

# PSY 202 DATA ANALYSIS PROJECT

## ASSIGNMENT DETAILS

20% OF FINAL GRADE

### ***What and Why?***

The goal of this assignment is to provide you the opportunity to apply your statistics toolbox to a real dataset. This assignment also will give you experience with using statistical software to prepare and analyze data and interpret output.

For this project, you will complete a class survey; prepare and analyze your data; formally report the results in APA style; and respond to a few brief essay questions regarding the project. Please note, you will need to make changes to the original data set in order to complete this assignment.

To complete the assignment, you will need the following files, in addition to the questions below (all of which can be found on Quercus):

- Data Set (in csv format)
- Data Guide
- Questionnaire Info
- Survey Export (so you can see what it looked like online)

### ***Notes on Analyses***

Although in the ‘real world’ you might make choices about removing outliers or not doing an analysis because the assumption of homogeneity or normality is violated, please do not remove anyone from the data file and report all analyses even if assumptions are violated.

You may need/want to do a test that was not explicitly part of the tutorial agendas. You have lots of resources at your disposal to figure out the best ways to do this test – all of the YouTube channels and the open textbook, as well as Google! Part of the goal of the tutorials was not to just walk you through the only tests you could possible use, but also to gain comfort and familiarity with solving a problem you may not know the answer to. You can do it!

### ***Basic Submission Details***

All answers should be typed into a Word document and uploaded to Quercus by the due date (specified on the syllabus) in one of the following formats: .doc, .docx, .pdf. No other file formats will be accepted. It is your responsibility to make sure that your assignment has been uploaded and submitted correctly.

Please take care to format your assignment (including typed responses, pasted figures and images, and pasted output) in an organized, easy-to-read manner.

Please do NOT submit analysis output unless specifically requested.

For explanations and discussion questions, complete sentences are expected unless otherwise specified.

***Honour Statement:*** Academic integrity is fundamental to learning and scholarship at the University of Toronto and beyond. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the U of T degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves.

This is an individual project. The work you submit is expected to be your own, and the submission will be run through TurnItIn. You may not work with any classmates to complete this assignment. Discussing or completing this project with the help of a classmate, peer, other professor, or paid service is academic dishonesty and will be treated according to UofT policy.

The University of Toronto treats cases of academic misconduct very seriously. All suspected cases of academic dishonesty will be investigated following the procedures outlined in the Code. The consequences for academic misconduct can be severe, including a failure in the course and a notation on your transcript.

# PSY 202 DATA ANALYSIS PROJECT

## **Part 0: Preparing Your Data**

1. Before you begin working with your data, you should set up your file with descriptions and value labels to make interpreting output and figures more straightforward.
2. Before you begin your data analysis, you need to examine your data file. Are there any features of the variables or data that you need to adjust in order to be able to analyze your data? If so, briefly state what changes were necessary and how you made them. (4 pts)
3. In your dataset, there are items from 4 unique questionnaires. Using the information from the *Questionnaire Info* document, compute all relevant scales/subscales for these measures. In addition, do a reliability analysis for each scale/subscale and indicate the quality of internal consistency for each; explain what this means for this measure, in words. Finally, compute and record the mean and standard deviation for each scale/subscale using the Descriptives function. (8 pts each)
  - a. Subjective Well-Being
  - b. Big 5 Personality Inventory
  - c. Need for Cognition
  - d. PANAS (Positive and Negative Affect) – Before Mindfulness Exercise
  - e. PANAS – After Mindfulness Exercise

## **Part 1: Getting to Know Your Data**

4. For each of the following variables, compute and record (when appropriate): minimum, maximum, mean, median, mode, and standard deviation. Which measure of central tendency is most appropriate for each variable? Why? (4 pts each)
  - a. Age
  - b. SocMedia
  - c. Caffeine
  - d. Exercise
  - e. Faction
  - f. SurveyLength
5. Which, if any, of the variables in number 3 have outliers? How do you know? Describe in words what this means for the variable. (4 pts)

## **Part 2: Planning for Analysis (3 pts each)**

6. For each of the following tests, (a) choose the specific variables that can be analyzed using this test, (b) identify which is the predictor and which is the outcome (if relevant), (c) and describe in words the research question you propose to test. There are many correct answers!

**Example** (you may not use this in your assignment):

*Independent samples t-test*

- (a) Animalpref and Attract
- (b) Animalpref is predictor, Attract is outcome.
- (c) Do cat people or dog people consider themselves more attractive?

- a. *Independent sample t-test*
- b. *Dependent sample t-test*
- c. *One-way ANOVA*
- d. *Two-way between-subjects ANOVA*
- e. *Correlation*
- f. *Simple Linear Regression*
- g. *Chi-Square Test of Goodness of Fit*
- h. *Chi-Square Test of Independence*

**Part 3: Analyzing and Reporting I (15 pts each)**

7. Looking back at part 2, choose 4 of your proposed analyses to do: (1) one type of t-test; (2) one type of ANOVA; (3) either correlation or simple linear regression; and (4) one type of chi-square analysis (4 total analyses).
  - a. For each analysis, paste the output into the document, and report your results in APA format. Your report should include both the APA-style statistical tag as well as a description of the findings in plain language. Include measures of effect size for all analyses.
  - b. In addition, for each analysis, create one data visualization that clearly displays your results. Paste the figure into the document, and explain in words what exactly this figure communicates.
  - c. Finally, using <http://powerandsamplesize.com/Calculators/>, G\*Power, or another resource of your choice, calculate the post hoc power for each analysis. Take a screenshot of your power calculation (so we can see the values you used) and paste it into the doc. Is this sufficient power? If not, how many participants would you need for a fully-powered analysis?
  - d. Note: the test does not have to be significant; you will get credit for reporting the result accurately, significant or not.

**Part 4: Analyzing and Reporting II (8 pts each)**

8. For each of the following research questions, (a) identify the variables and specific hypothesis test necessary to address it. (b) Run this test, and for each one, paste the output into the document and report your results in APA format (including measures of effect size).
  - a. Is the effect of the level of perceived responsibility on the willingness to help a classmate moderated by whether someone is a cat person or dog person?
  - b. Does the effect of a mindfulness exercise on negative emotions depend on whether you are born in Winter (Dec-Feb), Spring (Mar-May), Summer (Jun-Aug), or Fall (Sep-Nov)?
  - c. Is the number of Facebook friends someone has more strongly related to subjective social status or how attractive they think they are?

**Part 5: Reflection (20 pts)**

9. Looking back over the semester (lecture, tutorials, assignments), reflect briefly on your experience completing this project. Importantly, please focus on **your** personal experiences, behaviours, goals, practices, strategies, etc. – this is not an opportunity to review the instructor or the course (those come soon, I promise!). This reflection should be around 400-600 words.

In addition to a selection of the questions listed below, **everyone** should directly address the following: what did you learn that will help you in the future? This may be a specific skill, a way of thinking, a strategy for managing work, etc., and it may be relevant to future university courses, graduate school, work, or your personal life.

**Some other questions to consider might include (but are not limited to):** Now that it's over, what are your thoughts about this overall project? What are some interesting discoveries you made while working on this project, about yourself/the problem/the field? What were some of the most challenging moments, and how did you deal with them? What were some of the most powerful learning moments, and what made them so? What got in the way of your progress, if anything? What would you do differently, if you were to approach the same problem again? What were your greatest strengths? Your biggest areas for improvement? How will you use what you've learned in the future? If you plan on continuing to work with data, what in particular do you need to work on or continue to do in order to succeed?