

EXPLOROPLETH

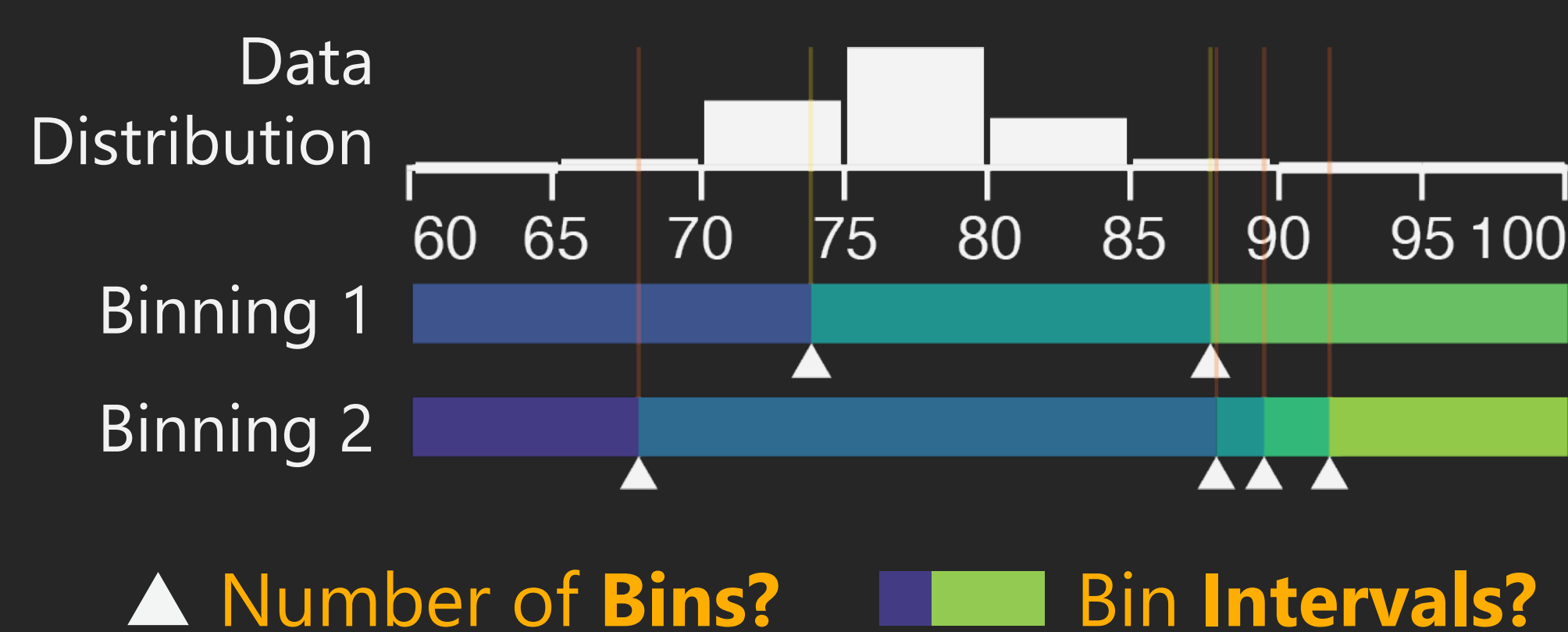
A Tour through the Data Binning Zoo for Choropleth Maps

1 Imagine you have to create a map of **Life Expectancy** across U.S. counties...

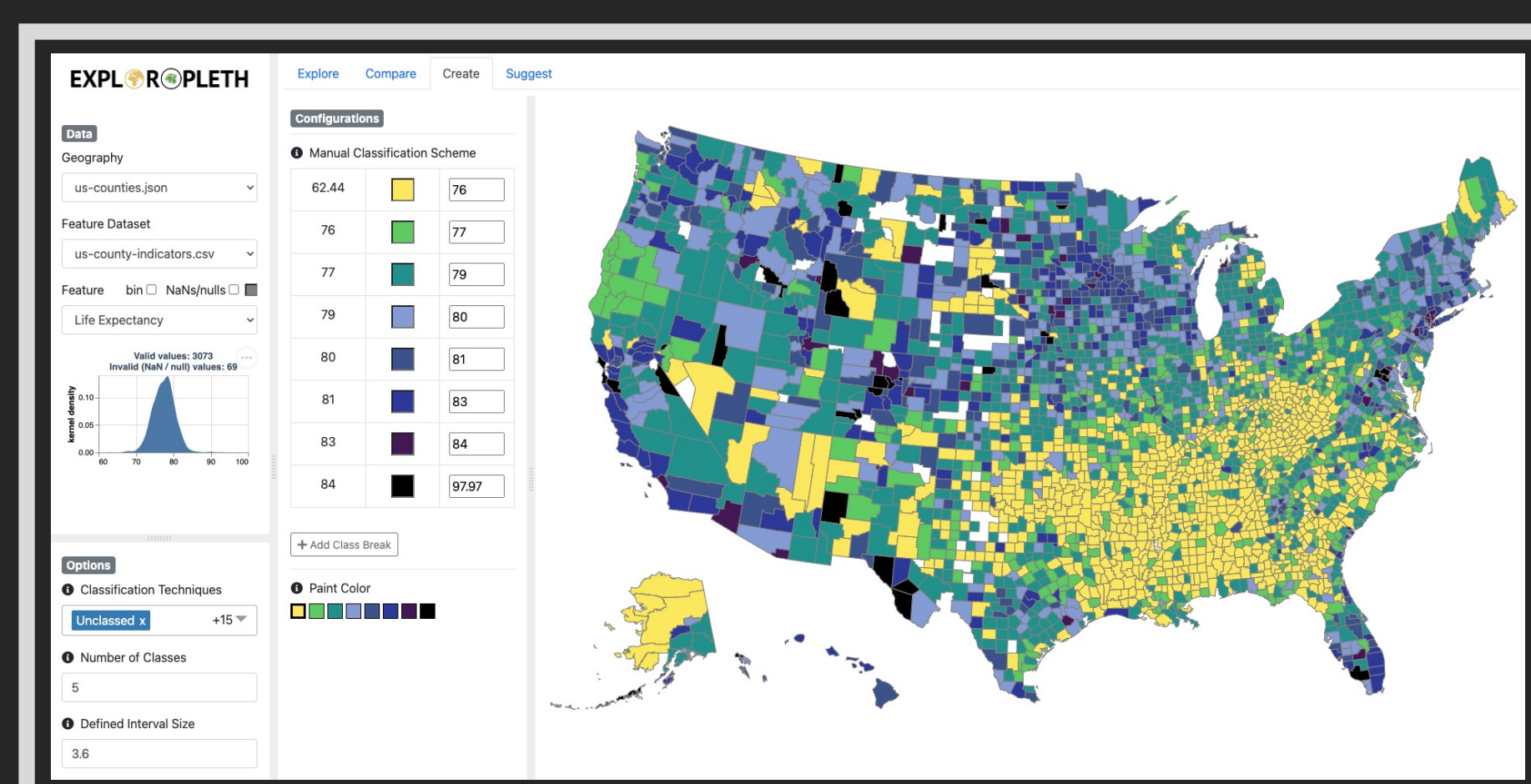
County	State	Life Expectancy
Oglala Lakota	SD	62.44 years
Los Angeles	CA	82.21 years
San Miguel	CO	97.97 years

Source: CDC, USA (2019)

2 How do you **bin / classify** the data?



3 We created **EXPLOROPLETH**



- A Explore 12+ binning methods.
- B Compare bin counts, sizes, intervals.
- C Create "custom" bins.
- D Suggest a "good" method.

