An intro to intro stat

You can do it!

Sara Stoudt

Assistant Professor of
Mathematics
Bucknell University

sas072@bucknell.edu
@sastoudt

Allison Theobold

Assistant Professor of
Statistics
Cal Poly

atheobol@calpoly.edu
@allisontheobold

Context

So You've inherited some intro stat materials...

How to put a gaise* spin on them (ONE STEP AT A TIME!)

*Guidelines for Assessment and Instruction in Statistics Education

GAISE GUIDELINES

1. Teach statistical thinking.
   - Teach statistics as an investigative process of problem-solving and decision making
   - Give students experience with multivariable thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and a purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
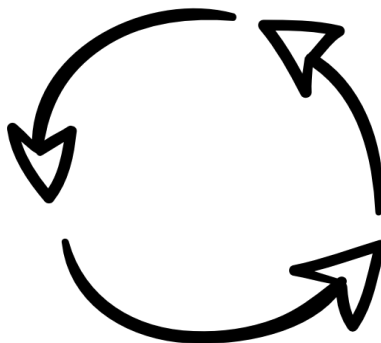6. Use assessments to improve and evaluate student learning.

Teach statistics as an investigative process of problem-solving and decision making.
(Try emphasizing communication)

For this exercise, you are to come up with a scientific/social question that interests you that you do not know the answer to and discuss what data you would need to collect in order to answer your question. Begin by stating your question and explaining why this question interests you. Then, describe the data needed to address your question by answering the following prompts.

- What would the setup be for the data collection (type of study)?
- What is the setting of the collection (where and when)?
- What/who is an observation?
- Who would collect the data?
- What information would you need to measure or collect? Are these quantitative or categorical variables?
- Give a rough estimate of how many observations you would need for a sense of scale. 10s? 100s? 1000s?

Teach statistics as an investigative process of problem-solving and decision making.
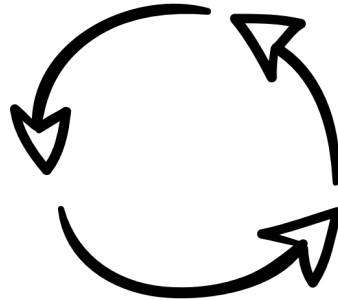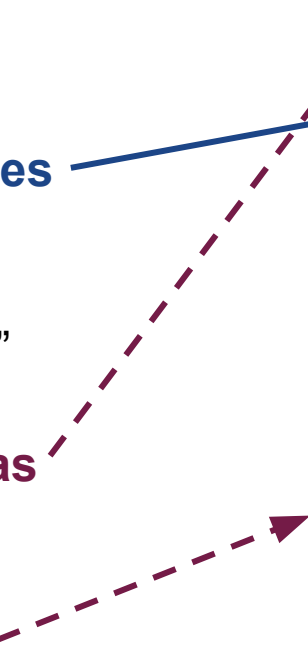## (Try emphasizing communication)

For this exercise, you are to come up with a scientific/social question that interests you that you do not know the answer to and discuss what data you would need to collect in order to answer your question. Begin by stating your question and explaining why this question interests you. Then, describe the data needed to address your question by answering the following prompts.

- Give a rough estimate of how many observations you would need for a sense of scale. 10s? 100s? 1000s?
- What could go wrong in the data collection that would lead to problems in answering your question?

Teach statistics as an investigative process of problem-solving and decision making.
(Try emphasizing communication)

For this exercise, you are to come up with a scientific/social question that interests you that you do not know the answer to and discuss what data you would need to collect in order to answer your question. Begin by stating your question and explaining why this question interests you. Then, describe the data needed to address your question by answering the following prompts.

- How well would the data from your study generalize to other similar situations?
- What is a related question that *can* be answered with the data you described?
- What is a related question that *cannot* be answered with the data you described?

Give students experience with multivariable thinking.

Try focusing on variability!

- Variation in a response **comes from many sources**

- Relationships "discovered" between two variables depend on how a study **has accounted for other variables that may influence that response**

- **Design studies** that effectively **control for the many variables that may influence the response** (e.g., blocking, inclusion criteria, direct control, etc.)

- Draw appropriate conclusions from an analysis, **describing the limitations of these conclusions**

Designing a study

Are crabs at higher latitudes are larger than ones at lower latitudes?

