



STAT 101 Introduction to Data Science

Course Description

This is an introductory course to the field of data science and roles of a data scientist. Topics will include: history of data science, ethics, available tools, methodology, visualization and pathways to data science careers. This course will expose students to applications in various fields and prepare them to be critical thinkers of future career paths.

Illinois Articulation Initiative (IAI) number: N/A

Credit and Contact Hours:

Lecture	3
Lab	0
Credit Hours	3

Prerequisites: Satisfactory placement test score or grade of “C” in Math 090 or equivalent.

Books, Supplies, and Supplementary Materials

A. Textbooks

O’Neil, Cathy. *Weapons of Math Destruction*. Broadway Books, NY.

Wickham, H. and Golemund, G., *R for Data Science*. Riley. <https://r4ds.had.co.nz/>

B. Other Required Materials

Excel Software, R free download

Methods of Instruction:

Hybrid, Online, Face-to-Face

General Education Student Learning Outcome

Applied Knowledge: Students draw from learning experiences/concepts to solve a variety of problems or challenges.

Course Learning Outcomes (CLOs)

1. Students recognize and explain the central role of data scientists.
2. Students apply critical thinking to understand statistical reasoning.
3. Students examine importance of ethics in research.

4. Students describe importance of methodology.
5. Students organize the cycle of data analysis: understand, prepare, model, evaluate.
6. Students present findings from analysis cycle.

Lesson Learning Outcomes (LLOs)

1. Students describe big data.
2. Students discuss examples of data science.
3. Students explain how data is used in a wide range of fields, including those that are traditionally part of the liberal arts.
4. Students describe the difference between data science tools available to them.
5. Students calculate the probability of basic events in real-life data.
6. Students calculate descriptive statistics on data from a variety of disciplines.
7. Students consider case studies.
8. Students discuss identity, privacy, ownership, and reputation.
9. Students identify a problem which needs data influence to solve.
10. Students explain possible solutions.
11. Students identify available data sources.
12. Students find a data set and prepare the data.
13. Students identify types of data.
14. Students find sources of data.
15. Students extract data from data sources.
16. Students explain storage of data.
17. Students illustrate structured versus unstructured data.
18. Students prepare solution documents.
19. Students justify solutions orally.
20. Students present visualizations.
21. Students identify building blocks of R.
22. Students identify workspace and files in R.
23. Students evaluate sequences of numbers in R.
24. Students evaluate vectors in R.
25. Students evaluate Missing Values in R.
26. Students evaluate subsetting vectors in R.
27. Students evaluate matrices and data frames in R.
28. Students evaluate logic in R.
29. Students evaluate functions in R.

TOPICAL OUTLINE/TIMELINE

Week	LLOs Learned	CLOs Learned
Week One	1,2	1
Week Two	3,4	1
Week Three	5,6	2
Week Four	7	2
Week Five	7	2,3,4
Week Six	8	3
Week Seven	8	3
Week Eight	9,10	4
Week Nine	11,12,13,14, 21, 22	5
Week Ten	11,12,13,14, 23	5
Week Eleven	15,16,17, 24,25	5
Week Twelve	15,16,17, 26	5

Week Thirteen	18, 27	6
Week Fourteen	18, 28	6
Week Fifteen	19,20, 29	6
Week Sixteen	19,20	6

Grading Assignments

Quizzes	11 quizzes at 5 points each = 55pts
Discussion Boards	13 weekly discussions at 5 point each + one 10point assignment = 75pts
Module Activity Projects	Various projects in intro module to module 5 = 125pts
Case Study Projects	15 each, in modules 6 to 10 = 75pts
Individual Projects	Various projects in midterm modules to module 10 = 110pts
Final Project	150 points
Peer Evaluation	Two evaluations at 25 points each = 50pts
TOTAL	640 points

Final Course Grading Scale

Grade	Percentage
A	90-100%
B	80-89%
C	70-79%
D	60-69%
F	lower than 60%

Faculty Commitment

Faculty members are committed to providing a quality learning experience through thoughtful planning, implementation, and assessment of course activities. They are also committed to being readily available to students throughout the semester by returning e-mails and phone calls within 48 hours and to returning graded course work within a week. Furthermore, they are committed to selecting appropriate course materials and making them available in an organized and timely manner.

