

# Experiment Results

## Introduction

Many times published experiments are not repeated due to various reasons (ie. limited resources, not as distinguished as original results, etc.). In addition, the experiments that are published in peer-reviewed journals are sometimes biased by the researcher's need for "significant results" instead of those that agree with natural law or known experiments which are usually considered "boring" or "uninteresting" and are subject to the file-drawer effect. This creates a problem of having too many remarkable studies published with non-realistic outcomes.

We've created a simple experiment to observe the variability of a single sample experiment versus multiple samples to highlight the importance of test reproduction in scientific experimentation. The experiment will be to flip a fair coin 20 times and count the number of times it lands on heads,  $X \sim \text{Binomial}(20, 0.5)$ . At such a low number of observations we would expect to see a good amount of variability from the mean value of 10. Each team member will individually repeat the experiment independently and then we will compare our individual results with that of the overall team outcome.

## Experiment

### Raw Data

The following table lists the results of each team member independently flipping a fair coin 20 times. An additional sample created from `rbinom()` is included using R's psuedo-random algorithm:

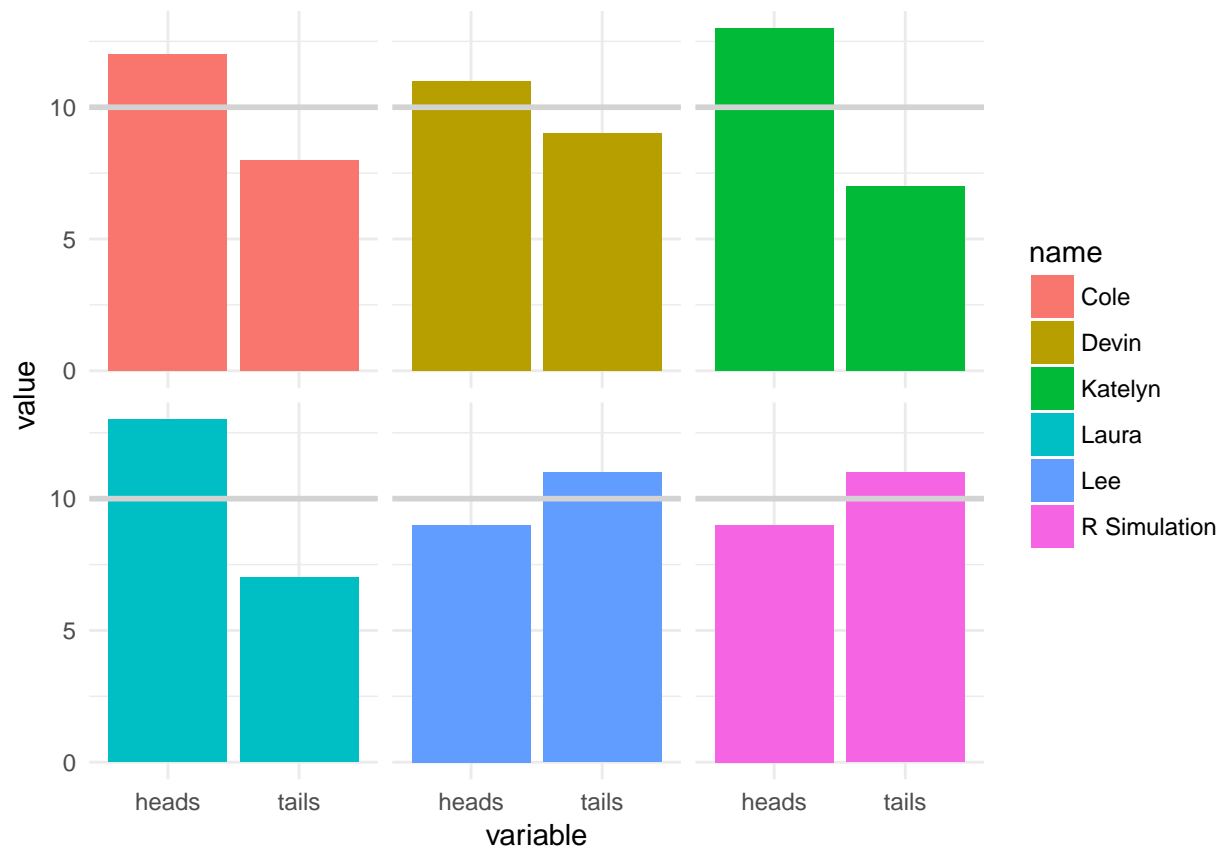


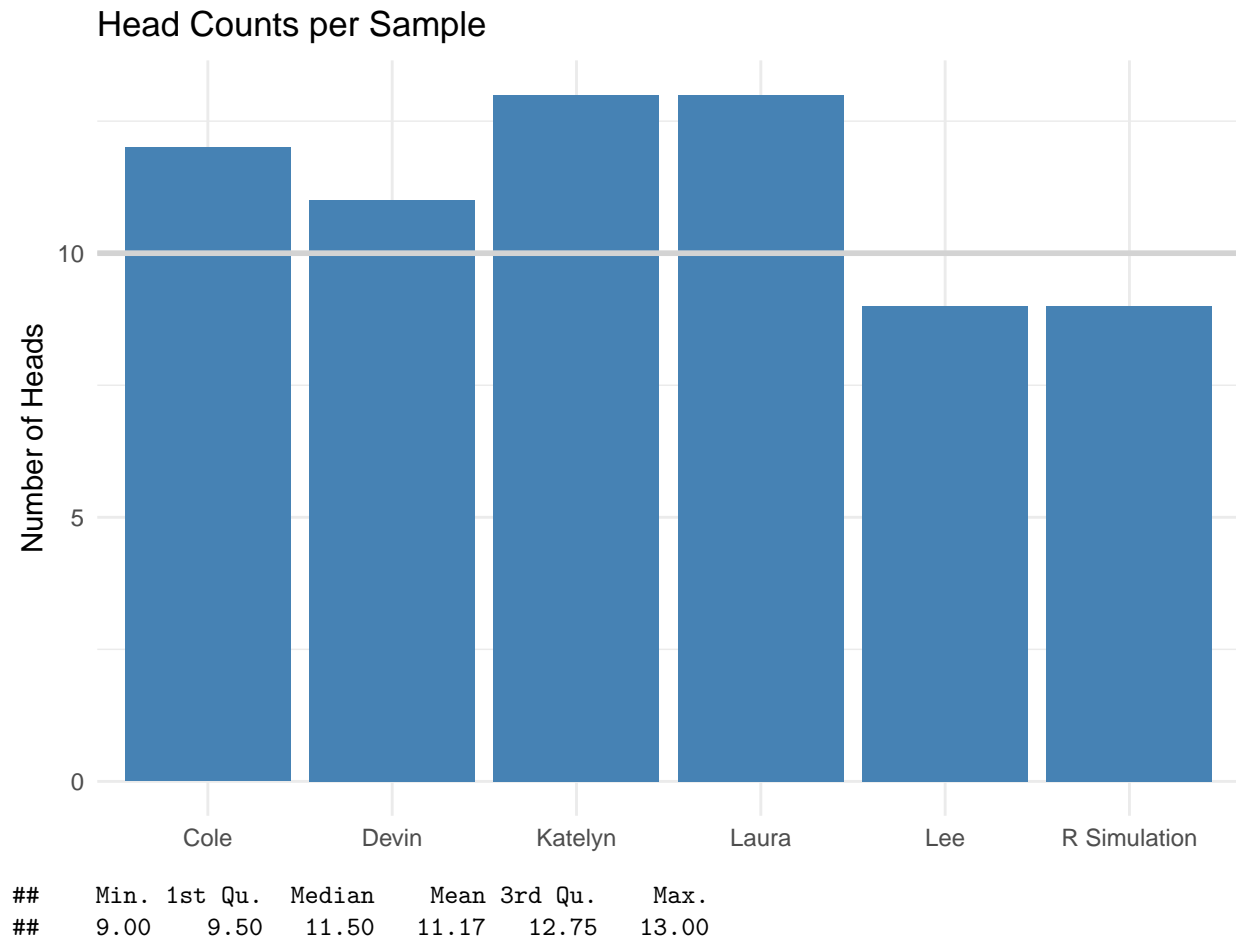
Table 1: Summary of Results

name	heads	tails	percentHeads	sampleError
