MIE237: Project Assignment (Winter 2026)

Proposal Due by February 27th

1 Introduction

In this group project, you will design and conduct an experiment and analyze and interpret the results using the concepts learned in the course. By doing so, you will gain hands-on experience in reviewing literature, experiment design, collection and analysis of real-world data in R; and scientific report writing.

Experiments are a key part of scientific and engineering inquiry because they allow hypotheses testing and the establishment of causal relationships. In planned experiments, we want to test the relationships between different variables of interest by manipulating one or more of the variables (the **independent variable(s)** or **factor(s)**) while keeping all other factors the same (as much as possible) and observing the changes in one or more outcome variables of interest (the **dependent variable(s)**). If all other factors are well controlled for, we can then attribute the changes in the dependent variable to the manipulation of the independent variable. For example, if you wanted to test the effect of studying alone or in groups (a categorical independent variable with two levels) on final grades in a probability course (a continuous dependent variable), you could design an experiment that randomly assigns group-studying to half a class, and studying-alone to the other half, and examine if there are any differences in the final grades for the two groups.

2 Project Description

Topic. Identify a topic that uses human participants and measures performance on a task. You need a task that is:

- Appropriate in difficulty (i.e., not trivially easy, not overly difficult).
- Something that is done or can be done by the general population or your classmates.
- Not risky, dangerous, illegal, sexually explicit, or costly to do.
- Straightforward for you to measure performance (e.g. number of correct answers, time to complete the task, etc.)

Make use of the literature to select your task and develop your hypothesis as to what you would expect. Your task does not need to be an original task and can, in fact, be a replication study. Your study must include the following:

- Two dependent variables of performance such as 1) number of correct answers and 2) time to complete the task.
- Two independent variables each with at least two different levels.

A **literature review** is a summary of published peer-reviewed research related to your topic. For example, one may decide to conduct an experiment that uses audio as a distraction. How do we know that sounds can distract from a task? For a literature review, you would find published research paper that provide evidence for that, and cite them. We can start by searching within Google Scholar or the UofT Library system for "audio distraction performance" and find a paper "Jahncke, H., Björkeholm, P., Marsh, J. E., Odelius, J., & Sörqvist, P. (2016). Office noise: Can headphones and masking sound attenuate distraction by background speech? *Work (Reading, Mass.), 55(3)*, 505–513. https://doi.org/10.3233/WOR-162421" in which the authors discuss how background speech can be one of the most disturbing sounds that can reduce performance. This can then be the basis for the design of an experiment of what type of sound to play to interfere with a person's performance. You should cite the references you use in this section. To avoid plagiarism follow the UofT guidance: How Not to Plagiarize by Margaret Procter, Using Quotations by Jerry Plotnick, Paraphrase and Summary by Jerry Plotnick. All references used to write the literature should be provided in the same format (e.g. APA, Chicago).

What is already known and has been published related to your selected topic? How have other researchers studied the topic? What type of experiments did they run? Identify and review relevant literature that should **form the basis for your hypothesis**. The literature review section should be about 2 pages in length.

Some example topics that could inspire experiments include:

- **Fitts' Law (Speed–Accuracy Tradeoff in Human Performance)**. Fitts' Law describes how the time required to point to or select a target depends on the target's size and its distance from the starting position, with smaller and more distant targets requiring more time and leading to more errors. In this experiment, participants complete a simple pointing task in which they are asked to use a mouse, stylus, or pen to mark a sequence of targets on a computer screen or on paper. Each participant is presented with multiple targets that vary systematically in size (e.g., large vs. small) and distance from the starting point (e.g., near vs. far), and is instructed to complete the task as quickly and accurately as possible. The dependent variables are (i) movement time, measured as the total time required to select all targets, and (ii) accuracy, measured as the number or proportion of targets successfully marked without error. The independent variables are target size and target distance, each with at least two levels. The hypothesis is that movement time will increase and accuracy will decrease as target size decreases and target distance increases, consistent with the speed–accuracy tradeoff predicted by Fitts' Law.
- **Detecting AI-Generated Media**. Show participants a sequence of images or short video clips and ask them to decide whether each item is real or AI-generated. The dependent variables are

(i) detection accuracy, measured by the number or proportion of correctly classified items, and (ii) decision time, measured as the time taken to make each judgment or to complete the task. One independent variable is the type of media shown, with two levels: still images and short videos. A second independent variable manipulates decision support: in the control condition, participants receive no additional guidance, while in the experimental condition, participants are provided with a short checklist highlighting common artifacts of AI-generated media (e.g., unnatural lighting, inconsistent shadows, distorted hands or facial features). The hypothesis could be that participants will have lower accuracy and longer decision times for videos than for images, and that providing a checklist will improve detection accuracy, potentially at the cost of increased decision time.

