

Week 3 DS Recitation Part 2

SIBDS 2024 @ Columbia

13 June, 2024

Data Import

Tasks:

1. Load the `FAS_pups.csv` dataset. Use both absolute and relative paths.

Your answer starts here:

2. Import the `FAS_pups.csv` again with column specifications by setting the `col_type` arguments. What happens if your specifications for column parsing aren't reasonable (e.g. character instead of double, or vice versa)?

Let's see what happens together:

Data Manipulation

Tasks:

1. In the `pups` data, do the following:
 - a. select the columns containing litter number, sex, and PD ears
 - b. Filter to include only pups with `sex == 1`
 - c. Filter to include only pups with `pd_walk` less than 11 and `sex == 2`
 - d. Create a variable that subtracts 7 from `pd_pivot`
 - e. Create a variable that is the sum of all the `pd` variables
 - f. Sort the variable `pd_eyes` in ascending order and `pd_walk` in descending order

Your answer starts here:

2. Write a chain of commands that:
 - Loads the `pups` data from `FAS_pups.csv`
 - Cleans the variable names
 - Filters the data to include only pups with `sex == 1`
 - Removes the `pd_ears` variable
 - Creates a variable that indicates whether `pd_pivot` is 7 or more days

Your answer starts here: