

A 21st Century Census Curated Data Enterprise:

*Pilot Demonstration on
Skilled Nursing Facilities*

Presentation to Census Bureau
October 31, 2022

Team: Vicki Lancaster, Dawen Xie,
Henning Mortveit, Samarth Swarup,
Aaron Schroeder, Sallie Keller,
Stephanie Shipp

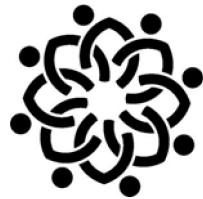


BIOCOMPLEXITY
INSTITUTE



Paradigm Shift for Census Bureau: Focus on Purpose and Use

With the accelerating pace of change in our society and economy, we must constantly innovate how we collect and report official statistics to understand who we are, and where we are going.



Equity and Inclusion



Counting Children



Family Units



Rural Communities



Local Businesses



Tribal Areas



Vulnerable Populations



Elderly Residential Care



The Unhoused



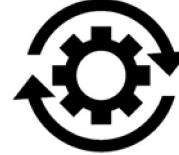
Migration



Climate Change



Gig Economy



Workforce Automation



Differential Undercount



Broadband Deployment



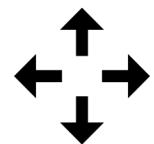
Emergency Preparedness



Economic Mobility



Commuting Patterns



Changing Demographics

Introducing the Curated Data Enterprise

“Linking across time, and surveys, supplemented with third-party and administrative data can help improve the coverage and completeness of Census Bureau data, innovating beyond a survey-alone model that has reached its scientific and practical limits.” (Keller et al. 2022)

We describe this new, innovative state as a 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.

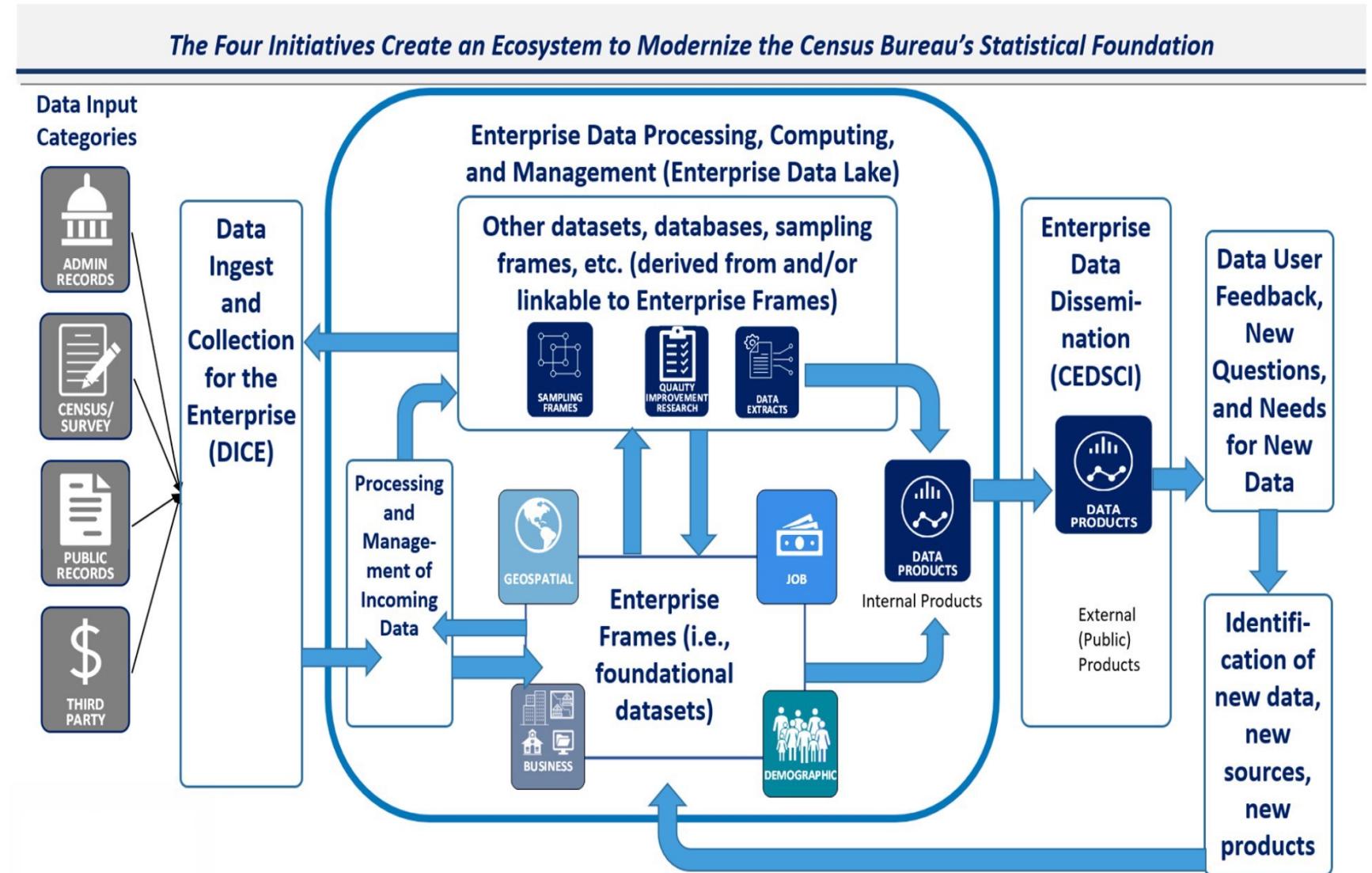
Exploits multiple data sources, from sample surveys, censuses, and other official administrative and third-party data, which can provide more robust, timely, and comprehensive measures when combined.

Building a Census Bureau Modern Data-Centric Ecosystem

Change is here!

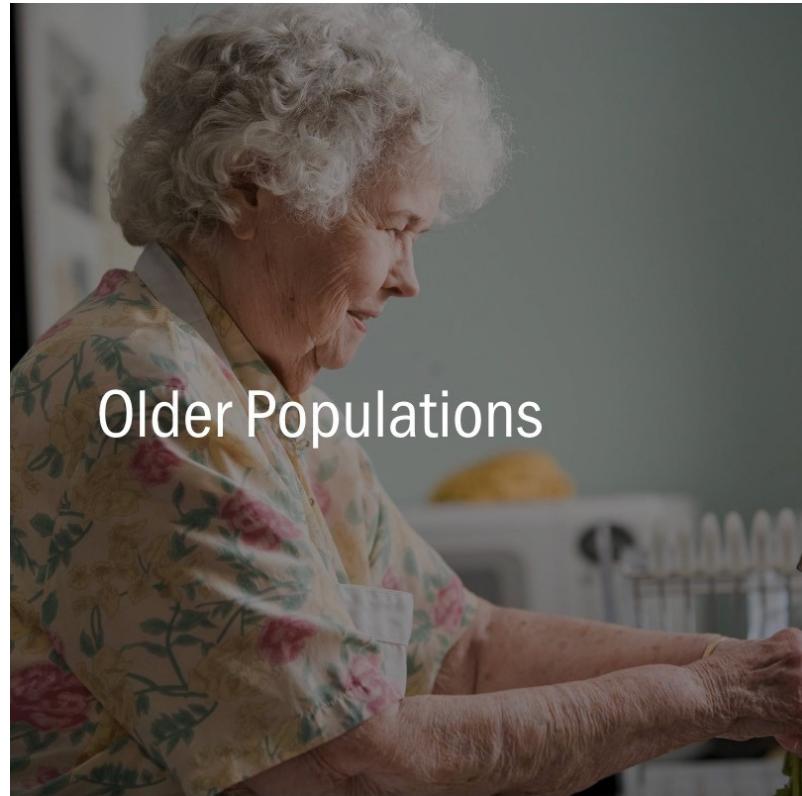
“Our changing culture and rapid changes in data and technology tell us that censuses and surveys alone, while still critical, can no longer answer society’s questions completely or quickly enough to satisfy the modern appetite for information.”

