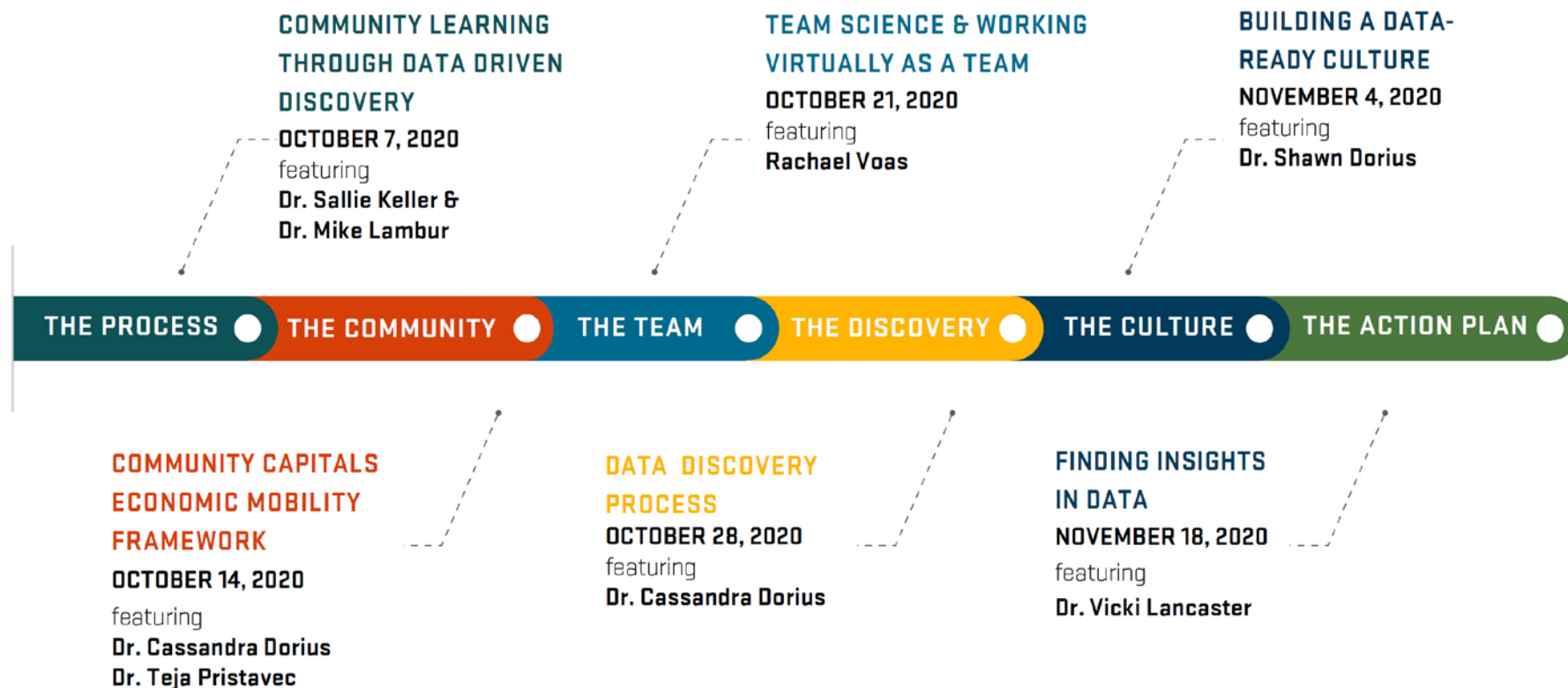COMMUNITY LEARNING THROUGH DATA DRIVEN DISCOVERY

DATA DISCOVERY WORKSHOP

Community Catalyst Webinar
Oregon State University

October 28, 2020

COMMUNITY CATALYST WEBINAR SERIES OVERVIEW

COMMUNITY LEARNING THROUGH DATA DRIVEN DISCOVERY
OCTOBER 7, 2020
featuring
**Dr. Sallie Keller &
Dr. Mike Lambur**

TEAM SCIENCE & WORKING VIRTUALLY AS A TEAM
OCTOBER 21, 2020
featuring
Rachael Voas

BUILDING A DATA-READY CULTURE
NOVEMBER 4, 2020
featuring
Dr. Shawn Dorius

THE PROCESS ● THE COMMUNITY ● THE TEAM ● THE DISCOVERY ● THE CULTURE ● THE ACTION PLAN ●

COMMUNITY CAPITALS ECONOMIC MOBILITY FRAMEWORK
OCTOBER 14, 2020
featuring
**Dr. Cassandra Dorius
Dr. Teja Pristavec**

DATA DISCOVERY PROCESS
OCTOBER 28, 2020
featuring
Dr. Cassandra Dorius

FINDING INSIGHTS IN DATA
NOVEMBER 18, 2020
featuring
Dr. Vicki Lancaster

MADE POSSIBLE BY:

IOWA STATE UNIVERSITY

Oregon State University

UNIVERSITY of VIRGINIA

Virginia Cooperative Extension
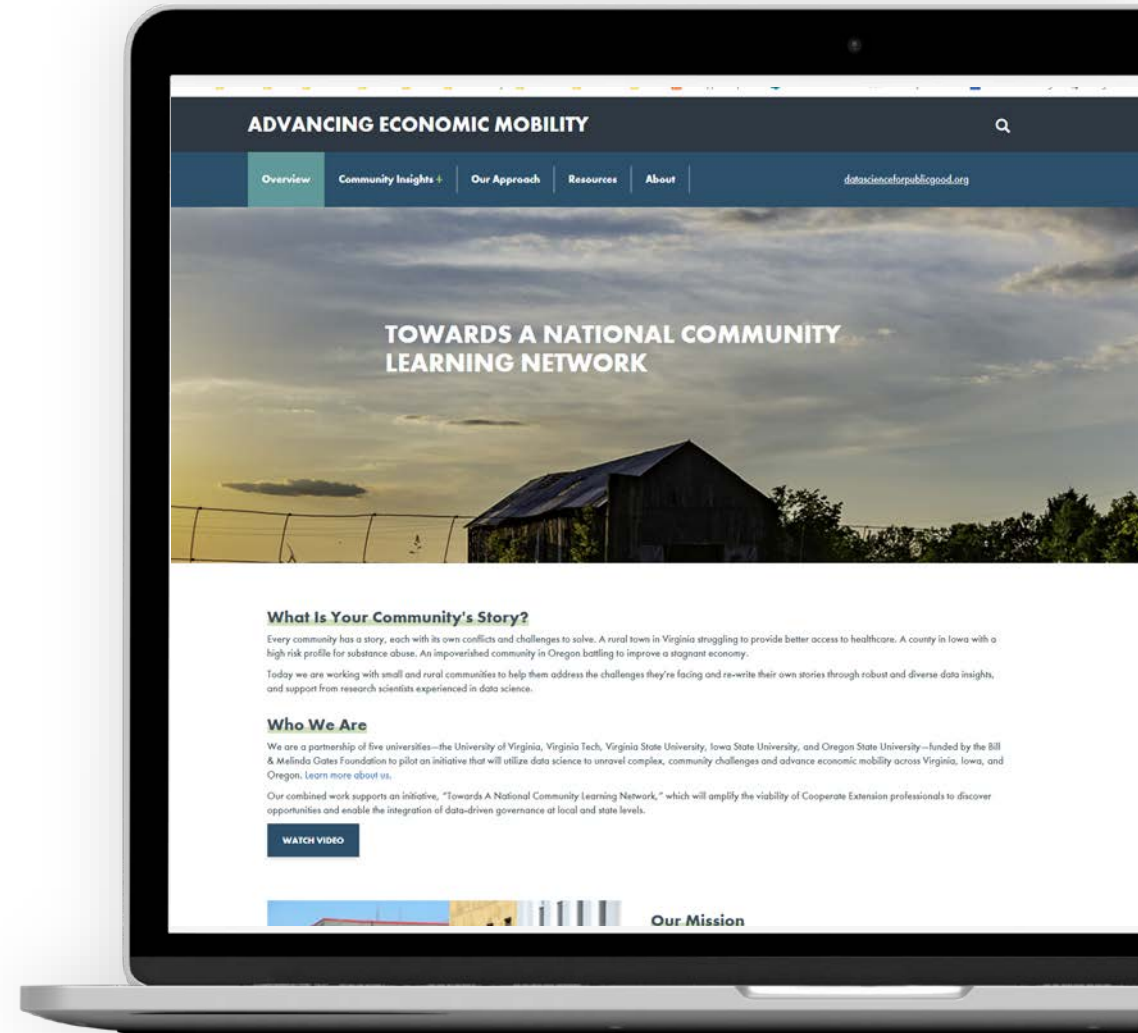Virginia Tech • Virginia State University

