

Data driven policy for pandemic response: Data, knowledge, and action well outside of the ivory tower



Jessie Tenenbaum, PhD
Duke University School of Medicine

**Chief Data Officer,
NC Dept of Health and Human
Services**



@jessiet1023

November 3, 2022

Disclaimer

“Photo” credit DALL-E

<https://labs.openai.com/>





NC DEPARTMENT OF HEALTH AND HUMAN SERVICES



[Children's Services](#)



[Deaf-Blind Services](#)



[Disability Services](#)



[Vital Records](#)



[Low-Income Services](#)



[Medicaid](#)



[Mental Health and Substance Abuse](#)



[Pregnancy Services](#)



[Refugee Services](#)



[Aging and Adult Services](#)



[Guardianship](#)

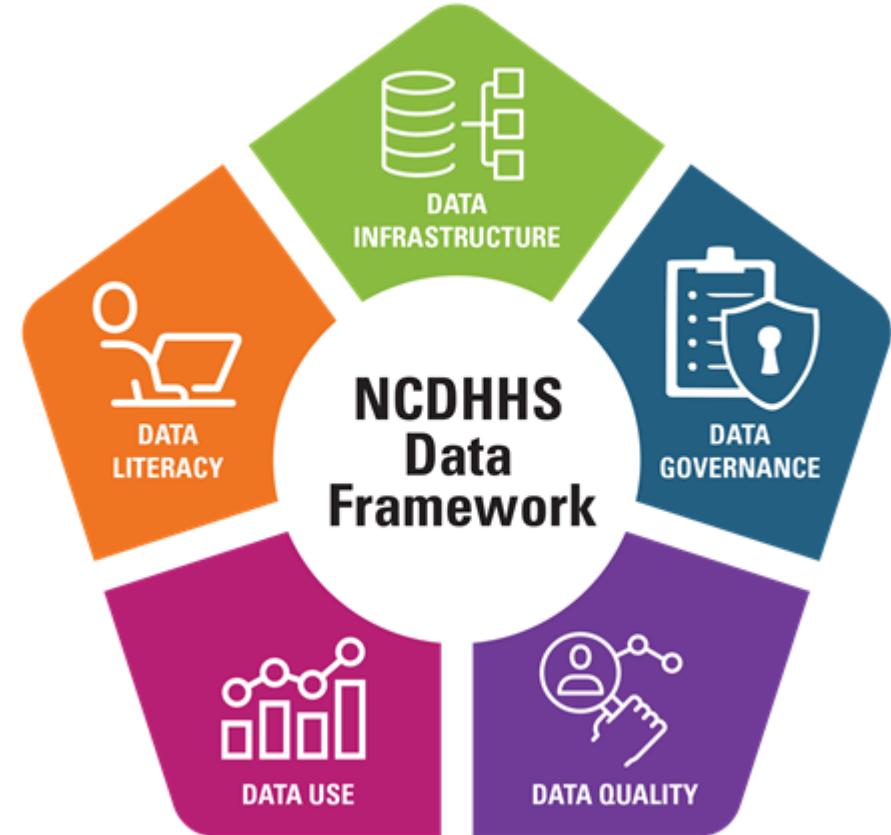


[Veteran's Services](#)

<https://www.ncdhhs.gov/>

NCDHHS Data Office

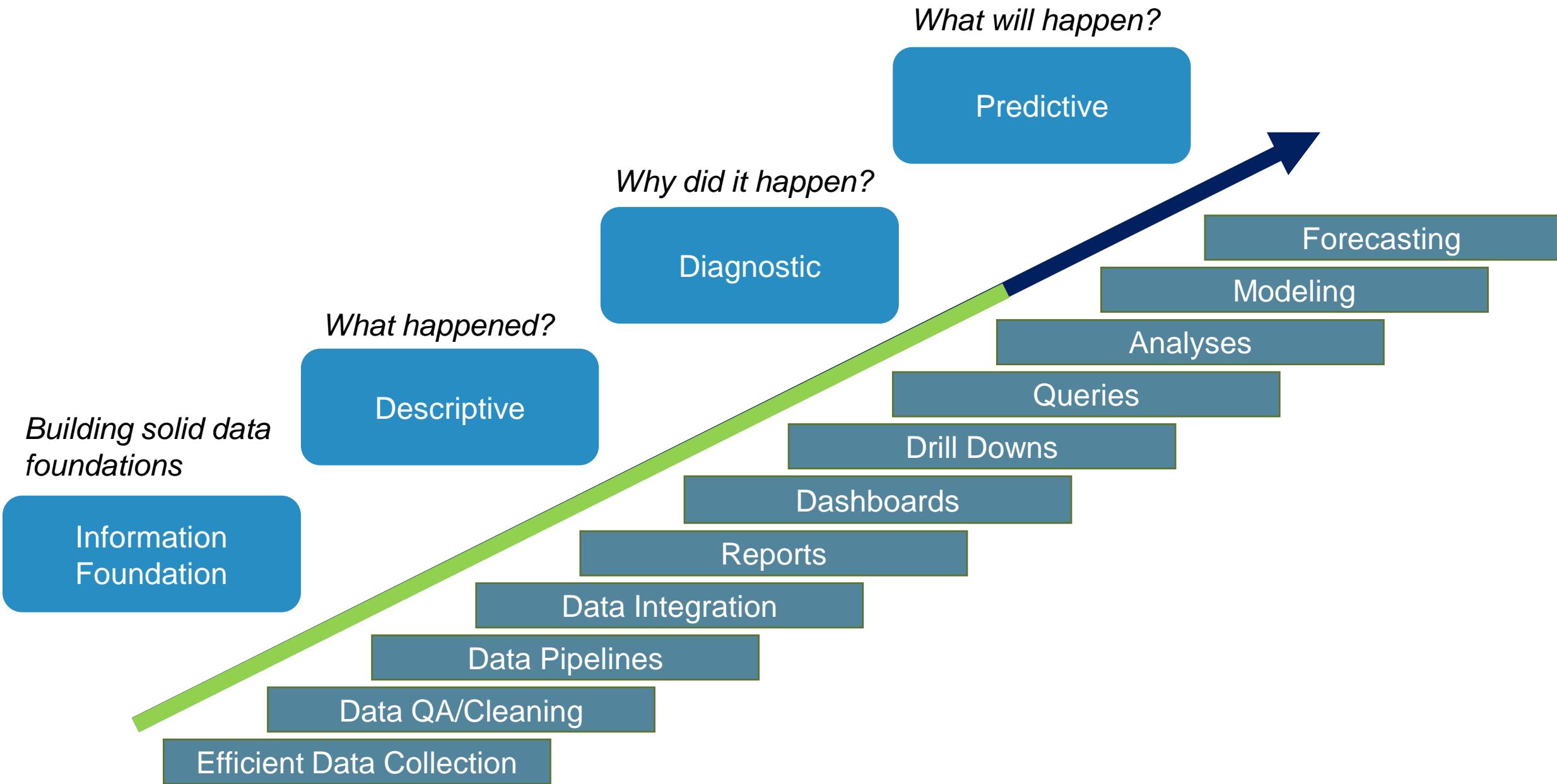
- Established 2019
- Led by Chief Data Officer
- Drive Department-wide cultural transformation
- Data as 1st class citizen, strategic business asset
- Data strategy and roadmap based on 5 data pillars



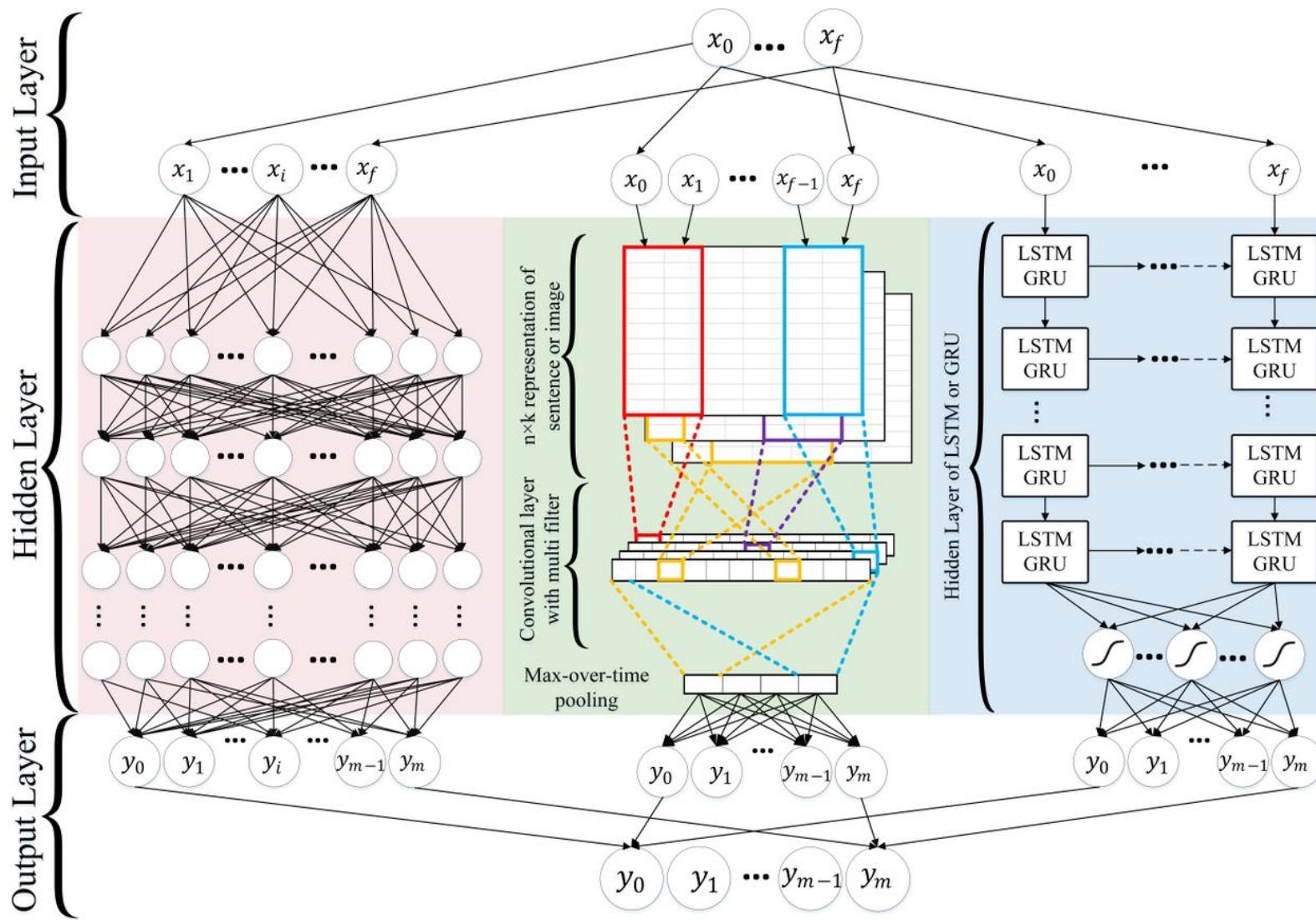
Data Strategy Framework

- **Data Infrastructure:** Technology used to store, exchange, and access data
- **Data Governance:** People, processes, and technology for data security, management, and access
- **Data Quality:** Tools and processes to ensure data are valid, accurate, and trustworthy
- **Data Literacy:** Workforce training across all levels of baseline knowledge
- **Data Use:** Reports, visualization, and analysis

Moving from descriptive to predictive



What academics are focused on



What most of the world needs

```
-- SQL Server Group By Example  
USE [SQL Tutorial]  
GO  
  
SELECT Education,  
       COUNT([EmpID]) AS [Total IDs]  
  FROM [Customer]  
 GROUP BY Education
```

	Education	Total IDs
1	Bachelors	4
2	Education	2
3	Graduate Degree	2
4	High School	1
5	Masters Degree	2
6	Partial College	1
7	Partial High School	3

Medicaid Utilization Dashboard

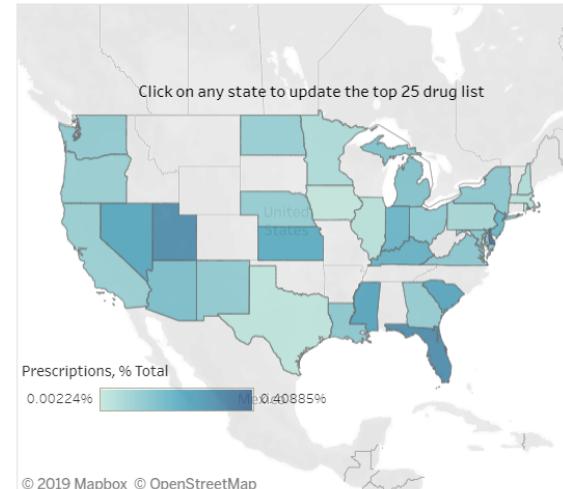
Choose Year-Quarter
2018-Q2

Choose NDC Description
OXYCODONE-ACETAMINOPHEN 10-325 MG TAB

Choose Utilization Type
Managed Care

OXYCODONE-ACETAMINOPHEN 10-325 MG TAB
Prescriptions
% of overall Prescriptions
(Utilization Type = Managed Care)