Michael Thieme, Oct. 5, 2022

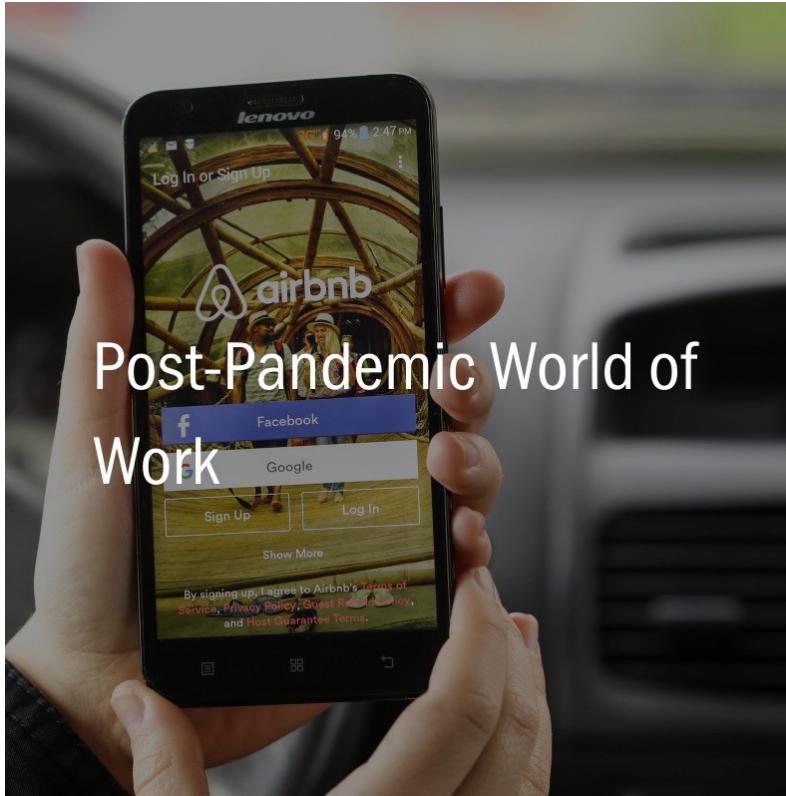


Use Cases - Compelling Examples of “Purpose and Use”

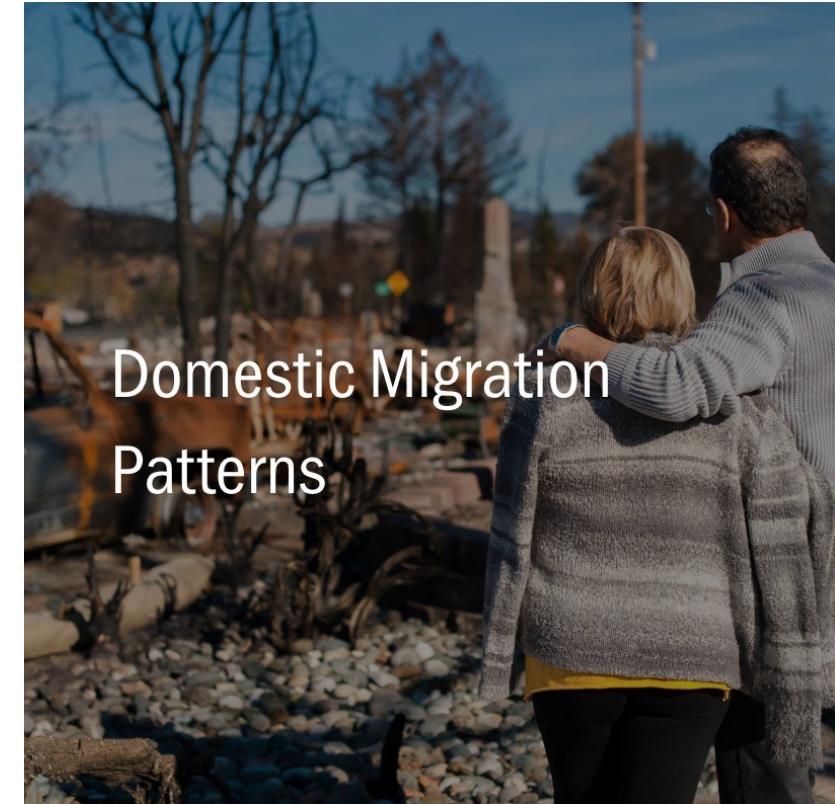
Use Case - a research method to identify, clarify, and organize requirements to create a system or overall system component, such as the Curated Data Enterprise



Older Populations



Post-Pandemic World of
Work



Domestic Migration
Patterns

Use Case—How resilient is a skilled nursing facility...

Climate Resilience

The capacity of an individual, community, or institution to dynamically and effectively respond to shifting climate impact circumstances while continuing to function at an acceptable level.¹

How **Resilient** is a **Skilled Nursing Facility** to an **Extreme Flood Event** ?

Goal: Quantify Skilled Nursing Facility (SNF) resilience to a severe climate event

Conduct a realistic appraisal of the skilled nursing facility that includes the community in which the SNF resides, the physical facility, and the ability of the nursing staff to get to work and function.

¹U.S. Department of Health and Human Services (Fall, 2014). Guenther, R., & Balbus, J. Primary Protection: Enhancing Health Care Resilience for a Changing Climate.

Overview - Curated Data Enterprise Demo

CDE Framework guides
the Use Case research

Purpose & Use: Questions & Working Hypotheses

Data Discovery [Inventory • Screening • Acquisition]

Data Ingestion & Governance

Data Wrangling

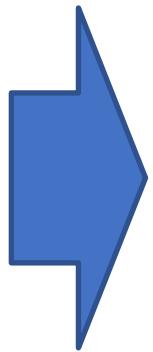
Profiling • Preparation • Linkage • Exploration

Fitness-for-Purpose

Statistical Analyses & Applications Development

Communication & Dissemination

Equity & Ethics Reviews



Census Linked Frames

Geography

People

Jobs

Businesses



Many other types of data...

Networks

Procedural

Climate

Travel

Policies

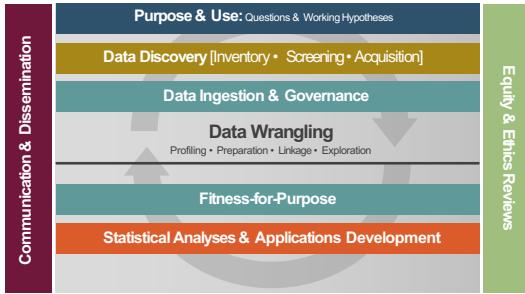
Administrative

Synthetic

Opportunity

Curated Data Enterprise (CDE) Purpose and Use

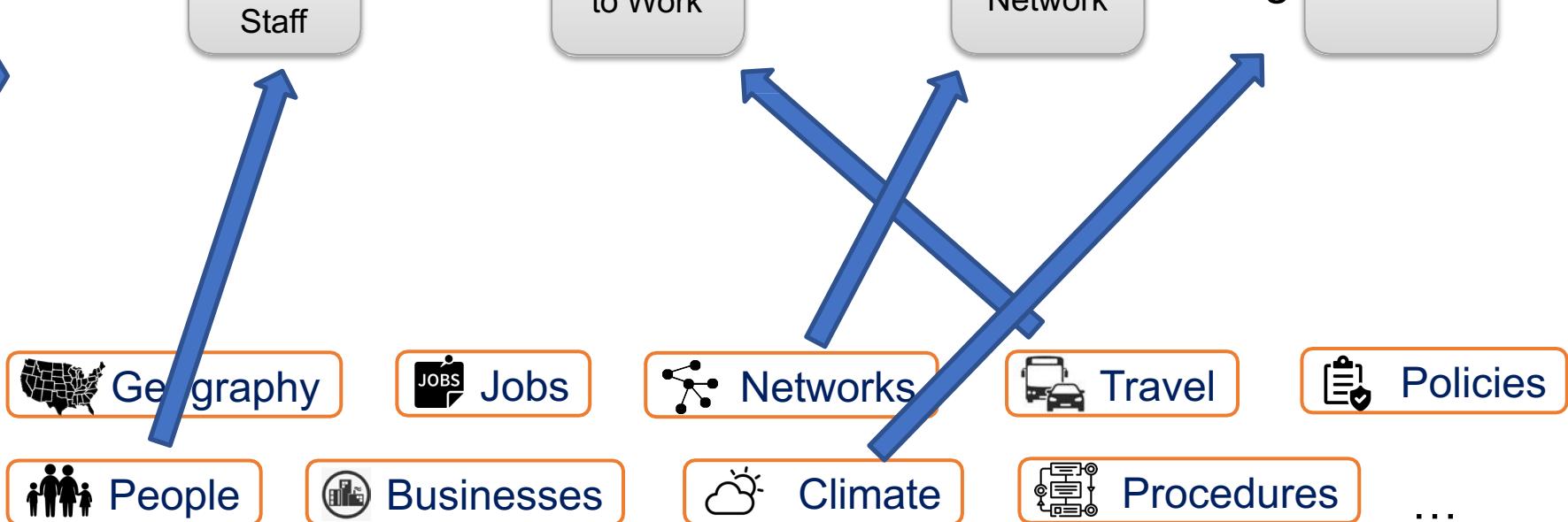
CDE Framework guides
the Use Case research



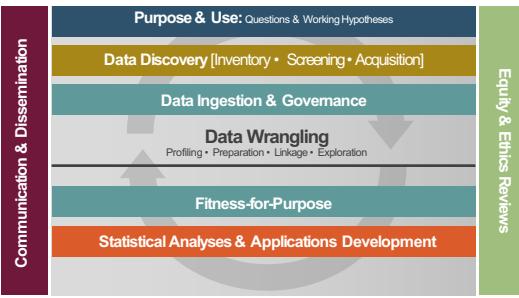
... which selects relevant data and composes into a workflow

Purpose of use
defines a context...

