



By: Githika Annapureddy

# Private vs. State Prison Spending

# Introduction

Research Question: How do private prisons and state prisons spend money differently? Does this uncover any underlying patterns amongst private or state prisons?

Context: In a post emancipation America, state governments established prison contracts with rich white former slave-owners. Former slave-owners used forced prison labor as a replacement to slave labor, continuing their cruel ways of money making. They treated prisoners worse than slaves – seeing them as replaceable. Private prisons were attractive to governments because it was cheaper to send prisoners there than to imprison them in state facilities. Today, conglomerates benefit from a high incarceration rate – making private prisons incompatible to a country with an equitable justice system.



# Data Wrangling

## Source

[Wikipedia of Private Prisons](#)

[Arizona Department of Corrections- Cost Report](#)

[Arizona Department of Corrections - Inmate Fact Sheet](#)

## Challenges

PDF → XLSX → CSV

## Tools

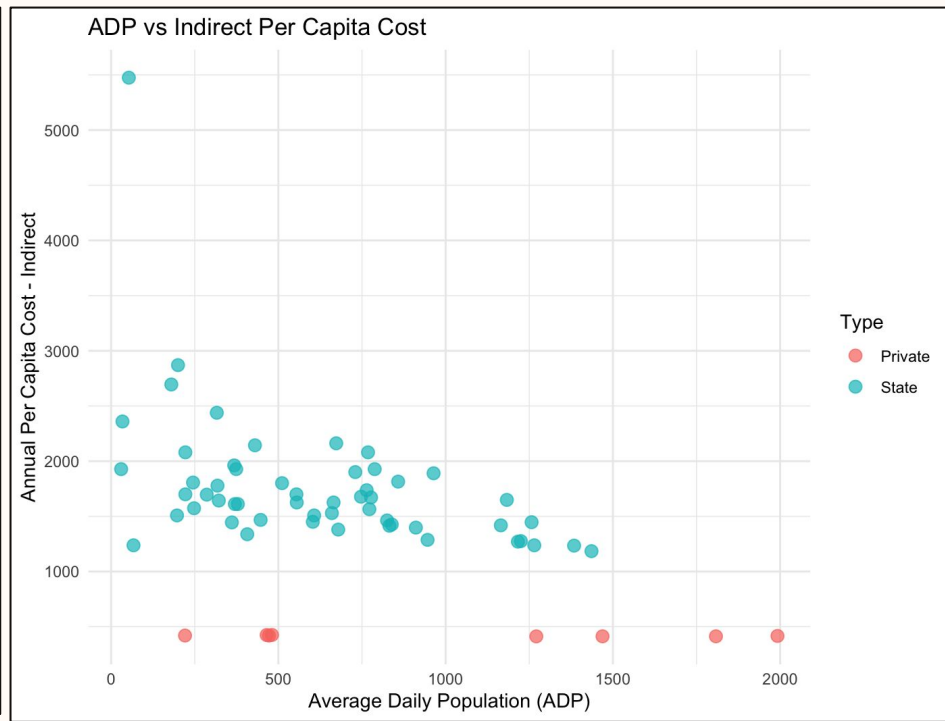
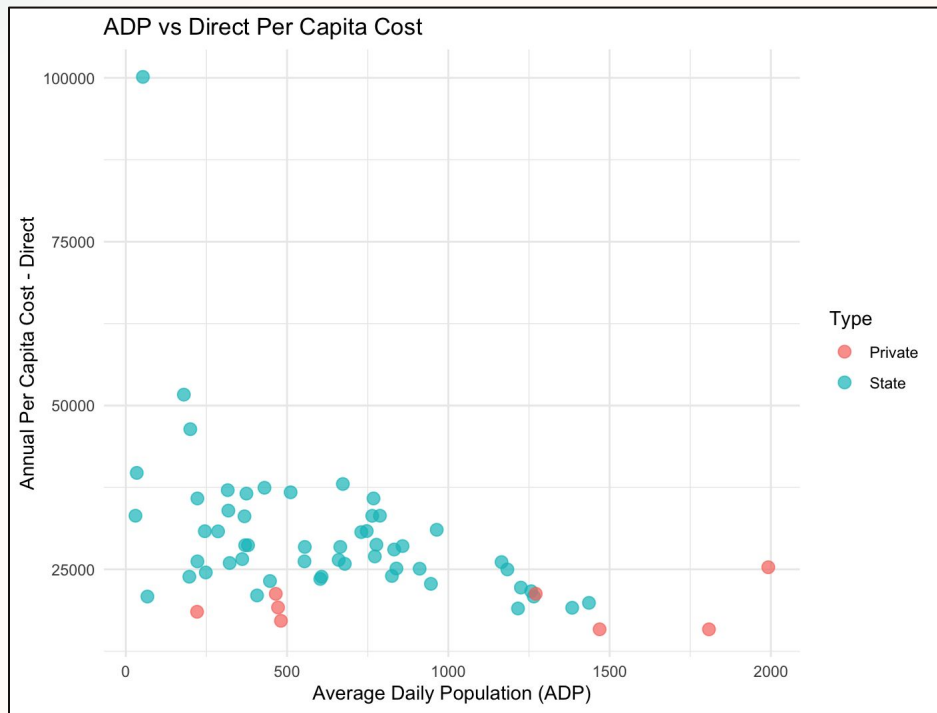
Adobe Acrobat, Excel, and R

