**Haas School of Business**
**University of California, Berkeley**

# UGBA 147 Advanced Business Analytics
# Syllabus Summer 2021 | 3 units

Welcome to UGBA 147 Advanced Business Analytics.

Successful business analysts, managers, and executives are increasingly relying on data-driven decisions to run their businesses, rather than relying on experience and intuition alone. This course teaches you the latest data analytic methods and decision methods now used by leading-edge business practitioners, going deep to understand their technical inner workings and going broad to realize their practical business applications.

- Business decision modeling, exploratory data analysis, cluster modeling, predictive modeling
- Data analytic methods, including machine learning and related approaches
- Introduction to R and Jupyter software for data analysis
- Real-world/real-data business practicum across a variety of industries:
  stock market, political fundraising, health care products, banking, electric utilities, real estate, hospitality, retail, call centers, telecom, transportation, and others

## Official Course Title & Description

UGBA 147 Special Topics in Operations and Information Technology Management
A variety of topics in manufacturing and information technology with emphasis on current problems and research.

**Class Meetings** Mon, Wed, Fri | 3:40-6:00pm PT | Synchronous Remote
Zoom https://berkeley.zoom.us/j/94099284785 passcode:147147

**Instructor** Richard Huntsinger
Phone: 408.219.6301
Email: rhuntsinger@berkeley.edu
Office: Zoom https://berkeley.zoom.us/j/488973358
Office Hours: Regular schedule (Zoom) – Mon, Wed, Fri 1:30-2:30pm PT
Drop in (phone, Zoom) – anytime as available
Appointment (phone, Zoom) – send text to instructor for appointment

**Assistant** Augustine Santillan
Email: asantillan@berkeley.edu
Office Hours: email assistant for appointment

## Recommended Preparation

UGBA 104 Intro Business Analytics, DATA C100 Principles & Techniques of Data Science, or equivalent

## Course Format & Norms

We meet Mondays, Wednesdays, and Fridays for synchronous remote method lectures and application demonstrations. Attendance is not taken, but you are encouraged to attend all meetings, if it's practical for you to do so. All meetings are recorded and will be available to you shortly afterward.

We use Zoom for our meetings. Join the meetings with your **video on**, if it's practical for you to do so. You are encouraged to speak out with questions or comments during meetings, but keep your audio muted when not speaking. Normally, we will not use Zoom's chat feature during meetings. The course works best when you ask lots of questions.

We use R on Haas JupyterHub to disseminate materials, and to do your lab and project assignments.

We use bCourses > Announcements to disseminate announcements and some other materials. You use bCourses > Assignments to submit links to your lab and project solutions on Haas JupyterHub. You can use bCourses > Discussion to post questions to the instructor and to other students.

## Textbook & Software

You will be provided method lecture materials, application demonstration materials, other study materials, and links to suggested readings. There is no required textbook for this course.

You will be provided access to Haas JupyterHub. Use the Google Chrome web browser directed to https://ugba147-2021-summer.haastech.org. There you will be prompted for your Berkeley email address and CalNet ID to verify your access privileges.

*Note: You are not expected to be already proficient in R and Jupyter; rather, you will be provided instruction and lots of examples on which to base your work. We leverage a variety of R functionalities that can often enable robust analysis with just a few lines of code.*

## Lab Assignments

You will conduct several lab exercises using R on Haas JupyterHub that allow you to explore specific topics covered in lectures. You are provided all the expected outputs for all the labs, but you write the necessary R code. Lab assignments with due dates are listed on bCourses and available on Haas JupyterHub. Each lab submission will normally be graded as full credit if mostly correct, half credit if not all correct but shows good faith effort. You may collaborate with anyone on labs assignments, but you must submit your own solutions. The lab assignments are intended to give you practice with specific important methods.

## Project Assignments

You will conduct several analyses using R on JupyterHub to explore stock market data, construct, evaluate, and tune your own data analytic models, and recommend an investment portfolio. You are provided some of the expected outputs for some of the analyses, but you write the necessary R code. Project assignments with due dates are listed on bCourses and available on Haas JupyterHub. Each

project submission will normally be graded as full credit if mostly correct, half credit if not all correct but shows good faith effort. You may collaborate with anyone on the project assignments, but you must produce and submit your own original solutions. The project assignments are intended to give practice applying the full data-to-decision process.

### Quizzes

You will take 3 online quizzes to demonstrate your working knowledge of business analytics. Quiz assignments with due dates (course end) are listed on bCourses. They will be scoped to take less than 1 hour each, but you will have 3 hours each to take them any time during the course so that you can do so at your convenience. You can use notes, internet resources, and other materials, but you are not allowed to collaborate with anyone on the quizzes.

