ESM 201 - Assignment 2

Background

In lecture and readings, we've discussed how crop yield can change over time, and how fertilizer use contributes to such trends. In this assignment, you will be working with two agricultural datasets from the United States Department of Agriculture. The first dataset ("grains.csv") contains data from the mid 1800s to present day on national production of corn, sorghum, barley, and oats:

year	crop	acres_planted	acres_harvested	production	yield
1936	Corn	101.96	67.83	1258.67	18.6
1936	Sorghum	13.36	2.79	30.27	10.8
1936	Barley	12.84	8.33	147.74	17.7
1936	Oats	41.93	33.65	792.58	23.6
1937	Corn	97.17	81.22	2349.43	28.9
1937	Sorghum	13.00	4.92	69.95	14.2
1937	Barley	12.35	9.97	221.89	22.3
1937	Oats	39.83	35.54	1176.74	33.1

Data from the USDA's Economic Research Service, 2019 Yearbook Tables.

- acres_planted: million acres
- acres_harvested: million acres
- production: million bushels
- yield: bushels per acre

The second dataset ("fertilizer.csv") contains information on fertilizer use for soybeans, corn, and wheat from 1964 to 2016:

year	crop	fertilizer	ppa
1964	Wheat	nitrogen	27
1964	Wheat	phosphate	27
1964	Wheat	potassium	34
1964	Soybeans	nitrogen	14
1964	Soybeans	phosphate	30
1964	Soybeans	potassium	37
1964	Corn	nitrogen	58
1964	Corn	phosphate	41
1964	Corn	potassium	41

- ppa: pounds applied per acre

Data from the USDA's Economic Research Service, 2019 Fertilizer Use.

Tasks

As with Assignment 1, we recommend doing any data visualization and analyses in R and RMarkdown; however, you're welcome to use whatever language or platform you wish. You will be turning in a pdf the completed sections to Gauchospace following this file name: **lastname_firstname_assignment02.pdf** before your section on **Wednesday the 13th or Thursday the 14th of March**.

1. Patterns in yield over time

In "grains.csv", there is a column for yield: bushels/acre. As discussed in Grassini et al. 2013 (from week 6), there are 6 possible statistical models to describe crop yield trends over time.

    a. Create a finalized plot (including title and caption) of yield vs. time for each of the grains in the dataset. (15 points)

    b. In 1-2 sentences for each grain, describe which statistical model describes the trend you see and justify your choice with the shape of the line you observe. (20 points)

    c. **Extra** Animate the plot(s) in any way you see fit. Save the plot as a gif as **lastname_firstname_extra1.gif** and upload this to Gauchospace. (5 points)

2. Fertilizer use

In "fertilizer.csv", pounds per acre for three different classes of fertilizers are tracked through time.

    a. Create a finalized plot (including title and caption) of fertilizer use vs time for each of the crops in the dataset. (15 points)

    b. Describe differences in fertilizer use for each crop. What does this suggest about limiting nutrients for each crop? (20 points)

    c. **Extra** Delve into the agricultural literature for one of these crops to figure out why these trends in fertilizer use might be occurring. Summarize your findings in 3-5 sentences and provide citations. (10 points)

3. Case study: corn yield and fertilizer use

    a. For corn, create a finalized plot (including title and caption) of yield vs fertilizer ppa, separated by the type of fertilizer. Make the color of the dots correspond to year. (15 points)

    b. Explain the relationship between yield and fertilizer you see in the data. Is there a positive linear relationship between fertilizer use and yield? (Hint: recall similar plots presented in lecture) (15 points)