Click on any state to update the top 25 drug list



Prescriptions, % Total
0.00224% - 0.40885%

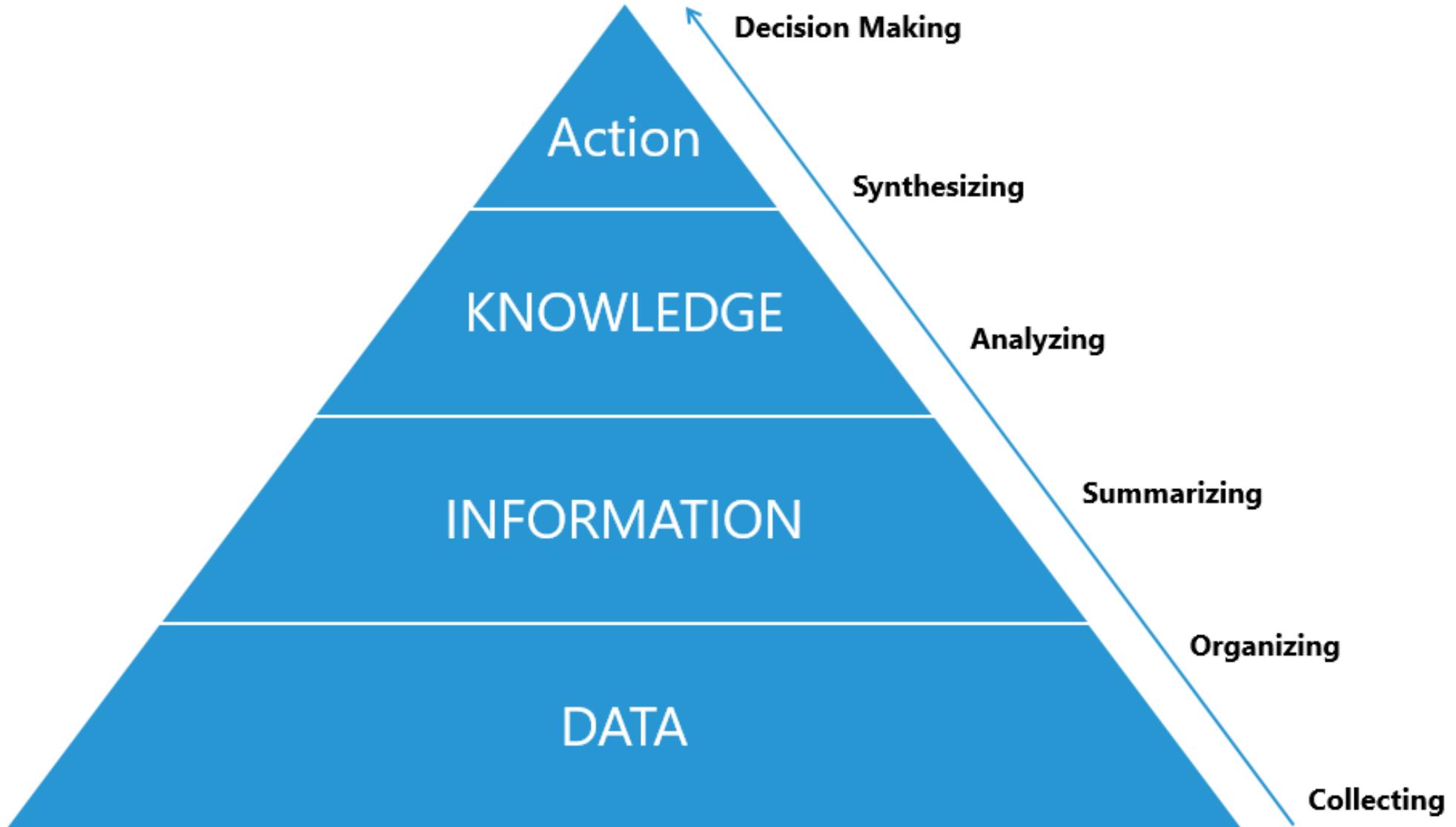
© 2019 Mapbox © OpenStreetMap

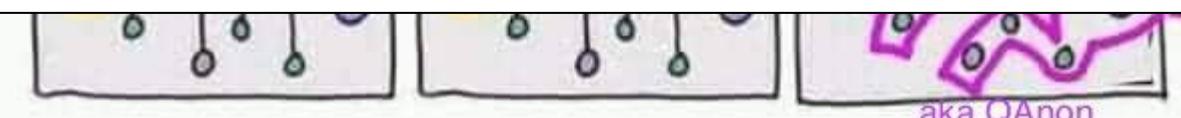
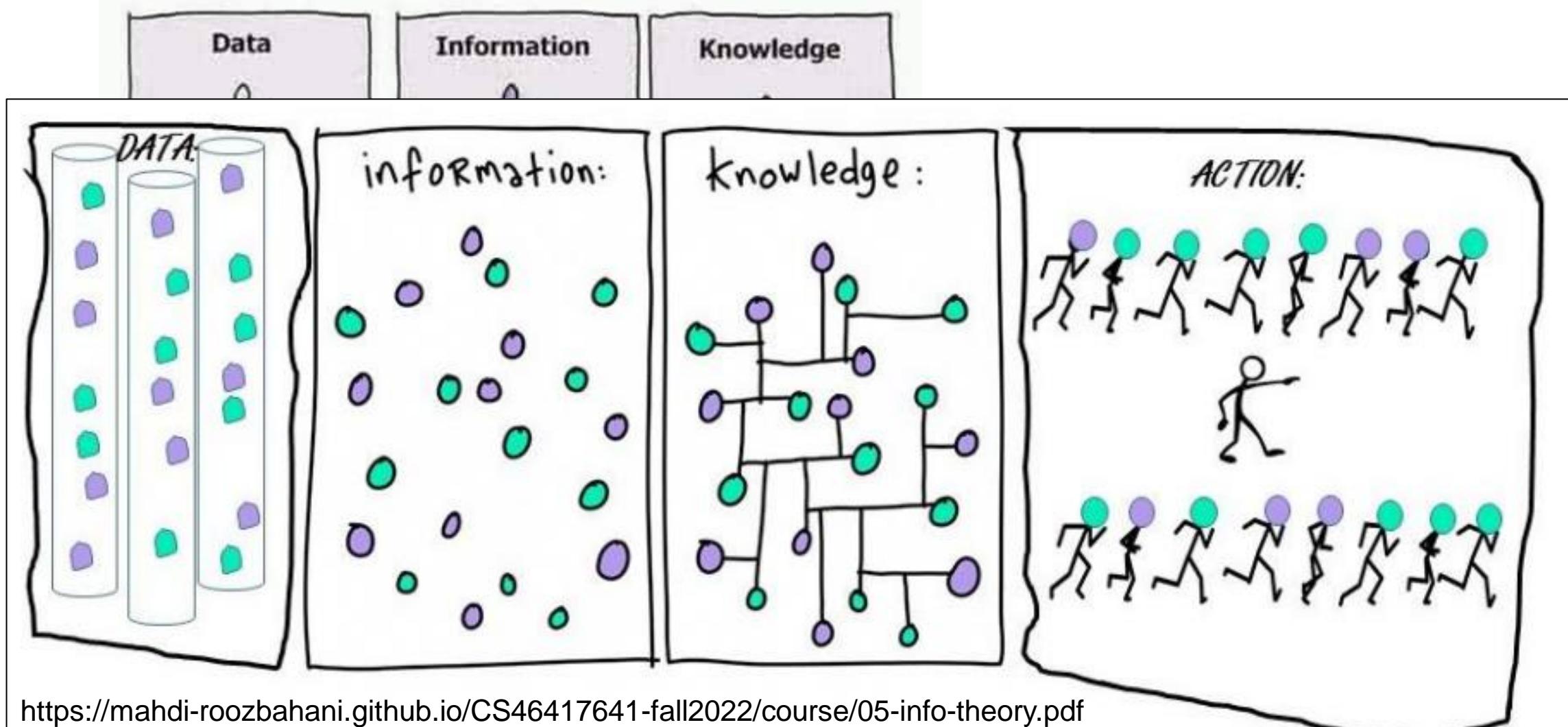
Choose Brand or Generic
Generic

Top 25 Generic NDC Descriptions (by \$)
State = All; Year-Quarter = 2018-Q2

NDC Description	Managed Care	Grand Tot..
METHYLPHENIDATE ER 36 MG TAB	\$31.4M	\$31.4M
BUPRENORPHINE-NALOXONE 8-2 MG SL ..	\$25.9M	\$25.9M
METHYLPHENIDATE ER 54 MG TAB	\$20.1M	\$20.1M
EPINEPHRINE 0.3 MG AUTO-INJECT	\$17.0M	\$17.0M
METHYLPHENIDATE ER 27 MG TAB	\$15.7M	\$15.7M
FLUTICASONE PROP 50 MCG SPRAY	\$15.7M	\$15.7M
METHYLPHENIDATE ER 18 MG TAB	\$12.8M	\$12.8M
BUDESONIDE 0.5 MG/2 ML SUSP	\$10.3M	\$10.3M
MEDROXYPROGESTERONE 150 MG/ML	\$10.1M	\$10.1M
ONDANSETRON HCL 4 MG/2 ML VIAL	\$9.9M	\$9.9M
ARIPIPRAZOLE 5 MG TABLET	\$9.1M	\$9.1M
DEXTRAMPHETYLAMPHETAMINE ER 20 MG CAP	\$8.6M	\$8.6M
FENTANYL 100 MCG/2 ML VIAL	\$7.6M	\$7.6M
EPINEPHRINE 0.15 MG AUTO-INJECT	\$7.4M	\$7.4M
DULOXETINE HCL DR 60 MG CAP	\$7.3M	\$7.3M
ONDANSETRON ODT 4 MG TABLET	\$7.3M	\$7.3M
DIVALPROEX SOD ER 500 MG TAB	\$7.1M	\$7.1M
GABAPENTIN 300 MG CAPSULE	\$6.9M	\$6.9M
HYDROCODONE-ACETAMINOPHEN 10-32..	\$6.6M	\$6.6M
ALBUTEROL SUL 2.5 MG/3 ML SOLN	\$6.6M	\$6.6M
POLYETHYLENE GLYCOL 3350 POWD	\$6.3M	\$6.3M
MONTELUKAST SOD 10 MG TABLET	\$6.1M	\$6.1M
LIDOCAINE 5% PATCH	\$6.0M	\$6.0M
PALiperidone ER 6 MG TABLET	\$4.9M	\$4.9M
CETIRIZINE HCL 10 MG TABLET	\$4.3M	\$4.3M

The central dogma of informatics: DIKA





Original source unknown

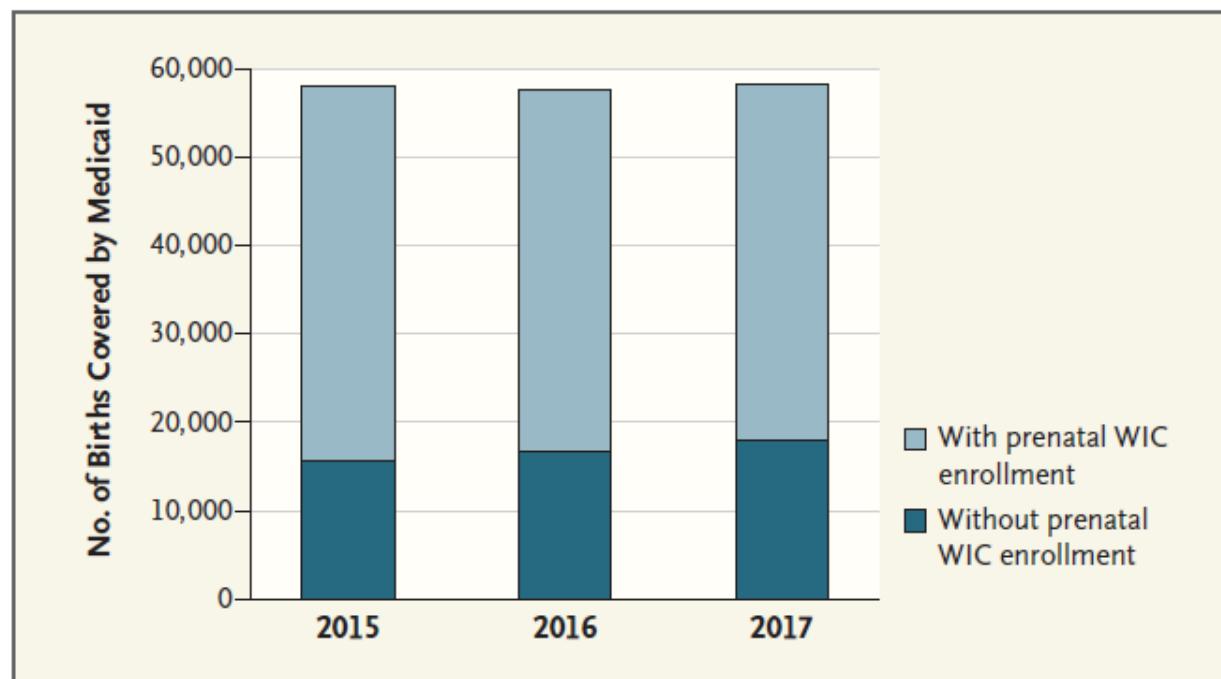


The NEW ENGLAND
JOURNAL of MEDICINE

Perspective FREE PREVIEW

Focusing on Population Health at Scale — Joining Policy and Technology to Improve Health

Aaron McKethan, Ph.D., Seth A. Berkowitz, M.D., M.P.H., and Mandy Cohen, M.D., M.P.H.