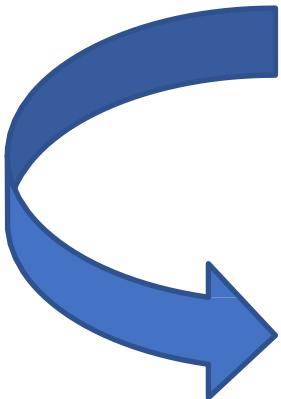
Are SNF Nursing Staff able to Commute to Work over the Road Network during Flooding ?



Curated Data Enterprise Components



The result forms new data products



Are SNF Nursing Staff able to Commute to Work over the Road Network during Flooding ?

 Expected number of SNF nursing staff

 Geography

 Jobs

 Networks

 Travel

 Policies

 People

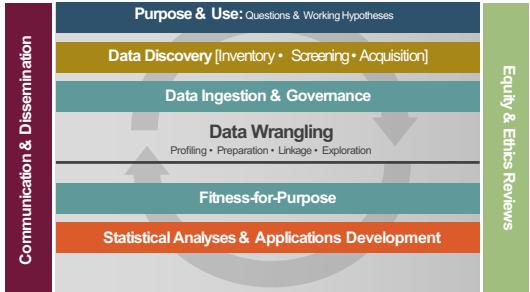
 Businesses

 Weather

 Procedures

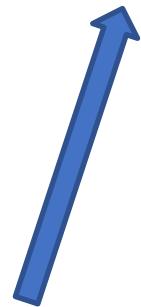
...

Curated Data Enterprise Process Over Time



...more complex measures can be computed

As more data products are created...



SNF Climate Resilience



SNF Nursing Staff



SNF Deficiencies



SNF Community Support



Geography



Jobs



Networks



Travel



Policies



People



Businesses



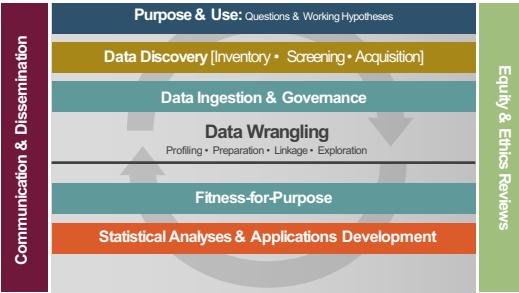
Weather



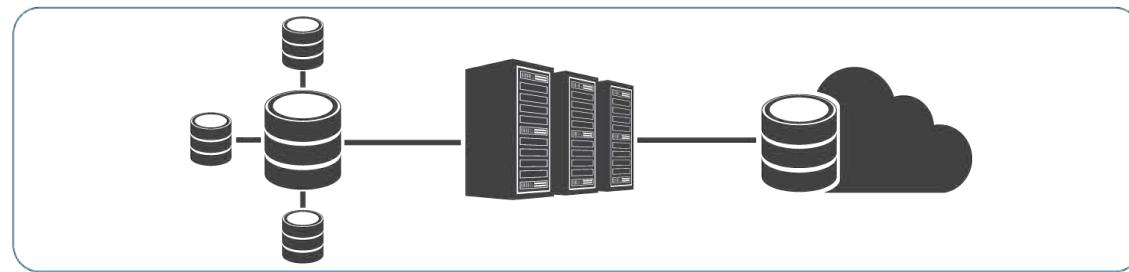
Procedures

...

Curated Data Enterprise Process Over Time



...facilitated and supported by high-performance computing



E(SNF Nursing Staff)



SNF Deficiencies



SNF Community Support



Geography



Jobs



Networks



Travel



Policies



People



Businesses



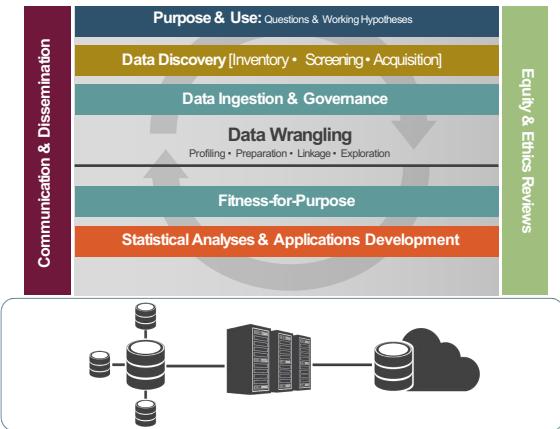
Weather



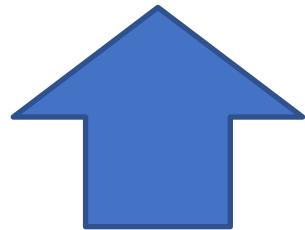
Procedures

...

Curated Data Enterprise Process Over Time



The CDE will provide a **platform** for innovative data use and growth through purpose-driven computational and statistical integration and analysis.



New data sources

New data products

New analyses

...



E(SNF Nursing Staff)



SNF Deficiencies



SNF Community Support



Geography



Jobs



Networks



Travel



Policies



People



Businesses



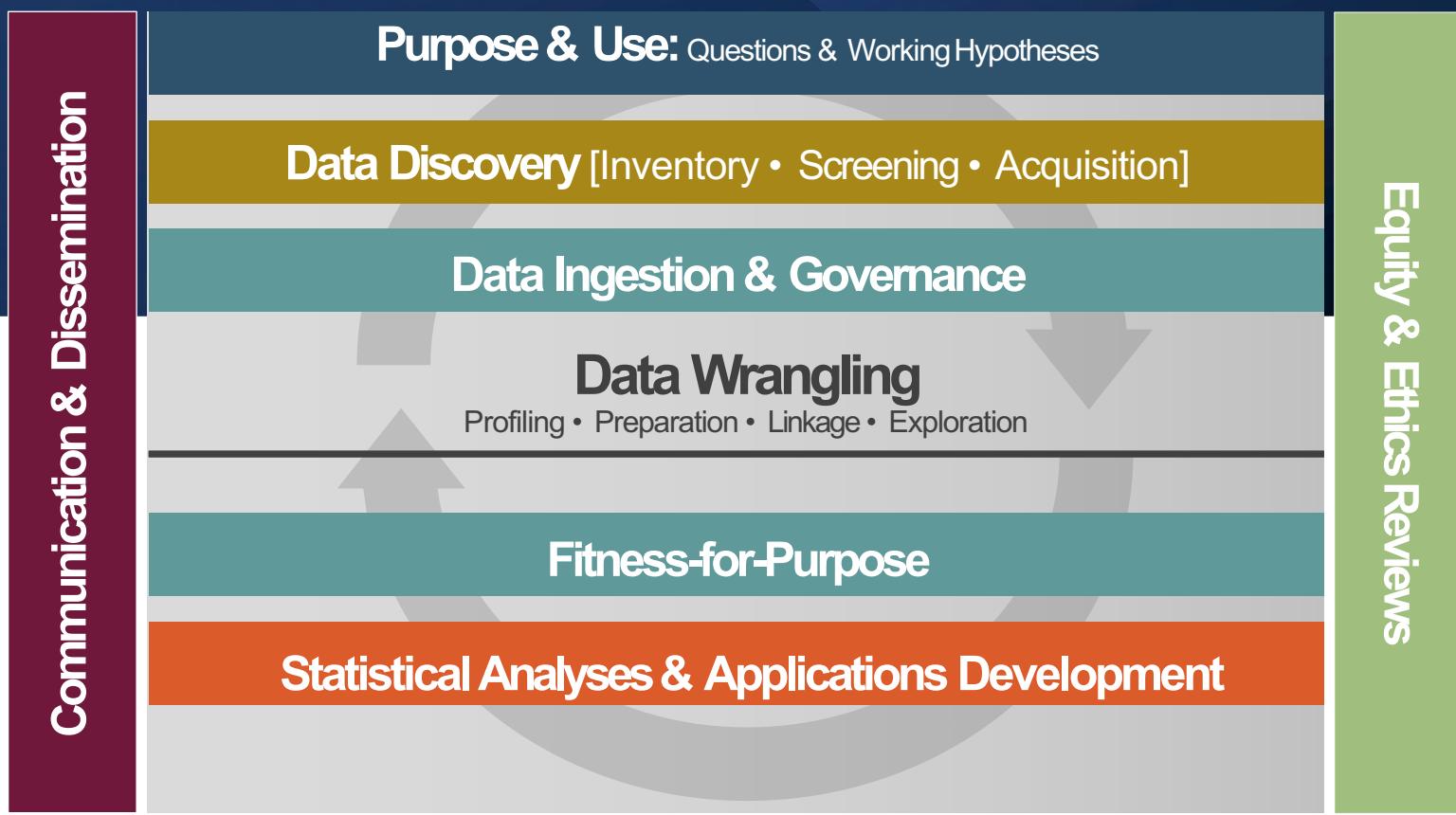
Climate



Procedures

...

