- Pandas
- Matplotlib
- Scipy
- Numpy
- Datetime

**Figure 1 - The Share of total real estate sales over the time frame examined.**

Sales from September 2018 to July 2020 broken down by category



*Share of total volume by Property Type.*

The figure above is simply here to illustrate that the Western Australian market is heavily weighted towards dwellings (houses) with only 13% of the market being shared between units and townhouses. This means that the sales of the latter two categories will only have a minor influence on the data following.

After grouping the data by year, month, date sold and adding a count of entries, I was able to produce the plot below by plotting the months against the counts which corresponded to sales volume. On the above figure and the ones following, I marked some key dates to see if they coincided with any major trends in the data. The dates are as follows:

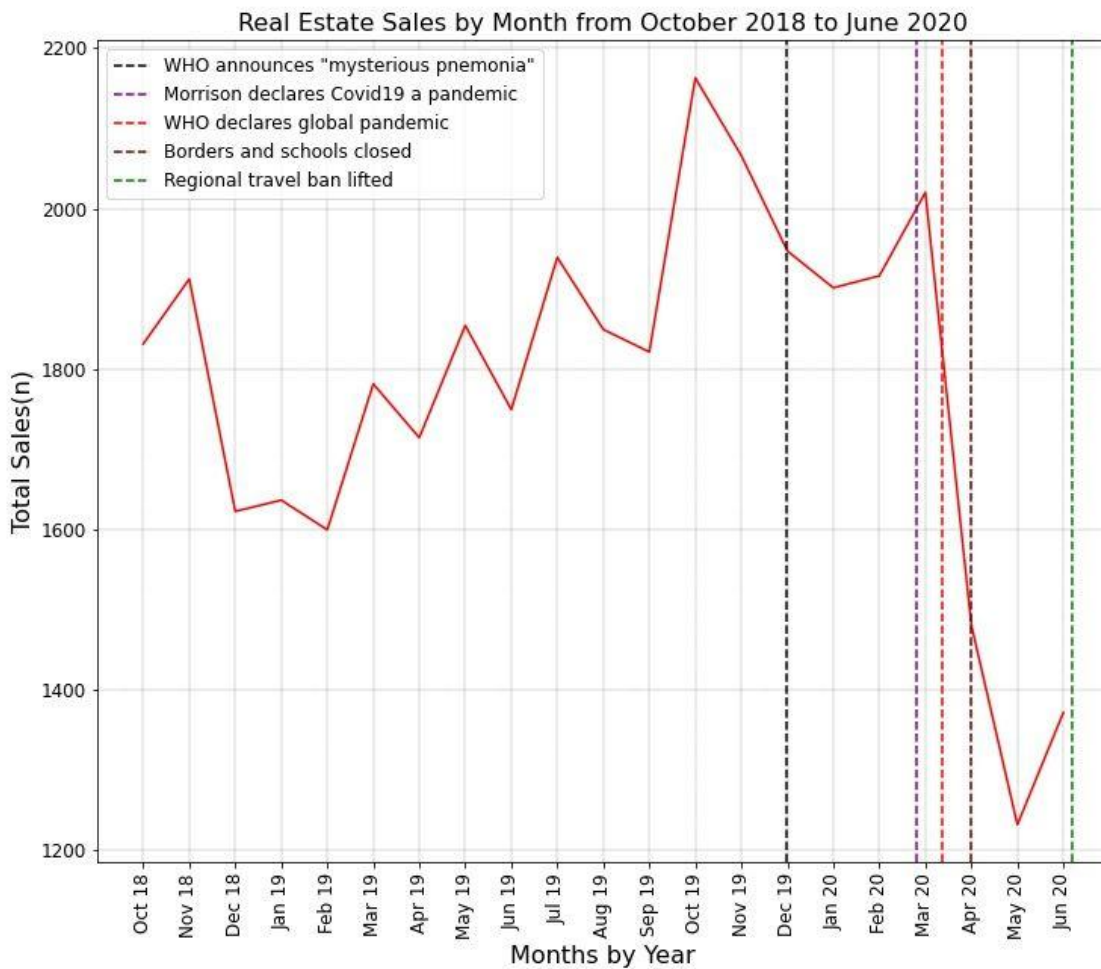
- Dec 31st 2019 - WHO announces a "mysterious pneumonia"
- Late February (Around the 23rd) 2020 - Scott Morrison Declares Covid 19 a pandemic
- March 11 2020 - WHO Declares Covid19 a global pandemic
- March 31 2020 - Regional borders and schools closed within Western Australia



- June 6 2020 - Regional travel ban lifted

The dates were sourced from various news articles from the ABC.