Medicaid-Covered Births with and without Concurrent Prenatal WIC Enrollment,
North Carolina.

In order to gain actionable knowledge, need 2 things:

1. The ability to integrate data between divisional silos (“Data Integration”)



2. The ability to identify and link the same individual from different datasets (“Entity Resolution”)

ID: 123456
Name: Waldo



ID: 123456
Name: Waldo

**We interrupt our regularly
scheduled Data Office business to
bring you...**

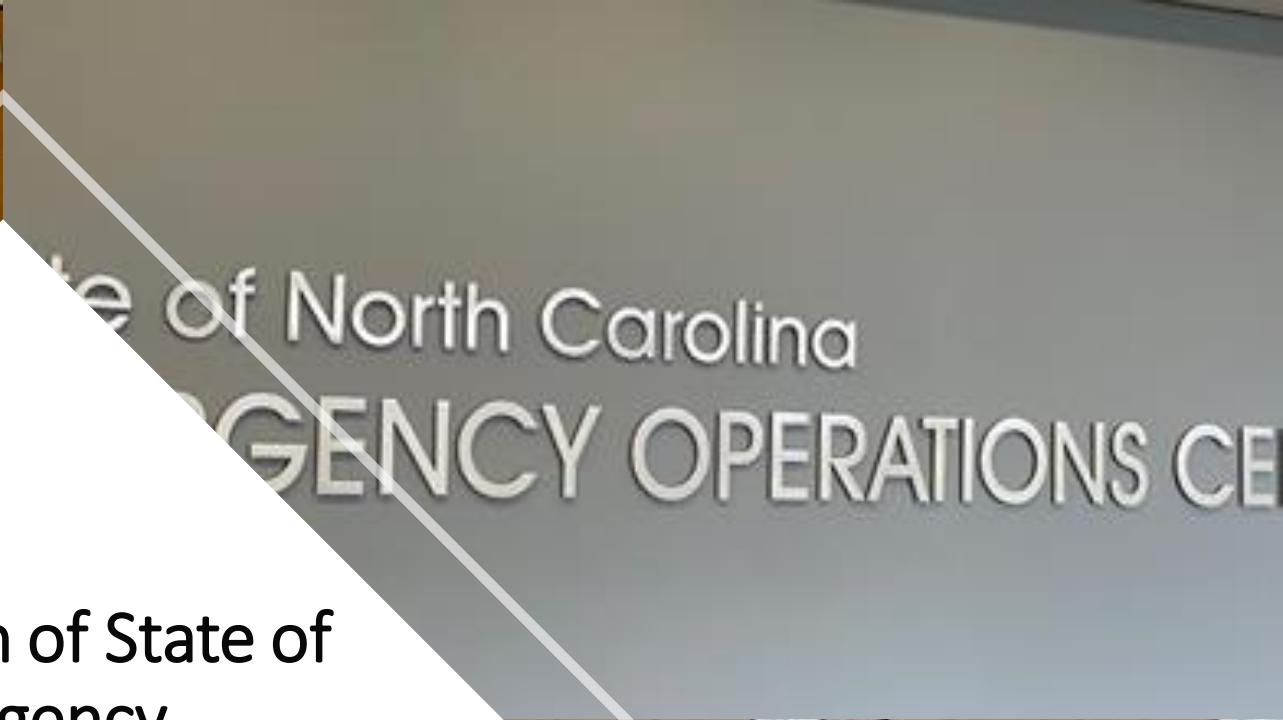


SITUATION ROOM
0702



SITUATION ROOM
0702

Declaration of State of Emergency *March 10, 2020*

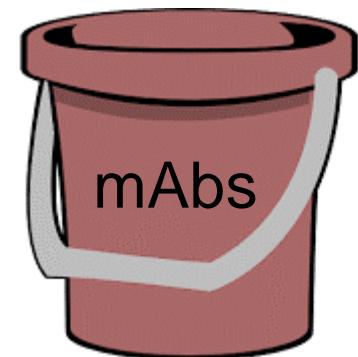
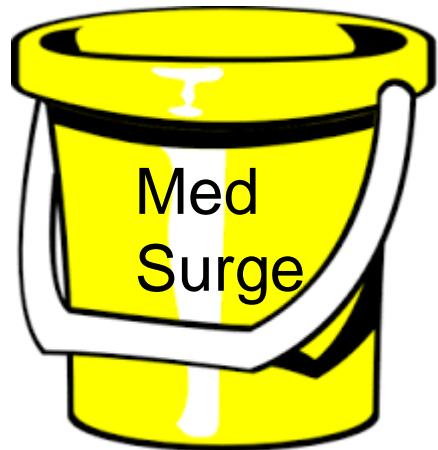


Early Actionable Questions to Enable Data Driven Policy

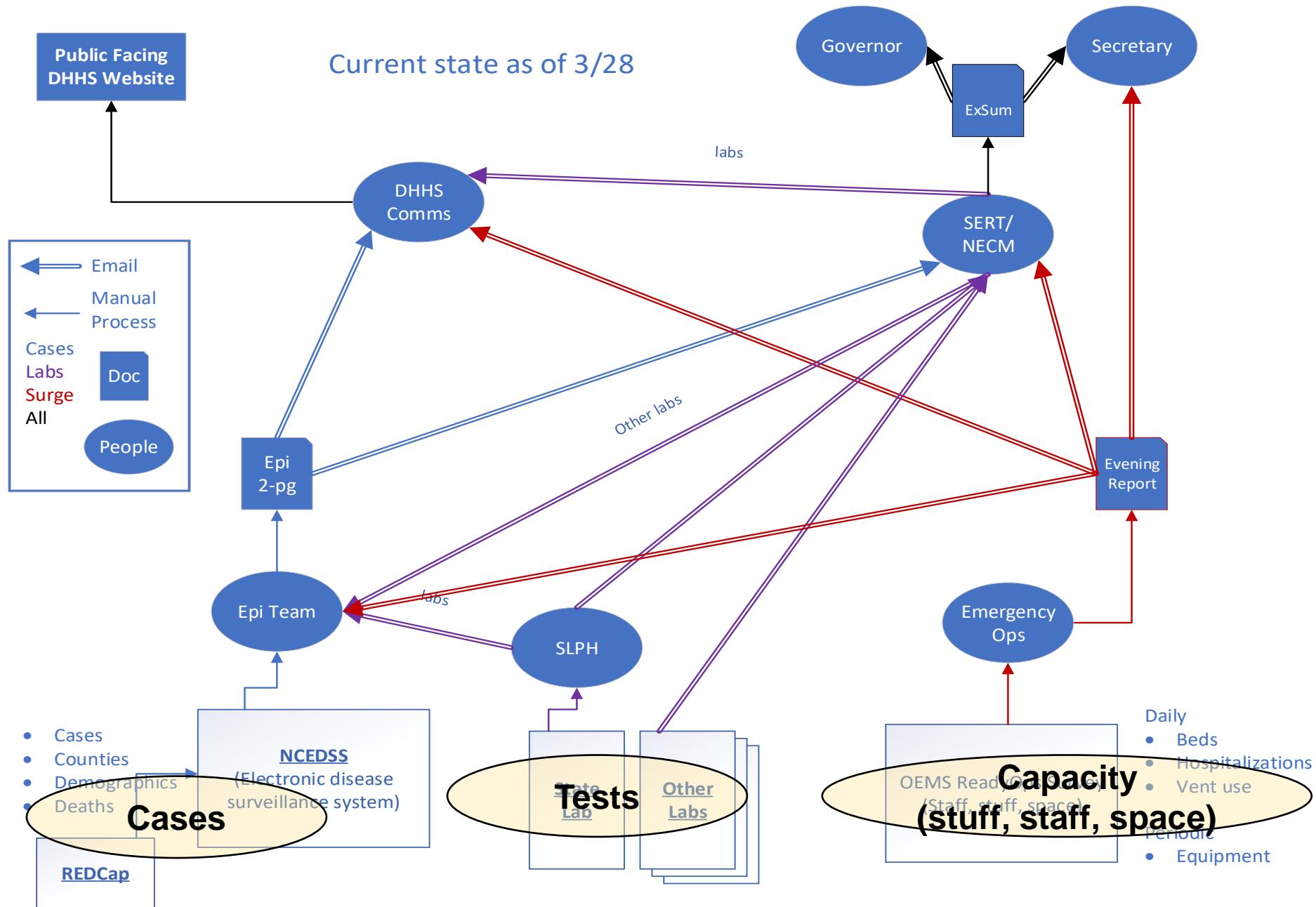
- How many cases will we see? When will our “epi curve” peak?
- How much (extra) PPE is needed, and where can we get it from?
- Will we run out of hospital beds? ICU beds? Ventilators?
- Should we shut down bars and restaurants? Schools? Businesses?
- How can we support families who are unable to work (either because workplace is shut down, or childcare is unavailable)?



