

# Lab 1: R Basics and the Tidyverse

*Jenni Putz*

*1/10/2020*

## R Basics

## Tidyverse

```
install.packages("pacman", repos = "http://cran.us.r-project.org")

##
## The downloaded binary packages are in
## /var/folders/30/xcfr30vj55q2bbs6d1m27fqw0000gn/T//RtmpjtE7zh/downloaded_packages
library(pacman)
p_load(tidyverse)
```

## Test Your Learning

The following questions use content learned in Lab 1. All questions refer to the starwars dataset that we used in lab.

1. Use pipes to make a new data frame to include characters with blue eyes and retain only the columns of name, gender, and homeworld.

```
new_df <- starwars %>% filter(eye_color == "blue") %>% select(name, gender, homeworld)
```

2. Create a new data frame from the starwars data that meets the following criteria: contains only the mass column and a new column called mass\_half containing values that are half the mass values. In this mass\_half column, there are no NAs and all values are less than 50. Hint: to filter out NA values use !is.na()

```
new_df <- starwars %>% select(mass) %>% filter(!is.na(mass)) %>% mutate(mass_half = mass/2) %>% filter(mass_half < 50)
```

3. Use group\_by() and summarize() to find the mean, min, and max mass for each homeworld.
4. How many characters are female?
5. Run a regression of height on mass and gender. Filter out Jabba the Hutt and filter out the NAs in gender. What are the null and alternative hypotheses for the coefficient on gendermale? Interpret the coefficient on gendermale. Is this significant at the 1% level? What about the 5% level?