

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a stylized tree structure.

AN ANALYSIS OF 2020 US POVERTY RATES BY COUNTY


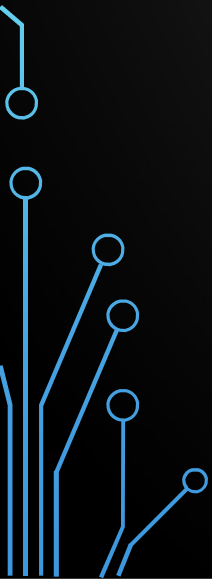

CAROLYN FIORE

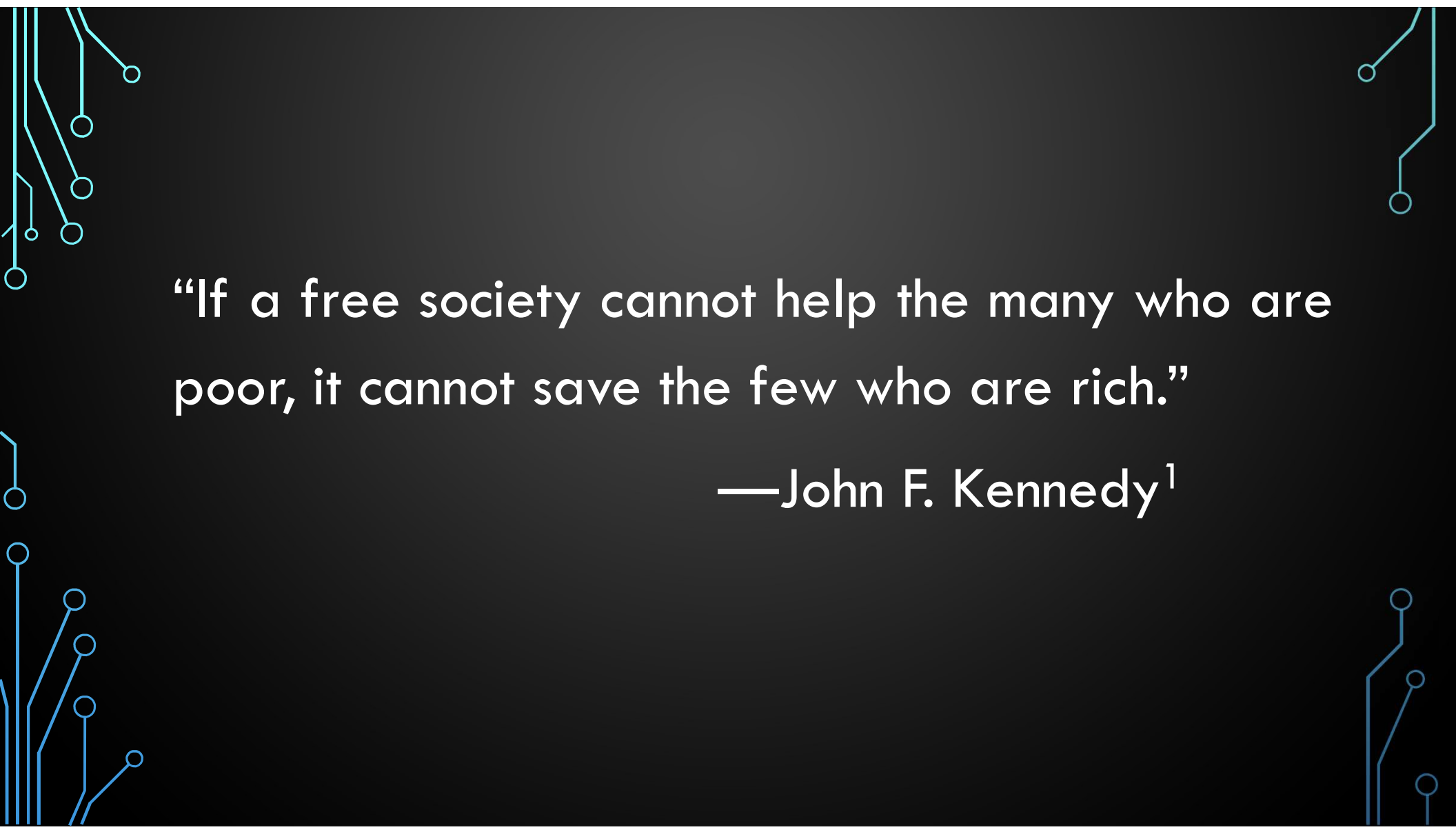
SHRUTI BIRADAR

KEVIN RUSSELL



AGENDA

- SUMMARY
 - METHODS
 - RESULTS
 - DISCUSSION
 - REFERENCES
- 
- 
- 

The image features a dark gray background with white, stylized circuit board traces in the corners. These traces consist of thin lines that branch out and terminate in small circles, resembling electronic components or nodes on a circuit. The traces are located in the top-left, top-right, bottom-left, and bottom-right corners, framing the central text.

