Comp309 Assignment 2 Vincent Yu 300390526

Objectives:

The goal of this assignment is to explore what kind of factors can affect the rental cost in Wellington and explore the relationship of these factors with the number of enrollment of our university.

.

Dataset selection

For this assignment, six datasets were selected to analyze and make predictions. The datasets selected were active-bonds, mean-rents, population, number of enrolments, GDP of NZ and the house value of Wellington. Rental cost should depend on the relationship between market supply and demand. The number of rental bonds can reflect how many houses are available in Wellington. The population of Wellington and the number of enrolment can reflect the demand for the rooms. Normally a large population would lead to higher rental cost. But if the number of available houses is much higher than the demand, it would cause the rental cost decrease. So these two datasets would be good evaluation factors. GDP shows the primary indicators used to gauge the health of a country's economy. Usually, the GDP rise would indirectly lead to a higher rental cost. So these datasets are valuable to explore which factor would affect the rental cost of Wellington.

Criteria of selecting datasets are datasets with less missing value, datasets have appropriate number of instances, the datasets should strongly correlate to the class.

Resources & data details:

rental-mean:

https://catalogue.data.govt.nz/dataset/rental-bond-data-by-region/resource/f53e86da-217f-49d9-86cb-cb56ad5cebd3

Active-bonds:

https://catalogue.data.govt.nz/dataset/rental-bond-data-by-region/resource/d477310f-47f1-4aa4-b9ce-84232badfa4b

Details:

These two files include the private bonds and the mean rental-cost starting from January 1993 to July 2018. These two datasets are counted monthly. It contains the data (mean rental cost and number of active bonds) of majorities cities in New Zealand.

Population:

http://nzdotstat.stats.govt.nz/wbos/Index.aspx?

DataSetCode=TABLECODE7501&_ga=2.46129063.281699752.15344666411077836335.1533857836& gac=1.218244843.1534213980.CjwKCAjw2MTbBRASE

iwAdYlpsamUN_pJAw8e7FAP27Rg2YNZXG_60uPM7EiyhS0iO6D8ztgBpOgrBhoC UdQQAvD_BwE#

Details:

This data set contains the population by regions of New Zealand. It contains the data of year 1996,2001 and from year 2007 to 2017. It contains 14 years data. The table shows the total population of the New Zealand. It also contains the population of several regions. (Auckland, Waikato, Bay Of Plenty region, Gisborne region, Hawke's Bay Region, Wellington etc).

GDP:

https://www.google.co.nz/publicdata/explore?ds=d5bncppjof8f9_&met_y=ny_gdp_mktp_cd&hl=en&dl=en

Details:

This dataset using the unit of billion US dollar. And it collect the data from year 1993 to year 2017.

Manipulation of Dataset:

Merge datasets:

For this assignment, we can use weka to merge multiple datasets into one and apply to a classifier(linear regression). It can be done by using "simple UCL" in weka. But for this assignment, it was done by excel. The datasets were integrated manually through excel. In the beginning, the dataset was changed counted yearly. But this would cause the number of instances to drop to around 20. So it led to a small dataset. It may cause the final result won't be very accurate.

Pre-processing:

Before import the data into a pipeline, the raw data still needed to be preprocessed. The biggest question was **missing values**. The number of enrollments only collected data from 2008 to 2017. For the missing value in the enrollment dataset, it is not a good idea to use linear regression. Because of the appearance, the number of enrollment has nothing to do with time. Since this, missing data for the number of enrollment were replaced by the nearest value. The house value set only hold five data, and it was hard to predict the missing value. Therefore this dataset was abandoned. The missing values in the GDP dataset were fixed by using linear regression. We can apply a filter called "relplacemissingvalue". But in this assignment, it turned out better use a linear regression in excel because excel can easily illustrate changes by plotting the graph.

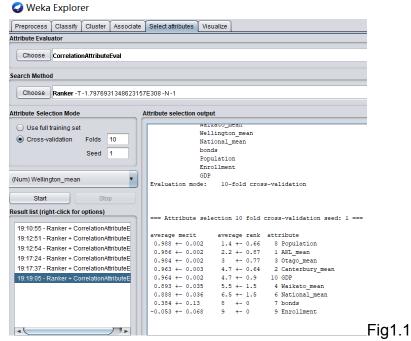
The table of datasets before and after manipulation could be found in the appendix.