Same Data Binned Differently **CHECK IT OUT**

Georgia Tech Visualization Lab



Arpit Narechania
anarechania3@gatech.edu

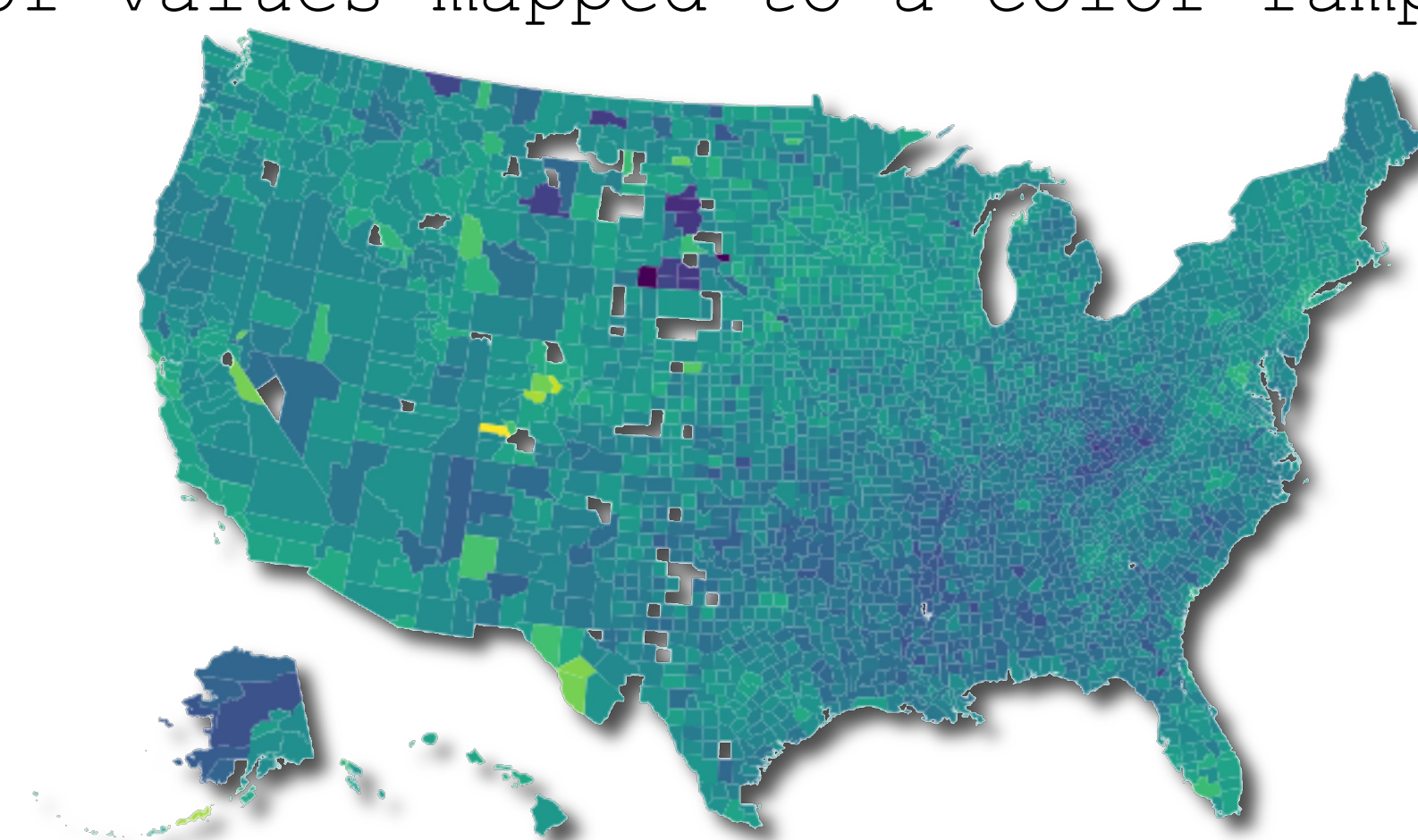
Alex Endert
endert@gatech.edu

Clio Andris
clio@gatech.edu

A **Explore** Life Expectancy in U.S. counties (62.44 – 97.97 years)

