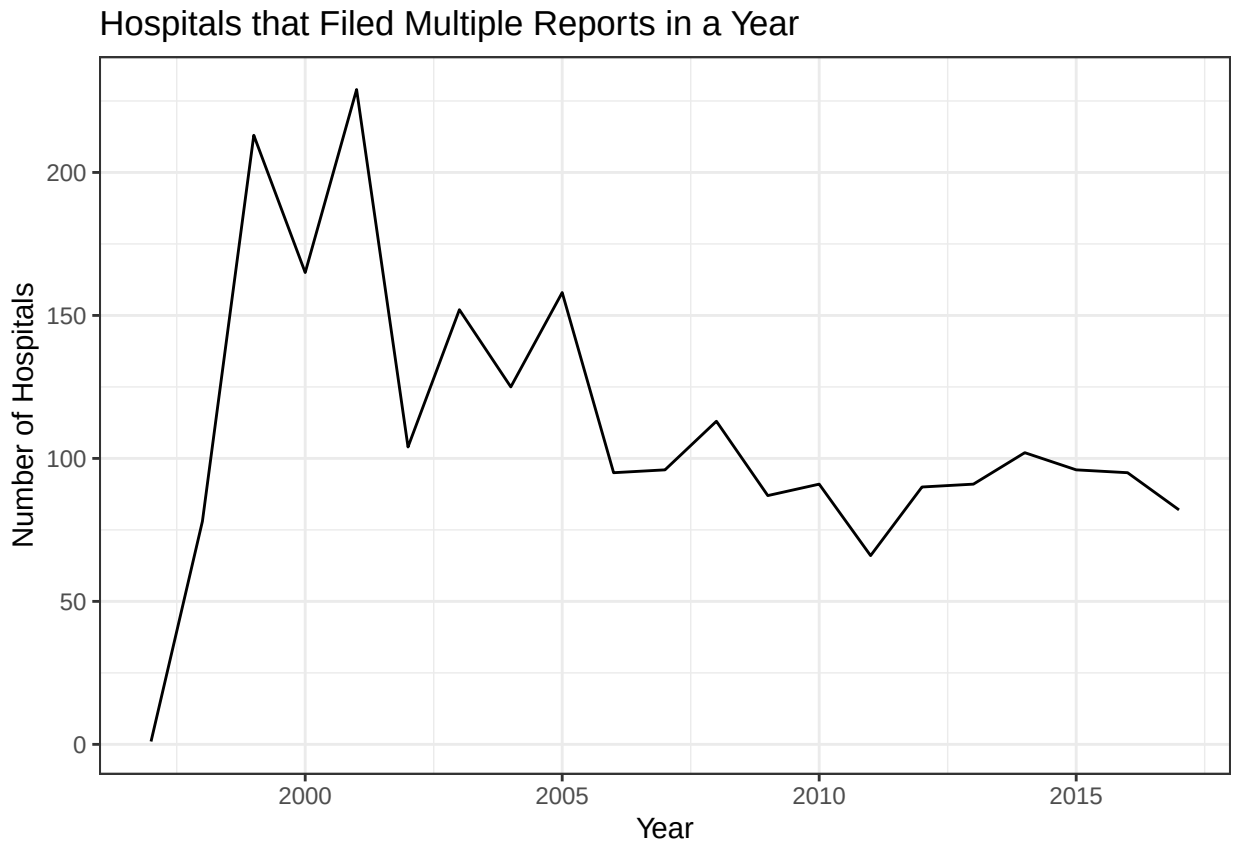
Bhasin-S-hwk2-3

Sachi Bhasin

Question 1

2,329 hospitals filed more than one report in the same year.

```
graph_2
```



Question 2

After removing/combining multiple reports, there are 9,323 unique hospital IDs in the data.

