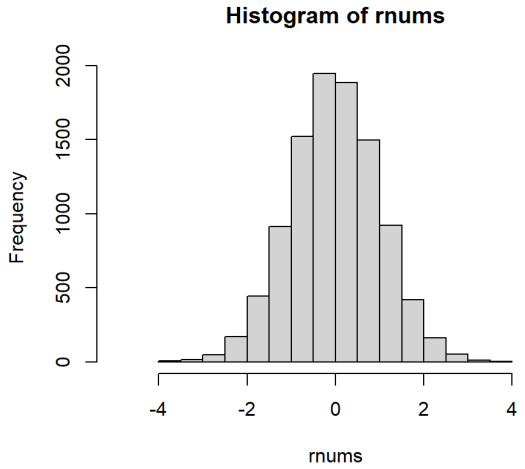
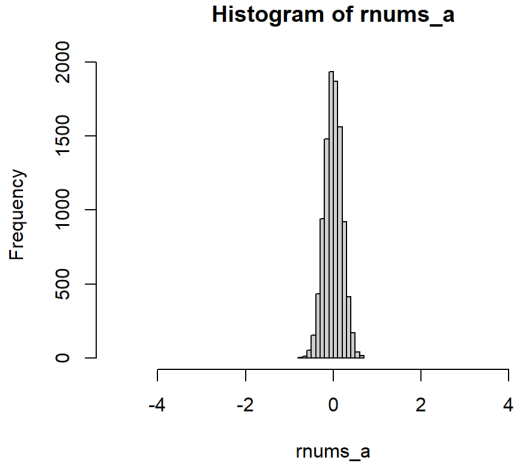
Homework 1 : Images and Explanations.

Problem 2(a). The comparison between the histogram for rnorm(10000, mean = 0, sd = 1) and rnorm(10000, mean = 0, sd = 0.2, and with xlim = c(-5, 5).



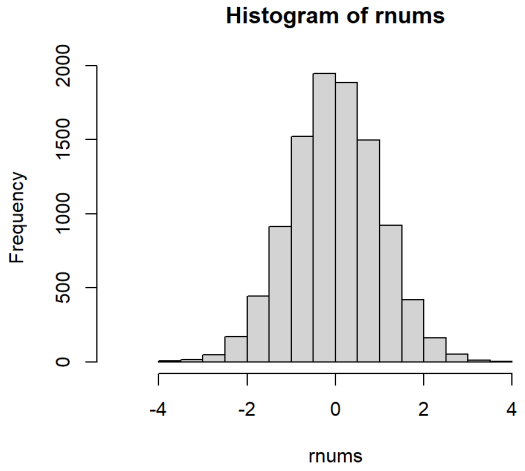
This histogram represents rnums = rnorm(10000, mean = 0, sd = 1) and was created with the following: hist(rnums, xlim = c(-5, 5)).



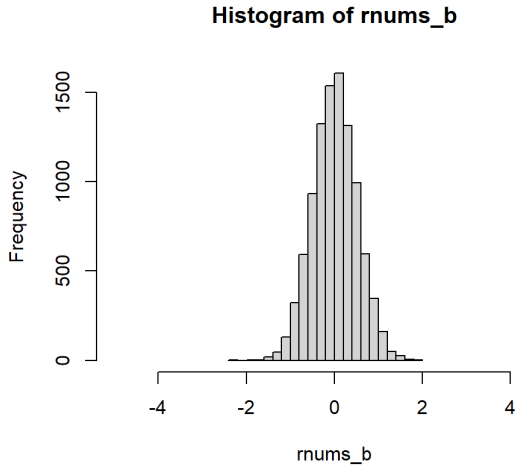
This histogram represents rnums_a = rnorm(10000, mean = 0, sd = 0.2) and was created with the following: hist(rnums_a, xlim = c(-5, 5)).

Comparison: Comparing the two histograms side-by-side, it is clear that the x range of the histogram with a standard deviation pf 0.2 is much steeper and narrower than that of the histogram with a standard deviation of 1.

