Homework 2 - Submission 3

ECON 470

Baran Pasa

Homework 2 Analysis

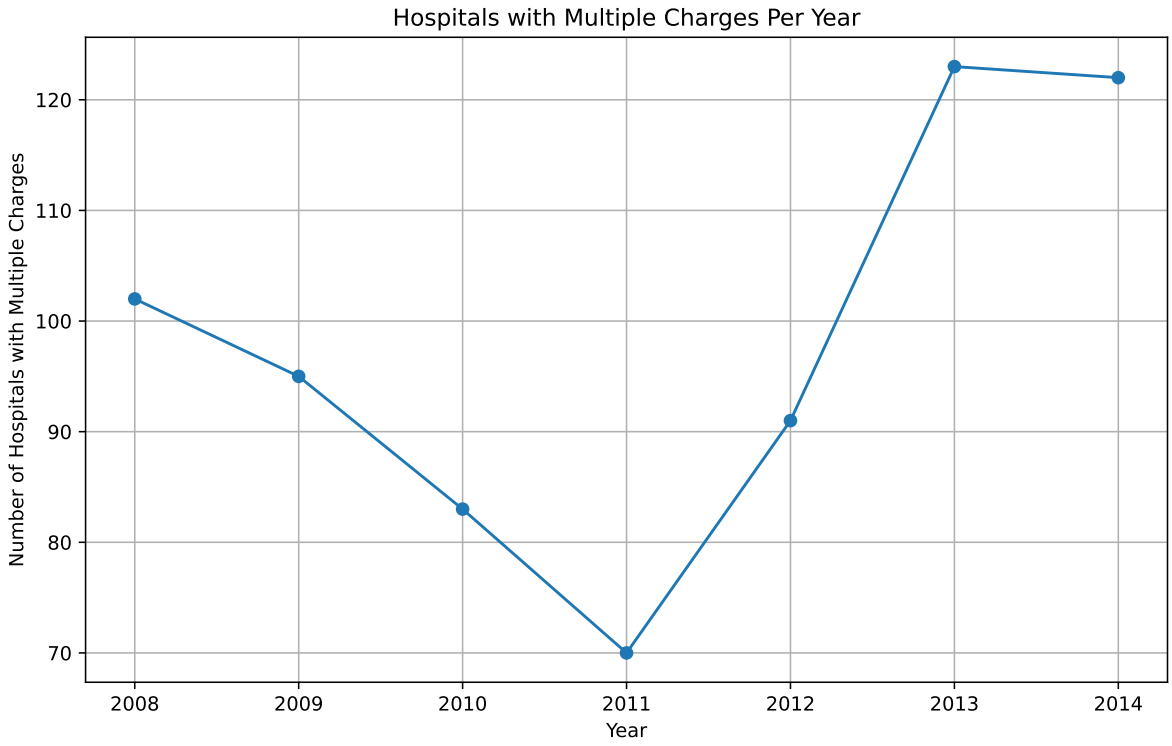
Link to Github

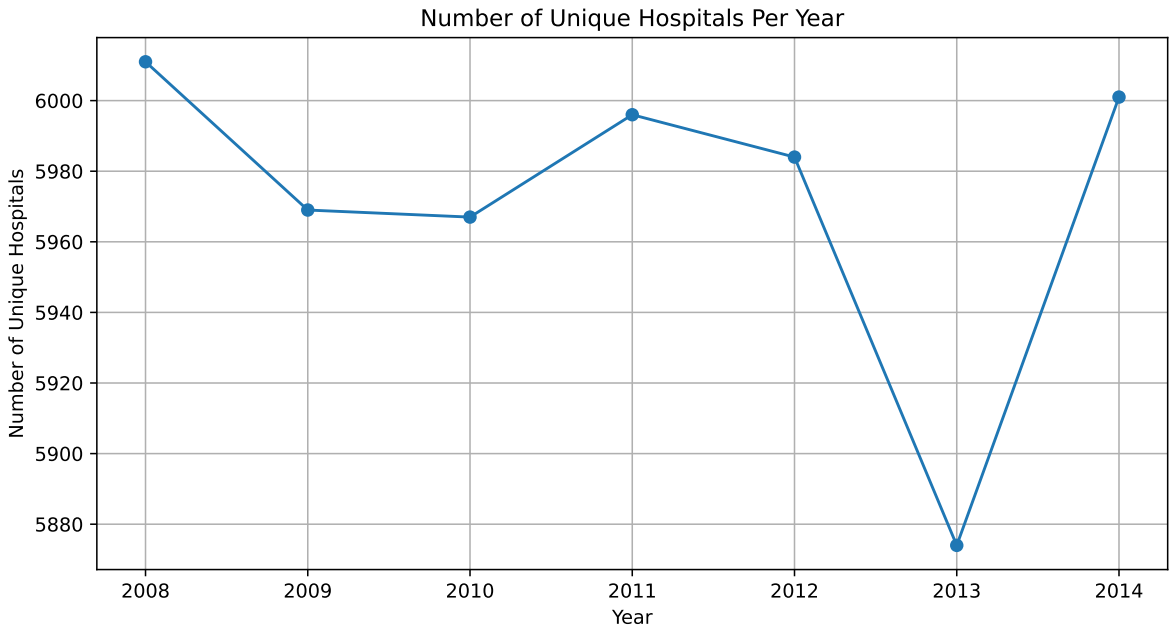
Part 0: Importing and cleaning the data

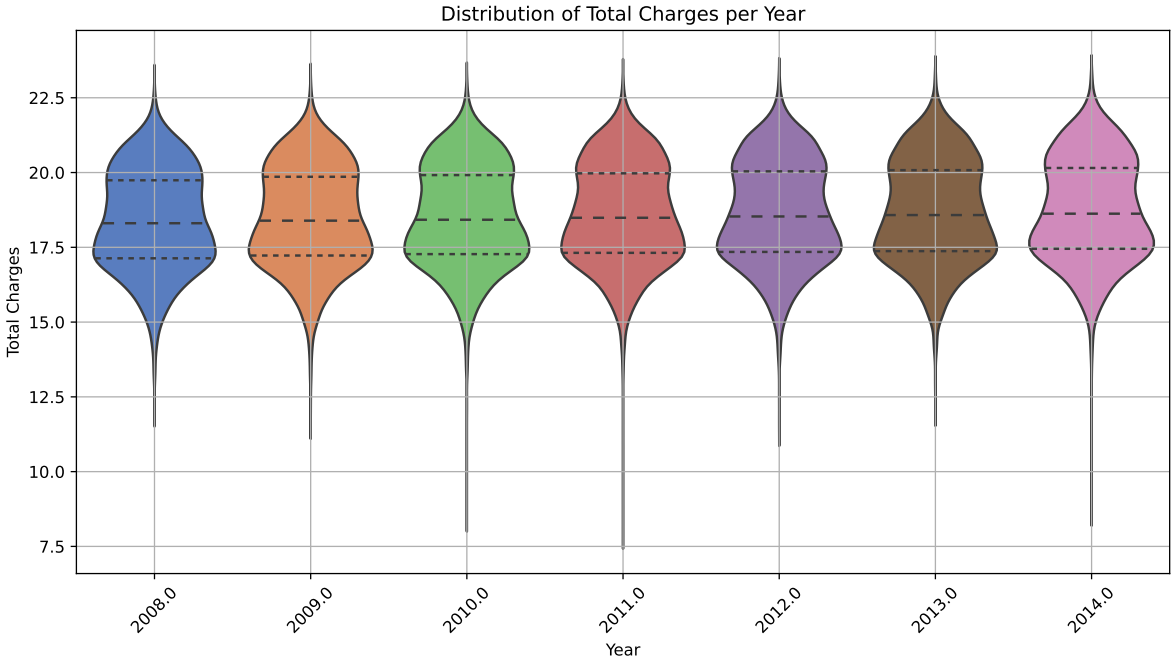
The data was cleaned in order to make summarizing and analysis easier. My cleaning process relied heavily on the code kindly provided to us by our TA Pablo. Additinally, data was only used from 2008 to 2015 due to my computer's lack of power.

