

# Most Influential Factors Impacting Income Inequality Across States in the US

Team Cris Jenner

Claire Raney, Rachel Collins, Irene Biju, Sophia Liu



After accounting for state, which variable accounts for the most variability in median household income?

**Education**

**Ethnicity**

**Gender**

**Age**



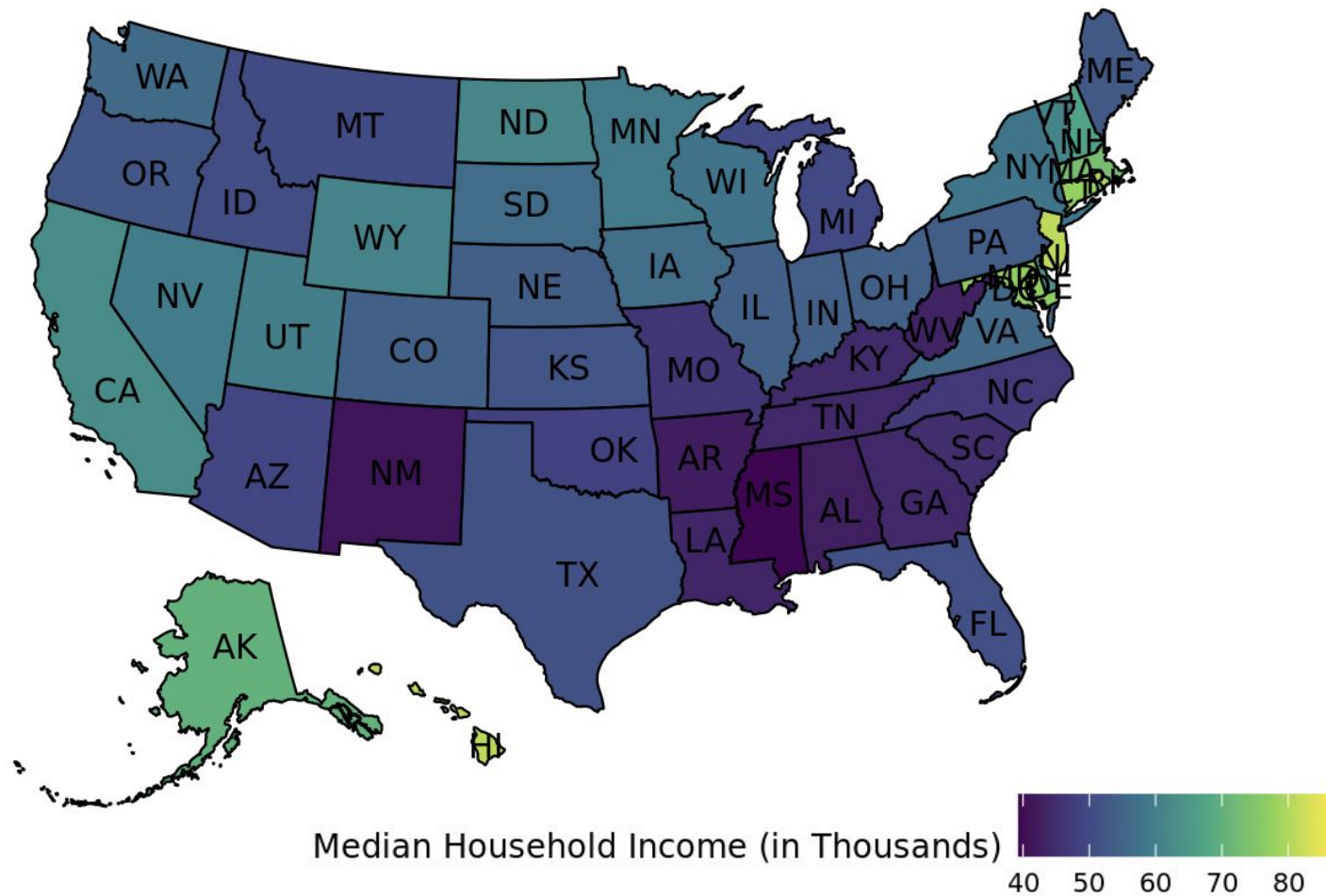
# Dataset

---

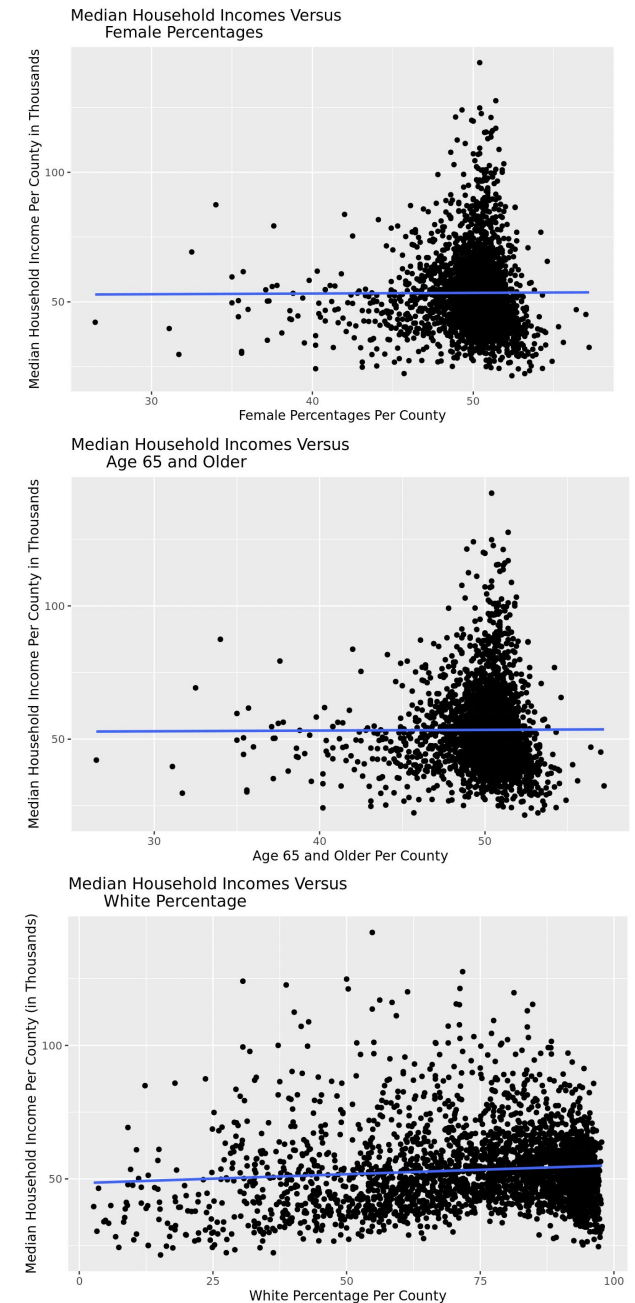
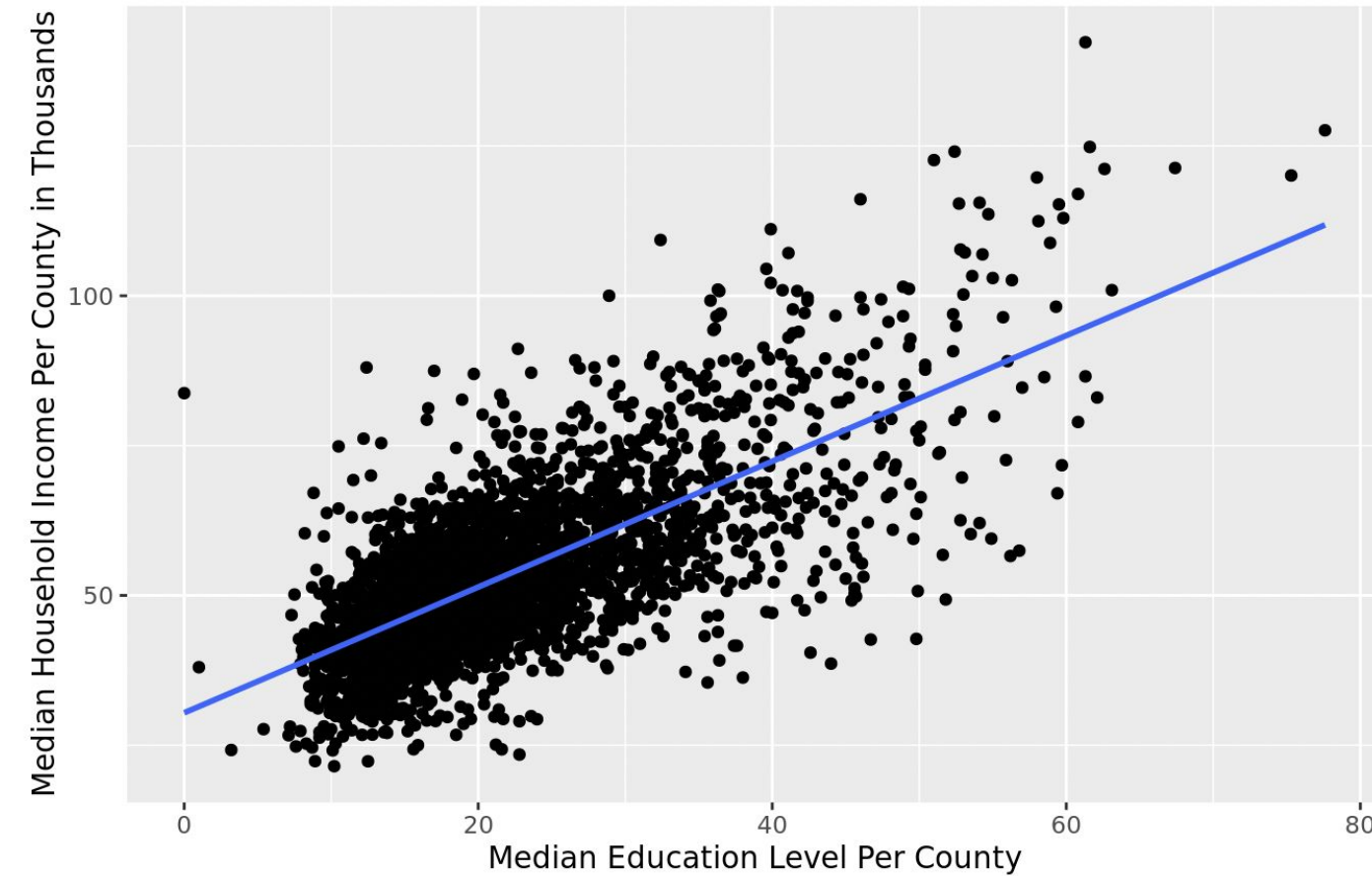
- Data was collected by the **US Census Bureau** from 2010 to 2019
- Data organized into counties more specifically as well as states; we focused on **states** as it gives us a more holistic picture of the US
- Data is tracking percentages (ex. % female in North Carolina)
- Variables we looked at:
  - Median household income
  - States
  - Gender
    - % female population
  - Ethnicity
    - % white population
  - Education Levels
    - % with a Bachelor's degree or higher
  - Age
    - % population that is 65 or older