- **Ordering Decisions in a Supply Chain**. This experiment examines how individuals make ordering decisions under demand uncertainty in a simple supply chain setting. Participants play an online single-period inventory (newsvendor) game, where they must choose an order quantity for a product before demand is realized, incurring holding costs for excess inventory and penalty costs for unmet demand. The independent variables include decision support, with levels such as no support versus access to a brief rule-of-thumb or cost summary (e.g., how holding and shortage costs trade off), and quantitative background, measured by whether participants have prior coursework in operations research or supply chain management. The dependent variables are (i) decision accuracy, measured by the regret or absolute deviation of the chosen order quantity from a benchmark order quantity computed by the experimenters (e.g., the cost-minimizing order quantity based on the known demand distribution), and (ii) decision time, measured as the time taken to select an order quantity. The hypothesis is that decision support and prior OR background will lead to lower regret and faster decisions, illustrating how information and training influence inventory decisions under uncertainty.

Dependent Variables (the outcome variables of interest). You should have at least two dependent variables of interest: such as response time and accuracy. These should be continuous variables.

Independent Variable (the variable that you are manipulating). For your study, there should be at most two independent variables. (More than two independent variables could be too complex of an experiment and require too much data collection). It is suggested that one of the independent variables be a categorical factor (e.g., 'Distraction Type') with two to three levels (e.g., No distraction vs. Distraction). If your experiment only involves a single independent variable, it should have at least three levels.

Covariates. You may also need to select covariates for your experiment depending on your hypothesis. A covariate is a variable that may impact your dependent variable, but you are not directly manipulating in the experiment. By measuring and accounting for covariates in your analysis you can attempt to separate out their effects from those of the independent variable which you are

interested in. For example, math aptitude or level of motivation may be covariates in a study about final grades in a probability course where your independent variable of interest is studying alone or in groups. By comparing grades of different study conditions across those with high aptitude separately from those who have low aptitude (or high motivation vs. low motivation), you can assess the effects of your covariate.

Experimental Method. You must also design a method for measuring the outcomes for your experimental conditions. This includes how you are manipulating your independent variable and how you are measuring your dependent variables (e.g., having the experimenter time a participant using a stopwatch). Your goal should be to control for factors that you are not particularly interested in testing by making your experimental conditions as similar as possible (other than the change in your independent variable), and by having clear definitions for what your dependent measures represent. In addition, you will need to decide whether you want to test the same participants on both levels of your independent variable or test a different group of participants for each level of your independent variable.

Participants (your experimental units). You will also need to decide on your participants, who should be representative of the study population you wish to examine. For between-subject experiments, **you should collect data from at least 8 participants per experimental condition** (e.g., for a 2×2 design requires at least 32 participants total). For within-subject experiments, you should collect data from at least 12 participants.

Methods and Analysis. For your analysis, you are asked to describe the data you have gathered using descriptive statistics for each of your experimental conditions. You will then need to compare the conditions using appropriate inferential statistics and to draw conclusions based on your analysis.

Results. Reporting of numbers, graphs, and statistical conclusions should be done first in a 'Results' section and then interpreted in a separate 'Discussion and Conclusions' section as follows.

Discussion and Conclusions. You will need to interpret what your results mean in terms of your overall research context. Are there any conclusions that can be drawn based on your experiment and what does your study contribute to the understanding of the problem? You should also discuss how your two dependent variables (e.g. speed and accuracy) might be related. You should also describe any limitations to your study that come about because of the way your experiment was designed or how your data was collected (e.g., your results might have varied due to having different experimenters operating the stopwatch). Finally, was what you actually did different than what you proposed?

3 Deliverables

There are two deliverables: a project proposal (20% of the project grade) due by **February 27th**, and a final report (80% of the project grade) due by **April 7th**. No late submissions are accepted.

3.1 Project Proposal

A good proposal is critical to ending up with a good final report. The project proposal should be submitted in PDF format via Quercus (one submission per group) and include the following sections:

Cover Page. Give a title for your project. Provide your names and student numbers.

Introduction. Provide background and clearly specify the research questions. Clearly specify the experiment goal.

Literature Review. Provide the literature review information that may be necessary for the reader to understand the research question, with proper external sources. This section should be about 2 pages in length and form the basis for your hypothesis.

Proposed Methods. Provides answers to the following questions:

- What data will be collected?
- What are the independent and dependent variables?
- Who are the study participants?
- What are the null and alternative hypotheses?
- How will the experiment be carried out? Provide a brief description of the experiment procedures, survey questions, etc.

References. Citations can be in any standardized format (e.g., APA, IEEE) as long as they are consistent throughout the document.
Recommendation on Tools: To manage your references efficiently and ensure consistent formatting, we recommend using LaTeX (e.g., via the online editor Overleaf). By using **BibTeX**, you can automatically generate citations and bibliographies in any standard style (APA, Chicago, IEEE) without manual formatting errors. You can find a quick tutorial on managing bibliographies in Overleaf here: `https://www.overleaf.com/learn/latex/Bibliography_management_with_bibtex`.

