

# Capstone Checkpoint Two

SURE 2024

Department of Statistics and Data Science  
Carnegie Mellon University

---

Jessica Lambert, Rohan Devraj, Jade Kay, Brooke Benton

# How are juvenile healthcare outcomes impacted by adult health-related practices?

- Examine how adult behaviors impact quality and length of life in children under 20
  - Intergenerational health impacts
  - Early intervention and prevention
  - Healthcare policy and education

# County Health Rankings Data

## Adult Behaviors

- Excessive Drinking
- Smoking
- Diet
- Physical inactivity
- Insufficient sleep
- Food Insecurity
- Uninsured
- STI's

## Juvenile Outcomes

- Child Mortality Rate
- Low Birthweight

## Data Cleaning Process

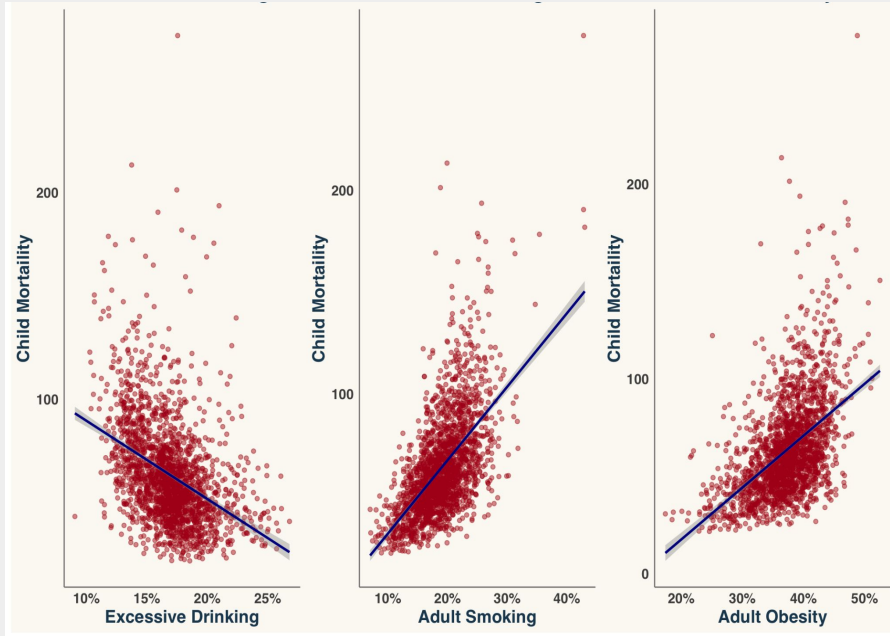
- Remove first row of variable description
- Remove “na” values
- Remove states mistakenly listed as counties
- Creation of binary response variable for insufficient sleep (high/low proportion of population)

# Plan of Action (Completed Steps)

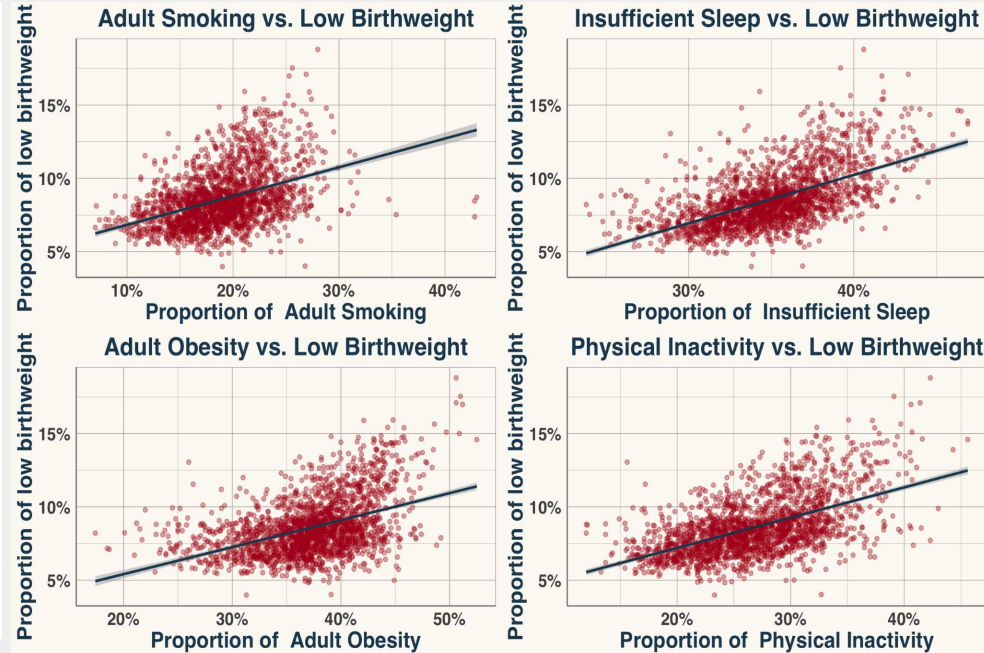
- Data wrangling/cleaning
- Choice of variables for analysis
- EDA
- Use lasso regression to determine which factors best predict low birthweight and child mortality
- Model using decision trees
- Creation of binary response variable for insufficient sleep

