# Better Data and Measurement for Local Public Policy Decisions

## Data Science for the Public Good Symposium 2022

Joseph Salvo, Institute Fellow
Social and Decision Analytics Division
Biocomplexity Institute & Initiative
University of Virginia

# Applied Demography: The NYC Perspective

1. Being proactive about getting people counted in the census.
2. Empower communities through access to quality data.
3. Apply scientific methods to data development and analyses in support of informed decisions.
4. Learn from colleagues, evolve, and mentor the next generation.

# Being Proactive Using Local Data and Knowledge: A Case in Point

## Simplified, the Decennial Census Needs To:

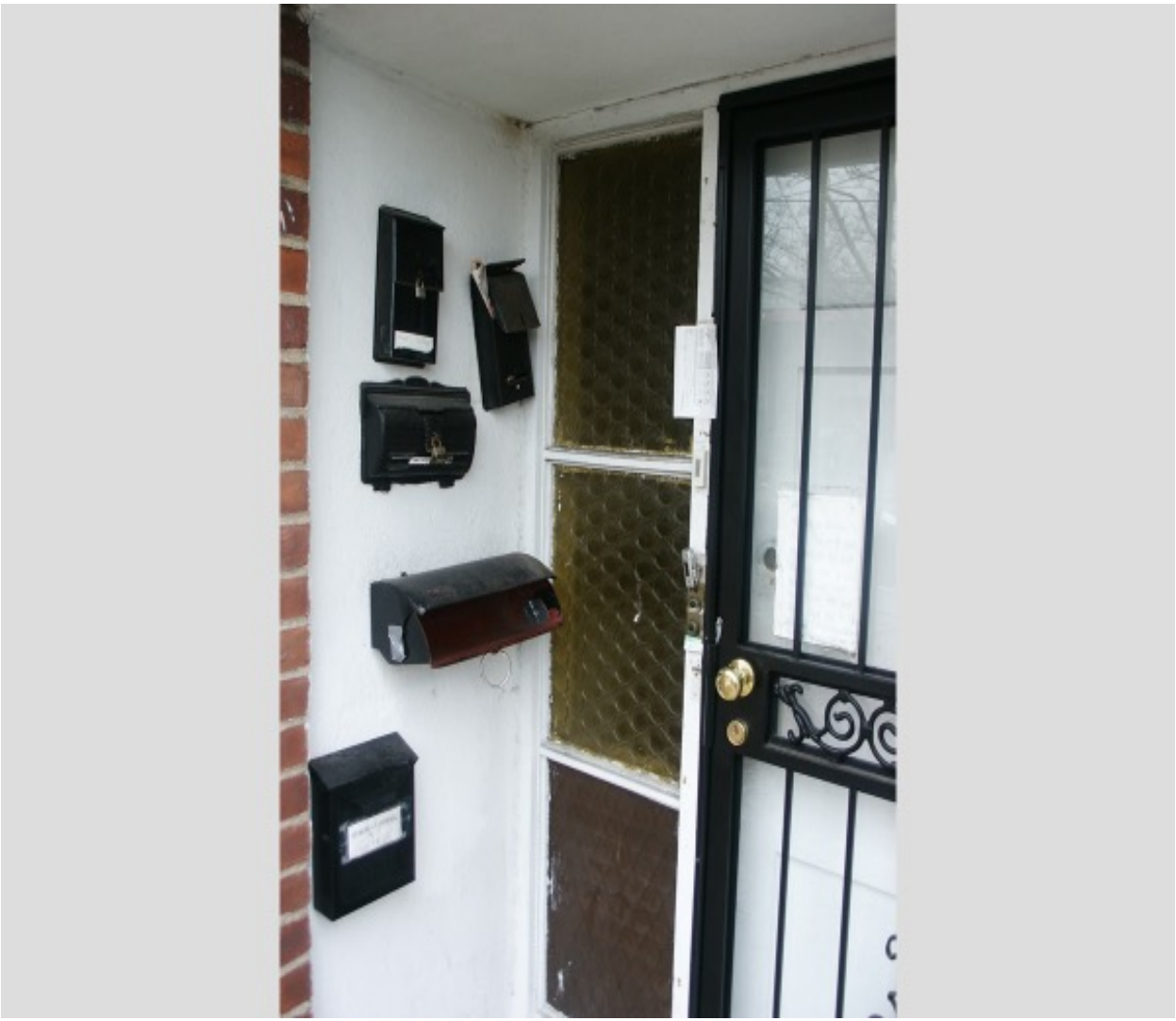
1. Create a list of addresses (NO address, NO mailing/questionnaire, NO count)
2. Check addresses in the field or in office (Address Canvassing)
3. Mail to addresses with a request to respond via the internet/mail back their questionnaire
4. Follow-up on all nonresponding households (NRFU)

# The Origins of the Local Update of Census Addresses (LUCA)Program

- The 1990 Census had an undercount of 244,000 persons in New York City, or 3.2 percent of the population.
- The LUCA program was created by Congress in 1994 to bring local knowledge to bear on the conduct of the census.
- LUCA authorizes the Census Bureau to share the list of addresses for the census from their Master Address File (MAF) with local governments. The locals must sign an agreement to keep the addresses confidential.
- Local governments can use a variety of data sources to update the address list.