3.2 Final Project

Cover Page. Provide name of the project, team member names and student numbers.

Introduction. Describe the overall goal of your study and question(s) you are addressing. Provide some background information regarding the topic that you have selected.

Literature Review. This section includes your literature review. You can revise and update your literature review and incorporate feedback received from the project proposal.

Methods. Describe the experimental procedure and how you collected data. Provide a precise description of the participants in your study and apparatus you have used. Provide a description for independent and dependent variables of interest in your study.

Analyses. Describe the analyses you conducted (using the methods covered in the lectures) on the data.

Results. The result section should include both descriptive and inferential statistics based on your data including analysis of the degree to which your data satisfies the assumptions of the procedures you use. Explain what the results mean in terms of the overall research question(s). Use tables and figures to effectively communicate your findings to your reader. Pay particular attention to making figures that clearly show the findings from your study.

Discussion. State the conclusions based on your statistical analysis and judgment. This might include ideas or recommendations on further study of the question(s) your study addressed. As part of your discussion, you should describe possible reasons for why you observe any differences (or lack of differences) between the different conditions, and discuss limitations of your study including biases associated with your data collection.

References. When you include references, you should cite them properly. You can adopt any citation style. But be consistent, and only use one style throughout your report.

4 Submission Details and Important Notes

Proposal. The project proposal (in PDF format) must be no longer than 4 pages, double spaced, with at least 1.5 line spacing, 1-inch margin, and 12-point font size (cover page and reference list excluded).

Final Report. Your final submission on Quercus should include the following files:

- A final report in PDF format with a maximum of 12 pages excluding the title page, and References
- A zip file that consists of all collected data and R code used for the project to ensure reproducibility.

- An attribution table along with an original authorship form must be handed together with the final report. We expect that all team members contribute at comparable levels.

Important Remarks:

- ***Each group must independently design and conduct its own experiment and collect its own data*** (members of other groups can be your participants). Using data from other groups, reusing data from past projects, fabricating or manipulating data, or presenting simulated or borrowed data as real observations is not permitted and constitutes academic misconduct. While discussion of general ideas is allowed, all data and analyses must be original to your group.
- You must use R for all analyses conducted in the project. This includes but is not limited to data manipulation, exploratory data analysis, and statistical inference.
- The results reported in the report should be fully reproducible on our end given the provided data and code. Submissions that report results inconsistent with the submitted code and data will receive zero marks.
- You are welcome and encouraged to ask questions regrading the choice of topic and conducting experiments and analysis during office hours or on Piazza.

5 Grading Rubrics

Section	**Criteria**	**Marks**
Introduction	• Research question clearly stated • Goal of the experiment is clear and reasonable	/ 5
Background	• Literature review provides a concise summary of what is known about the selected topic • Background information motivates the research question • Background information contains references to external sources to back up claims	/ 5
Methods	• Description of the data that you plan on collecting • Description and justification for independent variables and dependent variables chosen • Description of hypotheses • Description of proposed experiment procedures (participants, apparatus, tasks and procedure, experimental design)	/ 8
Formatting	• Correct grammar/spelling • Concise and well written • Follows formatting requirements including properly written references	/ 2
	Total	/ 20

Table 1: Project Proposal Rubric

Final Presentation and Data Supplement Rubric		
Section	**Criteria**	**Marks**
Introduction	• Research question clearly stated • Goal of the experiment is clear and reasonable	/ 5
Liteature Review	• Literature review provides a concise summary of what is known about the selected topic • Motivates research question and includes references	/ 5
Methods and Analyses	• Clear description of study setting and participants • Experiment design (apparatus, tasks, procedures) is clearly explained and justified as necessary • Clear definition and justification for independent and dependent variables • Null and alternate hypotheses are identified and justified • Statistical analysis plan is identified and justified	/ 20

Results	• Descriptive statistics of the data is analyzed and are presented intuitively
	• Inferential statistics are appropriately analyzed and are presented intuitively
	• Conclusions regarding the data are well supported by results from the statistical analysis
	• Data supplement:
	  ◦ Contains analysis and discussion on model assumptions
	  ◦ Figures add value to the analysis being conducted
	• Reported results are computationally reproducible using the provided code and data
		/ 20

Discussion	• Implications of results on answering the research question are identified and discussed • Implications of the experiment conclusions on the bigger picture of the background context are identified and discussed • Experiment strengths and limitations are adequately discussed • Potential future experiments that can shed further light on the background context are appropriately identified and discussed	/ 15
Formatting and Presentation Quality	• Correct grammar/spelling • Concise and well written • Follows formatting requirements including properly written references	/ 5
R Code	• Submitted R code is well-structured with comments.	/ 5
	Total	/ 80

Table 2: Project Report Rubric