### Professionalism

You should submit lab and project solutions **by their due dates**. If you are unable to submit them on time, then you should still submit them by course end.

You should submit quiz solutions **by course end**.

Participation is required. You can participate by a combination of class attendance with video on, office hour attendance, posting/answering questions on bCourses > Discussions, and/or posting/commenting on article reviews on bCourse > Discussions.

### Performance Evaluation

Your course grade will be evaluated based on meeting minimum performance standards and your relative performance within the class.

Lab Assignments: 20%

Project Assignments:
- Parts A, B, C, D, E: 15%
- Part F (All Together Now): 15%

Quizzes:
- Quiz 1: 10%
- Quiz 2: 10%
- Quiz 3: 10%

Professionalism:
- All lab and project assignments submitted by course end: 10%
- All lab and project assignments submitted by due dates: 5%
- Participation: 5%

### Instructor Biography [https://haas.berkeley.edu/faculty/richard-huntsinger](https://haas.berkeley.edu/faculty/richard-huntsinger)

## Extenuating Circumstances

Normally, extensions to assignment due dates are accommodated only for medical or other serious emergencies that could not have been mitigated by reasonable time management discipline. Per UC Berkeley and Haas policy, it is the student's responsibility to notify the instructor of such circumstances in advance if possible, or if not possible then as soon as possible thereafter.

## Accommodation for Disabled Students Program (DSP) Students

Official DSP students requesting special accommodations are subject to the following:
To receive an accommodation, you must visit the instructor during office hours (or by appointment, if necessary) in advance of the assignment due date to discuss the logistics of the requested accommodation.

## Accommodation for University-Required Travel

Members of an official university-sponsored organization (for example, an athletic team) that requires travel on the day of a scheduled exam are subject to the following:
To receive an accommodation, you must visit the instructor during office hours (or by appointment, if necessary) in advance of the due date to discuss the logistics of the requested accommodation.

## Requesting Score Corrections

Any request for score corrections must be made within 1 week of the return of the assignment.

## Academic Integrity

Any commission of academic dishonesty, including receiving or giving unauthorized assistance during an exam, will be immediately referred to the campus student conduct committee, with the instructor's recommendation of a final course grade of F and suspension from the university.

For more about academic integrity, see the following:

- https://groups.haas.berkeley.edu/AcademicAffairs/Bylaws/documents/Policy%20on%20Academic%20Dishonesty.pdf
- http://www.haas.berkeley.edu/Undergrad/academic_dishonesty.html
- http://sa.berkeley.edu/conduct/integrity
- http://www.haas.berkeley.edu/Undergrad/academic_guidelines.html

## Safety

See UC Coronavirus Information website at https://news.berkeley.edu/coronavirus/ for latest news and policy announcements.

See the UC Police Department website at http://ucpd.berkeley.edu/crime-prevention-strategies-and-services for crime prevention strategies and services.

If you have not already done so, you are urged to sign up for the "WarnMe" campus alert and warning system. To learn more visit http://warnme.berkeley.edu/.

| Week | Date | Methods & Applications | Lab Due | | Project Due | |
|---|---|---|---|---|---|---|
| 1 | Mon May 24 | Introduction | Mon May 24 | Lab 00 Try to Submit an Assignment | Wed Jun 2 | Project A Exploratory Data Analysis |
| | | Intuition, Statistics, & Data in Business | | | | |
| | | Data-to-Decision Methodology | | | | |
| | | Syllabus & Teaching Plan | | | | |
| | | *Primer for Jupyter* | | | | |
| | | *Primer for R* | | | | |
| | | Data | | | | |
| | | Data Retrieval | | | | |
| | | Data Manipulation | Wed Jun 2 | Lab 01A Data Manipulation | | |
| | | Data Exploration | | | | |
| | Wed May 26 | Descriptive Statistics | Wed Jun 2 | Lab 01B Statistics & Synthetics | | |
| | | Synthetic Variables | Wed Jun 2 | Lab 02 Cross-Tabulation | | |
| | | Cross-Tabulation | Wed Jun 2 | Lab 03 Data Visualization | | |
| | | Data Visualization | | | | |
| | | *Primer for R ggplot2* | | | | |
| | | *Primer for R rgl* | | | | |
| | | Political Campaign Fundraising | | | | |
| | Fri May 28 | Kernel Density Estimation with Probability Density | Wed Jun 2 | Lab 04 Kernel Density Estimation | | |
| | | Iowa Liquor Sales Trends | | | | |
| | | Data Representation | | | Mon Jun 7 | Project B Representation |
| | | Variable Types | | | | |
| | | Dummy Variables | | | | |
| | | Balancing | Fri Jun 4 | Lab 05 Balancing, Imputing, Aligning | | |
| | | Imputing | | | | |
| | | Aligning | | | | |
| 2 | Mon May 31 | HOLIDAY | | | | |
| | Wed Jun 2 | Principal Component Analysis | Mon Jun 7 | Lab 06 Principal Component Analysis | | |
| | | Cluster Analysis | Mon Jun 7 | Lab 07 Cluster Analysis | | |
| | | Cluster Analysis & Evaluation | | | | |
| | | Methodology | | | | |
| | | Model Performance Measures | | | | |
| | Fri Jun 4 | Model Construction Methods | | | | |
| | | Gaussian Mixture Model by Expectation-Maximization | | | | |
| | | Hierarchical Agglomeration | | | | |
| | | k-Means | | | | |
| | | Fortune 500 Workforce Diversity | | | | |
| | | NCHS Customer Segmentation | | | | |