BILL & MELINDA GATES foundation

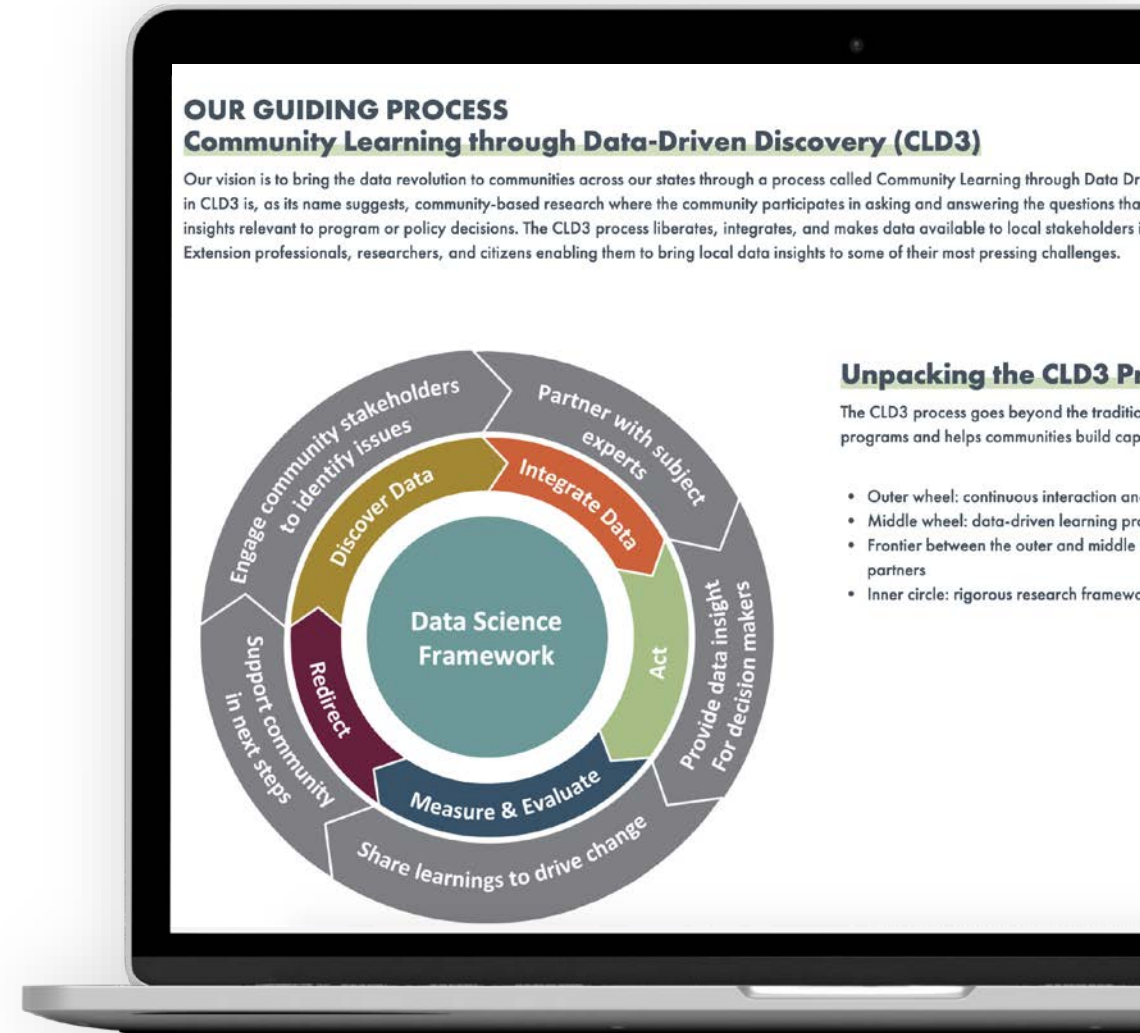USING DATA TO INFORM DECISIONS

TODAY WE WILL EXPLORE THE DATA DISCOVERY PROCESS:

1. HOW DOES IT WORK?

2. WHAT IS THE COMMUNITY'S CRITICAL QUESTION?

3. WHAT DATA ARE AVAILABLE?

4. WHY SHOULD YOU TRUST THE DATA?

5. WHAT ARE THE ETHICAL CONSIDERATIONS?

Our goal is to **empower** Cooperative Extension to **engage** with local stakeholders on **data-driven** approaches to economic mobility challenges **unique** to their communities.

datascienceforthepublicgood.org/economic-mobility/community-insights

Our **guiding process** is
Community
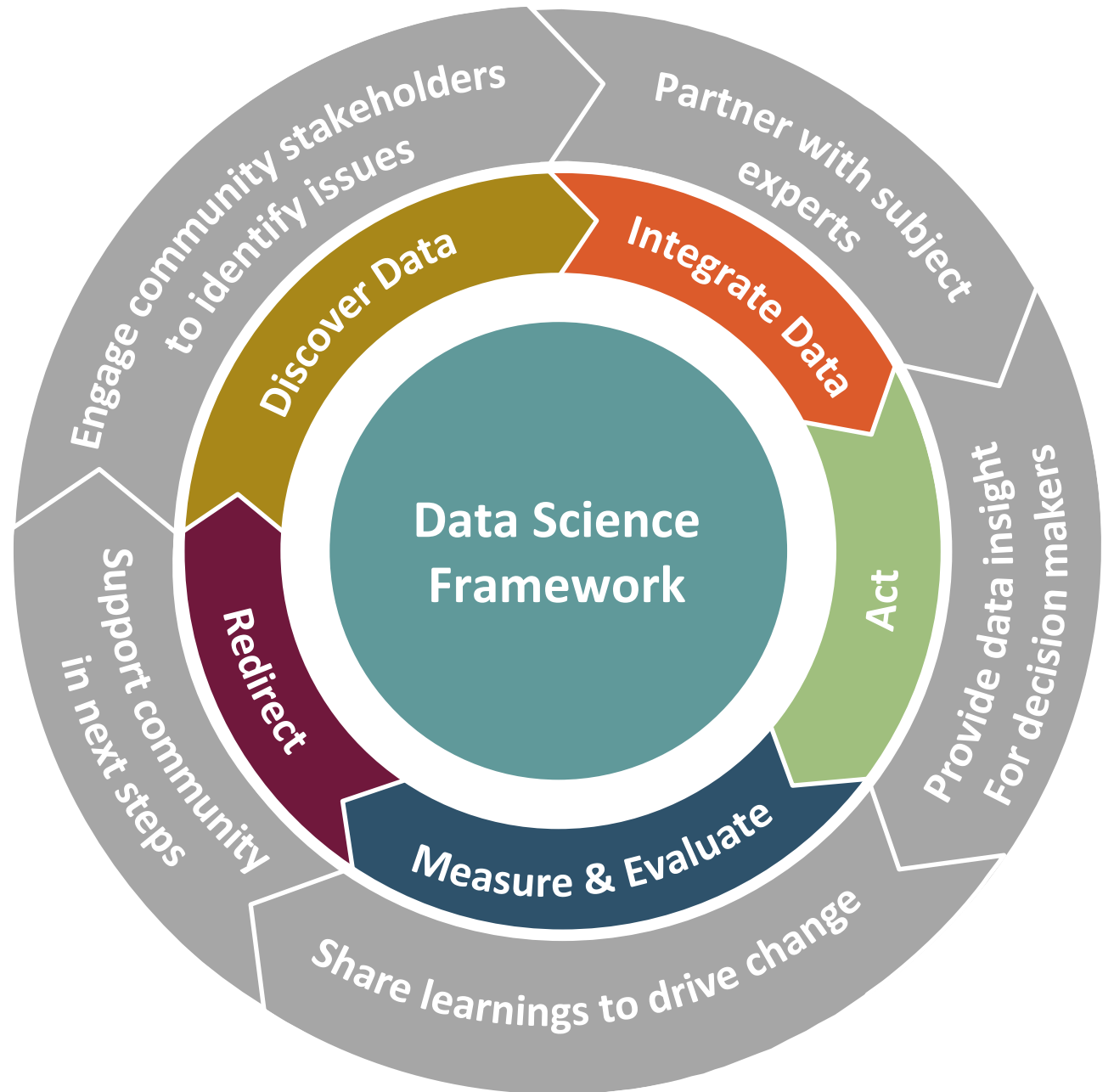Learning through
Data-Driven
Discovery

datascienceforthepublicgood.org/economic-mobility/research-framework

GUIDING PROCESS

Community Learning through Data Driven Discovery (CLD3)

- **Outer wheel**: continuous interaction and communication across stakeholders

- **Middle wheel**: data-driven learning process

- **Frontier** between these two wheels: active collaboration between all partners