# Administrative Records: 2000 LUCA

- Bell Atlantic Telephone Data
- Consolidated Edison Electrical Accounts
- Finance Department Real Property Address Directory
- Buildings Department Certificates of Occupancy
- Building Violations
- Fieldwork (evaluation and verification)



# THE CENSUS/NEW YORK; CITY POPULATION TOPS 8 MILLION IN CENSUS COUNT FOR FIRST TIME

By SUSAN SACHS MARCH 16, 2001

| | Total NYC | Bronx | Brooklyn | Manhattan | Queens | Staten Island |
|---|---|---|---|---|---|---|
| Total "Added" Addresses | 439,069 | 58,310 | 126,127 | 69,772 | 161,027 | 23,833 |
| As a Percent of Housing Units | 13.1 | 11.4 | 12.9 | 8.4 | 18.6 | 14.2 |

# Purpose and Use

- Rural Communities
- Family Units
- Vulnerable Populations
- Elderly Residential Care
- The Unhoused
- Local Businesses
- Tribal Areas
- Counting Children

- Differential Undercount
- Climate Change
- Changing Demographics
- Broadband Deployment
- Commuting Patterns
- Economic Mobility
- Gig Economy
- Equity and Inclusion
- Migration
- Emergency Preparedness
- Workforce Automation

# Making Sense of the Headlines

## Gig Workers

"Racial and ethnic differences stand out in the U.S. Gig workforce." *Pew Research Center*, December 2021.

"The Wild West: Gig workers on the front lines of the coronavirus pandemic lack basic worker protections." *Economic Policy Institute*, April 2020.

## Migration Patterns

"In 2020, fewer Americans moved, exodus from cities slowed." *Pew Research Center*, December 2021

"New census data shows a huge spike in movement out of big metro areas during the pandemic." *Brookings,* April 2022

## Nursing Homes

"COVID-19 crisis advances efforts to reimagine nursing homes," *JAMA Network*, October 2021.

"Covid awakened Americans to a nursing home crisis. Now comes the hard part." *Washington Post*, April 2021.

# Introducing the Curated Data Enterprise

The U.S. Census Bureau has begun modernization activities towards creating a **universal frame** that could form the foundation for a transformative capability to **integrate** the Bureau's demographic and economic **surveys and censuses with administrative and third-party data.**

This new innovative state is called the **Curated Data Enterprise (CDE)**.

CURATED DATA ENTERPRISE

*The Census Bureau of the Future*

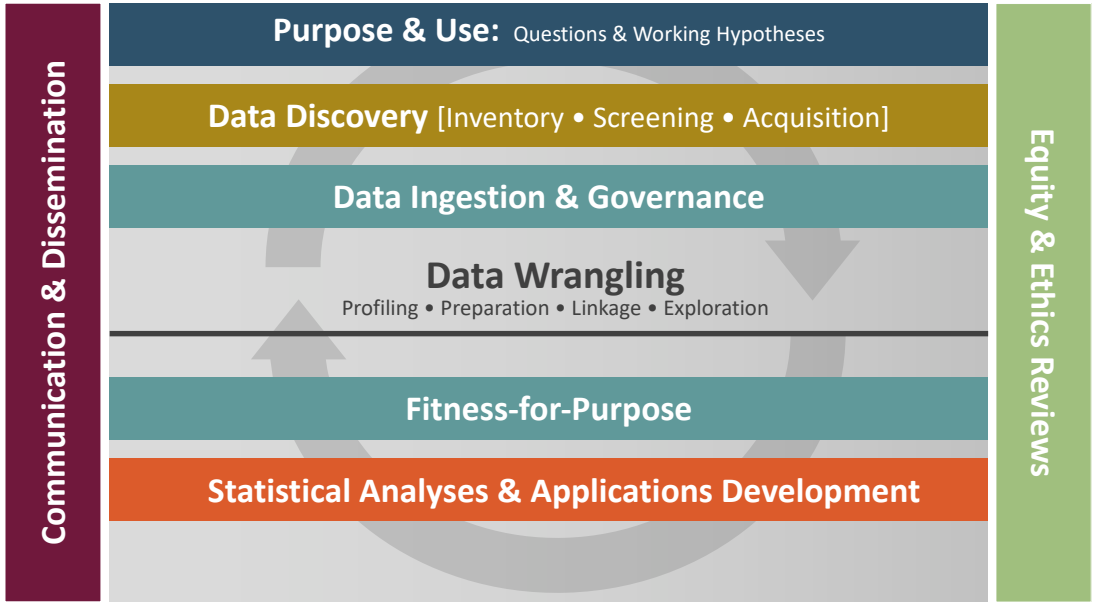
- Both an infrastructure and a continuously evolving ambition to **empower and enable** Bureau scientists and their data users to develop **new measures** of people, places, and the economy.
- Exploits **multiple data sources**, from sample surveys, censuses, and other official administrative and third-party data, which can provide **more robust, timely, and comprehensive measures** when combined.

