# Rules for In-Class Students:

**Exam 2 Guide**

IST 3420, Spring 2021 Zou

1. Exam 1 is 60 minutes long and will be conducted on Canvas on **April 7 during our normal class time**. Try to log in Canvas at least 5 minutes before the exam starts and make sure you have good internet connection.
2. The exam is a closed book exam. Textbooks, notes, internet, electronic devices, and other references are **NOT** allowed to be used during the exam.
3. We will use **Proctorio** for remote proctoring for our online exams. Google Chrome on a laptop or desktop computer, a microphone, webcam, and a stable internet connection will be needed to take exam 2. More details are available at <https://keeplearning.umsystem.edu/students/learning-remotely/taking-proctorio-tests>

# About the Exam

1. The exam contains three kinds of questions: (1) multiple selection; (2) programming syntax, and (3) short answers.
2. Use the exam guide to guide your review of the lab assignments, demo codes, slides, and reading materials.
3. During the exam, use your time properly. If you get stuck in one question, you need to move on and come back later.
4. Read each question carefully.

# Coverage of Exam 2

**Note: The exam may cover other issues NOT mentioned on this guide.**

|  |  |  |
| --- | --- | --- |
| **Module** | **Content to Cover** |  |
| #4: Cleansing and Manipulating Data | * Be able to use regular expression to manipulate strings in R * Be able to use base R and dplyr to cleanse and transform raw data * Understand basic data structure conversion functions * Be able to use dplyr to manipulate dataset such as sub-setting, grouping, merging data, making new variables * Be able to use tidyr::gather() to reshape dataset between wide format and long format | * Data Wrangling with dplyr and tidyr Cheat Sheet: <https://rstudio.com/wp-content/uploads/2015/02/data-wrangling-cheatsheet.pdf> |
| #5: Data Summarization and Visualization | * Understand the importance of data visualization for business analytics * Be able to use base R plot and ggplot2 to summarize and visualize data * Be able to choose appropriate tabular and basic graphic methods for different types of data (qualitative vs. quantitative) * Be able to use boxplot rule to detect outliers * Understand Simpson’s paradox and be cautious when analyzing data combining all groups | * Data Visualization with ggplot2 Cheat Sheet: <https://rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf> |

Graphical user interface, diagram, application

Description automatically generated