**COURSE\_NUM Module 3 and 4: Regression analysis (10% + 25%)**

***Due on Blackboard: Part 1 on DATE and Part 2+3 on DATE***

**Objective:** To conduct multivariable regression analysis to answer a research question aimed at understanding a policy problem, and interpret and communicate the results of the analysis in a policy brief.

**Grading:** This assignment is broken into two parts (worth 10% and 25%), in total worth 35% of your final grade. The grading rubric is on page 4 of this document.

**Assignment:** The assignment is broken into three parts.

***For Part 1 (due xx),*** you will identify your policy problem, articulate a research question that is appropriately answerable with multivariable regression analysis and that will answer or provide context for your policy problem, and identify an appropriate data source and analytical sample.

You will conduct a linear regression to demonstrate and/or examine whether there is a statistical association between your key independent and dependent variables.

***Part 2 (due xx)*** will involve conducting multivariable analysis.

***Part 3 (due xx)*** involves presenting your multivariable regression findings in the form of a policy brief for key stakeholders using both words, tables and/or charts/figures.

**Assignment submission checklist by xx (10%):**

* Part 1: Approximately 1-2 pages answering the questions for Part 1

**Assignment submission checklist by xx (25%):**

* Policy brief
* Your data if you are using data that *was not* one of the options provided on Blackboard, if it’s appropriate to share and not sensitive.

**Useful resources:**

* Module 3 and 4 lectures
* Module 1 workshop (creating means and charts/figures for Part 3, if presenting charts)
* Module 3 workshop (estimating regression models, creating a scatterplot and linear trend line)
* Readings: Module 4 textbook chapters 8 and 9

**Part 1: Research questions, target population identification, and feasibility test (10%)**

**Due by DATE**

For Part 1, answer the six questions below. The information in Part 1 will inform the regression analyses you conduct in Part 2 and the policy brief you write for Part 3.

This part should be brief (approximately 1-2 pages) and is primarily to demonstrate that you are able to answer your research question with the data you have available and that a preliminary examination of the data indicates statistical significance, setting you up for a successful analysis and policy brief.

1. *What is the policy/social problem you are aiming to inform? (1-2 paragraphs)*
2. *What is/are the research question/s that will shed light/inform the policy/social problem and that your analysis will answer?*
   * *In short, I am looking to see that your research question actually informs the problem, and that you will be able to answer the research question using a multivariable regression approach.*
   * *You may have multiple research questions.*
3. *What dataset will you use?*
   * *If you are not using the dataset provided on Blackboard (i.e., the NZ GSS SURF), please provide a link to the data (if publicly available).*
4. *What/who will be in your analytical sample and are there any sampling issues (e.g., some key population subgroups not able to be surveyed)?*
   * *I do not necessarily need a sample size here, but I need to know that you have selected the sample with the right attributes. In many cases, this will be every unit in your dataset, but if you’re wanting to understand a phenomenon among a certain subgroup, you will want to identify who that subgroup is.*
   * *If your sample will change for different analytical steps, please describe that process.*
5. *What is/are your dependent variable/s, your key independent variable/s, and potential covariates/control variables? Describe this in words and write out your final model as an equation.*
6. *Demonstrate that a preliminary association exists using both bivariate statistics and a linear regression. In 1-2 paragraphs, describe what you’ve found and whether it provides preliminary support for further investigation of your research question (i.e., feasibility check).*

**Part 2: Regression analysis**

In order to complete Part 3, you need to conduct a multivariable regression analysis using the data you identified in Part 1. You may also present conduct a linear regression analysis if that is an important step in informing your multivariable analysis. For example, you may want to demonstrate whether adjusting for other covariates partially or fully accounts for a significant association between two factors.

You will produce a regression table that you will present in your policy brief. You may run and present just one regression, or you may run and present multiple models.

**Part 3: Policy brief (25%)**

For Part 3 you will write a policy brief, consisting of between 1,500-2,000 words (not including table text or reference page).

In this brief you will outline the policy problem and how your research question will answer or shed light on our understanding of that policy problem in ways that can inform policy solutions.

You will use APA style in-text citations and references, and findings in the briefs should be written in third person (e.g., “The analyses provide evidence of...”) not first (not “I think the analyses provide evidence of...”) when appropriate.

Please read the grading rubric for more information on specific pieces of information that need to appear in the policy brief.

**Policy brief guidance**

You will need to organise your brief as follows:

Titles and subheadings

1. A clear and descriptive, but concise, title.
2. Subheadings throughout the brief to organise the main and secondary sections.

Summary paragraph

1. Limit this to 3-5 sentences.
2. It will need these pieces of key information:
   1. Identify the policy problem;
   2. What is important to find out in order to inform that problem ;
   3. What you examined (that taps into the “what is important to find out” above);
   4. What you found; and,
   5. Why this matters for the policy problem and/or brief policy recommendation/implication.
3. Although this paragraph should be the first paragraph of the brief, write this part last once you’ve written the main text.