Buckets of NC DHHS COVID-19 Data

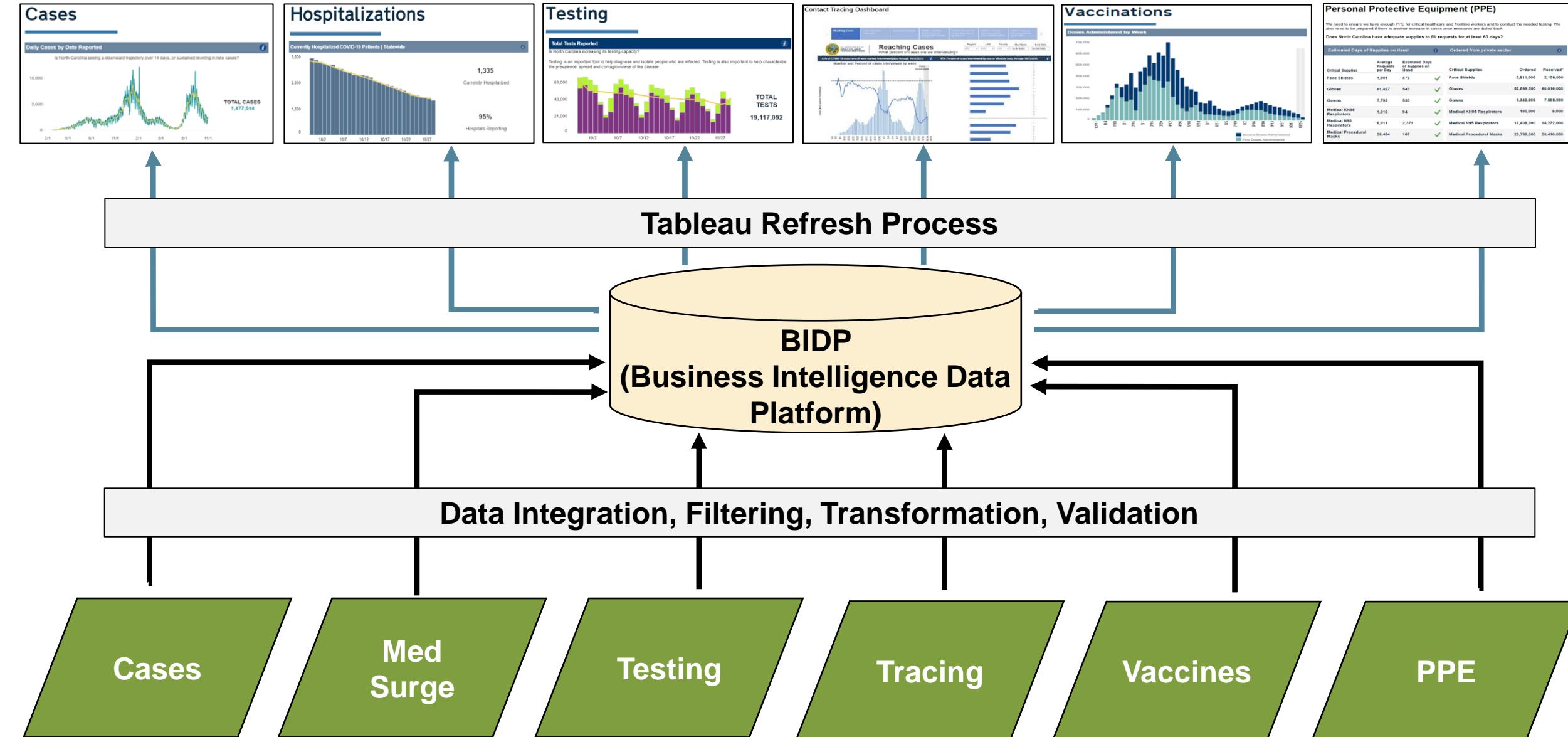


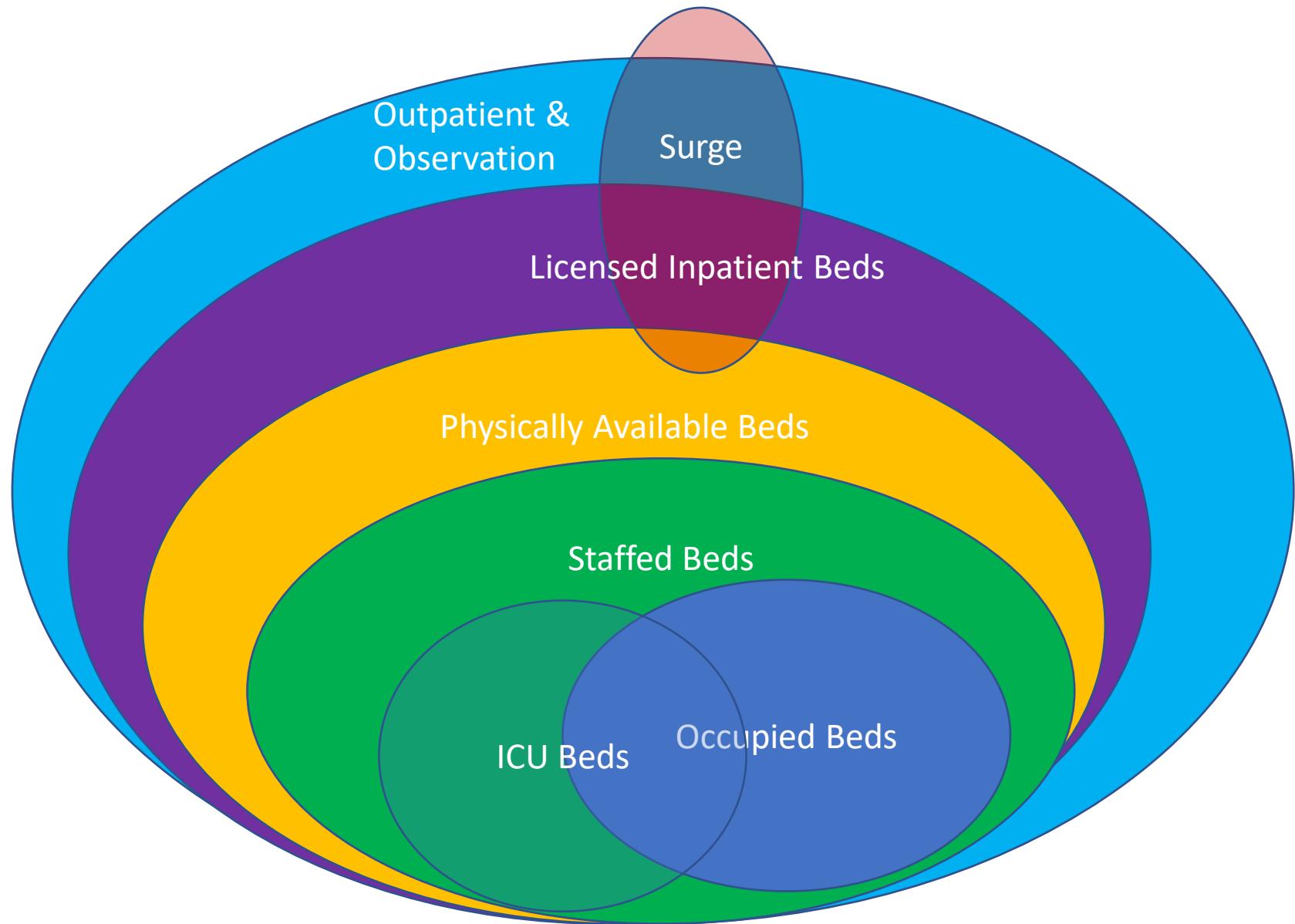
Baseline Data Flow (as of March 2020)



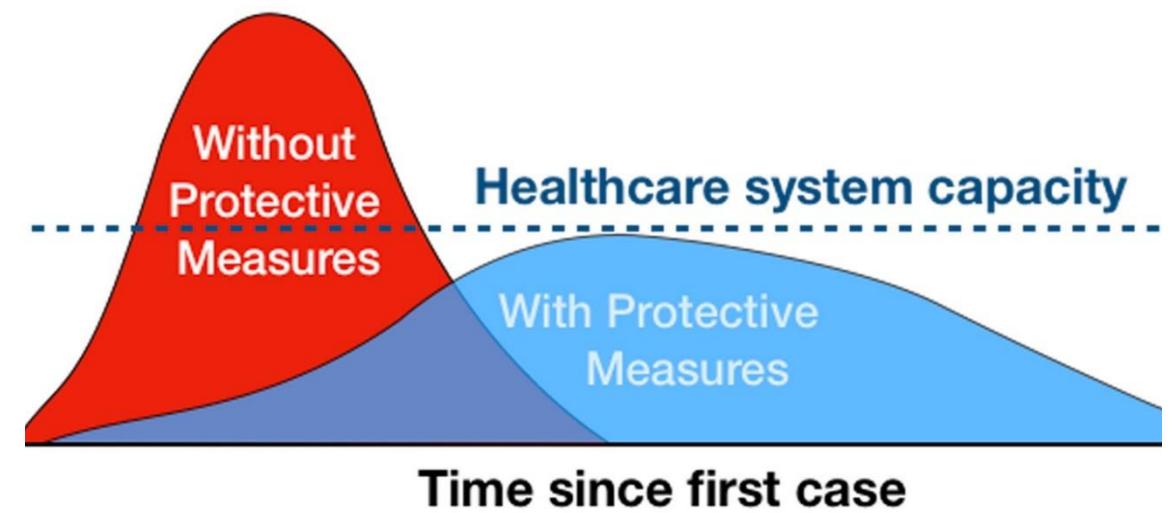
Today's Data Flow

Tableau Dashboards

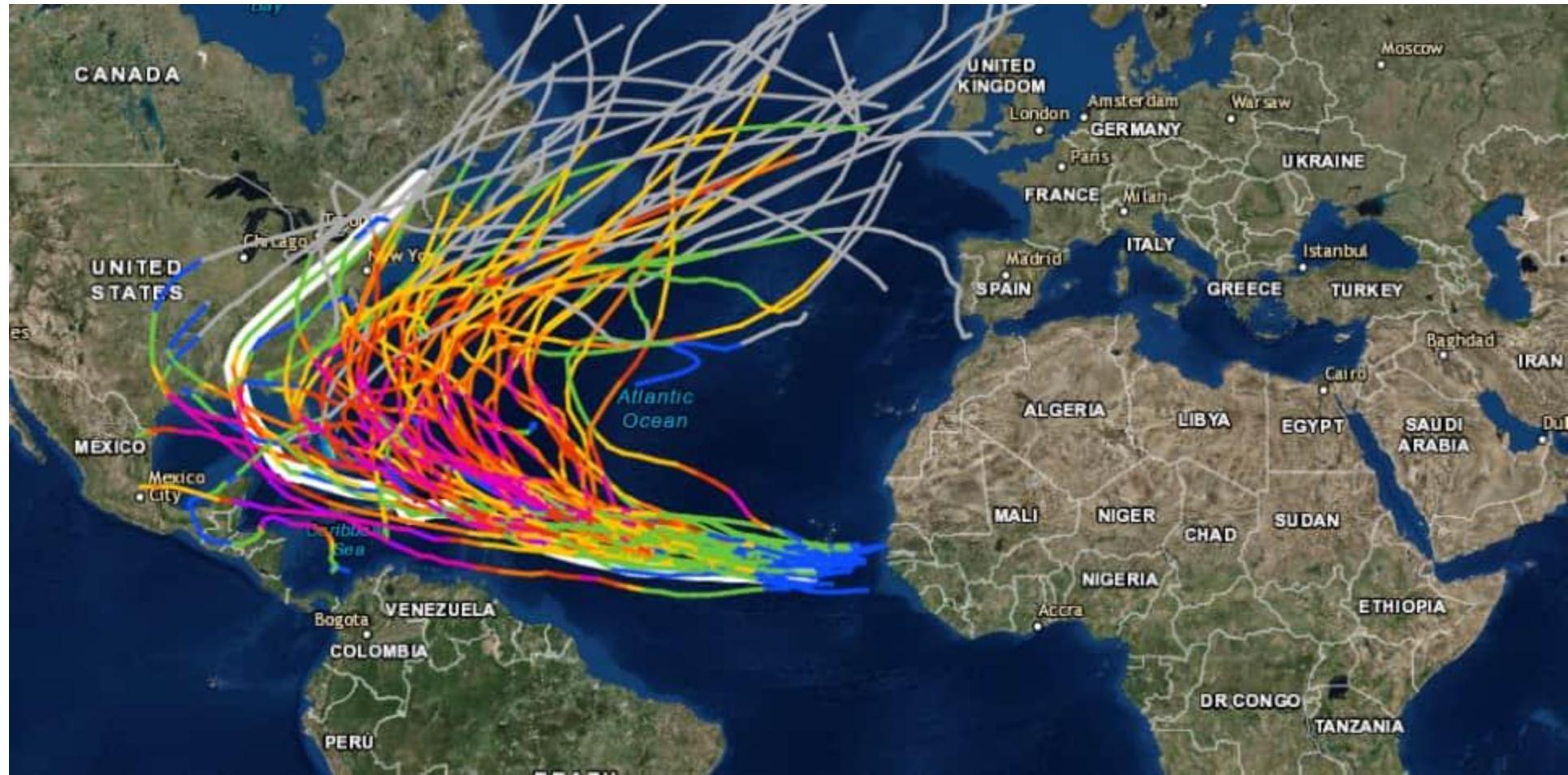




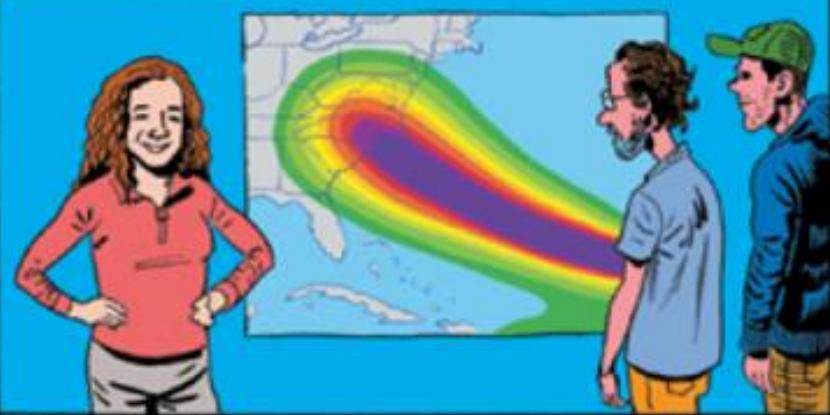
Data Modeling: Team science at work



Adapted from CDC / The Economist



IT'S JESSIE TENENBAUM, ASSISTANT PROFESSOR OF BIOSTATISTICS AND BIOINFORMATICS, WHO'S CURRENTLY CHIEF DATA OFFICER OF THE N.C. DEPARTMENT OF HEALTH & HUMAN SERVICES. HER METAPHOR FOR MODELS IS THOSE HURRICANE TRACKS WE ALL LOOK AT DURING LATE SUMMER AND EARLY FALL.



THEY DON'T TELL YOU EXACTLY WHEN OR WHERE THE STORM WILL COME ASHORE, OR WHAT WIND AND RAIN YOU CAN EXPECT, BUT THEY GIVE YOU A SENSE OF WHAT'S COMING SO YOU CAN PLAN.



YOU SEE THOSE MAPS, AND YOU SEE THE LIKELY PATH THE STORM IS GOING TO TAKE, BUT THEN YOU KIND OF SEE IT DIVERGE. IF THE OUTCOME IS DIFFERENT FROM THAT MEAN PRIMARY PATH THE MODEL PREDICTED, THAT DOESN'T MAKE THE MODEL WRONG. IT MEANS IT WAS MORE TOWARDS THE EDGES OF THE CURVE THE MODEL HAD PROJECTED.



