Analysis of Residential Home Foreclosure Patterns in Portland

Presented by:
Joe Chan
Tyler Kunter
Margaret Seiler

GIS II
Spring 2009

Background

Research Question

- Is there a geographic concentration of foreclosures in Portland?
- Can foreclosures be correlated with certain socioeconomic factors?

The Foreclosure Problem

- National foreclosure rates have been steadily increasing in 2008 due to poor macroeconomic conditions
- In Oregon, foreclosure rates have been rising but are still below the national average
- However, rising unemployment rates, softening of real-estate markets, and high-cost loans are contributing factors in rising foreclosure rates
Background

Relevance

- Concentrated foreclosures negatively affect the social and economic integrity of a neighborhood and metropolitan area
  - Decrease in property values
  - Loss of tax revenue
  - Vacant property may be associated with increased crime
  - Resident turnover

Risk Factors Identified from Literature Review

- Median income
- Employment status
- Education level
- Percent vacant housing
- Population density
- Percent low-credit score (FICO)
- Percent of high-cost loans
- Percent minority
Background

High Cost Conventional Loans by Race, 2005

Source: Federal Reserve Bank of San Francisco

Data Sources

Datasets

- RLIS Feb 2009
  Mapping boundaries (census, city, and neighborhoods)
- U.S. Census 2000
  Various socioeconomic variables at the census tract and block group levels
- Citydata.com
  Aggregated neighborhood level data
- RealtyTrac.com
  Foreclosure listings in Portland
Data Collection Methods

Methods Overview

I. Data Collection

- Compiled socioeconomic variables from U.S. Census website into a comma separate value (.csv) file
- Compiled foreclosure data from RealtyTrac.com into .csv
- Attribute join census data to mapping units to create socioeconomic dataset
- Geocoded foreclosure data to create a point dataset
- Imported all datasets into a geodatabase for analysis

II. Analysis

- Census tract analysis
  - Ripley’s K
  - Point density
  - Moran’s I
- Rasterize data layers and reclassification
- Neighborhood level analysis
  - Hot-spot analysis
  - Kriging
- Rasterize data layers and reclassification
In addition to the factors driving foreclosures nationally, are there socioeconomic variables that correlate to areas with high foreclosure rates in Portland specifically?

What do current areas of high foreclosures tell us in terms of age, income, place of birth, race and ethnicity, and other factors?

Can these socioeconomic information about these areas help us predict other census tracts that may be at risk for high foreclosure rates?
Raw Counts

Foreclosures per person
Foreclosures per square mile

Foreclosures per housing unit
per owner occupied unit

Ripley’s K Analysis
Point Density (half-mile window)

Reclassify

0-30 foreclosures in half-mile radius = 0
30-50 foreclosures = 1
50-72 foreclosures = 2

Majority filter

Convert raster to features
The per-unit density captures most of the clustering. The highlighted tracts had foreclosure rates of 1.5% or higher.

What do these census tracts have in common?

- Compared high-foreclosure census tracts to the Portland mean or median for more than 60 socioeconomic variables

- For 23 variables, more than 75% of the tracts were all above or all below the mean or median for Portland
Possible Correlation

- Below average educational attainment
- Below average now married
- Below average percent age 45 to 54
- Below average percent white
- Below average percent Korean
- Below average in same house as 5 years before
- Below average in a different county 5 years before
- Below average born in another state
- Below average foreign-born from North America or Latin America
- Below median income

Possible Correlation

- Above median age of home
- Above average Asian
- Above average Vietnamese
- Above average family size
- Above average in a different house in same county 5 years before
- Above average foreign-born
- Above average arrived in U.S. in past decade
- Above average foreign language at home
- Above average Asian language at home and poor English
- Above average Russian ancestry
Convert to raster, reclassify

Add up the 1’s with the raster calculator:

Darkest green: Values for 19 or more out of the 23 variables are above the city mean or median
Part II : Neighborhood Analysis
Neighborhood Level Analysis

• What neighborhood socioeconomic variable may be driving foreclosures?

• Such as the following variables:
  • Age
  • Poverty
  • Job
  • Birthplace (foreign born?)
  • Median Home Value

Process Outline

• Gather all Data (city-data.com)

  VERY TEDIOUS WORK....

• Prep the Neighborhood (RLIS)

  Edit to fit dataset (Delete/Merge)

  Join Foreclosures sites to dataset = COUNT

• Visualize Spatial Distribution

  Foreclosures

  Socioeconomic (S.E.) distributions

• Test Interpretations

  Hot Spot Analysis + Regression Testing

  Transform data if needed
Process Outline

- Rasterize "Contributors"
  Convert + Reclassify
- Create Test Surface
  Raster Calculator
- Create Prediction Surface
  Kriging
- Finally....Analyze Prediction to Site Relationship

Editing the Neighborhood
Finalized Neighborhoods

What’s Driving Foreclosures?

