Assignment 7
EDH7916 | Summer C 2020

Benjamin Skinner

Use the IPEDS data sets, `hd2007.csv`and `ic2007mission.csv`, to answer the questions below. You may need to look up and download the data dictionaries for each file. Click the "continue" button on this page to see the data and accompanying dictionary files. You can also use the supplementary lesson on getting higher education data for help.

You **do not** need to save the final output as a data file: just having the final result print to the console is fine. For each question, I would like you to try to pipe all the commands together. Throughout, you **should** account for missing values to the best of your ability by dropping them.

For each question, show your data work and then answer the question in a short (1-2 sentence(s)) comment.

Questions

NB To answer the questions, you will need to join the two IPEDS data sets using the common `unitid` key. Note that column names in `hd2007.csv` are uppercase (`UNITID`) while those in `ic2007mission.csv` are lowercase (`unitid`). There are a few ways to join when the keys don't exactly match. One is to set all column names to the same case. If you want to use `left_join()` starting with `hd2007.csv`, you can first use the the dplyr verb `rename_all(tolower)` in your chain to lower all column names. See the help file for `left_join()` for other ways to join `by` different variable names.

1. How many chief administrator names start with "Dr."?
   NB Many chief administrators are listed on more than one line due to branch campuses. Make sure to take this into account by keeping only distinct names.
2. **BONUS** How many chief administrator names end with the title "PH.D." or some variant?
3. Among those schools that give their mission statement:
   1. How many repeat their institutional name in their mission statement?

   2. How many use the word *civic*?
   3. Which top 3 states have the most schools with mission statements that use the word *future*?

   4. Which type of schools (public, private-non-profit, private-for-profit) are most likely to use the word *skill* in their mission statement?
4. Among the schools that closed in 2007 or 2008 and give a date with at least a month and year:
   1. Which has been closed for the longest time? How many months has it been from its close date to the beginning of this current month (1 February 2020)?

   2. How many days were there between the first school to close and the last?

Submission details

- Upload your script (`<lastname>_assignment_7.R`) to our CANVAS course site no later than 11:59 p.m. EDT on the due date.