What are Grids?

A discussion on grids and gridded census data.

New Jersey SDC meeting October 5th, 2023

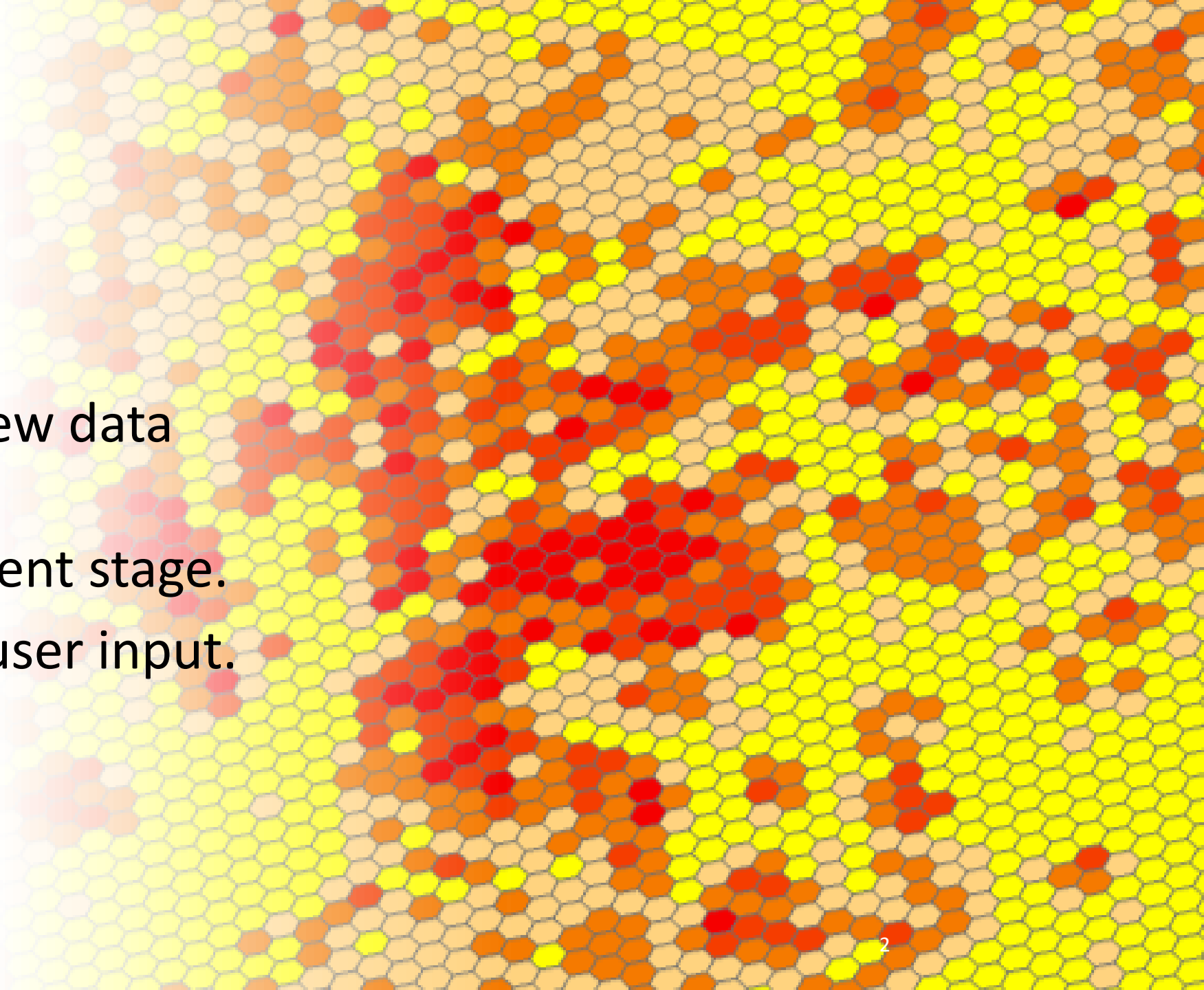
United States® Census Bureau

The following slides DO NOT contain Title 13 data.

U.S. Census Bureau GEOGRAPHY

Gridded Census Datasets

- Researching a new data product.
- Initial development stage.
- Gathering data user input.



Project Overview

Purpose:

- Expand our available product portfolio to **meet changing demands**.
- More accessible, interoperable, and relevant data.
- Producing **new and innovative** Census data products.
- Fulfillment of mission needs and **meeting stakeholder requests**.
- **Creating quality products** to ensure we meet today's data demands.
- Providing tools to **support informed, timely decisions**.

Expectations:

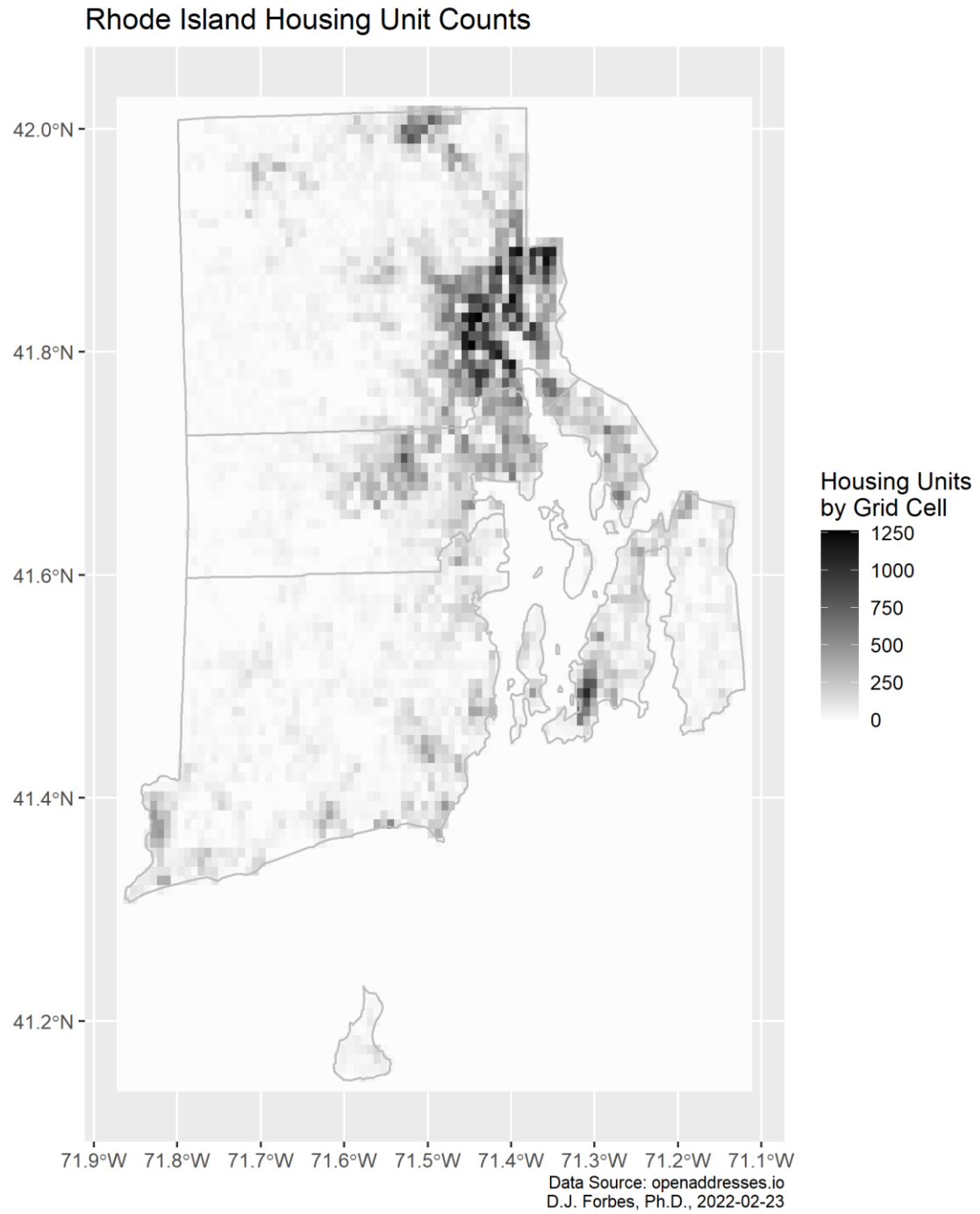
- Gridded data products will be **in addition to** current administrative/statistical units.
- We will produce gridded data products as a **standard annual public delivery**.

