Berkeley ExecEd

Data Science: Bridging Principles and Practice (Online) – Program Topics

The Data Science program curriculum covers the following topics:

Module 1 | Probabilistic Decision Making

We'll introduce the foundational concepts behind data science and analytics before exploring the fundamentals of data.

- Compare categorical vs. numerical data.
- Explore the basic ways that data reveal information.
- Learn from a healthcare example: HMO membership and doctor visits using aggregated data.
- Become acquainted with Jupyter Notebook, Python, and Panda.

Module 2 | Creating Sample Data

Learn the definitions of key survey terms as well as methods that use sampling to analyze the pros and cons of business decisions through the exploration of sampling, type I and type II errors, and control limits.

- Learn to define types of data samples, sampling variation, and quality.
- Identify and define foundational sampling concepts.
- Identify and mitigate bias when sampling data.
- Evaluate examples that illustrate joint, marginal, and conditional probability: Comcast, Google, and Nextag.

Module 3 | Testing Hypotheses

Making data-driven business decisions relies on well-articulated hypotheses that lend themselves to statistical tests. We'll cover the foundations of this approach, including statistical comparisons, confidence intervals, and margins of error.

- Identify the basic tenets of experimentation.
- Identify and discriminate between one-sided and two-sided statistical tests.
- Complete problem sets using the 4M model (Motivation, Method, Mechanics, and Message).
- Analyze an industry example: 24 Hour Fitness tests a new proprietary diet—testing between control and treatment groups.

Module 4 | Extrapolating Information from Sample Data

We'll explore the most common linear and curved patterns and understand different ways to fit data to linear models. A central application will be understanding market demand, price setting, and elasticities.

- Identify conditions for using and interpreting linear and curved patterns.
- Examine curved (non-linear) patterns as applied to vehicle weight and fuel efficiency.
- Complete problem sets using the 4M model for credit cards, crime, and housing prices in Philadelphia.

Data Science: Bridging Principles and Practice (Online) – Program Topics

The Data Science program curriculum covers the following topics:

Module 5 | Basic Regression Models

Simple regression analyses are at the heart of more elaborate data-driven business decision making. We'll focus on understanding how these models are used, the assumptions that make their use valid, and how to leverage these models to make better business decisions.

- Define and apply the simple regression model and identify conditions for its use.
- Apply and interpret prediction intervals.
- Identify three major problems that affect regression models: changing variation in data, outliers, and dependence among observations.
- Practice with a retail example: use regression modeling to determine the location of a franchise outlet.

Module 6 | Advanced Regression Models

Build on the basics to define the multiple regression model and explore different use cases.

- Discriminate between marginal and partial slopes.
- Articulate inference in the multiple regression model.
- Summarize the process of fitting and building a multiple regression model.
- Learn from a financial example: build a multiple regression model to explain the returns on Sony's stock.
- Practice with a human resources example: analyze salary data using MRM to identify gender imbalances.

Module 7 | Forecasting and Machine Learning

We'll demystify machine learning by mastering the fundamentals and studying different applications.

- Discriminate between supervised, semi-supervised, and unsupervised learning.
- Examine machine learning approaches, including the "bag-of-words" method for supervised learning.
- Practice forecasting by using time series regressions.
- Explore a cybersecurity example: machine learning for spam detection.

Module 8 | A/B Testing & Building Effective Data Science Teams

With the fundamentals and some of the most common tools under our belts, we'll dive deep into the competencies that define effective data science teams and show you how to build a data-driven culture in your organization. We will stress common pitfalls and strategies to work effectively with data scientists.

- Review the requirements for building effective data science teams.
- Continue the exploration of building a data-driven culture.
- Evaluate an advertising example: Rocket Fuel's conversion rate, benefit, ROI, opportunity cost, and A/B testing.