Homework 2

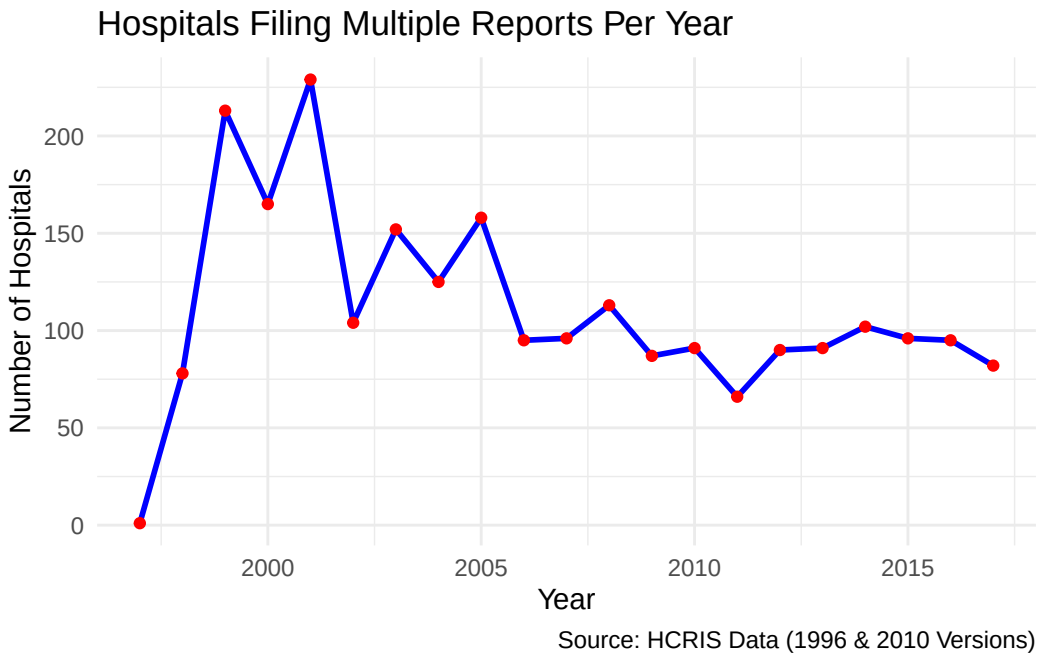
ECON 470, Spring 2025

Molly Catlin

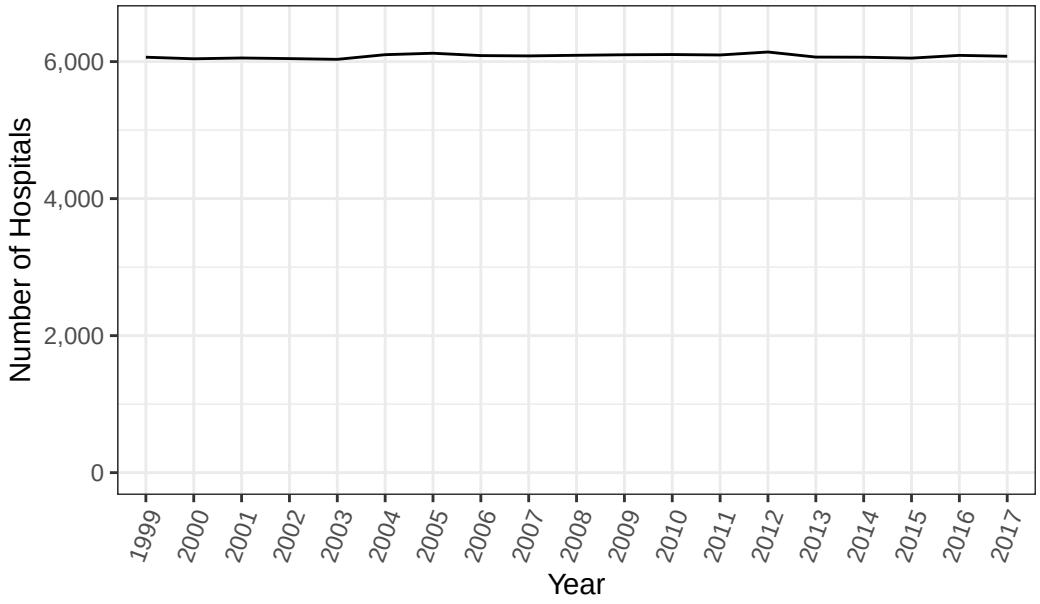
Here is a link to my repository: {https://github.com/mollyjc02/Homework_2.git}

1. How many hospitals filed more than one report in the same year? Show your answer as a line graph of the number of hospitals over time.

