**Details**

Here I estimate a series of models for medium-risk fires, fire size, injuries and deaths.

Predictors used in the model were organized into subgroups to simplify model selection for most models. The groups are summarized below.

**Time**: Year, with 2014 being year zero, was included in the analysis to allow for risk to vary over time.

**Base**: The base variables were population, number of males, and total number of housing units.

**Race**: includes a number of racial and ethnic categories.

**Age**: includes the age distribution of the population.

**House**: Includes a set of housing characteristics, including number of vacant houses, number of renters, number of households with more people than rooms, *number of 1 and two unit residences (“single-family homes”),* number of units that are part of a 10-unit or more block, number of units built before 1980, and number of mobile homes.

**Personal**: A list of personal and household characteristics: median household income, Social Vulnerability Index, *number of married people, number of unemployed, number of those not in the labor force,* and percent of adult smokers (determined at both the state and county level).

***Data***

I defined MED\_RISK fires as discussed in our earlier emails. The data used for this analysis is generated using a series of queries which are included as Appendix I.

As with low risk fires, geocoding is incomplete, and varies by department and year. This model adjusts for geocoding percentage at the department × year level, by including the geocoding percentage of all reported incidents (by department and year) as an offset to the model.

I used the data from 2007 to 2013 to estimate the model. The same filters used for low risk fires were used here. See the low-risk fires report for details.

*base*

I excluded any tract with any of the following characteristics:

* SVI < 0
* No reported median income
* No reported department size
* No reported County smoking data

*small.x*

All models excluded department × years where the department reported responding to fewer than 25 *incidents* that year. I also filtered out “outlier” years. “Outlier” years were defined as those years which fell at least 2 standard deviations below the mean number of incidents for the department. A specific definition is included in Appendix II. Note that the “two standard deviations” standard is based on at most 7 years per department of data reported as part of the study.

*giants*

The definitions of the *giants* filter is unchanged from before. Specific definition is included as part of Appendix II.

*random\_subset*

One third of the tracts are set aside to serve as a test set, while the remaining tracts are used as the training set. The partition for test versus training sets is the same as that used for the low risk analysis.

***Department Size × Region***

As before, I excluded all departments serving fewer than 10,000 people (sizes 0 – 2). All departments nationwide serving 1 million or more people were analyzed together. For some models (more details below) departments in the Northeast in size range 8 were combined with those in the size range 7.

***Models***

Most models were estimated using the techniques of generalized linear models (glm), and had the following basic form:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Where *yij* is the dependent variable being analyzed for the *i*th tract served by the *j*th department in year *t*, *g* is a linking function used in the glm analysis, *aijt* is a *known* offset value for the model, *xijt* are the predictors used for that observation, *βt* is the change per year, *β* are the parameter estimates, *ηj* is the department effect.

It is further assumed that

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

And what is reported in the model is . Specific values for the individual departments can be provided.

For the constant model, the department dummies were not included. For the department effects model, only the department effects (and a constant) were included.

A LASSO model was run for fires, deaths and injuries. The offset term described below was included as part of the LASSO analysis, and all the variables listed above were included. A large number of LASSO models were run, and compared using cross-validation over the training set. Two LASSO models were selected for evaluation against the test set: the model with the cross-validation minimum error (“min”), and the model with the largest cross-validated error within 1 standard error of the minimum (“1se”).

I grew two random forests for fires deaths and injuries. The first forest (rf.0500) consisted of 500 trees, and the second forest (rf.2500) consisted of 2500 trees. Both were run on all the variables listed above. Otherwise both random forests were the same. Neither LASSO nor random forest models included the department dummies to keep the models from being too large. No LASSO or random forest models were run for fire size. More work is needed before I can apply those techniques to the fire size data.

It is assumed that the number of fires, injuries or deaths in a tract follow a (overdispersed) Poisson process. An offset term is used to reflect the proportion of incidents that were geolocated by department. That should convert these estimates, based on geolocated fires, injuries or deaths, into a model of total fires, injuries or deaths.

For fire size, it is assumed that they follow a binomial model. No offset was used for the fire size models because the number of relevant fires was already included as part of the model.

A large number of different models were run. They fall into two main classes. In the first group, I effectively estimated separate models for each region × department size group of departments. This is the same as I did before. In the second group, I just included dummy variables representing department size and region. This tested the effectiveness of a more parsimonious model.

The setup and functions used in computing and analyzing the models are described in Appendix III.

***Results***

All models are estimated against the training set. Then the estimated model is used to predict number of fires (or percent of fires for fire size, or injuries or deaths, depending on the model) for each tract in the test set. Then for each model the Root Mean Square Error (RMSE) of the predicted value is calculated for the test set. Note that all models except the constant, LASSO and random forest models include a department dummy as part of the model estimate. All “Dummies” models, and the LASSO and random forest models also include dummies for department size and region, in addition to the variables listed. There are a few tracts in the test set associated with departments which do not appear in the training set and, therefore, for which no department dummy could be estimated. In those cases, the department dummy for that department was arbitrarily assigned a value of zero.

For comparison, I include three naïve models that estimated number of fires for the test set. The first is a constant model. It assumes that all tracts have the same (average) expected value per year. The second is a pure department-effect model. It assumes that all tracts served by a particular department have the same expected value per year. The third is a tract model that predicts that the value for a tract would be the same as the value for that tract experienced in the previous year. Since my data set did not include data for 2006, no estimate was made for 2007 for the tract model. Again, specific definition of the tract model is included in Appendix III.

In all cases below, I effectively restricted the test set to departments serving 100,000 people or more. This is consistent with the test set used in the low risk models.

Results are reported in Table 2 through Table 6, below.

Comparing the results to the averages per tract in the test set indicates that the models for fires, injuries and deaths are still over-dispersed relative to a Poisson model.

Overfitting of the models is apparent for all models, so care is required in selecting the model to be used.

There is little difference between the random forest models between 500 trees and 2500 trees. So for future work, random forests will be limited to 500 trees.

For fires, the random forest model handily beat all other models except for the TRACT model. That is particularly impressive considering that the random forest models do not include department effects. The TRACT naïve model did very well for fires, but poorly for all the other effects. As with low risk fires, the tract ID contains a lot of information, including department ID and all the census data, but for injuries and deaths contains little information about occurrence.

For deaths, the LASSO model is the best predictor.

