

# Data Science for Strategic Pricing

Instructors:	Jacob LaRiviere, Affiliate Professor & Senior Principal Researcher, Microsoft
Emails:	<a href="mailto:jariv@microsoft.com">jariv@microsoft.com</a>
Office Hours:	Wednesdays 1:30-2:30pm, Microsoft Teams and by appointment
Timing and Location:	Wednesdays 3:30-7:20pm, Microsoft Teams
Course Duration:	April 1 <sup>st</sup> - June 3 <sup>rd</sup>
Final Exam:	Thursday June 11 6:30-8:30pm
Website:	<a href="http://www.jacoblariviere.com/DS_pricing.html">http://www.jacoblariviere.com/DS_pricing.html</a>
Code of Honor:	The University of Washington Code of Honor applies in full to this course. <a href="https://econ.washington.edu/policy-academic-conduct">https://econ.washington.edu/policy-academic-conduct</a>

## Learning Objectives

---

This course is designed to rigorously cover the theory of industrial organization and strategic behavior of firms. We will cover both theoretical and empirical topics and focus on applications in the technology sector. We will develop both theoretical and empirical tools that will serve students in being competitive for quantitatively focused jobs in both tech and other sectors or be competitive graduate students in continuing your education. We will learn the R statistical computing language. At the end of the course you'll be able to perform supervised and unsupervised machine learning techniques in R with special attention paid to pricing and causal inference.

The course is structured to have both lecture and facilitated discussion. We anticipate lots of interaction and idea flows. It is going to be a fun, honest and stimulating classroom environment. This class should be thought of as a mechanism for your improvement as an economist, which increases your ability to think both critically and do interesting and important work that adds value to your endeavors. There will be guest lecturers by a few other economists working on strategic pricing and at the intersection of causal inference and machine learning at Microsoft.

## Course Requirements & Grading

---

Homework	35%
Midterm	30%
Final	35%
Class Participation/Attendance	10% (+/-)

Homework assignments will be both theory and empirical with data sets being provided by the instructors. Homework assignments will have two main themes. The first will be to have students practice the mechanics of the concepts covered in class. The goal of this theme is to cement the comprehension of the

material. The second theme of the homework assignments will be to foster creativity in thinking about the material. They will be graded on a full credit/partial credit/zero credit scale (e.g.: 100%, 50%, and 0%).

In some cases with empirical assignments, we may ask you to supplement the data sets on your own. You are encouraged to use R; any code that provided in solutions will be provided in R. The empirical assignments can span multiple weeks and steady progress is vital to not fall behind. These assignments will be similar to the types of problems you may face working with a firm like Microsoft, Amazon, Google, etc.

Exams will be of standard format: T/F, multiple choice, short answer, graphs and essays.

Since we are only meeting once a week, one can fall way behind by missing just one course. Accordingly, class attendance is mandatory, please email the instructors if you have a legitimate conflict so it does not hurt your participation score. Attendance combined with homework completion, and a thorough review of one's notes and homework assignments should adequately prepare a student for the exams. Note that there will be material on exams that is not covered in the book. **IMPORTANT:** If you will miss a midterm for a verifiable medical/legal/sports reason, notify us immediately to set up a make-up exam. Failure to do so will result in a zero grade for that exam. Unexcused absences for an exam will also result in a zero. If you have a condition that dictates special circumstances for exam taking, please notify us as early as possible.

Appropriate class participation is strongly encouraged. Grades can be increased or lowered by 10% based upon exceptional or unacceptable interaction in classroom, email, or office settings. The grading for this course will be curved. Do not be alarmed if the highest grade on a midterm is 80%; that 80% will earn an A and grades will be assigned from there.

**ADVICE:** You are strongly advised to actively engage in the course. You will find that in economics, a deep understanding of key concepts facilitates high grades. Deep understanding of concepts is most easily attained with class attendance and participation. Coming to class and not paying attention is virtually worthless; given the once a week format of this course, we strongly encourage you to prepare for class to minimize work catching up later on.

