Name: Yola Charara

Data Mining HW1

R script images results and their explanations

Problem 2

a) mean=0, sigma=0.2

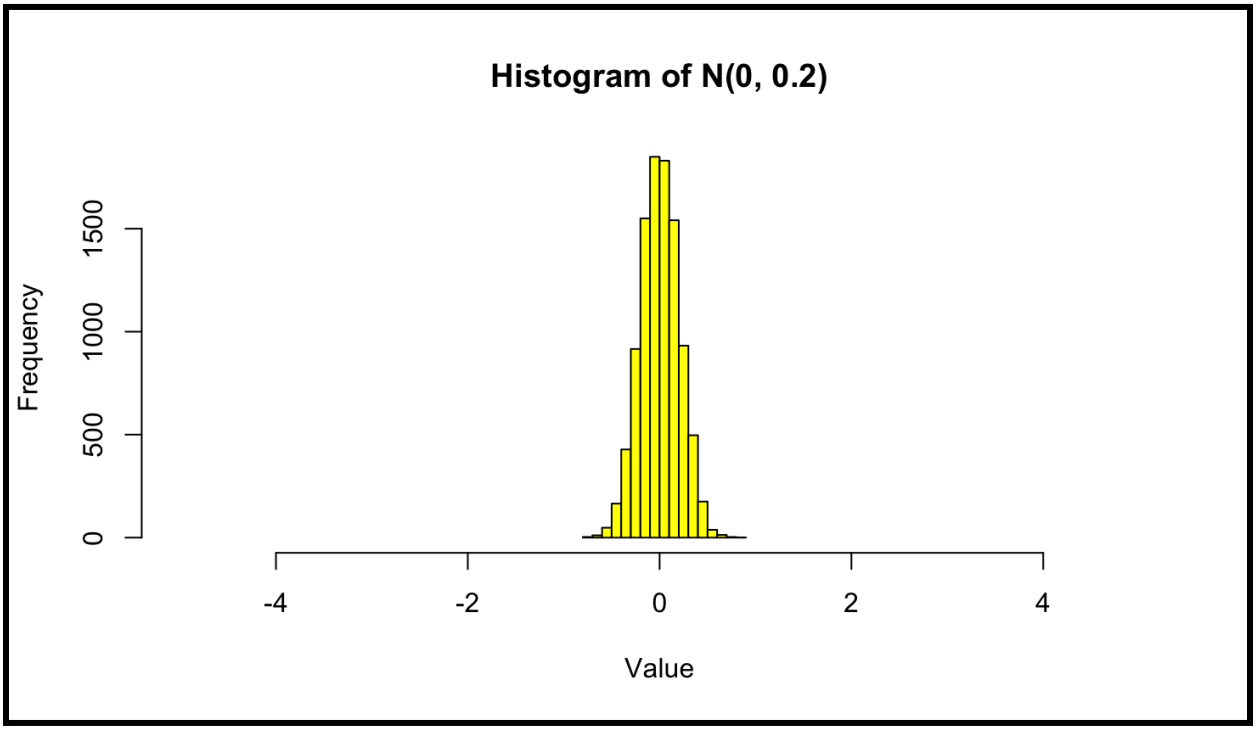

Figure (a) histogram has a normal distribution N with a mean of 0 and a standard deviation of 0.2. The small standard deviation results in a narrower and taller histogram because it keeps the values closer to the mean.