I ran cluster analyses on the fires, deaths and injuries models using the AGNES algorithm. The models that are italicized are all in the closest cluster to the actual outcomes. The best way to interpret this is that the italicized models are all similar to the actual outcome, but they are more similar to each other than to the actual outcome.

Table 2: Mean Square Errors of the Fire models tested.

|  |  | **Mean Fires** | **RMSE** | |
| --- | --- | --- | --- | --- |
| **Model Run** | **Predictors** | **Separate** | **Dummies** |
| constant |  | 1.96 | 4.2898 | |
| fx |  | 1.96 | *3.8521* | |
| tract |  | 1.96 | *2.9349* | |
| lasso.min | time, base, race, age, house, personal | 1.96 | *5.4329* | |
| lasso.1se | time, base, race, age, house, personal | 1.96 | *5.1042* | |
| rf.0500 | time, base, race, age, house, personal | 1.96 | *2.9066* | |
| rf.2500 | time, base, race, age, house, personal | 1.96 | *2.9066* | |
| mr.100000 | time | 1.96 | *3.8528* | *3.8550* |
| mr.010000 | base | 1.96 | *6.0926* | *4.0019* |
| mr.110000 | time, base | 1.96 | *6.0855* | *4.0033* |
| mr.001000 | race | 1.96 | 9.1421 | 7.4384 |
| mr.101000 | time, race | 1.96 | 9.1411 | 7.4365 |
| mr.011000 | base, race | 1.96 | 14.0619 | 18.1033 |
| mr.111000 | time, base, race | 1.96 | 14.0684 | 18.0627 |
| mr.000100 | age | 1.96 | *3.5807* | ***3.6102*** |
| mr.100100 | time, age | 1.96 | *3.5794* | *3.6134* |
| mr.010100 | base, age | 1.96 | *4.0325* | 8.3532 |
| mr.110100 | time, base, age | 1.96 | *4.0801* | 8.7797 |
| mr.001100 | race, age | 1.96 | *4.9517* | 6.2661 |
| mr.101100 | time, race, age | 1.96 | *4.9333* | 6.2540 |
| mr.011100 | base, race, age | 1.96 | *4.8144* | 13.8594 |
| mr.111100 | time, base, race, age | 1.96 | *4.7710* | 14.3395 |
| mr.000010 | house | 1.96 | *3.6017* | *4.1742* |
| mr.100010 | time, house | 1.96 | *3.6040* | *4.1763* |
| mr.010010 | base, house | 1.96 | *3.5931* | *4.3357* |
| mr.110010 | time, base, house | 1.96 | *3.5861* | *4.3348* |
| mr.001010 | race, house | 1.96 | *4.5524* | 5.4363 |
| mr.101010 | time, race, house | 1.96 | *4.5539* | 5.4312 |
| mr.011010 | base, race, house | 1.96 | *4.5188* | 5.9286 |
| mr.111010 | time, base, race, house | 1.96 | *4.5158* | 5.9245 |
| mr.000110 | age, house | 1.96 | *3.6150* | *4.1887* |
| mr.100110 | time, age, house | 1.96 | *3.6181* | *4.1992* |
| mr.010110 | base, age, house | 1.96 | *3.6307* | *3.8707* |
| mr.110110 | time, base, age, house | 1.96 | *3.6364* | *3.8759* |
| mr.001110 | race, age, house | 1.96 | *3.7425* | 4.4266 |
| mr.101110 | time, race, age, house | 1.96 | *3.7372* | 4.4276 |
| mr.011110 | base, race, age, house | 1.96 | *4.4387* | 5.2162 |
| mr.111110 | time, base, race, age, house | 1.96 | *4.4181* | 5.1980 |
| mr.000001 | personal | 1.96 | *5.3904* | 4.3304 |
| mr.100001 | time, personal | 1.96 | *5.3992* | 4.3367 |
| mr.010001 | base, personal | 1.96 | *3.9126* | 4.7478 |
| mr.110001 | time, base, personal | 1.96 | *3.9665* | 4.7706 |
| mr.001001 | race, personal | 1.96 | 7.5266 | 5.6002 |
| mr.101001 | time, race, personal | 1.96 | 7.5237 | 5.6153 |
| mr.011001 | base, race, personal | 1.96 | *4.3830* | 6.1931 |
| mr.111001 | time, base, race, personal | 1.96 | *4.4337* | 6.2262 |
| mr.000101 | age, personal | 1.96 | *3.9020* | 5.0191 |
| mr.100101 | time, age, personal | 1.96 | *3.8976* | 5.0339 |
| mr.010101 | base, age, personal | 1.96 | *3.7502* | 4.1935 |
| mr.110101 | time, base, age, personal | 1.96 | *3.7622* | 4.1918 |
| mr.001101 | race, age, personal | 1.96 | *4.2938* | 5.9742 |
| mr.101101 | time, race, age, personal | 1.96 | *4.2904* | 5.9802 |
| mr.011101 | base, race, age, personal | 1.96 | *4.1424* | 5.4903 |
| mr.111101 | time, base, race, age, personal | 1.96 | *4.1526* | 5.4922 |
| mr.000011 | house, personal | 1.96 | ***3.4203*** | 5.0597 |
| mr.100011 | time, house, personal | 1.96 | *3.4223* | 5.0871 |
| mr.010011 | base, house, personal | 1.96 | *3.7852* | 5.4390 |
| mr.110011 | time, base, house, personal | 1.96 | *3.7756* | 5.4675 |
| mr.001011 | race, house, personal | 1.96 | *4.0709* | 6.5501 |
| mr.101011 | time, race, house, personal | 1.96 | *4.0924* | 6.5815 |
| mr.011011 | base, race, house, personal | 1.96 | *4.8339* | 7.7235 |
| mr.111011 | time, base, race, house, personal | 1.96 | *4.8298* | 7.7502 |
| mr.000111 | age, house, personal | 1.96 | *3.6996* | 5.0109 |
| mr.100111 | time, age, house, personal | 1.96 | *3.6901* | 5.0014 |
| mr.010111 | base, age, house, personal | 1.96 | *4.1191* | 5.4157 |
| mr.110111 | time, base, age, house, personal | 1.96 | *4.1015* | 5.3867 |
| mr.001111 | race, age, house, personal | 1.96 | *4.1740* | 5.5691 |
| mr.101111 | time, race, age, house, personal | 1.96 | *4.1772* | 5.5584 |
| mr.011111 | base, race, age, house, personal | 1.96 | *4.9311* | 6.8501 |
| mr.base | time, base, race, age, house, personal | 1.96 | *4.9303* | 6.8467 |

