# Homework 2

Kendall Pollard

Here is the link to my GitHub Repository: https://github.com/kpollard8/Homework-2

Here are my answers for Homework 2. I do the coding in a separate `R` script, but here is the cleaned-up version. I run the analysis separately, save the workspace with only the summary stats, figures, and tables that I need, and then load the workspace in the final qmd. My analysis file with answers and code to all the questions is available in the analysis folder.

## Duplicate Report

1. How many hospitals filed more than one report in the same year?

The answer to quesiton 1 is

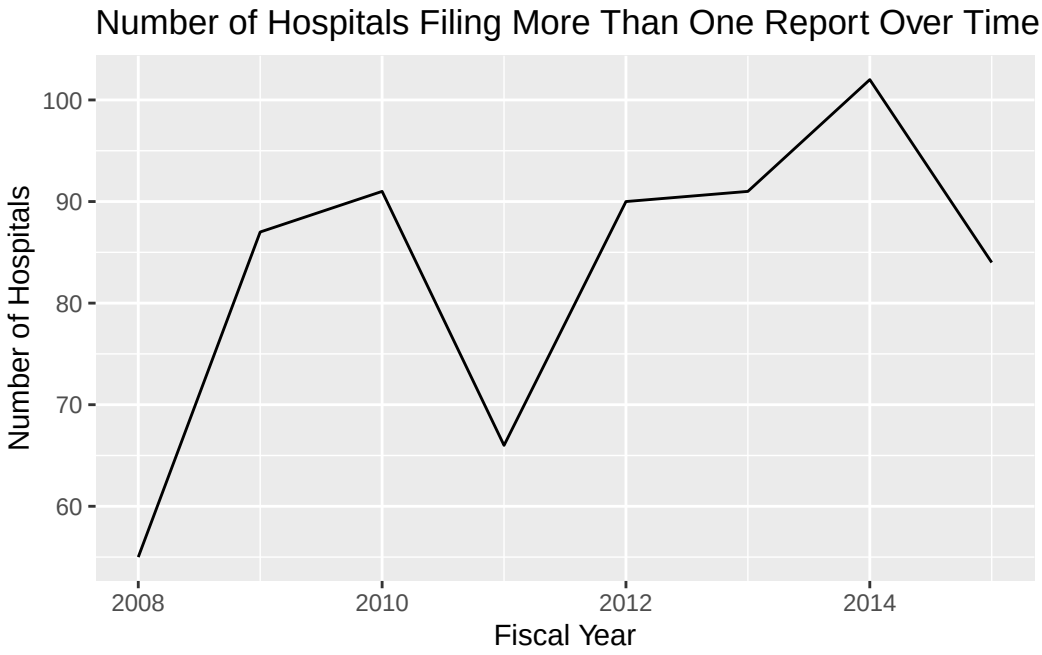


Figure 1: Hospitals with Mult Reports

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

The answer for question 2 is 6,747

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.

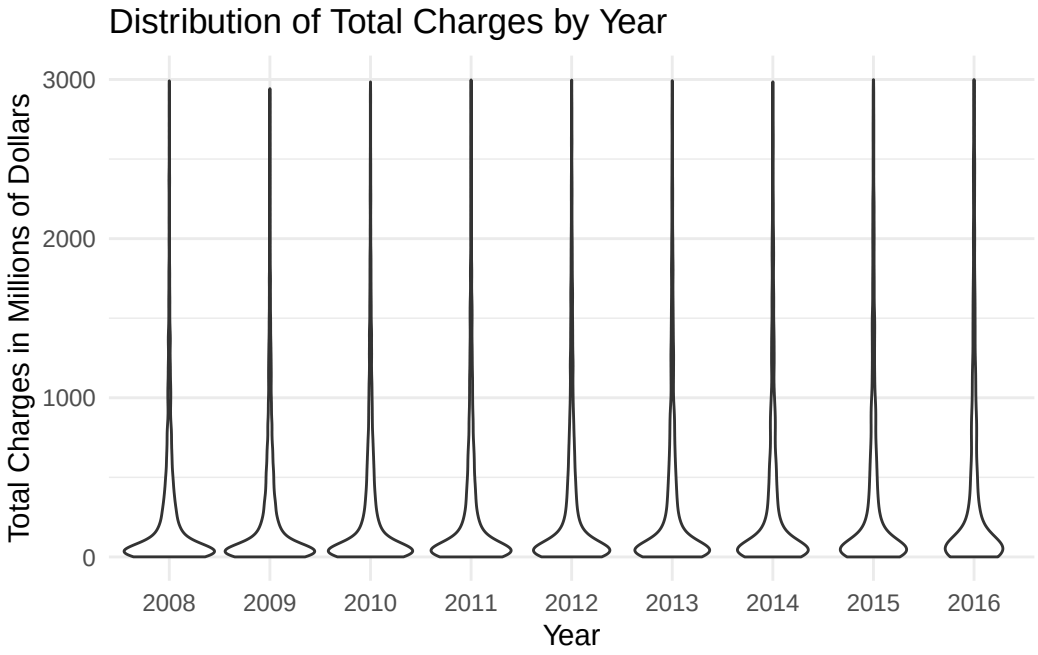


Figure 2: totalcharges

4. What is the distribution of estimated prices in each year?

I used the equation from class and filtered outliers/negative prices by removing prices lower than 0 and setting a custom upper limit.

