



BIVARIATE MAPPING

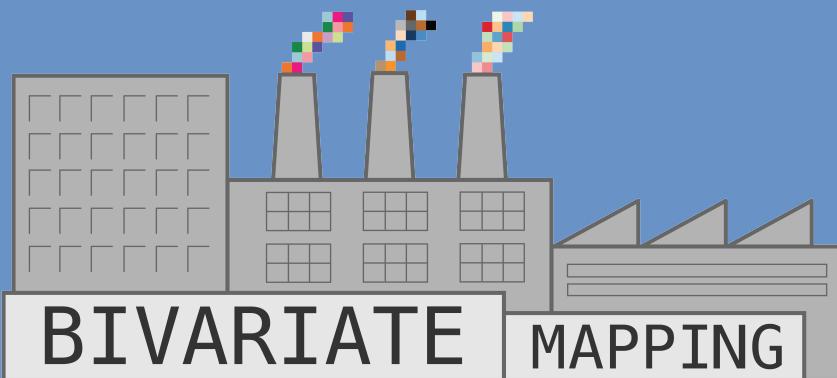
An Introduction and Case Study
Land Use / Land Cover with Visual Analytics

Ben Thornton

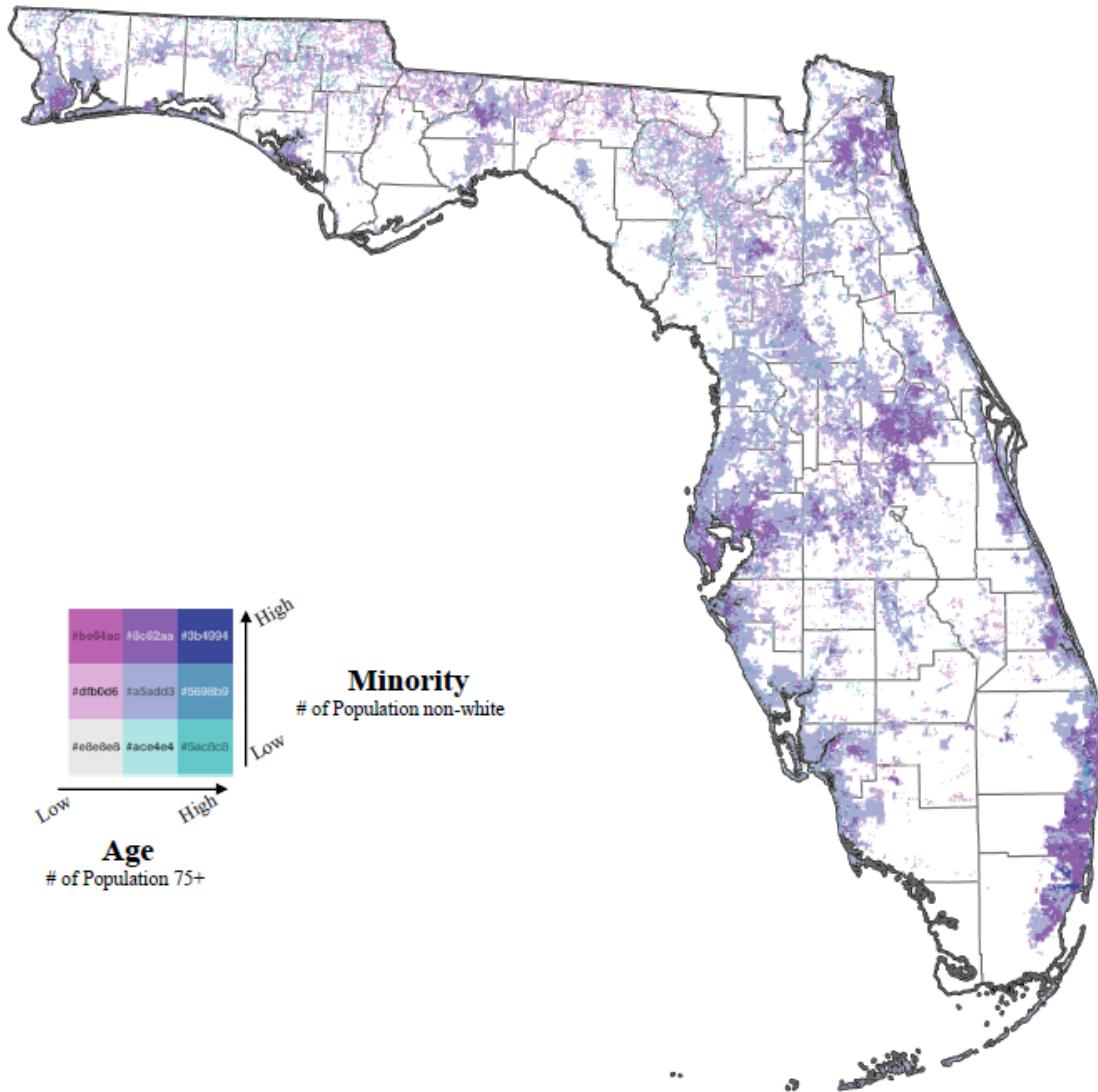
Contributors: Georgianna Strode, Evan Rau & Nathan Johnson

Advisors: Drs. Mesev, Yang, Morgan, & Trumbo

Background



Bivariate Analysis of Minority and Elderly (75+) Population



The first attempt ...

We want to see the
Interrelationships of
High elder and high
Minority

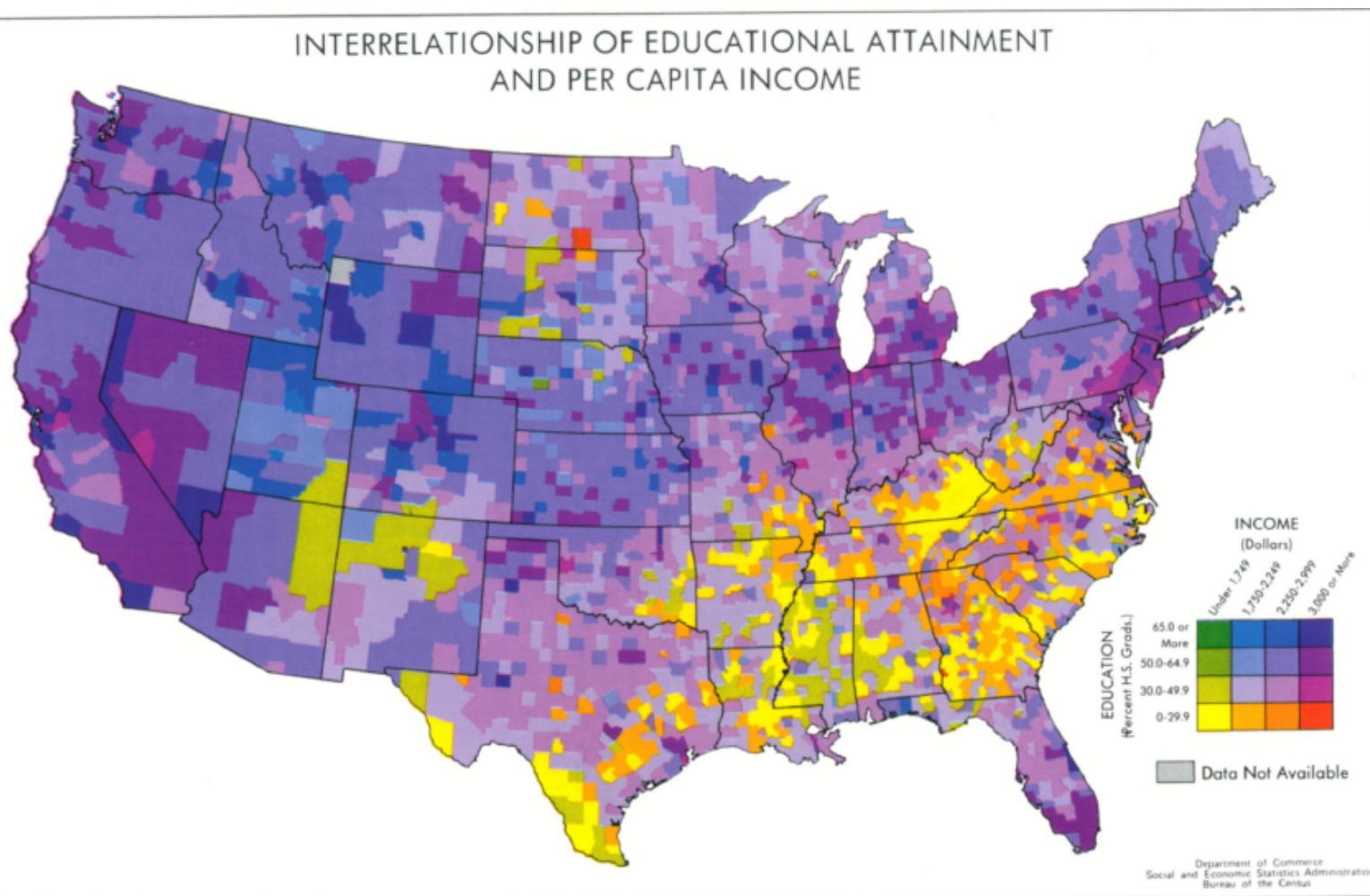
But ... it looks “mushy”

Can anyone tell the
difference between
Blue-violet and
Violet-blue?

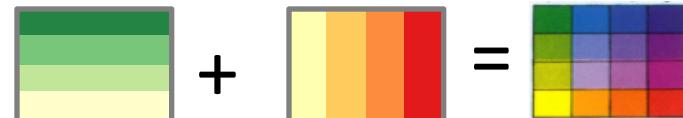
What is the problem?

- Scale?
- Color?
- Data?

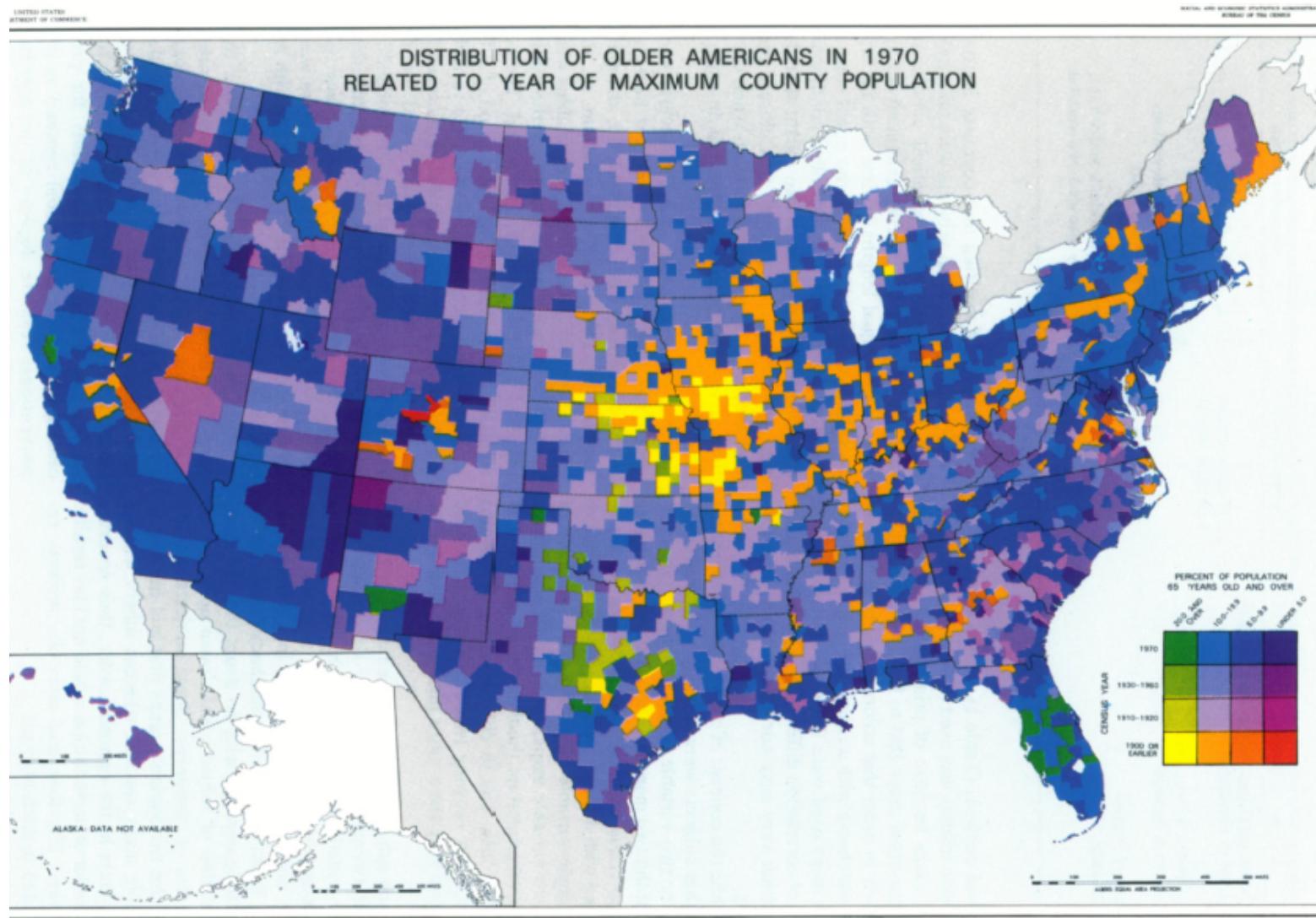
1976 Census Map



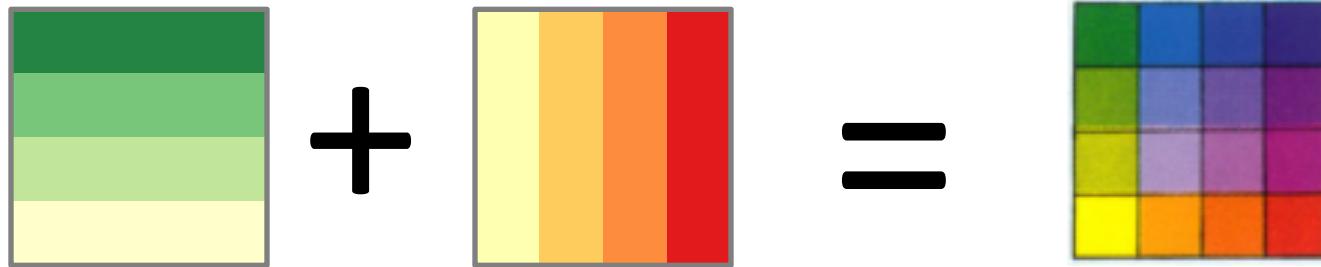
Method: “crossing” or “criss-cross” or “overlay”



1976 Another Sample Map



1976 Overlay Theory



C_{41} GREEN	C_{42}	C_{43}	C_{44}
C_{31} YELLOW GREEN	BLUSH	VIOLETS	DEEP
C_{21} GREEN-YELLOW	PALE	REDDISH	
C_{11} YELLOW	C_{12} ORANGE-YELLOW	C_{13} YELLOW-ORANGE	C_{14} ORANGE

In the 1980s, a simple color improvement was made by not using yellow in the lower left. However, this did not fully solve the problem.

