MIS 620: E-Business Infrastructure: Data Science and Big Data  
Spring 2016 Syllabus

Time: Thursday 4:00 PM – 6:40 PM, Education and Business Administration, Rm 443  
Instructor: Dr. Aaron C. Elkins  
Email: aelkins@mail.sdsu.edu  
Office: Student Services East (SSE), 3367  
Office Hours: Thurs 1:00 PM - 3:00 PM or by appointment

Course Description
In this course students will learn how to manage, analyze, and mine big data. Big data analysis requires the  
management of immense amount of data quickly, often hundreds of terabytes or larger in size. The demanding  
nature of big data requires new technologies and tools to distribute data storage and analysis across hundreds or  
thousands of separate compute nodes.

Through real-world examples with R, data science/machine learning, big data technologies (e.g., Hadoop), and  
Azure cloud resources, students in this course will learn how to identify and solve big data challenges faced by  
modern organizations.

Course Objectives
Upon completing this course, students will be able to use R, an open-source analytics platform, to conduct  
statistical and predictive (machine learning) analysis on data. Students will also understand the trade-offs of  
alternative technologies to determine the appropriate solution for their big data problems. In addition to  
analysis, students will also be able to design visualizations of their data and communicate results to  
management.

Program Learning Goals
MSIS students will graduate being able to:
  • Design and use technology-supported solutions to improve decision making and create value
  • Create value through the development of data, information or knowledge (DIK) strategies and the  
    management of processes and projects
  • Demonstrate business professional skills

Course Materials

Required Textbook
Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, 1st edition,  

Required Technology
This course will require a computer/laptop with the open-source R statistics software, which is available for  
Windows, Mac, and Linux, (https://www.r-project.org/) and a web browser for accessing and managing cloud  
services and virtual servers.

Paper and Writing Implements
Please come to class with paper and pens/pencils to facilitate in class activities and group discussions.

Last Updated 1/20/2016
Enrollment Information

Prerequisites
This course will be introducing data science and machine learning concepts that assumes students already possess knowledge from previous courses such as MIS 380 Data Management Systems or equivalent (Relational Database Management and SQL) AND MIS 301 Statistical Analysis for Business or equivalent

Adding and Dropping (Deadline to Add and Drop is February 2\textsuperscript{nd})
Add codes have been replaced with the Wait List. Students can enroll in any open seats through the fifth day of classes. For additional details visit: http://arweb.sdsu.edu/es/waitlist/students/index.html

You are responsible for dropping or adding classes by February 2\textsuperscript{nd} through your WebPortal account. The instructor may drop a student from the course for nonattendance or lack of required prerequisites.

Course Structure and Conduct
Class meetings will be primarily lecture and discussion based.

Readings and Assignments
All students are responsible for completing all reading and homework assignments prior to coming to class. The concepts and skills needed to learn data science and big data analysis using tools such as R, require practice and hands-on experience with the data outside of the classroom.

Weekly Practice Exam
Each class will start with a practice exam to assess your understanding of the week’s reading and/or assignment and will also serve as a study aid for midterm and finals exams.

Group Project
Students will design and complete a group project using concepts and skills learned from the course. The groups will be 3-4 students and will involve analyzing and visualizing data, and presenting the results at the end of the semester.

Computer Account and Blackboard
MIS 620 course Announcements and emails and will be distributed via BlackBoard (BB). I will use Blackboard for course announcements, facilitate discussions, to post your class Power Point (PP) lecture notes, for course updates, class-related urgent messages, and to post helpful hints about assignments and exams. You are responsible for checking your BB course site frequently. Please have your BB email forwarded to your primary email account. Check your MIS 620 email daily.

Students with Disabilities
If you are a student with a disability and believe you will need accommodations for this class, it is your responsibility to contact Student Disability Services at (619) 594-6473. To avoid any delay in the receipt of your accommodations, you should contact Student Disability Services as soon as possible. Please note that accommodations are not retroactive, and that accommodations based upon disability cannot be provided until you have presented your instructor with an accommodation letter from Student Disability Services. Your cooperation is appreciated.
**Academic Honesty**

The University adheres to a strict policy regarding cheating and plagiarism. These activities will not be tolerated in this class. Become familiar with the policy ([http://www.sa.sdsu.edu/srr/conduct1.html](http://www.sa.sdsu.edu/srr/conduct1.html)). Any cheating or plagiarism will result in a failing of the assignment or class and a disciplinary review by Student Affairs.

Examples of Plagiarism include but are not limited to:

- Using sources verbatim or paraphrasing without giving proper attribution (this can include phrases, sentences, paragraphs and/or pages of work)
- Copying and pasting work from an online or offline source directly and calling it your own
- Using information you find from an online or offline source without giving the author credit
- Replacing words or phrases from another source and inserting your own words or phrases
- Submitting a piece of work you did for one class to another class
- Copying results or answers from another students assignments.

If you have questions on what is plagiarism, please consult the policy¹ and the online-tutorial² to learn more.