Table 3: Mean Square Errors of the Size-2 models tested.

|  |  | **Mean Size 2 Fires** | **RMSE** | |
| --- | --- | --- | --- | --- |
| **Model Run** | **Predictors** | **Separate** | **Dummies** |
| constant |  | 0.3653 | 0.4275 | |
| fx |  | 0.3653 | 0.4051 | |
| tract |  | 0.3653 | 0.5279 | |
| lasso.min | time, base, race, age, house, personal | 0.3653 |  | |
| lasso.1se | time, base, race, age, house, personal | 0.3653 |  | |
| rf.0500 | time, base, race, age, house, personal | 0.3653 |  | |
| rf.2500 | time, base, race, age, house, personal | 0.3653 |  | |
| mr.100000 | time | 0.3653 | 0.4103 | 0.4100 |
| mr.010000 | base | 0.3653 | 0.4089 | 0.4088 |
| mr.110000 | time, base | 0.3653 | 0.4094 | 0.4090 |
| mr.001000 | race | 0.3653 | 0.4095 | 0.4096 |
| mr.101000 | time, race | 0.3653 | 0.4099 | 0.4098 |
| mr.011000 | base, race | 0.3653 | 0.4084 | 0.4086 |
| mr.111000 | time, base, race | 0.3653 | 0.4089 | 0.4088 |
| mr.000100 | age | 0.3653 | 0.4090 | 0.4086 |
| mr.100100 | time, age | 0.3653 | 0.4093 | 0.4087 |
| mr.010100 | base, age | 0.3653 | 0.4085 | 0.4081 |
| mr.110100 | time, base, age | 0.3653 | 0.4089 | 0.4083 |
| mr.001100 | race, age | 0.3653 | 0.4086 | 0.4084 |
| mr.101100 | time, race, age | 0.3653 | 0.4090 | 0.4086 |
| mr.011100 | base, race, age | 0.3653 | 0.4082 | 0.4080 |
| mr.111100 | time, base, race, age | 0.3653 | 0.4085 | 0.4082 |
| mr.000010 | house | 0.3653 | 0.4066 | 0.4066 |
| mr.100010 | time, house | 0.3653 | 0.4069 | 0.4067 |
| mr.010010 | base, house | 0.3653 | 0.4065 | 0.4065 |
| mr.110010 | time, base, house | 0.3653 | 0.4068 | 0.4067 |
| mr.001010 | race, house | 0.3653 | 0.4064 | 0.4065 |
| mr.101010 | time, race, house | 0.3653 | 0.4067 | 0.4067 |
| mr.011010 | base, race, house | 0.3653 | **0.4064** | 0.4064 |
| mr.111010 | time, base, race, house | 0.3653 | 0.4067 | 0.4066 |
| mr.000110 | age, house | 0.3653 | 0.4072 | 0.4066 |
| mr.100110 | time, age, house | 0.3653 | 0.4075 | 0.4068 |
| mr.010110 | base, age, house | 0.3653 | 0.4070 | 0.4065 |
| mr.110110 | time, base, age, house | 0.3653 | 0.4073 | 0.4066 |
| mr.001110 | race, age, house | 0.3653 | 0.4070 | 0.4066 |
| mr.101110 | time, race, age, house | 0.3653 | 0.4073 | 0.4067 |
| mr.011110 | base, race, age, house | 0.3653 | 0.4068 | **0.4064** |
| mr.111110 | time, base, race, age, house | 0.3653 | 0.4072 | 0.4066 |
| mr.000001 | personal | 0.3653 | 0.4097 | 0.4098 |
| mr.100001 | time, personal | 0.3653 | 0.4102 | 0.4100 |
| mr.010001 | base, personal | 0.3653 | 0.4086 | 0.4084 |
| mr.110001 | time, base, personal | 0.3653 | 0.4090 | 0.4086 |
| mr.001001 | race, personal | 0.3653 | 0.4093 | 0.4096 |
| mr.101001 | time, race, personal | 0.3653 | 0.4097 | 0.4098 |
| mr.011001 | base, race, personal | 0.3653 | 0.4083 | 0.4082 |
| mr.111001 | time, base, race, personal | 0.3653 | 0.4087 | 0.4084 |
| mr.000101 | age, personal | 0.3653 | 0.4090 | 0.4085 |
| mr.100101 | time, age, personal | 0.3653 | 0.4093 | 0.4087 |
| mr.010101 | base, age, personal | 0.3653 | 0.4085 | 0.4081 |
| mr.110101 | time, base, age, personal | 0.3653 | 0.4089 | 0.4083 |
| mr.001101 | race, age, personal | 0.3653 | 0.4087 | 0.4084 |
| mr.101101 | time, race, age, personal | 0.3653 | 0.4091 | 0.4085 |
| mr.011101 | base, race, age, personal | 0.3653 | 0.4084 | 0.4079 |
| mr.111101 | time, base, race, age, personal | 0.3653 | 0.4088 | 0.4081 |
| mr.000011 | house, personal | 0.3653 | 0.4069 | 0.4067 |
| mr.100011 | time, house, personal | 0.3653 | 0.4073 | 0.4069 |
| mr.010011 | base, house, personal | 0.3653 | 0.4068 | 0.4065 |
| mr.110011 | time, base, house, personal | 0.3653 | 0.4071 | 0.4067 |
| mr.001011 | race, house, personal | 0.3653 | 0.4069 | 0.4067 |
| mr.101011 | time, race, house, personal | 0.3653 | 0.4072 | 0.4069 |
| mr.011011 | base, race, house, personal | 0.3653 | 0.4068 | 0.4065 |
| mr.111011 | time, base, race, house, personal | 0.3653 | 0.4072 | 0.4067 |
| mr.000111 | age, house, personal | 0.3653 | 0.4075 | 0.4066 |
| mr.100111 | time, age, house, personal | 0.3653 | 0.4078 | 0.4068 |
| mr.010111 | base, age, house, personal | 0.3653 | 0.4072 | 0.4065 |
| mr.110111 | time, base, age, house, personal | 0.3653 | 0.4076 | 0.4067 |
| mr.001111 | race, age, house, personal | 0.3653 | 0.4074 | 0.4066 |
| mr.101111 | time, race, age, house, personal | 0.3653 | 0.4077 | 0.4068 |
| mr.011111 | base, race, age, house, personal | 0.3653 | 0.4073 | 0.4065 |
| mr.base | time, base, race, age, house, personal | 0.3653 | 0.4076 | 0.4066 |