Curated Data Enterprise (CDE) Framework - a guide to implementing use case



Purpose & Use [Questions & Working Hypotheses]

Can Skilled Nursing Facilities (SNF) workers get to work?

Are SNFs prepared for emergencies?

Can communities provide support to SNF during an emergency?

Skilled Nursing Facility (SNF) definition

A nursing facility with the staff and equipment to give skilled nursing care and/or skilled rehabilitation services and other related health service.

Source: CMS Glossary

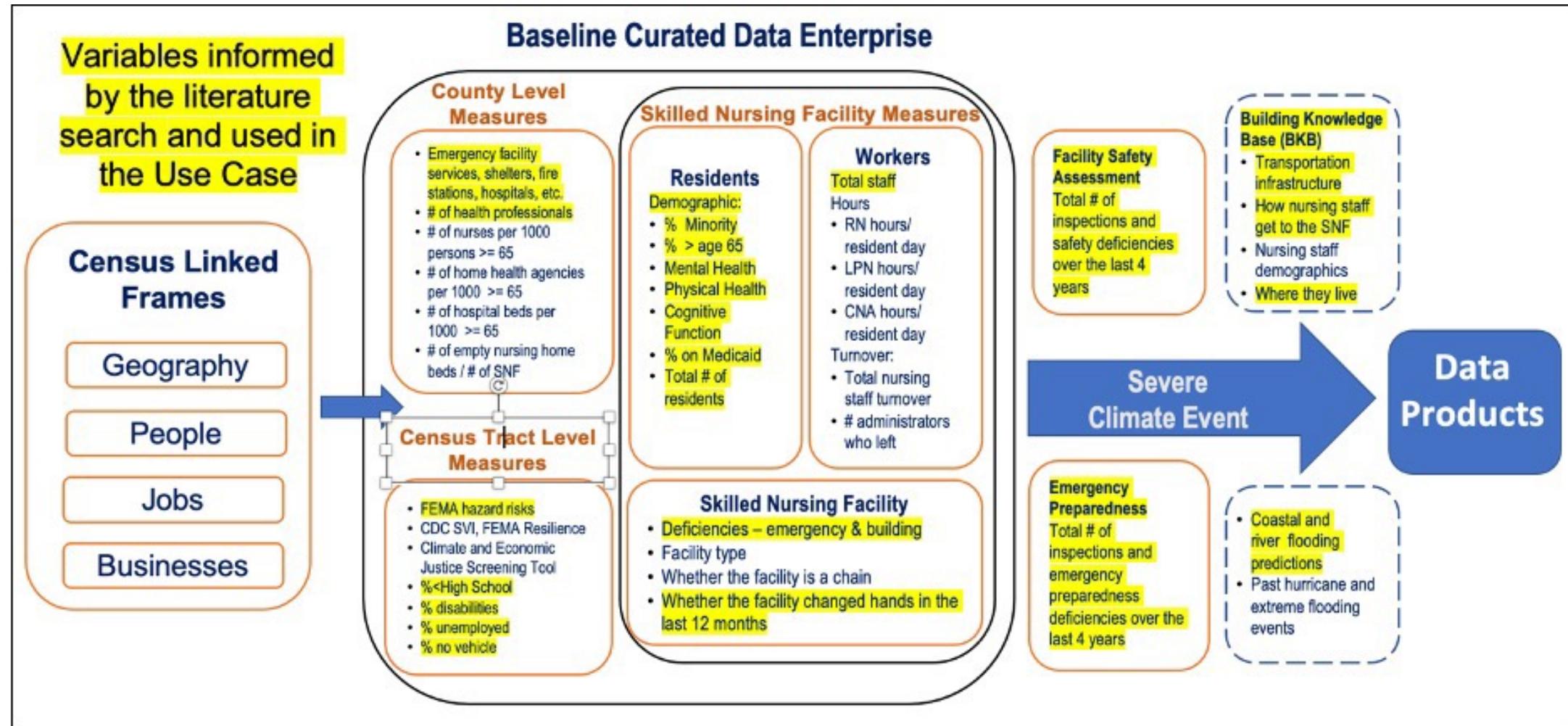
THE WHITE HOUSE



FACT SHEET: Protecting Seniors by Improving Safety and Quality of Care in the Nation's Nursing Homes

FEBRUARY 28, 2022 • STATEMENTS AND RELEASES

Conceptual Data Map Aligned to Purpose and Use



Data Discovery [Inventory-Screening-Acquisition]

SNF Resilience

Department of Health and Human Services

- Centers for Medicare and Medicaid Services (CMS)
 - Payroll Based Journal
 - Provider Information
 - Fire Safety Deficiencies
 - Inspection Dates
 - Ownership
 - Citation Lookup

LTC Focus Data sponsored by NIH

- Certification and Survey Provider Enhanced Reporting

Biocomplexity Institute synthetic data

Community Resilience

(County/City, Census tract)

Department of Commerce

- Census American Community Survey

Department of Homeland Security

- Federal Emergency Management Administration: National Risk Index for Natural Disasters
- Homeland Infrastructure Foundation-Level Data

Department of Health and Human Services

- Health Resources and Services Administration: Area Resource Files

Data Ingestion and Governance - GitHub

Folders, sub-Folders, and sub-sub-Folders

Documents		Source Code		Data	
Literature	Products	Discovery	Analyses	SNF***	Community
Deficiencies	Data Tables**	Subfolder: Discovery <ul style="list-style-type: none">▪ File: EDAshape.R Output: eda_emergency_preparedness_deficiency_index.pdf Output: eda_fire_life_safe_code_deficiency_index.pdf▪ File: EDA_flatviolin.R Output: eda_snf_deficiency_index_by_hazard.pdf▪ File: EDA_boxplots.R Output: county_population_indicator_resilience_box_plots.pdf▪ File: EDA_Choropleths_Resilience_Indicator_Variables.R Output: census_tract_population_16_and_over_unemployed_choropleth.pdf Output: census_tract_population_no_hs_diploma_choropleth.p Output: census_tract_population_with_disability_choropleth.pdf Output: county_old_age_dependency_choropleth.pdf	Subfolder: Analyses <ul style="list-style-type: none">▪ File: SNF_Deficiency_Index.R Output: va_snf_deficiency_indices_k_e.csv▪ File: VA_Population_Resilience_Index_Census_Tract.R Output: census_tract_population_resilience_index_choropleth.pdf Output: va_census_tract_population_resilience.csv▪ File: VA_Population_Resilience_Index_County.R Output: county_population_resilience_index_with_gini_choropleth.pdf Output: county_population_indicator_resilience_box_plots.pdf Output: va_county_population_resilience.csv▪ File: VA_Probability_of_Getting_to_SNF.R Output: va_snf_estimated_average_daily_nursing_staff_during_extreme_flood_event.csv Output: snf_estimate_daily_nursing_staff_during_climate_event.pdf	Facility*	Demographic
Emergency Preparedness	Curation Report & Presentations			Nursing Staff*	Climate Change Risk*
Nursing Staff Levels				Residents*	Resilience*
Ownership	Visualizations			Owners*	
Quality	Processes**				
Resilience					

https://github.com/uva-bi-sdad/census_cde_demo_2/

*Each folder contains a meta data folder with data dictionaries and technical documents.

**Contains additional subfolders.

***SNF is Skilled Nursing Facilities

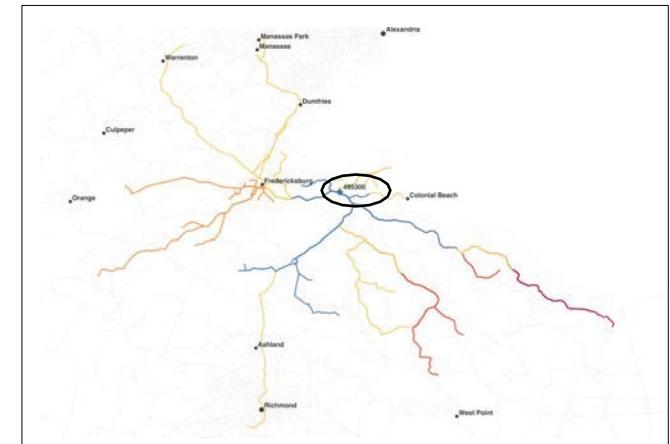
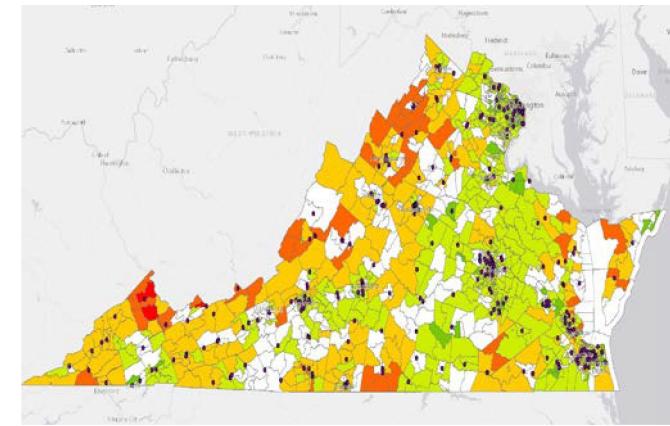
Data Wrangling and Exploration

RQ1: Can SNF workers get to work?