# Excessive drinking, adult smoking, and other adult health behaviors are associated with child mortality and low birthweight

## Child Mortality



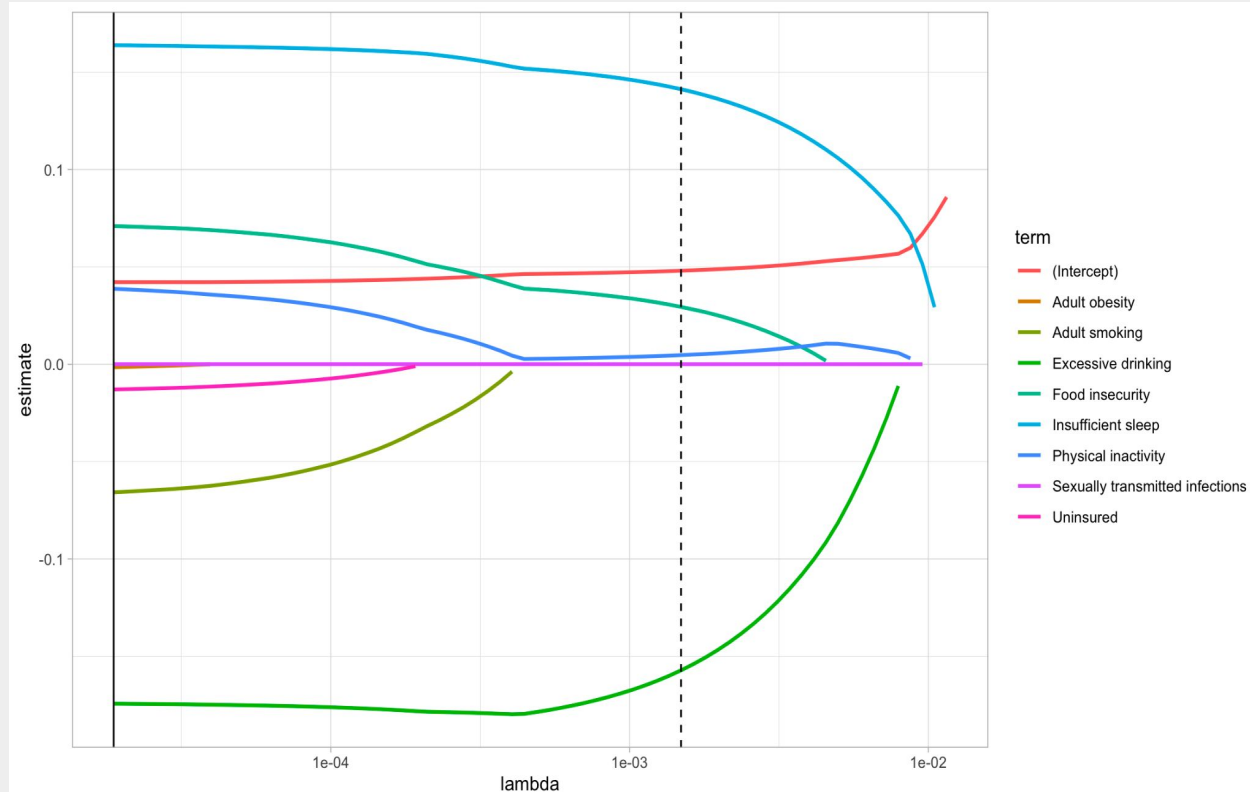
