

# GETTING STARTED



Please move to the inside of a row instead of sitting at the end.



Log into the desktop (unless you are using your own laptop) and make sure you know where the **RStudio** icon is (may be on the Desktop or in the Start Menu).



Make sure you have completed **all** of the course onboarding tasks listed on the course website (<https://slu-soc5650.github.io>)

CHRISTOPHER PRENER, PH.D.  
SPRING, 2018

WEEK 02  
LECTURE 01

SOC 4650/5650: INTRODUCTION TO GIS

---

# COURSE INTRODUCTION

# AGENDA

1. Front Matter
2. Introductions
3. What We Learn from Maps
4. Syllabus Overview
5. Defining GIS
6. What is a Workflow?
7. Our First Maps!
8. Back Matter

# 1 FRONT MATTER

## 1. FRONT MATTER

---

# ANNOUNCEMENTS



We'll start every class with "**Front Matter**" - goal is to share what we are covering, what due dates are coming up, and any announcements.



Lecture Prep 02 due **next Monday**; will be posted on Wednesday



If you have not already done so, please download the course data release onto your external storage device (thumb drive or external hard drive).

# **2 INTRODUCTIONS**

## 2. INTRODUCTIONS

---

# ABOUT CHRIS

- ▶ Assistant Professor of Sociology
- ▶ Coordinator, Sociology Honors Thesis
- ▶ Primary faculty mentor for the Computation Geospatial Science Minor (for undergrads)
- ▶ Former EMT and EMS Dispatcher



## 2. INTRODUCTIONS

---

# ABOUT CHRIS

### ▶ Things I teach:

- SOC 1120: Introduction to Sociology - Diversity & Health Emphasis
- SOC 3220: Urban Sociology & The Wire
- SOC 4930/5050: Quantitative Analysis

### ▶ SLU Data Science Seminar



## 2. INTRODUCTIONS

---

# ABOUT CHRIS

### ▶ Things I research:

- Paramedic work and the EMS system as a part of the social safety-net
- Relationship between urban space and first responder work
- Mental health outcomes, literacy, and discrimination
- Approaches to processing “big” and complex data



## 2. INTRODUCTIONS

---

# ABOUT BRANDON

- ▶ Course TA
- ▶ SOC 4650 Spring 2017 Alumni

Need Brandon's info

## 2. INTRODUCTIONS

---

# YOUR TURN!

- ▶ What is your name?
- ▶ What is your program/major?
- ▶ Have you taken GIS class or use GIS software before?
- ▶ Have use used R before?



# 3 WHAT WE LEARN FROM MAPS

ONE CANNOT UNDERSTAND  
SOCIAL LIFE WITHOUT  
UNDERSTANDING THE  
ARRANGEMENTS OF PARTICULAR  
SOCIAL ACTORS IN PARTICULAR  
SOCIAL TIMES AND PLACES. . .  
**SOCIAL FACTS ARE LOCATED.**

Andrew Abbot

■ “Of time and space: the contemporary relevance of the Chicago School”  
(1997)

### 3. WHAT WE LEARN FROM MAPS

---

## VONDERRIT MYERS, JR.



### 3. WHAT WE LEARN FROM MAPS

---

## VONDERRIT MYERS, JR.



### 3. WHAT WE LEARN FROM MAPS

---

# VONDERRIT MYERS, JR.





### 3. WHAT WE LEARN FROM MAPS

---

## A “BAD” CORNER



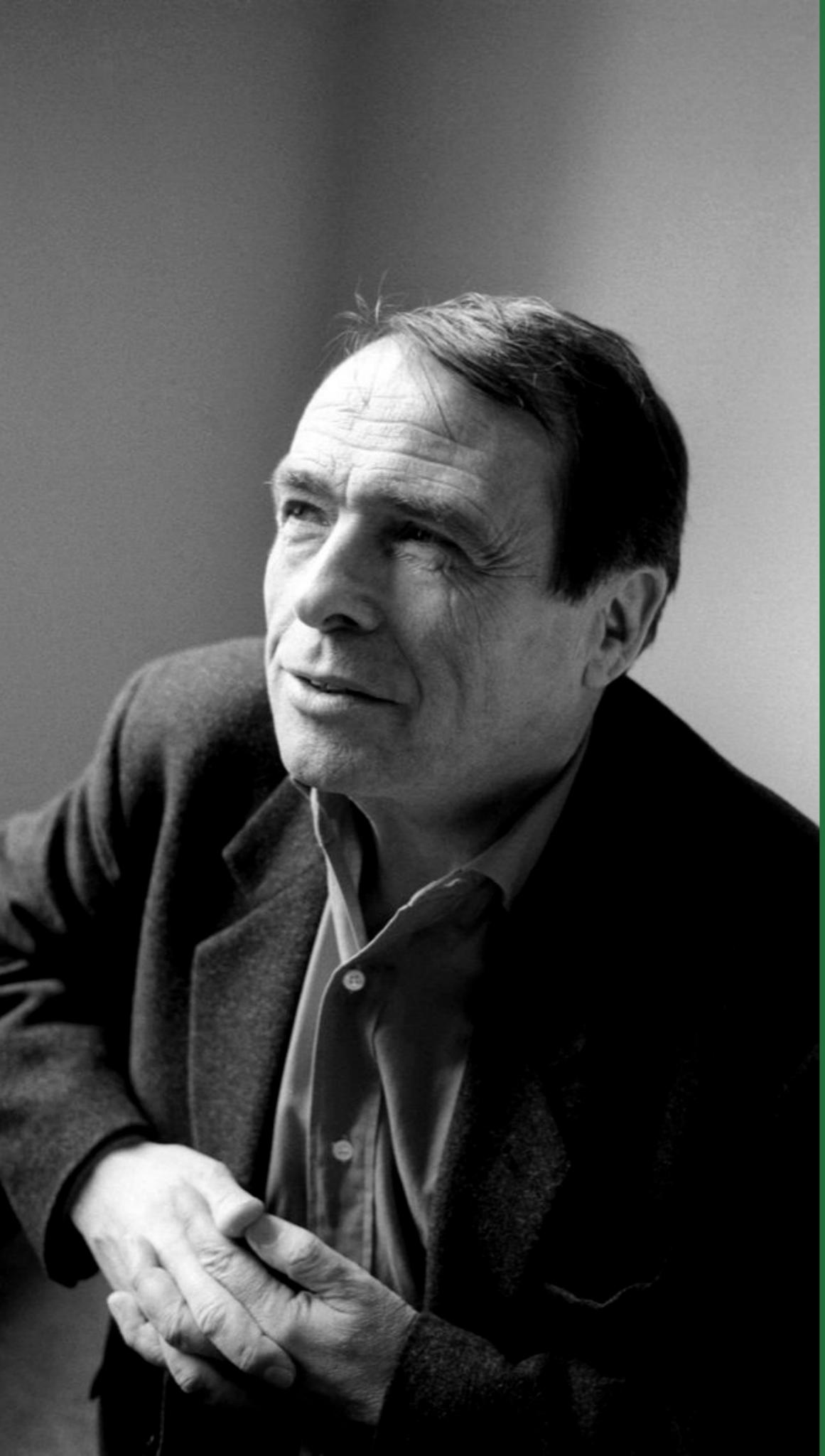
### 3. WHAT WE LEARN FROM MAPS

---

# VONDERRIT MYERS, JR.



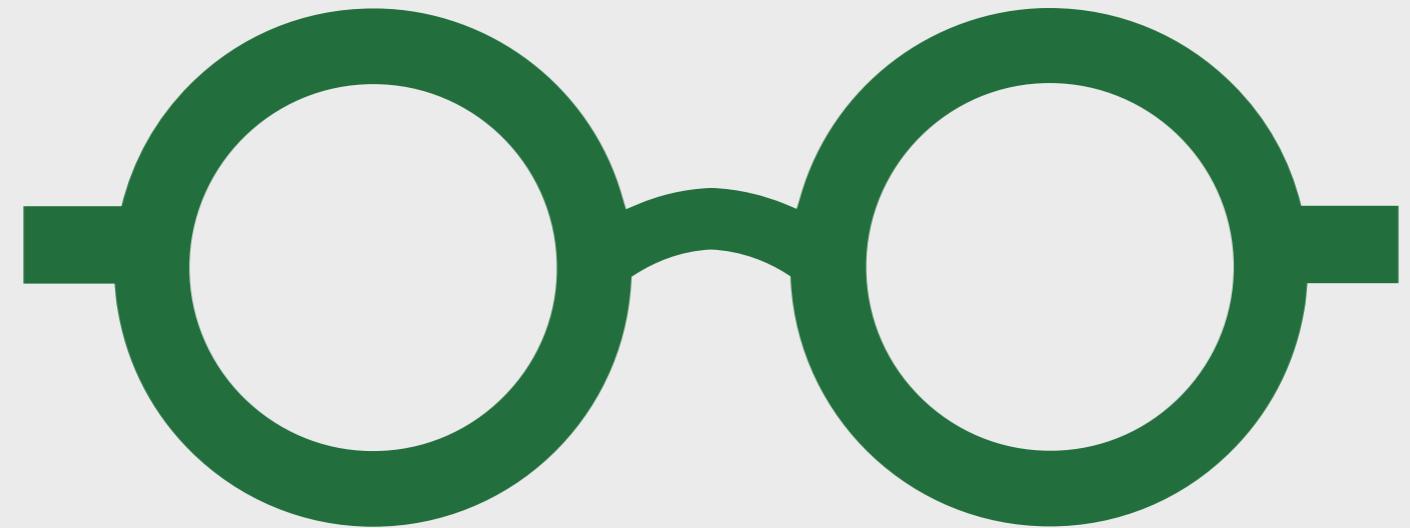
# **4 SYLLABUS OVERVIEW**



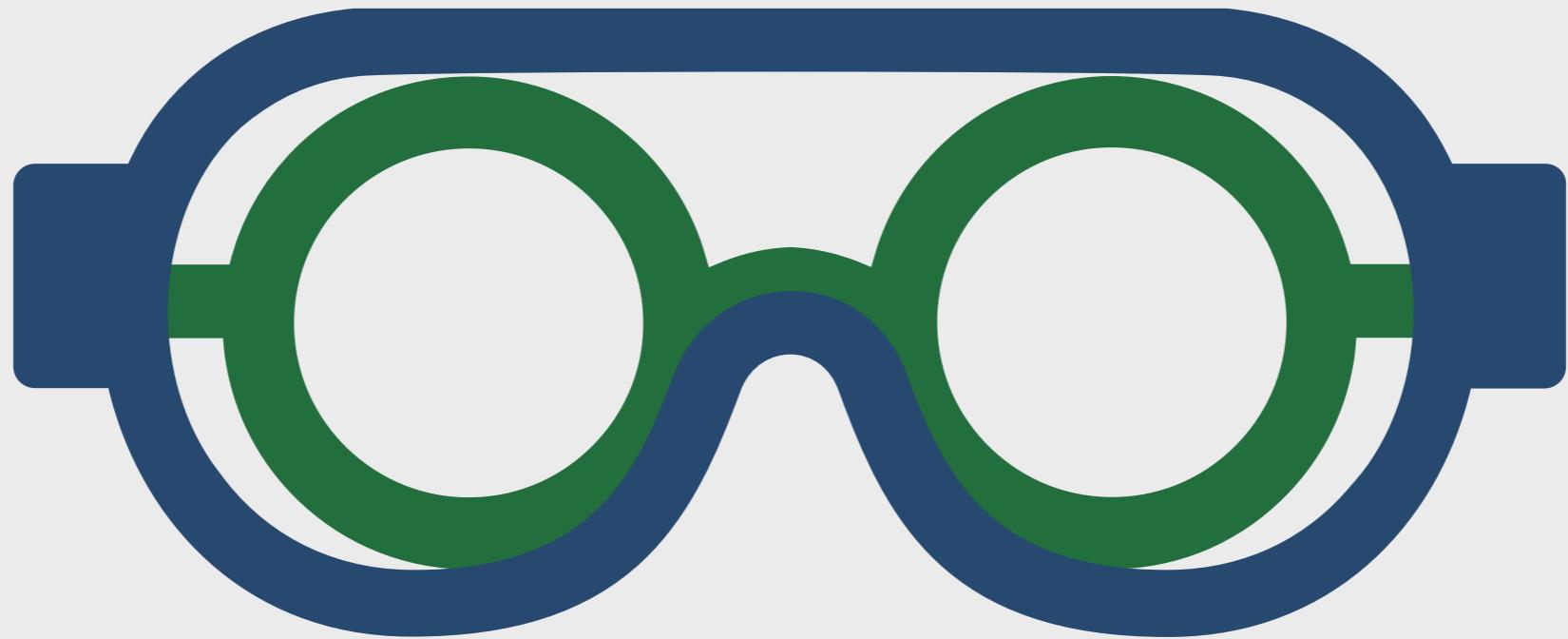
THE FUNCTION  
OF SOCIOLOGY,  
AS OF EVERY  
SCIENCE, IS TO  
REVEAL *empirically*  
THAT WHICH IS  
HIDDEN.

Pierre Bourdieu

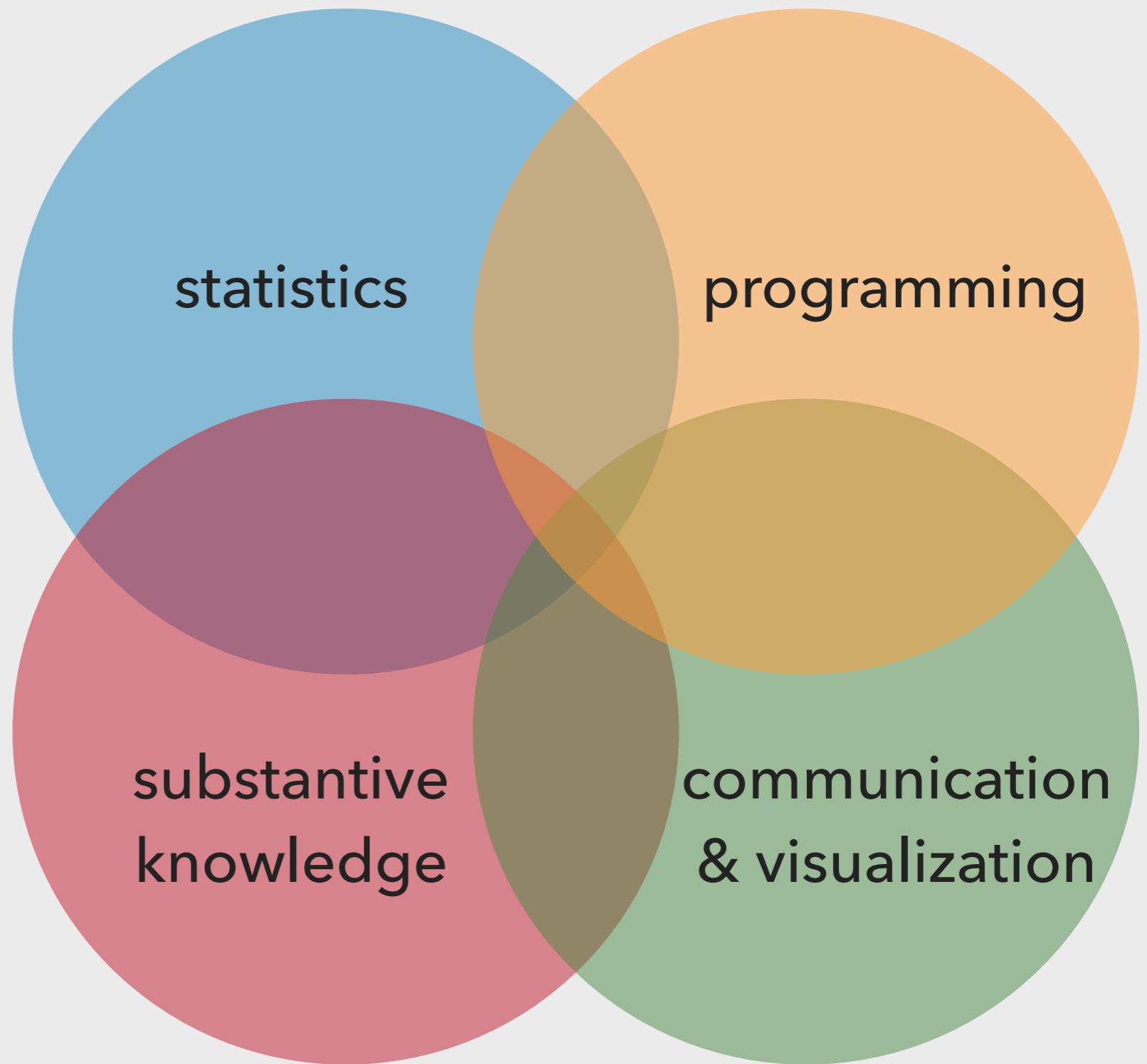
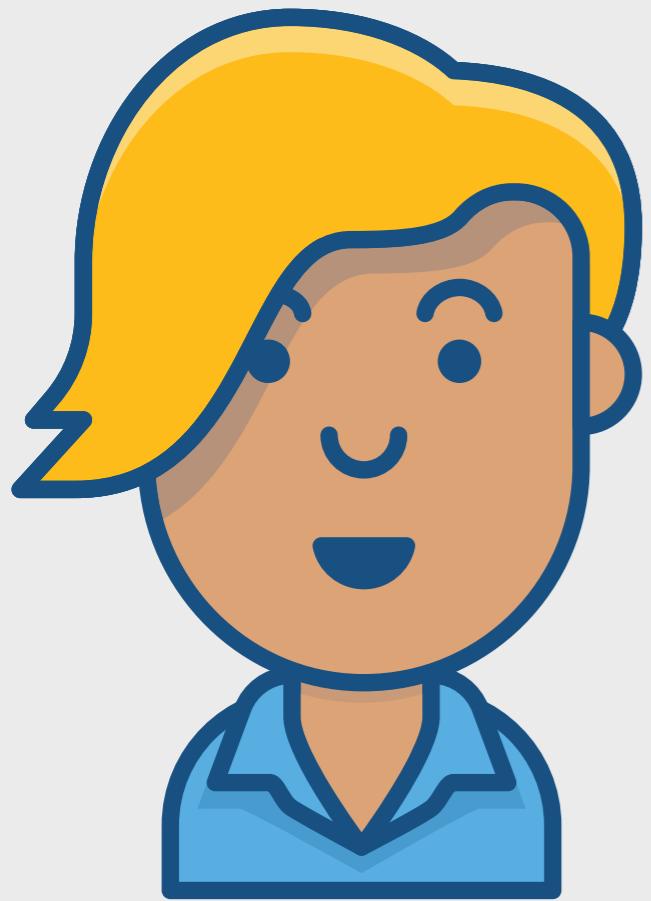
■ On Television  
(1996)



