Welcome to COGS 108!

Data Science in Practice

Shannon E. Ellis, Ph.D

Associate Teaching Professor
Cognitive Science Department
Halıcıoğlu Data Science Institute

Lectures: https://github.com/COGS108/Lectures-Ellis/tree/wi25

hello

my name is

Shannon Ellis

COGS 108 Instructional Staff: TAs

Yueyan

Yueying

Reuben

Tori

Jon

Tyler

Liz

Divya

COGS 108 Instructional Staff: IAs

Ricky

Shrieyaa

Coco

Ashley

Will

Sehee

Jorge

Nick

The (dreaded) waitlist

1. I know this matters to you and is a source of stress (and I hate that).
2. I do not have access to the system to let someone off of the waitlist
3. **I don't anticipate anyone getting off the waitlist this quarter**
    a. Last Winter, there were 700 students in this class. This quarter there are already 820 students enrolled. (~182 projects; ~22/TA; 20 *should* be the max)
    b. If students drop (not common), staff will enroll students from the waitlist
    c. Our staff (cogsadvising@ucsd.edu) take care of this
4. The waitlist settles after week 2.

Students on the waitlist have access to Canvas/all course materials. If you want/hope to be enrolled, please keep up with the material these first two weeks. **You will be expected to have completed the material if you are eventually enrolled.**

...so what exactly is data science?

Google

what is biology

All | Books | Videos | Images | News | More | Settings | Tools

The study of life, living things, & organisms

Google

what is biology

All | Books | Videos | Images | News | More | Settings | Tools

science | study | definition | lesson | introduction | presentation | organisms | education | dominant | life | chapter | biologist | concept | ecology | living

Introduction to Biology | What is ...
youtube.com

Biology? Biology is the study of life ...
slideplayer.com

What Is Biology? | Live Science
livescience.com

PPT - What is biology? PowerPoint ...
slideserve.com

What is biology and Why study biolo...
slideshare.net

Biology Areas of Study and Definition ...
youtube.com

Branches of Biology: What are the ...
golifescience.com

Biology - Definition and Examples ...
biologyonline.com

What is Biology?
studylib.net

Topic 1 - What is Biology?
slideshare.net

What is Biology? - Real Science-...
gravitaspublications.com · In sto...

Study of Biology - Presentation Biology
sliderbase.com

What is Biology - Introduction to ...
docsity.com

PPT - Chapter 1 Biology: The Study of ...
slideserve.com

What is biology? ☐ Bio means "life ...
slideplayer.com

Biology work, - Home | Facebook
facebook.com

What are the Branches of Biology ...
empoweryourknowledgeandhappytri...

PPT On The Introduction To Biology ...
myprivatetutor.my

PPT - What is BIOLOGY? PowerPoint ...
slideserve.com

1.1 Introduction to Biology - ppt video ...
slideplayer.com

What is biology? Concept, Importance ...
burrosabio.net

Biology - Wikipedia
en.wikipedia.org

Topic 1 - What is Biology?
slideshare.net

What is Biology? | Importanc...
swiftutors.com

Related searches

- study what is biology
- organisms what is biology
- what is biology pictures

Google

what is philosophy

All | Books | Videos | Images | News | More | Settings

Study of the fundamental nature of knowledge reality, existence, ethics

Google

what is philosophy

All | Books | Videos | Images | News | More | Settings | Tools

definition | introduction | knowledge | reality | activity | philosophical | mathematics | concept | nature | of education | study

What is PHILOSOPHY? (and what is it for ...
youtube.com

What is Philosophy by BVH Prasad ...
in.pinterest.com

What is philosophy?. - ppt download
slideplayer.com

What is philosophy?
slideshare.net

Philosophy PPT for KS3, KS4 ...
pinterest.co.uk

by Gilles Deleuze, Felix...
amazon.com

PPT - Chapter 1: What is Philosophy ...
slideserve.com

What is philosophy?
slideshare.net

What is Philosophy? - Philosop...
123philosophy.com

How to think like a philosopher - ...
quora.com

Deleuze and Guattari b...
en.wikipedia.org

1st Lesson. 16/9. Introduction. What is ...
philocyclevl.wordpress.com

What Is Philosophy for...
amazon.com

The philosophy of education. (Chapter...
en.ppt-online.org

Introduction to Philosophy (Beginners ...
carneades.org

What is Philosophy ? | Know-It-All
empoweryourknowledgeandhappytrivia...