Duke Magazine, Summer 2020

Story by Scott Huler, Illustrations by James Yamasaki

Geospatially Explicit Agent Based Model (ABM) Framework

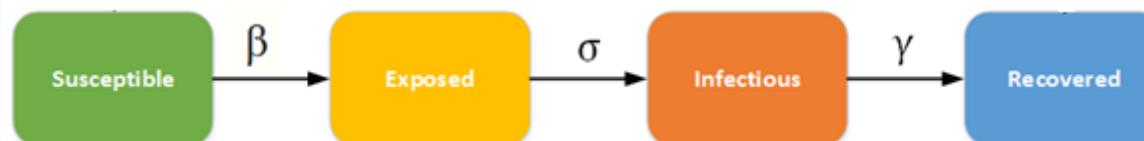
- ABM includes ~10.4 million agents
 - Based on the NC synthetic population: Dataset of anonymous synthetic persons and households, located geographically
 - Each agent represents a NC resident and has attributes (e.g., age)
- Each agent's physical location, health status, and life status are simulated daily



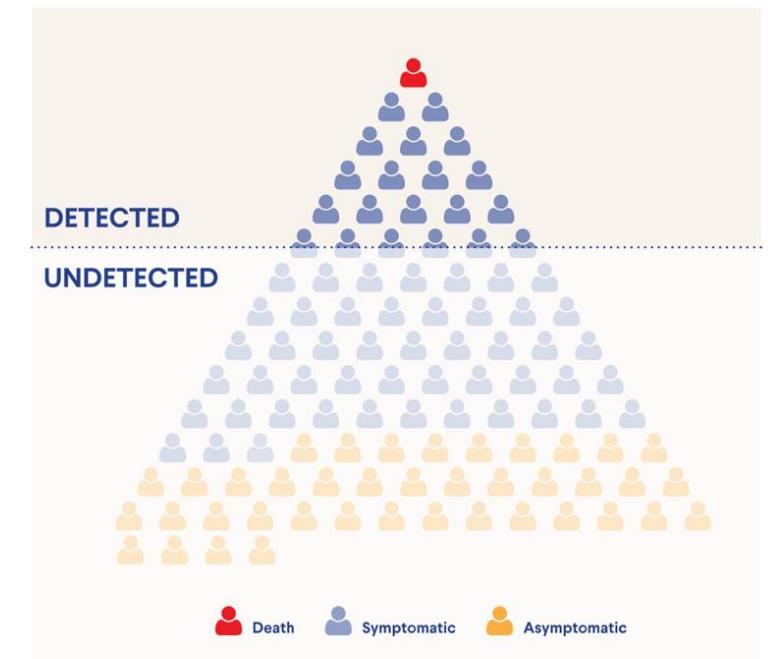
NC County Level SEIR* models

- Predict new SARS-CoV-2 infections for the next 30 days
- Leverage NC county-level reported case data while accounting for unreported infections

The COVID-19 Case Iceberg



*Susceptible-Exposed-Infectious-Recovered Model Diagram
<https://institutefordiseasemodeling.github.io/Documentation/malaria/model-seir.html>



June 10, 2020 Adapted from RTSL and CDC estimates for COVID-19 in the United States.

Source: Prevent Epidemics

<https://preventepidemics.org/covid19/science/insights/covid-19-case-iceberg/>

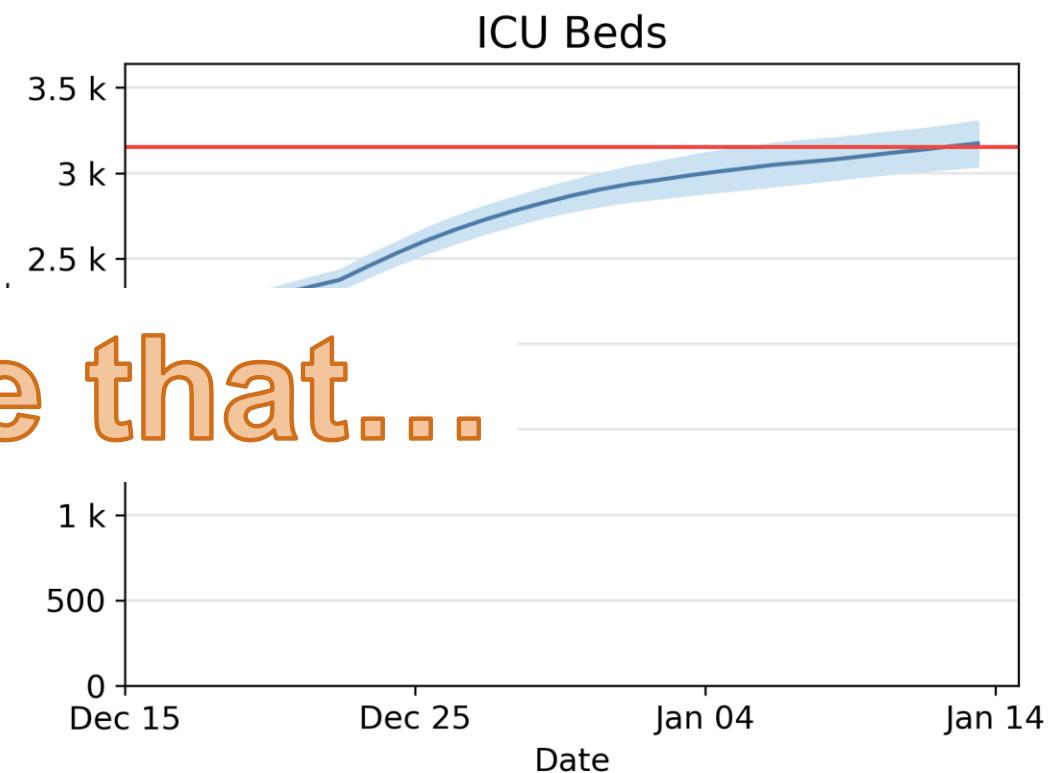
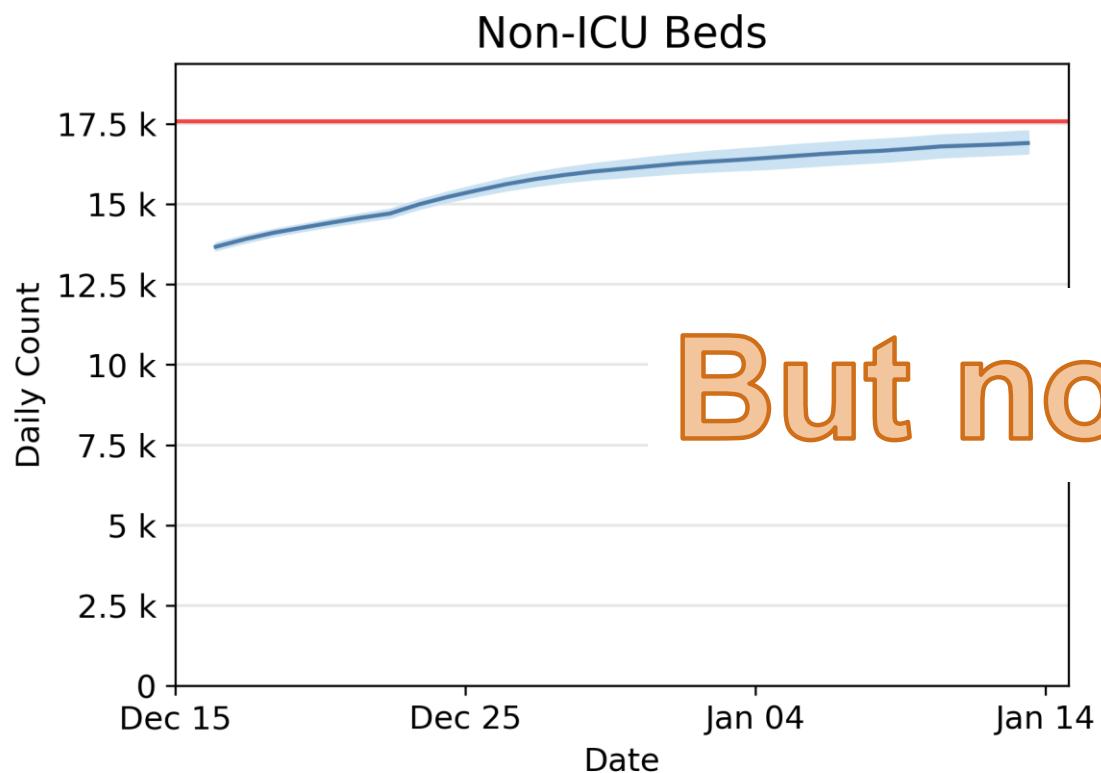
Model-Projected Patient Demand for Non-ICU & ICU Beds by Day - STATEWIDE

Statewide

Model-projected mean number of patients needing a hospital bed on a given day* statewide and by North Carolina flu monitoring region. Modeled scenarios consider an effective reproductive number (R_e) range** and a reported cases multiplier to account for unreported/undocumented infections at model initialization and assume 25% of hospitalized patients requiring ICU-level care. Regional demand values are based on the modeled hospitals in each region. See model documentation for additional detail and references.

$R_e = 1.2 - 1.4$

Updated Dec 17, 2020



But note that...

Light blue area: 25th and 75th percentiles of model output

Red line: Number of non-ICU operational beds or licensed ICU beds, as indicated (per NC DHHS, July 2020); scaled to account for rounding in forecasts

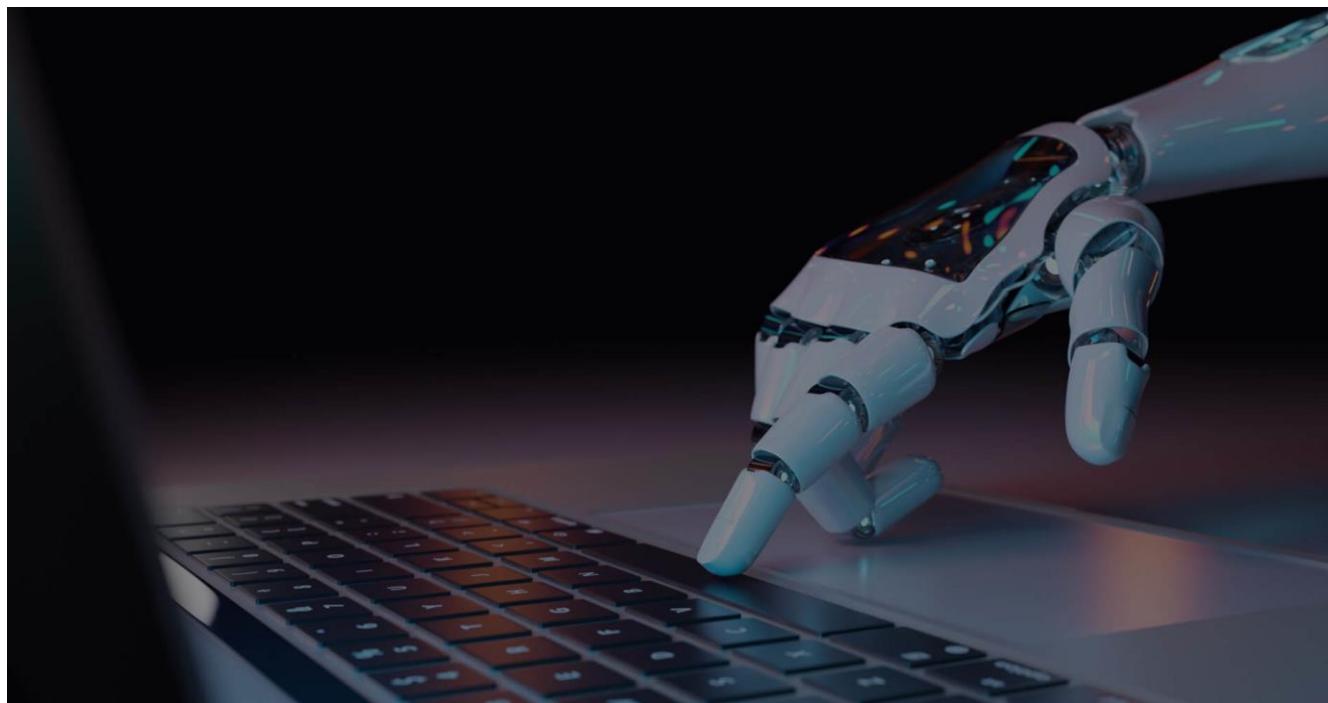
Parameter values and assumptions are subject to change as knowledge of COVID-19 evolves

*Based on 100 model runs and using a sample of 2.5 million agents for each run, with results scaled to the 10.4 million agents representing the North Carolina population

** R_e ranges are sampled from a uniform distribution

Notable Exception where AI has proven useful in Public Health

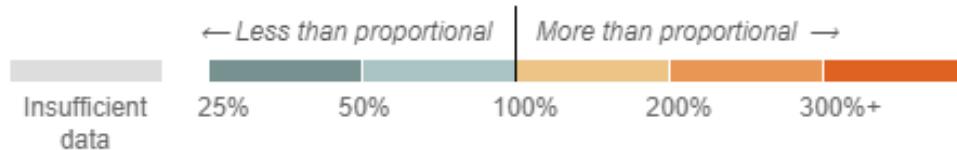
- RPA: Robotic Process Automation
- Software to perform mindless human tasks, often to make up for poorly designed and/or non-interoperable systems



