Introduction to Business Analytics

Harry J. Paarsch
Department of Economics
College of Business Administration
University of Central Florida
Orlando

University of Victoria
20 September 2017
What is Business Analytics?

- Business analytics is an exciting new field of data science informed by computer science—specifically, algorithmics, databases, and programming—as well as numerical methods from applied mathematics.

- Methods from statistics—specifically, statistical learning—are important in business analytics as are methods from applied econometrics and machine learning.

- What makes business analytics different from other fields of data science (such as bioinformatics) is that it makes use of tools from economics, the imperial discipline of business. In particular, models of incomplete information, such as the models of adverse selection (in auctions) or moral hazard (in contract theory) are especially relevant in business analytics.
Business Analytics Process Flow

1. Problem
2. Theory
3. Data
4. Training
5. Validation
6. Testing
7. Conclusions
8. Production
Process of Solving Problem in Business Analytics

1. Theory: in business analytics, you can often use the theory of incomplete information (for example, models of adverse selection or moral hazard) to put structure on the problem.

2. Data: the theory will then also inform you concerning data to gather and to organize data using, for example, a Spreadsheet as well as Python, and SQLite—perhaps even a NoSQL store.

3. Training, Validation, and Testing: theory can also inform you concerning which tools from statistical learning to implement in Python or R when train, validating, and testing empirical specifications.

4. Conclusions: only theory can provide an interpretation of empirical results required to make a business decision.

5. Production: having settled on an empirical specification, you can use C as well as Hadoop, Pig, Hive, and Spark and cloud computing (for example, AWS) to implement solution at scale.
Tools of Business Analytics

Business
Generates a problem from a domain

Economics
Incomplete information
Moral hazard
Adverse selection

Relational Algebra
Spreadsheet: Excel
DBMS: SQLite
NoSQL

Statistical and Machine Learning
Binary prediction: naïve Bayes, SVCs, …
GLMs: linear, logistic, Poisson regression
Survival: censoring & Cox’s PH model
Time series: smoothing, graduation, & ARMA
Ensembles: boosting, …

First-order asymptotics: LLN, …
Resampling: jackknife & bootstrap
Regularization: $L_1$ & $L_2$
Examples: $L_1$—LASSO & $L_2$—ridge regression as well as Hodrick–Prescott filter from graduation

Problem Theory Data Training Validation Testing Conclusions

Conclusions
$\LaTeX$/\BibTeX
Beamer

Testing
ROC Curve
RMSE

Validation
Cross-validation

Training

Production
C, Java
Hadoop, Pig, Hive, Spark
Cloud computing—AWS

Important Tools
Python: convenient, popular scripting language
R: lingua franca of data analysis in data science
Numerical methods: linear algebra, finding zero of a function, unconstrained and constrained optimization, quadrature, simulation
Parts of a Business Analytics Curriculum

1. Economics: multi-person decision theory, particularly under incomplete information;
2. Statistical Learning: theory of estimation and inference;
3. Numerical Methods: solving numerically for estimators using, for example, numerical linear algebra, but other methods, too;
4. Algorithmics: investigating complexity of algorithms to implement estimators and to conduct inference;
5. Databases: relational algebra and structured query languages;
Mathematics Required

First, and foremost, data science in general (so business analytics in particular) is a mathematical science. Trying to become a business analyst without some elementary training in mathematics would be like trying to become a dentist or a doctor without some basic training in biology. Experience has shown that you need to know the following:

1. Set Theory
2. Linear Algebra
3. Differential and Integral Calculus
4. Optimization Theory
5. Numerical Analysis
6. Graph Theory
Economics Required

1. Microeconomic Theory
   i) Deterministic Economic Decision Problems
      (a) Cost Minimization
      (b) Profit Maximization
      (c) Consumer Choice
      (d) Equilibrium
   ii) Theory of Incomplete Information
      (a) Adverse Selection
      (b) Moral Hazard

2. Game Theory
   a) Non-Coöperative Games of Incomplete Information
      i) Auctions
      ii) Hidden Action Model
   b) Mechanism Design
Probability and Statistics Required

1. Probability Theory: knowledge of different parametric models and their properties;
2. Estimation Theory: different estimation (training) strategies;
3. Inference Theory: methods to evaluating sampling variability;
4. Experimental Design: how to conduction A/B testing at scale;
5. Simulation Methods: how to use the computer when analytic methods are to arduous (or impossible) to perform at scale.
Numerical Methods Required

1. Numerical Linear Algebra
2. Find Zeros of Vector-Valued Functions
3. Unconstrained Optimization
4. Constrained Optimization, particularly Convex Optimization
5. Approximation Methods
6. Generating Pseudo Random Numbers
7. Quadrature, Cubature, Monte Carlo Methods
Computer Science Required

1. Algorithmics
2. Relational Algebra
3. Databases—SQL
4. Statistical Software—R
5. Programming—Python and C
6. Distributed and Parallel Computing—Scala, Spark
Ability to Argue and to Write

- Regardless of your focus, in order to be successful in business at any level, you will need to know how to argue and to write effectively.
- Having been successful in an introductory course in logic is very helpful.
- Having been successful in a basic course in scientific writing is essential.
- Having completed an honours (senior) thesis is a good indicator of the ability to argue and to write effectively.
In Summary

- As you can see, even though it could take a while, becoming a business analyst is relatively straightforward.
- Business analytics requires less formal training than some professions—for instance, accounting, law, or medicine, which require formal accreditation as well.
- Currently, however, few formal programs exist anywhere in the world.
- Many who practice business analytics have tried to develop their skills on-the-job.
- This takes time, and can result in considerable heterogeneity in outcomes.
- I have developed a master’s level program focusing on business analytics delivered by the Department of Economics of the College of Business Administration at the University of Central Florida in Orlando.
Program Description