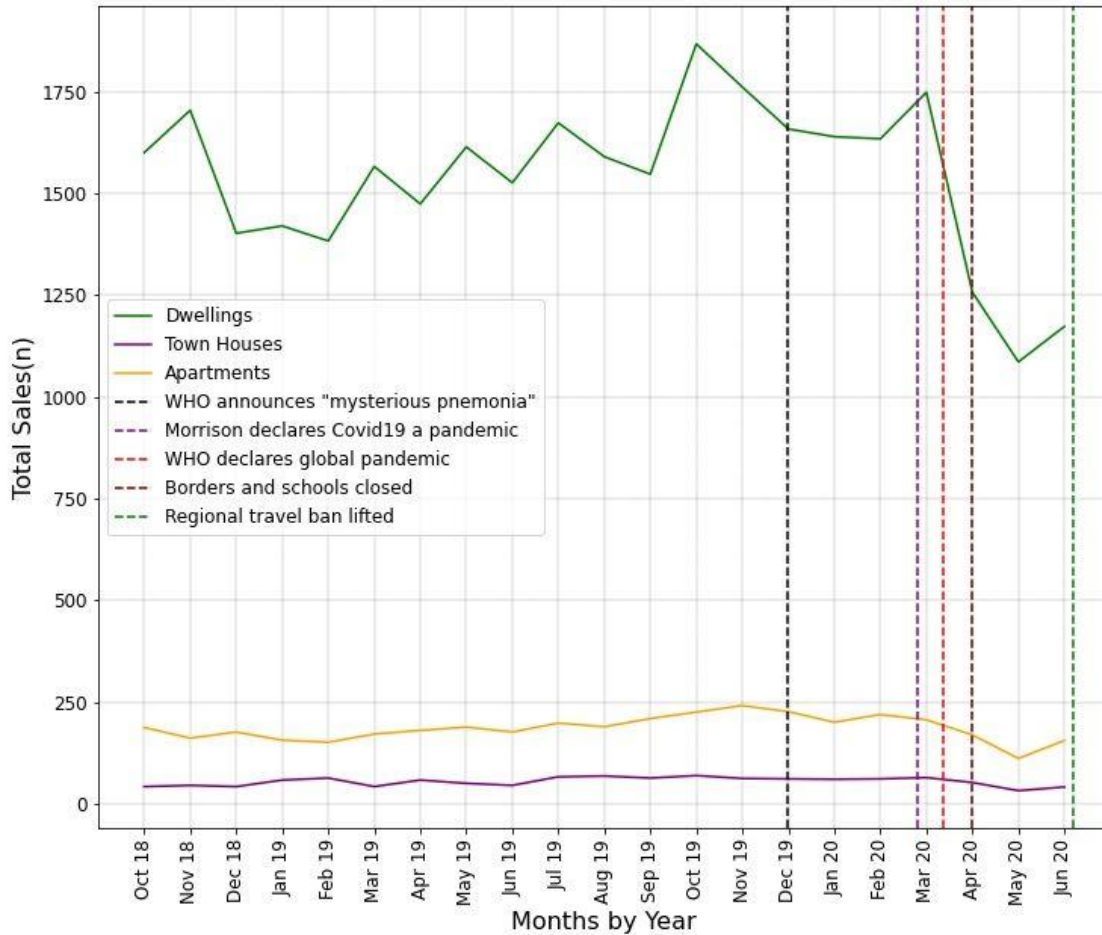
**Figure 2 - Total volume of real estate sales from Oct 2018 to June 2020**



*Total volume of real estate sales over the period examined*

This particular figure illustrates the trend followed by the total volume of sales during the period being examined. I chose this graph because I thought it would be interesting to explore the effect that Covid19 had on the overall volume of sales. The sharpest drop in volume appears to follow shortly after Scott Morrison declared Covid 19 a pandemic (Indicated by the purple intermittent vertical line) in late February 2020 (ABC News). While the correlation between the two can only be suggested as a best guess with the limited data on hand, the market does show a strong reaction not long after the announcement.

**Figure 3 - Volume of realestate sales from Oct 2018 to June 2020 by housing category**



*Total volume of realestate sales by category Oct 2018 to Jun 2020*

After breaking the dataset up into property type I was able to plot individual trend lines for each property type against the monthly periods. The above figure provides an insight into the influence the various housing categories had on the overall volume of sales.

When compared to the previous line graph an almost identical trend can be noticed in dwellings (green); this is likely because they make up the lions share of the total volume of sales. The market’s reaction had a less severe but noticeable effect on the sale of units (yellow), while the low volume of town house (purple) sales was barely affected at all.

**Figure 4 - mean and median house price by month**



Q4 2018	Q1 2019	Q2 2019	Q3 2019	Q4 2019	Q1 2020	Q2 2020
Baldivis	Baldivis	Baldivis	Baldivis	Baldivis	Baldivis	Baldivis
Canning Vale	Morley	Thornlie	Canning Vale	Canning Vale	Scarborough	Ellenbrook
Morley	Scarborough	Canning Vale	Scarborough	Scarborough	Canning Vale	Canning Vale
Dianella	Canning Vale	Ellenbrook	Dianella	Thornlie	Dianella	Dianella
East Perth	Dianella	Scarborough	East Perth	Ellenbrook	Morley	Butler
Gosnells	Como	Dianella	Ellenbrook	Rockingham	Rockingham	Scarborough
Perth	Gosnells	Maylands	Morley	Perth	Como	Duncraig
Thornlie	Maylands	East Perth	Como	South Perth	Maylands	Thornlie
Clarkson	Bayswater	Morley	Byford	Dianella	Duncraig	Wanneroo
Scarborough	South Perth	Secret Harbour	South Perth	Morley	Ellenbrook	Claremont