Feature selection

After the dataset imported into the pipeline through ARFF loader, we could connect

to "data-visualizer" and set the rental cost of Wellington as Y-axis and set the others as X-axis in turn. We desire a linear relationship between rental-cost of Wellington and other relationship. So from the graph, we can roughly find the most correlate feature, which should have a nearly linear slope from the graph. But there is a better way to do this by using explore on the main menu of weka which will be included later.

For the most uncorrelated datasets, we can use a filter called "remove attribute" in Weka, which could remove a specific attribute. In this assignment, it was also applied to another tool in Weka to explore which factor is most correlated. As mentioned before, it is better to use the method in "Explore". After importing a dataset, we could explore the correlation level by using "correlation-Attribute-Eval". It showed how likely the attribute correlate to the class.



The average merit represents the correlation coefficient, the amplitude closer to 1, the better. And a positive number means a positive linear relationship, negative coefficient means a negative linear relationship.

As fig 1.1 shown above the most correlate datasets were the population of Wellington, mean-rental-cost of Auckland, mean-rental-cost of Otago, mean-rental-cost of Canterbury, GDP, mean-rental-cost of Waikato and mean-rental-cost of New Zealand is the most correlate feature. The enrollment has no linear relationship with the rental cost of Wellington. And the bonds of Wellington has a weak linear relationship.

From this table, the enrollment feature was filtered out, due to low correlation. But in order to explore which feature could affect the number of enrollment, this feature was set to class later. Weka is a very good tool for data mining, it has pre-implemented "normalize" filter which could help normalize the dataset which could reduce even eliminate data redundancy. It was applied to the pipeline used in this assignment, but it doesn't vary the result much due to the small size of instances.

Explore importance of attributes:

In this part, used pipeline in weka to explore which feature is the most important, and the combination of attributes to have the best performance. At this stage we have two main goals to achieve, one is the smallest root mean squared error and the other one is the highest the correlation coefficient. Small root mean squared error means less difference between values predicted by the model and the values in the data, which means the model is more reliable. High correlation coefficient means the model is more correlate to the data.

Minimize squared error

$$\begin{split} &\sum_{i=1}^{n} \varepsilon_{i}^{2} = \sum_{i=1}^{n} (y_{i} - \beta_{0} - \beta_{1} x_{i})^{2} \\ &0 = \frac{\partial \sum_{i=1}^{n} \varepsilon_{i}^{2}}{\partial \beta_{0}} = \frac{\partial \sum_{i=1}^{n} (y_{i} - \beta_{0} - \beta_{1} x_{i})^{2}}{\partial \beta_{0}} \\ &= -2(n\overline{y} - n\beta_{0} - n\beta_{1}\overline{x}) \\ &\hat{\beta}_{0} = \overline{y} - \hat{\beta}_{1}\overline{x} \end{split}$$
 Fig 2

As fig 2 shown above, it illustrates the formula for calculating squared error. The root mean error can be calculated by taking the square root of the squared error.

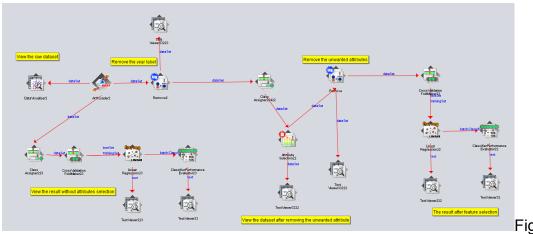


Fig 3

As fig 3 shown above, the pipeline is illustrated. After exploring by several trails, the best combination of datasets was the combination of the population of Wellington, GDP of New Zealand, the mean rental cost across New Zealand, and the active rental bonds in Wellington.

Result:

```
Scheme: LinearRegression22: LinearRegression
Options: -S 0 -R 1.0E-8 -num-decimal-places 4
Relation: RentalCost-weka.filters.unsupervised.attribute.Remove-Rl-weka.f

Correlation coefficient 0.9959
Mean absolute error 6.5905
Root mean squared error 7.7304
Relative absolute error 8.3702 %
Root relative squared error 8.7354 %
Total Number of Instances 26
```

Fig 4

The result for the best combination of attributes is shown in Fig 4. The less correlation attributes were filtered out. This process made the correlation coefficient increase. And it has the smallest root mean squared error.

