

# Appendix 3: Correlation Analysis

(See EDS Part II: Correlation)

## Looking at Correlations between Indicators

The pilot EDS model shown in this plan analyzes indicators that relate to physical conditions to demonstrate how statistical tests can be useful in indicator evaluation. The tests show changes in neighborhoods located in clusters 3 and 5 of Detroit. The data used for this testing came from federal Home Mortgage Disclosure Act (HMDA) data <sup>141</sup> from 1997 to 2002, City of Detroit Assessor's data from 2002, city of Detroit crime data for 2002 from Wayne State University CULMA,<sup>142</sup> geographic boundary shape files from the City of Detroit Planning & Development Department, and U.S. Census data from 2000. The numbers from the assessor's data were aggregated to the census tract level.

<sup>141</sup> The HMDA data used in the analysis was provided by the City of Detroit Planning & Development Department. The Home Mortgage Disclosure Act (HMDA) mandates that the Federal Financial Institutions Examination Council (FFIEC) compile aggregate lending data by census tract for each Metropolitan Statistical Area (MSA). The FFIEC organizes the loan information by location, age of housing stock, income level, and racial characteristics. The data are collected to "help the public determine if lending institutions are meeting the housing credit needs of their communities, to help public officials target community development investment, and to help regulators enforce fair lending laws". (Source: Federal Financial Institutions Examination Council. *History of HMDA*. <http://www.ffiec.gov/hmda/history2.htm>. Accessed April 2004. )

<sup>142</sup> Crime data for each census tract was compiled from the links provided on the City of Detroit Planning and Development Web site. The underlying data is provided by the City of Detroit Police Department from Wayne State University's Center for Urban, Labor and Metropolitan Affairs' Center for Urban Studies. (Detroit Planning & Development Department. *Census Tract 2000 Map*. <http://www.ci.detroit.mi.us/plandev/advplanning/cinfo/inter/Census/TractMap.htm>. Accessed April 26, 2004. )

Table A2 and Table A3 show the correlation analysis for this pilot EDS. Here, two sets of data were analyzed, one from 2000 and one from 2002. Table A2 uses 2002 data to show the relationships between specific tax assessor's and HMDA data. The chart using data from 2000 (Table A3) looks at the relationship between indicators from the U.S. Census and HMDA.

In Table A2, the correlation figures that are highlighted in grey show a strong relationship:

- Average true cash value – Mortgage approval rate
- Total delinquent tax – Total number of loan application
- Average delinquent tax – Average true cash value

In Table A3, the correlation table shows that the following sets of indicators are strongly related to each other. Both positive and negative relationships are present between the indicators:

- Median household income, Percent of the population below poverty, and Median amount of home purchase loans – Education and Home ownership
- Percent poverty, and Total number of loan applications – Median household income
- The total number of loan applications – Percent of poverty

The correlation values for these sets of variables are all greater than 0.6 (or in statistical terms  $R > 0.6$ ). The closer an R value is to 1.0, the stronger the relationship between two variables. This result shows that these indicators may be useful in an EDS model

because they have a strong relationship with each other.

## Regression Analysis: Two Variable Relationships

The following indicators provided the basis for the preliminary regression analysis: education, income, poverty, crime rate, and property value. These indicators were selected because of the strong relationships shown in the correlation tables and the assumptions of our model. This model demonstrates how the testing process can work and adopts some of Galster's findings—primarily, that median home purchase loan amount, number of loan applications, and mortgage loan approval rate are relevant leading indicators of Housing Type and Tenure and Status.<sup>143</sup> Despite the lack of a complete dataset for the same years for all indicators, the regression analysis yielded discernible relationships.

Table A4 shows how two of these leading indicators are associated with the key and supporting indicators.

## Property Values and Median Home Purchase Loan Amounts

The following graph (Figure A6) shows the relationship between median home purchase loan amount in 2001 and property value in 2002. The graph shows a positive relationship between property value and median home purchase loan amount. This means that a larger median home purchase loan amount for a neighborhood may indicate higher property values for a given census tract. The actual values, as shown by the dots, are more tightly clustered around the trend line when the median home purchase loan amounts are less than \$60,000. This means that once loan amounts exceed \$60,000, this indicator does not forecast property values as well. The more the points vary around the line drawn on the graph, the less loan amounts explain property values.

Overall, Figure A6 does not indicate that loan amounts predict property values well. Other variables may better explain property value, or loan amount may better predict a different supporting indicator for the key indicator “Status.” Refining the ability of a regression model to forecast key indicators may entail combining indicators, which requires more advanced regression techniques such as multivariate analysis. Multivariate analysis involves analyzing more than two variables and allows the model developer to see the effect of the combination of variables on a desired indicator. This test involves analyzing the relationship between more than just two indicators and developing performance measurements for the model to account for the accuracy of the forecasts. This avenue is worth investigating, especially if

<sup>143</sup> Galster, George, Chris Hayes, and Jennifer Johnson. 2004. *Identifying Robust, Parsimonious Neighborhood Indicators*. Working Paper.

