Crime Report Project

Nicholas Perez
Gabriel Bahr
Mission Objective

Predict Property Crime in the City of Austin, Texas

Steps
- Utilized data on years 2010-2016 provided by the City of Austin, Uniform Crime Report and US Census Data
- Analyzed crime data on 4 separate offenses, Robbery, Burglary, Theft, and Motor Vehicle Theft.
- Focused on specific Zip-Codes and Census Tract data that offenses were perpetrated in to predict an outcome.
**Terminology**

**Theft:**
Any time there in an unauthorized taking of property from another with the intent to permanently deprive that person of the property.

**Robbery:**
Essentially theft accomplished through the use of physical force or fear.

**Burglary:**
The unlawful entry into a structure, such as a home or business, with the intent to commit a crime inside.

**Vehicle-Theft:**
The theft or attempted theft of a motor vehicle.
Population in the City of Austin by ZIP

• Degrees of population change from 2010-2015
Predictive Models

Three Main Issues:

- Complexity of Predictive Policing
- Incredible amount of Variables
- Opportunity and Human Behavior

So where did we focus?

- Using the crimes committed in a specific place to pattern corresponding offenses
- Introduce variables which “might” have an impact on the rates of crime.
Change in Crime

- 2016-”2017”
  Ran the decision tree analysis using changes in the average crime rate between all 4 types for each year.
  Provided an Rsquare value of .502, mediocre but with this varied of data we consider that applicable
Robbery

Highest Zips= Cluster 2, 78758, 78753, 78741
Analyzing Census Tract Data

• Why Census Tract?
  • Gives Police Departments and Law Enforcement one level lower to analyze.
  • Ran a HP Forest analysis to allocate better areas of crime.
  • Utilized a 3 cluster target variable.

<table>
<thead>
<tr>
<th>Target</th>
<th>Target Label</th>
<th>Fit Statistics</th>
<th>Statistics Label</th>
<th>Train</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td><em>ASE</em></td>
<td>Average Squ.</td>
<td></td>
<td>0.033817</td>
<td>0.029293</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>DIV</em></td>
<td>Divisor for A.</td>
<td></td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>MAX</em></td>
<td>Maximum A.</td>
<td></td>
<td>0.645567</td>
<td>0.645567</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>NOBS</em></td>
<td>Sum of Fre.</td>
<td></td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>RASE</em></td>
<td>Root Avera.</td>
<td></td>
<td>0.183894</td>
<td>0.171151</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>SSE</em></td>
<td>Sum of Squ.</td>
<td></td>
<td>2.434811</td>
<td>1.054534</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>DISF</em></td>
<td>Frequency ...</td>
<td></td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>MISC</em></td>
<td>Misclassification</td>
<td></td>
<td>0.083333</td>
<td>0.083333</td>
</tr>
<tr>
<td>Cluster</td>
<td><em>WRONG</em></td>
<td>Number of ...</td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Interpreting Census Tract Data

• Allocating the target variables to the 3 clusters allowed us to achieve a reliable dataset.
• We wanted to focus on cluster 2 for this analysis since JMP designated this the smallest cluster. These zip codes included 78741, 78753, 78758. And had the highest frequency in EM.
Findings About Crimes

**City of Austin**
- Our cluster 2 and Zip Code data showed that zip codes in the 78500 area have some of the highest crime rates.
- NE side of the City of Austin according to our data shows the highest level of crime and most likely area where crime will occur in the next year.

**Crimes Over Time**
- Unfortunately, as Austin’s City limits and population increase the opportunity for property crime will also increase.
- We hope that with this data Austin will be able to allocate resources for fighting crime in these susceptible areas.