Table 4: Mean Square Errors of the Size-3 models tested.

|  |  | **Mean Size 3 Fires** | **RMSE** | |
| --- | --- | --- | --- | --- |
| **Model Run** | **Predictors** | **Separate** | **Dummies** |
| constant |  | 0.1229 | 0.3054 | |
| fx |  | 0.1229 | 0.3020 | |
| tract |  | 0.1229 | 0.3821 | |
| lasso.min | time, base, race, age, house, personal | 0.1229 |  | |
| lasso.1se | time, base, race, age, house, personal | 0.1229 |  | |
| rf.0500 | time, base, race, age, house, personal | 0.1229 |  | |
| rf.2500 | time, base, race, age, house, personal | 0.1229 |  | |
| mr.100000 | time | 0.1229 | **0.3043** | 0.3040 |
| mr.010000 | base | 0.1229 | 0.3046 | 0.3037 |
| mr.110000 | time, base | 0.1229 | 0.3046 | 0.3036 |
| mr.001000 | race | 0.1229 | 0.3047 | 0.3041 |
| mr.101000 | time, race | 0.1229 | 0.3048 | 0.3041 |
| mr.011000 | base, race | 0.1229 | 0.3049 | 0.3037 |
| mr.111000 | time, base, race | 0.1229 | 0.3049 | 0.3036 |
| mr.000100 | age | 0.1229 | 0.3060 | 0.3033 |
| mr.100100 | time, age | 0.1229 | 0.3061 | 0.3033 |
| mr.010100 | base, age | 0.1229 | 0.3065 | 0.3032 |
| mr.110100 | time, base, age | 0.1229 | 0.3066 | 0.3031 |
| mr.001100 | race, age | 0.1229 | 0.3059 | 0.3033 |
| mr.101100 | time, race, age | 0.1229 | 0.3061 | 0.3033 |
| mr.011100 | base, race, age | 0.1229 | 0.3068 | 0.3032 |
| mr.111100 | time, base, race, age | 0.1229 | 0.3070 | 0.3031 |
| mr.000010 | house | 0.1229 | 0.3044 | 0.3022 |
| mr.100010 | time, house | 0.1229 | 0.3045 | 0.3021 |
| mr.010010 | base, house | 0.1229 | 0.3052 | 0.3021 |
| mr.110010 | time, base, house | 0.1229 | 0.3053 | 0.3021 |
| mr.001010 | race, house | 0.1229 | 0.3050 | 0.3021 |
| mr.101010 | time, race, house | 0.1229 | 0.3050 | 0.3021 |
| mr.011010 | base, race, house | 0.1229 | 0.3054 | 0.3022 |
| mr.111010 | time, base, race, house | 0.1229 | 0.3055 | 0.3021 |
| mr.000110 | age, house | 0.1229 | 0.3071 | 0.3021 |
| mr.100110 | time, age, house | 0.1229 | 0.3073 | 0.3020 |
| mr.010110 | base, age, house | 0.1229 | 0.3074 | 0.3021 |
| mr.110110 | time, base, age, house | 0.1229 | 0.3077 | 0.3020 |
| mr.001110 | race, age, house | 0.1229 | 0.3071 | 0.3020 |
| mr.101110 | time, race, age, house | 0.1229 | 0.3073 | **0.3020** |
| mr.011110 | base, race, age, house | 0.1229 | 0.3075 | 0.3021 |
| mr.111110 | time, base, race, age, house | 0.1229 | 0.3079 | 0.3020 |
| mr.000001 | personal | 0.1229 | 0.3048 | 0.3037 |
| mr.100001 | time, personal | 0.1229 | 0.3048 | 0.3036 |
| mr.010001 | base, personal | 0.1229 | 0.3054 | 0.3031 |
| mr.110001 | time, base, personal | 0.1229 | 0.3054 | 0.3031 |
| mr.001001 | race, personal | 0.1229 | 0.3051 | 0.3036 |
| mr.101001 | time, race, personal | 0.1229 | 0.3051 | 0.3036 |
| mr.011001 | base, race, personal | 0.1229 | 0.3054 | 0.3031 |
| mr.111001 | time, base, race, personal | 0.1229 | 0.3054 | 0.3030 |
| mr.000101 | age, personal | 0.1229 | 0.3074 | 0.3031 |
| mr.100101 | time, age, personal | 0.1229 | 0.3075 | 0.3031 |
| mr.010101 | base, age, personal | 0.1229 | 0.3079 | 0.3029 |
| mr.110101 | time, base, age, personal | 0.1229 | 0.3081 | 0.3029 |
| mr.001101 | race, age, personal | 0.1229 | 0.3074 | 0.3031 |
| mr.101101 | time, race, age, personal | 0.1229 | 0.3076 | 0.3031 |
| mr.011101 | base, race, age, personal | 0.1229 | 0.3081 | 0.3029 |
| mr.111101 | time, base, race, age, personal | 0.1229 | 0.3084 | 0.3029 |
| mr.000011 | house, personal | 0.1229 | 0.3060 | 0.3023 |
| mr.100011 | time, house, personal | 0.1229 | 0.3061 | 0.3022 |
| mr.010011 | base, house, personal | 0.1229 | 0.3065 | 0.3023 |
| mr.110011 | time, base, house, personal | 0.1229 | 0.3066 | 0.3022 |
| mr.001011 | race, house, personal | 0.1229 | 0.3063 | 0.3023 |
| mr.101011 | time, race, house, personal | 0.1229 | 0.3063 | 0.3022 |
| mr.011011 | base, race, house, personal | 0.1229 | 0.3068 | 0.3022 |
| mr.111011 | time, base, race, house, personal | 0.1229 | 0.3069 | 0.3022 |
| mr.000111 | age, house, personal | 0.1229 | 0.3088 | 0.3021 |
| mr.100111 | time, age, house, personal | 0.1229 | 0.3090 | 0.3021 |
| mr.010111 | base, age, house, personal | 0.1229 | 0.3091 | 0.3021 |
| mr.110111 | time, base, age, house, personal | 0.1229 | 0.3094 | 0.3021 |
| mr.001111 | race, age, house, personal | 0.1229 | 0.3092 | 0.3021 |
| mr.101111 | time, race, age, house, personal | 0.1229 | 0.3094 | 0.3021 |
| mr.011111 | base, race, age, house, personal | 0.1229 | 0.3095 | 0.3021 |
| mr.base | time, base, race, age, house, personal | 0.1229 | 0.3098 | 0.3021 |