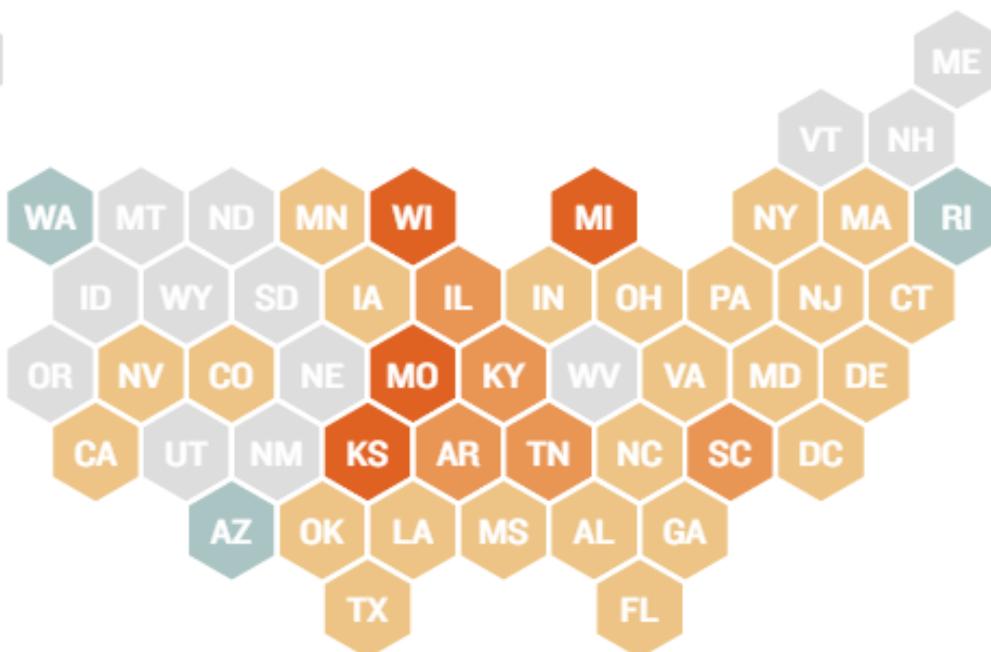
COVID disproportionately impacted historically marginalized populations

Deaths and Cases Disproportionately Affect African Americans In Most States

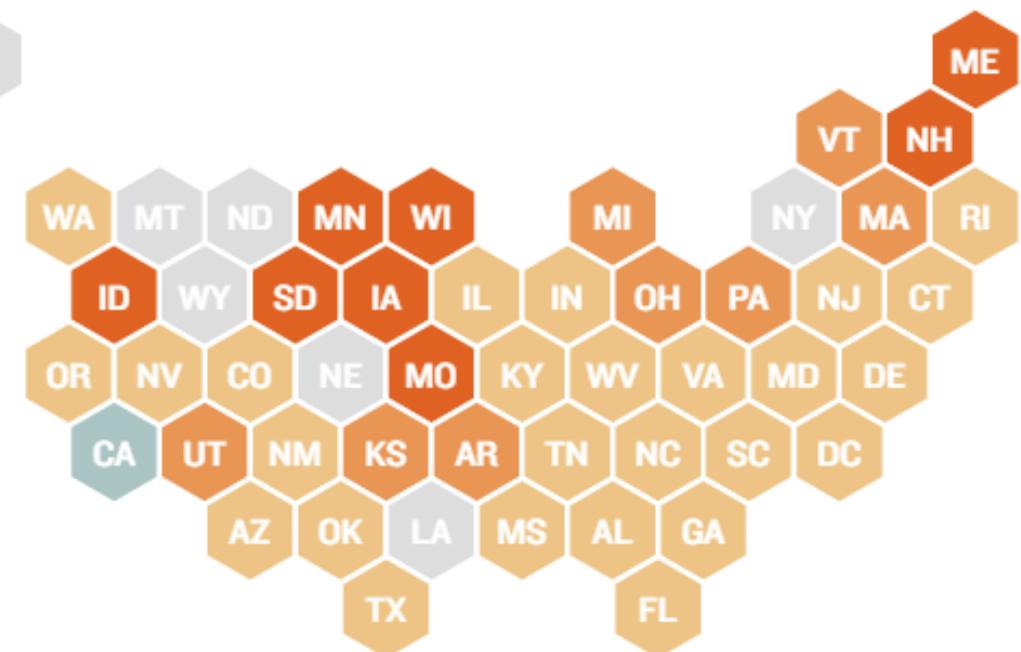
PERCENTAGE DIFFERENCE IN SHARE OF DEATHS/CASES, COMPARED WITH
AFRICAN AMERICANS' SHARE OF THE POPULATION



DEATHS



CASES



- What is the distribution of race/ethnicity among
 - COVID-19 cases?
 - COVID-19 deaths?
 - People tested for COVID-19?
 - People vaccinated against COVID-19?

Bloomberg

Prognosis One U.S. State's Laser Focus on Data Helps Shrink Racial Vaccine Gap

North Carolina has racial data for 99.6% of people who got a shot in the state

By [Angelica LaVito](#)
March 6, 2021, 8:30 AM EST



LIVE ON BLOOMBERG
[Watch Live TV >](#)
[Listen to Live Radio >](#)

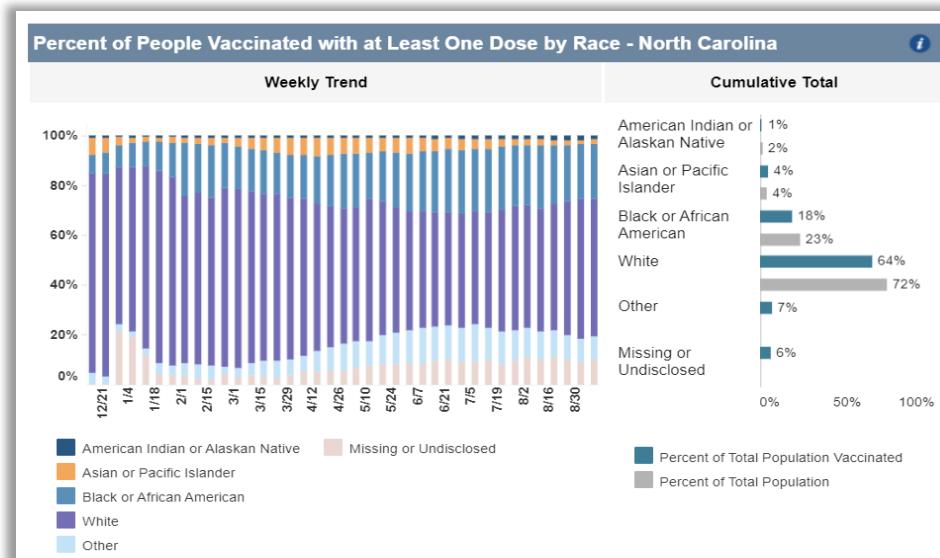
IG
Open an account and start trading in minutes

Forex trading involves risk. Losses can exceed deposits.

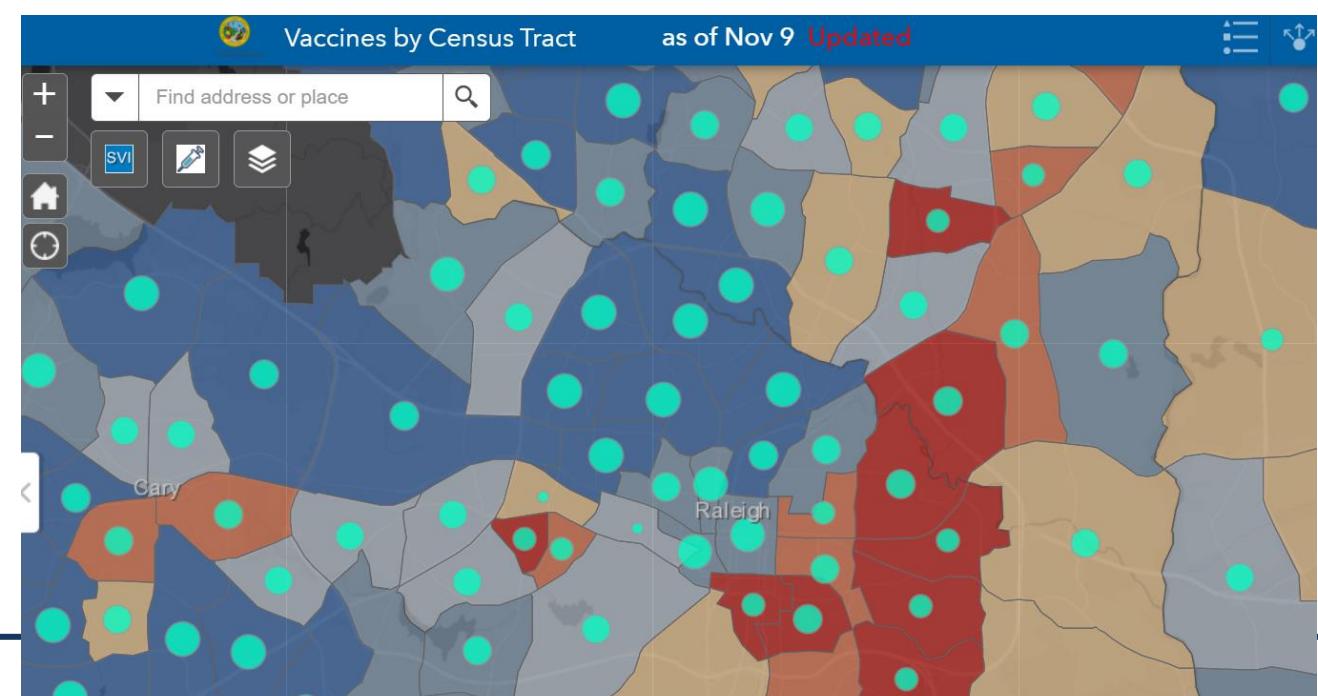
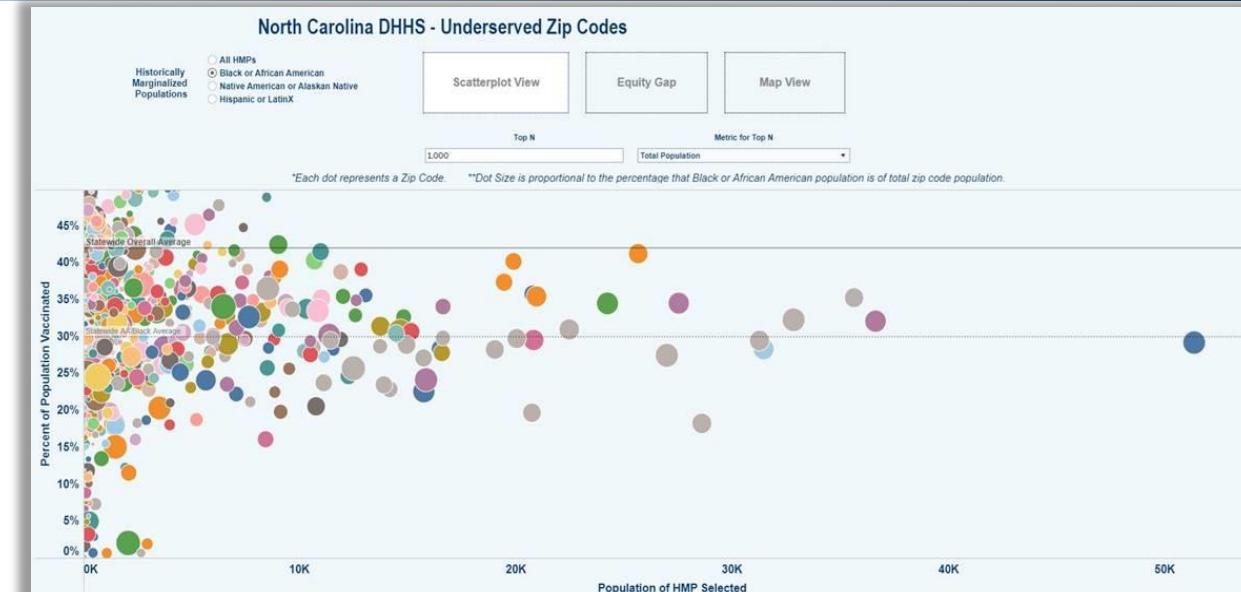
A health-care worker puts a bandage on man after he received the Moderna Covid-19 vaccine at the Durham County Department of Public Health in Durham, N.C., on Jan. 8. Photographer: Julia Wall/The News & Observer via AP

NC Leads the Country in Collecting Equity Data

- Data tools built to infuse equity data insights across entire vaccine operation

