CoreLogic Real Estate Data for Research

A CSCAR-MIDAS Workshop

CoreLogic Data

CoreLogic is a private company that provides data on different aspects of real estate

UM library bought parcel level data and maintains it

UM community has free, unlimited access. See here to get access:

https://search.lib.umich.edu/databases/record/43987

CoreLogic Data Files

| Data | Meta data and records count |

Deed sales → Michigan_Uni_**Deed**_KZA_85HRZB.zip (27 GB)

Deed_Layout_PropertyLevel_**Lat_Long**_11172016.xlsx

Michigan_Uni_Deed_KZA_85HRZB_**Count**.xlsx.

Tax → Michigan_Uni_**Tax**_AKZA_85HRQ5.zip (22 GB)

Tax_Layout_w_Property_Level_**lat_long**_w_code_01262017.xlsx

Michigan_Uni_Tax_AKZA_85HRQ5_**Counts**.xlsx

Foreclosure → Michigan_Uni_**FCL**_AKZA_85HRYY_Data.zip (6 GB)

FCL_Layout_Bulk_w_Code_Tables_04072017.xlsx

Michigan_Uni_FCL_AKZA_85HRYY_**Count**.xlsx

Records

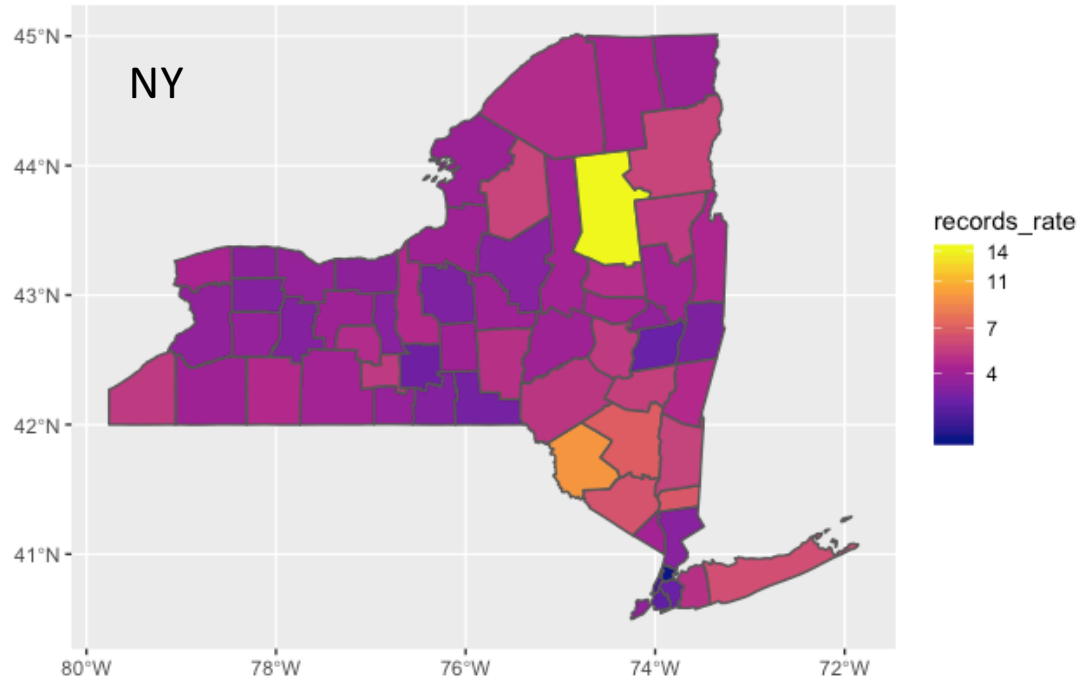|	
Deed sales	Total 3197 counties with 367,782,480 records
Tax	Total 3137 counties with 150,014,715 records (41% of deeds)
Foreclosure	Total 2893 counties with 37,487,908 records

The Tax data file mostly contains only one year of data (probably 2016)