I also wanted to explore the effect that Covid might have had on the best performing suburbs by volume. I had wanted to plot these in bar charts but didn't feel those bar plots added any value to the big picture. Not even a pandemic could shake Baldivis from its top spot; while Canning Vale, Scarborough and Dianella rarely fell out of the top 10 if ever. I started to wonder why anyone would willingly buy in Baldivis but thought better of it. After all, I myself made the mistake of buying in the same postcode as Armadale.

**Figure 5 - Scatter Plot Showing Correlation between number of permits issued and number of real estate sales for W.A.**

The r-squared is: 0.2547026417198125

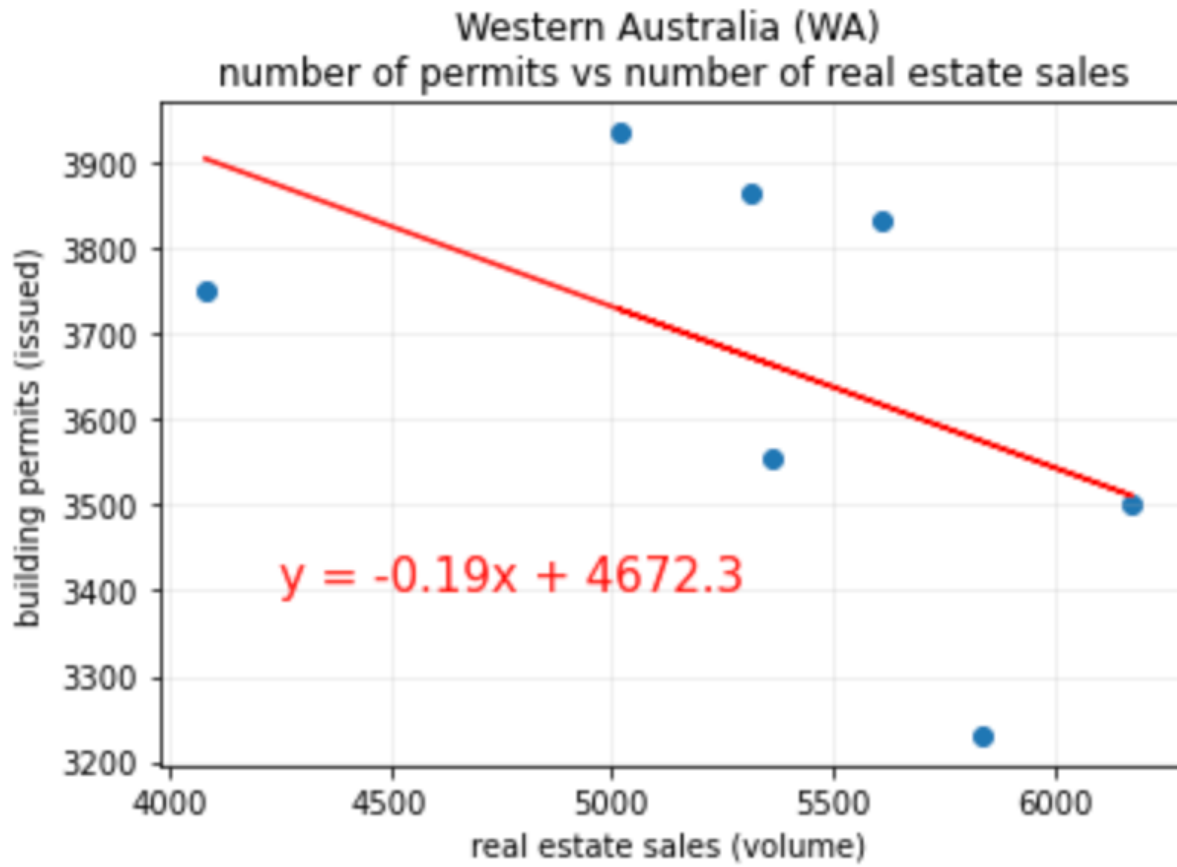
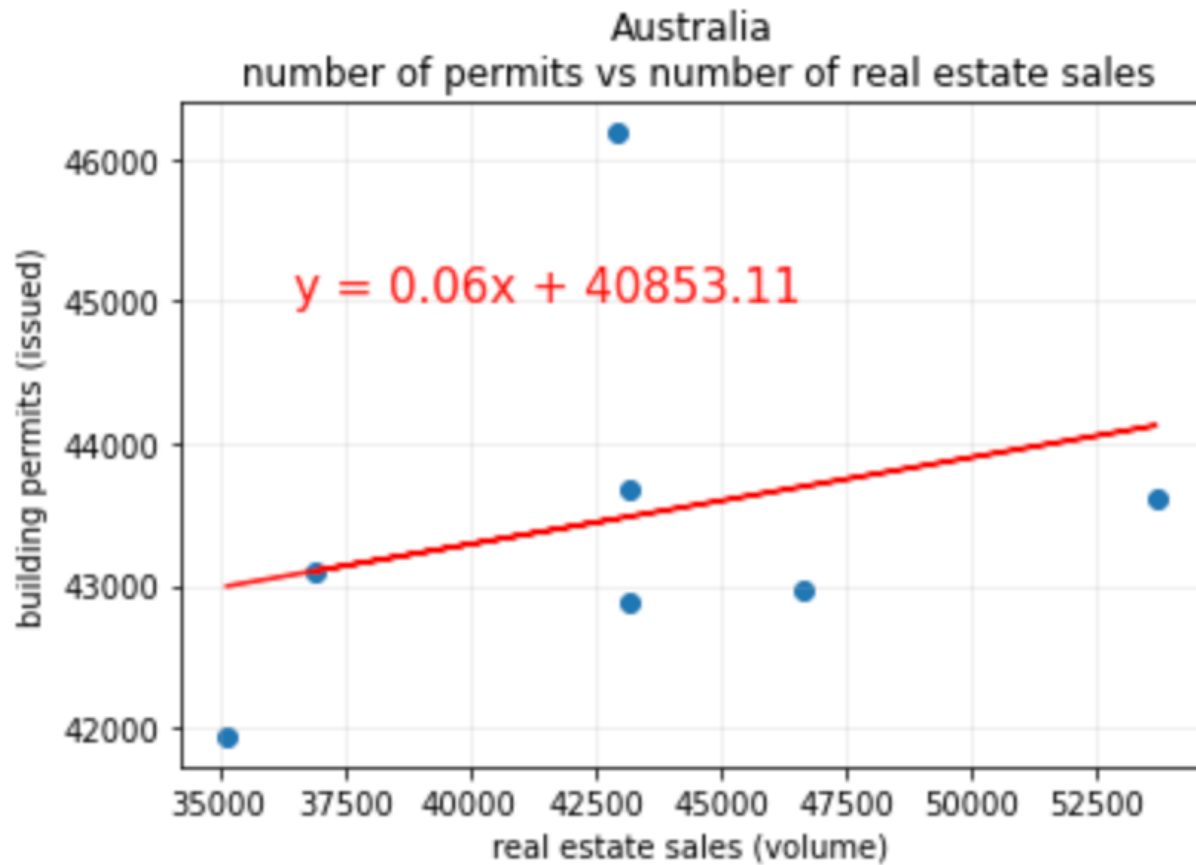


