

Pathways through the SDS Major at Smith College

Randi L. Garcia, Ph.D.
Smith College
2018 Joint Statistical Meetings
Vancouver, BC

SDS Program Goals

- Produce highly-skilled, versatile statisticians and data scientists
- Interdisciplinary
- Students learn statistics by **doing**, often in teams
- Inclusivity and diversity
- Statistical and data ethics

Brief History of SDS at Smith



Smith College SDS
@SmithCollegeSDS

Following

Can you tell we're excited to have a major in statistical and data sciences at
@smithcollege?



10:05 AM - 3 May 2016

The SDS Faculty

Visiting Assistant Professor/Mass
Mutual Faculty Fellows



Joint w/
MTH

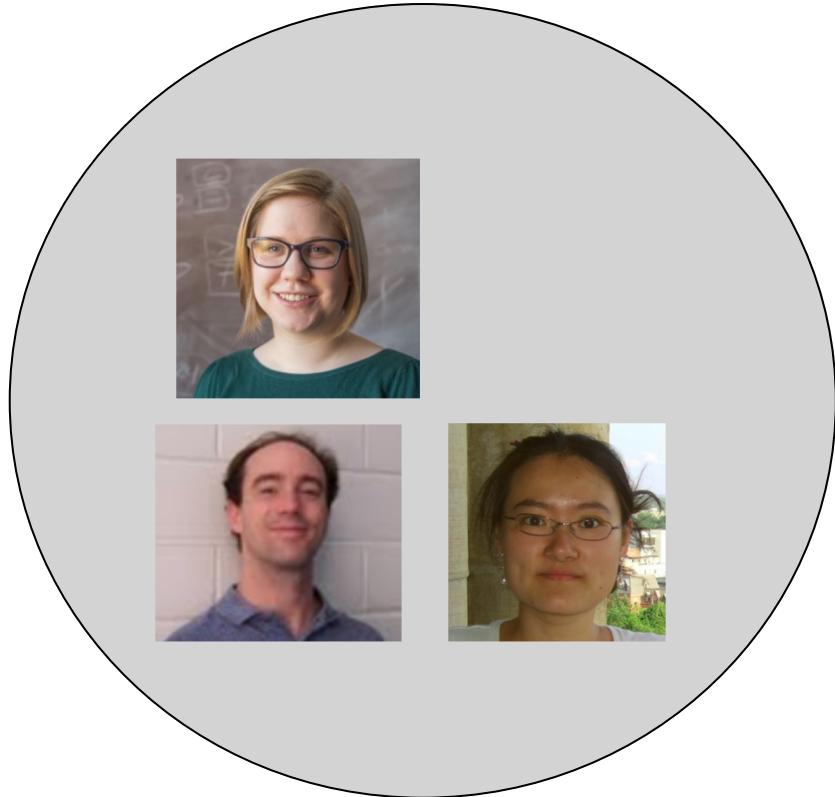


Joint w/
PSY



The SDS Faculty

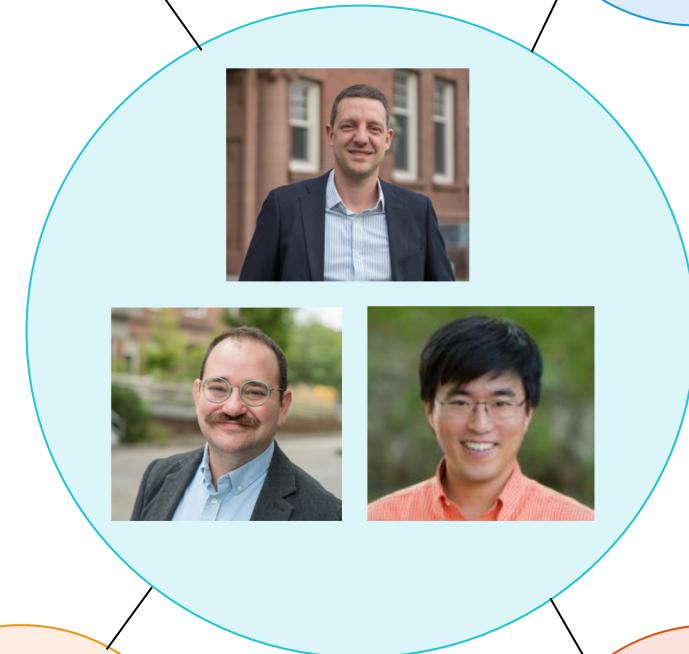
Visiting Assistant Professor/Mass
Mutual Faculty Fellows



