

Course Syllabus: Applied Statistics and Data Analysis - STAT 210

Offering Department	Statistics
Course Number	STAT 210
Course Title	Applied Statistics and Data Analysis
Academic Semester	Fall 2022/2023
Semester Start Date	08/28/2022
Semester End Date	12/13/2022
Class Schedule (Days & Time)	<p>Applied Stat.& Data Analysis - Section A</p> <p>Lecture STAT 210 Tue 09:30 - 11:00</p> <p>Applied Stat.& Data Analysis - Section B</p> <p>Lecture STAT 210 Mon 13:00 - 14:30</p> <p>Applied Stat.& Data Analysis - Section C</p> <p>Lecture STAT 210 Sun 08:00 - 09:30</p>

Instructor(s)				
Name	Email	Phone	Office Location	Office Hours
Joaquin Ortega Sanchez	JOAQUIN.ORTEGAS ANCHEZ@KAUST.EDU.SA		4297, 1, Al-Khawarizmi (bldg. 1)	Mon - Tue 3:00 - 4:00 pm

Teaching Assistant(s)	
Name	Email
TBA	TBA

Course Information

Course Description	Prerequisites: Advanced and multivariate calculus. For students outside STAT wishing to obtain an introduction to statistical method. No degree credits for STAT or AMCS majors. Provides fundamentals of probability and statistics for data analysis in research. Topics include data collection, exploratory data analysis, random variables, common discrete and continuous distributions, sampling distributions, estimation, confidence intervals, hypothesis tests, linear regression, analysis of variance, two-way tables and data analysis using statistical software.
Learning Outcomes	By the end of this course, the students are expected to have mastered the following: (1.) Statistical visualization methods (2.) Framework for statistical modeling of continuous and discrete-valued types of data (3.) Formal inferential procedures (4.) Various data analytic techniques
Textbook/Materials	Crawley, Michael J. (2015) Statistics, An Introduction Using R, Second Edition. J. Wiley & Sons. Weisberg, S. (2013). Applied Linear Regression. Wiley, 4th edition. ISBN: 978-1-118-38608-8. Additional references Dalggaard, P. (2008) Introductory Statistics with R. Springer. John Maindonald, W. John Braun (2010). Data Analysis and Graphics Using R - an Example-Based Approach, Third Edition, Cambridge University Press. Shahbaba, B. (2012). Biostatistics with R: An Introduction to Statistics Through Biological Data. Springer.
Method of Assessments	30.00% - Homework /Assignments 30.00% - Group Project(s) 20.00% - Exam 1 20.00% - Exam 2
Nature of the Assignments	Homework will complement the work in class. Solutions should be uploaded using the blackboard application on the due date. There will be two in-class exams; time will be outside class hours to allow for more time. Each student is required to work on a project in a group of three or four students and submit a final report. More information on the final project will be given in a separate document that will be posted on the Blackboard page for the course. Project grading will include the proposal, draft, final paper, and oral presentation.
Course Policies	Attendance is required. Assignments will be posted on the blackboard application. Solutions should be uploaded using the blackboard application on the due date. Late assignment submissions will not be accepted unless prior arrangements have been made (except in university-established cases of illness or emergency). All homework assignments must be neatly typed (LaTeX or RMarkdown are recommended). All projects and homework assignments are required. If you dispute your grade on any homework or project, you may request a regrade (from the TA/instructor for the assignments or the instructor for the exams) only within 48 hours of receiving the graded exam. An Incomplete (I) grade for the course will only be given under extraordinary circumstances such as sickness, and these exceptional circumstances must be verifiable. The assignment of an (I) requires first approval of the Dean and then a written agreement between the instructor and student specifying the time and manner in which the student will complete the course requirements. Collaboration and checking answers on assignments are allowed and encouraged. Of course, copying assignments is not tolerated. In brief, you are allowed to collaborate on all homework problems according to the following rules: you must first attempt to solve each problem on your own. If you get stuck, you can then talk to any student currently enrolled in the class about the issue, as well as the instructor or the TA. However, solutions and R code should not be exchanged (i.e., you still must work through the details of the problem after you have gotten help, write the final answers alone, and understand them thoroughly). The use of cellular phones in the classroom is not allowed.