Table 5: Mean Square Errors of the injury models tested.

|  |  | **Mean Injuries** | **RMSE** | |
| --- | --- | --- | --- | --- |
| **Model Run** | **Predictors** | **Separate** | **Dummies** |
| constant |  | 0.0628 | *0.3786* | |
| fx |  | 0.0628 | *0.3777* | |
| tract |  | 0.0628 | *0.5490* | |
| lasso.min | time, base, race, age, house, personal | 0.0628 | *0.3830* | |
| lasso.1se | time, base, race, age, house, personal | 0.0628 | *0.3791* | |
| rf.0500 | time, base, race, age, house, personal | 0.0628 | *0.3816* | |
| rf.2500 | time, base, race, age, house, personal | 0.0628 | *0.3812* | |
| mr.100000 | time | 0.0628 | *0.3780* | *0.3776* |
| mr.010000 | base | 0.0628 | *0.3814* | *0.3784* |
| mr.110000 | time, base | 0.0628 | *0.3816* | *0.3784* |
| mr.001000 | race | 0.0628 | *0.3874* | *0.3925* |
| mr.101000 | time, race | 0.0628 | *0.3875* | *0.3925* |
| mr.011000 | base, race | 0.0628 | *0.3962* | *0.4781* |
| mr.111000 | time, base, race | 0.0628 | *0.3953* | *0.4781* |
| mr.000100 | age | 0.0628 | *0.3773* | *0.3779* |
| mr.100100 | time, age | 0.0628 | *0.3775* | *0.3779* |
| mr.010100 | base, age | 0.0628 | *0.3813* | *0.3918* |
| mr.110100 | time, base, age | 0.0628 | *0.3814* | *0.3928* |
| mr.001100 | race, age | 0.0628 | *0.3769* | *0.3891* |
| mr.101100 | time, race, age | 0.0628 | *0.3768* | *0.3894* |
| mr.011100 | base, race, age | 0.0628 | *0.3814* | *0.4078* |
| mr.111100 | time, base, race, age | 0.0628 | *0.3811* | *0.4085* |
| mr.000010 | house | 0.0628 | 0.5522 | *0.3788* |
| mr.100010 | time, house | 0.0628 | 0.5541 | *0.3787* |
| mr.010010 | base, house | 0.0628 | 0.5733 | *0.3788* |
| mr.110010 | time, base, house | 0.0628 | 0.5775 | *0.3787* |
| mr.001010 | race, house | 0.0628 | 1.1627 | *0.3799* |
| mr.101010 | time, race, house | 0.0628 | 1.1654 | *0.3797* |
| mr.011010 | base, race, house | 0.0628 | 0.9225 | *0.3815* |
| mr.111010 | time, base, race, house | 0.0628 | 0.9254 | *0.3814* |
| mr.000110 | age, house | 0.0628 | 0.7159 | *0.3795* |
| mr.100110 | time, age, house | 0.0628 | 0.7155 | *0.3793* |
| mr.010110 | base, age, house | 0.0628 | 0.7082 | *0.3785* |
| mr.110110 | time, base, age, house | 0.0628 | 0.7082 | *0.3785* |
| mr.001110 | race, age, house | 0.0628 | 0.7995 | *0.3820* |
| mr.101110 | time, race, age, house | 0.0628 | 0.8011 | *0.3818* |
| mr.011110 | base, race, age, house | 0.0628 | 0.8089 | *0.3830* |
| mr.111110 | time, base, race, age, house | 0.0628 | 0.8087 | *0.3830* |
| mr.000001 | personal | 0.0628 | *0.3784* | ***0.3761*** |
| mr.100001 | time, personal | 0.0628 | *0.3790* | *0.3761* |
| mr.010001 | base, personal | 0.0628 | *0.3846* | *0.3778* |
| mr.110001 | time, base, personal | 0.0628 | *0.3847* | *0.3778* |
| mr.001001 | race, personal | 0.0628 | *0.3815* | *0.3788* |
| mr.101001 | time, race, personal | 0.0628 | *0.3823* | *0.3788* |
| mr.011001 | base, race, personal | 0.0628 | *0.3970* | *0.3825* |
| mr.111001 | time, base, race, personal | 0.0628 | *0.3982* | *0.3825* |
| mr.000101 | age, personal | 0.0628 | ***0.3754*** | *0.3820* |
| mr.100101 | time, age, personal | 0.0628 | *0.3755* | *0.3823* |
| mr.010101 | base, age, personal | 0.0628 | *0.3805* | *0.3775* |
| mr.110101 | time, base, age, personal | 0.0628 | *0.3806* | *0.3775* |
| mr.001101 | race, age, personal | 0.0628 | *0.3757* | *0.3824* |
| mr.101101 | time, race, age, personal | 0.0628 | *0.3758* | *0.3826* |
| mr.011101 | base, race, age, personal | 0.0628 | *0.3857* | *0.3800* |
| mr.111101 | time, base, race, age, personal | 0.0628 | *0.3861* | *0.3800* |
| mr.000011 | house, personal | 0.0628 | 0.9446 | *0.3786* |
| mr.100011 | time, house, personal | 0.0628 | 0.9890 | *0.3786* |
| mr.010011 | base, house, personal | 0.0628 | 0.8528 | *0.3793* |
| mr.110011 | time, base, house, personal | 0.0628 | 0.9016 | *0.3793* |
| mr.001011 | race, house, personal | 0.0628 | 2.6674 | *0.3825* |
| mr.101011 | time, race, house, personal | 0.0628 | 2.8300 | *0.3825* |
| mr.011011 | base, race, house, personal | 0.0628 | 2.1459 | *0.3850* |
| mr.111011 | time, base, race, house, personal | 0.0628 | 2.3017 | *0.3851* |
| mr.000111 | age, house, personal | 0.0628 | 0.9614 | *0.3847* |
| mr.100111 | time, age, house, personal | 0.0628 | 0.9599 | *0.3847* |
| mr.010111 | base, age, house, personal | 0.0628 | 1.1853 | *0.3848* |
| mr.110111 | time, base, age, house, personal | 0.0628 | 1.1909 | *0.3848* |
| mr.001111 | race, age, house, personal | 0.0628 | 1.3700 | *0.3851* |
| mr.101111 | time, race, age, house, personal | 0.0628 | 1.3842 | *0.3851* |
| mr.011111 | base, race, age, house, personal | 0.0628 | 1.5971 | *0.3868* |
| mr.base | time, base, race, age, house, personal | 0.0628 | 1.6196 | *0.3868* |

