EST 610 – Data Analysis for Technology, Policy and Innovation
Thursdays 6-8:50pm

Important Note: All times and deadlines in the syllabus are based on UTC -4 (time in New York). Every effort will be made to avoid changing the course schedule, but the possibility exists that unforeseen events will make syllabus changes necessary. It is your responsibility to check Blackboard for corrections or updates to the syllabus. Any changes will be clearly noted in course announcements or through Stony Brook email.

Contents
Instructor Information ................................................................. 2
Course Description: ................................................................. 2
Textbook: .............................................................................. 2
Helpful Books: ....................................................................... 2
Course Delivery Mode and Structure: ........................................ 2
Learning objectives: .................................................................. 4
How to Succeed in this Course: ................................................ 4
Course Schedule ....................................................................... 4
Grading, Attendance, and Late Work Policies............................. 6
  Letter Grades: ........................................................................ 7
  Attendance Policy: ............................................................... 7
  Late Work Policy: ................................................................. 7
Course and University Policies ....................................................... 7
  Student Accessibility Support Center Statement: ....................... 7
  Academic Integrity Statement: .............................................. 7
Critical Incident Management: .................................................... 8
    Understand When You May Drop This Course: ....................... 8
Incomplete Policy: .................................................................... 8
Course Materials and Copyright Statement: ............................... 8
    Online Communication Guidelines and Learning Resources: .... 8
Instructor Information
INSTRUCTOR: Dr. Thomas S. Woodson
E-MAIL: thomas.woodson@stonybrook.edu
OFFICE PHONE: (631) 632-9974
OFFICE HOURS: By appointment

Course Description:
Course title: Data Analysis for Technology, Policy and Innovation
Course catalog # and section: EST 610
Credit hours: 3
Prerequisites: Must be a PhD in Technology and Society or receive permission from the instructor

Welcome. I am excited to teach EST 610. I have taught this course over the past 5 years and each year it is an enriching experience for both me and the students. Data is all around us. How do you sort through the vast amount of information to make valid conclusions and inferences? This is an introductory data analysis course that will teach you things like how to describe your data, types of data, linear regression, and principal component analysis. You will get practical data analysis experience using a variety of data sets ranging from stock prices to health statistics. This course uses R, an open-source statistical software that is increasingly becoming the most used data analysis software. At the end of the course, you will be equipped fully equipped to analyze a variety of data using the most modern software tools.

Textbook:
- *The Art of Statistics: How to Learn from Data* by David Spiegelhalter

Helpful Books:
- An Introduction to Applied Multivariate Analysis with R by Brian Everitt and Torsten Hothorn
- Big Data Analytics with R, by Simon Walkowiak

Course Delivery Mode and Structure:
This is a hybrid online course, delivered both in the Blackboard learning management system (LMS) and in face-to-face meetings. Students must be mindful of all course expectations, deliverables and due dates, especially because the online portion of the course requires
significant time management. All assignments and course interactions will utilize internet technologies. See “Technical Requirements” section for more information. In Blackboard, you will access online lessons, course materials, and resources. EACH WEEK HAS 1 major LEARNING MODULES; with deadlines most days Thursday. Some variations will occur.

How We Will Communicate:
Course-related questions should be posted in the General Questions Forum in the course Discussion board. For personal/private issues, email me directly. If you use Blackboard’s email tool from the course site, it will automatically include your full name, course name and section when you send me an email. Please allow between 24–48 hours for an email reply. Your Stony Brook University email must be used for all University-related communications. You must have an active Stony Brook University email account and access to the Internet. All instructor correspondence will be sent to your SBU email account. Plan on checking your SBU email account regularly for course-related messages. To log in to Stony Brook Google Mail, go to http://www.stonybrook.edu/mycloud and sign in with your NetID and password.

Regular announcements will be sent from Blackboard. These will be posted in the course site and may or may not be sent by email.

Regular communication is essential in online classes. Logging in once a day, checking the discussion board and participating with your colleagues ensures that you are able to remain an active member of the class and earn full points for participation.