What is Philosophy? (Philosophical ...
youtube.com

Geoffrey Gorham ...
slideplayer.com

What is philosophy? (2) - Vanguard News
vanguardngr.com

What is Philosophy ? | What is ...
pinterest.com

What is Philosophy? | ...
blogs.lmu.edu

What Is Philosophy? | ...
sup.org

CURRICULUM AND PEDAGOGY: Chapte...
viola3343.blogspot.com

Look Up Aim High: Philosophy and ...
hamsareddy.blogspot.com

Related searches

knowledge what is philosophy

what is philosophy quotes

reality what is philosophy

Philosophy What is the...
studylib.net

Google

what is physics

All | Books | Videos | Images | News | More | Settings

Study of matter, energy, atoms, mechanics, and so on.

Google

what is physics

All | Books | Videos | Images | News | More | Settings | Tools

definition | math | presentation | biology | energy | lesson | matter | work | E=mc² formula | power | measurement | acceleration | introduc

What is Physics? - YouTube
youtube.com

PPT - What is Physics ? PowerPoin...
slideserve.com

Introduction What is Physics? | What is ...
pinterest.com

What is Physics ? - explained (Hindi ...
youtube.com

PPT - What is Physics? Power...
slideserve.com

What is Physics by IKPENI STEPHE...
flipsnack.com

What is Physics? The study of matter ...
slideplayer.com

Introduction to physics (video) | Khan ...
khanacademy.org

What is Physics?. - ppt video online ...
slideplayer.com

Physics in Everyday Life ...
youtube.com

An Introduction to String Theory: What ...
slimy.com

What is physics? (article) | Khan Academy
khanacademy.org

Physics - Lessons - Tes Teach
tes.com

1 Chapter 1 Measurement What is Phy...
slideplayer.com

What is physics, Primary science, Physics
pinterest.de

What is Physics? :: Physics
bellevuecollege.edu

1.1 What is Physics?
slideshare.net

scientific study of matter and energy ...
slideplayer.com

PPT - What is physics? PowerPoint ...
slideserve.com

Condensed-Matter ...
nap.edu

Definition of Physics - YouTube
youtube.com

Study Material for IIT JEE ...
askiitians.com

PPT - What is Physics? PowerPoint ...
slideserve.com

Scope of Physics: What is Physics ...
toppr.com

The study of Venn Diagrams?

Google
what is data science
All
Videos
Books
Images
News
More
Settings
Tools
infographic
machine learning
big data
diagram
statistics
artificial intelligence
definition
cartoon
process
predictive
chart
poster
engineering
visualization
pdf
What is Data Science - Introduction to ...
intellipaat.com
What is Data Science? – Dataquest
dataquest.io
What Is Data Science Defini...
datarobot.com
Data Science for Beginners. ...
towardsdatascience.com
What is Data Science?
WHAT IS DATA SCIENCE?
What Is Data Science, and What Does...
innoarchitech.com
What Is Data Science? A Beginner's ...
edureka.co
Why Is Data Science Certification W...
cognixia.com
What is Data Science?
datascience.berkeley.edu
A Complete Data Science Tutorial for ...
data-flair.training
What is Data Science? - Codeup
codeup.com
Data Science Interview Guide....
towardsdatascience.com
6 Amazing Data Science Applications ...
data-flair.training
How Data Science is Taking Over Mobil...
clevertap.com
Data Science and its scope in IT industry
digitalnest.in
Tips for Hiring Data Scientists from...
burtchworks.com
The Essential Data Science Venn Diagram
kdnuggets.com
Data Science?. by: Dr. Benjamin ...
towardsdatascience.com
difference between data science ...
quora.com
What on earth is data science?
kdnuggets.com
What is Data Science - Introduction to ...
intellipaat.com
Drew Conway's original DataSci venn diagram. (The OG)

What is data science?