b) mean=0, sigma=0.5

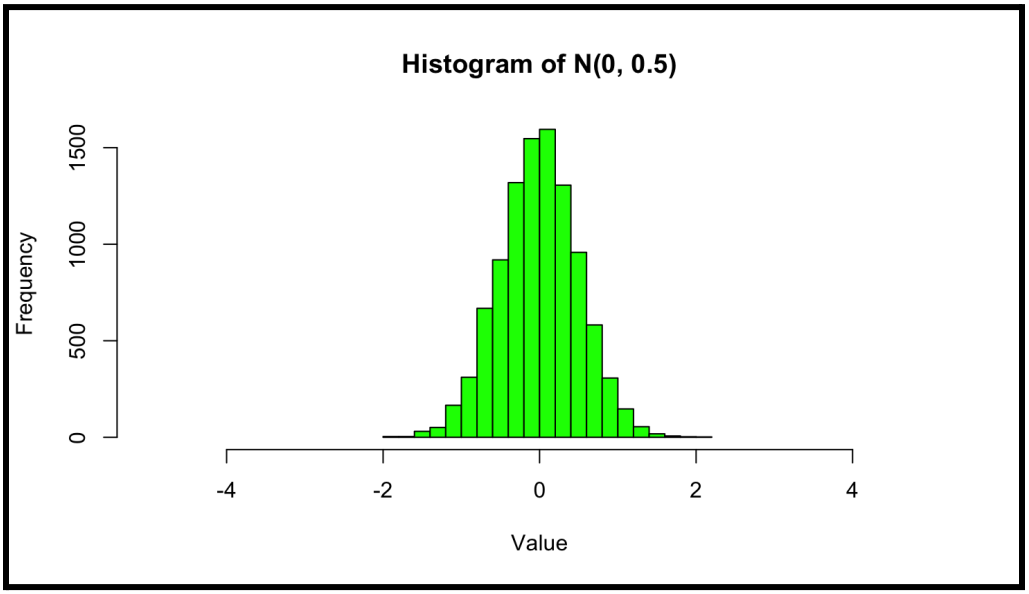

Figure (b) histogram has a normal distribution N with a mean of 0 and a standard deviation of 0.5. The larger standard deviation resulted in a wider histogram because it disperses the values away from the mean.

Problem 3

f) From the diabetes dataset, plot the mass attribute by class using ggplot2.

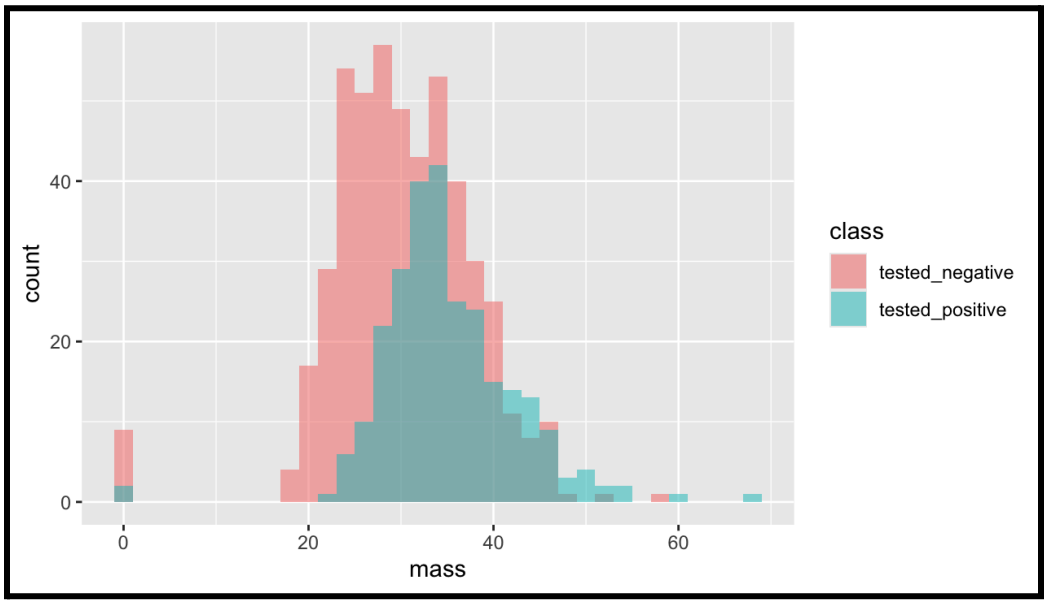

Figure (c) represents the overlaid histogram. This histogram overlaps both class groups' distributions of mass. This visual representation helps identify similarities and differences between classes.

