   FAIR WARNING: Due to my rearrangement of course material, we in spring 2020 have
not yet seen the material behind question 2.

1. What will R's output be from the following code? Write your answer as a comment to
   each line.

```
c(1, 0, 1, 0) &   c(1, 1, 0, 0)
c(1, 0, 1, 0) && c(1, 1, 0, 0)
x <- rnorm(10); length(x > 0)
```

2. Consider the dataset `elmhurst`, which records information on 50 students of Elmhurst
   College. The dataset has two numeric variables, `family_income` and the amount of
   money each student received in financial aid (as a gift) `gift_aid`. Both variables are
   in thousands of dollars. Be careful with the units.
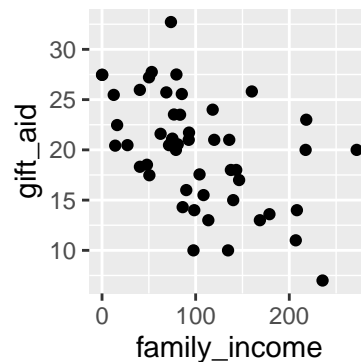
   (a) Provide R code to replicate the plot above, assuming the dataset is named `elmhurst`.

   (b) Does simple linear regression seem reasonable for these data? Explain.

   (c) Assume you estimated $\hat{\beta}_0 = 24$. Interpret this estimated coefficient in context of
       the data.

   (d) Does $\hat{\beta}_0$ make sense in the context of these data? Explain.

   (e) Assume you estimated $\hat{\beta}_1 = -0.04$. Interpret this estimated coefficient in context
       of the data.

(f) Predict the value of `gift_aid` for a student with `family_income` of $50,000$.

(g) Assume you calculated a 95% confidence interval for β_0 to be $(21.72, 26.92)$. Interpret this confidence interval in context of the data.

(h) Assume you calculated a 95% confidence interval for β_1 to be $(-0.06, -0.02)$. Interpret this confidence interval in context of the data.

(i) Fill in the blanks the function definition, which could be used with `boot::boot()` to calculate a confidence interval for the predicted amount of `gift_aid` a student would recieve if their `family_income` is $50,000$.

```
pred_y <- function(____, ___) {
    y <- ____[___, 1]
    X <- ____[___, -1]
    ____ <- optim(rexp(2), ll, method="L-BFGS-B", mx=X, y=y)$par

    --------------------
}
```

(j) Assume you calculated the above interval to be $(20.40, 23.93)$. Interpret this confidence interval in context of these data.

3. The bootstrap procedure has essentially two steps: resample, and on each resample, calculate something. What do we resample from?, how is the resampling done?, and what is being calculated on each resample?

4. What is the conceptual goal of the bootstrap procedure? Don't just say to calculate confidence intervals.

5. Supppose you calculated the following quantiles from some distribution.

```
##          5%          50%          95%
## -100.455730   -34.695169    -6.966642
```

(a) What percentage confidence interval would the 5% and 95% quantiles make?

(b) Which direction of skew does this distribution appear to have?

(c) Draw a picture of a density function that appropriately matches the skew from the previous question.