

Projects first in an interdisciplinary data analytics curriculum

Jessen Havill

Director, Data Analytics program

Professor of Computer Science

Benjamin Barney Chair of Mathematics

Denison University

`havill@denison.edu`

The Denison University logo, featuring the word "DENISON" in white, serif, all-caps font, centered within a solid red rectangular background.

DENISON

Outline

1. Context
2. Denison's Data Analytics major
3. Leading with problems and projects
4. Questions?

Denison University

- 30 minutes east of Columbus, Ohio
- liberal arts, ~40 majors
- undergraduate, independent, residential
- 2,250 students (50 states, 40 countries)
- 9:1 student/faculty ratio

The logo for Denison University, featuring the word "DENISON" in white, serif, uppercase letters centered within a solid red rectangular background.

DENISON

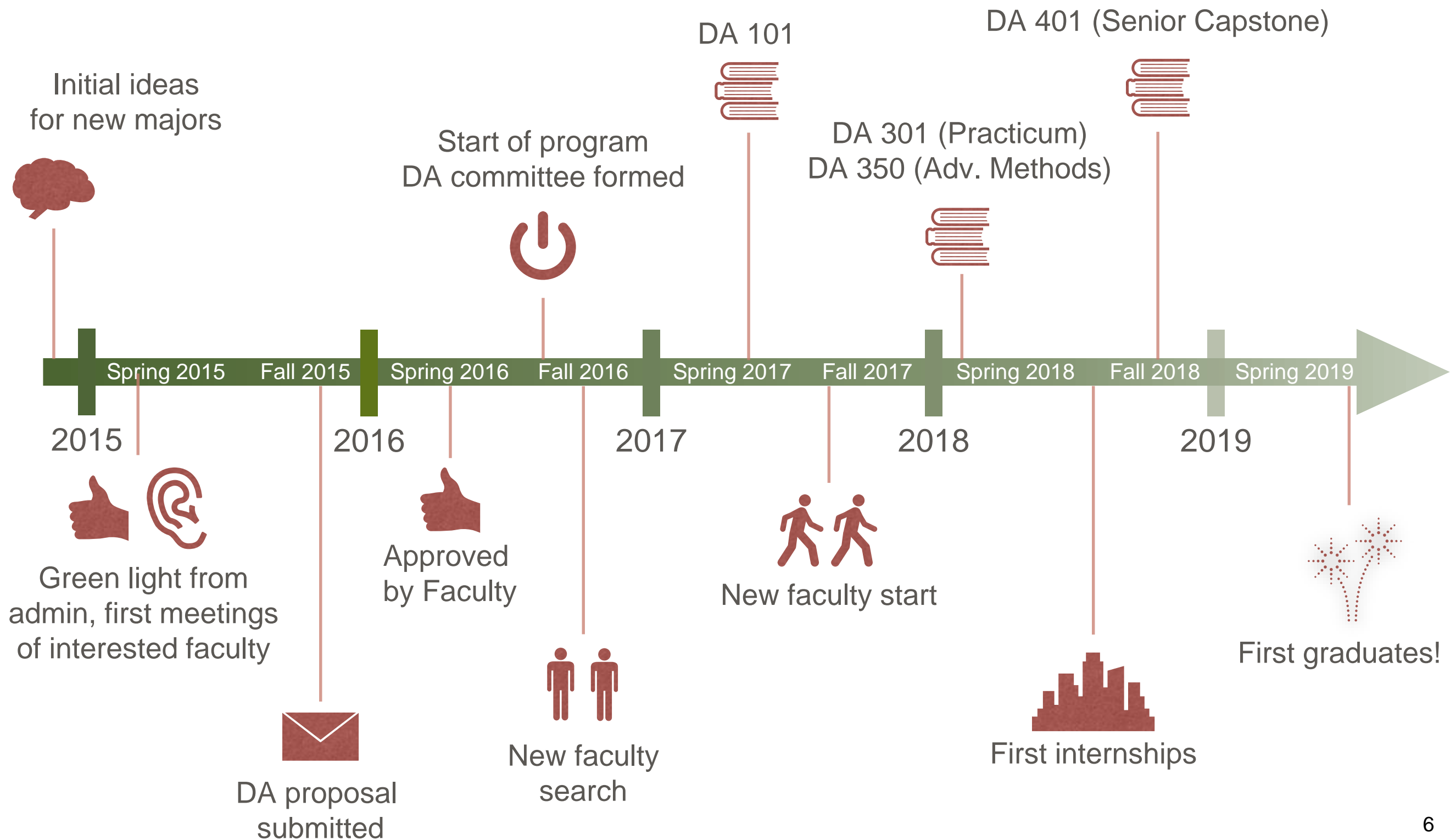
Data Analytics Major @ Denison

- attract *new* students to quantitative methods
 - breadth over depth
 - interdisciplinary vision
 - program independent of (but inv
- *leverage* liberal arts learning goals
 - encourage creativity
 - comfort with uncertainty
 - communication
 - ethical and social implications
 - disciplinary knowledge and connections among disciplines

Current program committee affiliations:

- Data Analytics (2)
- Biology
- Computer Science
- Economics
- Mathematics
- Philosophy
- Political Science
- Psychology

Data Analytics milestones



Data Analytics Major @ Denison

Four components of the major

1. Computer Science and Mathematics (4 courses)
2. Project-oriented Data Analytics courses (4 courses)
3. Choice of disciplinary focus (3 or 4 courses)
4. Summer internship

Data Analytics Major @ Denison

1. Computer Science and Mathematics (4 courses)
 - Discovering Computer Science (CS 111/112)
 - ❖ problem solving, programming, data storage/manipulation (Python)
