

Ec 101: Selected Topics in Economics: Business Analytics

Instructor: Alice Lin (Caltech CNS PhD, 2012)

Supervising faculty: Colin Camerer

Time: Tuesday 7-830pm, Fall 2012 (1.5-0-3.5 units)

Course intro

This course challenges students to examine and practice all parts of the data science pipeline for use in business decision making: problem specification, data munging [format synchronization], data exploration, statistical modeling, and presentation. This course requires basic knowledge of a scripting language such as R. Students should also have a basic familiarity with elementary probability (prerequisite Math 2A) and be comfortable with basic data manipulation.

By the end of the class students will be able to deconstruct a broad problem into logical units and pick a strategy for analysis based on the structure of the data. Students will learn techniques for gathering, transforming, and cleaning data from disparate sources while building a toolkit of analytical techniques. They will also gain an understanding of how to construct models and test the validity of results. Lastly the course will cover how to present both approach taken and results clearly.

Course Syllabus

Week 1: Traditional business analytics

- What questions are businesses asking?
- How does data inform business?
- How were questions answered?
- What is the monetary value of data?

Week 2: Big data

- What is big data and data science?
- How does big data impact business?
- What questions can we now answer with big data?

Assignment: Case study of 5 “Big Data” Companies. For each, answer: What makes them valuable? What technology enables? What is their business model?

Week 3: Overview of data lifecycle

- Cleaning
- Analytics
- Visualization

Week 4: Statistics

- Linear regression
- Logistic regression

Week 5: Dimensionality reduction

- PCA

Week 6: Classifiers

- Nearest neighbor, k-Means clustering
- Random Forest
- Naïve Bayes, SVM
- Model Validation
- LASSO, elastic nets

Week 7: Natural Language Processing

- Latent Semantic Analysis
- LDA

Week 8: Reinforcement Learning

Week 9: Data visualization and storytelling

- What is the important information to display?
- How can color, movement, and dimensions convey information?
- Examples (e.g. visual.ly)

Week 10: Final Project Presentations