Data Development for Graduate Research

Class Sessions

Friday (10-11 AM)

Instructor

Dr. Peter Houk

Phone

(671) 735-2188

Email

houkp@uguam.uog.edu

peterhouk@gmail.com

Office Location

UOG Marine Lab 102

Office Hours

Afternoons

Seminar Objectives

This seminar will introduce students to several tools and approaches that aid in the development of (ecological) datasets. Four main topics will be introduced, discussed, and practiced throughout the semester: 1) local database design and query, 2) graphical interpretations of datasets, 3) introduction to multivariate approaches for data development, and 4) univariate and multivariate confidence limits of collected data. These topics have been selected because they represent focal areas for graduate-level research across disciplines. Class sessions will be a mixture of short presentations, discussions, hands-on exercises, and group presentation.

Course Materials

Papers and book chapters will be provided, temporary licenses for all required software will be provided. In addition to peer-reviewed publications that will be discussed, chapters from several books and technical reports will be used.

Houk P (2010) Improving local capacity for coral reef monitoring data interpretation. Pacific Marine Resources Institute, Saipan, MP

Clarke KR, Warwick RM (2001) Change in marine communities: an approach to statistical analysis and interpretation. Primer-E Ltd, Plymouth Marine Laboratory

Anderson M, Gorley R, Clarke K (2008) PERMANOVA+ for PRIMER: Guide to software and statistical methods. PRIMER-E, Plymouth, UK

Crawley MJ (2007) The R boook. John Wiley and Sons Ltd, Chichester.

Published papers, book chapters, and technical reports on select topics

Grading

• Attendance and participation (50%)

• Group projects (20%)

• Facilitate discussion on selected topic (30%)

Course Schedule

8/19 Week 1 -   Introduction to seminar course, expectations, and background on data development themes.

Reading/Assignment –Work through chapters 1 - 2 of data analysis guidebook. Prepare additional queries on your own, prepare to facilitate discussion on your query.

8/26 Week 2 -   Discussion on chapters 1 - 2 and advanced pivot table use. Individual presentations of data queries to the class using pivot tables.

Reading/Assignment –Work through chapters 2 - 4 of data analysis guidebook. Prepare additional queries on your own, prepare to facilitate discussion on your query.

9/2 Week 3 -   Discussion on chapters 2 - 4 and advanced pivot table use. Individual presentations of data queries to the class using pivot tables. Hands on examples together if time permits.

Reading/Assignment – Each small group (3-4 people) will be given a new dataset to develop. Group will work together to first describe the components of the database, and draw basic queries and graphical summaries to interpret some main trends.

9/9 Week 4 – Discussions and presentation by groups on datasets.

Reading/Assignment – Complete two exercises on creating publication-quality summary graphs to present the findings of your data analyses. Temporary licenses of Sigmaplot software will be provided, but other programs are fine to use as well. We will use both Sigmaplot and introduce the package ggplot in R. You should download and install both R and R-Studio.

9/16 Week 5 – Show your graphs that were created in Sigmaplot. Discuss benefits and limitations of this software platform. Lecture introducing the basic principles of R software package. We will work together to begin introducing the basic concepts of R, and how to make simple plots using R and R-studio with several packages.

Reading/Assignment – Working in small groups, complete the short exercise introducing the basic concepts of R software package. Then, using materials provided to you in class, create several basic plots using the package ggplot in R. This will require you to download and install both R and R-studio on your computer.

9/23 Week 6 – Continue week 5. Hands-on exercise together in class to get a handle on basic R skills, and graphing.

Reading/Assignment – Continue week 5 work.

9/30 Week 7 – Continue week 5. Hands-on exercise together in class to get a handle on basic R skills, and introduce the concept of building a basic linear statistical model in R (t-test, ANOVA, regression, etc.).

Reading/Assignment – Chapter on linear models in R.

10/7 Week 8 – Building and graphing a linear statistical model using R and R-Studio.

In small groups, use data of your choosing and build and graph a simple statistical model. Validate your model. This requires an understanding of basic statistical tests.

10/14 Week 9 – Building and graphing a linear statistical model using R and R-Studio. Second example on the topic. Hands on exercise together in class, a bit more complicated model.

10/21 Week 10 – Lecture on statistical power and sampling effort.

Reading/Assignment – Two papers on calculating sampling effort for univariate and multivariate needs.

10/28 Week 11 – Discussion of sampling effort papers and approaches for calculations.

Reading/Assignment – Complete exercise on calculating univariate sampling effort.

11/4 Week 12 – Present back to class on independent power analyses regarding your selected datasets.

Reading/Assignment – Read two published studies that have used multivariate statistics.

11/11 Week 13 – No class Veteran's Day

Last group exercise on multivariate data analyses, discuss different sources of variance structures and their meaning (intra- and inter-site variation).

Class presentations of multivariate queries/paper discussion.

11/18 Week 14 – Class exercise on generating power curves for three example studies (with discussion).

Reading/Assignment – Using datasets given to you (or that you have), each group will conduct a power analysis.

11/25 Week 15 – No class Thanksgiving Holiday

12/2 Week 16 – Introductory talk on multivariate statistical approaches, discuss several examples of basic multivariate analyses and use.

Reading/Assignment – Read two published studies that have used multivariate statistics.

12/9 Week 17 – Last hands-on example of principle components analysis and principle components ordination. This includes an introduction to both types of analyses for environmental and ecological data. This aims to serve as a primer for interested students to read more on their own.