

# Essayopgave 1

## Linear Regression and Model-building

**Due: 28. February at 12:00 to Thomas via email (tleeper@ps.au.dk)**

*Consider the following hypothesis and analysis plan:*

Compared to non-democracies, democratic states produce better health outcomes for their citizens. This expectation is tested with a large-n analysis of a cross-section of countries.

*Your task:*

1. Write a research question, which could be addressed by the above-stated hypothesis and plan of analysis.
2. Write a short research paper (3–5 pages) that incorporates your research question, the above hypothesis, your operationalizations of “democracy” and “health outcomes,” the specification of your causal model with a discussion of included variables and how they address concerns about omitted variable bias, and a discussion of your results. You do not need to cite outside literature nor elaborate any theory to justify the provided hypothesis. The paper should include:
  - A brief discussion of the measures
  - Identification of possible confounding variables, which might bias the observed bi-variate relationship between X and Y
  - Graphical and/or correlational methods to test whether any of the putative confounding factors should be in the regression model
  - A comparison of your full model to a “reduced” model with fewer covariates than the full model (e.g., using `nestreg`)
  - A clear interpretation of the size of the relationship (if any) between X and Y
3. Include your complete Stata do-file for all included analyses as an appendix to the paper. (Note: Use a fixed-width font (e.g., `Courier`) for this appendix to aid readability.)

*Data:*

Visit the website for The Quality of Government Institute at the University of Gothenburg. There you will find the “Quality of Government Standard Cross-Section Data.”<sup>1</sup> You can download the Stata version of the data (.dta file extension) and a codebook describing the data.

*Variables:*

The hypothesis includes two constructs: democracy and health outcomes. Neither of these constructs has a single clear operationalization. You will have to choose variables included within the QoG dataset to operationalize each. You may want to consider, for democracy, the Freedom House/Polity index (`fh_polity2`, `fh_ipolity2`) or Economist Intelligence Unit Index of Democracy (`eiuiod`). For health outcomes, you may want to consider infant mortality (e.g., `wdi_mort`) or life expectancy (e.g., `wef_lifexp`). Justify your choice in any case.

---

<sup>1</sup>A direct link is: <http://www.qog.pol.gu.se/data/datadownloads/qogstandarddata/>.