

## Project for STAT 705

You will analyze a data set of your choosing, using a method (or combination of methods) we have discussed in this class, such as:

Two-Factor ANOVA, Multi-Factor ANOVA, Random Effects ANOVA, Mixed Effects ANOVA, Analysis of Block Design, Repeated Measures, Analysis of Covariance, ANOVA for Nested Designs, Piecewise Regression, Nonlinear Regression, Generalized Linear Models, Nonparametric Regression, or Regression Trees/Random Forests.

The data set could be one that is of personal or academic interest to you and that fits one of the analyses studied in STAT 705. It should be a real data set which you have not analyzed before and which has not been analyzed in a textbook (using the methods of STAT 705, anyway).

### Project Part I– Data Set Proposal/Description:

In Part I you should write a one- to two-page typed description of the data set you propose to study. You should include details about the response variable, about factors and/or predictor variables, and about the number of observations. If there are issues such as unbalanced data, comment on these. Discuss the source of the data set, and whether the data come from an experimental or observational study. In addition, please include a printout of the data set in an appendix (or if it is quite large, a selected part of the data set).

You should also include a general proposal for what sort of analysis you plan to do with this data set. If there are any hypotheses/research questions that are of interest from the beginning, you might mention those.

### Project Part II – Written Report:

You should write a concise report summarizing your analysis. The report should be five to eight (typed) pages, not counting any SAS/R output, graphs, etc., which you may wish to include as support or illustration for your analysis.

The style of the report is up to you, but the best reports will address many of the questions and details studied in class when we discussed the relevant type of analysis.

Some things to include (depending on the data set and choice of model) might be:

- An introduction and discussion of the data set itself (this is in part I)
- A statement of the fitted model, with any relevant interpretations
- Summary results for any relevant hypothesis tests or confidence intervals
- Discussion of variable selection or study design (if appropriate)
- A summary of model assumptions and any violations you may have seen
- Any remedial action you took to fix such violations
- Your overall conclusions about the data, based on your analysis

Part II of the project (the final project report) will be due *on or before Tuesday, April 27, 2021*. It will count for 10% of your final grade.