Lesson 3 Exercises

Stefanie Molin

April 14, 2017

Let's do some practice problems to challenge your understanding.

1. Query for two dataframes: (1) all AS in your office along with their employee IDs and (2) the accounts in the US and the AS employee ID associated with them. Use `dplyr` filtering joins to (a) preview the results that will be lost from dataframe (1) if you do an inner join on both tables and (b) preview the results that will remain in dataframe (1) if you do an inner join. (c) inner join dataframes (1) and (2) and confirm your results.
2. Pull in the first names of every employee *currently* working at [redacted] (cost centers US, NY, IL, SF), and, in a second dataframe, the first names of every employee that currently works at [redacted] but *not* in the US. Be sure to write a dynamic query, so that you only have to write one query! Use an `rbind()` to get the complete list of employee first names in a separate dataframe.
3. Using the two dataframes you queried for in `(2)` and set operations, (a) find all first names that are either in [redacted] (cost centers US, NY, IL, SF) **or** any other office, but **not** in both; (b) count how many people have each name, and sort it from most common to least common and by name alphabetically. Then, (c) flag and return the top 10 most common along with their counts, and (d) find the first names of the employees that are the only one in the company with that name, and (e) compare this result to the result from (a). (*Hint use* `setequal()`).