

## Assessment 3: CAPSTONE ASSIGNMENT

### Overview

- This assessment will require you to perform a statistical investigation on a dataset that will be described with this assignment on LearnJCU.
- Your statistical approach is bounded by the content and programs used in this subject.
- This investigation will be summarised as a formal report that can be understood by a general audience who is fluent in statistical concepts.
- Key findings of the report however, should also be written in a way that a non-statistical expert could appreciate.

### Structuring the Report

The report should have the following sections marked clearly:

#### Title:

In today's busy world, it is very important to make the most of your title. Make the title 'eye-catching', informative and an accurate representation of the contents of the report.

#### Executive Summary:

The executive summary provides a short sharp overview of the contents in the report and will be around 300-500 words. There will be five parts:

- I. Introductory statement: background to the study, important issue(s) the report addresses. (approximately 2 to 4 sentences)
- II. Purpose of the report: state the objectives and associated hypotheses (1 sentence with multiple bullet points or ~3 separate sentences.)
- III. Methodological approach: overview the data and analysis methods (1-2 sentences with multiple bullet points or ~3 separate sentences.)
- IV. Findings: list the main overarching findings from your statistical investigation (approx. 3 sentences)
- V. Conclusions and Implications: what conclusions can be drawn from your investigation? How can the findings in your report to deliver a benefit to people, things, systems or processes? (approx 2-3 sentences)

#### Introduction:

The introduction sets the scene for the investigative efforts. It provides motivation for the work and relevant background information and references that will enable the reader to put in context the key objectives and findings in your report. Address the important issues that have motivated your investigation. At the end of the introduction clearly state the key objectives of the paper.

#### Hints:

- I. Do not put any results from your investigation in the introduction.
- II. Do not discuss the data and methods in this section.
- III. Do not discuss your conclusions or key findings in the introduction.

IV. You should have three independent objectives.

#### Data:

This section should provide clear details about the data that have formed part of your investigation and **where** the data were sourced. This section will describe:

- The sample size of each variable
- The variables used in your analysis, and the variable type.
- Relevant summary statistics should be provided in a table.
- Any data pre-processing that was done prior to analysing your data. Pre-processing may include but is not limited to ways of handling missing data; transforming data.
- Constraint: your dataset must have at least one categorical variable and at least one quantitative variable.

Hint: it should be easy for another person to read this section and independently recreate your dataset. Provide enough detail for this to be done.

#### Methods:

This section should summarise the statistical methods that were used to investigate the objectives of the report and the software used to generate the results and check assumptions. To cite R-Studio type `RStudio.Version()` from the command line.

The methods should be appropriate to ensure the objectives and hypotheses are met. It is often helpful if the author lists the key R functions and associated arguments that generated the results. E.g. “The `lm` command with default settings for the arguments was used to produce a simple linear regression model between  $y$  and  $x$  in R-Studio”. It is important to provide sufficient details so that your methodology could be repeated by an independent person.

Hint:

Your investigation will comprise of

1. Three independent objectives
2. Each objective will have its own hypothesis or set of hypotheses to be tested (i.e. you need to be able to generate a p-value(s)). Make sure that you state the hypotheses.
3. Each objective will require a **different** type of statistical significance procedure to be performed eg chi-squared test, regression analysis, t-test. Note an ANOVA with its post-hoc tests counts as one procedure; similarly a regression analysis with associated significance tests counts as one procedure).
4. Where appropriate support your hypothesis test with a confidence interval(s).
5. State any assumptions that you needed to make in this section.
6. One method, must be a regression analysis.

### Results and Discussion:

This section provides details about the findings from your statistical investigations.

Hint:

1. Relevant graphical outputs should go in this section, especially graphs that (i) visualise the adequacy of assumptions and (ii) reinforce findings associated with the objectives/hypotheses. Each figure must be numbered with a caption and the figure must be referred to in the text. E.g. "Figure 1 shows ..."
2. Write your findings in sentence form and in brackets provide the test statistic, degrees of freedom and p-value in brackets.
3. You should not cut and paste statistical output from hypothesis tests here (statistical code and output should go in the appendix).

### Concluding Remarks:

Final remarks about the key findings of the investigations and the relevance of these findings. Make sure that you mention any limitations of your work here.

### References:

List the sources your investigation has drawn from. Note that all references should be referred to in the text.

### Format:

The assignment must be presented in 12 pt Time New Roman style font on A4 pages using single line spacing, double column format.

The assignment should not exceed six (6) A4 pages.

References can be listed in an appendix and do not form part of the page limit.

WARNING: only the first SIX pages will be assessed!!!

Appendices containing R code are accepted, but make sure that you put any important outputs in the body of the report. Markers will not sift through pages of appendices to search for a piece of output you refer to.

Appendices will not count toward the page limit.

### File Name Convention

Does your name appear in the file name? 1 Mark will be deducted from the overall assignment score, if your name is omitted from the file name. For example, if your name is Mary Smith, and you submitted this file as pdf, then you must call your file  
MarySmithMA5820Report.pdf

If Mary uses word, then the pdf extension will be replaced with the docx extension. The most important thing is that your name appears in the final name.

### Marking Scheme

Note: Assignments that are not formatted according to the directions will receive a mark of 0%

Criteria (each worth 20%)	0 Mark	1 Mark	2 Marks
1. Communication	Multiple improvements are required	Sectioning and content was generally acceptable  AND/OR  Content is written to an acceptable level  AND/OR  Some details are vague and prevent reproducible analyses.	Section and content across the report is appropriate  AND  Content is written to a high standard  AND  Analyses could be easily reproduced  AND  Figure numbers referred to in text
2. Objectives and associated hypotheses	Difficult to comprehend the meaning of stated objective(s)/hypotheses(s).	Are generally articulated well enough to be recognizable but could be improved.	Are precise, concise and where appropriate are supported by mathematical symbols.
3. Experimental design	Elements of experimental design are flawed  And/or  Constraint(s) have <u>not</u> been obeyed	The linkage between dataset, methods and objectives/hypotheses is recognisable but could better aligned.  AND  Constraints have been obeyed (i.e. regression analysis performed and dataset contains a categorical and	The dataset, methods and objectives/hypotheses are well aligned  AND  Constraints have been obeyed (eg regression analysis performed; dataset contains a categorical and quantitative variable)

		quantitative variable)	
4. Displaying and Summarising Data to support objectives/hypotheses	Inappropriate choice(s).	Some improvements would enhance the report.	Appropriate choice of graphs AND Appropriate summary statistics AND Well presented visuals (e.g. axes and labels clearly visible, figure numbers and captions)
5. Integration, evaluation, synthesis of statistical principles, methods and techniques.	Major errors in execution AND/OR comprehension AND/OR interpretation.	Minor errors in execution AND/OR comprehension AND/OR interpretation.	Exemplary execution of methods AND demonstrates full comprehension of associated statistical theory, assumptions and limiting factors. AND Student has clearly interpreted solutions in a highly articulate statistical and English language.