Course Title: Introduction to Data Science

Instructor: Julian Chan

Instructor email: julianchan@weber.edu

Course Number: Honors 2920

Credit Hours: 3

Course Overview: There are a lot of issues that relate to sustainability such as global warming, gender discrimination, education, income, pollution and mortality, but a lot of these issues leave us asking the question - to what extent are these claims true and by how much. You will learn data science methods used to answer these questions, and analyze the data for yourself.

Objectives and Outcomes

<table>
<thead>
<tr>
<th>Learning outcome</th>
<th>Course Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Demonstrate the computational foundations of data science and implement these programming concepts in python.</td>
<td>Ch: 1, 3-8</td>
</tr>
<tr>
<td>2) Demonstrate statistical knowledge and use tools such as regression, classification, hypothesis testing, comparing two populations, confidence intervals with applications in sustainability.</td>
<td>Ch: 9-17</td>
</tr>
<tr>
<td>3) Practice clear and compelling written and/or creative expression. This outcome requires students to explore economic systems or in relation to sustainability and data science in the form of a written assignment.</td>
<td>Written assignment related to a data set on sustainability with the studies described below. This written assignment requires students to use a significant number of course coverage, and also learning outcome 6.</td>
</tr>
<tr>
<td>4) Engage in critical thinking that is open-minded, objective, and as free as possible from prejudice and presupposition.</td>
<td>Ch: 2, 11-12, 15-17</td>
</tr>
<tr>
<td>5) Undertake the comprehension of abstract arguments and the ability to move between the general and the particular.</td>
<td>Ch: 2,3, 15-17</td>
</tr>
<tr>
<td>6) Encounter a variety of human experience, exploring both its universality and its diversity</td>
<td>This learning outcome includes both a data set (described below) and a written assignment. It explores a study (data set) in relation to the human experience in and sustainability.</td>
</tr>
</tbody>
</table>

Studies and data sets (related to sustainability and learning outcomes 3, 5 & 6):

<table>
<thead>
<tr>
<th>Study/data set</th>
<th>Description / study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex discrimination and income.</td>
<td>The data are the beginning salaries for all 32 male and all 61 female skilled, entry–level clerical employees hired by a bank between 1969 and 1977.</td>
</tr>
<tr>
<td>Education and future income</td>
<td>The data are incomes in U.S. dollars for 1,020 working Americans who had 12 years of education and 406 working Americans who had 16 years of education, in 2005.</td>
</tr>
</tbody>
</table>

Utah climate change data. Data collected by the Utah weather stations. This is used to answer the question to what extent has climate changed occurred, and how sure of our conclusions can we make?

Required Materials

- Access to a computer that can run the free software Python (and Jupyter notebook).

Course Coverage: The actual examples covered and specific models that result will depend upon the text and the instructor. The course will cover roughly the following materials depending on the class needs.

Chapters

- 1. Data Science
- 2. Causality and Experiments
- 3. Programming in Python
- 4. Data Types
- 5. Sequences
- 6. Tables
- 7. Visualization
- 8. Functions and Tables
- 9. Randomness
- 10. Sampling and Empirical Distributions
- 11. Testing Hypotheses
- 12. Comparing Two Samples
- 13. Estimation
- 14. Why the Mean Matters
- 15. Prediction
- 16. Inference for Regression
- 17. Classification

Assignments

- Homework (25%)
- Written Statistical Reports (30% - 10% each)
- Exams I & II (20% or 10% each)
- Final Exam (25%)

Students with Disabilities:

Any student requiring accommodations or services due to a disability should contact Services for Students with Disability (SSD) in room 181 of the Student Services Center or contact their office by phone (801) 626-6413 or online at http://www.weber.edu/ssd. SSD can also arrange to provide course materials (including the syllabus) in alternative format if necessary.

Extended Campus Closure Plans:
In the event of an extended campus closure due to natural disaster, epidemic, or other event, I will continue to provide instruction via the Canvas learning management system. I will be in email contact with further details if such an event occurs.