

ECON 470 Homework 2

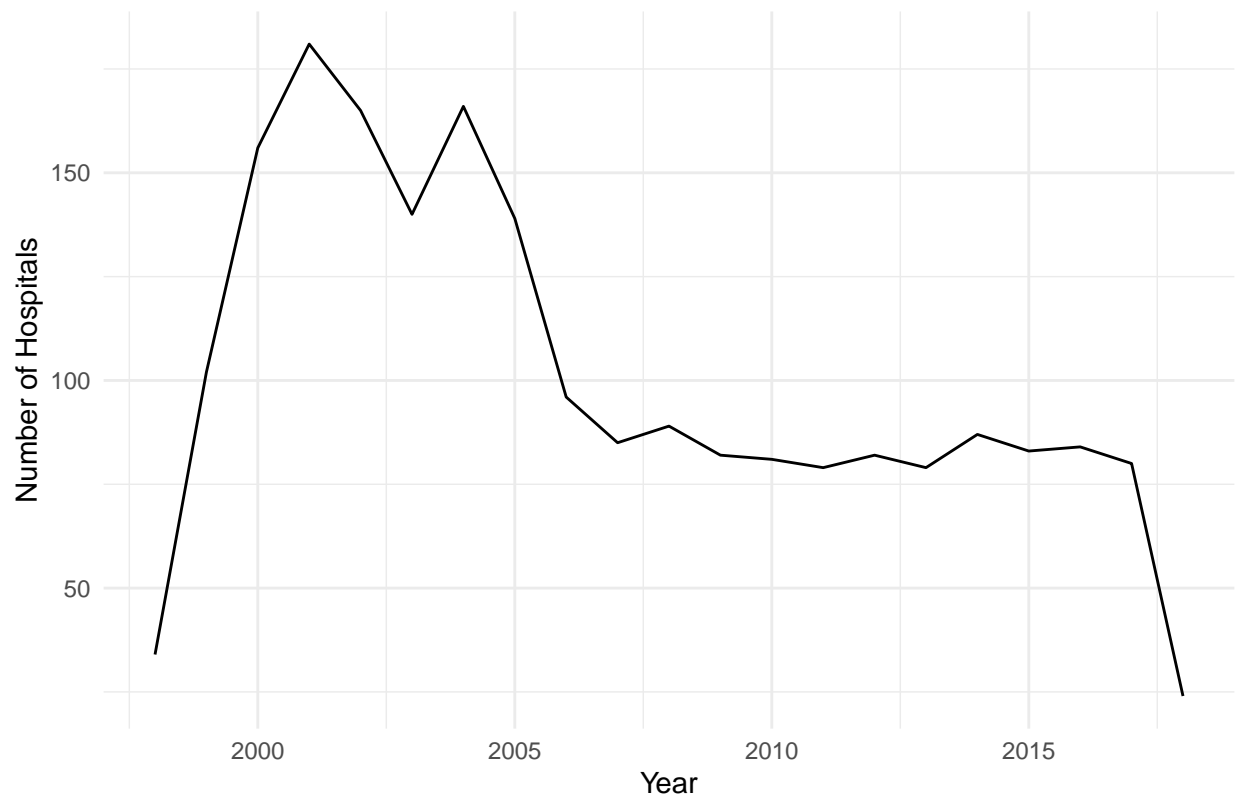
Virginia Sanson

2023-02-14

Question 1

2,114 hospitals filed more than one report in the same year, from 1997 to 2018.