Copyright © 2014 by Steven Geringer Raleigh, NC.
Permission is granted to use, distribute, or modify this image,
provided that this copyright notice remains intact.

http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram

David Donoho

This coupling of scientific discovery and practice involves the collection, management, processing, analysis, visualization, and interpretation of vast amounts of heterogeneous data associated with a diverse array of scientific, translational, and interdisciplinary actions.

Bradley Voytek

The study of how the quantification of observable phenomena can lead to human understanding of the processes giving rise to those phenomena—or even the ability to predict future outcomes absent human understanding—and why certain phenomena require more or less data to lead to human understanding and/or prediction accuracy

Rafael Irizarry

An umbrella term used by organizations to describe the processes used to extract value from data.

Wikipedia

A concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data.[3] It employs techniques and theories drawn from many fields within the context of mathematics, statistics, information science, and computer science.

Cogs 108 Definition

The scientific and ethical process of extracting value from data.

UC San Diego
Cognitive Science

50 Best Jobs in America

Awards

This report ranks jobs according to each job's Glassdoor Job Score, determined by combining three factors: number of job openings, salary, and overall job satisfaction rating.

	Job Title	Median Base Salary	Job Satisfaction	Job Openings
#1	Enterprise Architect	$144,997	4.1/5	14,021
#2	Full Stack Engineer	$101,794	4.3/5	11,252
#3	Data Scientist	$120,000	4.1/5	10,071

Highest Paying Jobs

Oddball Interview Questions

View Jobs

https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm

Data scientists ask interesting questions & answer them with data

The goal in COGS 108 is to *do* data science.

Course Objectives

- Formulate a plan for and complete a data science project from start (question) to finish (communication)
- Explain and carry out descriptive, exploratory, inferential, and predictive analyses in Python
- Communicate results concisely and effectively in reports and presentations
- Identify and explain how to approach an unfamiliar data science task

How we'll approach learning about *and doing* data science in COGS 108

Scheduling

Lecture: MWF 9-9:50 AM OR 2-2:50 PM (PETER 108)
Discussion Sections: M, W, F
Prof Office Hours: Fri 10-12 PM (starting week 2)

A04	M	2-2:50p	CSB	004
A05	M	3-3:50p	CSB	004
B01	W	10-10:50a	CSB	004
B02	W	11-11:50a	CSB	004
B03	W	12-12:50p	CSB	004
B04	W	1-1:50p	CSB	004
B05	W	3-3:50p	CSB	004
A06	W	4-4:50p	CSB	004
A01	F	11-11:50a	CSB	005
A02	F	12-12:50p	CSB	004
A03	F	1-1:50p	CSB	004
A07	F	4-4:50p	PETER	102
B06	F	5-5:50p	PETER	102
B07	F	6-6:50P	PETER	102

What are course staff up to in COGS 108?

Everyone | assisting in discussion section labs, project grading, answering questions, providing me with feedback

Roles | Lead TA (Reuben), grade management/executive assistant, assignments, labs, Q&A, lecture quizzes, discussion slides

When helping, our goal is to do just that. Help. Often, you'll be under time pressure to complete an assignment (which we understand), but we *really* care about your learning. So, <u>we all will prioritize guiding you to an answer to ensure learning</u> (rather than just giving an answer away).

