ECO 602
Analysis of Environmental Data

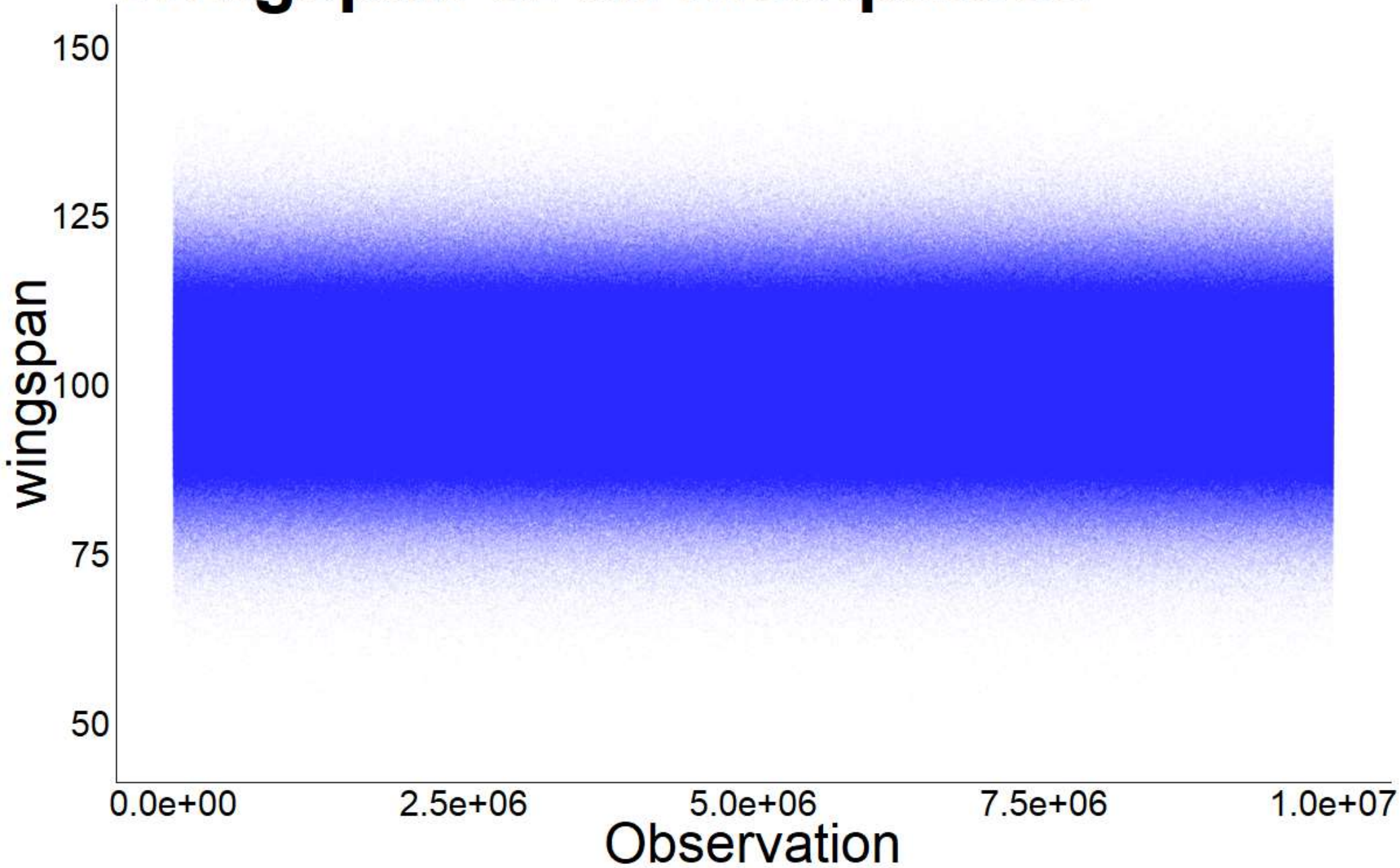
FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Today's Agenda

1. Quiz
2. Some clarifications from last lecture
3. More on populations and samples
4. Sampling: Assignment 1
5. Common data types
6. Graphical summaries
7. Assignment 1: group time

Wingspan of all mosquitoes



Announcements

Lecture notes on Moodle:

- I will (try my best to) post draft notes before class.
- I will post updated, revised, corrected, and/or beautified notes after.

Office hours in my office, room 311

- Wednesdays 9:00 – 11:00

Definition of ‘scale’: usage ambiguity

- Environmental data context: variable measurement scheme
- Other common meanings:
  1. Size, scope: large spatial scale
  2. Ratio of object size to model size: map scale, model trains