Technical Requirements:
This course uses Blackboard for the facilitation of communications between faculty and students, submission of assignments, and posting of grades and feedback. The Blackboard course site can be accessed at https://blackboard.stonybrook.edu

If you are unsure of your NetID, visit https://it.stonybrook.edu/help/kb/finding-your-netid-and-password for more information. You are responsible for having a reliable computer and Internet connection throughout the term. Caution! You will be at a disadvantage if you attempt to complete all coursework on a smart phone or tablet. It may not be possible to submit the files required for your homework assignments.

Students should be able to use email, a word processor, spreadsheet program, and presentation software to complete this course successfully.

The following list details a minimum recommended computer set-up and the software packages you will need to have access to, and be able to use:

- PC with Windows 8 or higher (we recommend a 3-year Warranty)
- Macintosh with OS 10.11 or higher (we recommend a 3-year Warranty)
- Intel Core i5 or higher
- 250 GB Hard Drive
- 8 GB RAM
• Latest version of Chrome, Firefox or Explorer; Mac users may use Chrome or Firefox. (A complete list of supported browsers and operating systems can be found on the My Institution page when you log in to Blackboard.)
• High speed internet connection
• Printer
• Word processing software (Microsoft Word, Pages, etc.)
• Speakers (either internal or external) or headphones
• Ability to download and install free software applications and plug-ins (note: you must have administrator access to install applications and plug-ins).
• Adobe Flash player with the latest update is crucial for playing multiple videos throughout the course

Technical Assistance:
If you need technical assistance at any time during the course or to report a problem with Blackboard you can:
• Phone: 631-632-9800 (client support, Wi-Fi, software and hardware)
• Submit a help request ticket: https://it.stonybrook.edu/services/itsm
• If you are on campus, visit the Walk-Up Tech Support Station in the Educational Communications Center (ECC) building.

Learning objectives:
Upon completion of the course, you will be able to:
1. Use R to analyze data
2. Do linear regression and logistic regression analyses
3. Cluster data with techniques like k-means and hierarchal clustering
4. Conduct anova and t-test
5. Know how to interpret the results of statistical analyses

How to Succeed in this Course:
• Complete all assigned readings in the course
• Practice R everyday
• Write your own code instead of just “copy and pasting” what I write.
• Ask for help if you have errors.
• Expect to spend 10-15 hours completing each homework assignment.

Course Schedule

<table>
<thead>
<tr>
<th>Course calendar subject to change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Week</td>
</tr>
<tr>
<td>------</td>
</tr>
</tbody>
</table>
| 2    | 3-Sep| Review Linear Algebra                                               | JW2      | Homework #1 due.  
Read sections 2.5 and 2.6. |
| 3    | 10-Sep| Distance measures and trees                                        |          | Homework #3 due.  
Read 12.1 to 12.4. |
| 4    | 17-Sep| How do you decrease the number of variables to analyze?            | JW 8     | Homework #2 due.  
Read 8.1 to 8.5. |
| 5    | 24-Sep| How do you test for normality and prepare the data for analysis?    | JW4      | Homework #4 due.  
Read 4.1-4.8 (pay close attention to 4.4-4.8)  
Don't worry about proofs. Understand main concepts |
| 6    | 1-Oct| Review/Midterm                                                      |          | Homework #5 due.  |
| 7    | 8-Oct| What is the difference between simultaneous confidence interval and Bonferonni interval? | JW5      |              
5.1-5.2, 5.4-5.5 |
| 8    | 15-Oct| How do you do an Anova?                                            | JW 6     | Homework #6 due.  
Read 6.1-6.7, 6.10 |
| 9    | 22-Oct| Ordinary least squares regression                                  | JW 7     | Project outline due  
Homework #7 due.  |
| 10   | 29-Oct| Ordinary Least Squares Regression II                               | 7.1-7.4,7.5-7.6  | Homework #8 due.  |
Grading, Attendance, and Late Work Policies

Assessment and Grading:
The main assignments for this class are homework assignment, tests and a project.

