STA305/1004 - Project

Due Dates: November 11, 2019, 23:59 (Proposal) ; Dec 6, 2019, 23:59 (Project)

For this project, each student - no groups allowed - will design, analyse, and communicate the results of a homemade ***replicated*** 2^k factorial experiment, where $k \geq 2$. Each student is required to submit **four** items:

0. Draft proposal of the experiment that you plan to conduct; (due: Nov. 11, 23:59 on Quercus)
1. `R Markdown` document used to author your report, along with necessary data file(s); (due: Dec. 6, 23:59 on Quercus)
2. pdf export of your report; (due: Dec. 6, 23:59 on Quercus) and
3. video presentation of the objectives, methods, results, and conclusions of your experiment. (due: Dec. 6, 23:59 on Quercus)

Detailed description and requirements are provided below.

Description.

Each student is asked to plan and perform a homemade factorial experiment. The experiment must be a replicated 2^k factorial experiment, where $k \geq 2$, and the number of replications is at least two. That is, the experiment should include 8 measurements at the minimum.

Students should start thinking about decide what they want to study. The number of possible topics is very large. We strongly recommend selecting a topic that you are interested in and will enjoy working on.

Submission requirements.

You are responsible for submitting the following via Quercus by the due dates. The grading rubric for the project is attached at the end of this document.

0. Draft Proposal of Experiment.

Due date: November 11, 23:59

- An R Markdown document (i.e., has .Rmd extension) and pdf export of the R Markdown document that states your experiment's objectives, methods, statistical analysis plan.
  - The objectives should be concise statements about what you are hope to learn, and the methods are your plan for collecting the data. In the example of the baseball experiment in the introductory video [add link] the objective could be stated as: "What are the effects of grip, and stance on ball speed."

  - The method could be stated as: "I will collect measurements of my ball speeds while varying grip and stance at two different levels each. I will vary my grip with placements of the index and middle fingers, open versus closed. The stance factor will be based on the height of the left leg during the wind-up motion. I will perform 5 repetitions in a randomised order since my arm may loosen up and/or tire out effecting the ball speeds. The order will be radomised using an R script prior to the measurement session. I will also randomly pick a baseball from a ball cart each time I pitch. It is well known that a baseball with worn out seams result in lower ball speeds. Finally, a friend will measure my ball speeds using a speed radar gun while I pitch following the radomised order."
  - The statistical analysis plan is a brief description of how you plan to analyse the data, including which summary statistics and data visualizations you plan to include.

- The draft will not be graded. If a draft with the content outlined above is submitted then you will receive full credit. TAs will give students feedback on their draft as soon as possible. The draft is an opportunity for you to receive early feedback on your ideas and plans so that you can incorporate the feedback into your final submission.

1. R Markdown document.

Due date: December 6, 23:59

You are required to write your report and perform your analysis in an R Markdown document using `R`. Submit an error-free `R Markdown` document (`.Rmd` file) that contains the `R` codes used to perform analysis. If your code requires extra files such as your observed data, you must include them as well. Please keep your data files and `.Rmd` file in the same directory to help accelerate the grading process. The grading TA must be able to run your notebook without an error on their machine after downloading the files as they are. (The TA will install extra packages if necessary.)

We recommend using `https://rstudio.cloud/` if you don't already have RStudio and/or LaTeX installed on your computer.

An introduction to R Markdown is available on the course web site.

2. PDF report.

Due date: December 6, 23:59

A maximum 4 page PDF report that contains the following sections. You are encouraged to use `R Markdown` to create your report. However, the report must not include any `R` codes. (Include `knitr::opts_chunk$set(echo=FALSE)` in your first code chunk to hide code chunks throughout your output PDF document.)

1. **Description (1 page maximum).** Include how and why you conducted the experiment. What do you hope to learn by doing this experiment?

2. **Analysis of data (2 pages maximum).** Include appropriate plots and calculations used to answer experimental question(s). These may include main effects and interactions, estimated variance of the effect; confidence intervals for the true values of effects; Lenth plot; or half normal plot.

3. **Conclusions (1 page maximum).** State your conclusions based on the results of your experiment in a paragraph or two.

