

# // FLATIRON SCHOOL

---

Data Science Immersive

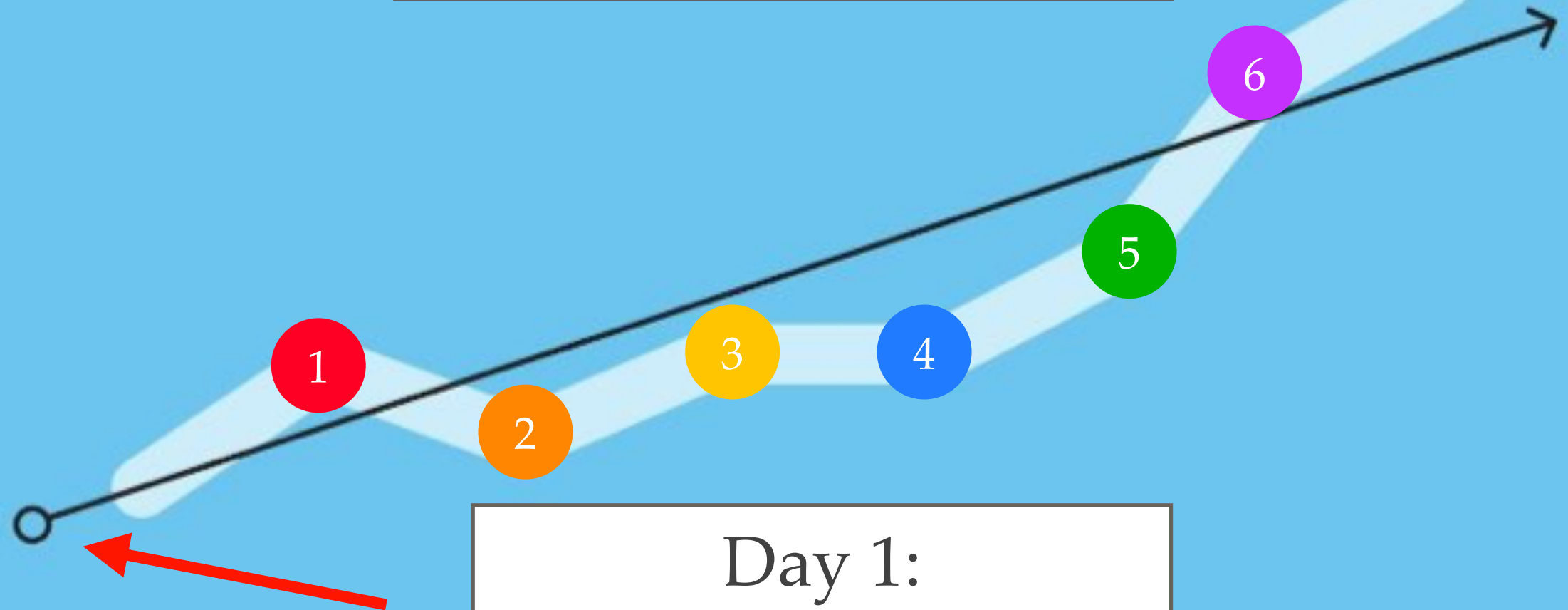
10.28.2019

---

## Welcome!

# Introductions

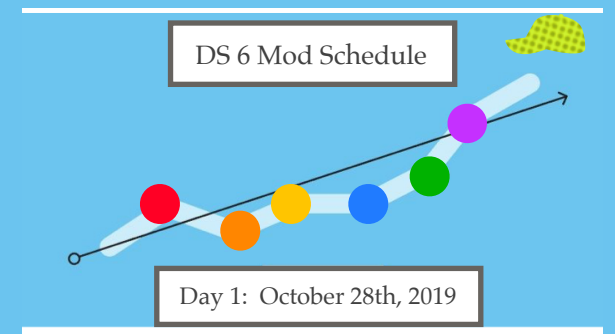
# Data Science 6 Module Structure



Day 1:  
October 28th, 2019

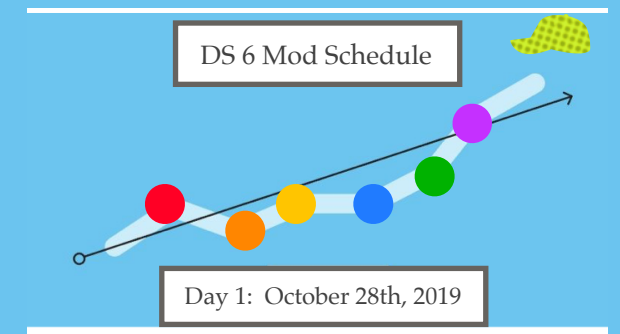
# ● Mod 1: Python for DS

- Terminal
- Python
- Numpy
- Pandas
- Git
- Matplotlib / Seaborn



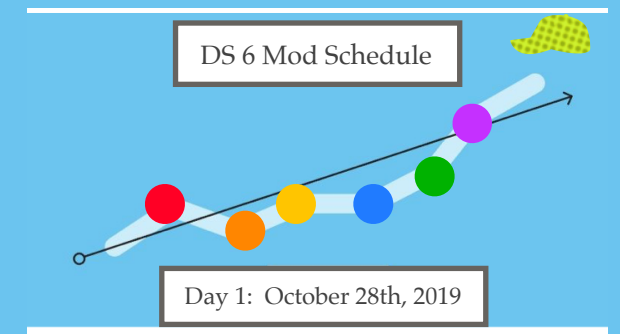
# ● Mod 2: Data Engineering

- SQL
- NoSQL
- JSON / XML
- APIS
- Web scraping (Beautiful Soup / Selenium)
- OOP



