Washington Home Loan Data Project 1

Introduction

For this project, we were given this scenario: A mortgage company (the client) wants to set up operations in Washington State and is ready to open two branch offices in different counties; find the best two counties. After looking over the data, I decided to create a scoring system, where I would make a variety of measurements and add them together to get a score I could use to rank the counties.

Data Description

Sources

We were given three data sets to consider. One contained information generated by banks regarding mortgages that is collected by the federal government for 2016 (HMDA.) Another was a national list of average monthly mortgage delinquency rates by county from January 2008 through June 2017, giving the percentage of mortgages that were either one or two months delinquent (Del 30), or three months or more delinquent (Del 90). The third consisted of two annual reports prepared at the University of Washington, one with data as of the end of March, 2016 (2016-Q1) and as of the end of December, 2016 (2016-Q4), which provided housing information that included the number of homes, how many existing homes were sold, and how many new permits were issued for building new homes.

In addition to the provide data, I went to the website providing the housing information and retrieved the report giving numbers as of December 2017 (2017-Q4.) I also went to the Washington Office of Financial

Management and retrieved annual population data for counties from 2000 through 2016 (Pop), with 2000 and 2010 giving the official census count and the other years giving the official estimate.

Selection

Del 30 and Del 90 include every county in the United States with relevant data, which is defined as having at least 1000 mortgages issued in 2016. For Washington, this meant only 10 of the state's 40 counties were included. A check showed that those 10 counties represented not quite 85% of all entries in HDMA. It therefore seemed reasonable to limit my analysis to just those 10 counties, as the remaining counties would not have a sufficient volume of mortgages to interest a mortgage company. These counties are: Benton, Clark, King, Kitsap, Pierce, Snohomish, Spokane, Thurston, Whatcom, and Yakima.

Metrics

I used the data to generate nine metrics. When necessary, I used a scaling factor so that the maximum metric for all counties was greater than or equal to .100 and less than 1. I used a smaller divisor on Del 90 than I used on Del 30, which had the effect of giving Del 90 (the percentage of extant mortgages that were delinquent by 3 or more months) slightly more weight than Del 30. In some cases, I used the complement of the metric so that larger values would be draws for the client.

Percent of population

For this metric, I used the 2016 population estimate in Pop. The population of each county was divided by the sum of the population for the ten counties. This gave me a measure of each county's relative size by population. It indicates that larger populations would likely mean more business for our client. This value was already in the desired range, and was not scaled.

Population trend

To calculate the population trend, I found the trend line in how the county's population has changed since 2000 and used the slope of that line as my metric. Basically, the slope gives the average change in population each year and provides a measure of how quickly a county's population is changing. This is useful, as

rapid increases will likely mean an increased number of homes being bought. To put the value into my desired range, the slope was divided by 100,000.

Percent of successful mortgages by population

This metric was calculated by taking the count of successful mortgages – those that were offered, approved, and accepted – and dividing it by the 2016 population estimate. This gives a measure of how often residents of the county change homes, another useful consideration for a mortgage company. Because this ratio was pretty low, I multiplied the value by 10.

Mean percent short-term and long-term delinquencies

The data in Del30 and Del90 varied little from month to month, so there was no real trend to look at. Instead, I took the mean of all the values given for a county. The Del30 means were divided by 10 to bring it into my desired range, and the Del90 means by 7. This value was then subtracted from 1, in effect getting the mean value of mortgages that were *not* in default. This was to keep with the idea that larger values would be of stronger interest to the client. The effect of the smaller scaling factor on Del90 was to lower this metric, as a higher incident of long-term default should negatively affect the final score.

Existing single-family home sales trend

From 2017-Q4, I used the reported count of single-family homes that went up for sale in each county between 2010 and 2017 inclusive and found the slope of the data's trend line. This gives a measure of how quickly the home sales market is changing, which bears directly on opportunities to write mortgages. This was scaled by dividing the value by 2000.

New single-family home permits trend

Also from 2017-Q4, I found the trend line in the number of permits to build new single-family homes given every year from 2007 through 2016. This is a useful measure of how quickly new homes are coming on to the market. This was scaled by dividing the value by 1000.

Percent single-family homes by population

My last metric is the number of single-family homes in the county divided by the population in the county. This gives a measure of how much of the population live in single-family homes rather than apartments and other multi-family buildings. I am undecided as to how useful this metric is, as it does not take into consideration how many of those homes are rentals, and how much of the population owns a condo or part of a duplex. Because it is a percentage, it was already in range and did not need to be scaled.

Generating the score

To generate the score, I used to evaluate the ten counties, I added all of the metrics together, and multiplied the result by 100.

Findings

The calculated scores for each of the ten counties, ranked high to low, is given in Table 1. A graphical representation of the final scores is given in Figure 1. The length of the bars matches the county's score, and each bar is shaded to show how much each metric contributed to that score. A full table of the calculated metrics is available in the appendix.

