

**University of Toronto**  
**Department of Statistical Sciences**

Course	Section	Session
STA305	L0101 & L0201	Winter 2024
Assessments	Due Date	Time
Project Report	April 5, 2024	11:59 pm

Date: 11 March 2024 (version v1)

**1. DOs**

	What to submit?	When due?	Where?
1	Project Report as a single PDF file (1)	5 April 2024	Report, Crowdmark
2	R/Rmd codes	5 April 2024	R/Rmd, Crowdmark
3	Outputs, data sets etc. as a PDF file	5 April 2024	Output, Crowdmark
4	Video Presentation file	5 April 2024	Quercus

**2. DON'Ts**

2.1 E-mail submissions will NOT be accepted.

2.2 SUBMIT A SINGLE PDF DOCUMENT FOR THE REPORT. DO NOT SUBMIT descriptive texts in one document and graphs/tables etc. in another.

**3. DETAILS in the following pages.**

# Project: General

## 1 An Overview

Each student (no groups allowed) will plan and perform a homemade factorial experiment, or carry out an analysis of data from an observational study available on websites, using the guidelines given in the following. This includes:

a) designing an experiment, collecting or generating the data, and analyzing the data in R (for factorial experiment),

OR

b) description and data searched from various sources [other than the data used in the course] and analysis in R (for observational study).

You are encouraged to choose a topic of your interest for this project.

The report should not be longer than 4 pages including tables and plots, and should be **submitted in one document**.

An example of report layout is given at the end of this document. Feel free to use this report as a template.

**Each student is required to submit the following four items** (*Where do you submit what?*):

### Crowdmark

1. Your final report as a single PDF file (due: **5 April 2024, 23:59 hr** on Crowdmark). See the question box: Report.

2. R/R Markdown codes file (due: **5 April 2024, 23:59 hr** on Crowdmark). See the question box: R/Rmd.

3. Supplementary materials(s) such as output file, along with necessary data files, etc. save them as PDF file(s) (due: **5 April 2024, 23:59 hr** on Crowdmark). See the question box: Output.

### Quercus

4. Video presentation of the objectives, methods, results, and conclusions of your experiment (due: **5 April 2024, 23:59 hr** on Quercus). See more on the format and size etc. in PART III.

## 2 Details

### 2.1 PART I. Guidelines for the project on factorial experiment and Observational Study

1a. Number of factors under investigation should be more than two. For example, you may develop a  $2^4$  factorial in blocks of size 8 or less, confound the interactions, preferably not the same set of interactions in all the replicates. Also use the common understanding “higher order interactions are either negligible or of lesser importance compared to the lower order interactions”. Main effects must not be confounded with blocks. You may also choose a fractional factorial design.

Other experimental designs could be a two factor factorials in Latin Squares, or as a split plot or strip plot design in RCBD or CRD, or even split-split plot design or other variants.

OR

1b. Alternatively, you can present an Observational Study (OS), using the data on websites, but not the study presented in the course.

2. Write a title for your experimental investigation or the observational study in the Report.

3. Write [Author: Your name, ID, specialization] in the Report.

### 2.2 PART II. Project Report

(Submit by 5 April 2024, 23:59 hr)

This document should have:

#### **Introduction/Description** (1/2 page maximum)

Introduce the need for this study, i.e., the rational behind this study, or, why and what did you plan to study/investigate? Write the objectives of the study. Cite any references used.

#### **Materials and Methods** (2 pages maximum)

Write in short the Materials and Methods you have pursued. This may include:

## **Experimental Design and Data**

Describe the experimental design used — details of the factorial designs you generated or acquired, features of the designs, comment on how this design align with the objectives you have set. Cite any references used. Details of the data. How did you get the data — randomly generated? found from a secondary source on the Internet? etc. Cite clearly the references. If you have arbitrarily inputted the data, write that as well.

## **Observational Study and Data**

Most likely you will be using study described on a website. Describe the source of the study giving its *url* — details of the study, time and location frame, list of variables: name and type of variables, treatment factor and a response variable. Cite any references used.

## **Statistical Analysis**

Details of the statistical methods used in the analysis. Cite any references used. R/RStudio recommended for computations save the codes as \*.Rmd. You may find it more convenient to run the \*.Rmd file on RStudio.cloud or JupyterHub, but they are not mandatory.

