1. In the chapter readings and the tutorials you’ve learned about different data types. Fill in the associated data types (e.g. character, factor, integer) with each type of variable.

|  |  |
| --- | --- |
| Variable | Data Type in R |
| Categorical variable |  |
| Continuous numerical variable |  |
| Discrete numerical variable |  |

1. Describe in your own terms what the pipe operator (%>%) is.
2. In the Language of Data Tutorial, you are tasked with creating a new variable in the email50 dataset, named num\_char\_cat. You were given the code below:

email50\_updated <- email50 %>%

mutate(num\_char\_cat = if\_else(num\_char < med\_num\_char,

"[below median]",

"[at or above median]")

)

# Calculate median number of characters: med\_num\_char

med\_num\_char <- median(email50$num\_char)

# Create num\_char\_cat variable in email50

email50\_updated <- email50 %>%

mutate(num\_char\_cat = if\_else(num\_char < med\_num\_char,

"[VALUE IF TRUE]",

"[VALUE IF FALSE]")

)

# Count emails in each category

email50\_updated %>%

count(num\_char\_cat)

Fill in the code with the correct inputs **and** explain what the code is doing.

1. At the end of the Language of Data tutorial the code displays a scatterplot with different lines for each type of program. In your own words, describe what each component of the code below is doing:

ggplot(data = hsb2, aes(x = science, y = math, color = prog)) +

geom\_point()

1. In the Types of Studies tutorial they have you find how many students of each gender (sex) were admitted, using the code below.

ucb\_admit %>%

group\_by(Gender) %>%

count(Admit)

This code differs from the code given in the first tutorial by one major step, what does that step do?

1. In the three tutorials you saw a variety of different ways to filter rows out of a data set. In the table below describe what each of these different methods does!

|  |  |
| --- | --- |
| Code | Description |
| filter(Admit == "Admitted") |  |
| filter(Gender == "Male",  Admit == "Admitted") |  |
| filter(state != "District of Columbia") |  |