Viewing Grades on Blackboard: Points and feedback for graded activities will be posted to the My Grades tab in the Tools area of Blackboard.

In this course, you will be assessed on the following:

<table>
<thead>
<tr>
<th>Activity/Assignment</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments (drop the two lowest)</td>
<td>30</td>
</tr>
<tr>
<td>Data Analysis Paper</td>
<td>20</td>
</tr>
<tr>
<td>Exam 1</td>
<td>20</td>
</tr>
<tr>
<td>Exam 2</td>
<td>20</td>
</tr>
<tr>
<td>Participation/other assignments</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>
Letter Grades:
Final grades assigned for this course will be based on the percentage of total points earned and are assigned as follows:

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Points</th>
<th>GPA/Points</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>94-100</td>
<td>4.0</td>
<td>Excellent work</td>
</tr>
<tr>
<td>A-</td>
<td>90-93</td>
<td>3.7</td>
<td>Nearly excellent work</td>
</tr>
<tr>
<td>B+</td>
<td>87-89</td>
<td>3.3</td>
<td>Very good work</td>
</tr>
<tr>
<td>B</td>
<td>83-86</td>
<td>3.0</td>
<td>Good work</td>
</tr>
<tr>
<td>B-</td>
<td>80-82</td>
<td>2.7</td>
<td>Mostly good work</td>
</tr>
<tr>
<td>C+</td>
<td>77-79</td>
<td>2.3</td>
<td>Mediocre work</td>
</tr>
<tr>
<td>C</td>
<td>73-76</td>
<td>2.0</td>
<td>Mediocre/poor work</td>
</tr>
<tr>
<td>C-</td>
<td>70-72</td>
<td>1.7</td>
<td>Poor work</td>
</tr>
<tr>
<td>D+</td>
<td>65-69</td>
<td>1.3</td>
<td>Very poor work</td>
</tr>
<tr>
<td>D</td>
<td>60-34</td>
<td>1.0</td>
<td>Very poor work</td>
</tr>
<tr>
<td>F</td>
<td>Less than 60</td>
<td>0.0</td>
<td>Failing work</td>
</tr>
</tbody>
</table>

Attendance Policy:
Attendance to this class is mandatory. Having more than 1 unexcused absence will impact your grade. If you miss a class, you are still expected to do all the readings and assignments for that week. Be on time to class. If you are often late to class, you will lose class participation points.

Late Work Policy:
Turn in your work by the due date in order to receive full credit. After the deadline you have the option to turn in work, but you will lose points. Late work will only be accepted with full credit given university sanctioned absences (i.e. Documented illness, or documented family emergencies)

Course and University Policies

Student Accessibility Support Center Statement:
If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Student Accessibility Support Center, ECC (Educational Communications Center) Building, Room 128, 631-632-6748. They will determine with you what accommodations, if any, are necessary and appropriate. All information and documentation is confidential. Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and Student Accessibility Support Center. For procedures and information go to the following website:
http://www.stonybrook.edu/ehs/fire/disabilities

Academic Integrity Statement:
Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person’s work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-
specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty please refer to the academic judiciary website at http://www.stonybrook.edu/commcms/academic_integrity/index.html.

**Important Note:** Any form of academic dishonesty, including cheating and plagiarism, will be reported to the Academic Judiciary.

**Critical Incident Management:**
Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of University Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students’ ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.

**Understand When You May Drop This Course:**
It is the student’s responsibility to understand when they need to consider withdrawing from a course. Refer to the Stony Brook Academic Schedule for dates and deadlines for registration:

**Incomplete Policy:**
Under emergency/special circumstances, students may petition for an incomplete grade. Circumstances must be documented and significant enough to merit an incomplete. If you need to request an incomplete for this course, contact me for approval as far in advance as possible.