The correlation coefficient is 0.9959, it is very close to 1. It means the model is very suit this classification. And the RMSE is only 7.7, it means the values predicted by the model is very close to the dataset.

```
Scheme: LinearRegression
Relation: RentalCost-weka.filters.unsupervised.attribute.Remove-Rl-weka.filters

Linear Regression Model

Wellington_mean =

0.0008 * Population +
-0.3714 * Otago_mean +
0.9452 * Waikato_mean +
0.3106 * National_mean +
-269.8059
```

Fig 5

Fig 5 illustrates the model found by Weka. The coefficient of population is very small because the values in population attributes are very large, and by linear regression it limits the coefficient to a very small number. We can predict the rental cost with this model. But maybe it won't be very accurate, because there are only four variables. Say we have a very large outlier in the mean-rental cost Waikato attribute, the rental cost of Wellington will be extremely large. And the result will be ridiculous. So if we want to have a accurate prediction we should collect more data and more correlated attributes.

Unfixed dataset:

А	В	U	U	E	г	U U	Н	l .	J	N
Year	AKL	Canterbury	Otago	Waikato	Wellington	National total	WellingtonBonds	Wellington Population	Wellington Enrollment	GDP
1993	199.0833	147.5	154.5833	134.1667	177.75	161.9166667	12750.33333	?	?	46.78
1994	216.8333	152.9166667	161.1667	138.6667	184.25	171.25	14108.58333	?	?	55.32
1995	246	163	168.8333	150.5833	194	185.3333333	15168.41667	?	?	63.92
1996	271.8333	172.4166667	169	162.4167	211.25	200.5	12750.33333	426900	?	70.14
1997	277.6667	176.5833333	163.6667	169.5833	222.5	207.0833333	14108.58333	?	?	66.07
1998	269.0833	174.6666667	161	171.1667	229.8333333	206.8333333	15168.41667	?	?	56.23
1999	263.75	171.5833333	179.0833	172	234.8333333	206.3333333	12750.33333	?	?	58.76
2000	266.1667	175.1666667	176.5833	170.3333	236.4166667	208.25	14108.58333	?	?	52.62
2001	275.1667	178.3333333	185.75	171.25	240.3333333	214.0833333	15168.41667	440200	?	53.87
2002	300.3333	191.75	200.8333	178.75	247.1666667	228.0833333	12750.33333	?	?	66.63
2003	324.9167	212.9166667	220.4167	194.1667	256.4166667	246.5833333	14108.58333	?	?	88.25
2004	330.8333	231.4166667	245.0833	210.6667	267.9166667	259.1666667	15168.41667	?	?	103.9
2005	333.5833	241.5833333	258.25	224.0833	277.9166667	268.4166667	12750.33333	?	?	114.7
2006	342.0833	254	261.4167	237.75	295.9166667	280.4166667	14108.58333	466300	?	111.6
2007	363.0833	267.9166667	273.25	253.9167	323.1666667	298.9166667	15168.41667	469300	?	137.3
2008	380.1667	280.4166667	283.75	265.3333	344.5	313	12750.33333	471800	21930	133
2009	382.0833	278	289.3333	268.9167	351.8333333	314.9166667	14108.58333	475600	22925	121.3
2010	394.1667	284.1666667	291.9167	272.6667	363.5	323.1666667	15168.41667	479400	22880	146.6
2011	414.75	302.8333333	310.0833	278.3333	365.25	334.4166667	12750.33333	483400	22560	168.5
2012	430.4167	329.5833333	317	283.5833	371.4166667	347.4166667	14108.58333	485100	21195	176.2
2013	444.8333	366.5833333	327	290.3333	379.5	361.6666667	15168.41667	486700	21480	190.5
2014	463.5833	394.9166667	341.8333	297.5	390.1666667	376.5833333	12750.33333	491400	21190	200.7
2015	493.6667	388.1666667	359	310.1667	400.0833333	394.6666667	14108.58333	496900	21450	175.6
2016	514.5833	385.1666667	385.1667	333.6667	420.5833333	412.9166667	15168.41667	504900	21950	185
2017	534.3333	367.8333333	415	357.3333	444.6666667	431.75	12750.33333	513900	?	?
2018	549.4286	374.1428571	424.2857	216.6667	477	261.5833333	45746.57143	510898.4848	?	?

