Codebook for SAT data

Constantin Manuel Bosancianu
manuel.bosancianu@outlook.com
Wissenschaftszentrum Berlin
Institutions and Political Inequality

February 26, 2019

The data contains information on the average SAT score in each US state, collected around 1990-1991 in the United States. The data set is taken from Lawrence C. Hamilton, "Statistics with STATA 9", Duxbury Press, 2009. The original data set had a few more variables, but for the sake of keeping things manageable, I excluded them from the data.

The following variables can be found in the data:

1. `state`: state name;

2. `region`: geographic region in the United States;

3. `pop`: population of the state (information from the 1990 Census);

4. `area`: land area, in square miles;

5. `density`: population density (number of people per square mile);

6. `metro`: metropolitan area population, as percentage of the entire state population;

7. `waste`: per capita solid waste, measured in tons;

8. `energy`: per capita energy consumption, measured in BTUs;

9. `toxic`: per capita toxins released in the environments, measured in pounds;

10. `green`: per capita greenhouse gas emissions, measured in tons;

11. `csat`: average composite SAT score (verbal plus math);

12. `vsat`: average verbal SAT score;

13. `msat`: average mathematics SAT score;

14. `percent`: percentage of high school graduates taking the SAT;

15. `expense`: expenditure per student on primary and secondary education, measured in USD;

16. `income`: median household income, expressed in 1,000s of USD;

17. `high`: percentage of adults with a high school diploma;

18. `college`: percentage of adults with a college degree.