## Methods

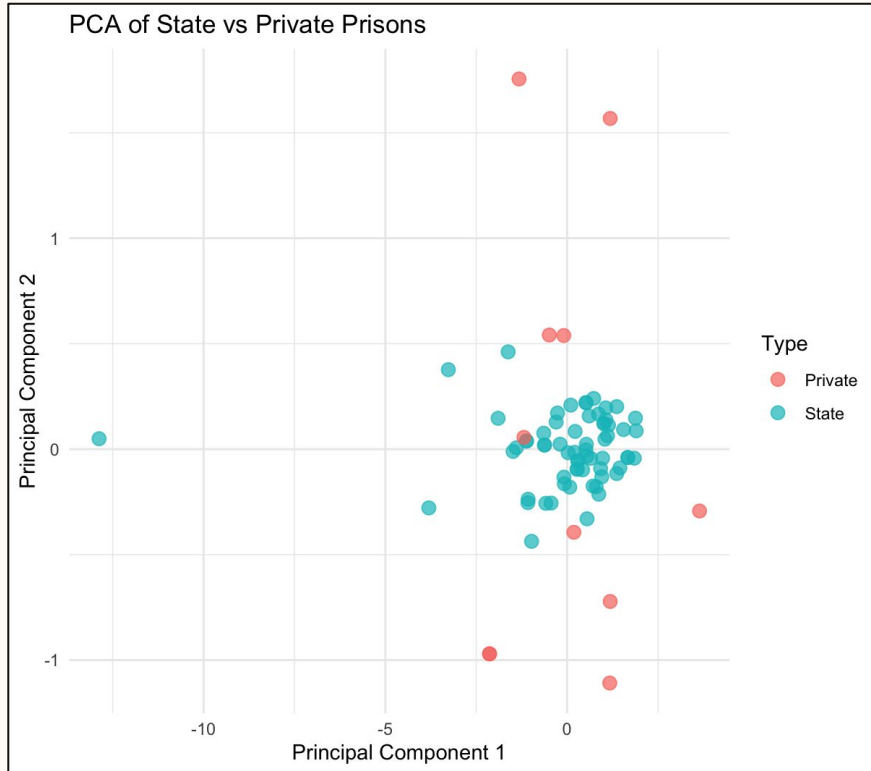
Extract and save certain lines of a csv to new csv