(Note: the cell filled with "?" are the missing values)

Fixed dataset:

1993 199 0833 147.5 154.5833 134.1667 177.75 161.9166667 12750.33333 412103.7251 21930 46 1994 216.8333 152.9166667 161.1667 138.6667 184.25 171.25 14108.58333 416059.7951 21930 65 1996 271.8333 172.4166667 169 162.4167 211.25 2005 12750.33333 426900 21930 76 1997 277.6667 176.5833333 163.6667 169.5833 222.5 207.0833333 14108.58333 427395.6297 21990 66 1998 269.833 174.66666667 161 171.1667 229.8333333 206.8333333 14108.58333 427395.6297 21990 66 1999 263.75 171.5833333 1779.9833 172 234.8333333 206.8333333 12750.33333 435383.8234 21930 56 2000 266.1667 175.1666667 176.5833 170.3333 236.4166667 208.25 14108.58333 47397.9202 21	А	D	U	ט	E	Г	G	П	l l	J	Γ
1994 216.8333 152.9166667 161.1667 138.6667 184.25 171.25 14108.58333 416059.7951 21930 55 1995 246 163 168.8333 150.5833 194 185.3333333 151.664.1667 420015.8651 21930 62 1997 277.6667 176.5833333 163.6667 169.5833 222.5 207.0833333 14108.58333 42799.6297 21930 66 1998 269.0833 174.6666667 161 171.1667 229.8333333 206.8333333 151.668.41667 431.389.7266 21930 56 1999 263.75 171.58333333 170.9333 226.43833333 206.3333333 12750.33333 435383.8234 21930 56 2000 266.1667 175.1666667 176.5833 170.3333 236.4166667 208.25 141.08.58333 43937.7902 21930 56 2001 275.1667 178.3333333 185.75 171.25 240.3333333 214.0833333 15168.41667 40200 21930	Year	AKL	Canterbury	Otago	Waikato	Wellington	National total	WellingtonBonds	Wellington Population	Wellington Enrollment	GDP
1995 246 163 168 8333 150 5833 194 185,3333333 15168 41667 42015 8651 21930 63 1996 271 8333 172 4166667 169 162 4167 211 25 205 12750 33333 426900 21930 66 1997 277,6667 176,5833333 16,6667 161 171 1667 229,8333333 206,8333333 15168,41667 431389,7266 21930 66 1999 263,75 171,5833333 179,0833 172 234,8333333 206,83333333 15168,41667 431389,7266 21930 56 2000 266,1667 175,1666667 176,5833 170,3333 226,4166667 202,25 14108,58333 439377,9202 21930 56 2001 275,1667 178,3333333 178,75 247,1666667 228,0333333 15168,41667 440200 21930 56 2001 275,1667 178,333333 178,75 247,1666667 228,033333 12750,33333 448112,4709 21930 56	1993	199.0833	147.5	154.5833	134.1667	177.75	161.9166667	12750.33333	412103.7251	21930	46.78
1996 271.8333 172.4166667 169 162.4167 211.25 200.5 12750.33333 426900 21930 77 1997 277.6667 176.5833333 163.667 169.5833 22.5 207.0833333 14108.58333 427395.6297 21930 66 1998 269.0833 174.6666667 161 171.1667 229.8333333 206.83333333 12750.33333 435388.8234 21930 56 2000 266.1667 175.16666667 176.5833 170.3333 236.4166667 208.25 14108.58333 439377.9202 21930 56 2001 275.1667 178.3333333 185.75 171.25 240.3333333 214.0833333 15168.41667 440200 21930 53 2001 275.1667 178.33333333 185.75 171.25 240.3333333 12750.33333 448112.4709 21930 56 2001 232.916667 212.9146667 194.1667 256.4166667 226.08333333 14108.58333 45296.726 21930 15 <td>1994</td> <td>216.8333</td> <td>152.9166667</td> <td>161.1667</td> <td>138.6667</td> <td>184.25</td> <td>171.25</td> <td>14108.58333</td> <td>416059.7951</td> <td>21930</td> <td>55.32</td>	1994	216.8333	152.9166667	161.1667	138.6667	184.25	171.25	14108.58333	416059.7951	21930	55.32
1997 277.6667 176.5833333 163.6667 169.5833 222.5 207.0833333 14108.58333 427395.6297 21930 66 1999 269.0833 174.6666667 161 171.1667 229.8333333 206.83333333 12750.33333 43588.2344 21930 56 1999 263.75 171.5833333 170.3333 236.4166667 208.25 14108.58333 439377.9202 21930 56 2001 275.1667 178.3333333 185.75 171.25 240.3333333 214.0833333 15168.41667 440200 21930 52 2002 300.3333 191.75 200.83333 178.75 247.1666667 228.0833333 1478.06 440200 21930 58 2003 324.9167 212.9166667 204.4167 194.1667 256.4166667 245.833333 14708.58333 452036.5967 21930 88 2004 330.8333 231.4166667 245.0833 217.9166667 256.9166667 1516667 455.94667 259.166667 156.64166	1995	246	163	168.8333	150.5833	194	185.3333333	15168.41667	420015.8651	21930	63.92