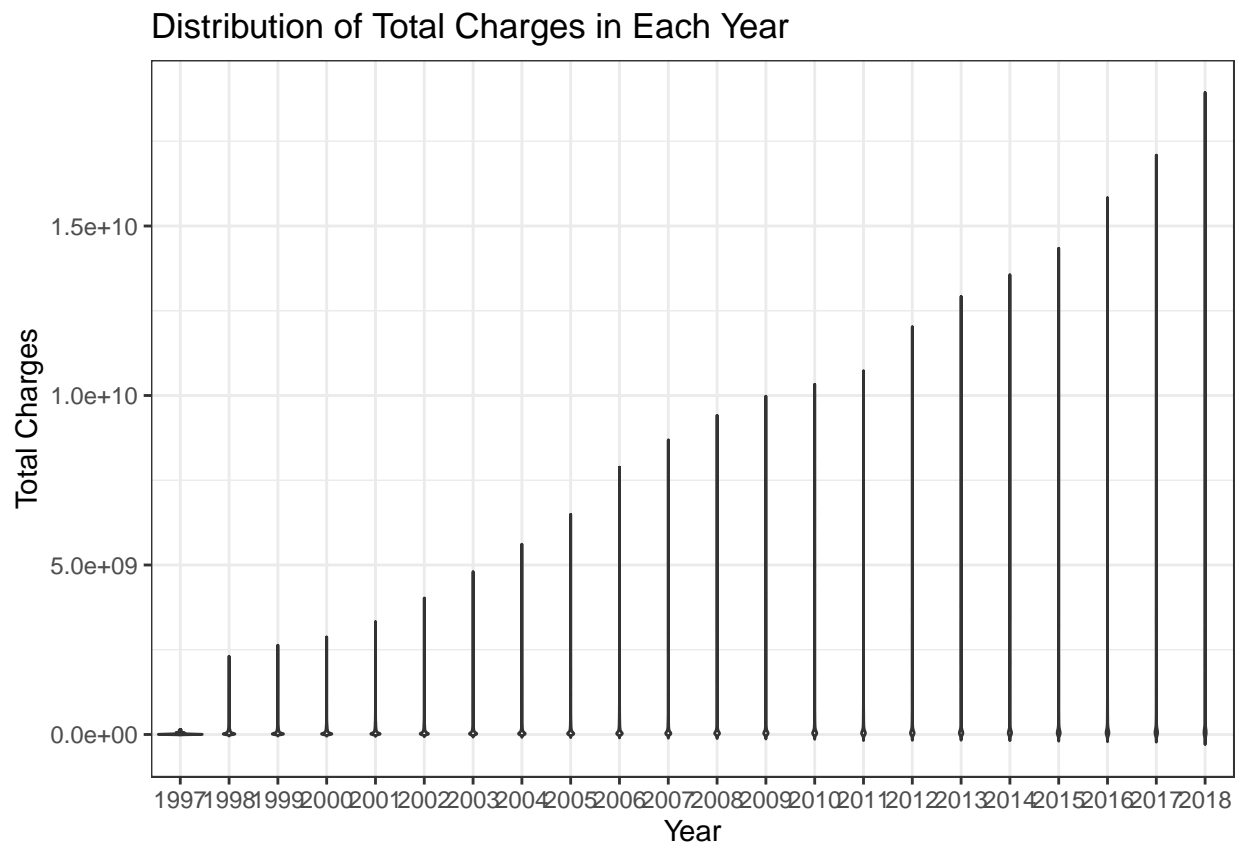
Hospitals with 2+ Filed Reports in One Year



