

University of Kentucky  
School of Information Science (SIS)

## LIS/ICT 661: Introduction to Data Science

Spring 2019 (January 9 to April 28)

### Instructor

Youngseek Kim

Assistant Professor

Office: 331 Little Library Building

Phone: (859) 218 – 2295

Email: [youngseek.kim@uky.edu](mailto:youngseek.kim@uky.edu) and Canvas Messages (Preferred)

Response Time: Within 24 hours during weekdays (expect a delay during weekends and holidays)

### Office Hours

Thursdays: 9:00 a.m. – 1:00 p.m.

and by appointments

*Virtual office hours are available during the regular office hours via Zoom*

### Course Information

This is an online course. Please visit <https://uk.instructure.com> (Canvas) for course homepage.

## COURSE DESCRIPTION

This course will provide a foundation in the area of data science based on data curation and statistical analysis. The primary goal of this course is for students to learn data analysis concepts and techniques that facilitate making decisions from a rich data set. Students will investigate data concepts, metadata creation and interpretation, general linear method, cluster analysis, and basics of information visualization. At the beginning, this course will introduce fundamentals about data and data standards and methods for organizing, curating, and preserving data for reuse. Then, we will focus on the inferential statistics: drawing conclusions and making decisions from data. This course will help students understand how to use data analysis tools, and especially, provide an opportunity to utilize an open source data analysis tool, *R*, for data manipulation, analysis, and visualization. Finally, in this course we will discuss diverse issues around data including technologies, behaviors, organizations, policies, and society.

## COURSE OBJECTIVES

Upon successful completion of this course, students should have developed some or all of the following areas of skills and knowledge:

- An understanding of how the nature of the data collection, the data itself, and the analysis processes relate to the kinds of inferences that can be drawn
- Understand the limitations of data sets based on their contents and provenance
- Knowledge of data organization, management, preservation, and reuse
- Knowledge of what statistical analysis techniques to choose, given particular demands of inference and available data
- Knowledge of general linear models and cluster analysis methods for statistical analysis
- Skills and knowledge in preparing data for analysis, including cleaning data, manipulating data, and dealing with missing data
- Skills in actually analyzing data using open source data analysis tools
- Skills in scripting for data manipulation, analysis, and visualization using *R*, *R-Studio*, and a variety of add on packages.

## COURSE MATERIALS

### ***Required Textbook:***

Jeffrey S. Saltz & Jeffrey M. Stanton (2018). An Introduction to Data Science (First Edition). Sage Publications. ISBN-13: 978-1506377537; ISBN-10: 150637753X

### ***Additional readings:***

For some weeks, there will be additional readings to the textbook, and they will be made available on the Canvas course site.

## COURSE CONDUCT

The course includes *online* lectures (PPT slides), *online* discussions and exercises, case studies, assignments, and project. As this is an online course with no set meeting times, the question arises when are assignments due each week. For the purposes of this class we will treat Monday as the first day of class each week. Also, readings should be completed by at least Wednesday in order to participate in online discussions. I will put up course notes (PPT slides) no later than midnight on Sunday of each week. Regarding the weekly discussion, you need to answer some discussion questions and discuss them (see more information below). The discussion board is asynchronous, meaning that you can join in the discussion whenever it is convenient for you to do so.

### ***Canvas as the Learning Platform of the Course***

This class uses Canvas as a required and main part of the course. Teaching materials (syllabus, course notes, discussions, assignments, resources, etc.) will be made available in the Canvas. All assignments should also be submitted to the Canvas. Students can check grading status and progress in the Canvas. Please visit the Canvas Information Pages at <http://www.uky.edu/canvas/> to learn about the Canvas. For technical support, call the Information Technology Services (ITS) at (859) 218-HELP (4357) or email [218help@uky.edu](mailto:218help@uky.edu).

### ***Communications:***

All course related communications (online discussions, queries on assignments, etc.) will occur within the Canvas. Please post your questions on the Canvas discussion board because other students may have the same questions and receive the benefits from answers. Important announcements will be made inside the Canvas. Students thus are required to check Canvas on a regular basis. Failure to receive such announcements cannot be used as an excuse for not being informed.

I welcome emails sent to my UKY.EDU email account and/or Canvas Messages account. I prefer to use the Canvas Messages in order to keep all course related emails in one place to facilitate communication; however, please feel free to send me any email message to my UKY.EDU email account. Please do not expect an immediate response on your email message. However, in ordinary circumstances, it is expected that the instructor will respond within 24 hours during weekdays. Please expect a delay during weekends and holidays.

## METHODS OF ASSESSMENT

Your final grade is determined by your performance on the items in the table below. Assignments and individual project will be assigned as the course continues. Lastly, we will have a weekly discussion.

Each assignment will be posted at least two weeks before its due date. Please submit your assignments and project through Canvas assignment drop-box. The table below shows the list of assignments, project, and participation and the points for each component.