## **Results and Discussion** (1 page maximum)

Describe and display your project findings in text, tables and figures along the lines of the objectives (including whether the results support or contradict what you expected). Discuss an overall canvas of the study. Cite any references used.

## **Conclusion** (1/2 page maximum including References)

Present in short the highlights of the study.

## **References** (within 1/2 page of the Conclusion)

List the references in alphabetical order .

## **Format and size of the document:**

Due date for all the submissions: **5 April 2024, 23:59 hr**

Recommended font size is 12pt, but not less than 10pt.

## Crowdmark

1. Report in only PDF, no more than 4 pages, submit this file on Crowdmark. This file should not contain R codes. Upload in the question box Report.
2. R/R Markdown codes file on Crowdmark. Upload in the question box R/Rmd.
3. Output file, along with necessary data files, etc. save them as PDF file(s) on Crowdmark. Upload in the question box Output.

## Quercus

4. Video presentation of the objectives, methods, results, and conclusions of your experiment (due: **5 April 2024, 23:59 hr** on Quercus). See more on the format and size etc. in Section 2.3.

## 2.3 PART III. Video presentation.

(Submit by **5 April 2024, 23:59 hr**, and on Quercus)

Your video must meet the following criteria:

- a* In the beginning of the video, you must clearly present your student ID along with yourself. The grading TA must be able to identify you and your student ID number. Failure to present your student ID will result in a 0 grade for the video presentation. We recommend that you update your Quercus profile with a picture where your face is clearly identifiable.
- b* The presentation should not exceed 5 minutes. Any video beyond 5 minutes will not be viewed by the grading TA, and will not be considered when marking. In the video you should describe the objectives, methods, results, and conclusions of your experiment.
- c* The video may be of any form, so be creative! For example, you may include clips of yourself conducting the experiment while describing the experiment - beware that the clips will also count towards your 5 minute limit.
- d* Notes on video submission. If you are using a Mac, the Quercus media recorder submission page may not work on your Safari browser. The recorder works fine with Chrome or Firefox on both Mac and PCs.
- e* Beware that the Quercus media recorder doesn't allow pauses but you are able to retake your videos as many times as you want.

*f* Quercus accepts media file uploads of size up to 500MB if you are uploading a file. The supported file types for playbacks are FLV, ASF, QT, MOV, MPG, MPEG, AVI, M4V, WMV, MP4, and 3GP. If you upload any other file types, the TAs may not be able to assign you a grade.

It would also be acceptable if you upload the video via a link, either an internal UofT link such as <https://play.library.utoronto.ca/watch/73e40a5e7c2e9e1dfe08f25b5ff0510b> (Mymedia) or YouTube, rather than uploading an MP4 file. Uploading an MP4 file can be time-consuming and take up a lot of space, whereas viewing the video assignment online would be more convenient.

Also, in the file name, include your surname or first name. Example: WangSTA305Proj

### 3 PART IV. Grading Scheme/Grading Rubric.

[Total Marks = 10+10+5=25] Report (submitted on Crowdmark) [Question box: Report]

DAE: Design and Analysis of Experiments. OS: Observational Studies

[Marks =10] 4.1 Description of the design including objectives of the study.

Grade (points)	Descriptions
Excellent (10)	DAE: Strong evidence of original thinking and a clear explanation of why and how they conducted the experiment. OS: Strong evidence of original thinking and a clear explanation of why and why they carried out the study.
Good (8)	DAE: Grasped the basics of designing a factorial study; a good explanation of how and why they did the experiment. OS: Grasped the basics of designing an observational study; a good explanation of why and how they did the study.
Adequate (6)	DAE: Understood the basics of designing a factorial study, but may not have designed a factorial experiment. Provided an adequate explanation of the design. OS: Understood the basics of observational study, but may not have an observational study to work with. Provided an adequate explanation of the observational study.
Marginal (4)	DAE: Some evidence of understanding the basic design of a factorial study. Provided a poor explanation of their design. OS: Some evidence of understanding the basic design of an observational study. Provided a poor explanation of their observational study.
Inadequate (2)	DAE: Little evidence of even a superficial understanding of a factorial design. Little explanation about how or why the design was chosen. OS: Little evidence of even a superficial understanding of an observational study. Little explanation about how or why the observational study was chosen.

