

## Econ 8330: Data Analysis from Scratch

Dr. Ben Smith

*What I cannot create, I do not understand*

Richard Feynman

This fall semester, we are offering the course Data Analytics from Scratch on Wednesday nights (starting at 6PM). This course is the most advanced data analytics course in the College of Business and very likely the university. As a student, you might ask yourself “why would I want to put myself in that kind of pain?” While the class is difficult, it offers advanced applied statistical and computer programming skills that the business community desperately needs. These skills allow companies to better understand and improve the performance of marketing campaigns, forecast sales and inventory, identify outlier events (e.g. detect fraud), and answer numerous other business questions involving data. By taking this course, you acquire skills that few others have but are necessary to properly analyze many types of data.

This course has the following learning objectives:

1. Students demonstrate an ability to programmatically clean, merge, extract, and manage data from raw files and databases
2. Students programmatically build and apply parametric regression estimators using established statistical theory and programmatic optimization techniques
3. Students programmatically visualize data using built-in libraries and non-parametric procedures
4. Students programmatically develop non-parametric procedures, including kernels, re-sampling methods and rank-based tests
5. Students write custom simulation code to derive the null hypothesis distribution in the absence of knowable distributional properties
6. Students write and apply elementary machine learning techniques including regression trees with aggregation methods

This list of learning objectives was developed in response to needs expressed by the business community. While these techniques sound scary, they are presented in a real world context (and there is no expectation that you know what they mean going into the course). Each class involves a lecture followed by a lab. The steps to write the procedure (in plain English) are laid out at the beginning of lab and the answer key is available if you get stuck. This class is difficult, but with effort it is completely doable.

As a result, **all** of the students who have successfully finished the course have reported using the skills learned in this class in their job and most have been promoted. Moreover, given the course’s focus on deep learning, students who successfully complete the class can easily augment their statistical and programming knowledge when confronted with new statistical situations.

The class culminates with a project for the UNO Athletics department and presentation; this class performs data analytics consulting for the Athletics department where each student will answer a real (and different) data-based question currently facing the Athletics department. There will be a poster presentation in December where you can display your results and show employers the advanced skills you could bring to their organization.