STAT 210
Applied Statistics and Data Analysis
FINAL PAPER PROJECT

This project gives you an opportunity to apply the techniques studied in the course and to have hands-on experience on using these models on a real data set. The project is not long and should not be elaborate. Try to work on a problem you find interesting and challenging.

Teams. Ideally, groups of three or four students should develop the projects. If for any reason this is not possible in your case, please talk to me before starting your work. It is very important that all members of the group share a similar workload and all take part on the presentation. **You should send me an email with the topic you plan to work on and the persons in the team by Sep. 30.**
There are three deliverables: a proposal, a final report and a presentation.

Proposal (6-8 pages, including figures, double spaced). **Due on October 21** but try to have this ready as soon as possible.

You are required to select a data set for which the methods covered during the course are appropriate tools for analysis (regression, analysis of variance/covariance, logistic regression, etc.). The proposal should include: (1) description of the dataset; (2) the scientific goals, specific hypotheses; (3) preliminary plots, (4) plan for analysis and modeling.

Final Report (12-15 pages including figures double spaced). **Due on Nov 30**. Again, try to have it ready before then.

Reports have to be complete but brief. Complete but brief. Complete but brief. Extract the signal, remove the noise.

Presentations. You will be asked to make a short presentation (15-20 min) on your project. This will take place on the date assigned for the final exam, so it will be on the week starting on Dec 13. Make sure you include all the relevant material in your presentation and remember that your audience may not know beforehand the problem you will talk about, so plan a brief but clear introduction to the problem and explain all the relevant facts before plunging into the analysis of the data.

Grading
This project is 30% of your final grade for the course, distributed as follows: 5% for the project, 15% for the report and 10% for the presentation.

Feel free to contact me with any questions (joaquin.ortegasanchez@kaust.edu.sa).

<u>DATA ANALYTIC STRATEGIES:</u>

- Perform adequate exploratory analysis of the data and provide a complete, yet succinct, presentation of the results.

- Clearly state the statistical model used when presenting model estimates.
- Clearly state the model building/selection/validation criteria used to address the scientific question(s) of interest.
- Perform adequate model diagnostics.
- Provide precise interpretations of the parameters in your model (or your estimates of those parameters) in the context of the scientific problem.

GENERAL GUIDELINES ON THE REPORT

Your final analysis should be presented in the form of a brief report between 12-15 double-spaced pages including relevant tables and figures. Your report should be structured as follows:

- Abstract - A brief summary of your basic findings
- Introduction - A brief introduction/motivation to the problem at hand, relevant details about the data, additional relevant scientific information from searching the web, for example, and what is to be addressed
- Statistical Methods - A discussion and justification of the methods you have used to analyze the data and how you went about analyzing the data. Don't forget to describe in some detail how and why the particular model was selected.
- Results - A presentation of the results of your analysis. Interpretations should include a discussion of statistical versus practical import of the results.
- Discussion - A synopsis of your findings and any limitations your study may suffer from. Present final conclusions in terms that non statisticians will understand. Quantitative and qualitative aspects should be discussed.

Your report should be succinct and to the point! It should be written in a language that is understandable to the scientific community.

IMPORTANT DATES

Sep. 30 **Email with team and topic.**
Oct. 21 **Project proposal.**
Nov. 30 **Final Report.**