STAT 430

Introduction to Statistical Computing with SAS

MTuWThF 9:30 – 10:50, EGR 3114

Instructor: Dr. Alexey B. Stepanov
E-mail: alex314@umd.edu
Office: MTH 4310

Course Description: This course is an introduction to statistical and graphical techniques of data analysis and their implementation in the SAS programming language/platform. The emphasis is on data analysis skills, but since one such important skill is justification of assumptions and understanding of the rationale behind analyses, the course develops ideas and explains concepts from statistical theory.

Prerequisite: STAT 400. The material needed is mostly definitions and concepts. But later in the course, some understanding of STAT 400 material will help you make sense of statistical simulation methods. (That material will be reviewed as needed.)

Textbook: Required: *Applied Statistics and the SAS Programming Language*, 5th ed., by R. Cody and J. Smith.

Recommended: *Probability and Statistics for Engineering and the Sciences, 8th ed.*, J. Devore

Chesapeake Project. As part of the Chesapeake Project, this course aims to bring the concepts of sustainability into the classroom. Such important sustainability ideas as Systems Thinking, Tragedy of the Commons and others will be explored on the basis of statistical analyses of real-world data sets. For more information on the Chesapeake Project and other sustainability initiatives, please visit `www.sustainability.umd.edu`.

Grading: The course grade will be based on a weighted average with

Homework	30%
Midterm Exam	20%
Final Exam	30%
Final Project	20%

Grades of A-, B-, C-, D- are guaranteed with averages exceeding 90%, 80%, 70%, 60%, respectively.

Homework: There will be 5-6 graded problem sets, **due on the date assigned**. These will involve writing and running small SAS programs and interpreting the sequence of data-analysis operations and outputs. While you can discuss homework problems in general, it is essential that the edited hard copies you hand in are your own work. **No homework grades will be dropped. Late homework will be accepted only in the case of a medical emergency, etc.** Only a random subset of problems will be graded.

Final Project: There will be a slightly more ambitious data-analysis term project (due Thursday, August 18, subject to change). You will be expected to produce a written report discussing some interesting questions about the data. The data and detailed project guidelines will be given.

Midterm Exam: There will be an in-class test (tentatively scheduled for Friday, 29 July), on basics of the SAS language and concepts underlying data-display and statistics in categorical data, two-sample comparisons, and simple linear regression (subject to change). You can bring a double-sided notebook sheet to the test as a memory aid. Except for your notebook sheet, the test is closed-book. You can use a calculator, but I will not ask for much arithmetic.

Final Exam: There will be an in-class final exam on Friday, August 19, 9:30-10:50. The Final Exam is closed-notes, closed-book.

Course Coverage:

Chapter 1: Basics of SAS syntax, all sections, with some reference to material in Chapters 12, 13.

Chapter 2: Data display and graphics, all sections.

Chapter 3: Categorical data, esp. sections A, C-E, G, I-L, N-Q

Chapter 5: all

Chapter 6: all

Chapter 7: sections A-E

Chapter 9: sections A-E, G-H