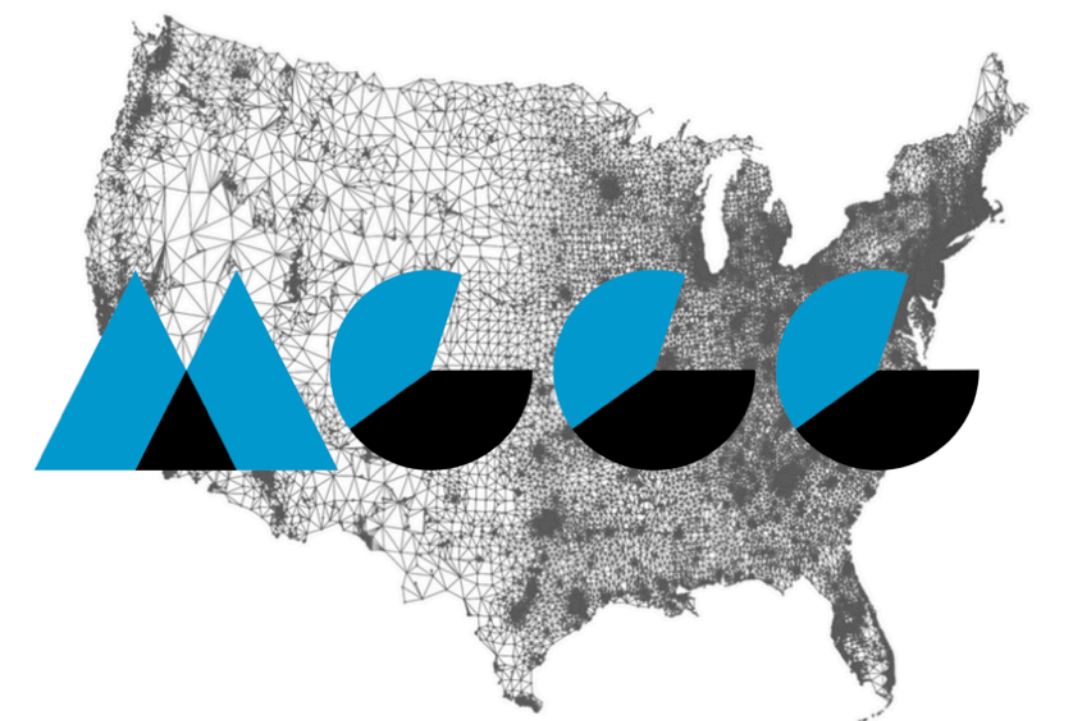
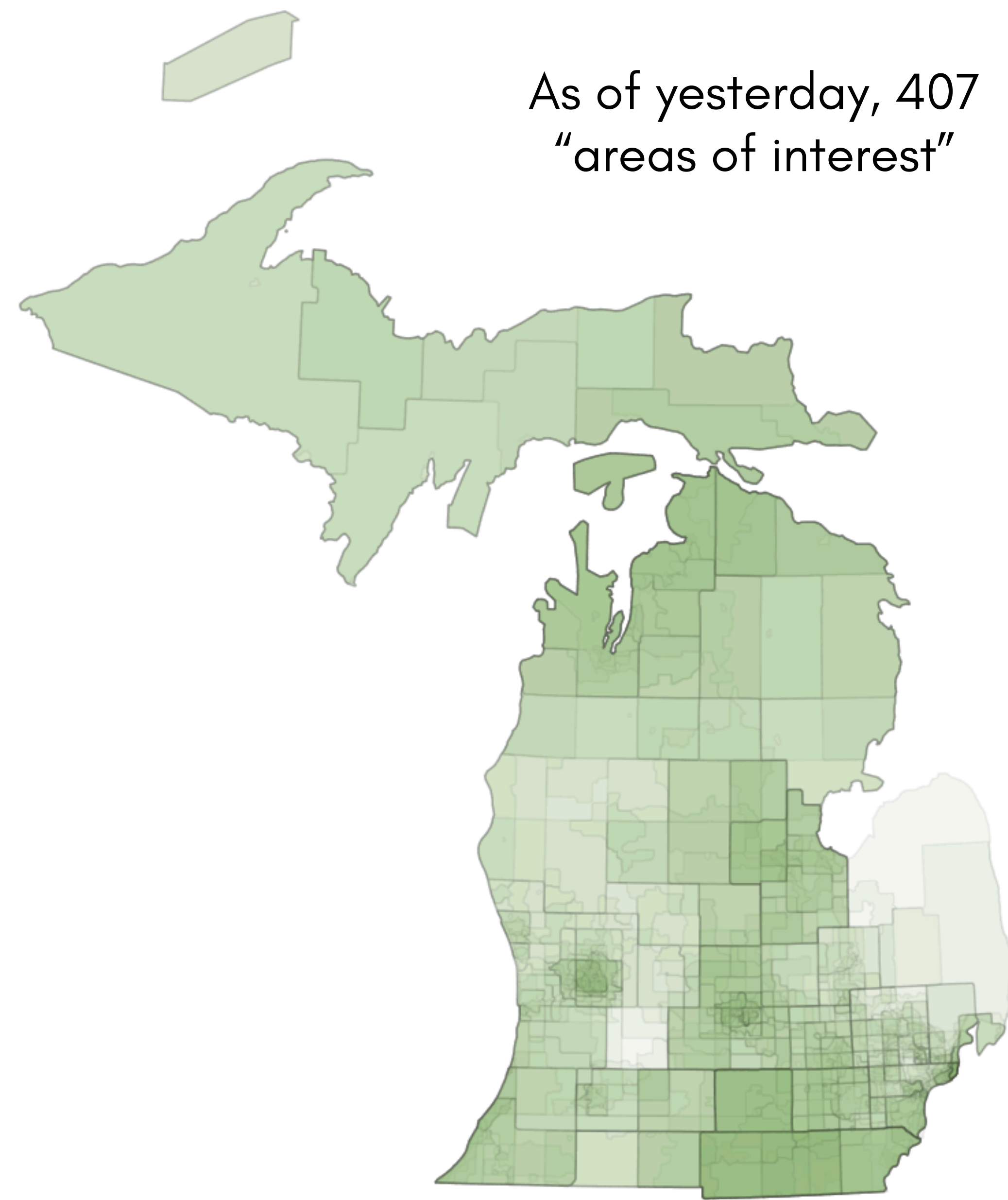