Problem 2(b). The comparison between the histogram for rnorm(10000, mean = 0, sd = 1) and rnorm(10000, mean = 0, sd = 0.5, and with xlim = c(-5, 5).



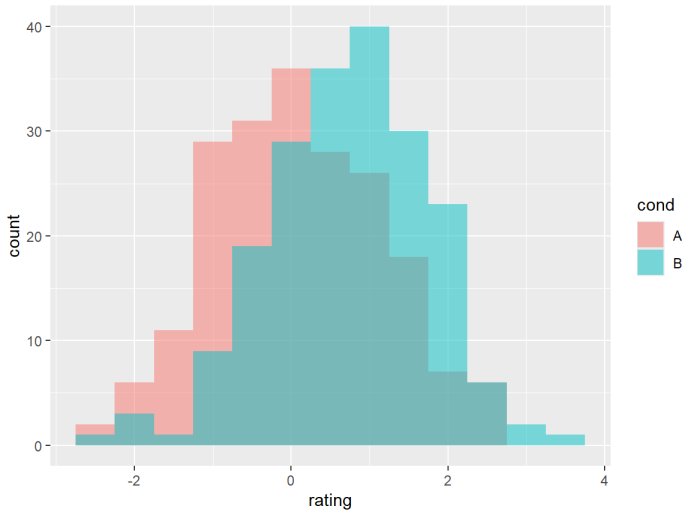
This histogram represents rnums = rnorm(10000, mean = 0, sd = 1) and was created with the following: hist(rnums, xlim = c(-5, 5)).



This histogram represents rnums = rnorm(10000, mean = 0, sd = 0.5) and was created with the following: hist(rnums, xlim = c(-5, 5)).

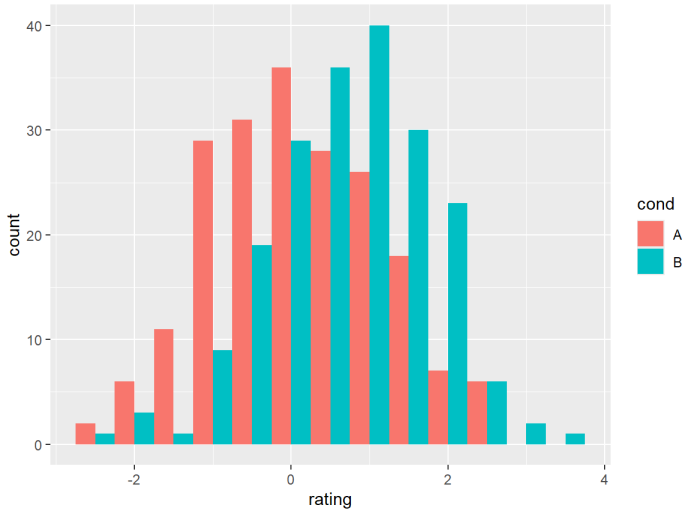
Comparison: Comparing the two histograms side-by-side, it is clear that the x range of the histogram with a standard deviation pf 0.5 is narrower than that of the histogram with a standard deviation of 1, with a slightly lower frequency for its peak.

Problem 3(b). Observing the overlaid histograms for dat and created with ggplot(dat, aes(x = rating, fill = cond)) + geom_histogram(binwidth = .5, alpha = .5, position = "identity").



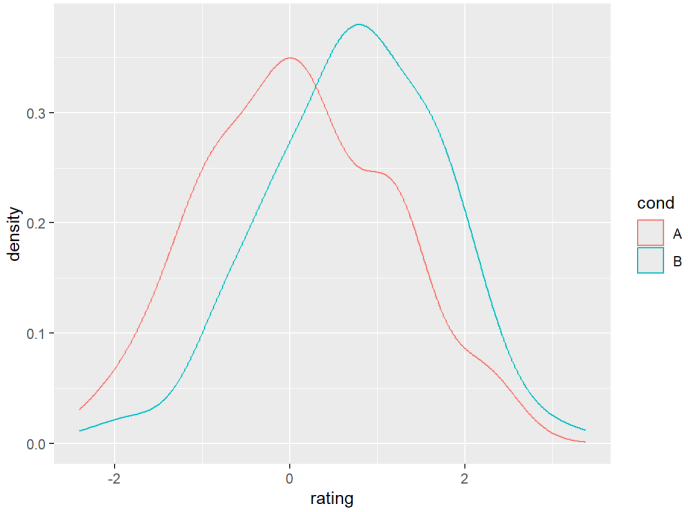
Observations: Where there is overlap between the two conditions the histogram is a dark green in color.