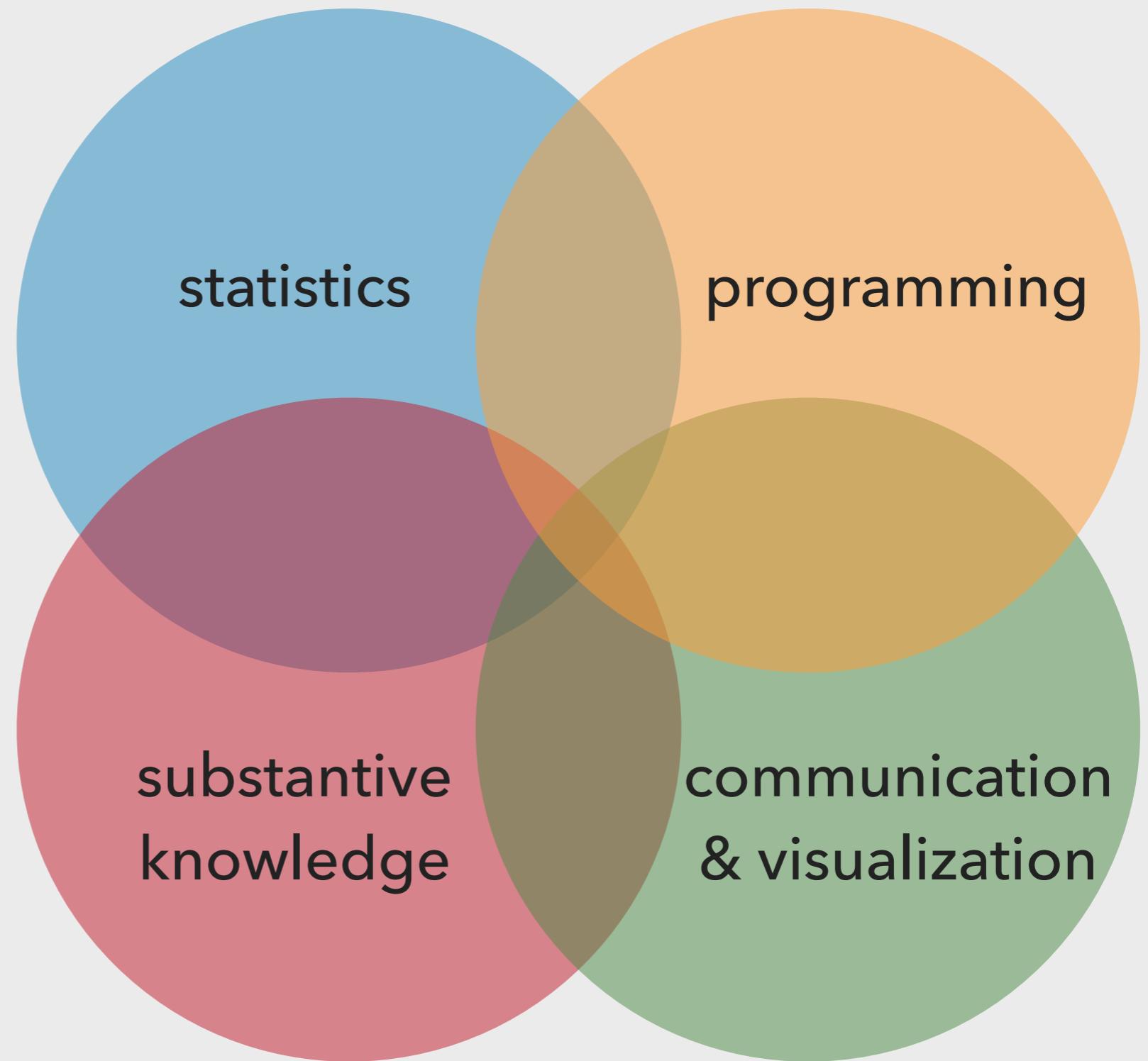
*seeing like a  
social scientist*



*seeing like a  
social scientist,  
thinking like a  
data scientist*



# data scientist



# *geospatial* data scientist

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science
2. Data Management
3. Data Visualization
4. Analysis Development
5. GISC and Public Policy
6. Research Synthesis

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science - Describe the concepts that form the foundation of GISc work.
2. Data Management
3. Data Visualization
4. Analysis Development
5. GISc and Public Policy
6. Research Synthesis

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science
2. Data Management - Perform basic data cleaning tasks using R, construct geo-databases using ArcCatalog for data organization and storage, and modify that data using ArcMap's geoprocessing tools.
3. Data Visualization
4. Analysis Development
5. GISc and Public Policy
6. Research Synthesis

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science
2. Data Management
3. Data Visualization - Create and present visualizations of spatial data using R and ggplot2, ArcMap, and other design tools.
4. Analysis Development
5. GISc and Public Policy
6. Research Synthesis

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science
2. Data Management
3. Data Visualization
4. Analysis Development - Apply techniques that make GISc work more reproducible, accurate, and collaborative using GitHub, R, Markdown, and other tools.
5. GISC and Public Policy
6. Research Synthesis

## 4. SYLLABUS OVERVIEW

---

# COURSE OBJECTIVES

1. Geographic Information Science
2. Data Management
3. Data Visualization
4. Analysis Development
5. GISc and Public Policy - Describe the ways in which public entities support GISc research and the ways in which GISc research supports public policy goals.
6. Research Synthesis

## **4. SYLLABUS OVERVIEW**

---

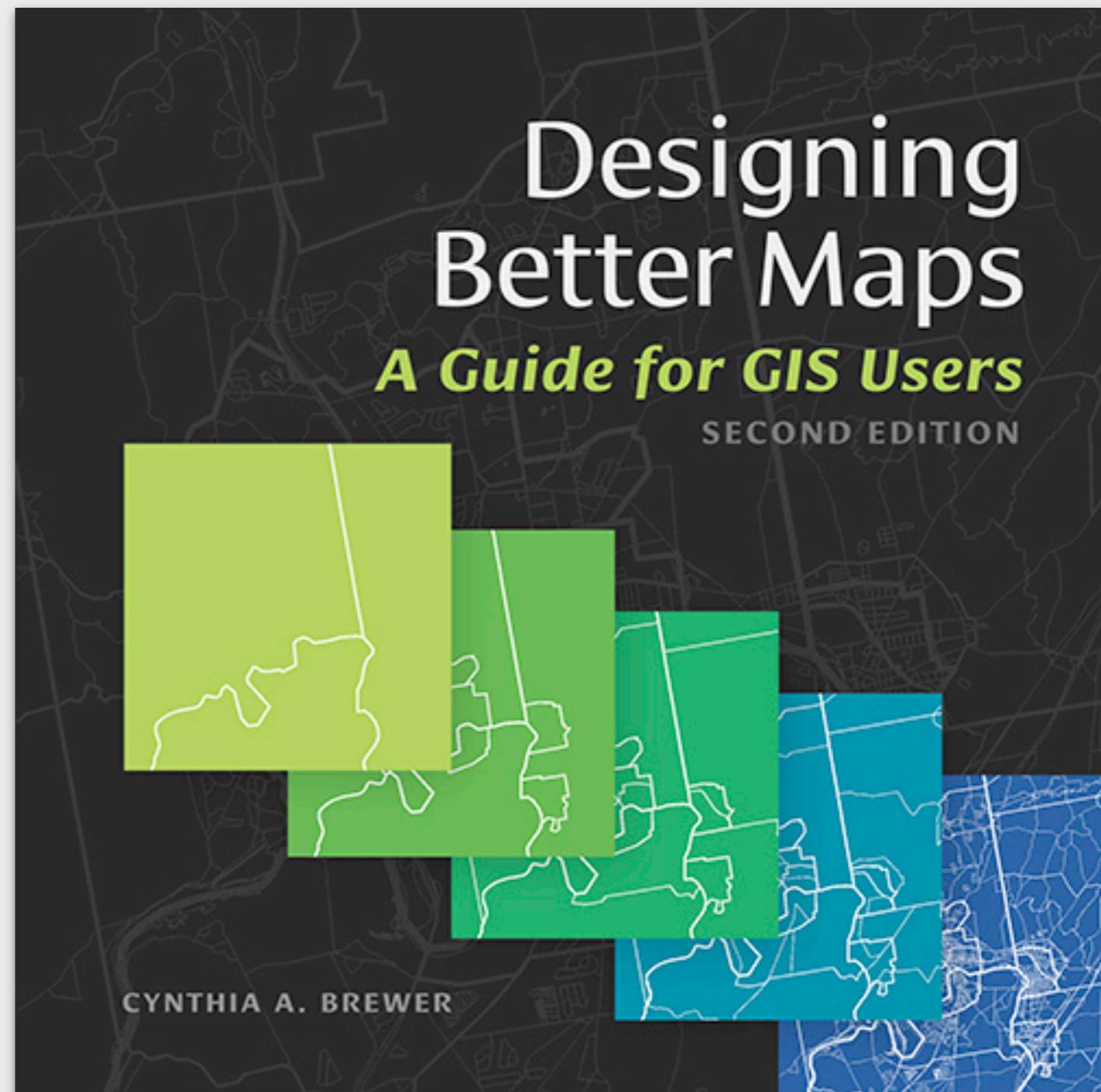
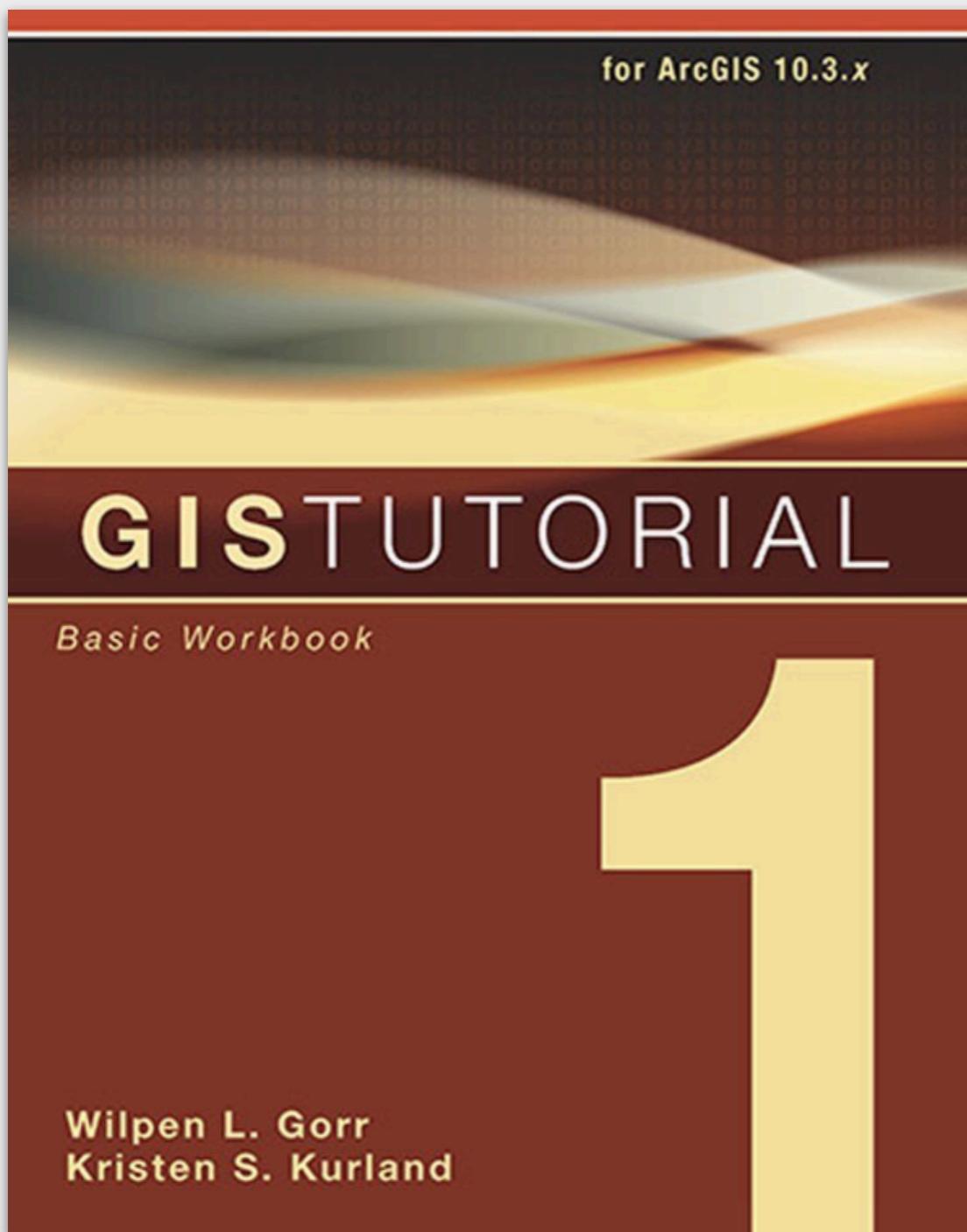
# **COURSE OBJECTIVES**

- 1. Geographic Information Science**
- 2. Data Management**
- 3. Data Visualization**
- 4. Analysis Development**
- 5. GISC and Public Policy**
- 6. Research Synthesis - Plan and implement a spatial data analysis project that utilizes the techniques described throughout the course.**

## 4. SYLLABUS OVERVIEW

---

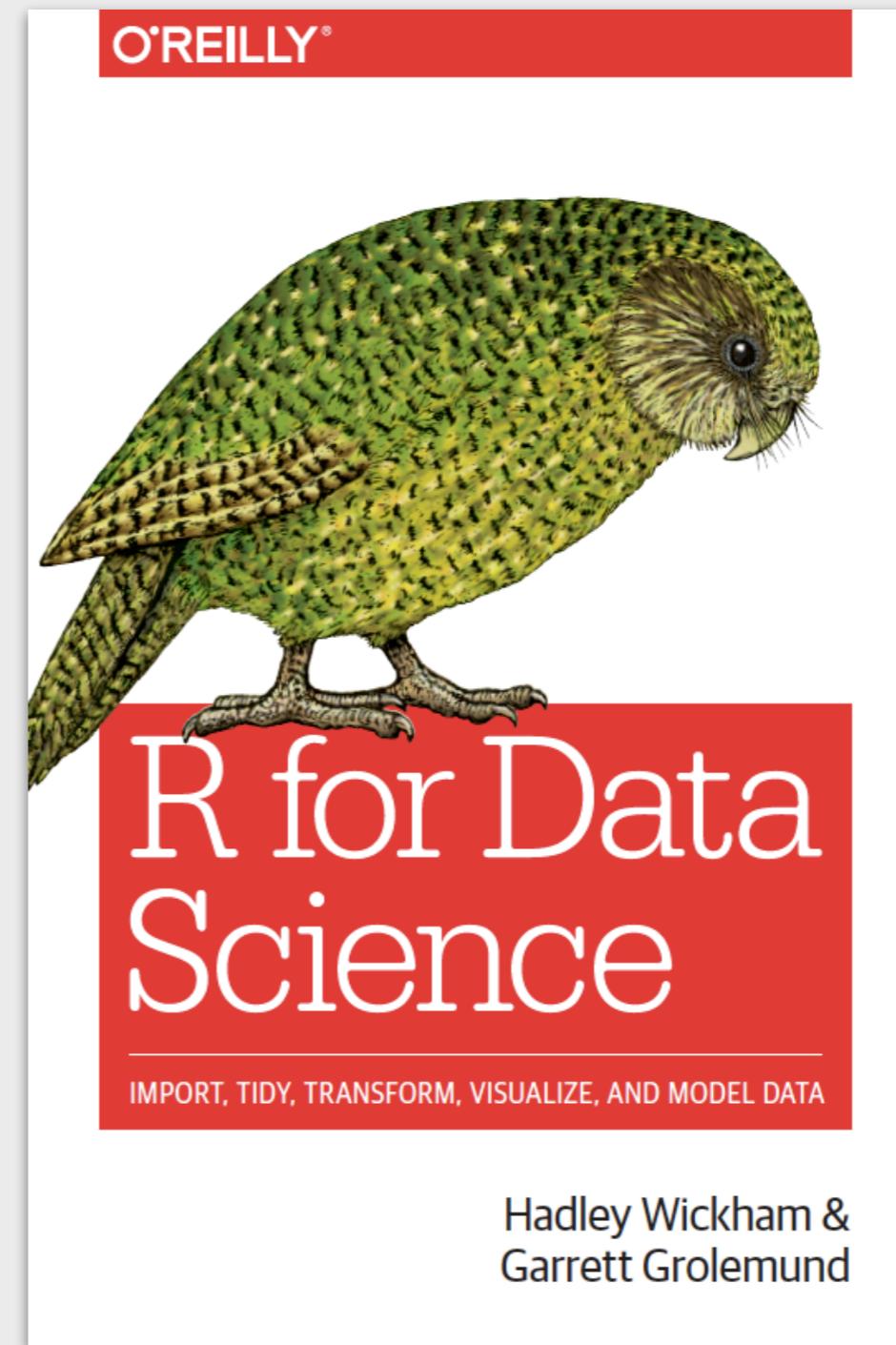
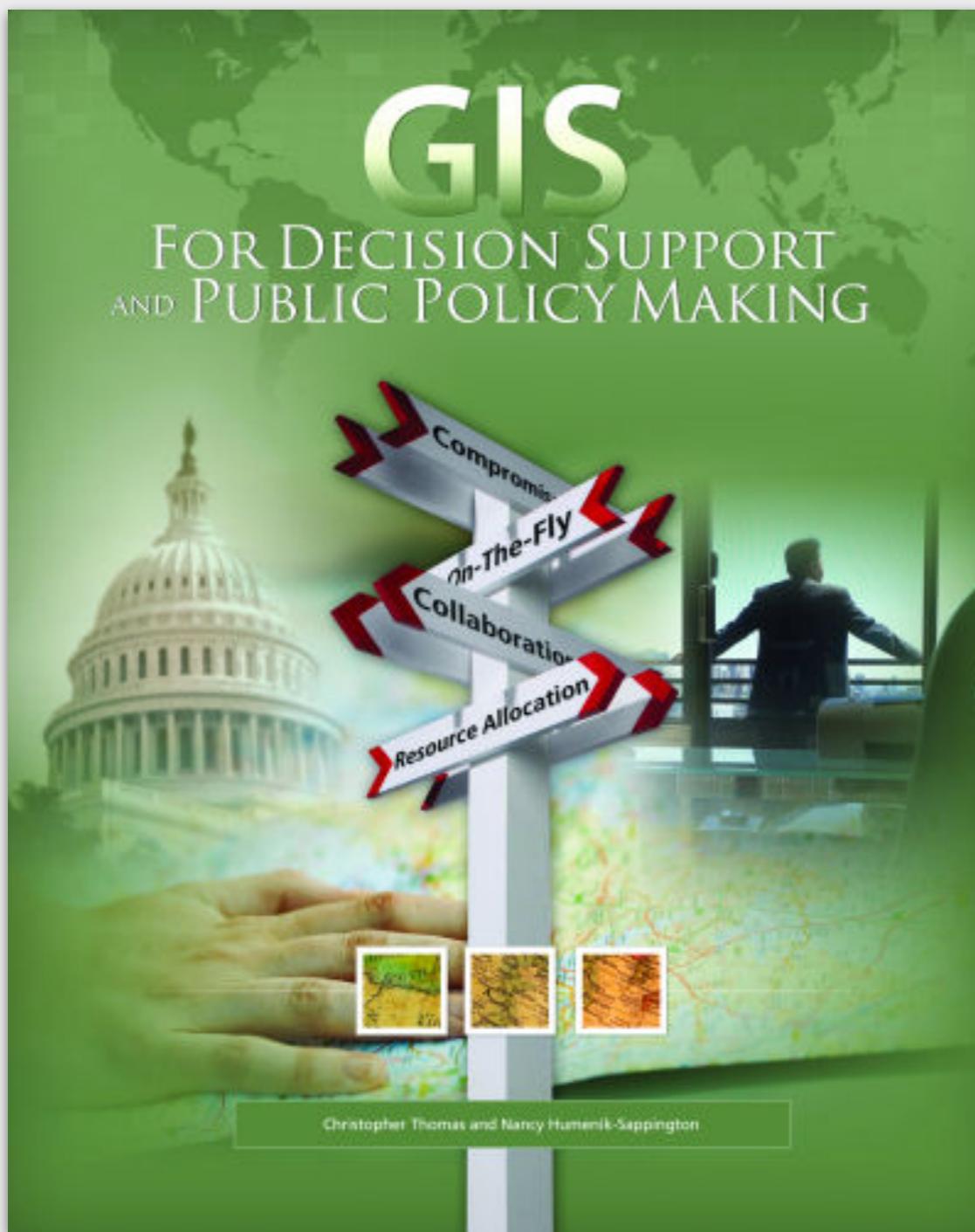
# BOOKS



