

## DSCI 601: Introduction to Data Science (3 Credit Hours)

Spring 2026: 01/05/2026 – 05/01/2026

### Course Staff

#### Developer and Instructor

Dr. Aihua Li

Assistant Professor in the School of Earth, Atmosphere, and Sustainability, and faculty affiliate in Data Analytics and Data Science Program, Ball State University

#### Facilitator: TBD

Coursera Response Staff: Coursera provides numerous staff who respond to Chat requests and provide a first-line response to email requests, 24 hours, 7 days a week. Your question may be answered directly by the Coursera staff, especially if it is a platform-related question. Any content-related questions are sent to the instructor/facilitators. If you have technical questions, you can email us using [ds.support@bsu.edu](mailto:ds.support@bsu.edu).

### Course Description

We reside in a world experiencing an explosion of information, with a rapid and exponential growth of data. This surge in data captures increasing interest across various fields. Data science involves the gathering of extensive data and the fusion of domain expertise, programming skills, mathematics, and statistical knowledge to derive meaningful insights. Given the breadth and depth of data science, this course aims to furnish you with a comprehensive theoretical foundation and framework to initiate your journey in this field. "Data" permeates every aspect of data science. The course is divided into five parts, each centered around core topics related to "data". The initial part introduces data ethics, outlining the ethical issues surrounding data collection, usage, and reporting. The second part delves into data collection, acquisition sources, and data structures. The third part focuses on cutting-edge research in Data Science, immersing you in the realm of data science. The fourth part acquaints you with basic data processing using programming, specifically in R, the prevailing data analytics tool. Here, you will gain familiarity with R fundamentals, execute basic data wrangling tasks, develop an understanding of data storage and management, and gain experience in data visualization. The fifth part of the course imparts fundamental knowledge of probability and statistics, preparing you to move to the next stage of exploration.

### Course Objectives

This course introduces a few fundamental topics in the field of data science including ethical data-driven decision making, data acquisition process, basic data management steps and an introduction to probability and statistics. Students will be able to:

- Outline and identify ethical concerns, such as informed consent, privacy, ownership, transparency, fairness etc., in the field of data science.

- Explain and describe data sources, acquisition ways and storage structures.
- List and state current research interests in the field of data science.
- Outline and discuss basic statistical knowledge, programming skills and data visualization techniques required in data science.
- Conduct basic data manipulation, data visualization and simulation with R.

## Course Format

- There is a total of 16 modules for the 16-week semester.
- The first part of the course (Modules 1-7) focuses on key aspects concerning data itself, including data ethics, data collection, and the potential applications of data. In the latter half, the course shifts its focus towards providing foundational statistical knowledge and programming skills for data analysis.
- Learners should allocate 6-9 hours for each module, which is the standard workload expected for a three-credit graduate course.
- R/RStudio in windows is the utilized programming language in the latter half of the course.
- There is no textbook required for the course. The course materials consist of readings from journal articles, PowerPoints, online and other resources. The readings will be given aligned with the topics of each module.
- In the course, the learning activities and assessment measures in each module will include some kinds of activities below.
  - **Lecture videos** cover basic theory and concepts.
  - **Lab demo videos** provide demonstrations for possible quizzes about R syntax and programming assignments. This kind of learning activity will occur in the latter part of the course.
  - **Reading assignments** are given as complementary materials to enhance your further understanding of concepts and skills.
  - **Practice labs** are to give your practices after the lab demo videos. This kind of learning activity will occur in the latter part of the course.
  - **Quizzes** are given to test concepts in lectures and/or in the lab demo videos for programming code expressions.
  - **Programming assignments** are given to assess your programming skills after the lab demo videos and practice labs. This kind of learning activity will occur in the latter part of the course.

## Course Content

"Data" permeates every aspect of data science. All the topics in data science are where data comes from, how ethical the data is, where to store the data, which field data can be applied in and how to analyze data. All these questions will be simply answered in the course.