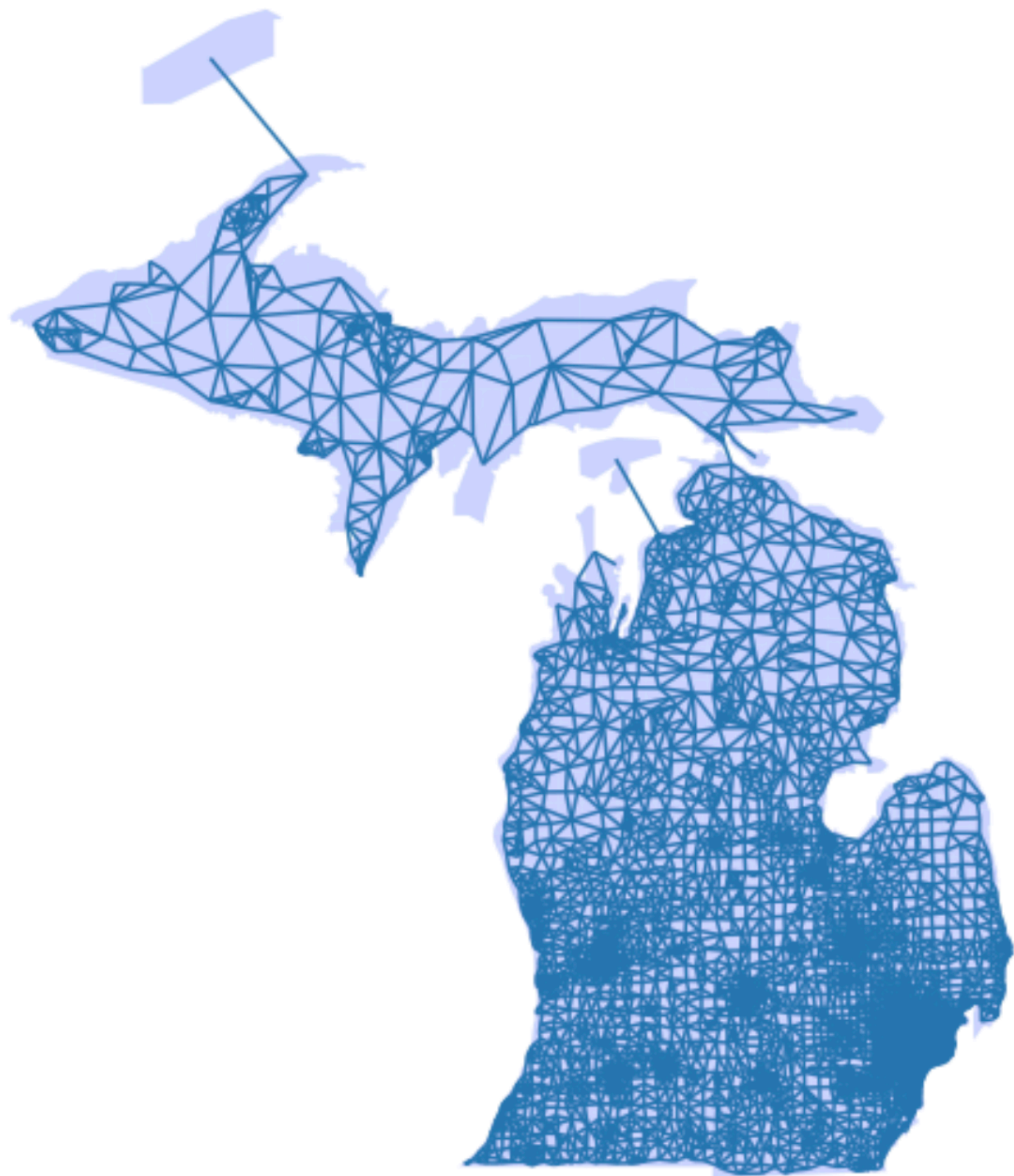


