1. **Credits**: 3
   a. Class Meets: **Tuesday and Thursday 1:30pm - 2:45pm, Old Main (MAIN) 406**

2. **Instructor**: Kyumin Lee, (435) 797-8420, kyumin.lee@usu.edu
   Office Hours: **2:45-3:45pm T/Th** at MAIN 401D, or by appointment

   **TA**: Santosh Kallala, santoshreddykallala@gmail.com
   Office Hours: **9-10am M/F** at MAIN 422

3. **Textbooks**: Course readings will be drawn from the following textbooks:
   a. Data Mining: Concepts and Techniques. Jiawei Han and Micheline Kamber. Morgan Kaufmann.
   b. Introduction to Data Mining. Pang-Ning Tan, Michael Steinbach and Vipin Kumar. Addison-Wesley.

4. **Specific Course Information**:
   a. Course Description: Introduce the theoretical foundations, algorithms, and methods of deriving valuable insights from data. Students will learn how to manage and analyze data at scale (e.g., big data). Specifically, the students will study big data management and processing techniques, data analytics, statistical methods and models, data visualization, and etc. Project required.
   b. Prerequisites: CS 2420 for undergraduate students or Graduate classification.

5. **Specific goals for the course**
   a. Course Objectives
      By the end of the semester you will be able to:
      i. Define and explain the key concepts and models relevant to data science.
      ii. Design, implement, and evaluate the core algorithms underlying an end-to-end data science workflow, including the experimental design, data collection, mining, analysis, and presentation of information derived from large datasets.
      iii. Apply "best practices" in data science, including facility with modern tools (e.g., Hadoop).
Mapped objectives in IDEA:
   i. Learning fundamental principles, generalizations, or theories
   ii. Learning to apply course material (to improve thinking, problem solving, and decisions)
   iii. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course

b. Student Outcomes:
   i. An ability to design, implement, and evaluate a computer-based system, process, component, or program to meet desired needs
   ii. An ability to use current techniques, skills, and tools necessary for computing practice

6. Brief list of topics to be covered
   a. Data Exploration
   b. Data Preprocessing
   c. Mining and Analytics
   d. Visualization
   e. MapReduce and Cloud Computing

Communication:
   All course announcements will be posted to the Google Group's course mailing list. If you have a question to discuss with everyone, please post it to the group! If you have a specific question to me, please send me an email with 5890 in the subject line.

Grading Policy:
The course grading policy is as follows:
   5%      Attendance and in-class discussion
   30%     Assignments (3 assignments)
   20%     Midterm
   15%     Practicum
   30%     Project

The grading scale is A:93-100, A-:90-92, B+:87-89, B:83-86, B-:80-82, C+:77-79, C:73-76, C-:70-72, D:60-69, F:0-59

Assignments:
There are four assignments. Each assignment is proportion to 8% of your grade. You will have total 4 late days during the semester. You can use up to 2 late days for each assignment without penalty. After you consume the total 4 late days or two late days for an assignment (whichever comes first), then you will get penalty proportion to extra late days (e.g., 10% off for the next late day, 20% for the next two late days and so on).
For example, you submitted your first assignment 2 days late. You will not get any penalty, but use 2 out of 4 late days. Or if you submit your first assignment 3 days later than due date, you will use 2 late days (again up to 2 late days for an assignment), and get 10% off penalty because of the third late day. For each assignment, we will NOT accept your solution more than 5 days late.

You may discuss an assignment with your colleague, but you should write a program and a report by yourself and should NOT copy and paste your colleague's solution. If you discussed an assignment with your colleague, explicitly report the colleague's name and what you discussed in your submission.

Exam:
There is a 'closed book' exam which will be held in class. You may bring one standard 8.5" by 11" piece of paper with any notes you think appropriate or significant (front and back). No electronic devices allowed.

Practicum:
As part of our exploration of Data Science, we're going to engage in a series of practicums. A practicum, according to Wikipedia, is "designed to give students supervised practical application of a previously or concurrently studied theory." For this course, what that means, is that each week we will tackle some practical application of Data Science - be it a tool, a framework, or some other artifact that will help you transition your theoretical foundation into practice. You will be responsible for one practicum over the course of the semester. Detailed information will be given to you in class.

