Data Science @ Winona State

Speaker: Chris Malone (cmalone@winona.edu)
Winona State University
8,500 students
- 7000 Winona
- 1000 Rochester
- 500 graduate
8,500 students
- 7000 Winona
- 1000 Rochester
- 500 graduate

Department
- Data Science: 1 | 25
- Statistics: 7 | 50
8,500 students
- 7000 Winona
- 1000 Rochester
- 500 graduate
Department
- Data Science: 1 | 25
- Statistics: 7 | 50

Scale to 30000
35 faculty / 325 majors
We are not data scientists
Current Undergraduate Statistics Program

- Started about 30 years ago
- 7 faculty
- Nearly half of our graduates attend grad programs
Current Undergraduate Statistics Program

- Started about 30 years ago
- 7 faculty
- Nearly half of our graduates attend grad programs

- Moto: “Innovative and Dynamic Curriculum”
Current Undergraduate Statistics Program

- Started about 30 years ago
- 7 faculty
- Nearly half of our graduates attend grad programs

  - Moto: “Innovative and Dynamic Curriculum”
    - Introductory Course
    - Introductory Year
Current Undergraduate Statistics Program

- Started about 30 years ago
- 7 faculty
- Nearly half of our graduates attend grad programs

Moto: “Innovative and Dynamic Curriculum”

- Introductory Course
- Survey Sampling
- Multivariate
- Nonparametric

- Introductory Year
- Study Design
- Un/Supervised Learning
- Professional Skills
We are not data scientists
Last Program Review – 2010

Key Questions

• Is the curriculum evolving quickly enough?
• What characteristics should drive change?
• Are we sufficiently training students for employment and/or graduate school?
Timeline

**USCOTS**
[Raleigh, NC]
Survey of skill-sets necessary for employment vs. graduate school
May 2011

**JSM 2012**
[San Diego, CA]
Present initial proposal of new major. Obtain feedback from professional community
August 2012

**MinneAnalytics**
[Minneapolis, MN]
Present proposed major to working data science professionals at a regional level
January 2013

**CSP**
[New Orleans, LA]
Present proposed major to working data science professionals at a national level
February 2013

**JSM 2013**
[Montreal, QC]
Present job posting outcomes and revised major. Obtain feedback from professional community
August 2013

**Campus Visit**
[Winona, MN]
Visit from data science expert to evaluate major
September 2013

**Interviews with Industry Reps**
Conducted personal interviews with 19 reps from industry to collect desired skill-set for entry level employment

**Job Posting / Revise Major**
Reviewed 130+ job postings for necessary skills. Seek additional input and revise major

**WSU Community**
Communication with administration and impacted Departments
January 2014

**Proposal Submitted**

**System-level approval**
Fall 2014

**Program Launch**
April 2015

**Curriculum Development / Program Implementation**
May 2016
• Started in 2012

• Models
  - ASA Data Analysis Competition of mid 1990s
  - COMAP: Mathematical Contest in Modeling
  - DATAFEST: Rob Gould, UCLA
Key Features of MUDAC

• Corporate partner motivates the problem, fields questions during the competition, and participates in the judging

• Teams of students have 24 hours to provide a solution

• Teams are judged/interviewed by corporate partner representatives, data science professionals, and faculty advisors

• Professional development for students (and advisors)

• On-site job fair
MUDAC 2016

- Teams from 17 different colleges/universities (currently limited to 2 teams per school)
- 100 students, 25 advisors, 40 working data science professionals assist with judging
- Budget: $12,500
MUDAC has convinced us that

- undergraduates can do data science
- strong demand exists for students who possess these skills
- developing a comprehensive undergraduate data science program should be done
We are not data scientists
Data Science

Computational

Analytical

Domain / Professional Skills
Computational (17 credits)

Computer Science
- (3) Intro to Computer Science
- (8) Year-long sequence in object-oriented, high-level programming
- (3) Database Course

Data Science
- (3) DSCI 325: Management of Structured Data
Analytical (18 credits)

Data Science
• (6) Year-long sequence in data science, e.g. software, basic programming, data management, visualization

Statistics
• (6) Year-long sequence of introductory statistics
• (3) Regression Analysis

Mathematics
• (3) Applied Calculus
Domain & Professional Skills (18 credits)

Interdisciplinary component
• (12) Any discipline, 300+ level courses only

Professional Skills
• (3) Career development, communication, etc.

Capstone / Internship
• (3) Substantial independent project or internship
We are not data scientists
Data Science

Computational

Analytical

Domain / Professional Skills
1. Balance
2. Develop skills
3. Students do

Data Science

- Computational
- Analytical
- Domain / Professional Skills
~ Thank You ~

Speaker: Chris Malone (cmalone@winona.edu)
Winona State University
(Very) Slow to Change

<table>
<thead>
<tr>
<th>Years Ago</th>
<th>Author</th>
<th>Statement / Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>ASA President Wilks</td>
<td>“As business and industry become more complex executives will depend more and more on scientific statistical methods for collecting, analyzing, and interpreting information for decision-making.”</td>
</tr>
<tr>
<td>33</td>
<td>Minton</td>
<td>Suggests training undergraduates is necessary to meet the demands of industry.</td>
</tr>
<tr>
<td>18</td>
<td>Higgins</td>
<td>Encourages the training of statisticians at the undergraduate level and proposes new curriculum</td>
</tr>
<tr>
<td>15</td>
<td>Moore</td>
<td>“our future depends on achieving a more prominent place in undergraduate education beyond the first methods course.”</td>
</tr>
</tbody>
</table>
DSIC Example

Voice Mail Preview:

The following message is brought to you by school reach.

Hello from asphyxiation 12 taken either packing schools will be attending so I was great to go back at 10:25 every now and preschool kids clamber be okay at 8:30.

Actual Voice Mail Message
DSIC Example

**Message per system**
The following message is brought to you by school reach. Hello from asphyxiation 12 taken either packing schools will be attending so I was great to go back at 10:25 every now and preschool kids clamber be okay at 8:30.

**Message per human ear**
The following message is brought to you by School Reach. Hello from Rushford Peterson schools. Due to the weather R P schools will be starting 2 hours late tomorrow. School will start at 10:25. There will be no am preschool. Kids Club will be open at 8:30.