IT Teaching Workshop: Teaching in a Flipped Classroom
BU Questrom School
May 1, 2018

Geoffrey Parker
Professor, Thayer School Engineering, Dartmouth College
Director, Master of Engineering Management Program
goeffrey.g.parker@dartmouth.edu
Some history

• Guest speaker debacle

• Removed powerpoint from ops mgmt
  – Same content in handout felt different to students
    • Lots of examples/problems
  – power of the 5 page memo (Tufte, Amazon)
2014 Analytics launch at Tulane

• Only a seminar room available ... impacted design
• 10 ... 15 ... 20 students by week #2. No more room...
• Used Florian Zettelmeyer’s material (1/2 or so in 2014... mostly replaced by 2018).
• Students added the social media API exercises
• Math PhD student built lots of exercises
Professor: Geoffrey Parker
Murdough 327
603-646-9075
goeffrey.g.parker@.dartmouth.edu

Assistant: Holly Bunker
Murdough MEM Space
603-646-3546
holly.a.bunker@dartmouth.edu

Teaching Assistants:
Austin Powell austinpowell2010@gmail.com
Alberto J. Ruiz Alberto.J.Ruiz.TH@dartmouth.edu
Ligin M. Solamen Ligin.M.Solamen.TH@dartmouth.edu
Irwin I. Tendler Irwin.I.Tendler.TH@dartmouth.edu

Class Location & Schedule:
Whittemore Labs, Murdough Center
Monday, Tuesday 10:15 – 11:45am
Assumptions

• Students already know how to code or are willing to learn
  – True at the Undergrad business, MBA, and Engineering Masters level

• Students have already seen linear models and know probability
  – Not always true; so incorporate into learning R

• Students tune out if lectured at ... not always true.
  – Industry expert lectures work pretty well (e.g., Thomson Reuters Labs CTO)
  – General knowledge lectures flop even from brand name organizations

• Projects generate enough value to justify the time
  – Probably true; students show projects in interviews (e.g., McKinsey 2018)
Faculty Role

• “Anjana: I’m a content curator” ... resonates!
• “Analytics Coach; let me Google that for you...”
  – Stack Overflow, Google, etc.
Q3 What prior work or school/project experience have you had that required modeling and/or data analysis?

Data Mining for Business Analytics
Data Analysis
Tableau
Modeling
Data Science
Research
Q4: What is your level of experience with the following?

<table>
<thead>
<tr>
<th>Tool</th>
<th>Strong Skills</th>
<th>Have Used But Not Expert</th>
<th>Some Familiarity</th>
<th>Have Not Used</th>
<th>Total</th>
<th>Weighted Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>3.13%</td>
<td>43.75%</td>
<td>37.50%</td>
<td>15.63%</td>
<td>32</td>
<td>2.66</td>
</tr>
<tr>
<td>Stata</td>
<td>0.00%</td>
<td>6.25%</td>
<td>0.00%</td>
<td>93.75%</td>
<td>30</td>
<td>3.88</td>
</tr>
<tr>
<td>Excel</td>
<td>43.75%</td>
<td>53.13%</td>
<td>3.13%</td>
<td>0.00%</td>
<td>17</td>
<td>1.59</td>
</tr>
<tr>
<td>Tableau</td>
<td>12.50%</td>
<td>3.13%</td>
<td>9.38%</td>
<td>75.00%</td>
<td>24</td>
<td>3.47</td>
</tr>
<tr>
<td>SQL</td>
<td>6.25%</td>
<td>34.38%</td>
<td>9.38%</td>
<td>50.00%</td>
<td>22</td>
<td>3.03</td>
</tr>
<tr>
<td>Python</td>
<td>3.13%</td>
<td>25.00%</td>
<td>18.75%</td>
<td>53.13%</td>
<td>17</td>
<td>3.22</td>
</tr>
<tr>
<td>SPSS/SAS</td>
<td>3.23%</td>
<td>6.45%</td>
<td>12.90%</td>
<td>77.42%</td>
<td>31</td>
<td>3.65</td>
</tr>
<tr>
<td>Matlab</td>
<td>25.00%</td>
<td>37.50%</td>
<td>12.50%</td>
<td>25.00%</td>
<td>32</td>
<td>2.38</td>
</tr>
</tbody>
</table>
Q5: What types of modeling and/or data analysis software are you hoping to learn?

- Learning Tools
- Python
- SQL
- Tableau

Q6: What are your expectations?

- Software
- Insight
- Skills
- Languages
- Data Analysis
- Real World
- Data Analytics
- Proper Methods
- Data Set
- SQL
- Gain
- Actionable
- Projects
Q7: What is the most interesting thing you’ve heard of/learned about data analytics in the past six months??
Course Structure (20 sessions)

- Part 1: Review linear/non linear models, learn R in process
- Part 2: Introduction to experimental design, visualization, social media, machine learning
- Part 3: Class Project Presentations (4 classes)
Course Software Stack

• Excel (legacy but widely used; assume students know)
• R (workhorse—could use Python instead; have tutorials)
• Shiny (R visualization application; library staff teaches)
• SQL (optional lab sessions; strongly encouraged)
• Tableau (optional; good for presentations)
Grading

• Class attendance and participation
• Individual assignments
• Group Project

... no tests ... decided the value add was too low

... But ... hard to assess individual performance

... So ... met with Thayer course designer to review
A Typical Class

• Set up for 10-15 minutes
• In-class exercise for 60 min
• Debrief for 10-15 min
Teaching Logistic Regression: Titanic Exercise

• Effect of age, gender, fare class (or price)
• Split data into training and test sets
• Predict survival for Rose (~96%) and Jack (~17%)
Experiments

• Put back into course - Uber analytics team request
• Cover random assignment to control and experimental groups
• Describe Factorial Designs
  – Main effects and interactions
    • T-tests, regression, plots
• Mississippi Flood of 1927, Dan Ariely experiments
Introduce some ML algorithms

• C4.5 (Decision Trees)
• k-Means (clustering)
• Support Vector Machines (SVM)
• Apriori
• Expectation Maximization (EM)

• PageRank
• AdaBoost
• k-Nearest Neighbors (kNN)
• Naive Bayes
• Classification and Regression Tree (CART)

See http://www.cs.uvm.edu/~icdm/
Machine Learning from Industry

- [https://www.youtube.com/watch?v=gn7rdaYkYc](https://www.youtube.com/watch?v=gn7rdaYkYc)
  - Susan Athey
  - Artificial Intelligence: The Economic and Policy Implications – Keynote
  - <start at 2:00 end around 14:30>

- [https://www.youtube.com/watch?v=9XjCTlQmzk8](https://www.youtube.com/watch?v=9XjCTlQmzk8)
  - Machine Learning Class (Session #1)
  - John Langford - Principal Researcher, Microsoft Research
  - <about the first 5 minutes>
Machine Learning from Industry

• What is supervised versus unsupervised learning?
• What are the industrial organization issues?
• How is innovation changing?
• What are the labor market implications?
Project Examples

• Airbnb – safety / cost screening tool
  – Merged Airbnb plus crime data

• Superbowl ad analysis
  – Python web scrape to bypass Twitter API limits

• Elon Musk impact on stock price
  – Pulled data from Twitter, Guardian news, Morningstar, Quandl, Morningstar, Bloomberg