

General Data: Building a Time Series Model - Michigan State

Charles Calloway, Connor Ward, Megan White

Introduction

Founded in 1837, Michigan became the 26th state in the US. Although the size of its land is ranked 22nd of the 50 states, the inclusion of the Great Lakes pushes that ranking to 11th in terms of total area. The state's name is derived from “michi-gama,” a Chippewa word that means “large lake.” The Great Lakes continue to make Michigan unique in that Michigan is the only state that is split into two separate parts.

Michigan's largest city, Detroit, has become well-known throughout the world for hosting many automotive companies (Cultural Life). Michigan has become one of the country's leading tourist destinations due to its borders on four of the five Great Lakes and its many wilderness attractions. Its residents are primarily urban and are concentrated around the industrialized areas. This state prides itself in its diversity and strong education system (Michigan).

Economy

Similar to a national trend, manufacturing in Michigan has weakened recently, but this still continues to be a main contributor to the state. Their leading manufacturers include motor vehicles, fabricated metal products, chemical products, and processed foods. Michigan is the home of companies such as Kellogg Company and The Dow Chemical Company.

Originally, Michigan's economy was dependent on lumbering and mining, but the automotive business has dominated since its introduction to the area in the 1920s. Because Michigan stays loyal to this source of income, their unemployment levels tend to fluctuate with the price of oil. The state suffered great economic loss in the early 2000s when gas prices peaked and the overall US economic growth slowed. Michigan's government has implemented programs in response to these fluctuations in order to promote the service sector of their economy, specifically in their tourism and education (Michigan).

Real Estate Market

Throughout history, Michigan's population has given the state a metropolitan atmosphere. As of 2019, the median home value was \$167,178 with over 3.9 million homes and apartments. For the most part, homes in this state are at least 20 years old or older. Seventy-two percent of these homes are single-family, 13% are apartments, and the remaining 15% consist of townhomes or mobile homes (Michigan Appreciation Rate Trends and Housing Market Data).

More than a decade after the housing crisis in 2008, Michigan is still recovering. The median household income was \$66,000 in 1999, and, at only \$56,000 in 2016, is still working its way back up. Housing ownership also peaked early on with an ownership rate of 77.2% in 2000, but

only an ownership rate of 72.9% in 2017. This is largely due to the unemployment rates in the state. Home sales stabilized in 2019 as more homes entered the market and unemployment fell in line with the national rate.

Independent Variables

Crimes Committed. This variable was selected to determine if criminal acts negatively impact the value of single-family residential houses. The expectation is crimes committed will have a negative correlation to the dependent value as people prefer to live in a safer location. However, crimes committed produced a positive coefficient of 0.0205 which suggests that an increase in criminal activity benefits the value of single-family residential homes. The P-value is 0.04538, which signifies the value is statistically important.

Population. Population was selected to determine if increases in population will have a positive effect on the value of single-family residential houses. The expectation is the population will have a positive correlation to the dependent variable as the demand for houses depends upon the number of available residents. Population produced a positive coefficient of 0.0893 which suggests that an increase in population will increase the value of single-family residential houses. The P-value is 1.43707E-17 indicating that the population in the state does significantly impact value.

Marriage Rate. The marriage rate was expected to have an overall positive correlation to the dependent value. However, some marriages have no effect or a negative effect on the value of single-family residential homes depending on prior housing status for the newlywed couple. Marriage produced a negative coefficient of (-2,315) which suggests that the combination of households does not increase the value of single-family residential houses. The P-value of 0.00199, signifies the value is not as statistically important as other variables such as population, but it has a greater significance than the number of crimes committed.

Tornadoes. The number of tornadoes was selected to determine if natural disasters significantly affected the value houses. The expectation is tornadoes will negatively correlate to the dependent value as natural disasters frequently destroy houses beyond repair. However, tornadoes produced a positive coefficient of 99.49 which indicates that tornadoes increase the value of single-family residential houses. The P-value is 0.2478, which signifies that the value is not statistically important as it is above the 5% threshold. While this is contradictory to our expectations, the significance does make sense considering the moderate climate of the state.

