

WELCOME

How to maximize your DSI experience

ICEBREAKER

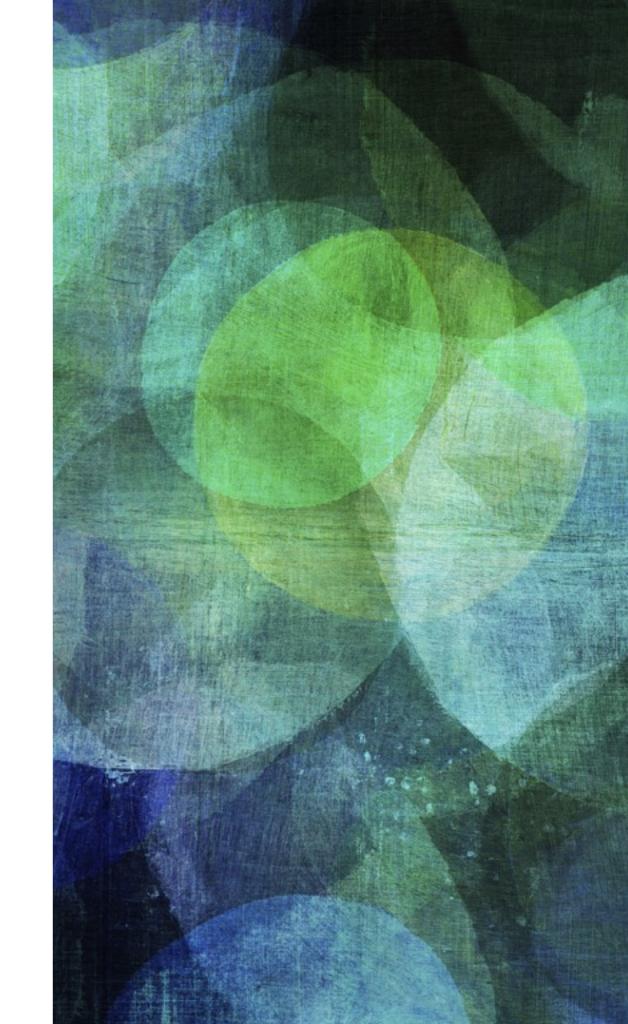
Name

History

Spirit animal

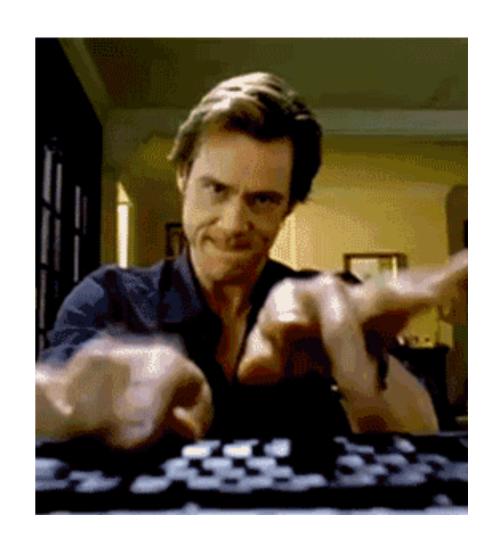
Quirky habit

Life goal



GENERAL LEARNING HACKS

- ➤ Communicate regularly: seek help early
- Understand your learning styles
 - ➤ Visual-Auditory-Kinetic
 - ➤ Early vs. late study sessions
- ➤ Manage your time
 - ➤ Focus: Pomodoro technique
 - Divide and conquer
 - Schedules and prioritized checklists
- ➤ Take great notes
- ➤ Record lectures





DSI SPECIFIC TIPS

- ➤ Comment your code!
- Keep your Github repo organized
 - ➤ Easy to navigate and grade
 - ➤ Portfolio of model examples for reference
 - ➤ Ready to share with prospective employers
- ➤ Get familiar with documentation!!
- ➤ Google and class resources first... then other students... then us
- ➤ Master Slack and keep the community hoppin'
- ➤ Margin in everything
- ➤ Your deliverables must be client-worthy

LOGISTICS

- ➤ Slack is how we communicate
 - ➤ Text is backup
 - ➤ No emailing
 - ➤ Let all three of us know if you're late, absent, or a deliverable will be late
- ➤ Github is how we share resources and you submit deliverables
- ➤ Schedules will change often
- ➤ Build margin, be early and ahead!

66

And in the end, the love you take is equal to the love you make.

PYTHON REVIEW

- 1. How do you create an empty list named "mylist"?
- 2. What will the following code return? 5 > 3 or 5 < 3
- 3. What will be stored in the "nums" object? nums = range(10)
- 4. How do you check the type and the length of the "nums" object?
- 5. How do you return the last number in the "nums" object?

PYTHON REVIEW

- 6. Slice the "nums" object to return a list with the numbers 2, 3, 4.
- 7. What is the difference between nums.append(10) and nums + [10]?
- 8. How do you divide 3 by 2 and get a result of 1.5?
- 9.Import the "math" module, and then use its "sqrt" function to calculate the square root of 1089.
- 10.What type of object is created by this code? d = {'a':10, 'b':20, 'c':30}

PYTHON REVIEW

- 11.In the "d" object, what are "a", "b", and "c" called? What are 10, 20, and 30 called?
- 12. How do you return the 10 from the "d" object?
- 13. How do you change the 30 to a 40 in the "d" object?
- 14.From the "people" object, return Brandon's state only: people = {'Alice': ['Washington', 'DC'], 'Brandon': ['Arlington', 'VA']}
- 15.Define a function "calc" that takes two variables, "a" and "b", and returns their sum.

OVERVIEW OF THE DATA SCIENCE WORKFLOW

- 1. Identify the problem
- 2. Acquire the data
- 3. Parse the data
- 4. Mine the data
- 5. Refine the data
- 6. Build a data model
- 7. Present the results

DATA SCIENCE WORKFLOW **IDENTIFY THE PROBLEM** ☐ Identify business/product objectives Identify and hypothesize goals and criteria for success Create a set of questions for identifying correct data set **ACQUIRE THE DATA** ☐ Identify the "right" data set(s) ☐ Import data and set up local or remote data structure Determine most appropriate tools to work with data PARSE THE DATA Read any documentation provided with the data ☐ Perform exploratory data analysis Verify the quality of the data MINE THE DATA Determine sampling methodology and sample data Format, clean, slice, and combine data in Python ☐ Create necessary derived columns from the data (new data) REFINE THE DATA ☐ Identify trends and outliers Apply descriptive and inferential statistics □ Document and transform data BUILD A DATA MODEL Select appropriate model □ Build model Evaluate and refine model PRESENT THE RESULTS Summarize findings with narrative, storytelling techniques Present limitations and assumptions of your analysis Identify follow up problems and questions for future analysis