United States® Census Bureau

U.S. Census Bureau GEOGRAPHY

What are grids?

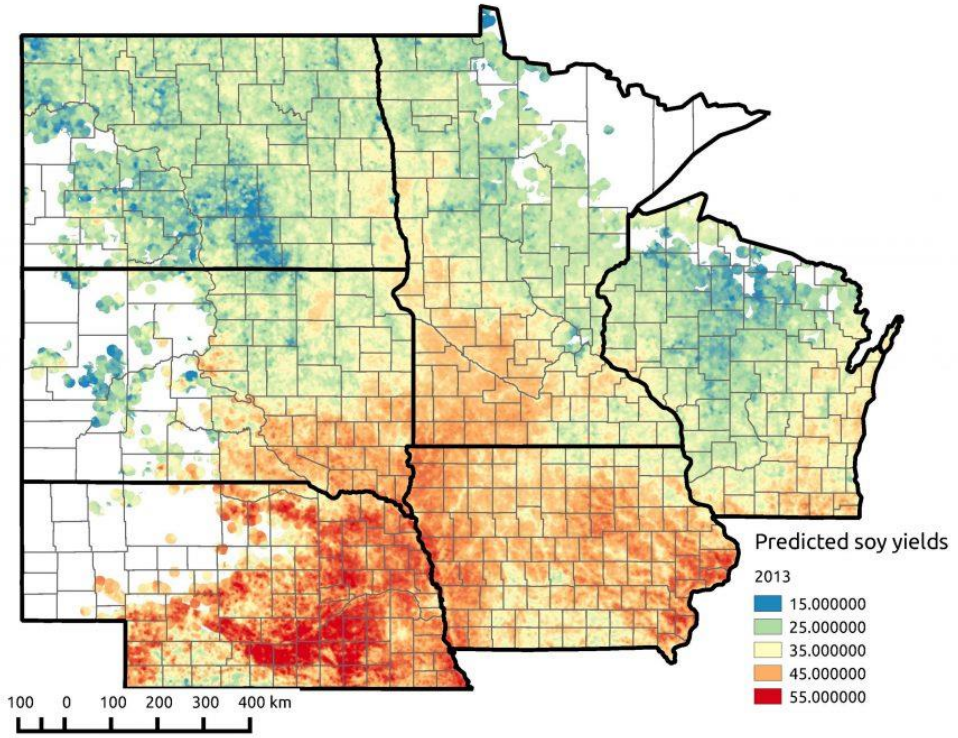
- Grids are characterized by a **pattern of regular repeated shapes** that fit together without gaps or overlaps.
- Can be any shape that **does not create gaps or overlaps**, such as squares, rectangles, triangles, quadrilaterals, or hexagons.
- **GEO has received inquiries** from external partners for Census data provided in a gridded format.
- **Grids have several advantages** in comparison to administrative units.



United States® Census Bureau

What are gridded data?

- Gridding data is to assign data values to grid cells.
- Grids are always aggregates because they use single values to describe (represent) a two-dimensional area of the Earth's surface.
- Examples:
  - Counts (number of housing units)
  - Sums (economic output in dollars, expected crop yield)
  - Averages (temperature, rainfall, property value)
  - Other (Ranks, presence/absence, etc)



United States® Census Bureau

U.S. Census Bureau GEOGRAPHY

Grids Advantages

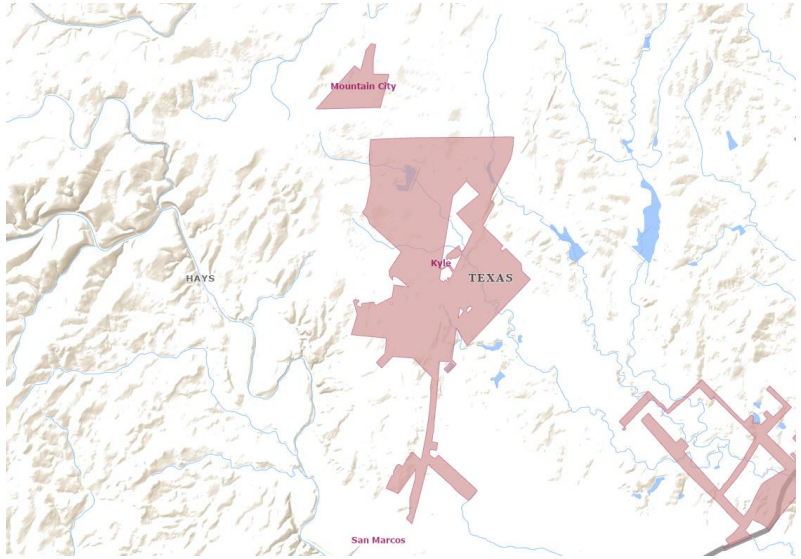
- **Uniform area** (densities).
  - Uniform area supports comparative analysis.
- **Consistent** and unchanging.
  - Supports change detection and time series analyses.
- **Easily integrated** with other gridded data.
  - Provide better support for disaster response.
- Provide an alternative to (but **do not replace**) administrative/statistical boundaries.
- May assist in **small area analysis** and estimation.
  - Regions of interest are more easily defined and remain consistent.
- Potential to serve as **consistent work units** for internal data processing or sampling.

