Data Assignment

Econometrics I

February 17, 2021

Instructions

a) The data for this assignment is present in the file dataset.csv.

b) You must submit your code as .R file and a pdf report containing your results.

c) Institute plagiarism rules apply.

Question 1

Education is considered as one of the most important tool that helps in economic and social mobility. Education is included in the UN's Sustainable Development Goals 2030. Given the importance of education, a significant portion of our children are still out of school. Gross Enrolment Ratio or GER measures this number. It is the ratio of total children enrolled in the schools and the total number of children in a territory. Through this assignment we'll study GER and school infrastructure of Indian states.

Description of variables:

STATE: State name

Year: Year for which data is reported

DW: percentage of schools having drinking water facility

EL: percentage of schools with electricity facility

BT: percentage of schools with boys toilets

GT: percentage of schools with girls toilets

CA: percentage of schools with computer facilities

GER: Gross-enrolment ratio of the state

a) Find the average GER of India from 2013-14 to 2015-16.

b) What percentage of schools on an average are electrified?

c) What percentage of schools on an average have drinking water facility?

d) What percentage of schools on an average have boys toilets?

e) What percentage of schools on an average have girls toilets?

f) What percentage of schools on an average have computer labs?

g) Report variance for the above variables.

h) Create a histogram for GER, drinking water availability, difference of boys and girls toilet availability.

i) Considering your findings from the above exercise, what do you think about the state of school infrastructure and enrollment in India?

Question 2

Now, create divide all the states into three groups (High, Medium, Low) according to their literacy rate in 2011. You can find the literacy rates here: `https://www.census2011.co.in/literacy.php` Add this column to your data and name it LIT.

a) Find mean GER, mean boys toilet availability, girls toilet availability, drinking water availability, electricity availability for these three groups.

(Note: For better representation you can create a table for these values)

b) From the findings, do you observe any pattern among the three groups in terms of relation between enrolment and school infrastructure? Explain.

Now, let's divide these states according to the geographical and administrative criteria into the following groups:

North-east and Hilly States: HIMACHAL PRADESH, ,UTTARAKHAND, MANIPUR, SIKKIM, MIZORAM, ASSAM, ARUNACHAL PRADESH, NAGALAND, TRIPURA, MEGHALAYA

UTs and City states: DELHI, CHANDIGARH, DAMAN & DIU, PUDUCHERRY, GOA, DADRA & NAGAR HAVELI, ANDAMAN & NICOBAR ISLANDS, JAMMU KASHMIR, LAKSHADWEEP

Southern States: KARNATAKA, MAHARASHTRA, TAMIL NADU, TELANGANA, KERALA, ANDHRA PRADESH, ODISHA

Other Major states: HARYANA, GUJARAT, PUNJAB, UTTAR PRADESH, RAJASTHAN, MADHYA PRADESH, JHARKHAND, BIHAR, CHHATTISGARH, WEST BENGAL

c) For GER, create a table displaying comparison between mean GER and variance of GER between different groups. Do this for all other variables.

d) **Question 3**

a) Set up a regression for GER, using all the covariates except state name and year.

(i) Estimate and report the regression coefficients

(ii) Estimate and report the variance of regression coefficients

(iii) Estimate and report the $\hat{\sigma}^2$.

b) Incorporate dummy variables in the above regression equation for the groups made in 2 (a) and carry out the below exercise.

(i) Estimate and report the regression coefficients.

(ii) Estimate and report the variance of regression coefficients.

(iii) Estimate and report the $\hat{\sigma}^2$.

c) Incorporate dummy variabes in the above regression equation for the groups made in 2 (c) except state name and year and estimate regression.

(i) Estimate and report the regression coefficients

(ii) Estimate and report the variance of regression coefficients

(iii)Estimate and report the $\hat{\sigma}^2$.