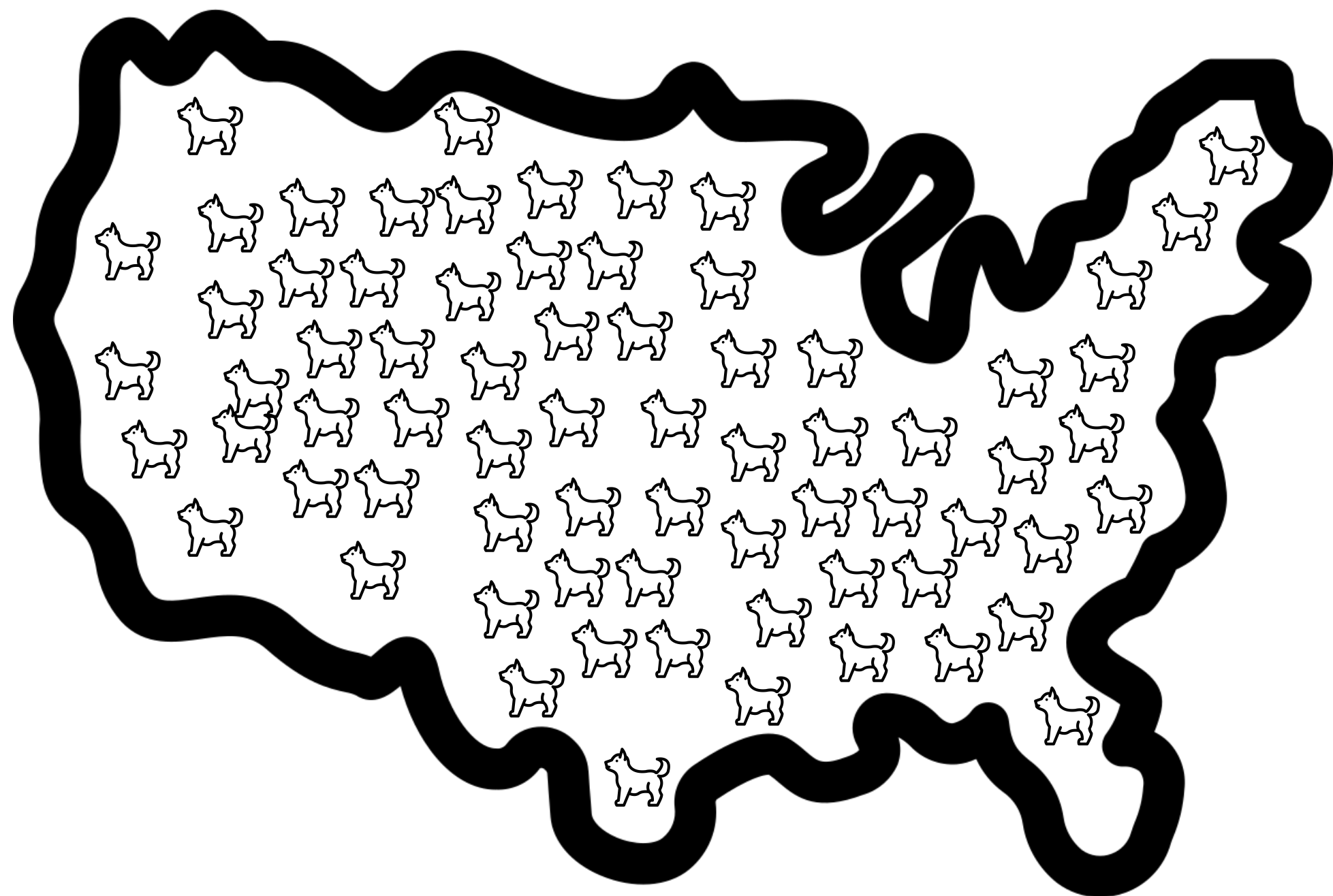


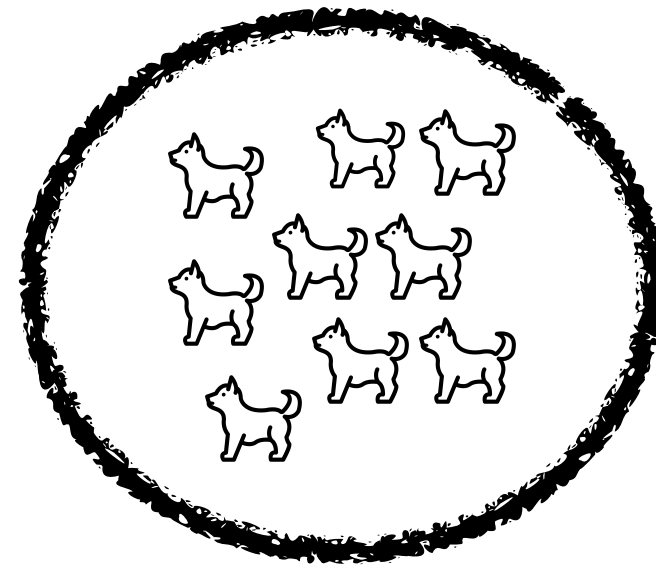
Describing Data

Manoel Horta Ribeiro