Noticed Visual Correlation to Foreclosures from

- % Below Poverty Level
- % Foreign Born
- Home Value
- Job Type (% job type per Neighborhood)
% Below Poverty Level

- Areas of mid to high poverty show relation to foreclosures sites

% Foreign Born

- Areas of mid to high % of foreign born residents show relation to foreclosures
% Median Home Value

- Areas of lower home values follow trend of foreclosure
- Least Correlation

Job Types Analysis

Created two base job types

- Blue collar:
  Construction, Production, Transportation, and Service Jobs

- White collar:
  Sales/Office, Management, Entertainment, Computer/Math, Education, Architects, Engineers
### Employment Trends by Industry in Oregon

<table>
<thead>
<tr>
<th>Industry</th>
<th>Total Employed (thousands)</th>
<th>1-mo.</th>
<th>3-mo.</th>
<th>12-mo.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1,654.9</td>
<td>14.5</td>
<td>10.2</td>
<td>4.7</td>
</tr>
<tr>
<td>Trade, Transportation &amp; Utilities</td>
<td>317.5</td>
<td>11.7</td>
<td>12.5</td>
<td>6.9</td>
</tr>
<tr>
<td>Government</td>
<td>301.4</td>
<td>3.2</td>
<td>0.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Educational &amp; Health Svcs.</td>
<td>223.6</td>
<td>11.6</td>
<td>0.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Professional &amp; Business Svcs.</td>
<td>185.4</td>
<td>19.6</td>
<td>12.9</td>
<td>6.6</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>175.9</td>
<td>28.1</td>
<td>25.4</td>
<td>12.4</td>
</tr>
<tr>
<td>Leisure &amp; Hospitality</td>
<td>168.2</td>
<td>15.6</td>
<td>9.6</td>
<td>3.6</td>
</tr>
<tr>
<td>Financial Activities</td>
<td>96.7</td>
<td>26.4</td>
<td>11.1</td>
<td>7.0</td>
</tr>
<tr>
<td>Construction</td>
<td>82.5</td>
<td>39.3</td>
<td>24.8</td>
<td>17.3</td>
</tr>
<tr>
<td>Other Services</td>
<td>61.1</td>
<td>-3.8</td>
<td>3.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Information</td>
<td>34.7</td>
<td>-8.8</td>
<td>-15.6</td>
<td>-4.9</td>
</tr>
<tr>
<td>Natural Resources &amp; Mining</td>
<td>81</td>
<td>0.0</td>
<td>-17.5</td>
<td>-8.0</td>
</tr>
</tbody>
</table>

Source: Bureau of Labor Statistics

### Blue Collar

![Blue Collar image](image-url)
Creating Collar Difference

- High Correlation to "Blue" collar neighborhoods
- Lower Correlation to "White" collar neighborhoods
- Highest Correlation
Data Analysis

- Data Testing
  - Regression Testing – Minitab + Excel
  - Multi-linear Regression Testing – Minitab
  - Histograms + Scatterplots - Minitab + Excel

- Transform data from % to # of people
  - (Data x Pop.)

R² = .837
R² = .627
R² = .770
Final Dataset

- **Raster Calculator**
  
  Used formula based on regression results
  
  \[
  \text{Base} = 100\% \\
  (\text{Col Diff} \times 60) + (\text{Foreign} \times 20) + (\text{H Val} \times 10) + (\text{Poverty} \times 10)
  \]
  
  Get "Preference Raster"

- **Kriging**
  
  Ordinary Kriging – Trend adjusted
  
  Create Prediction Map

- **Analyze**
  
  Overlay Foreclosures and Interpret Predictions

---

**Images**

- Poverty Raster
- Foreign Raster
- Foreclosures Raster
- Collar Raster
- Value Raster
Foreclosure “Preference”

Kriging Predictions

Matching Foreclosures to Predictions
Conclusions

Results Discussion

• Based on the foreclosure risk factors identified in various literature there is some correlation found at the Census tract and neighborhood level

  At the census tract level:
  • % minority (though nationally affected groups didn’t stand out)
  • migration factors
  • education

  Based on the neighborhood analysis (Most to Least):
  • Collar Type
  • % Foreign Born
  • % Below Poverty
  • Home Value

Conclusions

Relevance

• High-risk areas could be targeted for aid and outreach

• Correlating variables could raise social justice or neighborhood planning issues

• Developers and investors can think of where to look next for cheap foreclosed properties
Limitations

Data Issues

• Dated socioeconomic variables from the U.S. Census 2000
• Did not have access to key financial variables (credit scores and loan types)
• Information used was about whole block groups, tracts or neighborhoods, not actual households in foreclosure
• Variables singled out for analysis were chosen subjectively
• Temporal resolution: six-month window of auctions only provides a snapshot

References


Questions?