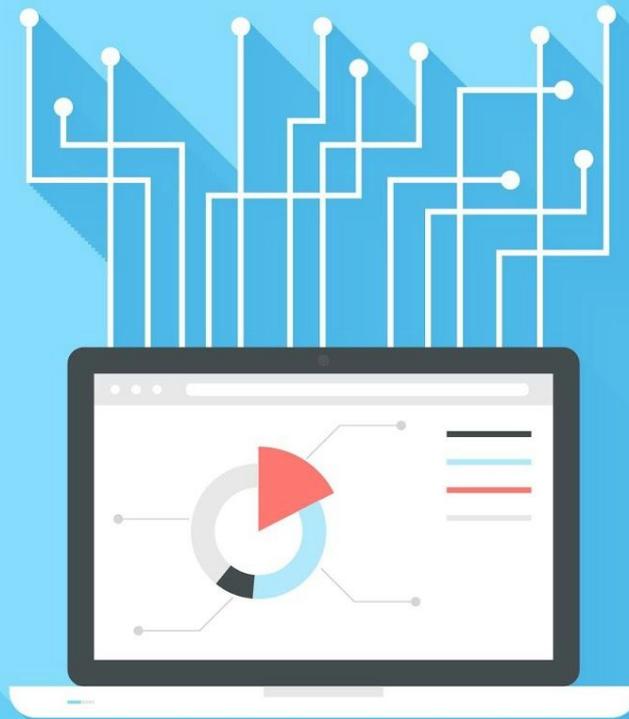


sliderule
your online learning hub



FOUNDATIONS OF DATA SCIENCE

Syllabus

Introduction

Data Science is one of the fastest growing fields of this decade. There is an explosion of data fuelled by cheap and ubiquitous storage of everything from personal and health records, every single action on millions of websites, mobiles, sensors, business transactions and so on. We now need technologies that help us make sense of this data, and become more intelligent in our decisions. That is the mandate of the field of Data Science.

Learning Data Science is perhaps the best career investment you can make in 2015. According to [LinkedIn](#), Statistical Analysis & Data Mining were the hottest skills that got recruiters' attention last year.

SlideRule's [Foundations of Data Science](#) is an introductory workshop for learning Data Science. The objective of the course is to emphasize learning by doing and help you build a portfolio as you learn.

How It Works

When you enroll in this course,

- You will get a high-quality online curriculum, which contains the best Data Science resources handpicked from the web and structured into a logical curriculum by a Data Science expert.
- A Student Advisor will talk with you about your learning objectives and preferences for a mentor, and find the best 1-1 mentor match accordingly. They will support you throughout the course.
- As you work through the curriculum and projects at your own pace, you will have 1-1 video calls with your mentor every week. You can ask them questions about the curriculum, get project feedback or career advice.
- In addition to your weekly mentor calls, you'll have access to weekly live Office Hours and an exclusive community of mentors and students for discussions with the broader community.
- Once you finish the course material and submit the Capstone Project, you will get a Certificate of Completion.

Syllabus

Each module will cover a key aspect of Data Science and have a combination of materials: lectures, theory, coding exercises, reading/viewing exercises, and optional materials. The recommended time allocation is based on a total of 100 hours of work, and can be scaled according to student needs.

Probability & Statistics (14 hours)

Probability & Statistics is at the heart of data science. This section will give you the required foundation in these topics.

Topics Covered:

1. Random variables and distributions
2. Statistical studies
3. Descriptive statistics
4. Dependent & independent events
5. Regression
6. Inferential statistics

Basics of R Programming (4 hours)

Get started with the R statistical software - its many libraries & packages make it a great tool for beginners.

Topics Covered:

1. Installing R and RStudio
2. Shortcuts, common commands and syntax quirks
3. Basic data visualization with the ggplot2 package

Exploratory Data Analysis (20 hours)

Learn how to form an intuition about a data set before getting into any formal methods through Exploratory Data Analysis (EDA) and basic visualizations.

Topics Covered:

1. EDA vs classical & Bayesian approaches
2. Histograms and frequency polygons
3. Box-plots, quartiles, scatter plots, heat maps etc.

Data Visualization (8 hours)

Data Visualization is a vast field, and can be a specialization unto itself. This elective covers the fundamentals of visual perception and good design, and some powerful visualization techniques.

Topics Covered:

1. Visual perception
2. Graphic design
3. Advanced visualization techniques like index charts, horizon graphs, parallel coordinates, maps, hierarchies, networks etc.

Data Wrangling (11 hours)

Learn Data Wrangling - the process of converting data from a raw form into another format that allows for more convenient analysis.

Topics Covered:

1. The split-apply-combine paradigm
2. Web-scraping and APIs
3. R-based tools and packages for data wrangling - reshape2, plyr, dplyr, Rvest etc.

Analytics Techniques (45+ hours)

Learn how to apply analytics to real-world applications from examples like Moneyball, eHarmony, Twitter, IBM Watson, and Netflix.

Topics Covered:

1. Linear & logistic regression
2. Trees
3. Clustering
4. Text analytics

Capstone Project (25 hours)

Now it's time to bring it all together. Pick a data set in your area of interest, tidy it up, explore and summarize it to form an intuition and then use formal methods to build a model and uncover insights!

Looks good? [**Enroll Here**](#)

or email us at founders@mysliderule.com with any questions.

sliderule

www.mysliderule.com