Homework2 Submission2

https://github.com/modugbe/homework2

Moyo Odugbemi

Question 1

Hospitals that filed more than one report in the same year is plotted. The data is first combined and then the unique duplicates are filtered out for this graph.
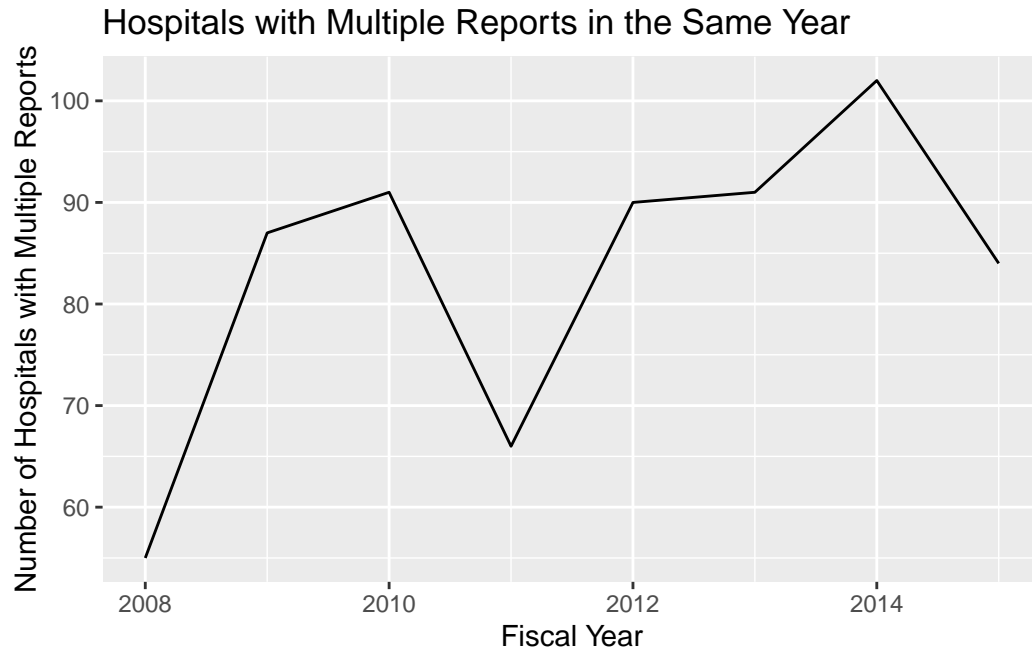

Figure 1: Hospitals with duplicate reports

Question 2

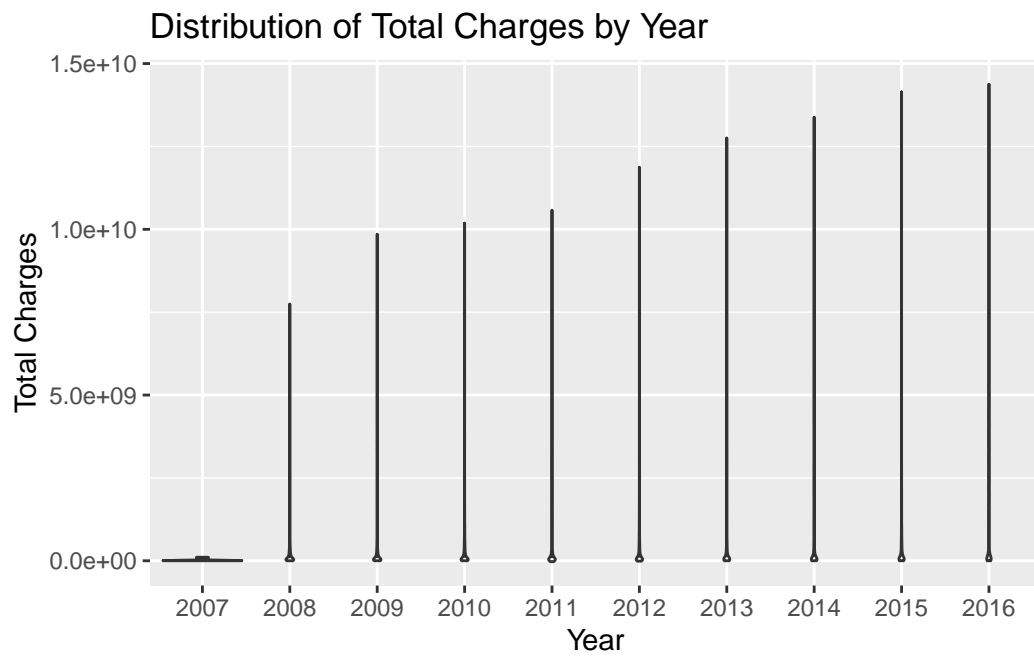The number of unique hospital IDs: 6486 .

