Homework 2
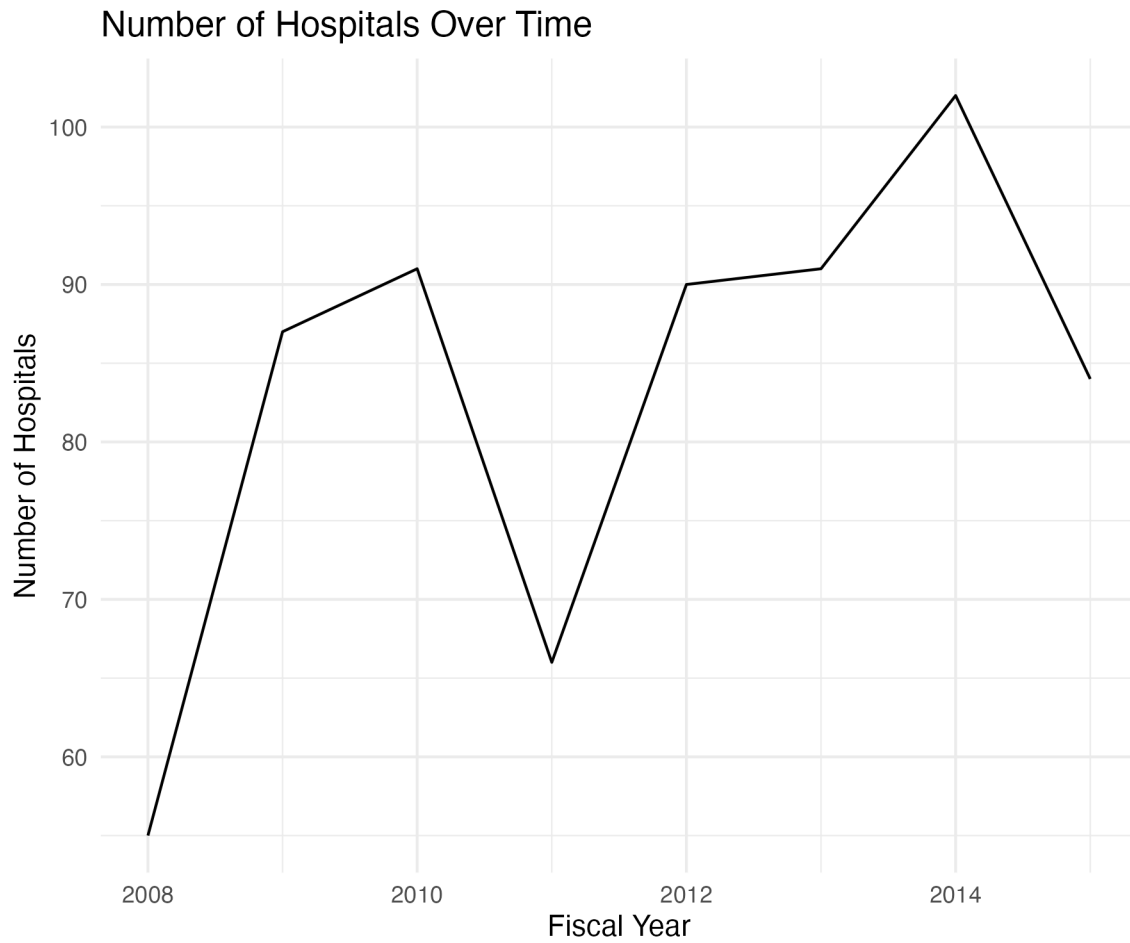
Megan Zheng

In this assignment, I wil be analyzing the HCRIS data to pull informtion from their reports and estimate ATEs. This is only a second submission, so the analysis and formatting of this file is still in progress. Also, I apologize for the lack of work in the document, there is more in the repository that I wasn't able to figure out how to add to the quarto file. We continue on!

The GitHub repository for this work is available here. Enjoy!

