Intro to Data Science

Why are we here?

1. Intro to the Codeup experience
2. Overview of Data Science
    - What is *Data Science*?
    - How industry utilizing Data Science
    - How you can get started learning this field
3. Hands-on workshop introduction to programming in Python
    - Intro to the Python programming language
    - Example problems, practice exercises, and guidance
    - How to get better at Python (Homework assignments & recommended resources)

Why Codeup?

1. Focus on student outcomes
2. Placement services and quality of network
3. Immersion works. Full-time, live, in-person instruction for 5 months works.

What is Data Science?

- The application of the scientific method to infer outcomes from data.
- Interdisciplinary applied science at the intersection of
    - Programming
    - Probability and Statistics
    - Domain Expertise
- A broad description of approaches ranging from business analysis and visualizations to machine learning and deep neural network analysis.
- An increasingly accessible field

Isn't data science just statistics?

- "Future of Data Analysis", Tukey 1962, `https://projecteuclid.org/euclid.aoms/1177704711`
- "50 Years of Data Science" by Donoho, `https://courses.csail.mit.edu/18.337/2015/docs/50YearsDataScience.pdf`

What *isn't* Data Science?

- Only statistics or only mathematics: let the computer compute and the people think
- Magic, inaccessible, or free from technical and ethical scrutiny

5 Ways to get the most out of this workshop

1. Engage with the material, engage with others
2. Ask questions
3. What you get out of this is a function of what you put into it
4. Treat each other with excellence
5. Try things out! Run code! Experiment with the material!

5 Takeaways you will get from today's workshop

1. An introduction to the Codeup experience
2. Overview of the fundamentals of Data Science
3. Intro to programming with Python
4. Instructional materials and your own copies of the prepared learning environments
5. Homework to keep practicing your craft

What we will *not* cover today

- Excel, R, SAS, SPSS, or other statistical analysis tools
- Everything you need to know about Python or statistics
- 5 months worth of practice, exercises, training, and mentoring
- Perceptrons, neural networks, and deep learning (too much for a 4 hour intro)

The five kinds of questions Data Science can answer

1. How many or how much of something
2. Is this observation A or B (or C or D)
3. What groupings exist in the data already?
4. What's most likely to happen next?
5. Is this thing weird?