# Research Components of the CDE

Goal: Illuminate the needed capabilities of the CDE and support its development

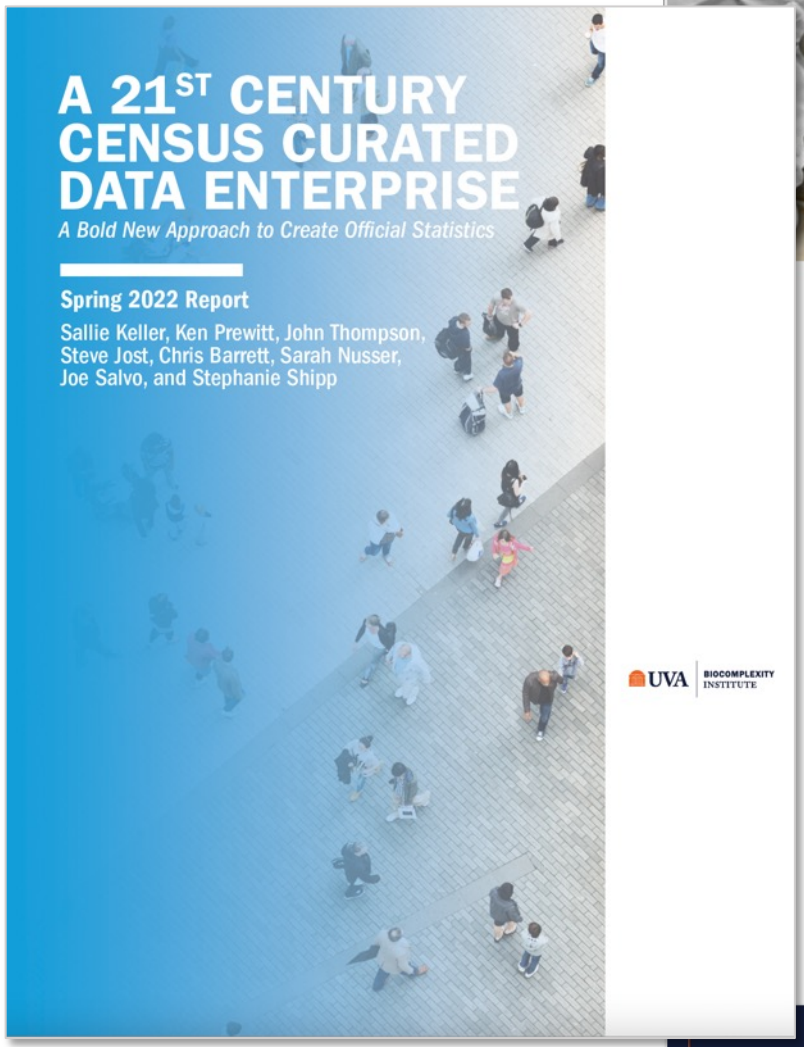
- Start with how to capture purpose and use
- Develop information collection processes for each box in the CDE framework
- Give insights into standards and automation
- Develop theory and methods