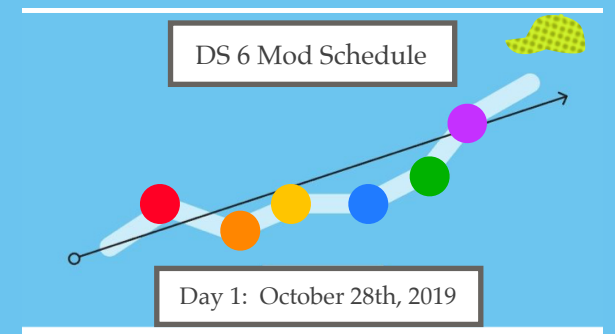
# ● Mod 3: Probability, Sample, AB Testing

- Combinatorics and Probability
- Distributions
- Central Limit Theorem
- Hypothesis Testing
- Bayes



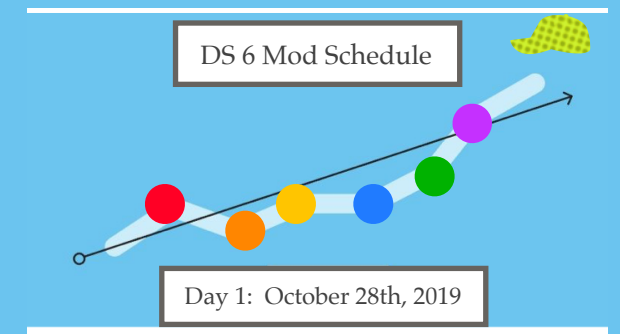
# ● Mod 4: Statistical Modeling

- Linear Regression
- Multiple Linear Regression
- Cost Functions / Gradient Descent
- Time Series



# ● Mod 5: Machine Learning

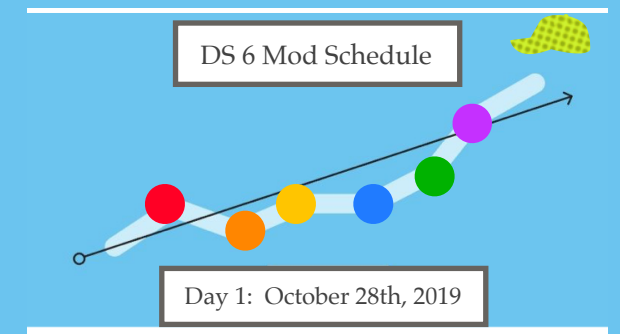
- Classification
  - Logistic Regression
  - Decision Trees
  - Ensemble Methods
  - Bagging and Boosting
  - SVM
  - PCA
- Clustering





# ● Mod 6: NLP and Deep Learning

- Natural Language Processing
  - NLTK, word embeddings, LSTM, Recurrent NN
- Deep Learning
  - CNN
  - Keras
  - Tensor Flow
- AWS
- Recommender Systems



---

# Projects

---

- Module 1: End of mod project with data provided
- Module 2: No project
- Mods 3,4,5: End of mod projects with your own data
- Module 6: Capstone

# Career Services

- Introduction in first few weeks
- Return in week 11
- Stay with you until you get a job.
- Career Search Module on Learn.

# Presentations

- Project Deliverables
- 5 minute presentations
  - Practice and feedback with instructors
  - In-house
  - To the Flatiron community

---

# Assessments

---

- Administered at the end of each module except 6.
- 1 - 2 hours, open book

---

# Your New Favorite Website

---

- <https://github.com/learn-co-students/chicago-102819>

---

# Learn.co labs

---

- Instructors deploy prior to lectures
- Get increasingly difficult and take more time
  - Unstructured time is all about labs
- Lab review sessions

---

# Networking

---

- It is crucial!
- You already have a built-in network at Flatiron
  - Introduce yourselves to other disciplines
- Attend events, both on and off campus



---

# Technical Sophistication

---

- What to do when you get stuck in a lab
  - First resource: Google it. We all do it.
    - Learn to love Stack Overflow
  - Second resource: your cohort
  - Third resource: your instructors

What types of problems  
do data scientists tackle?

What industries  
do data scientists work in ?

# Imposter Syndrome

What is it?

# Groups of Three

Take 5 minutes and outline what you think are the phases of data science project. Include as little or as much detail as you please.

# CRISP-DM