Katelyn	13	7	0.65	3
Laura	13	7	0.65	3
Devin	11	9	0.55	1
Cole	12	8	0.60	2
Lee	9	11	0.45	1
R Simulation	9	11	0.45	1

The individual samples have errors that range from 1 to 3 counts. If Katelyn and Laura were to individually publish their findings, it would seem like they have found significantly different results than what was to be expected. Also, we see that R's psuedo-random generator produced results indistinguishable from an actual team member (Lee).

### Team Results

Here we are interested in all of the sample results taken together. The 6 independent trials are binomially distributed with  $n = 20$  and  $p = 0.5$ .



Our sample mean and sample error are :

- $\bar{x} = 11.2$
- $\sigma_{\bar{x}} = 0.476$

At a confidence level of 0.95, our intervals are 10.2 and 12.1.

## Conclusion

By taking all of the samples together we have significantly decreased our sample error. In fact, no individual experiment could produce an error less than 0.476. However the expected value of the experiment should've produced a mean of  $\mu = 10$ . Even with 6 total samples, our experiment was still 3 standard deviations away from the expected value. To increase the accuracy of the experiment, we could increase the number of samples by asking more people to repeat the experiment independently. We will avoid including more coin flips per sample to increase accuracy because that would in return increase the likelihood of getting a significant effect by chance alone.

For more information on this experiment, visit [The Experiment Experiment](#)