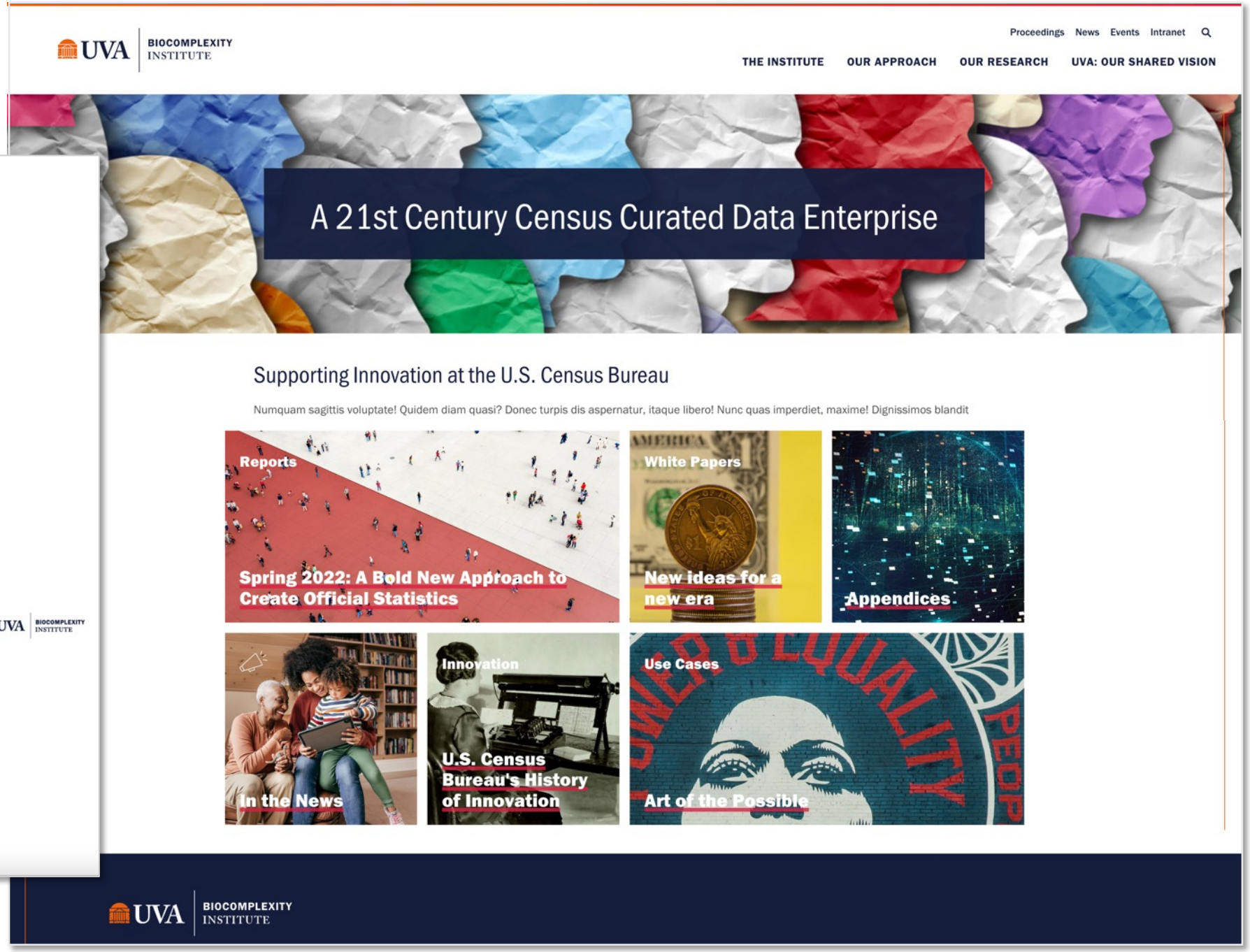
Curated Data Enterprise Framework




*Curation applied at every step*

https://biocomplexity.virginia.edu/institute/divisions/social-and-decision-analytics/census



# Thinking Like a "Local"

More than 70 percent of the 31,600 Places in the U.S. have fewer than 4,000 persons.

Local Planners as Consumers of data: Knowledge of population and the local economy permit an assessment of decisions and strategies adopted by local officials.

Local Planners as Directors: Decisions and strategies affect the quality-of-life of residents.

# Reacting to the Pandemic: Questions and Strategies

Thinking Like a “Local” - Three Key Knowledge Areas for the local town planner:

1. Local economy/business: How are residents able to sustain themselves economically?
2. Are residents reacting to the pandemic “with their feet?” What is the impact on population?
3. What about those who are most vulnerable, e.g., persons in nursing homes?

# Use Case - Gig Economy in Post-Pandemic World of Work

## Art of the Possible

- Complete picture of work:
  - GIG workers
  - Remote workers
  - Vulnerable workers and equity issues

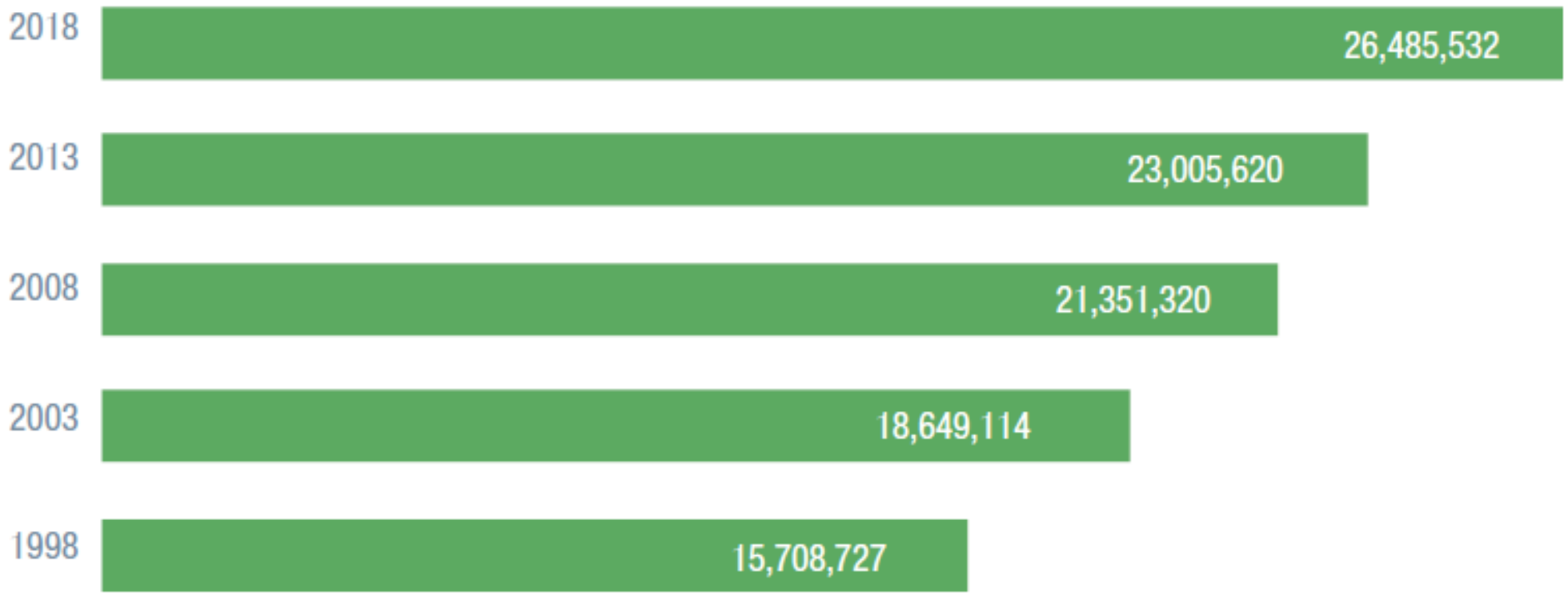
## Data Sources

- Labor Force, Occupations and Industries (ACS, CPS, LEHD)
- IRS, USPS, other administrative data
- State records that require authorization

## CDE Capability to Exercise

- Obtaining flexible geographic granularity
- Testing the limits of data partnerships.

