

STAT 324: Discussion 1

Ralph Trane

Week 1, Spring 2020

Objective

This exercise is meant to get you started with `RStudio`, and to show you just a few of the things we can do using `R`. It is also an exact copy of Homework 1, except from the specific data set you're using, so hopefully you'll find it to be a good use of your time.

This is NOT meant to teach you everything there is to know about `R`, nor do I expect you to totally comfortable with the things you see here. It is my hope that `R` is demystified a bit, and you get a sense of how valuable it is.

Getting Started

Download the polio data `polio.csv` and the pdf-file `intro_to_R.pdf` from Canvas. Create a new folder in your STAT324 directory, and place the data there. In `RStudio`, create a new `R Markdown` document, and save it to this folder.

All the code you need is in the pdf-file. Your job is to try to make sense of it all, and copy/paste the right parts to answer the questions below.

Questions

Using the polio data (`polio.csv`) from Canvas, answer the following questions:

Question 1: What variables are included in the data set?

Question 2: How many observations are in the data set?

Question 3: print the data from Wisconsin 1993

Question 4: which state has the lowest and highest rate, respectively, when we average over the entire period?

Question 5: plot the mean rate for the entire country against the year. Does this plot bring any interesting question(s) to mind?

Question 6: the large-scale use of the polio vaccine started in February 1954. Add a vertical line to your plot at the year 1954