**Table A2**

## Correlations of Indicators for Census Tracts, 2002

	Mortgage Approval Rate	Total # of Loan Application	Median Amount of Home Purchase Loan	Serious Crime Rate	Average True Cash Value	Sum True Cash Value	Average Delinquent Tax	Total Delinquent Tax
Mortgage Approval Rate	1							
Total # of Loan Application	0.256	1.000						
Median Amount of Home Purchase Loan	0.224	0.174	1.000					
Serious Crime Rate	0.445	-0.113	0.102	1.000				
Average True-cash-value	0.757	0.102	0.252	0.524	1.000			
Sum True-cash-value	0.634	0.725	0.327	0.308	0.603	1.000		
Average Delinquent Tax	0.315	0.170	0.272	0.066	0.604	0.286	1.000	
Total Delinquent Tax	0.046	0.712	-0.007	-0.124	-0.026	0.479	0.308	1.000

**Table A3**

## Correlations of Indicators for Census Tracts, 2000

	Vacancy Rate	% of Owner-occupied Houses	Percent Associate Degree or Above	Median Income of Household	Percent Poverty	Approval Rate	Total # of Loan Application	Median Amount of Home Purchase Loans
Vacancy Rate	1							
% of Owner-occupied Houses	-0.665	1.000						
% Associate Degree or Above	-0.332	0.369	1.000					
Median Income of Household	-0.542	0.712	0.744	1.000				
Percent Poverty	0.596	-0.718	-0.647	-0.874	1.000			
Approval Rate	-0.405	0.284	0.276	0.411	-0.511	1.000		
Total # of Loan Application	-0.538	0.701	0.445	0.639	-0.606	0.086	1.000	
Median Amount of Home Purchase Loans	-0.246	0.133	0.789	0.596	-0.554	0.327	0.242	1.000

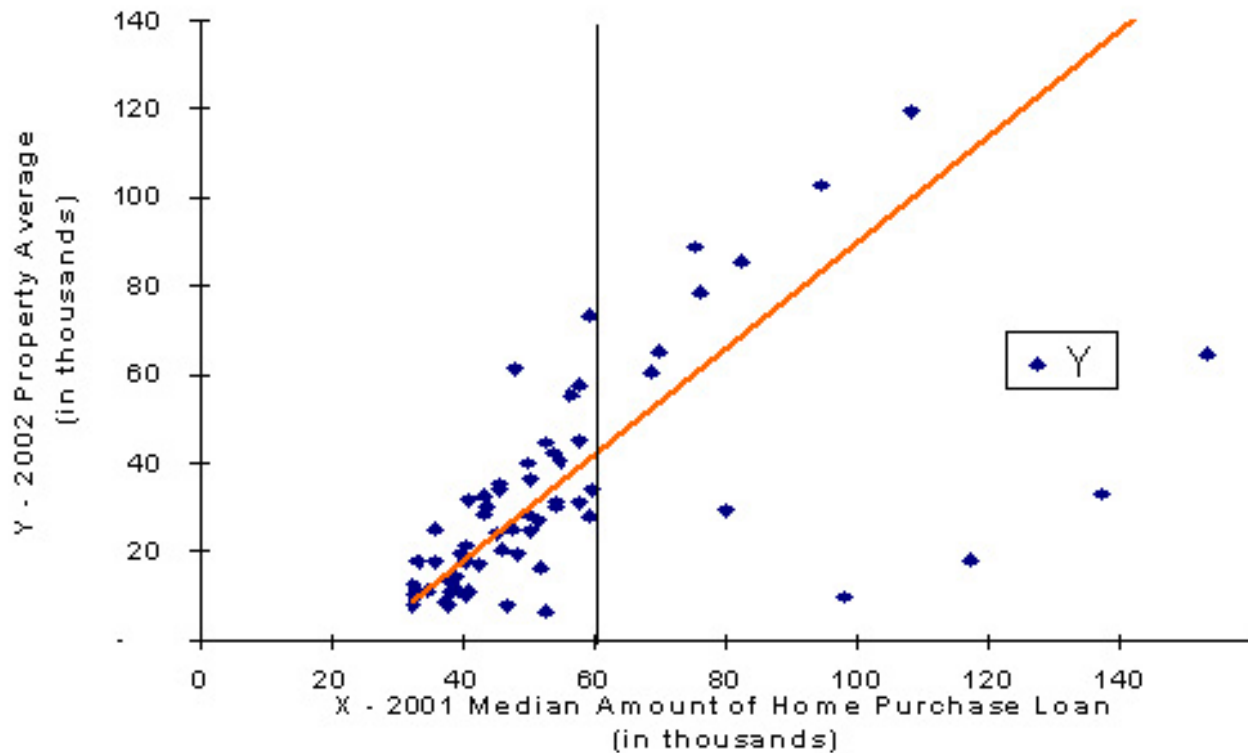
**Table A4**

## Indicators Used in Analysis

Key Indicators	Leading Indicators	Supporting indicators
Housing Type and Tenure	Number of mortgage loan applications	Home ownership rates
Status	Median amount of home purchase loans	Education; property values

**Figure A6**

Relationship between Property Value and Home Purchase Loan Amount



Source: City of Detroit P&DD, HMDA

Figure A6: Graphs the relationship between property value and home purchase loans..

the primary function for an EDS application is its forecasting capabilities.