Explanation of the Results
-- Trumbo, 1981

What's wrong with the overlay?
There is no clearly defined focal axes.

Which color should the reader be looking at?

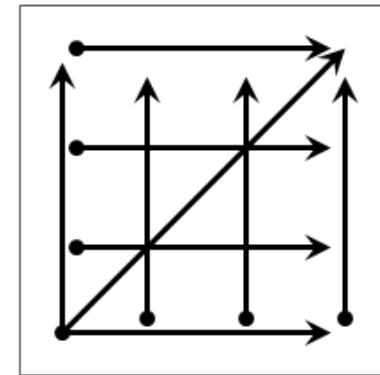
Potential problems with overlay method

1. **Unpredictable colors may need editing** - Different hues mix in different ways, making for complex combinations of colors especially between primary, secondary and tertiary color possibilities. Hence the need for a post-overlay manual method (Morgan).

2. **Won't work on categorical data** - Assumption that both y and x are continuous variables and will not work if one axis is categorical (Morgan).

3. **No focal point** -- Focus is on everything:

- Relationship of x and y (diagonal)
- Preserves univariate range in both x and y
- Shows highs and lows of x and y (Strode)



Where is focal point?



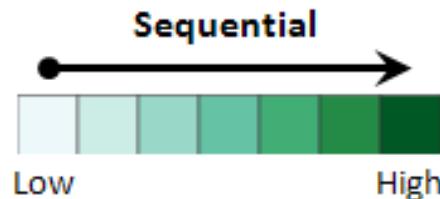
1981 Trumbo's Four Principles

- I. Order – For data ordered quantitatively, the colors chosen to represent the data should be perceived as preserving the order. Colors should have an orderly progression in hue, saturation, and brightness.
 - II. Separation – Important differences in values should be represented by colors perceived as different.
-
- I. Rows and Columns -- Colors should be distinct so that the corners stand out. Rows and columns representing the univariate data should be in a visual sequence.
 - II. Diagonal – If the interaction between variables is important, then the principal diagonal should be the focal point. The data should be divided into three classes: near or on the diagonal and those skewed to one side or the other.

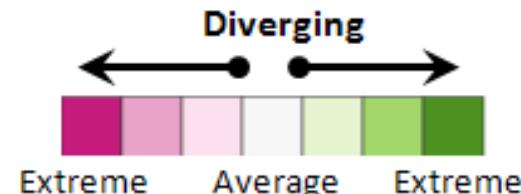
Trumbo, 1981. "A Theory for coloring Bivariate Statistical Maps."
The American Statistician 35(4):220-226.

1981 Trumbo Principles I and II

I Order



II Separation



Description range, progression, gradient

change, difference

Emphasize high, more, most

highs and lows, change,
distinct, difference, standard deviation,
unexpected, outlier, extremes

De-emphasize low, less, least

average, common, normal,
regular, expected, unchanged, middle

Principles III and IV

III Rows and Columns

Colors should be **distinct** so corners stand out

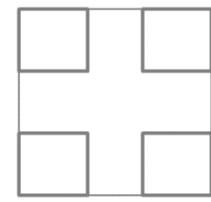
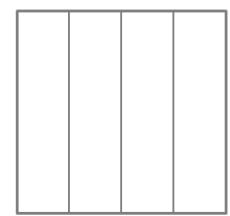
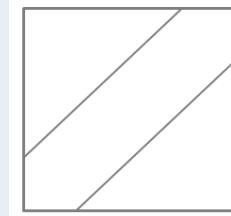
Univariate data should be in visual **sequence**

IV Diagonal

If interaction is important, diagonal should be focal point.

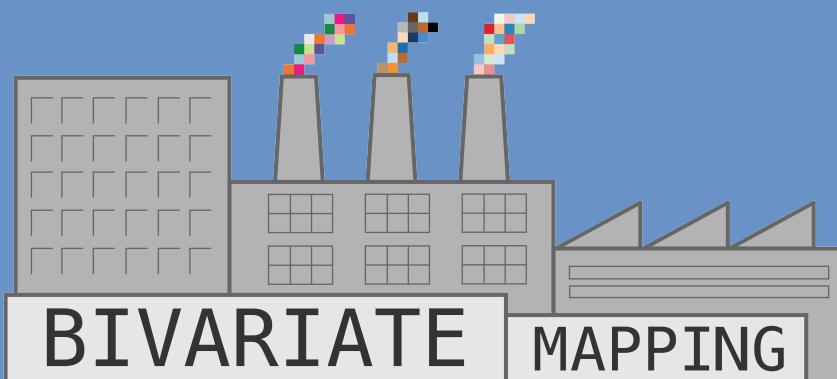
Data divided into 3 classes: on diagonal, above or below diagonal

Three Types of Questions a Bivariate Map Can Answer

Inquiry Formula	Sample Question	Focal Diagram
low/high of __ and low/high of __	Where are high elder populations with low income?	
range of __ within low/high of __	What is the range of education among high earners?	
relationship of __ and __	What is the relationship between income and education?	

(Trumbo, 1981)

Proposed Method



BIVARIATE MAPPING

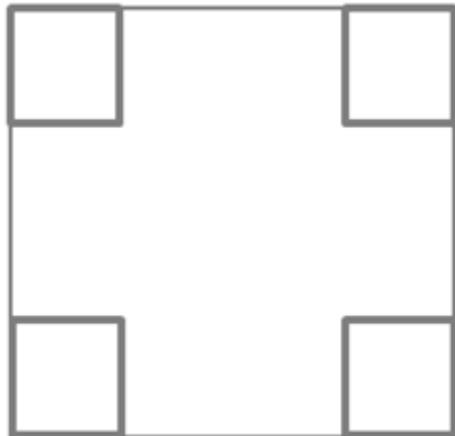
Focal Diagram Corners

Inquiry Formula: low/high of x and low/high of y

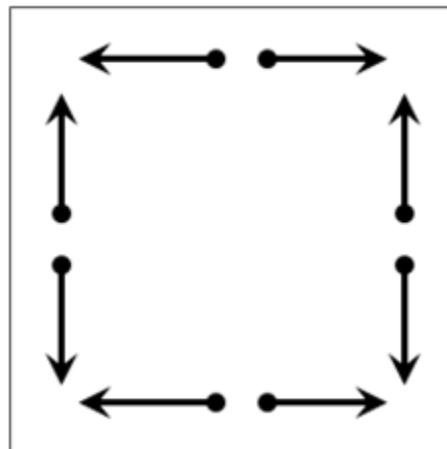
Sample Questions:

- Where are areas of high income and low education?
- Where are areas of low population density and high crime?
- Where are areas of high public transportation and high food deserts?

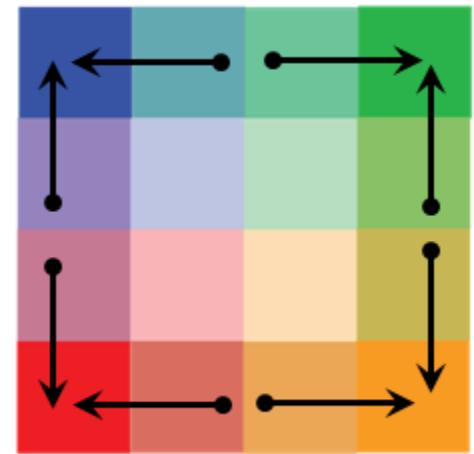
How It Works: multiple complementary diverging color schemes highlight the distinct corners while minimizing the interior



Focal Areas



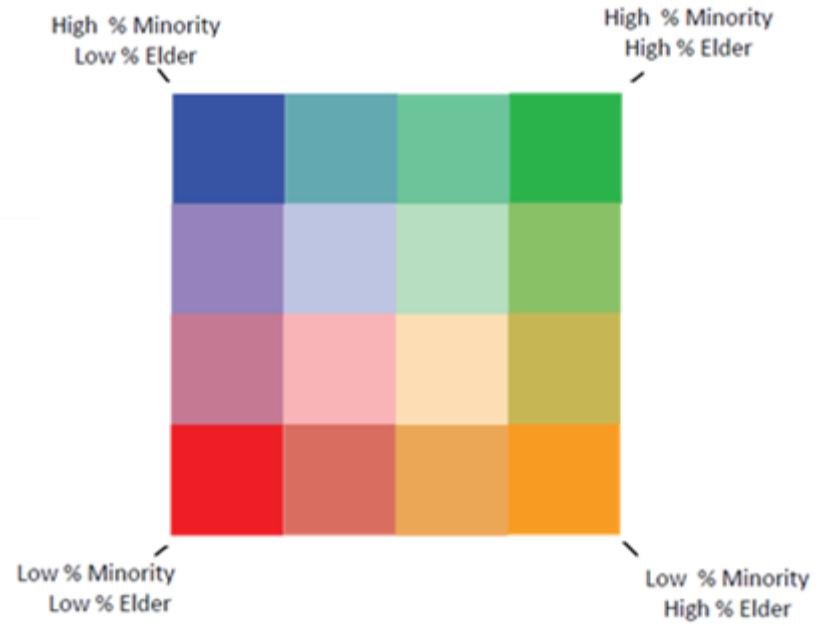
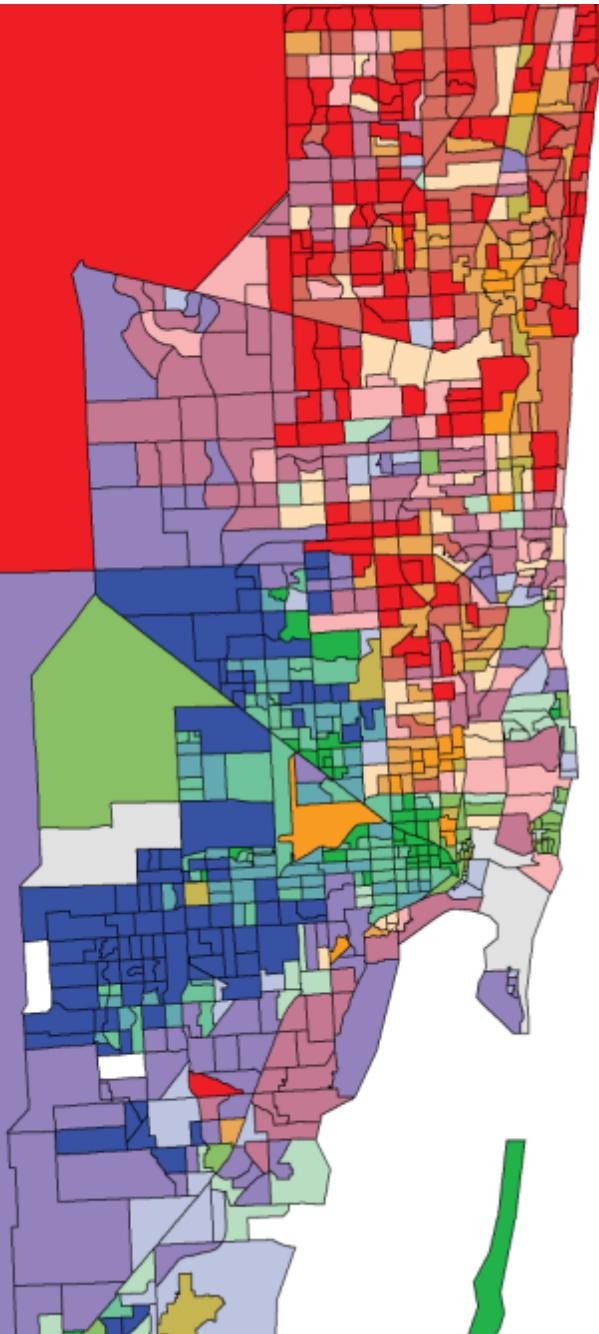
Focal Axes



With Color

Sample Corner Legend

Percent Elder and Minority



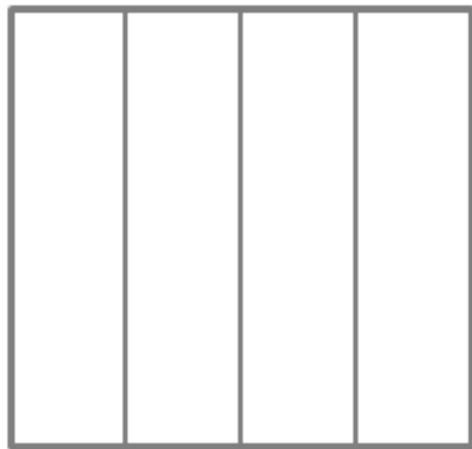
Focal Diagram Range

Inquiry Formula: range of y within low/high of x

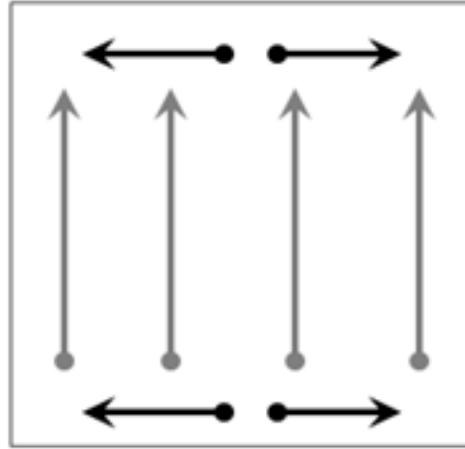
Sample Questions:

- What are the ranges of education among high incomes?
- About how many votes were cast in areas with strong Obama support?
- What are the income levels in areas of high foreclosures?

