

Term Project (Individual)

Topic: Socioeconomic factors on HIV/AIDS: a cross-country comparative study

- Your project cannot exceed 10 pages [double-line spacing (in Word, line spacing, use 1.5), font size = 12] (including everything except the cover page).
- The project should be submitted on or before December 22, 2020 (6:00pm). Points will be deducted for late submission (20 percent per day).

Data set: aids.xlsx

(Raw data extracted from Gapminder (<http://www.gapminder.org/>) and related websites)

The description of the data is given in the Appendix.

(Do not use extra data, except possibly the geographic location classification)

Perform a regression analysis and use *Aids* as the dependent variable.

(In English)

In your report,

- Clearly identify your research objectives and select a pool of potential predictors.
- In the model building process
 - State your assumptions of the model
 - Provide your variable selection method
 - Include diagnostics techniques
 - Give reasons to support your choice of the final model
- Present your final model with some explanations to help readers to relate your statistical findings to the research problem and objectives
- Identify outliers and search for appropriate underlying reasons
- Insert tables and graphs only if they are essential
- Give the list of references

Appendix: Description of the data.

Variable Name	Description
Country	
Prevalence of HIV/AIDS among people 15-49 years (%)	Percentage of people between 15 and 49 years-of-age living with HIV/AIDS
Agriculture, contribution to economy (% of GDP)	Agriculture includes forestry, hunting, and fishing, as well as cultivation of crops and livestock production. Value added is the net output of a sector after adding up all outputs and subtracting intermediate inputs. It is calculated without making deductions for depreciation of fabricated assets or depletion and degradation of natural resources.
Children per woman (total fertility rate)	Total fertility rate. The number of children that would be born to each woman with prevailing age-specific fertility rates.
General government expenditure on health as percentage of total government expenditure	General government expenditure on health as percentage of total government expenditure
Infant mortality rate (per 1,000 live birth)	The probability that a child born in a specific year will die before reaching the age of one, if subject to current age-specific mortality rates. Expressed as a rate per 1,000 live births
Improved sanitation facilities, urban (% of urban population with access)	The percentage of the population with at least adequate access to excreta disposal facilities, such as a connection to a sewer or septic tank system, a pour-flush latrine, a simple or improved ventilated improved pit latrine, that can effectively prevent human, animal, and insect contact with excreta
Prevalence of current tobacco use among adults (≥ 15 years) (%) female	Prevalence of current tobacco use among adults (≥ 15 years) (%) female
Prevalence of current tobacco use among adults (≥ 15 years) (%) male	Prevalence of current tobacco use among adults (≥ 15 years) (%) male
Urban population (% of total)	