Figure 6 - Scatter Plot Showing Correlation between number of permits issued and number of real estate sales.

The r-squared is: 0.07987105331653124



*Number of Permits in Australia vs Total Real Estate Sales*

## Permits & construction value - Sylvia Broadbent

Through my career as a building surveyor I know that councils are required to submit their statistical data monthly. I started with the Building Commission which is where they have to provide the data to. Apparently, they just forward it to the ABS (Australian Bureau of Statistics). On the ABS website they have a collection of data of all states and surprisingly the data from WA starts at 1970!

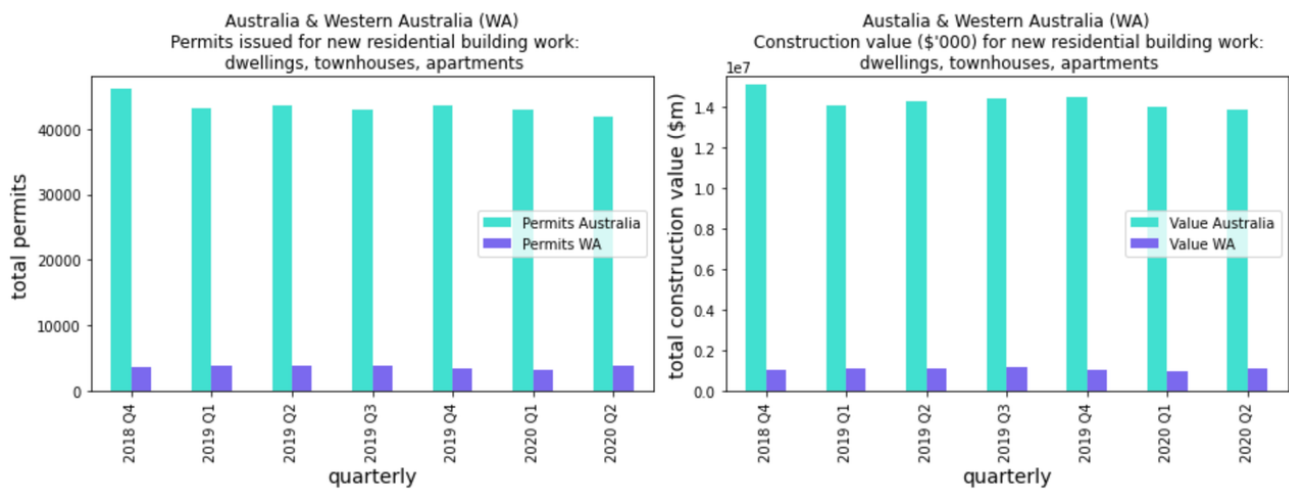
After finding the right files I could export them as excel files. Then it was just a matter of changing them to a csv format and extracting the relevant information by creating a Pandas DataFrame in Jupyter.