manoel@cs.princeton.edu

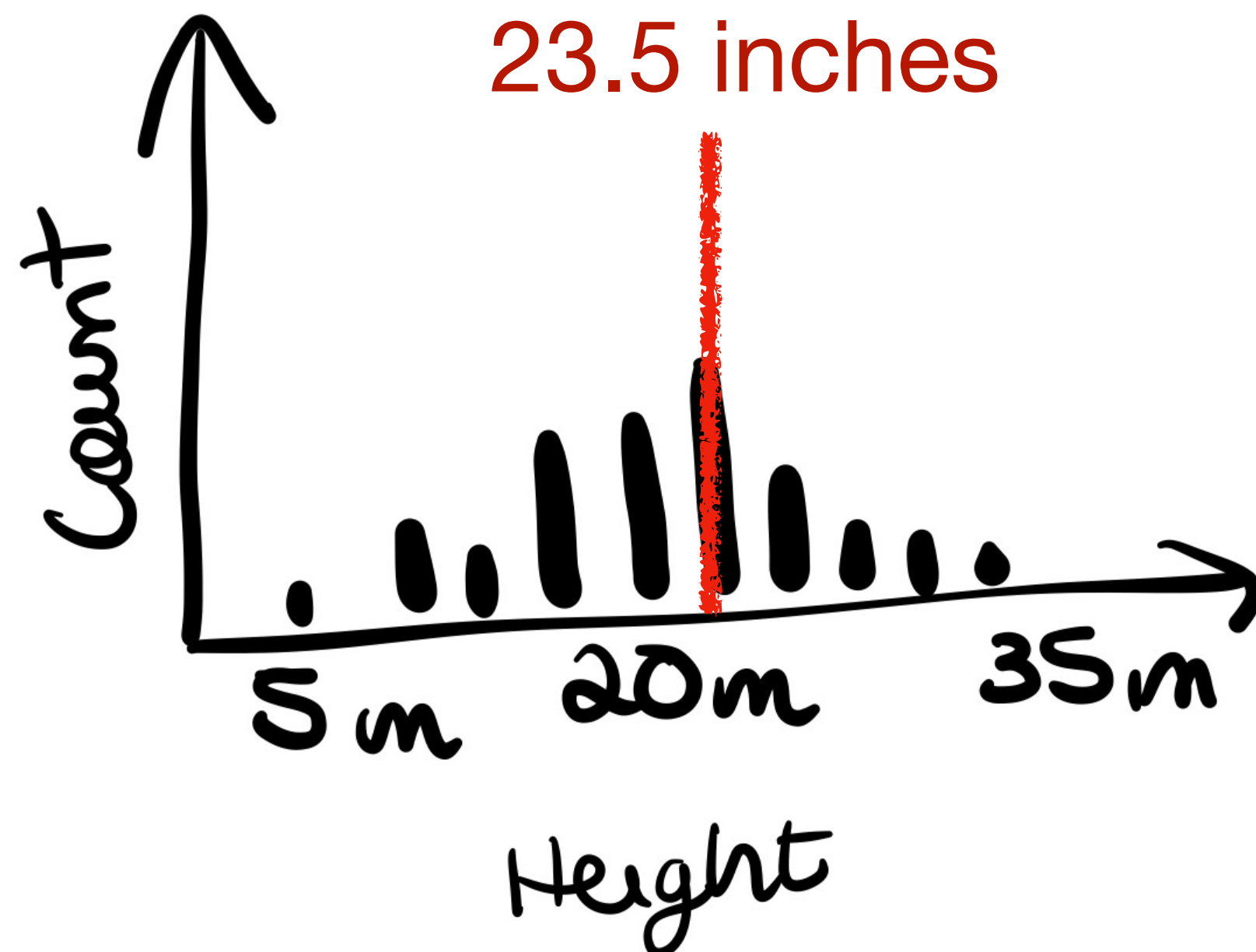
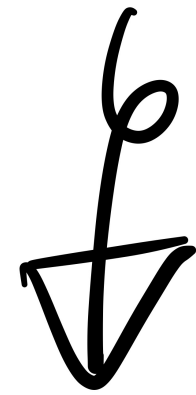
COS 424 / Fall 2026