There are a total of 3220 counties and 138.45 million housing units in the entire US (Wikipedia)
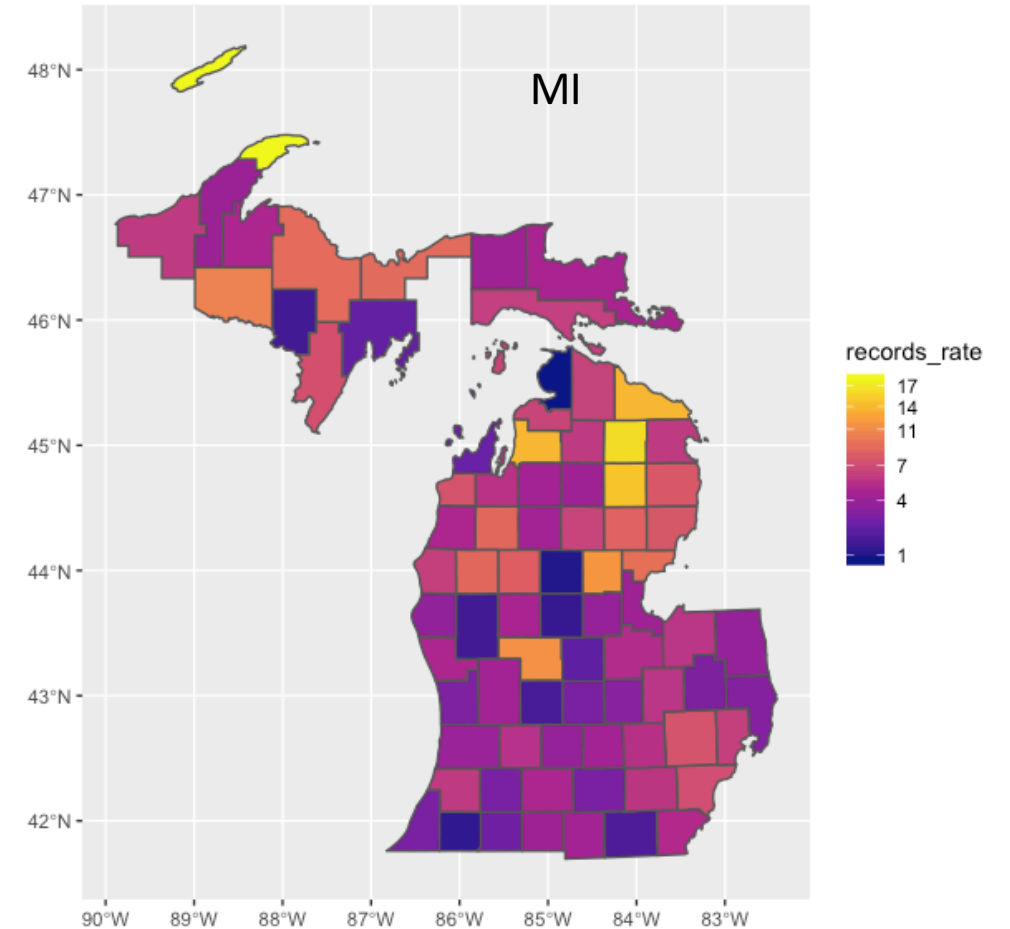
Coverage (deeds)

Total number of records for approximately per household (6 persons) based on 2010 census
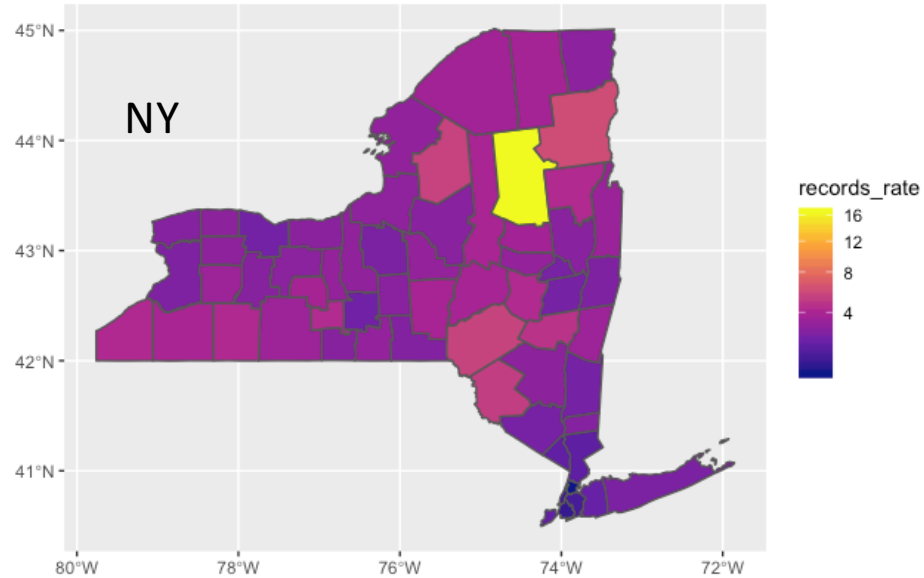
Assume that each family owns one parcel