Student Commitment

For every credit hour a student is enrolled in, they should expect to spend at least 2 hours outside of class studying, working on assignments, and preparing for class each week of the fifteen-week semester. For example, for this four credit-hour class, students can expect to spend four hours per week in class actively engaged in learning the material by participating in face-to-face classes or viewing lectures and instructional material online. In addition, students should expect to spend another eight hours per week outside of class completing homework and assignments, posting to discussion boards online, or studying for quizzes and tests. This means students should spend a minimum of 12 hours per week engaged in achieving the learning outcomes for this course. If you are not achieving your desired results in this class, you should consider increasing your prep time outside of class, in addition to using available resources such as instructor office hours and tutoring services.

By registering for this course, you commit yourself to active participation in course activities as well as the submission of all assignments and exams on time. Furthermore, you commit to accessing the course site and checking your JJC e-mail several times a week.

TOPICAL OUTLINE

Unit	Topic	Assignments
1	Data science in the world <ul style="list-style-type: none"> • Research the field of data science • Describe how data science relates to the world and multiple areas of study 	Orientation quiz Homework Project: Short research paper describing data science and its influence in a particular area of interest.
2	Data science tools <ul style="list-style-type: none"> • Describe the common tools used in the data science field. 	Homework Project: Short research paper describing common data science tools
3	Descriptive statistics <ul style="list-style-type: none"> • Define descriptive statistics • Describe the need for descriptive statistics • Review the basic concepts of descriptive statistics 	Quiz: Topics to date Homework Project: Complete a project using descriptive statistics.
4	Data analysis <ul style="list-style-type: none"> • Locate websites and other areas of study with available data. • Define quick steps needed to use the provided data ethically and legally • Define the technologies needed to analyze the data 	Homework Project: Short analysis of real-world data
5	Data analysis 2 <ul style="list-style-type: none"> • Locate more complex data sets • Perform more complex analysis 	Homework Project: Longer analysis of real-world data
6	Ethics in Research <ul style="list-style-type: none"> • Define ethics • Identify the importance of ethical practices in research 	Quiz: Topics to date
7	Ethics in Research 2 <ul style="list-style-type: none"> • Review and evaluate the ethics of various case studies 	Homework Project: Short research paper describing the ethics of a case study
8	Midterm exam	
9	Real-world need <ul style="list-style-type: none"> • Identify statistics needed to address real-world need in hypothetical situations • Identify visualizations needed to address real-world need in hypothetical situations 	Homework Project: Use real-world problem-solving techniques to identify a hypothetical situation. -Swirl Activity 1,2 in R

10	<p>Cycle of Data Analysis - Understanding</p> <ul style="list-style-type: none"> • Identify the cycle of data analysis • Define the four parts of the cycle: understand, prepare, model, evaluate 	<p>Homework Project: Outline research topic -Swirl Activity 3 in R</p>
11	<p>Cycle of Data Analysis - Prepare</p> <ul style="list-style-type: none"> • Describe the concept of preparing data • Define the importance of preparing data • Practice preparing data using common data preparation tools and techniques 	<p>Quiz: Topics to date</p> <p>Homework Project: prepare data for research topic -Swirl Activity 4,5 in R</p>
12	<p>Cycle of Data Analysis - Model</p> <ul style="list-style-type: none"> • Describe the concept of modelling data • Define the importance of modelling data • Practice modelling data using common tools and techniques 	<p>Homework Project: model data using a common data modelling technique for research project -Swirl Activity 6 in R</p>
13	<p>Cycle of Data Analysis - Evaluate</p> <ul style="list-style-type: none"> • Describe the concept of evaluating data • Define the importance of evaluating data • Practice evaluating data using common data evaluation tools and techniques 	<p>Homework Project: evaluate data using a common data evaluation technique for research project -Swirl Activity 7 in R</p>
14	<p>Visual data displays</p> <ul style="list-style-type: none"> • Describe the importance of presenting data visually • Implement a visual data presentation using a common data visualization tool 	<p>Quiz: Topics to date -Swirl Activity 8 in R</p>
15	<p>Presentations of data analysis</p> <ul style="list-style-type: none"> • Describe the importance of presenting data via an oral presentation • Present a data analysis project as an oral presentation. 	<p>Homework Project: Present an in-class data analysis project, using both oral and visual techniques. -Swirl Activity 9 in R</p>
16	<p>Presentations of data analysis 2</p> <ul style="list-style-type: none"> • Describe the importance of presenting data via an oral presentation • Present a data analysis project as an oral presentation. 	<p>Homework Project: Present an in-class data analysis project, using both oral and visual techniques.</p>
17	Final presentation	