 - ❖ two project-oriented “flavors”: natural sciences, social sciences
 - Data Systems (CS 181) **NEW**
 - ❖ data wrangling, databases, client/server, web (Python, SQL)
 - Calculus (Math 123)
 - Applied Statistics (Math 242)
 - ❖ calculus-based probability, experimental design, statistical inference, hypothesis testing, regression (R)

Data Analytics Major @ Denison

2. Project-oriented Data Analytics courses (4 courses)

- Introduction to Data Analytics (DA 101) **NEW**
 - ❖ DA cycle, types of data, wrangling, summarizing, visualizing, basic predictive, communication, ethics, social effects, open science (R)
- Practicum in Data Analytics (DA 301) **NEW**
 - ❖ end-to-end group projects with campus or community clients
- Advanced Methods in Data Analytics (DA 350) **NEW**
 - ❖ machine learning, regression, network analysis
- Senior Capstone (401) **NEW**
 - ❖ individually-designed project in student's interest area

Data Analytics Major @ Denison

3. Choice of disciplinary focus (3 or 4 courses)

- Economics
- Political Science
- Sociology
- Biology
- Physics
- Psychology
- Philosophy

Future possibilities

- Data Journalism
- Communication
- Linguistics/Textual Analysis
- Sports Analytics

Data Analytics Major @ Denison

4. Summer internship

- summer after junior year
- networking with alumni
- support of Knowlton Center for Career Exploration

Student interest in Data Analytics

- 75 declared majors so far
 - 28 juniors ('19), 28 sophomores ('20), 19 first years ('21)
 - informal survey suggests 25-30 majors per year is likely
 - 35% female
 - about half are interested in pursuing an Economics focus
- DA 101 enrollments
 - Spring 2017: 77 students in 4 sections
 - Fall 2017: 56 students in 3 sections
 - Spring 2018: 60 students in 3 sections

Projects-first approach

language feature or math topic or technique (“tools”)