**Course Materials and Copyright Statement:**
Course material accessed from Blackboard, SB Connect, SB Capture or a Stony Brook Course website is for the exclusive use of students who are currently enrolled in the course. Content from these systems cannot be reused or distributed without written permission of the instructor and/or the copyright holder. Duplication of materials protected by copyright, without permission of the copyright holder is a violation of the Federal copyright law, as well as a violation of Stony Brook’s Academic Integrity.

**Online Communication Guidelines and Learning Resources:**
Maintain Professional Conduct Both in the Classroom and Online: The classroom is a professional environment where academic debate and learning take place. I will make every effort to make this environment safe for you to share your opinions, ideas, and beliefs. In return, you are expected to respect the opinions, ideas, and beliefs of other students—both in the face-to-face classroom and online communication. Students have the right and privilege to learn in the class, free from harassment and disruption. The course follows the standards set in the Student Code of Conduct, and students are subject to disciplinary action for violation of that code. If your behavior does not follow the course etiquette standards stated below, the grade you receive for a posting may suffer. I reserve the right to remove any discussion messages that display inappropriate language or content.

**Online Etiquette:**
- Offensive language or rudeness will not be tolerated. Discuss ideas, not the person.
• Avoid cluttering your messages with excessive emphasis (stars, arrows, exclamations).
• If you are responding to a message, include the relevant part of the original message in your reply, or refer to the original post to avoid confusion.
• Be specific and clear, especially when asking questions.
• Use standard punctuation and capitalization. Using all UPPERCASE characters gives the appearance of shouting and makes the message less legible;
• Remember that not all readers have English as their native language, so make allowances for possible misunderstandings and unintended discourtesies.

**Online Classes Require Better Communication:**
It is important to remember that we will not have the non-verbal cues that occur in a face-to-face classroom. I cannot see the confused, frustrated, or unhappy expressions on your face if you encounter problems. You MUST communicate with me so that I can help. To make the experience go smoothly, remember that you’re responsible for initiating more contact, and being direct, persistent, and vocal when you don’t understand something.

**My Role as the Instructor:**
As the instructor, I will serve as a “guide” in our online classroom. While I will not respond to every post, I will read what is posted, and reply when necessary. Expect instructor posts in the following situations:
• To assist each of you when it comes to making connections between discussion, lectures, and textbook material.
• To fill in important things that may have been missed.
• To re-direct discussion when it gets “out of hand.”
• To point out key points or to identify valuable posts.

**Student Resources:**
• Create data visualization: [http://www.creativebloq.com/design-tools/data-visualization-712402](http://www.creativebloq.com/design-tools/data-visualization-712402)
• DASL (The Data and Story Library): [http://lib.stat.cmu.edu/DASL](http://lib.stat.cmu.edu/DASL)
• Statlib-Datasets Archive [http://lib.stat.cmu.edu/datasets/](http://lib.stat.cmu.edu/datasets/)
• University of California, Los Angeles Case Studies [http://www.stat.ucla.edu/cases/](http://www.stat.ucla.edu/cases/)
• U.S. Census Bureau
  http://www.census.gov
• Stats in the news, from George Mason University:
  http://www.stats.org/

**Online statistics textbooks and software:**
• Computing for Data Analysis
  https://www.coursera.org/course/compdata
• Data Analysis
  https://www.coursera.org/course/dataanalysis
• Rice virtual lab in statistics
  http://onlinestatbook.com/rvls.html
• SISA simple interactive statistical analysis
  http://www.quantitativeskills.com/sisa/

**Online resources for R:**
• The main R project site:
  www.r-project.org
• An R online textbook
  Kickstarting R: http://cran.r-project.org/doc/contrib/Lemon-kickstart/
• Website for the Sarkar book:
  http://lmdvr.r-forge.r-project.org/figures/figures.html
• Quick R website (many helpful “how to” pages)
  http://www.statmethods.net/
• A U. of Wisconsin Stats professor’s site (B. Yandell)
  http://www.stat.wisc.edu/~yandell/software/
• Book on Regression and Anova on the R site:
  http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf