**PSYC372 APA Statistical Write-Up Module 2: Multiple Regression**

# Introduction to the Task

In this assignment, you will conduct write-ups of the data cleaning / assumption testing and results of two multiple regressions. You will write up the data cleaning / assumption testing only for one of the studies and the results only for the second study.

**Resources for assignment**:

* **Section A data file.sav** is to be used for **Section A (Data Cleaning and Assumption Testing) only.**
* **Section B data file.sav** is to be used for the **Hierarchical Regression analysis only.** The data files for the assignment are directly below the assignment in Moodle.
* Lecture 2 (Effect Sizes – background information)
* Lecture 3 (Data Cleaning and Assumption Testing – background information)
* Lecture 4, slides 18-99.
* Practice quiz (regression). The tutorial quiz (Moderation and Mediation) will also build relevant skills.
* Zoom session (live) on multiple regression procedure and writing up a report – see announcements for details.

**Word length, submission and due date.**

The required word count is in the 700-900 word range. Ideally, you should write as succinctly as possible but ensure that you address all the requirements in sufficient detail. The title page, page numbers, and required table **are** counted for word count purposes. Penalties will apply if assignments exceed the word count limit by more than 10%. Assignments that fall below 700 words will not be penalised for word count but marks may be lost for missing information or a lack of detail.

**Due date: Sunday, 8th August, 2021 at 23h59.**

Please submit your assignment through the Assignment Submission Portal, under the heading “Statistical write-up PSYC472” in the **Assessments** block on Moodle. This assignment contributes 20% to your final mark in the unit.

# General requirements

Write up only the sections of the report specified below.

The report must be in APA 7 format, including a title page and page numbering. Much of the unit materials remains in APA 6 but we will provide information about key differences between APA 6 and APA 7 in the announcements and Zoom session 2. The changes have very limited impact on the write-up for this assignment.

The report must have a title page, with a title, student name and student number. At bottom left of the title page, you should indicate the word count of the assignment.

**You should use two-tailed tests, a critical alpha level of .05 for null hypothesis significance testing and an alpha of .001 for the assumption testing.**

**SECTION A (Data Cleaning and Assumption Testing) (20%)**

In this section of the assignment, you will conduct data cleaning / assumption testing for a multiple regression study. Read the information about the study below and then follow the instructions for writing up the section.

**Information about the study**

Variables used in the study:

Child age (continuous independent variable).

Anger and hostility (independent variable): continuous scale measuring perception of anger / hostility. A high score indicates a high level of anger / hostility.

Chaos (independent variable): continuous scale of household chaos. A high score indicates a high level of household chaos.

Child behaviour (continuous dependent variable): High scores on the scale indicate a high level of maladjusted behaviour.

**INSTRUCTIONS FOR SECTION A:**

**Write a section of a report with a level-1 heading ‘Section A’ and a level-2 heading ‘Data Cleaning and Assumption Testing’, following the instructions below.**

**You must use the following data file: Section A data file.sav**

**Follow the sequence below and write up the analysis. Do not change the data file.**

1. Examine all variables except age for univariate outliers. Use histograms to detect univariate outliers and an appropriate z-score cut-off for confirmation (as per lecture 3). Report how many outliers there are per variable on the basis of your analysis of the z-scores. If a variable has no outliers, report zero. Note whether there was any potential discrepancy between likely outliers in the histograms and confirmed outliers as per the z-scores.
2. Examine all the independent variables for multivariate outliers. Report formally on the Mahalanobis Distance test (α = .001) mentioning the family id, critical value, observed value and alpha for each case you identify as a multivariate outlier. Determine the critical Mahalanobis Distance value using the table of values in the Chi-square distribution provided at the end of this document.
3. Using the appropriate Tabachnick and Fidell (2007) rule/s, determine the minimum sample size for this study and as a whole and for the individual predictors. Briefly comment on whether the minimum sample size has been met.
4. Report briefly on multicollinearity and independence of residuals. Provide statistical evidence to support your arguments.

**SECTION B (Hierarchical Regression – Results section) (80%)**

In this section of the assignment, you will run a hierarchical regression. Read the information about the study below and then follow the instructions for writing up the section.

