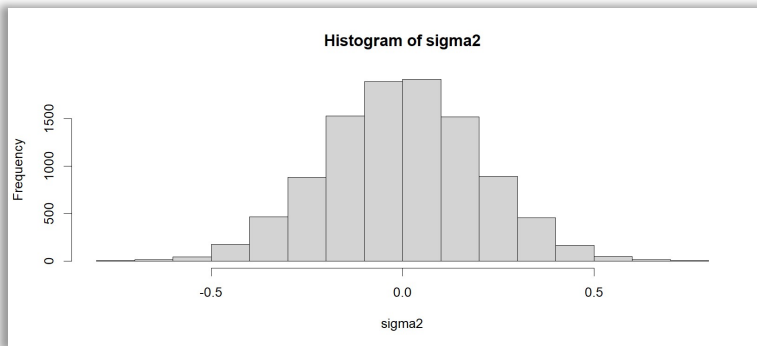


Ricky Parcels
9/20/25
DSC411
HW#1

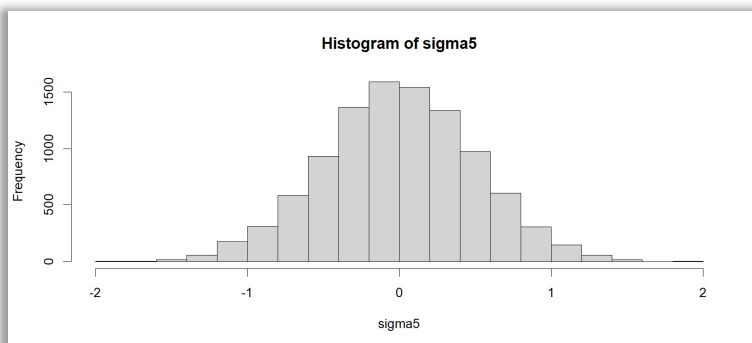
2. a) $mean = 0, sigma = 0.2$

The histogram shows a taller, narrower spread. Most values are between -0.5 and 0.5. The SD is small (0.2) which means most values are close to the mean (0).



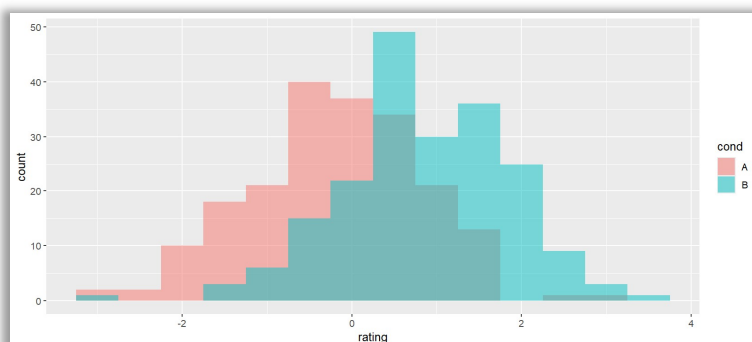
b) $mean = 0, sigma = 0.5$

The histogram shows a shorter, wider spread. Most values are between -1 and 1. The SD is larger (0.5) which means the values are spread out away from the mean (0) further.