Make multiple csv files with each table from the original

# EDA



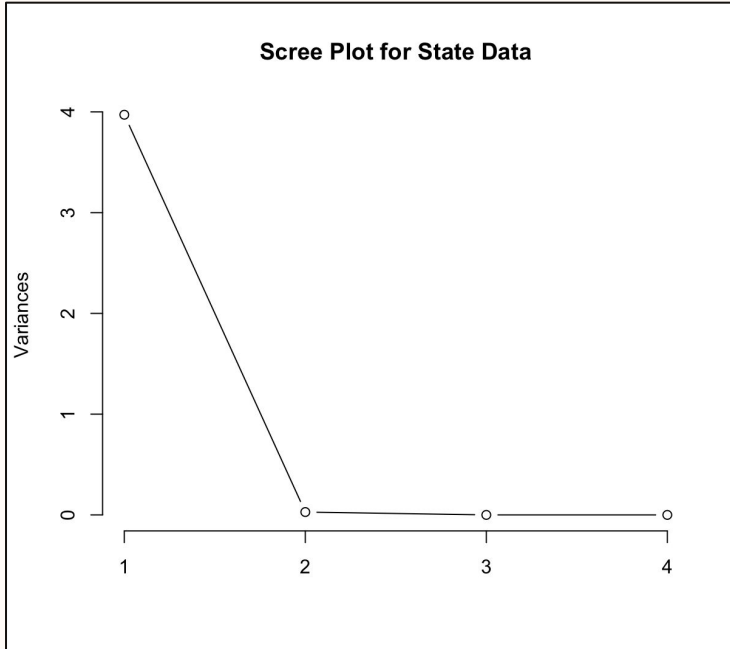
# Modeling - PCA



Does not show clear separation between Private and State

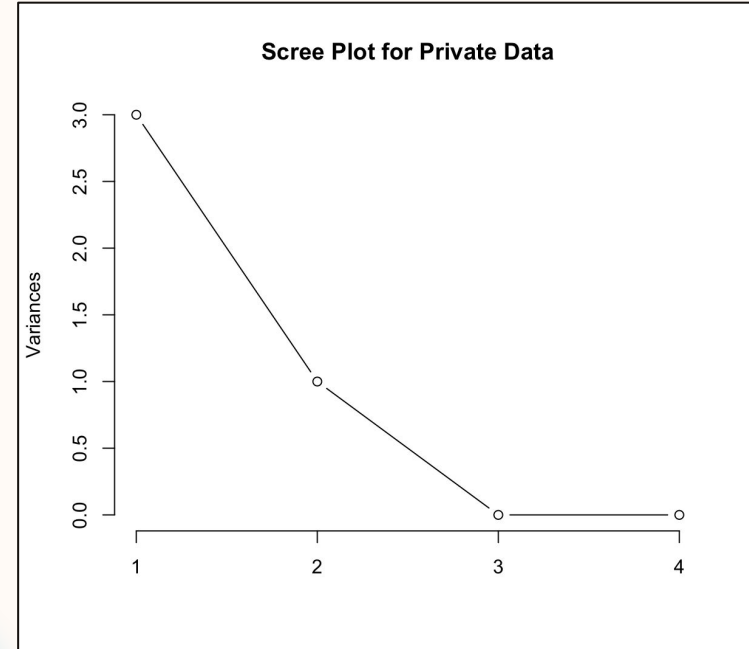
Private is more spread out, whereas State is clustered to one area

# Modeling - PCA



**PC1 (State) and PC1 (Private):**

- Annual Per Capita Cost - Total
- Daily Per Capita Cost
- Annual Per Capita Cost - Direct
- Annual Per Capita Cost - Indirect