Where is "**Here**"? Part 1

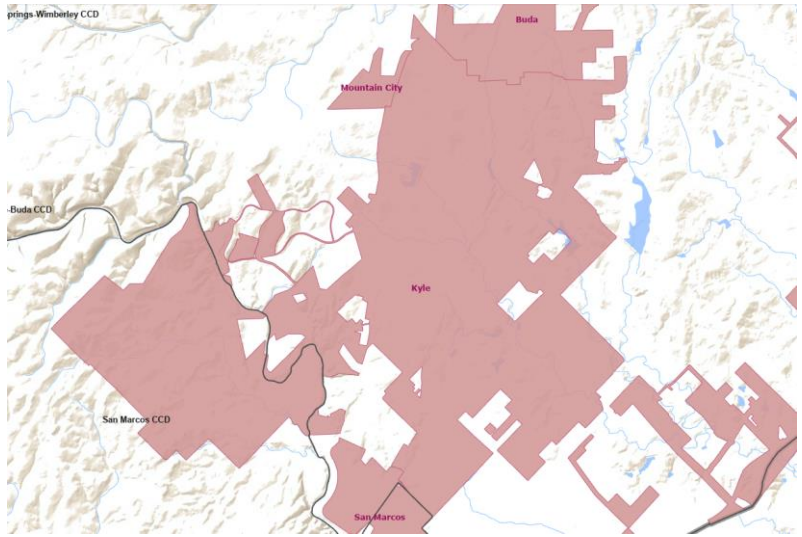
- Sometimes this question can be meaningfully answered with a word linked to an administrative or statistical boundary (e.g. "here in Maryland", "here in Philadelphia", "here in Census tract 7").
- But sometimes those units aren't appropriate because:
  - boundaries may have changed (illusory stability—in fact change is the norm for administrative units);
  - comparisons are difficult or hard to intuitively grasp owing to units radically different physical areas;
  - these units do not actually encapsulate the spatial extent of many important events and phenomena.

Boundary Changes

Historical population		
Census	Pop.	%±
1890	779	—
1910	742	—
1920	744	0.3%
1930	606	−18.5%
1940	874	44.2%
1950	888	1.6%
1960	1,023	15.2%
1970	1,629	59.2%
1980	2,093	28.5%
1990	2,225	6.3%
2000	5,314	138.8%
2010	28,016	427.2%
2020	45,697	63.1%
2022 (est.)	57,470	25.8%
U.S. Decennial Census[16][failed verification]		2020[3]



