Homework2 Submission3

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.

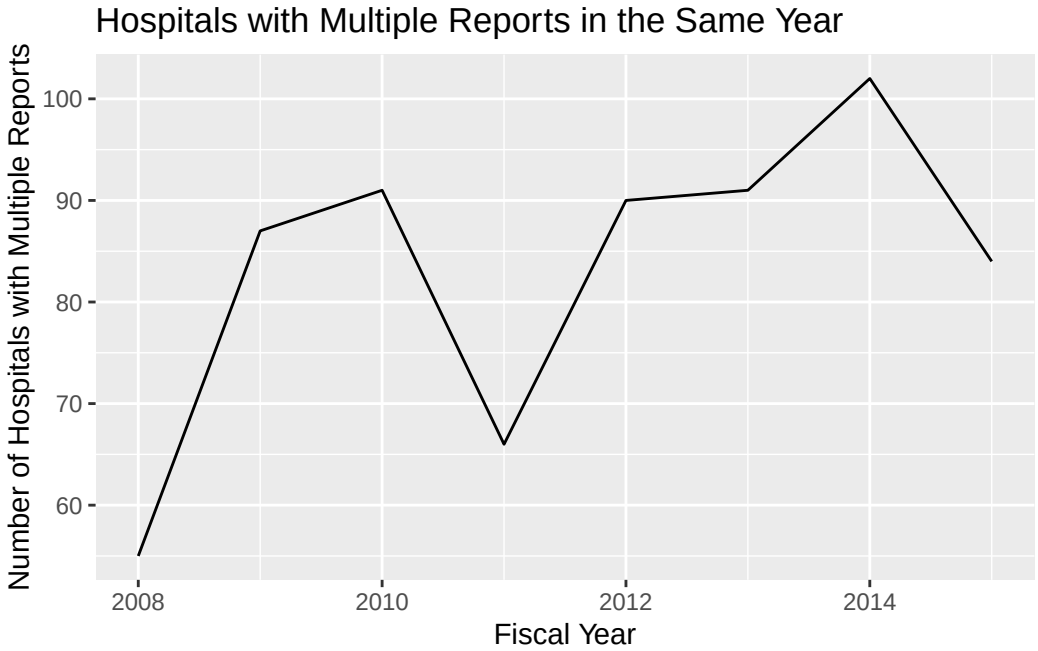

Figure 1: Hospitals with duplicate reports

Question 2

The number of unique hospital IDs: 6486 .