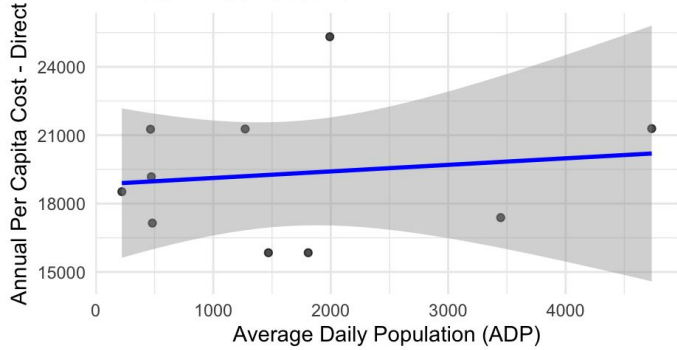


**PC2 (Private):**

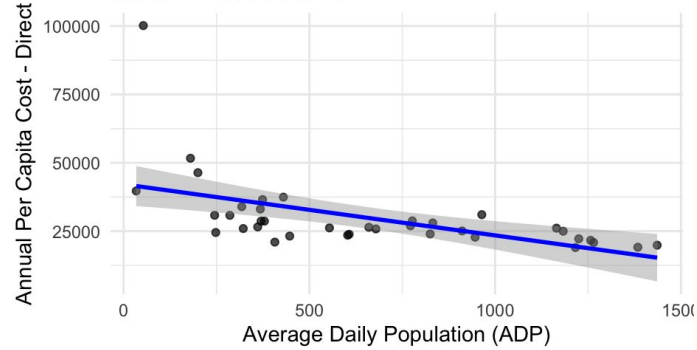
- Annual Per Capita Cost - Indirect
- Annual Per Capita Cost - Direct
- Daily Per Capita Cost
- Annual Per Capita Cost - Total