Joint w/
MTH



Joint w/
PSY



Joint w/
CSC



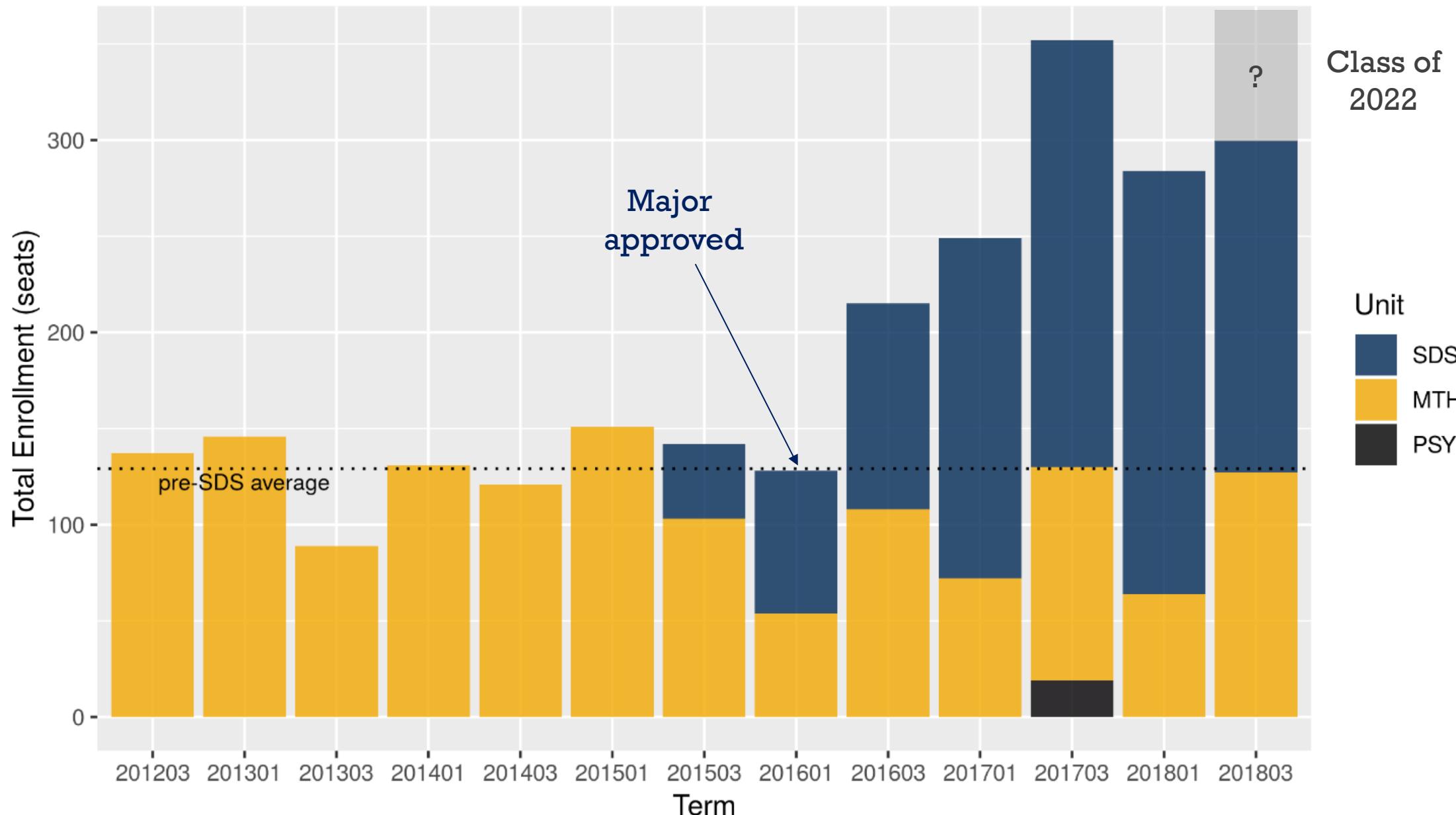
Joint w/
GOV

Art?
English?

This
could be
you!