example problems and/or applications

problem or question



principle/technique/algorithm



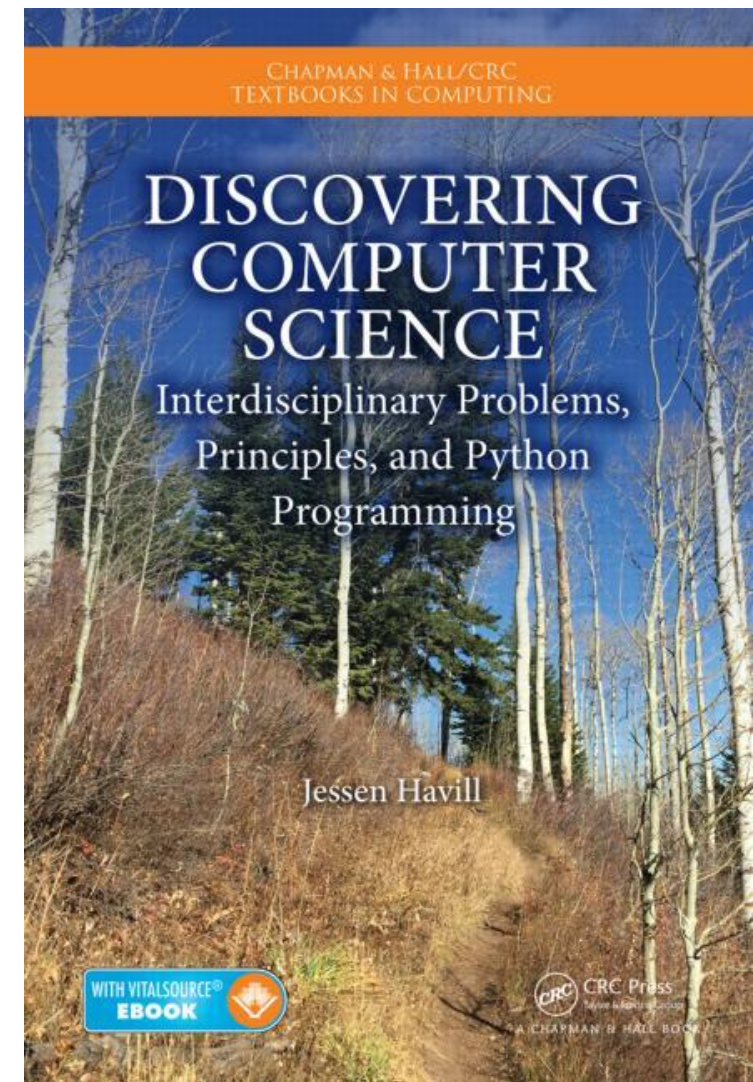
tools and language features

Projects-first approach

- goal 1: enlarge the tent
 - appeal to wider audience with questions from interest areas
 - involve the expertise of others as first-class partners
 - strive for better gender equity
- goal 2: deepen learning
 - project topics connect to prior knowledge
 - enables a natural spiral approach to DA concepts
 - emphasis on solving problems (and data analytics cycle)