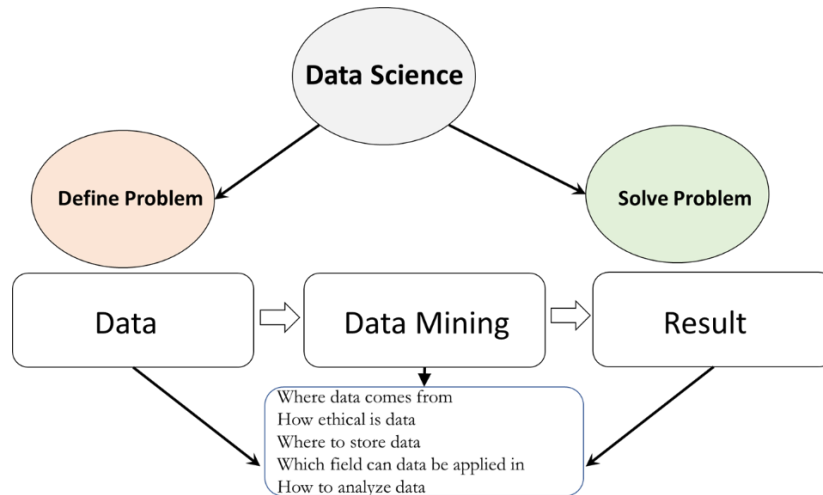


Figure 1 Data Science Workflow

Therefore, the course focuses on “data” and “data analysis” two primary topics which include five parts to respond the questions asked above.

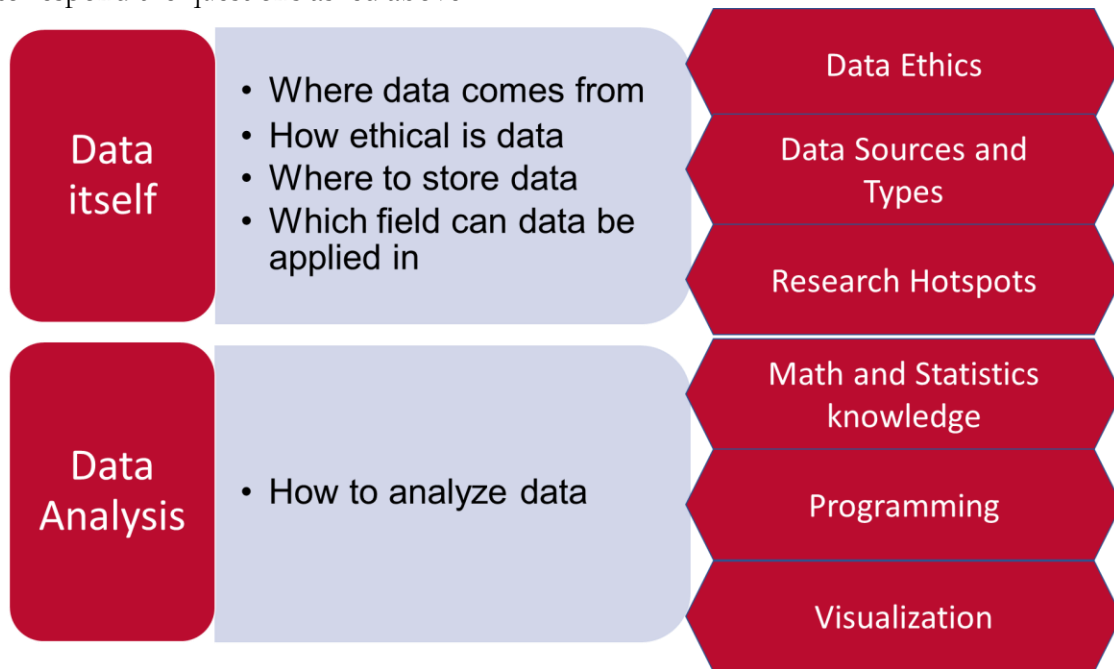


Figure 2 Contents structured in the course.

The modules for each part are listed below for 16-week sessions.

**Course Modules (for 16-week session):**

**Part 1: Introduce Data Ethics in Data Science.**

- Module 1: Informed Consent and Data Ownership
- Module 2: Privacy, Transaction Transparency and Anonymity
- Module 3: Data Validity and Algorithmic Fairness
- Module 4: Societal Consequences and Code of Ethics

**Part 2: Data Sources and Types**

Module 5: Data Sources and Data Acquisition

Module 6: Data Types and Structures

**Part 3: Math Preliminaries and Introduction to Data Wrangling and Data Visualization in R**

Module 7: Introduction to Productivity Tools

Module 8: Math Preliminaries and Their Operations in R

Module 9: Introduction to Database

Module 10: Introduction to Data Wrangling

Module 11: Introduction to Data Visualization

**Part 4: Introduction to Probability and Simulation in R**

Module 12: Probability, Events and Their Types

Module 13: Conditional Probability and Probability Theorem

Module 14: Monte Carlo Simulation

**Part 5: Research Topics in Data Science**

Module 15: Research Topics in Data Science

Module 16: Final exam

Course Assignments

**Grading Summary**

Assignment Type	Percent of Final Grade
Auto-graded Quizzes	50%
Graded Peer Review Discussions	10%
Programming Assignments	20%
Final Exam	20%