# Individual Regression

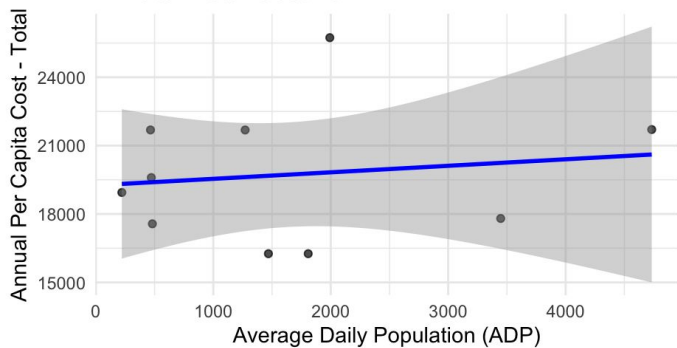
Private: Direct Cost vs ADP



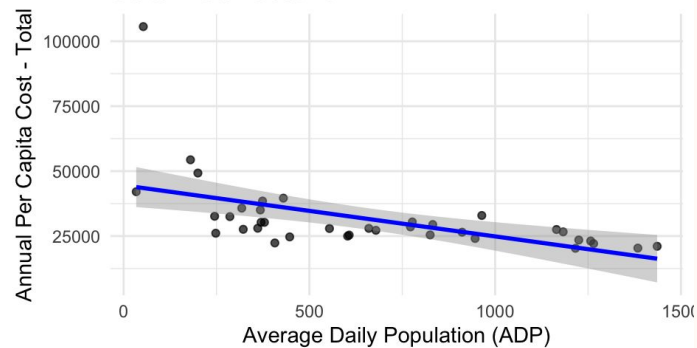
State: Direct Cost vs ADP



Private: Total Cost vs ADP



State: Total Cost vs ADP

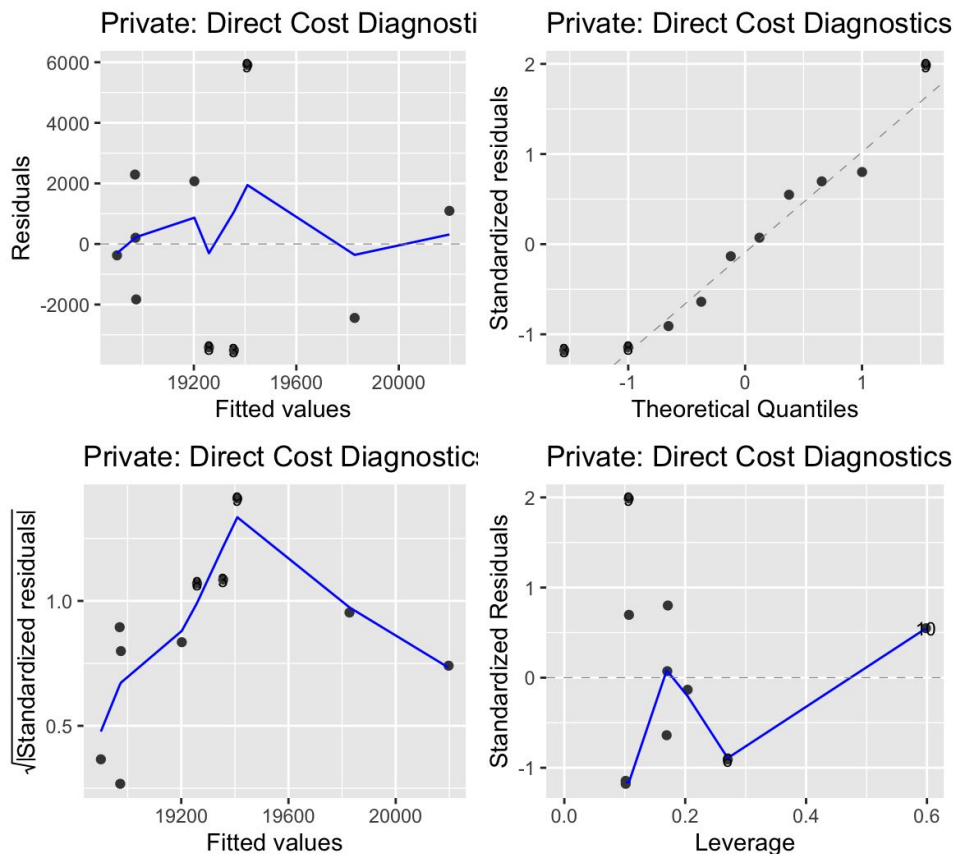


**Private:**  
No clear relationship

**State:**  
Negative relationship

Per Capita costs  
decreases as  
Population increases

# Regression: Private



## Residuals (top left)

- Non-randomness suggests heteroscedasticity

## Q-Q Plot (top right)

- Diagonal means normal distribution