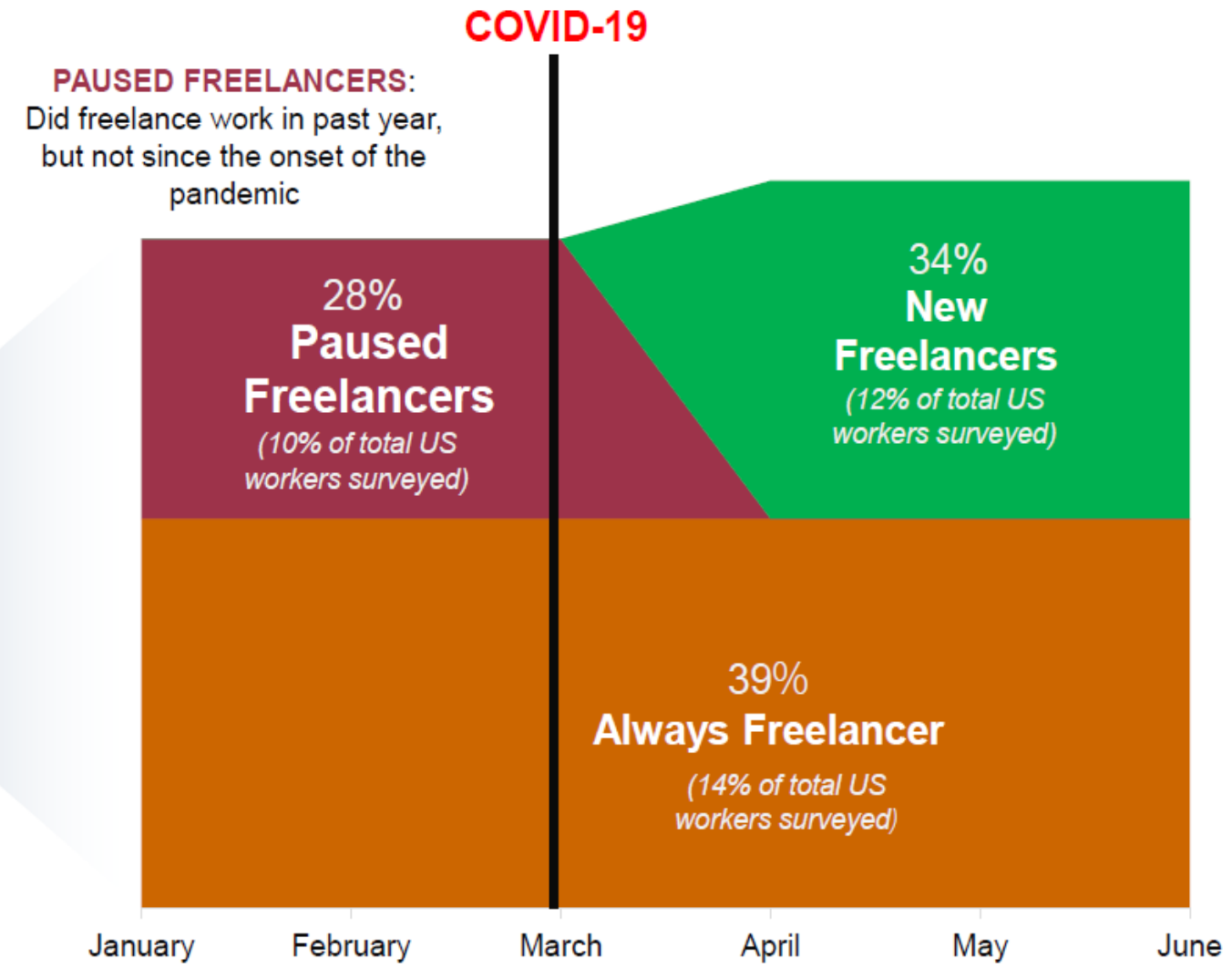
## NUMBER OF NONEMPLOYER FIRMS[3]




1 *Data for small employer firms is sourced from US Census Bureau's Statistics of US Businesses. Data for nonemployer firms is sourced from US Census Bureau's Nonemployer Statistics.*
2 *Small employer firms are businesses with 1–499 employees.*
3 *Sourced from US Census Bureau, Nonemployer Statistics.*

# AMIDST COVID-19, FREELANCING OFFERS CONSISTENCY AND NEW CAREER OPPORTUNITIES

**36% of the American workforce are freelancers**




**NEW FREELANCERS:** Started doing freelance work as of March 2020, at the onset of the pandemic

**ALWAYS FREELANCERS:** Did freelance work in the past year and still freelancing during the pandemic



S101: BEHIND THE SCENES – CLASSIFICATION QUESTION; S105: PRE/POST COVID FREELANCER CODING
Base: Total Freelancers n=2132, Always FLs n=827, Paused FLs n=591, New FLs n=714

Percent of workers with the ability to work from home by income percentile, 2018
**(Bureau of Labor Statistics)**

| **Income percentile** | **Percent** |
|---|---|
| Bottom 25 | 9.2% |
| 25-50 | 20.1% |
| 50-75 | 37.3% |
| Top 25 | 61.5% |

Source: Adapted from Gamio, 2020.

