Have I lost my Marbles?

Unit 2 - Lab 3

Directions: Follow along with the slides and answer the questions in **BOLDED** font in your journal.

Some background...

- Let's assume we have a bag of 100 marbles.
- There are 23 blue marbles, and the rest are green .
- **What is the probability of selecting one blue marble from this bag?**
- In this lab, we will be *estimating* **probabilities** by drawing marbles from the bag and recording our responses.

But wait! I didn't bring a bag of marbles...

- No problem! You can **simulate** one in RStudio!
- Let's first create a vector for just the blue marbles

```
blue <- rep("blue", times = 23)
```

- **What do you think the `rep()` function does?**

Finishing up the bag

- Create a vector for the green marbles on your own.
  - **Write down your command in your DS Journal.**
- Next, combine the `blue` and `green` vectors into one new vector and name it `marbles`.
- **Write down your command in your DS Journal.**

Now we can create our first sample!

- If we just want to draw ONE marble out, we can run the following code:

```
sample1 <- sample(x = marbles, size = 1)
```

- **What color marble did you draw?**
- **If we wanted a sample of 10 marbles instead of just one, how could we revise the code?**
- Run your revised code and **write down the sample of 10 marbles in your DS Journal.**

But wait!

- We can select marbles one after the other, or put the marble back into the bag each time we draw.
- Within the `sample()` function, there is an option called `replace`, which we can set to either `TRUE` or `FALSE`.

To replace or not

- Let's take a sample of 150 marbles. Use the `replace` option - first set it equal to `FALSE`, and then try it with `TRUE`.

```
sample(x = marbles, size = 150,
       replace = FALSE)
```

- **What happens when `replace = FALSE`? Why do you think this happened?**
- **What happens when `replace = TRUE`? Which is better for this scenario?**

Back to those 10 marbles...

- Rerun your code for the sample size of 10 (don't forget to create a new name for this sample - maybe `sample2`?) and include the `replace` argument and **write down the sample in your DS Journal.**
- **What percent of the 10 marbles were blue?**

$$EstimatedProbability = \frac{\#ofBlueMarbles}{\#ofTotalDraws}$$

- **How does your answer compare to your neighbor's?**

Sample probability vs. theoretical probability

- **How does your answer from the previous slide compare to the theoretical probability of selecting a blue marble?**
- Let's run our code a few more times, but instead of just drawing 10 marbles from our bag, let's draw out 1000.
- Obviously, we don't want to write out all 1000 marble colors, so we can simply tally up the number of blue marbles that our sample produced by using the `tally()` function.

Sample probability vs. theoretical probability

- **In your sample of 1000, how many marbles were blue? What is your estimated probability?**
- **As a class, report your estimated probabilities. Examine them and make a guess about which value was the *typical* estimated probability.**
- Note: The estimated probabilities are all close to the theoretical probability.
- The larger the sample size, n, the closer the estimated probability will be to the theoretical.

For future study...

- **How would you design a simulation to find the probability that a second marble is blue?**