## 4. SYLLABUS OVERVIEW

---

# BOOKS



## 4. SYLLABUS OVERVIEW

# READINGS

<https://chris-prener.github.io/SSDSBook/>

The screenshot shows a web browser window with the URL <https://chris-prener.github.io/SSDSBook/> in the address bar. The page content is titled "Sociospatial Data Science" by Christopher Prener, Ph.D., dated 2018-01-06. It features three hexagonal icons: a blue one for "Sociospatial Data Science" with a flask icon, a green one for "SOC 4650 & 5650" with a location pin icon, and an orange one for "SOC 4930 & 5050" with a calculator icon. The left sidebar lists chapters from "Preface" to "Advanced Git". A footer note at the bottom states: "This text is a companion text for both of my research methods courses at Saint Louis University:".

Sociospatial Data Science

Preface

License

1 Introduction

I First Steps

2 Approaching These Courses

3 "Good Enough" Research Practices

4 Protecting Your Work

5 Getting Help

II Data Science Toolkit

6 Opinionated Tools

7 Markdown

8 Basic Git

9 Advanced Git

Published with bookdown

**Sociospatial Data Science**

*Christopher Prener, Ph.D.*

2018-01-06

**Preface**

Sociospatial Data Science

SOC 4650 & 5650

SOC 4930 & 5050

This text is a companion text for both of my research methods courses at Saint Louis University:

## 4. SYLLABUS OVERVIEW

# READINGS

<http://eres.slu.edu>

The screenshot shows a web browser window displaying the Saint Louis University Ereserves course reserves page. The URL in the address bar is [eres.slu.edu](http://eres.slu.edu). The main title is "SAINT LOUIS UNIVERSITY". Below the title, course information is listed: "SOC5650 - Introduction to Geographic Information Science (Spring 2018) - Prener" and "SOC4650 - Introduction to Geographic Information Science (Spring 2018) -". There are three tabs at the top: "Course Info" (selected), "Documents", and "Page Management". Below the tabs are links: "Open/Close All Folders", "Download All As Zip", and "Help Opening Documents". A table lists course materials organized by week:

Title	Primary Author	Format	Size (KB)
Week 01		Folder	-
Week 02		Folder	-
Week 03		Folder	-
Week 04		Folder	-
Week 05		Folder	-
Week 06		Folder	-
Week 10		Folder	-
Week 12		Folder	-

Questions? Comments? Please contact the system managers.  
Docutek ERes v5.7 - © 2000-2018 Docutek, a SirsiDynix Company.  
Docutek ERes and DocuFax are trademarks of Docutek, a SirsiDynix Company.

## 4. SYLLABUS OVERVIEW

# WEBSITE

<https://slu-soc5650.github.io>

The screenshot shows a web browser displaying the course website. The address bar shows the URL <https://slu-soc5650.github.io>. The page has a green header bar with the text "SLU SOC 4650 & SOC 5650". On the left sidebar, there is a logo for "SOC 4650 & 5650" and a "Core Documents" section. Below the sidebar, there is a blue button labeled "DOWNLOAD SYLLABUS". The main content area features a section titled "Introduction to Geographic Information Science" with a "Enrollment Open" status. This section contains text about the course being taught at Saint Louis University for Spring 2018. To the right of this section is a green hexagonal icon with a location pin and the text "SOC 4650 & 5650". Below the main content area, there are sections for "Course Description", "Course Objectives", and "Topic Index".

SLU SOC 4650 & SOC 5650

SOC 4650 & 5650  
slu-soc5650/Core-Documents

[DOWNLOAD SYLLABUS](#)

**About**

- Course Description
- Course Objectives
- Core Documents
- Acknowledgements

**Course Resources**

- Course Onboarding
- Course Preview
- Lecture 01 - Course Introduction
- Final Project
- Topic Index
- Package Index
- License

**Introduction to Geographic Information Science**

**Enrollment Open**

SOC 4650 & 5650 is the Department of Sociology and Anthropology's undergraduate and graduate introductory GIS course at Saint Louis University. It will next be taught for Spring 2018 and there is currently **one open seat**. If you have an interest in the course, you are encouraged to enroll **as soon as possible** - enrollment will close soon! This site is for sections taught by Assistant Professor Christopher Prener.

**Course Description**

This class introduces both the theoretical and technical skills that constitute the nascent field of Geographic Information Science (GISc). Techniques introduced include data cleaning and management, map production and cartography, and the manipulation of both tabular and spatial data. The impacts of GISc on public policy, and the effects of public policy on GISc, are also discussed. The course incorporates a wide variety of social, economic, health, urban, meteorological, and environmental data. These data are mapped at a variety of extents, from the City of St. Louis to the St. Louis Metropolitan region, Missouri, all United States counties, and all U.S. states.<sup>1</sup>

**Course Objectives**

This course has six intertwined objectives:

SOC 4650 & 5650

## 4. SYLLABUS OVERVIEW



<https://github.com/slu-soc5650>

The screenshot shows the GitHub organization page for `slu-soc5650`. At the top, there's a header bar with various icons and links. Below it, the organization's name and a green hexagonal profile picture are displayed. The profile picture contains a location pin icon and the text "SOC 4650 & 5650". The organization's name is "Saint Louis University - Introduction to Geographic Information Science". Below the name, it says "SOC 4650/5650 Course Repositories". There are links for "St. Louis, MO", "https://slu-soc5650.git...", and an email address "chris.prener@slu.edu". A navigation bar below the profile picture includes "Repositories 88" (which is highlighted), "People 12", "Teams 34", "Projects 0", and "Settings".

**Pinned repositories**

- Core-Documents**  
Course Syllabus and Reading List
- finalProject**  
Citizens' Service Bureau Data Analysis  
• Stata • 1

Customize pinned repositories

Search repositories... Type: All Language: All New

**slu-soc5650.github.io**  
Content for Course Website  
• JavaScript Updated 4 days ago

**Core-Documents**  
Course Syllabus and Reading List

Top languages

- Stata • HTML • Shell
- JavaScript • Python

People 12 >

## 4. SYLLABUS OVERVIEW



<https://slu-soc5650.slack.com>

slu-soc5650 Chris Prener All Threads Channels

#\_news 22 2 Course-wide announcements

Dropbox [SOC5650 Data Release.zip](#) Shared with Dropbox 1

Second, the Electronic Reserves site is now functional but needs some additional attention - I wasn't able to get in and organize it to reflect this semester's syllabus because of the server issues after the fire. The Week-01 readings are in there, though you need to be careful to read the correct John Logan article. Double check the Reading List! You can skip his other article and the Intro from J. Scott Long's book since they are not listed on this year's Reading List. I will get this cleaned up this weekend.

Speaking of this weekend - I am heading to a wedding in New York City tomorrow and will have more limited presence on Slack and email. Thanks for your patience - I'll be back in full force on Monday!

Finally, I just wanted to remind everyone that the course onboarding checklist needs to be completed by Monday. <https://slu-soc5650.github.io/getting-started/>

[slu-soc5650.github.io](#) **SLU SOC 4650 & SOC 5650** SOC 4650 & SOC 5650: Introduction to Geographic Information Science SOC 4650 & SOC 5650

1 This includes the course preview videos, Lecture Prep 01, and the readings for Lecture 01!

One more announcement for today - the password for the Electronic Reserves is "space"!

Diana 5:19 PM joined #\_news along with 5 others.

Thursday, January 18th

**Chris Prener** 1:51 PM Reminder - the password for electronic reserves:

**Chris Prener** One more announcement for today - the password for the Electronic Reserves is "space"! Posted in #\_news | Jan 17th 1

**Carter Alcock** 2:13 PM joined # news along with 5 others.

## 4. SYLLABUS OVERVIEW

---

# SOFTWARE



ArcGIS



R Programming  
Language

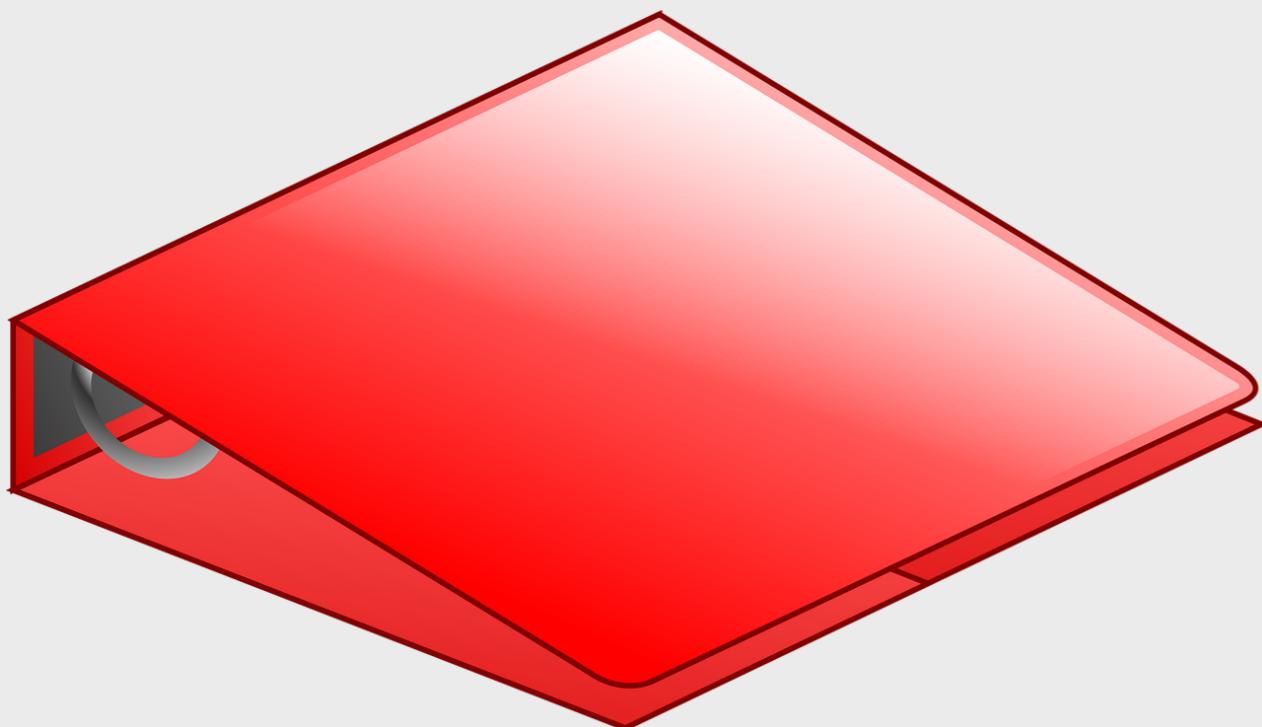


RStudio

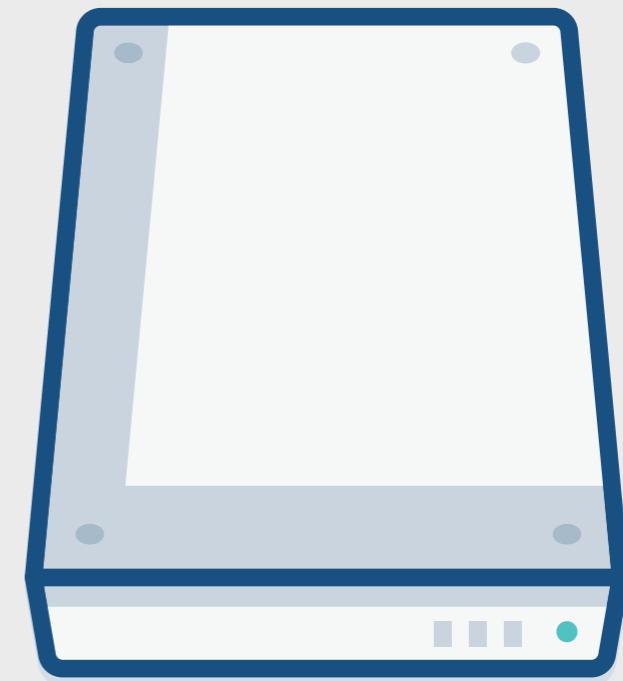
## 4. SYLLABUS OVERVIEW

---

# OTHER MATERIALS



Binder



External storage

## 4. SYLLABUS OVERVIEW

---

# ASSIGNMENTS

1. Participation	10%		100
2. Lecture Preps	6%	$4 * 15 =$	60
3. Labs	15%	$10 * 15 =$	150
4. Problem Sets	28%	$35 * 8 =$	280
5. Final Project	41%		410
			<hr/> <hr/> 1,000

