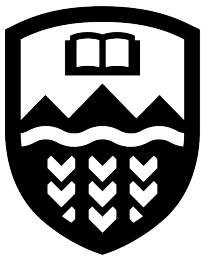

UNIVERSITY OF ALBERTA

PSYCH 213 - Winter 2025

Course Project I:

Testing Your Own Hypothesis

Instructor: **Dr. Jeffrey M. Pisklak**

Objective

Go into the real world and collect a sample of at least 50 measurements of a single variable. Then analyze and test the data you collect using R.

What to Measure

You are free to measure anything you like, provided that your data meets the following criteria:

- It must be quantitative in nature.
- It must exhibit some natural variation.
- It must not involve personal information about other individuals.
- It must not violate the student code of behaviour.
- It must not be a dataset you found online.
- It must not involve dice rolling (let's aim for something more engaging).

Measurement Ideas

Here are a few examples of things you could potentially measure. Feel free to pick one of these or choose your own.

- Number of red M&M's in 50+ bags of M&M's.
- Temperature readings of a cup of coffee/tea cooling down (in degrees Celsius).
- Time spent on different apps or websites.
- Your score on games of Tetris.
- Time between sips of a drink.
- The number of words inside pages of a textbook.
- Time spent waiting (e.g., in line, at traffic lights) per instance.
- Typing speed at repeated attempts of a typing test.
- Reaction time to a stimulus (in seconds) – Use a simple online tool to test your reaction time repeatedly (e.g., pressing a key when an image appears).
- The calories per serving size for items in your fridge.
- Time between vehicles passing.
- The amount of time you spend looking at your phone (there are apps that track this).
- Song duration in a playlist (in seconds or minutes).

- The length of your writing utensils (pens and pencils) in mm or cm.
- Your heart rate at random points throughout the day.
- The number of characters in 100 different text messages.
- Time between phone checks (in minutes).

Originality is encouraged but not necessary!

Data File

Save your data as a CSV file. **The file should contain only a single column of measurements with a clear and concise descriptive header in the first row.** There should be no other columns in the data file. You will lose marks if you ignore this basic instruction.

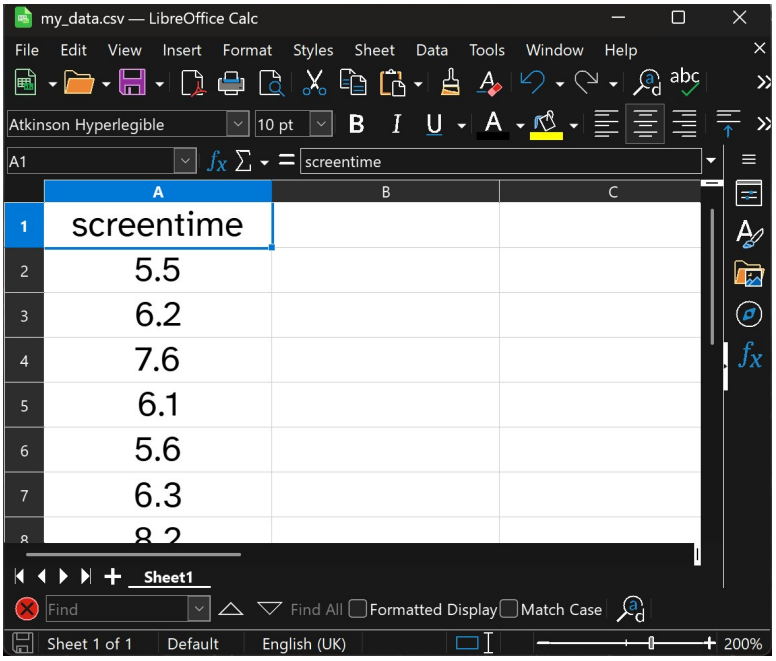

Figure 1: Example CSV file data.

How to Make a CSV File

All spreadsheet programs (e.g., Libre Office Calc, Microsoft Excel, Google Sheets) allow you to *save* or *export* data as a CSV file with just a few clicks. In fact, you don't even need spreadsheet software to create a CSV file. If you're unsure how to create a CSV file, here are some basic instructions: Click Here. Alternatively, you can try searching online for a quick tutorial specific to your software—there are plenty of step-by-step guides available!

- Tip: Avoid long or overly descriptive names in your column header. For example, if you're tracking Tetris scores, name the column `"score"` or `"tetris_score"` rather than some-

thing like `"Score Playing Tetris Over 50 Attempts"`. Short, clear names enhance readability, boost efficiency, reduce the chance of errors, and make it easier to identify mistakes.

Data Analysis using R

Using the data you collected, write R code in a Colab Notebook that accomplishes the following tasks:

1. Create a Histogram

- Use `ggplot` to visualize the distribution of your data as a histogram.
- Choose appropriate binning to represent the data effectively.
- Add clear labels, ensuring the x-axis label accurately describes what was measured.
- The plot should look professional—avoid unnecessary, strange, or distracting elements, as they may result in a penalty.

2. Save the Plot

- Include code to save your histogram as either:
  - A high-resolution image (at least 300 dpi), or
  - A vector-based image.
- For details on how to do this, refer to our coverage of this in the lectures or look up information about the function `ggsave()`.

3. Summary Statistics

- Ensure that when you run the code in your notebook, it displays a dataframe that includes the following summary statistics:
  - `N` = Sample size
  - `med` = Median
  - `iqr` = Interquartile range
  - `m` = Mean
  - `ss` = Sum of squared deviations
  - `var` = Sample variance
  - `s` = Sample standard deviation

  - `se` = Standard error
- Note: There are many ways to achieve this, but the simplest and most elegant is likely using the `|>` operator with the `summarise()` function.

4. Q-Q Plot

- Generate a Q-Q plot to assess the whether your data is normally distributed.
- As with the histogram, save this plot as a high-res image or vector graphic.

5. One-Sample t-Test

- Conduct a one-sample t-test on your data *without* using the `t.test()` function or equivalent functions from other R packages.
- Test whether your sample mean differs from a reasonable reference value (x). This could be:
  - A commonly accepted average (e.g., 6 hours per day for phone screentime, 3 cups per day for coffee consumption).
  - A theoretically meaningful value (e.g., 0 if testing whether an event happens at all).
  - A value supported by research or logical reasoning.
- Clearly state your hypotheses in a markdown/text cell:
  - **Null Hypothesis** (H_0): The true mean of the variable equals the reference value.
    $H_0 : \mu = x$
  - **Alternative Hypothesis** (H_1): The true mean of the variable is different from the reference value.
    $H_1 : \mu \neq x$
  - If you have a specific directional expectation, you may use a one-tailed test instead.
- The following results should be **CLEARLY** displayed when your notebook is run:
  - Test statistic (`t`)
  - Degrees of freedom (`df`)
  - P-value (`p`)
  - 95% confidence interval (`95_CI`)

Code

- Please ensure that your Colab notebook and R code are well-organized and run without issue when your supplied data file is used.
- Click here for guidelines on formatting code.

Project Report

Fill out the questions in `project_report.docx` with the results from your analysis. It is a standard Microsoft Word file: Click here to download it.

If you do not have Microsoft Word on your computer, you can use Google Docs to edit it. All written responses can be answered in just one or two sentences.

Submission Instructions

Submit the following files:

1. The CSV file containing the data you collected.
2. A copy of your notebook (`.IPYNB`) file.
3. A copy of your histogram as a high-res or vector-based graphic.
4. A copy of your Q-Q plot as high-res or vector-based graphic.
5. A completed version of `project_report.docx`.

Due Date

This is due Friday March 7th by midnight.