## Low birthweight



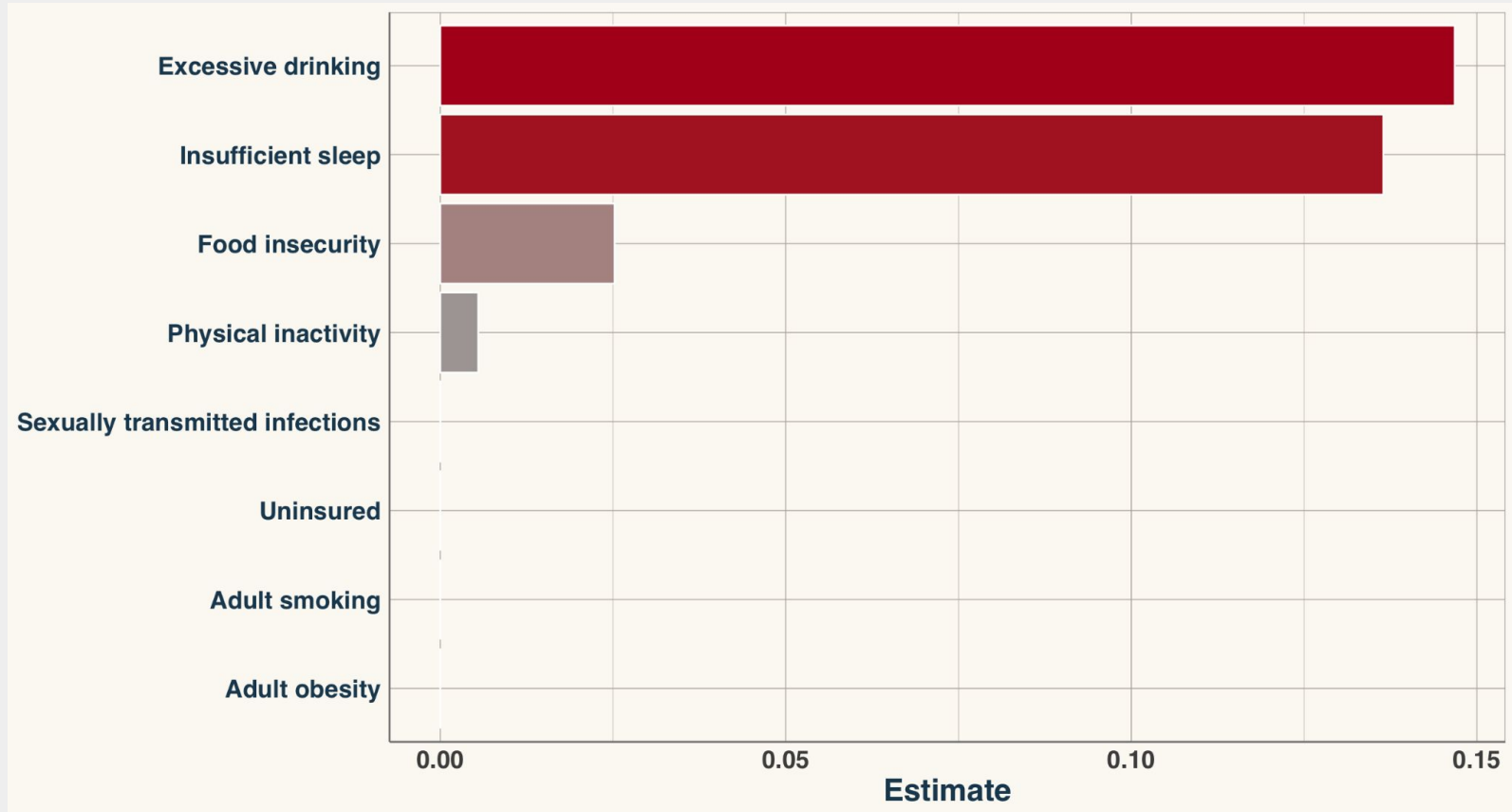
# Lasso Regression for predicting low birthweight

Variable selection and shrinkage penalty for preventing overfitting and greater interpretability