1998 269 0833 174 6666667 161 171.1667 229 8333333 206 8333333 15168 41667 431389 7266 21930 56 1999 263 75 171.5833333 179 0833 172 234 8333333 206 3333333 12750 33333 435383 8234 21930 52 2001 275 1667 175 1666661 176 5833 170 3333 236 4166667 208.25 14108 58333 43937 7900 21930 52 2001 275 1667 178 33333333 185 75 171 25 240 33333333 15168 41667 440200 21930 52 2002 300.3333 191 75 200 8333 178.75 247.1666667 228.0833333 12750 33333 448112 4709 21930 66 2004 330.8333 231.4166667 220.41667 256.4166667 259.166667 15168.41667 459060 7226 21930 16 2005 333.5833 241.4167 237.75 295.9166667 259.166667 12750.33333 459884.8485 21930 11	1996	271.8333	172.4166667	169	162.4167	211.25	200.5	12750.33333	426900	21930	70.14
1999 263.75 171.5833333 179.0833 172 234.8333333 206.3333333 12750.33333 435383.8234 21930 58 2000 266.1667 175.1666667 176.58333 236.4166667 208.25 14108.58333 439377.9202 21930 52 2001 275.1667 178.3333333 185.75 171.25 240.3333333 214.0833333 15168.41667 440200 21930 53 2002 300.3333 191.75 200.8333 178.75 247.1666667 228.0833333 12750.33333 448112.4709 21930 66 2003 324.9167 212.9166667 226.4166667 226.4166667 226.4166667 245.0833333 14108.58333 452036.5967 21930 88 2004 330.8333 241.566667 224.0833 227.9166667 259.1666667 15168.41667 455960.7226 21930 11 2006 342.0833 205.253.9167 237.75 295.9166667 280.4166667 14108.58333 456830 21930 11	1997	277.6667	176.5833333	163.6667	169.5833	222.5	207.0833333	14108.58333	427395.6297	21930	66.07
2000 266 1667 175 1666667 176 5833 170 3333 236 4166667 208 25 14108 58333 439377 9202 21930 52 2001 275 1667 178 3333333 185 75 171 25 240 3333333 15168 414667 440200 21930 55 2002 300 3333 191 75 200 8333 178 75 224 7.1666667 228 0833333 12750 33333 448112 4709 21930 66 2004 330 8333 231 4166667 220 4167 194 1667 256 4166667 245 6833333 14108 58333 452036 5967 21930 88 2004 330 8333 231 4166667 245 0833 210.6667 259 166667 259 1666667 15168 41667 455960 7226 21930 11 2005 333 5833 241 5833333 258 25 224.0833 277.9166667 268 4166667 1408 58333 459848 8485 21930 11 2007 363 0833 267.9166667 237.55 253 9167 323 166667 298 9166667 14108 58333 471800	1998	269.0833	174.6666667	161	171.1667	229.8333333	206.8333333	15168.41667	431389.7266	21930	56.23
2001 2751667 178 3333333 185.75 171.25 240.3333333 1214.0833333 15168 41667 440200 21930 53 2002 300.3333 191.75 200.83333 178.75 247.1666667 228.0833333 12750.33333 448112 4709 21930 66 2003 324.9167 212.9166667 220.4167 194.1667 256.4166667 246.5833333 14108.58333 452036.5967 21930 86 2004 330.8333 231.4166667 245.0833 210.6667 267.9166667 256.9166667 15168.41667 455960.7226 21930 16 2005 333.5833 241.5833333 258.25 224.0833 277.9166667 268.4166667 12750.33333 468000 21930 11 2007 363.0833 257.9166667 237.55 259.9166667 280.4166667 14108.58333 46800 21930 13 2007 363.0833 267.9166667 237.55 253.9167 323.166667 298.9166667 14108.58333 471800	1999	263.75	171.5833333	179.0833	172	234.8333333	206.3333333	12750.33333	435383.8234	21930	58.76
