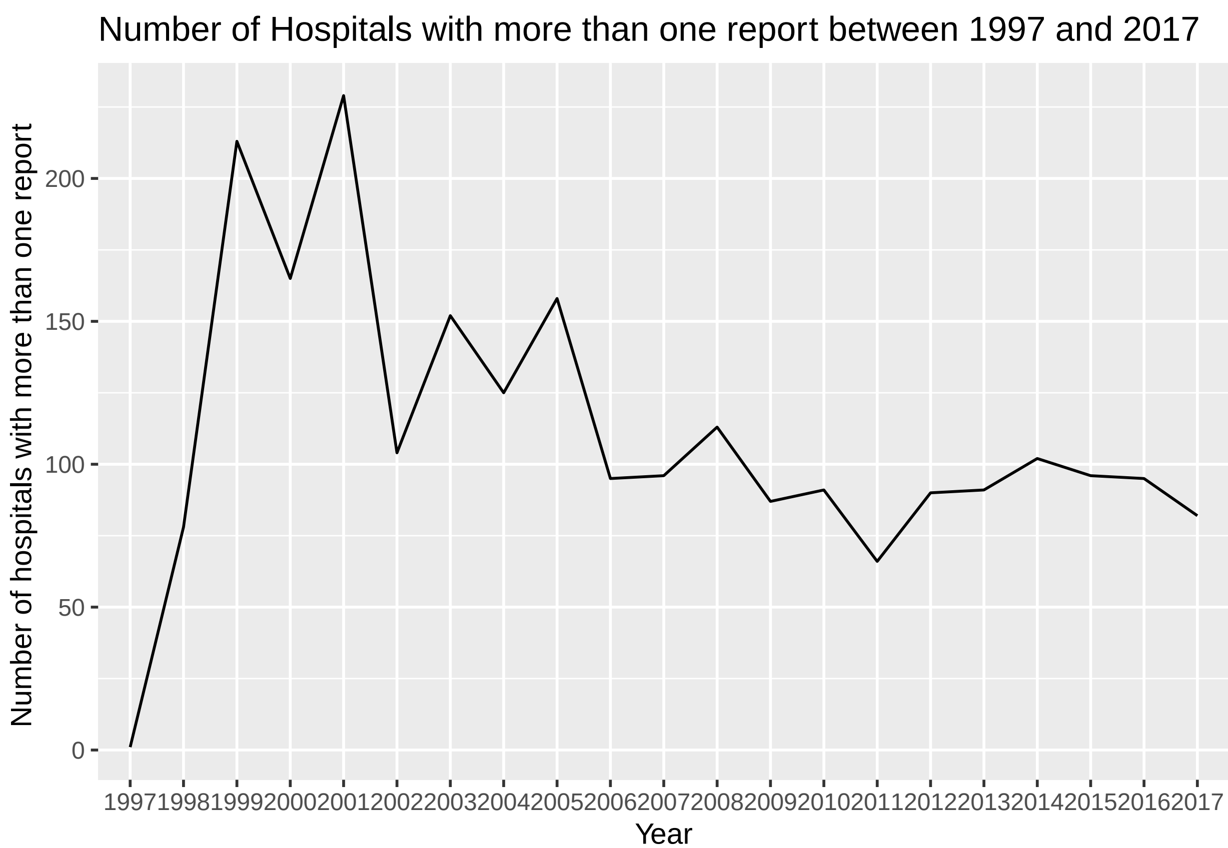
Joseph-J-hwk2-3

Justin Joseph

2023-02-15

Contents

0.0.1 1. Number of hospitals that filed more than one report in the same year

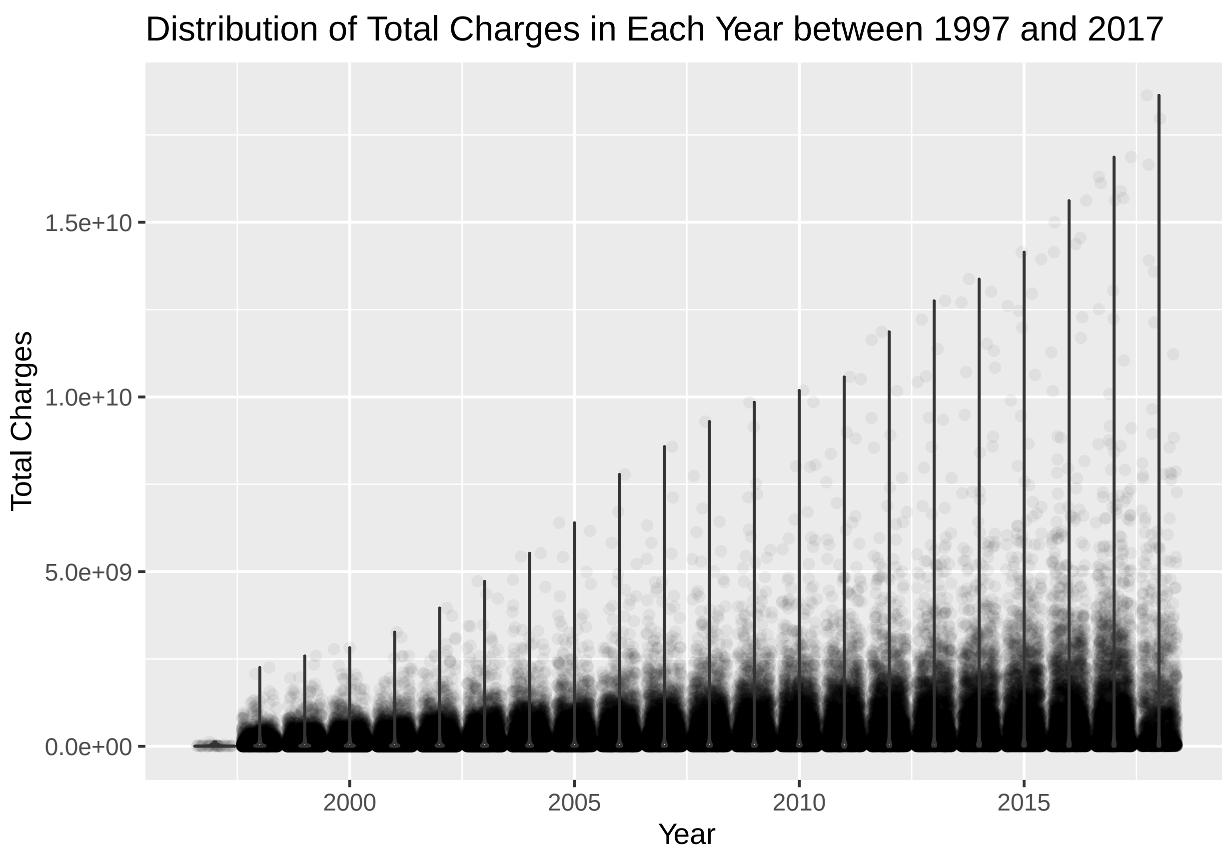


0.0.2 2. Number of unique hospitals

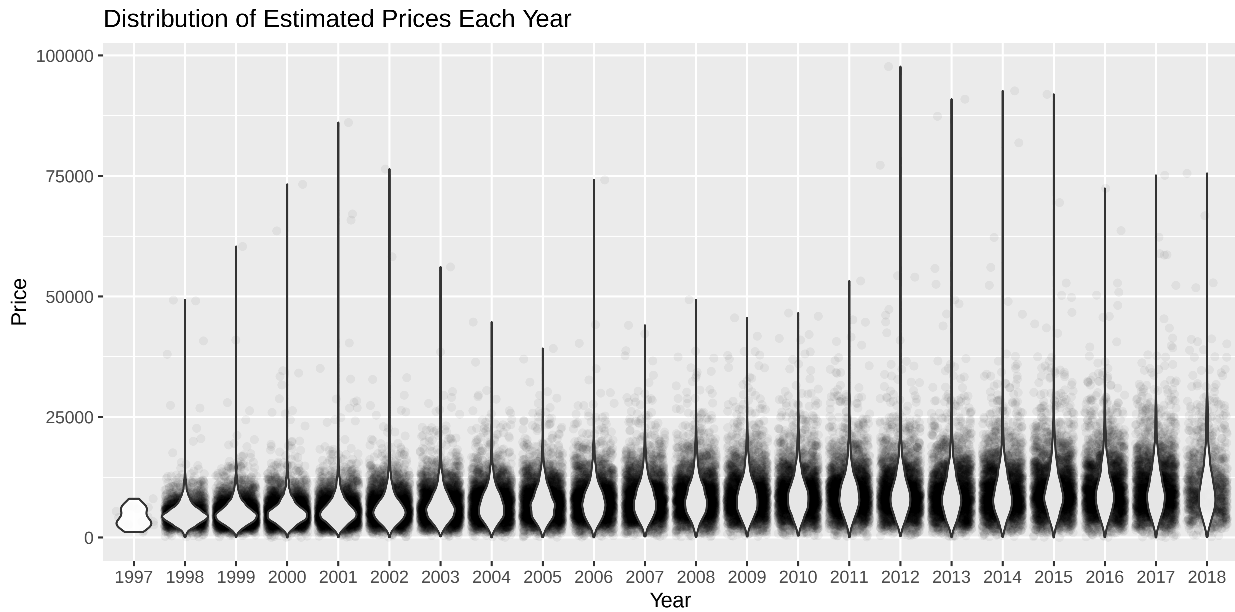
```
## [1] 9323
```

After removing/combining multiple reports, there are 9,323 unique hospitals that exist in the data set

0.0.3 3. Distribution of total charges



0.0.4 4. Distribution of estimated prices in each year



0.0.5 5.

Table 1: The average price among penalized versus non-penalized hospitals

penalty	Average_Price
FALSE	9560.413
TRUE	9896.308

0.0.6 6.

Table 2: The average price among treated/control groups dependent among bed size stratafied by quartiles

group	Q1	Q2	Q3	Q4	mean_price
Control	0	0	0	1	12367.332
Control	0	0	1	0	9848.404
Control	0	1	0	0	8525.607
Control	1	0	0	0	7696.470
Treatment	0	0	0	1	12068.479
Treatment	0	0	1	0	10132.315
Treatment	0	1	0	0	8721.033
Treatment	1	0	0	0	8286.338

0.0.7 7. Find the average treatment effect using each of the following estimators

	Inverse Variance	Mahalanobis	Inverse Propensity	Simple Regression
Average Treatment Effect	193.8313	193.8313	193.8313	179.1706