# Summary of aggregation process for COI submissions

Duchin, for MICRC, August 6, 2021





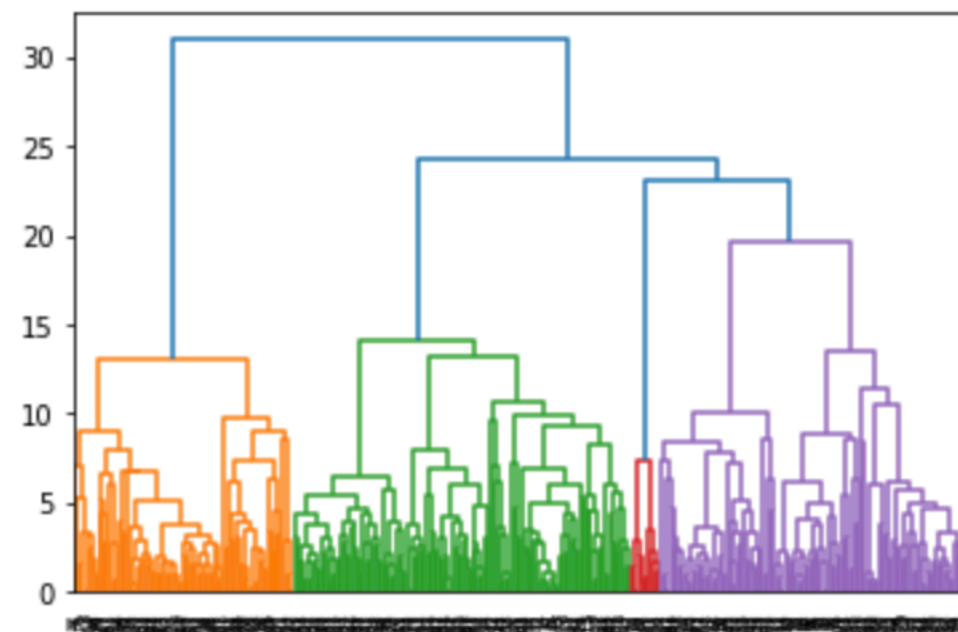
As of yesterday, 407  
"areas of interest"

# Geographic clustering

```
number_of_cois = coi_data.shape[0]
coi_total_dissimilarities = np.zeros((number_of_cois,number_of_cois))
for i in range(number_of_cois):
    for j in range(i+1,number_of_cois):
        first_map_plan = coi_data.iloc[i,4:]
        second_map_plan = coi_data.iloc[j,4:]
        first_map_plan = set(first_map_plan[first_map_plan>0].index)
        second_map_plan = set(second_map_plan[second_map_plan>0].index)
        coi_total_dissimilarities[i,j] = matching_distance_between_maps(first_map_plan,second_
map_plan)
    print("Finished row " + str(i) + " out of " + str(number_of_cois))

coi_total_dissimilarities = coi_total_dissimilarities + np.transpose(coi_total_dissimilaritie
s)
```

```
In [112]: dend = hierarchy.linkage(squareform(coi_total_dissimilarities),'average')
hierarchy.dendrogram(dend)
plt.show()
```



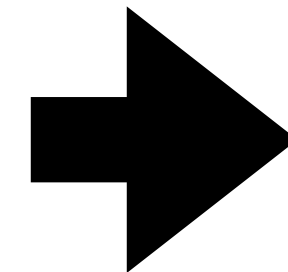
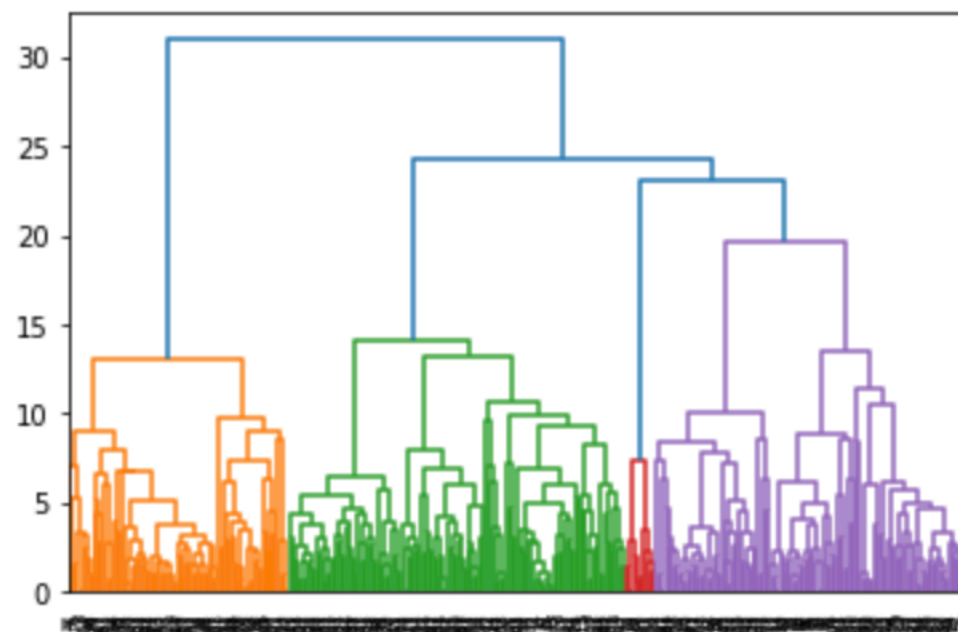


# Geographic clustering

```
number_of_cois = coi_data.shape[0]
coi_total_dissimilarities = np.zeros((number_of_cois,number_of_cois))
for i in range(number_of_cois):
    for j in range(i+1,number_of_cois):
        first_map_plan = coi_data.iloc[i,4:]
        second_map_plan = coi_data.iloc[j,4:]
        first_map_plan = set(first_map_plan[first_map_plan>0].index)
        second_map_plan = set(second_map_plan[second_map_plan>0].index)
        coi_total_dissimilarities[i,j] = matching_distance_between_maps(first_map_plan,second_
map_plan)
    print("Finished row " + str(i) + " out of " + str(number_of_cois))

coi_total_dissimilarities = coi_total_dissimilarities + np.transpose(coi_total_dissimilaritie
s)
```

```
In [112]: dend = hierarchy.linkage(squareform(coi_total_dissimilarities),'average')
hierarchy.dendrogram(dend)
plt.show()
```