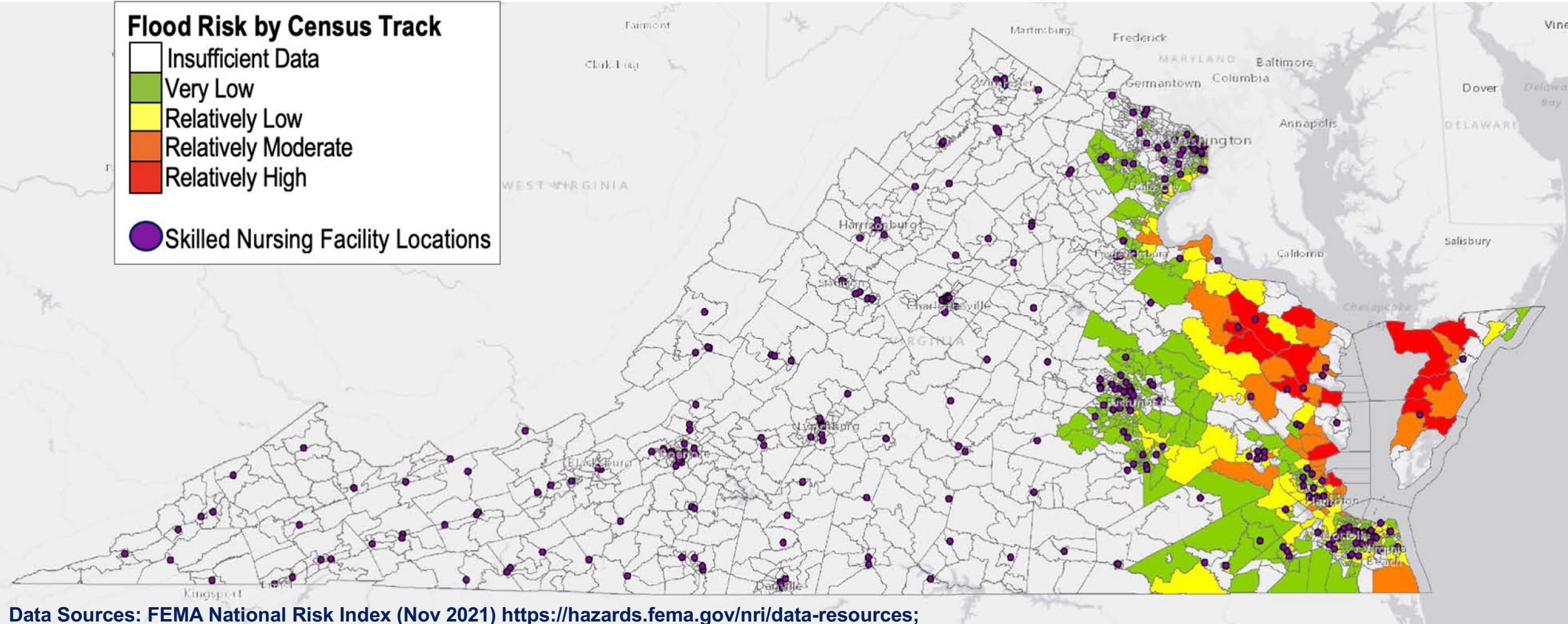
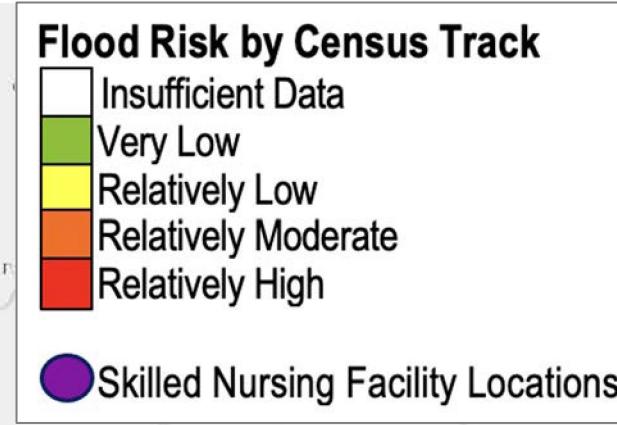
Identifying Flood Risk and SNF Workers' Transportation Routes at Census Tract Level

RQ1: Can SNF workers get to work during a flooding event?

- Data discovery – flooding risks, nursing home locations, road networks
- Compute the likely commuting route for each synthetic worker - Identify workers, home locations, nursing home locations, and routes
- Overlay the river and coastal flooding by census tracts
- Estimate flooding risk along the transportation routes
- For each SNF, there is a flood risk; treat that value as a proxy for the probability of a worker not getting to work
- Calculate risk for each SNF worker as the percent of workers that cannot make it to work during a flooding event

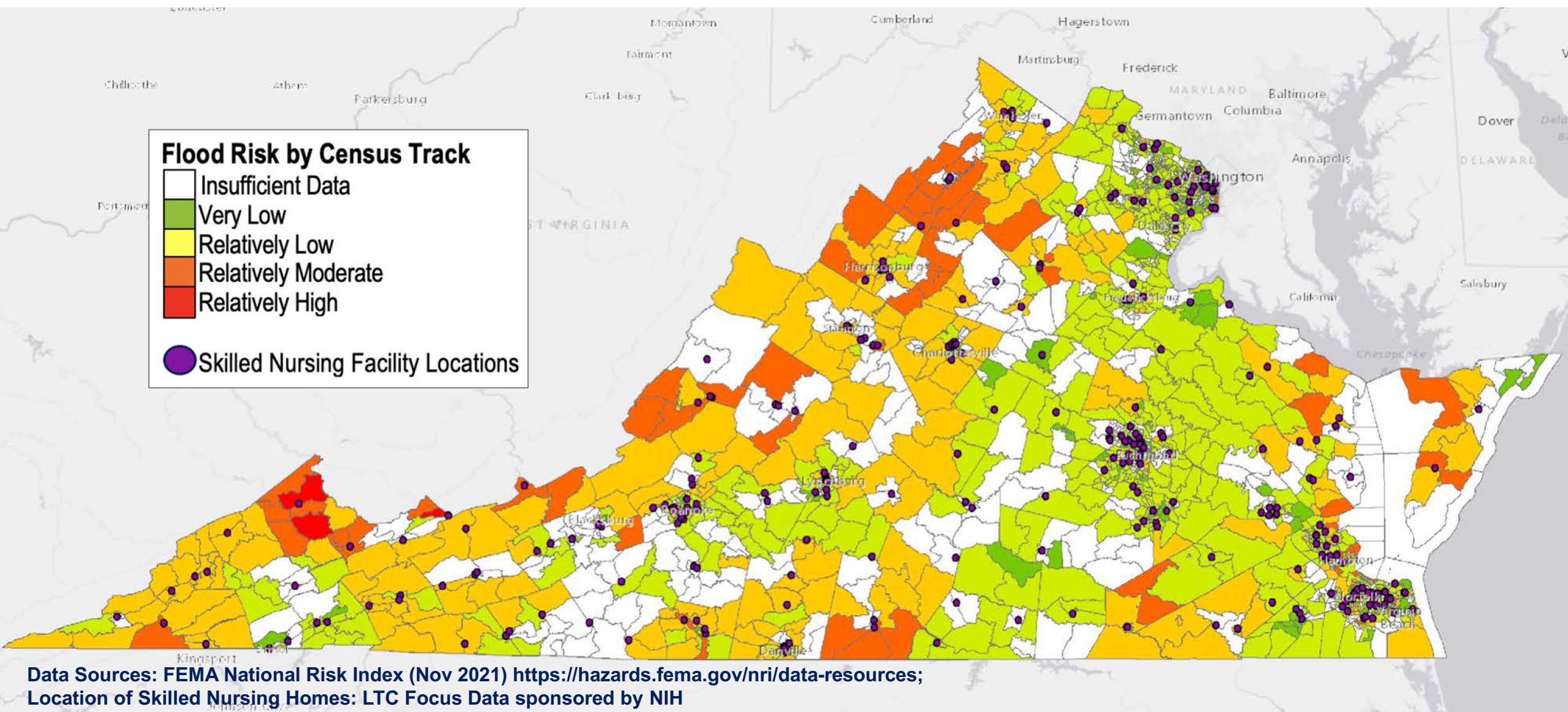


Coastal Flooding Risk in Virginia & SNF Locations



Data Sources: FEMA National Risk Index (Nov 2021) <https://hazards.fema.gov/nri/data-resources>;
Location of Skilled Nursing Homes: LTC Focus Data sponsored by NIH

River Flooding Risk in Virginia & SNF Locations



Risk Scores – Can SNF workers make it to work?