“Kyle, TX”: 2010

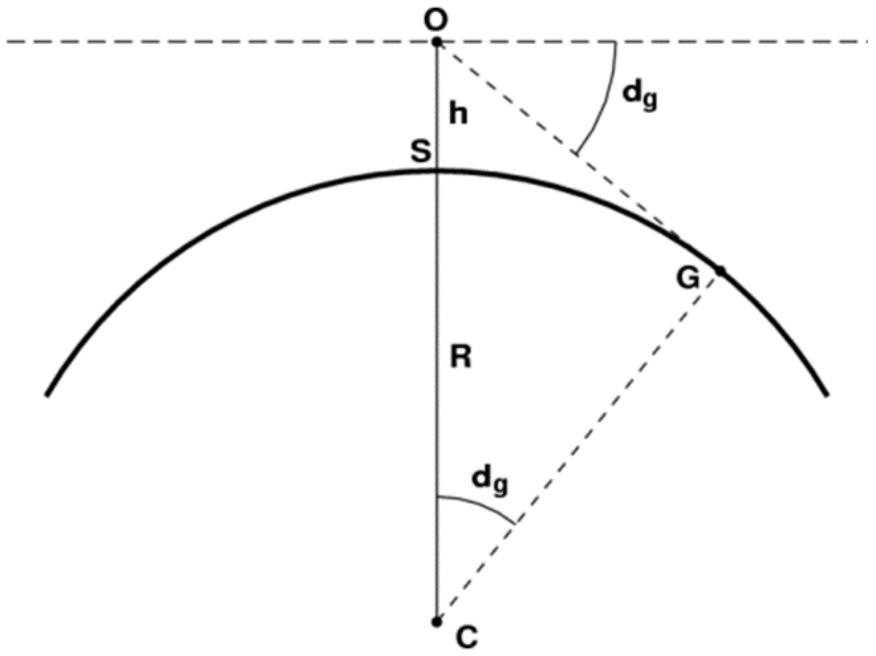


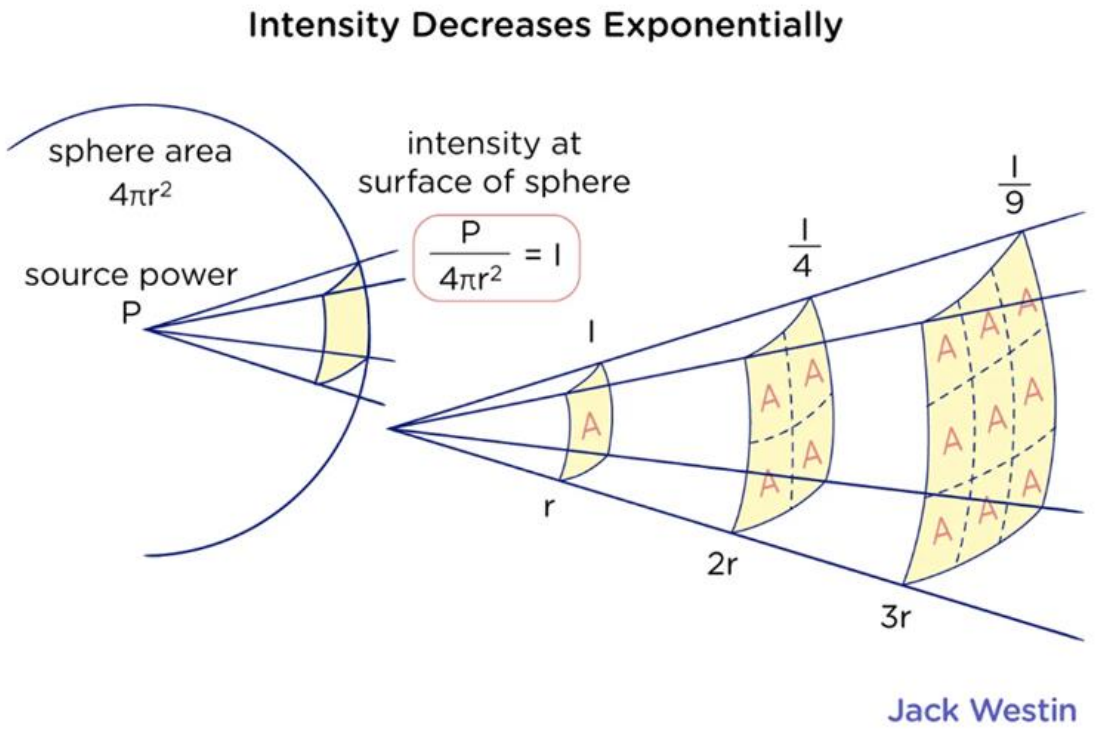
“Kyle, TX”: 2020

Where is "**Here**"? Part 2

- Sometimes this question can meaningfully be answered by referring to a coordinate or location.
- But sometimes coordinates are not appropriate or workable because:
  - Continuous nature of coordinate data can make for expensive computation.
  - Integration of different datasets in this format can be difficult.
  - Coordinates are in essence an analog format.
  - It is much easier to find patterns and meaning (defining "here") at the human scale by treating events as impacting small areas such as grid cells.

