

For this lab, you will investigate a question of your choosing which can be answered through data collection/experimentation and the analysis of one of the following:

1. an ANOVA with at least 2 factors, or at least 1 factor and 1 block factor.
2. a linear regression model with a minimum of 3 distinct predictor variables.
3. a logistic regression model with a minimum of 3 distinct predictor variables.

You may also use data found from an online source, but not one in which the model you are using was already used to analyze the data.

Your report should be thorough, and use many of the techniques we have discussed in class for whichever type of analysis you choose. Your write-up should include (but not necessarily be limited to) the following:

- A clearly stated problem or question that can be investigated through the statistical analysis of data. Why are you personally interested in this question, or why is it important? Give a decent introduction.
- Explain how the data was collected. Did you conduct an experiment or implement a survey? Was the data found using an online source?
- Are there flaws in the sampling methodology (i.e. how the data was collected)? If you conducted an experiment, were certain variables appropriately controlled for or randomized?
- Describe the model you will be using to analyze the data. What are the model assumptions?
- Estimates and interpretations of model parameters. What do these estimates tell you about your data, and about the variables you are studying?
- Are model assumptions met?
- Can your model be simplified/reduced? Remember: the best statistical models are ones that explain/fit the data well, but are also reasonably simple.
- Conclusions and observations about your model. Do you think it was a good model? Did it seem a good fit for the data? Did it tell you anything interesting or useful? What are your take-aways?