

## Term Information

Effective Term Spring 2020

## General Information

Course Bulletin Listing/Subject Area Statistics  
Fiscal Unit/Academic Org Statistics - D0694  
College/Academic Group Arts and Sciences  
Level/Career Undergraduate  
Course Number/Catalog 4911  
Course Title Data Analytics Capstone  
Transcript Abbreviation DA Capstone  
Course Description A teamwork-based synthesis of the Data Analytics major curriculum through the analysis of data supplied by a partnering institution. Prepares students for the complexity of data analysis they will encounter outside of the university in a mentored setting.  
Semester Credit Hours/Units Fixed: 4

## Offering Information

Length Of Course 14 Week  
Flexibly Scheduled Course Never  
Does any section of this course have a distance education component? No  
Grading Basis Letter Grade  
Repeatable No  
Course Components Lecture  
Grade Roster Component Lecture  
Credit Available by Exam No  
Admission Condition Course No  
Off Campus Never  
Campus of Offering Columbus

## Prerequisites and Exclusions

Prerequisites/Corequisites 4620 or permission of instructor.  
Exclusions  
Electronically Enforced Yes

## Cross-Listings

Cross-Listings

## Subject/CIP Code

Subject/CIP Code 27.0501  
Subsidy Level Baccalaureate Course  
Intended Rank Senior

## Requirement/Elective Designation

Required for this unit's degrees, majors, and/or minors

## Course Details

### Course goals or learning objectives/outcomes

- Use statistical models and concepts to analyze and draw conclusions from data.
- Apply computer science principles relating to data representation, retrieval, programming and analysis.
- Use critical thinking skills to identify and solve problems, make decisions, assess value propositions supported by data, and generate logical syntheses of information from data.
- Transfer knowledge gained from one area to problems and data in another.
- Communicate findings and their implications.

### Content Topic List

- Introduction to data subject area
- Exploratory data analysis
- Data ethics
- Computing workflows for data analysis
- Problem identification
- Project scoping and planning
- Statistical modeling and analysis
- Model criticism and refinement
- Report writing
- Presentations

### Sought Concurrence

No

## Attachments

- stat4911\_Capstone.pdf  
*(Syllabus. Owner: Hans,Christopher M)*
- DACoreCurriculumMap.pdf: curriculum map  
*(Other Supporting Documentation. Owner: Hans,Christopher M)*

## Comments

- Deorah: I have uploaded the curriculum map for the DA major core curriculum. *(by Hans,Christopher M on 02/26/2019 03:19 PM)*
- 02/20: Chris: We will need a curriculum map for this new course. If there is a curriculum map for DA, use that one. If not, please identify this course in the STAT curriculum map so that we know which program goals are met at what level -- introductory, intermediate, or advanced. *(by Haddad,Deborah Moore on 02/20/2019 07:03 PM)*

**COURSE REQUEST**  
4911 - Status: PENDING

Last Updated: Haddad,Deborah Moore  
02/26/2019

**Workflow Information**

Status	User(s)	Date/Time	Step
Submitted	Hans,Christopher M	01/15/2019 05:55 PM	Submitted for Approval
Approved	Lee,Yoonkyung	02/20/2019 04:34 PM	Unit Approval
Revision Requested	Haddad,Deborah Moore	02/20/2019 07:03 PM	College Approval
Submitted	Hans,Christopher M	02/26/2019 03:19 PM	Submitted for Approval
Approved	Lee,Yoonkyung	02/26/2019 04:55 PM	Unit Approval
Approved	Haddad,Deborah Moore	02/26/2019 05:57 PM	College Approval
Pending Approval	Nolen,Dawn Vankeerbergen,Bernadette Chantal Oldroyd,Shelby Quinn Hanlin,Deborah Kay Jenkins,Mary Ellen Bigler	02/26/2019 05:57 PM	ASCCAO Approval

# Syllabus | STAT 4911

## Data Analytics Capstone

---

### Course Staff

#### Instructor:

- Email:
- Office:
- Office Hours:

#### Teaching Assistant:

- Email:
- Office: DALC, Pomerene Hall 151
- Office Hours:

### Location and Time

The course will consist of two 110-minute lectures for a total of 4 credit hours.

- **Lecture:** Tuesday and Thursday, 12:45 PM - 2:35 AM, Pomerene Hall (PO) 0161

### Prerequisites

STAT 4620: Introduction to Statistical Learning, or permission of instructor.

