

**DSA/ISE 5103**  
**INTELLIGENT DATA ANALYTICS**  
Tuesdays and Thursdays 3:00p – 4:15p  
Gould Hall 155  
Fall 2021

**Instructor:** Charles Nicholson, Ph.D.  
**Email:** [cnicholson@ou.edu](mailto:cnicholson@ou.edu)  
**Office hours:** T/R 10:30a-12:00p via Zoom  
**Facebook:** facebook.com/ou.analytics  
**Website:** oklahomaAnalytics.com

**GTA:** Robert Sandel  
**Email:** [robertsandel@ou.edu](mailto:robertsandel@ou.edu)  
**Office hours:** TBD  
**Office hour location:** TBD

**Zoom links:**

- Instructor office hours:

<https://oklahoma.zoom.us/j/92199869128?pwd=VFRiZjBBbGdJemN2RTBqc3ZYTEFkUT09>

- Passcode: OfficeHrs

**Course description:** Intelligent Data Analytics is an approach to addressing real-world data intensive problems that integrates human intuition with data analysis tools to best draw out meaningful insights. Topics include problem approach and framing, data cleansing, exploratory analysis and visualization, dimension reduction, linear and logistic regression, decision trees, and clustering, among others. Students will be introduced to a powerful open source statistical programming language (R) and work on hands-on, applied data analysis projects.

**Student outcomes:** You will demonstrate mastery in 4 areas:

- (1) **Defining and framing Analytics Problems,**
- (2) **Understanding and coping with Data,**
- (3) **Selecting and using appropriate Analytical Tools,**
- (4) **Discovering and communicating the Insight.**

**Textbooks**

Max Kuhn and Kjell Johnson. 2013. *Applied Predictive Modeling*. Springer

<http://libraries.ou.edu/access.aspx?url=http://link.springer.com/book/10.1007/978-1-4614-6849-3>

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. 4th printing. <http://www-bcf.usc.edu/~gareth/ISL/>

Trevor Hastie, Rober Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Ed. <http://statweb.stanford.edu/~tibs/ElemStatLearn/>

**Software Resources**

R <http://www.r-project.org/>  
RStudio IDE <http://www.rstudio.com/ide/>  
R for Beginners [http://cran.r-project.org/doc/contrib/Paradis-rdebuts\\_en.pdf](http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf)  
Introduction to R <http://cran.r-project.org/doc/manuals/R-intro.pdf>

**Grading Policy:** All assignments must be uploaded to course website by the due date and time indicated for that assignment on the course website. Ensure your familiarity with the system in advance, and allot a sufficient amount of time for assignment submission *and any technical difficulties that may arise*.

Assignments should be submitted in advance of an anticipated absence.

**Late assignments will be penalized.** See following table for breakdown:

Submission is:	Maximum grade possible:
On time + up to 10 min late	100%
Up to 12 hours late	80%
12 to 24 hours late	60%
More than 24 hours late	0%

**Slack day:** Each student will be allocated *two slack days* per semester. A slack day is a 24-hour extension for a homework assignment (not a course project or other assessment task). The slack day may be invoked only once, but it may be used as the student sees fit. To invoke the slack day please provide a comment in the submission location when the work is submitted. Note: for team assignments, each “slack day” value is equal to 24 hours / # of students on the team. For example, if there are two students, and one student uses their slack day, there will be a 12-hour extension. If all students invoke their slack days, then there will be a 24-hour extension.

**Grading Details:**

*Percentages of course grading requirements are as follows:*

<b>Homework</b>	60%
<b>Project</b>	
<ul style="list-style-type: none"> <li>• Project Team Selection and Proposal</li> <li>• Project: Initial data analysis (12%)</li> <li>• Project: Initial draft (12%)</li> <li>• Project: Critique (6%)</li> <li>• Project: Presentation (35%)</li> <li>• Project: Final report (35%)</li> </ul>	30%
<b>Participation</b>	10%

**Grading scale**

Total grade percentages in the following intervals result in the associated letter grade:  
 [90-100] ← A; [80-90) ← B; [70,80) ← C; [60,70) ← D; [0,60) ← F

**Homework:** Assignments consist of a variety of written and programming problems. Students must perform the majority of work in R, however this can be supplemented with other tools (e.g., Excel, Python). For each assignment please read the submission requirements outlined in the instructions.

**Class Project:** The class project will be developed during the last half of the semester. Instructions, guidelines and due dates will be posted on the course website. The project has several deliverables. Make sure you are aware of the deadlines for each.

- Project team selection and proposal.
- Initial data analysis.
- Initial draft.
- Peer critique.
- Powerpoint slides and recorded video.
- Final report.