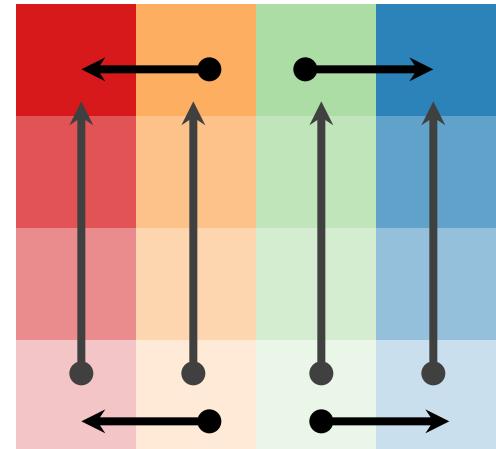
How It Works: diverging color scheme organizes multiple sequential schemes



Focal Areas



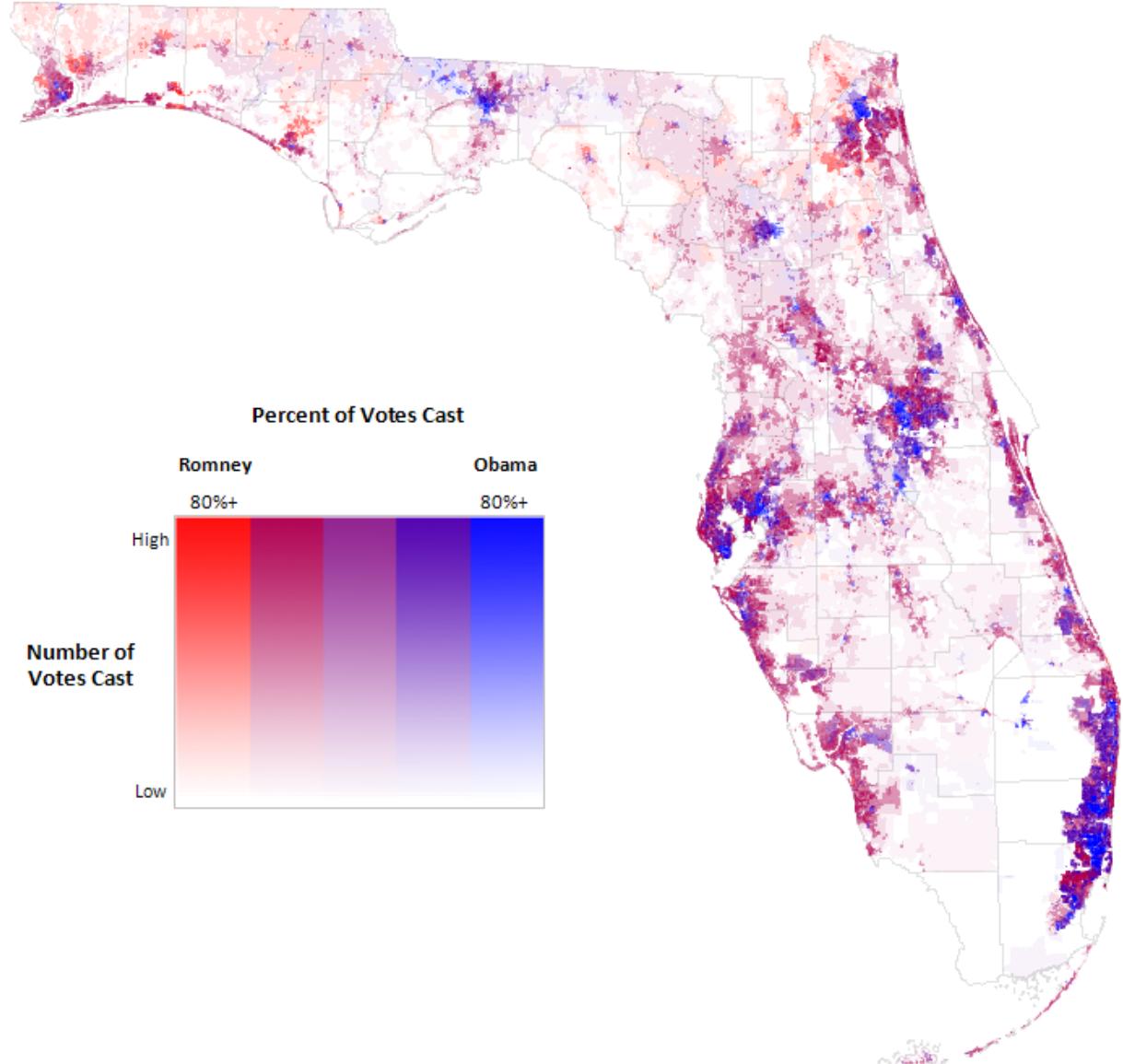
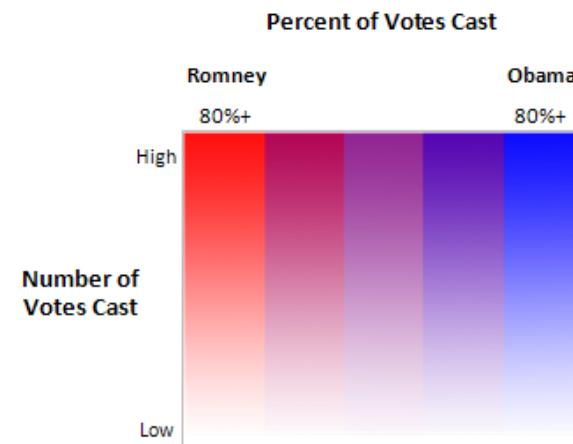
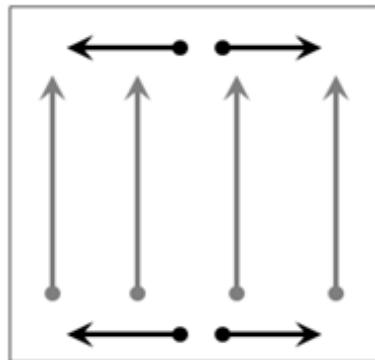
Focal Axes



With Color

Sample Range Map

Florida 2012 Presidential Election



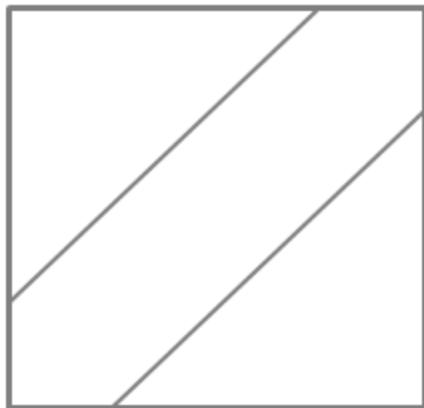
Focal Diagram Diagonal

Inquiry Formula: relationship of x and y

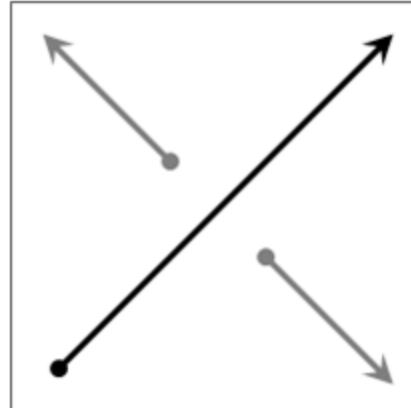
Sample Questions:

- What is the relationship between income and education?
- Are tobacco sales and food deserts related?
- Is there a relationship between population density and public transportation?

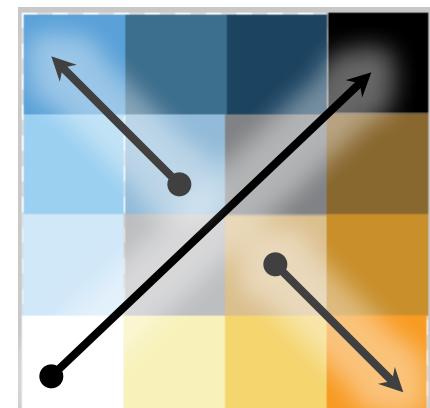
How It Works: sequential color scheme on diagonal with two complementary diverging schemes on the opposite diagonal to create three identifiable areas



Focal Areas



Focal Axes



With Color

Ostwald Color Model

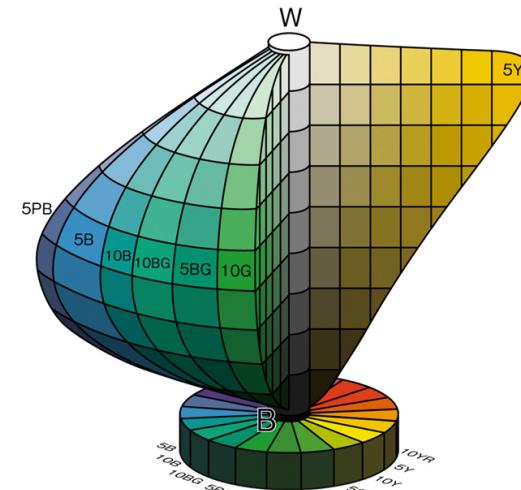
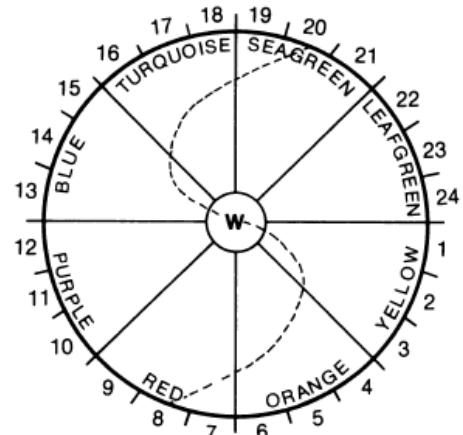
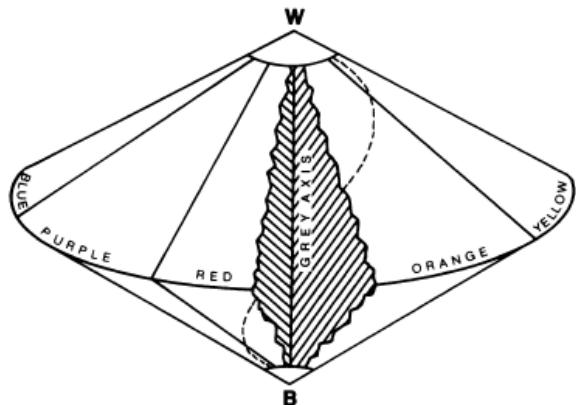


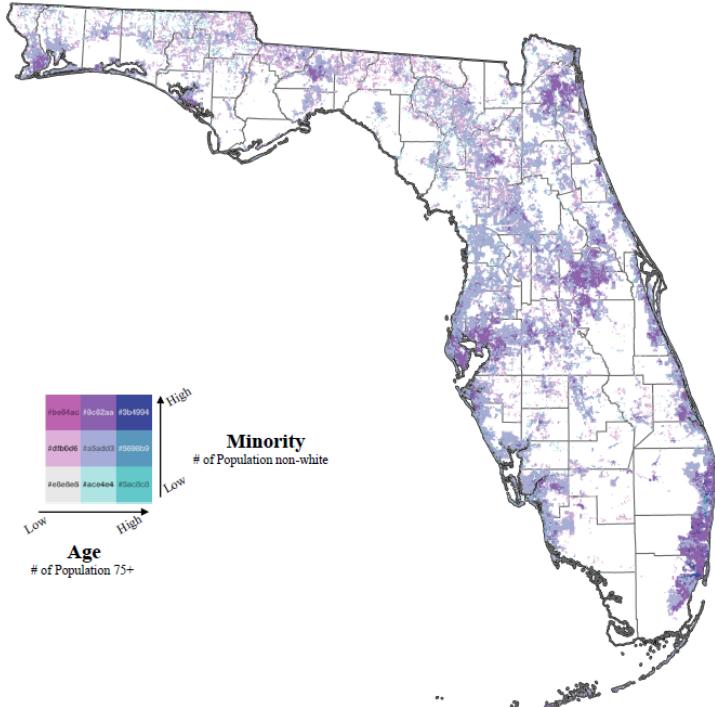
Figure 5. Side and Top Views of Ostwald Model

