

Department of Statistics, The Chinese University of Hong Kong  
STAT 5102 Regression in Practice (Term 1, 2018–19)

GROUP PROJECT

Specifications:

1. This is a group project. Each group should be comprised of four registered students in this course.
2. Each group should submit a report with the following format requirements:
  - At most 12 pages (these 12 pages, however, need not include the cover page and the reference);
  - The report should be saved in .pdf format;
  - Page layout: 1-inch on each side of A4 sheets with 1.5 line spacing and
  - Font: Times News Roman with size 11.
3. Each group is required to submit the final report on a later agreed due date. Points will be deducted for late submission.

Data source:

The dataset is obtained from World Bank:

<http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=world-development-indicators#>

The dependent variable (quantity of interest) is Mortality rate (under-5).

A zipped dataset can also be found on Blackboard for easy download.

Required content:

In your report, you are advised to

1. clearly identify your research objectives and select a pool of potential predictors;
2. state your assumptions that are imposed on the model;
3. clarify your variable selection method;
4. include details on model diagnostics;
5. provide reasons to support your choice of the final model and/or
6. present your final model with justifications to help readers relate your statistical findings to the research problem and objectives.

On Alternative Topics:

Students may choose other datasets of interest for this project. Groups that prefer to analyse alternative datasets should submit brief proposals by 7/Nov in which main project objectives and lists of variables considered should be included.