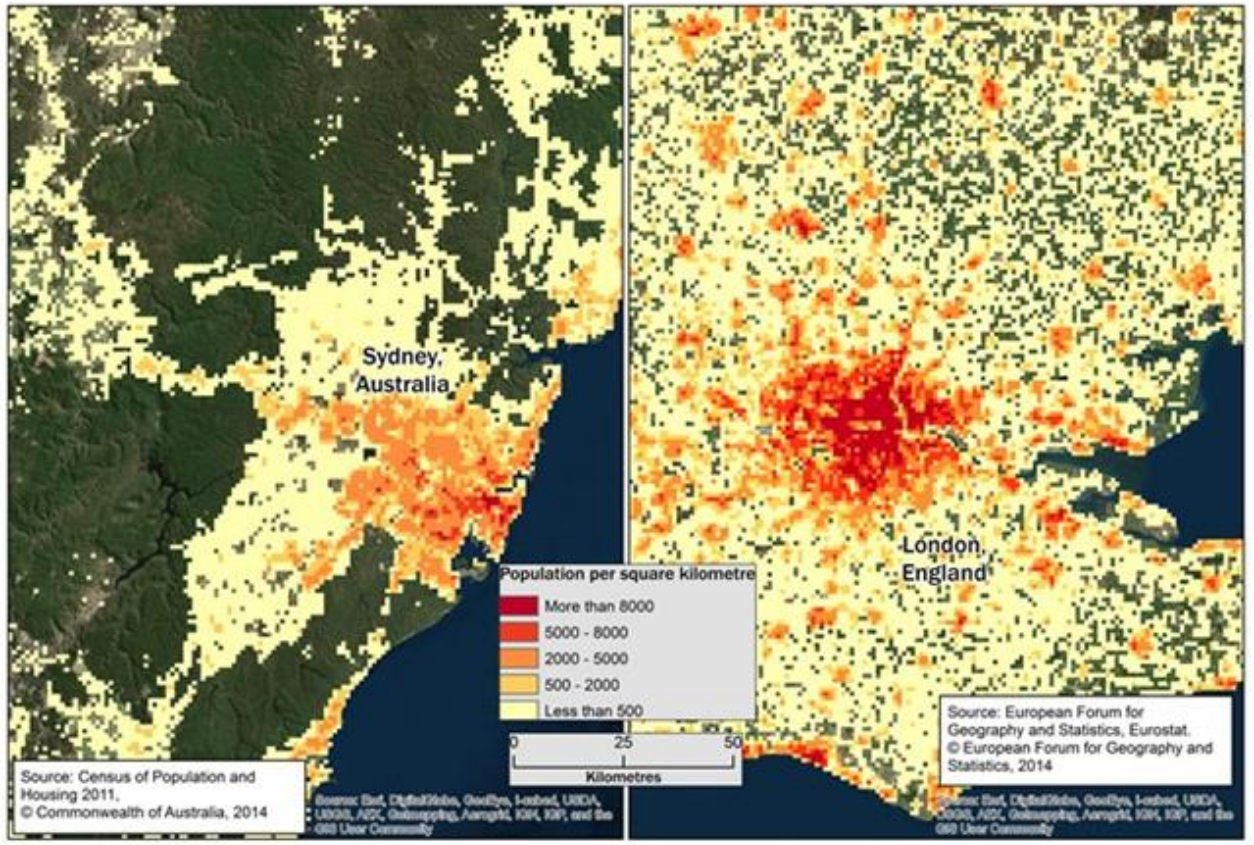
Size: 200m-5km Grids are Human Scale



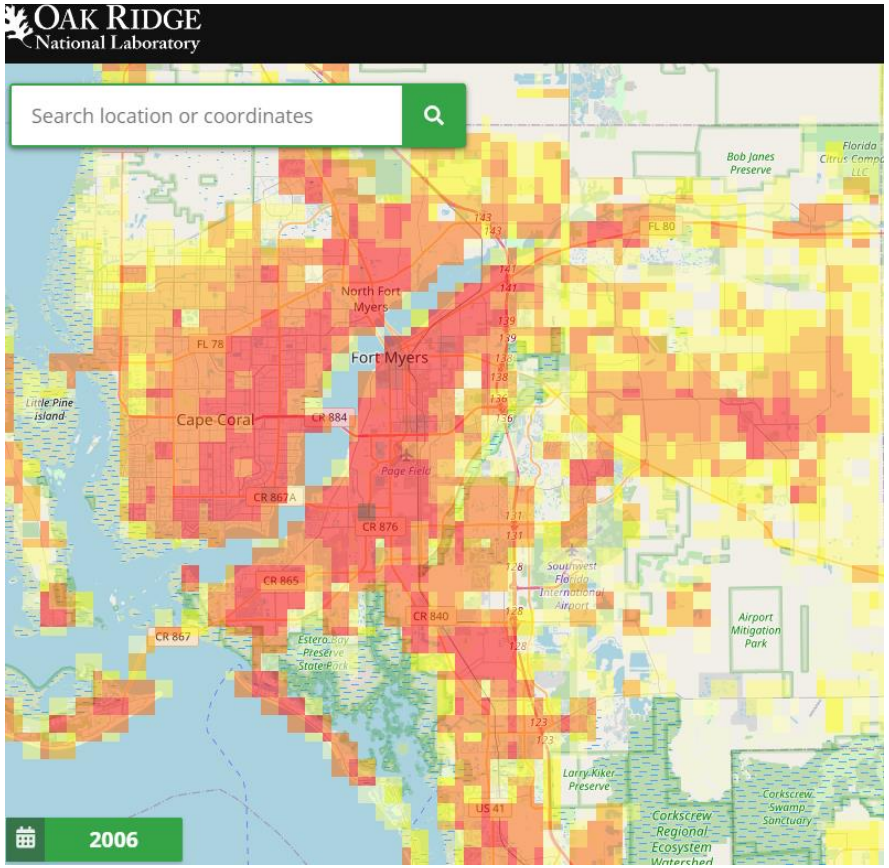


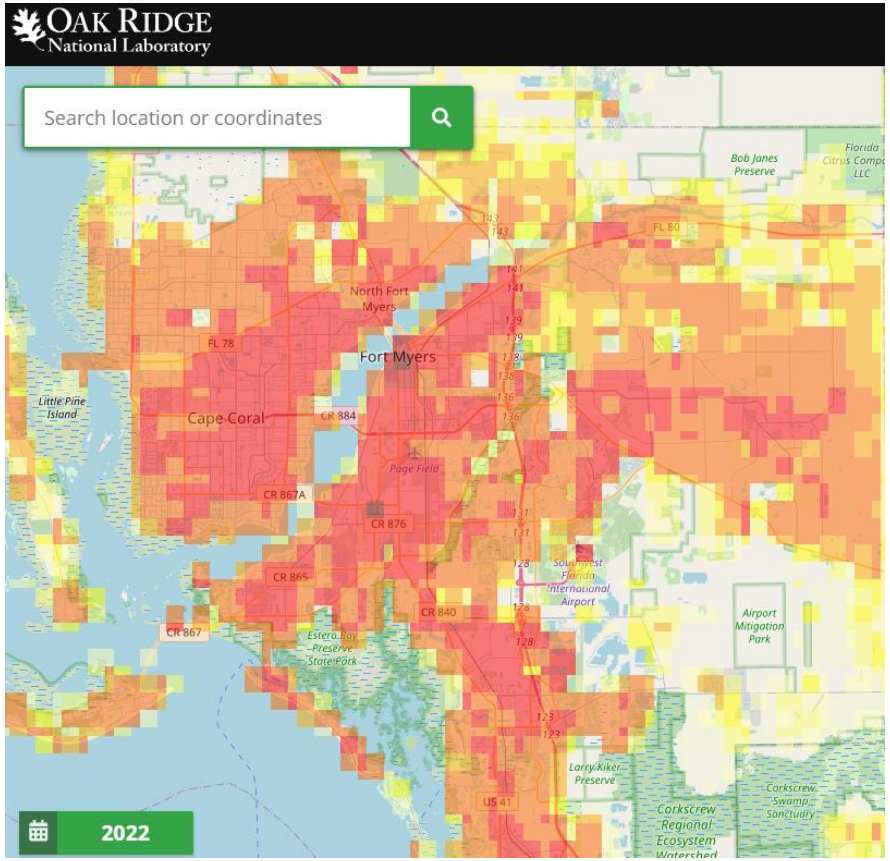
International Comparisons

Figure 5. Population Density 1km² Grid 2011 - Sydney and London



Comparisons Over Time

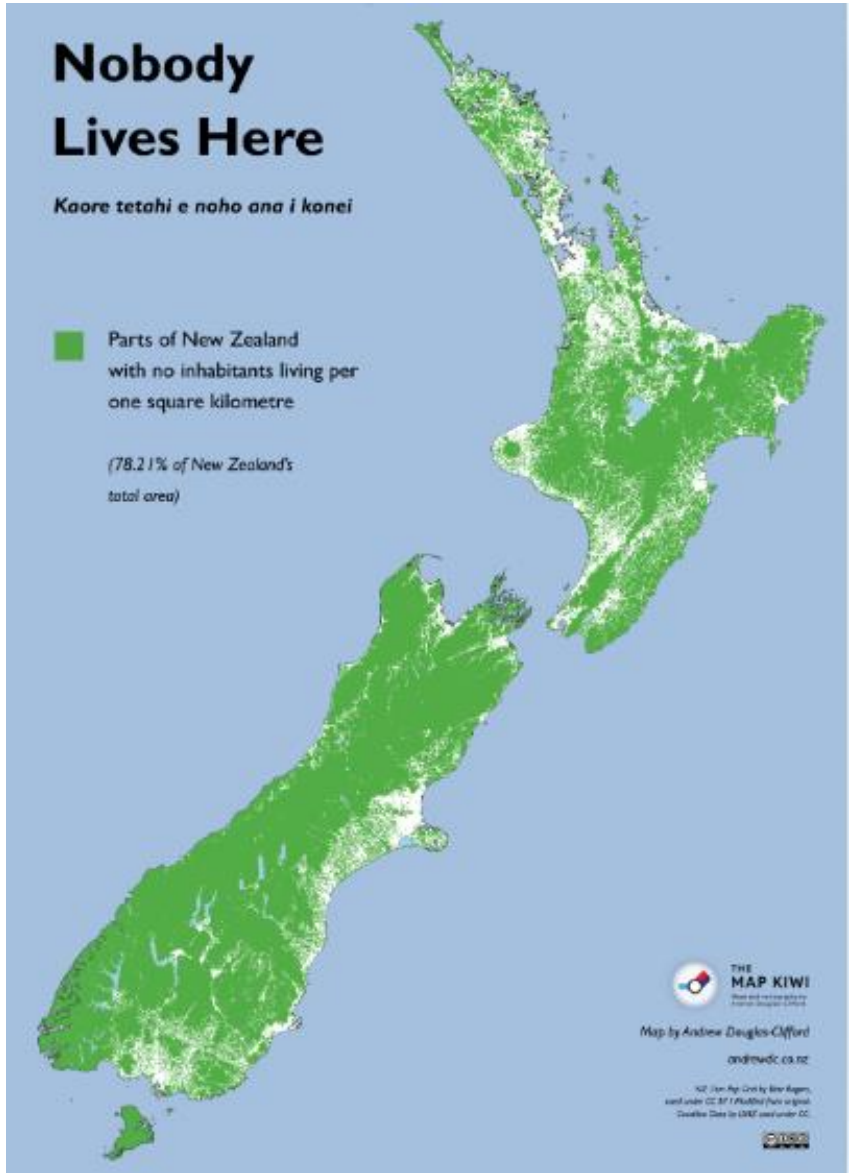




United States® Census Bureau

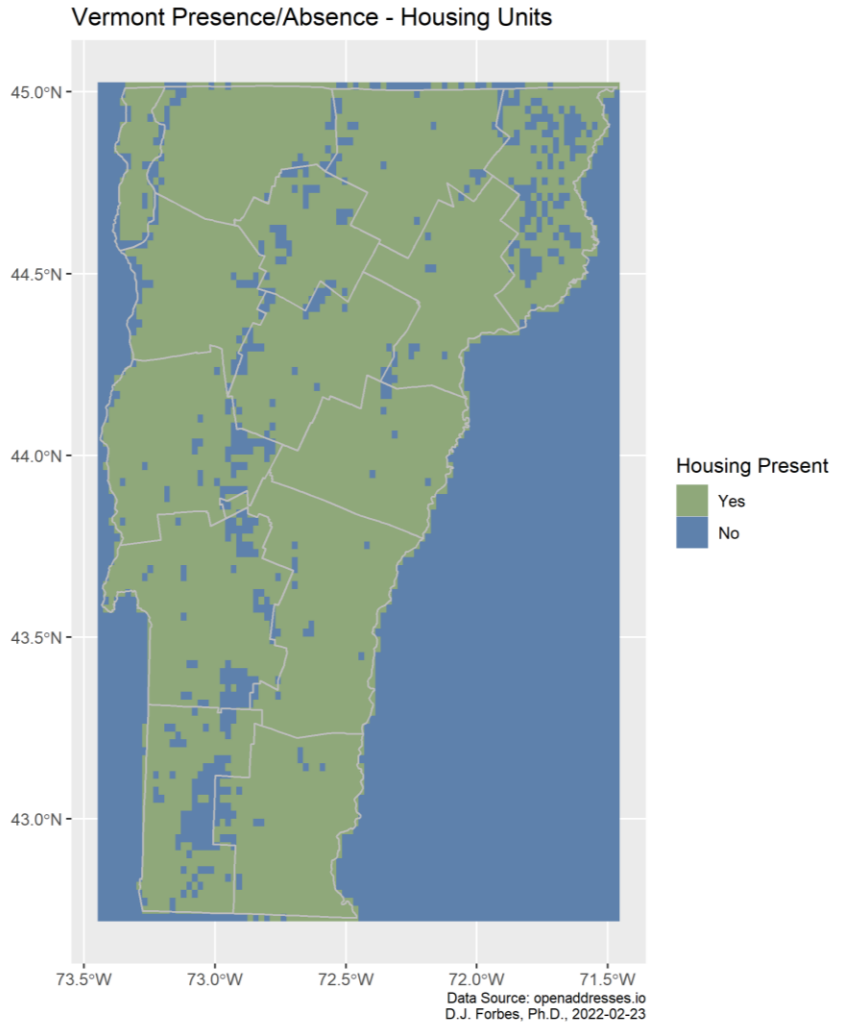
U.S. Census Bureau GEOGRAPHY

Representing Null or Zero Values



Current Work

- Engaging with the statistical community.
  - Disclosure avoidance planning.
  - Statistical content discussions.
- Integrating grids into existing production processes.
- Developing testing and quality control procedures.
  - Identify best practices for process and product metrics.
- Planning for storage and serving grids to the public.
  - Formats (geopackage, shps, raster vs. vector, etc.)
    - Have yet to prototype/test raster formats in MTDB.
  - Tiling.
  - How many versions of grids do we support?
- Staffing and resource planning.
  - How many staff are required for these activities?



Key Questions for our data partners!

- How should we communicate with YOU, our data users to solicit input?
- We want your feedback on data use-cases and requirements:
  - Statistical data (contents) -
    - What census and survey data should be disseminated?
    - Decennial Census Data? American Community Survey data? Economic data?
    - Minimal vs. maximal datasets –
      - What's does the "minimum viable dataset" look like for your needs?
    - Integration of other datasets – what would you "mashup" with census data?
  - Geospatial formats and preferences (container) -
    - Raster vs Vector.
    - Shape and Size(s) of Grid Cells.
    - Equal area vs equal dimensions.

