Practice Problems on Causality

1. What concept is key to understanding and reasoning about causality?
2. Describe the fundamental problem of causal inference.
3. Give one example of the fundamental problem of causal inference.
4. What do we mean by "no causation without manipulation"?
5. What key feature of randomized experiments allows the researcher to assess the *average* causal effect of the treatment?
6. What are two potential problems with randomized experiments?
7. What is internal validity? External validity?
8. What is the key difference between a randomized experiment and an observational study?
9. What key assumption allows researchers to draw causal inferences from an observational study.
10. Briefly summarize a before-and-after design.
11. Briefly summarize a difference-in-differences design.
12. What kind of value is `FALSE`?
    A. character
    B. logical
    C. binary
    D. numeric
13. Translate the following statement using R's logical values (i.e., `TRUE` and `FALSE`) and operators (i.e., `!`, `==`, `&`, and `|`): "True or false is not false."
14. Read the `resume.csv` data set into R using the `read.csv()` function. Data are available here. The following questions assume that you assign this data set to the object `resume`.
15. Use the `head()` function to look at the first six rows of `resume`.
16. Use the `dim()` function to find the dimensions of `resume`.
17. Use the `summary()` function to obtain a summary of each variable in `resume`.

18. `resume` contains two binary variables, `sex` and `call`. (Remember that you'll need to use `resume$sex` and `resume$call` to access these variables–they are hidden in a data frame.) Create a table that compares the number of female applicants to the number of male applicants who did and did not receive a call back. Be sure to label the rows as `sex` and the columns as `call`.

19. Use indexing with logical vectors to find the callback rate for fictitious *female* job applicants in `resume`. Now repeat for *black females*. There are several ways to do this, see *QSS* section 2.2.3 and 2.2.4.

20. Take a look at the following nested `ifelse()` statement:

```
social$type <- ifelse(social$yearofbirth <= 1943 & social$primary2004 == 1, "Senior Voter",
                 ifelse(social$yearofbirth <= 1943 & social$primary2004 == 0, "Senior Non-voter",
                    ifelse(social$yearofbirth > 1943 & social$primary2004 == 1, "Non-senior Voter",
                       "Non-senior Non-voter")))
```

    What does this code do? See Table 2.3 in *QSS* for the variable descriptions. Recall that the `ifelse()` function takes three arguments. The first argument is a logical statement that can either be `TRUE` or `FALSE`. If it is `TRUE`, `ifelse()` returns the second argument. If it is `FALSE`, `ifelse()` returns the third argument.