Project:
In this term project, you will apply data science techniques that you learned from this course to your project. The detailed information regarding the term project will be announced in class. You will present and may demonstrate your project in December 8 and 10.

Add policy:
The last day to add this class is September 21 (5:00 PM). Attending this class beyond that date, without being officially registered, will not be approved by the Dean's Office. Students must be officially registered for this course. No assignments or tests of any kind will be graded for students whose names do not appear on the class list.

Drop policy:
The last day to drop this class without notation is September 21 (5:00 PM).

Withdrawal Policy and "I" Grade Policy:
Students are required to complete all courses for which they are registered by the end of the semester. In some cases, a student may be unable to complete all of the coursework because of extenuating circumstances, but not due to poor performance or to retain financial aid. The term 'extenuating'
circumstances includes: (1) incapacitating illness which prevents a student from attending classes for a minimum period of two weeks, (2) a death in the immediate family, (3) financial responsibilities requiring a student to alter a work schedule to secure employment, (4) change in work schedule as required by an employer, or (5) other emergencies deemed appropriate by the instructor.

**Learning Aids:**
Lecture notes and other useful information will be available in electronic form on the class's section of the Canvas system. Please check the class's news and notes sections on a regular basis.

The Computer Science Department is a member of the Microsoft’s DreamSpark program. Through this program, students in CS courses can obtain and use a number of Microsoft's operating and software packages. If you are interested in downloading any of this software for your use, please follow the directions found on the department’s website.

**Academic Integrity — “The Honor System”:**
Each student has the right and duty to pursue his or her academic experience free of dishonesty. The Honor System is designed to establish the higher level of conduct expected and required of all Utah State University students.

The Honor Pledge: To enhance the learning environment at Utah State University and to develop student academic integrity, each student agrees to the following Honor Pledge: "I pledge, on my honor, to conduct myself with the foremost level of academic integrity." A student who lives by the Honor Pledge is a student who does more than not cheat, falsify, or plagiarize. A student who lives by the Honor Pledge:

- Espouses academic integrity as an underlying and essential principle of the Utah State University community;
- Understands that each act of academic dishonesty devalues every degree that is awarded by this institution; and
- Is a welcomed and valued member of Utah State University.

**Plagiarism and Cheating:**
Plagiarism includes knowingly "representing, by paraphrase or direct quotation, the published or unpublished work of another person as one's own in any academic exercise or activity without full and clear acknowledgment. It also includes the unacknowledged used of materials prepared by another person or agency engaged in the selling of term papers or other academic materials." The penalties for plagiarism are severe. They include warning or reprimand, grade adjustment, probation, suspension, expulsion, withholding of transcripts, denial or revocation of degrees, and referral to psychological counseling.

This course adheres to the cheating policy for courses in the Department of Computer Science posted on the bulletin board outside the CS office on the 4th floor of Old Main and posted online at http://cs.usu.edu/htm/cheating-policy/.
**Students with Disabilities:**
Students with ADA-documented physical, sensory, emotional or medical impairments may be eligible for reasonable accommodations. Veterans may also be eligible for services. All accommodations are coordinated through the Disability Resource Center (DRC) in Room 101 of the University Inn, (435)797-2444. Please contact the DRC as early in the semester as possible. Alternate format materials (Braille, large print, digital, or audio) are available with advance notice.

**Sexual Harassment:**
Sexual harassment is defined by the Affirmative Action/Equal Employment Opportunity Commission as any "unwelcome sexual advances, requests for sexual favors, and other verbal or physical conduct of a sexual nature." If you feel you are a victim of sexual harassment, you may talk to or file a complaint with the Affirmative Action/Equal Employment Opportunity Office located in Old Main, Room 161, or call the AA/EEO Office at 797-1266.

**Academic Freedom and Professional Responsibilities (Faculty Code):**
Academic freedom is the right to teach, study, discuss, investigate, discover, create, and publish freely. Academic freedom protects the rights of faculty members in teaching and of students in learning. Freedom in research is fundamental to the advancement of truth. Faculty members are entitled to full freedom in teaching, research, and creative activities, subject to the limitations imposed by professional responsibility. Faculty Code Policy #403 further defines academic freedom and professional responsibilities: USU Policies Section 403