Descriptive Statistics



100 dogs



Summarize the quantitative features of the data you have!

23.5 inches is the average height of dogs in our sample

Measures of Central Tendencies

- Mean

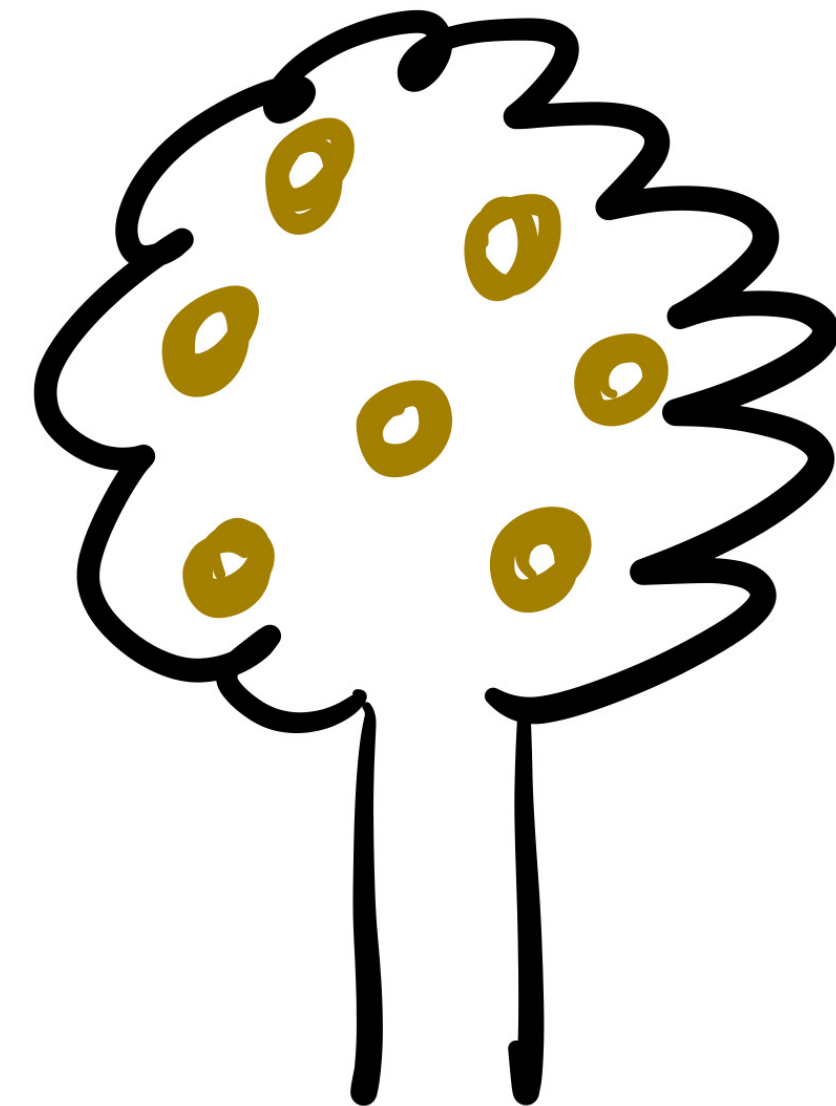
- *Arithmetic:* $(x_1 + x_2 + \dots + x_n)/n$
- *Geometric:* $\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$

