**ENVS 10: Introduction to Environmental Statistics**

**Instructor:** Michael Cox  
**Office location:** 105 Fairchild  
**Email:** Michael.e.cox@dartmouth.edu  
**Class periods:** 10A – Tuesdays and Thursdays 10:00-11:50  
**Classroom:** 061 Carson  
**Office hours:** Wednesdays from 2 to 4pm (includes X-hours)  
**X-hours:** Wednesdays 3:00-3:50

**Course objective and background**

This course is designed to introduce students to the basic statistical methods necessary to conduct and understand statistical analyses of environmental issues and problems. Methods and subjects include: measurement, descriptive statistics, graphs, basic probability, correlation and regression, and basic inferential statistics (hypothesis testing, confidence interval construction, effect size calculation). There is a strong emphasis on the application of these and other techniques to socio-environmental datasets that describe current environmental problems and issues. No previous experience with mathematics or particular software is required. Students will also be required to complete a final project in which they use statistical techniques to address an environmental research question. Because of the large overlap in material covered, no student may receive credit for more than one of the courses ECON 10, ENVS 10, GOVT 10, MATH 10, PSYC 10, Social Sciences 10, or SOCY 10 except by special petition.

Examples of potential research questions include:

1) How quickly has the concentration of CO\textsubscript{2} been increasing in the earth’s atmosphere?  
2) What mixes of experimental treatments are most effective in increasing crop growth?  
3) What technologies and gear types lead fishermen to catch what types of fish in a small-scale fishery?  
4) Are more sustainable businesses also more profitable, and if so, under what conditions?  
5) How does urbanization affect water quality?  
6) What types of crops are associated with increased rates of farmer pesticide poisoning events?

At the end of this course you will be able to apply statistical techniques to answer questions like these. As the list implies, the range of such questions is quite broad.

**Prerequisites:**

ENVS 3 or ENVS 2 or permission of the instructor

**Basic course information:**

By the end of this course you will be able to:

- Determine which basic statistical methods are most appropriate for different types of data.  
- Conduct these statistical methods and interpret their results.  
- Read, understand, and digest articles that use statistical methods.
In each week we will cover (1) a new statistical method/topic and (2) apply this approach to an environmental dataset. The following activities will take place most Tuesdays and Thursdays:

Tuesday: Discussion of previous week’s lab assignment
Discussion of the day’s reading assignment
Work in Stata

Thursday: Quiz
Discussion of the day’s reading assignment
Work in Stata
Introduction to that week’s lab assignment

Finally, for the sake of our sanity, each class will be divided into two parts, with a ten-minute break in the middle.

Readings, software and statistical resources


We will be also using Stata to conduct statistical analyses throughout the term. Stata is available for free on the Dartmouth Network and is also available on most public computers on campus. See the following link for more information: http://www.dartmouth.edu/comp/soft-comp/software/statistics/stataintro.html

In addition, there are many online resources to help you learn Stata.


Stata’s own website, with video tutorials: http://www.stata.com/links/video-tutorials/

UCLA’s Stata resources site: http://www.ats.ucla.edu/stat/stata

Princeton’s Stata page: http://data.princeton.edu/stata

Consult Mr. Hua by email at Jianjun.hua@dartmouth.edu or make an appointment to meet with him.

Also, I like the following statistics blog/website: http://fivethirtyeight.com/

Finally, we will be analyzing several datasets in this class. Each of these is available in the “Files” link on the Canvas site for this class.
Assignments and grading

Lab assignments: 42%

Each week, except for weeks 1 and 2, a lab assignment will be due on Tuesday that was handed out the previous Thursday. These are to be done on your own, although you may discuss your work with your classmates. The lab assignments will ask you to conduct a basic analysis of an environmental dataset. During the first portion of each Tuesday class, we will discuss these assignments as a group. Each assignment is worth 6 points (6% of your grade).

Quizzes: 21%

At the beginning of each Thursday class period, except for the first and last ones, you will take a quiz to test your comprehension of the material assigned for that week. Each quiz is worth 3 points.

Class participation: 7%

Class participation will primarily be evaluated based on you demonstrating that you have done the reading prior to coming to class.

Final project: 30%

<table>
<thead>
<tr>
<th>Draft 1:</th>
<th>5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft 2:</td>
<td>5%</td>
</tr>
<tr>
<td>Presentations:</td>
<td>5%</td>
</tr>
<tr>
<td>Final report:</td>
<td>15%</td>
</tr>
</tbody>
</table>

There will be a final project in which you will work to analyze your own datasets. Group sizes may range from 1 to 3 (larger groups will be expected to do more). Group members may self-select based on common research interests (e.g. energy policy, climate change, agricultural development) that we will discuss during the first week. There are several mandatory meetings and sub-assignments, which are indicated on the schedule below. My expectations for the two drafts are as follows:

**Draft 1:**

- A one-page description of the research question that you are intending to address and the dataset that you will use to address this question.
- A list of relevant hypotheses that you are thinking of testing in your analysis.
- A copy of the dataset itself.
- An in-class presentation on the status of your project

**Draft 2:**

- A one-page description of the analytical procedures that you are planning on conducting in order to answer your research question, including a list of the variables involved in each procedure, and the hypotheses being tested by each procedure.
- A copy of your dataset (particularly if this has changed since the first draft)
- An in-class presentation on the status of your project