Table 6: Mean Square Errors of the fire-death models tested.

|  |  | **Mean Deaths** | **RMSE** | |
| --- | --- | --- | --- | --- |
| **Model Run** | **Predictors** | **Separate** | **Dummies** |
| constant |  | 0.0044 | *0.0831* | |
| fx |  | 0.0044 | *0.0832* | |
| tract |  | 0.0044 | *0.1176* | |
| lasso.min | time, base, race, age, house, personal | 0.0044 | *0.0831* | |
| lasso.1se | time, base, race, age, house, personal | 0.0044 | *0.0831* | |
| rf.0500 | time, base, race, age, house, personal | 0.0044 | *0.0855* | |
| rf.2500 | time, base, race, age, house, personal | 0.0044 | *0.0852* | |
| mr.100000 | time | 0.0044 | ***0.0832*** | *0.0832* |
| mr.010000 | base | 0.0044 | *0.0832* | *0.0832* |
| mr.110000 | time, base | 0.0044 | *0.0833* | *0.0832* |
| mr.001000 | race | 0.0044 | *0.0837* | *0.0832* |
| mr.101000 | time, race | 0.0044 | *0.0838* | *0.0832* |
| mr.011000 | base, race | 0.0044 | *0.0842* | *0.0838* |
| mr.111000 | time, base, race | 0.0044 | *0.0844* | *0.0838* |
| mr.000100 | age | 0.0044 | *0.0873* | *0.0832* |
| mr.100100 | time, age | 0.0044 | *0.0971* | *0.0832* |
| mr.010100 | base, age | 0.0044 | *0.0846* | *0.0835* |
| mr.110100 | time, base, age | 0.0044 | *0.0845* | *0.0835* |
| mr.001100 | race, age | 0.0044 | *0.1149* | *0.0833* |
| mr.101100 | time, race, age | 0.0044 | *0.1092* | *0.0833* |
| mr.011100 | base, race, age | 0.0044 | *0.0862* | *0.0838* |
| mr.111100 | time, base, race, age | 0.0044 | *0.0860* | *0.0838* |
| mr.000010 | house | 0.0044 | *0.0834* | *0.0832* |
| mr.100010 | time, house | 0.0044 | *0.0835* | *0.0832* |
| mr.010010 | base, house | 0.0044 | *0.0834* | *0.0832* |
| mr.110010 | time, base, house | 0.0044 | *0.0836* | *0.0832* |
| mr.001010 | race, house | 0.0044 | *0.0836* | *0.0832* |
| mr.101010 | time, race, house | 0.0044 | *0.0843* | *0.0832* |
| mr.011010 | base, race, house | 0.0044 | *0.0838* | *0.0832* |
| mr.111010 | time, base, race, house | 0.0044 | *0.0842* | *0.0832* |
| mr.000110 | age, house | 0.0044 | *0.0898* | *0.0832* |
| mr.100110 | time, age, house | 0.0044 | *0.1267* | *0.0832* |
| mr.010110 | base, age, house | 0.0044 | *0.0876* | *0.0832* |
| mr.110110 | time, base, age, house | 0.0044 | *0.0938* | *0.0832* |
| mr.001110 | race, age, house | 0.0044 | *0.0857* | *0.0832* |
| mr.101110 | time, race, age, house | 0.0044 | *0.0857* | *0.0832* |
| mr.011110 | base, race, age, house | 0.0044 | 0.4781 | *0.0832* |
| mr.111110 | time, base, race, age, house | 0.0044 | 0.1750 | *0.0832* |
| mr.000001 | personal | 0.0044 | *0.0833* | *0.0832* |
| mr.100001 | time, personal | 0.0044 | *0.0833* | *0.0832* |
| mr.010001 | base, personal | 0.0044 | *0.0833* | *0.0833* |
| mr.110001 | time, base, personal | 0.0044 | *0.0833* | *0.0833* |
| mr.001001 | race, personal | 0.0044 | *0.0835* | *0.0832* |
| mr.101001 | time, race, personal | 0.0044 | *0.0837* | *0.0832* |
| mr.011001 | base, race, personal | 0.0044 | *0.0835* | *0.0834* |
| mr.111001 | time, base, race, personal | 0.0044 | *0.0838* | *0.0834* |
| mr.000101 | age, personal | 0.0044 | *0.0930* | *0.0833* |
| mr.100101 | time, age, personal | 0.0044 | *0.1030* | *0.0833* |
| mr.010101 | base, age, personal | 0.0044 | *0.0855* | *0.0833* |
| mr.110101 | time, base, age, personal | 0.0044 | *0.0867* | *0.0833* |
| mr.001101 | race, age, personal | 0.0044 | *0.1126* | *0.0833* |
| mr.101101 | time, race, age, personal | 0.0044 | 0.1883 | *0.0833* |
| mr.011101 | base, race, age, personal | 0.0044 | *0.0911* | *0.0833* |
| mr.111101 | time, base, race, age, personal | 0.0044 | *0.1078* | *0.0833* |
| mr.000011 | house, personal | 0.0044 | *0.0837* | ***0.0832*** |
| mr.100011 | time, house, personal | 0.0044 | *0.0840* | *0.0832* |
| mr.010011 | base, house, personal | 0.0044 | *0.0837* | *0.0832* |
| mr.110011 | time, base, house, personal | 0.0044 | *0.0845* | *0.0832* |
| mr.001011 | race, house, personal | 0.0044 | *0.0838* | *0.0832* |
| mr.101011 | time, race, house, personal | 0.0044 | *0.0844* | *0.0832* |
| mr.011011 | base, race, house, personal | 0.0044 | *0.0944* | *0.0832* |
| mr.111011 | time, base, race, house, personal | 0.0044 | 307.3813 | *0.0832* |
| mr.000111 | age, house, personal | 0.0044 | 0.4783 | *0.0832* |
| mr.100111 | time, age, house, personal | 0.0044 | 1.85E+25 | *0.0832* |
| mr.010111 | base, age, house, personal | 0.0044 | *0.0940* | *0.0833* |
| mr.110111 | time, base, age, house, personal | 0.0044 | *0.1398* | *0.0833* |
| mr.001111 | race, age, house, personal | 0.0044 | 0.0959 | *0.0832* |
| mr.101111 | time, race, age, house, personal | 0.0044 | 0.7341 | *0.0832* |
| mr.011111 | base, race, age, house, personal | 0.0044 | *0.0932* | *0.0832* |
| mr.base | time, base, race, age, house, personal | 0.0044 | *0.1434* | *0.0833* |