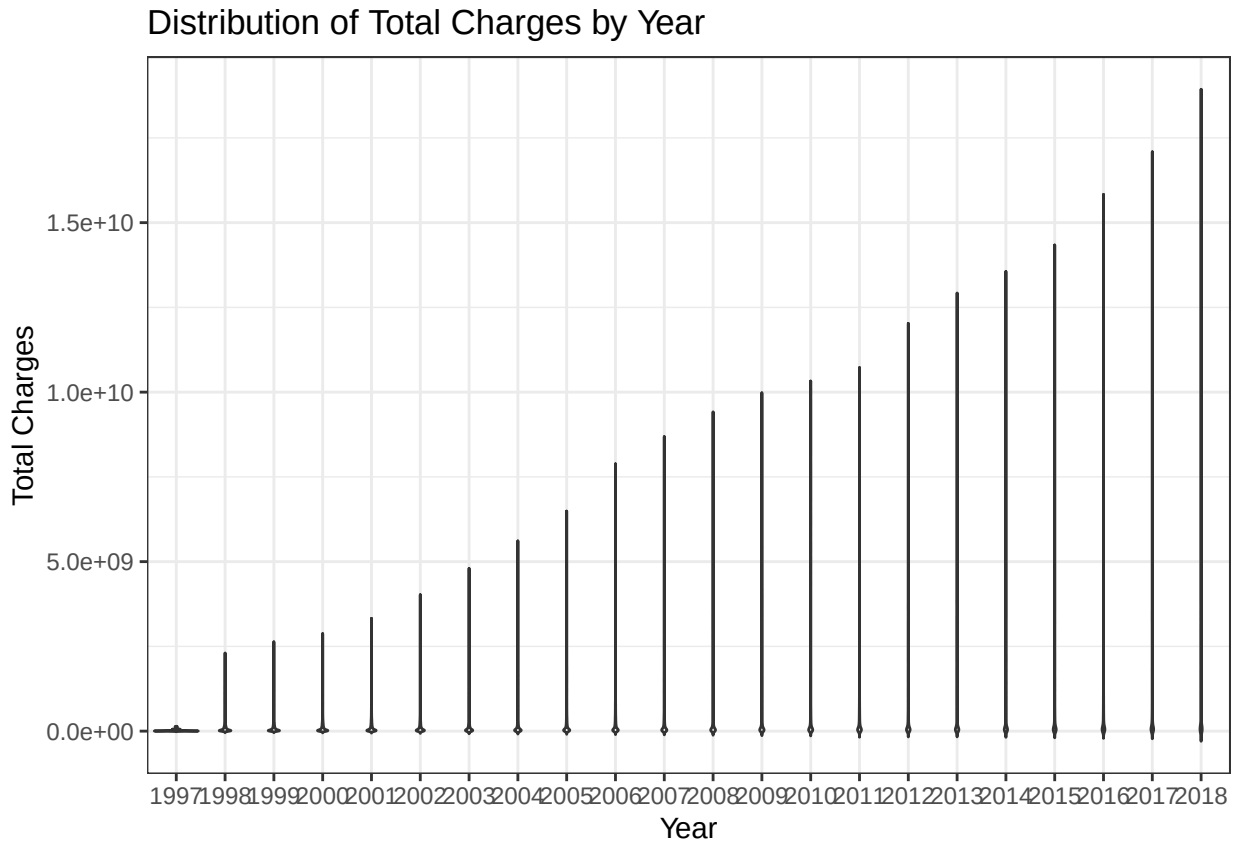
```
unique_provider_numbers
```

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

Question 3

```
graph_3
```

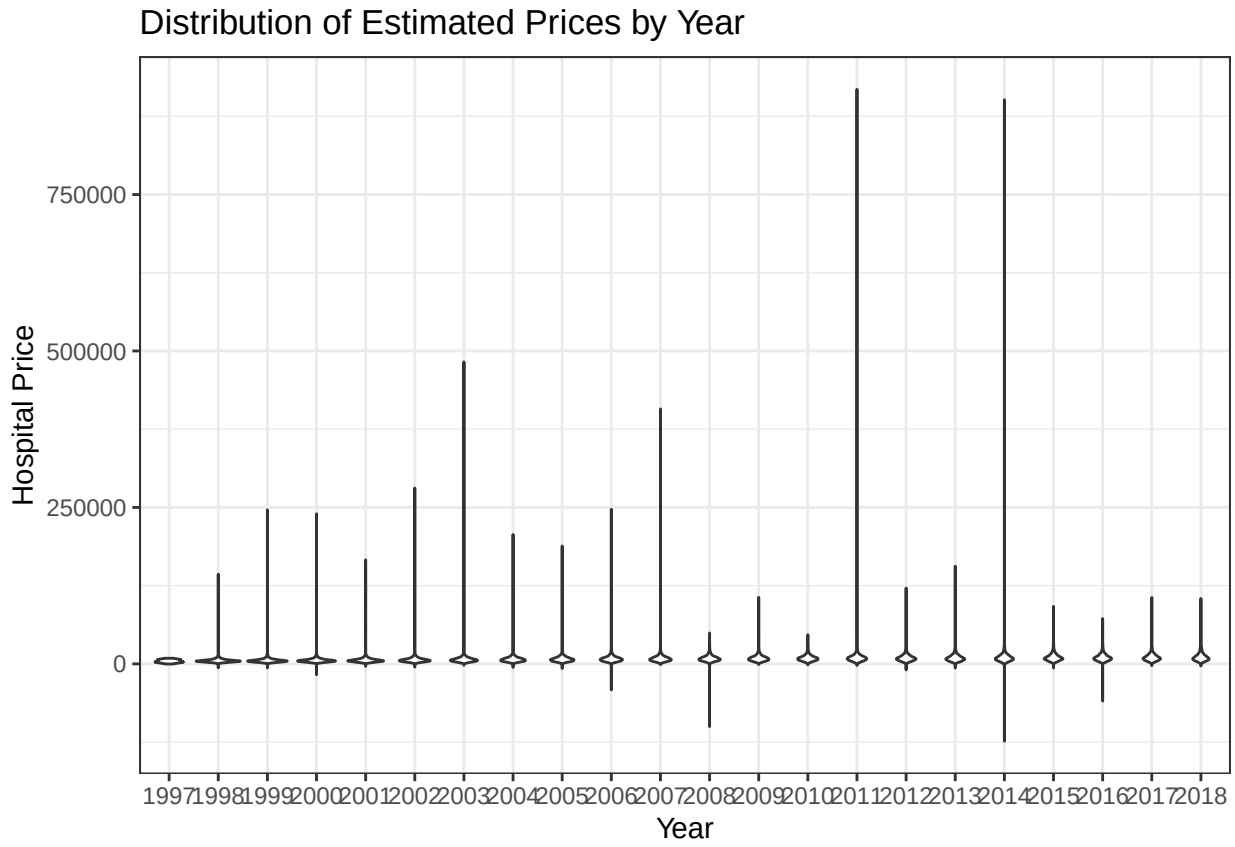
```
## Warning: Removed 4748 rows containing non-finite values (‘stat_ydensity()‘).
```



Question 4

```
graph_4
```

```
## Warning: Removed 63663 rows containing non-finite values (‘stat_ydensity()‘).
```



Question 5

Before calculating the average price among penalized versus non-penalized hospitals, prices were filtered to be positive and the penalties were filtered to below 100,000 to get rid of outliers.

```
table_5
```

```
## # A tibble: 2 x 2
##   penalty  price
##     <dbl>  <dbl>
## 1       0  9791.
## 2       1 10235.
```

Question 6

```
table_6
```

```
## # A tibble: 8 x 3
## # Groups:   penalty [2]
##   penalty quartile avg_price
##     <dbl>    <int>     <dbl>
## 1       0        1     8482.
## 2       0        2     8361.
## 3       0        3    10521.
## 4       0        4    11749.
## 5       1        1     7653.
## 6       1        2    10833.
## 7       1        3     9339.
## 8       1        4    12435.
```

Question 7a

```
summary(inv_var)
```

```
##                   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      24724 -none- numeric
## index.control      24724 -none- numeric
## index.dropped          0 -none- NULL
## weights            24724 -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        123620 -none- numeric
## version                1 -none- character
## estimand               1 -none- character
```

Question 7b

```
summary(Maha)
```

```
##                   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      24724 -none- numeric
## index.control      24724 -none- numeric
## index.dropped          0 -none- NULL
## weights            24724 -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        123620 -none- numeric
## version                1 -none- character
## estimand               1 -none- character
```

Question 7c

```
reg.ipw
```

```
##
## Call:
## lm(formula = price ~ penalty, data = year_2012, weights = ipw)
##
## Coefficients:
## (Intercept)      penalty
##      9775.2        286.5
```

Question 7d

```
mean(pred1_alt-pred0_alt)
```

```
## [1] 443.5711
```

8 The results are identical for the nearest neighbor matching with inverse variance distance and Mahalanobis distance as well as the inverse propensity weighting. The simple linear regression was similar to the other estimators.

9 I think we have estimated a causal effect of the penalty by matching with inverse variance distance and Mahalanobis distance as well as inverse propensity weighting. Also, running the simple linear regression was another method for eliminating potential confounding variables, suggesting a causal effect.

10 I found working with this data challenging but easier than homework 1 as I am getting more comfortable trouble shooting and working with this application. I learned how to create a dummy variable and quartiles for a data set. It has been very aggravating to troubleshoot the error I keep getting with the propensity score.

save.image("Hwk2_workspace.Rdata")

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

0.1 Including Plots

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.