

SYLLABUS

On Campus

Data

Science



Table of Contents

3	Program Overview
4	Why Data Science?
5	Curriculum Overview
6	- Module 1: Introduction to Data with Python and SQL
7	- Module 2: Statistics, AB Testing and Linear Regression
8	- Module 3: Machine Learning
9	- Module 4: Big Data, Deep Learning and Natural Language Processing
10	- Module 5: Final Project: Data Science Advanced Project
11	Contact Us

Overview

FLATIRON SCHOOL'S DATA SCIENCE PROGRAM

This program will provide you with the knowledge, skills, and experience to get a job in the field of Data Science – which requires a mix of software engineering, statistical understanding, and the ability to apply both skills in new and challenging domains.

Over 15 challenging weeks at Flatiron School, you'll learn how to gather data, apply statistical analysis to answer questions with that data, and to make your insights and information as actionable as possible. Our pedagogy ensures not only job readiness for today's market, but the aptitude and skills to keep learning and stay relevant in the industry in the years ahead.

What will you learn?

- How to retrieve data from outside sources and organize data using Python
- Create beautiful visualizations to present key findings
- Explore data and write down multiple hypotheses for further analysis of the data
- How to perform A/B tests
- Build machine learning API that outputs results of an analysis
- Apply and use Big Data
- Learn presentation techniques to better share conclusions about approach and analysis to key stakeholders

When and where does the course meet?

- Classes meet 5 days a week, full-time; Monday-Friday from 9 a.m to 6 p.m
- Classes will typically begin at 9 a.m., with a 60-minute break for lunch around 12:30 p.m., followed by continued lectures and lab work on-campus through the late afternoon

Why Data Science?

WHY IS THIS COURSE RELEVANT?

More than ever before, industries are capturing data on a variety of topics, behaviors, and trends. Without data science, this information stays stuck, without a story to tell or insights to share. In order to determine business goals, more and more companies are looking to data scientists to fill in the gaps and find opportunities never before considered.

Over the last four years, the rise of job opportunities in the field of Data Science has increased tremendously.



Note: The chart above offers a 7-day rolling mean of all Indeed job posts that featured "data science" or "data scientist" in the title across the world as a percentage of all job posts between January 1, 2014 and November 16, 2017. The data was pulled using Imhotep, Indeed's open source analytics platform.

As this area of expertise has grown, the positions within the field have become more nuanced. After completing our Data Science program, you'll not only have the skills to be able to secure a job as a Data Scientist, but can also consider pursuing other data-focused positions like:

- Data Analyst
- Data (Science) Consultant
- Business Analyst
- Data Engineer

Curriculum Overview



From Python to Machine Learning, our 15-week data science program gives you the breadth and depth of knowledge needed to become a well-rounded data scientist. You'll also leave with an understanding of how to discover new techniques as your career progresses.

Every 3 weeks, you'll be introduced to a new module that builds off prior learnings while allowing you enough time to dive into each area for a thorough understanding of the subject matter.

<i>Module</i>	<i>Clock Hours</i>	<i>Module Length</i>	<i>Projects</i>
Module 1: Introduction to Data with Python and SQL	105	3 weeks	Mod 1 Project
Module 2: Statistics, AB Testing and Linear Regression	105	3 weeks	Mod 2 Project
Module 3: Machine Learning	105	3 weeks	Mod 3 Project
Module 4: Big Data, Deep Learning and Natural Language Processing	105	3 weeks	Mod 4 Project
Module 5 : Data Science Advanced Project	105	3 weeks	Final Project

Program Total: 525 hours, projects, and homework

Getting Started

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

The data science program moves quickly and our passionate students embrace that challenge. To make sure you will be successful and able to keep pace with the class, we require you to demonstrate some data science knowledge prior to getting admitted. To help you get up to speed, we provide a free introductory Bootcamp Prep course.

