STAT 632: Linear and Logistic Regression

This syllabus (class) may be updated/modified due to the uncertainty created by the ongoing COVID surge.

Instructor: Dr. Joshua Kerr
Office: North Science 243
Email: `joshua.kerr@csueastbay.edu`

Lecture: VBT 137

Office Hours: TuTh 3:00pm – 4:00pm; also via Zoom (`https://csueb.zoom.us/j/97411601270`), or by appointment

Website: Course materials will be posted on Blackboard.

Textbook: Simon Sheather. *A Modern Approach to Regression with R*. Springer, 2009.
Free electronic version: `http://library.csueastbay.edu/home`
Data sets and R code: `http://gattonweb.uky.edu/sheather/book/`

Additional Reference: James, G., Witten, D., Hastie, T., and Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*. Springer, 2013.
Free PDF version: `https://www.statlearning.com/`

Software:
R, can be downloaded here `https://www.r-project.org/`
RStudio, can be downloaded here `https://www.rstudio.com/`
RStudio Cloud, `https://rstudio.cloud/`

Course Topics:

- Simple linear regression
- Multiple linear regression, matrix notation
- Model diagnostics and transformations
- Variable selection
- Logistic regression
- Cross-validation: training and test set
- Regularization: ridge regression and LASSO*

- Decision trees and random forests*

- Other Topics*

*Indicates optional topics

Grading: There will be two exams, each worth 25% of your grade. For the project you will need to find a data set of interest, and then conduct a regression analysis using that data set. For the project you are also encouraged to use a modern method (e.g., random forests, LASSO).

- 50% Two Exams (25% each)
- 20% Homework
- 30% Project

Grading Rubric: In general, I grade on a standard partition of grades

A: [100%, 90%]   B: (90%, 80%]   C: (80%, 70%]   D: (70%, 60%]   F: (60%, 0%)
allowing for "+" and "-" depending on where in the above ranges the grade falls.

Policy on Late Assignments: Late homework will generally not be accepted. However, your lowest scoring homework will be dropped. I may agree to extensions on due dates if you are experiencing an emergency or illness.

Student Learning Outcomes: Upon successful completion of this course, students will be able to:

- Apply statistical methodologies, including (a) simple and multiple linear regression, (b) model diagnostics and transformations, and (c) logistic regression.
- Derive and understand basic theory underlying these methodologies.
- Use R and RStudio to analyze data sets, estimate statistical models, and conduct model diagnostics.
- Understand and apply modern techniques such as cross-validation, regularization, decision trees and random forests.
- Communicate statistical concepts clearly and appropriately to others.

Common Syllabus Items: Items such as policies on academic dishonesty, disability, and handling emergency situations can be found under "University Policies" on Blackboard.

A Note on Discrimination, Harassment, and Retaliation (DHR):
California State University East Bay is committed to a community free from sexual assault and violence. Title IX and CSU policy prohibit discrimination, harassment and retaliation, including Sex Discrimination, Sexual Harassment or Sexual Violence. CSUEB encourages anyone experiencing such behavior to report their concerns immediately. CSUEB has

both confidential and non-confidential resources and reporting options available to you. **As a faculty member, I am required to report all incidents and thus cannot promise confidentiality.** I must provide our Title IX coordinator and or the DHR Administrator with relevant details such as the names of those involved in an incident. For confidential services, contact the **Confidential Advocate at 510-885-3700** or go to the Student Health and Counseling Center. For 24-hour crisis services call the Bay Area Women Against Rape (BAWAR) hotline at 510-845-7273. For more information about policies and resources or reporting options, please visit the following websites: `https: //www.csueastbay.edu/diversity/title-ix/`