

02 Worksheet

David Gerard

2018-12-07

Read in the data from Case Study 1.1.2 from the `Sleuth3` R package.

```
library(Sleuth3)
data("case0102")
head(case0102)
```

1. Create a boxplot for `Salary` vs `Sex`
2. Calculate the mean salary for just the females.
3. Calculate the mean salary for just the males.
4. What is the difference in average salaries between males and females?
5. Now we will walk through simulating the randomization distribution.
 - a. The function `sample()` will permute the the elements of a vector. Try it out on `case0102$Sex`

```
sample(case0102$Sex)
```

- b. Take one permutation of `Sex` and recalculate the average salary difference between males and females.
- c. To automate this, consider the following code using the `replicate()` function

```
sample_vec <- replicate(n = 100,
  expr = {
    a <- rnorm(1)
    b <- rnorm(2)
    diff <- b - a
    diff
  }
)
```

```
head(sample_vec)
```

Whatever is placed within the braces “{}” will be run `n` times. Anything printed (in this case `diff`) will be returned and placed in `sample_vec`. Now modify this code to create a permutation distribution.

- d. Plot the permutation distribution and compare to the observed value calculated in 4.