Some of the columns required changing the type from object to integer and from string to Datetime. Then sorting the information into quarters for comparison and plotting. I didn't use any average or median values as there is not enough information to make an accurate representation. The dataset needs to be larger for that.

	dwelling	townhouses	apartments	dwelling_\$\$'000	townhouses_\$\$'000	apartments_\$\$'000	total_permits	total_value		permits_au	value_au
2018Q4	3068	240	247	910726	56850	84852	3555	1052428	2018Q4	46181	15107274
2019Q1	2821	299	815	824553	69611	218410	3935	1112574	2019Q1	43091	14062628
2019Q2	2996	343	526	878667	74793	171036	3865	1124496	2019Q2	43681	14315655
2019Q3	2770	265	799	826658	66966	294430	3834	1188054	2019Q3	42972	14390276
2019Q4	2897	268	336	877361	67937	119569	3501	1064867	2019Q4	43616	14496008
2020Q1	2718	220	293	830906	53003	89497	3231	973406	2020Q1	42878	13992118
2020Q2	2878	196	676	851247	52387	208546	3750	1112180	2020Q2	41943	13845277

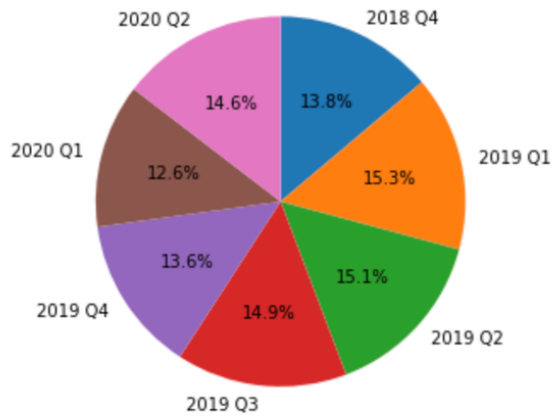
All information presented below shows the number of permits on the left and the value of permits in \$'000 on the right.

First I checked out if there was any major differences between the permits issued Australia wide compared to WA per quarter. Australia wide the best quarter is 2018 Q4 (46,181) with not many variation between the other quarters, while in WA the best quarter is 2019 Q1 (3,935) sloping down to the lowest number in 2020 Q1 (3,231 = 82% of 2019 Q1) with a significant increase in 2020 Q2 (3,750).

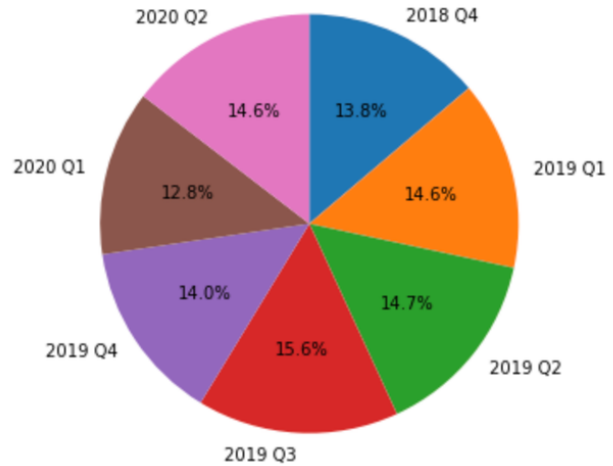


The following is a pie chart showing just the information for WA of each quarter. The quarters are going clockwise. You can see that there is little variation in the size of the slices with the smallest slice being Q1 this year with 12.6% vs the largest slice being 15.3%.

