Factors of Denver's Residential Real Estate

Principal Investigator: Andrew Peterson (acpeterson7@wisc.edu)

This study examines the changes in housing prices in Denver, Colorado. In order to determine the variables that are most important, I performed multivariate regression to isolate and quantify variables that are economically and statistically important. The results from the regression were useful but did not provide any information on housing demand. As a part of my analysis, I also tracked the population growth and average home prices. There appears to be a strong correlation between population growth and the mean home price.

Analysis

Regressions

I found data from Denver's government website to perform this regression. The data was purely cross-sectional. The data covered 46 variables. I reshaped the data to the nine most important variables: total value, assessed value, land area, the area above ground, basement area, finished basement area, number of bedrooms, number of full bathrooms, number of half-bathrooms, and years since built. I performed two regressions. The first regression was to determine an equation for the assessed value. The second regression was intended to predict the total value based on these factors. I found that both regressions had p-values of zero, therefore all variables were statistically significant. The regression on assessed value had a much lower intercept than the regression on total value. The substantial difference between these regressions lead me to believe that housing demand was high.

Population

Housing demand can be driven by many factors, but the two most important are income and population. I gathered data from the US Census to examine changes in the population of the state of Colorado. The dataset only provided estimates from 2010 to 2019. According to Figure 1, the population of colorado has been growing over the period. The population growth could be an explanation for increased housing demand. If housing demand were in fact increasing, the sales price would also be increasing.

Evolution of Price

I used data from Zillow to perform a time-series analysis. The Zillow Home Value Index (ZHVI) is "a smoothed, seasonally adjusted measure of the typical home value and market changes across a given region and housing type" (Zillow). I used state, county, and city data tables to track how average housing prices changed from 2000 to the present. In order to do this, I downsampled the data to yearly estimates. I found the results to be interesting because the average price at the state level matched the county and city levels. There was no differentiation.

This is apparent in Figure 2. The only explanation I can think of for this behavior is that an overwhelming majority of property transactions are taking place in the city of Denver. My reasoning stems from the fact that transactions in Denver would impact the county and the state as well.

Conclusions and directions for future research

Residential real estate in Denver is clearly becoming more expensive. The fact that the value of the average home has doubled in the last ten years can be a signal for any number of economic events or opportunities. Given that the market has already seen explosive growth over the last ten years, investing in the market may not be a good idea. The market has signs of a housing bubble. At the very least, the housing market in Denver, Colorado will stop growing at the current rate. I started this project hoping to confirm my idea that Denver has real estate opportunities at low costs. That is not the conclusion I reached. The future of this research will lead to other cities in the United States. Tampa, Florida is an interesting place. I will apply many of the same methods to analyze that market.

Appendix

Regression 1

OLS Regression Results

Dep. Variable:	ASSESS_VALUE	R-squared:	0.688
Model:	OLS	Adj. R-squared:	0.688
Method:	Least Squares	F-statistic:	4.390e+04
Date:	Sun, 08 May 2022	Prob (F-statistic):	0.00
Time:	15:00:49	Log-Likelihood:	-1.7588e+06
No. Observations:	159141	AIC:	3.518e+06
Df Residuals:	159132	BIC:	3.518e+06
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	3.294e+05	2394.428	137.550	0.000	3.25e+05	3.34e+05
land_sqft	0.6775	0.013	52.223	0.000	0.652	0.703
area_above_ground	26.2615	0.092	286.192	0.000	26.082	26.441
basement_area	-1.8574	0.115	-16.179	0.000	-2.082	-1.632
finished_basement_area	11.1856	0.133	84.067	0.000	10.925	11.446
BED_RMS	-6433.8830	59.443	-108.236	0.000	-6550.390	-6317.376
FULL_B	3285.7108	72.627	45.241	0.000	3143.364	3428.058
HLF_B	3448.5447	93.827	36.754	0.000	3264.646	3632.443
CCYRBLT	-166.9975	1.234	-135.276	0.000	-169.417	-164.578

Omnibus:	173499.450	Durbin-Watson:	0.951
Prob(Omnibus):	0.000	Jarque-Bera (JB):	83800146.692
Skew:	4.918	Prob(JB):	0.00
Kurtosis:	114.987	Cond. No.	4.59e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.59e+05. This might indicate that there are
strong multicollinearity or other numerical problems.