- **Inner circle**: rigorous research framework to guide the data science

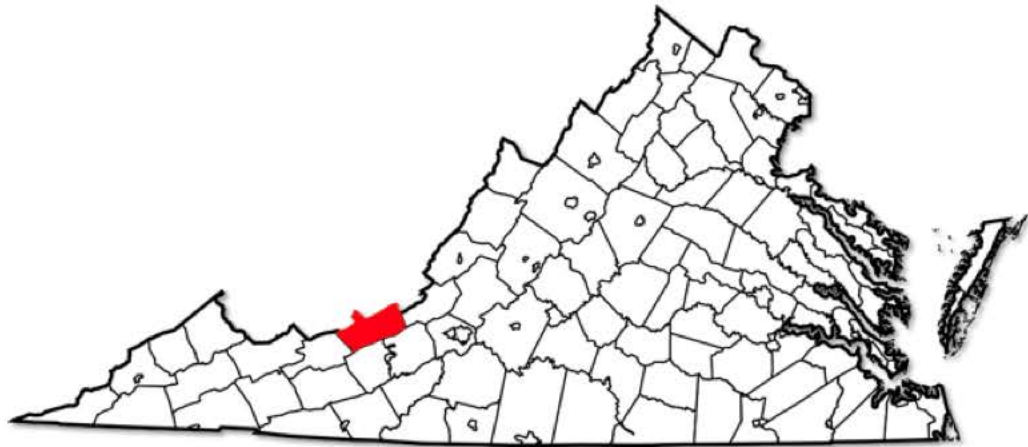COMMUNITY LEARNING THROUGH DATA DRIVEN DISCOVERY

EXAMPLES

COMMUNITY LEARNING THROUGH DATA DRIVEN DISCOVERY

GILES COUNTY, VA

BACKGROUND

POPULATION: 16,837

SIZE: 360 SQ. MILES

UNEMPLOYMENT RATE: 4.7%

ANNUAL BUDGET: $53 MILLION

RACE: 95% WHITE*

*WHITE ALONE, NOT HISPANIC OR LATINO

GILES COUNTY, VIRGINIA

ISSUE

- In 2015, Giles County officials realized reactive services make up 61% of the county's operating budget

DATA DISCOVERY: WHICH REACTIVE SERVICES ARE BEING UTILIZED?

- General District Court, Circuit Court, Jail
- Community Service, Day Report
- 911 Calls – Animal Control, Assaults, B&E, Civil Disturbance, Domestic Problem, Drunk in Public, EMS, Fire, Fight, Larceny, Mental Subject
- Department of Social Services – Child Protection, Emergency Removals, Foster Care Placements, Investigations, Relative Placements, Service Cases

TEAM

- Local government, Cooperative Extension professionals, community stakeholders

Engage community stakeholders to identify issues

Partner with experts

Discover Data

Integrate Data

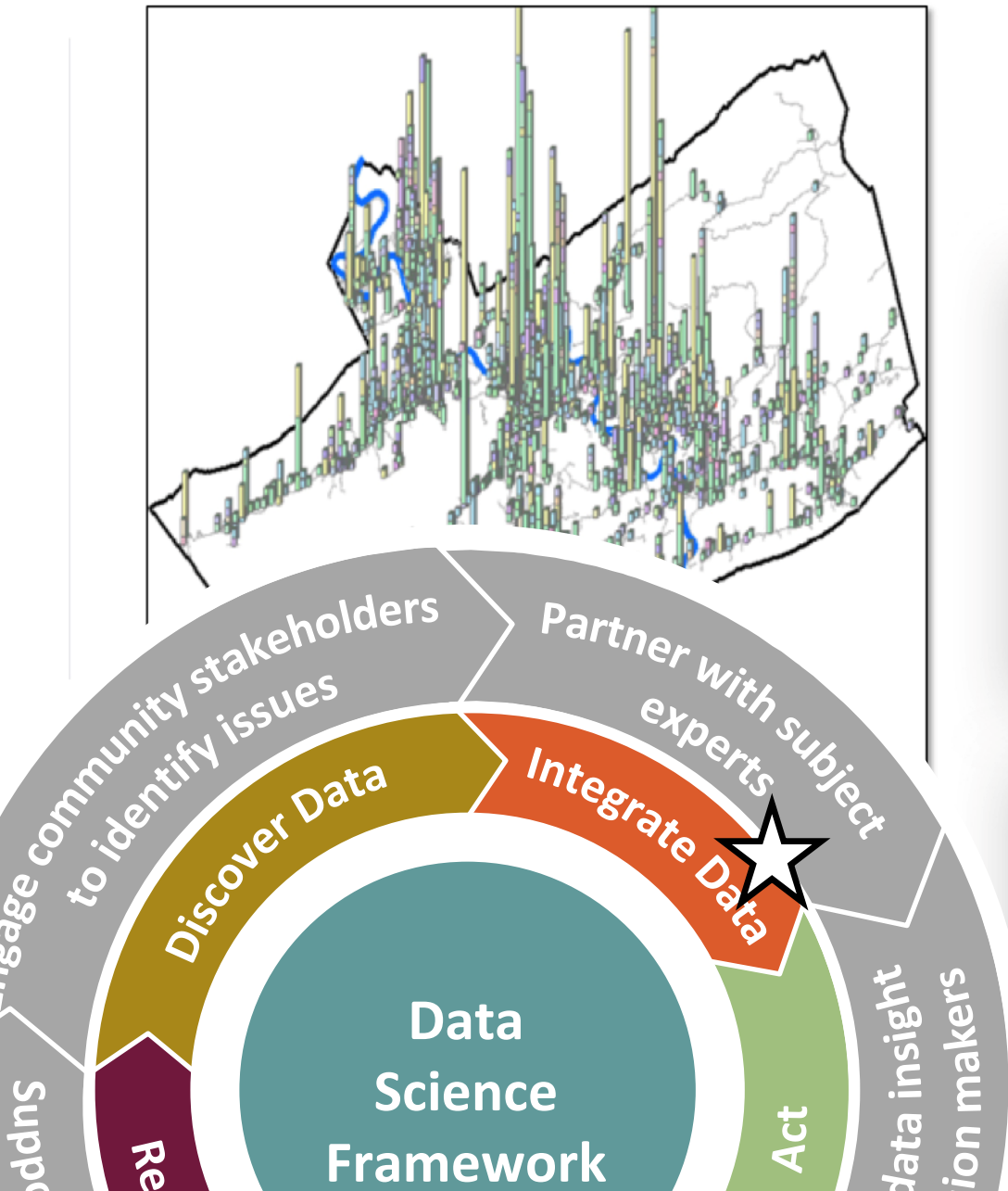Support change in ...

Redirect

Data Science Framework

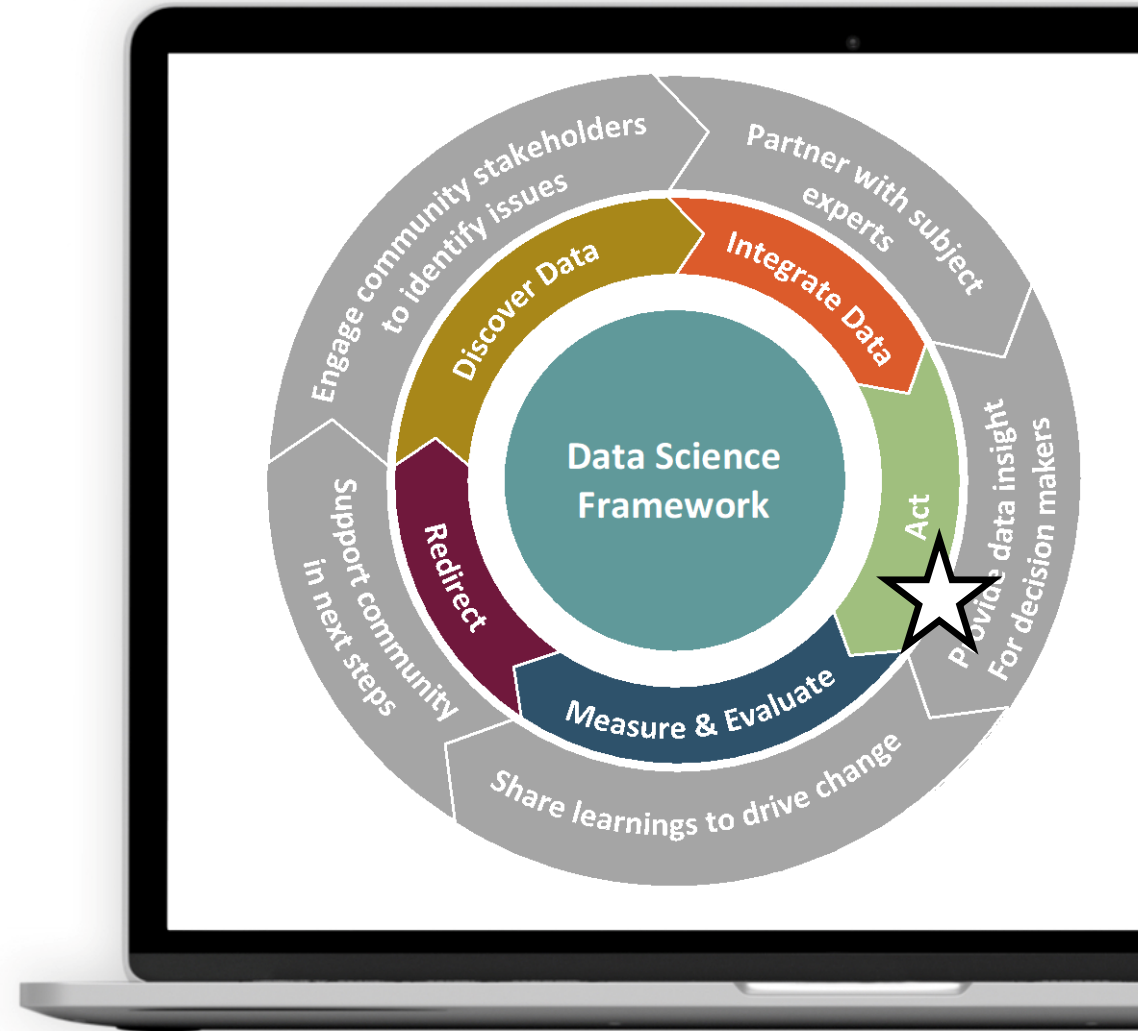# LINKING SERVICE REQUESTS TO LOCATION

FINDINGS

- 1% of the locations utilized 25% of the services
- 5% of the locations utilized 55% of the funding
- 10% of the locations utilized 71% of funding

GIS locations of the 5,279 reactive service requests in 2014 across the county. Color indicates the service and bar height indicates number of services at that location.