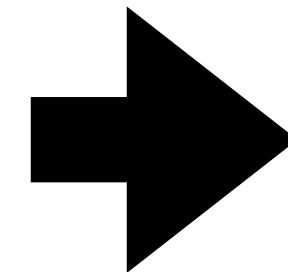
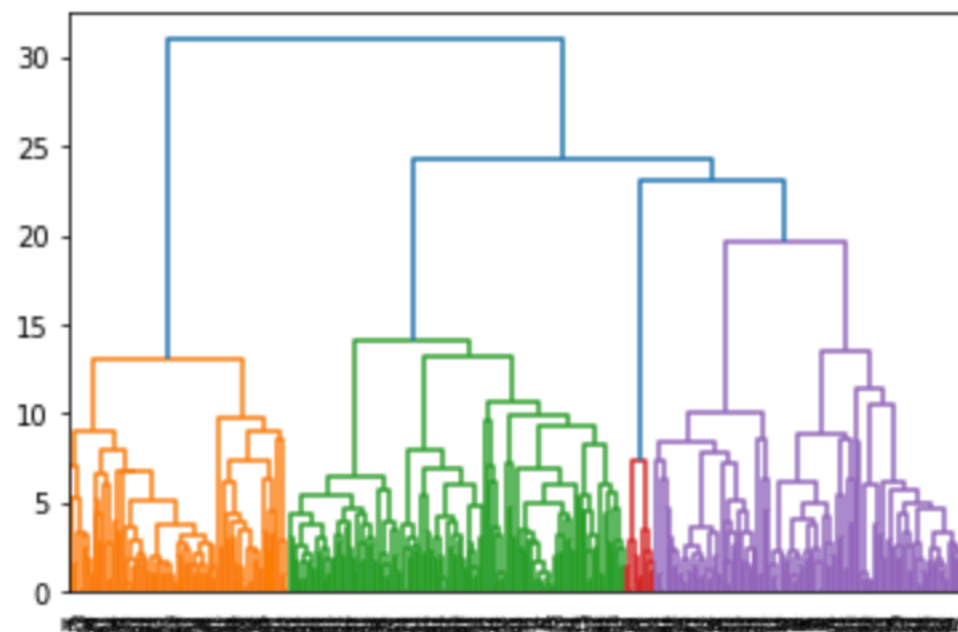


# Geographic clustering

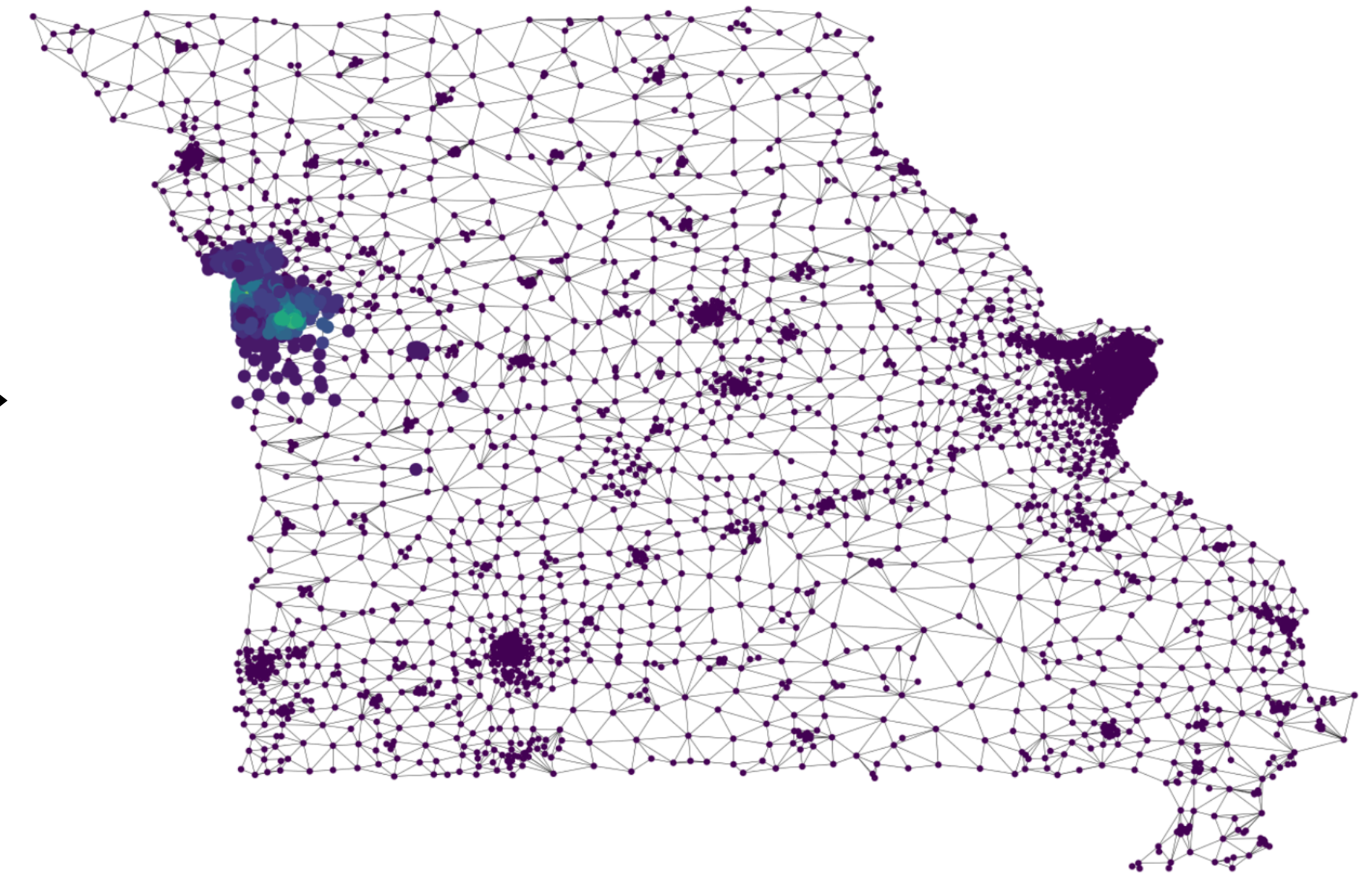
```
number_of_cois = coi_data.shape[0]
coi_total_dissimilarities = np.zeros((number_of_cois,number_of_cois))
for i in range(number_of_cois):
    for j in range(i+1,number_of_cois):
        first_map_plan = coi_data.iloc[i,4:]
        second_map_plan = coi_data.iloc[j,4:]
        first_map_plan = set(first_map_plan[first_map_plan>0].index)
        second_map_plan = set(second_map_plan[second_map_plan>0].index)
        coi_total_dissimilarities[i,j] = matching_distance_between_maps(first_map_plan,second_
map_plan)
    print("Finished row " + str(i) + " out of " + str(number_of_cois))

coi_total_dissimilarities = coi_total_dissimilarities + np.transpose(coi_total_dissimilaritie
s)
```

```
In [112]: dend = hierarchy.linkage(squareform(coi_total_dissimilarities),'average')
hierarchy.dendrogram(dend)
plt.show()
```