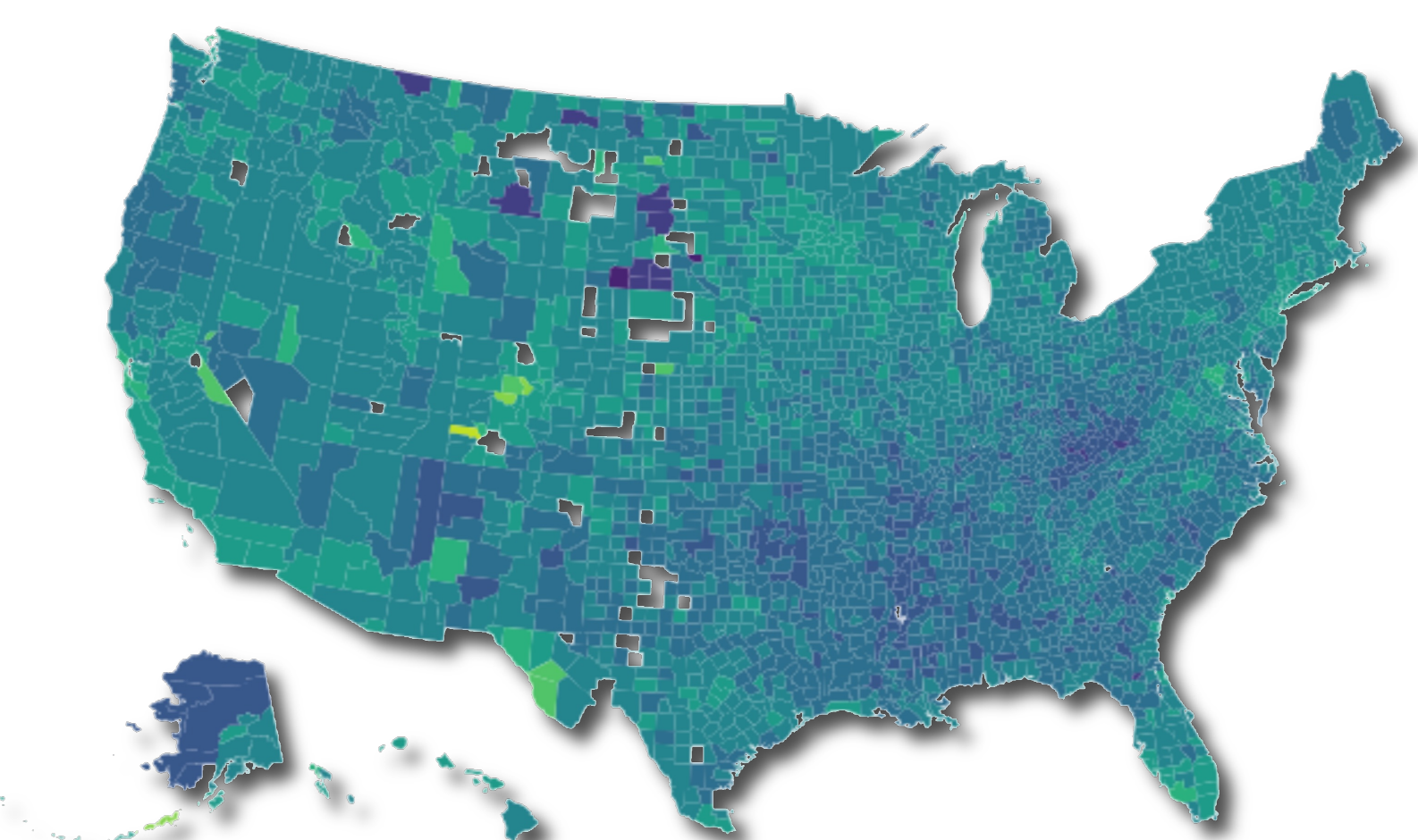
Continuous

No discrete bins; continuous range of values mapped to a color ramp.



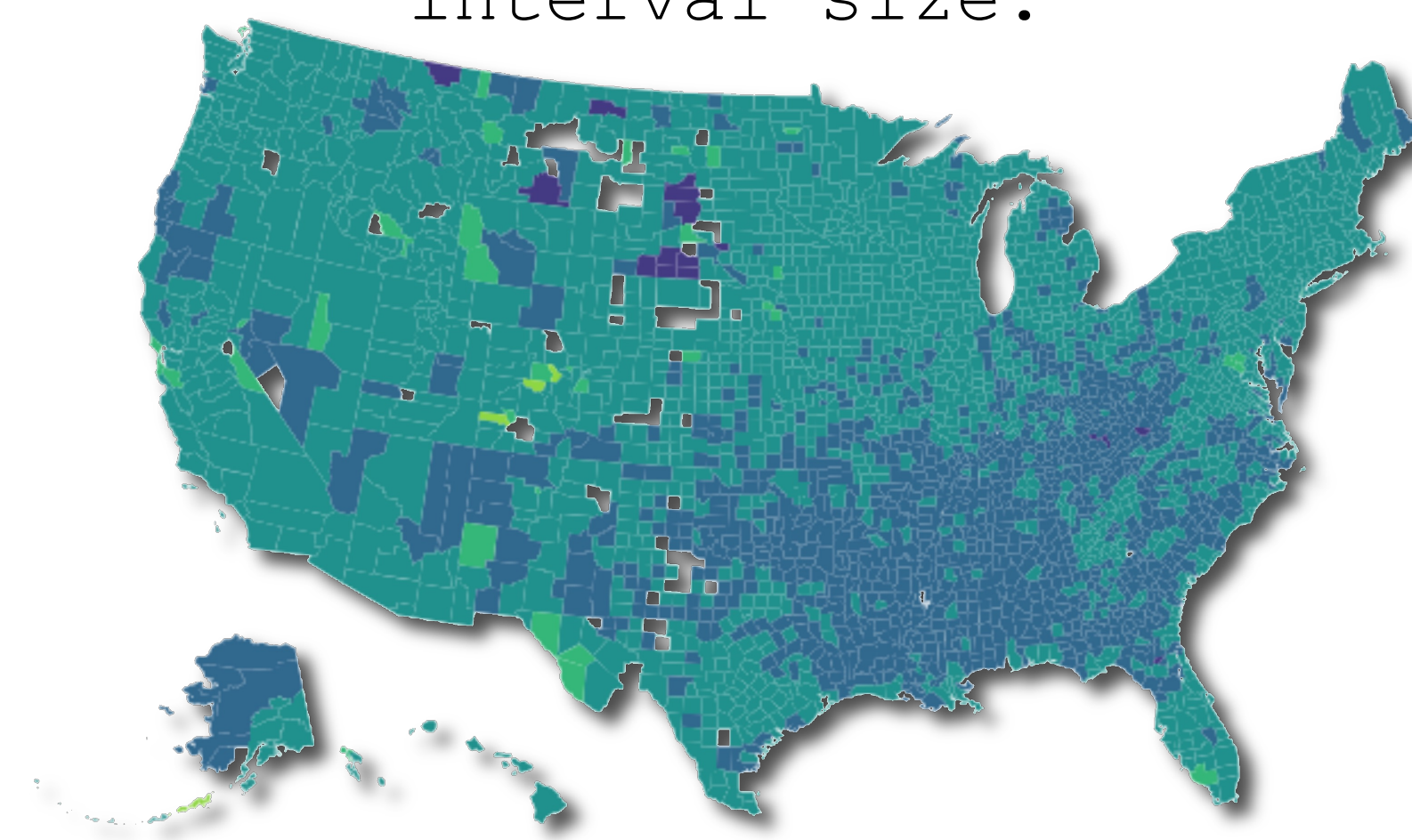
Defined Interval

Each bin is dependent on the specified interval size.