“If a free society cannot help the many who are poor, it cannot save the few who are rich.”

—John F. Kennedy¹

An abstract graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network diagram. The lines are vertical and horizontal, with some diagonal connections, and the circles are placed at various points along these lines.

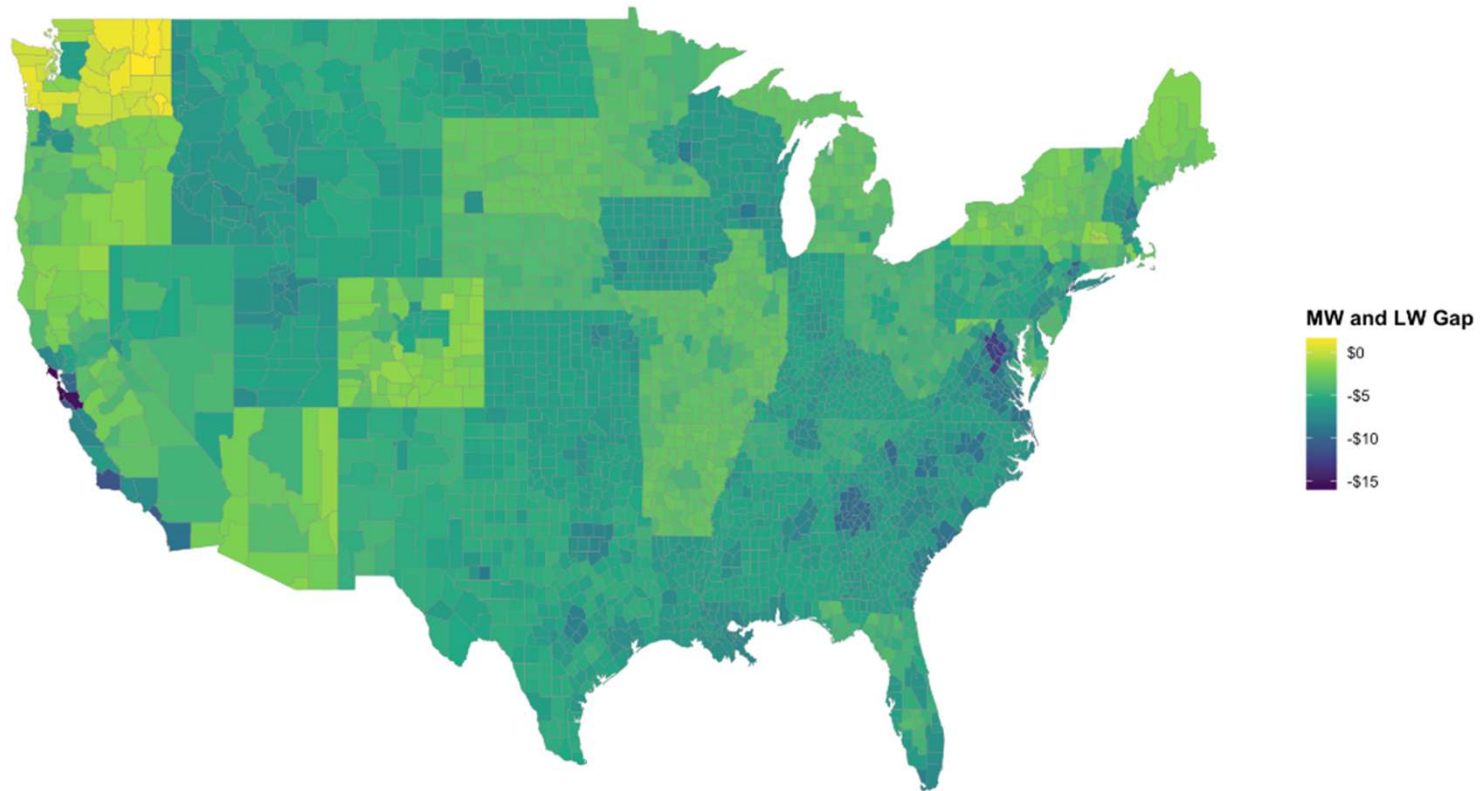
SUMMARY

SUMMARY

- The 2020 official U.S. poverty rate in was 37.2 million, 11.4% of population²
- 8.6% of Americans did not have any health insurance for the entire year, despite the onset of the COVID-19 pandemic²
- Federal Minimum Wage initiated at \$0.20 on October 24, 1938. In the 83 years since then, it has been increased 23 times, currently at \$7.25.³
- 21 states currently maintain the federal minimum wage. Five states currently have no minimum wage laws, and 2 (Georgia and Wyoming) have state minimum wage lower than the federal requirement (\$5.15/hr).⁴


