BST 219
Core Principles of Data Science

Lecture 16: Advanced Data Wrangling continued
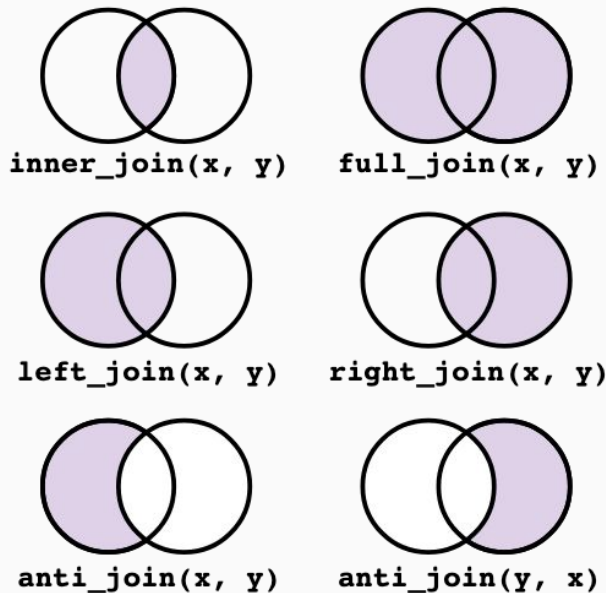October 24, 2024

Recipe of the Day!

Butternut Squash Pasta with Walnuts and Sage

~ When you realize tomorrow is Friday ~

Agenda

- Announcements
  - Lab this week!
  - Homework 2 is due 10/25

- Continue the advanced data wrangling module
  - Reshaping data
  - Combining tables
  - Dates and times

inner_join(x, y)   full_join(x, y)

left_join(x, y)   right_join(x, y)

anti_join(x, y)   anti_join(y, x)

Using the `gapminder` dataset, create a new dataframe containing only the `country`, `year`, `life_expectancy`, and `fertility` columns. Use the `pivot_longer()` function to reshape the dataset so that the life expectancy and fertility columns are combined into a single value column called "`value`". The column names should be stored in a new column called "`health_indicator`".

Make sure to run this code first

```r
# Load necessary libraries
library(dslabs)
library(tidyr)
library(dplyr)

# Load the gapminder dataset
data("gapminder")
```