CONSIDERING OPPORTUNITIES: TAKING ACTION

- The map identified clusters where positive change could occur through more efficient delivery of reactive services

- Cooperative Extension is helping replicate approach in another county in SW Virginia

GILES COUNTY F.O.C.U.S. PROGRAM
"FOCUS ON FAMILIES UTILIZING SERVICES"
GEOGRAPHIC DATA-DRIVEN SERVICE INTERVENTIONS

COUNTY DAY REPORTER CENTER AS AN ALTERNATIVE TO INCARCERATION

- 31% reduction of recidivism rate

INMATE WORK PROGRAM

- Annual increase of 15,000 hours of community service

TANF WORK PROGRAM

- 44% reduction in families on TANF
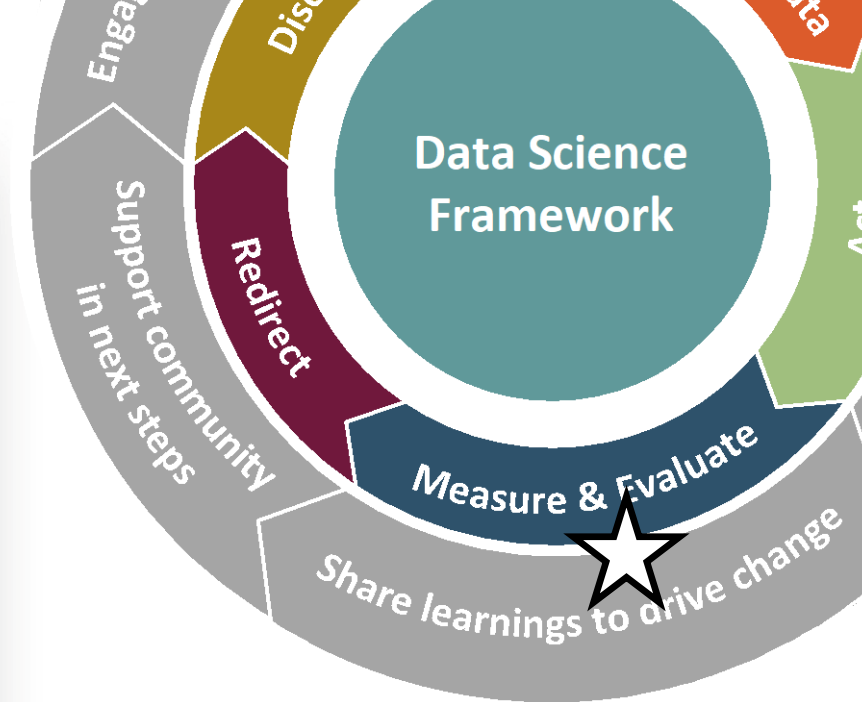
SUMMER YOUTH WORK PROGRAM

- 50 students and Over 90 local businesses are involved

ENHANCED PRE-KINDERGARTEN EDUCATION

- County continues to add new slots to accommodate more children
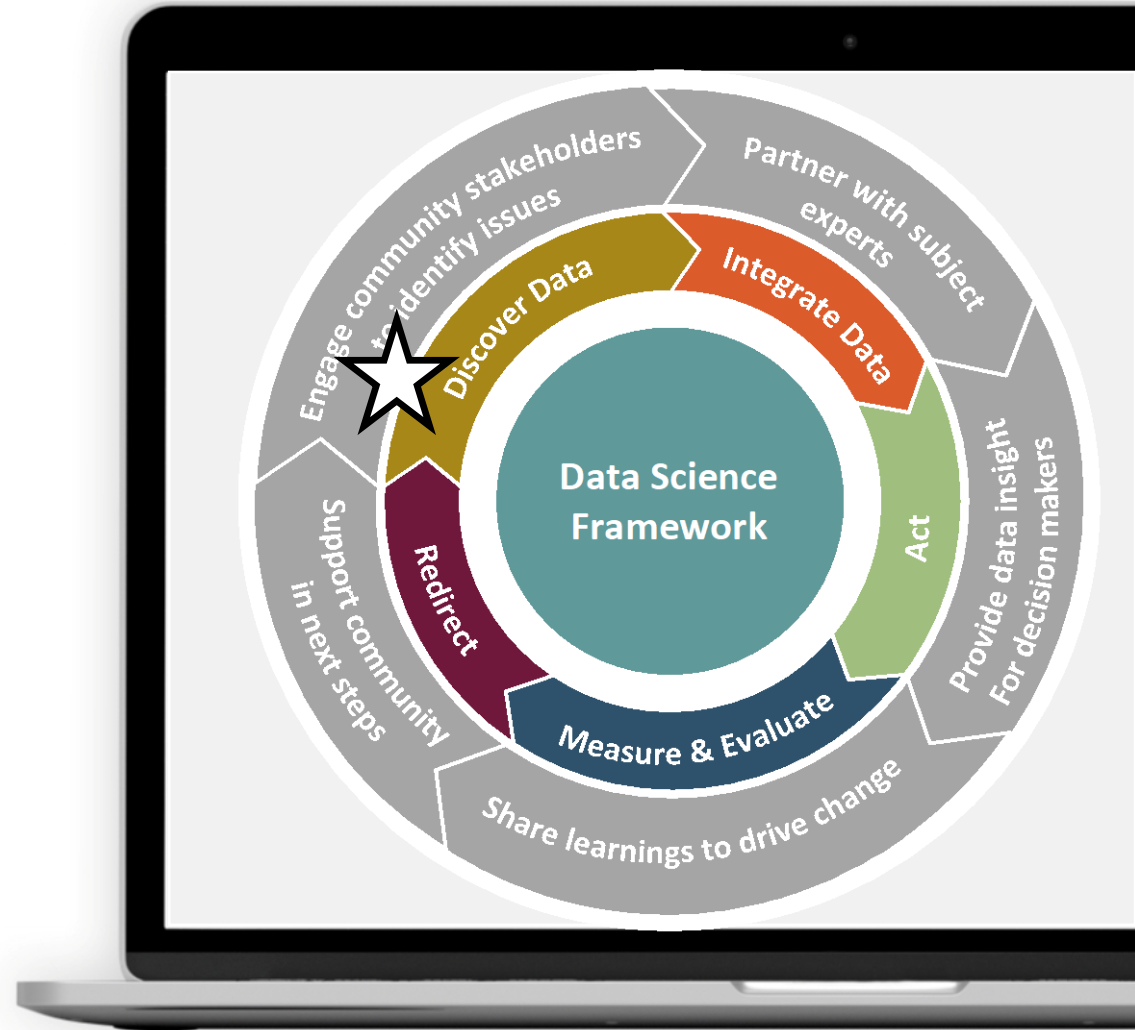
ACCESS TO COMMUNITY COLLEGE EDUCATION

