EECS 731: DATA SCIENCE
FALL 2016

Professor: Nicole Beckage
Office: Eaton 3022
Office Hours: T 12:30pm – 1:30pm, R 11:00am-12:00pm, and by appointment
E-Mail: beckage@ku.edu
Course Website: https://csel.cs.colorado.edu/~nibe3229/ds/index.html
Forum Website: piazza.com/ku/fall2016/eecs731
Lecture Hours and Venue: TR 2:30pm – 3:45pm in Learned 1136

Description: Data Science draws on tools from computer science and statistics to answer disciplinary scientific questions by integrating different analysis techniques or types of data. This course aims to introduce students to tools from computer science and statistics used to analyze data. What questions can we answer with a particular data set? How can we collect data that is sufficient to study our topic of interest? We consider how to construct predictive and explanatory models of data as well as the limitations of models. A large component of this course is the process and evolution of a research question using data science. The course is project focused but includes exams related to understanding the statistical framework underpinning data science. We will cover theoretical topics of statistics and probability as well as applied topics related to real-world data science. We focus on scientific thinking in the context of data. This course aims to define data science in terms of using data to ask and support scientific ideas and hypotheses.

Additional Textbook: Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani An Introduction to Statistical Learning.
PDF: www-bcf.usc.edu/~gareth/ISL.

Homework: There will be 3-5 homework assignments throughout the semester covering topics central to data science. These homeworks will be submitted electronically and are due by 11:55pm central time of the day listed on the schedule. Homework will be graded based on the correctness as well the clarity of the solution. No late homework will be accepted.

Quizzes: There will be 4-8 statistics quizzes throughout the semester. Your lowest quiz scores will be dropped.

Exams: There will be an exam on Thursday, October 13th. Exams will consist of an in-class portion as well as a take-home portion, which will be due the following Thursday. The final exam information is to be announced.

Final Project: The final project will be a research project. This project can be done either as a team or individually. If the project is a team project, the team
can have no more than 5 people (3-4 preferred). Teams will be required to give a short (7-10 minute) presentation during the last two sessions of class. All projects require a final paper. More details of the project will be discussed throughout the course.

**Course Grading:** Your course grade will be calculated based on the breakdown below.

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homeworks</td>
<td>20%</td>
</tr>
<tr>
<td>Quizzes</td>
<td>10%</td>
</tr>
<tr>
<td>Exams (incl. final)</td>
<td>40%</td>
</tr>
<tr>
<td>Final Project</td>
<td>30%</td>
</tr>
</tbody>
</table>

**Academic Misconduct:** All work submitted for credit must be the student’s own and is subject to the provisions of KU policies. Students should review the university policy on Academic conduct at: [http://policy.ku.edu/governance/USRR#art2sect6](http://policy.ku.edu/governance/USRR#art2sect6).

**Accommodation Procedure:** The Academic Achievement and Access Center (AAAC) coordinates academic accommodations and services for all eligible KU students with disabilities. If you have a disability for which you wish to request accommodations and have not contacted the AAAC, please do so as soon as possible. They are located in 22 Strong Hall and can be reached at 785-864-4064 (V/TTY). Information about their services can be found at [www.access.ku.edu](http://www.access.ku.edu). Please contact me privately in regard to your needs in this course.

**Nondiscrimination:** The University of Kansas prohibits discrimination on the basis of race, color, ethnicity, religion, sex, national origin, age, ancestry, disability, status as a veteran, sexual orientation, marital status, parental status, retaliation, gender identity, gender expression and genetic information in the University’s programs and activities. Please contact the University’s Title IX Coordinator at IOA@ku.edu with any inquiries.

**Religious Observances:** Should the examination schedule for this course conflict with your mandated religious observance, please contact me at the beginning of the semester so that we can schedule a make-up exam at a mutually acceptable time. In addition, students will not be penalized for absence from regularly scheduled class activities which conflict with mandated religious observances. Students are responsible for initiating discussion with the instructor to reach a mutually acceptable solution.
Tentative Schedule

Week 1: What is Data Science?
Week 2: Data Science in Python and R
Week 3: Getting Data
Week 4: Probability Theory
Week 5: Data Visualization
Week 6: Central Tendencies
Week 7: Hypothesis and Inference
Week 8: Midterm
Week 9: Cost Functions and Optimization
Week 10: Model Validation
Week 11: Predictive Modeling
Week 12: Generative Modeling
Week 13: Linear Regression
Week 14: Logistic Regression
Week 15: Classification
Week 16: Classification
Week 17: Projects