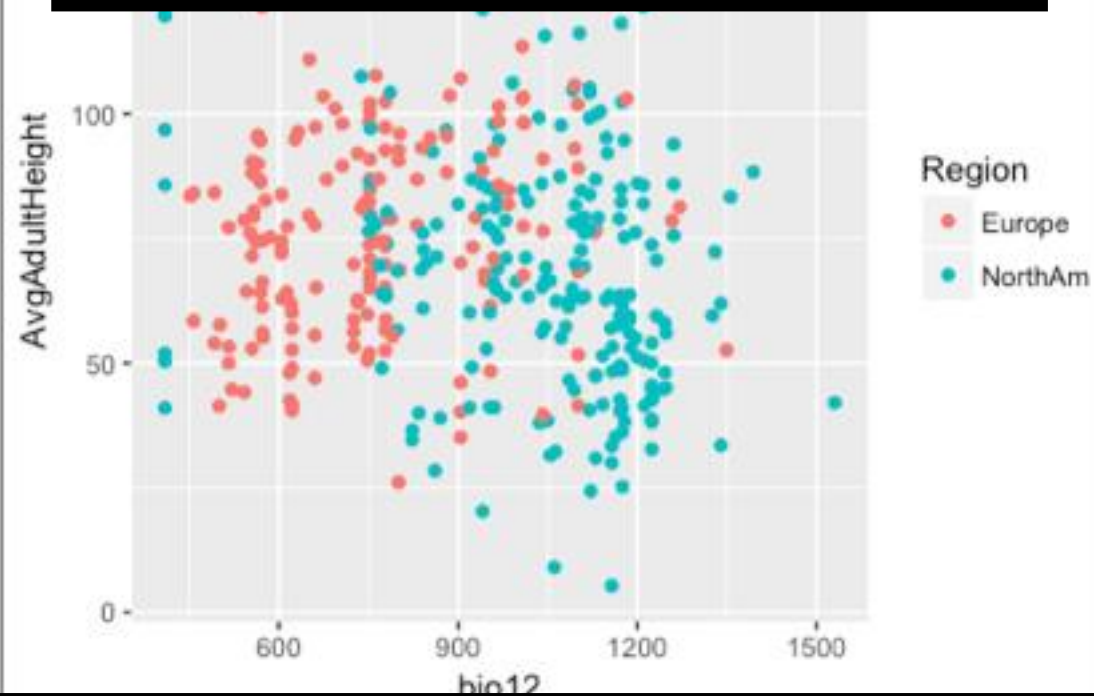
CS 111

- population modeling → iteration, accumulators, for loops, list mutation
- Brownian motion → Monte Carlo → PRNG
- text/genome analysis → linear read/search, string methods
- data analysis → accumulators, regression, covariance
- fractals in nature → divide & conquer, recursion
- searchable database → sort/search, quadratic
- each “unit” followed by a partner project with `discovercs.denison.edu`

