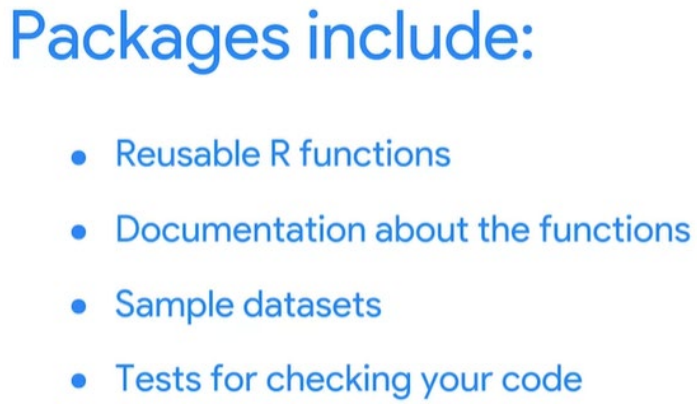
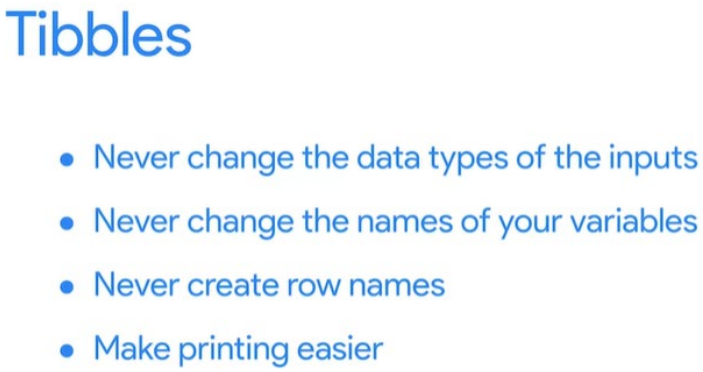
Week 1

- RStudio includes an R console pane for executing commands, a source editor pane for writing code, and an environment pane for managing loaded data. RStudio does not include a Command pane.
- Pipe (R): a tool in R for expressing a sequence of multiple operations, represented with “%>%”



- tidyr:a package used for data cleaning to make tidy data
- readr: used for importing data
- dplyr: offers a consistent set of functions that help you complete some common data manipulation tasks



- Tibbles can make printing easier. They also help you avoid overloading your console when working with large datasets. Tibbles are automatically set to only return the first ten rows of a dataset and as many columns as it can fit on the screen.
- A geom is the geometric object used to represent your data. In this case, the function geom_point() tells R to represent your data with points. Each geom function in ggplot2 takes a mapping argument. The mapping argument defines how variables in your dataset are mapped to visual properties, and is always paired with the aes() function. The x and y arguments of the aes() function specify which variables to map to the x-axis and the y-axis of your plot. In this

case, you want to map the variable “flipper_length_mm” to the x-axis, and the variable “body_mass_g” to the y-axis.