SUMMARY

Gap b/w Minimum Wage-Living Wage for Single Workers by County





GOALS

- Explore the demographics of American poverty and find variables that are correlated with the poverty rate
 - Create a classification model to identify predictors for poverty rate (dependent variable: pov_pct)
 - Data based on family unit of 1 adult, 0 children
- 

DATA SETS

- Poverty Status In The Past 12 Months By Sex By Age, 2020 (U.S. Census) [1] ⁵
- MIT Living Wage Calculator Data, 2020 (Dr. Amy Glasmeier, MIT) [2] ⁶
- State Minimum Wage Data, 2020 (Bureau of Labor Statistics – BLS) [3] ⁷
- Local Area Unemployment, 2020 (BLS) [4] ⁸
- Gross Domestic Product, 2020 (Bureau of Economic Analysis – BEA) [4] ^{9,10}
- Consumer Price Index, 2020 (BLS) [5] ¹¹
- Current State Government Trifectas (Ballotpedia) ¹²

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network diagram. The lines are vertical and horizontal, with some diagonal connections, and the circles are placed at various points along these lines.

METHODS

COLLECTING AND TIDYING DATA

[1] Poverty Status Dataset: all 8 dataset CSVs imported, tidied, and joined – mutations required to calculate desired variables such as percentage of population demographics and county poverty percentage

[2] Living Wage Dataset: imported, mutation required to add State Minimum Wage Dataset [3] and calculate Minimum Wage/Living Wage gap

[4] GDP Dataset consisted of 4 separate datasets which required significant research into documentation, tidying, and joining to create one dataset

[5] CPI Dataset: most difficult dataset to tidy and join. Required significant research into documentation and county conversion

[6] State Government Dataset: created based on 2020 election results

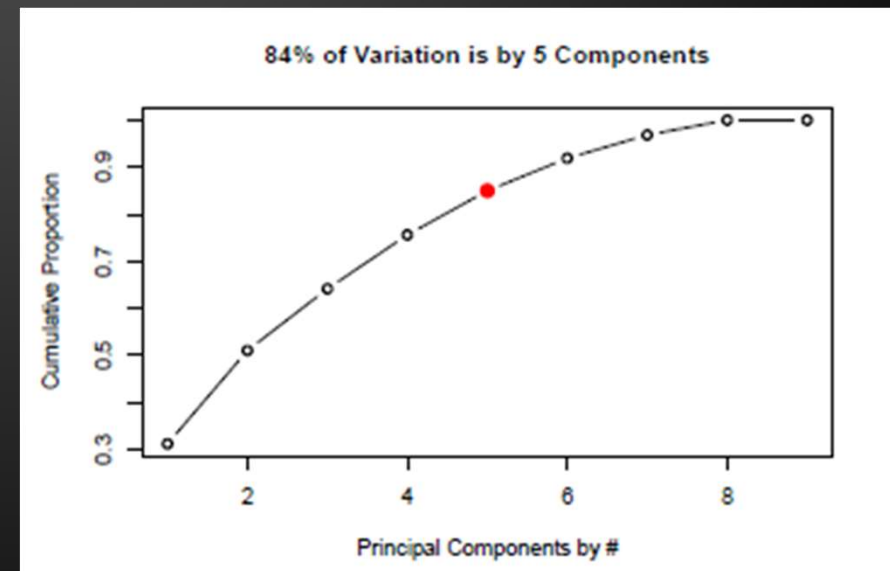
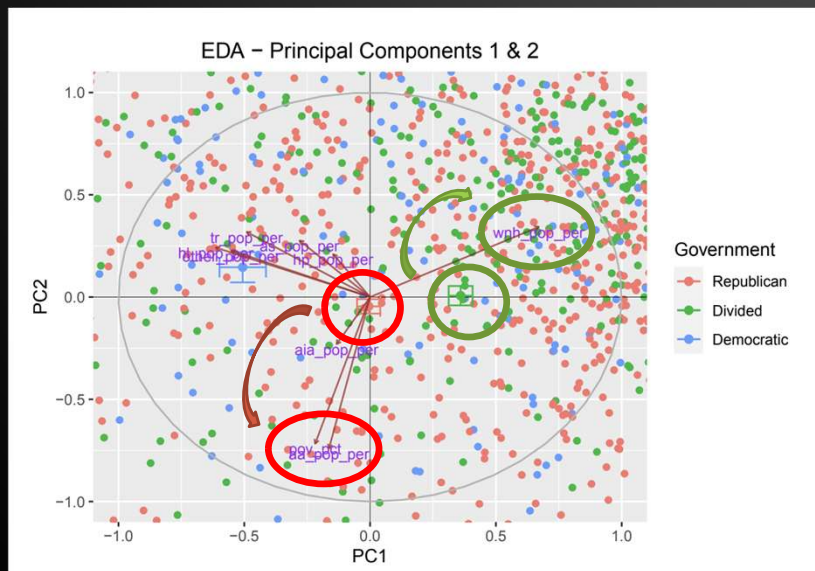
Once imported, all datasets underwent additional tidying, mutation to group by FIPS code (counties), NA values removed, and data validated before EDA.