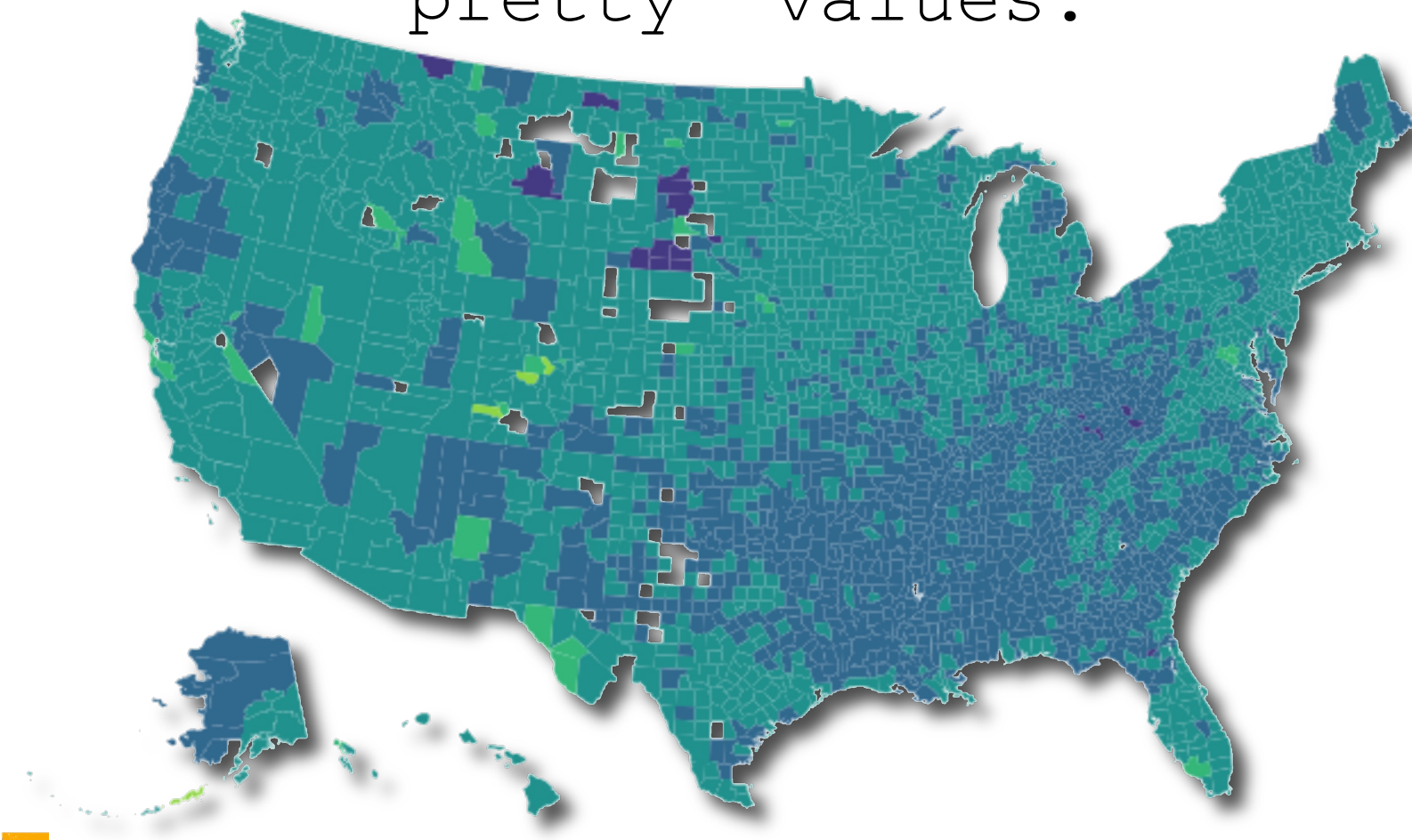
Equal Interval

Each bin has the same interval size.



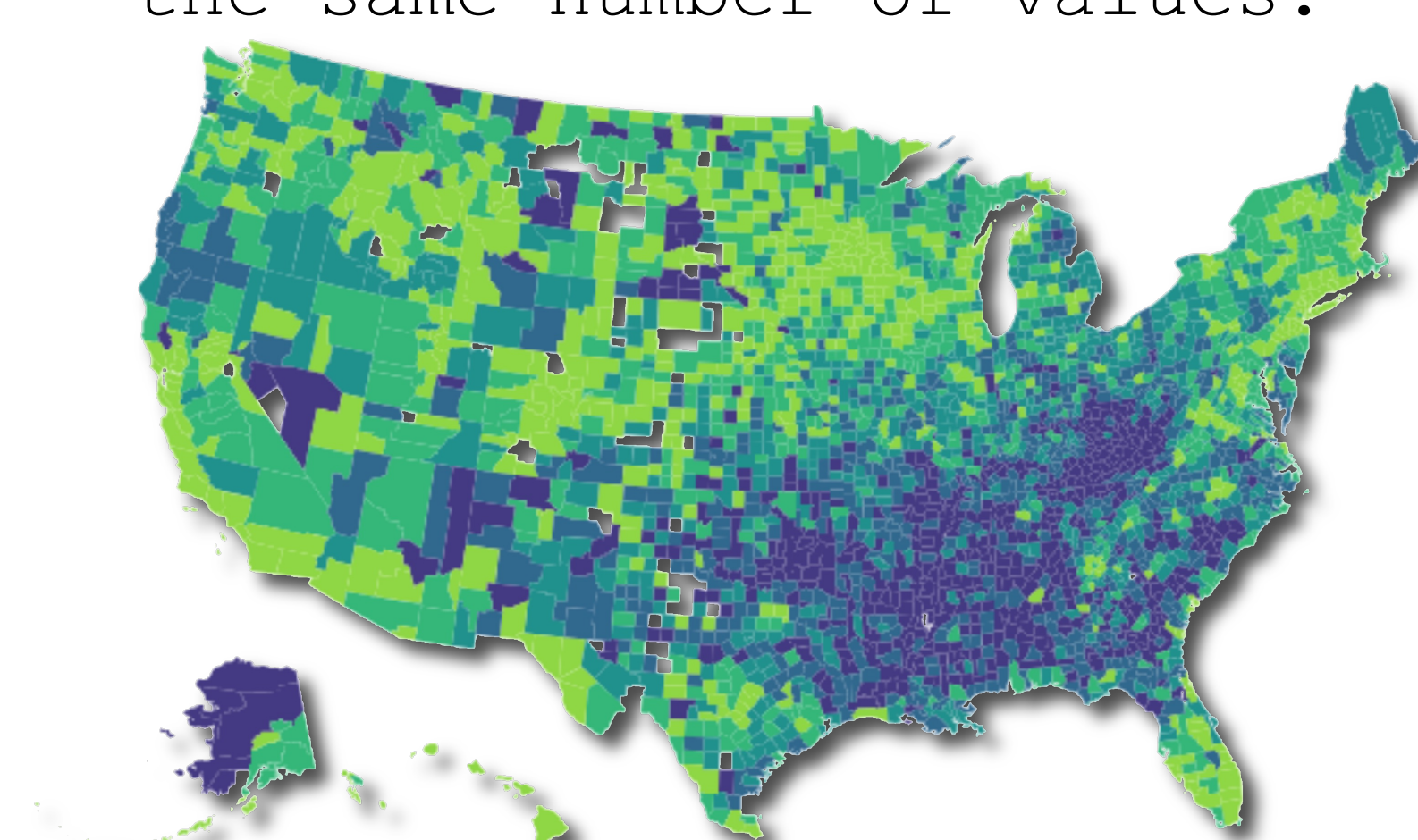
Pretty Breaks

Each bin break is rounded off into "pretty" values.