COGS 108: General Plan

Week	Topic(s)
1	Data Science & Version Control
2	Data Intuition & Wrangling
3	Ethics & DS Questions
4	Data Viz
5	Exploratory Data Analysis
6	Inference
7	Text Analysis
8	Machine Learning
9	Geospatial Analysis
10	Data Science Communication & Jobs

Programming Prerequisite

- MAE 8 - MATLAB
- CSE 8A or 11 - Python/Java
- COGS 18 - Python
- DSC 10 - Python
- (CSS 1 - Python)

Bottom line: **we will assume programming knowledge.**

Python will be used for all labs/projects/assignments.

No programming experience (or you forget it all)?

- *Preferred option*
  - Take a programming course first
  - COGS 18 : Introduction to Python
- *Can't wait?*
  - Use online sites like *codecademy.com* or *LearnPython.org*
  - *Python Data Science Handbook*

Course links

GitHub	https://github.com/COGS108	lecture/section slides & final projects
datahub	https://datahub.ucsd.edu	assignment & lab completion & submission
Q&A (EdStem/Ed)	https://edstem.org/us/courses/70671/discussion	questions, discussion, and regrade requests
Canvas	https://canvas.ucsd.edu/courses/62380	grades
Anonymous Feedback	Submit via Google Form	if I ever offend you, use an example you hate, or to provide general feedback that you don't want your name attached to

General grading:

	% of Total Grade
Pre-course Survey (#finaid)	1
(8/9) Weekly Quizzes (lecture content)	8
(8) Discussion Labs (technical)	16
(4) Assignments	32
Final Group Project	43
(1) Project Review*	5
(1) Project Proposal*	9
(2) Project Checkpoints*	10
(1) Final Report*	15
(1) Final Video*	3
(1) Project Survey	1

* indicates group submission

Lecture attendance is *not* required, but is incentivized

Google Forms & Kahoot! will be used for in-class **lecture** Q&A

- Being present for >= 2/3 of the in class exercises will net 0.5% bonus to the final grade
- Most (but not all) days will have an activity
- To get credit for lecture attendance, you must participate in a given day's activities

<u>Notes</u>:

- you can attend *either* lecture on a given day to get credit. If you attend both, you'll only get credit for attending one day's lecture.
- For Google Forms, you'll have to use your UCSD email for login

Why Google Form questions/Kahoot! in COGS 108?

- There are a whole lot of you!
- Checks understanding
- Provides me with feedback
- Aids in critical thinking & allows for application of concepts
- Give you all a break from listening to me (we humans need this!)

Weekly Lecture Quizzes:

- (9) weekly quizzes (first one due Sunday of this week)
- Goal: to help you keep on top of the material covered in lecture
- Why?: experience + student feedback
- How:
  - Taken on Canvas
  - Single Attempt
  - ~10 Questions
  - Timed : 15 minutes
  - Posted by Friday @ 3PM (after each week of lecture)
  - Meant to test concepts from that week's lecture

Lecture quizzes will be due on Sundays by 11:59 PM.

No late submissions. Lowest quiz score will be dropped.

All deadlines Sun at 11:59 PM

Week	Quiz	Discussion Lab	Lecture Quiz	Assignment
1	Data Science, Python, & Version Control	--	Q1	--
2	Data Intuition	D1	Q2	A1 - python, Pre-course survey
3	Wrangling, Data Ethics	D2	Q3	--
4	DS Questions & Data Visualization	D3	Q4	A2 - pandas/viz, Project Review*
5	Exploratory Data Analysis	D4	Q5	Project Proposal*
6	Inference	D5	Q6	A3 - Inference
7	Text Analysis	D6	Q7	Chkpnt #1: Data*
8	Machine Learning	D7	Q8	--
9	Nonparametric Analysis & Geospatial Analysis	D8	Q9	A4 - NLP/ML; Chkpnt #2: EDA*
10	Data Science Communication & Jobs	--	--	--

Final Project(Report*, Video*, Survey): due Wed of finals week by 11:59 PM

*indicates group submission. All other assignments/surveys are completed & submitted individually.

(4) Assignments

Assignments are completed individually and graded programmatically.