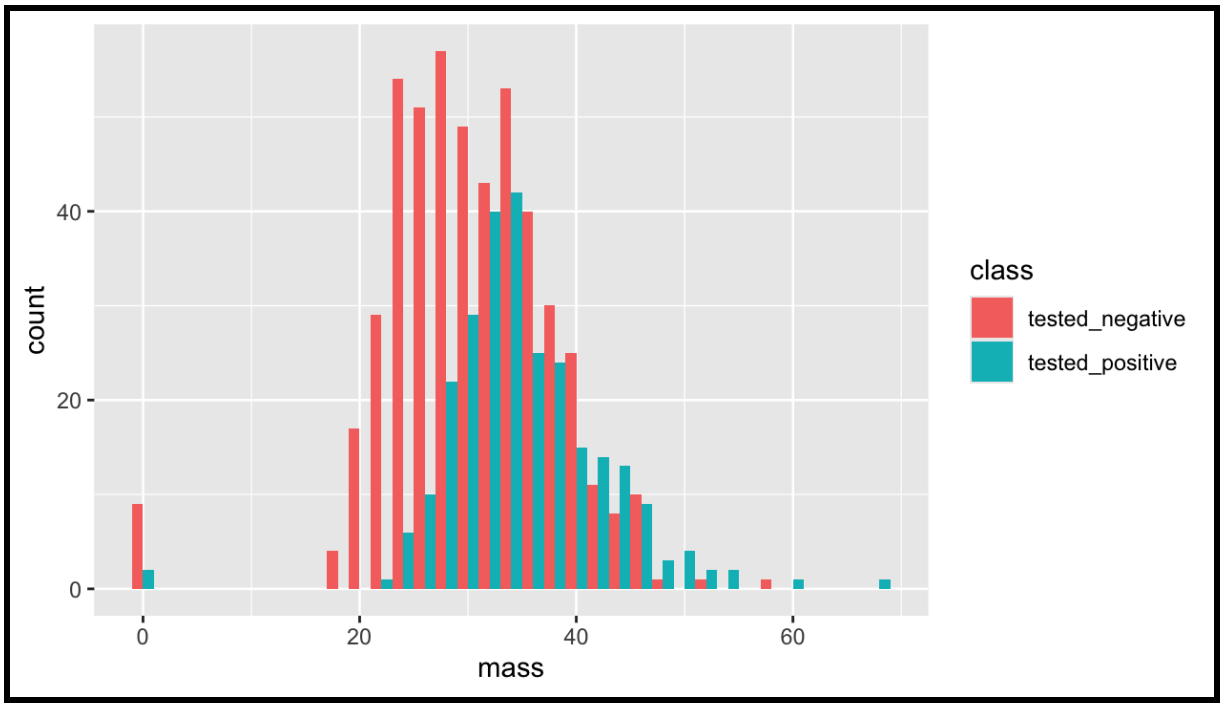

Figure (d) represents the interleaved histogram. It is a bar representation showing the different classes in a more clear and simple illustration.