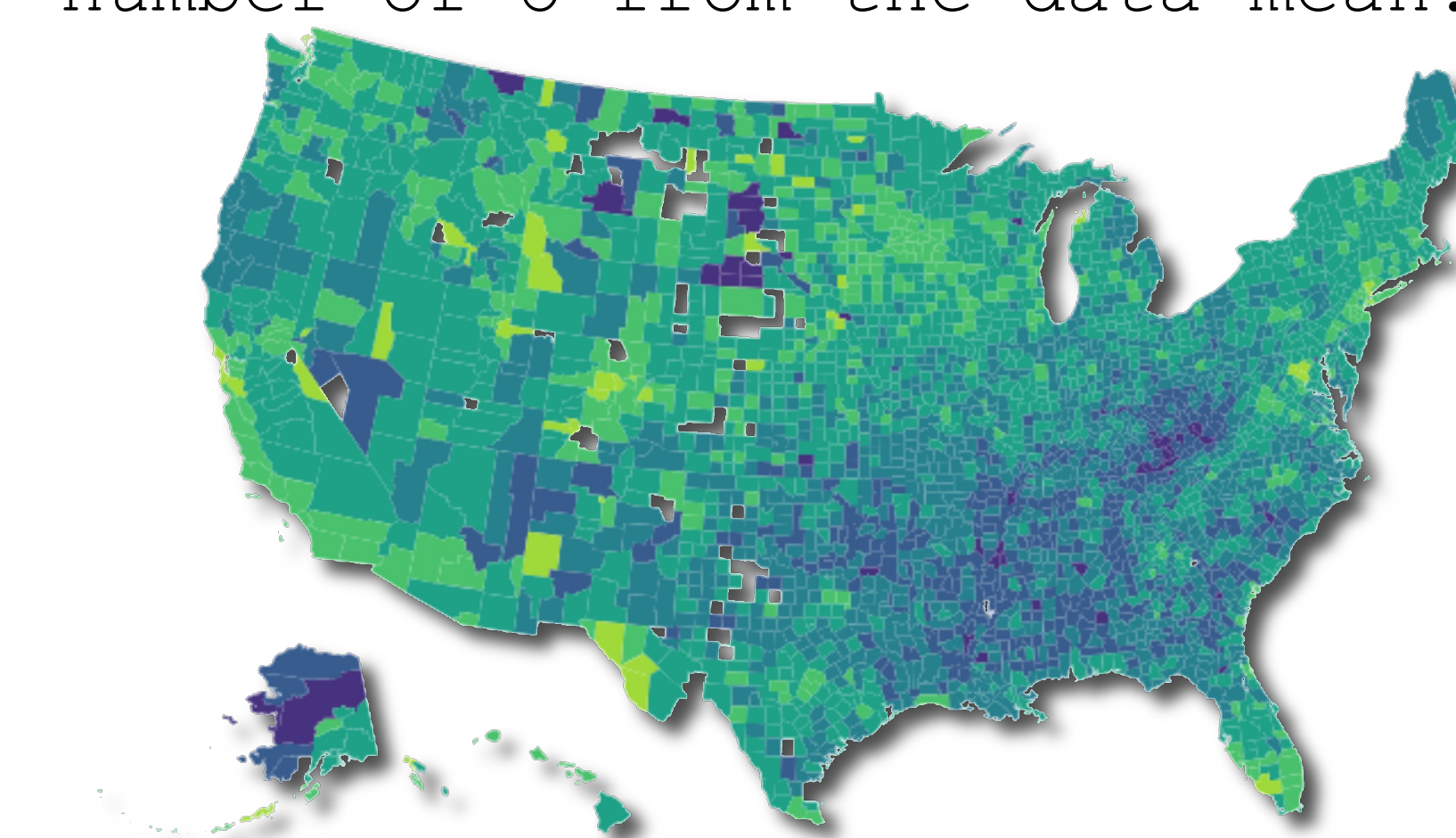
Quantile

Each bin has approximately the same number of values.



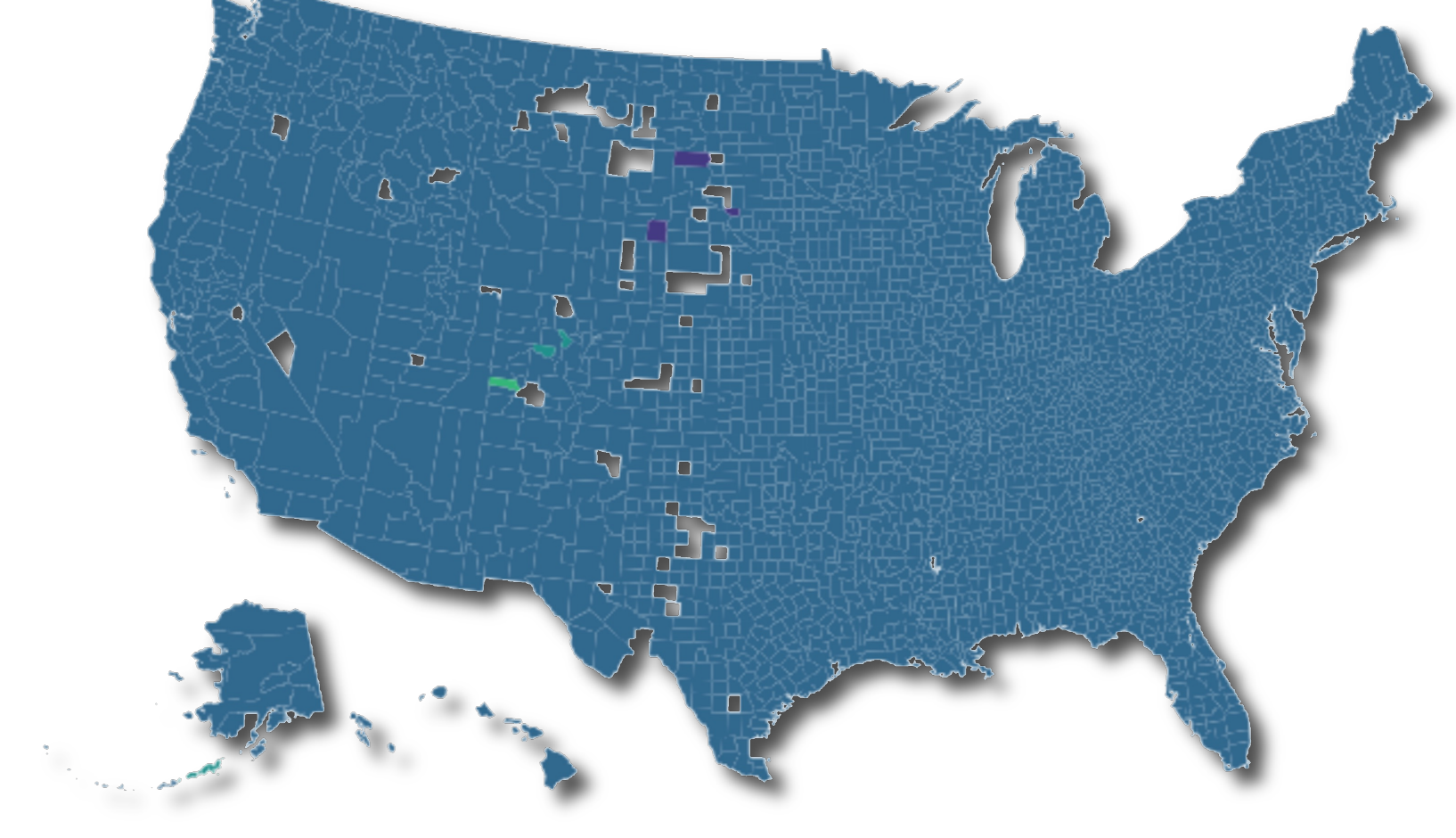
Standard Deviation (σ)

Each bin contains data within a certain number of σ from the data mean.