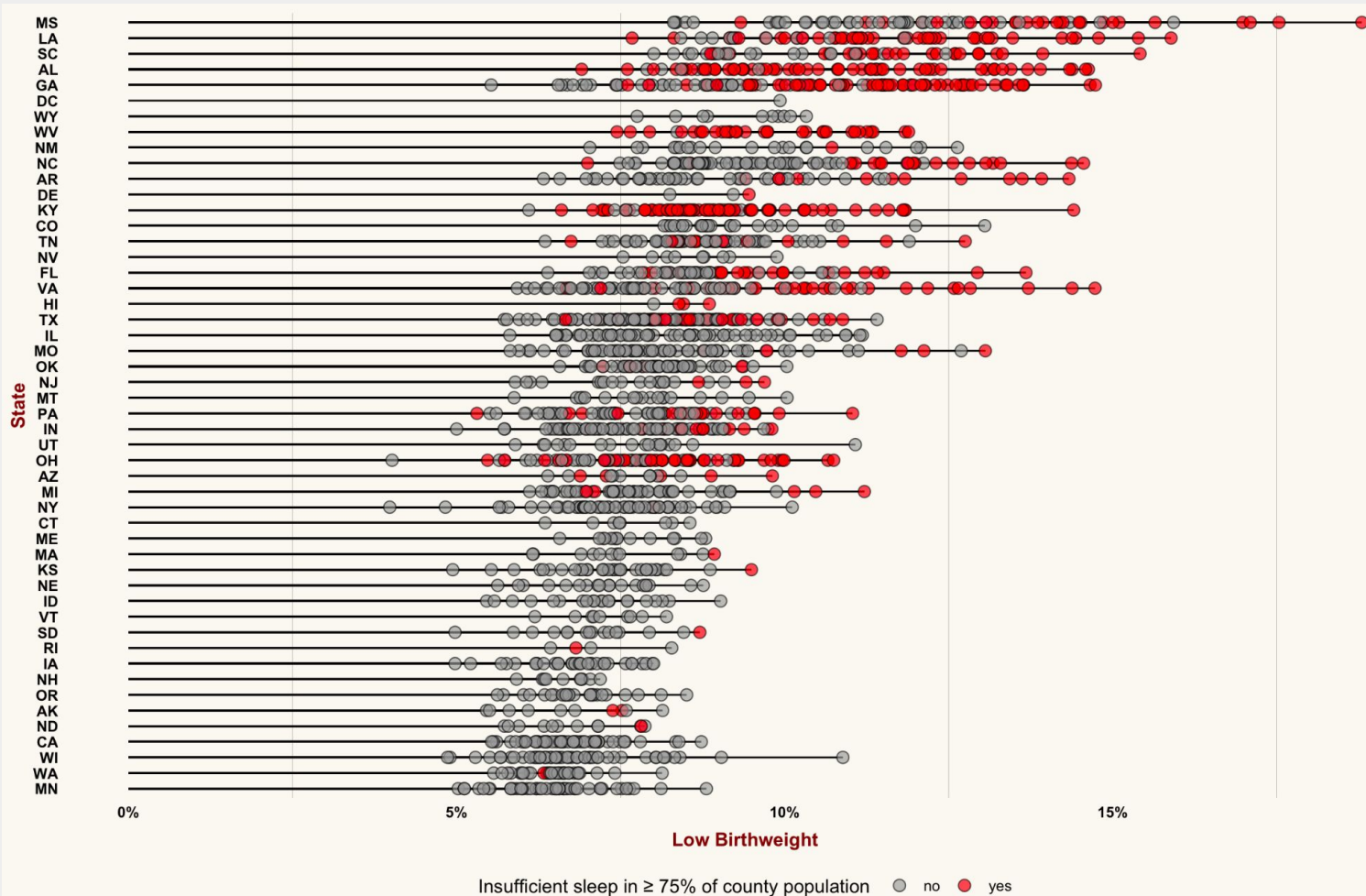
- $\lambda$  was picked with 10-fold cross validation (cv.glmnet)
- Solid line on the left is the  $\lambda_{\text{min}}$  that gives the minimum mean cross-validated error.
- Dashed line is the largest value of  $\lambda$  such that the cross-validated error is within one standard error of the minimum mean cross-validated error