-- from Trumbo 1981

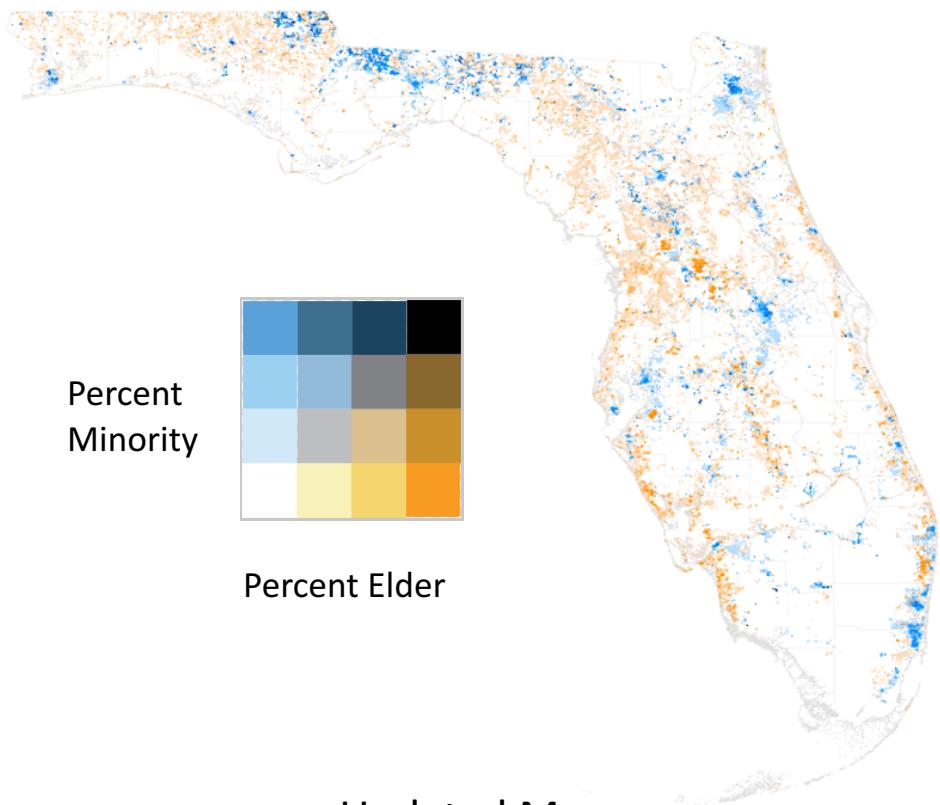
Sample Diagonal

Florida Elders and Minorities

Bivariate Analysis of Minority and Elderly (75+) Population

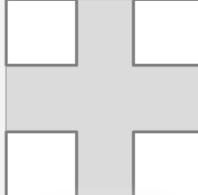
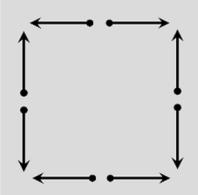
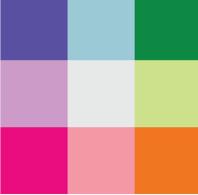
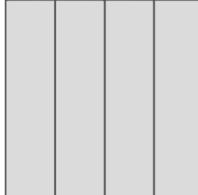
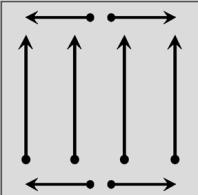
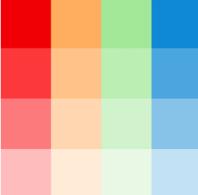
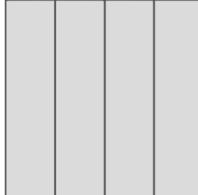
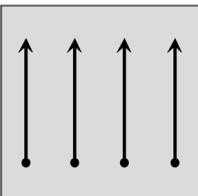
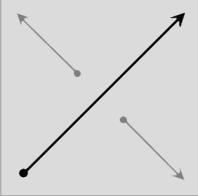


Original Map



Updated Map

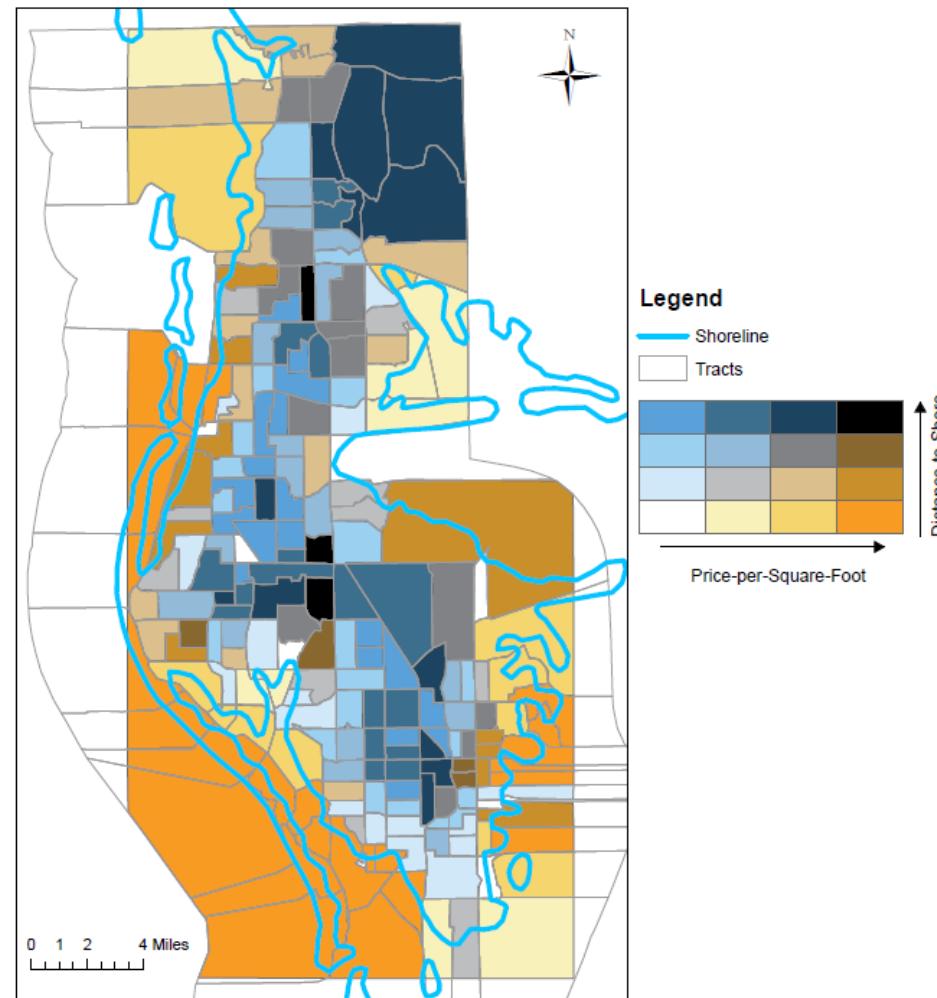
TYPES OF BIVARIATE MAPS

Focal Model	Inquiry Syntax & Simple Question	Focal Areas	Focal Axes	Sample Color Palette
Corners	<p>low/high of x and low/high of y</p> <p>Where are areas of high income and low education?</p>			
Range	<p>Diverging</p> <p>range of y within low/high of x</p> <p>What is the range of education among high earners?</p>			
	<p>Qualitative</p> <p>range of y within category</p> <p>What is the range of education within -- categories?</p>			
Diagonal	<p>relationship of x and y</p> <p>What is the relationship of income and education?</p>			

Based on Trumbo's Four Principles (1981)

Price-per-Square-Foot by Distance-to-Shoreline

Pinellas County, FL Case Study



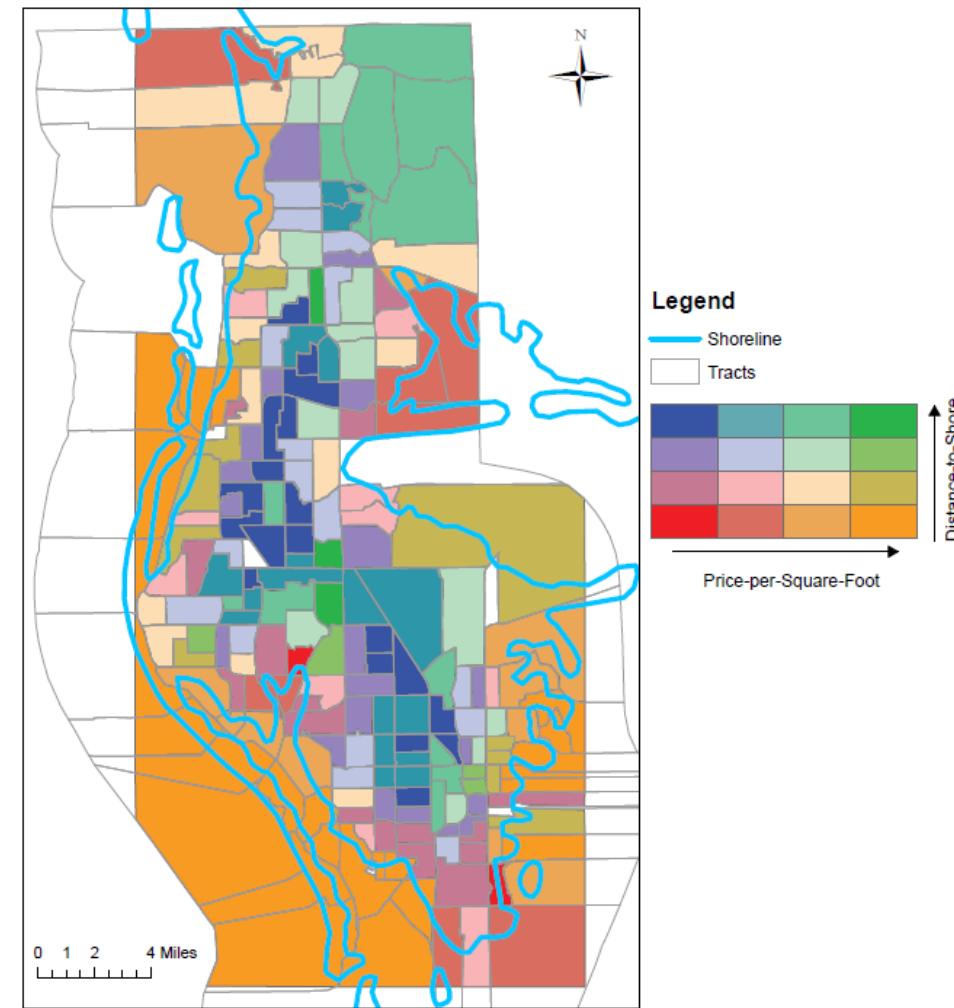
Data Sources:

Census tract boundaries from Pinellas County GIS Group
Parcel sales data from Pinellas County Property Appraiser
Shoreline generalized from boundary from U.S. Geological Survey
Classification method: 4-class quantile (Manual)

Created By:
John Derek Morgan, PhD, GISP
11/17/15

Price-per-Square-Foot by Distance-to-Shoreline

Pinellas County, FL Case Study



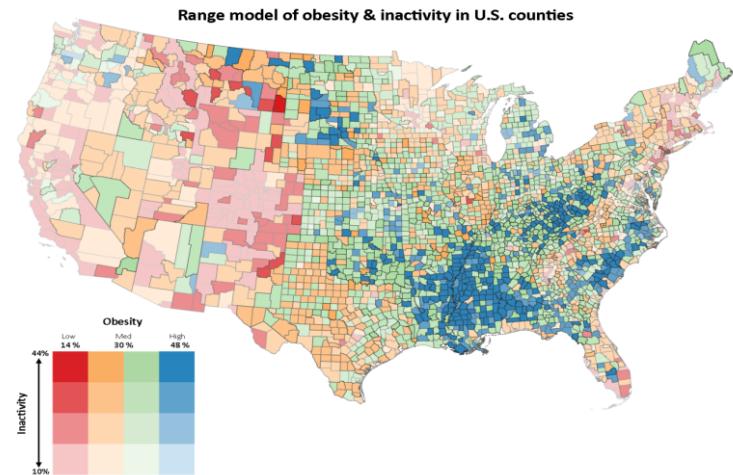
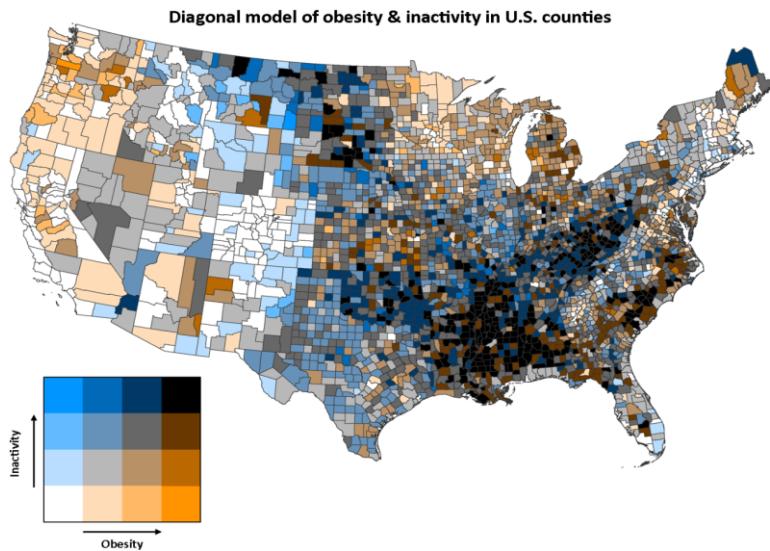
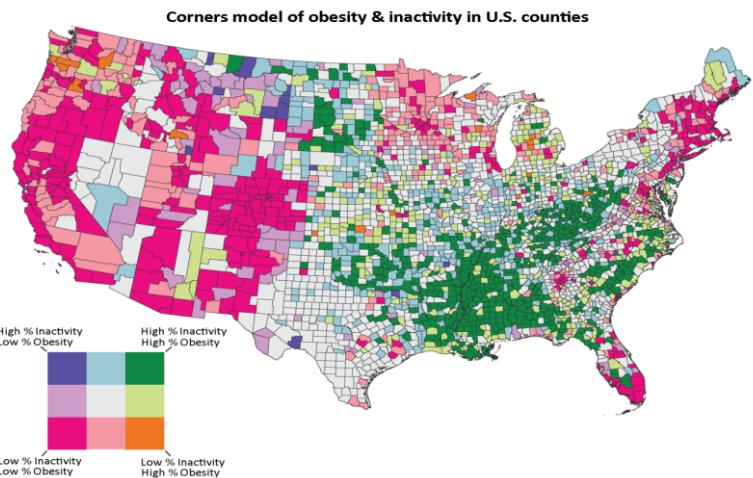
Data Sources:

Census tract boundaries from Pinellas County GIS Group
Parcel sales data from Pinellas County Property Appraiser
Shoreline generalized from boundary from U.S. Geological Survey
Classification method: 4-class quantile (Manual)