- Median

- *The point splitting the sample in half*

- Mode

- *The most common value in the sample*



Jan: 100 apples
Feb: 200 **+100%**
Mar: 220 **+10%**
Apr: 440 **+100%**

Avg (Arithmetic) Growth: **70%**

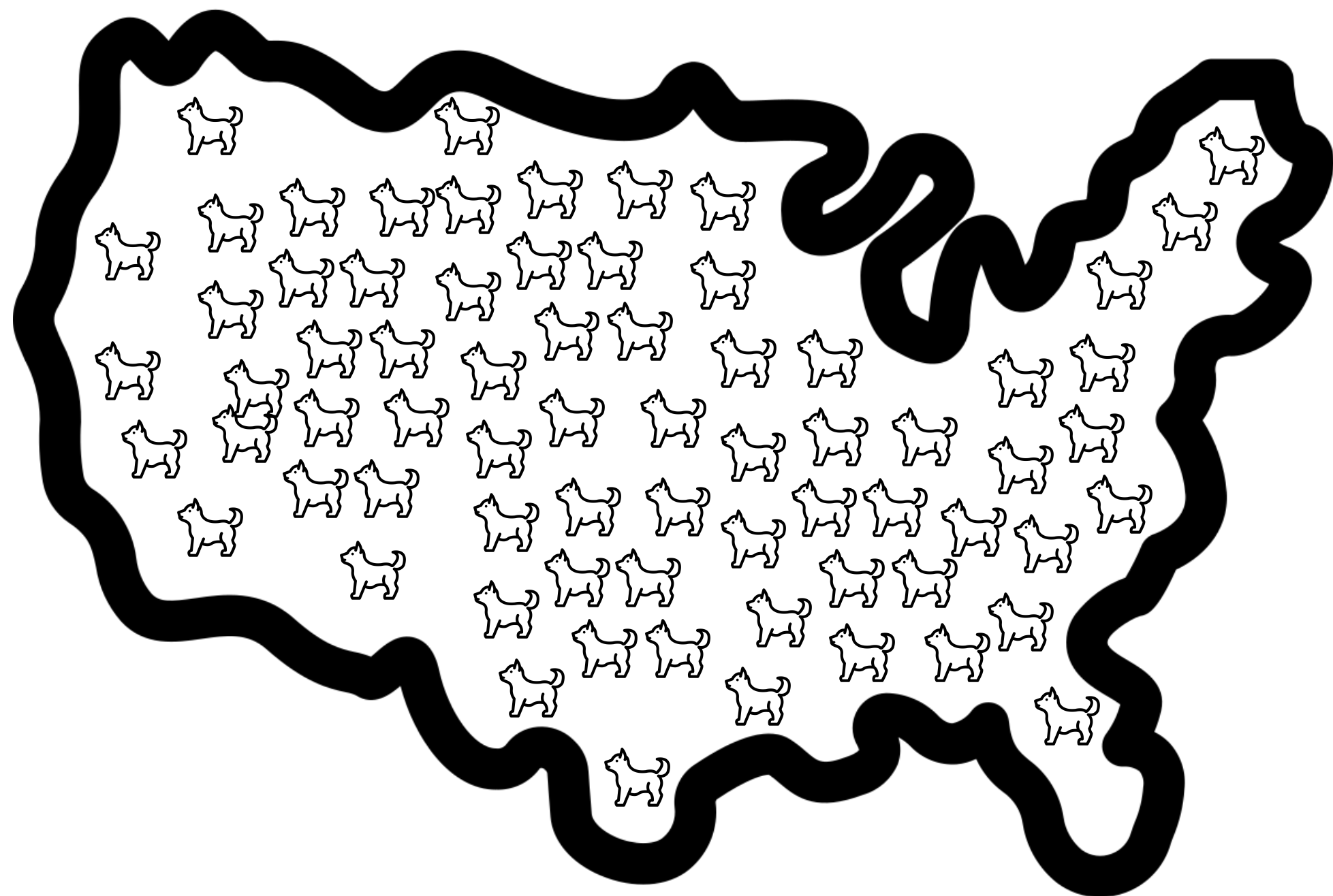