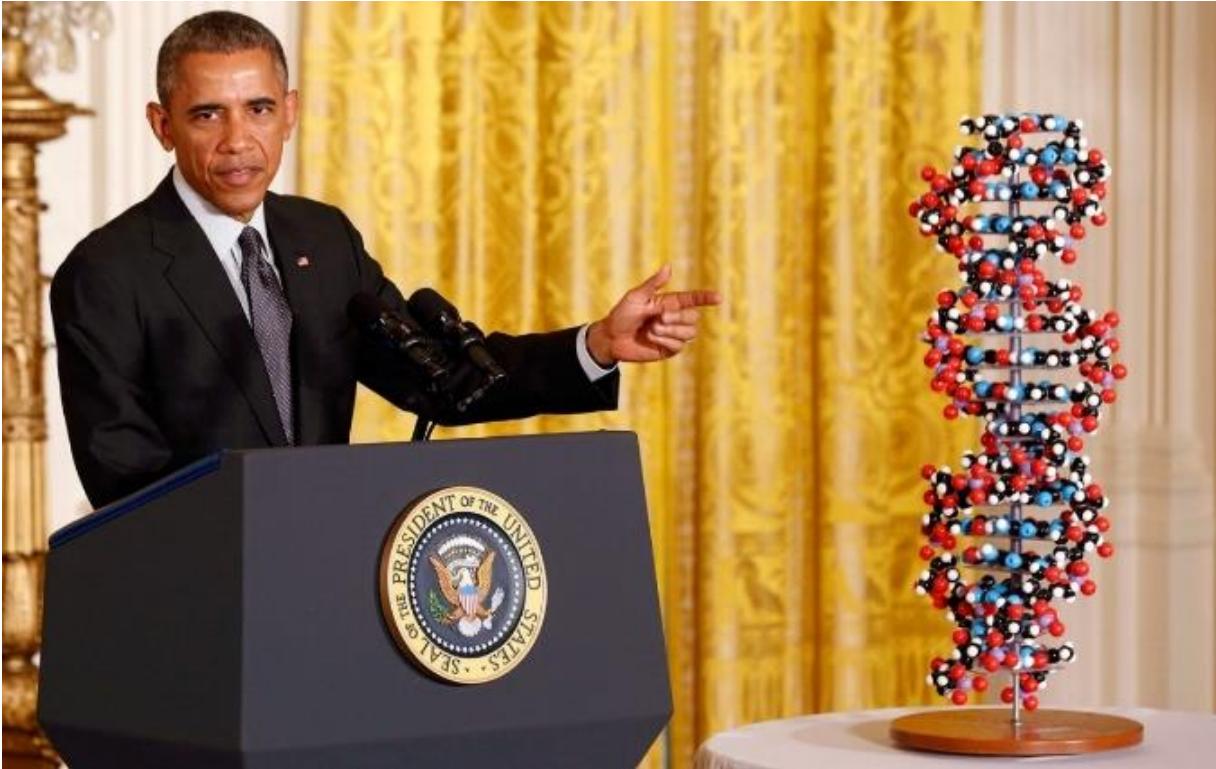


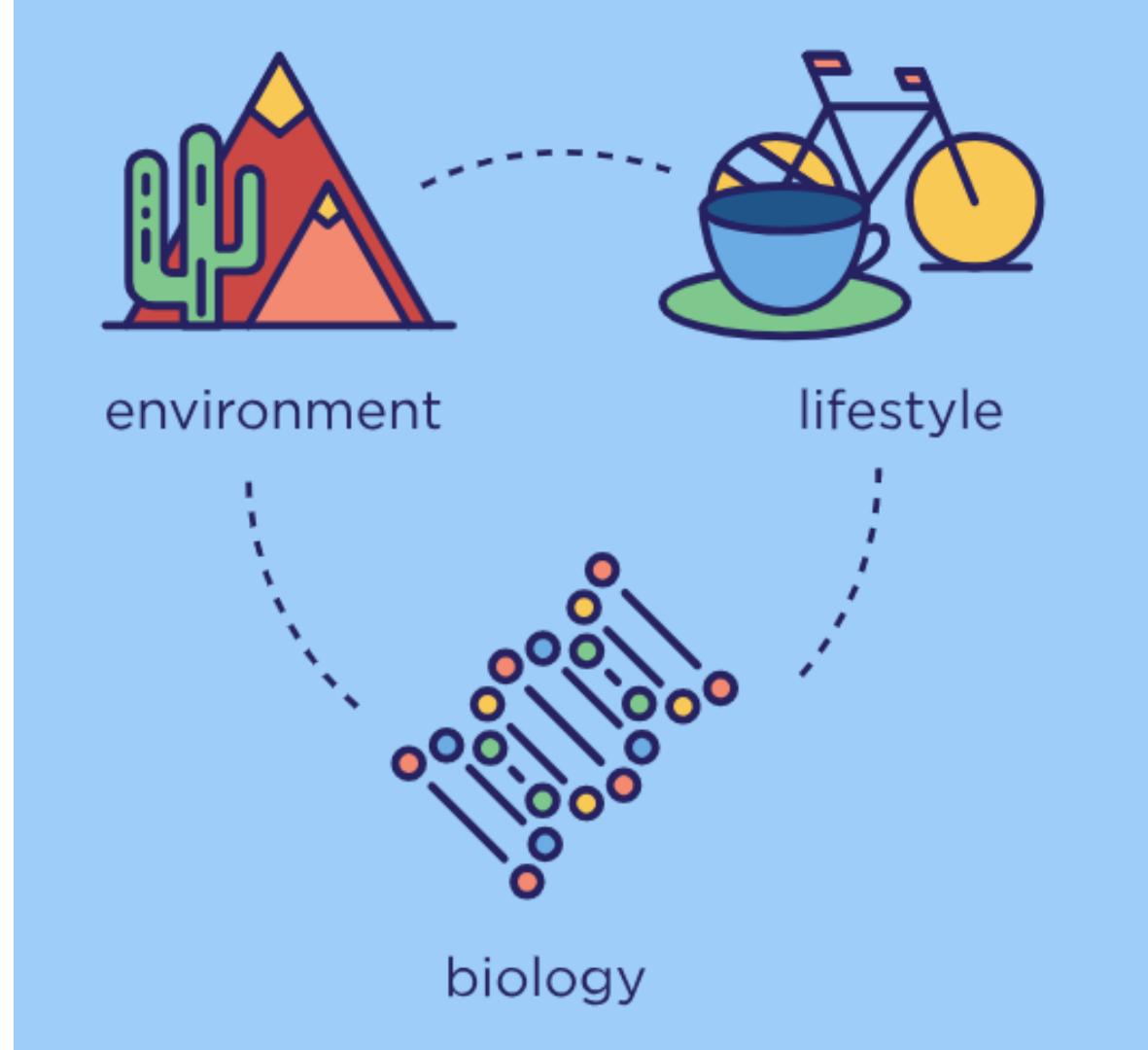
<https://covid19.ncdhhs.gov/dashboard>



**We now return to our regularly
scheduled Data Office business...**

Remember this pic?





<https://allofus.nih.gov/>

Beyond DNA



Childhood
experiences



Housing



Education



Social support



Family income



Employment



Our communities



Access to health
services

<https://www.publichealthnotes.com/social-determinants-health-sdh/>

“...North Carolina will implement a groundbreaking program ... to pilot evidence-based interventions addressing issues like housing instability, transportation insecurity, food security, and interpersonal violence and toxic stress.”

HEALTH

RELATED TOPIC

MEDICAID | BE

| CASE MANAG

CMS A

Medic

Outco

Seema Verma

OCTOBER 24, 2018

10.1377/hblog20181024.406020

Homelessness is a medical condition

- Hospital pays \$1,000 per individual per month for housing.
- Significant drop in ER use among individuals given housing options.



To Reduce Chronic Homelessness, A Chicago Hospital Is Treating It As A Medical Condition

February 21, 2019

By [Jeremy Hobson](#)



Texas lawmakers look to Uber, Lyft to transport Medicaid patients

By Allie Morris, Austin Bureau | March 12, 2019 | Updated: March 12, 2019 3:17 p.m.



Screening Questions

Food

1. Within the past 12 months, did you worry that your food would run out before you got money to buy more? (Y/N)
2. Within the past 12 months, did the food you bought just not last and you didn't have money to get more? (Y/N)

Housing/Utilities

3. Within the past 12 months, have you ever stayed: outside, in a car, in a tent, in an overnight shelter, or temporarily in someone else's home (i.e. couch-surfing)? (Y/N)
4. Are you worried about losing your housing? (Y/N)
5. Within the past 12 months, have you been unable to get utilities (heat, electricity) when it was really needed? (Y/N)

Transportation

6. Within the past 12 months, has a lack of transportation kept you from medical appointments or from doing things needed for daily living? (Y/N)

Interpersonal Safety

7. Do you feel physically and emotionally unsafe where you currently live? (Y/N)
8. Within the past 12 months, have you been hit, slapped, kicked or otherwise physically hurt by anyone? (Y/N)
9. Within the past 12 months, have you been humiliated or emotionally abused by anyone? (Y/N)

NCCARE360 Closed Loop Referral Platform

Dashboard Clients Reports My Networks

Sara Jones

Jane Smith CONSENT ACCEPTED

DOB 1/2/1984 (Age 35) | TEL 123-456-7890 | EMAIL jane.smith@email.com
ADDRESS 99 Main Street, New York, NY 10130 | HOUSEHOLD 4 | HOUSEHOLD INCOME \$32,000
PREFERRED METHOD OF CONTACT Message | BEST TIME TO CONTACT Afternoon

VIEW/EDIT PROFILE TAKE ACTION

Mount Kisco Interfaith Food Pantry Food Pantry
300 East Main Street, Mount Kisco, NY 10549
Today: 11:30 - 10:00 pm | OPEN NOW
0.3 mi
914-610-5187
info@mountkiscofoodpantry.com ADD

Mount Kisco Interfaith Food Pantry Food Pantry
300 East Main Street, Mount Kisco, NY 10549
Today: 11:30 - 10:00 pm | OPEN NOW
0.3 mi
914-610-5187
info@mountkiscofoodpantry.com ADD

Mount Kisco Interfaith Food Pantry Food Pantry
300 East Main Street, Mount Kisco, NY 10549
Today: 11:30 - 10:00 pm | OPEN NOW
0.3 mi
914-610-5187
info@mountkiscofoodpantry.com ADD

Mount Kisco Interfaith Food Pantry Food Pantry
300 East Main Street, Mount Kisco, NY 10549
Today: 11:30 - 10:00 pm | OPEN NOW
0.3 mi
914-610-5187
info@mountkiscofoodpantry.com ADD

Search... SERVICE TYPE: Food

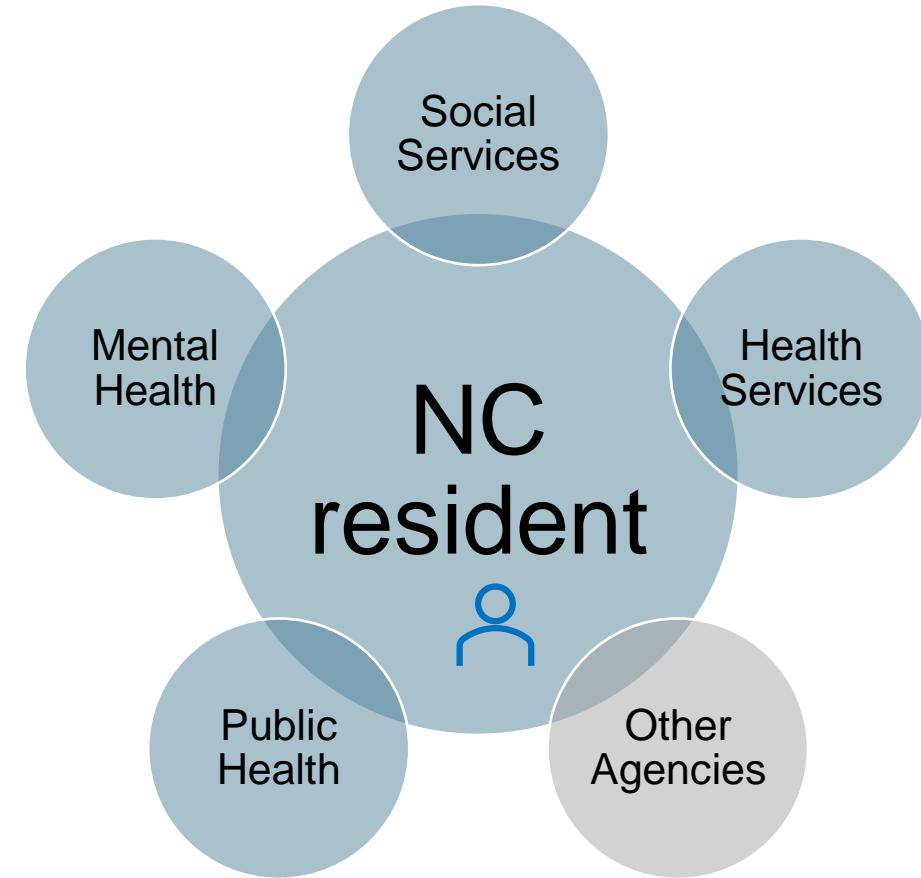
17 Results for Food within 25 miles of Northern Westchester Hospital, 400 E Main St, Mt Kisco, NY 10549 Filters

SEARCH THIS AREA

MacBook Pro

Whole Person Health

- Goal: link data to facilitate a “Whole Person Health” view of the people we serve.
 - Real-time individual level
 - Aggregate analysis to inform policy
- Requires the ability to integrate data across divisional silos
- Which requires ability to link records between systems