**Final report:** A write-up of your analysis and results based on a template that will be handed out later in the term.
To conduct this project you’ll need to find and explore your own data. This website offers a variety of ways to explore and visualize datasets that are relevant to this course: http://www.google.com/publicdata/

In addition, the following links are helpful sources of public social-environmental data (NOTE: an increasing amount of socio-environmental data is explicitly spatial and analyzable via a geographic information systems, or GIS, software package. If you are serious about analyzing environmental data I would strongly suggest you take a GIS course to complement this course):

CCAFS Dataverse: https://thedata.harvard.edu/dvn/dv/CCAFSbaseline

CGIAR agricultural trials database: http://www.agtrials.org/

Earth Policy Institute: http://www.earth-policy.org/data_center/

LTER network: https://portal.lternet.edu/nis/home.jsp


U.S. Environmental Protection Agency (EPA): http://www.epa.gov/datafinder/


New Hampshire socio-environmental data: http://des.nh.gov/onestop/
http://www.nh.gov/epht/ehdin/

United Nations data: http://data.un.org/


Vermont socio-environmental data: http://www.anr.state.vt.us/site/html/maps.htm
http://healthvermont.gov/tracking/metadata.aspx

World Resources Institute: http://www.wri.org/resources/data_sets
http://cait2.wri.org/

Course policies

**Academic honor:** The Dartmouth Academic Honor Principle applies in this class (see [http://www.dartmouth.edu/~uja/honor/](http://www.dartmouth.edu/~uja/honor/)). Students are expected to conduct their own work for the individual quizzes and lab assignments.

**Student Needs:** Students with disabilities enrolled in this course and who may need disability-related classroom accommodations are encouraged to make an appointment to see me before the end of the second week of the term. All discussions will remain confidential, although the Student Accessibility Services office may be consulted to discuss appropriate implementation of any accommodation requested.

**Religious observances:** I realize that some students may wish to take part in religious observances that occur during this academic term. Should a religious observance conflict with your participation in the course, please come speak with me before the end of the second week of the term to discuss appropriate accommodations.
## Course schedule:

<table>
<thead>
<tr>
<th>Date</th>
<th>Reading</th>
<th>Topic and activity</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 6</td>
<td></td>
<td>Introduction to statistics and to Stata</td>
<td></td>
</tr>
<tr>
<td>Jan 8</td>
<td>Ch. 1 &amp; 3</td>
<td>Graphical methods of summarizing data</td>
<td></td>
</tr>
<tr>
<td>Jan 13</td>
<td>Ch. 4</td>
<td>Descriptive statistics</td>
<td></td>
</tr>
<tr>
<td>Jan 15</td>
<td>Ch. 5</td>
<td>Bivariate data analysis</td>
<td>Quiz 1</td>
</tr>
<tr>
<td>Jan 20</td>
<td>Ch. 6</td>
<td>Probability</td>
<td>Lab 1: summarizing data</td>
</tr>
<tr>
<td><strong>Jan 22</strong></td>
<td><strong>NO CLASS</strong></td>
<td><strong>NO CLASS</strong></td>
<td><strong>Quiz 2</strong></td>
</tr>
<tr>
<td>Jan 27</td>
<td>Ch. 7</td>
<td>Population distributions</td>
<td>Lab 2: bivariate analysis</td>
</tr>
<tr>
<td>Jan 29</td>
<td>Ch. 8</td>
<td>Sampling distributions, central limit theorem</td>
<td>Quiz 3</td>
</tr>
<tr>
<td>Feb 3</td>
<td>Ch. 9</td>
<td>Point estimation and confidence intervals</td>
<td>Lab 3: distributions</td>
</tr>
<tr>
<td>Feb 5</td>
<td>Ch. 10</td>
<td>Hypothesis testing</td>
<td>Quiz 4, <strong>Project 1st draft</strong></td>
</tr>
<tr>
<td>Feb 10</td>
<td>Ch. 11</td>
<td>Hypothesis testing to compare groups</td>
<td>Lab 4: hypothesis testing</td>
</tr>
<tr>
<td>Feb 12</td>
<td>Ch. 12</td>
<td>Analysis of categorical data</td>
<td>Quiz 5</td>
</tr>
<tr>
<td>Feb 17</td>
<td>Ch. 13</td>
<td>Simple linear regression</td>
<td>Lab 5: Categorical analysis</td>
</tr>
<tr>
<td>Feb 19</td>
<td>Ch. 14</td>
<td>Multiple regression</td>
<td>Quiz 6</td>
</tr>
<tr>
<td>Feb 24</td>
<td>Ch. 15</td>
<td>Analysis of Variance</td>
<td>Lab 6: regression analysis</td>
</tr>
<tr>
<td>Feb 26</td>
<td>TBD</td>
<td>Other topics: cluster analysis, time series</td>
<td>Quiz 7, <strong>Project 2nd draft</strong></td>
</tr>
<tr>
<td>Mar 3</td>
<td>None</td>
<td>In-class work on final projects</td>
<td>Lab 7: ANOVA, TBD</td>
</tr>
<tr>
<td>Mar 5</td>
<td>None</td>
<td>In-class work on final projects</td>
<td></td>
</tr>
<tr>
<td>Mar 10</td>
<td>None</td>
<td>Presentations of final projects</td>
<td></td>
</tr>
<tr>
<td>Mar 17</td>
<td>None</td>
<td>No class</td>
<td>Final reports due</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>