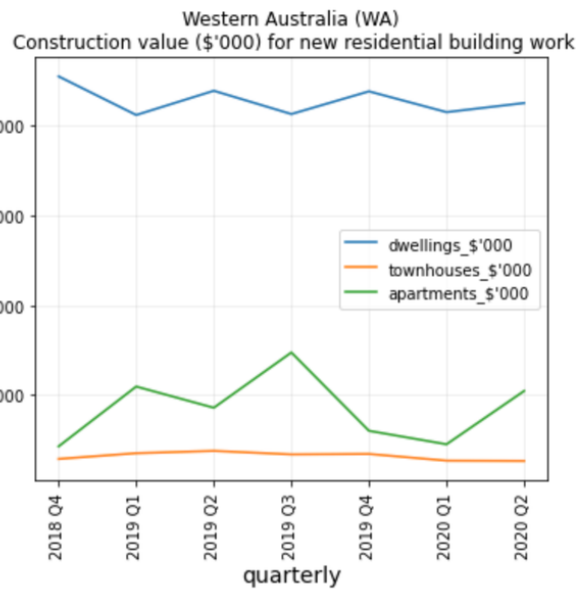
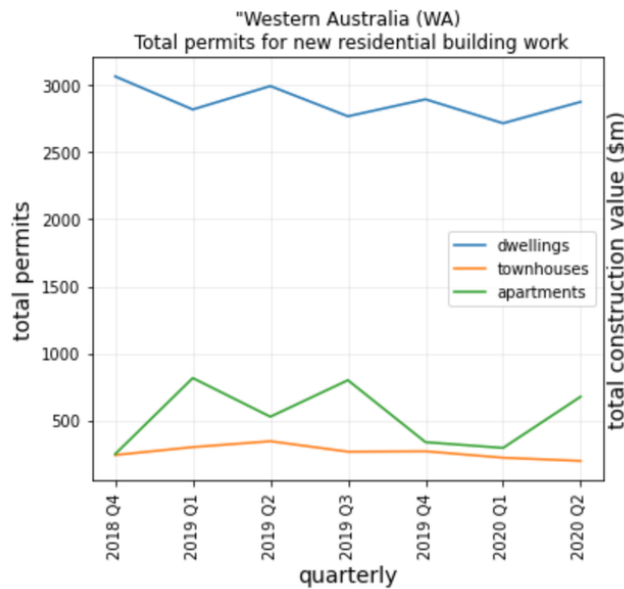
Western Australia (WA)  
Total permits for new residential building work:  
dwellings, townhouses, apartments



Western Australia (WA)  
Construction value (\$'000) for new residential building work:  
dwellings, townhouses, apartments



This line graph is showing just the information for WA of each quarter separated by type of residential construction. Dwellings are the largest parts of residential permits followed by apartments and then townhouses. When the demand for dwellings go up the demand for apartments go down and vice versa. This can be explained that the total number of dwellings remains stable whether that is the demand for dwellings or apartments.



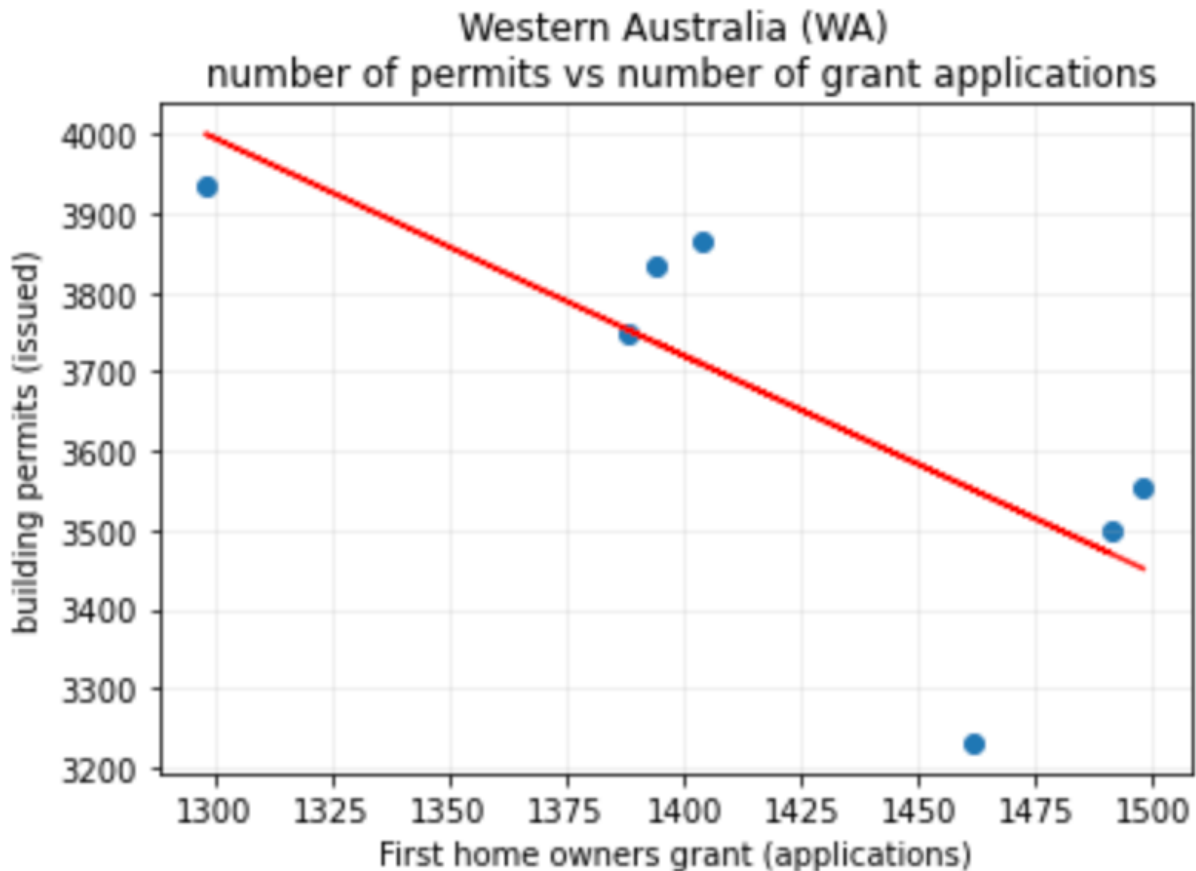
You can also see that there is not much difference between the amount of permits issued versus the value it represents. We can conclude that the average value per permit does not vary much, which means that the distribution of high spec homes versus low spec homes is fairly equal.