| Week | Date | Methods & Applications | Lab Due | | Project Due | |
|---|---|---|---|---|---|---|
| 3 | Mon Jun 7 | Classification | Mon Jun 14 | Lab 08 Classifier Evaluation | Mon Jun 21 | Project C Classification |
| | | Lending Tree Loan Risk | | | | |
| | | Classification & Evaluation | | | | |
| | | Methodology | | | | |
| | | Confusion Matrix | | | | |
| | | Model Performance Measures: accuracy, etc | | | | |
| | | Business Performance Measures | | | | |
| | | Cross-Validation | | | | |
| | Wed Jun 9 | Model Construction Methods | | | | |
| | | Naïve Bayes | Mon Jun 14 | Lab 09 Naïve Bayes | | |
| | | SATO Customer Churn Intervention | | | | |
| | Fri Jun 11 | Model Tuning for Classifier | | | | |
| | | Forward Variable Selection | | | | |
| | | Backward Variable Selection | | | | |
| | | Hyper-Parameter Selection | | | | |
| | | Model Construction Methods continued | | | | |
| | | Support Vector Machine | Wed Jun 16 | Lab 10 Support Vector Machine | | |
| 4 | Mon Jun 14 | Scania Predictive Maintenance | | | | |
| | | Perceptron | Mon Jun 21 | Lab 11 Neural Network | | |
| | | Neural Network | | | | |
| | | Facial Recognition (maybe) | | | | |
| | Wed Jun 16 | Logistic Regression | | | | |
| | | Decision Tree | | | | |
| | | k Nearest Neighbors | | | | |
| | | Credit Card Fraud Detection | | | | |
| | Fri Jun 18 | Multinomial Classification | Mon Jun 21 | Lab 12 Multinomial Classification | | |
| | | Multinomial Forms | | | | |
| | | One vs Many | | | | |
| | | One vs One | | | | |

| Week | Date | Methods & Applications | Lab Due | | Project Due | |
|---|---|---|---|---|---|---|
| 5 | Mon Jun 21 | Regression | Mon Jun 28 | Lab 13 Regression | Mon Jun 28 | Project D Regression |
| | | Regression & Evaluation | | | | |
| | | Model Performance Measures: RMSE, MAPE, etc | | | Mon Jun 28 | Project E Deployment |
| | | Business Performance Measures | | | | |
| | | Cross-Validation | | | | |
| | | Model Construction Methods | | | | |
| | | Linear Regression | | | | |
| | | More Regression | | | | |
| | Wed Jun 23 | Call Center Scheduling | | | | |
| | | Model Tuning for Regressor | | | | |
| | | Ensemble Methods | Mon Jun 28 | Lab 14 Ensemble Methods | | |
| | | Bagging | | | | |
| | | Boosting | | | | |
| | | Stacking | | | | |
| | Fri Jun 25 | Special Data Types | Wed Jun 30 | Lab 15 Special Data Types | | |
| | | Text Data | | | | |
| | | Hotel Customer Satisfaction | | | | |
| | | Time Series Data | | | | |
| | | Electricity Demand | | | | |
| 6 | Mon Jun 28 | Social Network Data | | | | |
| | | Descriptive Statistics & PageRank | | | | |
| | | Collaborative Filtering | | | | |
| | | Enterprise Software Marketing | | | | |
| | Wed Jun 30 | CATCH-UP DAY | | | Fri Jul 2 | Project F All Together Now |
| | Fri Jul 2 | Project Presentations | | | | |