Type	Topic	Points	Sub Total
	Assignment #1: Developing Research Questions	5.0	35.0
	Assignment #2: Descriptive Statistics Exercise	10.0	
	Assignment #3: Regression/Modeling Exercise	10.0	
	Assignment #4: Cluster Analysis Exercise	10.0	
Project	Final Project: Data Analysis Report	30.0	30.0
Participation	Topical Presentation	5.0	35.0
	Participation & Discussions (2 points per each week)	30.0	
<b>Total: 100.0</b>			

### ***Online Discussions:***

Students are expected to participate in the weekly discussion (or exercise) via Canvas Discussions. I will pose several discussion questions (and/or a few exercises), so you can have the discussions based on the questions (and/or exercises). In addition, there will be a course content discussion thread, where you can talk about each week's course material.

The discussion will start on Monday morning, and it will end on the following Sunday night. I will leave the discussion forum open later, but I am going to evaluate your postings during the week ONLY. You can earn up to 2.0 points for each week's discussion (and/or exercises). You can earn up to 0.5 point for any valuable posting regarding the discussion questions which are posed for each week, your own question regarding each week's course content, answers for the questions posed by other students. This means that you need to post at least 4 posting a week in order to get 2.0 points for that week.

A valuable posting or quality posting includes a substantive and thoughtful contribution to each week's discussion topics, during that week. No credit will be given for posts that occur after the week. A quality posting is both substantive (in most instances this means at least 100 words) and thoughtful ("I agree with the author" only is not a credit-worthy response). Also, please write each discussion posting concisely (100 to 150 words – 1 or 2 paragraphs). I encourage you to complete your discussion posts and other work in Word and then paste it to Canvas. If you compose online and there is a technology-related failure, you will likely lose your work.

### ***Grade Expectations:***

Grades are based on the quality of the submitted work, not upon how well others performed. The following are grade expectations and divisions.

Grade	Score (Percentage)	Expectation
A	90% - 100%	Exceptional Achievement
B	80% - 89%	High Achievement
C	70% - 79%	Average Achievement
E	0% - 69%	Failing

### ***Late assignment policy:***

Some of the assignments will be discussed in following week's online discussions after the assignments are due. Most assignments will help build a base for future assignments and the project. Thus all assignments should be turned in on time as specified. An overdue assignment will get a penalty of 20% of total points for each day late. No assignment and project will be accepted after five days except 'excused

absences' ([see more information below](#)).

***Re-Grading Requests:***

The grade for each assignment is recorded in the Canvas before the assignment is returned to the student. It is a student's responsibility to ask questions or request re-grading of an assignment within five business days from the time the assignment is returned. No re-grading requests will be accepted after the five business day period.

**ATTENDANCE POLICY**

This class is a community whose success depends on everyone's participation, and it is vital for you to attend class online in order to be successful. If you miss any assignment or weekly discussion, you will be allowed to make up that work only if the absence is officially excused. You will be asked to provide official written documentation for excused absences in order for your absence to be excused. If you know ahead of time that you will be absent from class with an excused absence, please discuss this with your instructor and turn in any assignments ahead of time. Excuses for university-sponsored activities must be made prior to such absences. For any emergency situation that arises, email your instructor as soon as you know about the situation when possible. No make-up work is available for online discussions or assignments unless approved in advance by your instructor. Students have one week to complete a missed assignment or online discussion due to an excused absence upon their return without penalty.

***Excused Absences:***

Students need to notify the professor of absences prior to class when possible. Senate Rules 5.2.4.2 defines the following as acceptable reasons for excused absences: (a) serious illness, (b) illness or death of family member, (c) University-related trips, (d) major religious holidays, (e) interviews for graduate/professional school or full-time employment post-graduation, and (f) other circumstances found to fit "reasonable cause for nonattendance" by the professor.

Students anticipating an absence for a major religious holiday are responsible for notifying the instructor in writing of anticipated absences due to their observance of such holidays no later than the last day in the semester to add a class. Two weeks prior to the absence is reasonable, but should not be given any later. Information regarding major religious holidays may be obtained through the Ombud (859-257-3737, [http://www.uky.edu/Ombud/ForStudents\\_ExcusedAbsences.php](http://www.uky.edu/Ombud/ForStudents_ExcusedAbsences.php)).

In situations where a student's total EXCUSED absences exceed 1/5 (or 20%) of the class periods scheduled for the semester, students are strongly encouraged to withdraw (take a "W") from the class as per university policy. If a student has excused absences in excess of one-fifth of the class contact hours for that course, the student shall have the right to receive a 'W', or the Instructor of Record may award an 'I' for the course if the student declines to receive a 'W.'

Per Senate Rule 5.2.4.2, students missing any graded work due to an excused absence are responsible: for informing the Instructor of Record about their excused absence within one week following the period of the excused absence (except where prior notification is required); and for making up the missed work. The professor must give the student an opportunity to make up the work and/or the exams missed due to an excused absence, and shall do so, if feasible, during the semester in which the absence occurred.

