

R Practice 4: dplyr

Ryan Safner

11/13/2018

Intro to dplyr syntax

1. Load the `gapminder` and `tidyverse` packages.

`select()`

1. Make a data frame containing the columns `year`, `lifeExp`, `country` from the `gapminder` data, in that order.
2. Select all variables, from `country` to `lifeExp`.
3. Select all variables, except `lifeExp`.
4. Rename `continent` to `cont`.

`arrange()`

1. Order by `year`.
2. Order by `year`, in descending order.
3. Order by `year`, then by life expectancy.

Piping, `%>%`

Note: think of `%>%` as the word “then”!

1. Combine `select()` Task 1 with `arrange()` Task 3.

`filter()`

1. Only take data with population greater than 100 million.
2. Of those, only take data from Asia.

`mutate()`

1. Make a new variable that is GDP instead of gdp per capita.
2. Make a new variable for GDP that is in millions
3. Make a new population variable that is the population in millions.

`summarize()`

1. Get the average GDP per capita
2. Get the average, minimum, maximum, and standard deviation for GDP per capita.
3. Get the average for GDP per capita, Life expectancy, and population

`group_by()`

1. Track the change in average GDP per capita over time. Hint, first group by year.
2. Get the average GDP per capita by continent.