HARVARD
Online

Data Science for Business

Syllabus

Data Science for Business moves beyond the spreadsheet and provides a hands-on approach for demystifying the data science ecosystem and making you a more conscientious consumer of information. Starting with the questions you need to ask when using data for decision-making, this course will help you know when to trust your data and how to interpret the results.

Modules		Case Studies	Takeaways	Key Exercises
Module 1	**The Data Science Shift**	Carvana: Good Data and Bad Buys	• Apply the steps of the Data Driven Decision Framework • Identify the benefits that data science brings to business problems	• Translate business problems into data hypotheses • Explore and describe datasets • Use visualizations to generate hypotheses
Module 2	**Data Wrangling**	Fannie Mae: Identifying Investments	• Relate the quality of data with the the quality of the conclusions • Assess the quality of data • Guide decisions for merging tables and managing missing data	• Prepare and clean data for analysis • Examine data dictionaries • Design table joins • Identify solutions for managing missing data
Module 3	**Visualization**	StockX: Drawing Demand	• Incorporate visualizations throughout the data science process • Interpret charts and graphs • Develop questions from visualizations • Design visualizations for clear communication with maximal impact	• Critique existing charts and identify methods of improvement • Generate insight with graphs • Design visualizations to express data clearly
Module 4	**Time Series Forecasting**	NICU beds: Creating Capacity	• Connect yesterday's data with tomorrow's prediction • Evaluate temporal patterns in data • Match the time scale with the business problem • Select appropriate smoothing techniques for time series forecasting	• Determine when time series analysis is useful and informative • Select appropriate methods for exponential smoothing

Modules	Case Studies	Takeaways	Key Exercises
Module 5 — Linear Regressions	Bark Gift Shop: Motivating Managers ATO Pictures: Marketing Movies	• Interpret linear regression results • Extend intuition into analysis • Apply advanced methods to gain sophistication and insight to your understanding.	• Identify relationships between variables • Write hypotheses • Explain the parts of a linear model, including interactions and dummy variables • Interpret linear regression results
Module 6 — Logistic Regressions and Machine Learning	Carvana and Fannie Mae	• Differentiate linear and logistic regression • Conceptualize Machine Learning • Evaluate model fit	• Complete a confusion matrix • Interpret results from logistic regression, CART, random forest, lasso, and neural networks • Select a model to guide decisions

Learning requirements: In order to earn a Certificate of Completion from Harvard Online and Harvard Business School Online, participants must thoughtfully complete all 6 modules, including satisfactory completion of the associated quizzes, by stated deadlines.