Conduct Week 1 – Working in R

Replace *data* with the name of the dataframe and *var* with the name of the variable. For example, “hist(*data*$*var*)” might become “hist(student\_data$gre\_score)”

Things we will learn:

Part A

* Assigning something to an object (e.g., assigning a new dataframe to a dataset name): *data* <- …
* Reading in a dataset with read.table(). For example:
  + binge\_data <- read.table("https://minusthemath.com/data/binge.csv", sep=",", header=TRUE)
* Creating a nicely-formatted qualitative variable with factor(). For example:
  + binge\_data$sex <- factor(binge\_data$sex, levels = c(1,2), labels = c("Male", "Female"))
* Creating a histogram: hist(*data*$*var*)
  + For a qualitative variable: barplot(prop.table(table(*data*$*var*)))
* Creating a kernel density plot: plot(density(*data*$*var*))
* Creating a boxplot: boxplot(*data*$*var*)
* Install a package (only need to do once on your computer, until you reinstall R): install.packages("*package*")
* Load a package (need to do once per session you have R open): library(*package*)
* Creating a violine plot:

library(vioplot)

vioplot(*data*$*var*)

* Creating a stripplot:

library(lattice)

stripplot(~*var*, data = *data*, jitter = TRUE)

Part B

* Piping: |>
* Using mutate(), as describe in Ch. 4 of <https://r4ds.hadley.nz/>. Some examples:
  + *data* <- *data* |> mutate(*new\_var* = *old\_var* / 1000)
  + *data* <- *data* |> mutate(*per\_capita\_var* = *raw\_var* / *population\_var*)
* Creating a scatterplot: plot(*data*$*var1*, *data*$*var2*)