Assumption: Use the combined value of maximum river flood + maximum coast flood risk along worker's transportation route as the worker's probability of not getting to work during a flooding event

Annual average risk across Virginia: The 283 SNF used in our research have a total average daily nursing staff of 12,609. We estimate that ~10,005 (95% CI: 9,013, 10,700 or 79% will be able to get work.

Percent Loss: The individual SNF nursing staff percentage loss ranges from a high of 52% to a low of 7%, with a median of 19%.

Case Studies – Flood Risk in Virginia

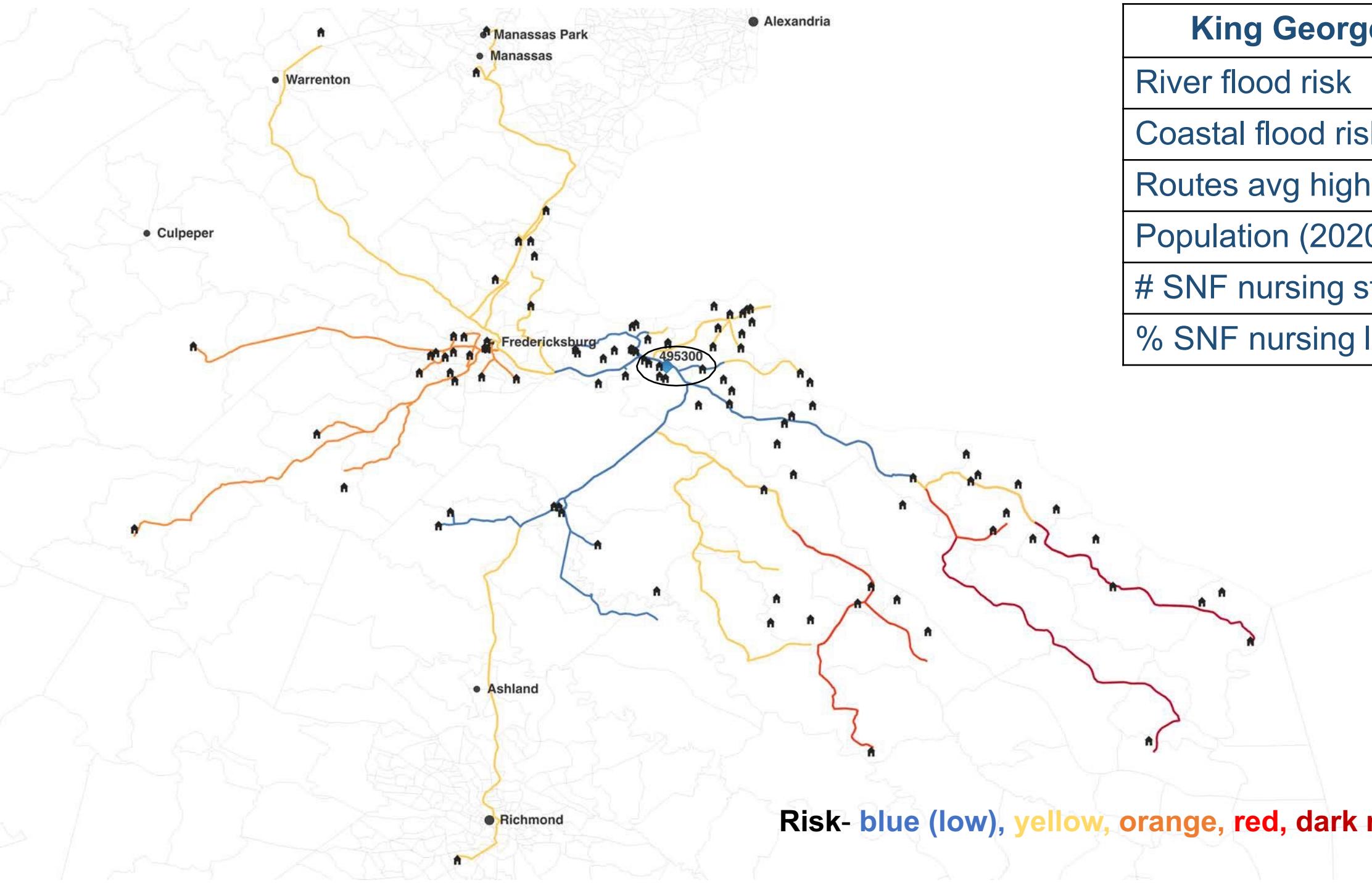
1. King George, VA, 41 nursing staff employees; **mixed risk routes**
2. Warsaw, VA, 35 nursing staff employee; **most routes have high risks**
3. Hopewell, VA, 48 nursing staff employees; **most routes have low risks**



Case Study 1: King George, Virginia

Located near Potomac River and the opening to the Chesapeake Bay





King George 495300

River flood risk	3.5
Coastal flood risk	6.2
Routes avg high risk	21.7
Population (2020)	4601
# SNF nursing staff	41
% SNF nursing loss	24

Risk- blue (low), yellow, orange, red, dark red (high)

Data Wrangling and Exploration

RQ2: Are SNF prepared for emergencies?

Identifying SNF Resilience

Computations

- SNF Emergency Preparedness/fire safety deficiency index as a proxy for SNF resilience

Are SNF prepared for emergencies?

CMS Skilled Nursing Home Deficiencies as a proxy for preparedness

Emergency Preparedness (EP) Deficiencies – 4 elements

- Emergency plans
- Communication plan
- Training and testing
- Policies and procedures

Life Safety Code (LSC) Deficiencies - Fire protection requirements

- Designed to provide a reasonable degree of safety from fire, smoke, and panic
- Construction
- Protection
- Operational features

EP Deficiency Index

Calculate

Calculate a separate Emergency Preparedness (EP) deficiency index for each Skilled Nursing Facility.

Normalize

Normalize the number deficiencies over the past four years by the number of SNF inspections over the past four year (2018 – 2022).

Average

Average the number of days it took to resolve each deficiency.

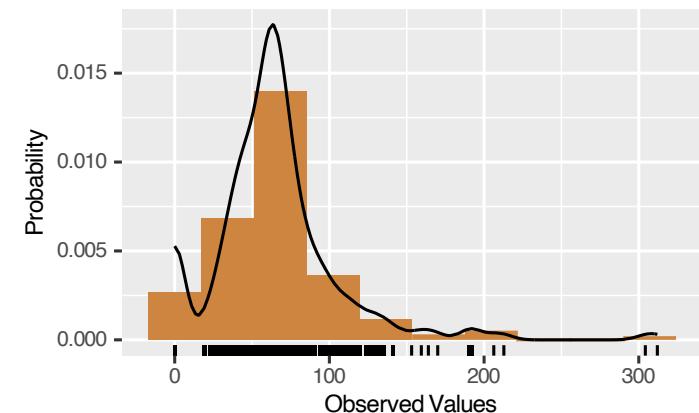
Transform

Transform the letter inspection rating for scope and severity to a numerical weight using the CMS technical guide, Care Compare Nursing Home Five-Star Quality Rating System; then we took the average.

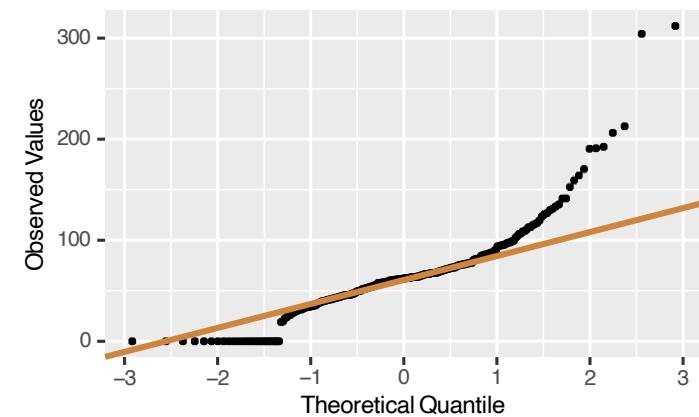
Add

Add 1 through 3 to get the deficiency index for EP and LSC.

