**Part A (Proposal)**

**Weight:** 10%

**Length: 750-1000 words (not including code samples in appendix)**

**Group Assessment**

**Submitted: On Canvas in PDF format by *one group member*. Be sure to follow the naming convention defined at the front of this brief.**

**Task**

In this Assessment item, you will work in a team to produce a project proposal for your data analysis project. You will do this by following this sequence of steps:

1. Define a broad research area of interest (e.g. public health, climate change, demographic change, finance etc.) and form groups accordingly.
2. Work to define a set of well specified research questions for your broad area of interest. (Note: at least some of these should be actionable – see Part B).
3. Look for a range of datasets that might help to answer their questions.
4. Refine at least one research questions so that it can be answered by a **Regression Model**. (Note: you can ask other questions too, but you must produce at least one regression model in Assessment 2B.)
5. Write a proposal that summarises the following:
   1. the rationale and stakeholders for the project,
   2. the research questions,
   3. the range of datasets examined as well as those chosen for the analysis (include details about how you merged the different datasets and an assessment on whether the granularity of the data sources is sufficient to answer your research questions),
   4. the regression modelling techniques to be employed and,
   5. any issues that you anticipate might arise in carrying out the project.
   6. Include an Appendix that contains code samples demonstrating the data acquisition and merger processes that you have used to date.

**Assessment Criteria: Part A**

|  |  |  |  |
| --- | --- | --- | --- |
| **SLO** | **CILO** | **Assessment Criteria** | **Weight** |
| 2 | 1.4 | Clarity in articulating the research questions along with a well defined proposal for making the invisible visible for a specified set of stakeholders. | 30% |
| 4 | 1.2 | Level of expertise using key R functionality demonstrated in the process of data acquisition, and creativity in solving the problem of finding and merging datasets that can answer the research questions. | 50% |
| 3 | 4.2 | Eloquence and robustness of the argument used to justify the proposal. | 20% |
| **Sub Total** | | | **100** |
| **Total (10%)** | | | **/10** |

# Research Questions

Selected research questions for chosen stakeholders, with blurb

One selected for regression analysis

How have weather events effected cereal crop yield production in Australia?

What factors influence the yield of cereal crops grown in a given location of Australia?

How will climate change factors affect cereal crop yield production?

How do different farming practices around the world effect cereal crop yield production?

# Background

List of factors that affect crop yield – Marco

# Data collected

Table of data that lists variables, source – Rato, All to add the sources they have collected

# Proposed methodology

Procedure/techniques to be used to perform regression modelling – Hayden, Wei Lin

# Project risks

Any potential issues or difficulties

# Appendix

R code for data acquisition and joining