Created By:
John Derek Morgan, PhD, GISP
11/17/15

Summary Example

- 1 Data Set
- 3 Maps
- 3 Questions/Stories



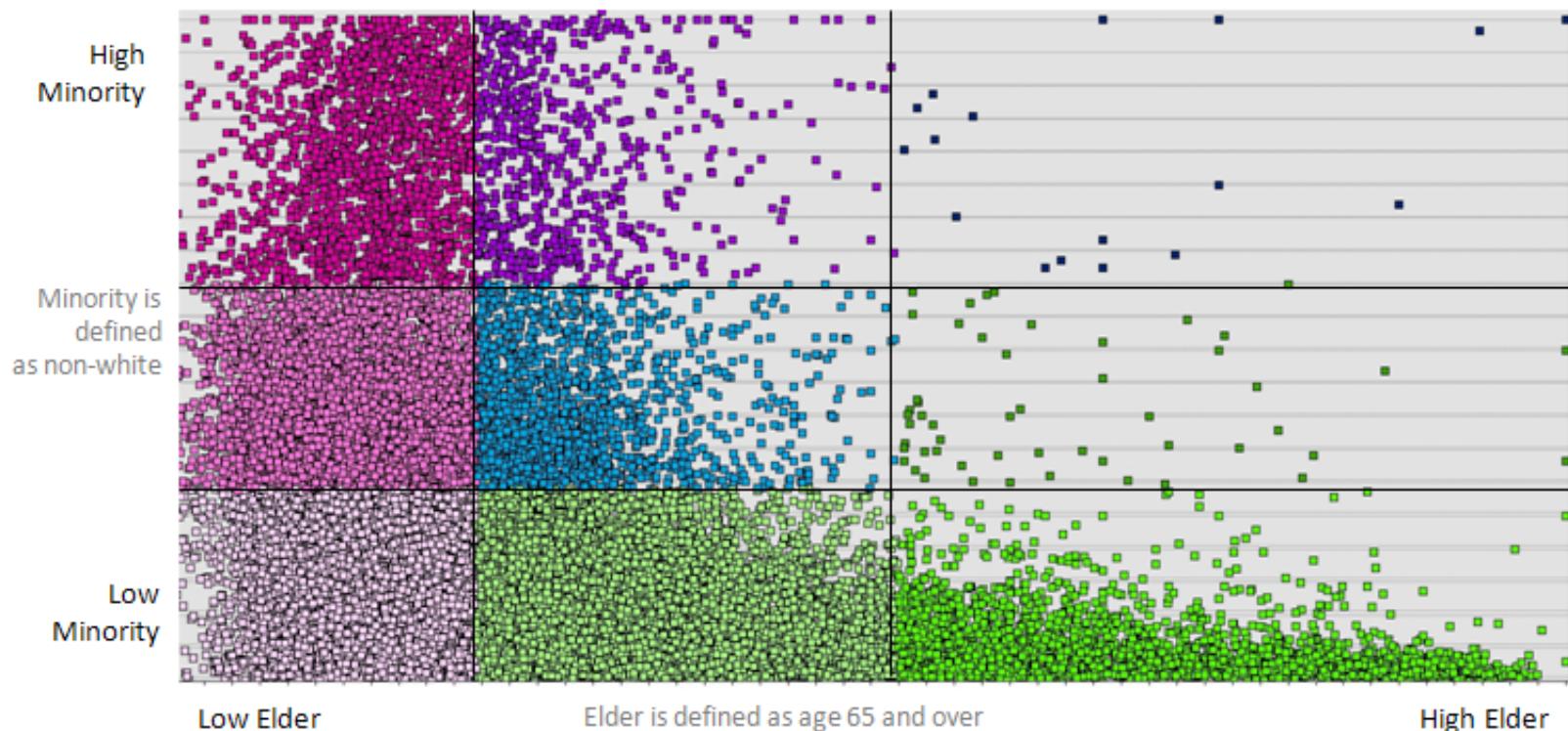
Bivariate Statistical Legends



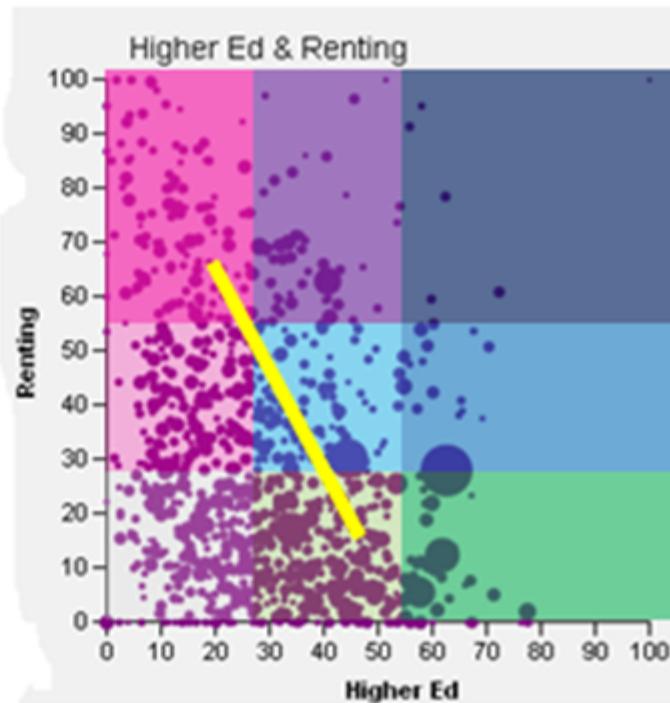
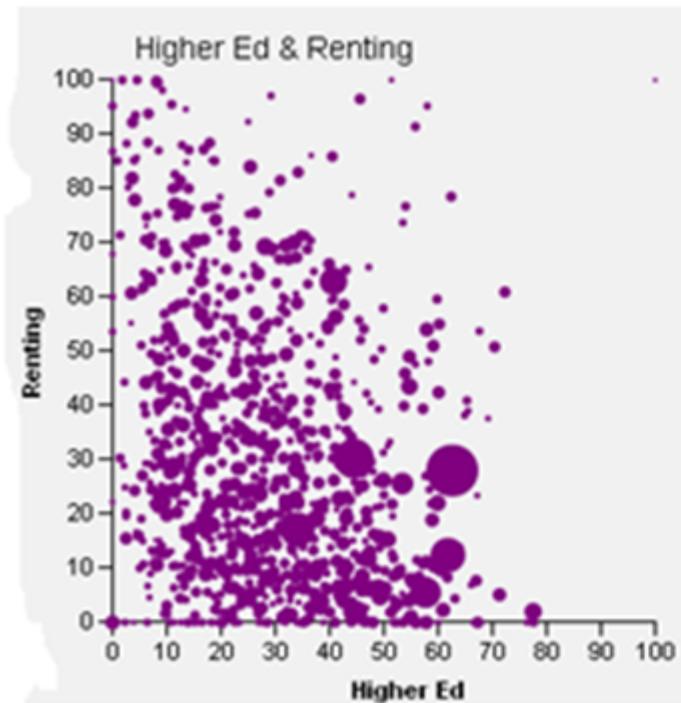
Scatterplots

Florida Elders and Minorities

Per 1-km Grid Cell
As Percentage of Total Population
3-class Jenks Natural Breaks



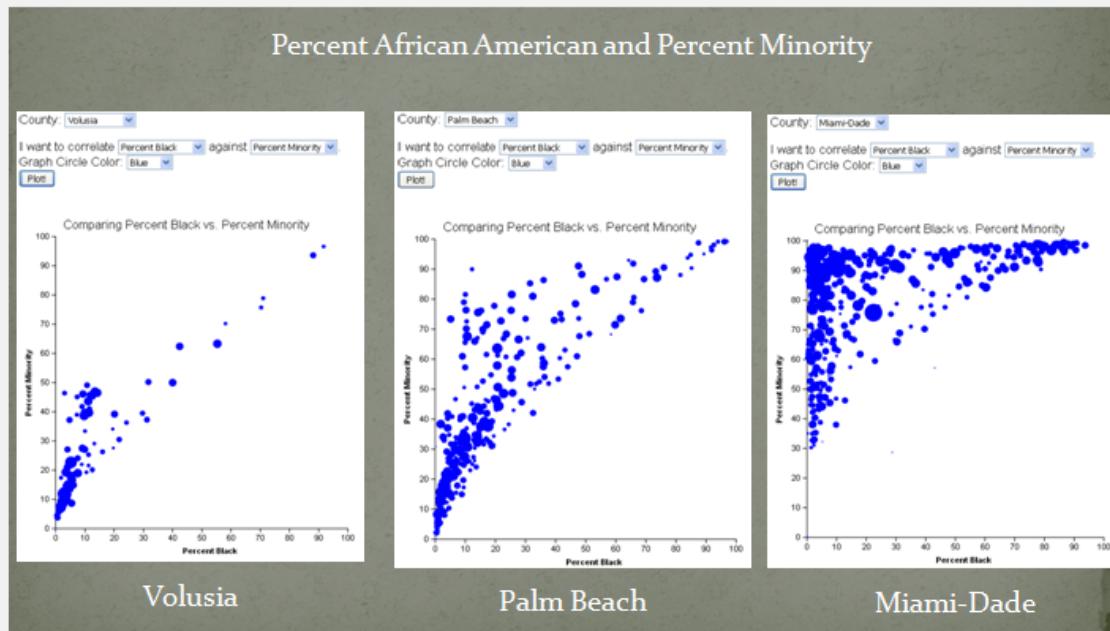
Turn a scatterplot into a legend



Scatterplots Show Relationships

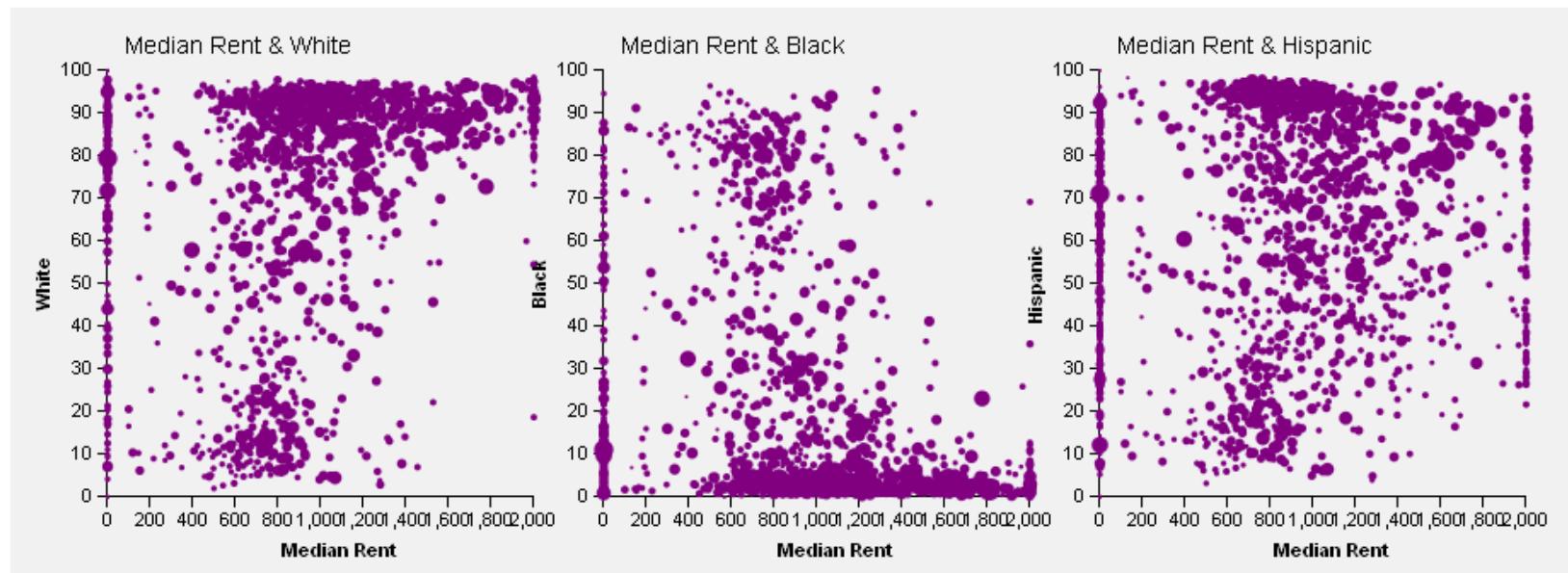
The scatterplots below show the distribution of percentages of African American and Minorities across 3 counties. Each county has its own pattern that tells a different story. Miami-Dade County has a high percentage of minorities that are not African American. Volusia County has a much different distribution.