Regression 2

OLS Regression Results

Dep. Variable:	TOTAL_VALUE	R-squared:	0.688
Model:	OLS	Adj. R-squared:	0.688
Method:	Least Squares	F-statistic:	4.390e+04
Date:	Sun, 08 May 2022	Prob (F-statistic):	0.00
Time:	15:01:01	Log-Likelihood:	-2.1831e+06
No. Observations:	159141	AIC:	4.366e+06
Df Residuals:	159132	BIC:	4.366e+06
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	4.739e+06	3.45e+04	137.548	0.000	4.67e+06	4.81e+06
land_sqft	9.7478	0.187	52.220	0.000	9.382	10.114
area_above_ground	377.8656	1.320	286.193	0.000	375.278	380.453
basement_area	-26.7253	1.652	-16.179	0.000	-29.963	-23.488
finished_basement_area	160.9472	1.914	84.068	0.000	157.195	164.700
BED_RMS	-9.257e+04	855.295	-108.237	0.000	-9.43e+04	-9.09e+04
FULL_B	4.727e+04	1044.991	45.239	0.000	4.52e+04	4.93e+04
HLF_B	4.962e+04	1350.027	36.753	0.000	4.7e+04	5.23e+04
CCYRBLT	-2402.8024	17.763	-135.274	0.000	-2437.617	-2367.988

Omnibus:	173499.013	Durbin-Watson:	0.951
Prob(Omnibus):	0.000	Jarque-Bera (JB):	83799306.613
Skew:	4.918	Prob(JB):	0.00
Kurtosis:	114.987	Cond. No.	4.59e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.59e+05. This might indicate that there are
strong multicollinearity or other numerical problems.

Figure 1

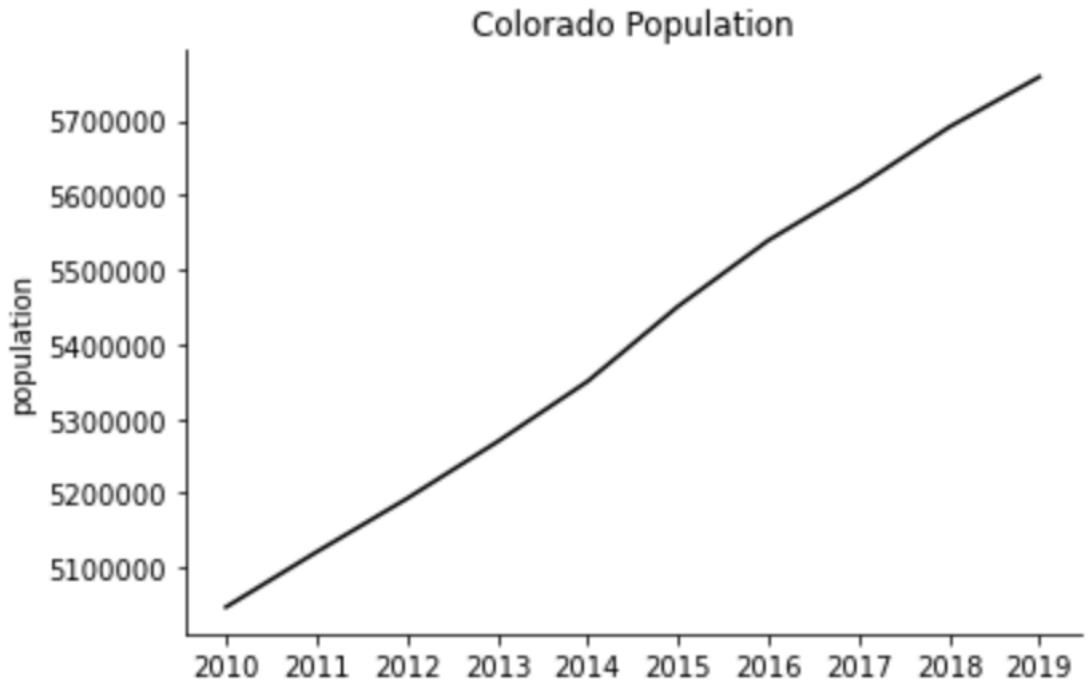

Figure 2