1. You don't need to simplify answers for this problem, they can be left as unreduced fractions with multiple terms. Two boxes have colored balls in them. Box A has 3 blue and 7 red balls in it, box B has 2 blue and 1 red.

   (a) Three balls are selected without replacement from box A and one ball is selected from box B. What is the chance that at least one of these balls is blue?

   (b) A random box is selected, and then 10 balls are selected with replacement all from that random box. Find the chance that there are exactly 4 blue balls in the 10 draws.

2. In an exercise class, participants are timed running a mile. The times (in minutes) are summarized in the table below, but some values are missing.

   (a) Fill in the table and draw the corresponding histogram. If you can't figure out the table, just put in anything and draw a histogram. More than one answer may be possible.

Time	Percent	Height of bar
5-6	10	
6-7	20	
7-8	25	
8-10		15
10-15		

   (b) Assuming that the percent in each bar is spread evenly for every interval, find the 45th percentile as accurately as possible.

   (c) Do you think the percent of people who took an above average amount of time is less than 50%, about 50%, or more than 50%? Explain briefly (one line is OK).

3. Suppose that the data from the previous problem is in a data frame called `data` which has two columns: `miletime` with the data used for the table on the previous page, and a logical vector called `smoker` that is TRUE for everyone who is a current smoker and FALSE for everyone else.

   (a) Write code to find the mean time to run a mile for both smokers and non-smokers.

   (b) Using the same data frame as in a), write code using `ggplot()` to make a histogram for the mile times for just the smokers, with the same class intervals as on the previous page (you can assume there aren't any values at the endpoints of those intervals if that makes it easier to think about).

   (c) In making the histogram in part b) did you ever use an argument called `y`? Whether you used it or not, what does that argument mean in this context?

4. I have a box with 40 tickets: 37 marked $\boxed{1}$, 1 marked $\boxed{2}$, 1 marked $\boxed{3}$, and 1 marked $\boxed{38}$. For each part of this question, use normal approximation if you think it is appropriate (look up area on table or use calculator). If you don't think it's appropriate say why not and if possible find the chance in some other way.

   (a) In 30 draws, find the chance that the sum of draws is 30 or less.

   (b) In 40 draws, find the chance that the sum of the odd numbers is 43 or more.

   (c) Write a line of code in R that will answer the following question as accurately as possible. In 50 draws, find the chance that the number of times I draw a $\boxed{1}$ is 43 or more.