Question 3

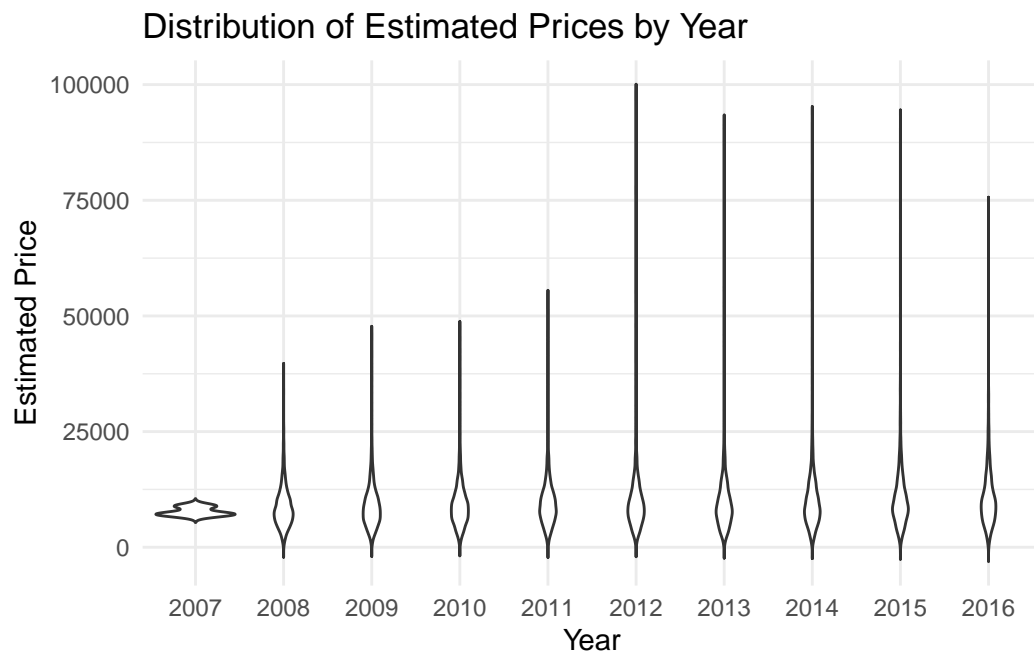Figure 2: Distribution of total charges in each year

Question 4

Figure 3: Distribution of estimated prices in each year

Question 5

Table 1: Average Price by Penalty

Penalty	Mean_Price
Penalized	9,896.31
Non-Penalized	9,560.41

Question 6

Table 2: Average Price by Quartile

quartile_1	quartile_2	quartile_3	quartile_4	penalty	average_price
0	0	0	1	FALSE	12,355.606
0	0	0	1	TRUE	12,068.479
0	0	1	0	FALSE	9,856.928
0	0	1	0	TRUE	10,127.130
0	1	0	0	FALSE	8,510.959
0	1	0	0	TRUE	8,690.891
1	0	0	0	FALSE	7,684.240
1	0	0	0	TRUE	8,318.709

Question 7

Estimate for nearest neighbor (variance)

Quarto is doing something weird where I cannot get it to display the results of the ATE on the pdf. When ran in the terminal, the estimate is 240.45

Estimate for nearest neighbor (mahalanobis)

When ran in the terminal, the estimate is 240.45

Estimate for propensity score

Estimate for linear regression

Question 8

The treatment effect calculated with inverse variance and Mahalanobis are identical. I am struggling with IPW and the dummy variables for linear regression

Question 9

I don't think this is a causal effect of the penalty because it is estimated based on beds only and other variables are not controlled for.

Question 10

Learning new packages to install for matching (update: this package is the one that messes with the select function). Running the R-code for some reason doesn't diplay the graphs until I preview the qmd.