3. Video presentation.

Due date: December 6, 23:59

A video presentation of your study. Your video must meet the following criteria:

- 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 identifable.

- 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.

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.

Submission Instructions.

1. Submit your report and analysis including the PDF document, `.Rmd` file, and any associated data files under *Project: Report*.

2. Submit your video presentation under *Project: Video*. You may use Quercus's built-in media recorder or upload your own video file. The uploading process may take a few minutes.

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.

- Beware that the Quercus media recorder doesn't allow pauses but you are able to retake your videos as many times as you want.

- 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.

Grading Rubric

	Excellent (5)	Good (4)	Adequate (3)	Marginal (2)	Inadequate (1)
PDF Report (20).					
Description.	Strong evidence of original thinking and a clear explanation of how and why they conducted a replicated factorial experiment.	Evidence of original thinking and a mostly clear explanation of how and why they conducted a replicated factorial experiment.	An adequate explanation of how and why the experiment was conducted, and a replicated factorial experiment conducted.	A factorial experiment was not conducted, although the experiment that was conducted is appropriate for the stated objectives. There is little evidence of understanding a factorial experiment.	Little evidence to no evidence of experimental design or the analysis of experimental data.
Analysis.	Appropriate data analysis was conducted to answer objectives of experiment including appropriate calculations and plots.	Almost all of the data analysis was conducted to answer objectives of experiment. Most of the calculations and plots are appropriate.	Most of the data analysis was conducted to answer objectives of experiment. Some calculations and plots may be superflous or inappropriate.	Some of the data analysis was conducted to answer objectives of experiment, although there is no statistical evidence to support all of the objectives. Some calculations and plots may be superflous or inappropriate.	Most of the data analysis conducted does not help answer objectives of experiment. Several calculations and plots are superflous or inappropriate.
Conclusions.	All the conclusions are clearly stated, and supported by statistical analysis in the context of the experiment.	Almost all the conclusions are clearly stated even if a few are not clear. The clearly stated conclusions are supported by statistical analysis.	Some of the conclusions are stated, some may be missing or unclear. The stated conclusions are supported by statistical analysis.	Some of the conclusions are stated, some may be missing, and none are supported by statistical analysis in the context of the experiment.	None of the conclusions are stated, and none are supported by statistical analysis in the context of the experiment.
Organization.	Very well organized with clear headings and sections. Excellent flow from one section to the next with tables and plots carefully tuned and placed.	Well organized with clear headings and sections. Flow from one section to the next might be lacking.	Generally well organized but some sections were muddled. Appropriate tables and plots were used but might be poorly presented.	Sections unclear and no attempt to flow from one topic to the next. Some tables and plot might have fundamental flaws in their presentations.	Difficult to read the report. For example, the report does not contain headings, figures are far away from where they are referenced in the text. Missing required parts.

Grading Rubric

	Excellent (5)	Good (4)	Adequate (3)	Marginal (2)	Inadequate (1)
R Markdown notebook (5).					
Appropriate use of built-in `R` functions.	Appropriate `R` functions are used correctly to perform the intended tasks. Entire notebook runs without an error with necessary packages installed. Reproduces the same numerical results presented in the report.	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.	`R` functions are often 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.	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.	Contains major errors and does not reproduce the result presented.
Video presentation (5).					
Presentation.	Information is presented in a logical and interesting sequence. Experiment objectives, methods, results, and conclusions clearly stated, repeated appropriately, and strongly supported throughout the presentation. Clearly audible voice throughout the video.	Information is presented in a logical sequence. Experiment objectives, methods, results, and conclusions clearly stated and supported throughout the presentation. Audible voice throughout the video.	Presentation jumps around topics making it difficult to follow. Experiment objectives, methods, results, and conclusions are stated but minimally supported through out the presentation. Voice is not audible in some parts of the video.	Presentation has no sequence of information and audience cannot understand the presentation. Experiment objectives, methods, results, and conclusions are not explicitly stated and need to be guessed. Majority of the video is not audible.	A video presentation was submitted with the student ID presented, but the video is not audible throughout the presentation.