# Friday, October 26, 2018

- **Warm-up**

B is the correct answer – explain what is wrong with scenarios I and II|

A basketball players has a 70% free throw percentage. Which plan could be used to simulate the number of free throws she will make in her next five free throw attempts?

I. Let 0, 1represent making the first shot, 2, 3 represent making the second shot, ..., 8, 9 represent making the fifth shot. Generate five random numbers 0-9, ignoring repeats.

II. Let 0, 1, 2 represent missing a shot and 3, 4, ..., 9 represent making a shot. Generate five random numbers 0-9 and count how many numbers are in 3-9.

III. Let 0, 1, 2 represent missing a shot and 3, 4, ..., 9 represent making a shot. Generate five random numbers 0-9 and count how many numbers are in 3-9, ignoring repeats.

a. I only
b. II only
c. III only
d. II and III
e. I, II, and III

- **Sampling & sample statistics**

Social Objective: I will participate in the class activities.

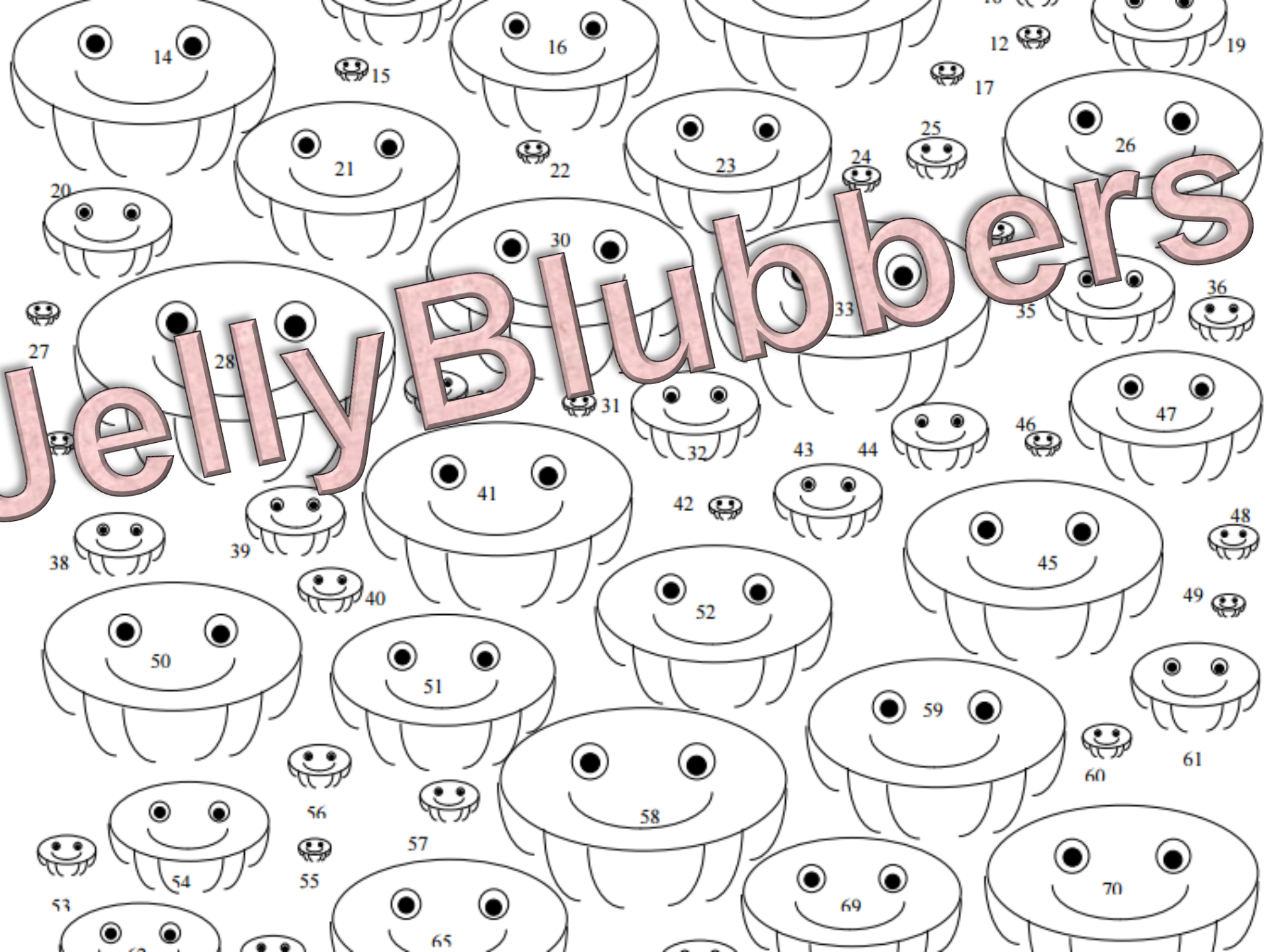
Content Objective: I will apply different types of random sampling to my class.

Language Objective: I will describe the types of random sampling clearly both verbally and in writing (in my notes).

# Why sample?

## Idea 1: Examine a Part of the Whole

- The first idea is to draw a sample.
  - We'd like to know about an entire population of individuals, but examining all of them is usually impractical, if not impossible.
  - We settle for examining a smaller group of individuals—a sample—selected from the population.
- Sampling is a natural thing to do. Think about sampling something you are cooking—you taste (examine) a small part of what you're cooking to get an idea about the dish as a whole.