**Turnitin**

Students agree that by taking this course all required papers may be subject to submission for textual similarity review to Turnitin.com for the detection of plagiarism. All submitted papers will be included as source documents in the Turnitin.com reference database solely for the purpose of detecting plagiarism of such papers. You may submit your papers in such a way that no identifying information about you is included. Another option is that you may request, in writing, that your papers not be submitted to Turnitin.com. However, if you choose this option you will be required to provide documentation to substantiate that the papers are your original work and do not include any plagiarized material.

**Assessments and Grading**

Course grades will be assigned in accordance with San Diego State University policy (see Graduate Bulletin, pp. 62-64). Graduate grades shall be: A (outstanding achievement, available for the highest accomplishment), B (average, awarded for satisfactory performance), C (minimally passing), D (unacceptable for graduate credit, course must be repeated), F (failing).

**Table 1. Your course grade will be based on the following weighted components**

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Assignments</td>
<td>25%</td>
</tr>
<tr>
<td>Group Project</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm Exam</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>25%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

**Grade and Course Point Appeals**

If you are dissatisfied with a grade, you are welcome to discuss it with me during office hours. However, we will review the exam questions in class when I return the exams to you.

---

¹ [https://studentaffairs.sdsu.edu/srr/conduct1.html](https://studentaffairs.sdsu.edu/srr/conduct1.html)
If you wish to appeal a grade or a potential mistake about course points (e.g. attendance, participation etc.) or a missing lab assignment you must follow the following procedure:

a) **Type** your appeal argument neatly on paper, clearly stating if the discrepancy in grades points relates to attendance or participation or lab assignments, etc.

b) Submit your argument to me, **within two (2) weeks** of the possible grade or point discrepancy.

**Mid-Term and Final Exams**

All Midterm and Final exams will be CLOSED BOOK. Additionally, no electronic devices of any kind may be used (including cell phones, notebooks, laptops. or hand held PCs, or PDAs) during any exams.

**Missed Exams**

If you have to miss an exam due to an emergency or other important matter contact me via email and discuss the problem with me before the scheduled exam AND you provide written documentation to support that such an emergency occurred.

Additionally, the make-up exam must be taken **within 2-3 days** after the regular scheduled exam time. You will receive a different exam of similar difficulty.

**Assignments**

You may turn in your homework either in class or to BlackBoard. Assignments will be posted at the end of class and will be due one or before the next class meeting.

**Participation / Attendance**

Your participation in our MIS 620 class discussions constitutes an important learning component of your education. Your student participation grade will depend primarily on your in-class participation and contribution to the class discussion. Quality of participation, followed by quantity of participation is the expectation. Participation will also depend on outside discussions and threads created and used on BlackBoard.

**Grade of Incomplete**

A grade of Incomplete (I) indicates that a portion of required coursework has not been completed and evaluated in the prescribed time period due to unforeseen, but fully justified, reasons and that there is still a possibility of earning credit. It is your responsibility to bring pertinent information to the instructor and to reach agreement on the means by which the remaining course requirements will be satisfied. The conditions for removal of the Incomplete shall be reduced to writing by the instructor and given to you with a copy placed on file with the department chair until the Incomplete is removed or the time limit for removal has passed. A final grade is assigned when the work agreed upon has been completed and evaluated. An Incomplete shall not be assigned when the only way you could make up the work would be to attend a major portion of the class when it is next offered. Contract forms for Incomplete grades are available at the Office of the Registrar website

# Tentative Course Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: January</td>
<td>Course and R Introduction</td>
</tr>
<tr>
<td>2:</td>
<td>Introduction to Big Data Analytics</td>
</tr>
<tr>
<td>3:</td>
<td>Data Analytics Lifecycle</td>
</tr>
<tr>
<td>4:</td>
<td>Basic Data Analysis</td>
</tr>
<tr>
<td>5:</td>
<td>Clustering (Unsupervised Learning)</td>
</tr>
<tr>
<td>6:</td>
<td>Cloud Computing</td>
</tr>
<tr>
<td>7:</td>
<td>Classification Algorithms (Supervised)</td>
</tr>
<tr>
<td>8:</td>
<td>Classification Algorithms (Continued) and MIDTERM Review</td>
</tr>
<tr>
<td>9:</td>
<td>MIDTERM</td>
</tr>
<tr>
<td>10:</td>
<td>Text Analysis</td>
</tr>
<tr>
<td>11:</td>
<td>Distributed Files Systems and Algorithms (Big Data)</td>
</tr>
<tr>
<td>12:</td>
<td>Big Data Topics Continued &amp; Visualizations</td>
</tr>
<tr>
<td>13:</td>
<td>Communicating Results and Data Privacy/Ethics</td>
</tr>
<tr>
<td>14:</td>
<td>Project Presentations</td>
</tr>
<tr>
<td>15:</td>
<td>Project Presentations and FINAL Review</td>
</tr>
<tr>
<td>16: May</td>
<td>FINAL EXAM</td>
</tr>
</tbody>
</table>

Changes to the course schedule, if any, will be announced in class and posted on BlackBoard.