Data Importying and Tidying Homework

These problems give you some additional practice importing and tidying data for analysis in R.

On Canvas, I’ve posted a dataset us\_avg\_tuition.xlsx that gives the average tuition cost for each state. In particular, it shows the average tuition and fees for a year of study at four-year public institutions yearly from 2004 until 2015.

1. Read the data into R and turn it into a tidy dataset.
2. Calculate the average tuition in each state across all years in the data, and sort it from highest to lowest. Do you notice any patterns?
3. Plot the average tuition in each state using a bar chart, and arrange it from highest to lowest. **Hint:** you’ll need to read the documentation for geom\_bar to figure out what options to use. And to arrange the states in your plot, you’ll need to look up the reorder option in ggplot.
4. Does the average tuition cost decrease for any state? Which states have the largest/smallest increases over this period? To adjust for the fact that states have very different tuition costs in 2004, calculate this in terms of percentage change. **Hint:** Think about how arrange() combined with first() and last() would be helpful to calculating this change.
   1. Now consider which states have the largest/smallest increases in *absolute* value (dollar amount) over this period. Are the states with the highest *percent* change the same as the states with the highest *absolute* change?
5. Let’s get a closer look at a few states and their tuition trends. Explore the data and choose five states that have an interesting pattern. Justify why you chose these states.
   1. Next, plot each state’s tuition change over time using a line graph. What can you conclude from the graph?

BONUS QUESTION:

Show the tuition trends over time for all 50 states all in one image, with each state in its own little graph. Make the line for the state that you are from (or pick a random state if you’re not from the US) a different color than all the other states.