

# MODELING HOSPITAL READMISSION DRIVERS BY STATE

- Sheena Sharma

## WHY:

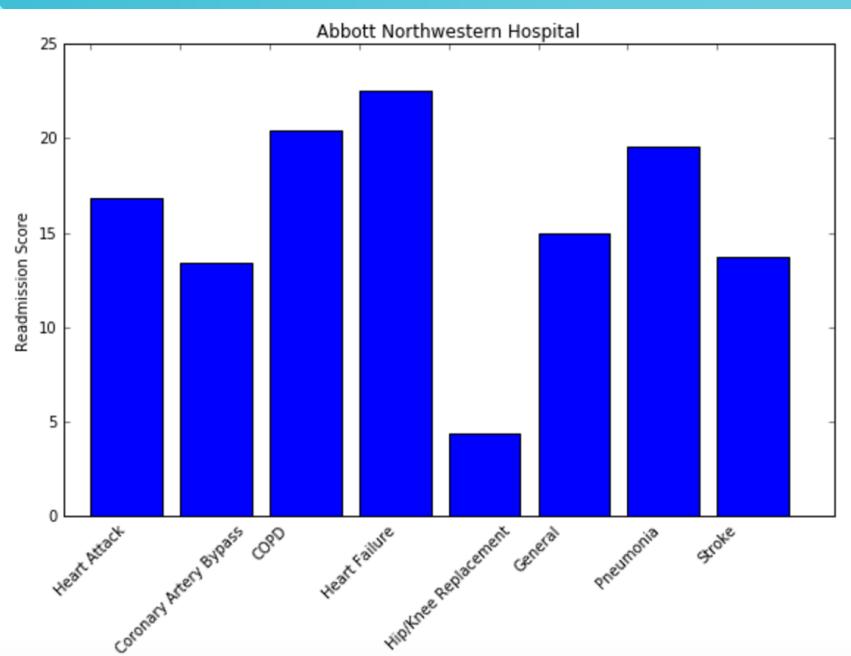
BECAUSE INSURE COMPANIES WANT TO KNOW WHERE READMISSIONS ARE MORE LIKELY (BECAUSE THE INSURANCE COMPANY LOSES MONEY) AND WHICH MARKET TO TARGET NEXT

## SOLUTION:

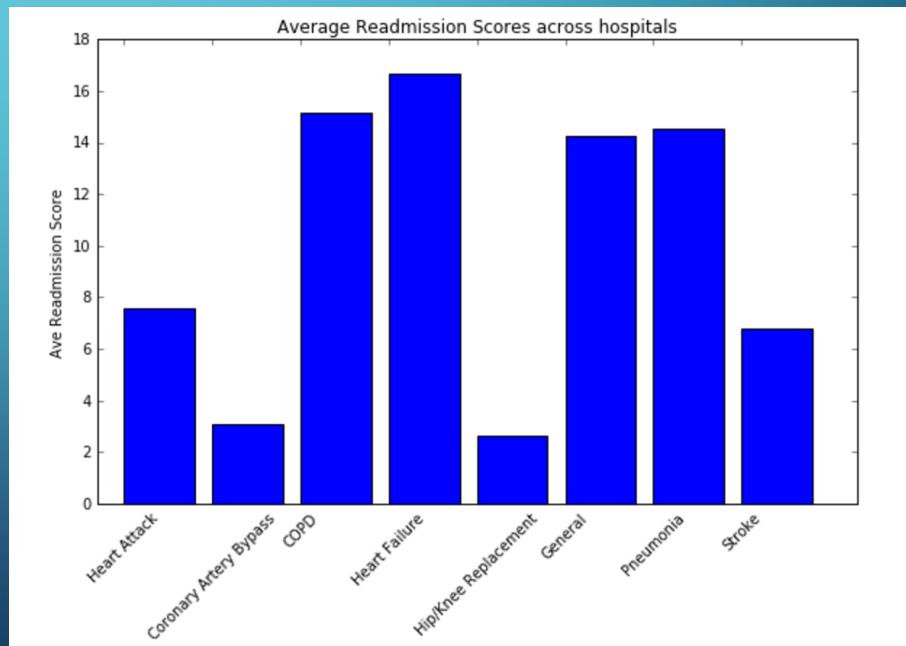
- TO PREDICT NATIONAL READMISSION RATES
- CAN WE BUILD A PER STATE MODEL

## DATASETS:

- MEDICARE DATASETS
- FEATURES INCLUDE – SPENDING PER CLAIM, PATIENT SURVEYS, COMPLICATION RATES DURING ADMISSION, INFECTION RATES DURING ADMISSION