Mortgage Interest Rate. The mortgage interest rate influences the number of mortgages issued to consumers. If mortgage rates decreased, the expectation held that the value of homes would rise due to an increase in demand. A positive coefficient of 274,339 and a P-value of 4.62587E-06 supports the fact that changes in the mortgage interest rate greatly affects the value of single-family home values. This is the model's most impactful independent variable.

Value of Construction. The value of new construction that is put into place should share a positive relationship with the dependent variable, single-family residential values. While the P-value of 1.15134E-34 for this variable indicates significance, the coefficient was expected to be

greater than 0.0755. Overall, the assumption was that the value of construction would add more value to the homes than indicated in the regression.

Death Rate. The expectation for this variable is that as the death rate increases, more homes become available for sale which impacts the values of home. The findings supported the expectation with a coefficient of 3788. The death rate also had a P-value of 2.19932E-15 further emphasizing the significance that death rates have on the value of homes in Michigan.

Unemployment Rate. The unemployment rate has an impact on homebuyers' incomes. As incomes fluctuate, so does the housing market and the demand for homes. This would have a great impact on the value of homes. In fact, the unemployment rate has a negative impact on the value of home as shown with the coefficient of -116,739. While it does have a negative impact, it does not have as much significance as other variables, having a P-value of only 0.00027.

State GDP. Real estate contributes to overall GDP, so if GDP is increasing, real estate values can be expected to do the same. Because real estate impacts GDP, it is expected to add little value and have very little significance. The coefficient of 1.50069E-05 supports this, but the P-value of 3.19402E-08 does not. The state's GDP did not add much value, but it is a significant variable. In fact, the GDP is the variable that ranked second when measuring variables in the regression that have the most significance.

Birth Rate. The birth rate expectation is that as more children are born the need for larger housing will increase and slowly increase the average cost of a single-family home. However, the birth rate produced a negative coefficient of -1843, showing that birth rates do not positively impact the value of single-family residential homes. The P-value is .00040483, which signifies the variable is considered to have significance when determining value.

The Final Model

Michigan's multiple regression model formula is " $Y = -884,943.20 + (274,339.65 * \text{mortgage interest rate}) - (2,315.71 * \text{marriage rate}) + (0.08 * \text{value of construction}) + (3,787.65 * \text{death rate}) + (0.02 * \text{crime rate}) + (0.09 * \text{population}) + (99.49 * \text{tornadoes}) - (116,739.92 * \text{unemployment rate}) + (0.00 * \text{state GDP}) - (1,843.8 * \text{birth rate})$ ".

In order to run the most accurate model, the data had to be adjusted. Years 1996 and 2019 were removed due to missing and/or inaccurate information. A natural log was also run, but it did not provide much variation, so the original regression was used in this analysis. Although the multiple regression model is statistically important and shares a strong positive relationship to the value of single-family homes, the independent variable's signage damages the credibility of the model.

The model's "Multiple R" is 0.9811, which represents a strong positive relationship between the variables. The "Adjusted R Square" is 0.9611, which means that 96% of the dependent variables are explained by the independent variables (Linear Regression Analysis in Excel). The model's "F" value of 652.36 represents the overall significance of the model. The "Significance F" value is 2.3369E-174, which signifies the regression is reliable as it is beneath the 5% cut-off range.