Question 1

How many hospitals filed more than one report in the same year? Show your answer as a line graph of the number of hospitals over time.
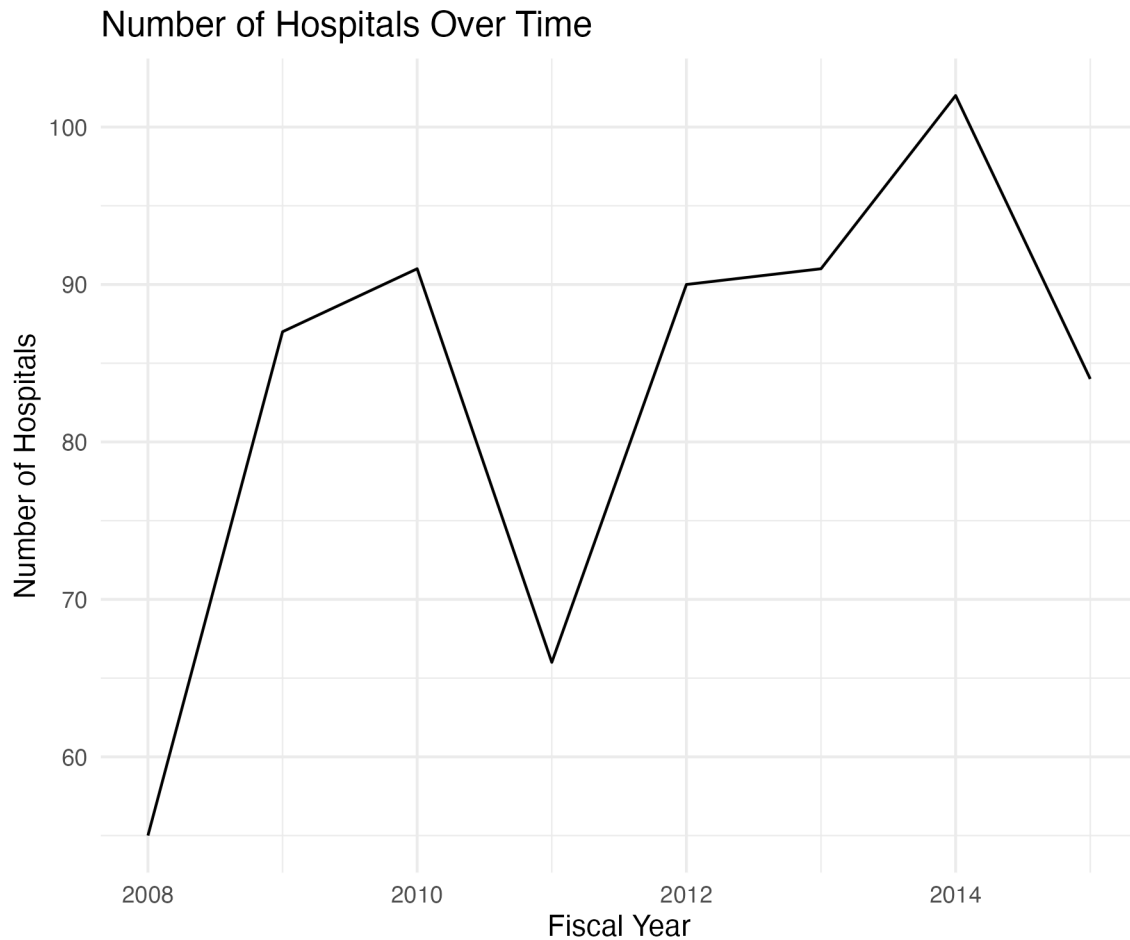
Number of Hospitals Over Time

