

In this assignment, you will practice some data manipulation tasks using the `dplyr` package and its built-in `starwars` data set. Please complete the following tasks, with `dplyr` functions being the default approach.

1.) Create a table that consists of the original data, but only includes species that appear at least three times within the data set. Order the table by species. Print the full table - you can pipe to `print(n = Inf)` to achieve this.

2.) Suppose a fan is interested in the distribution of which types of characters or from different planets, so they'd like to create a frequency table of the co-occurrence of these two variables. Complete this task in two different ways.

a.) Use `mutate()` to create a new identification variable for each character that combines their homeworld and species into one string (the `paste()` command is your friend here).

b.) Use `group_by()` and count the number of occurrences.

List the top 3 most common variables – there are different ways to do this in both a) & b).

3.) Calculate the average height for all female characters that appeared in *Attack of the Clones*.

4.) Suppose you go deep down a Reddit StarWars rabbit hole and find a fan theory that posits that the smallest (i.e. least mass) masculine and feminine (i.e. gender) characters on each planet (i.e. homeworld) always have yellow eyes. Devise a way to test this theory.