## Textbook & Readings

---

*An Introduction to Statistical Learning w/ Applications in R.* James, G., Daniela, W., Hastie, T., Tibshirani, R. Available free and legally from:

<http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Sixth%20Printing.pdf>

*Lectures on Pricing.* McAfee, R. Preston. Available free and legally from:

<http://www.mcafee.cc/Courses/BEM116/PDF/LectureNotes.pdf>

*Reinventing the Bazaar: A Natural History of Markets.* John McMillan, 2003. (recommended)

We will also assign reading from selected academic journal articles and other sources, which will be announced on a rolling basis.

## Course Outline

---

*NOTE: This outline is subject to change at the discretion of the instructors.*

\* means required reading.

Date	Topic	Reading
<b>April 1</b>	Core Pricing Concepts <ol style="list-style-type: none"> <li>Deriving Demand</li> <li>Single good case and optimal price</li> <li>Elasticities and the Lerner formula</li> <li>Selling <math>n</math> goods</li> <li>Intro to value-based pricing</li> <li>Introduction to R</li> </ol>	McAfee Ch. 1 & 2 James et. al. 2.3 McMillan Ch. 1-3 (recommended)
<b>April 8</b>	Value based pricing (price discrimination) <ol style="list-style-type: none"> <li>Product versions</li> <li>Eligibility and direct</li> <li>Bundling</li> <li>Timing</li> <li>Two-part tariffs</li> <li>Demand curve modelling</li> <li>Regression (OLS)</li> </ol>	McAfee Ch. 3 <a href="#">Varian's notes</a> on price discrimination <a href="#">Determinants of Store Level Price Elasticity</a>
<b>April 15</b>	Empirical methods for practical pricing <ol style="list-style-type: none"> <li>Econometric methods</li> <li>Model Fit (MSE)</li> <li>Regression Interpretation</li> <li>Parametric/non-parametric</li> <li>Overfitting</li> <li>Empirics of demand modelling</li> </ol>	Hastie Ch. 1 and 3.1-3.3 <a href="#">Measuring Customer's Reaction to Price</a>
<b>April 22</b>	Empirical methods for practical pricing II <ol style="list-style-type: none"> <li>Causality</li> <li>Bias</li> <li>Machine Learning (ML)</li> <li>Cross-Validation</li> </ol>	Lecture notes will be provided. Hal Varian <a href="#">"Causal Inference in the Social Sciences"</a>
<b>April 29</b>	Data science methods for pricing <ol style="list-style-type: none"> <li>Cross-Validation cont...</li> <li>LASSO/Ridge</li> <li>Supervised/Unsupervised</li> <li>Prediction vs interpretation</li> <li>Logit</li> <li>Logit Demand</li> </ol>	Lecture notes will be provided.
<b>May 6</b>	First half: midterm  Second Half: <ol style="list-style-type: none"> <li>Regression Trees</li> <li>K-means clustering</li> </ol>	Hastie 4.3, 8.1
<b>May 13</b>	Trees and Forests <ol style="list-style-type: none"> <li>CART</li> <li>Cross-Validation</li> <li>Random Forest</li> <li>Bagging</li> <li>Heterogeneous Treatment Effects</li> </ol>	Notes will be provided <a href="#">"Big Data: New Tricks for Econometrics"</a>
<b>May 20</b>	Game Theory & Pricing <ol style="list-style-type: none"> <li>Nash Equilibrium</li> <li>Quality Differentiation</li> <li>Hotelling line</li> </ol>	Notes will be provided

	<ul style="list-style-type: none"> <li>d. Endogenous entry</li> <li>e. Life Time Value</li> <li>f. Customer Inertia</li> </ul>	
<b>May 27</b>	<p>Additional Pricing</p> <ul style="list-style-type: none"> <li>a. Cross-Validation (revisited)</li> <li>b. Premium (quality differentiation)</li> <li>c. Transfer Pricing and double marginalization.</li> <li>d. Intertemporal elasticity.</li> </ul>	Notes will be provided
<b>June 3</b>	<p>Double ML</p> <ul style="list-style-type: none"> <li>a. Causality</li> <li>b. Omitted Variable Bias</li> <li>c. Double ML procedure</li> </ul>	Notes will be provided
<b>June 11</b>	<p>6:30pm-8:20pm</p> <p><b>Note:</b> this is a Thursday.</p>	Final Exam