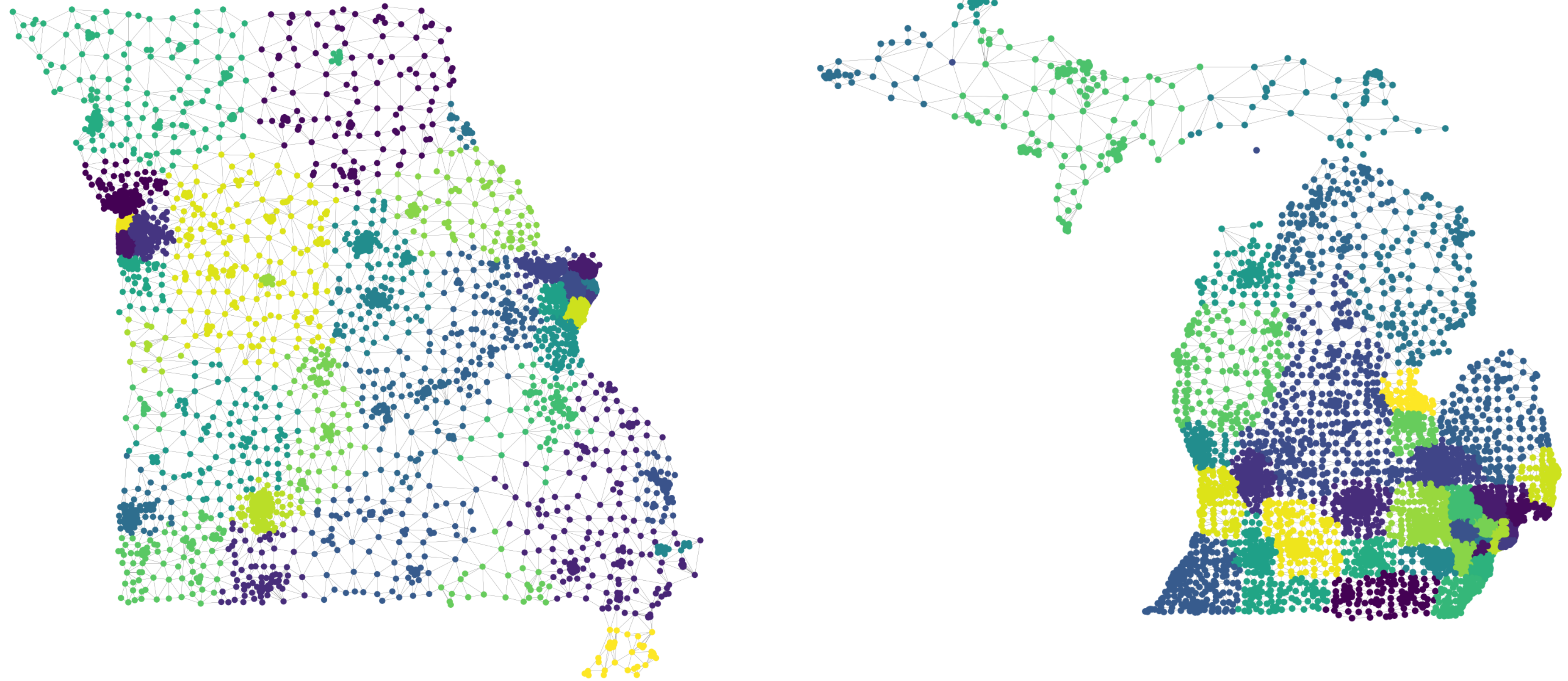


Cluster 1 contains 90 submissions





## Geographic clustering II – spectral methods



*Michigan/Missouri partitioned into 40 regions. We used a "spectral clustering" technique that avoids splitting up geographic units that usually appear together in COI submissions.*

# Narrative clustering

	submission_text	area_text	area_name
13234-1	Single-family homes, multigenerational	Romanelli West	Romanelli West is outlined by these streets:\nN: 64th Terrace\nS: Gregory/71st\nE: Ward Parkway\nW: State Line Road\nOne-family homes. \nMultigenerational, a mix of single owners, families and older owners.\n
13234-1	Single-family homes, multigenerational	Romanelli West	Romanelli West is outlined by these streets:\nN: 64th Terrace\nS: Gregory/71st\nE: Ward Parkway\nW: State Line Road\nOne-family homes. \nMultigenerational, a mix of single owners, families and older owners.\n
16818-1	Gentrification, family issues, autonomy from Downtown KC. Our neighborhood is truly a residential area, but tends to be over-taken by the Crossroads and Downtown areas, which causes many issues amongst residents and city officials.	Westside Neighborhood-KCMO	This community sits just West of Downtown Kansas City, and has many diverse residents, including Hispanic, Black and White citizens.

Hyde Park, plus Rockhill and Southmoreland neighborhoods north of 47th and east of Oak in order to include the census tract in which the southernmost part of Hyde Park is located, is a diverse community of older (100+ years) houses with some duplexes and apartments (many in tall buildings on Armour Blvd and some in smaller apt. buildings in north and south Hyde Park and Southmoreland). Many of the apt. buildings on Armour Blvd were 100% Section 8 until Mac Properties started turning them into market rate apartments probably less than 10 years ago, and only two 100% Section 8 buildings remain. Mac is building 3 new apt. buildings on Armour (also market rate), and there is already not enough off-street parking. Although a nursing home and a group home for people with developmental disabilities have been returned to large single family homes in the last several

This community has mostly older houses (100+ years). "Urban pioneers" moved to Central Hyde Park (Armour to 39th) in the '70s and rehabbed houses, returning houses that had been divided into apartments after WWII back into single family homes. The neighborhood association became hyper vigilant about "density," but there have always been apartments and duplexes throughout Hyde Park, mostly in North and South Hyde Park. In the last decade or so more houses in N and S Hyde Park seem to have become owner-occupied, and more have small children. Historically, real estate agents steered home buyers away, citing

Narrative clustering options include various TDF/IDF, SBERT (sentence-based), word2vec, and more NLP packages

## What is the goal?

We are preparing a report identifying several dozen “**emergent COIs,**” each supported by multiple public submissions and identified through clustering processes

We will have titles, narratives, and scale descriptions for each, as well as the ability to link back to the supporting submissions

You can decide how to negotiate these in your Congressional, Senate, and House plans