How should we measure a Fiddler crab's "size"?

● Carapace width

What are some potential sources of variation in the size of Fiddler crabs?

● Geography
● Climate
● Habitat
● Species
● Diet

Bergmann's rule

1 cm

Massachusetts

Virginia

Georgia

Sources of Variation diagram

We collected ~30 male, adult Minuca pugnax from thirteen marshes from Florida to Massachusetts and measured their carapace width with calipers. Water and air temperature data were collected from monitoring programs (i.e., LTER, NERR sites), nearby weather stations, and ocean buoys for 2016.

Observed Variation in: Carapace Width	**Sources of Explained Variation**	**Sources of Unexplained Variation**
Inclusion Criteria: Atlantic Coast **Controlled by Design:** Geography?	Water temperature Air temperature	Unknown Food Resources

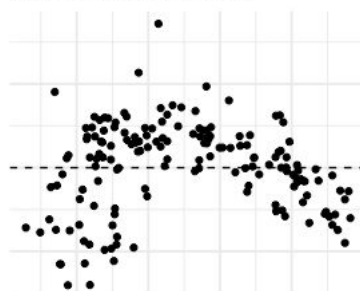# Focus on conceptual understanding.
(Try going visual)

https://drawdata.xyz/

Match each figure containing a scatterplot and regression line to its residual plot. Explain your reasoning in one sentence for each.
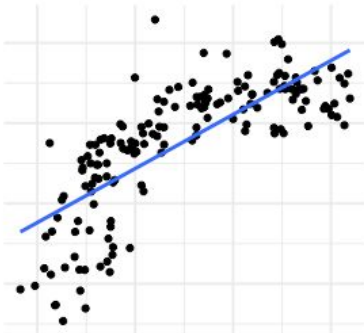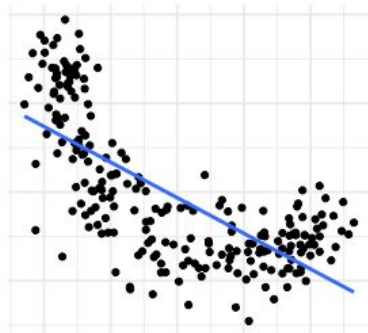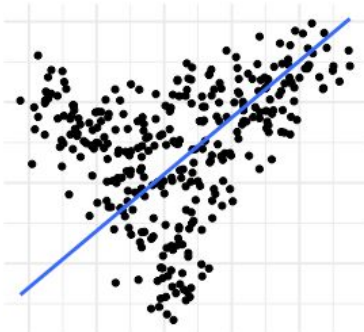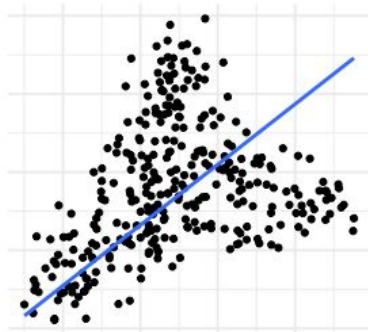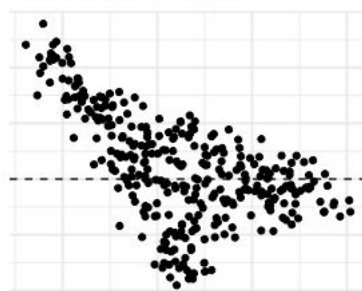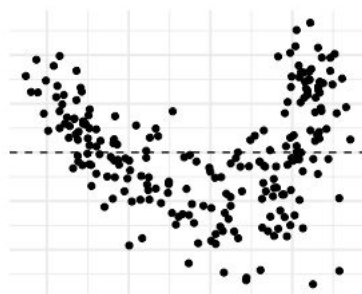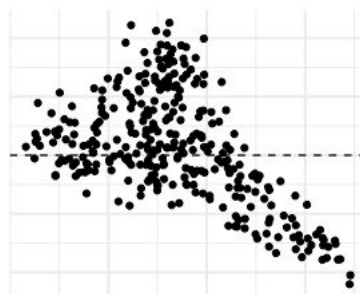
Focus on conceptual understanding.
(Try going visual)