The model's statistically important variables are the intercept, mortgage interest rate, marriage rate, value of construction, death rate, crimes committed, population, unemployment rate, state GDP, and birth rate. However, the independent variables with the most significant coefficients are the intercept, mortgage interest rate, and the unemployment rate with the following values of -884,943.20; 274,339.65; and -116,739.92 respectively.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.98115513							
R Square	0.962665389							
Adjusted R Square	0.961189713							
Standard Error	3699.719151							
Observations	264							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	10	89293894839	8929389484	652.355384	2.3369E-174			
Residual	253	3463044214	13687921.8					
Total	263	92756939053						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-884943.1991	79256.84495	-11.16551132	8.4616E-24	-1041030.425	-728855.973	-1041030.425	-728855.973
Mortgage interest rate	274339.6532	58587.97015	4.682525313	4.6259E-06	158957.3951	389721.9113	158957.3951	389721.9113
Marriage Rate (# married per 1,000 population)	-2315.70547	741.0525607	-3.124886941	0.00198611	-3775.123122	-856.2878176	-3775.123122	-856.2878176
Value of Construction Put in Place	0.075491373	0.00525061	14.3776374	1.1513E-34	0.065150901	0.085831845	0.065150901	0.085831845
Establishment Death Rate	3787.650333	447.8206626	8.457962414	2.1993E-15	2905.719125	4669.581541	2905.719125	4669.581541
Crimes Committed - MI	0.020466247	0.010176852	2.011058765	0.04538001	0.00042411	0.040508384	0.00042411	0.040508384
Population - MI	0.089331792	0.009715806	9.194480973	1.4371E-17	0.070197632	0.108465953	0.070197632	0.108465953
Unemployment rate - MI	99.48619135	85.885636	1.158356577	0.2478108	-69.65567508	268.6280578	-69.65567508	268.6280578
unemployment rate	-116739.924	31629.0194	-3.690911897	0.00027365	-179029.6342	-54450.21382	-179029.6342	-54450.21382
State GDP	1.50069E-05	2.62918E-06	5.707814572	3.194E-08	9.829E-06	2.01847E-05	9.829E-06	2.01847E-05
Establishment Birth rate	-1843.375388	514.233504	-3.584704951	0.00040483	-2856.099036	-830.651739	-2856.099036	-830.651739

Specific Data

Ann Arbor

Ann Arbor's demographic information reveals a population of 119,303 while the median age is 27½. The city's ethnicity consists of White (68.6%), Asian (15.9%), Black (6.89%), Hispanic (4.38%), Two or More Races (3.83%), Other (0.201%), Alaska Native (0.183%), and Native Hawaiian (0.0235%)(Ann Arbor, MI).

The median household income is \$61,247 with a poverty rate of 22.1%, which both are subpar compared to the United States with an income of \$61,937 and poverty rate of 13.1% (Ann Arbor, MI).

The three largest industries in Ann Arbor are Education (32.7%), Health Care (15.3%), and Professional Services (10.4%). Ann Arbor's largest university is the University of Michigan followed by Concordia University. While the most popular majors are General Business Administration & Management, General Computer & Information Sciences, and Experimental Psychology (Ann Arbor, MI).

For Ann Arbor, creating a residential pricing model used the independent variables such as: square footage, number of bathrooms, the presence of a fireplace, the presence of a garage, presence of a basement, the age of house, and number of stories.

After running eight scenarios, the model with the best predictive ability was “ $Y = 284,766.95 + (69.93 \times \text{square footage})$ ”. The independent variables garage, and number of bathrooms were removed due to the coefficient producing the wrong sign. The other independent variables such as fireplace, age of the house, and number of stories were removed because the values were insignificant.

Regression Statistics								
Multiple R	0.618853427							
R Square	0.382979564							
Adjusted R Square	0.376618529							
Standard Error	57997.3176							
Observations	99							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	2.02518E+11	2.02518E+11	60.20711075	8.68098E-12			
Residual	97	3.26278E+11	3363688849					
Total	98	5.28796E+11						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	284766.9485	18150.61598	15.68910657	2.31974E-28	248743.0001	320790.8969	248743.0001	320790.8969
Square Footage	69.92677063	9.01196671	7.759324117	8.68098E-12	52.04051135	87.81302991	52.04051135	87.81302991

The model’s “F” value is 60.21 which represents the overall significance of the model. The next key value is the “Significant F” at 8.68098E-12, which denotes that the regression model is reliable as it falls beneath the 5% cut-off range. Both the intercept and square footage coefficients are statistically important as their P-values fall below the 5% cut-off range and contain the correct sign.

