The dataset we explore is a combination of 2018 census population from koordinates and the deprivation index from he university of otago website

The data set contains 1052 homes with 15 variables analysed. Notable variables include Capital value which is the value of the property, Bedrooms, Bathrooms and age of the property.
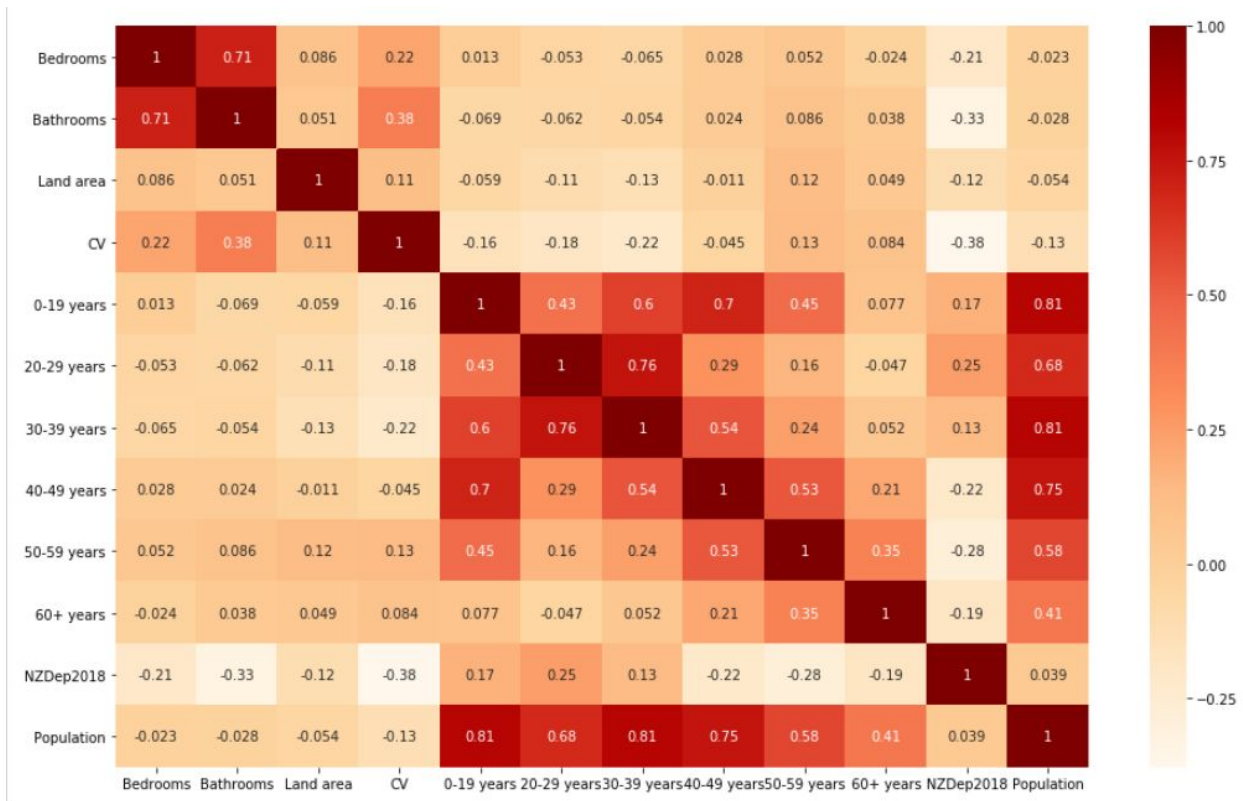
The years are columns of the number people who are in the age bracket that have lived in the house.

After plotting the data we see there are many data which have low or moderate correlation.

For the initial exploration of the data some descriptive statistics were acquired.
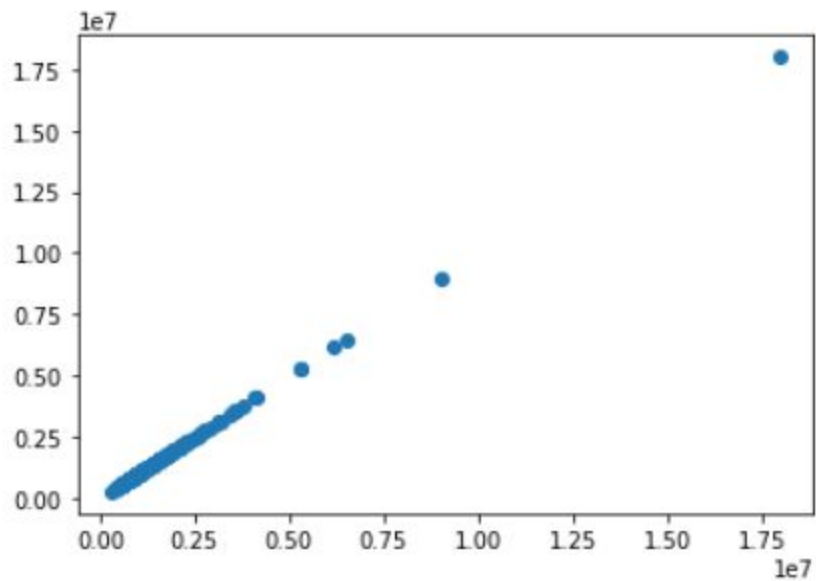
	Bedrooms	Bathrooms	CV	Latitude	Longitude	SA1	0-19 years	20-29 years	30-39 years	40-49 years	50-59 years	60+ years	SA12018_code	NZDep2018	Population
count	1048.000000	1048.000000	1.048000e+03	1048.000000	1048.000000	1.048000e+03	1048.000000	1048.000000	1048.00000	1048.000000	1048.000000	1048.000000	1.048000e+03	1048.000000	1048.000000
mean	3.779580	2.074427	1.388544e+06	-36.894561	174.799026	7.006332e+06	47.544847	28.915076	27.00000	24.131679	22.597328	29.353053	7.006332e+06	5.065840	179.799618
std	1.167894	0.992904	1.184422e+06	0.128426	0.117991	2.583920e+03	24.713408	20.993232	17.93158	10.956798	10.212455	21.810055	2.583920e+03	2.912027	71.087298
min	1.000000	1.000000	2.700000e+05	-37.265021	174.317078	7.001130e+06	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000	7.001130e+06	1.000000	3.000000
25%	3.000000	1.000000	7.800000e+05	-36.950873	174.722226	7.004426e+06	33.000000	15.000000	15.00000	18.000000	15.000000	18.000000	7.004426e+06	2.000000	138.000000
50%	4.000000	2.000000	1.080000e+06	-36.893409	174.798612	7.006334e+06	45.000000	24.000000	24.00000	24.000000	21.000000	27.000000	7.006334e+06	5.000000	174.000000
75%	4.000000	3.000000	1.600000e+06	-36.856280	174.880943	7.008390e+06	57.000000	36.000000	33.00000	30.000000	27.000000	36.000000	7.008390e+06	8.000000	207.750000
max	17.000000	8.000000	1.800000e+07	-36.177655	175.492424	7.011028e+06	201.000000	270.000000	177.00000	114.000000	90.000000	483.000000	7.011028e+06	10.000000	789.000000

To explore correlation a head map was used the right side scale represents the degree of correlation red being high and white being none. The value 1 represents strong correlation and -1 is no correlation.

This graph shows us there is a strong correlation between the number of bedrooms and bathrooms and some correlation between CV and bedrooms and bathrooms.

I used linear regression to model the data.


```
In [179]:  model.score(test_x, test_y)
Out[179]:  1.0
```