## Scale-Location Plot (bottom left)

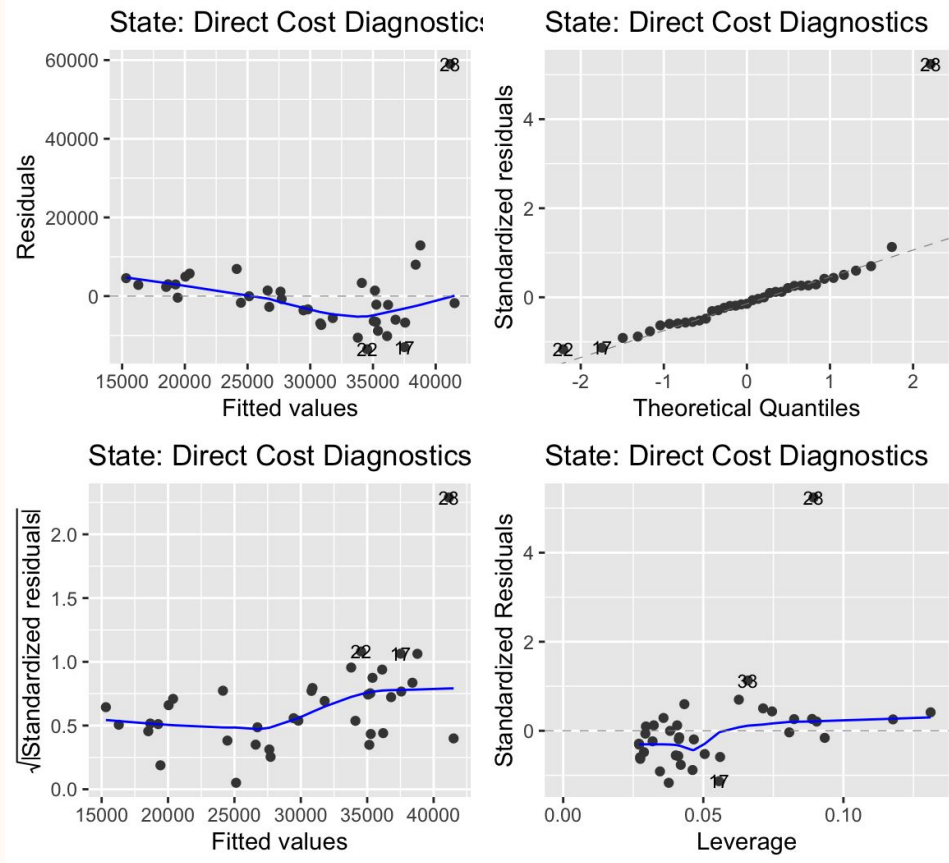
- Non-flat suggests heteroscedasticity

## Residuals vs. Leverage (bottom right)

- Inflection points suggest one point has high influence



# Regression: State



## Residuals (top left)

- Non-randomness suggests a linear relationship may not accurately capture relationship

## Q-Q Plot (top right)

- Diagonal means normal distribution

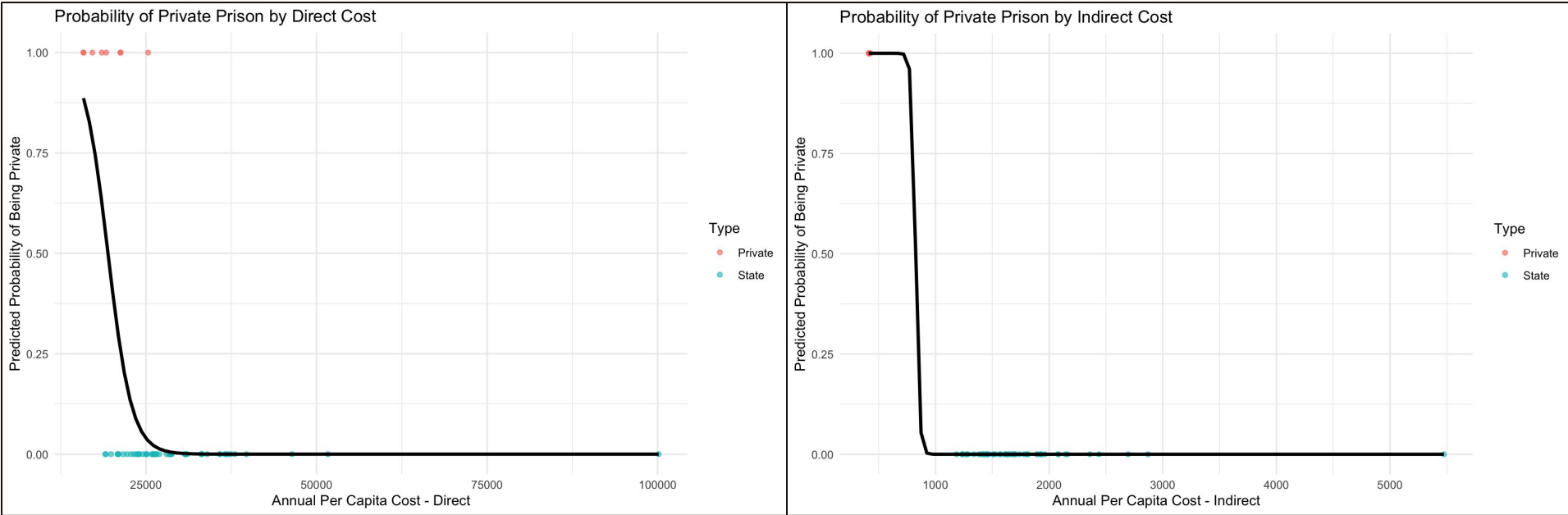
## Scale-Location Plot (bottom left)

