

# Courses Related to Data Science

# Edx.org - Spark Series - UC BerkeleyX

- ALL REQUIRE PYTHON EXPERIENCE - [here python course](#)
- **Introduction to Spark**  
Learn the fundamentals and architecture of Spark, the leading cluster-computing framework among professionals. Length 2 weeks Free Cert \$49
- **Big Data Analysis with Spark**  
Learn how to apply data science techniques using parallel programming in Spark to explore big data. Length 4 weeks Free Cert \$99
- **Distributed Machine Learning with Spark**  
Distributed Machine Learning with Spark  
Learn the underlying principles required to develop scalable machine learning pipelines and gain hands-on experience using Spark. Length 4 weeks Free Cert \$99
- **Advanced Distributed Machine Learning with Spark**  
Learn how to develop and deploy distributed machine learning pipelines and gain the expertise to write efficient, scalable code in Spark. Length 4 weeks Free Cert. \$99

## edx.org - [The Analytics Edge](#) - MIT

- This is a great next level course.
- Covers a number of domains and techniques
- Includes a private (course students) competition
- Very practical techniques even if you have gone through John's Hopkins, you will learn a lot
- R based mostly

# Edx.org - Data Analysis for Life Sciences - HarvardX

- **Data Analysis for Life Sciences 1: Statistics and R**

An introduction to basic statistical concepts and R programming skills necessary for analyzing data in the life sciences. Length 4 weeks Cost Free Cert \$50

- **Data Analysis for Life Sciences 2: Introduction to Linear Models and Matrix Algebra**

Learn to use R programming to apply linear models to analyze data in life sciences. Length 4 weeks Cost Free Cert \$50

- **Data Analysis for Life Sciences 3: Statistical Inference and Modeling for High-throughput Experiments**

A focus on the techniques commonly used to perform statistical inference on high throughput data. Length 4 week Cost Free Cert \$50

- **Data Analysis for Life Sciences 4: High-Dimensional Data Analysis**

A focus on several techniques that are widely used in the analysis of high-dimensional data. Length 4 week Cost Free Cert \$50

# [Edx.org](https://edx.org) - Applications of Linear Algebra - [DavidsonX](#)

- [Applications of Linear Algebra Part 1](#)
- Learn to use linear algebra in computer graphics by making images disappear in an animation or creating a mosaic or fractal and in data mining to measure similarities between movies, songs, or friends. Length 4 weeks Cost Free Cert N/A
- [Applications of Linear Algebra Part 2](#)
- Explore applications of linear algebra in the field of data mining by learning fundamentals of search engines, clustering movies into genres and of computer graphics by posterizing an image. Length 5 weeks Cost Free Cert N/A

# Coursera.org - Andrew Ng Machine Learning

Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human-level AI. In this class, you will learn about the most effective machine learning techniques, and gain practice implementing them and getting them to work for yourself. More importantly, you'll learn about not only the theoretical underpinnings of learning, but also gain the practical know-how needed to quickly and powerfully apply these techniques to new problems. Finally, you'll learn about some of Silicon Valley's best practices in innovation as it pertains to machine learning and AI.

Use Octave to develop algorithms to implement a number of algorithms - this course opens up the world of data science and machine learning to you in a very intimate way.

Week 1: Linear Regression with One Variable,  
Linear Algebra Review

Week 2: Linear Regression with Multiple Variables,  
Octave/Matlab Tutorial

Week 3: Logistic Regression, Regularization

Week 4: Neural Networks: Representation

Week 5: Neural Networks: Learning

Week 6: Advice for Applying Machine Learning, Machine Learning System Design

Week 7: Support Vector Machines

Week 8: Unsupervised Learning, Dimensionality Reduction

Week 9: Anomaly Detection, Recommender Systems

Week 10: Large Scale Machine Learning

Week 11: Application Example: Photo OCR

# Coursera.org - Johns Hopkins Data Specialization

This specialization includes 9 courses and 1 Capstone Project Coursework is using R

Purchase Specialization · ~~\$470 USD~~ **\$423 USD** 10% off,

- [The Data Scientist's Toolbox](#)  
Current session: May 9 — Jun 13.
- [R Programming](#)  
Current session: May 9 — Jun 13
- [Getting and Cleaning Data](#)  
Current session: May 9 — Jun 13
- [Exploratory Data Analysis](#)  
Current session: May 9 — Jun 13
- [Reproducible Research](#)  
Current session: May 9 — Jun 13
- [Statistical Inference](#)  
Current session: May 9 — Jun 13
- [Regression Models](#)  
Current session: May 9 — Jun 13
- [Practical Machine Learning](#)  
Current session: May 9 — Jun 13
- [Developing Data Products](#)  
Current session: May 9 — Jun 13
- [Data Science Capstone](#)  
Upcoming session: May 30 — Jul 25

Coursera.org - [Mastering Statistics in R](#)



<https://www.datacamp.com/>

R and Python Learning

# Coursera.org - [Managing Big Data with Mysql](#)

Great intro to Python with Big Data

# Kaggle.com - [Tutorials](#)

Great resources for getting started with Data Science