## Tentative Schedule

Week beginning (Monday)	Topic	Notes	Due (subject to change)
23-Aug	Introduction to ISE 5103; Intro to Analytics Intro to R and RStudio Project Understanding	First day of course: 8/25 Assign: HW #1	Install R and RStudio on your computers Download the textbooks
30-Aug	Data Understanding: visualizations, outlier analysis, missing value mechanisms	Assign: HW #2	HW #1 (R and R Studio): 1-Sep @ 11:59 pm (Wednesday)
6-Sep	Data Understanding: Dimension Reduction (PCA, LDA, t-SNE)	Assign: HW #3 <i>Labor Day: 6-Sep</i>	HW #2 (Data Understanding): 8-Sep @ 11:59 pm (Wednesday)
13-Sep	Data Preparation: dealing with outliers, missing value imputation, feature engineering, and transformation	Assign: HW #4	HW #3 (Dimension Reduction): 15-Sep @ 11:59 pm (Wednesday)
20-Sep	Principles of Modeling: performance metrics (I) and resampling Multiple Linear Regression (MLR): OLS	Assign: HW #5	HW #4 (Data Preparation): 22-Sep @11:59 pm (Wednesday)
27-Sep	OLS and OLS variants: PCR, PLS, Ridge, Lasso, Elasticnet, MARS, SVM-R		
4-Oct	OLS variants continued	Assign: HW #6 <i>Fall Student Holiday: 10/8</i>	HW #5 (MLR I): 6-Oct @11:59 pm (Wednesday)
11-Oct	Principles of Modeling: performance metrics (II) Classification: Logistic Regression	Check Project information!	**Form Project Teams
18-Oct	Classification: Decision Trees, Random Forests, Boosted Trees, SVM, Neural Networks		**Project Teams/Proposal: 22-Oct @11:59 pm (Friday)
25-Oct	Classification continued	Assign: HW #7	HW #6 (MLR II): 27-Oct @11:59 pm (Wednesday)
1-Nov	Classification continued		**Project Data Analysis: 31-Oct @11:59 pm (Sunday)
8-Nov	Clustering: k-means, PAM, hierarchical, density-based clustering	Assign: HW #8	HW #7 (Classification): 9-Nov @11:59 pm (Tuesday)
15-Nov	Association Mining		HW #8 (Unsupervised Learning): 19-Nov @11:59 pm (Friday)
22-Nov	Unsupervised Learning: Autoencoding	<i>Thanksgiving: 11/24-11/28</i>	**Project Report 1 <sup>st</sup> draft: 23-Nov @11:59p (Tuesday)
29-Nov	Project discussion	Conduct Peer Review	**Project Report Critique: 1-Dec @ 11:59 pm (Wednesday)
6-Dec	Open Topics	<i>Finalize projects</i>	**Recorded Presentations 8-Dec (Wednesday) @ 11:59p
13-Dec		No final exam!	**Final Project: 15-Dec @ 11:59p (Wednesday)

**Attendance, Participation, Assessment:** Participation is evaluated and counts toward your final grade. You are expected to view all videos and complete all reading assignments. Also, you will be expected to participate in online discussions and attempt all Self-Check Quizzes.

### Professional behavior

While many of the tools necessary for the completion of this course will be discussed in the course content or via guided assignments, nevertheless **considerable work outside of class in learning R or researching analytical techniques is expected.**

### Academic honesty

Cheating, plagiarism, or any act of dishonesty will NOT be tolerated. This policy applies to all parties involved in the incident. Never take credit for anyone else's intellectual property, be it on an exam or homework assignment. This includes, but is not limited to, copying from another student's paper, copying from a paper from a previous semester, using forbidden information on exams, and copying from published writings. Additionally, you are strictly prohibited from using online services such as Chegg.com to complete

homework, quizzes, or exams in this course. Students are responsible for knowing the requirements of the Academic Misconduct Code, available at <http://integrity.ou.edu/>.

### **Reasonable accommodation policy**

Students requiring academic accommodation should contact the Disability Resource Center for assistance at (405) 325- 3852 or TDD: (405) 325-4173. For more information please see the Disability Resource Center website <http://www.ou.edu/drc/home.html>. Any student in this course who has a disability that may prevent him/her from fully demonstrating his/her abilities should contact me personally as soon as possible so we can discuss accommodations necessary to ensure full participation and facilitate your educational opportunities.

### **Title IX Resources and Reporting Requirement**

For any concerns regarding gender-based discrimination, sexual harassment, sexual assault, dating/domestic violence, or stalking, the University offers a variety of resources. To learn more or to report an incident, please contact the Sexual Misconduct Office at 405/325-2215 (8 to 5, M-F) or [smo@ou.edu](mailto:smo@ou.edu). Incidents can also be reported confidentially to OU Advocates at 405/615-0013 (phones are answered 24 hours a day, 7 days a week). Also, please be advised that a professor/GA/TA is required to report instances of sexual harassment, sexual assault, or discrimination to the Sexual Misconduct Office. Inquiries regarding non-discrimination policies may be directed to: Bobby J. Mason, University Equal Opportunity Officer and Title IX Coordinator at 405/325-3546 or [bjm@ou.edu](mailto:bjm@ou.edu). For more information, visit <http://www.ou.edu/eoo.html>.

### **Religious Observance**

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty.

### **Adjustments for Pregnancy/Childbirth Related Issues**

Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact your professor or the Disability Resource Center at 405/325-3852 as soon as possible. Generally, modifications will be made where medically necessary and similar in scope to accommodations based on temporary disability. Please see [www.ou.edu/content/eoo/faqs/pregnancy-faqs.html](http://www.ou.edu/content/eoo/faqs/pregnancy-faqs.html) for commonly asked questions.

### **Mental Health Support Services**

If you are experiencing any mental health issues that are impacting your academic performance, counseling is available at the University Counseling Center (UCC). The Center is located on the second floor of the Goddard Health Center, at 620 Elm Rm. 201, Norman, OK 73019. To schedule an appointment call (405) 325-2911. For more information please visit: <http://www.ou.edu/ucc>

### **Pandemic-Related Protocols**

The university encourages masking indoors. The university strongly encourages masking for all individuals in high-density settings, such as classrooms and at special events. [www.ou.edu/together/university-masking-policy](http://www.ou.edu/together/university-masking-policy)

The university strongly encourages the entire OU community to get vaccinated, particularly those residing in congregate housing. COVID-19 vaccines are available to all students, faculty, and staff on all three OU campuses at no cost to the individual: [www.ou.edu/together/vaccine](http://www.ou.edu/together/vaccine)

***Students are responsible for any changes/additions to this syllabus announced during the semester.***