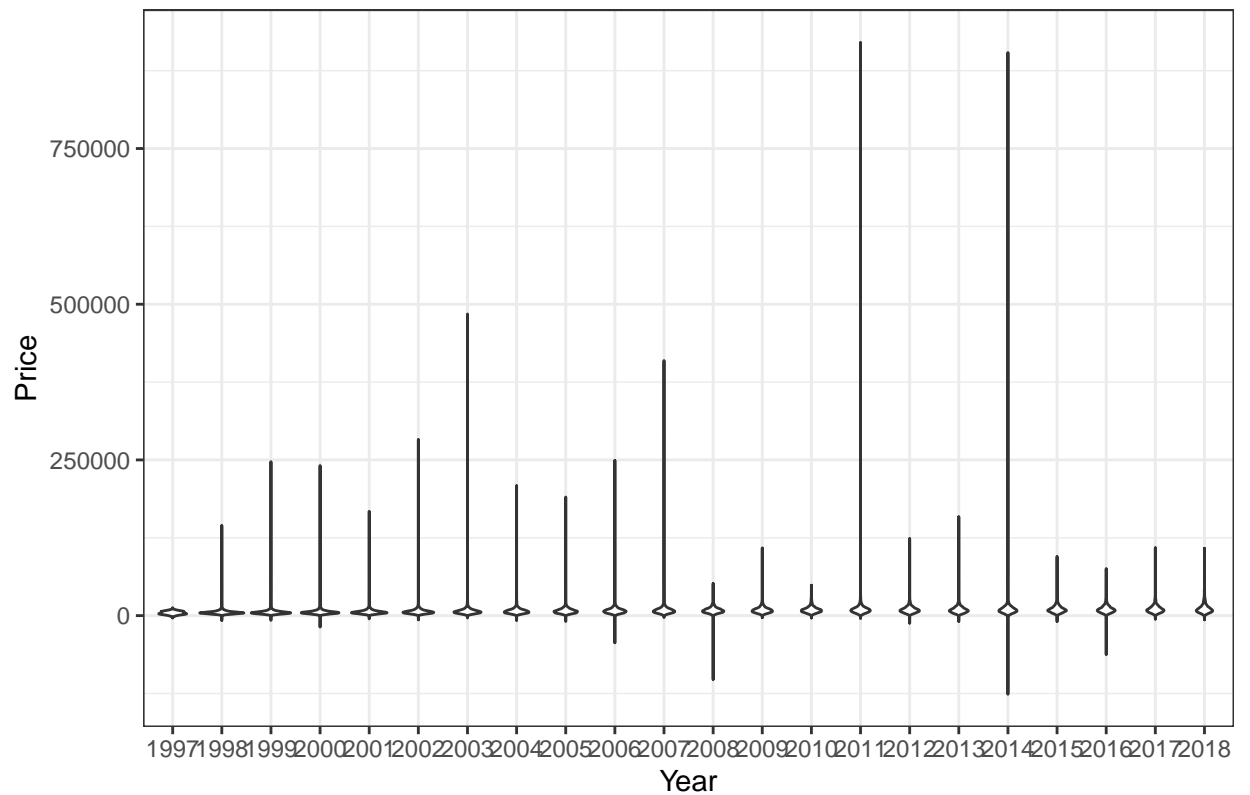
Question 2

The total number of unique hospital IDs in the dataset is 9323.

Question 3



Distribution of Estimated Prices in Each Year



Question 4

Question 5

```
## # A tibble: 2 x 2
##   penalty price
##   <dbl> <dbl>
## 1      0 9752.
## 2      1 10235.
```

Question 6

```
## # A tibble: 8 x 3
## # Groups:   quartile [4]
##   quartile penalty avg_price
##   <int> <dbl> <dbl>
## 1      1      0      NaN
## 2      1      1      NaN
## 3      2      0  10032.
## 4      2      1   6680.
## 5      3      0   8004.
## 6      3      1  10079.
## 7      4      0  11076.
## 8      4      1  11264.
```

Question 7

Part A Estimate... 325.68 AI SE..... 682.79 T-stat..... 0.47699 p.val..... 0.63337

Original number of observations..... 673 Original number of treated obs..... 85 Matched
number of observations..... 673 Matched number of observations (unweighted). 24812

##	Length	Class	Mode
## est	1	-none-	numeric
## se	1	-none-	numeric
## est.noadj	1	-none-	numeric
## se.standard	1	-none-	numeric
## se.cond	1	-none-	numeric
## mdata	4	-none-	list
## index.treated	24812	-none-	numeric
## index.control	24812	-none-	numeric
## index.dropped	0	-none-	NULL
## weights	24812	-none-	numeric
## orig.nobs	1	-none-	numeric
## orig.wnobs	1	-none-	numeric
## orig.treated.nobs	1	-none-	numeric
## nobs	1	-none-	numeric
## wnobs	1	-none-	numeric
## caliper	0	-none-	NULL
## ecaliper	0	-none-	NULL
## exact	0	-none-	NULL
## ndrops	1	-none-	numeric
## ndrops.matches	1	-none-	numeric
## MatchLoopC	124060	-none-	numeric
## version	1	-none-	character
## estimand	1	-none-	character

