

# Project Paper 1

Dr. Kris Pruitt

15 March 2024

## The Response Variable

In this paper, we have decided to study the response variable `peopleSum`, which we have created by summing all of the columns in the data that measure how many people work in each industry in a county. Importantly, this sum is distinct from the total amount of men employed in each county, which are counted by different columns. The amount of men working in each county and the amount of people working in each county is distinct. `peopleSum` sums across the columns indicating Persons Workin in Mining, Agriculture, Commerce, Manufacturing, Ocean Navigation, Inland Navigation and “Proffesions” (lawyers, politicians etc.). As such, it represents the total amount of people in the county who were employed in 1840. The response variable is distributed with a strong right skew, with most counties having under 10,000 people working, but a few outliers with more than 20,000 people working

