DATA SCIENCE: BRIDGING PRINCIPLES AND PRACTICE

Program Topics

The Data Science: Bridging Principles and Practice program curriculum covers the following topics:

**Module 1 | Foundations of Data Science**

You'll be introduced to the basic statistical concepts behind data science and analytics and to the main objectives of descriptive, predictive, and prescriptive analytics. Discover different types of data - categorical and numerical data - and the basic ways in which data reveal information. Cover the basics of probability theory that are essential to better understand the ideas behind data generation and interpretation. Includes a basic tutorial on Jupyter Notebooks, an interactive open-source platform we will use for analysis.

**Module 2 | The Science of Surveys for Business Decisions**

We will cover the definitions of key survey terms as well as methods by which sampling is used to analyze the pros and cons of business decisions through the exploration of sampling, type I and type II errors, and control limits. Problem sets will be done using Jupyter Notebooks to explore the properties of samples.

**Module 3 | Hypothesis Testing for Business Decisions**

Making data-driven business decisions relies on well-articulated hypotheses that lend themselves to statistical tests. We will cover the foundations of this approach including statistical comparisons, confidence intervals, and margins of error. Applications of statistical tests in Jupyter Notebook will be part of the self-directed work.

**Module 4 | Extrapolating Information from Sample Data**

We will explore the most common linear and curved patterns and understand different ways to fit data to linear models. A central application will be to understand market demand, price setting, and elasticities. More advanced analyses using Jupyter Notebook will build on previous work you've done.

**Module 5 | Regression Analysis for Business Decisions**

Simple regression analysis and multiple regression models are at the heart of more elaborate data-driven business decision making. We'll focus on understanding the ways in which these models are used, the assumptions that make their use valid, and how to leverage these models to make better business decisions. You will continue building your Jupyter Notebook toolkit using simulated real-world data.

Visit the program page at: https://executive.berkeley.edu/programs/data-science-bridging-principles-and-practice
Module 6 | Forecasting and A/B Testing

Learn the two of the most ubiquitous uses of data science and analytics: forecasting and A/B testing. Specifics include the analysis of variance, time series regressions, and the design and execution of simple and more complex A/B testing procedures. We will go over one of the most media-covered A/B tests done in Silicon Valley and discuss the importance of using regressions to tease out time variation and heterogeneous treatment effects.

Module 7 | Machine Learning

Explore some of the fundamental machine learning methods and how they apply to business decisions. We'll demystify this revered topic and explain when machine learning techniques are most effective, challenge the assumptions about machine learning, and point out important caveats regarding when ML models are not the models you want to use. Using Jupyter Notebooks you will review some off-the-shelf packages and go through examples of models in practice.

Module 8 | Building Effective Data Science Teams

With the fundamentals and some of the most ubiquitous tools behind us, we will wrap with a deep dive into the suite of competencies that define effective data science teams and how to build a data-driven culture in your organization. Common pitfalls will be stressed, and strategies to work effectively with data scientists will be laid out. You will complete a final project on Jupyter Notebooks to solidify your comfort with real-world data.