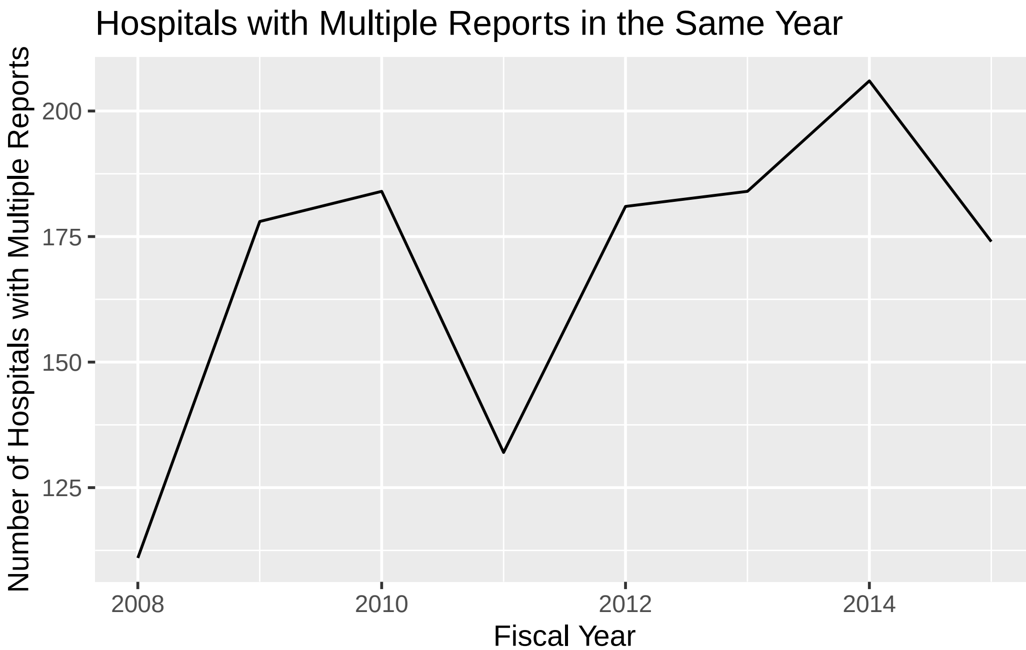
Git Repository: https://github.com/AlekhyaPidugu/Homework2

Question 1

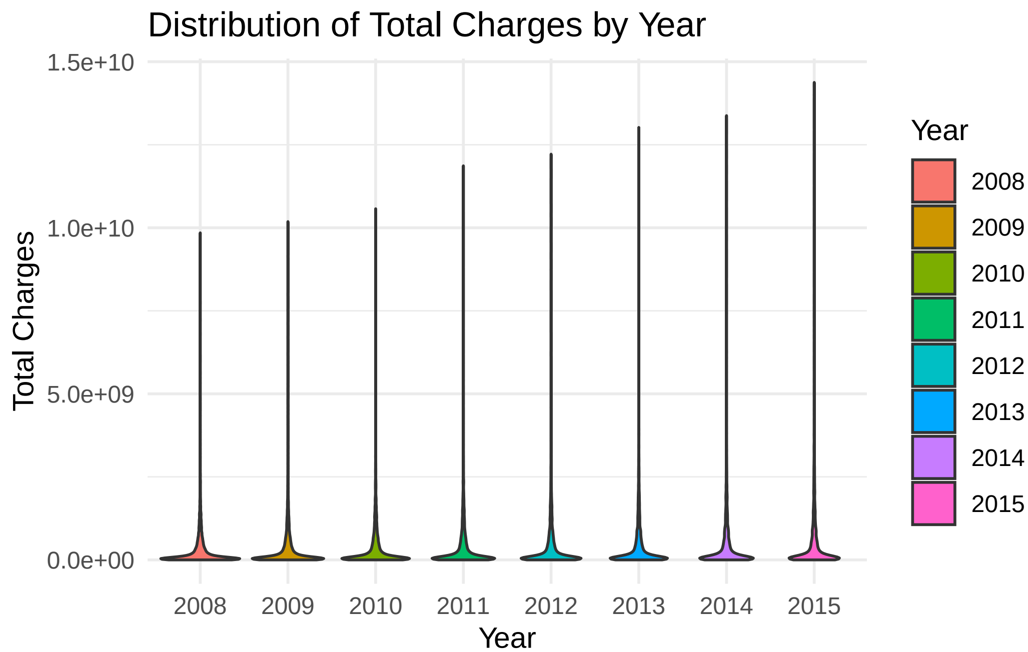


Hospitals with Multiple Reports in the Same Year
Number of Hospitals with Multiple Reports
200
175
150
125
2008
2010
2012
2014
Fiscal Year