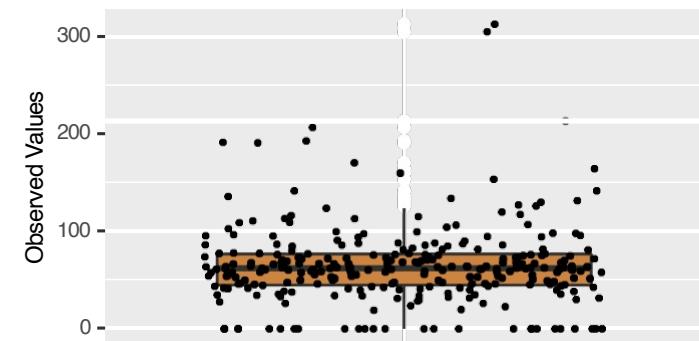
Histogram



Q-Q Plot



Box Plot



DATA SOURCE: The Centers for Medicare & Medicaid Services

Fire Safety Deficiencies <https://data.cms.gov/provider-data/dataset/lfjz-ge4w>

Inspection Dates <https://data.cms.gov/provider-data/dataset/svdt-c123>

Data Wrangling and Exploration

Can the community provide support to SNF
during emergency events?

Identifying Community Assets

Computations

- Community Resilience Indices
- Shelter Facilities and Emergency Services Providers per 10K people
- Medical Professional Shortages per 10K people

RQ3: Can community support SNF during emergencies?

County Resilience Indices

Using American Community Survey and local data to measure resiliency.

We summed the county level percentages for the Census variables:

- Fraction Employed
- Fraction with No Disability
- Fraction with a High School Diploma or Greater
- Fraction of Households with at least one vehicle, and
- Reverse GINI Index – so all indicators are in positive direction



Community Resilience Indicator Analysis:

County-Level Analysis of Commonly Used Indicators from Peer-Reviewed Research

2020 Update



Homeland
Security

Argonne
NATIONAL LABORATORY

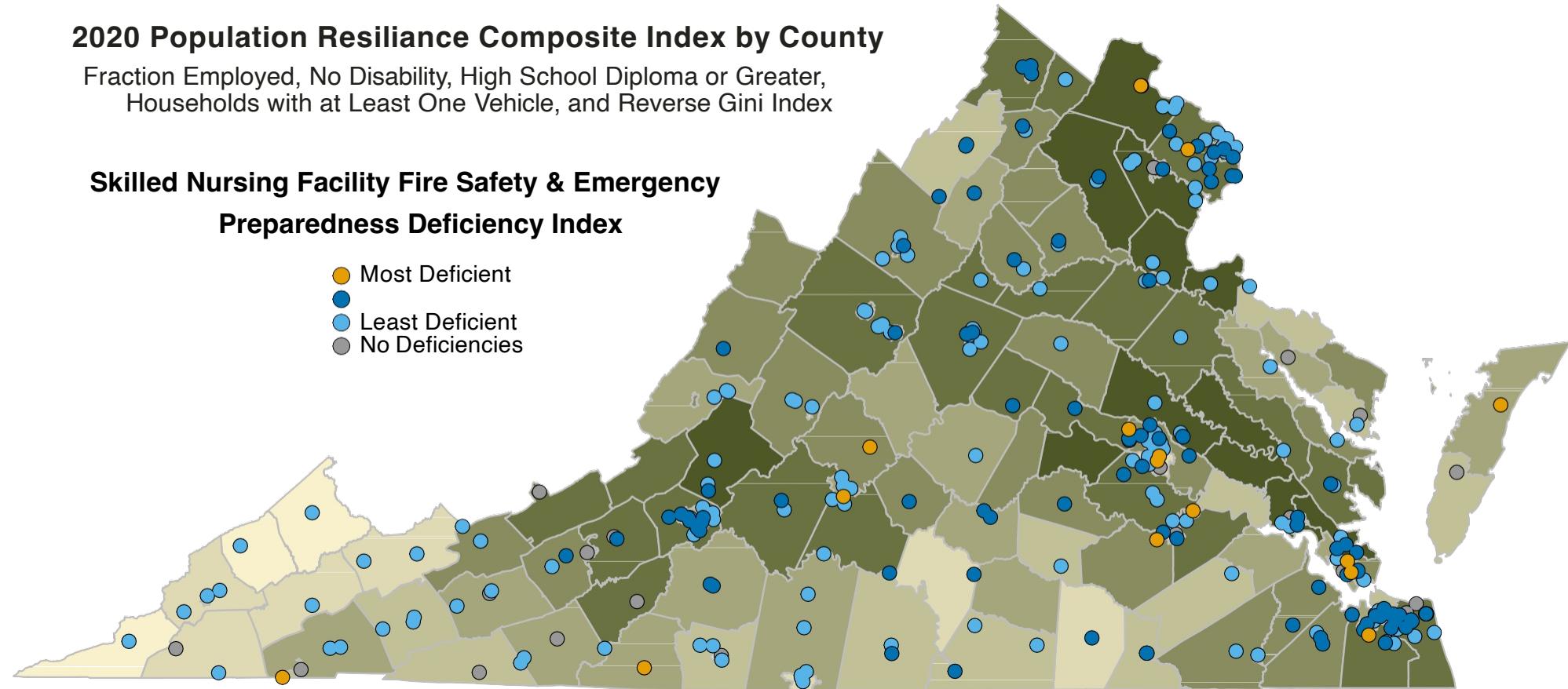
SNF Deficiency & County Resilience Indices

2020 Population Resiliance Composite Index by County

Fraction Employed, No Disability, High School Diploma or Greater,
Households with at Least One Vehicle, and Reverse Gini Index

Skilled Nursing Facility Fire Safety & Emergency Preparedness Deficiency Index

- Most Deficient
- Least Deficient
- No Deficiencies



Population Resiliency Composite Index
(0=Least Resilient, 1=Most Resilient)



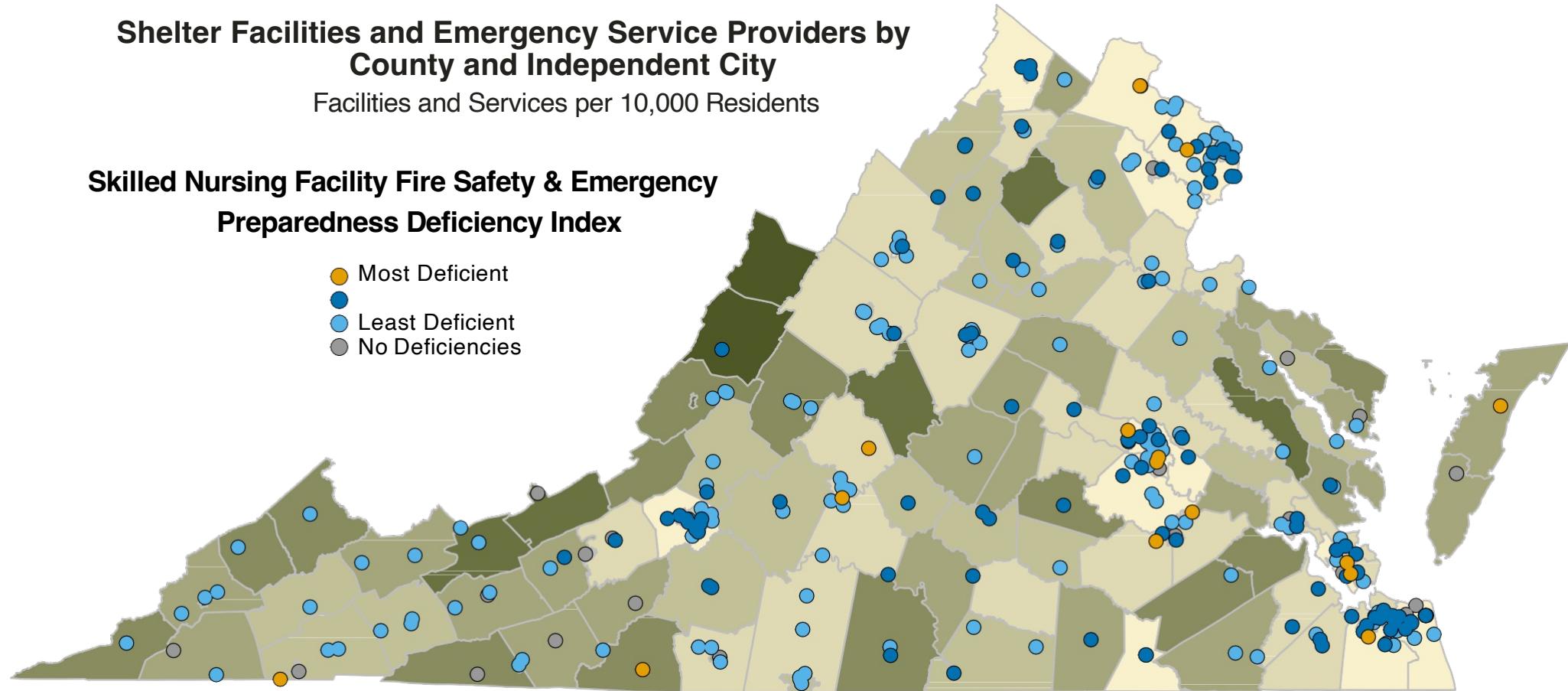
Virginia Emergency Facility Services & Shelters