MODULE 1

Introduction to Data with Python and SQL

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

Our first module introduces the fundamentals of Python for data science. You'll learn basic Python programming, how to use Jupyter Notebooks, and will be familiarized with popular Python libraries that are used in data science, such as Pandas and NumPy. Additionally, you'll learn how to use Git and Github as a collaborative version control tool. To organize your data, you'll learn about data structures, relational databases, ways to retrieve data, and the fundamentals of SQL for data querying for structured databases. Furthermore, you'll learn how to access data from various sources using APIs, as well as perform Web Scraping. Finally, we'll conclude with a heavy focus on visualizations as a way to go from data to insights. At the end of this module, students will use their newly learned skills to collect, organize and visualize data, with the goal to provide actionable insights.

What is Covered in Module 1?

- Variables
- Booleans and Conditionals
- Lists
- Dictionaries
- Looping
- Functions
- Data Structures
- Data Cleaning
- Pandas
- NumPy
- Matplotlib/ Seaborn for Data Visualization
- Git/Github
- SQL
- Accessing Data through APIs
- Web Scraping

Statistics, AB Testing and Linear Regression

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
----------------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------

Having learned how to gather and explore data with Python and SQL you can now go deeper into analyzing that information with statistics. In this module, you'll learn about the fundamentals of probability theory, where you will learn about probability principles such as combinations and permutations. You will go on and learn about statistical distributions and how to create samples when distributions are known. By the end of this module, you will be able to apply this knowledge by running AB tests. Additionally, you'll learn how to build your first (and important) data science model: a linear regression model.

What is Covered in Module 2?

- Combinatorics
- Probability Theory
- Statistical Distributions
- Bayes Theorem
- Sampling Methods
- Hypothesis Testing
- AB Testing
- Linear Regression
- Model Evaluation

Machine Learning

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

Module 3 is all about machine learning, with a heavy focus on supervised learning. To start, you'll go a little deeper into regression analysis, learning about extensions to linear regression, and a new form of regression: logistic regression. In building regression models, you'll learn about penalization terms, preventing overfitting through regularization and using cross validation to validate regression model.

Next, you'll learn how to build and implement the most important machine learning techniques. You'll learn about classification algorithms such as Support Vector Machines and Decision Trees. Additionally, you'll learn how to build even more robust classifiers using ensemble methods such as Bagged and Boosted Trees, and Random Forests.

What is Covered in Module 3?

- Linear Algebra
- Logistic Regression
- Maximum Likelihood Estimation
- Cost Function Optimization
- Pipeline Building
- Hyperparameter Tuning
- Grid Search
- Scikit-Learn
- Gradient Descent
- K Nearest Neighbors
- Decision Trees
- Ensemble Methods

MODULE 4

Big Data, Deep Learning and Natural Language Processing

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
----------------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------

After a full module on supervised learning, this module focuses on a variety of advanced Data Science techniques. You'll start with learning about unsupervised learning techniques such as clustering techniques and dimensionality reduction techniques. Next, you'll be introduced to threading and multiprocessing to be able to work with big data. In doing so, you'll learn about PySpark and AWS, and how to use those tools to build a recommendation system. Next, you'll get an in-depth overview of deep learning techniques, learning about densely connected neural networks, enabling high-performing classification performance. You'll learn how to use regular expressions in Python and how to manage string values, analyze text and perform sentiment analysis.

What is Covered in Module 4?

- Dimensionality Reduction
- Clustering
- Time Series Analysis
- Neural Networks
- Big Data
- Natural Language Processing
- Text Vectorization
- Natural Language Toolkit
- Regular Expressions
- Text Classification
- Recommendation Systems

MODULE 5

Final Project

Prework	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

In your final project, you'll work individually to create a large-scale data science and machine learning project. This final project provides an in-depth opportunity for you to demonstrate your learning accomplishments and get a feel for what working on a large-scale data science project is really like.

You will pitch three different ideas and then decide on your final project with your instructors. Instructors advise on projects based on difficulty and feasibility given the course's time constraints.

Upon project completion, you'll know how to construct a project that gathers and builds statistical or machine learning models to deliver insights and communicate findings through data visualisation and storytelling techniques.

Contact Us

For more information, please check out our website at www.flatiron-school.com or contact us at admissions@flatiron-school.com