### Course Materials

#### Readings

There is no required textbook for the course. A list of freely available readings will be posted to the course website.

#### Computing

The R programming language (<https://www.r-project.org/>) and RStudio IDE (<https://www.rstudio.com/>) will be used throughout the course. Both R and RStudio are free software that you are encouraged to install on your personal machines.

#### Website

- <https://carmen.osu.edu>

## Course Description

A teamwork-based synthesis of the Data Analytics major curriculum through the analysis of data supplied by a partnering institution. Prepares students for the complexity of data analysis they will encounter outside of the university in a mentored setting.

## Learning Outcomes

Upon successful completion of the course, students will be able to...

1. *Use* statistical models and concepts to analyze and draw conclusions from data.
2. *Apply* computer science principles relating to data representation, retrieval, programming, and analysis.
3. *Use* critical thinking skills to identify and solve problems, make decisions, assess value propositions supported by data, and generate logical syntheses of information from data.
4. *Transfer* knowledge gained from one area to problems and data in another.
5. *Communicate* findings and their implications.

## Grading

### Grading Scale

A	A-	B+	B	B-	C+	C	C-	D+	D	D-
93%	90%	87%	83%	80%	77%	73%	70%	67%	63%	60%

## Assessments and Course Activities

For each individual assignment or project, specific policies and directions will be included with the release of an assignment, or in advance of a project.

- **Participation:** Occasionally there will be in-class activities which are turned-in for a completion grade.
- **(RQ) Reading Quizzes:** A number of short readings will be assigned throughout the semester. Short, comprehension quizzes will be completed in Carmen. These readings are intended to help students better understand the larger data science world.
- **(CQ) Concept Quizzes:** There will be **five** in-class quizzes on high-level course concepts. These quizzes are meant to help prepare students for the type of concept questions seen during an interview process.
- **(PC) Programming Challenges:** There will be **five** “programming challenges” meant to help prepare students for the type of technical questions seen during an interview process.
- **(PP) Practice Project:** There will be **one** practice *group* project that involves analyzing publicly available data. The project may have several smaller assignments including but not limited to: project proposals, EDA reports, progress reports, final reports, peer reviews, and presentations.
- **(MP) Midterm Projects:** There will be **two** midterm *group* projects that involve analyzing data provided by a partner institution. Each project will have several smaller assignments including but not limited to: project proposals, EDA reports, progress reports, final reports, peer reviews, and presentations.
- **(FP) Final Project:** There will be **one** final *group* project that involves analyzing data provided by a partner institution. The final project will have several smaller assignments including but not limited to: project proposals, EDA reports, progress reports, final reports, peer reviews, and presentations.

## Grade Components

Type	Percentage
Participation	5
Readings	5
Concept Quizzes	5
Programming Challenges	5
Practice Project	5
Midterm Project I	20
Midterm Project II	20
Final Project	35

## Outline of Topics and Activities

Week	Main Activity	Deadlines
1	Overview	
2	Practice Project	RQ1
3	Practice Project	PP
4	Midterm Project I	RQ2
5	Midterm Project I	PC1, CQ1
6	Midterm Project I	PC2, CQ2
7	Midterm Project II	MP1
8	Midterm Project II	RQ3, PC3, CQ3
9	Midterm Project II	MP2
10	Spring Break	
11	Final Project	RQ4
12	Final Project	PC4, CQ4
13	Final Project	RQ5
14	Final Project	PC5, CQ5
15	Final Project	FP
16	End of Course	

## Course Policies

### Email Communications

In order to protect your privacy, all course e-mail correspondence must be done through a valid Ohio State University name.## account. The subject line of the email must begin with **[STAT 4911]** followed by an informative message.

If an email is sent between 9:00 AM Monday and 11:59 PM Thursday, and you follow the above directions, the staff will make an effort to respond within 24 hours. Homework questions sent the same day a homework is due run the risk of not receiving a response before the homework is due. Plan accordingly and consider office hours as a more reliable alternative.

### Data Analytics Learning Center

Graduate teaching assistants (GTAs) for STAT 3201, 3202, 3301, 3302, 3303, 4620, and **4911** will hold their office hours in the Data Analytics Learning Center (DALC) in Pomerene 151. The hours during which the

GTA and grader for our course will hold office hours in PO 151 can be found at the top of the syllabus. You can meet with the GTAs for our course in the DALC during their office hours to discuss questions you have about the course material, homework assignments, R, etc.