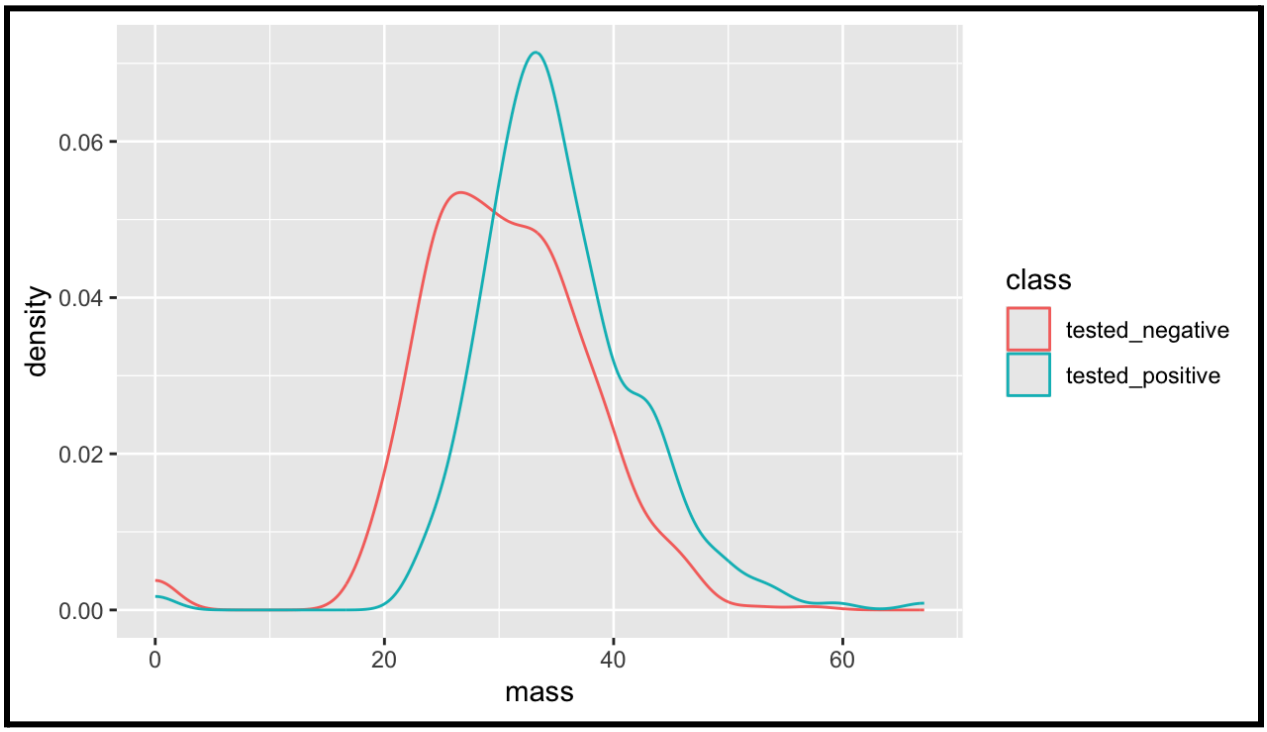

Figure (e) is the density plot. This plot represents each class in a curve distribution.

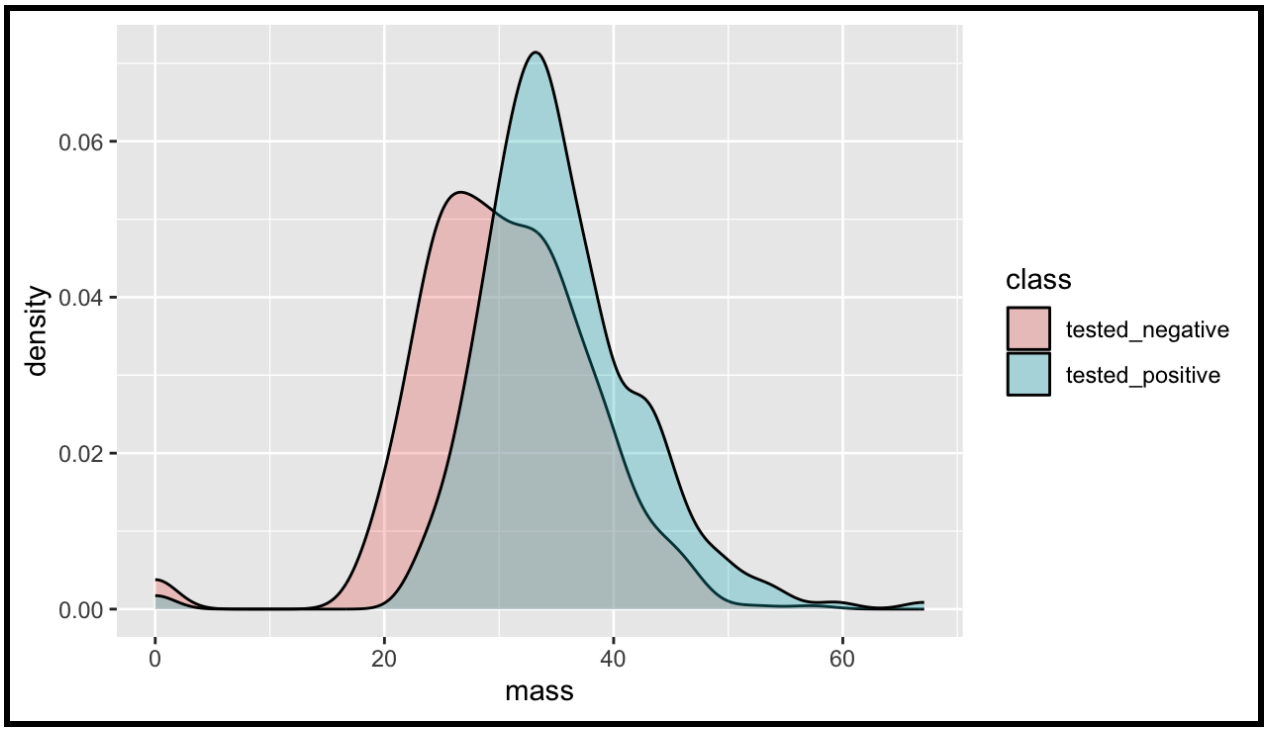

Figure (f) shows the density with fill plot. This graph also represents each class in a curve distribution, but this plot has an additional shaded overlap which helps to better illustrate the differences between classes.