Context dependence: statistical populations and sampling units

1. Can be confusing and subtle.
2. We'll revisit the improved hypothetical fish example:
    1. Bullhead standard length

Bullhead in Massachusetts

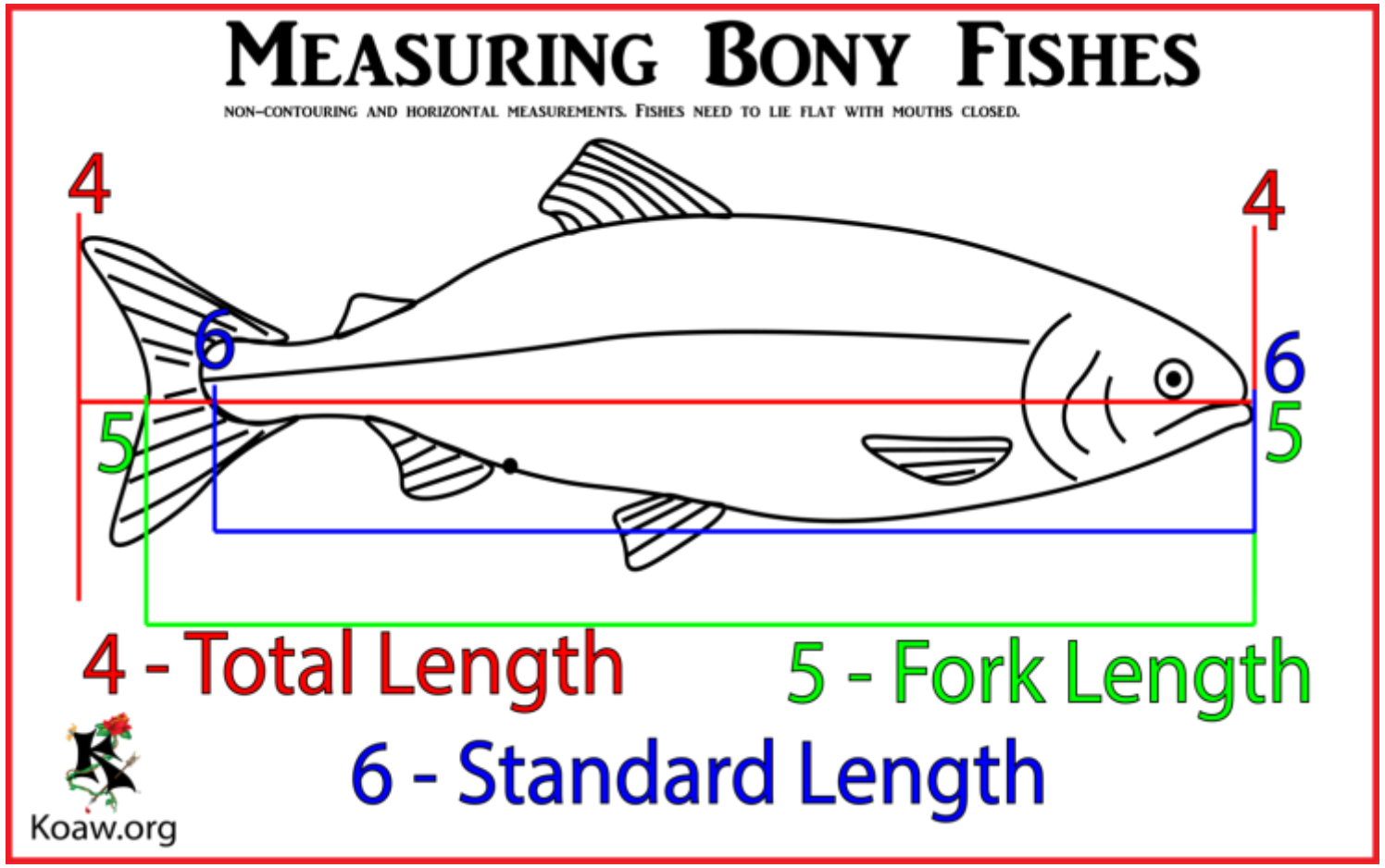


Image source https://www.koaw.org/measuring-fishes

Distribution of bullhead standard length in Lake Wyola

- Statistical pop. = all bullhead in Lake Wyola
- Ecological pop. = bullhead species range
- Sampling unit = individual bullheads
- Variable = standard length
  - What are the units of measurement and the scale of the variable?

Distribution of bullhead standard length in Massachusetts

- Statistical pop. = all bullhead in MA
- Ecological pop. = bullhead species range
- Sampling unit = individual bullheads
- Variable = standard length

Variability of bullhead standard length among Massachusetts lakes

- Statistical pop. = all lakes in MA
- Ecological pop. = bullhead species range
- Sampling unit = individual lake
- Variable = mean standard length in a lake

Sampling: statistical population context

1. Sampling scheme: plan for collecting data
2. Best case scenario:
    1. Independent samples: randomization
    2. Balanced sampling
    3. Best case rarely happens, especially in environmental and observational studies.

Sampling: implications for analysis

1. The frequentist paradigm assumes independence of observations
2. Violation of the independence assumption can inflate evidence against a null hypothesis
3. Many analyses can account for non-independence

Sampling

Randomized sampling: every observation is completely independent of all others

1. How would you randomize bullhead sampling in a lake?

Stratified random sampling:

1. Randomized sampling within treatment levels:
2. Lakes example: what are the levels?

The data and metadata concepts

- What is metadata?
- Why is it important?
- Can you find metadata for your assignment 1 datasets?

Assignment 1: populations, samples, sampling units

In your dataset and systems, what is/are the

- Ecological populations?
- Statistical populations?
- Sampling units?
- Sampling schemes?

What about the data types?

What is the sampling scheme?

Survey of data types

1. Continuous and discrete examples?

Survey of data types

1. Nominal and ordinal examples?

Survey of data types

1. Other data type examples?

For next week:

Exploratory Data Analysis

McGarigal Chapter 3