Avg (Geometric) Growth: **64%**

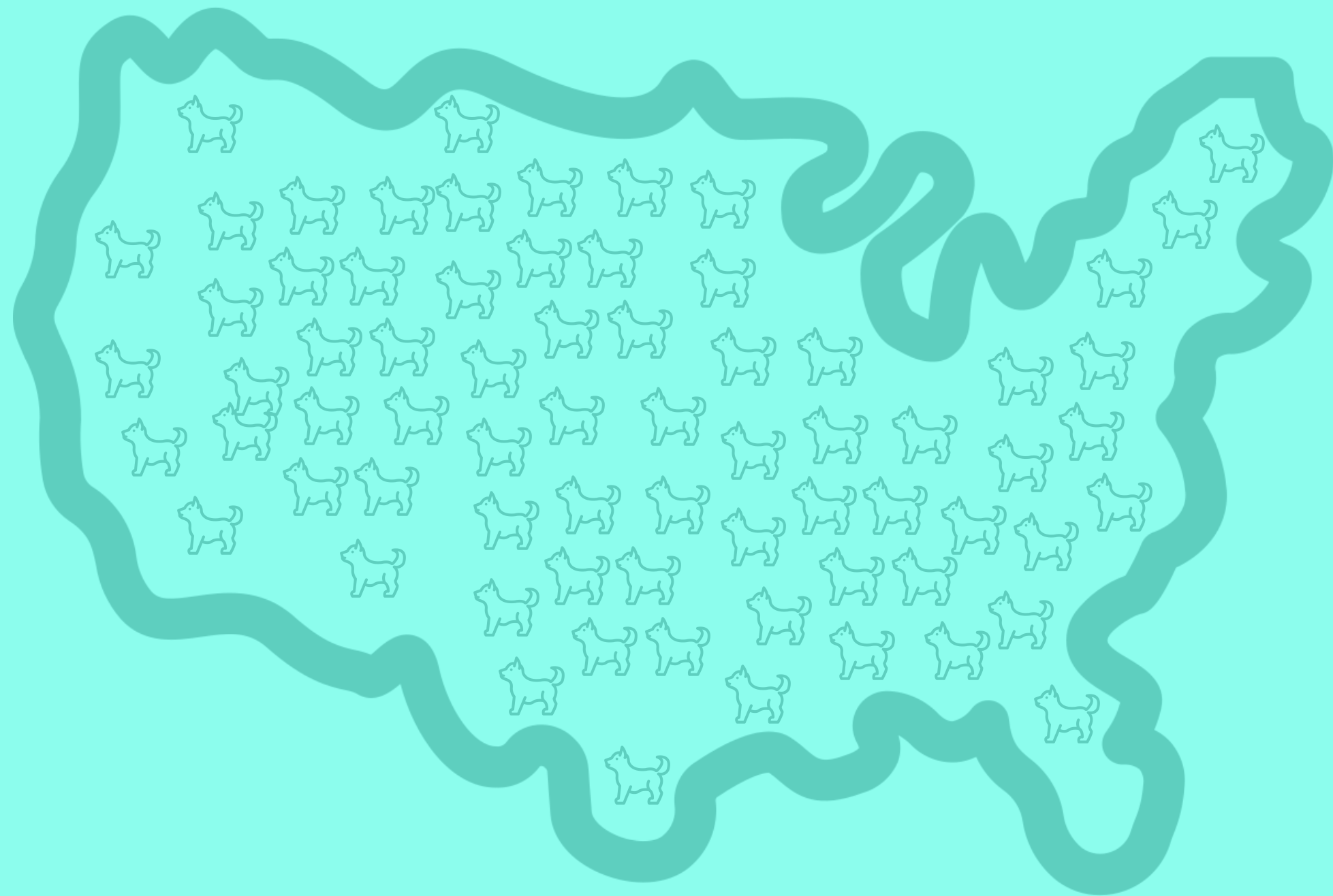
$$100 \times (1.64)^3 \simeq 440$$

Problems with the mean...</

Robust Statistics

Dispersion





The **estimator**

Key idea:

Unbiasedness

Consistency

Optimality

Is

Pearson Correlation

