

BST 210 HOMEWORK #4 – Group Project

Due 11:59pm, Tuesday, October 22, 2019

***Please be sure to submit your group assignment by 11:55ish pm (or before) to prevent any glitches in the upload from precluding your timely submission.**

***Please work well in advance, getting help during office hours and labs, as there will be no extensions given for this assignment, outside of extreme, extenuating circumstances which must be communicated in advance to the primary instructor.**

Group Project:

For this data analysis group project, you should work with 1 to 3 colleagues (groups of size 2 to 4, no singletons). Groups will submit one solution and group members will receive the same grade.

“The Heart and Estrogen/progestin Replacement Study (HERS) is a randomized, double-blind, placebo-controlled trial designed to test the efficacy and safety of estrogen plus progestin therapy for prevention of recurrent coronary heart disease (CHD) events in postmenopausal women. The participants are postmenopausal women with a uterus and with CHD as evidenced by prior myocardial infarction, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty, or other mechanical revascularization or at least 50% occlusion of a major coronary artery. Between February 1993 and September 1994, 20 HERS centers recruited and randomized 2763 women. The design, analysis, and conduct of the study are controlled by the Steering Committee of Principal Investigators and coordinated at the University of California, San Francisco. HERS was the largest trial of any intervention to reduce the risk of recurrent CHD events in women with heart disease and the first controlled trial to seek evidence of the efficacy and safety of postmenopausal hormone therapy to prevent recurrent CHD events.”

- NCBI, US National Library of Medicine, NIH

Students, please get your hands dirty with this dataset—have fun! Perhaps there are interesting questions you can derive as a group that you wish to answer from a statistical and real-world standpoint. The dataset contains 36 fields, a number of which have potential to serve as outcome variables for sub-analyses. The sky is the limit on analysis, but remember your ultimate goal for this assignment: a brief, concise write-up in the form of an abstract, statistical methods section, and some visual displays. Support each other and work as a team.

Use basic statistics, applied regression, and your knowledge of model building and model selection, confounding and effect modification, and regression diagnostics to develop and assess an appropriate model (or models) to address the research questions you develop. Use logic, common sense, and good statistical methods to do this.

Write up your findings by developing:

- A title
- A 250-word abstract matching a standard journal style (e.g., Background, Methods, Results, Conclusions)
- One to two paragraphs describing your statistical methods
- Up to five tables and/or figures (you may need fewer--no multi-panel figures)
- One to three paragraphs summarizing your results (including referring to your tables and figures)

All of this should be appropriate for submission as part of a manuscript. Use appropriate number of decimal places, labeling of variables, footnotes, nice layouts of abstracts and tables, and a good writing style, like you were submitting your work to a journal. Look at a few articles in journals for ideas. Higher grades will be assigned to strong analyses, perhaps with some interesting or challenging methods or findings, but also answering your overall research question/s well, having a professional appearance and writing style, and having nice tables, figures, and appropriate statistical methods and summarization of results. Focus your summarization of results to only include the most important information, not all the technical details, as you would see in a manuscript. (Excessive, less relevant material would be considered a negative in such a setting.)

You will submit your write-up in .pdf format online, through Canvas. You may also attach at the end *a limited number* of computer outputs (with comments) to help justify your approach and findings, or possibly to include some details not included in your summarizing of results text. However, the grading of this assignment will focus on your findings as described above, and only a limited number of additional pages should be submitted.