- Flat suggests constant variance

## Residuals vs. Leverage (bottom right)

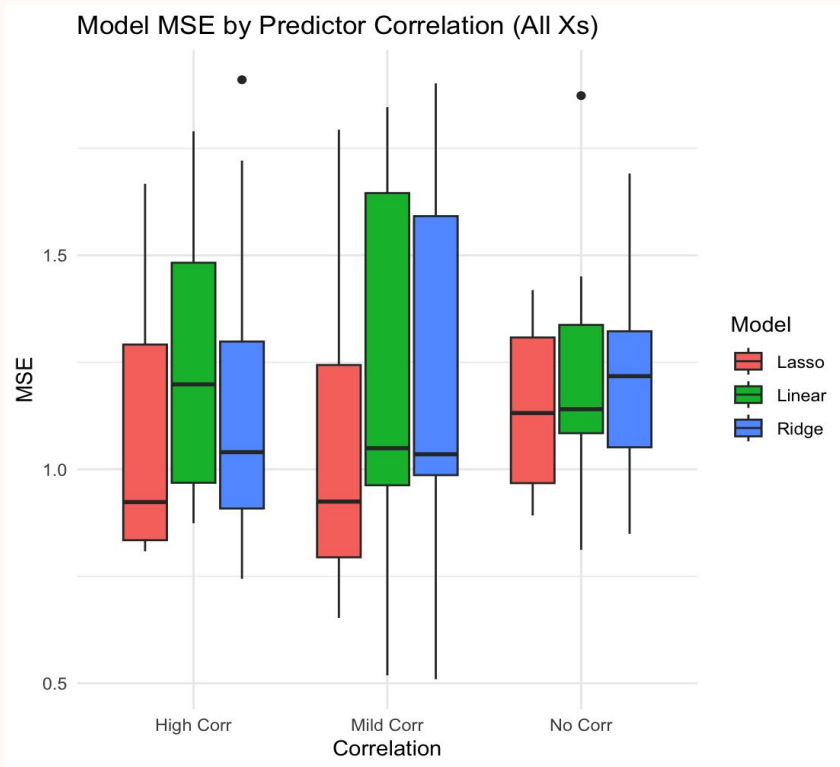
- Inflection points suggest one point has high influence

# Combined Regression



Likely overfit due to small amount of data. Model picked up on general trend, higher per capita costs are incurred by state prisons.

# Monte Carlo Simulation



My Example:

$x_1 \rightarrow x_{15}$

$$Y = 1 * x_2 + 2 * x_3 + 3 * x_4 + 4 * x_5 + \text{noise}$$

Class Example:

$x_1, x_2$

$$Y = 4 * x_2 + \text{noise}$$

MSE still not significantly different

# Summary/Reflection

## Model Summary

- As population increases, the average cost per inmate goes down in state-run facilities
  - Economies of scale — larger prisons can spread out fixed costs more effectively
- For private facilities, the cost per inmate stays the same or even increases slightly as the population grows.
- Private prisons have lower costs, likely because they aim to maximize profits.
- Challenges: Far less data for private than state data, potentially skewing results.

## Implications

- Questions the level of care of incarcerated people in private prisons receive.
- No clear benefit to privatization, state prisons seem to budget better.
- Could be the start of using data driven evidence to lean governments away from private prisons
- There is not much data from prisons, so analyzing what is publicly available is all we can do.

## I learned

- Reproducible repo: relative paths, make dependencies easy to download, explain steps in comments/README
- Structuring a data science project: importance of EDA, how to choose appropriate model



# Thanks!

# Questions?