Problem 3(c). Observing the interleaved histograms for dat and created with ggplot(dat, aes(x = rating, fill = cond)) + geom_histogram(binwidth = .5, position = "dodge").



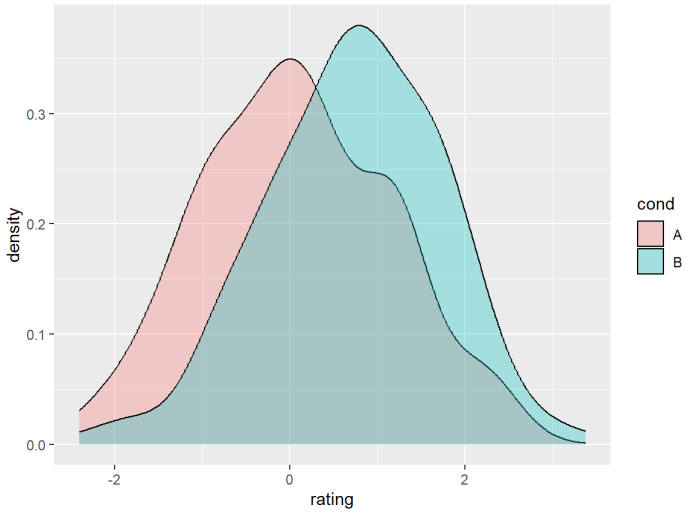
Observations: The two conditions have gaps between their bars where the other condition is visible. It is easier to make out the values where they overlap than in the previous overlaid histogram.

Problem 3(d). Observing the density plots for dat and created with ggplot(dat, aes(x = rating, colour = cond)) + geom_density().



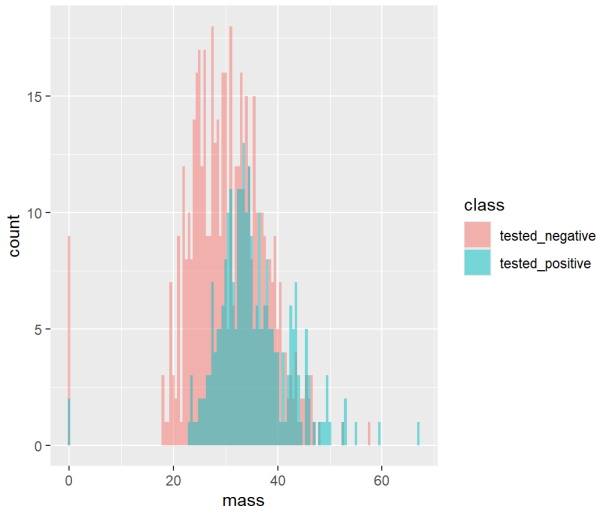
Observations: The plot of the two conditions shows their density and it is easy to read the values for the conditions and see where they overlap.

Problem 3(e). Observing the density plots with transparent fill for dat and created with ggplot(dat, aes(x = rating, fill = cond)) + geom_density(alpha = .3).