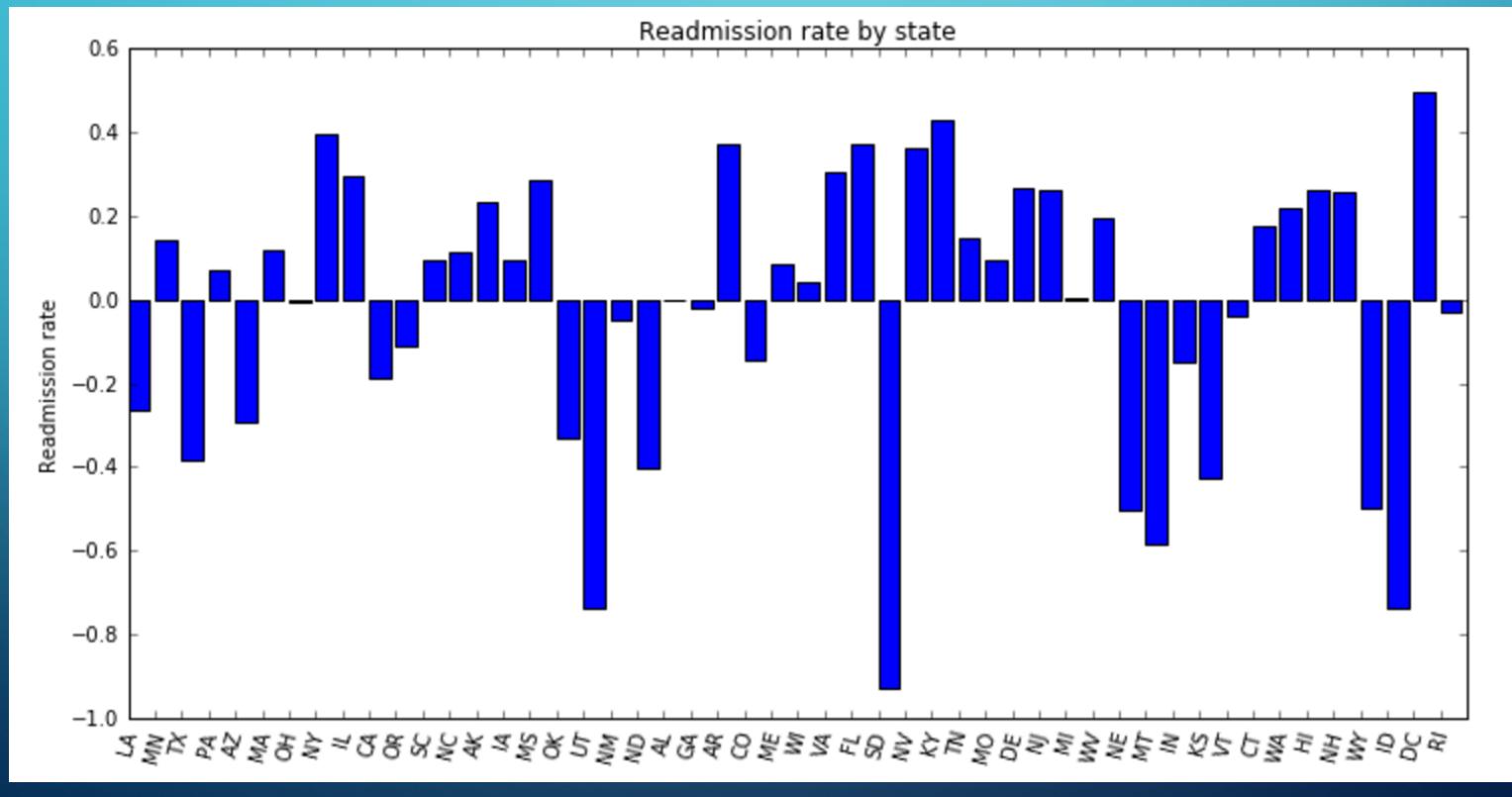


HF is one of the  
biggest contributors to  
readmissions

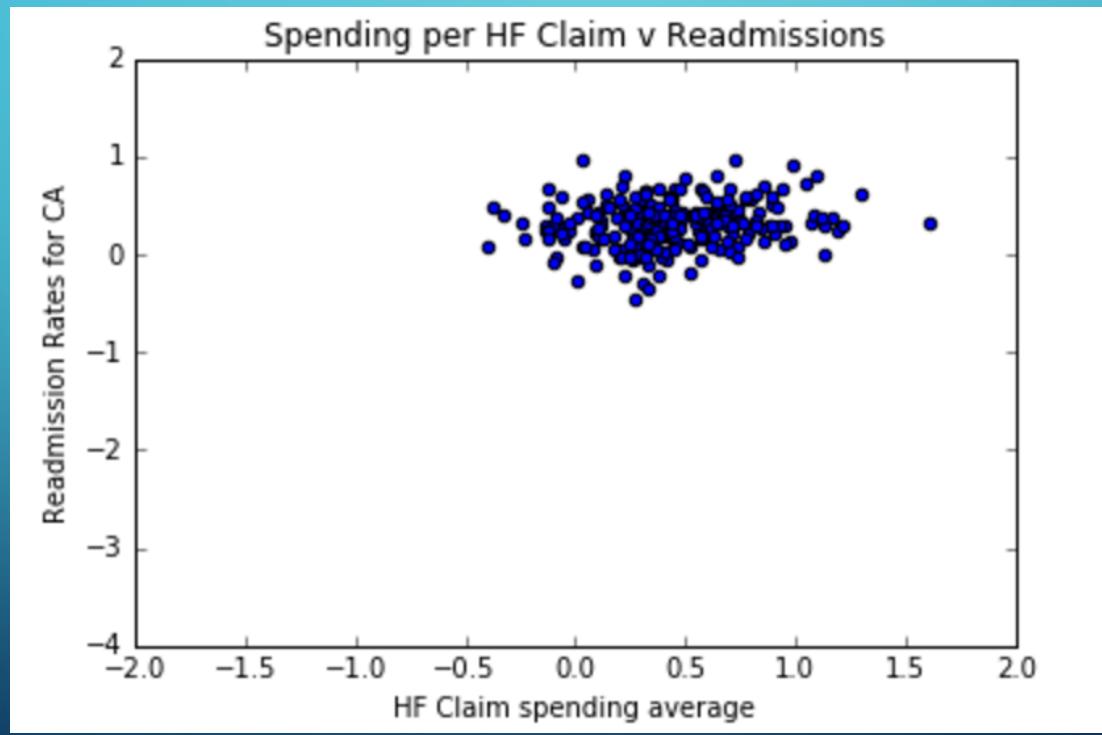


Huge state variance in readmission rates

This data is normalized



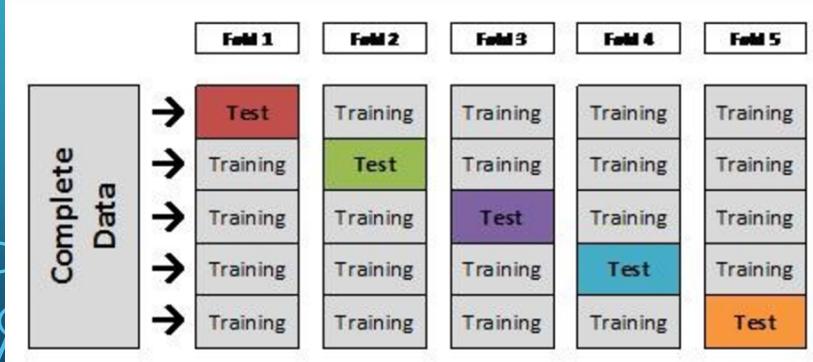
- There are many features (spending, patient surveys, complication rates, infection rates)
- Plotting one feature at a time vs readmission rates doesn't show any strong relationships.



## Linear Regression

- Built model for every state.
- Extracted important features (coefficients)

Training -> Cross Val -> Test



### Evaluation:

- Mean squared error of cross val

```
[-0.36911283256167149, -0.68437532927319467, -0.35120620301582382546505, -0.40797319847691715, -0.29463322117701524, -0.6406, -4.4492853845445994, -7.140983615411483, -1.226539949421232755434237, -0.51710105849513444, -0.31488612063966015, 32831, -1.9962351757242203, -1.4012877878314927, -0.1554477720983644310190966, -1.9523841397227695, -0.18873944890468436872634, -2.0832845294908102, -2.1345650333150474, -9.233244262659378647459685, -26.879970832809477, -2.8359635366050804, 5422801, -0.26895008256254299, -0.033159789025653474, -0.5862, -0.39892801138499967]
```

- $R^2$  of cross val
- Future: Scramble labels and get mse – compare to data with actual labels and plot differences.