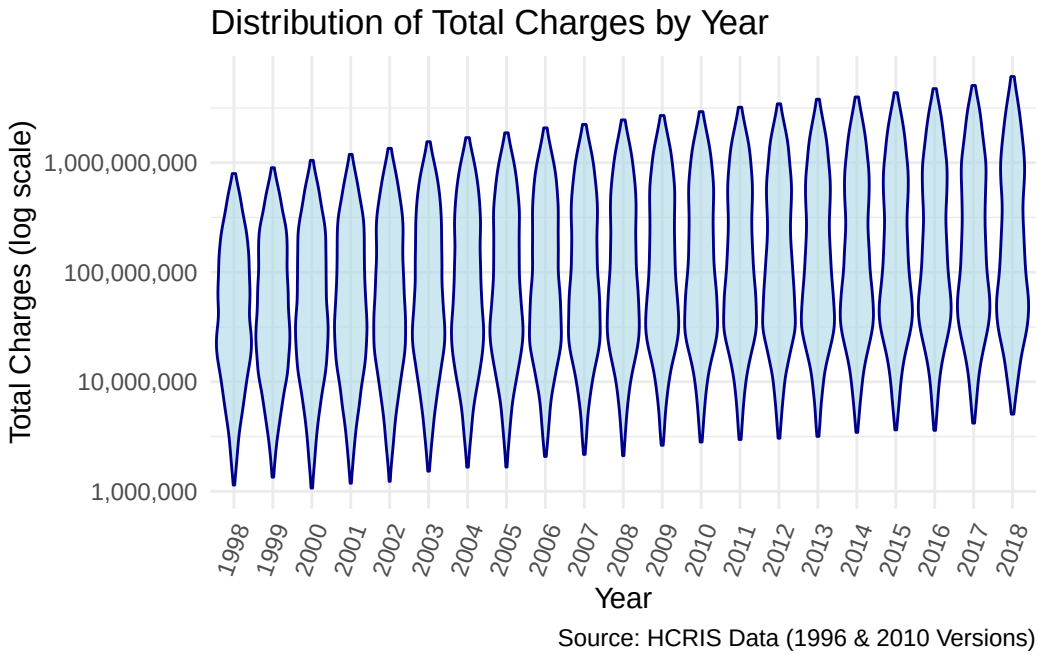
```
Warning in geom_point(color = "red", linewidth = 2): Ignoring unknown
parameters: `linewidth`
```



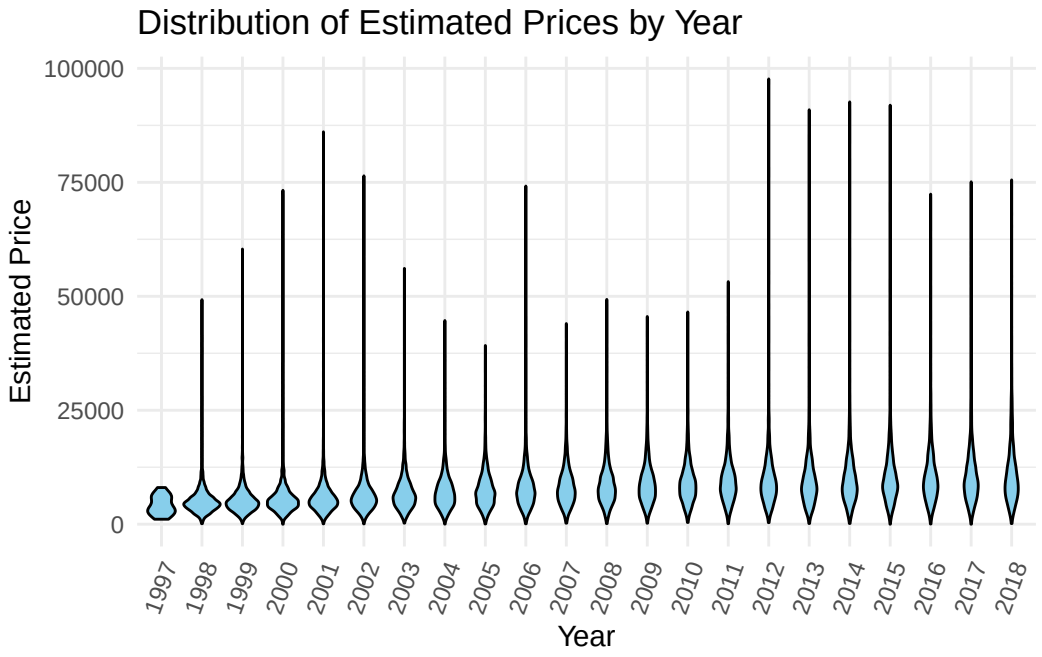
2. After removing/combining multiple reports, how many unique hospital IDs (Medicare provider numbers) exist in the data?



3. What is the distribution of total charges (tot_charges in the data) in each year?



4. What is the distribution of estimated prices in each year?



For the rest of the assignment, I have included only observations in 2012. So I am now dealing with cross-sectional data in which some hospitals are penalized and some are not.

5. Calculate the average price among penalized versus non-penalized hospitals.

The average price among penalized hospitals is 9896.31, while the average price among non-penalized hospitals is 9560.41.

6. Split hospitals into quartiles based on bed size and provide a table of the average price among treated/control groups for each quartile.

Table 1: Average Price by Treatment Status and Bed Size

Bed Quartile	No Penalty	Penalty
Q1	7834.979	7802.316
Q2	8327.120	9083.821
Q3	9356.467	10144.617
Q4	10633.655	10971.422

7. Find the average treatment effect based on quartiles of bed size using each of the following estimators: nearest neighbor matching with inverse variance distance, nearest neighbor matching with Mahalanobis distance, inverse propensity weighting, and simple linear regression.

Table 2: ATE Estimates

Method	ATE Estimate
Nearest Matching (Inverse Variance)	505.7106
Nearest Matching (Mahalanobis Distance)	505.7106
Inverse Propensity Weighting (IPW)	505.7106
Linear Regression	505.7106