- 81 new students enrolled

Data Science Framework

Engage

Disc

Data

Support community in next steps

Redirect

Act

Measure & Evaluate

Share learnings to drive change

TWO YEARS LATER: THE RESULTS

DATA DRIVEN DISCOVERY PROCESS

Adopted from materials provided by Dr. Sallie Keller and team at
The Social Decision Analytics Division,
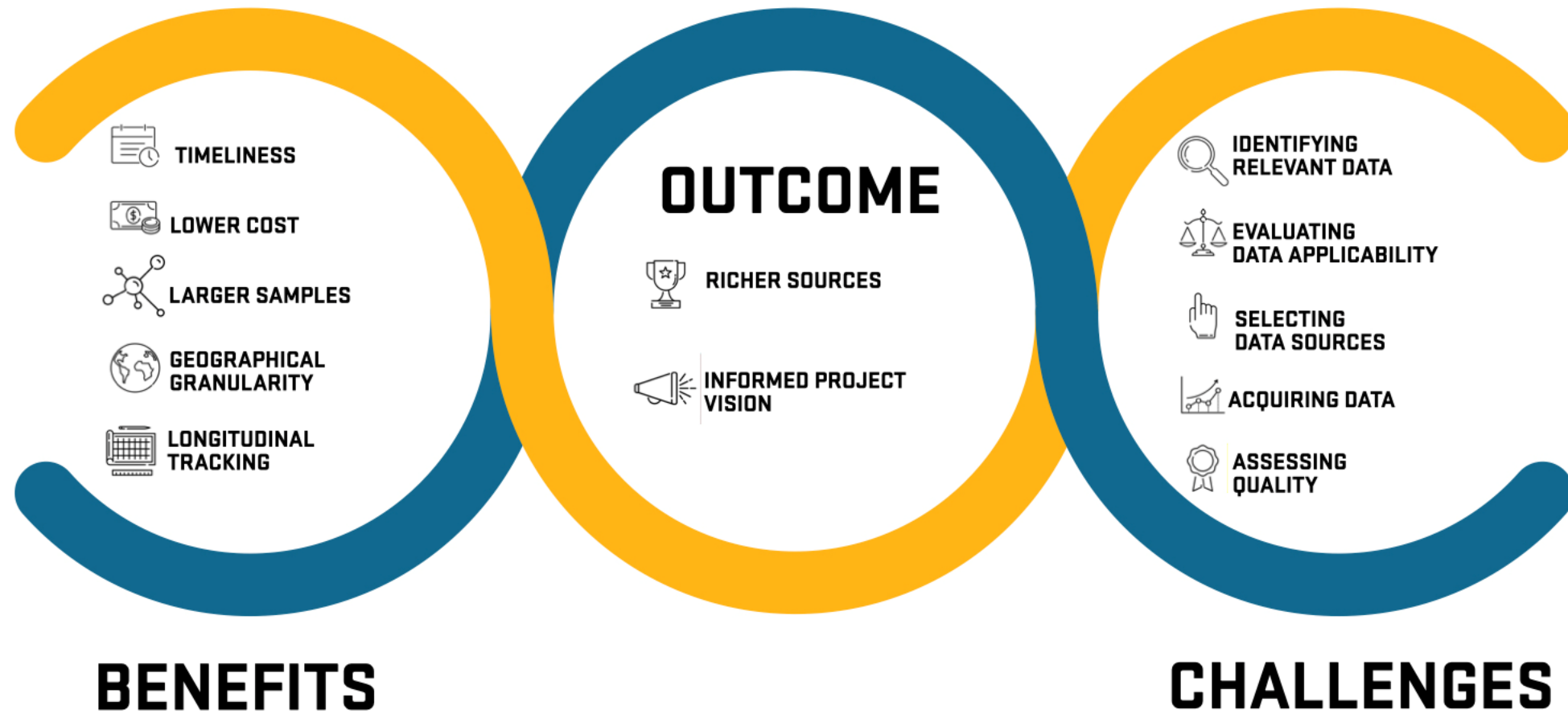Biocomplexity Institute, University of Virginia

DATA DISCOVERY
PROCESS & DEFINITIONS

- **DATA DISCOVERY** is the open-ended and continuous process by which potential datasets are identified

- **DATA INVENTORY** refers to the broadest, most far-reaching 'wish list' of information pertaining to your 'big' question

- **DATA SCREENING** is an evaluative process by which eligible datasets are sifted from the larger pool of candidate datasets

- **DATA ACQUISITION** is the method of physically acquiring the data, which can involve cooperative or purchase agreements, web-scraping, and file-transfer protocols
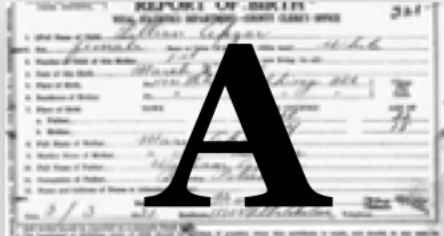
DATA DISCOVERY

DATA INVENTORY

DATA SCREENING

DATA SOURCE

DATA ACQUISITION

PROJECT'S DATA ASSETS

DATA DISCOVERY: BENEFITS & CHALLENGES

TIMELINESS

LOWER COST

LARGER SAMPLES

GEOGRAPHICAL GRANULARITY

LONGITUDINAL TRACKING

OUTCOME

RICHER SOURCES

INFORMED PROJECT VISION

IDENTIFYING RELEVANT DATA

EVALUATING DATA APPLICABILITY

SELECTING DATA SOURCES

ACQUIRING DATA

ASSESSING QUALITY

BENEFITS

CHALLENGES

Accessing Alternative Sources of Data for Education Research. Ziemer, Shipp, et al, 2016.

DATA INVENTORY

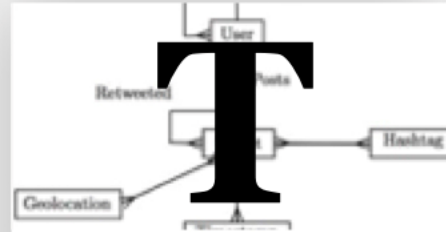## CONSIDER ALL RELEVANT DATA SOURCES:
LOCAL, STATE, FEDERAL

DESIGNED DATA

- CENSUS SURVEYS
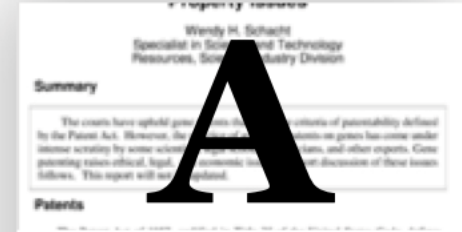- CDC SURVEYS
- BRFSS SURVEYS

ADMINISTRATIVE DATA

- BIRTH RECORDS
- ER VISIT RECORDS
- WASTE MANAGEMENT

OPPORTUNITY DATA

- GPS
- SOCIAL MEDIA

PROCEDURAL DATA

- BILLS
- SUPREME COURT RULINGS
- LEGAL BRIEFS

BEFORE YOU BEGIN

- Clearly identify the problem BEFORE you begin

- Meet stakeholders where they are at.

- Now take it a step further with the DD process.

- Share back what worked and what didn't so we can provide better support in the future

WHAT IS THE CRITICAL QUESTION

Lee County Community Needs Assessment

- According to the *Iowa Department of Human Services (DHS),* Lee County ranks 99 of 99 counties in child neglect, and 79 of 99 for child abuse. The overall county risk ranking is 91 out of 99 with a risk assessment score of 9.49

- The *2019 County Health Rankings and Roadmaps for Lee County* data reveals Lee County's Social and Economic Factors ranks the bottom 99 of 99 counties, including lower high school graduation, higher unemployment, disconnected youth, children eligible for free and reduced price lunch, children in singe parent households, injury deaths, and violent crimes.

- *Health factors*- Lee is ranked 99 out of 99 for health behaviors such as excessive drinking, alcohol impaired deaths, drug overdose deaths, and food insecurity

What environmental factors are influencing child maltreatment as well as the family risk factors continuously leading to maladaptive family behaviors and norms over time?

How can Lee County better **identify**, **monitor**, and **support** strong and resilient families?

DATA DISCOVERY
LET'S TRY IT FOR OUR PROBLEM!

ISSUE

How can Lee County better identify, monitor, and support strong and resilient families?

STEP ONE: DATA INVENTORY

Team will brainstorm dataset 'wish list' related to identifying and monitoring strong & resilient families in Lee County

STEP TWO: DATA SCREENING

Teams will brainstorm screening criteria used to assess data

DATA INVENTORY EXERCISE
GOAL: WISHLIST OF DATA SET IDEAS

TOWARDS A NATIONAL COMMUNITY LEARNING NETWORK

BRAINSTORM ACTIVITY
45 MINUTES—RECORD ALL IDEAS!

What **information** and **data** would be useful to help Lee County better identify, monitor, and support strong and resilient families?

DATA ENVIRONMENT: Strong & Resilient Families

What themes emerge from problem elaboration?

PROBLEM

How can Lee County better identify, monitor, and support strong and resilient families?

THEMES

- Child neglect
- Child abuse
- Teen births
- Poverty
- Social and economic concerns
- Poor health outcomes
- ADD RESILIANCE

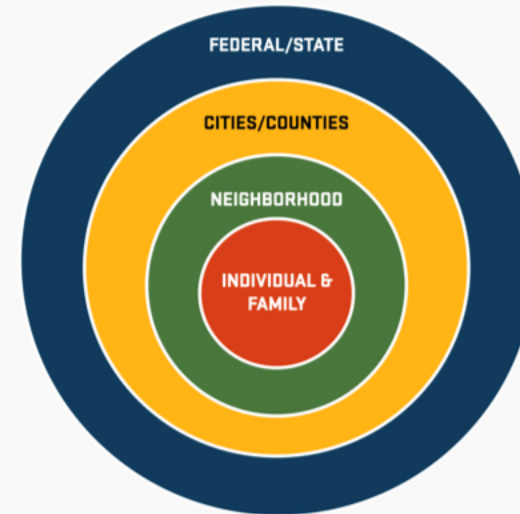**Who is engaged in or impacted by this work?**

ENGAGED

- City governments
- Community leaders
- Community agencies/programs
- State agencies/programs
- Families
- **ADD 10 sectors**

IMPACTED

- Children, families, adults
- Employers
- Medical professionals
- Criminal justice system
- Agency partners
- Community members
- School districts
- Everybody!

What internal or external factors affect this space?

