## Statistics

**STAT 430  **Topics in Applied Statistics  **credit: 3 OR 4 hours.**  
Formulation and analysis of mathematical models for random phenomena; extensive involvement with the analysis of real data; and instruction in statistical and computing techniques as needed. 3 undergraduate hours. 4 graduate hours. May be repeated with approval. Prerequisite: STAT 410 or STAT 420; or consent of instructor.

<table>
<thead>
<tr>
<th>CRN</th>
<th>Type</th>
<th>Section</th>
<th>Time</th>
<th>Days</th>
<th>Location</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>55664</td>
<td>Lecture-Discussion</td>
<td>2GR</td>
<td>08:00 AM - 08:50 AM</td>
<td>MWF</td>
<td>103 - Transportation Building</td>
<td>Dalpiaz, D</td>
</tr>
</tbody>
</table>

Credit Hours: 4 hours  
Basics of Statistical Learning  
Restricted to Graduate - Urbana-Champaign.  
For up-to-date information about statistics course registration, please see our registration update pages: go.illinois.edu/StatisticsRegistration  
**TOPIC: Basics of Statistical Learning**  
Description: This course introduces machine learning techniques for prediction, classification, and clustering. There is an emphasis on resampling methods in model building, especially cross validation. Topics include model selection, nonparametric regression, logistic regression, decision trees, random forests, support vector machines, dimension reduction and cluster analysis. Prerequisite: A course in linear regression, such as STAT 420 or STAT 425.

<table>
<thead>
<tr>
<th>CRN</th>
<th>Type</th>
<th>Section</th>
<th>Time</th>
<th>Days</th>
<th>Location</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>55666</td>
<td>Lecture-Discussion</td>
<td>2UG</td>
<td>08:00 AM - 08:50 AM</td>
<td>MWF</td>
<td>103 - Transportation Building</td>
<td>Dalpiaz, D</td>
</tr>
</tbody>
</table>

Credit Hours: 3 hours  
Basics of Statistical Learning  
Restricted to Undergrad - Urbana-Champaign.  
For up-to-date information about statistics course registration, please see our registration update pages: go.illinois.edu/StatisticsRegistration  
**TOPIC: Basics of Statistical Learning**  
Description: This course introduces machine learning techniques for prediction, classification, and clustering. There is an emphasis on resampling methods in model building, especially cross validation. Topics include model selection, nonparametric regression, logistic regression, decision trees, random forests, support vector machines, dimension reduction and cluster analysis. Prerequisite: A course in linear regression, such as STAT 420 or STAT 425.

<table>
<thead>
<tr>
<th>CRN</th>
<th>Type</th>
<th>Section</th>
<th>Time</th>
<th>Days</th>
<th>Location</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>46976</td>
<td>Lecture-Discussion</td>
<td>ID</td>
<td>01:00 PM - 03:50 PM</td>
<td>M</td>
<td>126 - Grad Sch of Lib &amp; Info Science</td>
<td>Stodden, V</td>
</tr>
</tbody>
</table>

Credit Hours: 4 hours  
Introduction to Data Science  
Restricted to Statistics or Statistics & Computer Science major(s) or minor(s).  
For up-to-date information about statistics course registration, please see our registration update pages: go.illinois.edu/StatisticsRegistration  
**TOPIC: Introduction to Data Science**  
Description: This class meets with IS 490 section ID (CRN 68848) and CS 398 section ID (CRN 69240). Please see IS 490 section ID for more information. This course is intended to introduce students to modern programs and technologies that are useful for organizing, manipulating, analyzing, and visualizing data. We start with an overview of the R language, which will become the foundation for your work in this class. Then we'll move on to other useful tools, including working with regular expressions, basic UNIX tools, XML, and SQL. We'll also cover supervised and unsupervised statistical learning techniques made possible by recent advances in computing power. This course is very computer-oriented, so it's very important to take the time outside of class to learn by doing to explore the software we'll be covering in class, and try out new skills on real datasets in the homework assignments. [This information was updated on 6/26/2017 and may have changed since then. Please LIS490IDS on https://ischool.illinois.edu/academics/courses/catalog#400level for the most updated description.]  
Restrictions: The STAT 430 section is restricted to Statistics students only. All other students would register for IS 490 section ID (CRN 68848) or CS 398 section ID (CRN 69240). Priority registration is restricted to students majoring in Statistics or Statistics & Computer Science. Students minoring in Statistics are not able to register during the priority registration period. Please see our registration update pages for further details: go.illinois.edu/StatisticsRegistration
Credit Hours: 3 hours
Foundations of Data Science
Restricted to Statistics or Statistics & Computer Science major(s) or minor(s). Not intended for students with Freshman class standing.
For up-to-date information about statistics course registration, please see our registration update pages: go.illinois.edu/StatisticsRegistration Students MUST register by August 30 at 4 pm. Registration in this course after that point will not be permitted.

TOPIC: Foundations of Data Science Description: This class is an asynchronous, online course. This class meets with INFO 490 section RB (CRN 65222) and IS 490 section RB (CRN 68792). Please see INFO 490 section RB for more information. This course will build a practical foundation for data science by teaching students basic tools and techniques that can scale to large computational systems and massive data sets. Students will first learn how to work at a Unix command prompt before learning about source code control software like git and the github site. Next, the Python programming language will be covered, with a focus on specific aspects of the language and associated Python modules that are relevant for Data Science. Python will be introduced and used primarily via the IPython (or Jupyter) Notebooks, and will cover the Numpy, Scipy, Matplotlib, Pandas, Seaborn, and scikit_learn Python modules. These capabilities will be demonstrated through simple data science tasks such as obtaining data, cleaning data, visualizing data, and basic data analysis. Students must have access to a fairly modern computer, ideally that supports hardware virtualization, on which they can install software. This class is open to sophomores, juniors, seniors and graduate students in any discipline. Restrictions: Not intended for students with Freshman class standing. The STAT 430 section is restricted to Statistics students only. All other students would register for INFO 490 section RB (CRN 65222) or IS 490 section RB (CRN 68792). Priority registration is restricted to students majoring in Statistics or Statistics & Computer Science. Students minoring in Statistics are not able to register during the priority registration period. Please see our registration update pages for further details: go.illinois.edu/StatisticsRegistration