Total enrollment in courses now managed by SDS, 2012-2018

107, 136, 192, 201, 220, 235, 290, 291, 293, 320, 364, 390



The Major in Statistical and Data Sciences

❑ Core Courses

- ❑ SDS 192: Intro to Data Science
- ❑ MTH/SDS 220 or SDS 201: Intro Stats
- ❑ CSC 111: Intro to Computer Science
- ❑ MTH/SDS 291: Multiple Regression
- ❑ MTH 211: Linear Algebra

- ❑ Programming Depth
- ❑ Statistics Depth
- ❑ Communication
- ❑ Application Domain
- ❑ Capstone

CORE (5 courses)

SDS/MTH 220

SDS/MTH 291

CSC 111

SDS 192

MTH 211

SDS/MTH 290: Design and Analysis

SDS 390: Topics in SDS

SDS 293: Machine Learning

CSC 212: Data Structures

SDS/CSC 235: Visual Analytics

SDS/CSC 109: Communicating with Data

SDS 236: Data Journalism

Stat Ethics

PSY 364

ECO 311

EGR 389

BIO 334

NSC 318

(etc.)

SDS 410

**Statistical Depth
(1 course)**

**Programming Depth
(1 course)**

**Communication
(1 course)**

**Application Area
(1 course)**

**Capstone
(1 course)**

Tenth Elective Course

Any course on previous slide

Five College Statistics Course

CSC 220: Advanced Programming

SDS/CSC 352: Parallel and Distributed Computing

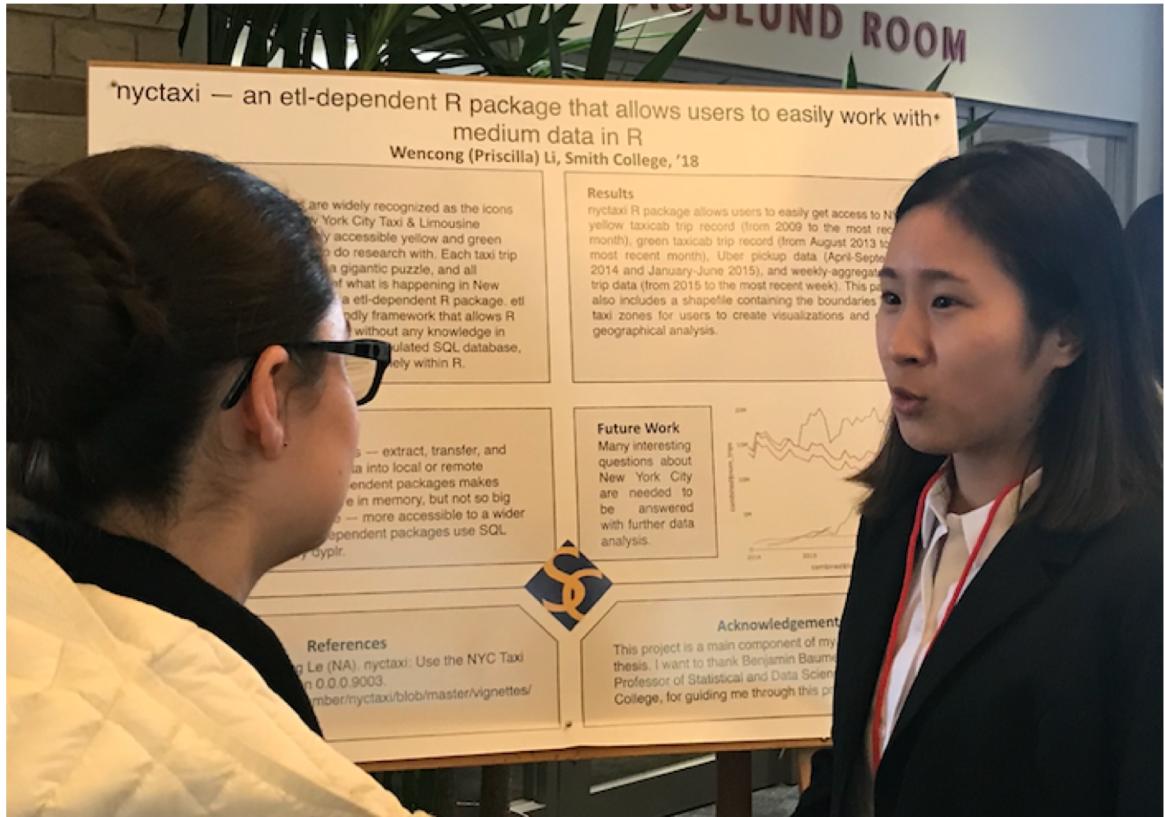
SDS/MTH 246: Probability

CSC 252: Algorithms

CSC 290: Artificial Intelligence

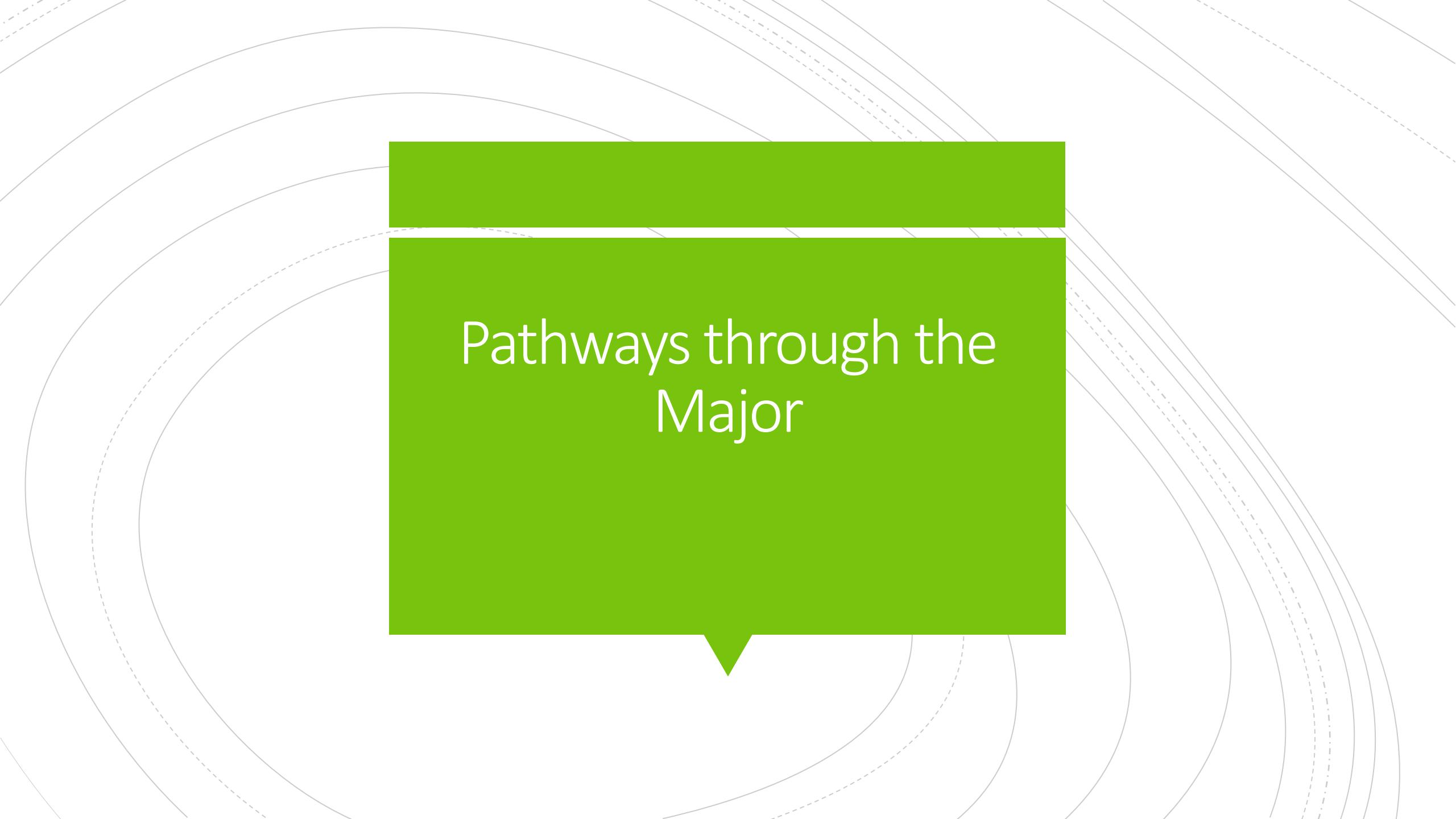
MTH 353: Advanced Topics in Discrete Mathematics

Honors Theses



Class Projects





Pathways through the Major

First Years Dabbling in SDS

- Intro to Data Science and Communicating with Data have no cap, no pre-requisites
- Two versions of Intro Stats: 220 requires calculus and 201 has no pre-requisites
 - Students less-prepared in math can progress through the major
- Interest in math might come via data science
- Exposure to data science before intro stats can be helpful
 - “Math people” and “data people”

CORE (5 courses)