FEDERAL/STATE
CITIES/COUNTIES
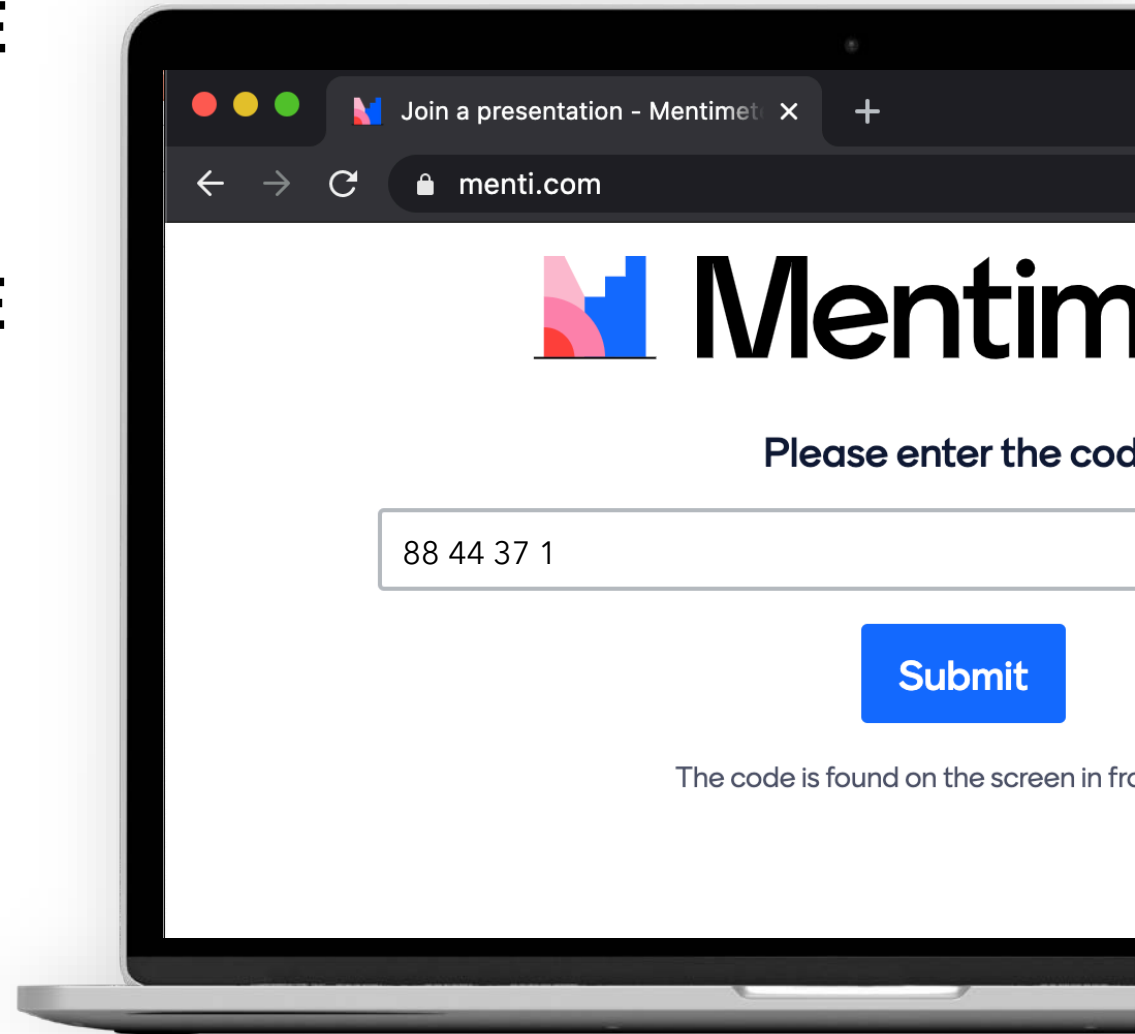NEIGHBORHOOD
INDIVIDUAL & FAMILY

- Laws and regulations
- Programs and formal supports
- Informal supports
- State or community resources
- Stakeholder preferences
- Individual and family characteristics

What is the desired outcome of this process?

Are any patterns emerging? Do these suggest the type of data domains you should be considering?

If you had access to a database on strong & resilient families, how might you use it to improve decision making and programming?

GO TO www.menti.com

ENTER CODE 88 44 37 1

DATA MAP: Strong & Resilient Families

What are your levels of analysis? What are the relationships or connections between each domain?
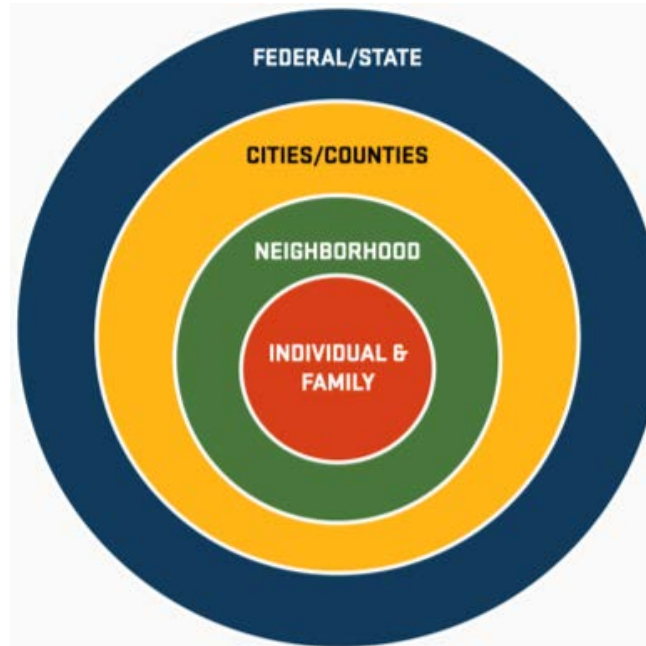
Domain: Substance Use and Health

- Use of opioids, methamphetamines, marijuana, alcohol, cigarettes/e-cigarettes
- Drug related mortality
- Drug related crime
- Drug distribution
- Mental health
- Physical health
- Risk taking
- Deaths of despair

Domain: Social Isolation

- Community characteristics
- Church & religious involvement
- Individual & community social capital
- Community & political engagement

Domain: Trauma and Coping

- Adverse Childhood Experiences (ACES)
- Incarceration, probation, juvenile detention
- Mortality
- Crime

FEDERAL/STATE
CITIES/COUNTIES
NEIGHBORHOOD
INDIVIDUAL & FAMILY
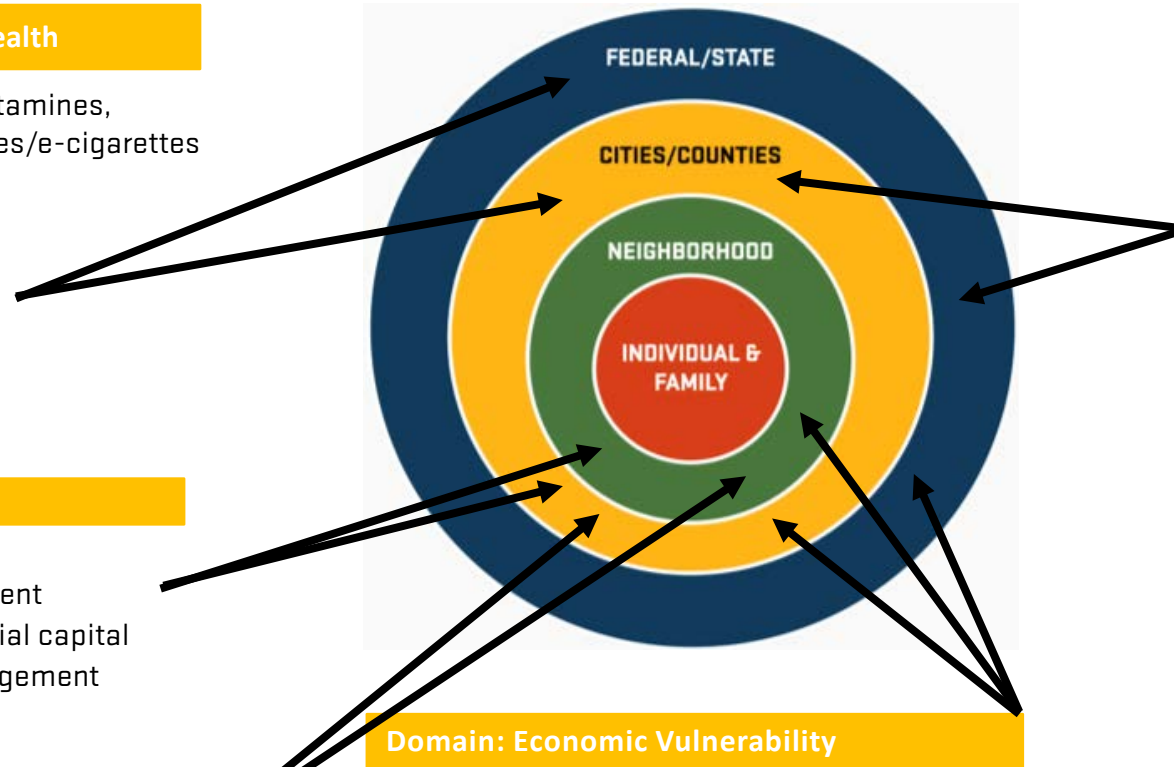
Domain: Economic Vulnerability

- Occupation
- Income
- Housing
- Employment
- Community resources
- Economic system
- General economic wellbeing

Domain: Family & Individual Characteristics

- Family characteristics
- Child/teenager characteristics
- Household characteristics
- Births and fertility
- Sexual Orientation and gender
- Immigration status
- Language

Domain:

Domain:

DATA MAP: Strong & Resilient Families

What are your levels of analysis? What are the relationships or connections between each domain?