JellyBlubbers

- Opinion polls are examples of sample surveys, designed to ask questions of a small group of people in the hope of learning something about the entire population.
  - Professional pollsters work quite hard to ensure that the sample they take is representative of the population.
  - If not, the sample can give misleading information about the population.

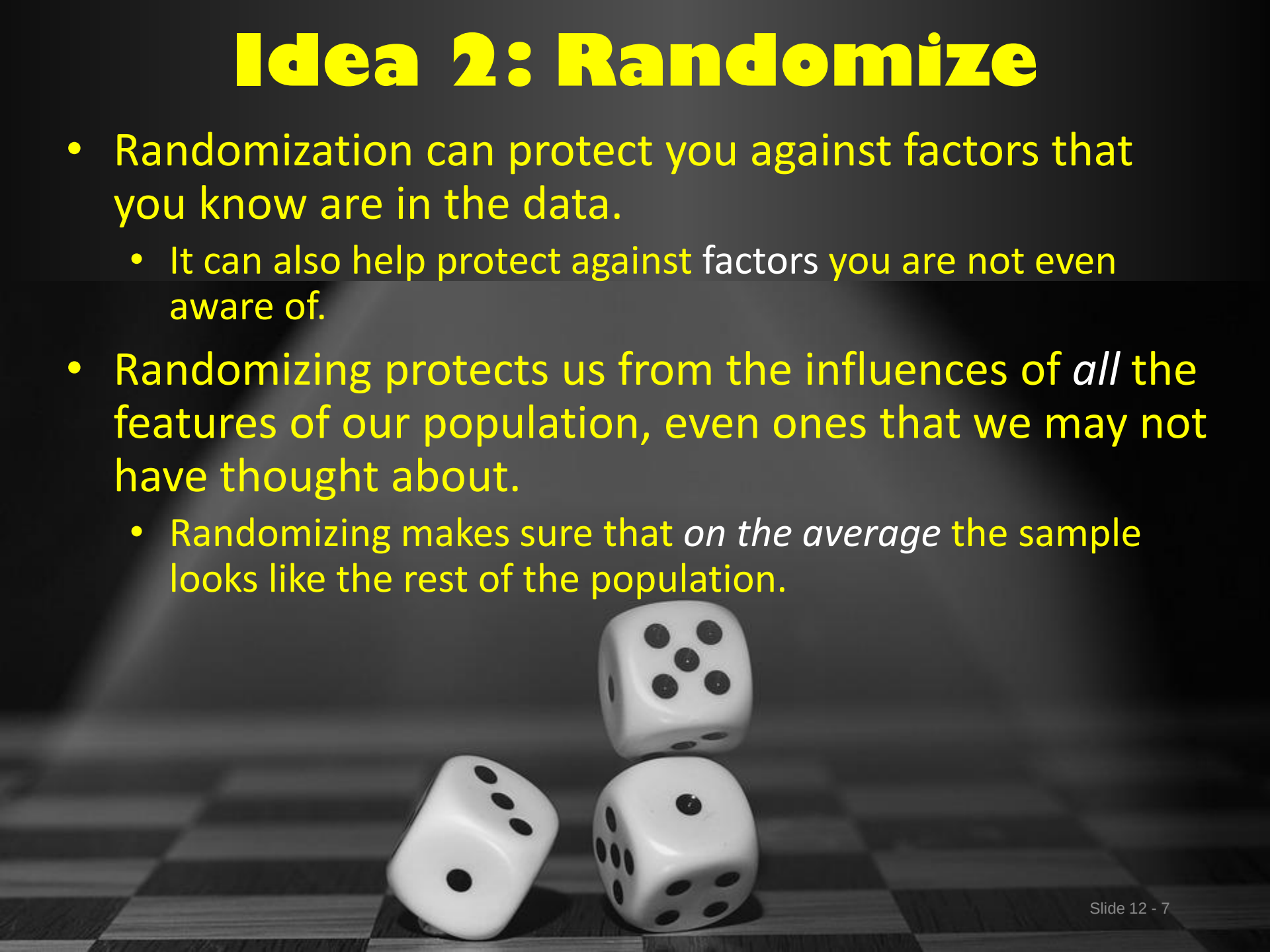
# Bias



- Sampling methods that, by their nature, tend to over- or under- emphasize some characteristics of the population are said to be biased.
  - Bias is the bane of sampling—the one thing above all to avoid.
  - There is usually no way to fix a biased sample and no way to salvage useful information from it.
- The best way to avoid bias is to select individuals for the sample *at random*.
  - The value of deliberately introducing randomness is one of the great insights of Statistics.

# Idea 2: Randomize

- Randomization can protect you against factors that you know are in the data.
  - It can also help protect against factors you are not even aware of.
- Randomizing protects us from the influences of *all* the features of our population, even ones that we may not have thought about.
  - Randomizing makes sure that *on the average* the sample looks like the rest of the population.



- Not only does randomizing protect us from bias, it actually makes it possible for us to draw inferences about the population when we see only a sample.
- Such inferences are among the most powerful things we can do with Statistics.
- But remember, it's all made possible because we deliberately choose things randomly.

**Ask yourself: "Did each individual in the population have an equal chance of being included in the sample?"**

# Idea 3: It's the Sample Size

- How large a random sample do we need for the sample to be reasonably representative of the population?
- It's the **size of the sample**, not the **size of the population**, that makes the difference in sampling.
  - Exception: If the population is small enough and the sample is more than 10% of the whole population, the population size *can* matter.
- The *fraction* of the population that you've sampled doesn't matter. It's the *sample size* itself that's important.