Relationships between African American and Minorities in Three Counties



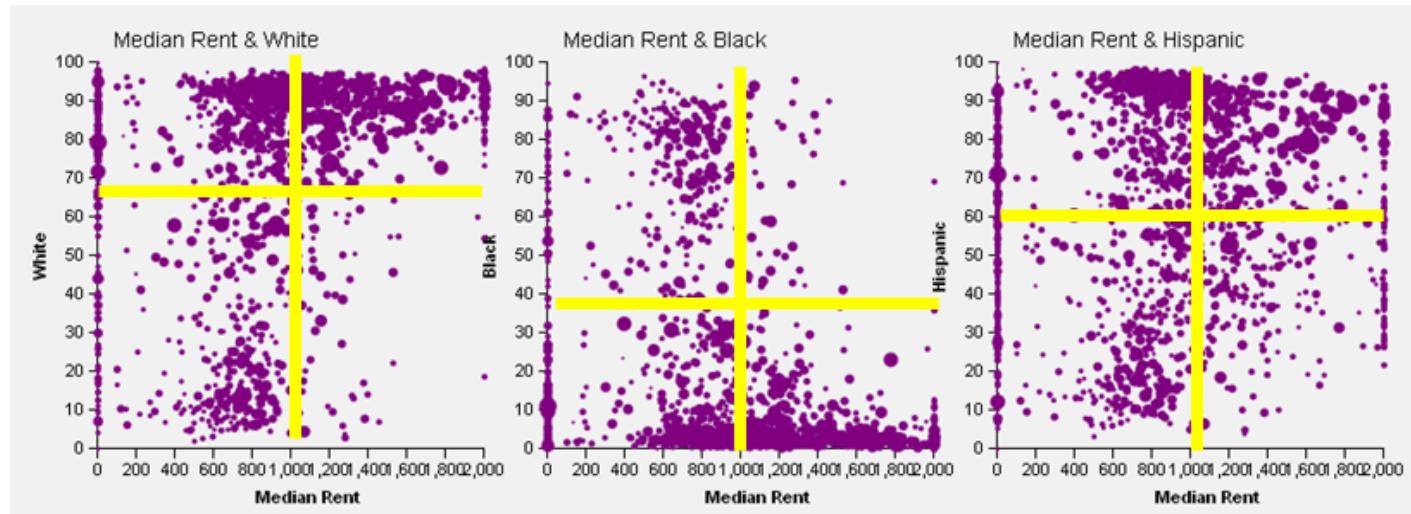
DIY Scatterplot Maker – Sample Results

Sample Results

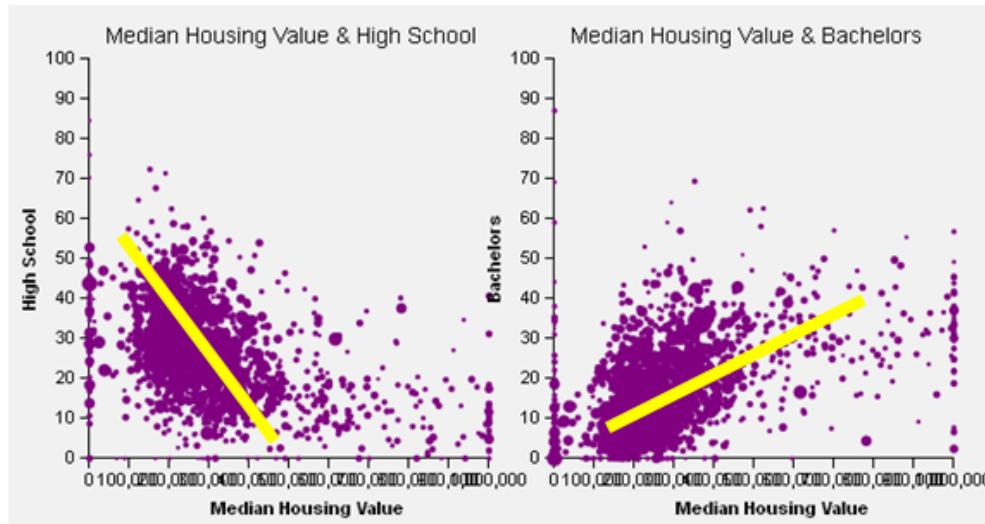


What Can We Learn?

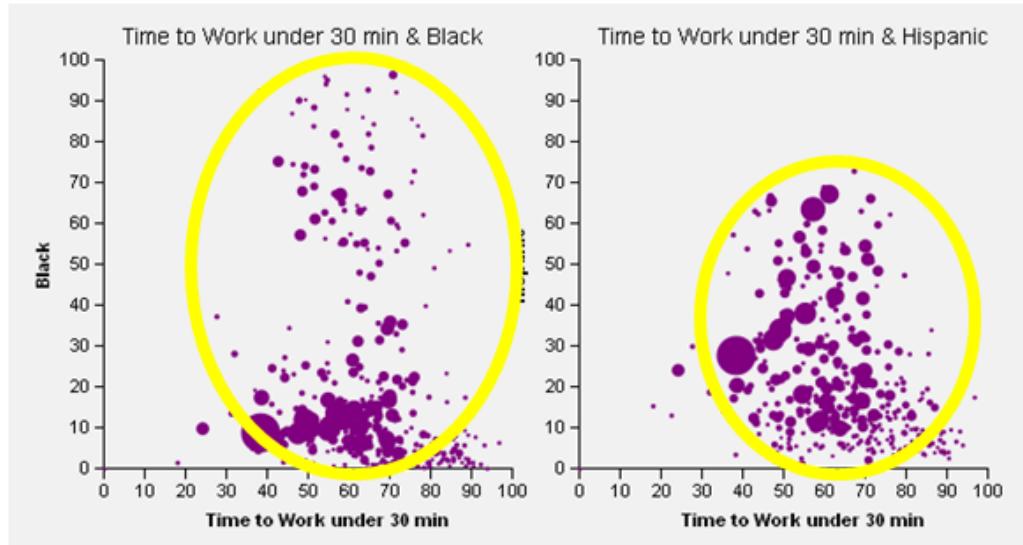
The "center" shows the median values



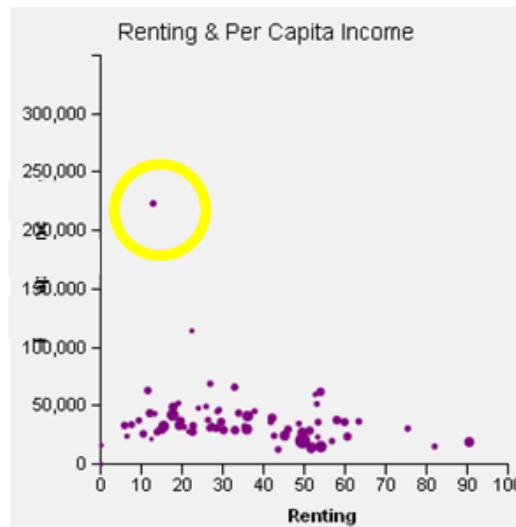
The "direction" shows the trend



The "spread" shows variability. A wider range shows more variability than a compact area.

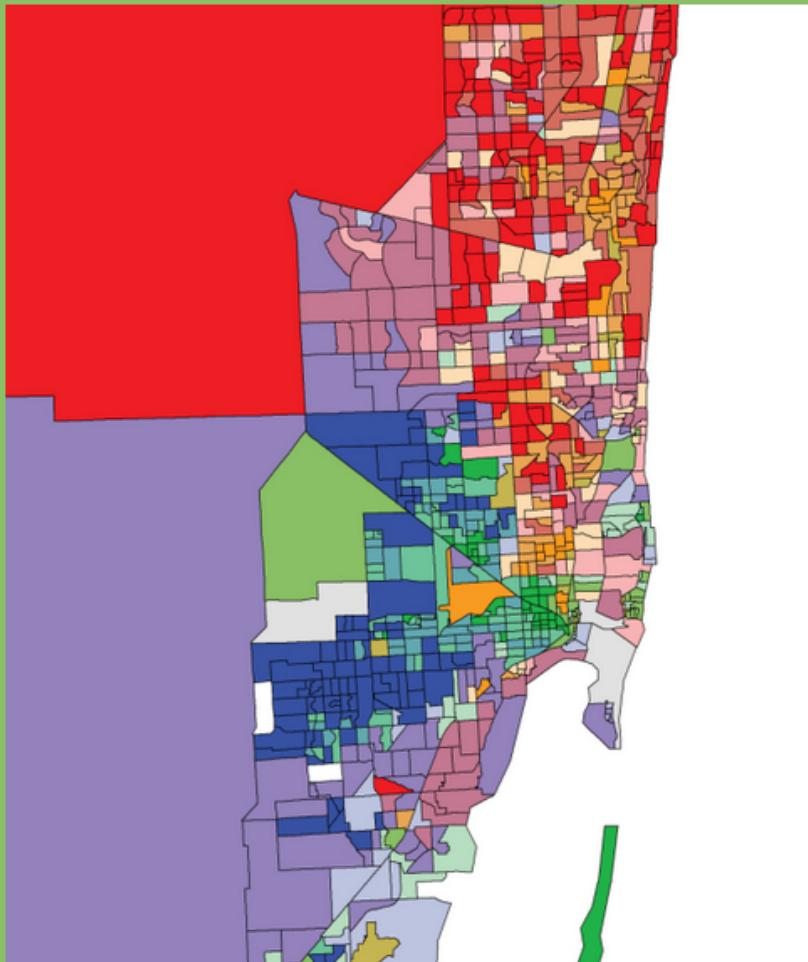


The "outliers" are anomalies. Should they be examined further?

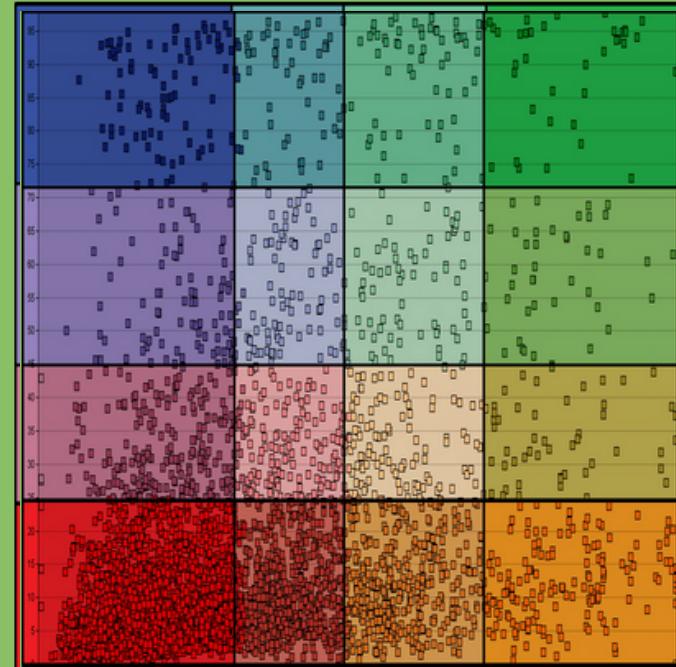


Hispanics and Renters in Miami-Dade County

Legend(s)



Many Hispanic
Fewer Renting



Many Hispanic
Many Renting

Fewer Hispanic
Fewer Renting

Fewer Hispanic
Many Renting

Traditional Legend

Scatterplots Legend

The **traditional** legend shows the colors used and the classification divisions. This representation does not provide any information on data distribution and the map

<http://freac.fsu.edu/scatterplots>

The Project Scatterplots Examples DIY Scatterplot Maker Acknowledgments

DIY Scatterplot Maker - Try it Yourself!

Explore 2010 census block group data in your county with this easy-to-use web tool.

- Geographic Unit is Census Block Group
- Economic Data is Shown in Dollar Values
- All Other Data is Shown by Percentages
- Circle size is scaled according to the total population of the Block Group

County:

I want to visualize:

and -

Race: White African American Asian American Indian Native
 Other Race

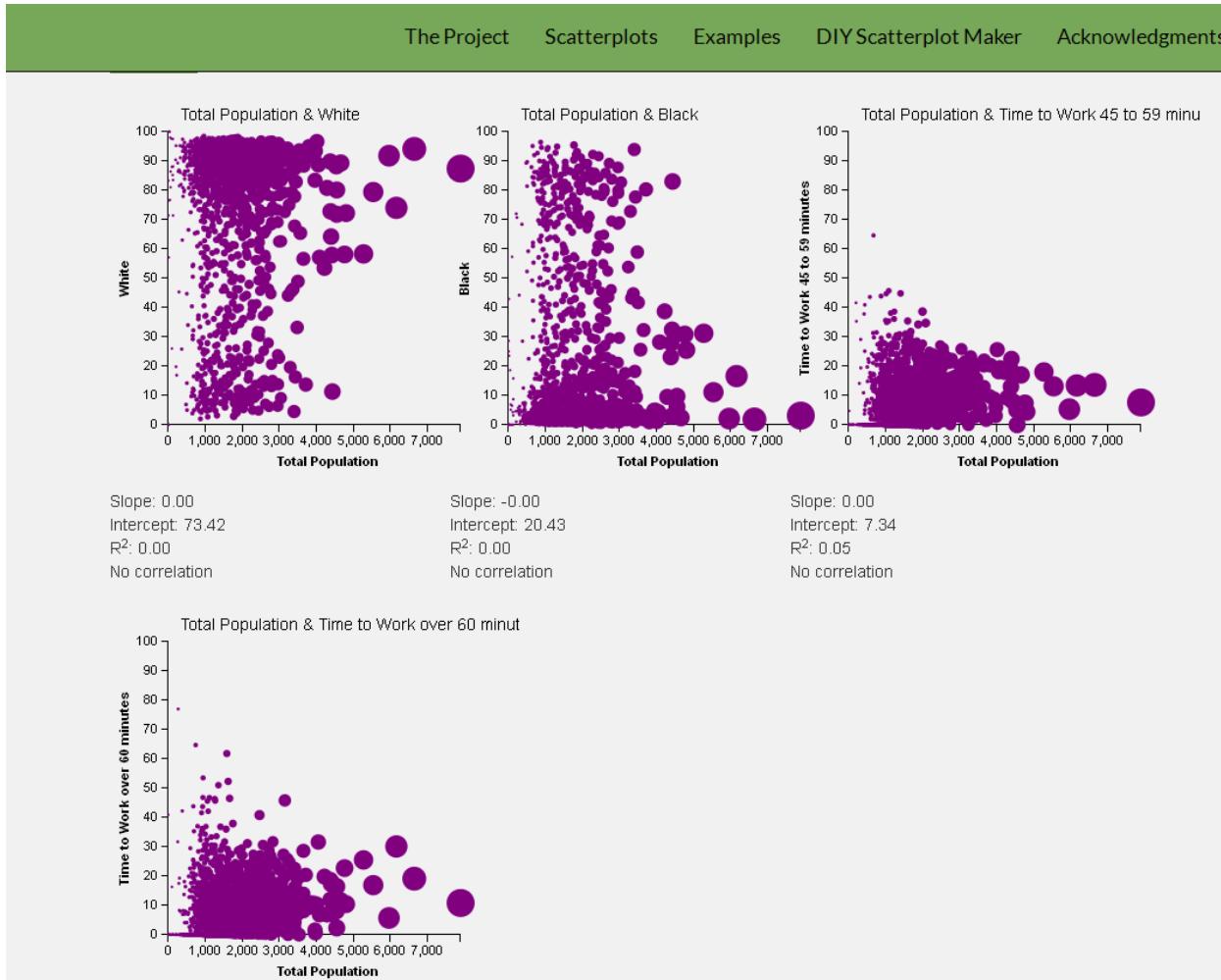
Hispanic: Hispanic Not Hispanic

Education: High School Some College Bachelors Higher Ed

Economics: Per Capita Income Median Housing Value Median Rent

...