# Excessive drinking and insufficient sleep are the strongest predictors of low birthweight

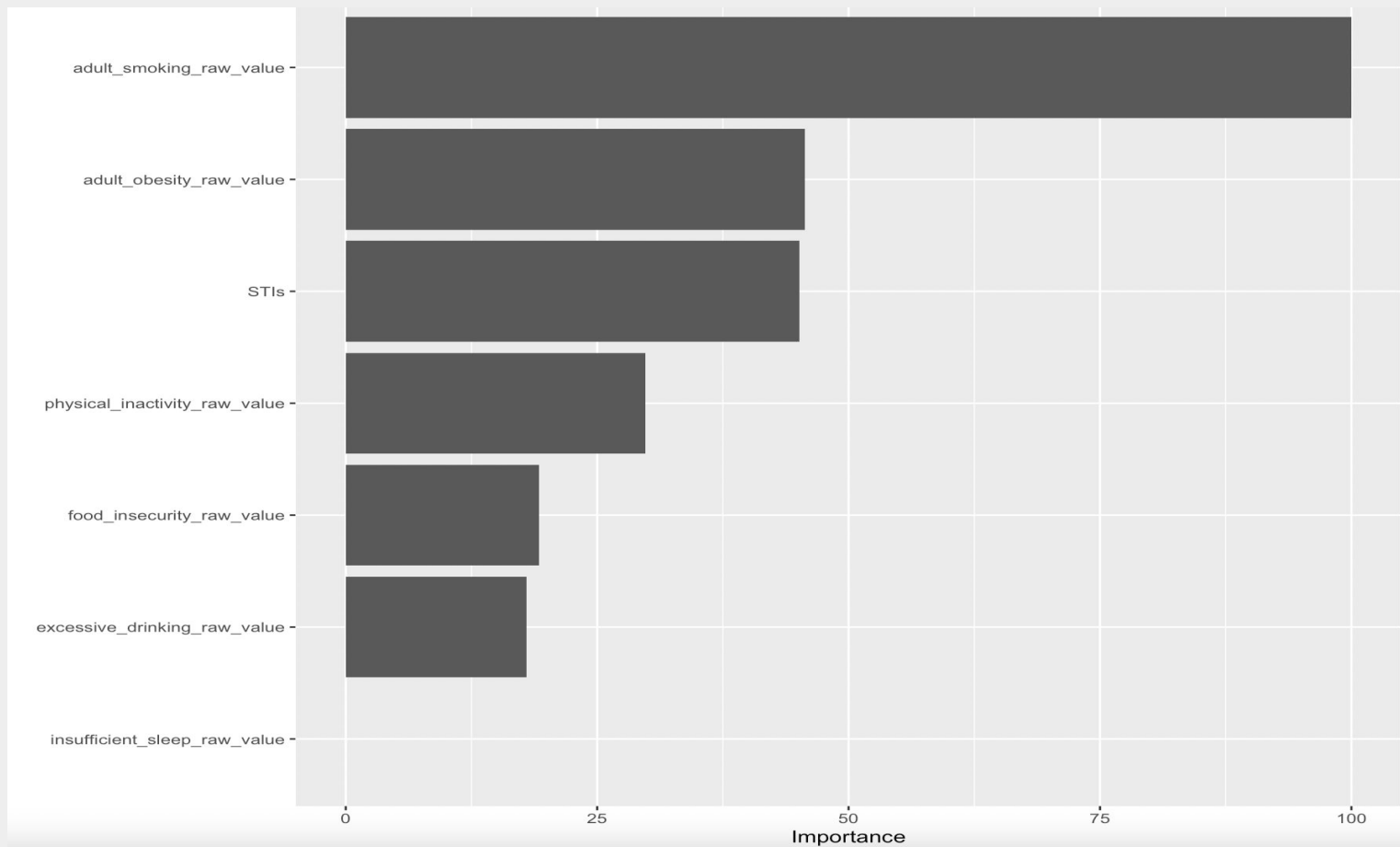


# Across the United States





# Variable Importance Graph



# Future Steps

- Editing aesthetics of graphs
- Finalize inferences/takeaways from models
- Finish report and poster
- Investigating potential limitations
  - Gender with excessive drinking
  - Parent to child data
- Investigating future work
  - Intervention programs e.g. (where and who)