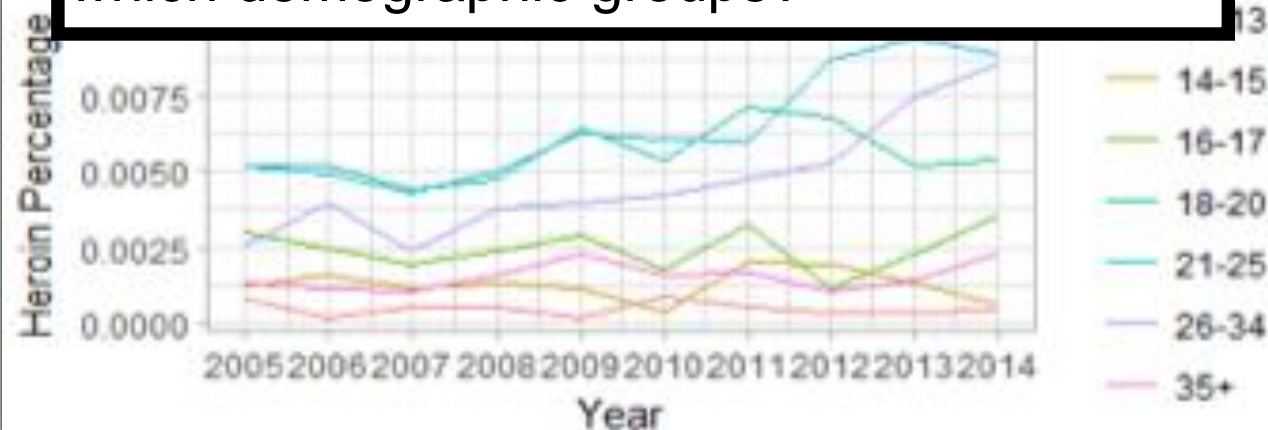


DA 101 Projects

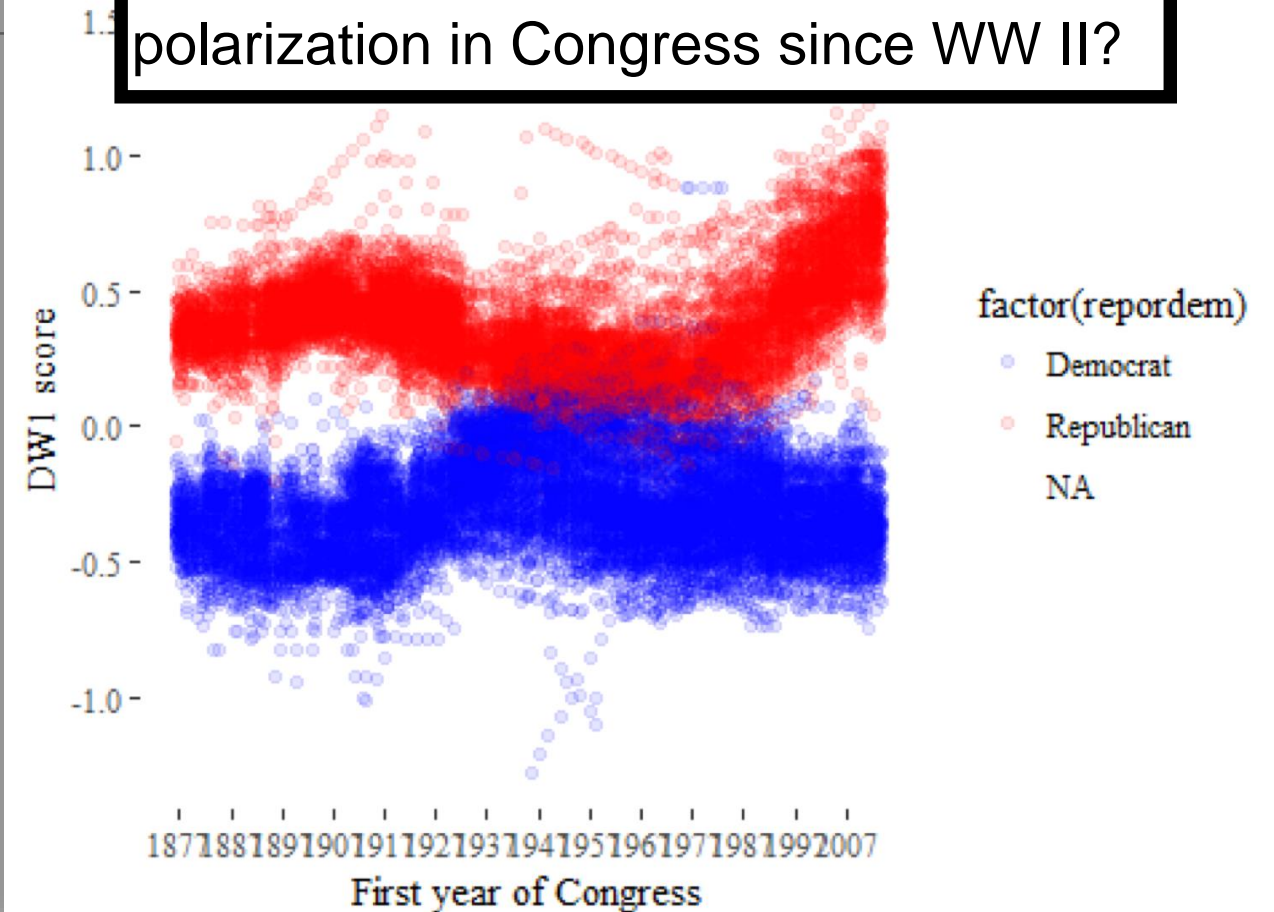
What variables impact the health of garlic mustard plants (weeds)?



On which drug types should the government focus prevention efforts? For which demographic groups?



What are the causes and correlates of polarization in Congress since WW II?



Regression EWP-WP for both halves of the season

```
##
## Call:
## lm(formula = minus2 ~ minus1, data = seasondata)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.12    -0.05    -0.02    -0.01    0.01
## Coefficients:
## (Intercept)
## minus1
## Residuals
## Multiple R-squared:
## F-statistic
```

Which player stats in the NBA best predict point scored/allowed in the past? Can we use this result to predict team success?

DA 301: the “Denison Practicum”

- end-to-end group projects for campus “clients”
 - Evaluate the usefulness of player and line-up metrics for men’s basketball to plan for next season.
 - Recommend changes to the campus library service model based on 20 years of reference requests.
 - Analyze energy consumption and recommend ways to cut costs.
 - Analyze the course scheduling time slot and classroom usage, and recommend improvements.
 - What factors might help us predict student attrition?

Challenges

- projects-first approach
 - finding projects
 - careful selection of projects and their sequencing
 - may sacrifice some depth
- program development
 - rolling out 301, 350, 401
 - internships
 - program space
 - teaching resources
 - continue campus and community outreach