BST 219
Core Principles of Data Science

Lecture 20: Maps continued
November 7, 2024

Recipe of the Day!

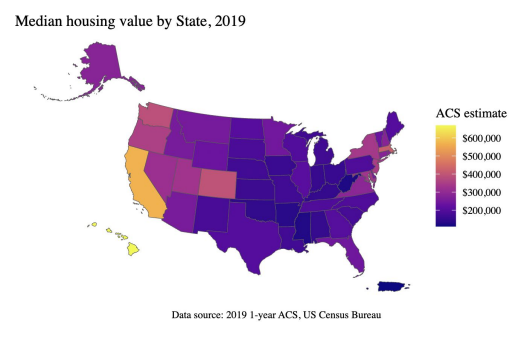
Cinnamon Rolls

Halloween Fun

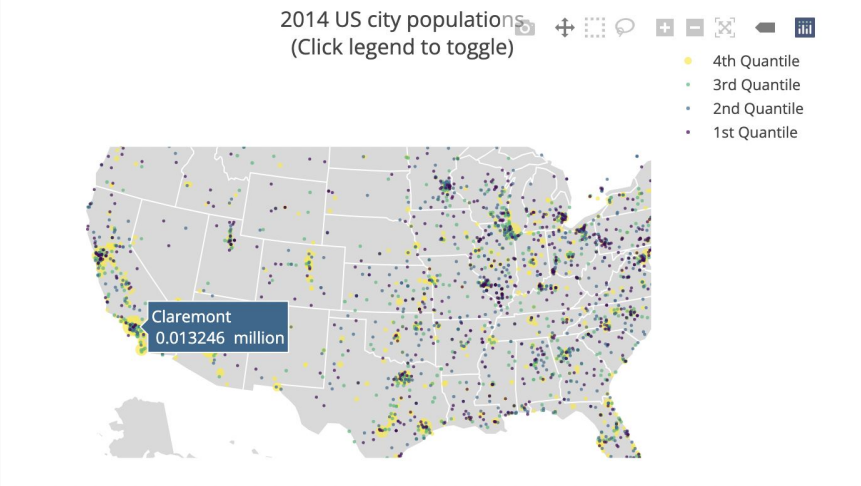
Agenda

- Announcements
  - No lab this week
  - Homework 3 is due 11/8
  - Midterm 11/8 - 11/17
  - Project preference due 11/12
    - Fill out this form

- Continue maps



Tidycensus package
Source



Bubble Maps
Source

Special Announcements

- Heather's **11/12** office hour will be moved to **12-1pm**
- The 11/14 lecture will be moved to **11/13, 12:30-2pm via Zoom**
- Lecture on **11/26** will be held via **Zoom**
  - Will be on a special topic that will be stand alone and not part of an assignment
- Heather's office hour on **11/26** will be **Zoom only**
- The TFs will not hold office hours the week of Thanksgiving (the week of 11/25)

Coding Question of the Day!

The data frame we created on Tuesday (11/5) contains a column called **`cases`**. The values of this column are the cumulative number of reported cases in a particular country on a particular day. Use the **`lag`** function to add a new column to the data frame that is the number of new cases for each day. For example, if there were 10 reported cases on January 22nd and 15 reported cases on January 23rd, the number of new cases would be 5 (15-10). Name this column **`new_cases`**.

Bonus challenge: create a new column that contains the seven-day rolling average of new cases. Name this column **`new_cases_7dayavg`**. I would suggest using the **rollmean** function from the **zoo** package.