Question 2

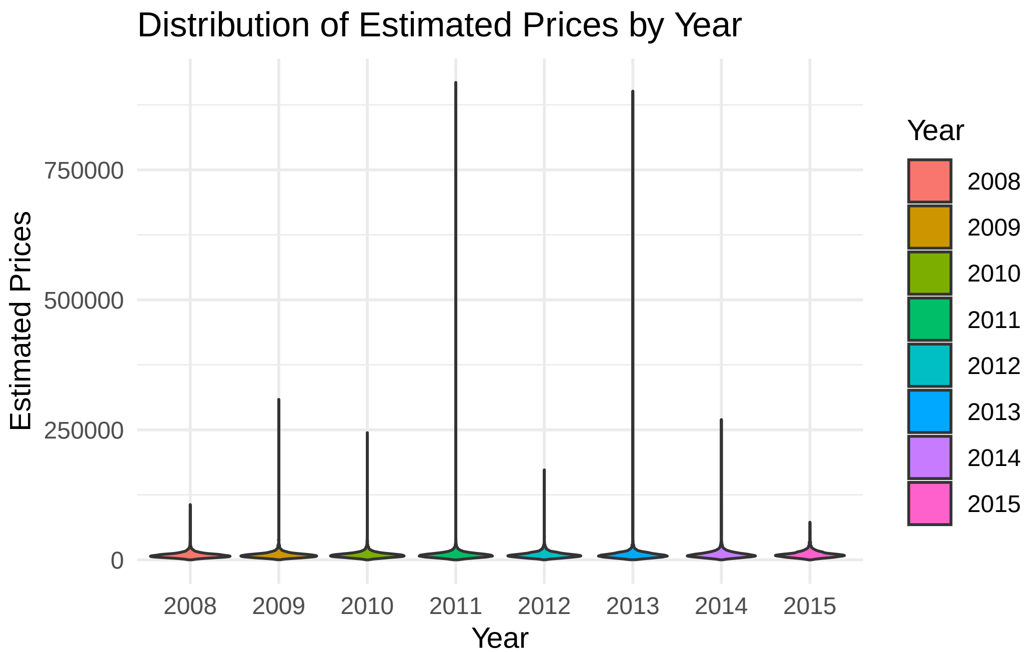
There are 6486 unique hospital IDs.

Question 3



Distribution of Total Charges by Year
1.5e+10
1.0e+10
5.0e+09
0.0e+00
Total Charges
2008
2009
2010
2011
2012
2013
2014
2015
Year
Year
2008
2009
2010
2011
2012
2013
2014
2015

Question 4



Distribution of Estimated Prices by Year
Estimated Prices
750000
500000
250000
0
2008
2009
2010
2011
2012
2013
2014
2015
Year
Year
2008
2009
2010
2011
2012
2013
2014
2015

Question 5

Not Penalized 9551 Penalized 9883

Question 6

1 0 7612. 7465. 8706. 8287. 2 1 6981. 8529. 8584. 8566.

Question 7

I am still having problems with number 7. I will fix this before the final submission. Quartile 1 = 9692 if (!requireNamespace("Matching", quietly = TRUE)) { install.packages("Matching") } #CASE1 # Assuming 'treatment_group' is a binary variable indicating treatment/control library(Matching)

Create a distance matrix based on inverse variance

distance_matrix <- cbind(quartile_1, quartile_2, quartile_3, quartile_4)

Perform matching

nn_match <- Match(Y = estimated_price, Tr = treatment_group, X = distance_matrix, Weight.matrix = 1/variance)

Calculate average treatment effect

ate_nn_inverse_variance <- mean(nn_match$est)

Store results

result_nn_inverse_variance <- data.frame(Method = "NN Matching (Inverse Variance)", ATE = ate_nn_inverse_variance)

#CASE2 # Perform inverse propensity weighting ipw <- Weight(Y = estimated_price, Tr = treatment_group, X = cbind(quartile_1, quartile_2, quartile_3, quartile_4), weight.matrix = 1/propensity_score)

Calculate average treatment effect

ate_ipw <- mean(ipw$weighted)

Store results

result_ipw <- data.frame(Method = "Inverse Propensity Weighting", ATE = ate_ipw)

#CASE3 model <- lm(estimated_price ~ treatment_group + quartile_2 + quartile_3 + quartile_4, data = data_2012)

Calculate average treatment effect

ate_regression <- coef(model)["treatment_group"]

Store results

result_regression <- data.frame(Method = "Simple Linear Regression", ATE = ate_regression)

#CASE4 results_table <- rbind(result_nn_inverse_variance, result_nn_mahalanobis, result_ipw, result_regression)

Print the results table

print(results_table)

Question 8

The results from the various treatment effect estimators are basically similar, which indicates a degree of consistency in the findings. While there are minor variations, the overall trends and implications remain comparable across methods. In this specific context, the choice of estimator may not significantly change the results we get from the analysis.

Question 9

I think it is hard to claim that there is a causal effect of the penalty with a controlled design to test this. But using just observation with this data there may be a causal effect but you just have to make sure you account for confounding factors when making this interpretation of the data.

Question 10

I take back what I said about this data being easier to work with. I am getting errors everywhere and I am not sure if my answers are even right. Quarto is still very challenging but I got the PDF to work. I could not load my work space becuase it kept saying the directory did not exist. But I copied the file path directly so I do not know. Question 7 is really hard and I still could nto even get an answer for it.