Data Science and Curriculum Overview

FLATIRON SCHOOL

Today's Setup

- Linkedin Profile.

- Slack.

- Github account.

- Data science work environment.

- In the future you will want to find a website, podcast or some other means to follow the latest trends in data science.

What to Expect from Flatiron

- We are not here to give a grade of pass/fail.

- Supportive environment.

- Make you the best data scientist you can be!
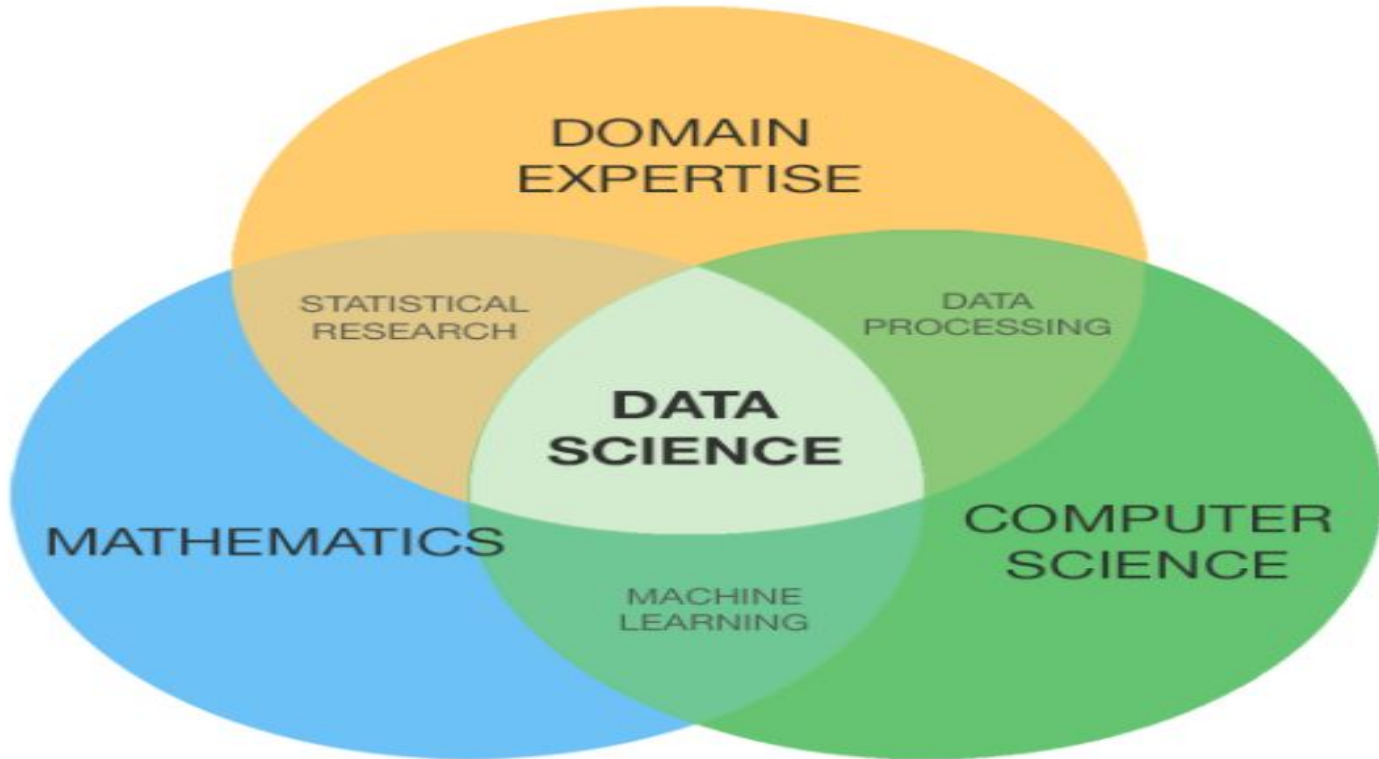
What we expect from you.

- Be respectful of members of your cohort and all Flatiron staff.

- Work hard and focus on self-development.

- Have fun!

- Blogs.

- **Self care.**

  - https://www.webmd.com/sleep-disorders/benefits-sleep-more#1

Why Become a Data Scientist?

What is Data Science

Can I become a Data Scientist ?

DIFFERENT TYPES OF DATA SCIENTIST

Types of Data Scientist

- **Data Analyst/Junior Data Scientist**
  - Junior data scientists doing a lot of number crunching and data cleaning.
  - Usually work on short-term projects.
  - Works with some Python, Tableau, SQL, and Excel.

Types of Data Scientist

- **Data Engineer**
  - Strong software engineering skills.

  - Develops, constructs, tests and maintains architectures, such as databases and large-scale processing systems used by data scientist/analyst.

  - Hadoop, Mapreduce, NoSQL, SQL, Python, Scala, Java, etc.

Types of Data Scientist

- **Statistician**
  - Formal Mathematics and/or Statistics training.

  - Focus on creating models.

  - Design surveys or experiments to collects data.

  - Some programming experience usually in R, Matlab, Mathematica or Python.

Types of Data Scientist

- **Data Scientist**
  - Give insight to the business process using metrics gained from Machine Learning and Statistical Analysis.

  - Capable of deploying apps using cloud based platform.

  - SQL, Spark Proficient with Python.

Types of Data Scientist

- **Machine Learning Engineer**
  - Similar to data scientist, but with a more specific and deeper focus on machine learning, deep learning, and reinforcement learning.
  - Good grasp of Linear Algebra and Calculus.
  - Fluent in Python.

Flatiron Curriculum

Module 1 - Python for Data Science.

Module 2 - Data Engineer for Data Science.

Module 3 - Probability, Sampling and AB

Testing.

Module 4 - Statistical Modeling.

Module 5 - Machine Learning and Big Data.

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.

- 13 Days.

- **Mini-Project.**

Module 2: Data Engineering for Data Science

- SQL Databases.

- Object Orientation.

- JSON and XML.

- APIs.

- HTML, CSS, & Web Scraping with Beautiful Soup.

- More SQL Practice.

- No-SQL.

- 10 Days

Module 3: Probability, Sampling, and A/B Testing

- Combinatorics and Probability.

- Statistical Distributions.

- Distributions and Sampling.

- Statistical Power and Anova.

- In Depth A/B 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

- A topic of your choice.

- Aim for project that you are capable of completing in the given time.

- Demonstrate the skills you have acquired to potential employers.

- Aim to deploy project as a web application.

DON'T
GIVE UP.
GREAT
THINGS
TAKE TIME.

Questions?