MSG500/MVE190, Linear Statistical Models   **LP2 2016**

Preliminary Course Outline

Instructor: Rebecka Jörnsten, `jornsten@chalmers.se`
Class homepage: http://www.math.chalmers.se/Stat/Grundutb/GU/MSG500/A16/
Email empty email with subject header "MSG500" to me first week of classes to be put on the mail-roster (used for class-updates).

Week	Topics	Chapters	Exercises
w44	Introduction, Basic Stats, Linear models Diagnostics and matrix formulation	1:1-7,9. 11, 12:1-4	MiniAnalysis1
w45	Multiple regression. Diagnostics and testing	2, 3, 4, (6), 11, 12, 13.1	Lab1
w46	Dummy variables, ANCOVA. Model selection and testing.	9 + notes	MiniAnalysis2
w47	Model Selection	7, 13 + notes.	Lab2
w48	Bootstrap. Cross-validation.	7 + notes.	MiniAnalysis 3
w49	Regularized regression	13 + notes.	Lab3
w50	Weighted least squares, non-linear model, GLM In-class presentations	8, 10, 12, 15	MiniAnalysis4

Labs and MiniAnalysis tasks make up 10 % of the final grade and are mandatory!
You will present MiniAnalysis tasks in class on Fridays (with the exception of MiniAnalysis 1 which will be presented on Tue Nov 8th).
On Labs and MiniAnalyses you can work in teams of 2 (or on your own if you prefer) but not the same team for more than one task.

The in-class final (January 12th) makes up 60 %.
The individual project and in-class presentation is worth 30 % of the final grade (to be handed in the week of the final exam).

Text: J.O. Rawlings, S.G. Pantula, D.A. Dickey. Applied Regression Analysis (available online).
Software: R. R is free - download for windows, linux and mac available. I will write the labs for R, but the data sets we will work with will be available in tab-del files so you can use another software package if you want. However, I strongly urge you to use R since I will be providing demo codes from the lectures using this language.
Here's a good place to start: https://www.r-bloggers.com/how-to-learn-r-2/

Project - Project Proposal

The project proposal is due December 1. However, the sooner you get a proposal to me, the sooner I can give you feedback on it. Don't wait until the last minute.

The proposal should contain the following;

a) a brief description of the data you intend to analyze.

b) an indication of the source of the data set.

c) the objectives of your investigation.

d) an overview of the analyses you *anticipate* completing.

Note, this is a regression class. Choose a data set that fits with the class syllabus (not time series data). There are data set repositories on the web (see e.g. `http://lib.stat.cmu.edu/`, `http://www.statsci.org/datasets.html`). Some data sets will not be accepted: these include the "body fat data", and "the boston housing data". You are required to provide the data source and projects will be run through URKUND so do not plagiarize projects from other classes or previous years....

Lab reports

Lab reports should be typed! Unless the lab contains an analytical problem, do not hand in handwritten material.

The report should contain the following;

a) Description of the methods used. Be brief - don't repeat what's in the text, just the key elements.

b) Discuss your results. Results without discussion are not graded.

c) Include only the crucial plots and graphs, don't go for quantity.

d) Label all plots and graphs.

e) Conclusions: what is the take-home message.

f) You can discuss programming problems with your fellow classmates, but do all the work yourselves.

For help with computing; Modern applied statistics with Splus (Venables and Ripley) is a good text.

Some online tutorials are available at

https://www.r-bloggers.com/how-to-learn-r-2/

http://www.uga.edu/strata/software/pdf/RTutorial.pdf

http://cran.r-project.org/doc/manuals/R-intro.html or google for more....

I use RStudio - highly recommended!

Linear Statistical Models   LP2 2016

MiniAnalysis

Labs are data analysis tasks where you have to do some coding yourself and write a report. MiniAnalysis is a lab where I provide most of the code (or ask you to adapt demo codes) and the main purpose of the task is for you to apply the methods and discuss and interpret the results.

MiniAnalysis tasks will be presented in class. I will choose teams at random to present their results. It is mandatory to prepare an in-class presentation (5-10 slides to be sent to me ahead of class or brought on a USB stick or available online through dropbox, sharelatex etc). If don't attend class when MiniAnalysis tasks are presented you will be required to write a report (which is a lot more work!).

It is mandatory to present at least once.

a) Prepare 5-10 slides.

b) Discuss your approach and results.

c) Include only the crucial plots and graphs, don't go for quantity.

d) Conclusions: what is the take-home message.

Don't think of MiniAnalysis as an exam or a test. The idea is for you and the rest of the class to focus on results and interpretation and to get a discussion going.