2002 300.3333 191.75 200.8333 178.75 247.1666667 228.0833333 12750.33333 448112.4709 21930 66 2003 324.9167 212.9166667 226.4166667 256.4166667 246.5833333 14108.58333 452036.5967 21930 18 2004 330.8333 231.4166667 245.0833 210.6667 259.166667 15168.41667 455960.7226 21930 10 2005 333.5833 224.5833333 258.25 224.0833 277.9166667 259.166667 12750.33333 459884.8485 21930 11 2006 342.0833 254 261.4167 237.75 295.9166667 280.4166667 14108.58333 459884.8485 21930 11 2007 363.0833 267.9166667 273.25 253.9167 323.1666667 280.4166667 14108.58333 466300 21930 11 2008 380.1667 280.4166667 283.75 265.3333 344.5 313 12750.33333 471800 21930 12 </td <td>2000</td> <td>266.1667</td> <td>175.1666667</td> <td>176.5833</td> <td>170.3333</td> <td>236.4166667</td> <td>208.25</td> <td>14108.58333</td> <td>439377.9202</td> <td>21930</td> <td>52.62</td>	2000	266.1667	175.1666667	176.5833	170.3333	236.4166667	208.25	14108.58333	439377.9202	21930	52.62
2003 324 9167 212 9166667 220 4167 194 1667 256 4166667 246 583333 14108 58333 452036 5967 21930 88 2004 330 8333 231 41666667 245 0833 210 6667 259 166667 15168 41667 455960 7226 21930 10 2005 333 5833 241 5833333 258 25 224 0833 277 9166667 268 4166667 12750 33333 45988 48465 21930 11 2007 363 0833 267 9166667 273 25 253 9167 323 1666667 298 9166667 15168 41667 469300 21930 11 2008 380 1667 280 4166667 280 4166667 298 9166667 15168 41667 469300 21930 12 2008 380 1667 280 4166667 280 4166667 280 416667 469300 21930 13 2009 382 0833 278 289 3333 286 9167 3518333333 314 9166667 14108 58333 475600 22925 12 2010 394 1667 284 1666667 291 9197 <td>2001</td> <td>275.1667</td> <td>178.3333333</td> <td>185.75</td> <td>171.25</td> <td>240.3333333</td> <td>214.0833333</td> <td>15168.41667</td> <td>440200</td> <td>21930</td> <td>53.87</td>	2001	275.1667	178.3333333	185.75	171.25	240.3333333	214.0833333	15168.41667	440200	21930	53.87
2004 330.8333 231.4166667 245.0833 210.6667 267.9166667 259.1666667 15168.41667 455960.7226 21930 10 2005 333.5833 241.5833333 258.25 224.0833 277.75 295.9166667 268.4166667 12750.33333 459884.8485 21930 11 2007 363.0833 257.9166667 273.25 253.9167 323.166667 298.916667 14108.58333 466300 21930 13 2008 380.1667 280.4166667 283.75 265.3333 344.5 313 12750.33333 471800 21930 13 2009 382.0833 278 289.3333 268.9167 351.8333333 314.9166667 14108.58333 475600 22925 12 2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 14108.58333 475600 22925 12 2011 414.75 302.8333333 310.0833 278.33333 365.25 334.4166667 1479400 22880 </td <td>2002</td> <td>300.3333</td> <td>191.75</td> <td>200.8333</td> <td>178.75</td> <td>247.1666667</td> <td>228.0833333</td> <td>12750.33333</td> <td>448112.4709</td> <td>21930</td> <td>66.63</td>	2002	300.3333	191.75	200.8333	178.75	247.1666667	228.0833333	12750.33333	448112.4709	21930	66.63
2005 3335833 2415833333 258 25 2240833 277.9166667 268.4166667 12750.33333 459884.8485 21930 11 2006 342.0833 254 261.4167 237.75 295.9166667 280.4166667 14108.58333 468300 21930 11 2007 363.0833 267.9166667 283.75 253.9167 323.1666667 298.9166667 15168.41667 469300 21930 13 2008 380.1667 280.4166667 283.75 265.3333 344.5 313 12750.33333 471800 21930 13 2009 382.0833 278 289.3333 268.9167 351.8333333 314.9166667 14108.58333 475600 22925 12 2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 15168.41667 479400 22880 14 2011 414.75 302.8333333 317 283.5833 371.4166667 12750.33333 483400 22560 16	2003	324.9167	212.9166667	220.4167	194.1667	256.4166667	246.5833333	14108.58333	452036.5967	21930	88.25
2006 342.0833 254 261.4167 237.75 295.9166667 280.4166667 14108.58333 466300 21930 11 2007 363.0833 267.9166667 273.25 253.9167 323.1666667 298.9166667 15168.41667 469300 21930 13 2008 380.1667 280.4166667 283.75 265.3333 344.5 313 12750.33333 471800 21930 21930 22925 12 2009 382.0833 278 289.3333 268.9167 351.8333333 314.9166667 14108.58333 475600 22925 12 2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 15168.41667 479400 22880 14 2011 414.75 302.8333333 310.0833 278.3333 365.25 334.4166667 12750.33333 485100 22195 12 2012 430.4167 329.5833333 371.4166667 347.4166667 14108.58333 485100 21195 17<	2004	330.8333	231.4166667	245.0833	210.6667	267.9166667	259.1666667	15168.41667	455960.7226	21930	103.9