0.0.8 8.

The first three models of the estimate of the average treatment effect using Inverse variance, mahalanobis, and inverse propensity weighting are equivalent. However, simple regression provides different estimate because of the difference in comparison to matching vs a regression in modeling the effect of penalty on price. The first three use in simple terms a weighing mechanism/matching in order to control for bed size (some obersverations may carry more weight as a match than others) while simple regression does not have a special weight for each observation and just controls for bed size.

0.0.9 9.

Yes I believe I was able to estimate a causal inference. In the first three estimators we are doing are best to match the control group observations to the treatment observations so we can deduct to the best of our ability the effect being in the penalty group has on our price and limiting the confounding variable of bed size. Bigger bed size hospitals may just have larger prices or more penalties thus by matching based on bed size we are able to make a much stronger inference to the association between penalty and price. However, even though we matched for bed size there could be other factors that could effect the average treatment effect of price based on penalty for example location that may effect the independent variable and/or dependent variable. Ultimately, since we can never really observe the counter factual world where individuals who are in the treatment group are the same as the control group and observe the outcome this is pretty good estimate and attempt, even though their may be other confounding variables, at understanding the effect of penalty on price.

0.0.10 10.

So my experience at the end of this project is much better than when i started. When i began this journey with this module I was very lost concerning the material and understanding what the estimators were. However, after going to office hours I feel much more confident about them and could explain to a friend. In terms of the code, cleaning up the code was much easier for me and in the beginning I did have some trouble with the estimators but through group work I was able to find my mistakes. Lastly, I am doing much better at formatting by having my rmarkdown only call variables in stead of running whole lines of code which is awesome.