

Welcome to...

DATA 311!

Instructor: Scott Wehrwein (He/him/his)

↑
call me this!



Plan for today:

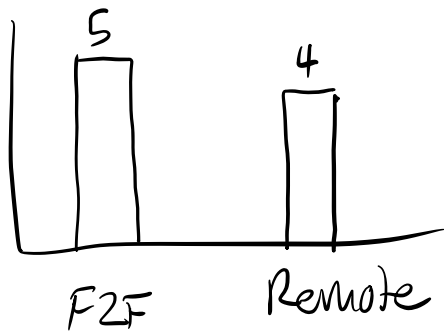
Logistics

- Course webpage/syllabus
 - Assessment
 - Policies
 - Textbook
 - Inclusive learning environment
- Flexibility and feedback

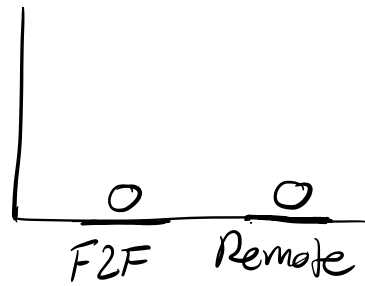
Introduction

- What is **data** science?
- What is **data**?
- Course overview

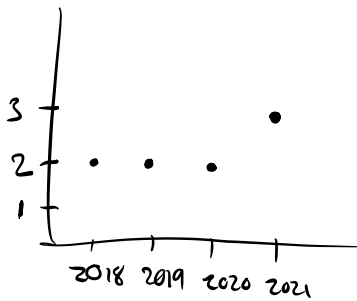
Quarters Taught



DATA 311 Taught



Family Size



Stats

organize, find patterns

applying info from data

prediction

analyze results

Solve problems

ethics of data collection
+ usage

presenting results
(to non-DSists)

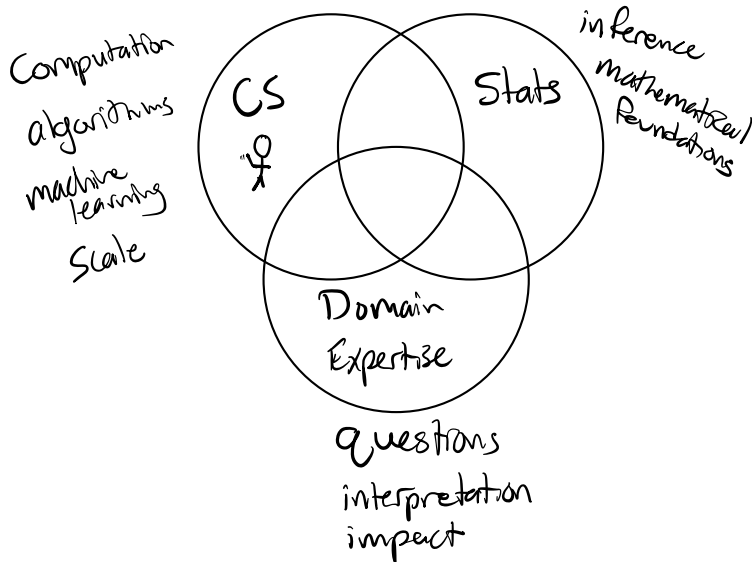
biases in data
lurking variables

What is data science?

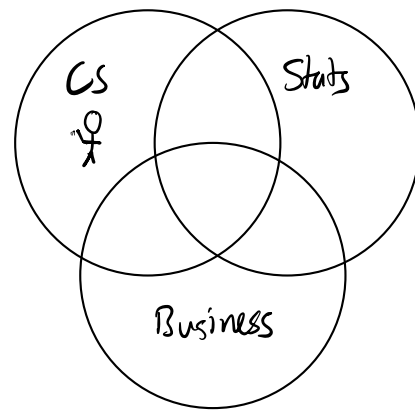
The process and/or study of extracting insights from data.

Premise: "data \neq information \neq understanding \neq wisdom."
- Clifford Stoll

The tools:



The Practitioners:



Solve problems

find patterns

predict things

optimize

get data

filter, sample, clean

Should we?

What do data scientists do?

Answer questions; test hypotheses

How much do daily temperatures vary between Dharm and Seattle?

Are movie budgets correlated with critic ratings?

Explore datasets to find trends, patterns, and insights

What impacts did COVID have on _____?

Make predictions

Given academic transcripts, can we identify students at risk of dropping out ahead of time so we can provide support when they need it?