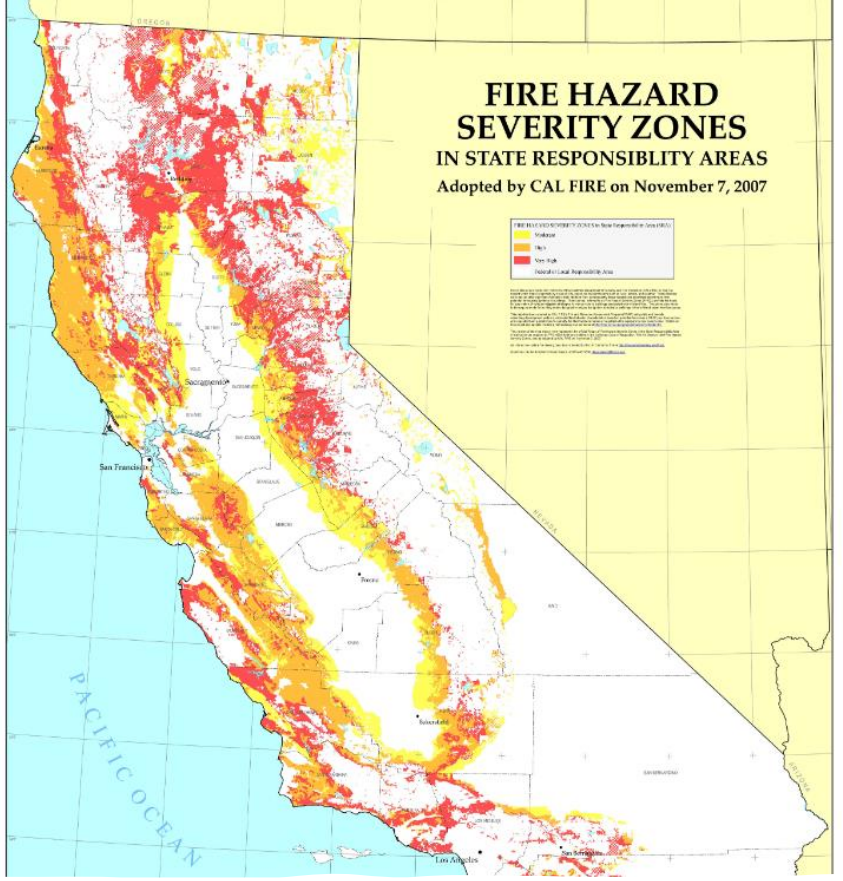
United States® Census Bureau

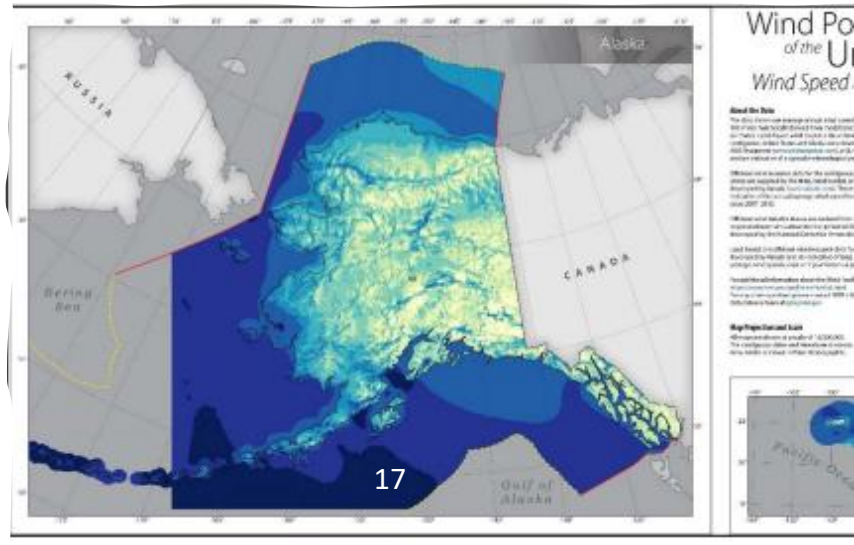
U.S. Census Bureau GEOGRAPHY

Discussion

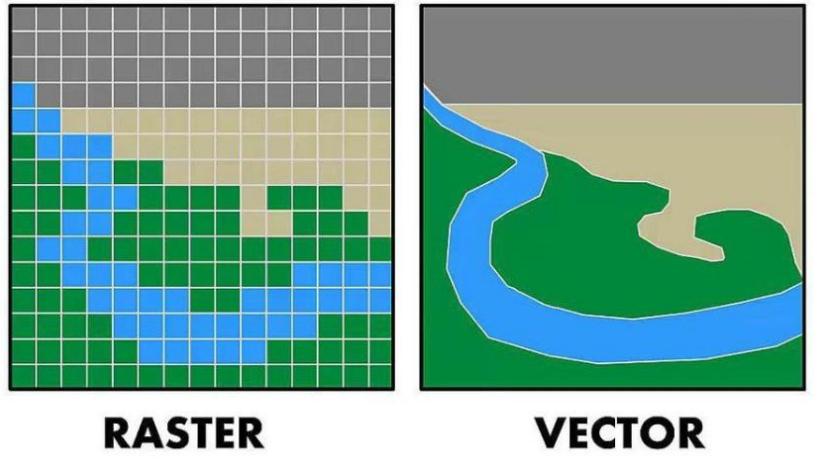
Current grids

- Do agencies in your state use gridded data?
- If yes, what kind of gridded data?
  - Earth Observation/Remotely Sensed/Satellite Imagery
  - Elevation data
  - Land Use/Land Cover data
  - Transportation/Traffic
  - Population/Housing/Demographics