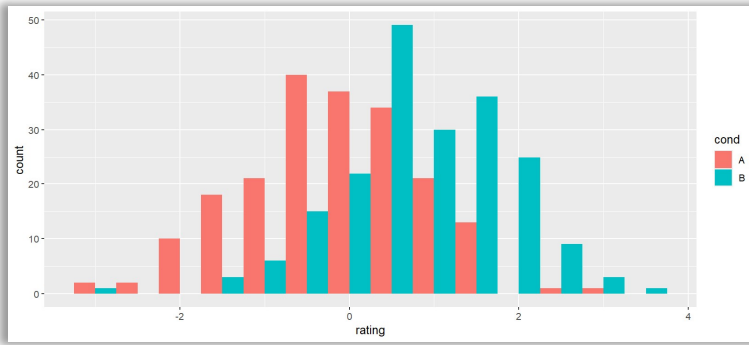


3. For "dat":

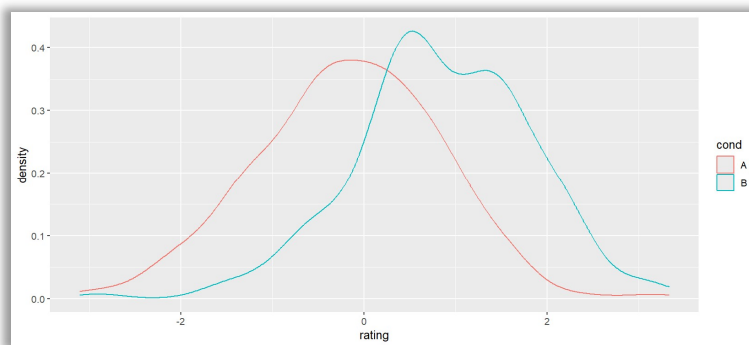
b) Overlaid Histograms



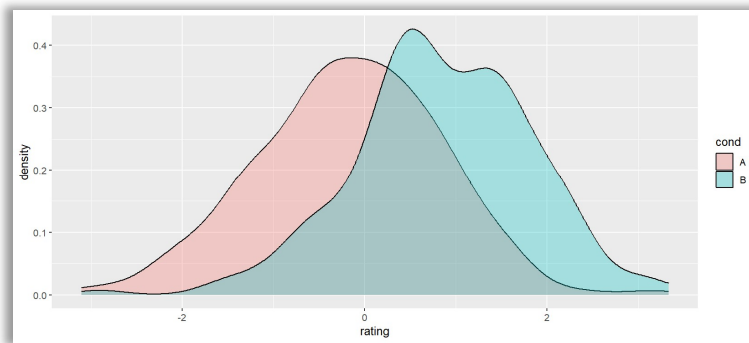
c) Interleaved histograms



d) Density plots

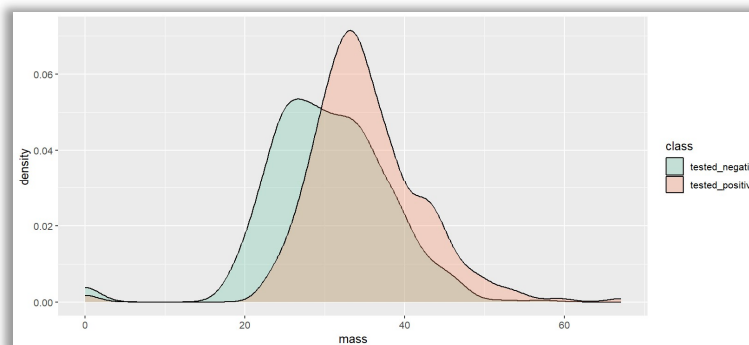
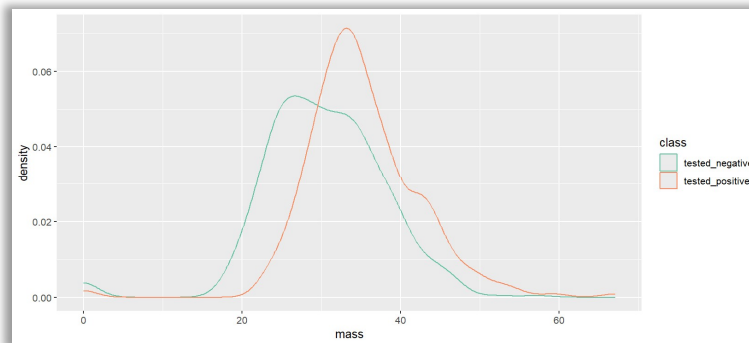
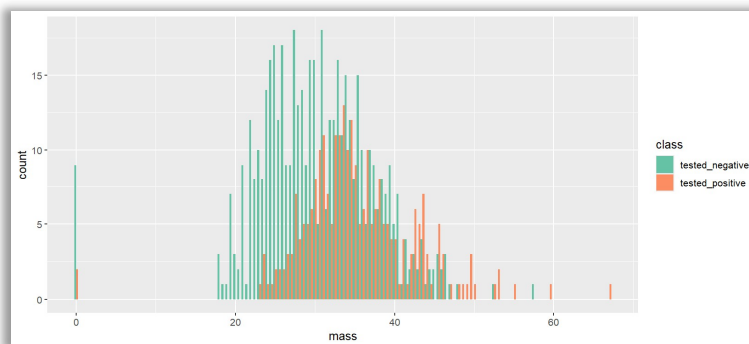
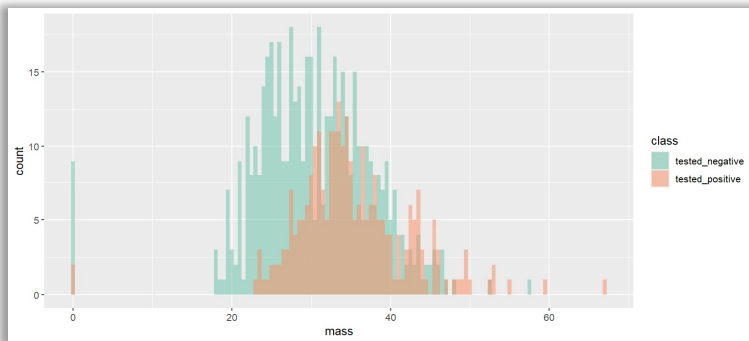


e) Density plots with semitransparent fill



f) For “diabetes”:

Overlaid histograms, interleaved histograms, density plots, & density plots w/ semitransparent



4. a) `passengers %>% drop_na() %>% summary()`

This operation removes rows without values, then summarizes each column (attribute) based on different statistics such as min, max, 1st quartile, 3rd quartile, mean, median, mode, class, length, etc.

b) `passengers %>% filter(Sex == "male")`

This operation filters down the dataset to show only those with “male” attribute.

c) `passengers %>% arrange(desc(Fare))`

This operation reorders the rows in descending order, based on the attribute of fare paid.

d) `passengers %>% mutate(FamSize = Parch + SibSp)`

This operation creates a new column (attribute) named FamSize and uses the value of Parch + SibSp to populate it.

e) `passengers %>% group_by(Sex) %>% summarise(meanFare = mean(Fare), numSurv = sum(Survived))`

This operation finds the mean of the Fare attribute, the sum of the numSurv attribute, then organizes them by sex (male/female)