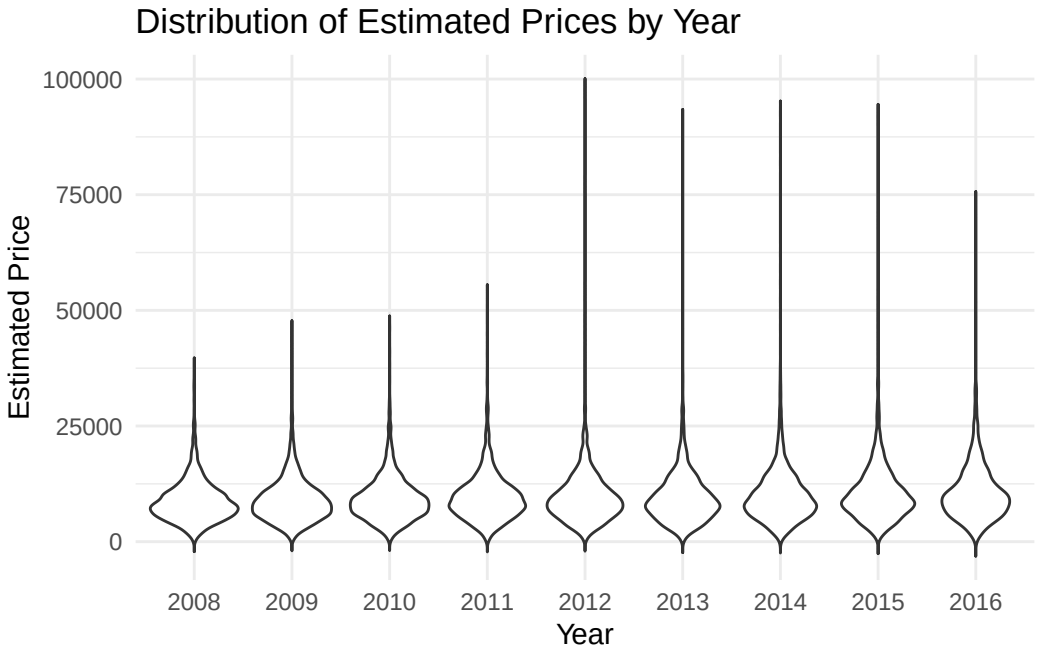


Figure 3: prices

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

Average price among penalized: 9,896.31

Average price among non-penalized: 9,560.41

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

Table 1: Bed size quartlies

| Quartile | Penalized | Non-Penalized |
|---|---|---|
| 1 | 8,318.709 | 7,684.240 |
| 2 | 8,690.891 | 8,510.959 |
| 3 | 10,127.130 | 9,856.928 |
| 4 | 12,068.479 | 12,355.606 |

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

Table 2: ATE 4 different ways

| Estimator | ATE |
|---|---|
| Nearest Neighbor (Inverse Variance Distance) | 190.1416 |
| Nearest Neighbor (Mahalanobis Distance) | 190.1416 |
| Inverse Propensity Weighting | 190.1416 |
| Simple Linear Regression | 190.1416 |

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

The numbers are all identical. This means that even with these different esitmators, the ATE is clearly around 190. This outcome is interesting because that means that all these estimators measured the exact same ATE, although I hypothesized that they would be similar with some variance.

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

I don't think we have established a causal effect because it is just bed size; hosptials of the same bed size could have differing other variables we are using various estimators to find an average treatment effect, not necessarily a causal effect. With ATEs that are statistically significant, we can say there is very high correlation/association, but still it's difficult to say it was directly causal. However, now that all the ATEs are the exact same, I'm wondering if we can say a causal effect because it has been reinforced 4 times.

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.

This time around, I was definitely more comfortable working with VSCode (Quarto, the setup, etc.), but there were a few hiccups. In the second attempt, moving from submission 1 to submission 2 was a little challenging with making sure my Quarto documents and the workspaces they were accessing were aligned.

I learned how to create violin plots which were not as hard as I thought they'd be. One thing that aggravated me was figuring out how to call all the variables/graphs/tables into the Quarto document. It took a lot of time to figure out how to correctly call them.