

EDS 241
Assignment 1

Due on 1/21/22 at 17:00

Turn in your Markdown pdf on Gauchospace in the “Assignment Turn in Area”

The data for this assignment come from CalEnviroScreen 4.0, a mapping and data tool produced by the California Office of Environmental Health Hazards Assessment (OEHHA). The data are compiled and constructed from a variety of sources and cover all 8,035 census tracts in California. Source: <https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-40>

The full data are contained in the file CES4.xls, which is available on Gauchospace (note that the Excel file has three “tabs” or “sheets”). The data is in the tab “CES4.0FINAL_results” and “Data Dictionary” contains the definition of the variables.

For the assignment, you will need the following variables: **CensusTract**, **TotalPopulation**, **CaliforniaCounty** (the county where the census tract is located), **LowBirthWeight** (percent of census tract births with weight less than 2500g), **PM25** (ambient concentrations of PM2.5 in the census tract, in micrograms per cubic meters), and **Poverty** (percent of population in the census tract living below twice the federal poverty line).

- (a) What is the average concentration of PM2.5 across all census tracts in California?
- (b) What county has the highest level of poverty in California?
- (c) Make a histogram depicting the distribution of percent low birth weight and PM2.5.
- (d) Estimate a OLS regression of LowBirthWeight on PM25. Report the estimated slope coefficient and its heteroskedasticity-robust standard error. Interpret the estimated slope coefficient. Is the effect of PM25 on LowBirthWeight statistically significant at the 5%?
- (e) Suppose a new air quality policy is expected to reduce PM2.5 concentration by 2 micrograms per cubic meters. Predict the new average value of LowBirthWeight and derive its 95% confidence interval. Interpret the 95% confidence interval.
- (f) Add the variable Poverty as an explanatory variable to the regression in (d). Interpret the estimated coefficient on Poverty. What happens to the estimated coefficient on PM25, compared to the regression in (d). Explain.
- (g) From the regression in (f), test the null hypothesis that the effect of PM2.5 is equal to the effect of Poverty