### First home owners grant vs Building permits

We expected the applications for construction grants to directly affect the number of building permits issued. As shown by the Linear regression model the value shows a moderate negative correlation.



The r-squared is: 0.5971298445529775



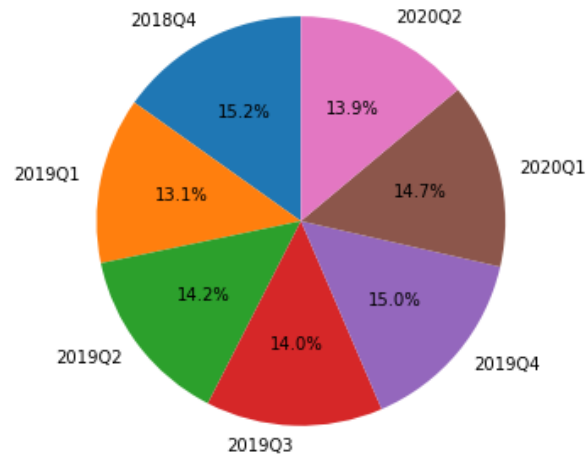
## First home owners grant - Lauren Prins

On the 15th of every month, Revenue WA collects data on First Home Owner Grants and publishes it on [www.wa.gov.au](http://www.wa.gov.au) for public online use. The data is broken down by month and backdates to July 2000. The downloadable excel file displays the total amount of applications made for the grant, and the number of grants that have been paid. It's further broken down into 'Established Homes' and 'New Dwellings'.

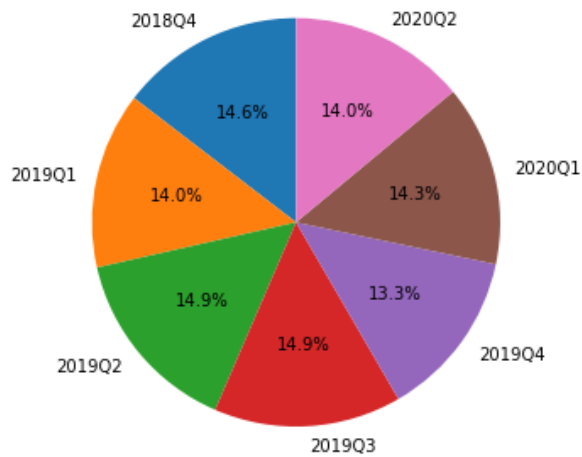
To begin analysis on the data I was able to save the excel tabs as csv files and upload them to Jupyter. I used Pandas to merge the two csv's together and changed the months listed into the quarters that we needed as an index. From there I could begin presenting the data through various graphs.

I used Matplotlib to create the 2 pie charts below which are displaying the 'Total applications made' and 'total applications paid' by the Commonwealth for Western Australia. I chose these graphs to see if there was a significant increase or decrease in any of the 7 quarters for either topic but after looking at the charts you can see it is evenly spread.

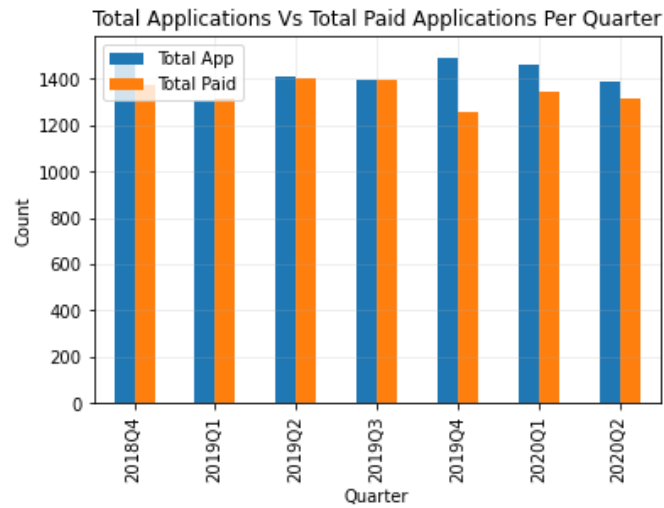
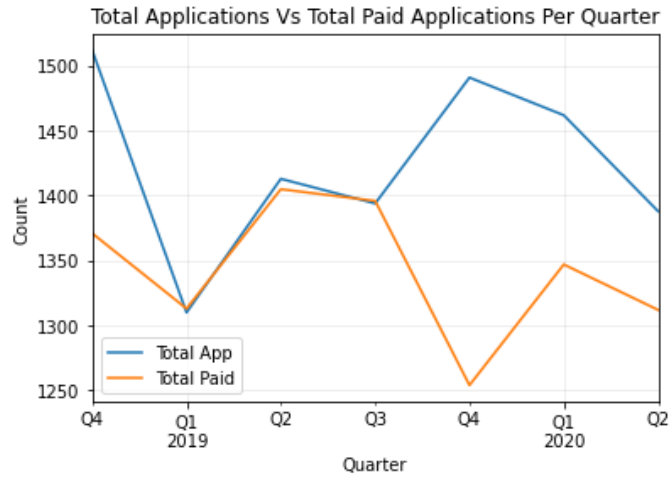
Total Applications made in WA for First Home Owner Grants