Match each sampling distribution to its sample setup. Explain your reasoning in one sentence for each.
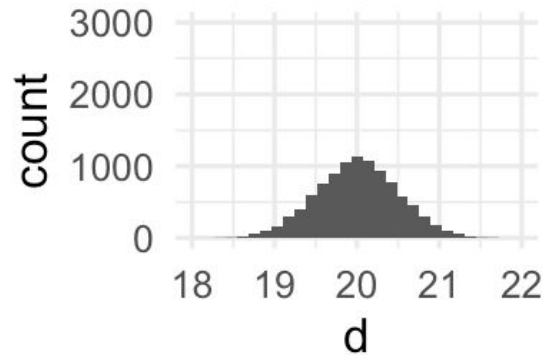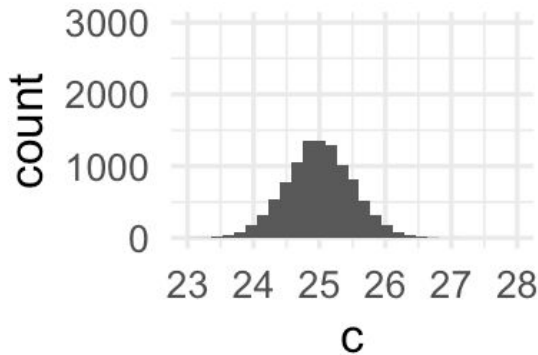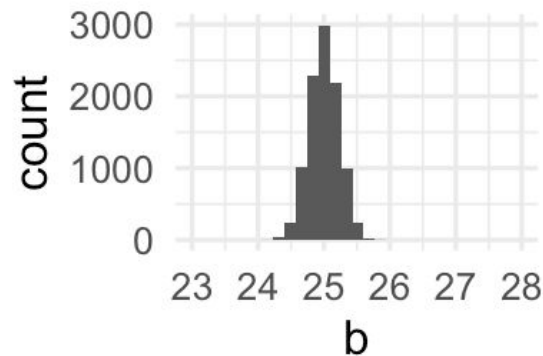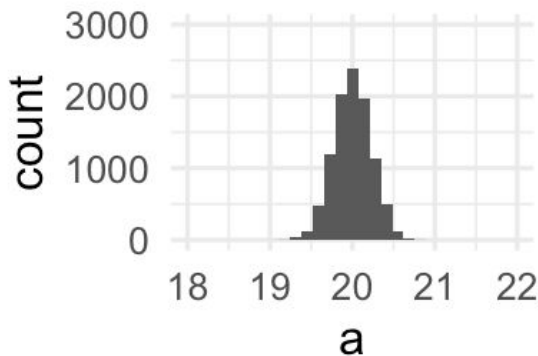
Setup 1: convenience sample, sample size = 100

Setup 2: convenience sample, sample size = 500

Setup 3: random sample, sample size = 100

Setup 4: random sample, sample size = 500

Helpful fact: the population distribution's mean is 20.

# Integrate real data with a context and purpose.

Awesome Public Datasets

NOTICE: This repo is automatically generated by apd-core. Please **DO NOT** modify this file directly. We have provided a new way to contribute to Awesome Public Datasets. Join the slack community for more communication.

- ✅ I am well.
- ❓ Please fix me.

