

dplyr (Part One) - Exercises

Exercises

mpg

- Load the `dplyr` package – either by loading the specific package or the `tidyverse` package as a whole.
- Make sure you have the most recent version of `dplyr` (1.0.2) to ensure smooth coding.
- Load the `ggplot2` package as well if you haven't done so yet. Store the default dataset `mpg` (Fuel economy data from 1999 to 2008 for 38 popular models of cars) in a data frame.
- Change the name of the following variables: `displ` to displacement, `cyl` to cylinders, `trans` to transmission, `fl` to fueltype.
- Drop the column `class`.
- Create a new column `distance_diff`, which shows the difference between `cty` and `hwy`. What's the highest difference in absolute terms?

- Create another new column `fivegal_distance`, which indicates how far one could drive on the highway with 5 gallons in their tank. How many columns do you have in your dataset by now?
- Bonus question: Which model(s) would get you the furthest on the highway with one gallon in the tank?

starwars

- Load the starwars dataset (part of `dplyr`) and store it in a data frame.
- Drop all columns after `species`.
- Create a new column `height_m`, listing the character's height in meters. Drop `height` afterwards.
- Keep the following variables: `name`, `mass`, all `_color` variables as well as all numeric variables.
- Set the height variable as the first column in your dataset. Put all the `_color`-variables at the end of the dataset. Try to find a neater way to do that than just manually typing all the `_color`-variables. Tip: Check out the `relocate()`-helpfile.

- Bonus question: Which characters were born more than 100 years before the Battle of Yavin?