Total Paid Applications in WA for First Home Owner Grants

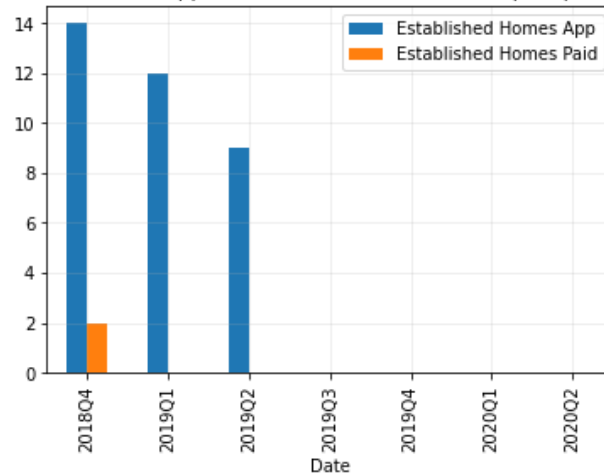


I then chose to combine that data to create a line and bar graph which showed the difference per quarter, between applications made and applications paid. As you can see in this graph, there is a lot of variance both within and between both variables.

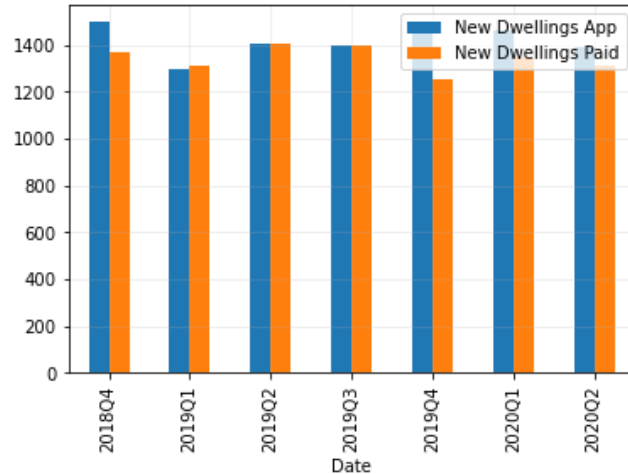


I found it interesting that the excel document informs the viewer that the grant guidelines changed in 2019 and the FHOG would only be available for New Dwellings. I displayed this information in the first graph below where you can see a sudden decline and a flat line from 2019 in applications and paid grants to established homes. The second graph shows a steady amount of paid grants for new dwellings.

Established home applications vs Established home paid per Quarter



New dwelling applications vs New dwelling paid per Quarter



## Conclusion

### Did we find what we expected to find?

- As a group, we were all hoping to find a significant downturn across all data sets for Q1 and Q2 of 2020. For Sylvia and Lauren, there was no significant change on construction permits or first home owner grants, and between both data sets there was a moderate positive linear relationship of 0.59.
- The total volume of sales, average house price and median house price decreased since the announcement of the pandemic. There was no shown correlation with this and the number of permits, as the r squared value is 0.08.
- Share value and volume dropped after March 2020. Although prices saw a recovery, the % change in traded volume kept negative.

---

## Limitations of the study/ Difficulties experienced

- A lot of real estate information is paid which limited the amount and type of data that is accessible to us.
- Several construction grants were announced late May, starting in June. There was no information available yet so that is a factor that will have had an impact on the residential construction market.
- Accessibility to the most recent data which would better provide information to the effects of Covid.
- It would be good to have more access to grant data sets, for WA and AU.
- The Normal test was not possible as it needs a minimum of 8 samples and our dataset has 7 samples (7 quarters). We also didn't remove any outliers as there are not enough points left. The linear regression model is limited that it is not provided to predict values but to see if a correlation exist between our several investigations. The outcome will be more reliable with more data.

## If we had more time

- Extend our data time to include more previous years to compare to, and expand with Q3 and Q4 of 2020 so see if there is a recovery.
- Compare WA's performance against other states and possibly the world.
- Add commercial as well as residential development.
- Include the rest of 2020 to see data trends past the apparent recovery of WA.
- Explore trends by suburb in more detail.

---

## Questions??