Use the R code (coverage_map_1.R) to generate for any state you like

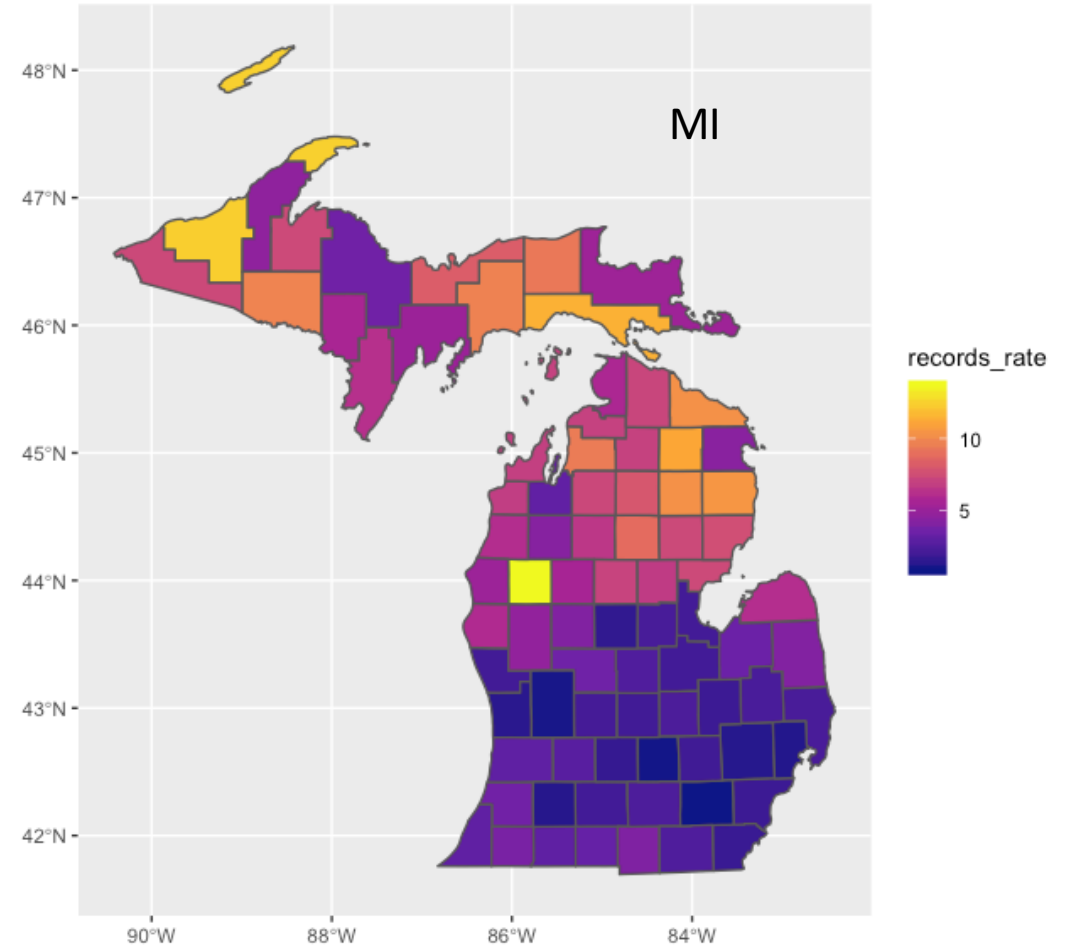Variation is real or missing data

Coverage (Tax)

Total number of records for approximately per household (6 persons) based on 2010 census

Color scales in the two maps are different

Use the R code (coverage_map_1.R) to generate for any state you like

Fields

Deed   179

Tax   134

Foreclosure   97

Most fields are empty

Run the first cell of the notebook to see different fields

FIPS and APN are the two common fields across the three data

FIPS is a five digit code, the first two identify state and the next three a county within the state

e.g. FIPS code for Washtenaw county is 26161 and for Wayne is 26163

APN (Assessor's parcel number), it is unique at least in a county (I think)

Field description and codes

Each metadata excel file can have more than one worksheet

For example, Deed_Layout_PropertyLevel_Lat _Long_11172016.xlsx has two sheets

Property indicator is a field with width 2 and type 'Alpha' that can take any of the several values indicated in sheet 2

PROPERTY INDICATOR	Alpha	2	CLGX generated code to easily identify types of property (Residential, Commercial, Vacant etc)

PROPERTY INDICATOR	10	Single Family Residence / Townhouse
PROPERTY INDICATOR	11	Condominium (residential)
PROPERTY INDICATOR	20	Commercial
PROPERTY INDICATOR	21	Duplex, Triplex, Quadplex)
PROPERTY INDICATOR	22	Apartment
PROPERTY INDICATOR	23	Hotel, Motel
PROPERTY INDICATOR	24	Commercial (condominium)
PROPERTY INDICATOR	25	Retail
PROPERTY INDICATOR	26	Service (general public)
PROPERTY INDICATOR	27	Office Building
PROPERTY INDICATOR	28	Warehouse
PROPERTY INDICATOR	29	Financial Institution
PROPERTY INDICATOR	30	Hospital (medical complex, clinic)
PROPERTY INDICATOR	31	Parking
PROPERTY INDICATOR	32	Amusement-Recreation
PROPERTY INDICATOR	50	Industrial
PROPERTY INDICATOR	51	Industrial Light
PROPERTY INDICATOR	52	Industrial Heavy
PROPERTY INDICATOR	53	Transport
PROPERTY INDICATOR	54	Utilities
PROPERTY INDICATOR	70	Agricultural
PROPERTY INDICATOR	80	Vacant
PROPERTY INDICATOR	90	Exempt
PROPERTY INDICATOR	00	Miscellaneous