## 4. SYLLABUS OVERVIEW

---

# GRADING

1. Participation ●

2. Lecture Preps ●

3. Labs ●

4. Problem Sets ○

5. Final Project ○



## 4. SYLLABUS OVERVIEW

---

# GRADING

1. Participation ○

2. Lecture Preps ●

3. Labs ●

4. Problem Sets ●

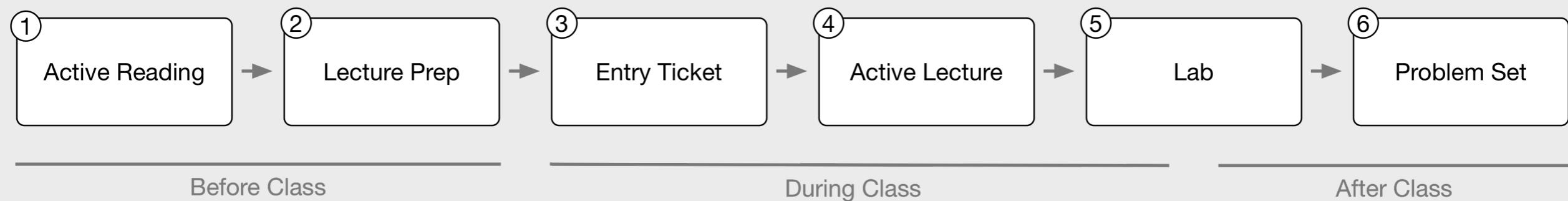
5. Final Project ●

within 24-hours	-15%
24 to 48-hours	-30%
48 to 72-hours	-45%
> 72-hours	-100%

## 4. SYLLABUS OVERVIEW

---

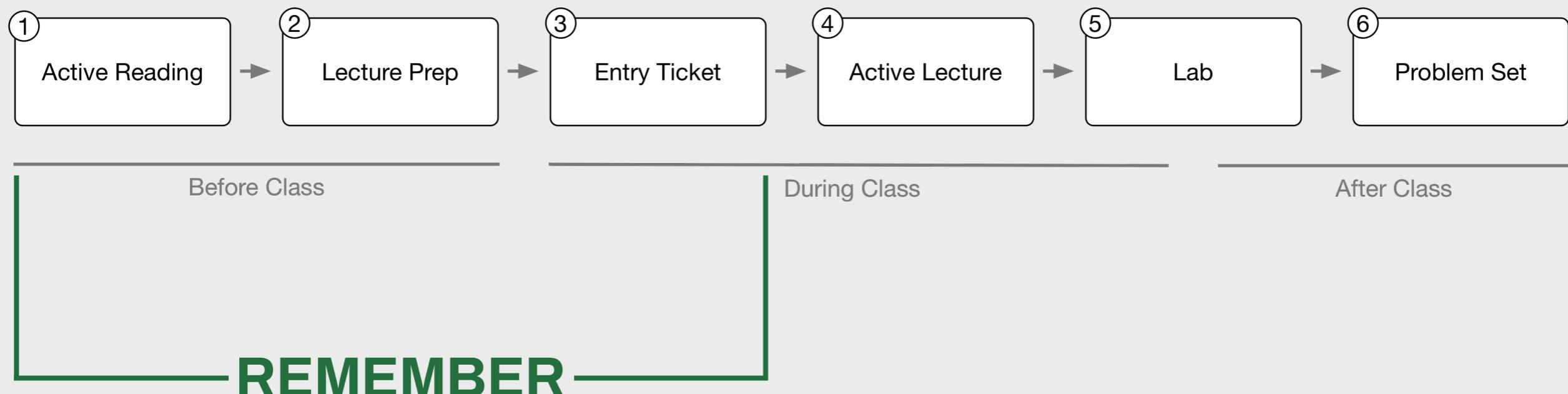
# COURSE FLOW



## 4. SYLLABUS OVERVIEW

---

# COURSE FLOW



Before Class

During Class

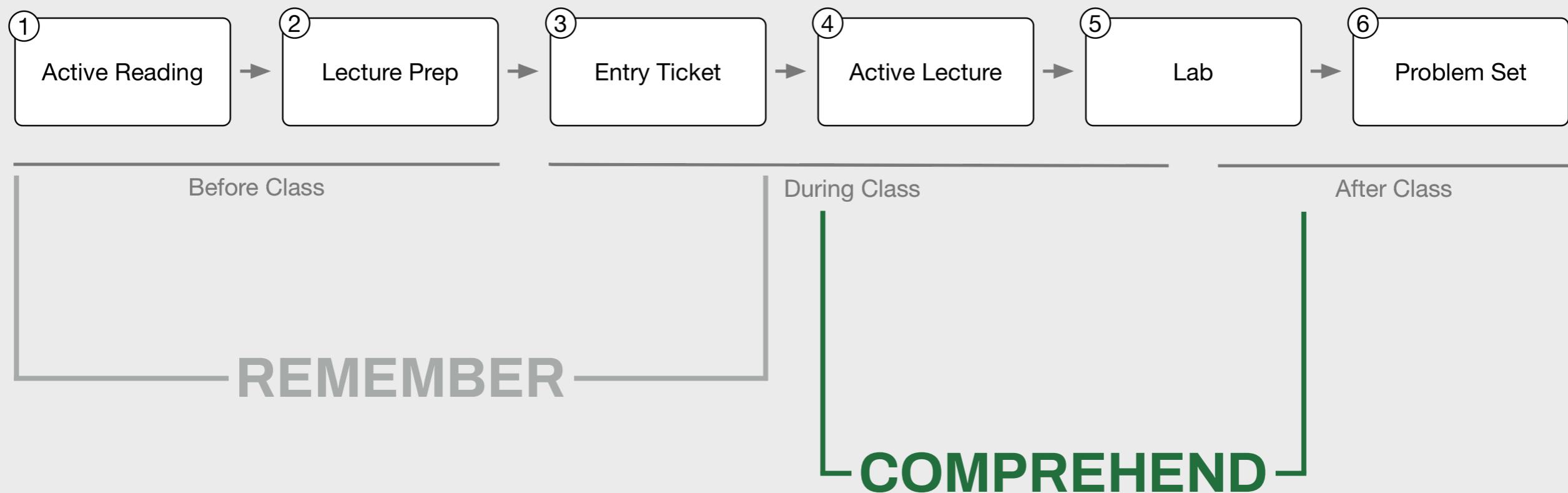
After Class

**REMEMBER**

## 4. SYLLABUS OVERVIEW

---

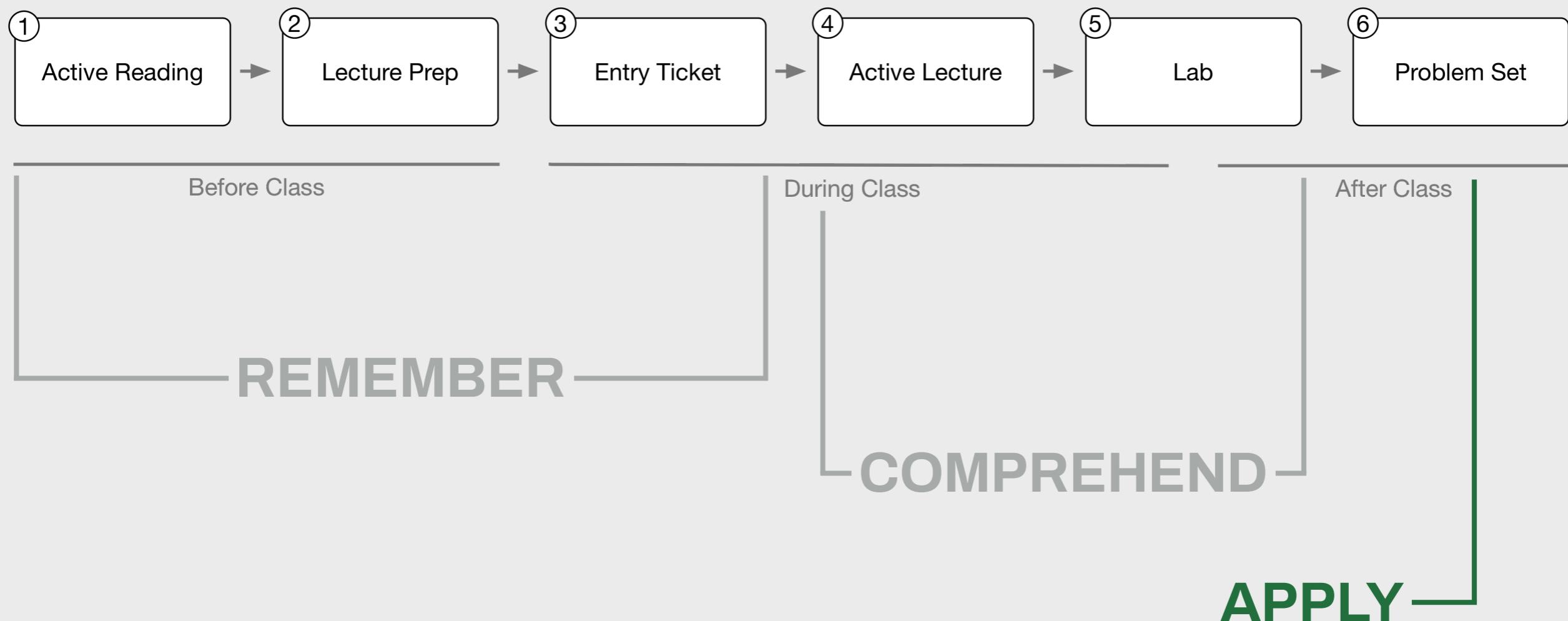
# COURSE FLOW



## 4. SYLLABUS OVERVIEW

---

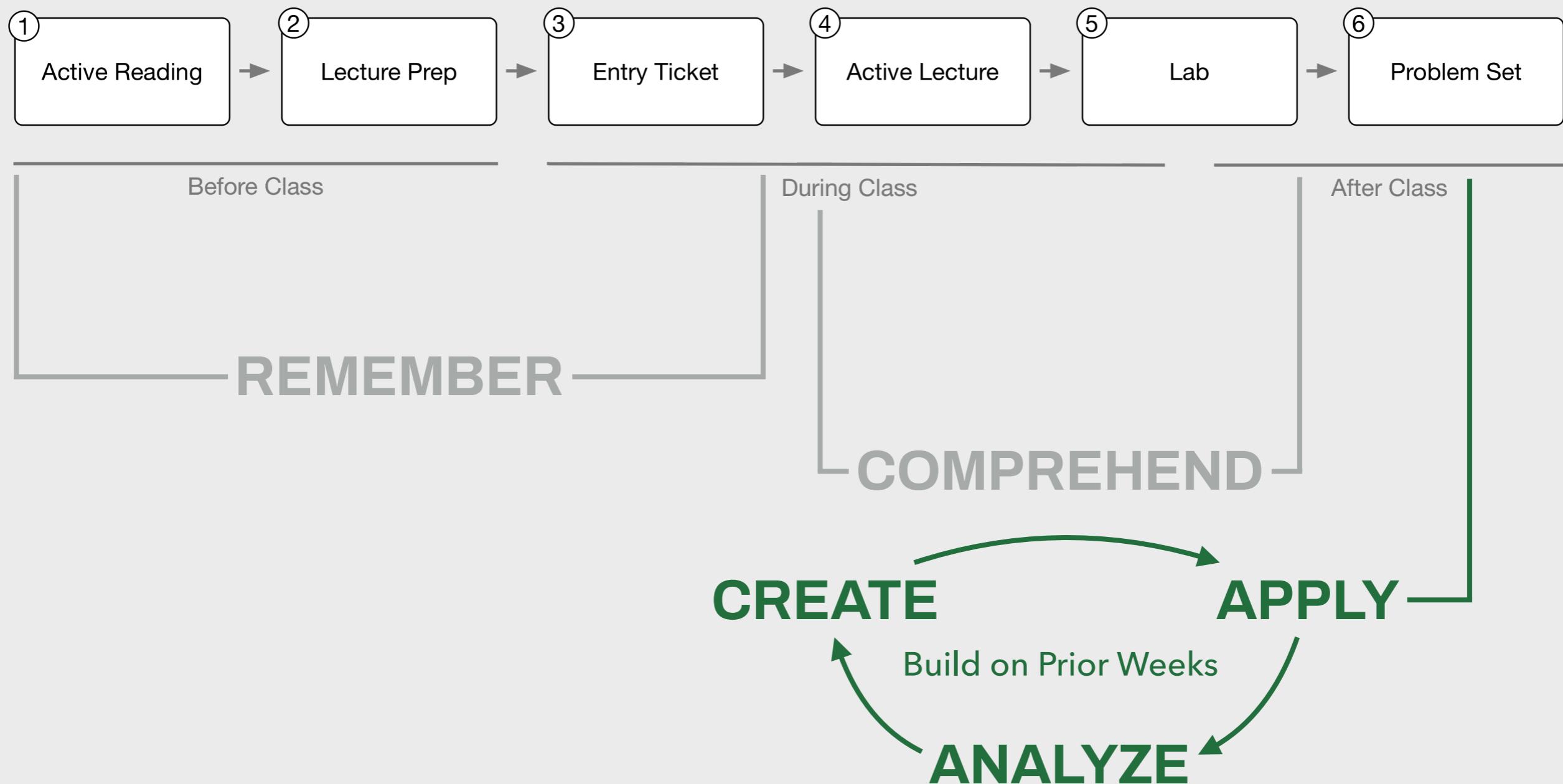
# COURSE FLOW



## 4. SYLLABUS OVERVIEW

---

# COURSE FLOW

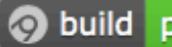


## 4. SYLLABUS OVERVIEW

---

# STAYING CURRENT

## ggplot2

build error  coverage 74% CRAN 2.2.1



**Overview**

ggplot2 is a system for declaratively creating graphics, based on [The Grammar of Graphics](#). You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

## Core-Documents

semester spring 2018 release updated version v18.4.0 last commit today repo size 1.4 MB



**Repository Contents**

This repository contains files for:

- 3600MorrisseySchedule.pdf

## 4. SYLLABUS OVERVIEW

---

# STAYING CURRENT

Branch: master ▾ [Core-Documents / NEWS.md](#) Find file Copy path

 chris-prener update NEWS with recent changes 78d26a3 12 minutes ago

1 contributor

21 lines (13 sloc) | 1.1 KB Raw Blame History   

## Core-Documents v18.4.0

- Added Brandon's office hours to `syllabus.pdf`
- Added the `3600MorrisseySchedule.pdf` file, which represents a sample week for the Spring 2018 semester. The R Seminars do not occur any week, but do occur on several weeks throughout the semester.

## Core-Documents v18.3.1

- Adjust dates of GISC and Public Policy readings for weeks 2 through 7
- Adjust chapters for [Sociospatial Data Science](#) on Week 2

## Core-Documents v18.3.0

## 4. SYLLABUS OVERVIEW

---

# GETTING HELP



**Chris Prener** 6:11 AM

I am trying to figure out why I am getting an error message with my histogram that says to "pick better value with `binwidth`". What is causing that error?

I've included a reprex here that reproduces this:



