1. Click on the arrows in the delimiter option. What are the other possible delimiter types that RStudio recognizes? How does the data preview change when you select each of the other delimiters? How do you know that you have the correct delimiter (based on the data preview)?

Comma – each comma will be one column,

Semicolon – each semicolon will be a column,

Whitespace – each white space will be a column,

In this data, each data is separate by a comma so I should choose comma delimiter.

1. What do you notice about this data frame?

This data fame shows the # of entries that show on the screen out of entire entries.

1. What do each of the columns represent? What do each of the rows represent?

Each column represent a variable, and each row represent a person

1. How many variables are in this data set? How many cases?

9 variable

1. Which variables are quantitative variables? Report the unit of measurement for each variable.

Exerany; hlthplan; smokke100; height; weight; wtdesire; age

1. Which variables are categorical variables?

Genhlth; gender

1. Collect at least four summaries of the data, and describe in words what this information tells you.

The mean of height is 67.2

The mean of weight is 170

The standard deviations of height is 4.13

The standard deviations of weight is 40.1

The maximum of height is 93

1. Calculate the median weight and height, and compare these to the means (i.e., which is higher?).

Median of height is 67, median of weight is 165

1. Which measure of center may be more appropriate for weights? for heights? Explain your rationale.

The center should be median because it’s tallest in the center of the data.

1. Describe the shape of this distribution. Are there any unusual features? Do you think a bell-shaped curve would be a good model for this distribution? Explain your reasoning.

The data is mainly between 0.010 to 0.75, but there still have some number after 0.75. bell-shaped curve can be a good model because it’s hard to show some small data.