SDS/MTH 201

SDS/MTH 291

CSC 111

SDS 192

MTH 211

SDS 293: Machine Learning

CSC 212: Data Structures

SDS/CSC 109: Communicating with Data

ECO 311

SDS 410

**Statistical Depth
(1 course)**

**Programming Depth
(1 course)**

**Communication
(1 course)**

**Application Area
(1 course)**

**Capstone
(1 course)**

Double Majors

- First exposure to data analysis might be in a domain application course
- Taken intro statistics in their domain department
 - They probably didn't use R
- Also have two minors if students realize they love data too late!
- Depending on the domain, working up to linear algebra can be a (positive) challenge

Psychology

CORE (5 courses)

PSY 201

SDS/MTH 291

CSC 111

SDS 192

MTH 211

SDS 390: Topics in SDS

SDS 293: Machine Learning

SDS 236: Data Journalism

PSY 364

SDS 410

**Statistical Depth
(1 course)**

**Programming Depth
(1 course)**

**Communication
(1 course)**

**Application Area
(1 course)**

**Capstone
(1 course)**

Linking to Computer Science and Math

- Students from computer science and math easily find their way to SDS.
- Computer science students exposed to data in SDS 192 or SDS 220
- Stats track in the Math Department
- Students also find their way to computer science and math through SDS.

Computer Science

CORE (5 courses)