CREATION OF NEW VARIABLES

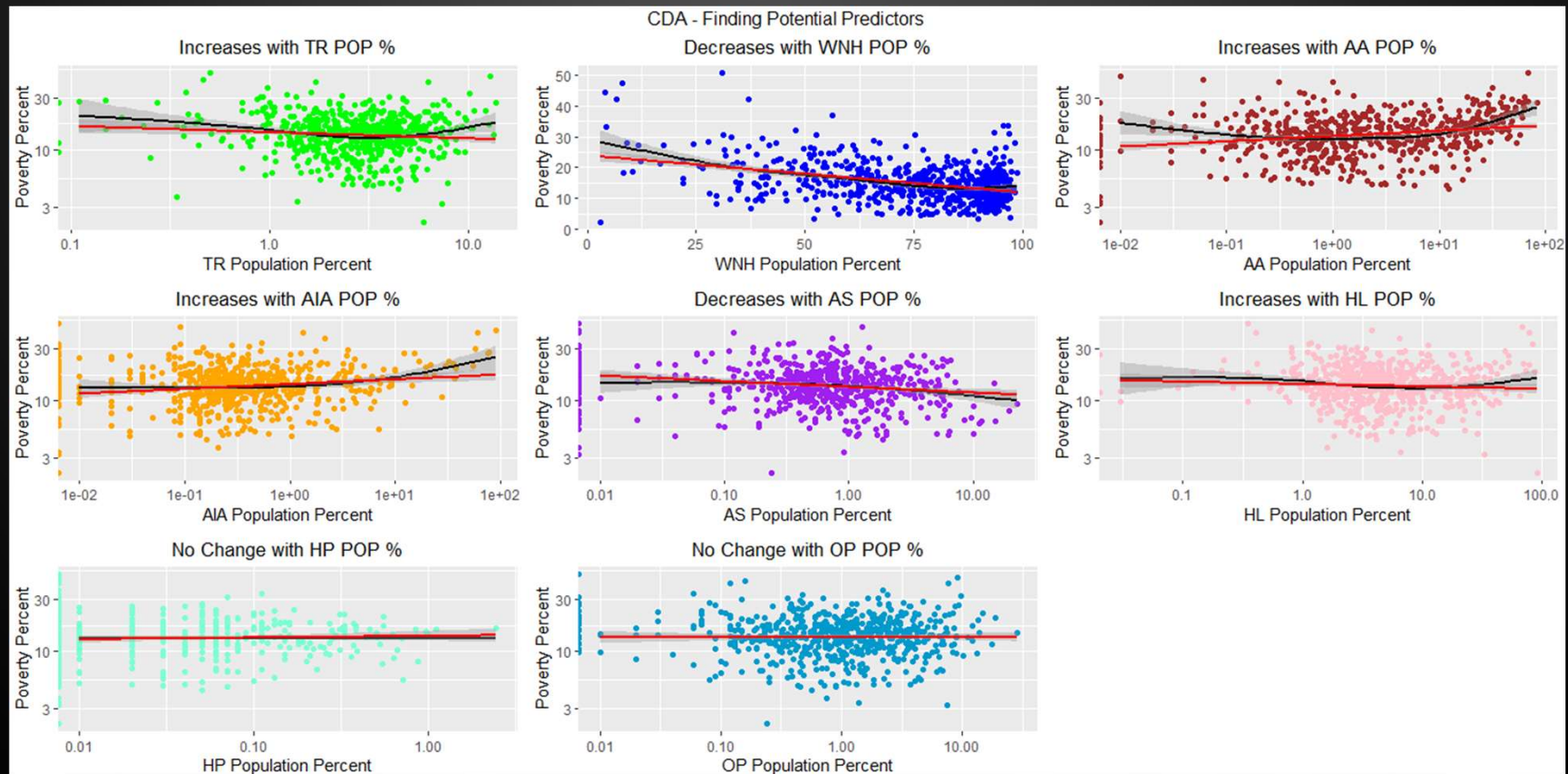
- County MW/LW Gap
 - Missing values imputed (3 counties / 3,173 counties)
 - For each county:
 - County Living Wage (1 Adult, 0 Children) [2] – State Minimum Wage [3]
 - Ranged from \$16/hour gap to \$1/hour surplus (WA)
- County Poverty Percentage
 - For each county:
 - $(\text{sum: \# of people living below poverty line, all ethnicities}) / (\text{sum: total population of county, all ethnicities})$
 - Ranged from 0% to 59.32%, with the density between 10-23%
- Ethnicity Poverty Percentage
 - For each county, for each of 8 ethnicities:
 - $\text{Count of XX Ethnicity in county living below poverty line} / \text{total population of XX Ethnicity in county}$
- Gender Poverty Percentage
 - For each county, for respondents categorized Male and Female:
 - $\text{Count of (Fe)Males in county living below poverty line} / (\text{total sum of people in county living below poverty line})$
- Government Control
 - Whether a state government is fully controlled (all three branches) by one party

