BST 219
Core Principles of Data Science

Lecture 14: Introduction to Data Visualization
October 17, 2024

Recipe of the Day!

Crockpot White Chicken Chili

It’s sweatshirt weather!

Agenda

- Announcements
  - Lab this week!
    - Will finish last 2 questions of last week's lab
    - The rest of the time will be an office hour
  - Homework 2 is due 10/25
- Continue the visualization module!
  - Visualization principles
  - ~~Maps~~ (we'll come back to this!)
- Finish R basics and start the advanced data wrangling module!

Coding Question of the Day!

Using the **`gapminder`** dataset, make a line graph of life expectancy over time.
Include one line per country, and another line with the global average life expectancy.
Be sure to label the global average line and to make it stick out among the country-specific lines. Use the code below to calculate the crude global average for each year.

Hints:

1. Use alpha to help make the global average line stick out.
2. Use 2 **`geom_line`** layers.
3. Note that **`avg`** is a data frame.

Make sure to run this code first

```
library(dslabs)
library(dplyr)
library(ggplot2)
data("gapminder")

avg <- gapminder %>%
  group_by(year) %>%
  summarize(global_avg = mean(life_expectancy))
```