# Use Case - Domestic Migration

## Art of the Possible

- Close to real-time estimates using external data

## Data Sources

- ACS residence one year ago (in-movers); county gross migration; net migration as a residual
- Cellphone detailed records
- Postal changes of address
- Commercial mover data

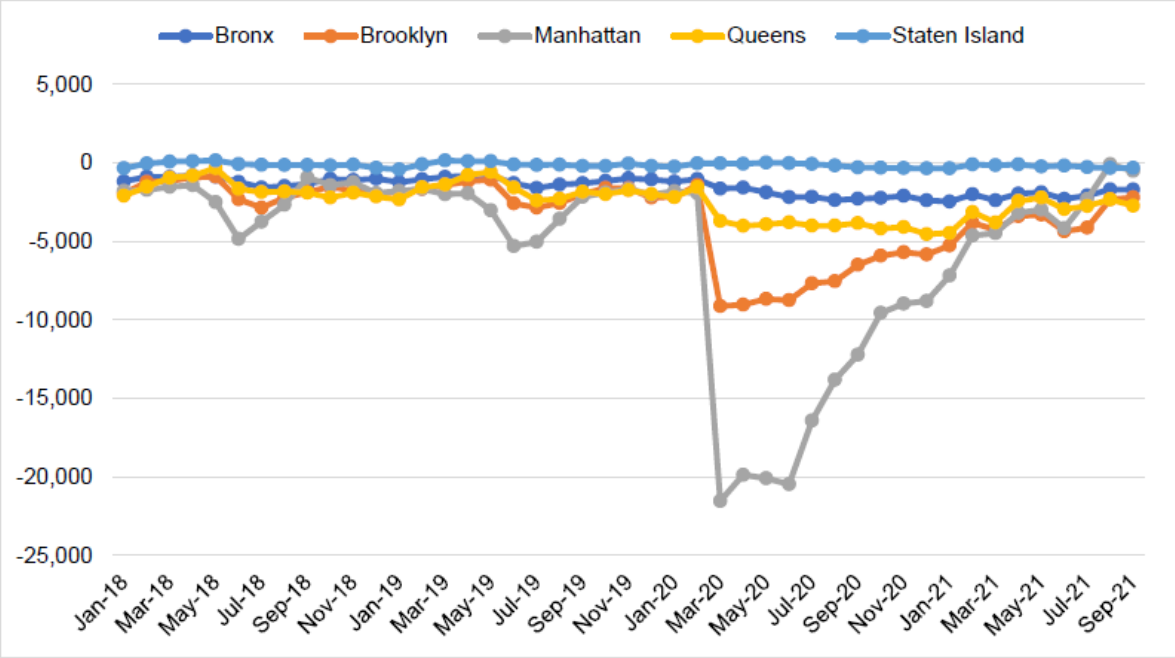
## CDE Capability to Exercise

- Overcoming barriers to the use of novel external data sources, e.g., digital records
- Build effective models to select best measures for different geographic areas (variables and overall)

# Postal Service Data, Covid-19 and Migration: New York City

**Net residential movement was highest in areas with higher incomes**

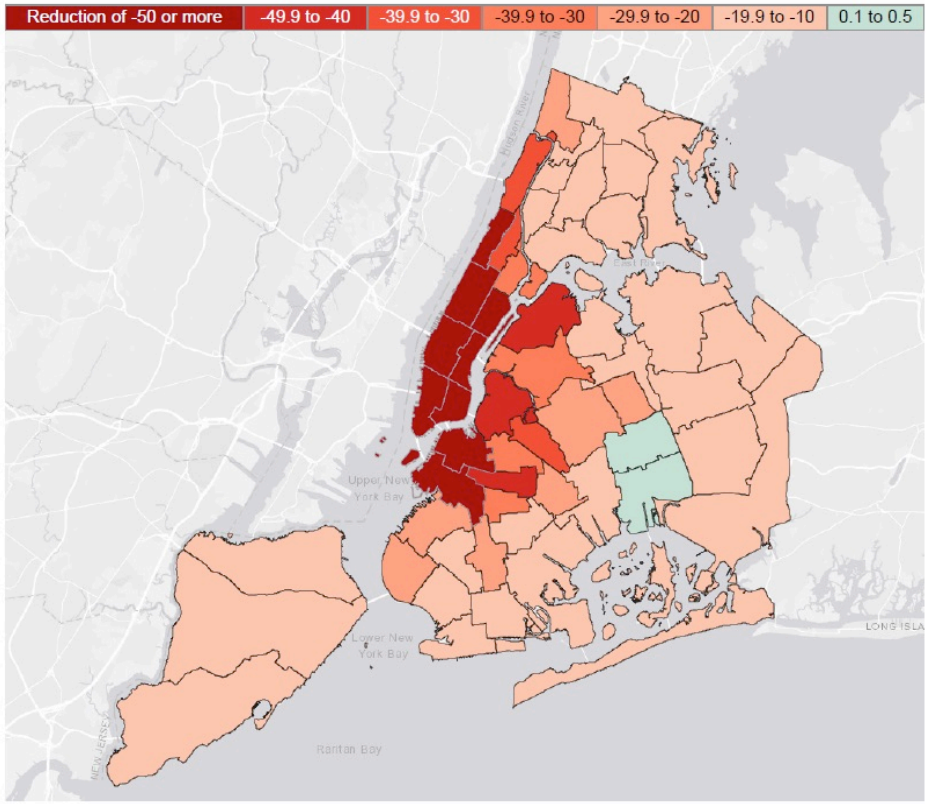
Chart 3. Monthly Net Residential Moves by Borough, 2018-2021