EDA

- PCA Results
- Scree Plot



CANDIDATE PREDICTORS



CHOSEN PREDICTORS

- Demographics (Percent of Population)
 - “American Indian/Alaskan”
 - “White Hispanic/Latino”
 - “African American”
 - “Two or More Races”
- MW/LW Gap
- CPI
 - Communication Commodities
 - Household Furnishings And Supplies
 - Other Personal Services
 - Education
 - Transportation
 - Medical Care
- Null Hypothesis: None of the variables included in the model (see left) will influence the classification of counties as “Impoverished.”
- Alternative Hypothesis: one or more of these measurements do influence the classifying of Impoverished.
- $\alpha = 0.01$

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network diagram. The lines are vertical and horizontal, with some diagonal connections, and the circles are placed at various points along these lines.

RESULTS

MODEL FITTING

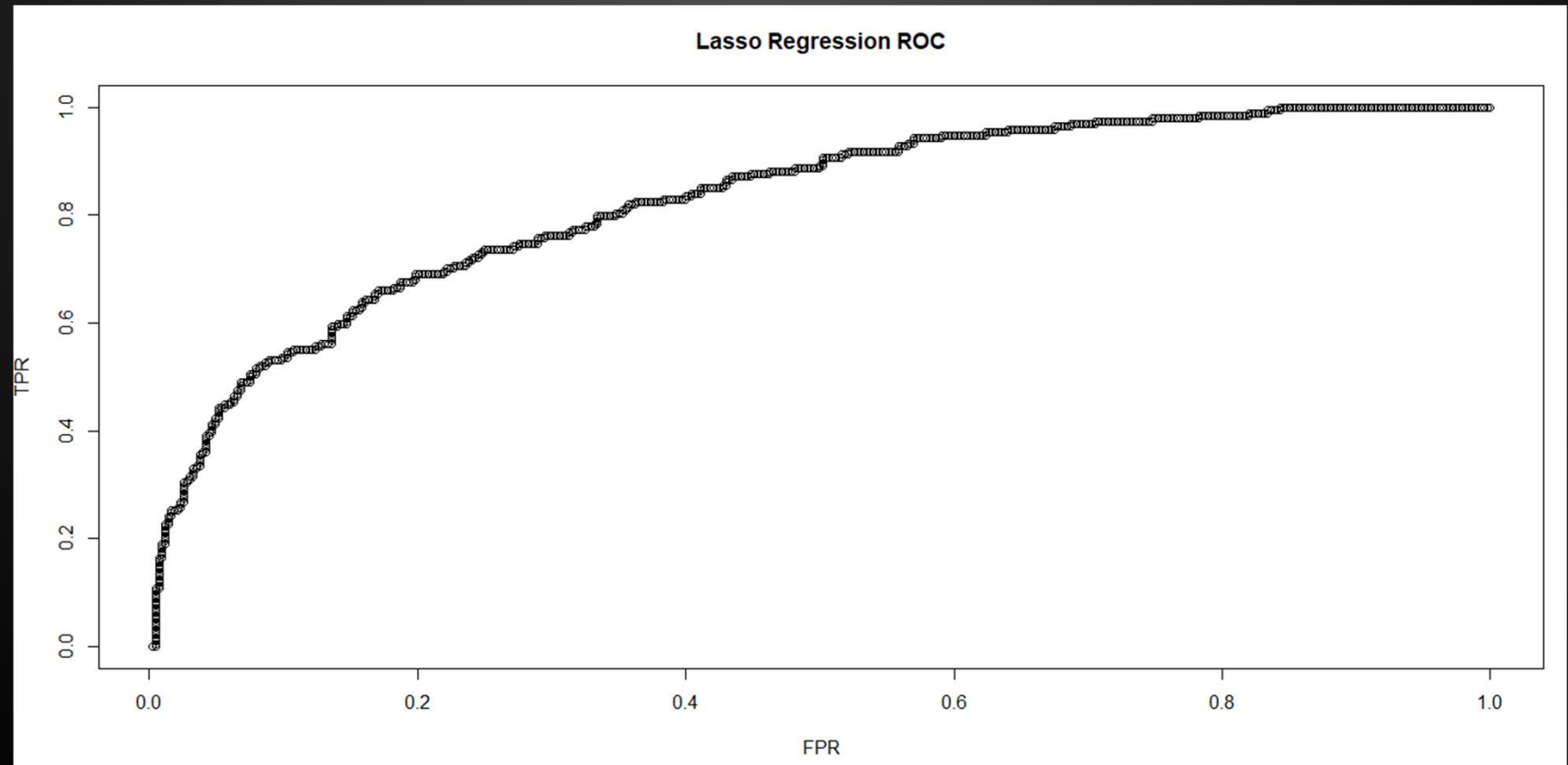
- Logistic Regression

- The fitting for logistic regression did not converge, suggesting that there are predictor variable(s) that perfectly separate the dependent variable.
- Penalized regression implemented to filter noise

- Lasso Regression

- Lasso Regression extracted MW/LW gap, cty_pop, unemployed, ue_rate, GDP, CPI_Dairy_And_Related_Products, CPI_Education_And_Communication_Commodities, CPI_Education_And_Communication_Services, CPI_Electricity, CPI_Gasoline_Unleaded_Regular, CPI_New_Vehicles, CPI_Other_Personal_Services, CPIUtility(Piped)_Gas_Service, and gov_party_id as selected predictors.