**Chris Prener** 6:11 AM

added this Plain Text snippet ▾

```
1 library(ggplot2)
2
3 # assign data to data frame
4 data <- mpg
5
6 # plot highway mpg
7 ggplot(data = data, mapping = aes(x = hwy)) +
8   geom_histogram()
9 #> `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



**Chris Prener** 6:11 AM

<https://i.imgur.com/VmcLq7w.png> (17 kB) ▾



# 5 DEFINING GIS

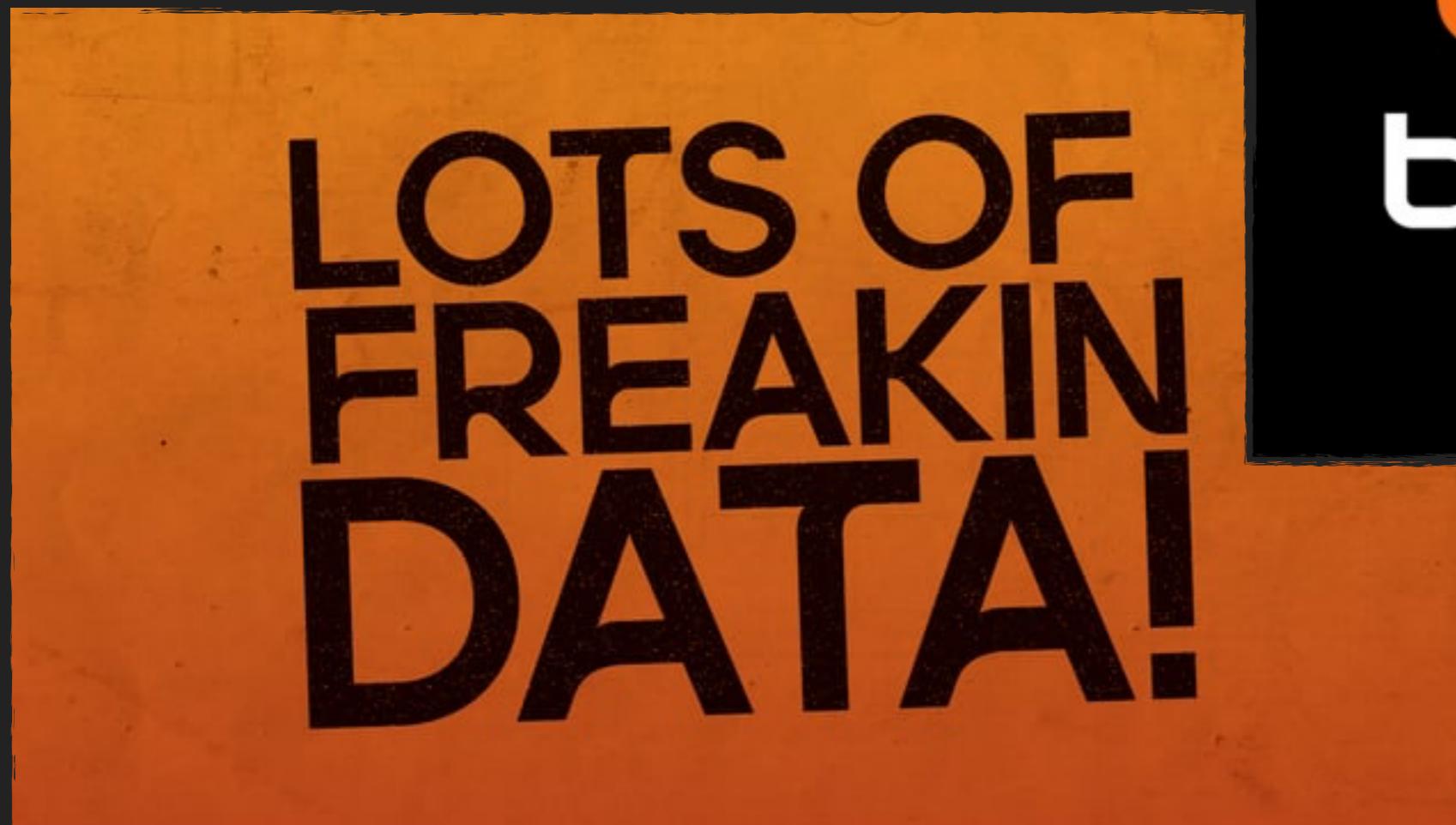
# a FIVE IMPORTANT TRENDS

## 5. DEFINING GIS

# TRENDS: SMARTPHONES



## TRENDS: BIG DATA



## 5. DEFINING GIS

# TRENDS: DATA SCIENCE



## 5. DEFINING GIS

# TRENDS: DEMAND FOR GIS



Christopher Prener

@chrisprener

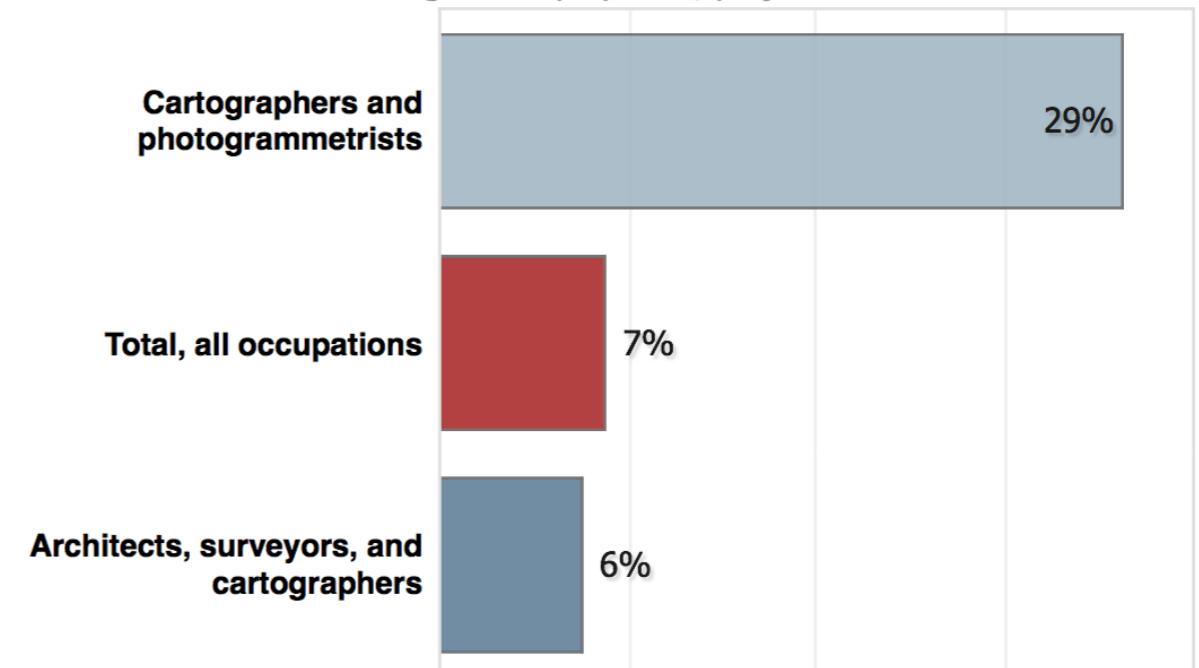
Twitter and other platforms allow users to embed spatial data in posts #soc4650 #soc5650 #gis

1:50 PM - 11 Jan 2017 from Morrissey Hall



### Cartographers and Photogrammetrists

Percent change in employment, projected 2014-24



Note: All Occupations includes all occupations in the U.S. Economy.

Source: U.S. Bureau of Labor Statistics, Employment Projections program

# TRENDS: OPEN SOURCE SOFTWARE



pro tip: 'data' is the  
plural form of 'datum'!