Question 2

After removing/combining multiple reports, how many unique hospital IDs (Medicare provider numbers) exist in the data?
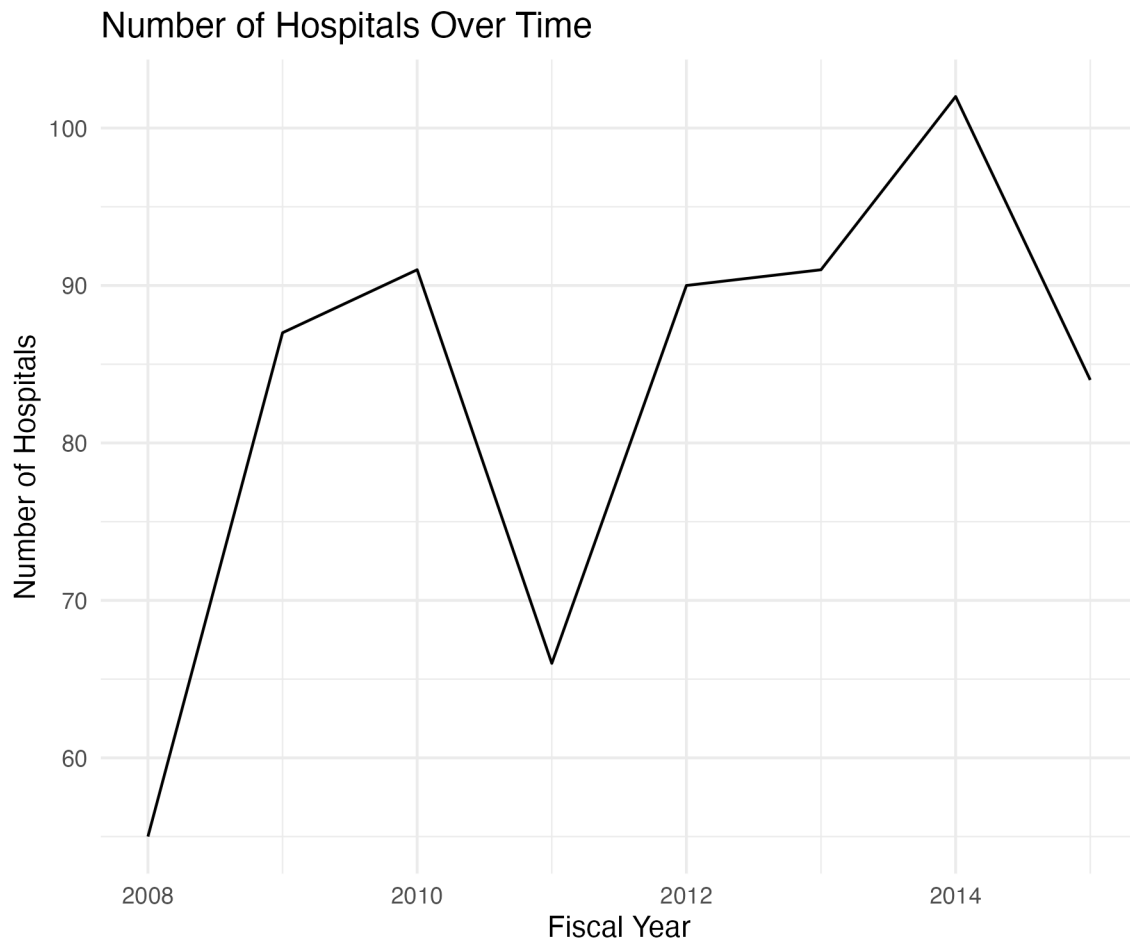
Question 3

What is the distribution of total charges (tot_charges in the data) in each year? Show your results with a "violin" plot, with charges on the y-axis and years on the x-axis. For a nice tutorial on violin plots, look at Violin Plots with ggplot2.

Number of Hospitals Over Time

Question 4

What is the distribution of estimated prices in each year? Again present your results with a violin plot, and recall our formula for estimating prices from class. Be sure to do something about outliers and/or negative prices in the data.

Number of Hospitals Over Time

Question 5

Calculate the average price among penalized versus non-penalized hospitals.

```
Penalized Hospitals Mean Price:   9543.74
```

Non-Penalized Hospitals Mean Price: 9162.81

Question 6

Split hospitals into quartiles based on bed size. To do this, create 4 new indicator variables, where each variable is set to 1 if the hospital's bed size falls into the relevant quartile. Provide a table of the average price among treated/control groups for each quartile.

Question 7

Find the average treatment effect using each of the following estimators, and present your results in a single table.

Question 8

With these different treatment effect estimators, are the results similar, identical, very different?

Question 9

Do you think you've estimated a causal effect of the penalty? Why or why not? (just a couple of sentences)

Question 10

Briefly describe your experience working with these data (just a few sentences). Tell me one thing you learned and one thing that really aggravated or surprised you.