Rmarkdown

Logical is True false

Double is decimal

#Load packages

library(dplyr)

#Load Data

scoobydoo <- readr::read\_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2021/2021-07-13/scoobydoo.csv')

# Chapter 4

# Naming variables: snake\_case, CamelCase, dot.case (don't do this because you dont want you confuse your varialbe with a function)

##Functions

###Format: Functions name( set\_of\_arguments\_inputs)

seq(1,10) # seq(to = 1, from =10)

seq(from =10) # has a default value if I do not specify a to value

seq(1,10, by = 2) #prints odd numbers

seq(2, 10, by = 2) #prints even numbers

seq(1,1000, length.out = 15) # Length.out is now many # to print

(b <- seq(1, 1000, length.out = 15)) # highlight the sentence and shift + parent

#Chapter 5 Data Transformations/Wrangling data

##package: dplyr

##functions: filter, arrange, group\_by, mutate, summarize, select

#number of observations %>% ctrl +shift+ M

num\_obs\_2 <- scoobydoo %>%

group\_by(format) %>%

#summarize(count = n())

mutate(count = n())

# mutate tacks on to the end of a dataset

# summarize condenses dataset

# arrange default is least to greatest unless I put a - sign in front of count

# filter will keep all the variable and reduce the number of rows

# select keeps the rows, reduces the variables

num\_obs <- scoobydoo %>%

group\_by(format) %>%

summarize(count = n()) %>%

arrange(-count)

df <- scoobydoo %>% #dataframe

filter(format == "TV Series") %>% # removes any rws that are not a TV series

select(c(series\_name, network, season, imdb)) %>%

group\_by(series\_name, season) %>%

na.omit() %>%

summarize(avg\_imdb = mean(imdb))