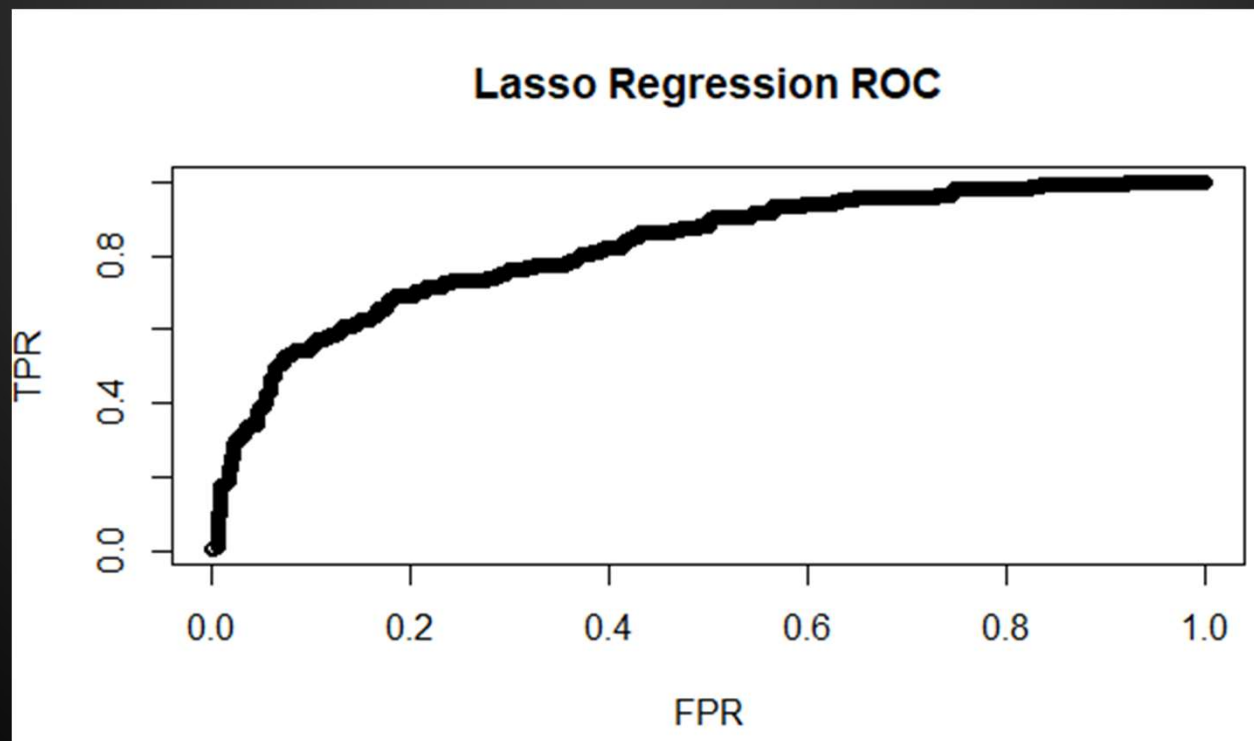
LASSO REGRESSION



CONFUSION MATRIX - LASSO

```
## Confusion Matrix and Statistics
##
##               Reference
## Prediction      Impoverished Not Impoverished
##   Impoverished             383             88
##   Not Impoverished         45             106
##
##               Accuracy : 0.7862
##               95% CI : (0.7518, 0.8178)
##   No Information Rate : 0.6881
##   P-Value [Acc > NIR] : 3.079e-08
##
##               Kappa : 0.4697
##
##   Mcnemar's Test P-Value : 0.0002707
##
##               Sensitivity : 0.8949
##               Specificity : 0.5464
##               Pos Pred Value : 0.8132
##               Neg Pred Value : 0.7020
##               Precision : 0.8132
##               Recall : 0.8949
##               F1 : 0.8521
##               Prevalence : 0.6881
##               Detection Rate : 0.6158
##               Detection Prevalence : 0.7572
##               Balanced Accuracy : 0.7206
##
##               'Positive' Class : Impoverished
##
```

ELASTIC NET REGRESSION – RIDGE AND LASSO



CONFUSION MATRIX – ELASTIC NET

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction   Impoverished Not Impoverished
## Impoverished      383          84
## Not Impoverished   45         110
##
##              Accuracy : 0.7926
##              95% CI : (0.7586, 0.8238)
##      No Information Rate : 0.6881
##      P-Value [Acc > NIR] : 3.547e-09
##
##              Kappa : 0.4887
##
##  Mcnemar's Test P-Value : 0.0008207
##
##      Sensitivity : 0.8949
##      Specificity : 0.5670
##      Pos Pred Value : 0.8201
##      Neg Pred Value : 0.7097
##      Precision : 0.8201
##      Recall : 0.8949
##      F1 : 0.8559
##      Prevalence : 0.6881
##      Detection Rate : 0.6158
##      Detection Prevalence : 0.7508
##      Balanced Accuracy : 0.7309
##
##      'Positive' Class : Impoverished
##
```