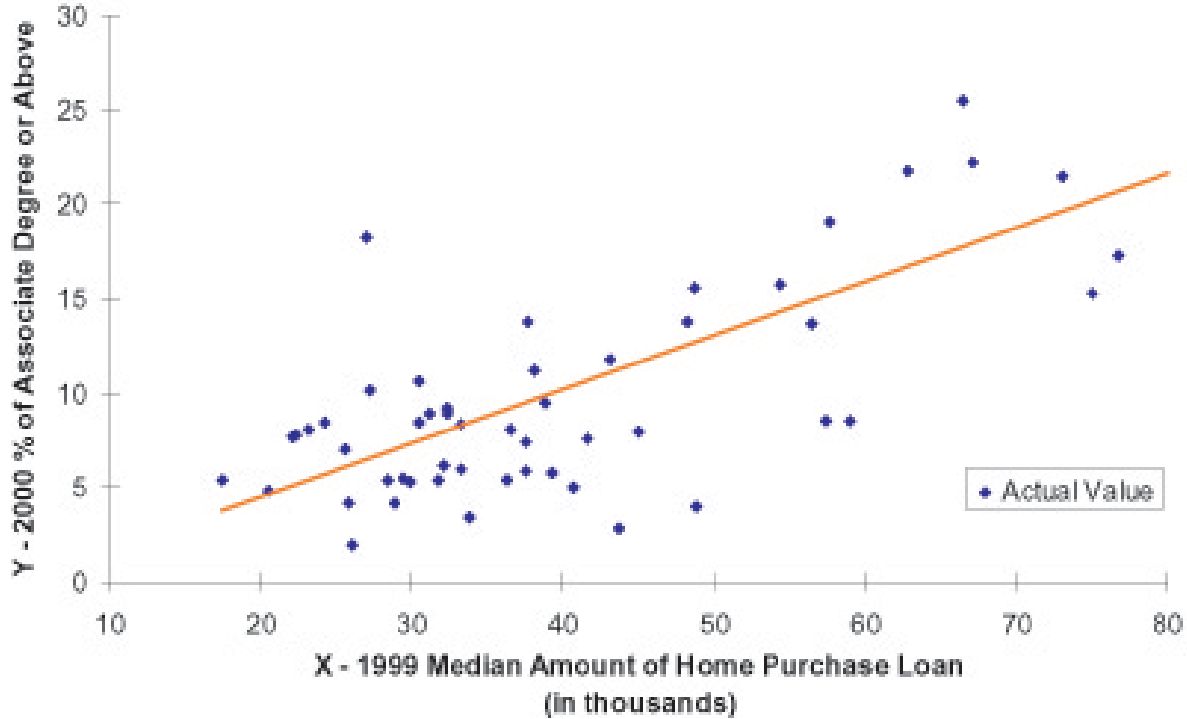
### Education Level and Median Home Purchase Loan Amounts

In Figure A7, the graph illustrates the positive relationship between education level (the percent of the population with an associate's degree or higher) and the median home purchase loan amount. A larger median home purchase loan amount for a neighborhood is likely to mean that the education level for that area is higher. Again, the actual values as shown by the dots are not tightly clustered around the trend line, which is what the predicted value would be. A trend in the relationship is definitely evident, but how well the education level for a neighborhood can be predicted by using the median amount of home purchase loans requires additional analysis.

The regression study illustrated above does not build a predictive model, but does begin to explore the relationship between leading and supporting indicators. Ideally, a predictive model consists of indicators that forecast the desired number accurately. This means

**Figure A7**

Relationship between Level of Educational Attainment and Home Purchase Loan Amount



Source: City of Detroit P&DD , U.S. Census 2000

Figure A7: Shows the strong relationship between education and home loan amounts.

that the model would need to be validated based on the numbers that are forecasted and whether they indeed predict the conditions they have set out to predict (such as home ownership rates or level of education). Models are often adjusted to account for other factors and as more datasets become available, the usefulness of feeding additional information into the model should be analyzed. A predictive model should account for as many values of the predicted value as possible, using other data that are already accessible. The concern is not with the degree of causality between the different indicators, but with how to put together information that will yield a value for an indicator that is unknown.

