

State of FL TaxStats

Team Members:

Matthew Sanabria

Kevin Medina

Significance/Usefulness TaxStats

For each city in Florida, this application will provide users with the highest, average, and lowest for each of the following metrics:

1. Property tax
2. Just/market value
3. Land value
4. Taxes as a percent of just value

In addition, the application will provide users with the population for each city and the percent of out of state owners.

There are many uses for this application, from analytics to simple research. A prospective homeowner can use this application to conduct their own research and due diligence when searching for an area to live. The government can use this data for analytical purposes and for building reports. In addition, a real estate agent can use this as a tool for convincing buyers to purchase/not purchase in a particular area based on the statistics provided.

TaxStats will only use a small portion of the data fields from the Department of Revenue's database. However, additional value can be provided by eventually incorporating other data fields in future development.

Schemas

There is no current existing taxonomy that meets the needs of TaxStats. Therefore, we must create a taxonomy (class hierarchy) using protégé to include the following data properties:

1. State
2. City
3. Coordinates (Longitude & Latitude)
4. Total Population
5. Parcel ID
6. Owner State
7. Percent of Out of State Owners
8. Postal Code

9. Property Tax
10. Just/Market Value
11. Land Value
12. Taxes as a Percent of Just Value

Databases

The database used for this application comes directly from the Florida Department of Revenue's Tax Data Portal, which can be seen using the following link:

<ftp://sdrftp03.dor.state.fl.us/Tax%20Roll%20Data%20Files/2014%20Final%20NAL%20-%20SDF%20Files/>

A legend of each data field can be seen using the following link:

ftp://sdrftp03.dor.state.fl.us/Tax%20Roll%20Data%20Files/About2014NAL_SDF_TPP_Files_wo_TOC.pdf

The data is provided in a comma separated value (CSV) format that can be easily incorporated into our database for computational purposes. The tax data provided is recent (2014) and published by the Department of Revenue annually, which makes the application seemingly perpetual to users—the database simply needs to be updated with new tax data as it is published each year.

In addition, the geonames database will be used. More specifically, the data related to the United States. This data can be downloaded as a daily text file from <http://download.geonames.org/export/dump/>. The database provides fundamental information for cities in the United States. The key data used in this project are:

1. City Name
2. Postal Code
3. Latitude
4. Longitude
5. Population

Tools and Languages

1. Protégé for taxonomy
2. Geonames API for city and zip code population data
3. Geonames web client for JAVA
4. Eclipse IDE

Inputs and Outputs

Inputs are tax and geonames databases. The outputs will be the tables and figures that present key statistics for each city. Figure I (below) depicts the output provided by the application for each city:

	City Name		
Population			
% of Out of State Owners			
	Highest	Average	Lowest
Property Tax			
Just/Market Value			
Land Value			
Taxes as % of Just Value			

Figure I: Sample Output Table

Project Development

To develop our project two sets of databases will be used. The first set of data will be accessed from the Florida Department of Revenue's tax database. This will be accessed using file I/O in Java. From this dataset key tax statistics will be extracted out.

The second database used will be the geonames database. This will be done using the geonames web client Java API. The API will be used to extract information (i.e. population from the cities in Florida).

For each city, the population and owner's state will be used to calculate the percent of out of state owners. The application will also extract/calculate the highest, average, and lowest property tax, just market, and land values. In addition, the property tax and just value will be used to calculate the highest, average, and lowest taxes as a percent of just value. Finally, all of this information will then be presented to the user in pop-up window from the Java application (as seen in Figure I).

Future Development

In the future, this application can be further developed to include more granular statistics. For example, data at the postal code level can be provided using the Department of Revenue's database. In addition, using this database we can add features such as distribution of land use; highest, average, and lowest property tax for each type of land use; list of historical properties, etc. Eventually, this data can also be displayed onto an interactive map using Geonames, including longitude, latitude, and elevation data.

Using the ontology/ taxonomy created with protege, coupled with a reasoner, this application can be used to extract similar properties from other states. Our goal is to have this application provide the same statistics nationwide.