Question 3

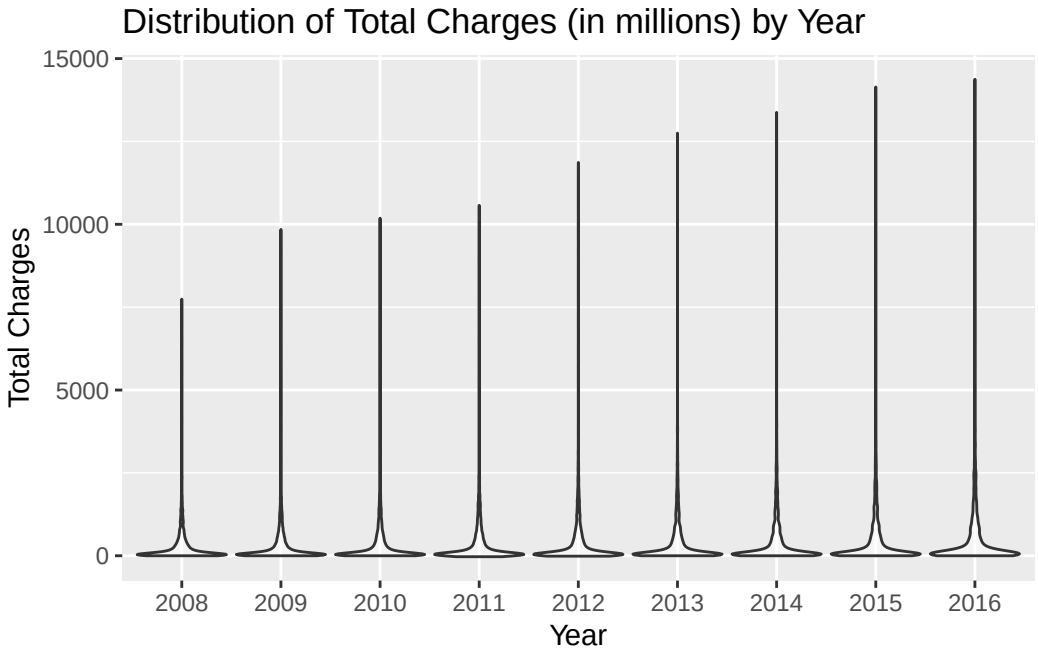

Figure 2: Distribution of total charges in each year

Question 4

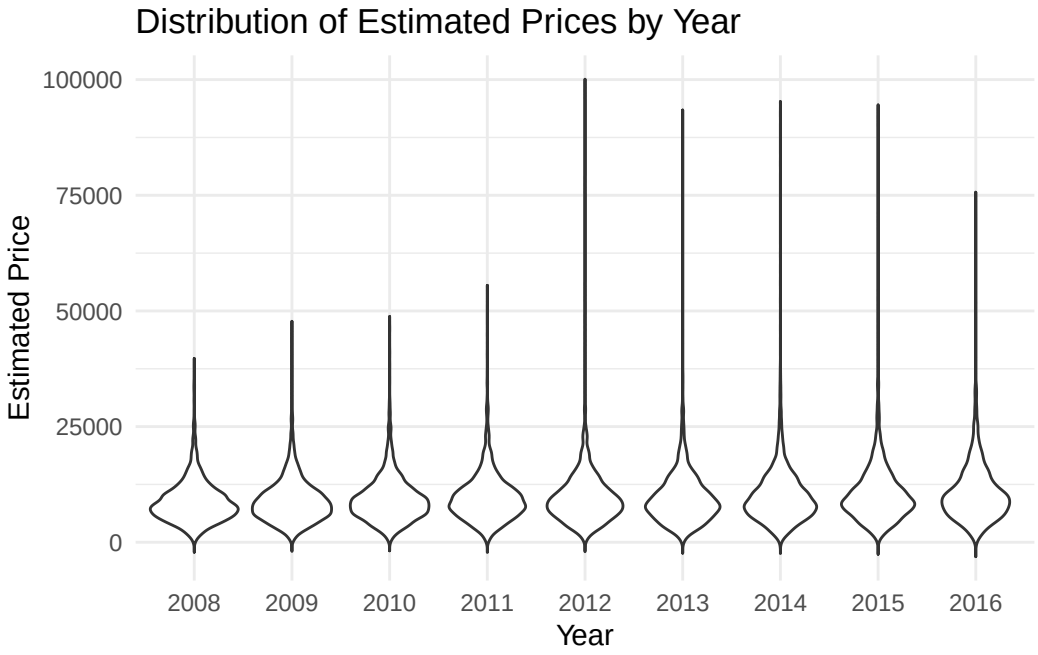

Figure 3: Distribution of estimated prices in each year

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

bed_size_quartile	penalty	average_price
1	FALSE	7,684.240
1	TRUE	8,318.709
2	FALSE	8,510.959
2	TRUE	8,690.891
3	FALSE	9,856.928
3	TRUE	10,127.130
4	FALSE	12,355.606
4	TRUE	12,068.479

Question 7

Estimate for nearest neighbor (variance)

Table 3: Average Effect Estimates Using different estimators

Estimator	ATE
Nearest Neighbor (Inverse Variance Distance)	199.5281
Nearest Neighbor (Mahalanobis Distance)	199.5281
Inverse Propensity Weighting	199.5281
Simple Linear Regression	199.5281

Question 8

The treatment effects are identical.

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.