Midterm coursework

Introduction to Quantitative Research Methods (PUBLG100A/B)

- The coursework will be posted on Moodle on 3 November 2017, and is due on 8 November 2017. Please follow all designated SPP submission guidelines for online submission as detailed on the PUBLG100A/B Moodle page. Late submission results in an automatic fail.
- This is an assessed piece of coursework (worth 25% of your final module mark) for the PUBLG100A/B module; collaboration and/or discussion of the coursework with anyone is strictly prohibited. The rules for plagiarism apply and any cases of suspected plagiarism of published work or the work of classmates will be taken seriously.
- As this is an assessed piece of work, you may not email/ask the course tutors or teaching fellows questions about the coursework.
- Along with the coursework itself, the datasets for the coursework can be found in the PUBLG100A/B Midterm Coursework folder on Moodle.
- The coursework consists of four sections; you must complete each part of each section to achieve full marks.
- Where appropriate, answers should be written in complete sentences. Be sure to answer all parts of the questions posed and interpret the results.
- PLEASE SUBMIT YOUR TYPE-WRITTEN ANSWERS IN ONE DOCUMENT. CREATE AT THE END AN APPENDIX SECTION CONTAINING ALL R CODE NEEDED TO REPRODUCE YOUR RESPONSES (you do not need to include the code that failed to run, but just the cleaned-up version. Your code has to work when we run it). FAILURE TO INCLUDE THE R CODE MEANS THAT THE COURSEWORK WILL BE MARKED INCOMPLETE (fail).
- You may assume the methods you have used (e.g. difference in means, linear regression, etc) are understood by the reader and do not need definitions, but you do need to explain the intuition of these methods in some cases.
- Round all numbers to two digits after the decimal point.
- Do not copy and paste *any* brute R output (e.g. `summary(lm(y ~x))`) into your answers. Create a minimally formatted table, e.g. with the `screenreg` command as seen in class. If that does not work, re-create by hand such a table.
- Assign every table and figure a title and a number and refer to the number in the text when discussing a specific figure or table.
- All variable names in the coursework are written in *italics*.

1) (X Points)

In the development politics sector, the effectiveness of deworming was long debated. Intestinal helminths, tapeworms, infect more than a quarter of the population worldwide. Deworming drugs are relatively inexpensive. Development practitioners long debated deworming programs but the benefits on health were not known. Therefore, a program in Kenya randomly treated students with deworming drugs and not others. The students health level was measured on a continuous scale.

1. Come up with a null and an alternative hypothesis.
2. How would you test the experiment? Describe which method you would use and how you would decide which hypothesis is more credible.

2) (X Points)

You analyse whether peacekeepers help improve political stability in the countries they were sent to. Your research associates have measured political stability in 30 post-conflict countries, 15 of which are countries were peacekeepers had been deployed. Political stability is measured on a scale from 0 to 100 where larger values correspond to more stability.

The values for stability in countries without peacekeepers are: 58, 13, 41, 6, 30, 46, 1, 10, 7, 49, 23, 38, 46, 15, and 23.

The values for stability in countries with peacekeepers are: 63, 54, 45, 60, 22, 33, 30, 61, 63, 62, 33, 20, 52, 58, 45.

1. Compute the appropriate measures of central tendency and dispersion.
2. Your theory suggests that peacekeeping facilitates state-building. Formulate a null hypothesis and an alternative hypothesis from your theory and calculate the difference in means.
3. Carry out the null hypothesis test assuming that the critical value of t to reject the null at an alpha level of 0.05 is 2.05. Interpret the result.

   Hint: The formula of the standard error for the difference in means is:

$$\sqrt{\frac{\sigma_{Xa}^2}{n_a} + \frac{\sigma_{Xb}^2}{n_b}}$$

   where a and b correspond to the two groups in the data.

3) (X Points)

A team of researchers sets out on an ambitious project. They want to test the effects of absolute and relative wealth on well-being. To do so, they randomly sample 1000 sub-national regions out of all regions in the world. For each region, they measure satisfaction with life (well-being) on a 0-100 scale where 100 is the happiest. To capture absolute wealth, they collect data on GDP/capita in each region in 1000 US dollars. Relative wealth is measured using the Gini-index. It ranges from 0-100, where 100 implies that all wealth is in the hands of one person and 0 implies perfect equality. They produced a dataset containing the three variable (uploaded on our website) and fitted a linear model to test the relationships. The results are in table 1.

- *Life Satisfaction* 0-100 (100 most happy)
- *Gini Goefficient* 0-100 (100 = 1 person owns everything, 0 - everyone owns the same)
- *GDP/captia* in 1000 US dollars

Table 1: Regression on Life Satisfaction

	Model 1
(Intercept)	66.26
	(0.60)
Gini coefficient	−0.52
	(0.01)
GDP/capita	1.74
	(0.02)
R^2	0.90
Adj. R^2	0.90
Num. obs.	1000
RMSE	5.17

Questions

1. Formulate hypotheses for both wealth variables and justify them.
2. Create scatter plots for the relation between absolute wealth and happiness and relative wealth and happiness. Discus the scatter plots.
3. Interpret the regression table.
4. Discuss the intercept in technical and substantive terms.
5. Should we predict life satisfaction given a Gini coefficient of 85 based on our model? If so, what is the prediction? If not, why?
6. Is the relationship between absolute wealth and happiness well described by a line?