Data Exam One Instructions

36-402 Advanced Methods for Data Analysis

Due Friday, April 2, 2021 at 3pm EDT
(grace period until Sunday, April 4, 2021 at 8pm EDT)

Data Exam Instructions

The data exam is an exam. That means:

- **You must not communicate at all with your classmates, your friends, or anybody other than the professors and the TAs about the Data Exam,** until one week after the exam is due. This rules applies to Piazza, online discussions, emails, Stack Overflow, and so on. Evidence of illegal communication will be taken seriously and handled according to the University's Policy on Academic Integrity.

- **Leave yourself plenty of time to submit the Exam.** Unexpected PDF formatting problems, weird R errors, and spontaneous Internet failures all happen, usually at the worst possible time. If you're using an Rmd (as we recommend), knit it regularly to ensure it works, and submit to Gradescope well in advance of the deadline. The grace period is intended to save you if problems pop up—but it can only help you if you try to submit earlier than 7:58pm on Sunday.

- During office hours, you can ask questions about any course material covered in class (including theory, concepts, and previous homework solutions and class examples) as well as questions about exam rules and format. However, any attempts to link this material to exam questions would not be valid as questions—including asking whether "this approach" or "what I've done here" would be OK or not for an exam question.

- Whereas the HWs are part of your learning, the data exam is part of an examination, so you need to make the connections between course material and solving a data analysis problem yourselves. Please be mindful that the TAs are limited in their ability to provide help during the data exam, and that they are supposed to step back if pressed for exam-specific answers.

- Piazza posts will be private by default. Only clarifying questions about the exam instructions should be public.

- You are of course allowed to consult your class notes, your homework assignments, or any of the other 402 course materials. You are permitted to use the Internet as a resource, but you may not use it to communicate with anybody else or to ask questions about your assignment.

- Everything you write must be in your own words, and you must explicitly cite any sources used other than the class materials. If you do not clearly indicate when materials come from other sources, you are committing plagiarism.

- The instructions will include a list of specific questions/issues that must be addressed in your report. You can raise other issues as you see fit. But for the listed issues, you **must do the following**: In your final PDF, mark the sentence that corresponds most closely to **each** question and write the question number that is being addressed. An example format is given in the template Rmd file.

- We will send any important announcements or clarifications about the exam via Canvas. Be sure to check your Andrew email regularly.

The Data Exam will be graded according to the posted rubric, which sets out the sections the report should contain and what should be contained in each section. You should **read the rubric first.**

Submission Checklist

When you submit your data exam, check that you have met these requirements:

1. Your exam is a PDF file formatted according to the rubric requirements.

2. You have checked all the requirements in the rubric and answered all the assigned questions.

3. You have marked your answers to all assigned questions, according to the instructions in the template.

4. In Gradescope, you marked what pages of your PDF contain each section of the report.

5. The sections of your report are marked with clear headings in large font.

You must submit via Gradescope.

Useful Resources

- A guide to R Markdown and all its various options: `https://rmarkdown.rstudio.com/lesson-1.html`

- A guide to all the knitr options, for your R Markdown submission, can be found here: `https://yihui.org/knitr/options/`

- To prevent R output (such as messages printed by code) from appearing in your PDF, use the `results = 'hide'` chunk option.

- To hide R warning messages, use the `warning = FALSE` chunk option.