An introduction

1. Identify the social or policy problem and explain why it’s a problem, using cites/references.
2. What population is primarily affected (e.g., age, socioeconomic status, countries, regions)?
   1. *Provide concrete numbers using indicator data where appropriate.*
3. Explain how your research question will shed light on this policy problem.

Methods

1. *Data:* Provide information on your data, including the source, time frame, population sample, how it was collected, and anything else that would allow us to understand your dataset.
2. *Sample:* Provide information on how your *analytical* sample was selected or what it consists of, including the sample size.
3. *Variables:* What is the outcome(s) variable(s) you are examining and how is it coded, and (potentially) how is the question asked if it’s central to the research? What are the focal independent variables? What are the other covariates that you are adjusting for?
4. *Analysis:* Describe the methodological approach or steps you’re taking.

Findings

1. This section will primarily focus on the results of your regression analysis.
2. However, you may also want to present some descriptive (univariate, bivariate) statistics (in charts/graphs) if they are important for understanding your policy problem.
   1. For example, if you were looking at what sociodemographic characteristics are associated with COVID-19 vaccine hesitancy in a multivariable regression, you might want to present the univariate statistics showing the rates of vaccine hesitancy in your sample.
3. Present a regression table.
4. Translate the findings that directly answer your research question(s) and/or sheds light on the policy problem (i.e., stick to discussing the coefficients that matter for your research question).

Discussion/Policy implications

1. Summarise very briefly the overall findings.
2. What do these findings tell us about the policy problem? This is different than when you describe the findings in the prior section—in the discussion section you can connect the findings back to informing the policy problem. In short, how does what you found inform what we know about the problem?
3. What implications do these findings have for policies aimed at addressing the problem? And/or, are you able to offer any specific policy recommendations based on what you found?
4. You can also discuss the limitations of your research and what future research might address.

In-text citations and reference page

1. Use in-text citations and have a stand alone reference page on the last page of your brief.
2. Use APA style references. At the end of this document is a citation style tip sheet.
3. The reference page does not count against your word count.

**Part 1: Regression analysis assignment grading (10%)**

**Defining research questions and sample**

\_\_\_ Policy/social problem is clearly identified, and research question informs policy/social problem identified.

\_\_\_ Research question is defined clearly enough to be answerable with linear/multivariable analysis, with dependent and independent variables identified.

\_\_\_ Dataset identified is appropriate to answer the research question, and analytical sample is defined and appropriate for answering the research question, and any sampling issues noted.

**Feasibility check**

\_\_\_ Dependent and independent variables identified answer the research question. Covariates/control variables identified are appropriate.

\_\_\_ Dependent and independent variables are described appropriately in text and as an equation.

\_\_\_ Bivariate statistics are calculated correctly and presented appropriately.

\_\_\_ Linear regression is conducted correctly and output presented.

\_\_\_ Description of bivariate statistics and linear regression describe whether there is any preliminary support for further examination of the research question.

**Grading feedback**

**Total marks** \_\_\_ / 10

**Part 2 + 3: Regression analysis assignment grading (25%)**

**Part 2+3: Multivariable analysis and interpretation**

\_\_\_ Multivariable regression models are conducted correctly, including variables modelled appropriately (e.g., categorical variables included as dummy variables) and multiple covariates included.

\_\_\_ Multivariable regression models are conducted in a way that answers the research question(s).

\_\_\_ Multivariable regression results, including the interpretation of coefficients and statistical significance, is interpreted and discussed correctly.

**Part 2+3: Statistical presentation**

\_\_\_ Inclusion of tables, charts/graphs/figures informs the research question and/or context for the study.

\_\_\_ Regression table includes the appropriate information, including coefficients, statistical significance, model fit statistics, standard errors/confidence intervals, and sample size. Labelling of variables is clear, and table title descriptive.

**Part 2+3: Content**

\_\_\_ Summary paragraph contains key information (as outlined in policy brief guidance).

\_\_\_ Introduction clearly identifies policy/social problem and makes the case for how the research questions and subsequent analyses will inform the problem.

\_\_\_ Methods include appropriate information.

\_\_\_ Description of findings informs the research question.

\_\_\_ Discussion section elaborates on how the findings inform the research question and policy/social problem.

\_\_\_ Implications or recommendations for policy are discussed and are informed by the findings.

**Part 2+3: Organisation**

\_\_\_ Presentation is aesthetically well-organised (e.g., titles and subheadings, some white space, bullet-points if appropriate) and the appropriate length (1,500-2,000 words).

\_\_\_ Writing style is easy to understand, free of grammar/spelling errors.

\_\_\_ In-text citations and references are used and in APA style.

**Grading feedback**

**Total marks** \_\_\_ / 25