FISH 604

Homework #3 (Individual Project)

TOTAL: 50 points

1. Write a short draft paper describing your student project for this class
Papers are due Friday, Oct 8

For your individual class project, you should define and describe a "manageable" research problem. Choose a problem that is appropriate for a short class project. The problem should be amenable to be addressed through some type of regression approach (see Appendix 1 for general suggestions on formatting your data); that is, you should identify a problem or question that can be addressed by modelling a response (univariate or possibly multivariate) as a function of a set of explanatory variables, but the emphasis of this homework is not on the modelling.

You should already have or be able to easily acquire the necessary data and either have done or at least have it ready for an exploratory analysis. Make sure you choose a manageable data set. For the introductory draft paper, use the following format:
- Preliminary title
- Introduction / Background (10 points) – Provide a brief background (i.e. some context for your specific class project) and a rationale.
- Goals / objectives (15 points) – What is (are) the question(s) you would like to answer? What is your overall goal? What are some specific objectives for the class project? If appropriate, you may want to include specific hypotheses or a conceptual model. A conceptual model is a representation of how you think your explanatory variables and your response(s) are linked. For example a simple graph depicting (hypothesized) links and explaining the nature of the links or a verbal description of what you think the relationships are. You should clearly articulate the specific questions and hypotheses that you are trying to get at with the help of a suitable statistical model. However, the emphasis in this section should NOT be on the statistical modelling itself but on the context for why and how you may eventually want to fit a statistical model. Setting up this context will help you (and me) identify an appropriate model. In a typical scientific paper, this section may be the last paragraph or two of your introduction, before you describe your methods / approach. It can be short, but should be carefully crafted to set the stage for the analyses to come.
- Methods: Data description (20 points) – Provide a brief description of the available data and how it will help address your question(s). What type of data do you have? What is the structure of the data (response or "dependent" variable(s) and exploratory or "independent" variables, temporal/spatial extent, categorical/continuous/counts, etc)? How were the data obtained? Describe the study area and experimental design or sampling design, as appropriate. Provide sources if you did not collect the data! Include a preliminary exploratory analysis consisting of a brief description of some of the main characteristics of the data. This should include, if appropriate, at least one or two summary plots suitable to your data (e.g. map, time series plot, pairwise scatterplot(s) of your explanatory variables and/or response vs explanatory variables, etc.) and/or one or

two tables summarizing key aspects of the data. In a typical scientific paper, a concise version of the data description may be the first section of the 'Materials and Methods' section, while some of the descriptive figures (maps, time series, etc) may eventually go in the results section.

- <u>References</u>

(5 point for overall structure and layout)

<u>Paper length:</u> No more than 2-3 pages of text plus a few figures/tables to illustrate the nature of your data are generally sufficient. I will provide feedback and you can re-write this section and expand it into a full-fledged paper for the final project report, towards the end of the semester (which will include results, conclusions, etc.).

The rationale for writing this paper is several-fold:
- to get you to think clearly about what your goals & objectives are
- to communicate your goals and objectives and the nature of your data to others
- to help me and others understand your specific objectives and your data to better help you

Please write in complete sentences and <u>structured paragraphs with clear topic sentences</u>, as you would for a report or scientific paper. Figures and tables should be <u>numbered and include stand-alone captions</u> that allow readers to understand what is being shown. Note that table captions typically appear <u>above</u> the table, while figure captions appear <u>below</u> the figure. You can embed figures in the text or include them at the end of the document (Tables first, figures second). You should refer to each figure and table in the text (in the order in which they are numbered). Tables and figures are used to support the text, but do not describe the table or figure in the text (that's what the caption is for). Include key references, but focus the background on what is relevant to the class project and keep it to a minimum (I don't want to read your full thesis or dissertation proposal!). If you have a reference that tackles a similar problem and may include relevant methods, please include it.

Appendix 1: Formatting your data

For analyses in this class, you should be able to format your data to fit into a single table of values with <u>each row corresponding to one 'observation' and each column corresponding to a variable</u>. Variables may include numeric values and categories and should include at least a response variable (or several), and one or more (and perhaps many) explanatory variables (= independent variables or predictor variables). These may include but are not limited to:
- Aspects of the experimental design:
    - year, day of year, time of day, etc. for temporal data
    - latitude, longitude, location, etc. for spatial data (may include a 3rd dimension such as depth in the ocean or altitude/elevation
    - replicate number, treatment, experimental 'block', etc
- Environmental covariates:
    - Continuous or categorical measures that you took at the time of sampling and may be unique to each sample or may apply to groups of samples (e.g. they may be associated with a station, but may be replicated for multiple samples taken at each station
    - Continuous or categorical measures that you later associated with a particular sample (such as satellite derived observations or model-predicted environmental conditions etc.). These may similarly be unique to each individual sample or may apply to multiple samples (e.g. all samples taken in a given year may be associated with some measure of average temperature conditions over the entire study region). You may also want to include several different versions of the same variable, for example you may have sampled once a month over the course of a year. You may want to include 'month' (1-12), season (Spring, summer, fall, winter) and/or Julian Day (days since Jan 1) so you can capture seasonal trends in one of several ways for comparison.
    - Lagged environmental conditions: Think about the appropriate temporal/spatial scale over which to measure the environmental conditions that affect your response. Examples:
        - If you study the year-class strength of some fish species, annual year class success as measured by the abundance of one-year old fish may be influenced by conditions measured during the preceding fall or winter
        - If you study (daily or annual) growth of fishes, growth between time *t-1* and time *t* may in part depend on fish size at time *t-1*.
        - If you're studying processes in the coastal ocean that may be influenced by runoff, you may want to include a measure of precipitation over the previous week or longer

The example below is based on field observations, where Chlorophyll *a* (Chl.a) was measured *in situ* at a number of locations and water depths and the goal may be to build a predictive model for Chlorophyll *a* concentrations.

Case.ID / Sample	Chl.a	Station	Lat	Long	Location	Stratum	Year	Month	Season	DOY
1	12.6	FB1	58.25	-134.9	FunterB	surface	2018	3	Spring	86
2	9.2	FB1	58.25	-134.9	FunterB	mid	2018	3	Spring	86
3	21.4	FB2	58.19	-134.9	FunterB	deep	2018	3	Spring	87

…

Case.ID / Sample	Depth	Temperature	Salinity	PAR	Weather	Fishery	Precip (lagged)
1	2	8.6	26.3	81	rainy	open	23
2	14	3.8	30.6	52	rainy	open	23
3	45	2.9	31.4	12	rainy	open	23