Motivating Questions (Examples)

- **What % of Medicaid beneficiaries have been vaccinated?**
- **Which children in foster care have prescriptions for >4 psychotropic medications?**
- **Who receives regular food assistance referrals and therefore may benefit from SNAP but is not enrolled?**
- **What % of people experiencing homelessness have been vaccinated?**
- **What is the relationship between early grade outcomes (e.g., third grade reading) and different early childhood conditions (e.g., early learning, health, housing, child welfare)?**

In order to achieve Whole Person Health, need 2 things:

1. The ability to integrate data between divisional silos (“Data Integration”)



2. The ability to identify and link the same individual from different datasets (“Entity Resolution”)

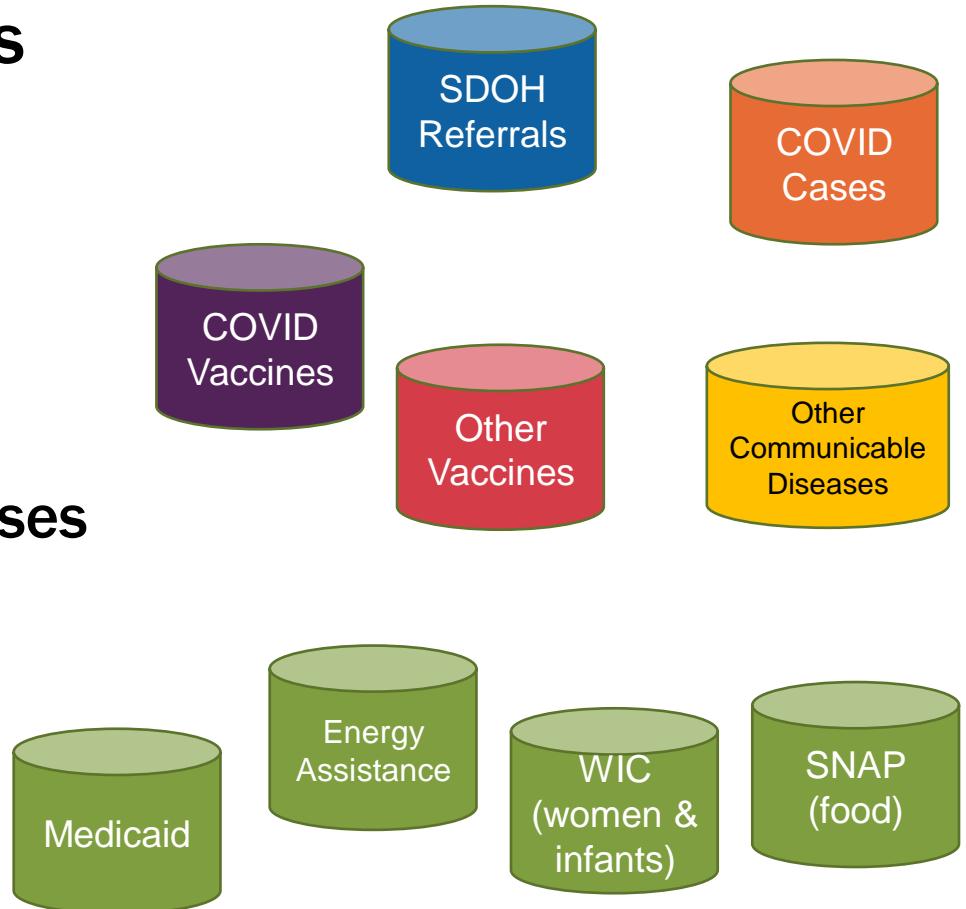
ID: 123456
Name: Waldo



ID: 123456
Name: Waldo

Answering those questions is currently difficult at best

- Data live in silos that do not share IDs
- Probabilistic match is possible, but
 - Labor-intensive
 - Prone to error
 - Example: initial approach for post-vax cases



Options for “solving” entity resolution

1. Each system uses its own ID, probabilistic “fuzzy match” between systems as needed
2. One universal ID, e.g. state-wide or National Health Identifier
3. Somewhere in between
 - a. Multiple “standard” identifiers
 - b. Map them to each other!
 - c. Refer to that mapping for efficient data integration



Map each ID to a universal unique identifier...

 Angie Taylor
12 Main St.
ID-123

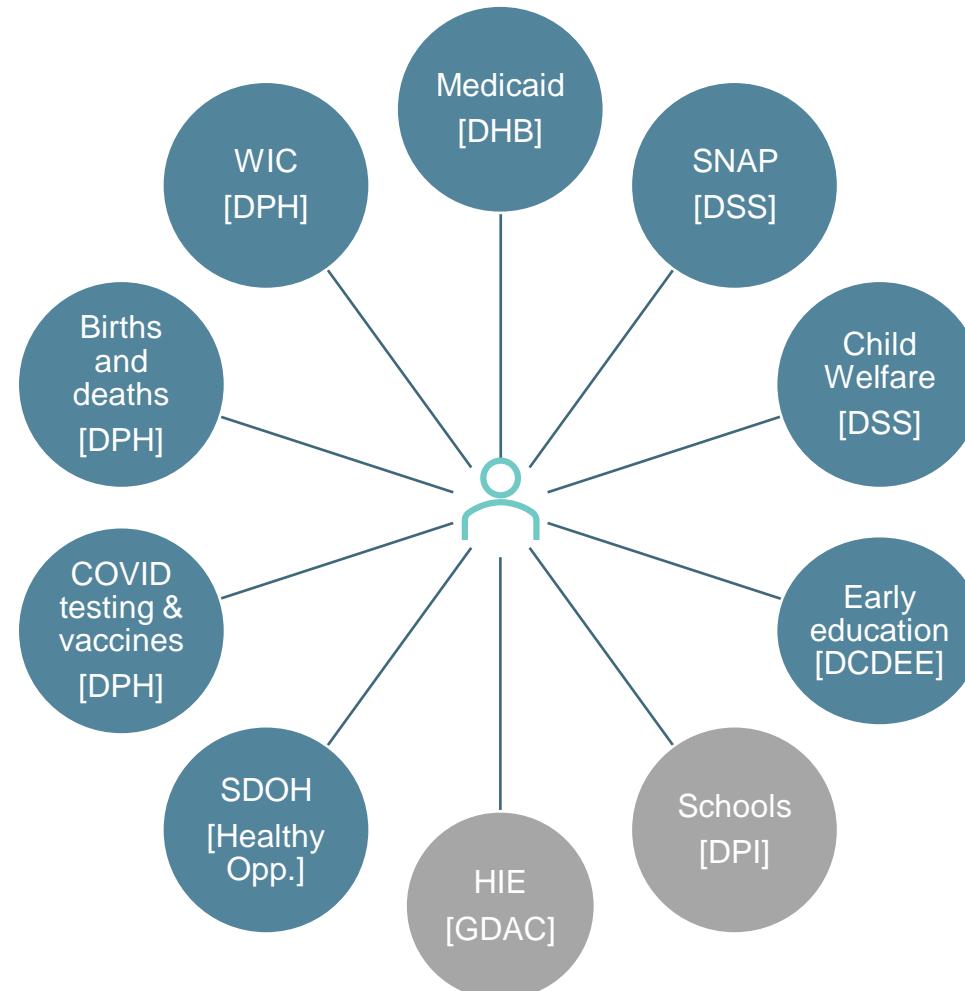
 Angela Taylor
21 Main St.
ID-XYZ

 A. Dale Taylor
21 Main St.
ID-QRS

“IDHub”

Source	Source ID	Unique ID
Medicaid	123	100
SNAP	XYZ	100
Immunizations	QRS	100

...enabling Whole Person Health



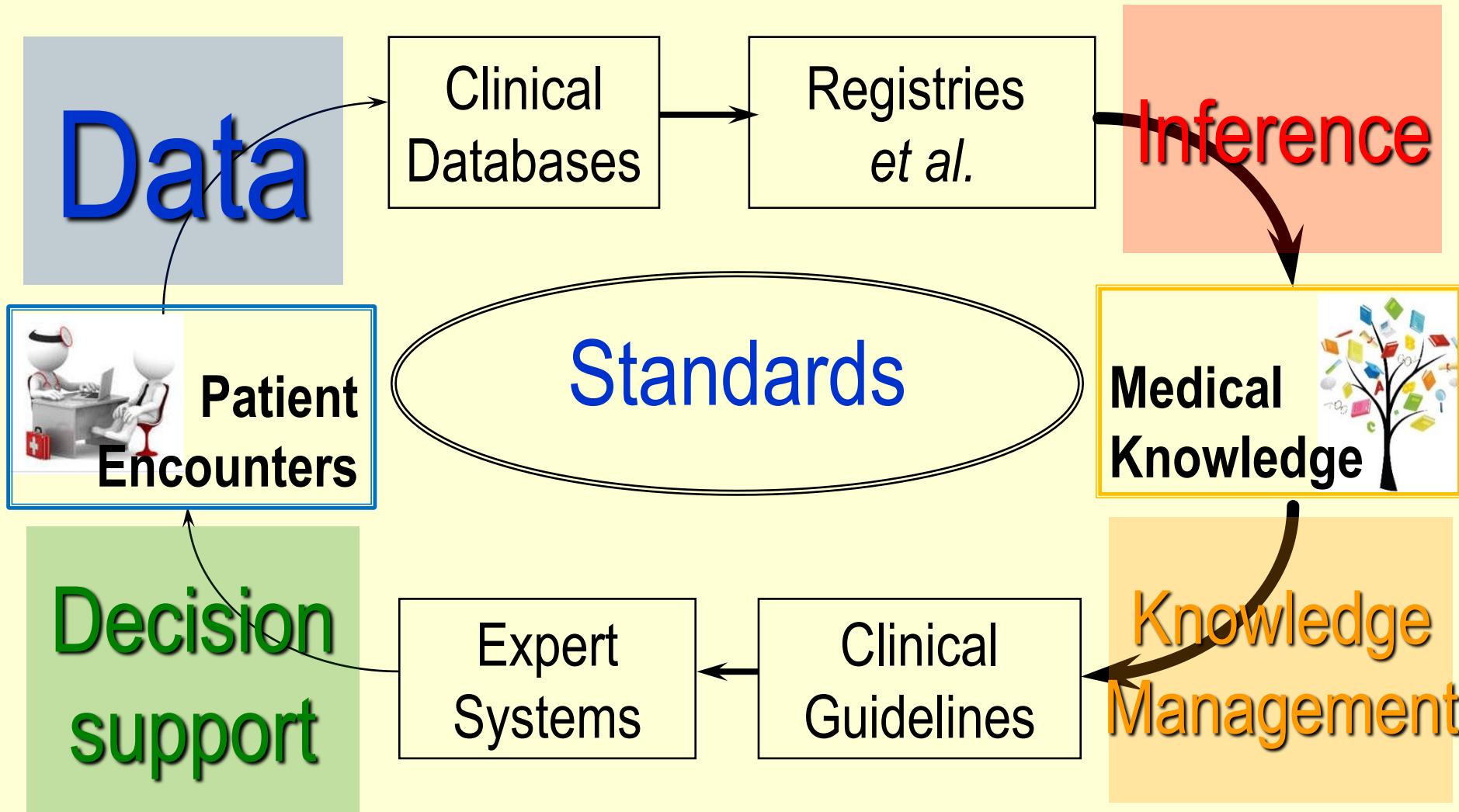
Thanks!

@jessiet1023

Foundations for Learning Health System

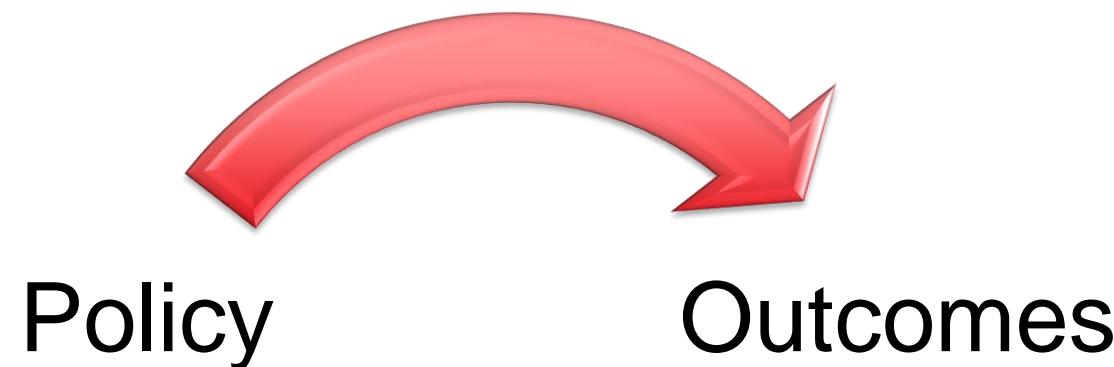
From “Evidence-based Medicine” to “Practice-based Evidence”

Adapted from Chris Chute



The Learning DHHS

Policy impacts outcomes



Outcomes inform policy