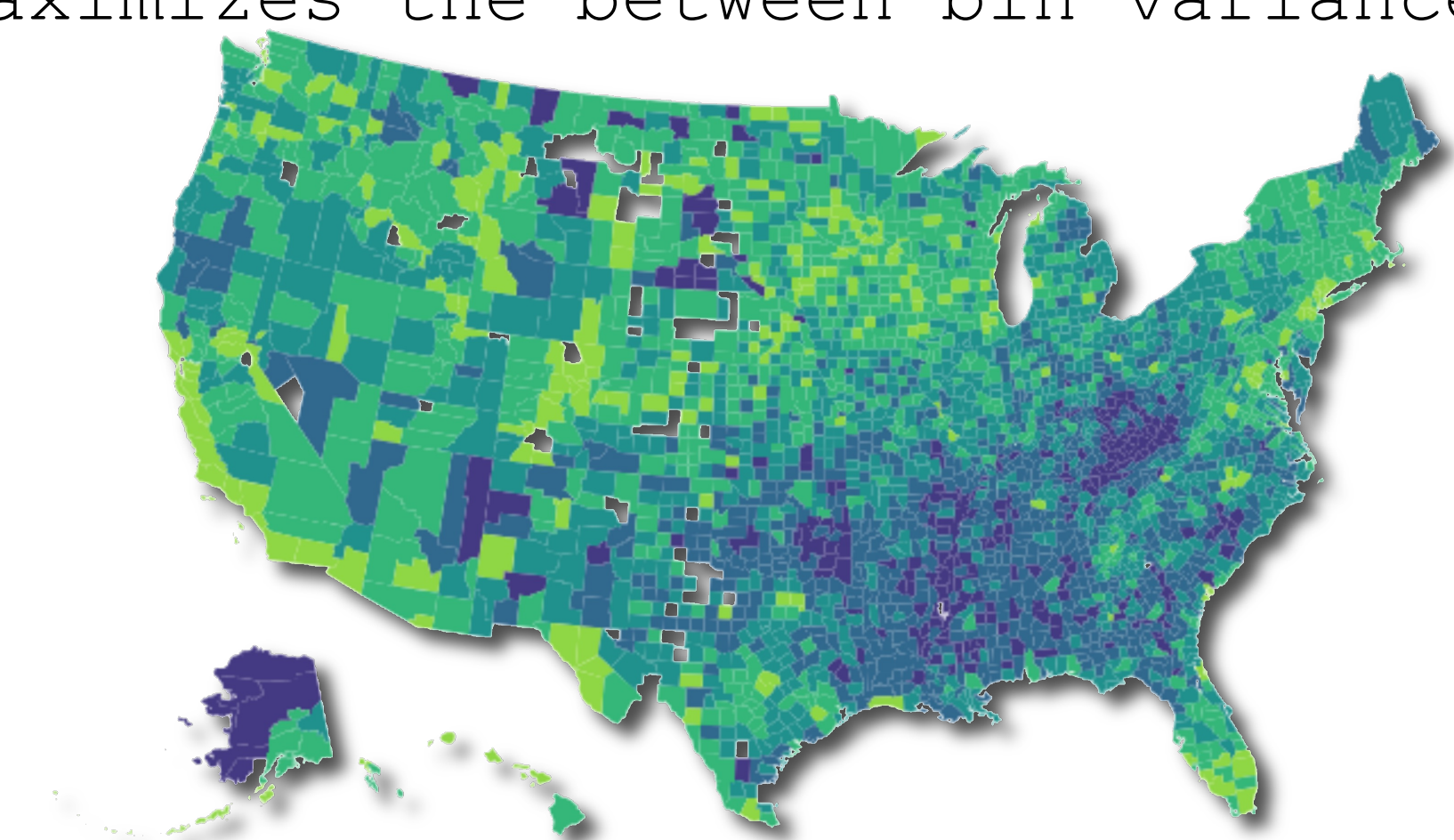
Maximum Breaks

Bins are based on maximum differences between successive values.