Shelter Facilities and Emergency Service Providers by County and Independent City

Facilities and Services per 10,000 Residents

Skilled Nursing Facility Fire Safety & Emergency Preparedness Deficiency Index

- Most Deficient
- Least Deficient
- No Deficiencies



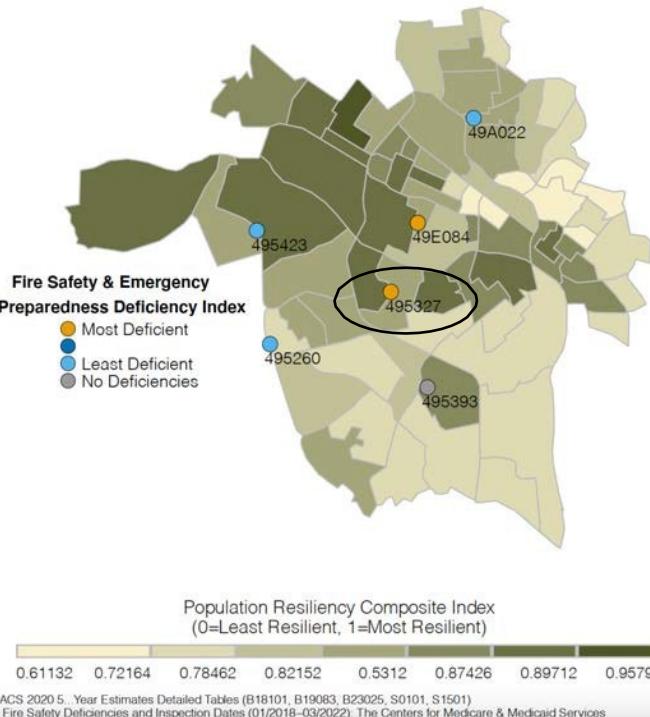
Number of Facilities & Services/10,000 Residents



If we take a closer look, what can we learn?

2020 Richmond Resiliency Composite Index by Census Tract

Fraction Employed, No Disability, High School Diploma or Greater, Households with at Least One Vehicle, and Reverse Gini Index



At the SNF level, no facility changed ownership in the last year.

- 3 for-profit, 2 not-for-profit, and 1 government owner.

Relatively high riverine flood risk and low coastal flood risk.

- On average 18% of the SNF workers will not make it to work if an extreme flooding event occurs.

SNF 495327 had emergency preparedness deficiencies that were wide in scope but not immediate jeopardy; they failed to:

- Conduct tests and exercises for requirements (EP);
- Establish policies and procedures for sheltering (EP); and
- Test and maintain the automatic sprinkler and fire alarm systems (LSC).

City of Richmond SNFs

	495260	495327	495393	495423	49A022	49E084
Maximum Riverine Flood Risk	13.80	15.51	14.57	16.39	14.41	14.48
Maximum Coastal Flood Risk	2.91	2.33	1.66	2.28	1.62	2.30
% Nursing Staff Loss if Flood Event	18%	20%	17%	21%	15%	17%
SNF Deficiency Index	Low	High	None	Low	Low	High

Data Ingestion and Governance - GitHub

Folders, sub-Folders, and sub-sub-Folders

Documents		Source Code		Data	
Literature	Products	Discovery	Analyses	SNF***	Community
Deficiencies	Data Tables**	Subfolder: Discovery <ul style="list-style-type: none">▪ File: EDAshape.R Output: eda_emergency_preparedness_deficiency_index.pdf Output: eda_fire_life_safe_code_deficiency_index.pdf▪ File: EDA_flatviolin.R Output: eda_snf_deficiency_index_by_hazard.pdf▪ File: EDA_boxplots.R Output: county_population_indicator_resilience_box_plots.pdf▪ File: EDA_Choropleths_Resilience_Indicator_Variables.R Output: census_tract_population_16_and_over_unemployed_choropleth.pdf Output: census_tract_population_no_hs_diploma_choropleth.p Output: census_tract_population_with_disability_choropleth.pdf Output: county_old_age_dependency_choropleth.pdf	Subfolder: Analyses <ul style="list-style-type: none">▪ File: SNF_Deficiency_Index.R Output: va_snf_deficiency_indices_k_e.csv▪ File: VA_Population_Resilience_Index_Census_Tract.R Output: census_tract_population_resilience_index_choropleth.pdf Output: va_census_tract_population_resilience.csv▪ File: VA_Population_Resilience_Index_County.R Output: county_population_resilience_index_with_gini_choropleth.pdf Output: county_population_indicator_resilience_box_plots.pdf Output: va_county_population_resilience.csv▪ File: VA_Probability_of_Getting_to_SNF.R Output: va_snf_estimated_average_daily_nursing_staff_during_extreme_flood_event.csv Output: snf_estimate_daily_nursing_staff_during_climate_event.pdf	Facility*	Demographic
Emergency Preparedness	Curation Report & Presentations			Nursing Staff*	Climate Change Risk*
Nursing Staff Levels	Visualizations	Subfolder: Discovery <ul style="list-style-type: none">▪ File: EDAshape.R Output: eda_emergency_preparedness_deficiency_index.pdf Output: eda_fire_life_safe_code_deficiency_index.pdf▪ File: EDA_flatviolin.R Output: eda_snf_deficiency_index_by_hazard.pdf▪ File: EDA_boxplots.R Output: county_population_indicator_resilience_box_plots.pdf▪ File: EDA_Choropleths_Resilience_Indicator_Variables.R Output: census_tract_population_16_and_over_unemployed_choropleth.pdf Output: census_tract_population_no_hs_diploma_choropleth.p Output: census_tract_population_with_disability_choropleth.pdf Output: county_old_age_dependency_choropleth.pdf	Subfolder: Analyses <ul style="list-style-type: none">▪ File: SNF_Deficiency_Index.R Output: va_snf_deficiency_indices_k_e.csv▪ File: VA_Population_Resilience_Index_Census_Tract.R Output: census_tract_population_resilience_index_choropleth.pdf Output: va_census_tract_population_resilience.csv▪ File: VA_Population_Resilience_Index_County.R Output: county_population_resilience_index_with_gini_choropleth.pdf Output: county_population_indicator_resilience_box_plots.pdf Output: va_county_population_resilience.csv▪ File: VA_Probability_of_Getting_to_SNF.R Output: va_snf_estimated_average_daily_nursing_staff_during_extreme_flood_event.csv Output: snf_estimate_daily_nursing_staff_during_climate_event.pdf	Residents*	Resilience*
Ownership	Processes**			Owners*	
Quality					
Resilience					

https://github.com/uva-bi-sdad/census_cde_demo_2/

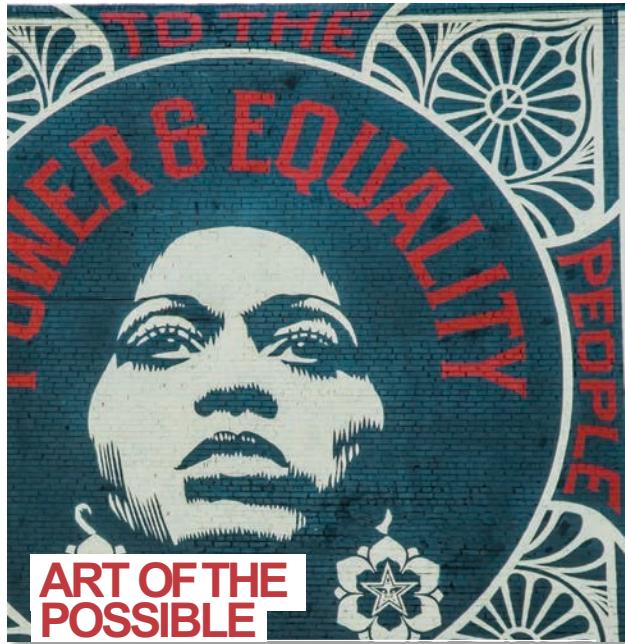
*Each folder contains a meta data folder with data dictionaries and technical documents.

**Contains additional subfolders.

***SNF is Skilled Nursing Facilities

Use and Reuse of CDE Components

Resolve once, use many times



Use cases inform CDE component development and reusability

- Data Sources
- Functions
- Applications

Data sources

- Synthetic population estimates
- SNF location and characteristics
- Baseline community emergency preparedness assets

Functions

- Creating transportation routes
- Computing the probability of flood risks
- Estimating SNF resiliency & community resilience

Applications

- Curation - Organize and document data to enable the discovery and reuse in research and data production
- Wrangle & integrate many types and sources of data
- Understand differences in similar measures

Discussion