**Information about the study (Research Questions)**

We are interested in the performance of salespeople in high pressure environments. Previous research in the USA suggests locus of control, confidence in the job and extraversion predict performance. However, recent research focusing on individual predictors has also identified motivation and manipulative behaviour as predictors of performance. There has been no research on an overall model using all these predictors. In this three-step hierarchical regression, we are interested in:

1. Testing the original model that locus of control, confidence in the job and extraversion predict performance (model 1).
2. Adding motivation to model 1 and testing the model (model 2).
3. Adding manipulative behaviour to model 2 and testing the model (model 3).
4. Evaluating the impact of all the individual predictors.

Variables used in the study (all continuous):

Locus of Control: High scores indicate an internal locus of control.

Extraversion: High scores indicate a high level of extraversion.

Confidence: High scores indicate high levels of confidence in the job.

Motivation: High scores indicate a high level of motivation.

Manipulative Behaviour: High scores indicate that the person is skilled at manipulation of others.

Performance: High scores indicate high levels of performance.

**INSTRUCTIONS FOR SECTION B**

**Write a section of a report with a level-1 heading ‘Section B’ and another level-1 heading ‘Results’, following the instructions below.**

**You must use the following data file: Section B data file.sav.**

**Follow the sequence below and write up the analysis. Do not change the data file. Assume that data has been cleaned and all assumptions have been met – do not conduct data cleaning or assumption testing. Cross-check that you are not using the Section A file.**

**Note about Effect Sizes:** In your assessment of effect sizes, use the information provided on slide 33 of lecture 2 (also included at the end of this document). You may use this slide for *R, r, and sr*

(*sr* = semi-partial correlation and *sr*2 = semi partial correlation squared or percentage of variance explained by a predictor).

**Note about rounding:** Round information about significance (*p*) to three decimal places. Otherwise, round to two decimal places. Express percentage of variance explained as a whole number rounded to two decimal places e.g. 39.25%.

1. **HYPOTHESES** (15%)

Provide hypotheses appropriate to a hierarchical multiple regression using a level-2 heading ‘Aims and Hypotheses’. The hypotheses must be based upon the research questions (above). You should provide hypotheses that cover each model you will test. Also, provide hypotheses for the individual predictors **in model 3 only (i.e. the model that includes all the IV’s).**

1. **RESULTS** (45%)

The **Results** section (under a Level 1 heading) should include (a) a *clear* description of the statistical results in relation to each finding and (b) references to **one** tables that complements the written description.

1. **Run a hierarchical multiple regression analysis in SPSS**. There should be three models as explained in the information about the study above, with one variable added in each of models 2 and 3.
2. Create a table in APA 7 format that provides information about all three models but not the individual predictors. The table should show *R*, *R*-squared, adjusted *R*-squared, *R*-square change and the *p*-value for each model (significance). This should be the **only** table (or figure) in the assignment.
3. Write a description of what is presented in the table, including an introduction to the table. Generally avoid repeating data values in your text that have already been shown in the table unless you are emphasizing key findings. For example, you might report adjusted *R2* and effect sizes in-text. If you refer to a test statistic in your discussion, it is essential to use the appropriate APA format (for example, italics). Aim to write a succinct but clear description of the results based on the assumption that the reader does not fully understand the table. You should explain your findings as regards each model as a whole, effect size for the model, interpretation of effect size (e.g. small), and percentage of variance explained. Do not compare the models at this stage.
4. Assess the independent variables (predictors) using the information in SPSS for **model 3 (the model that contains all the variables).** Discuss the significance (and direction) of the relationship between each individual predictor and the DV (with the predictive effects of all other variables removed). Discuss the percentage of variance explained and the effect size (e.g. small) of each predictor and also compare the relative impact of the predictors. Provide detailed statistical information to support your analysis as there is no table describing the findings for the predictors. Supporting statistical information is usually placed in brackets and must use statistical symbols / appropriate formatting.
5. Briefly compare the three models and indicate whether models 2 and 3 are better fits than model 1.
6. **DISCUSSION** (20%)

Under a Level 1 heading, titled **Discussion,** briefly discuss your study findings.

1. State the outcome of the hypothesis testing using both the overall MR models and the individual predictors. Ensure that you cover each hypothesis explicitly.
2. With due regard to discussions concerning NHSTand effect sizes in the lectures, suggest a model (based on only the variables used in this study) that should be used in future research and comment briefly on the extent to which the model would be a high quality predictor of performance.