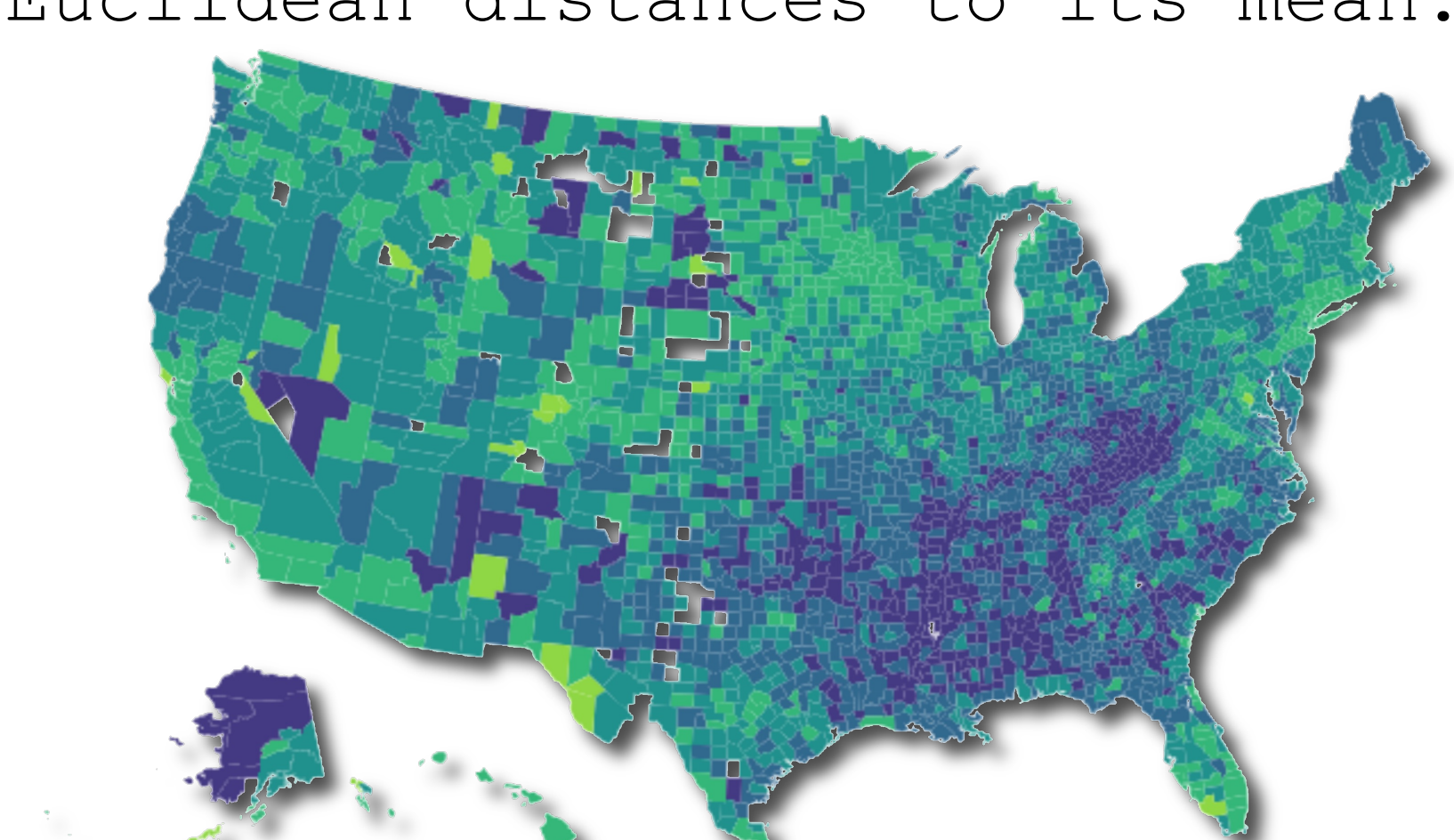
Natural Breaks

Each bin minimizes the within-bin and maximizes the between-bin variance.



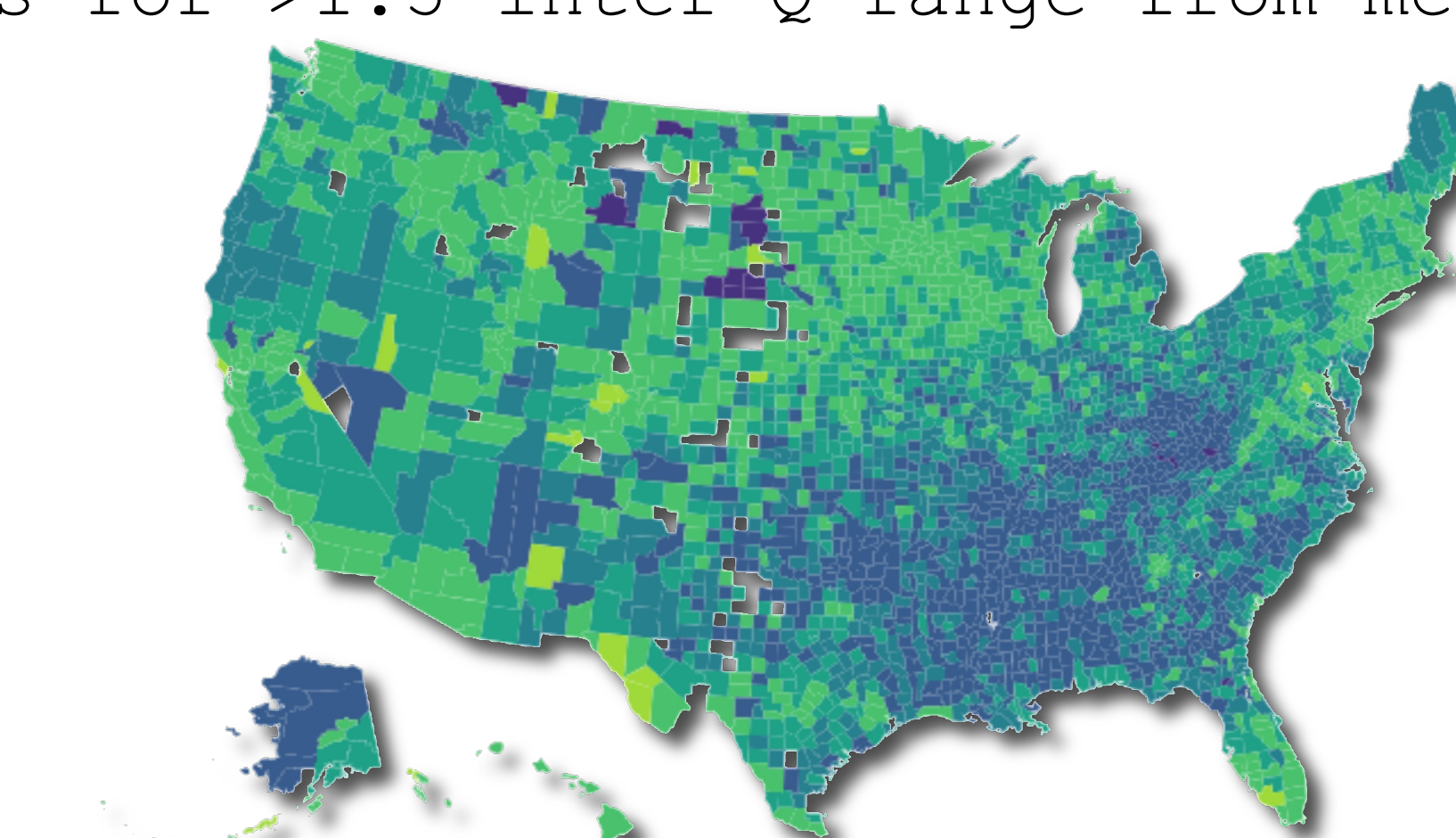
CK-Means

Each bin minimizes the sum of squared Euclidean distances to its mean.