```
word
(Intercept)
fips
aa_pop_per
tr_pop_per
hl_pop_per
aia_pop_per
employed
unemployed
ue_rate
livingwage
X1A_OC_Diff
GDP
CPI_Durables
CPI_Education_And_Communication_Services
CPI_Electricity
CPI_Fuels_And_Utillities
CPI_Meats_Poultry_Fish_And_Eggs
CPI_Medical_Care
CPI_Motor_Vehicle_Insurance
CPI_New_And_Used_Motor_Vehicles
CPI_New_Vehicles
CPI_Nonalcoholic_Beverages_And_Beverage_Materials
CPI_Other_Goods_And_Services
CPI_Other_Personal_Services
CPI_Recreation
CPI_Tuition_Other_School_Fees_And_Childcare
CPI_Utility_(Piped)_Gas_Service
gov_party_id
```

The image features a dark gray background with the title 'SUPPORT VECTOR MACHINE' in white, uppercase letters. The corners of the image are decorated with light blue circuit-like patterns, consisting of lines and small circles, resembling a stylized electronic board.


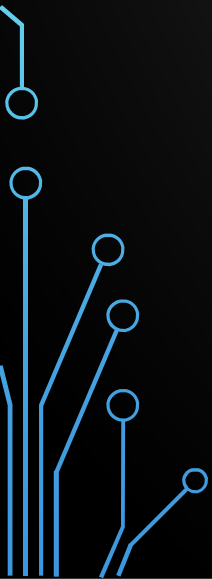

SUPPORT VECTOR MACHINE

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network, set against a dark grey background.

DISCUSSION



SUMMARY


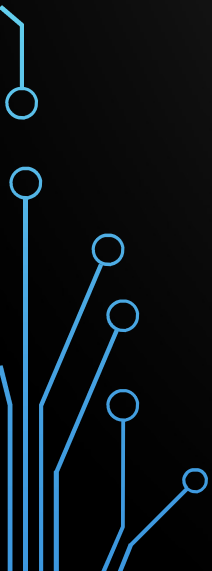

- Predictors in model based on population data and consumer data
 - Higher minority composition in a county does correlate to higher poverty
 - Consumer metrics correlated with poverty are primarily necessities
 - These results can be used to target financial aid policies and social programs, ideally decreasing poverty nationwide
 - Although clearly not a blanket solution, legislating an increase of the minimum wage to be at least the living wage would relieve a significant amount of financial pressure
- 
- 
- 

LIMITATIONS

- Missing Data
 - Insufficient Data:
 - Alaska
 - Hawaii
 - Puerto Rico
 - Living Wage Data:
 - Okaloosa County, FL
 - St. Martin's Parish, LA
 - Currituck County, NC
- Data does not include any information on use of social programs (i.e. WIC, SNAP, unemployment benefits, COVID-19 stimulation checks, etc.)
- Unclear whether data captures homeless population



FUTURE WORK

- Family size
 - Age and gender data
 - Dominant industry in each county
 - Inclusion of welfare program/benefit use
 - Predictive model
- 
- 
- 

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network diagram. The lines are vertical and horizontal, with some diagonal connections, and the circles are placed at various points along these lines.

REFERENCES

REFERENCES

1. Kennedy, J. F. (1963). *Inaugural Address*. Library of Congress. [Online]. Available: https://www.loc.gov/static/programs/national-recording-preservation-board/documents/JFK_essay.pdf [accessed 7 April, 2022].