**Appendix I**

This query creates a “parcel\_risk” intermediate table, and is aggregated from the CoreLogic data. It counts the number of parcels (fields ending in “\_n”) and units (fields ending in “\_u”) for various categories of land uses by census tract. In particular, it counts single-family residences (“sfr\_\_”), apartment complexes (“apts\_\_”), and other medium risk structures (“mr\_\_”).

SELECT p.state\_code,

p.cnty\_code,

"substring"(p.census\_tr, 1, 6) AS tract,

sum(

CASE

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'Yes' AND (p.risk\_category <> 'high' OR p.risk\_category IS NULL) THEN 1

ELSE 0

END) AS apts\_n,

sum(

CASE

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'Yes' AND

(p.risk\_category <> 'high' OR p.risk\_category IS NULL) AND p.bld\_units IS NULL THEN 1::double precision

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'Yes' AND

(p.risk\_category <> 'high' OR p.risk\_category IS NULL) AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS apts\_u,

sum(

CASE

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'No' AND (p.risk\_category <> 'high' OR p.risk\_category IS NULL) THEN 1

ELSE 0

END) AS mr\_n,

sum(

CASE

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'No' AND

(p.risk\_category <> 'high' OR p.risk\_category IS NULL) AND p.bld\_units IS NULL THEN 1::double precision

WHEN luse.risk\_category = 'Medium' AND luse.residential = 'No' AND

(p.risk\_category <> 'high' OR p.risk\_category IS NULL) AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS mr\_u,

sum(

CASE

WHEN luse.risk\_category = 'Low' AND luse.residential = 'Yes' THEN 1

ELSE 0

END) AS sfr\_n,

sum(

CASE

WHEN luse.risk\_category = 'Low' AND luse.residential = 'Yes' AND p.bld\_units IS NULL THEN 1::double precision

WHEN luse.risk\_category = 'Low' AND luse.residential = 'Yes' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS sfr\_u,

sum(

CASE

WHEN p.land\_use LIKE '2%' THEN 1

ELSE 0

END) AS com\_n,

sum(

CASE

WHEN p.land\_use LIKE '2%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '2%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS com\_u,

sum(

CASE

WHEN p.land\_use LIKE '3%' THEN 1

ELSE 0

END) AS ind\_n,

sum(

CASE

WHEN p.land\_use LIKE '3%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '3%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS ind\_u,

sum(

CASE

WHEN p.land\_use LIKE '4%' THEN 1

ELSE 0

END) AS vacant\_n,

sum(

CASE

WHEN p.land\_use LIKE '4%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '4%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS vacant\_u,

sum(

CASE

WHEN p.land\_use LIKE '5%' THEN 1

ELSE 0

END) AS agr\_n,

sum(

CASE

WHEN p.land\_use LIKE '5%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '5%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS agr\_u,

sum(

CASE

WHEN p.land\_use LIKE '6%' THEN 1

ELSE 0

END) AS gov\_n,

sum(

CASE

WHEN p.land\_use LIKE '6%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '6%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS gov\_u,

sum(

CASE

WHEN p.land\_use LIKE '7%' THEN 1

ELSE 0

END) AS rec\_n,

sum(

CASE

WHEN p.land\_use LIKE '7%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '7%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS rec\_u,

sum(

CASE

WHEN p.land\_use LIKE '8%' THEN 1

ELSE 0

END) AS trans\_n,

sum(

CASE

WHEN p.land\_use LIKE '8%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '8%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS trans\_u,

sum(

CASE

WHEN p.land\_use LIKE '9%' THEN 1

ELSE 0

END) AS oth\_n,

sum(

CASE

WHEN p.land\_use LIKE '9%' AND p.bld\_units IS NULL THEN 1::double precision

WHEN p.land\_use LIKE '9%' AND p.bld\_units IS NOT NULL THEN p.bld\_units

ELSE 0::double precision

END) AS oth\_u

FROM parcels p LEFT JOIN "LUSE\_swg" luse ON p.land\_use = luse."Code"

GROUP BY p.state\_code, p.cnty\_code, "substring"(p.census\_tr, 1, 6)

This query constructs the medium risk data table I use from the constituent queries. All other queries are as defined earlier.

WITH f AS (

SELECT cf.year,

cf.geoid,

count(\*) AS tot\_fires,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL THEN 1

ELSE 0

END) AS med\_risk,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL AND cf.fire\_sprd IS NOT NULL THEN 1

ELSE 0

END) AS mr\_1,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL AND cf.fire\_sprd IN ( '3', '4', '5' ) THEN 1

ELSE 0

END) AS mr\_2,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL AND cf.fire\_sprd = '5' THEN 1

ELSE 0

END) AS mr\_3,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL THEN cf.ff\_inj + cf.oth\_inj

ELSE 0

END) AS injuries,

sum(

CASE

WHEN cf.struc = 'Y' AND cf.risk = 'Med Risk' AND cf.geoid IS NOT NULL THEN cf.ff\_death + cf.oth\_death

ELSE 0

END) AS deaths

FROM coded\_fires cf

WHERE cf.year > 2006 AND cf.year < 2014 AND cf.version = 5.0 AND NOT (cf.inc\_type = '112' AND cf.year > 2007)

GROUP BY cf.year, cf.geoid

), d AS (

SELECT i.year,

d\_1.firecares\_id AS fd\_id,

d\_1.fd\_size,

sum(i.incidents) AS incidents,

sum(i.incidents\_loc) AS located,

sum(i.fires) AS dept\_fires

FROM dept\_incidents i

JOIN bg.gov\_units d\_1 ON i.state = d\_1.state AND i.fdid = d\_1.fdid

WHERE i.year > 2006 AND i.year < 2014

GROUP BY i.year, d\_1.firecares\_id, d\_1.fd\_size

), c AS (

SELECT casualties\_fire.geoid,

casualties\_fire.year,

sum(

CASE

WHEN casualties\_fire.sev <> '5' AND (casualties\_fire.type = 'ff' OR casualties\_fire.aid\_flag = 'N') THEN 1