***Verification of Absences:***

Students may be asked to verify their absences in order for them to be considered excused. Senate Rule 5.2.4.2 states that faculty have the right to request "appropriate verification" when students claim an excused absence because of illness, or death in the family. Appropriate notification of absences due to

University-related trips is required prior to the absence when feasible and in no case more than one week after the absence.

## **ACADEMIC INTEGRITY**

Per University policy, students shall not plagiarize, cheat, or falsify or misuse academic records. Students are expected to adhere to University policy on cheating and plagiarism in all courses. The minimum penalty for a first offense is a zero on the assignment on which the offense occurred. If the offense is considered severe or the student has other academic offenses on their record, more serious penalties, up to suspension from the University may be imposed.

Plagiarism and cheating are serious breaches of academic conduct. Each student is advised to become familiar with the various forms of academic dishonesty as explained in the Code of Student Rights and Responsibilities. Complete information can be found at the following website: <http://www.uky.edu/Ombud>. A plea of ignorance is not acceptable as a defense against the charge of academic dishonesty. It is important that you review this information as all ideas borrowed from others need to be properly credited.

Senate Rules 6.3.1 (see <http://www.uky.edu/Faculty/Senate/> for the current set of Senate Rules) states that all academic work, written or otherwise, submitted by students to their instructors or other academic supervisors, is expected to be the result of their own thought, research, or self-expression. In cases where students feel unsure about a question of plagiarism involving their work, they are obliged to consult their instructors on the matter before submission.

When students submit work purporting to be their own, but which in any way borrows ideas, organization, wording, or content from another source without appropriate acknowledgment of the fact, the students are guilty of plagiarism.

Plagiarism includes reproducing someone else's work (including, but not limited to a published article, a book, a website, computer code, or a paper from a friend) without clear attribution. Plagiarism also includes the practice of employing or allowing another person to alter or revise the work, which a student submits as his/her own, whoever that other person may be. Students may discuss assignments among themselves or with an instructor or tutor, but when the actual work is done, it must be done by the student, and the student alone.

When a student's assignment involves research in outside sources or information, the student must carefully acknowledge exactly what, where and how he/she has employed them. If the words of someone else are used, the student must put quotation marks around the passage in question and add an appropriate indication of its origin. Making simple changes while leaving the organization, content, and phraseology intact is plagiaristic. However, nothing in these Rules shall apply to those ideas, which are so generally and freely circulated as to be a part of the public domain.

Please note: Any assignment you turn in may be submitted to an electronic database to check for plagiarism.

## **ACCOMMODATIONS DUE TO DISABILITY**

If you have a documented disability that requires academic accommodations in this course, please make your request to the University Disability Resource Center. The Center will require current disability documentation. When accommodations are approved, the Center will provide me with a Letter of Accommodation which details the recommended accommodations. In order to receive accommodations in

this course, you must provide me with a Letter of Accommodation from the Disability Resource Center (DRC). The DRC coordinates campus disability services available to students with disabilities. Contact the Disability Resource Center, Jake Karnes, Director at 859-257-2754 or [jkarnes@email.uky.edu](mailto:jkarnes@email.uky.edu). Their web address is <http://www.uky.edu/DisabilityResourceCenter>.

## TECHNOLOGY INFORMATION & RESOURCES

Distance Learning Students are expected to have a minimum level of technological acumen and the availability of technological resources. Students must have regular access a computer with a reliable Internet connection and audio capabilities. Please see our student Technical Requirements and Recommendations at <http://www.uky.edu/elearning/technical-requirements>. Please be certain that your computer and/or browser allow you to view Adobe Reader documents (.pdf). Microsoft Office and other software products are free for students: <http://download.uky.edu/>.

As your instructor, I am your first go-to person for technology problems. If you need more immediate assistance, please contact Information Technology Services (ITS).

### ***Information Technology Services (ITS):***

<http://www.uky.edu/its/>; 859-218-4357

### ***Library Services & Distance Learning Services:***

<http://libraries.uky.edu/DLLS>

Carla Cantagallo, Distance Learning Librarian

Phone Number: (859) 218-1240

Email: [carla@uky.edu](mailto:carla@uky.edu)

Interlibrary Loan Service: <http://libraries.uky.edu/ILL>

For more resources about online classes and student resources, visit <http://www.uky.edu/ukonline/>

The School of Information Science has a page with a comprehensive list of technology resources here:

<http://ci.uky.edu/sis/students/techtips>

### ***Military Members and Veterans:***

We recognize the complexities of being a member of the military community and also a student. If you are a member of the military or a military veteran or dependent, please inform your instructor if you are in need of special accommodations. Drill schedules, calls to active duty, mandatory training exercises, complications with GI Bill disbursement, and other unforeseen military and veteran related developments can complicate your academic life. If you are aware of a complication, we will work with you and put you in contact with university staff members who are trained to assist you. Please contact the Coordinator of the University of Kentucky Veterans Resource Center at (859) 257-1148 for additional assistance. Visit <http://www.uky.edu/veterans> for more available resources.