Suitability scores by county

County	Score
King County	419.03
Clark County	327.61
Pierce County	320.82
Spokane County	320.08
Snohomish County	314.38
Whatcom County	299.51
Benton County	298.60
Kitsap County	296.56
Thurston County	283.01
Yakima County	247.00

Table 1

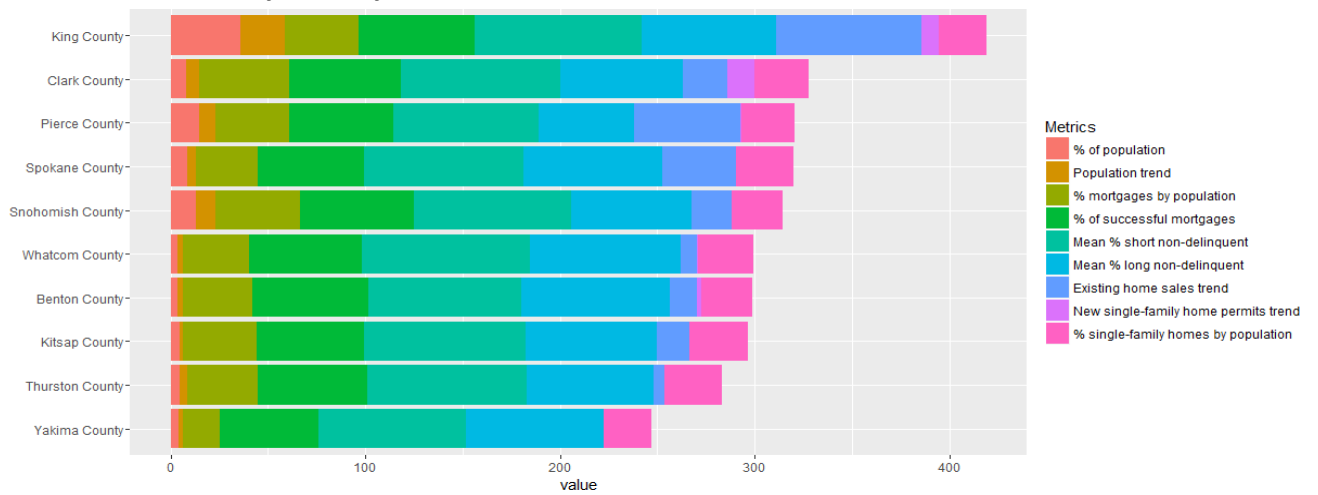**Breakdown of suitability score by metric**

Figure 1

Conclusions

Not surprisingly, King County topped the list: it is, after all, the most populous county in Washington by far, and home to the state's largest metropolitan area. The second-ranked county is Clark; while it is rather small in terms of geography, its main city of Vancouver sits just north of the Columbia River from Portland, Oregon, and has become an important residential district for people who work in Portland. Clark County ranked first in the number of successful mortgages given per population, and in the trend of giving more building permits for single family homes.

Based on these scores, I would recommend that the client establish branch offices in King County and Clark County, because they show the most promise for growth of their housing markets. While I did not consider the cost of homes in my evaluation, the housing data from the University of Washington shows that King County has the highest median home prices in the state. Because mortgage companies earn a percentage of the mortgages they write, King County represents a large amount of potential profit. Home prices in Clark County are close the state average, but that county shows great potential for increasing future growth as population climbs and new homes become available.

Appendix

Detailed breakdown of metrics used to build scores

County	Score	Raw	% Pop	Pop Trend	Approved /Pop
King County	419.03	4.1903	0.3603	0.2243	0.3803
Clark County	327.61	3.2761	0.0788	0.0672	0.4638
Pierce County	320.82	3.2082	0.1438	0.0855	0.3823
Spokane County	320.08	3.2008	0.0836	0.0485	0.3178
Snohomish County	314.38	3.1438	0.1321	0.0998	0.4343
Whatcom County	299.51	2.9951	0.0362	0.0280	0.3391
Benton County	298.60	2.9860	0.0324	0.0305	0.3541
Kitsap County	296.56	2.9656	0.0442	0.0178	0.3800
Thurston County	283.01	2.8301	0.0463	0.0411	0.3603
Yakima County	247.00	2.4700	0.0423	0.0186	0.1938

County	% Approved	1 - Short	1 - Long	Sold Trend	Permits Trend	SF / Pop
King County	0.5948	0.8619	0.6904	0.7449	0.0903	0.2431
Clark County	0.5745	0.8191	0.6306	0.2269	0.1391	0.2760
Pierce County	0.5337	0.7448	0.4882	0.5464	0.0060	0.2774
Spokane County	0.5448	0.8161	0.7133	0.3874	-0.0049	0.2941
Snohomish County	0.5827	0.8109	0.6190	0.2158	-0.0136	0.2628
Whatcom County	0.5798	0.8618	0.7751	0.0866	0.0020	0.2865
Benton County	0.5980	0.7859	0.7658	0.1376	0.0234	0.2585
Kitsap County	0.5533	0.8258	0.6774	0.1774	-0.0130	0.3026
Thurston County	0.5624	0.8199	0.6523	0.1260	-0.0716	0.2935
Yakima County	0.5030	0.7591	0.7095	0.0515	-0.0502	0.2423

Score – The raw score multiplied by 100

Raw Score – The sum of the nine metrics.

% Pop – The population of the county as of 2016 relative to the total 2016 population of all ten counties.

Pop Trend – The average change in the county's population over time. Positive values indicate an increase, and larger values indicate a faster change.

Approved / Pop – The number of approved mortgages divided by the 2016 population of the county.

% Approved – The number of mortgages approved in 2016 divided by the number of mortgages started in 2016.

 1 - Short – The complement of the average percent of existing mortgages that were in default by one or two months.

1 - Long – The complement of the average percent of existing mortgages that were in default by three months or more.

Sold Trend – The average change in how many homes were sold in the county over time. Positive values indicate an increase, and larger values indicate a faster change.

Permits Trend – The average change in how many building permits were issued for new single-family homes. Positive values indicate an increase while negative values indicate a decrease. Larger values indicate a faster change.

SF / Pop – The number of single family homes in the county as of the end of 2016, divided by the county's population in 2016.