Additional Information	<p>The course will be taught as a flipped classroom. There will be no formal lectures. Lectures have been pre-recorded and are available through the Blackboard page for the course. Students are required to watch these videos as scheduled and work on the problems that will be considered during the weekly sessions. These problems are different from the homework and will be posted on the Blackboard page.</p> <p>The class is divided into three sections with different schedules. Students should register for one section and will only attend one session per week. These meetings are not formal lectures. The instructor will answer doubts and questions and will focus on problems related to the topics covered in the videos. Students are expected to have solved all the weekly problems before the session and participate in the discussions. Students should be aware that this modality requires more individual work from the student than the traditional model of lectures. Changing groups is not allowed unless the instructor approves the change.</p> <p>The materials used in this course are copyrighted. By materials, I mean all materials generated for this class including videos, exams, course notes, computer code, and examples. You do not have the right to copy the material or distribute them, unless I expressly grant permission.</p> <p>This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular I reserve the right to change due dates or the methods of assessment. Official announcements will ALWAYS be those made in class.</p> <p>*Required Knowledge</p> <p>Undergraduate mathematics including calculus, multivariate calculus, and basic matrix algebra.</p>
-------------------------------	---

Tentative Course Schedule (Time, topic/emphasis & resources)		
Week	Lectures	Topic
1	Sun 08/28/2022 Mon 08/29/2022 Tue 08/30/2022	Introduction to data analysis. Introduction to R computing. Using R.
2	Sun 09/04/2022 Mon 09/05/2022 Tue 09/06/2022	Basic functions. Objects and data. Useful functions.
3	Sun 09/11/2022 Mon 09/12/2022 Tue 09/13/2022	Graphs in R. Quantile plots.
4	Sun 09/18/2022 Mon 09/19/2022 Tue 09/20/2022	One sample problems.
5	Sun 09/25/2022 Mon 09/26/2022 Tue 09/27/2022	Hypothesis tests. Comparing two populations, two-sample t-test, paired t-test. Analysis of contingency tables.
6	Sun 10/02/2022 Mon 10/03/2022 Tue 10/04/2022	Comparing population proportions. Comparing many populations (ANOVA). F-test. Multiple testing. Introduction to experimental design.
7	Sun 10/09/2022 Mon 10/10/2022 Tue 10/11/2022	Introduction to experimental design. Simple linear regression. Model formulation. Formal inference. Exam 1 to be given this week (outside of class hours)
8	Sun 10/16/2022 Mon 10/17/2022 Tue 10/18/2022	Mid-semester break. Simple linear regression

Week	Lectures	Topic
9	Sun 10/23/2022 Mon 10/24/2022 Tue 10/25/2022	Multiple linear regression. Estimations. Model diagnostics.
10	Sun 10/30/2022 Mon 10/31/2022 Tue 11/01/2022	Multiple linear regression. Variable selection. Data analysis and interpretation.
11	Sun 11/06/2022 Mon 11/07/2022 Tue 11/08/2022	Nonparametric function estimation. Kernel smoothing. Nonparametric regression.
12	Sun 11/13/2022 Mon 11/14/2022 Tue 11/15/2022	Logistic regression. Application to classification.
13	Sun 11/20/2022 Mon 11/21/2022 Tue 11/22/2022	Regression models for count data.
14	Sun 11/27/2022 Mon 11/28/2022 Tue 11/29/2022	Generalized linear models (GLM). Power analysis and sample size calculations for GLM. Exam 2 to be given this week (outside of class hours)
15	Sun 12/04/2022 Mon 12/05/2022 Tue 12/06/2022	Resampling methods and bootstrap.
16	Sun 12/11/2022 Mon 12/12/2022 Tue 12/13/2022	Oral presentations

Note

The instructor reserves the right to make changes to this syllabus as necessary.