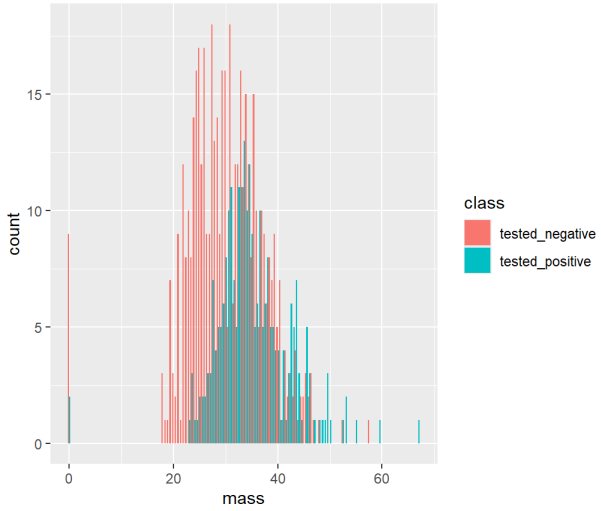
Observations: The plot of the conditions with transparent fill lines is very similar to the previous density plot, but with the colored in area of the two conditions and the different colored overlap it is a little easier to read.

Problem 3(f)(b). Observing the overlaid histograms for diabetes and created with ggplot(diabetes, aes(x = mass, fill = class)) + geom_histogram(binwidth = .5, alpha = .5, position = "identity").



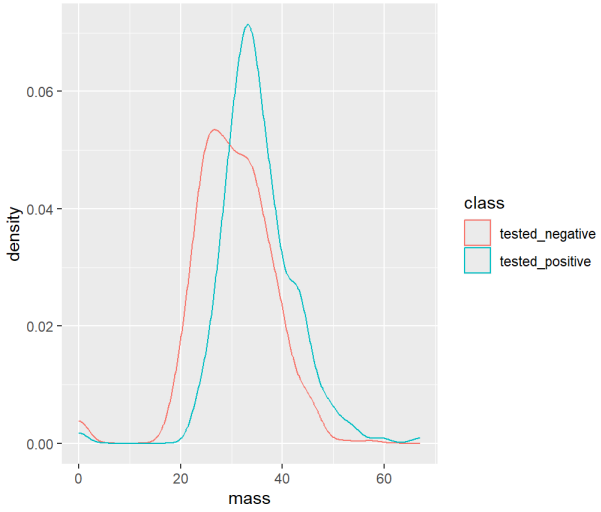
Observations: Where there is overlap between the two classes the histogram is a dark green in color and the original tested_positive color appears very little outside of this area.

Problem 3(f)(c). Observing the interleaved histograms for diabetes and created with ggplot(diabetes, aes(x = mass, fill = class)) + geom_histogram(binwidth = .5, position = "dodge").



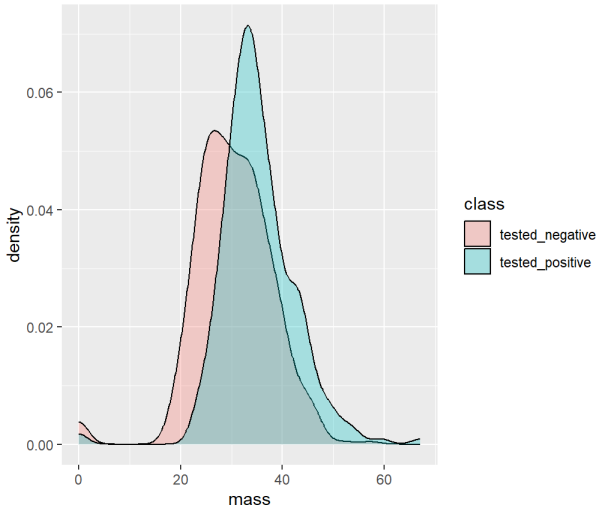
Observations: The segmented nature of this histogram makes it easier to interpret the otherwise overshadowed values for the tested_positive class.

Problem 3(f)(d). Observing the density plots for diabetes and created with ggplot(diabetes, aes(x = mass, colour = class)) + geom_density()



Observations: The density plots created from the diabetes classes is easy to read and interpret the data with regards to mass and density.

Problem 3(f)(e). Observing the density plots with transparent fill for diabetes and created with ggplot(diabetes, aes(x = mass, fill = class)) + geom_density(alpha = .3).



Observations: The density plots with transparent fill for diabetes is also easy to read and interpret the data with regards to mass and density. The fill color underneath classes and the dark color where the two classes overlap help increase its readability.