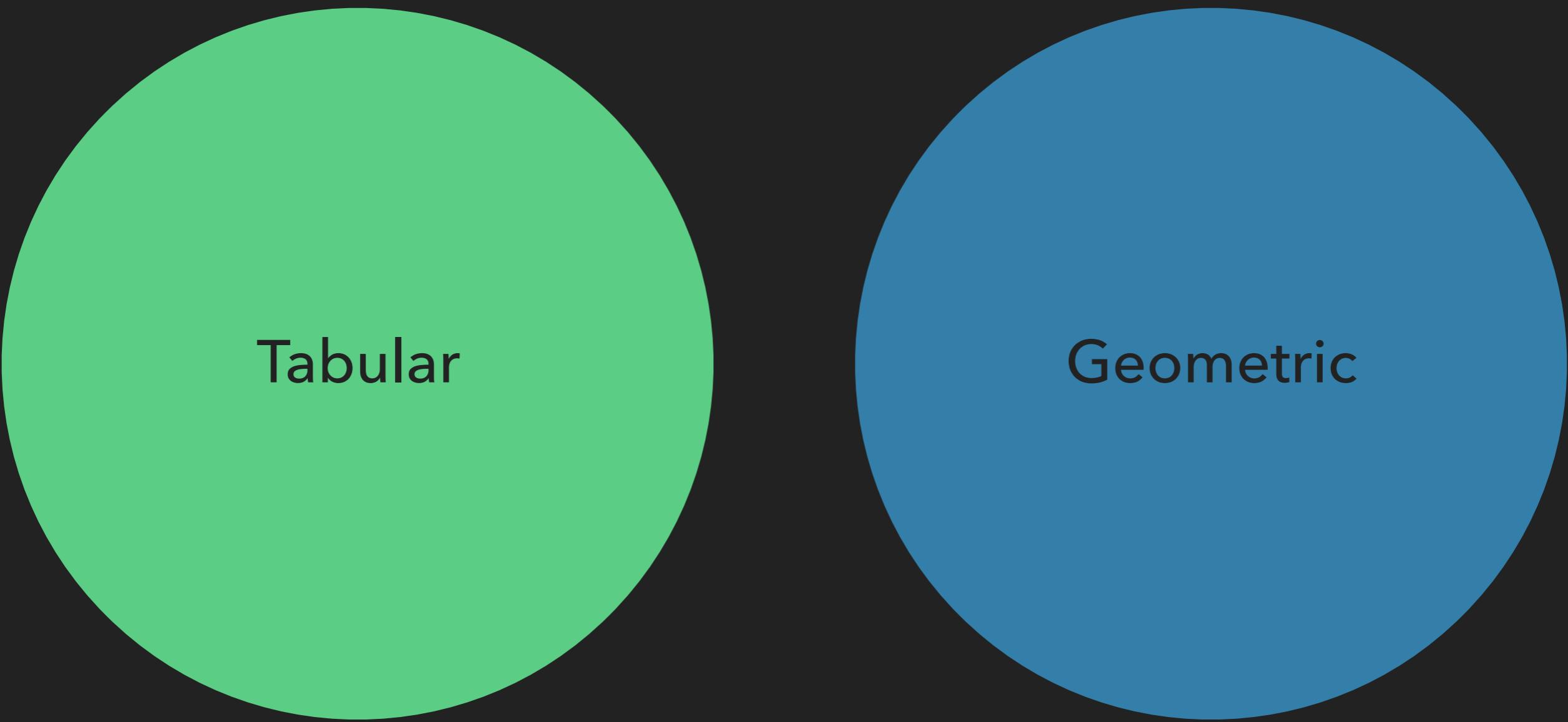
# b WHAT ARE SPATIAL DATA?

# WHAT ARE SPATIAL DATA?

- ▶ Data that can be **represented spatially**
  - They therefore have some type of **spatial reference** that locates data in two- or three-dimensional space
- ▶ Data that can (typically) be analyzed using **statistical techniques**
  - There is field at the intersection of geography and statistics known as **spatial statistics**

# WHAT ARE SPATIAL DATA?

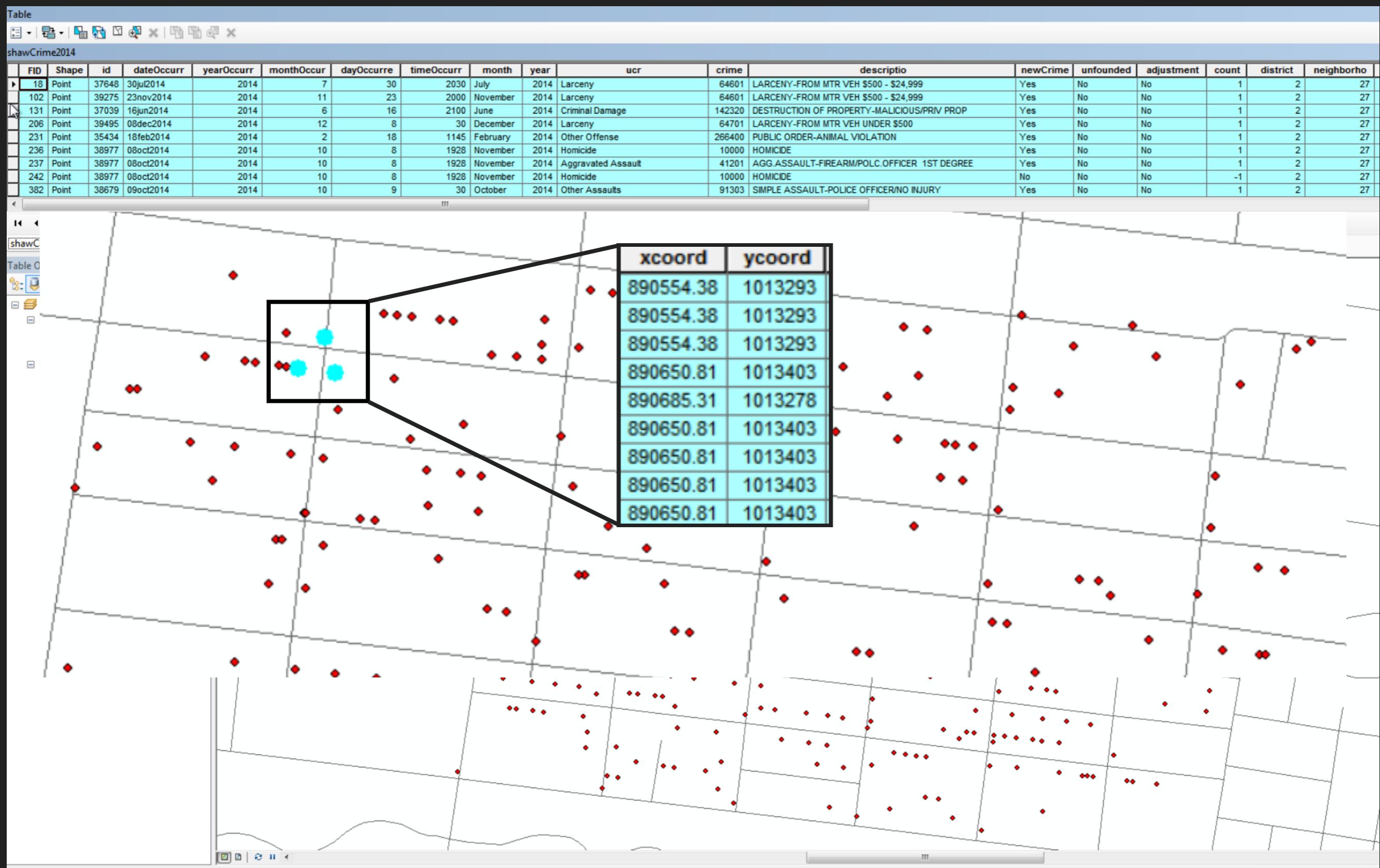
A combination of two forms of data



Tabular

Geometric

# WHAT ARE SPATIAL DATA?

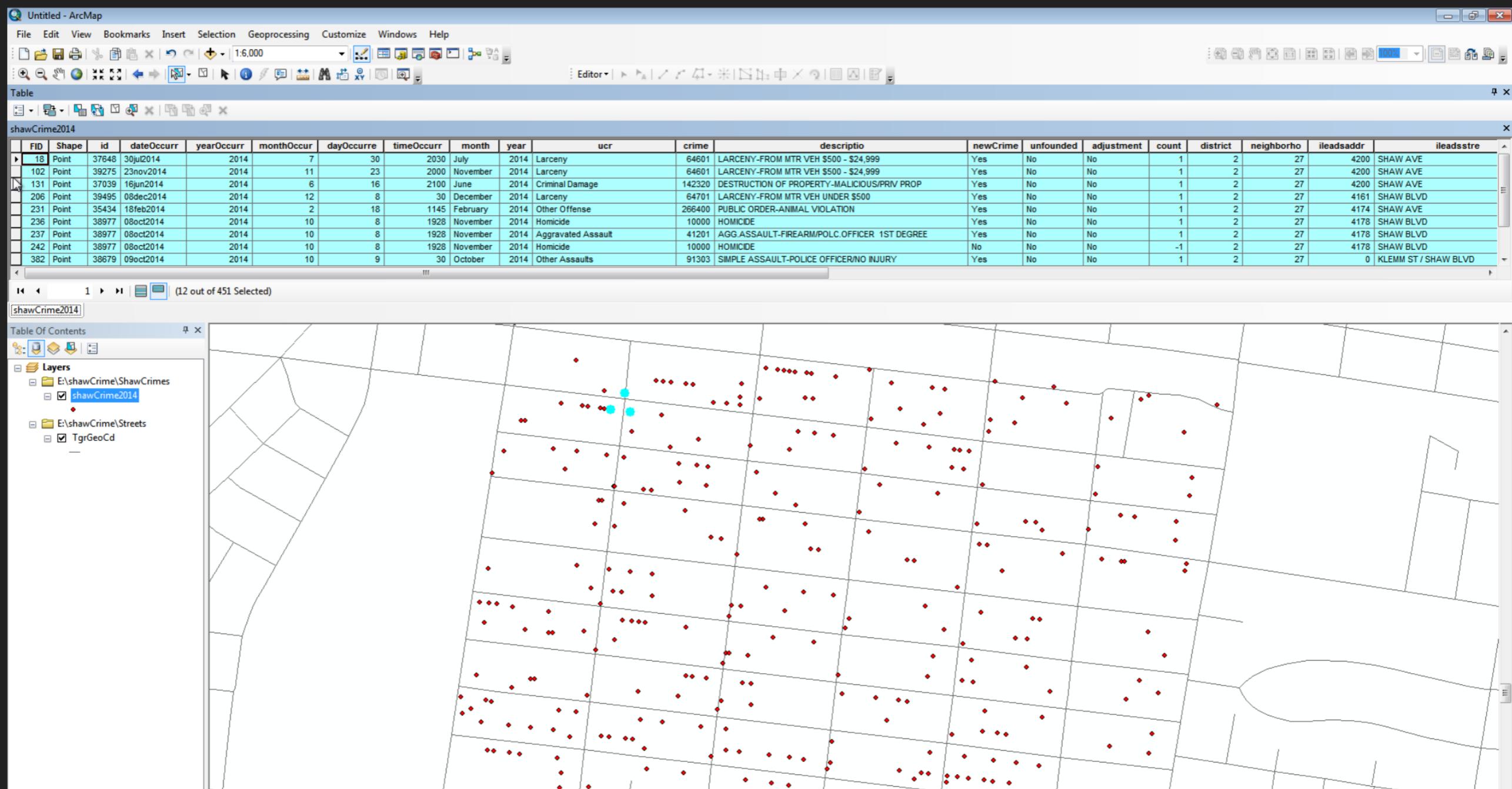


# C WHAT ARE GIS AND GISC?

## 5. DEFINING GIS

# ALPHABET SOUP

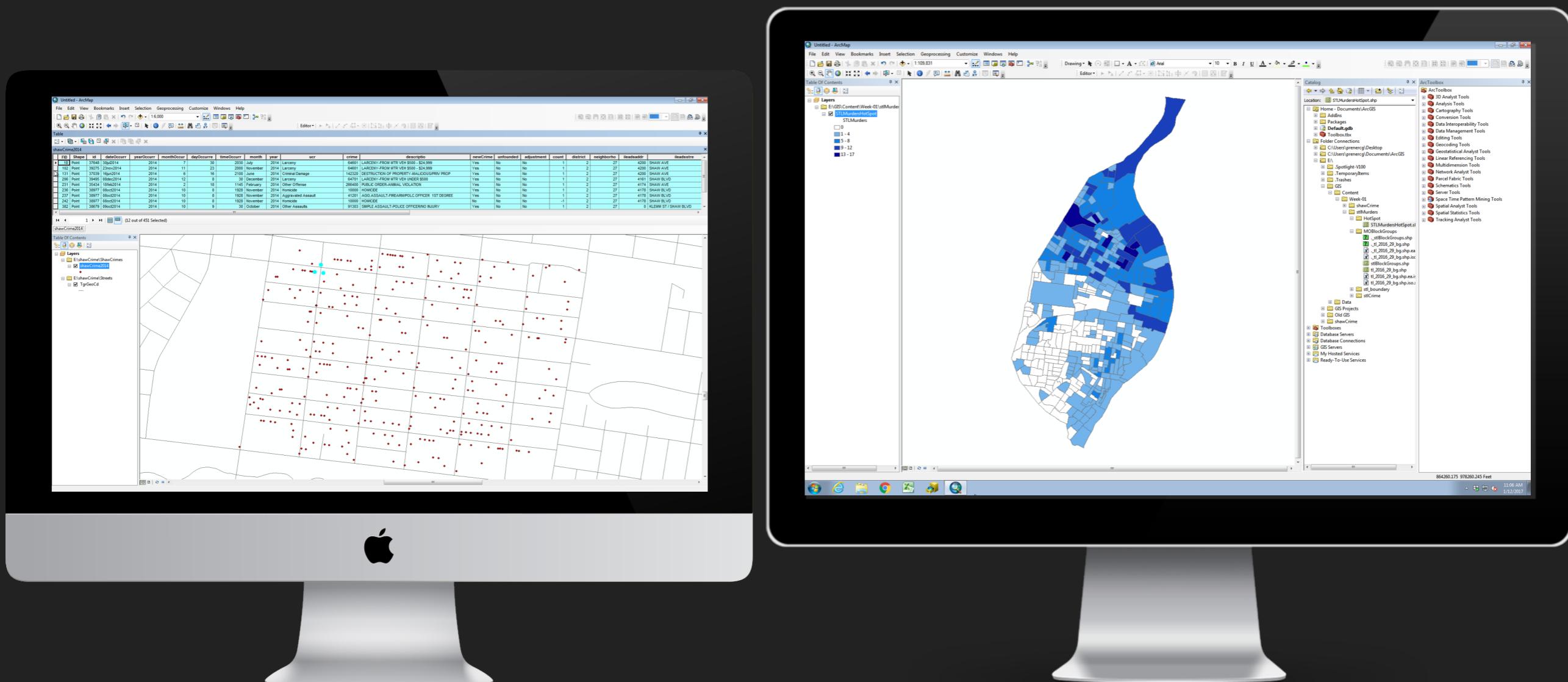
GIS = geographic information system



## 5. DEFINING GIS

# ALPHABET SOUP

GIS = geographic information system

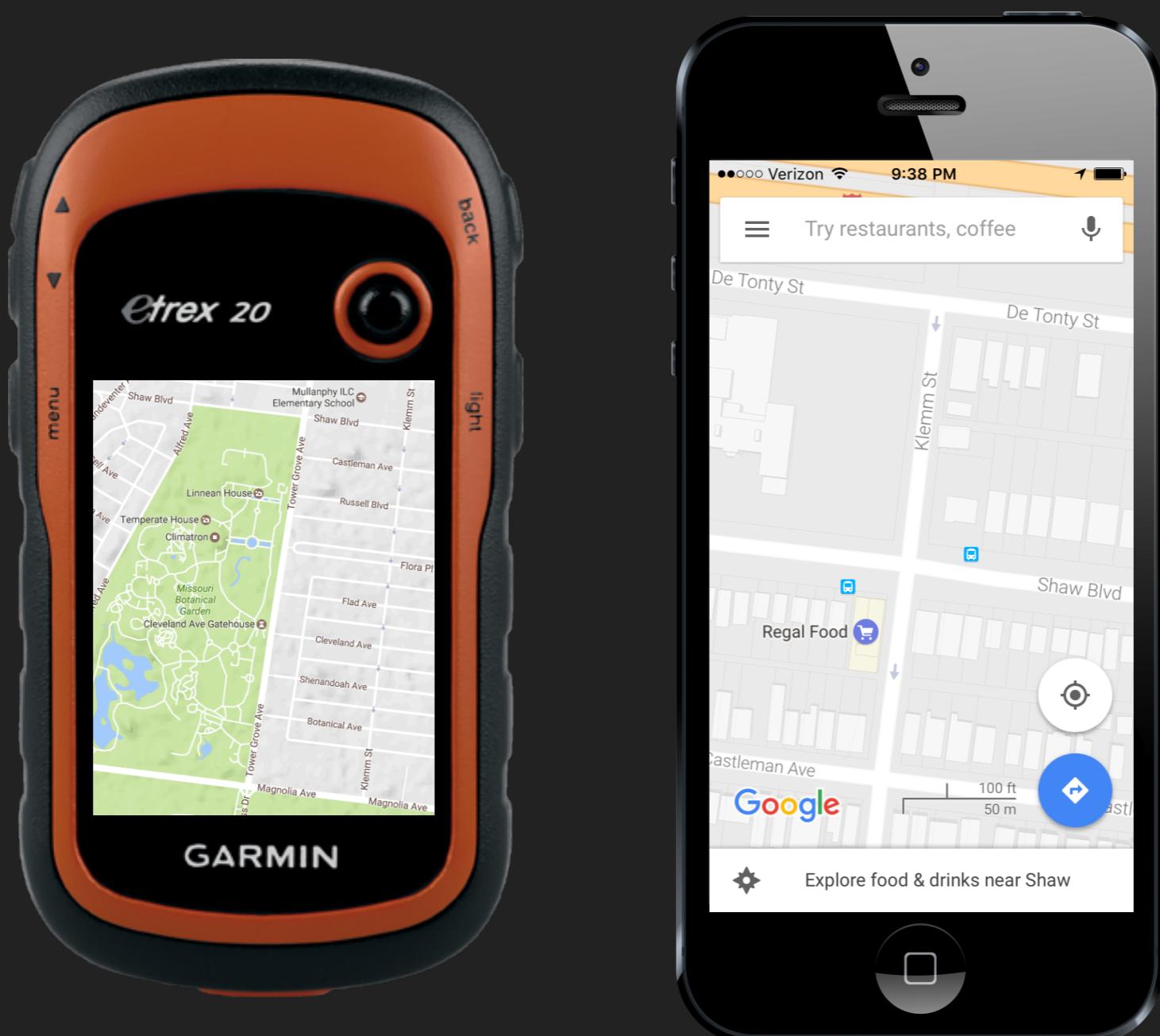


## 5. DEFINING GIS

---

# ALPHABET SOUP

GPS = global positioning system

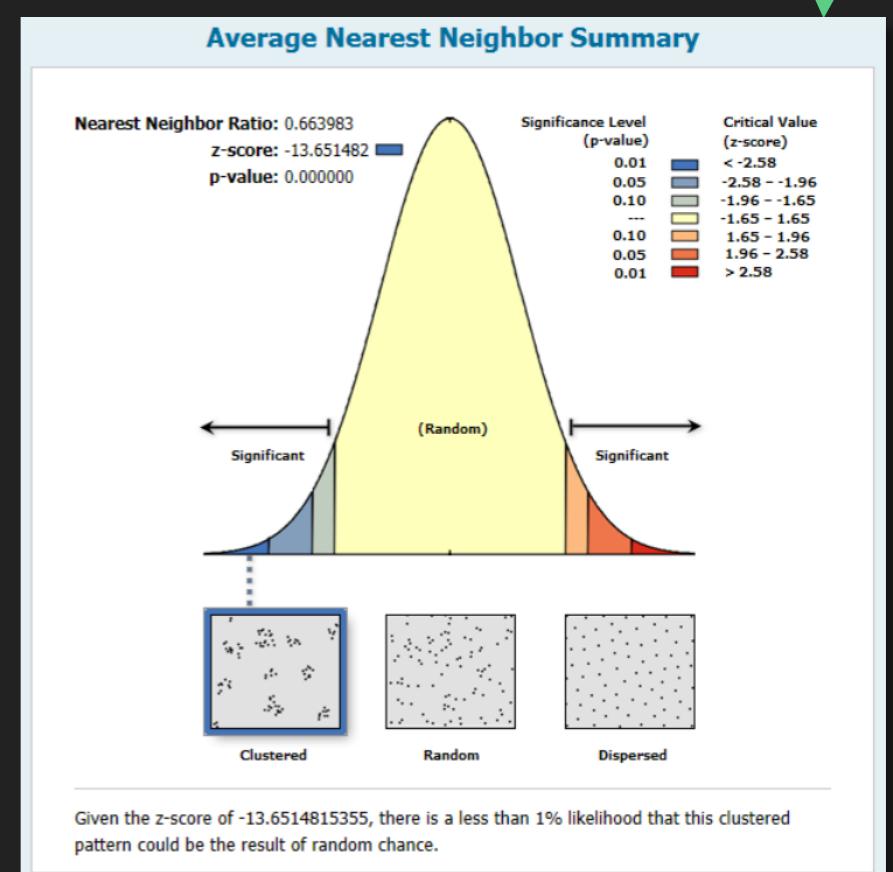
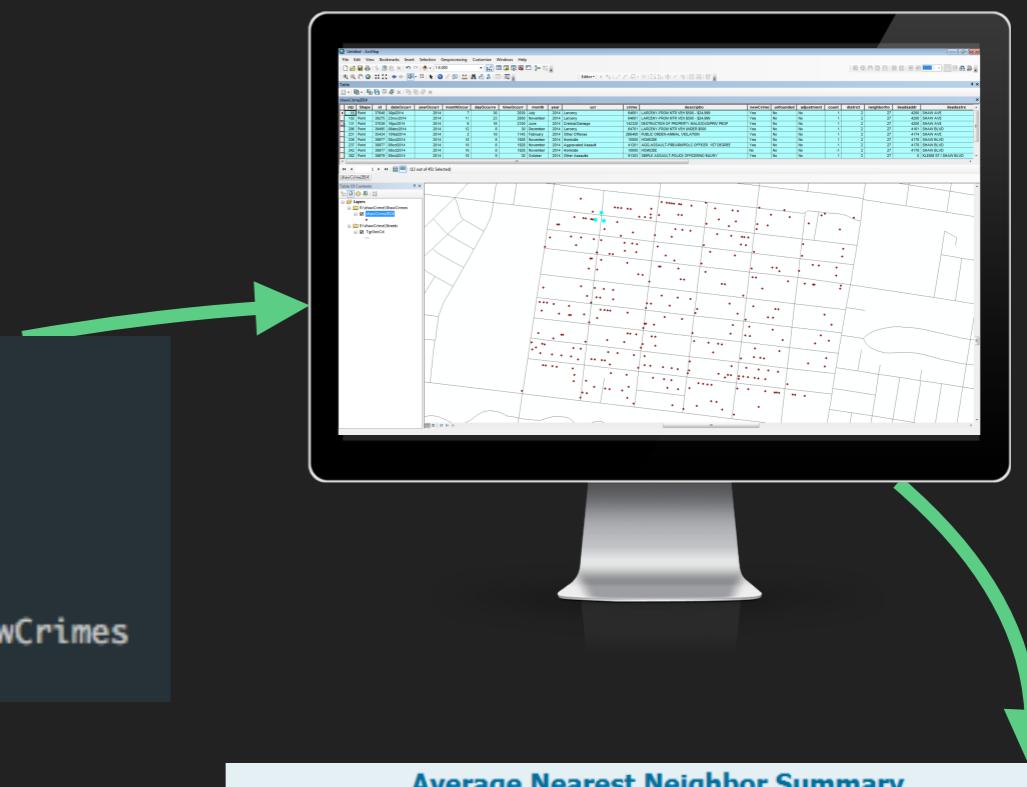
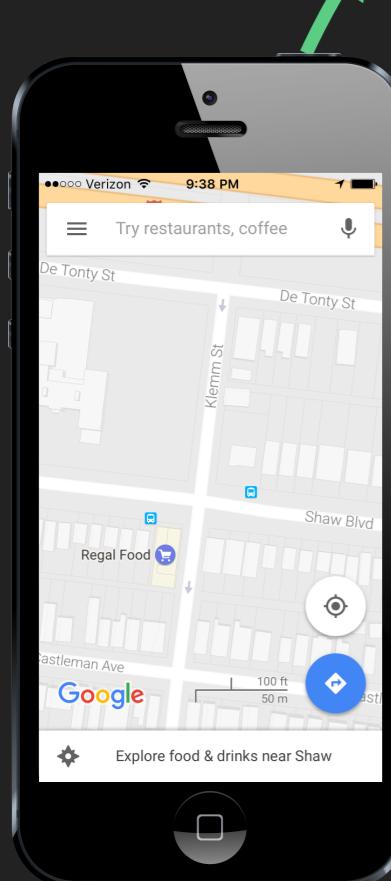


## 5. DEFINING GIS

# ALPHABET SOUP

GISc = geographic information science