2. U.S. Census Bureau. (2021). "Income, Poverty and Health Insurance Coverage in the United States: 2020". Publication: 14 September, 2021. [Online]. Available: <https://www.census.gov/newsroom/press-releases/2021/income-poverty-health-insurance-coverage.html#:~:text=The%20official%20poverty%20rate%20in,million%20more%20than%20in%202019> [accessed 17 March, 2022].
3. US Department of Labor. (n.d.). "History of Federal Minimum Wage Rates Under the Fair Labor Standards Act, 1938 - 2009". Wage and Hour Division. [Online]. Available: <https://www.dol.gov/agencies/whd/minimum-wage/history/chart> [accessed 24 March, 2022].
4. US Department of Labor. (2021). "State Minimum Wage Laws". Wage and Hour Division. [Online]. Available: <https://www.dol.gov/agencies/whd/minimum-wage/state>. [accessed March 12, 2021].
5. American Community Survey, U.S. Census Bureau. "Poverty Status In The Past 12 Months By Sex By Age". Publication: 2020. [Online]. Multiple datasets disaggregated by race/ethnicity available: <https://data.census.gov/cedsci/table?q=United%20States&t=Income%20and%20Poverty%3ARace%20and%20Ethnicity&g=0100000US,%240500000&y=2020&tid=ACSDT5Y2020.B17001B> [accessed 22 March, 2022].
6. Glasmeier, A. K., Department of Urban Studies and Planning, Massachusetts Institute of Technology. "MIT Living Wage Calculator." [Online] Available: <https://livingwage.mit.edu/> [accessed 26 February, 2022].
7. Department of Labor. "Changes in Basic Minimum Wages in Non-Farm Employment Under State Law: Selected Years 1968 to 2021". [Online]. Available: <https://www.dol.gov/agencies/whd/state/minimum-wage/history> [accessed 27 February, 2022].
8. Bureau of Labor Statistics. "Local Area Unemployment Statistics". Publication: 2 March, 2022. [Online]. Available: <https://www.bls.gov/lau/laucnty20.xlsx> [accessed 25 March, 2022].
9. Bureau of Economic Analysis. "2020 GDP Dataset". [Online, three files]. Available: <https://download.bls.gov/pub/time.series/cu/cu.series>, <https://download.bls.gov/pub/time.series/cu/cu.data.0.Current>, <https://download.bls.gov/pub/time.series/cu/cu.item> [accessed 23 March, 2022].
10. Bureau of Economic Analysis. "Real Gross Domestic Product, by County, 2017-2020." [Online]. Available: <https://www.bea.gov/sites/default/files/2021-12/lagdp1221.xlsx> [accessed 21 March, 2022].
11. Bureau of Labor Statistics. "Average Price Data – Consumer Price Index (CPI)." [Online]. Available: <https://www.bls.gov/cpi/tables/supplemental-files/home.htm> [accessed 13 March, 2022].
12. Ballotpedia. "Current State Government Trifectas". [Online]. Available: https://ballotpedia.org/Gubernatorial_and_legislative_party_control_of_state_government [accessed 23 March, 2022].

A decorative graphic on the left side of the slide, consisting of a network of light blue lines and small circles, resembling a circuit board or a neural network, set against a dark grey background.

QUESTIONS