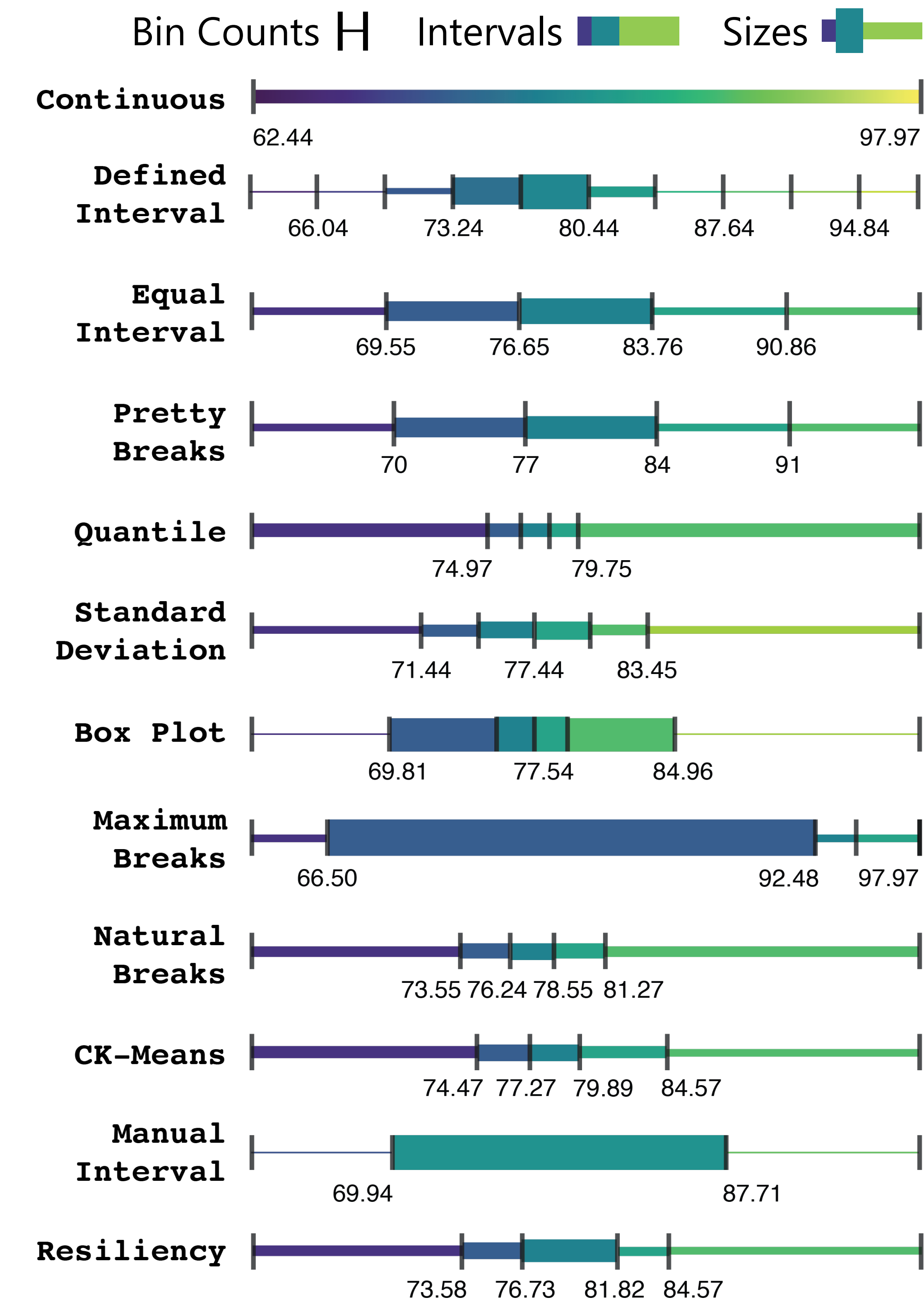


Box-Plot

Six bins: four quartiles (Q) and two bins for >1.5 inter-Q range from median.



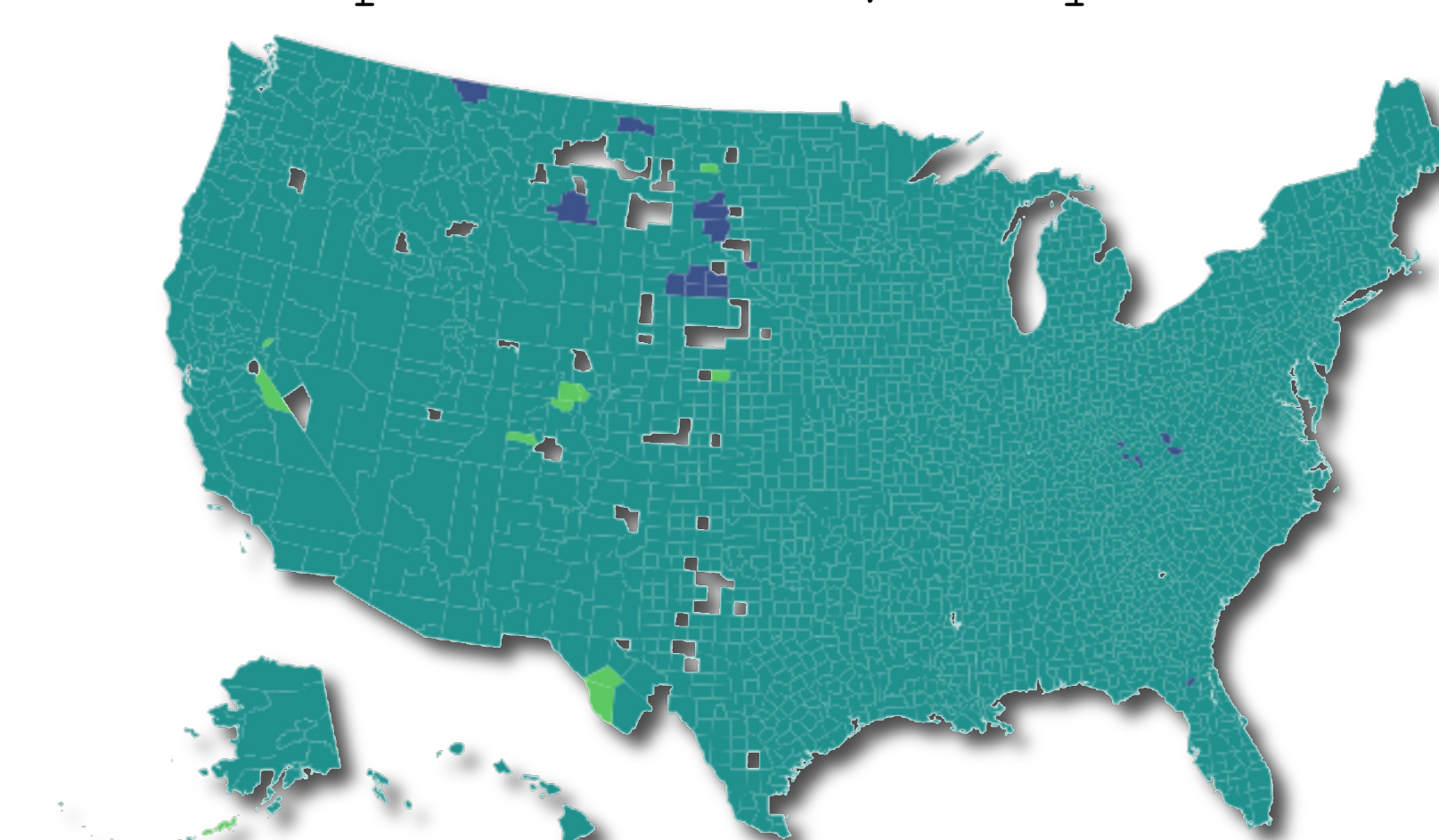
B **Compare** binning methods



C **Create** "custom" bins

Manual Interval

Per user preferences, requirements.



D **Suggest** a "good" method

Resiliency

Regions consistently fall in the same bin across binning methods.

