

For the project you should do the following:

1. Find an interesting topic that you like and know more about than we do, so that you can teach us something new during your presentation. Please don't choose topics that might offend anyone or be shocking to anyone (not too medical, religious, or political, that is).
2. Find a data set on the chosen topic, for which a regression analysis would be appropriate. It should contain between 5 and 10 predictors and between 50 and 100 rows. The data set should not be older than 10 years old.
3. Run a quick regression analysis on your data set, make a PDF file with your code and relevant output, and send me the data set and the PDF file for confirmation and approval by Wednesday, **November 13, 2024**, at 10 PM. Send it to stat410f24@mail.com with "Project" in the subject line.
4. Regression model(s) should be chosen from the following list. The number of spots is limited, so it is to your advantage to email me earlier rather than later. Before choosing a regression, please check that it is still available.
 - 1) General linear regression (2 spots)
 - 2) Box-Cox transformation + gamma regression (0 spots)
 - 3) Binary logistic + probit + complementary log-log (0 spots)
 - 4) Cumulative logit + probit + complementary log-log (1 spot)
 - 5) Generalized logit model for nominal response (2 spots)
 - 6) Poisson regression (1 spot)
 - 7) Zero-truncated Poisson (2 spots)
 - 8) Zero-inflated Poisson (2 spots)
 - 9) Hurdle Poisson (1 spot)
 - 10) Negative binomial regression (0 spots)
 - 11) Zero-truncated negative binomial (2 spots)
 - 12) Zero-inflated negative binomial (2 spots)
 - 13) Hurdle negative binomial (2 spots)
 - 14) Beta regression (0 spots)
 - 15) Zero-inflated beta regression (2 spots)
 - 16) One-inflated beta regression (2 spots)
 - 17) Zero-one-inflated beta regression (2 spots)
 - 18) Random slope and intercept model for normal response with covariance structure for error + generalized estimating equations models (2 spots)
 - 19) Generalized linear regression with mixed effects (2 spots)
5. Prepare a **5-minute** PowerPoint presentation to talk about the background, data, and regression analysis. Complete SAS and R codes and relevant outputs must be included on your slides. Also, include the title page, use a nice colorful slide design, insert slide numbers, and add a wonderful thank-you slide.
6. Email me your slides for approval **at least one day** before your presentation. Send to stat410f24@mail.com with "Slides" in the subject line.

7. Give your PowerPoint presentation at one of the sessions on December 4 and 9 (see the schedule of presentations for details).
8. When listening to presentations of others, you will be required to fill out a questionnaire (see the posted file). Your participation will be graded based on how well you complete the questionnaires. Presenters' grades will be based on your answers to those questionnaires. Please prepare 20 questionnaires. You will be listening to 20 presentations and filling out 20 questionnaires.
9. Write a **project report**. It should contain a gorgeous cover page with an image related to the content of your project, and two or three pages of write-up, consisting of an introduction, background, data description, results, and conclusion. Complete SAS and R codes and relevant outputs must be included in an appendix. A list of references should be added at the very end. It should contain the link to your data source and two or three books/articles on the topic. The report will be due in pdf format and sent to stat410f22@mail.com by 10 PM on **the day of your presentation**.
10. Your score for the project will be based on: slides (25 points), presentation (25 points), attendance (25 points), and report (25 points).