Sample Output



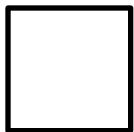
Bivariate Statistical Legends



Case Study Land Use and Land Cover

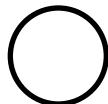
Goal: Visualize both Land Use and Land Cover data on one map.

Land Use data from the DOR and Land Cover data from the NLCD 2011 and is added to the USNG Spatial Data Model for uniformity



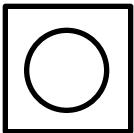
Land Use

Normalized to USNG Grid Cells



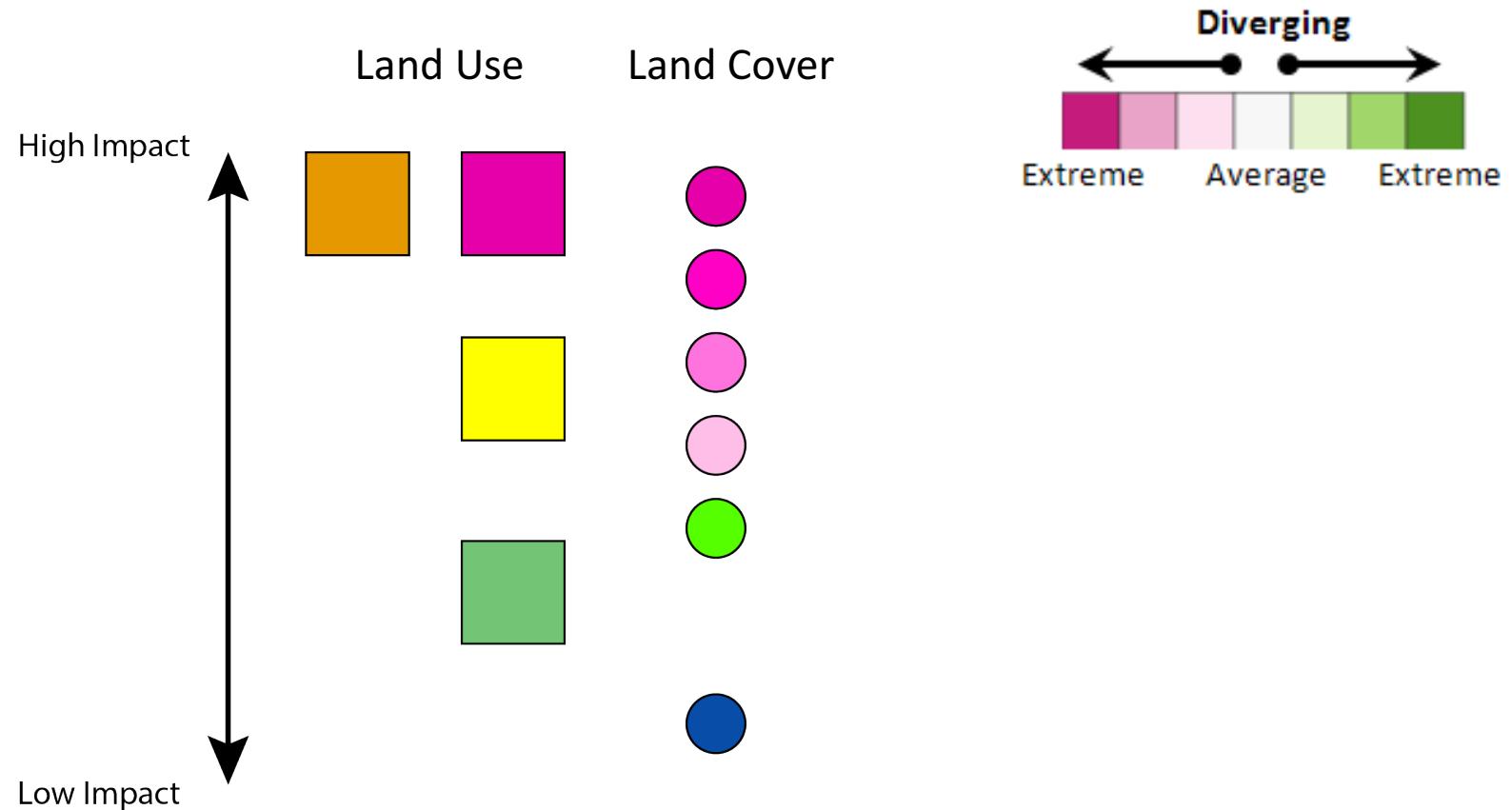
Land Cover

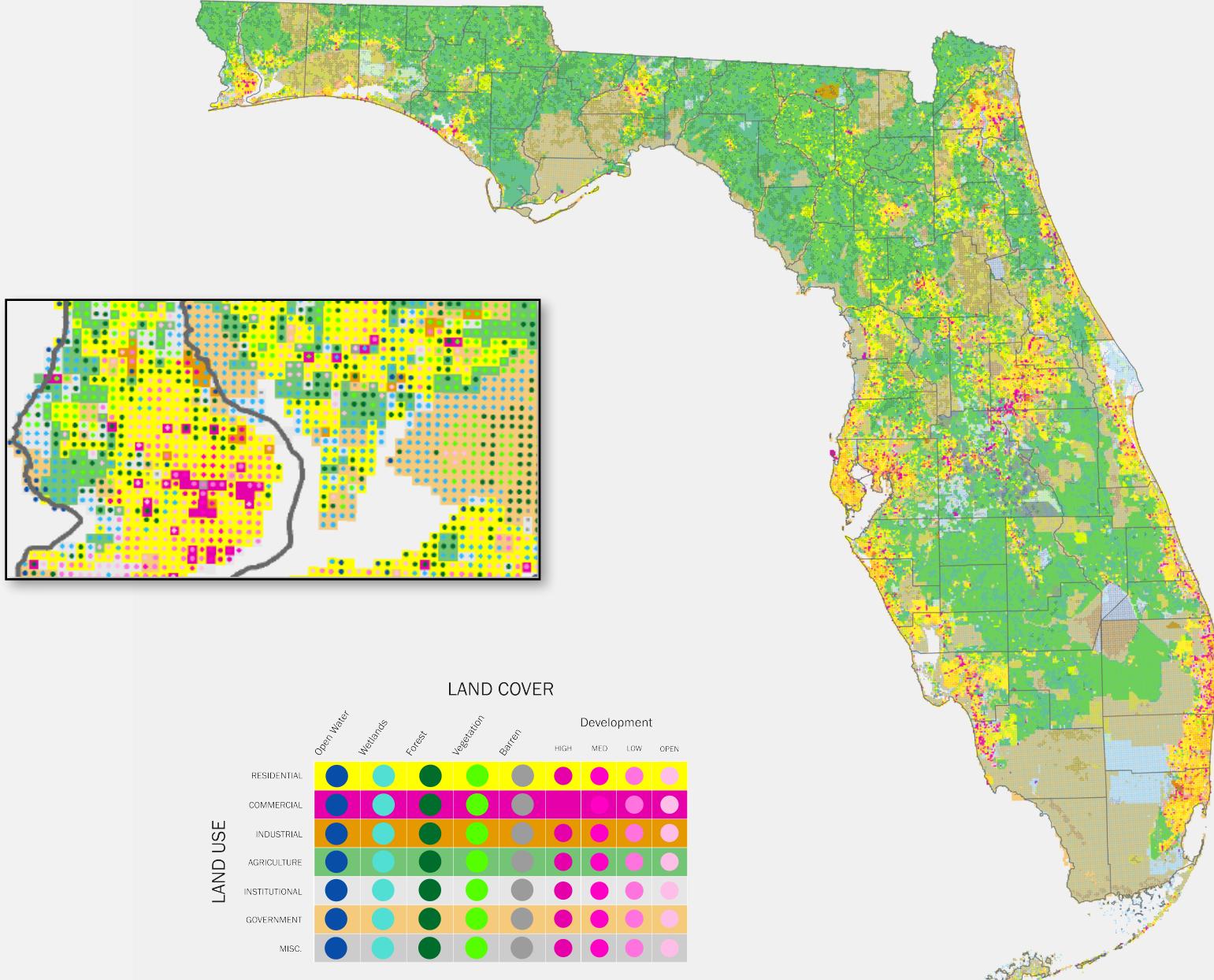
Represented by point in center of each Grid Cell

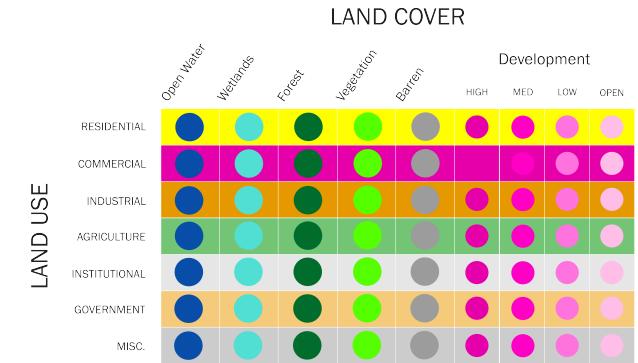
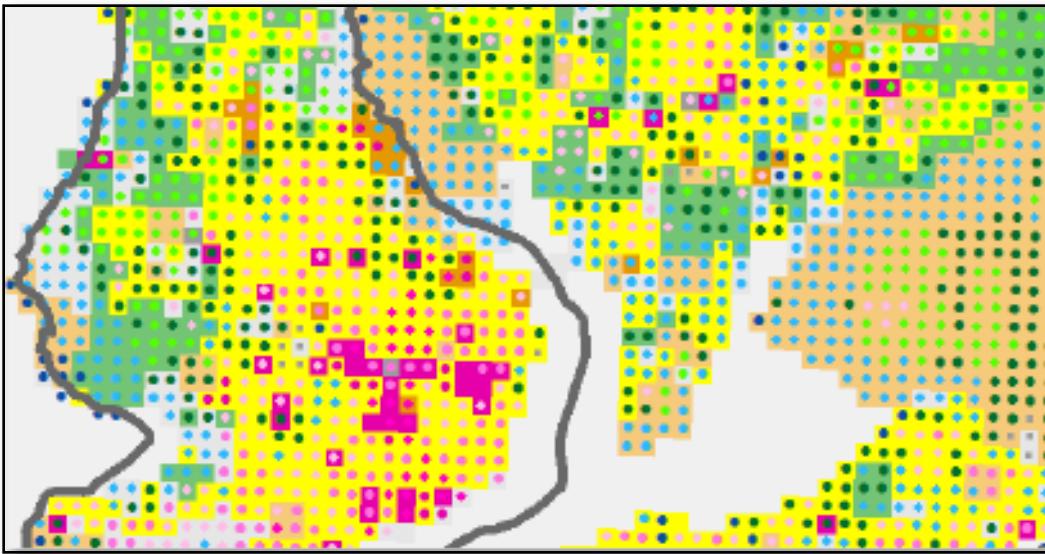


