Exercise 7: Data Manipulation

Marcy Shieh

10/22/2020

Data manipulation exercises

Please submit the exercise as a R file and upload it to your `ps-exercises` Git repository.

1. Create a dataframe in R based on the table below.

Justice	State	Position	Replacing	Year confirmed	Senate confirmation vote	Nominated by
Clarence Thomas	GA	Associate Justice	Thurgood Marshall	1991	52-48	George H.W. Bush
Ruth Bader Ginsburg	NY	Associate Justice	Byron White	1993	96-3	Bill Clinton
Stephen Breyer	MA	Associate Justice	Harry Blackmun	1994	87-9	Bill Clinton
John Roberts	MD	Chief Justice	William Rehnquist	2005	78-22	George W. Bush
Samuel Alito	NJ	Associate Justice	Sandra Day O'Connor	2006	58-42	George W. Bush
Sonia Sotomayor	NY	Associate Justice	David Souter	2009	68-31	Barack Obama
Elena Kagan	MA	Associate Justice	John Paul Stevens	2010	63-37	Barack Obama
Neil Gorsuch	CO	Associate Justice	Antonin Scalia	2017	54-45	Donald Trump
Brett Kavanaugh	MD	Associate Justice	Anthony Kennedy	2018	50-48	Donald Trump

2. Download `justices.csv` from the `ps811` GitHub repository. The `justices.csv` file contains Martin-Quinn scores (a measure of ideology) for justices from 1937 to 2019.

For the data manipulation questions below, use pipes (`%>%`).

3. Merge the `justices.csv` and `SCDB_2020_01_justiceCentered_Citation.dta` datasets using one of the `join` functions. Before performing the merge, check that the variable names you want to merge are the same in both datasets.

    - If the names are different, you will need to rename the variable names for one of the datasets so you can merge the two datasets. Make sure the values in the variable names that you would like to merge are formatted the same way.
    - For example, if you want to merge every Roberts vote with his Martin-Quinn score in that particular term, you will need to make sure that both datasets format Roberts' name correctly. An easy way to do this is to put the justice variable in a `table()`.

4. Filter to justices with Martin-Quinn scores.
5. Find the mean Martin-Quinn score for each term in your dataset.
6. Find the mean decision direction for each term in your dataset. Rescale the decision direction variable so it is analagous to the Martin-Quinn score.
    - Hint: the SCOTUS database decision direction scores are currently 1 (conservative), 2 (liberal), 3 (unspecified). You want to change it to -1 (liberal), 0 (unspecified), and 1 (conservative).
7. Compare the mean Martin-Quinn scores and vote directions. Are they similar or are they different?

Brainstorm final project

Please submit the following questions in an R Markdown document. (It does not need to be in the `papaja` template.)

1. What question(s) are you interested in?
2. What are your independent and dependent variables?
3. How do you plan to measure the variables?
4. What data will you need to collect? Which dataset(s) will you use?
5. What methods will you use to analyze the data?

Submit

Email me (mshieh2@wisc.edu) the link to your `ps811-exercises` repository when you are done.