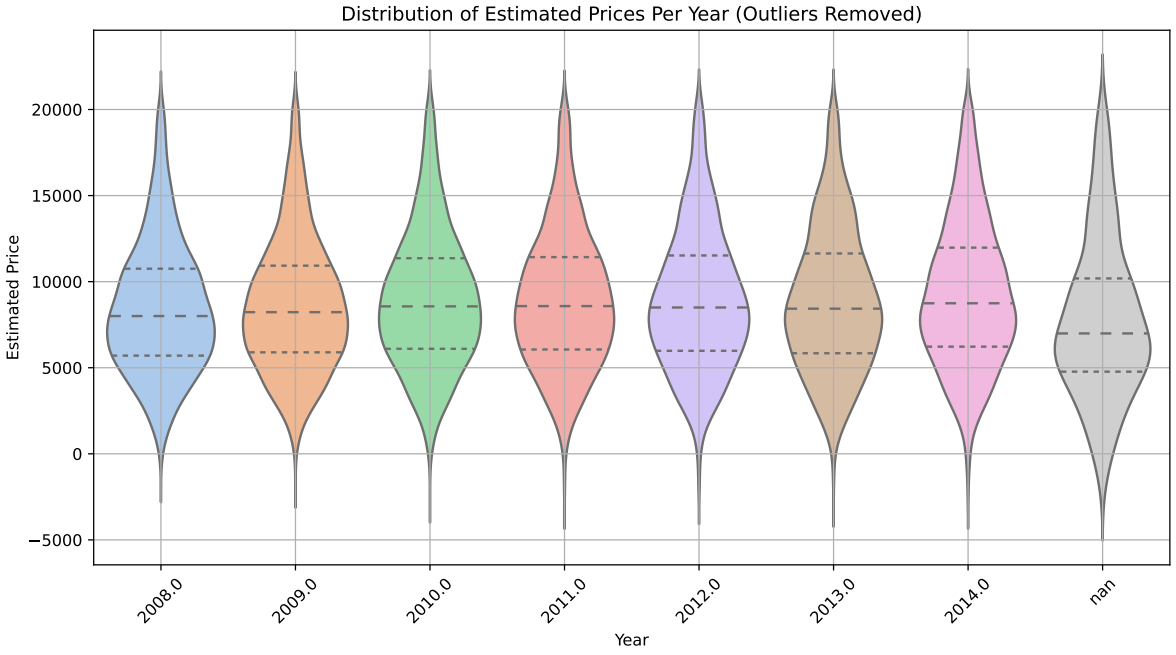
Part 1 Summarizing the Data

The graphs below provide a small summary of the data.


Hospitals with Multiple Charges Per Year
Number of Hospitals with Multiple Charges
120
110
100
90
80
70
2008
2009
2010
2011
2012
2013
2014
Year


Number of Unique Hospitals Per Year
Number of Unique Hospitals
6000
5980
5960
5940
5920
5900
5880
2008
2009
2010
2011
2012
2013
2014
Year





As you can see, only the years 2008 through 2014 are visible. You can also see that there is distribution of NaN on the final graph. This is an issue that needs to be furhter addressed, and is most likely occuring to due to faulty code.

Part 2: Estimating ATE

```
/Users/baranpasa/Library/Mobile Documents/com~apple~CloudDocs/Desktop/Emory/Junior Year/Junio


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide

/Users/baranpasa/Library/Mobile Documents/com~apple~CloudDocs/Desktop/Emory/Junior Year/Junio


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide

/Users/baranpasa/Library/Mobile Documents/com~apple~CloudDocs/Desktop/Emory/Junior Year/Junio


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide

/Users/baranpasa/Library/Mobile Documents/com~apple~CloudDocs/Desktop/Emory/Junior Year/Junio


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide

Average price for penalized hospitals in 2012: 10721.38
Average price for non-penalized hospitals in 2012: 9423.59
 Quartile  Penalized_Mean_Price  Non_penalized_Mean_Price
        1               8008.29                   7456.69
        2              10186.61                   8522.48
        3              11260.07                   9470.05
        4              13518.17                  12254.50

/opt/anaconda3/lib/python3.11/site-packages/causalinference/estimators/matching.py:100: Futu
```

```
`rcond` parameter will change to the default of machine precision times ``max(M, N)`` where M
To use the future default and silence this warning we advise to pass `rcond=None`, to keep us
```

	INV	MAH	IPW	OLS
ATE	1313.77	1313.77	1313.77	1313.77
SE	280.21	280.21	280.11	279.96

	INV	MAH	IPW	OLS
ATE	1313.77	1313.77	1313.77	1313.77
SE	280.21	280.21	280.11	279.96


```
/opt/anaconda3/lib/python3.11/site-packages/causalinference/estimators/weighting.py:23: Futu

`rcond` parameter will change to the default of machine precision times ``max(M, N)`` where M
To use the future default and silence this warning we advise to pass `rcond=None`, to keep us

/opt/anaconda3/lib/python3.11/site-packages/causalinference/estimators/ols.py:21: FutureWarni

`rcond` parameter will change to the default of machine precision times ``max(M, N)`` where M
To use the future default and silence this warning we advise to pass `rcond=None`, to keep us
```

display(Markdown(results.to_markdown()))