

Big Data in your pocket: call it a smartphone

Mina Mila

Dept of Entomology and Plant Pathology
North Carolina State University, Raleigh, NC, USA

almila@ncsu.edu

Description: How can you think like Leonardo Da Vinci in the 21st century? Leonard Da Vinci, one of the greatest minds of all times set the principles for most of modern sciences in an era that very little information was available. Nowadays plenty of information is available. Data come in all forms, volumes, and means. It is universally accepted that we live in the era of the information revolution. How can you think like Leonardo Da Vinci in today's world? This course is intended as an introduction to Data Science and its applications. It is designed for any undergraduate level students with little background in math and statistics. This is a 3-credit hour survey course.

Learning Objectives: Successful completion of this course will enable students to describe the types of data generated in different disciplines, compare and contrast how data were collected in the past, present and future, describe how developments in technology changed the role of data in innovation, describe how visualization is a major tool to summarize vast amount of data, and describe how large amounts of data will change the way we will generate knowledge in the future.

Textbook: No textbook is required for this course. The course is designed to build on information provided in the lectures with additional readings of the scientific literature. Such readings will be provided prior to the lectures.

Schedule: Monday to Friday for 2 weeks (June 11th – June 24th, 2017). 4 hours each day. There will be some extra meetings beyond in-class time to discuss final project requirements and progress.

Credits: Students will receive a transcript with the equivalent of 3 units (US) or 4.5 ECTS credits (Europe).

Laptop: Since the course is designed to be hands on, students will be required to bring a laptop. You should also buy in advance a European plug converter. Most laptop chargers should accommodate the European 220 Volts, however a European plug converter will still be needed.

Grading Policy: The final grade in this course will be computed based on the following percentages. Exams are 75 minutes at the beginning of indicated lectures.

Exam 1: 15%
Exam 2: 15%
Exam 3: 15%
Participation: 15%
Project: 40%

Unbounded Prometheus

Grading Scale:

100-98	=	A+
97.9-91	=	A
90.9-89	=	A-
88.9-87	=	B+
86.9-81	=	B
80.9-79	=	B-
78.9-77	=	C+
76.9-71	=	C
70.9-69	=	C-
68.9-67	=	D+
66.9-61	=	D
60.9-59	=	D-
<59	=	F

Assignments: There will be assignments for some lectures. Assignments should be brought in class the day they are due and they count towards the 15% of participation grade.

Final project: A final project will be assigned and it will be due 3 weeks after the completion of the course. Possible topics for the final project will be presented and discussed in class by the instructor. Students are encouraged to propose their own projects that might be related to something they have already been working on or something that they have always found interesting/appealing. Such proposal for final projects will have to be approved by me before the end of the course.

Tentative Lecture Schedule

Day 1: All about data

Data definition, Examples of data, Types of data in social sciences and methods of collection- Types of data in engineering and methods of collection – Types of data in life sciences and methods of collection

Day 2: Data and models

What is a model? Relationship between data and models in different disciplines, models in society

Day 3: Explore the world around us via data and models

Class will visit an archeological (plenty in Kavala!) or other location of interest. Class will be divided in teams and will be assigned to a brief exercise on information, data, and models related to the location visited. Exercise outcomes will be discussed back in class.

Day 4: Where is today's Data? Look at your smartphone to find some

EXAM 1 (first 75 mins of lecture)

Big Data definition. Big data and social media with examples: Does Facebook know you better than you do? Which jacket you want to buy? Can DNA tell when I will get sick? What questions your smartphone can or cannot answer

Day 5: Big Data and models

Redefining the relationship between data and models; Case study I - Big Data in Medical applications; Case study II –Big Data in Ag Business

Day 6:

Opportunities and Challenges with Big Data for Science and Society

EXAM 2 (last 60 mins of lecture)

Day 7: Data Visualization

How to put a million data in a picture, Storytelling

Day 8: Introduction to Tableau, Practice with tableau (use data to extract information and create a model)

Day 9: Practice with Tableau (storytelling)

Day 10 Misuses of data Viz, Critique on students' vizs,

EXAM 3 (60 mins)

Wrap-up course