Midland

According to the most recent US census estimates, Midland, Michigan’s population is around 41,800. While Midland may be considered the 29th largest city in the state when measuring population, it is the 5th largest city in terms of land size with a land area of 36.06 square miles.

Midland has been recognized as a great place to raise a family as it has an excellent education system, a safe community, and numerous opportunities to find and keep employment. The median income for households is \$61,076 as of February 2020 which is above the state’s 2018 average of \$54,938 (U.S. Census Bureau QuickFacts).

In order to construct the most accurate residential pricing model, various independent variables were considered: square footage, number of bedrooms, number of bathrooms, the presence of central air, the presence of a fireplace, and the presence of a garage.

Regression Statistics								
Multiple R	0.528508187							
R Square	0.279320904							
Adjusted R Square	0.244589381							
Standard Error	61715.35253							
Observations	88							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	4	1.22525E+11	30631343263	8.042287862	1.57173E-05			
Residual	83	3.16129E+11	3808784738					
Total	87	4.38655E+11						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	124438.5944	43390.59229	2.867870379	0.005237602	38136.45146	210740.7373	38136.45146	210740.7373
Sq. Ft.	61.70196542	15.37083069	4.014224519	0.000130153	31.1300014	92.27392944	31.1300014	92.27392944
Bedrooms	-6848.040632	12302.23962	-0.556649915	0.579264468	-31316.70043	17620.61917	-31316.70043	17620.61917
Bathrooms	4958.082628	9811.444274	0.505336675	0.614661691	-14556.48542	24472.65067	-14556.48542	24472.65067
Central Air	9735.175241	23811.38864	0.408845338	0.683706176	-37624.71908	57095.06956	-37624.71908	57095.06956

After running nine scenarios, the model that provided the best predictive ability excluded data concerning the fireplace and garage as those variables were the least significant. The best model is as follows: $Y = 124,438 + (61.70 \times \text{square footage}) - (6,848.04 \times \# \text{ bedrooms}) + (4,958.08 \times \# \text{ bathrooms}) + (9,735.18 \times \text{central air})$.

The presence/absence of a garage had zero impact on the listing price of home, so it was removed from the final model. The presence/absence of a fireplace was also excluded from the final model as there were not enough observations to provide an accurate representation. When compared to other models, this one provided one of the highest adjusted R square of 24.46%. With only 88 available observations that fit the criteria, the variables in any of the models did not have much significance. In order for the model to provide a more accurate representation of variables that significantly impact the listing price of homes in Midland, Michigan.

Detroit

Detroit has a population of 672,681 with a median age of 35. The ethnicity ratio in the city is Black (77.3%), White (10.8%), Hispanic (7.84%), Other (5.06%). The median household income of Detroit is \$31,283 and it has a poverty rate of 37.9%. This is almost half of the nation's median income of \$61,937 and double the nation's poverty rate of 13.1% (Detroit, MI).

The three largest industries in Detroit are manufacturing (17%), Health Care & Social Assistance (17%), and Accommodation & Food Services (9.39%). The median house value is \$51,600 which is 0.225 times smaller than the nation's average cost of \$229,700. Wayne State University is the largest university in the city with 58.9% of the city awarded degrees (Detroit, MI).