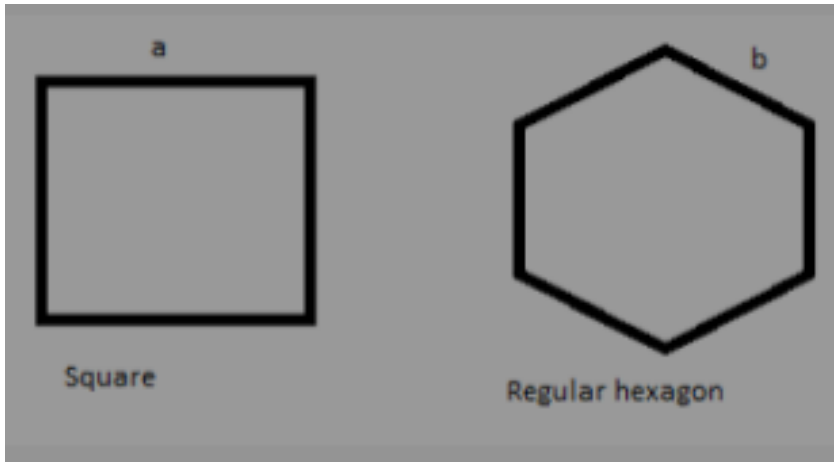




Grids – geospatial

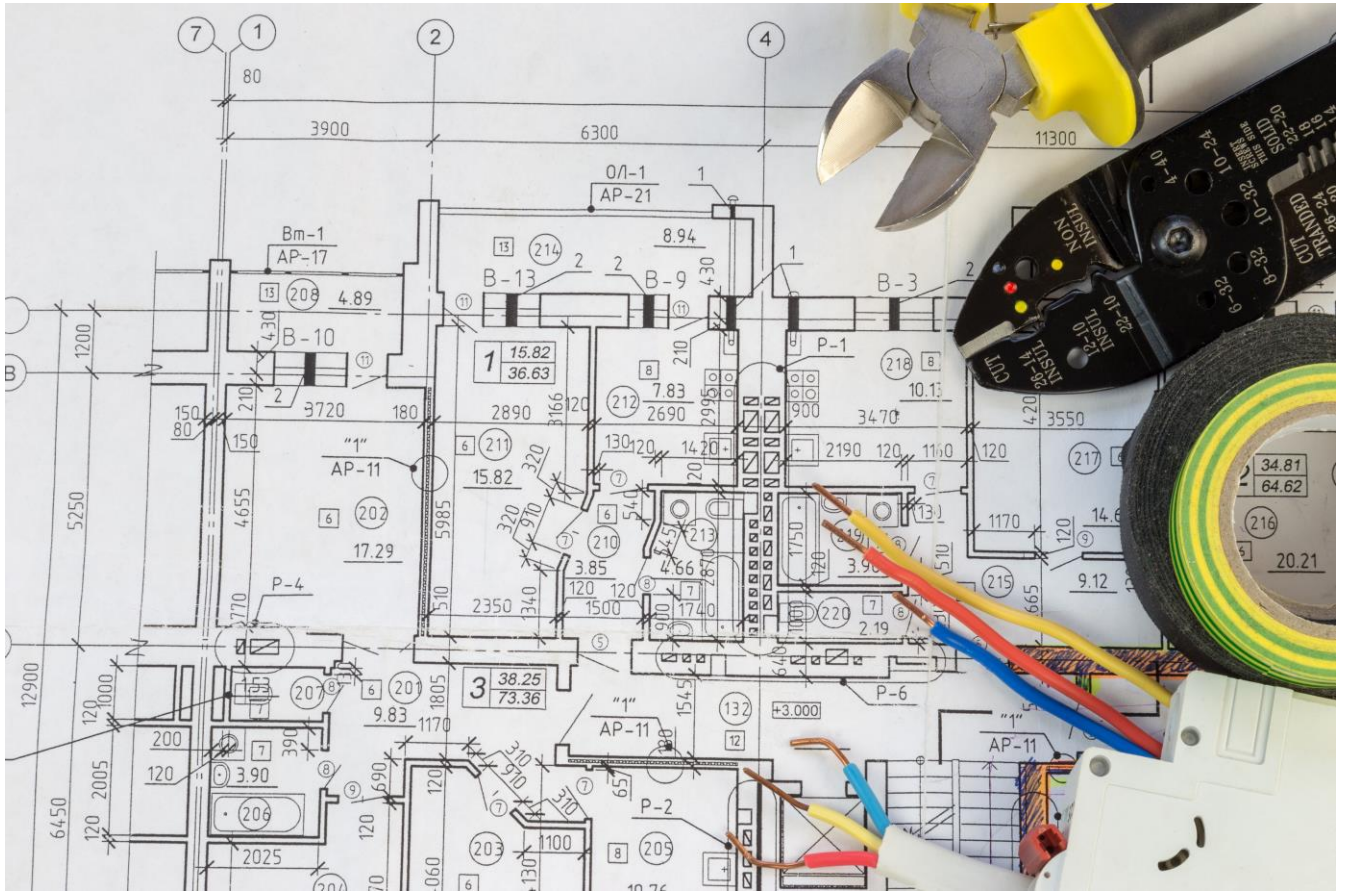


- For those that use grids, what type of grid are you using?
  - Raster data?
  - Shape
    - Rectangle/quadrilateral? Hexagonal?
  - Size?
    - 1km? 5km?
  - Existing standard (e.g. DGGS or National Grid) or self-generated grid?



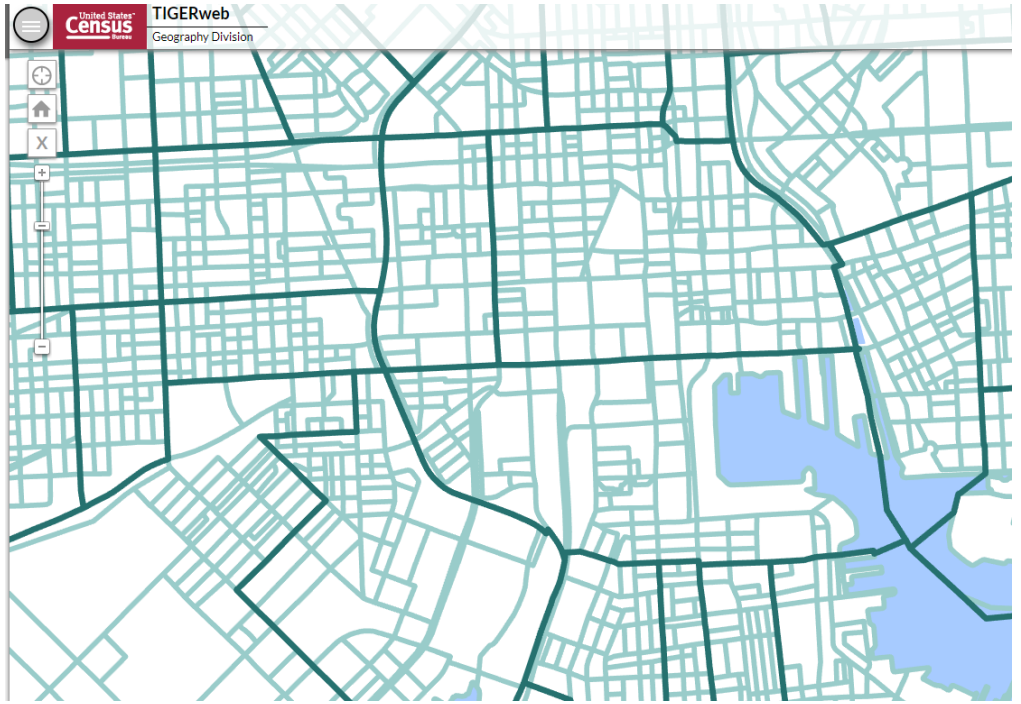
Grids – use-cases

- If we produced grids, how would you use them?
- What use cases do you have for grids?
- What statistical data might be useful in a grid?



Grids – comparative geographies

- What is the smallest geography you frequently use?
  - Tracts? Block groups? Blocks? Other?
  - What statistical data do you use with these geographies?
    - Demographic/Housing characteristics or counts.
    - Business, Jobs, Employment characteristics or counts.
    - Socio-economic characteristics or counts.
    - Other?
  - What drawbacks or limitations do you find with these geographic units?



United States® Census Bureau

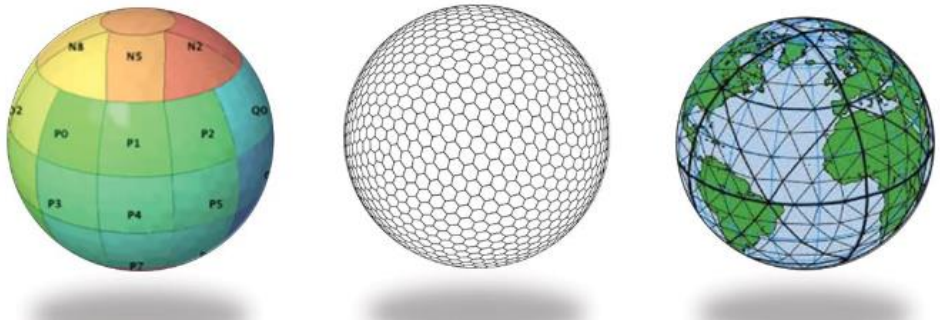
U.S. Census Bureau GEOGRAPHY

Final Discussion

- What final thoughts do you have about grids?
- What are your gridded data needs?
- What requirements would you like to communicate to us?

- Please share any final thoughts with us at **geo.grids@census.gov**

DGGS Advantages



- Each cell has unique address (vs continuous coordinate systems).
- Greatly simplifies the definition of “here” in studies.
- Data of any type can be held in these cells.
- Integration of data is on-demand and virtually free (computationally).
- Cells are hierarchically organized allowing for rapid aggregation and decomposition.
- Works at every latitude including polar regions.
- OGC DGGS suite is leading candidate for adoption as global standard.