TidyTuesday
A weekly data project in R from the R4DS online learning community

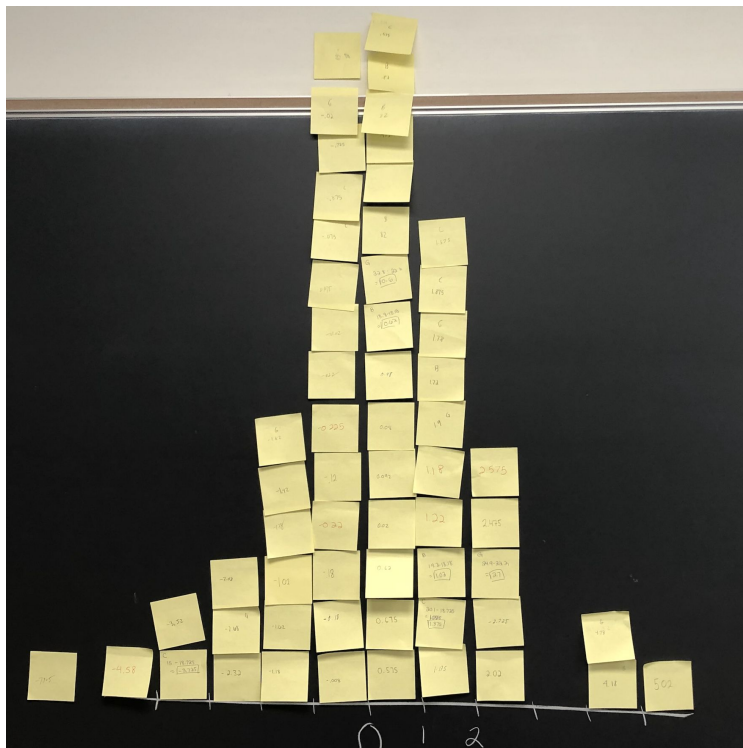variables   observations   values

lter datasampler

OpenIntro

DASH - Youth Risk Behavior Surveillance System (YRBSS): High School – Including Sexual Orientation   Youth Risk Behaviors

SHARE
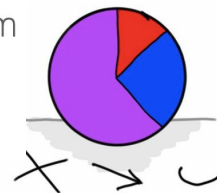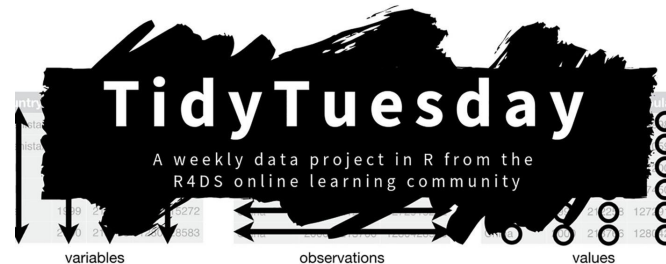
Gender Census 2022

Foster active learning. (Try getting tactile)

- Jar of pennies: mean year
  - Modern Dive Chapter 8
- Dolphin Therapy: randomization
  - Rossman SERJ 2008

- Effect of sample size on sampling distribution
- Effect of convenience sample on sampling distribution
- Histogram v. barchart
- motivate computer simulation (this is a pain to do by hand)

Use technology to explore concepts and analyze data.
(Try R)

R Studio Cloud

OpenIntro Labs for R and tidyverse

- Instructor account

- Classroom workspace

- Weekly assignments with code skeleton provided

- Notecards for common tasks

- Students work in groups

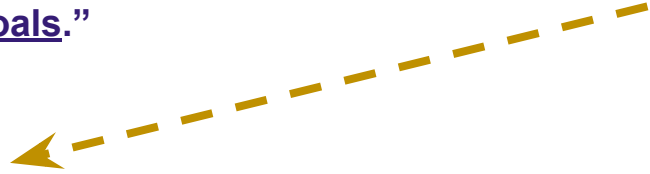# Use technology to explore concepts and analyze data.
(Try ISLE)
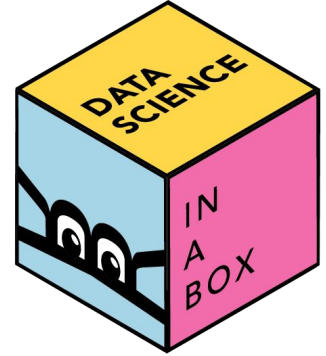
Integrated Statistics Learning Environment

Use assessments to improve and evaluate student learning.

"Students will value what you assess; therefore, assessments need to be <u>aligned with learning goals</u>."

Focus on Conceptual Understanding!

"Assess and <u>give feedback</u> on students' statistical thinking <u>as they progress</u> through the course."

"How would you explain what a p-value is to someone in your major who's never taken a stats class."

Formative vs. Summative Assessment

- monitor learning
- frequent
- low-stakes
- feedback
- help students improve their learning processes
- help faculty improve teaching

- evaluate learning
- midterms and finals
- high-stakes
- evidence for evaluation
- give a definitive measure of learning at a particular timepoint

QUESTIONS?

sas072@bucknell.edu
@sastoudt

atheobol@calpoly.edu

@allisontheobold