Part B Estimate... 325.68 AI SE..... 682.79 T-stat..... 0.47699 p.val..... 0.63337

Original number of observations..... 673 Original number of treated obs..... 85 Matched
number of observations..... 673 Matched number of observations (unweighted). 24812

##	Length	Class	Mode
## est	1	-none-	numeric
## se	1	-none-	numeric
## est.noadj	1	-none-	numeric
## se.standard	1	-none-	numeric
## se.cond	1	-none-	numeric
## mdata	4	-none-	list
## index.treated	24812	-none-	numeric
## index.control	24812	-none-	numeric
## index.dropped	0	-none-	NULL
## weights	24812	-none-	numeric
## orig.nobs	1	-none-	numeric
## orig.wnobs	1	-none-	numeric
## orig.treated.nobs	1	-none-	numeric
## nobs	1	-none-	numeric
## wnobs	1	-none-	numeric
## caliper	0	-none-	NULL
## ecaliper	0	-none-	NULL

```
## exact          0 -none- NULL
## ndrops         1 -none- numeric
## ndrops.matches 1 -none- numeric
## MatchLoopC     124060 -none- numeric
## version        1 -none- character
## estimand       1 -none- character
```

Part C

Received an error: “error occurred in group 1: provider_number =”010001”, year = 2012.”

Attempted this code: `logit.reg <- glm(penalty ~ obs_2012_fquartile_1 + obs_2012_fquartile_2 + obs_2012_fquartile_3 + obs_2012_fquartile_4, data = final.hcris.data, family = binomial)` `final.hcris.data <- final.hcris.data %>% mutate(ps = predict(logit.reg, type = ‘response’)) %>% filter(ps>0 & ps<1)`

`final.hcris.data <- final.hcris.data %>% mutate(ipw = case_when(penalty == 1 ~ 1/ps, penalty == 0 ~ 1/(1-ps), TRUE~NA_real_))` `view(final.hcris.data)`

`mean.t1 <- final.hcris.data %>% filter(penalty==1) %>% dplyr::select(price, ipw) %>% summarize(mean_y=weighted.mean(price, w=ipw))` `mean.t0 <- final.hcris.data %>% filter(penalty==0) %>% dplyr::select(price, ipw) %>% summarize(mean_y=weighted.mean(price, w=ipw))` `mean.t1mean_y - mean.t0mean_y` `reg.ipw <- lm(price ~ penalty, data=final.hcris.data, weights=ipw)` `reg.ipw`

Part D

Attempted this code: `obs_2012 <- obs_2012 %>% filter(penalty==1)` `reg1 <- lm(price ~ beds + mcaid_discharges + ip_charges + mcare_discharges + tot_mcare_payment, data=reg1.dat)` `reg0.dat <- lp.vars %>% filter(penalty==0)` `reg0 <- lm(price ~ beds + mcaid_discharges + ip_charges + mcare_discharges + tot_mcare_payment, data=reg0.dat)` `pred1 <- predict(reg1,new=lp.vars)` `pred0 <- predict(reg0,new=lp.vars)` `mean(pred1-pred0)`

Question 8

While I do not have the answers to 7c or 7d, I believe the results will return as similar but not identical.

Question 9 I predict that we cannot be estimating a causal effect of the penalty, and more of a correlation instead.

Question 10 My experience working with this data set was a lot better than last homework’s, however some parts are definitely still time-consuming. I learned how to create a violin plot in R; my previous work with violin plots was very basic ones in Python instead. However, I still wish I could make it a more aesthetically pleasing graph with graphics on density. Question 7 continued to aggravate me as I could not understand why I continued to get errors for parts C and D, so there is definitely a knowledge gap of what I wanted to do and what I could.