// FLATIRON SCHOOL

Data Science Syllabus

Module 1: Python for Data Science
Getting started with Data Science
Git, Github and Being Part of a Data Science Team
Importing and Statistical Analysis of Data
Python Libraries: Numpy and Pandas
Data Cleaning in Pandas
Improving your Projects & Results
Linear Relationships & Prediction
Multiple Regression and Model Validation
13 days Mini-Project

Module 2: Data Engineering for Data Science
SQL Databases
Object Orientation
JSON and XML
APIs
HTML, CSS, & Web Scraping w Beautiful Soup
More SQL Practice
No-SQL
10 days

Module 3: Probability, Sampling, and AB Testing
Combinatorics and Probability
Statistical Distributions
Distributions and Sampling
Statistical Power and Anova
In Depth AB Testing
Bayesian Statistics
Resampling and Monte Carlo Simulation
11 days Mini-Project

Module 4: Statistical Modeling
A Complete Data Science Project
Linear Algebra
Calculus, Cost Function and Gradient Descent
Extensions to Linear Models
Introduction to Logistic Regression
In-depth Logistic Regression
Time Series Visualization and Trends
Basic Time Series Models
9 days Mini-Project

Module 5: Machine Learning and Big Data
Ensemble Methods
Support Vector Machines
Principal Component Analysis
Clustering
Pipelines
Operationalizing Code and AWS
Big Data in PySpark
Developing a Recommendation System in PySpark
12 days Mini-Project

Module 6: Deep Learning & NLP
Graph Theory
Foundations of Natural Language Processing
Introduction to Deep Learning
Multi-Layer Perceptrons
Regularization and Optimization
Introduction to Convolutional Neural Networks
Convolutional Neural Networks Continued
Deep NLP - Word Embeddings
7 days

Final Project - 2.5 weeks