- The program is eleven months long; successful completion will earn you a Master of Science in Economic with focus on business analytics.
- The program is basically free in that each admitted, qualified candidate who is eligible to be a Graduate Teaching Assistant will be given a tuition waiver as well as a stipend.
- The economics department is a congenial place to work.
- The University has good plant and equipment.
- Orlando is a pleasant place in which to live.
First Semester Courses

1. **Microeconomic Theory I**: introduction to deterministic economic decision problems as well as equilibrium;

2. **Mathematical Economics**: how to cast economic decision problems using mathematics and then how to solve them using basic mathematics;

3. **Operations Research**: how to solve decision problems using mathematical programming on a computer;

4. **Introduction to Business Analytics**: how to embed the material from the above courses in a business ecosystem.
Second Semester Courses

1. **Databases**: how to use the relational algebra and a structured query language to organize data;

2. **Econometrics**: how to use econometrics and statistics to train, validate, and test empirical specifications in data science;

3. **Microeconomic Theory II**: introduction to the theory of incomplete information and game theory—in other words, multi-person decision theory;

4. **Behavioral Economics**: investigating deviations from neoclassical theory as well as how to exploit them using methods of business analytics.
The capstone project represents the culminating academic experience of the master’s program.

It provides students with a forum in which to develop, carry out, and write up research of a well-defined problem in business analytics using the tools developed in the program.

Students will be required to pose a relevant, important problem in business analytics; develop the necessary economic theory to provide an interpretation of the empirical specification developed; gather and organize the relevant data; train, validate, and test the empirical specification; and write a report in which this research and the conclusions are presented in a convincing manner.
The two-course capstone sequence prepares students for the initial assignment that virtually every business analyst gets during the first month on the job: take an ambiguous problem; put interpretable structure on the problem using theory; gather and organize data; train, validate, and test the empirical specification; formulate the conclusions; and write-up the research in a concise, effective way.
Preparation

The following courses would be the ideal preparation for the program:

1. **Computer Science**: Introduction to Computer Science;
2. **Mathematics**: Calculus I, II, and III as well as Linear Algebra;
4. **Economics**: Intermediate Microeconomics, Mathematical Economics, Basic Econometrics, and Game Theory;
## Course Numbers at Different Universities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Colby</td>
<td>CS152</td>
<td>MA121</td>
<td>MA122</td>
<td>MA253</td>
<td>SC212</td>
<td>EC223</td>
<td>EC336</td>
<td>EC293</td>
<td>ECON 379</td>
<td>ECON 420</td>
<td></td>
</tr>
<tr>
<td>McGill</td>
<td>COMP 202</td>
<td>MATH 140</td>
<td>MATH 141</td>
<td>MATH 222</td>
<td>MATH 123</td>
<td>ECON 257D1</td>
<td>ECON 257D2</td>
<td>ECON 230</td>
<td>MATH 125</td>
<td>ECON 337</td>
<td>ECON 420</td>
</tr>
<tr>
<td>Queen's</td>
<td>CISC 101</td>
<td>MATH 123</td>
<td>MATH 124</td>
<td>MATH 112</td>
<td>ECON 250</td>
<td>ECON 310</td>
<td>ECON 255</td>
<td>ECON 351</td>
<td>ECON 455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toronto</td>
<td>CSC108</td>
<td>MAT135</td>
<td>MAT136</td>
<td>MAT237</td>
<td>MAT221</td>
<td>ECO220</td>
<td>ECO206</td>
<td>ECO210</td>
<td>ECO227</td>
<td>ECO316</td>
<td></td>
</tr>
<tr>
<td>Western</td>
<td>CS 1026A/B</td>
<td>MA 0110</td>
<td>MA 1225</td>
<td>MA 1229</td>
<td>EC 2122</td>
<td>EC 2150</td>
<td>EC 2141</td>
<td>EC 2123</td>
<td>EC 2151</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UVic</td>
<td>CSC 110</td>
<td>MATH 100</td>
<td>MATH 101</td>
<td>MATH 200</td>
<td>MATH 211</td>
<td>ECON 245</td>
<td>ECON 246</td>
<td>ECON 313</td>
<td>ECON 350</td>
<td>ECON 365</td>
<td>ECON 450</td>
</tr>
<tr>
<td>UBC</td>
<td>CPSC 103</td>
<td>MATH 105</td>
<td>MATH 106</td>
<td>MATH 200</td>
<td>MATH 221</td>
<td>ECON 325</td>
<td>ECON 326</td>
<td>ECON 303</td>
<td>ECON 420</td>
<td>ECON 425</td>
<td>ECON 421</td>
</tr>
<tr>
<td>SFU</td>
<td>CMPT 102</td>
<td>MATH 157</td>
<td>MATH 158</td>
<td>MATH 232</td>
<td>STAT 270</td>
<td>STAT 285</td>
<td>ECON 302</td>
<td>ECON 331</td>
<td>ECON 435</td>
<td>ECON 431</td>
<td></td>
</tr>
<tr>
<td>UCF</td>
<td>ECO 4443</td>
<td>MAC 2311</td>
<td>MAC 2312</td>
<td>MAC 2313</td>
<td>MAS 3105</td>
<td>STA 3032</td>
<td>STA 2023</td>
<td>ECO 3101</td>
<td>ECO 3410</td>
<td>ECO 4412</td>
<td>ECO 4400</td>
</tr>
</tbody>
</table>