SOURCE: NYC Comptroller's Office analysis of United States Postal Service, "Change of Address Stats," https://about.usps.com/who/legal/foia/library.htm.

Note: Includes temporary and permanent moves.

Map 1: Change in Net Residential Moves from Pre-Pandemic Trend in 2020 (Change in Net Residential Moves per 1,000 Residents in 2020, as Compared to 2019)




SOURCE: NYC Comptroller's Office analysis of United States Postal Service, "Change of Address Stats," https://about.usps.com/who/legal/foia/library.htm.

Note: Per-capita figures based on 2020 decennial census data. Includes temporary and permanent moves.

# Use Case – Vulnerable Populations in Nursing Homes

## Art of the Possible

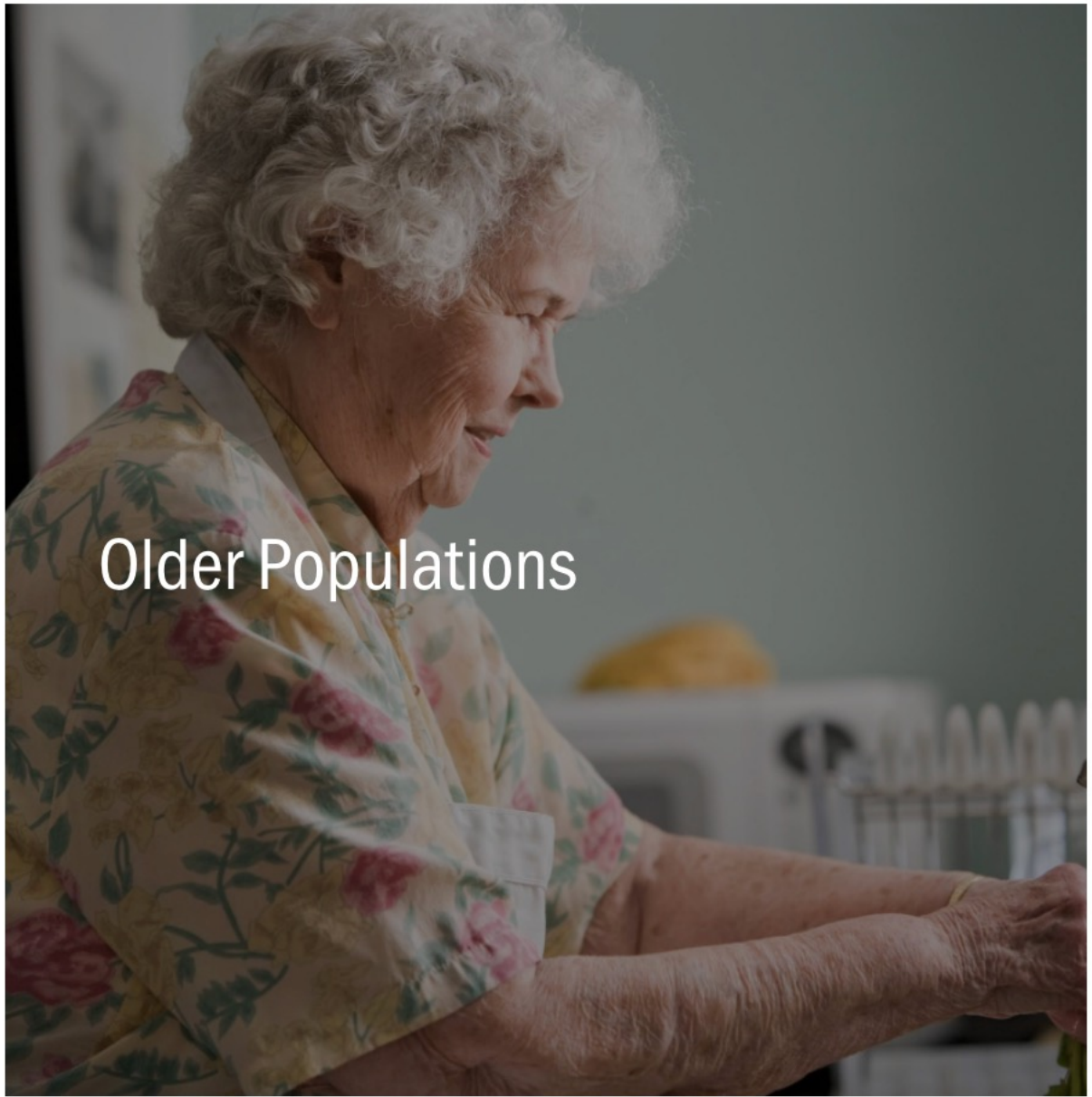
- Comprehensive picture of elder care
  - Residents
  - Staff
  - Owners