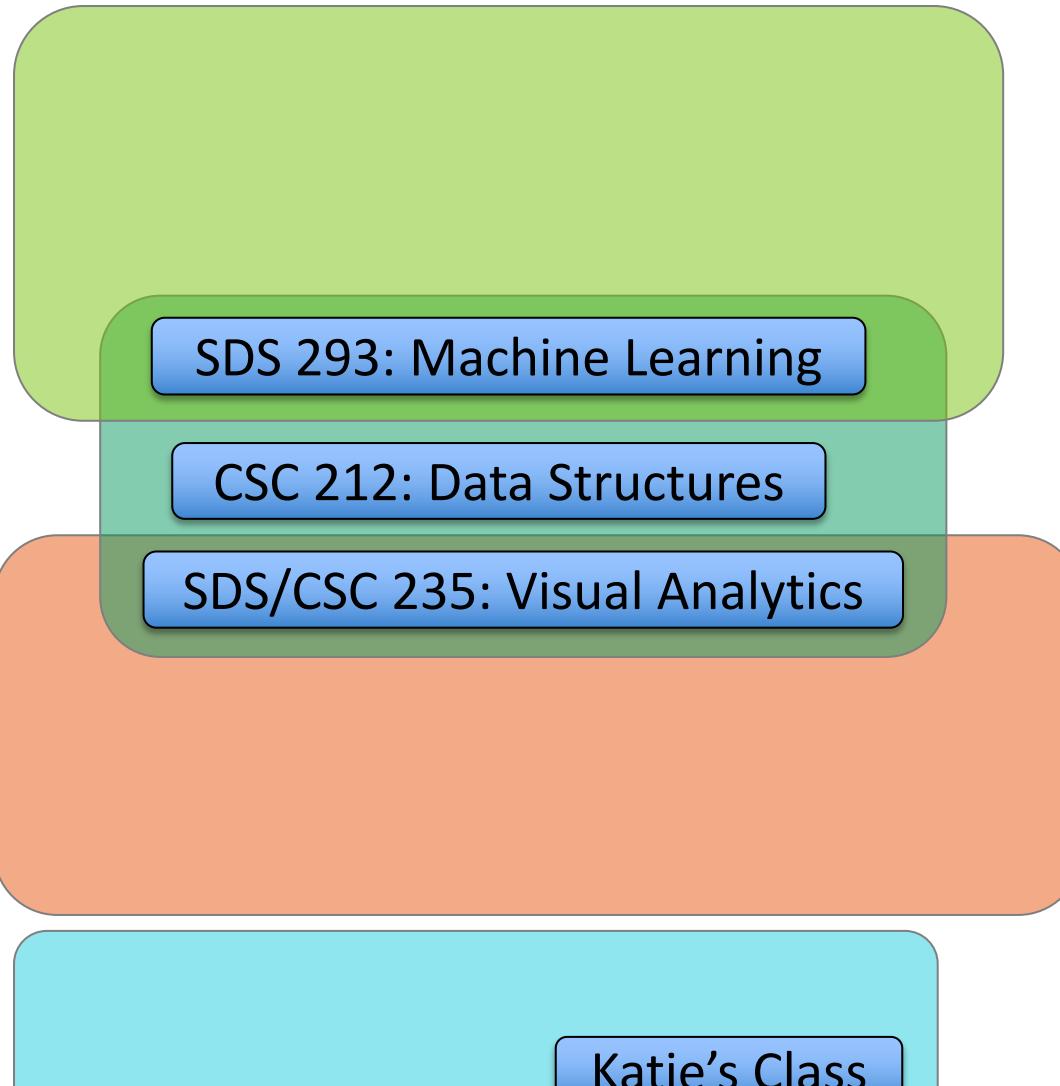
SDS/MTH 220

SDS/MTH 291

CSC 111

SDS 192

MTH 211



**Statistical Depth
(1 course)**

**Programming Depth
(1 course)**

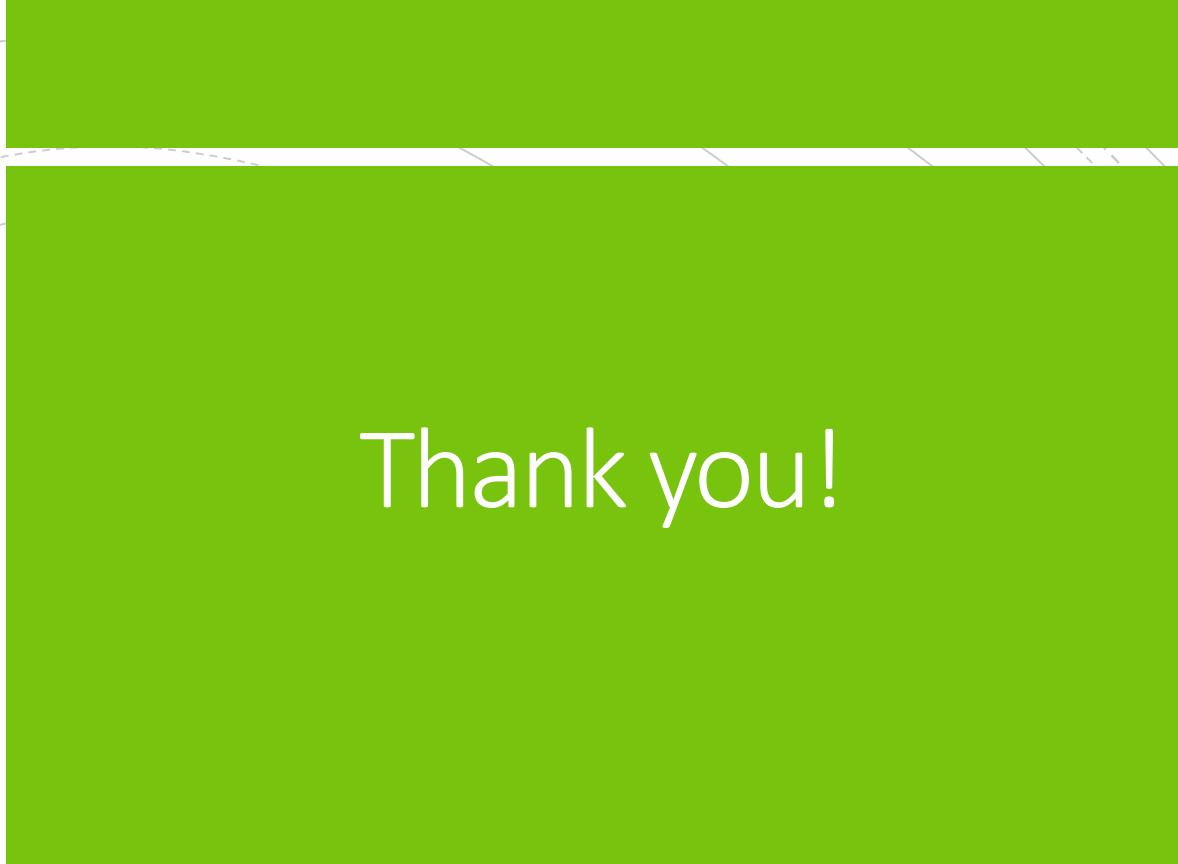
**Communication
(1 course)**

**Application Area
(1 course)**

**Capstone
(1 course)**

Challenges and Next Steps

- Managing the growth
- Hiring joint with government (and maybe the humanities next?)
- Curriculum tweaks (linear algebra before ML?, clarifying differences between intro courses, etc.)
- Richer offering of communication courses
- Creating a track for students interested in graduate studies in statistics and biostatistics.
- A sustainable model for giving students experiential learning opportunities, possibly through the capstone



Thank you!