# Exploratory Data Analysis

Household Incomes Per State



Median Household Incomes Versus Education Levels



- To visualize median household income across the US by state, we made a map
- We found that only **education level** seemed to positively correlate with median household income as income vs. education was the only graph that had a fairly strong linear fit; the other best-fit lines were almost completely flat and weak



# Modeling

Highest: New Jersey

Lowest: Mississippi

R-Squared: 0.2923676

```
Call:
stats::lm(formula = income ~ State, data = data)
```

Coefficients:

(Intercept)	StateAL	StateAR	StateAZ	StateCA	StateCO
68491.5	-24916.6	-26254.8	-19501.5	-777.9	-9696.7
StateCT	StateDC	StateDE	StateFL	StateGA	StateHI
10709.9	17928.5	-2503.5	-17201.3	-21261.8	7937.1
StateIA	StateID	StateIL	StateIN	StateKS	StateKY
-11231.9	-17569.8	-11835.2	-12624.9	-15626.8	-23609.3
StateLA	StateMA	StateMD	StateME	StateMI	StateMN
-23617.5	9254.1	9050.1	-14922.7	-16683.1	-6835.0
StateMO	StateMS	StateMT	StateNC	StateND	StateNE
-20395.4	-28878.4	-17507.5	-20072.0	-6016.1	-13796.5
StateNH	StateNJ	StateNM	StateNV	StateNY	StateOH
816.7	14395.7	-24768.5	-8371.8	-6006.9	-12691.2
StateOK	StateOR	StatePA	StateRI	StateSC	StateSD
-19780.8	-14538.4	-11435.4	7622.9	-22883.2	-13870.1
StateTN	StateTX	StateUT	StateVA	StateVT	StateWA
-21324.5	-15643.9	-5431.9	-7734.7	-8731.7	-9098.0
StateWI	StateWV	StateWY			
-10185.6	-23599.2	-6662.8			

```
glance(model_baseline)$r.squared

[1] 0.2923676
```

# Model Analysis

## *Additive Models*

Added Variable	R-Squared
Education	→ 0.6129233 ←
Percent Female	0.2956188
Percent Age 65 and Older	0.3746036
Percentage White	0.2985728

## *Conclusions*

Education, along with state, account for the most variability in household income

# Discussion

- Limitations
  - Only median values
  - Percentages
  - Only focused on certain subsets of categories
- Future Steps
  - Expand analysis to focus on specific counties
  - Making more explanatory variables more nuanced
    - Bachelor's degree vs. high school degree vs. no degree
    - Specific ethnic minorities