- These are meant to get you practice programming around the topics covered in class.
- The first two are much simpler than the following two and should take less time.
- You will have to look some stuff up on your own. This is by design.
- Instructions must be followed to receive credit.
- You'll have the opportunity to practice in discussion section.

Assignments will be due on Sundays by 11:59 PM.

7 free late days; 75% credit if submitted w/n 5d after deadline after that.

Used on assignments and/or labs.

Assignment Submission @ Datahub: https://datahub.ucsd.edu

DATA SCIENCE / MACHINE LEARNING PLATFORM

UC San Diego

Information Technology Services - Educational Technology Services Help Options

UC San Diego Jupyterhub (Data Science) Platform

Before Fri: log onto datahub. (If you have Canvas access, you have datahub access)

All deadlines Sun at 11:59 PM

Week	Quiz	Discussion Lab	Lecture Quiz	Assignment
1	Data Science, Python, & Version Control	--	Q1	--
2	Data Intuition	D1	Q2	A1 - python, Pre-course survey
3	Wrangling, Data Ethics	D2	Q3	--
4	DS Questions & Data Visualization	D3	Q4	A2 - pandas/viz, Project Review*
5	Exploratory Data Analysis	D4	Q5	Project Proposal*
6	Inference	D5	Q6	A3 - Inference
7	Text Analysis	D6	Q7	Chkpnt #1: Data*
8	Machine Learning	D7	Q8	
9	Nonparametric Analysis & Geospatial Analysis	D8	Q9	A4 - NLP/ML; Chkpnt #2: EDA*
10	Data Science Communication & Jobs	--	--	--

Final Project (Report*, Video*, Survey): due Wed June 9th of finals week by 11:59 PM

*indicates group submission. All other assignments/surveys are completed & submitted individually.

Group Projects: the main focus of COGS 108

Assigned groups of 4-5 Individuals

Carry out the entire data science process on a topic of your choosing

All deadlines Sun at 11:59 PM

Week	Quiz	Discussion Lab	Lecture Quiz	Assignment
1	Data Science, Python, & Version Control	--	Q1	--
2	Data Intuition	D1	Q2	A1 - python, Pre-course survey
3	Wrangling, Data Ethics	D2	Q3	--
4	DS Questions & Data Visualization	D3	Q4	A2 - pandas/viz, Project Review*
5	Exploratory Data Analysis	D4	Q5	Project Proposal*
6	Inference	D5	Q6	A3 - Inference
7	Text Analysis	D6	Q7	Chkpnt #1: Data*
8	Machine Learning	D7	Q8	--
9	Nonparametric Analysis & Geospatial Analysis	D8	Q9	A4 - NLP/ML; Chkpnt #2: EDA*
10	Data Science Communication & Jobs	--	--	--

Final Project(Report*, Video*, Survey): due Wed June 9th of finals week by 11:59 PM

*indicates group submission. All other assignments/surveys are completed & submitted individually.

Discussion Section

Goals:

- help with technical aspects of the course
- Provide time to work on labs
- Get questions answered

Students will complete and submit weekly labs (low stakes), where you will get to practice the material presented in class before seeing it on an assignment (higher stakes).

If you're able to fly through the labs, you never *have* to attend section.

If you're struggling, you should absolutely attend section.

Discussion Sections will start Mon of week 2 (note: Mon holidays)

All deadlines Sun at 11:59 PM

Week	Quiz	Discussion Lab	Lecture Quiz	Assignment
1	Data Science, Python, & Version Control	--	Q1	--
2	Data Intuition	D1	Q2	A1 - python, Pre-course survey
3	Wrangling, Data Ethics	D2	Q3	--
4	DS Questions & Data Visualization	D3	Q4	A2 - pandas/viz, Project Review*
5	Exploratory Data Analysis	D4	Q5	Project Proposal*
6	Inference	D5	Q6	A3 - Inference
7	Text Analysis	D6	Q7	Chkpnt #1: Data*
8	Machine Learning	D7	Q8	--
9	Nonparametric Analysis & Geospatial Analysis	D8	Q9	A4 - NLP/ML; Chkpnt #2: EDA*
10	Data Science Communication & Jobs	--	--	--

Final Project(Report*, Video*, Survey): due Wed June 9th of finals week by 11:59 PM

*indicates group submission. All other assignments/surveys are completed & submitted individually.