Regression Statistics								
Multiple R	0.60472564							
R Square	0.3656931							
Adjusted R Square	0.32224743							
Standard Error	23381.1636							
Observations	79							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	5	2.3008E+10	4601531878	8.41724939	2.51231E-06			
Residual	73	3.9908E+10	546678810					
Total	78	6.2915E+10						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	77989.1103	14716.0204	5.29960601	1.1831E-06	48660.12387	107318.097	48660.1239	107318.097
SF	24.5521807	7.46491339	3.2890108	0.00154905	9.674629812	39.4297317	9.67462981	39.4297317
# of baths	2485.69728	4156.6563	0.59800405	0.55168803	-5798.506775	10769.9013	-5798.5068	10769.9013
bedrooms	1485.96539	3470.34744	0.42818923	0.66977406	-5430.427078	8402.35787	-5430.4271	8402.35787
Finished basement	5495.60696	3302.13502	1.66425871	0.10034783	-1085.538643	12076.7526	-1085.5386	12076.7526
Fireplace	13081.5645	10877.8491	1.20258742	0.23302218	-8597.957527	34761.0865	-8597.9575	34761.0865

Detroit's multiple regression model formula is "Y = 77,989.11 + (24.55*square footage) + (2,485.70*# baths) + (1,485.97*bedrooms) + (5,495.61*finished basement) + (13,081.57*fireplace)."

The model's "F" value is 8.41 explains the overall significance of the model. The model has a Significance F value of 2.5123E-06. Because it is below the 5% alpha value, the regression model is reliable. The intercept and square footage are the only two coefficients that are significant because they have a P-Value of less than 5% with the correct sign.

Conclusion

Michigan is home to the headquarters of three major automobile companies, making it a hub for manufacturers. This keeps the economy of the state afloat which, in turn, impacts the real estate market. Overall, the most significant variables that impact the values of single-family residential homes in Michigan include the death and marriage rates as well as the population and the state's GDP. The value of homes were not impacted, however, by natural disasters such as tornadoes and the number of crimes committed in Michigan.

Bibliography

“Ann Arbor, MI.” *Data USA*, datausa.io/profile/geo/ann-arbor-mi/.

“Detroit, MI.” *Data USA*, datausa.io/profile/geo/detroit-mi/.

Glazer, Sidney, and Randall J. Schaetzl. “Cultural Life.” *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 17 Jan. 2020, www.britannica.com/place/Michigan/Cultural-life#ref279884/.

“Linear Regression Analysis in Excel.” *Excel Tutorials, Functions and Formulas for Beginners and Advanced Users - Ablebits.com Blog*, 4 Dec. 2019, www.ablebits.com/office-addins-blog/2018/08/01/linear-regression-analysis-excel/.

“Marriages and Marriage Rates Michigan and United States Occurrences.” *Marriages and Marriage Rates*, www.mdch.state.mi.us/pha/osr/marriage/Tab3.1.asp.

“Michigan.” *Britannica*. <https://www.britannica.com/place/Michigan>

“Michigan Appreciation Rate Trends and Housing Market Data.” *NeighborhoodScout*, www.neighborhoodscout.com/mi/real-estate/.

“Michigan Natural Disasters and Weather Extremes.” *USA.com*, www.usa.com/michigan-state-natural-disasters-extremes.htm/.

Pinho, Kirk. “Southeast Michigan Housing Market Will Take a Hit as Economic Fears Grow.” *Crain's Detroit Business*, 15 Mar. 2020, www.craigslist.com/real-estate/southeast-michigan-housing-market-will-take-hit-economic-fears-grow.

“Real Estate’s Impact on the US Economy.” *The Balance*, 10 Feb. 2020. <https://www.thebalance.com/how-does-real-estate-affect-the-u-s-economy-3306018>

“State of Michigan.” *SOM - State of Michigan*, www.michigan.gov/.

“You Are HereMSP Media Center Statistics & Reports.” *MSP - 1997-2006 Crime Statistics*, www.michigan.gov/msp/0,4643,7-123-1586_3501_4621-25744--,00.html.

“U.S. Census Bureau QuickFacts: Midland City, Michigan.” *Census Bureau QuickFacts*, www.census.gov/quickfacts/fact/table/midlandcitymichigan/PST040219.