## Data Sources

- Group Quarters (GQ) population-based on decennial census, ACS
- CMS administrative data
- BLS, CDC, NIH surveys

## CDE Capability to Exercise

- Develop cross-agency partnerships
- Integrating data with different units of analysis, timing, geography



# Successful Data Discovery and Acquisition will Require Partnerships with Federal, State, and Local Agencies, the Private Sector, and Non-Profit Organizations

Federal

- Centers for Disease Control and Prevention (CDC)
- Centers for Medicare and Medicaid Services (CMS)
- Agency for Healthcare Research and Quality's (AHRQ)

States

- State Department of Public Health divisions that are responsible for oversight and maintaining data about Skilled Nursing Home Facilities
- State Department of Social Services or Public Health divisions that are responsible for oversight and maintaining data about CCRC and assisted living facilities

Foundations

- Kaiser Family Foundation
- American Health Care Association
- National Center for Assisted Living

| Exhibit 2. Kaiser Family Foundation - Providers and Service Use Indicators by state and overall[19] |
| --- |
| *Aggregates CMS Nursing Home data by state.*<br>○ Total Number of Certified Nursing Facilities<br>○ Total Number of Residents in Certified Nursing Facilities<br>○ Average Number of Certified Nursing Facility Beds<br>○ Average Nurse Hours per Resident Day in All Certified Nursing Facilities<br>○ Distribution of Certified Nursing Facilities by Ownership Type<br>○ Distribution of Certified Nursing Facilities by Certification Type<br>○ Distribution of Certified Nursing Facility Residents by Primary Payer Source<br>○ Distribution of Certified Nursing Facilities by Chain-Owned Affiliation<br>○ Distribution of Certified Nursing Facilities by Hospital-Based Status<br>○ Certified Nursing Facility Occupancy Rate<br>○ Percent of Certified Nursing Facilities with Resident Groups and Family Groups<br>○ Average Number of Deficiencies per Certified Nursing Facility<br>○ Percent of Certified Nursing Facilities Receiving a Deficiency for Actual Harm or Jeopardy<br>○ Percent of Certified Nursing Facilities with Deficiencies<br>○ Percent of Certified Nursing Facilities with Top Ten Deficiencies |

# Applying the CDE in Response to a Climate Event

## Baseline Curated Data Enterprise

**Census Linked Frames**

- Geography
- People
- Jobs
- Businesses

**County Level Measures**

- # of nurses for every 1000 persons >= 65
- # of home health agencies for every 1000 persons >= 65
- # of hospital beds for every 1000 persons >= 65
- # of empty nursing home beds divided by the # of nursing homes

**Census Tract Level Measures**

- CDC Social Vulnerability Index
- Climate and Economic Justice Screening Tool climate disadvantaged census tract designation

**Skilled Nursing Home Measures**

**Residents**

Demographic:

- Percentage
- Minority
- Percentage > 65
- Mental Health
- Physical Health
- Cognitive Function
- Percentage on Medicaid
- Total # of residents

**Workers**

Hours:

- RN hours/ resident day
- LPN hours/ resident day
- CNA hours/ resident day

Turnover:

- Total nursing staff turnover
- # of administrators who have left

**Owners**

- Facility type
- Whether the facility is a chain
- Whether the facility changed hands in the last 12 months

A comprehensive picture of nursing homes

**Facility Safety Assessment**
Total # of safety deficiencies over the last 3 years

- Building Knowledge Base (BKB)
- Transportation infrastructure
- How workers get to the nursing home
- Worker demographics
- Where workers live

**Severe Climate Event**

**Data Products**

**Emergency Preparedness**
Total # of emergency preparedness deficiencies over the last 3 years

- Hurricane, flooding, and riverine flooding predictions
- Past hurricane and extreme flooding events

***Building on the baseline to address specific questions***

# What is the "Art of the Possible?"

From the Local Planner's Perspective

## Goal is to

Forge a path that leads to evidence-based decisions to mitigate shorter-term impacts related to the pandemic and devise longer-term strategies to promote economic growth

## Based on

A timely and accurate view of the current demographic and social characteristics of the population and the economic conditions of individuals and businesses

## Using

Data products that are actionable at a local level

Strive for progress,
not perfection.
Anonymous
quotefancy

*“Unless someone like you cares a whole awful lot, nothing is going to get better. It’s not."*

*from The Lorax by Dr. Seuss*

For additional information please contact Joseph Salvo at ckp7gd@virginia.edu