PSY652, Unit 4, In class activity – Exploring Correlation

Dataframe: wtloss.csv

Description of Variables: In Unit 4 handout

1. Build inside your WeightLoss\_Notebook for this activity.
2. At the bottom of your notebook, add a second level header called: Get correlation from correlation between y-hat and y. Then add a code chunk just below this header (INSERT > R). Use cor.test to obtain the correlation between the predicted value of lbslost (we added this to our wtloss dataframe with add\_predictions, it’s called pred) and lbslost. Compare this to the correlation between lbslost and caldef.

cor.test(wtloss$pred, wtloss$lbslost)

1. Take a look at the regression model output from mod1 (the regression of lbslost on caldef). Look at the standardized slope (labeled Std. Beta). How does it compare to the correlation? The standardized beta is simply the regression coefficient that you obtain if y and x (including all x’s when we get to MLR) are z-scores (i.e., standardized scores, mean = 0, std = 1).
2. Add a second level header called: Get correlation through a SLR on z-scores. Then add a code chunk. First, create z-scores for both lbslost and caldef. Use mutate for this, and the zscore function from the mosaic package. Second, regress the z-score of lbslost on the z-score of caldef. Compare the regression slope from this model to the correlation. Also notice that the Beta and Std Beta in this output are the same, why is this the case?

wtloss <- wtloss %>%

mutate(caldef.z = zscore(caldef), lbslost.z = zscore(lbslost))

mod1\_z <- lm(lbslost.z ~ caldef.z, data=wtloss)

ols\_regress(mod1\_z)

1. You can also get the correlation if you have the unstandardized slope (labeled Beta in the model output) and the standard deviation of y and x (we have this from our summary statistics). The formula is below. Calculate the correlation using this information.