```
2 shawCrimes <- import("shawCrime2016.csv")
3
4 shawCrimes %>%
5   mutate(ucr = case_when(crimes <= 19999 ~ 1,
6                         crimes >= 20000 & crimes <= 29999 ~ 2,
7                         crimes >= 30000 & crimes <= 39999 ~ 3,
8                         crimes >= 40000 & crimes <= 49999 ~ 4)) -> shawCrimes
9
```

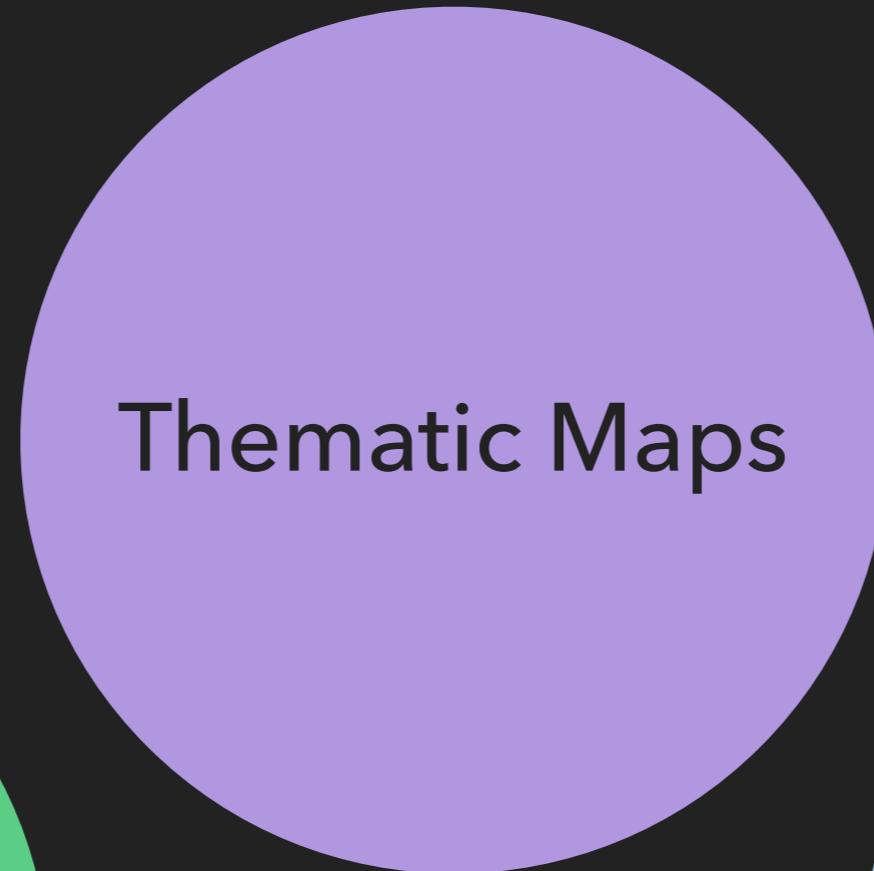


d **WHAT ARE  
SPATIAL DATA  
USED FOR?**

# THREE MAJOR USES



Reference Maps



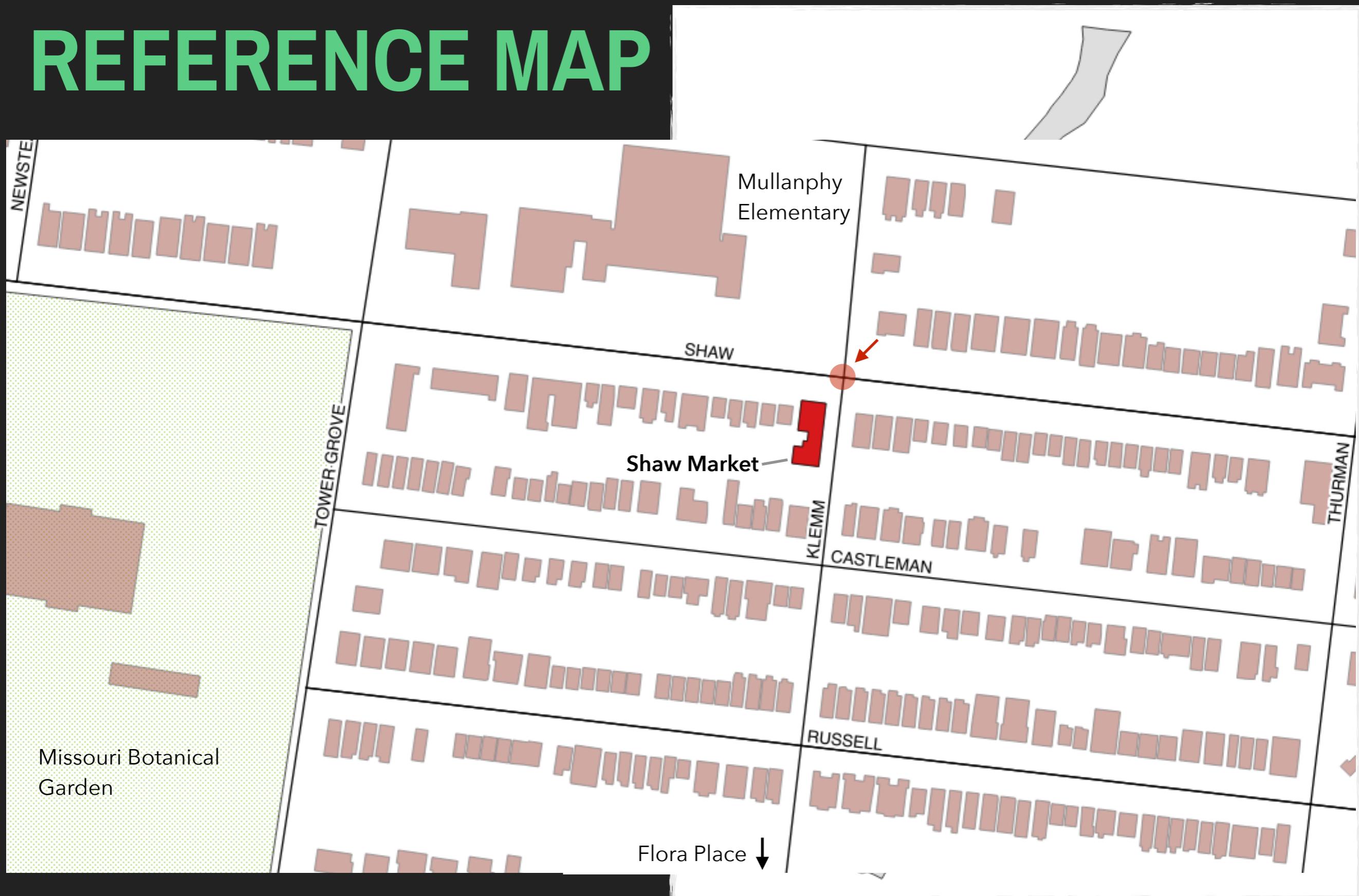
Thematic Maps



Inferential Maps

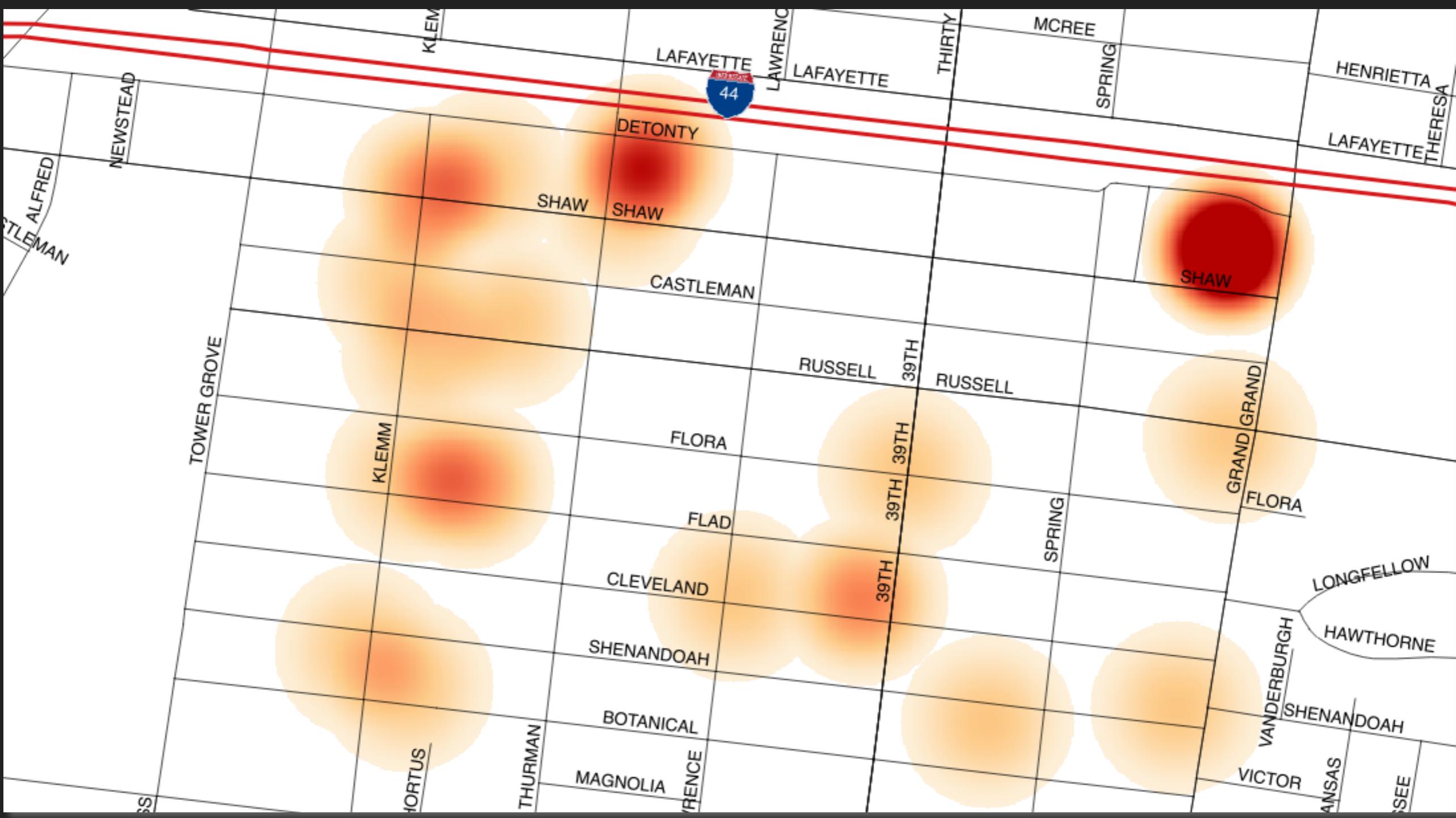
## 5. DEFINING GIS

# REFERENCE MAP



## 5. DEFINING GIS

# THEMATIC MAP



**EVERYTHING IS RELATED TO  
EVERYTHING ELSE, BUT  
NEAR THINGS ARE MORE  
RELATED THAN DISTANT  
THINGS.**

**William Tobler**

■ “A computer movie simulating urban growth in the Detroit region”  
(1970)

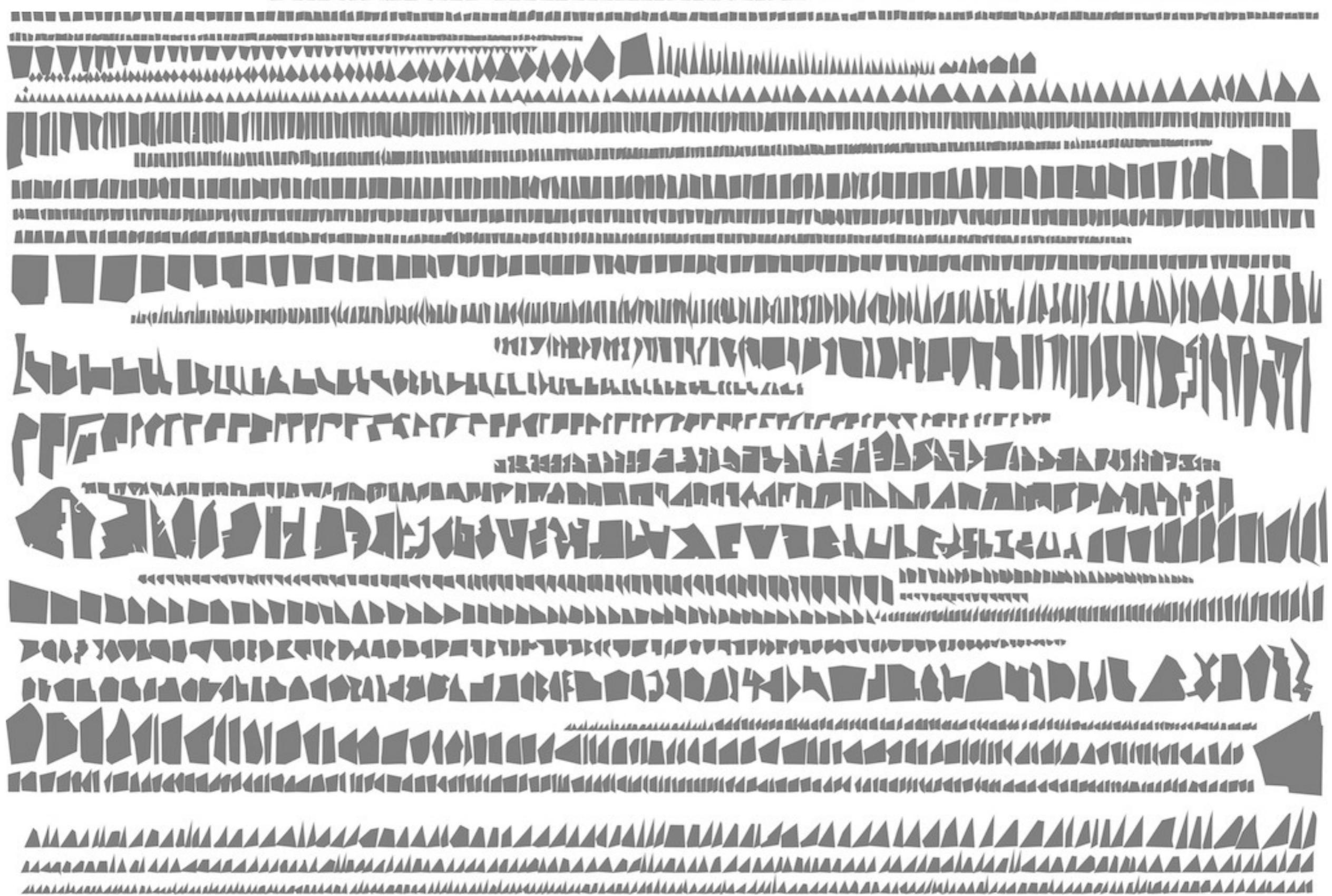
# PARIS

Armelle Caron



# PARIS

Armelle Caron



# PUTTING SPACE IN ITS PLACE

INTUITIVE MODEL



STATISTICAL MODEL



# PUTTING SPACE IN ITS PLACE

GIS MODEL



STATISTICAL MODEL

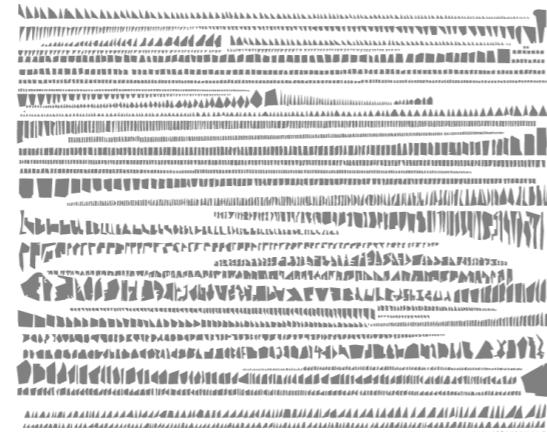


Armelle Caron  
armellecaron.fr

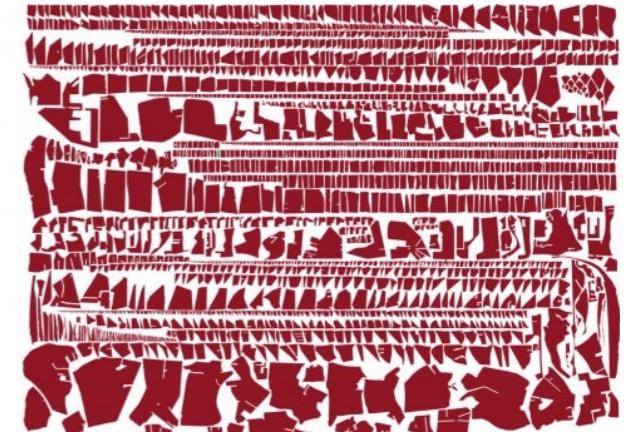
# LES VILLES RANGÉES



paris / paris rangé



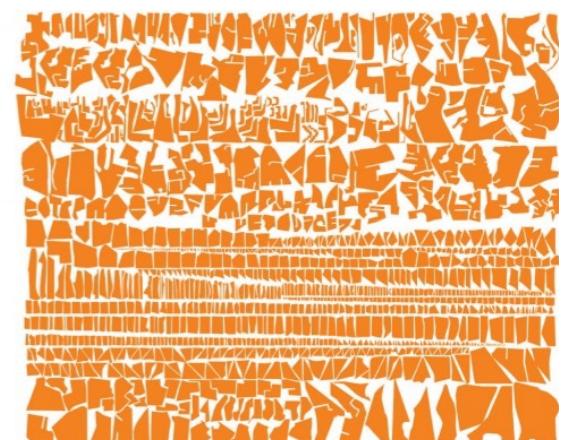
bordeaux / bordeaux rangé



tamarac / tamarac rangé



montpellier / montpellier rangé



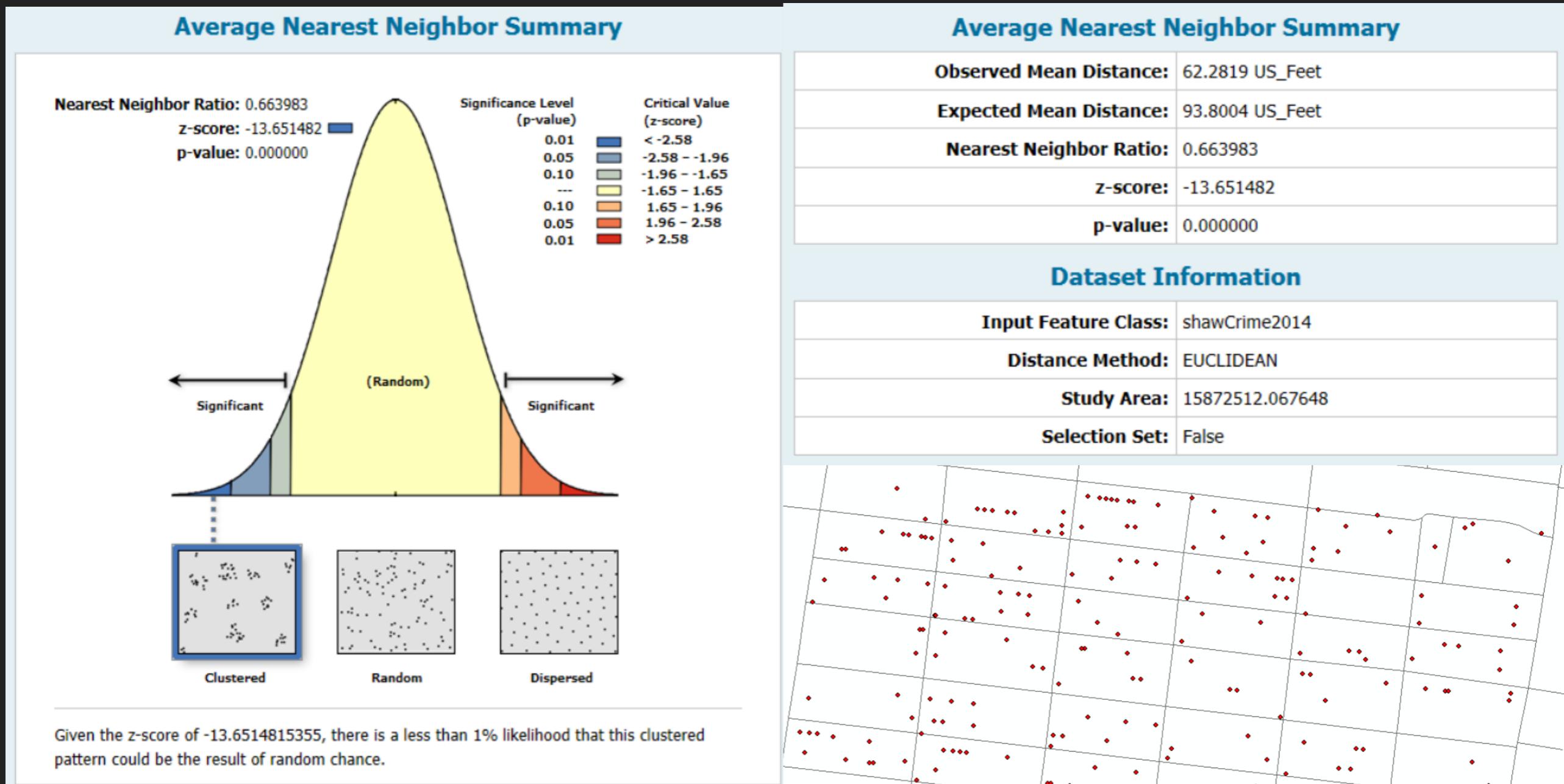
berlin / berlin rangé



new york / new york rangé

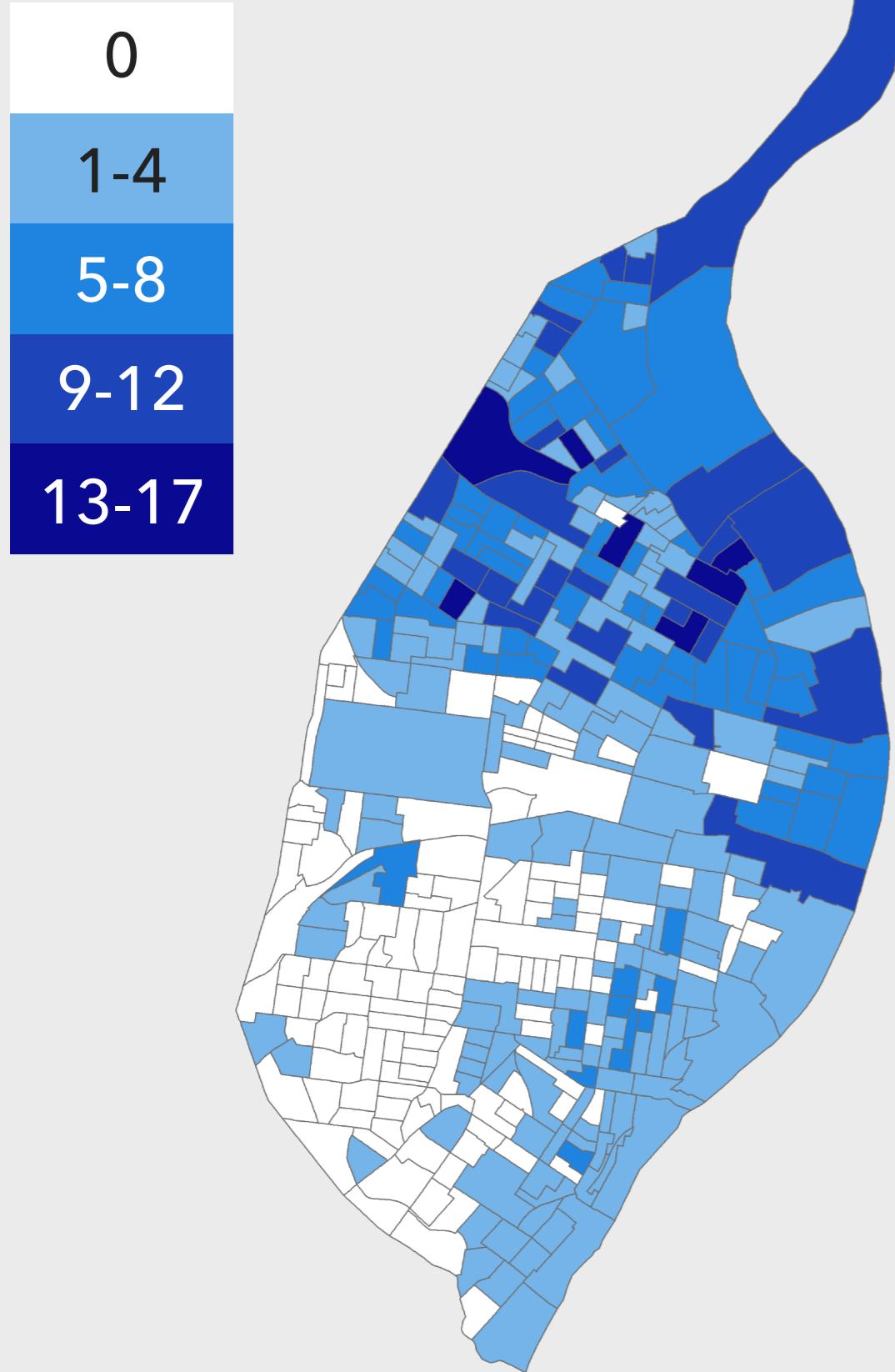


