STAT 201: Midterm 1

Your name

TYPE THE HONOR CODE PLEDGE IN THE CORRESPONDING AREA:

"I, <your name>**, have neither given nor received unauthorized aid on this assignment."**

A Pell Grant is a need-based federal grant for undergraduate students paying for college. Students are automatically considered for a Pell Grant award when they submit a FAFSA.

Each case in the dataset `pell` represents the Pell Grant award to a given school in a given year.

- `state`: code for the state/territory where the school is located
- `award`: total award amount in USD
- `recipient`: total number of recipients
- `name`: name of college/university
- `session`: ID for the meeting where award amount was decided
- `year`: year of the award (1999-2017)

We also have a dataset called `states` that provide information on the states and territories of the US. The variables are:

- `Name`: name of the state/territory
- `Abbreviation`: state/territory shortcode
- `Type`: political division ("state" or "territory")

Exercise 1

Let's clean and wrangle the data a bit.

- Remove any case where at least one of the following is true:
  - The school is missing information about the award amount
  - The school had less than 1 student receiving a Pell Grant
  - The school receive an award of 0 dollars but had more than 0 students receive a Pell grant
- Create a new variable called `award_pp` that represents the award amount per person for each school in that year
- Create a new variable called `time_range` that takes the value:
  - "1999-2004" if the award was granted in 1999-2004
  - "2005-2009" if the award was granted in 2005-2009
  - "2010s" if the award was granted in the 2010s

Store your cleaned and wrangled data frame back into `pell`.

Exercise 2

Now let's bring in the information about states vs territories. Combine the two datasets into a single dataset called `pell2` that retains the observations about Pell grants awards for which we also have information about the type of political division (i.e. state or territory).

I suggest you render here!

Exercise 3

Make a visualization the shows the distribution of the award amount per person for each political division, separated by the different time/year ranges. Have informative labels and titles. Interpret what you see in terms of the center and spread both across time and across the two political divisions.

Answer:

I suggest you render here!

Exercise 4

Display a beautiful table that shows for each of the 50 U.S. states (not including Washington, D.C.) in the 2010s:

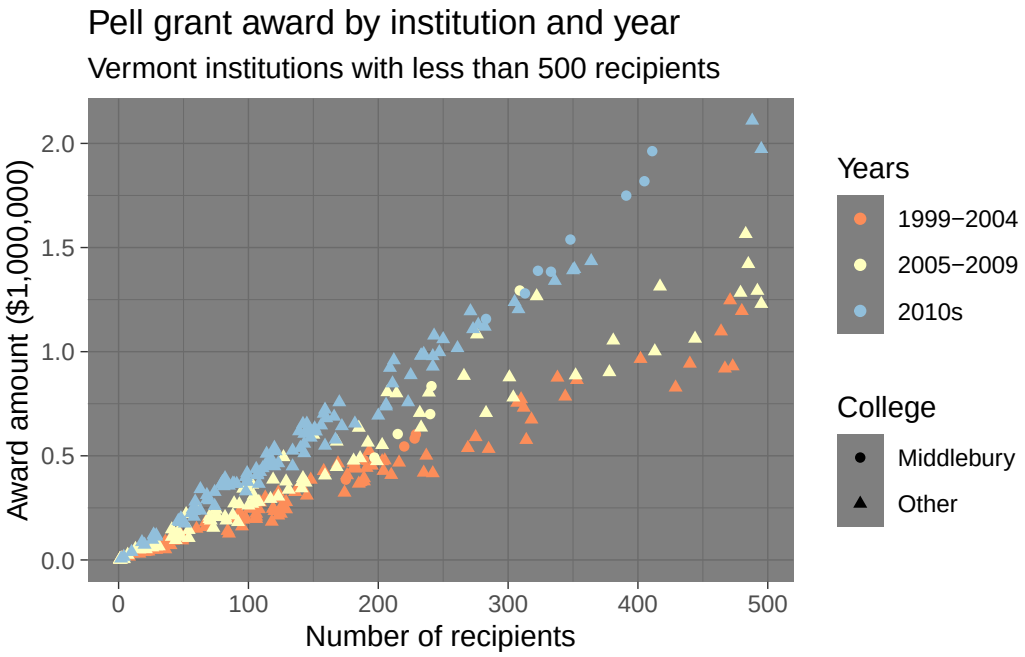
- the statewide average award per person, and
- the proportion of all money awarded to U.S. states during this time period that the state received

Your resulting table should display the state name spelled out in full, the statewide average award per person, and the proportion of total award received.

Lastly, your table should only display information for the five states that had the highest statewide average award per person.

Exercise 5

Re-create the following plot and interpret it:



Answer:

I suggest you render here!