Course Confusion

- If something in lecture, a section workbook, or an assignment is unclear:
  - *ask in class*
  - *ask during section*
  - *post on Ed*
  - *ask a classmate*
  - *come to office hours*

I look at Ed first every day. Then email. I have 820+ students. Please use Ed when possible. (I do answer email and Canvas messages, but please spare me.)

CLASS CONDUCT

In all interactions in this class, you are expected to be respectful. This includes following the UC San Diego principles of community.

This class will be a welcoming, inclusive, and harassment-free experience for everyone, regardless of gender, gender identity and expression, age, sexual orientation, disability, physical appearance, body size, race, ethnicity, religion (or lack thereof), political beliefs/leanings, or technology choices

At all times, you should be considered and respectful. Always refrain from demeaning, discriminatory, or harassing behavior and speech. Last of all, **take care of each other**.

If you have a concern, please speak with Prof. Ellis, your TAs, or IAs. If you are uncomfortable doing so, the OPHD and/or CARE are wonderful resources on campus.

Academic Integrity Policy

Don’t cheat.

Group work is required (projects) and encouraged throughout to help you learn.

COGS 108 is meant to be collaborative. We have a lot of students. Let’s harness that and learn from one another.

But, copying + pasting someone else’s work without understanding is not learning.

AI Policy

LLMs (and the Internet generally) are great.

Learning is important.

Use resources to help you learn...not to do your work.

If it helps you complete a task *and you understand what it provides you*, go ahead and use it.

But be sure to *think* about the output. Check it. Edit it.

And, be sure to cite it if you do.

https://github.com/COGS108/Overview/blob/wi24/COGS108-Syllabus.md#policy-on-using-ai-programming-assistance

A message for **first-gen students, transfer students, and those who don't have older siblings/friends who have attended UCSD/university**

If you are struggling, come to office hours. Ask questions on Ed. Reach out to me to ask for better approaches. Your classmates ARE doing this. And, you're not alone.

If you need a bit longer on something b/c you fell sick, a family thing came up, work called you in for an extra shift, etc., you have 7 free late days.

If something takes you beyond the 7 free late days, we can chat. Your classmates ARE doing this.

What COGS 108 logistics questions do you have?

Why this course?

You are going to be analyzing lots of data because you're studying to be a:

cognitive scientist
data scientist
computer scientist
neuroscientist, biologist, or chemist
social scientist (linguist? psychologist?)
statistician or biostatistician
CEO/small business owner
something else really awesome

I’m excited to have you all in COGS 108!

COGS 108

COGS 108

COGS 108

300
200
100
0
count
CG
CS
MA
DS
EC
EN
PC
BI
VA
BE
IS
PS
PY
CM
MC
SI
CE
CH
LN
UN
Major

Survey (link also on canvas; #finaid)

due 11:59 PM
Sunday after week 2

COGS 108 Student Survey (Winter 2025)

This survey must be completed by the Sunday following Week 2. It is being used as our #finaid survey.

This survey has three parts. The first part helps me get to know you better. It also collects data that will be anonymized and then used to demonstrate data science in class.

The second part collects your GitHub username. If you don't yet have one, you'll be directed where to go to get one.

The third part of the survey is collecting demographic information. THIS SECTION IS OPTIONAL - refusing to fill it out does NOT affect your grade. We use this to better understand how we're supporting/teaching students in this class.

If any of these data are used/displayed in class, the data will be anonymized. Please answer as truthfully as possible. How you respond will NOT affect how you do in this class. Also, many questions are NOT required. Please do not answer anything that makes you uncomfortable.