Domain: Substance Use and Health

- Use of opioids, methamphetamines, marijuana, alcohol, cigarettes/e-cigarettes
- Drug related mortality
- Drug related crime
- Drug distribution
- Mental health
- Physical health
- Risk taking
- Deaths of despair

Domain: Social Isolation

- Community characteristics
- Church & religious involvement
- Individual & community social capital
- Community & political engagement

Domain: Trauma and Coping

- Adverse Childhood Experiences (ACES)
- Incarceration, probation, juvenile detention
- Mortality
- Crime

Domain: Family & Individual Characteristics

- Family characteristics
- Child/teenager characteristics
- Household characteristics
- Births and fertility
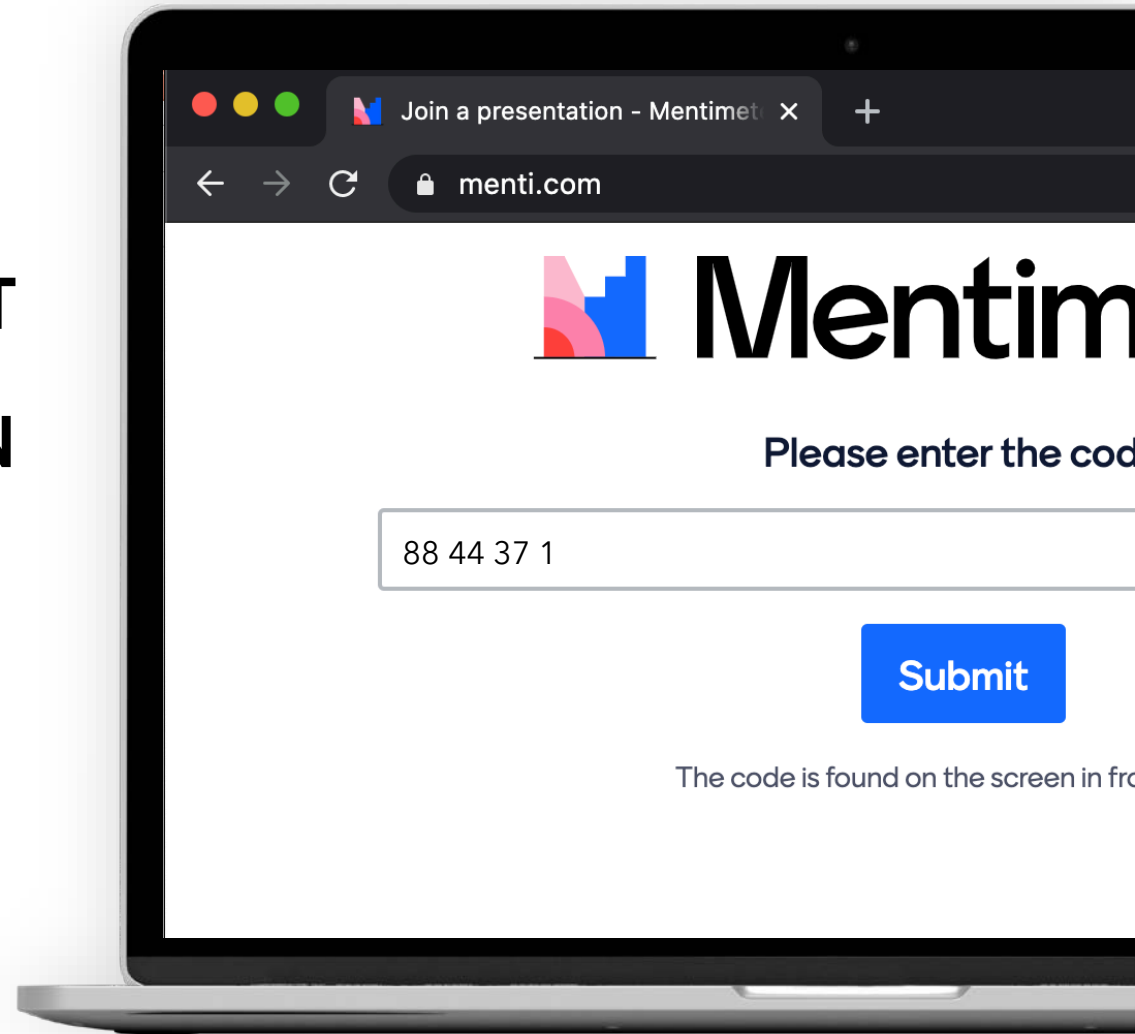- Sexual Orientation and gender
- Immigration status
- Language

Domain:

Domain:

Domain: Economic Vulnerability

- Occupation
- Income
- Housing
- Employment
- Community resources
- Economic system
- General economic wellbeing
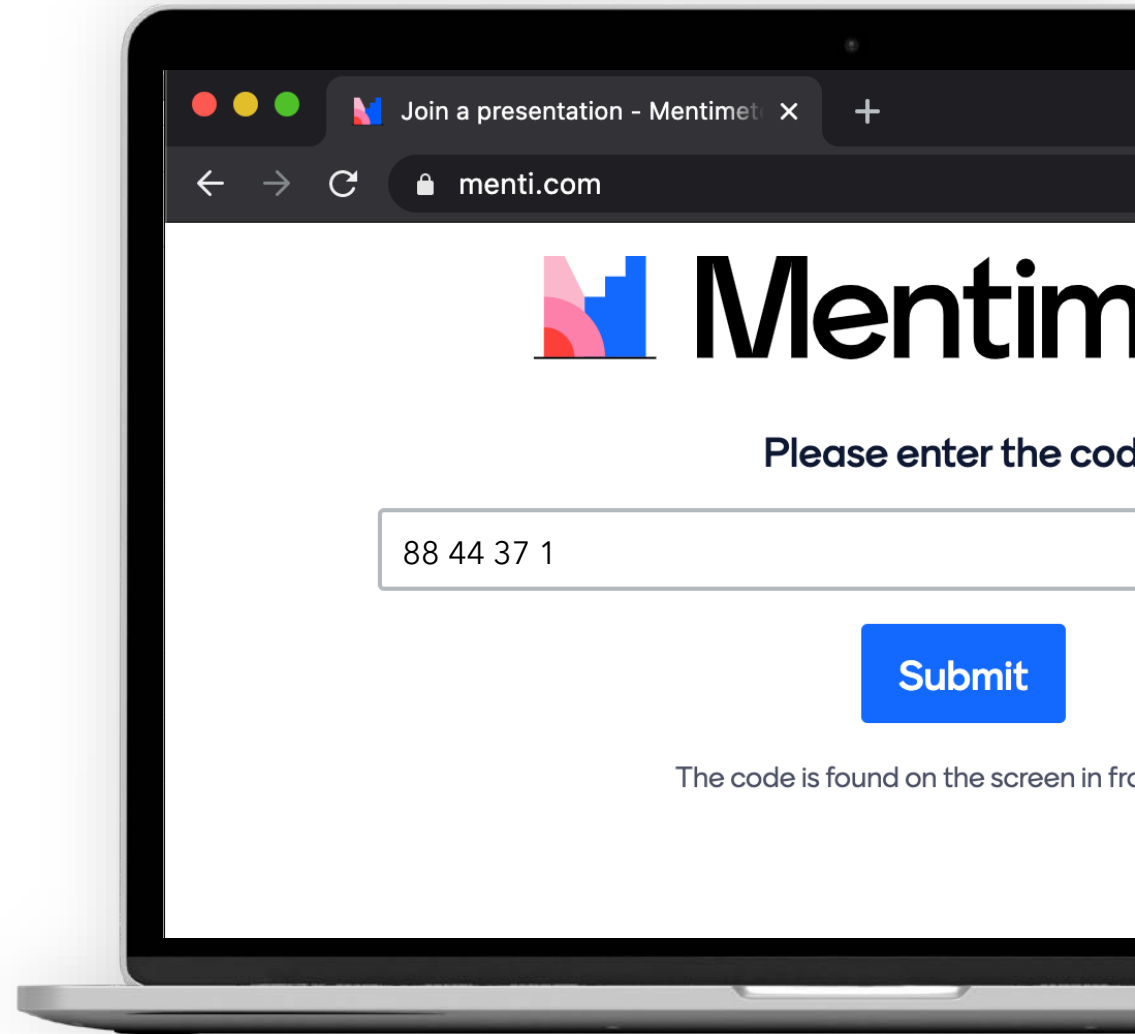
WHAT TYPES OF INFORMATION SHOULD WE CONSIDER TO HELP US BETTER UNDERSTAND AND SUPPORT STRONG AND RESILIENT FAMILIES IN LEE COUNTY?