ELSE 0

END) AS injuries,

sum(

CASE

WHEN casualties\_fire.sev = '5' AND (casualties\_fire.type = 'ff' OR casualties\_fire.aid\_flag = 'N') THEN 1

ELSE 0

END) AS deaths

FROM casualties\_fire

WHERE casualties\_fire.risk = 'Med Risk'

GROUP BY casualties\_fire.geoid, casualties\_fire.year

)

SELECT tr.year,

tr.geoid,

tr.region,

tr.state,

tr.fc\_dept\_id AS fd\_id,

d.fd\_size,

d.incidents AS dept\_incidents,

d.dept\_fires,

CASE

WHEN d.incidents > 0 THEN d.located / d.incidents

WHEN d.incidents = 0 AND d.located > 0 THEN 'Infinity'

ELSE 'NaN'

END AS f\_located,

f.med\_risk,

f.mr\_1,

f.mr\_2,

f.mr\_3,

c.injuries,

c.deaths,

CASE

WHEN acs."B25002\_002E" > 0 THEN acs."B01001\_001E" / acs."B25002\_002E"

WHEN acs."B25002\_002E" = 0 AND acs."B01001\_001E" > 0 THEN 'Infinity'

ELSE 'NaN'

END AS ave\_hh\_sz,

acs."B01001\_001E" AS pop,

acs."B02001\_003E" AS black,

acs."B02001\_004E" AS amer\_es,

acs."B02001\_005E" + acs."B02001\_006E" + acs."B02001\_007E" + acs."B02001\_008E" AS other,

acs."B03003\_003E" AS hispanic,

acs."B01001\_002E" AS males,

acs."B01001\_003E" + acs."B01001\_027E" AS age\_under5,

acs."B01001\_004E" + acs."B01001\_028E" AS age\_5\_9,

acs."B01001\_005E" + acs."B01001\_029E" AS age\_10\_14,

acs."B01001\_006E" + acs."B01001\_007E" + acs."B01001\_030E" + acs."B01001\_031E" AS age\_15\_19,

acs."B01001\_008E" + acs."B01001\_009E" + acs."B01001\_010E" + acs."B01001\_032E" + acs."B01001\_033E" + acs."B01001\_034E" AS age\_20\_24,

acs."B01001\_011E" + acs."B01001\_012E" + acs."B01001\_035E" + acs."B01001\_036E" AS age\_25\_34,

acs."B01001\_013E" + acs."B01001\_014E" + acs."B01001\_037E" + acs."B01001\_038E" AS age\_35\_44,

acs."B01001\_015E" + acs."B01001\_016E" + acs."B01001\_039E" + acs."B01001\_040E" AS age\_45\_54,

acs."B01001\_017E" + acs."B01001\_018E" + acs."B01001\_019E" + acs."B01001\_041E" + acs."B01001\_042E" + acs."B01001\_043E" AS age\_55\_64,

acs."B01001\_020E" + acs."B01001\_021E" + acs."B01001\_022E" + acs."B01001\_044E" + acs."B01001\_045E" + acs."B01001\_046E" AS age\_65\_74,

acs."B01001\_023E" + acs."B01001\_024E" + acs."B01001\_047E" + acs."B01001\_048E" AS age\_75\_84,

acs."B01001\_025E" + acs."B01001\_049E" AS age\_85\_up,

acs."B25002\_001E" AS hse\_units,

acs."B25002\_003E" AS vacant,

acs."B25014\_008E" AS renter\_occ,

acs."B25014\_005E" + acs."B25014\_006E" + acs."B25014\_007E" + acs."B25014\_011E" + acs."B25014\_012E" + acs."B25014\_013E" AS crowded,

acs."B25024\_002E" + acs."B25024\_003E" + acs."B25024\_004E" AS sfr,

acs."B25024\_007E" + acs."B25024\_008E" + acs."B25024\_009E" AS units\_10,

acs."B25024\_010E" AS mh,

acs."B25034\_006E" + acs."B25034\_007E" + acs."B25034\_008E" + acs."B25034\_009E" + acs."B25034\_010E" AS older,

pcl.apts\_n AS apt\_parcels,

pcl.mr\_n AS mr\_parcels,

acs."B19013\_001E" AS inc\_hh,

svi.r\_pl\_themes AS svi,

acs."B12001\_001" - (acs."B12001\_003" + acs."B12001\_012") AS married,

acs."B23025\_005" AS unemployed,

acs."B12001\_007" AS nilf,

sm.adult\_smoke AS smoke\_st,

sc.smoking\_pct AS smoke\_cty

FROM bg.tract\_years tr

LEFT JOIN f ON tr.geoid = f.geoid AND tr.year = f.year

LEFT JOIN d ON tr.fc\_dept\_id = d.fd\_id AND tr.year = d.year

LEFT JOIN bg.svi2010 svi ON tr.geoid = ('14000US' || svi.fips)

LEFT JOIN bg.acs\_est acs ON tr.geoid = acs.geoid AND

CASE

WHEN tr.year < 2008 THEN 2008

WHEN tr.year > 2012 THEN 2012

ELSE tr.year

END = (acs.year - 2)

LEFT JOIN bg.sins sm ON tr.state = sm.postal\_code AND sm.year = 2010

LEFT JOIN bg.sins\_county sc ON "substring"(tr.geoid, 8, 5) = sc.fips

LEFT JOIN c ON tr.geoid = c.geoid AND tr.year = c.year

LEFT JOIN parcel\_risk pcl ON tr.geoid = '14000US' || pcl.state\_code || pcl.county\_code || pcl.tract

WHERE tr.year > 2006 AND tr.year < 2014

The LUSE\_swg table used in the first query above has the following structure.

|  |  |  |
| --- | --- | --- |
| **Column** | **Type** | **Description** |
| Code | text | CoreLogic Land Use Code |
| Description | text | Description of CoreLogic Land Use Code |
| risk\_category | text | Risk category (Low, Medium, High) |
| residential | text | Whether the code is associated with a residential use (Yes, No). |