Data Analytics @ UNC

Vinayak Deshpande
New MBA elective @UNC
MBA706, Data Analytics: Tools and Opportunities

Instructor: Adam Mersereau

Course Goals:
Data Analytics: Tools and Opportunities" prepares students to lead data-driven organizations. Emphasizing fundamental data analysis concepts such as visualization, data quality, and model validation, the course will introduce students to key methodological tools including linear and logistic regression, clustering, tree-based models, and nearest neighbor prediction through hands-on work with data and software. Through case discussions and a guest speaker presentation, students will critically evaluate strategic opportunities and management challenges that arise with data-driven business models.
What does the course cover?

1. Data Mining Tools

2. Key Data Analytics / Data Mining Concepts
   - Visualization
   - Data preparation
   - Feature selection
   - Model validation
   - Overfitting
   - Experimentation
   - Unstructured data

3. Data Analytics Strategy
   - The business case for data analytics
   - Managing analytics capabilities
Class sessions

1. Moneyball, overview
2. Data visualization
3. Data cleaning and transformation
4. Predictive modeling & linear regression
5. Classification & logistic regression
6. Tree models
7. Neural nets, k nearest neighbors
8. Cluster analysis
9. Unstructured data and text mining
10. Big Data
11. Managing with analytics
12. Experimentation
13. Term project presentations
14. Term project presentations
Can't you just tell me about Data Analytics? Why do I have to do it?

It is important to understand Data Analytics even if you never intend to do it yourself.

But...

you can't really understand Data Analytics without rolling up your sleeves and trying it yourself.
Software: SAS JMP Pro

Why JMP Pro?

- I want to teach you something new and real."
- It is robust and well-supported.
- It does (pretty much) everything we'll need.
- It is free through UNC.
- It runs on Windows and Mac.
Textbook

Data Mining for Business Intelligence by Shmueli, Patel, and Bruce (SPB), 2nd edition.

• Solid but terse discussions of data mining models.
• Useful as a reference.
• Sparse on business insights, strategy.

Data Science for Business by Provost and Fawcett

• Has better discussion of business insights & strategy
Cases and Projects

• Managing with Analytics at Procter & Gamble, Harvard case 9-613-045.
• Capital One Financial Corporation, Harvard case 5-704-467.
• Initial Shipment Decisions for New Products at Zara by Gallien, Mersereau, Garro, Dapena Mora and Novoa Vidal.
• A Step-by-Step Guide to Smart Business Experiments, by Anderson & Simester.

Student Projects: Data analytics applications, challenges, and solutions

Data for projects:
• UNC Kenan-Flagler Donor data
• Data provided by a local company
• Data available from Textbook
## Other Data Science initiatives at UNC

- **Data Science Short Courses**
  - **Introduction to Data Science**
  - **Large-scale Data Networks**
  - **Introduction to Machine Learning**
  - **Managing Big Data**
  - **Hadoop for Huge Data Sets**
  - **Introduction to Data Visualization**
  - **Social Network Analysis:**
    - **Description and Inference**
    - **Data Studies Using SAS**
    - **Predictive Analysis**

[http://www.odum.unc.edu/odum/contentSubpage.jsp?nodeid=698](http://www.odum.unc.edu/odum/contentSubpage.jsp?nodeid=698)
RENCI Vision
We strive to be a preeminent national leader in R&D in the data and cyber technologies that power scientific discovery and business innovation. This is the science of cyberinfrastructure. We work to make RENCI a sought-after partner by those working to solve complex, multidisciplinary problems that require advanced and integrated cyberinfrastructure and data technologies to design solutions.

A Focus on Data
Data is the currency of the knowledge economy and 21st century cyberinfrastructure must enable scientists, educators, businesses, and governments to use big data to advance science, solve problems, and ignite economic innovation. We build, test and deploy data technologies for: Medicine and Genomics, Environmental Sciences, Data Management

http://renci.org/about/