2007 363.0833 267.9166667 273.25 253.9167 323.1666667 298.9166667 15168.41667 469300 21930 13 2008 380.1667 280.41666667 283.75 265.3333 344.5 313 12750.33333 471800 21930 21930 200 382.0833 278 289.3333 288.9167 351.8333333 314.9166667 14108.58333 475600 22925 12 2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 15168.41667 479400 22880 14 2011 414.75 302.8333333 310.0833 278.3333 365.25 334.4166667 12750.33333 483400 22560 16 2012 430.4167 329.5833333 371.4166667 347.4166667 14108.58333 485100 21195 17 2013 444.8333 366.5833333 327 290.3333 379.5 361.666667 15168.41667 486700 21480 15 2014 463.5833 <td< td=""><td>2005</td><td>333.5833</td><td>241.5833333</td><td>258.25</td><td>224.0833</td><td>277.9166667</td><td>268.4166667</td><td>12750.33333</td><td>459884.8485</td><td>21930</td><td>114.7</td></td<>	2005	333.5833	241.5833333	258.25	224.0833	277.9166667	268.4166667	12750.33333	459884.8485	21930	114.7
2008 380.1667 280.4166667 283.75 265.3333 344.5 313 12750.33333 471800 21930 2009 382.0833 278 289.3333 268.9167 351.8333333 314.9166667 14108.58333 475600 22925 12 2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 15168.41667 479400 22880 14 2011 414.75 302.8333333 310.0833 278.3333 365.25 334.4166667 12750.33333 483400 22560 18 2012 430.4167 329.8833333 317 283.5833 371.4166667 347.4166667 14108.58333 485100 21195 17 2013 444.8333 366.5833333 327 290.3333 379.5 361.6666667 15168.41667 486700 21480 19 2014 463.5833 394.9166667 349.333 297.5 390.1666667 376.5833333 12750.33333 491400 21190 20 <tr< td=""><td>2006</td><td>342.0833</td><td>254</td><td>261.4167</td><td>237.75</td><td>295.9166667</td><td>280.4166667</td><td>14108.58333</td><td>466300</td><td>21930</td><td>111.6</td></tr<>	2006	342.0833	254	261.4167	237.75	295.9166667	280.4166667	14108.58333	466300	21930	111.6
2009 382 0833 278 289 3333 268 9167 351.8333333 314.9166667 14108 58333 475600 22925 12 2010 394 1667 284 1666667 291 9167 272 6667 363.5 323 1666667 15168 41667 479400 22880 14 2011 414.75 302 8333333 310 0833 278 3333 365.25 334 4166667 12750 33333 483400 22560 16 2012 430 4167 329 5833333 317 283 5833 371 4166667 347 4166667 147 4108 58333 485100 21195 17 2013 444 8333 365 5833333 327 290 3333 379.5 361.6666667 15168 41667 486700 21480 19 2014 463 5833 394 9166667 347 383 297.5 390 1666667 376 5833333 12750 33333 491400 21190 20 2015 493 6667 381 1666667 347 3833 394 666667 14108 5833 496900 21450 17 2016	2007	363.0833	267.9166667	273.25	253.9167	323.1666667	298.9166667	15168.41667	469300	21930	137.3
2010 394.1667 284.1666667 291.9167 272.6667 363.5 323.1666667 15168.41667 479400 22880 14 2011 414.75 302.833333 310.0833 278.3333 365.25 334.4166667 12750.33333 483400 22560 16 2012 430.4167 329.5833333 317 283.5833 371.4166667 14108.58333 485100 21195 17 2013 444.8333 366.5833333 327 290.3333 379.5 361.6666667 15168.41667 486700 21480 19 2014 463.5833 394.9166667 341.8333 297.5 390.1666667 376.5833333 12750.33333 491400 21190 20 2015 493.6667 388.1666667 351.667 400.0833333 394.6666667 14108.58333 496900 21450 17 2016 514.5833 385.16667 385.1667 420.58333333 412.9166667 15168.41667 504900 21950 20 2017 534	2008	380.1667	280.4166667	283.75	265.3333	344.5	313	12750.33333	471800	21930	133