GO TO www.menti.com

ENTER CODE 88 44 37 1

WHAT DATA DO YOU KNOW OF THAT WOULD HELP US **IDENTIFY**, **MONITOR**, AND **SUPPORT** STRONG AND RESILIENT FAMILIES IN LEE COUNTY?

GO TO www.menti.com

ENTER CODE 88 44 37 1

DATA SCREENING: DEVELOP PROJECT SPECIFIC CRITERIA

WHAT INFORMATION **ABOUT** THE DATASETS WILL HELP YOU DETERMINE THEIR FIT?

HOW FINE GRAINED A GEOGRAPHY IS NEEDED?
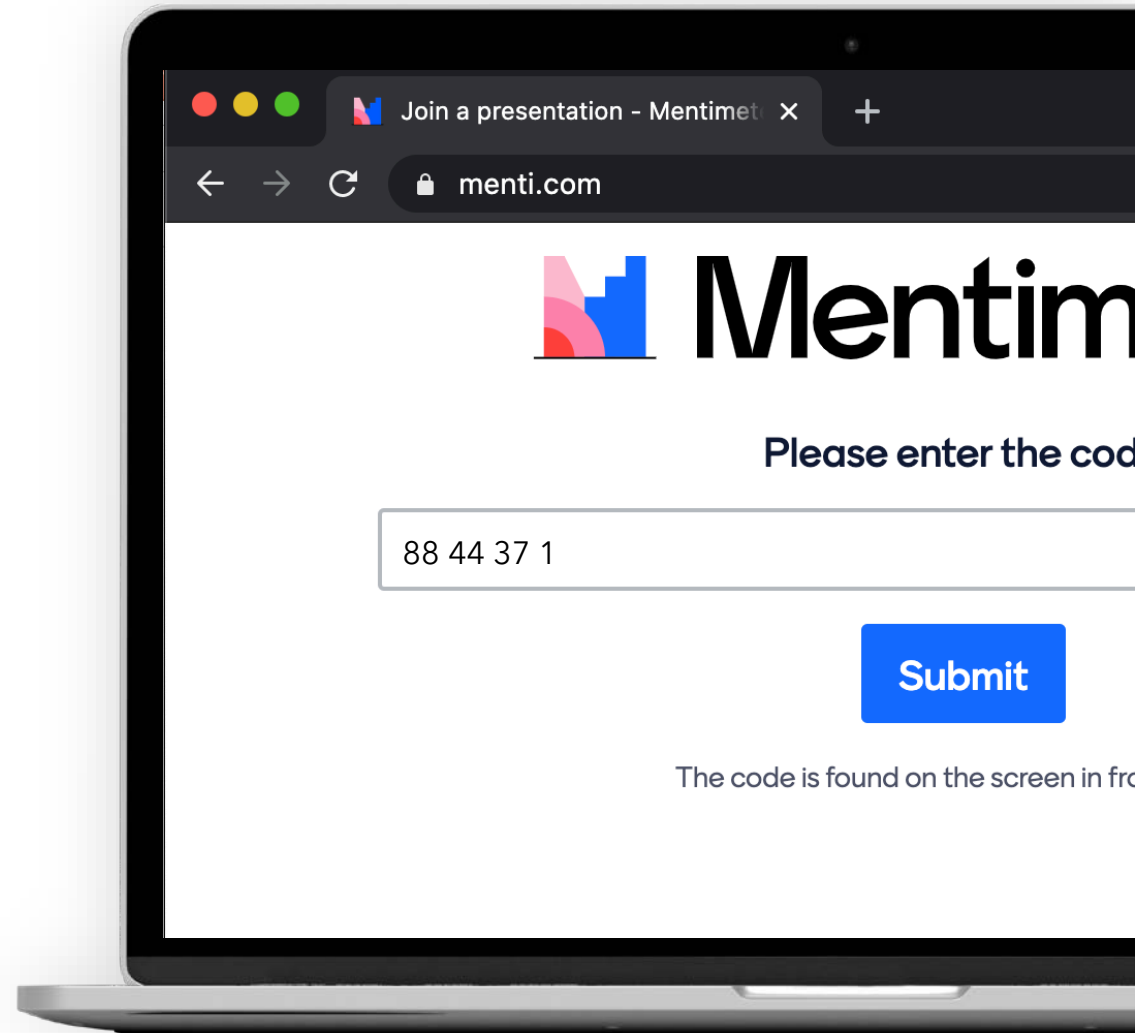

HOW FREQUENTLY ARE THE DATA COLLECTED?

WHAT IS THE DATA QUALITY? ARE THERE GAPS OR OTHER CONCERNS?

HOW DIFFICULT ARE THE DATA TO ACCESS? ARE LEGAL AGREEMENTS REQUIRED? IS THERE A FEE FOR USE?
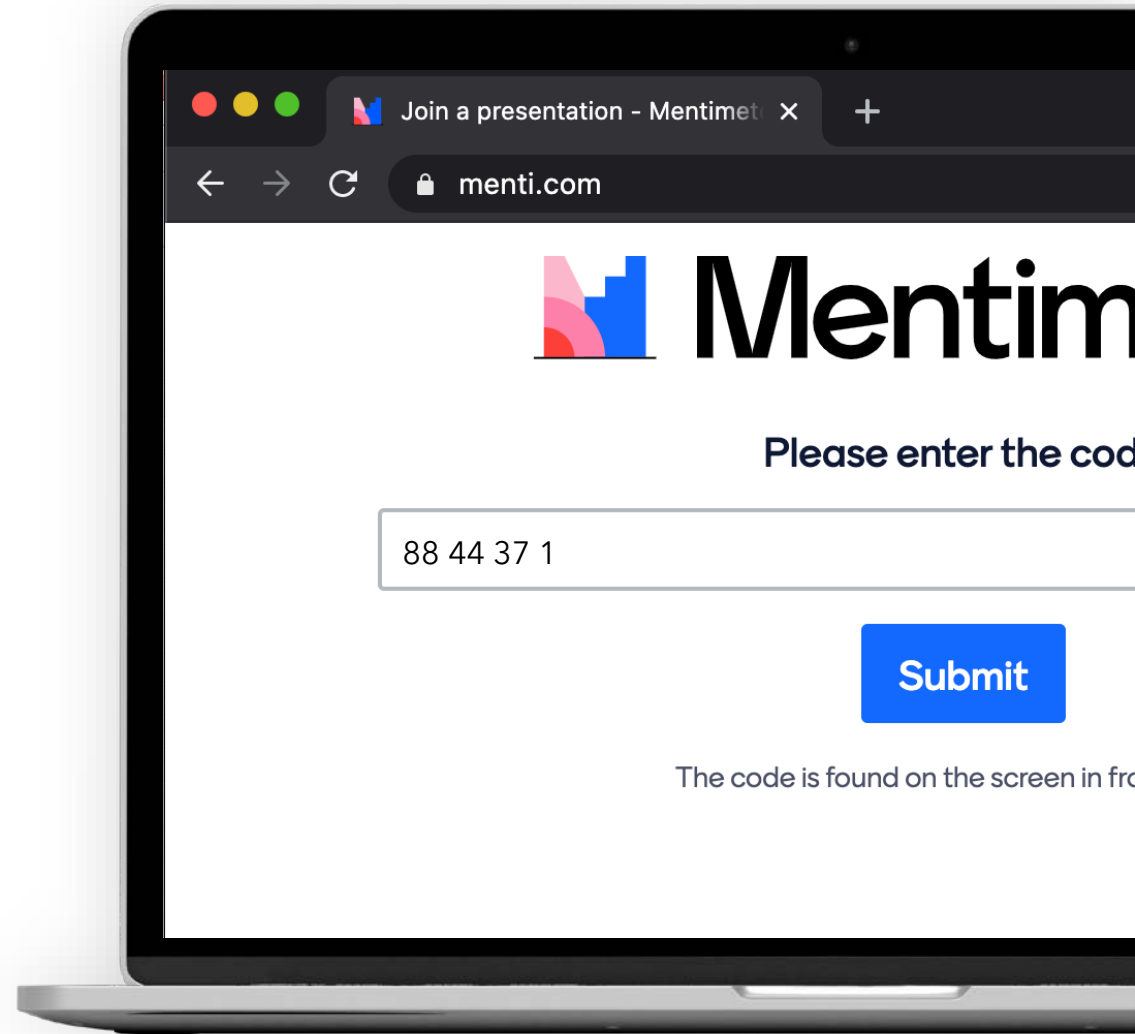
What information about the data will help determine their fit?

GO TO www.menti.com

ENTER CODE 88 44 37 1

What ethical considerations should we keep in mind related to this project?

GO TO www.menti.com

ENTER CODE 88 44 37 1

THANK YOU FOR JOINING US!

WRAPPING UP

QUESTIONS?

WE ARE HERE TO HELP!

https://datascienceforthepublicgood.com/economic-mobility

CONTACT:

CASS DORIUS

CDORIUS@IASTATE.EDU