You are welcome to stop by the DALC when it is open even if it is not currently being staffed by the GTA for our course, e.g. if you are looking for a place to study or work on an assignment for one of the supported courses. If the DALC is staffed by a GTA for another Statistics course when you stop by, they will help you if possible, but may not be able to answer all of your questions.

A complete list of hours during which the DALC will be staffed by GTAs for Statistics Department courses can be found at <https://data-analytics.osu.edu/dalc>.

In rare situations due to last minute emergencies, the GTA assigned to the DALC may not be able to attend their office hours. If the DALC is closed when the schedule indicates it should be open, we recommend waiting for a few minutes. If no one shows up in a reasonable amount of time, please email your instructor to let us know about the problem. You can also contact your GTA to see about arranging a make-up time to meet.

### **Grade Disputes**

If you feel a homework, exam, lab, or project was graded incorrectly, you have **one week** from the date it was returned to discuss it with the course instructor. After one week, grading is final except for exceptional circumstances. You may not simply ask for a re-grade, but instead must justify to the staff why the grading was done incorrectly. Also, by disputing any grading, you agree to allow the course staff to review the entire assignment or exam for other errors missed during grading.

### **Attendance**

You are expected to attend all lectures. Failure to do so may not have a direct effect on your course grade, but will likely have a significant indirect effect. Any known or potential extracurricular conflicts should be discussed in person with the instructor during the first week of class, or as soon as they arise.

### **Academic Misconduct**

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term “academic misconduct” includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct <http://studentlife.osu.edu/csc/>.

### **Special Accommodations**

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: [slds@osu.edu](mailto:slds@osu.edu); 614-292-3307; [slds.osu.edu](http://slds.osu.edu); 098 Baker Hall, 113 W. 12th Avenue.

## **Policy Changes and Updates**

Within reason and with advance notice, the instructor reserves the right to make any changes that are considered academically advisable. Such changes, if any, will be announced in class. Please note that it is your responsibility to attend the class and keep track of the proceedings. These changes are likely to contain minor changes to the schedule, however no changes will be made to the exam dates.

---

## Data Analytics Major: Core Curriculum Map

### Outcomes from the Core Courses

- M.1** Students will demonstrate an understanding of and ability to apply computer science principles relating to data representation, retrieval, programming and analysis.
- M.2** Students will demonstrate an understanding of and ability to apply mathematical and statistical models and concepts to detect patterns in data, as well as draw inferences and conclusions supported by the data.
- M.3** Students will demonstrate critical thinking skills associated with problem identification, problem solving and decision-making, assessing value propositions supported by data, and generating a logical synthesis of information from data.
- M.4** Students will demonstrate the ability to apply knowledge gained from one area to problems and data in another.
- M.5** Students will demonstrate the ability to communicate findings and their implications, and to apply them effectively in organizational settings.

### **Program outcomes (*B=beginner, I=intermediate, A=advanced*)**

<b>Course</b>	<b>M.1</b>	<b>M.2</b>	<b>M.3</b>	<b>M.4</b>	<b>M.5</b>
Math 1151		B			
Math 1152		B			
Math 2568		I			
CSE 2221	B				B
CSE 2231	I		B		B
CSE 2321	I	B	B		B
CSE 2421/3430	B		B		B
CSE 3241	A		I		I
STAT 3201		B			B
STAT 3202		B	B		B
ISE 3230		I	I		I
STAT 3301	B	I	I	B	I
STAT 3302	B	A	I	B	I
STAT 3303		I	A	I	I
STAT 4620	B	A	I	B	I
CSE 5242	A		A	B	A
CSE 5243	A	I	A	I	A
CSE 5544 or ISE 5760	CSE: A ISE: I	ISE: I	A	I	A
STAT 4911 (Capstone*)	A	A	A	A	A

\*Students in the Biomedical Informatics, Social Science Analytics, and Data Visualization Specializations use Stat 4911 to satisfy the major’s capstone requirement. Students in the Computational Analytics Specialization may satisfy the capstone requirement by taking either Stat 4911 or a CSE 591x capstone course. Students in the Business Analytics Specialization satisfy the capstone requirement by taking BUSADM 3630.05 and 3632.05.