2011 414.75 302.8333333 310.0833 278.3333 365.25 334.4166667 12750.33333 483400 22560 16 2012 430.4167 329.5833333 317 283.5833 371.4166667 347.4166667 14108.58333 485100 21195 17 2013 444.8333 366.8833333 327 290.3333 379.5 361.6666667 15168.41667 486700 21480 19 2014 463.5833 394.9166667 341.8333 297.5 390.1666667 376.5833333 12750.33333 491400 21190 20 2015 493.6667 388.1666667 359 310.1667 400.0833333 394.66666667 14108.58333 496900 21450 17 2016 514.5833 385.1666667 385.1667 333.6667 420.5833333 412.9166667 15168.41667 504900 21950 20 2017 534.3333 367.8333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950	2009	382.0833	278	289.3333	268.9167	351.8333333	314.9166667	14108.58333	475600	22925	121.3
2012 430.4167 329.5833333 317 283.5833 371.4166667 347.4166667 14108.58333 485100 21195 17 2013 444.8333 366.5833333 327 290.3333 379.5 361.6666667 15168.41667 486700 21480 19 2014 463.5833 394.9166667 341.8333 297.5 390.1666667 376.5833333 12750.33333 491400 21190 20 2015 493.6667 388.1666667 359 310.1667 400.0833333 394.66666667 14108.58333 496900 21450 17 2016 514.5833 385.1666667 385.1667 333.6667 420.58333333 412.9166667 15168.41667 504900 21950 20 2017 534.3333 367.8333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950 20	2010	394.1667	284.1666667	291.9167	272.6667	363.5	323.1666667	15168.41667	479400	22880	146.6
2013 444 8333 366 5833333 327 290 3333 379.5 361 6666667 15168 41667 486700 21480 19 2014 463 5833 394 9166667 341 8333 297.5 390 1666667 376 5833333 12750 333333 491400 21190 20 2015 493 6667 388 1666667 359 310 1667 400 0833333 394 6666667 14108 58333 496900 21450 17 2016 514 5833 385 1666667 385 1667 333 6667 420 58333333 412 9166667 15168 41667 504900 21950 2017 534 3333 367 8333333 415 357 3333 444 6666667 431.75 12750 33333 513900 21950 20	2011	414.75	302.8333333	310.0833	278.3333	365.25	334.4166667	12750.33333	483400	22560	168.5
2014 463.5833 394.9166667 341.8333 297.5 390.1666667 376.5833333 12750.33333 491400 21190 20 2015 493.6667 388.1666667 359 310.1667 400.0833333 394.6666667 14108.58333 496900 21450 17 2016 514.5833 385.1666667 385.1667 336.667 420.5833333 412.9166667 15168.41667 504900 21950 20 2017 534.3333 367.8333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950 20	2012	430.4167	329.5833333	317	283.5833	371.4166667	347.4166667	14108.58333	485100	21195	176.2
2015 493.6667 388.1666667 359 310.1667 400.0833333 394.6666667 14108.58333 496900 21450 17 2016 514.5833 385.1666667 385.1667 333.6667 420.5833333 412.9166667 15168.41667 504900 21950 2017 534.3333 367.8333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950 20	2013	444.8333	366.5833333	327	290.3333	379.5	361.6666667	15168.41667	486700	21480	190.5
2016 514.5833 385.1666667 385.1667 333.6667 420.5833333 412.9166667 15168.41667 504900 21950 2007 2017 534.3333 367.83333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950 2007	2014	463.5833	394.9166667	341.8333	297.5	390.1666667	376.5833333	12750.33333	491400	21190	200.7
2017 534.3333 367.8333333 415 357.3333 444.6666667 431.75 12750.33333 513900 21950 20	2015	493.6667	388.1666667	359	310.1667	400.0833333	394.6666667	14108.58333	496900	21450	175.6
	2016	514.5833	385.1666667	385.1667	333.6667	420.5833333	412.9166667	15168.41667	504900	21950	185
2018 549.4286 374.1428571 424.2857 216.6667 477 261.5833333 45746.57143 510898.4848 21950 213	2017	534.3333	367.8333333	415	357.3333	444.6666667	431.75	12750.33333	513900	21950	205.6
	2018	549.4286	374.1428571	424.2857	216.6667	477	261.5833333	45746.57143	510898.4848	21950	213.26

(Note: the data in red are fixed by linear regression, the data in blue are fixed by choosing the nearest value)