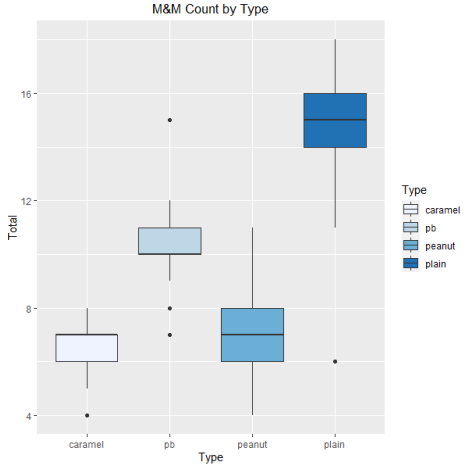
Mini Project 3
Kristina White

I hypothesized the bags on plain M&Ms contain more M&Ms than caramel, peanut, and peanut butter M&Ms. This could be relevant when purchasing M&Ms and choosing which type purchase, especially if the consumer values quantity.

The spreadsheet we used to collect data contained columns for M&M type, how many of each color were in the bag, the total number of M&Ms in the bag, the weight of each bag, and the scale used to measure the weight. Each row contains the information for one bag. We kept track of which scales were used to measure the weight of each bag because different scales can give slightly different weights for the same object. Additionally, we used drop-down menus for the columns for the type of M&Ms and the scale used to avoid people putting in, for example "peanut butter" or "pb," or "Peanut Butter," or "C" or "c," because R would treat these as different values when they actually represent the same value.

We did not randomize our design, however there was randomization present in the way the M&M bags were filled. M&Ms of a certain type are randomly mixed in a vat and dropped into each individual bag. We collected our data by taking four bags (one of each type) to ensure each group had the same sample size. Then we weighed each bag, keeping track of the scale, and counted the number of M&Ms of each color in each bag, as well as the total number of M&Ms in each, by hand.

Looking at the preliminary analysis of the data, the caramel and peanut M&MS had mean values for the total number of M&Ms in the bag between 6.5 and 7. For peanut butter, this value increased to 10.3, and plain had the highest average number of M&Ms per bag at 14.7.




I used the following three orthogonal contrasts to perform statistical analyses:

contr1 <- c(-1/3, -1/3, -1/3, 1)

contr2 <- c(1, -1/2, -1/2, 0)

contr3 <- c(0, 1, -1, 0)

I chose the first contrast comparing the caramel, peanut, and peanut butter M&Ms to the plain M&Ms since I am interested in whether there is a higher count of M&Ms in bags of plain M&Ms than for the other types. The other two contrasts comparing peanut butter and peanut to caramel and peanut butter and peanut to each other, respectively, are necessary to create a full orthogonal set for analysis. Looking at the summary of the linear model containing the contrasts, the coefficients for each contrast are significant. The coefficient for the first contrast (the contrast of interest) is 5.09434. Since this value is positive, and I used a positive value in the plain position in the first contrast, it is reasonable to conclude that plain M&Ms have a higher average count than the other types. The data collection was fairly efficient, but it could be improved by deciding ahead of time whether to count half or quarter M&Ms rather than having to make that decision once counting has already commenced.