Composite Bivariate Visualization Technique

Case Study Color Choices

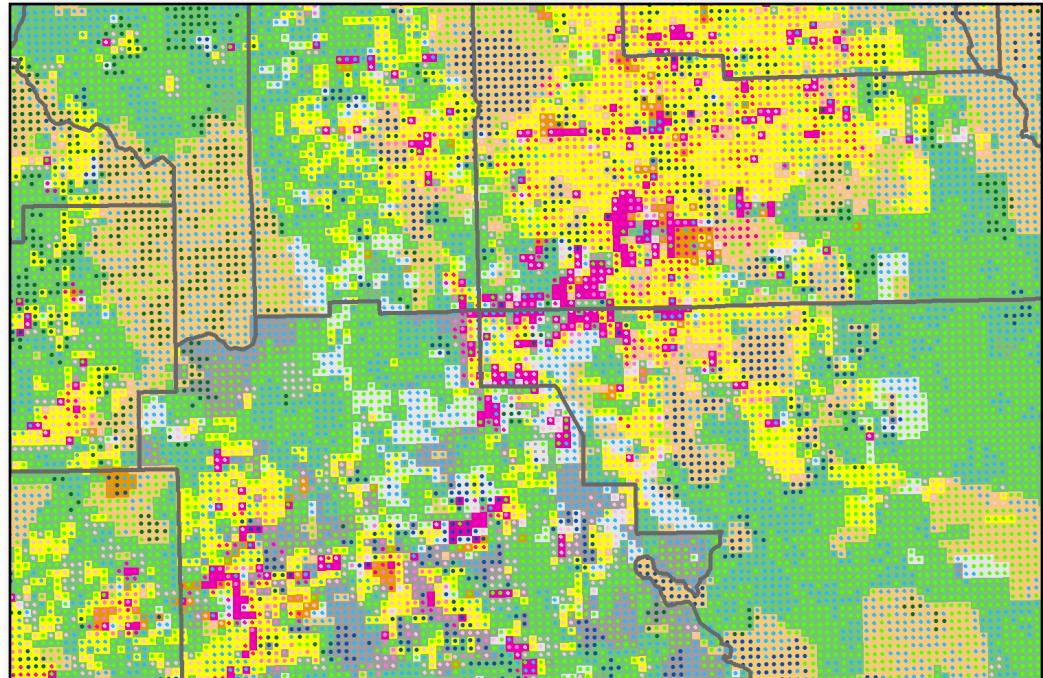






Pensacola, FL Area

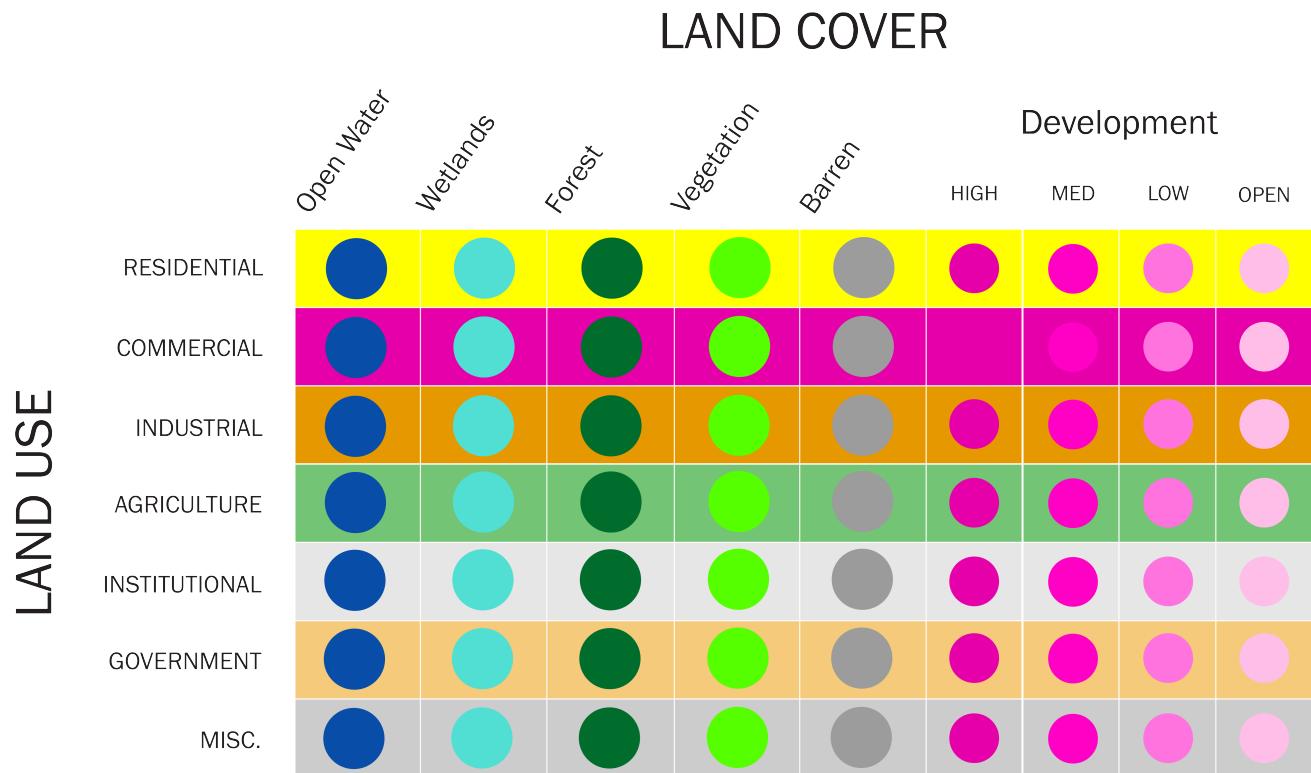
Orlando, FL Area



Visualization 1 Legend

Bivariate visualization

Data is categorical on both axes



Using D3.js

Javascript Programming Library allowing for the manipulation of documents using data.

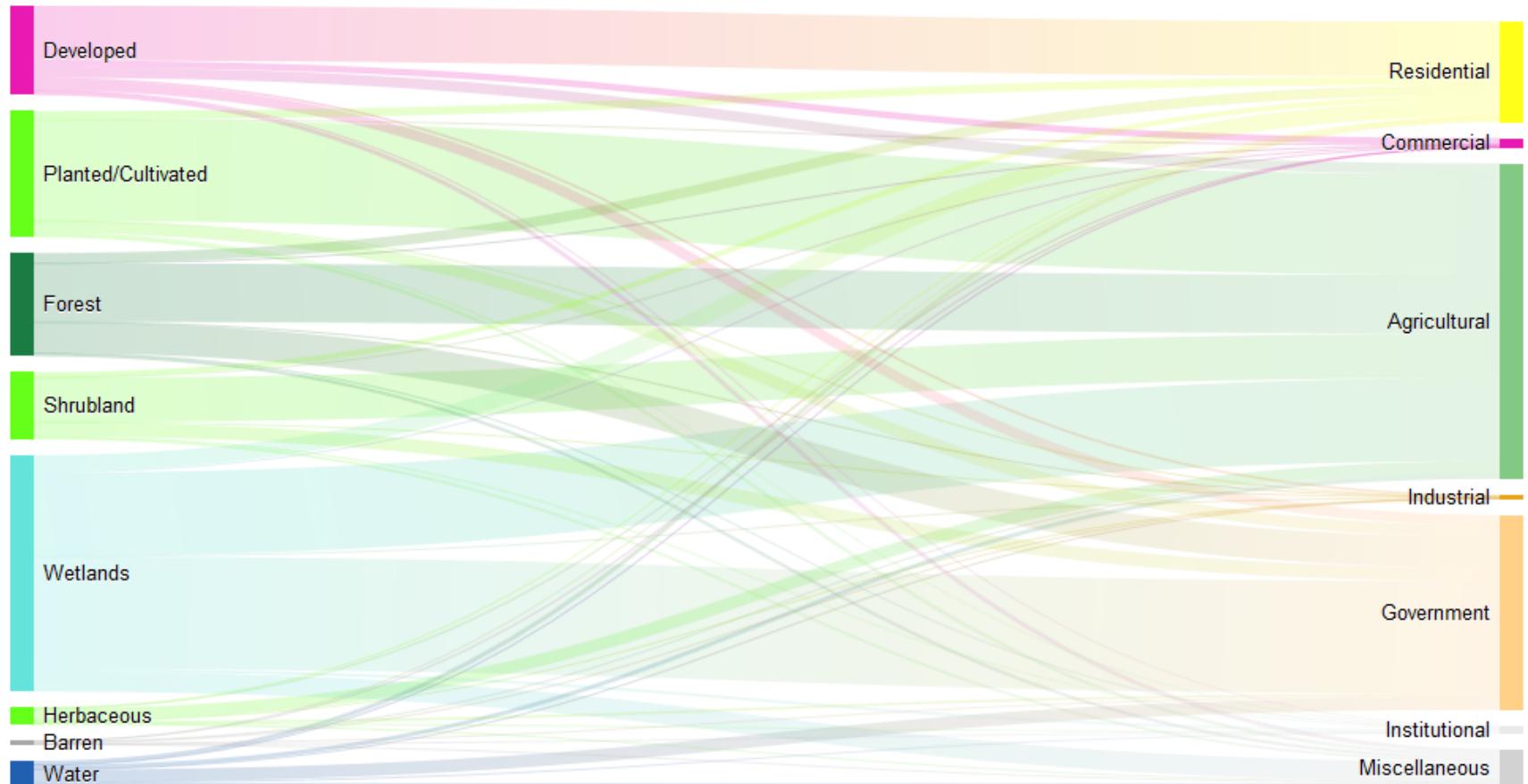
Using a Sankey Visualization and our Land Use/Land Cover data, we are able to us D3.js to draw the flows and interrelationships between all nodes

[D3js.org](http://d3js.org)



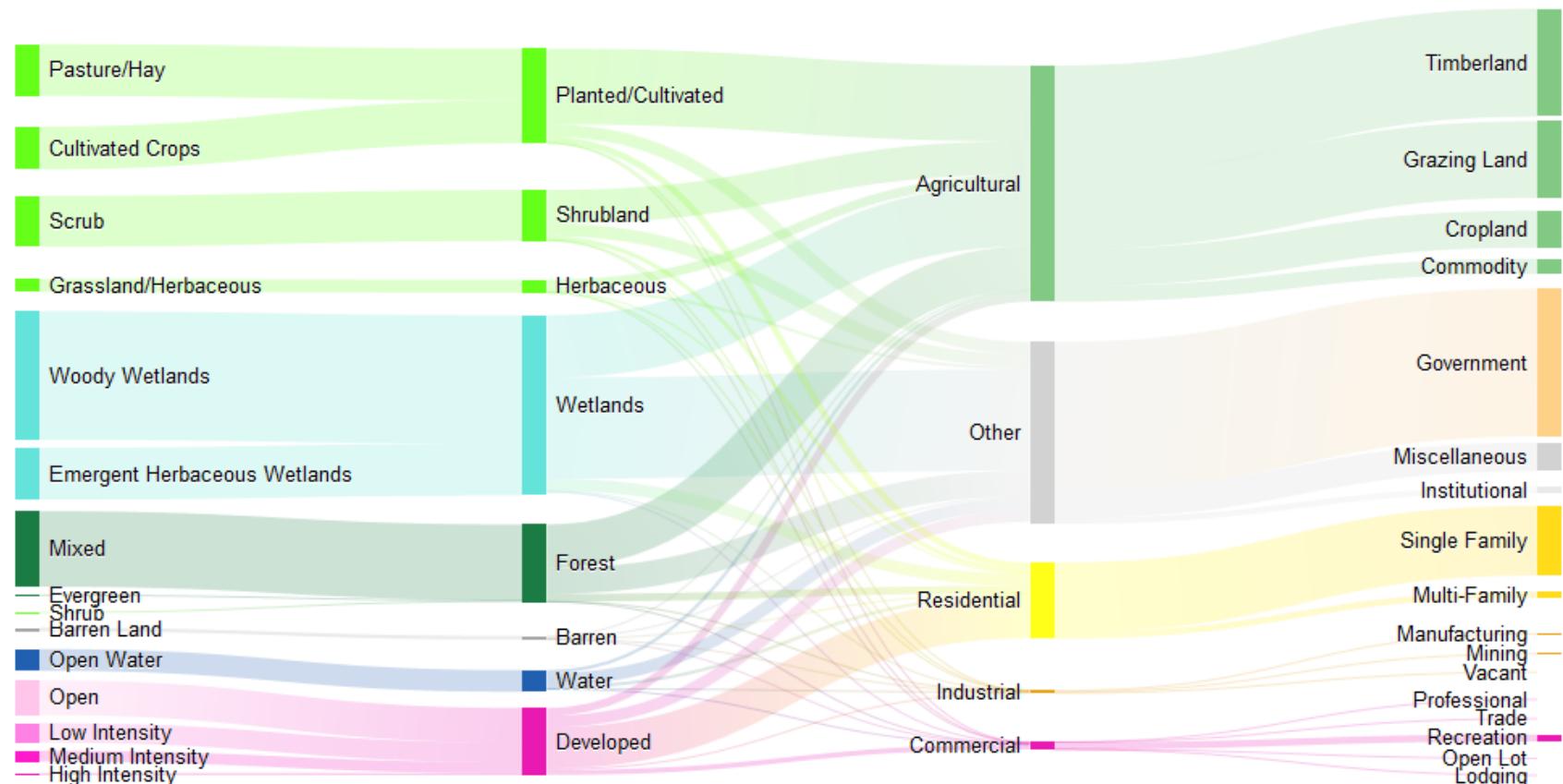
Visualization 2 Sankey Diagrams

Primary Classifications



Visualization 2 Sankey Diagrams

Secondary Classifications



Visualization 3 Statistical Legend

Work in Progress

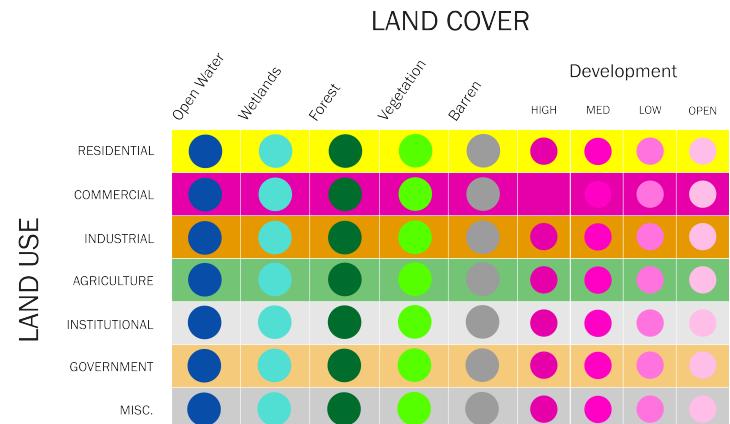
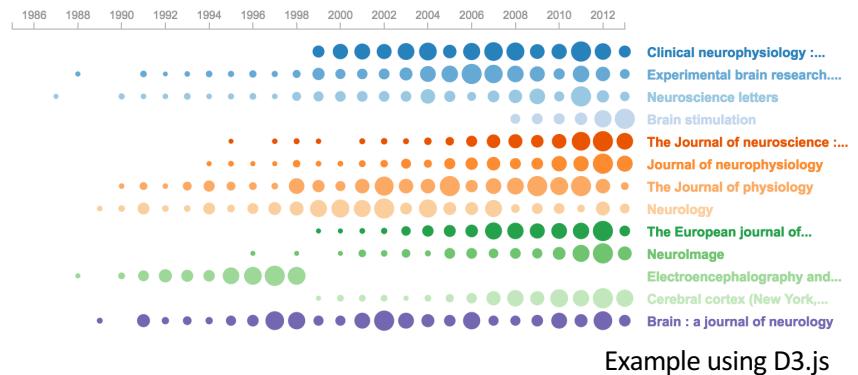
Can't simply add a scatterplot to a legend

Though it is categorical on 2 axes, there is a third dimension!

Legend changes based on land area on map

- Size of circles
- Opacity of circles
- Quantitative numbers within circles (% or #)

Transcranial Magnetic Stimulation



You can use the data!

www.fl-usng-gis.org

- Land Use / Land Cover data available for download
- Guides on the use of the USNG Spatial Data Model

Check out **Georgianna Strode's** Talk at **4:30pm** in Room 114 to learn even more on how you can use it.

Questions or Comments?



Thank You!



Ben Thornton

GIS Researcher/Analyst

Florida Resources and Environmental Analysis Center

✉ bwt13b@my.fsu.edu

🐦 @91degN

💻 91degreesnorth.io