[Marks =10] 4.2 Analysis of the data (including methods, 2 pages maximum)

DAE: Include appropriate plots and calculations such as: main effects and interactions; estimated variance of the effect (if replicated); confidence intervals for true values of effects (if replicated); Lenth plot; or half normal plot.

OS: Present screening of covariates, computation of propensity scores, approaches for the adjustment for treatment factor effect, you may use the R codes discussed during the lectures.

Grade (points)	Analysis of the data (DAE, OS)
Excellent (10)	Strong evidence of data analysis skills. Probably used R to do calculations and plots, but calculations and plots might also be done neatly by hand.
Good (8)	Good evidence of data analysis skills. Appropriate calculations were done, and maybe appropriate plots were included.
Adequate (6)	Understood the basics of required data analysis.
Marginal (4)	Some evidence of understanding the basic data analysis required, but might not have carried out all the appropriate calculations and plots.
Inadequate (2)	Little evidence of even a superficial understanding of the data analysis required to analyse a factorial design/OS.

[Marks =5] **4.3 Conclusions.** (1/2 page maximum including references)

What conclusions can you make based on the results of your experiment/OS? Write a paragraph or two outlining these conclusions.

Grade (points)	Conclusion
Excellent (5)	Conclusions are highly appropriate given the experiment/OS conducted. Clearly written.
Good (4)	Conclusions are appropriate given the experimental/OS context. Writing is good.
Adequate (3)	Some conclusions are appropriate; other obvious conclusions might be missing.
Marginal (2)	Some evidence that there was an understanding of the basic conclusions, but several obvious conclusions not stated.
Inadequate (1)	Little evidence of even a superficial understanding of the conclusions that can be drawn from the experiment/OS.

[Total Marks = 5 ] **4.4 R/Rmd** (submitted on Crowdmark) [Question box: R/Rmd]

**Assessment objective: Appropriate use of built-in R functions.**

Grade (points)	R Markdown notebook.
Excellent (5)	Appropriate R functions are used correctly to perform the intended tasks. Entire notebook runs without an error.
Good (4)	Appropriate R functions are used but may contain mistakes in their usage. May contain errors but they do not interrupt the analysis steps. Reproduces similar results as presented in the report but some numeric results may be different.
Adequate (3)	Most R functions are used inappropriately and do not perform the intended tasks. Contains errors that interrupt some parts of the analysis steps. Produces conflicting results for minor parts of the conclusions presented.
Marginal (2)	Most R functions are used inappropriately and do not perform the intended tasks. Contains errors that interrupt the analysis steps but requires only minor fixes. Produces conflicting results for most of the conclusions presented.
Inadequate (1)	Contains major errors and does not reproduce the result presented.



[Total Marks = 10 ] **4.5 Video presentation** (submitted on Quercus) [Quercus Quizzes: Video Presentation].

**Assessment objective: Effective presentation of study objectives, methods, results, and conclusions.**

Grade (points) Excellent (10)	Presentation Information is presented in a logical and interesting sequence. Experiment/OS objectives, methods, results, and conclusions clearly stated, repeated appropriately, and strongly supported throughout the presentation. Clearly audible voice throughout the video.
Good (8)	Information is presented in a logical sequence. Experiment/OS objectives, methods, results, and conclusions clearly stated and supported throughout the presentation. Audible voice throughout the video.
Adequate (6)	Presentation jumps around topics making it difficult to follow. Experiment/OS objectives, methods, results, and conclusions are stated but minimally supported through out the presentation. Voice is not audible in some parts of the video.
Marginal (4)	Presentation has no sequence of information and audience cannot understand the presentation. Experiment/OS objectives, methods, results, and conclusions are not explicitly stated and need to be guessed. Majority of the video is not audible.
Inadequate (2)	A video presentation was submitted with the student ID presented, but the video is not audible throughout the presentation.

**STA305H1F Sec L0101 , W2024**

## Sample Project Report

*Student's Name*

*Student's ID*

OPTIONAL: *Student's Specialization or Major/Minor)*

*December ??, 2024*

### **Introduction**

add Introduction/Description here.

### **Materials and Methods**

add experimental designs/OS and statistical analyses here.

### **Results and Discussion**

add results and discussion here.

### **References**

add references if any.