# INFERRENTIAL ANALYSES

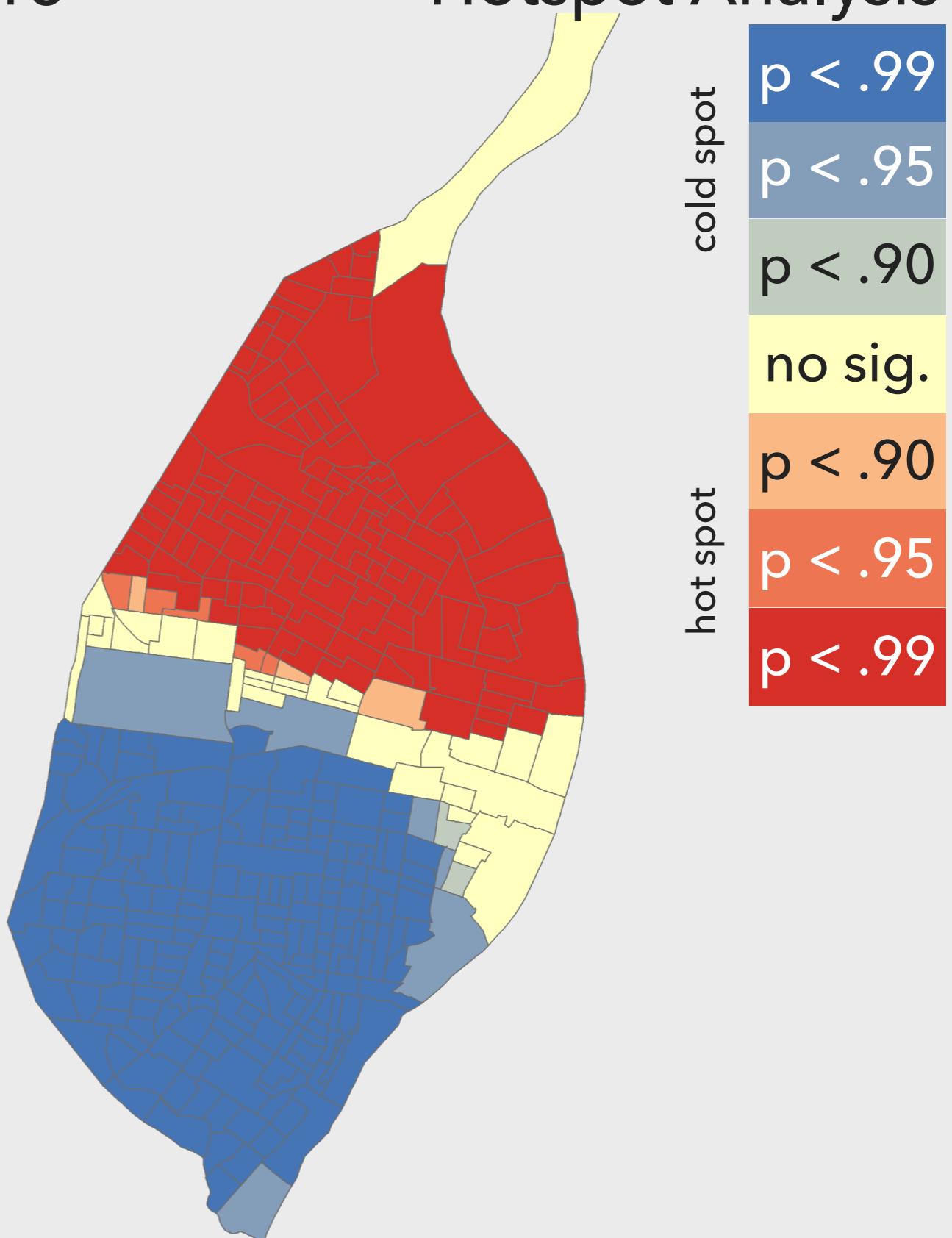


# INFERENTIAL MAPS

Murders in St. Louis, 2008-2015



Hotspot Analysis



# **6 WHAT IS A WORKFLOW?**

## 6. WHAT IS A WORKFLOW?

---

# WORKFLOWS SOLVE SPECIFIC PROBLEMS



## 6. WHAT IS A WORKFLOW?

---

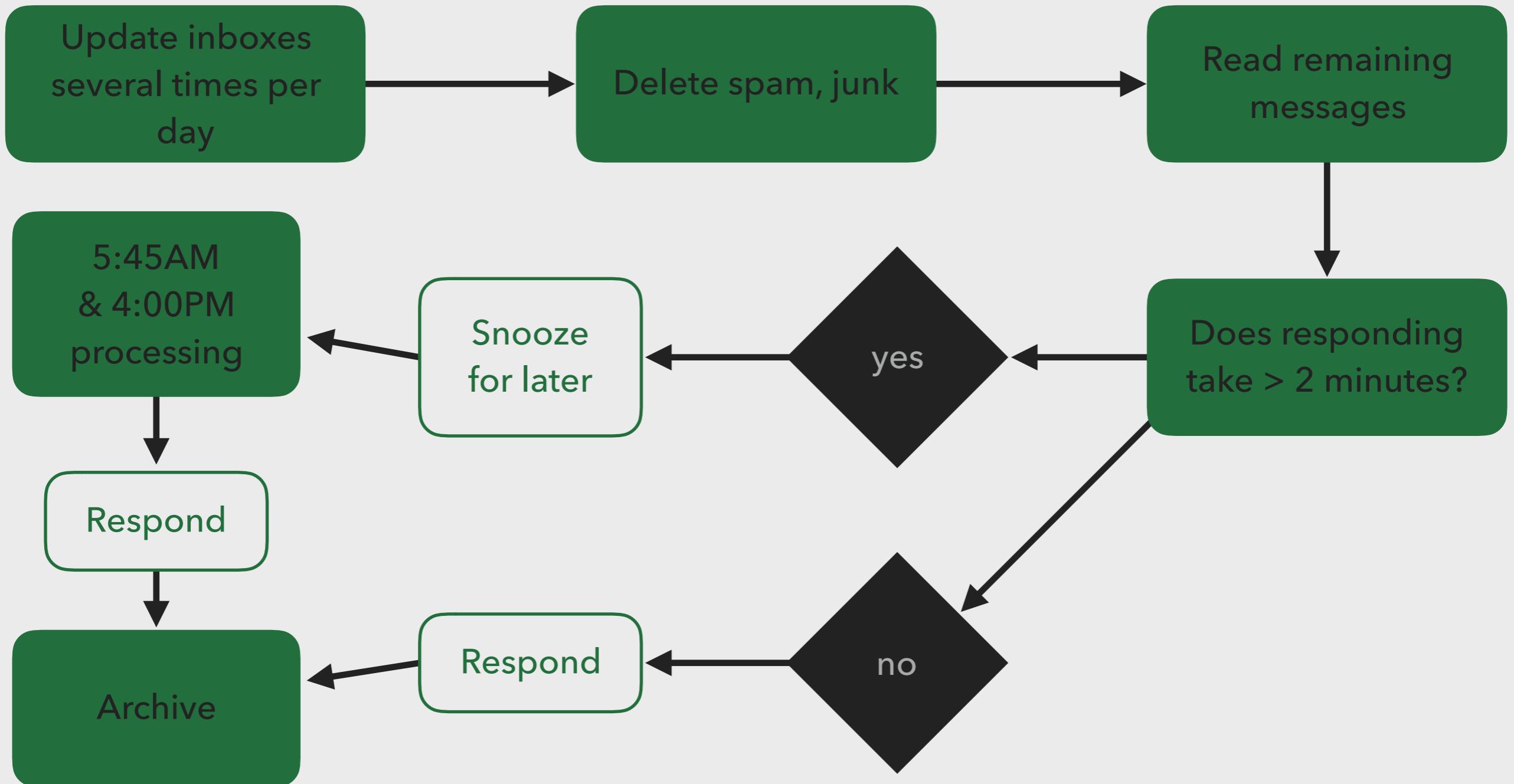
# WORKFLOWS SOLVE SPECIFIC PROBLEMS

There are two types of people in this world...



## 6. WHAT IS A WORKFLOW?

# WORKFLOWS SOLVE SPECIFIC PROBLEMS

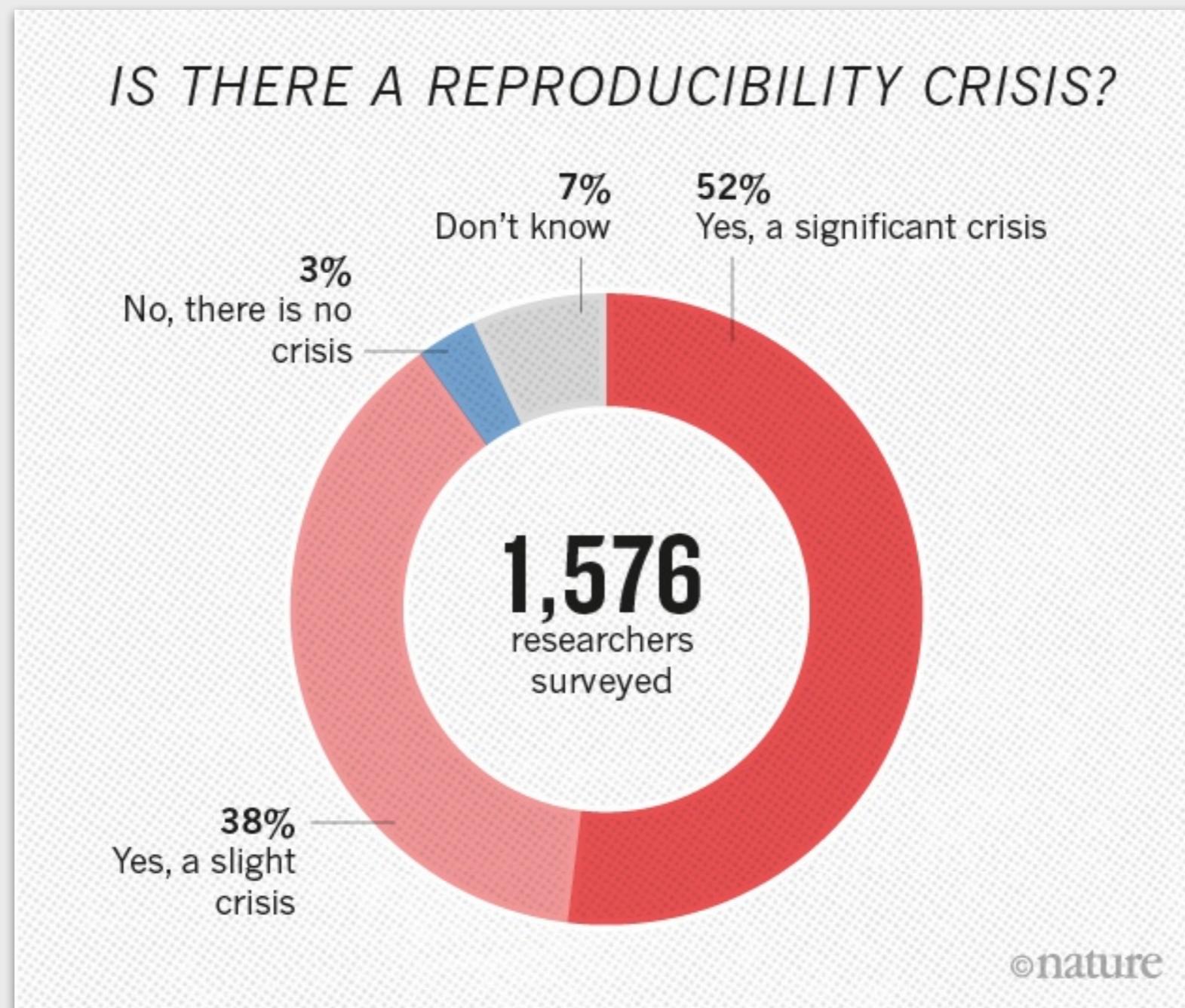


## 6. WHAT IS A WORKFLOW?

---

# WE HAVE A REPRODUCIBILITY PROBLEM

“1,500 scientists lift the lid on reproducibility” (Baker 2016)

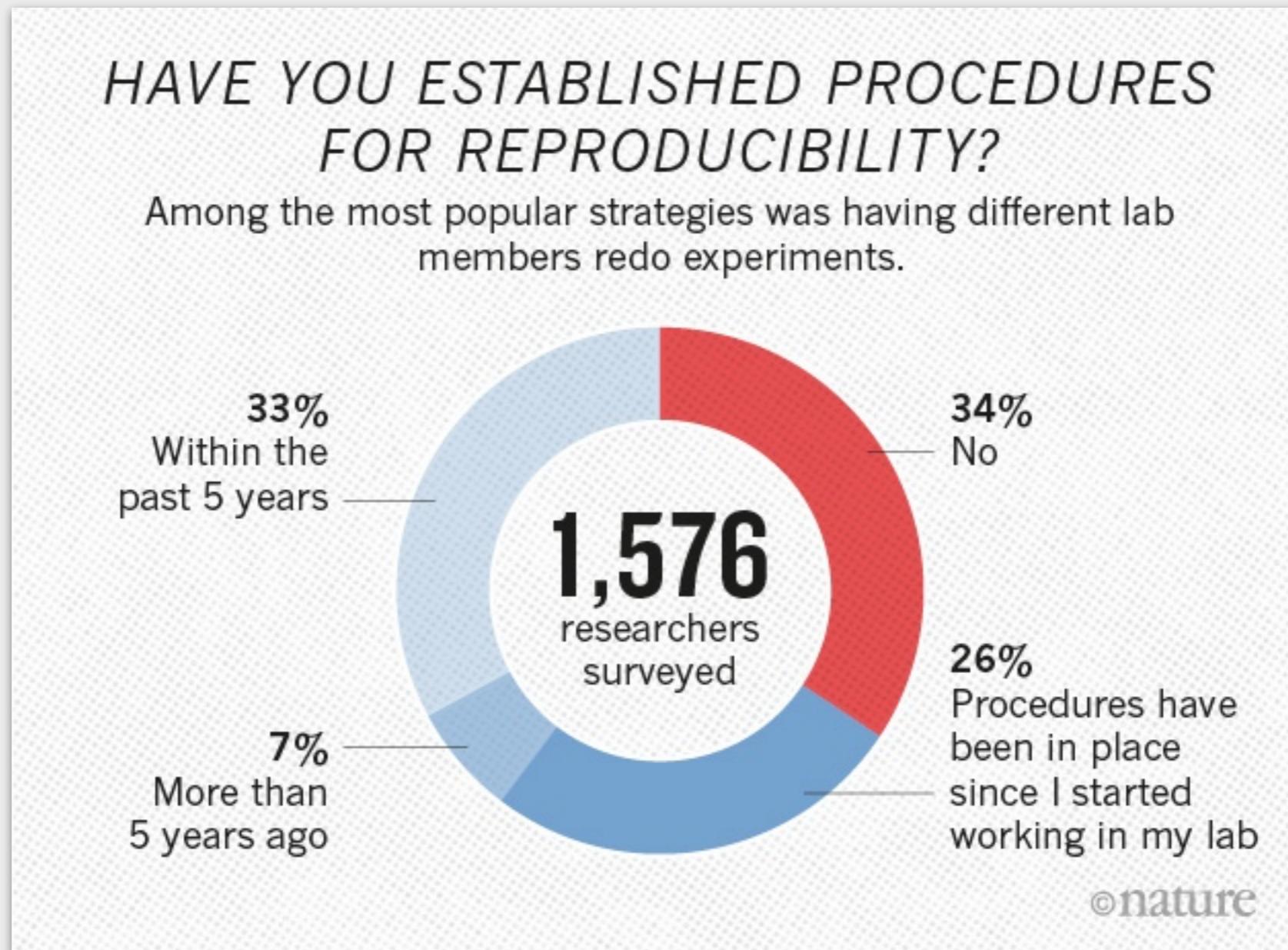


## 6. WHAT IS A WORKFLOW?

---

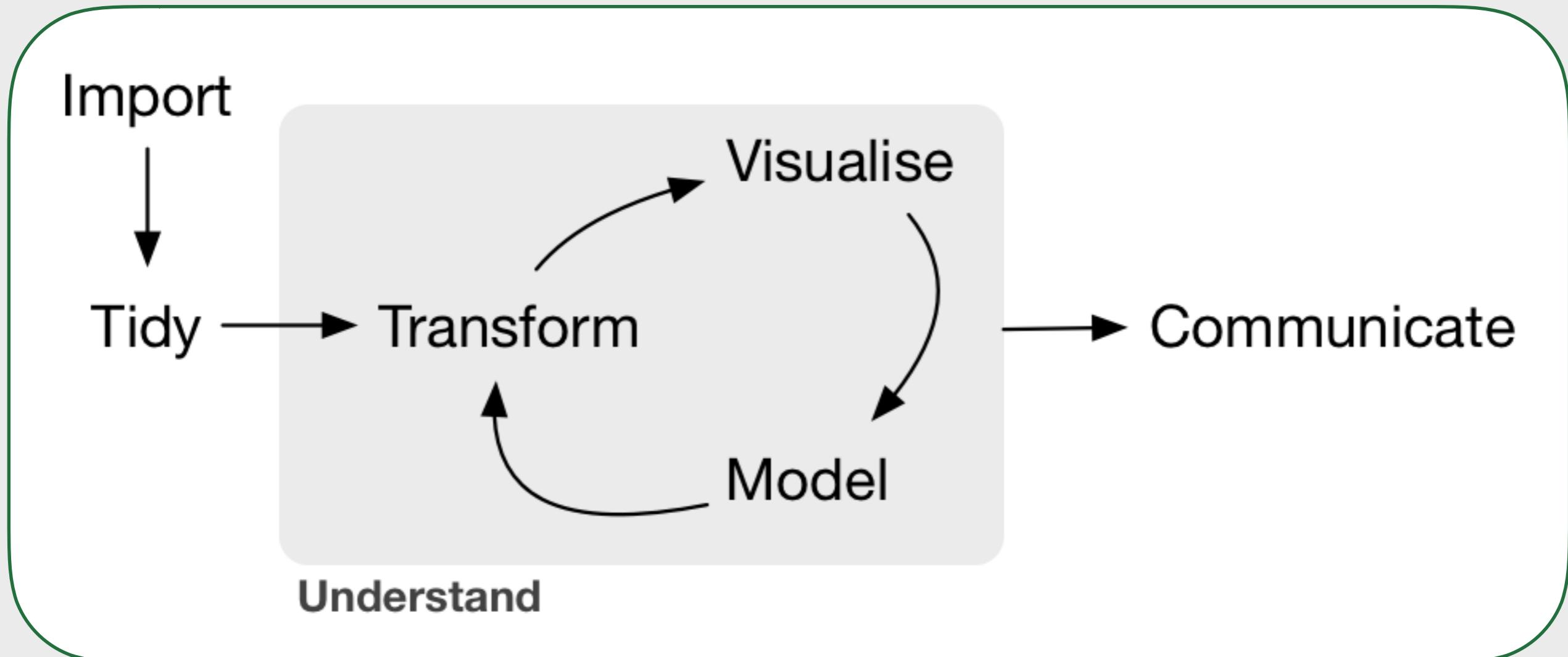
# WE HAVE A REPRODUCIBILITY PROBLEM

“1,500 scientists lift the lid on reproducibility” (Baker 2016)



## 6. WHAT IS A WORKFLOW?

# HIGH LEVEL WORKFLOW



**FOR  
EACH  
STEP:**

1. Plan
2. Organize
3. Document
4. Execute

7 OUR FIRST  
MAPS!

# 8 BACK MATTER

## 8. BACK MATTER

---

# AGENDA REVIEW

2. Introductions

3. What We Learn from Maps

4. Syllabus Overview

5. Defining GIS

6. What is a Workflow?

7. Our First Map!

## 8. BACK MATTER

---

# REMINDERS



We'll end every class with "**Back Matter**" - goal is to review what we covered, what due dates are coming up, and any announcements.



Lecture Prep 02 due **next Monday**; will be posted on Wednesday



If you have not already done so, please download the course data release onto your external storage device (thumb drive or external hard drive).