**Grading Scale**

A standard grading scale will be used:

<b>A</b> 93% - 100%	<b>B+</b> 88% - 89.99%	<b>C+</b> 78% - 79.99%	<b>D+</b> 68% - 69.99%	<b>F</b> 0% - 59.99%
<b>A-</b> 90% - 92.99%	<b>B</b> 82% - 87.99%	<b>C</b> 72% - 77.99%	<b>D</b> 62% - 67.99%	
	<b>B-</b> 80% - 81.99%	<b>C-</b> 70% - 71.99%	<b>D-</b> 60% - 61.99%	

Table 1. Grade distribution and calculation of final grade, with examples of grades and final calculation.

Assignment Type	Percent of Final Grade [Weight of grade (W)]	Assessment Grade (G) [a score value 0% - 100 you get]	Contribution to final grade W * G
Auto-graded Quizzes	50%	85%	43%
Graded Peer Review Discussions	10%	75%	8%
Programming Assignments	20%	75%	15%
Final Exam	20%	80%	16%
		Average=79%	Sum=81%
Final Grade (add the contributions)		C+	B- (Your final grade)

## Course Policies

### Attendance/Lateness/Late work Policy:

- Your preparedness, participation, and completion of weekly activities on time will significantly impact your course grade, and hence regularity and timeliness are desirable in order to complete this course successfully. It is recommended that you log into your course 3 to 4 times a week and check if you complete all assignments, quizzes, tests, and any other activities by the stated due date.
- There is no penalty for late submissions; however, you are expected to complete the assignments by the due date to avoid becoming overwhelmed in the later weeks. Please try to finish the peer-reviewed assignments as close to the suggested due dates as possible to ensure there are enough peer reviewers for your work.

### Assignment Submission Policy:

- Students are responsible for all material submitted correctly for assignments. Please make sure you submit the correct version of your documents.

### Feedback Policy:

- Questions of a personal nature, such as questions about your grade or anything personal intended for the instructor, please use your BSU email. The instructor will typically respond to messages within 24 hours of receiving them during weekdays (Monday-Friday) and within 48 hours (about 2 days) on the weekend.
- Emails must include the **COURSE NUMBER AND SECTION NUMBER** in the subject line in order to be answered.
- When requesting help to fix a code error, please send your original code and data (if necessary) to the support team at ds.support@bsu.edu. Please **state the error clearly**.

### Grading Policy:

- It is my policy that appropriate evaluation of your academic performance is an integral part of your learning experience. In the absence of mistakes, fraud, bad faith or incompetence, I

will be the key decision-maker on the assignment of grades. Please make your documents and/or graph readable and organized clearly.

## Course Withdrawal Statement

You can withdraw from this course by the withdrawal period. For details, see <https://www.bsu.edu/about/administrativeoffices/registrar/registration-activities/withdraw-from-classes> as well as Withdrawal Procedures in the current graduate catalog.

## Academic Ethics and Plagiarism Policy

Honesty, trust, and personal responsibility are fundamental attributes of the university community. Academic dishonesty and other forms of academic misconduct threaten the foundation of an institution dedicated to the pursuit of knowledge and will not be tolerated. To maintain its credibility and reputation, and to equitably assign evaluations of scholastic and creative performance, Ball State University is committed to maintaining a climate that upholds and values the highest standards of academic integrity. Visit the VPAA's [academic integrity website](#) (Students tab) for resources on understanding academic integrity, citing sources properly, and avoiding inadvertent academic dishonesty. To learn about BSU's academic integrity expectations and students' rights, please read the University [Student Academic Ethics Policy](#).

Academic integrity violations include giving or receiving an unfair academic advantage (cheating), presenting someone else's ideas or work as your own (plagiarism), and falsifying academic records. Unless otherwise indicated, you must work independently by yourself. Check with me if you are unsure whether something constitutes academic dishonesty. Examples of academic integrity violations include but are not limited to:

- Using resources not authorized by the faculty member (including devices, AI tools, hidden notes, and open books)
- Using commercial study websites to find answers to graded assignments (Chegg, Course Hero, StudyPool, OneClass, etc.).
  
- Without the instructor's written permission, giving away, buying, or selling graded assignments class notes, exams, study guides, or other course materials to other students or to third-party vendors (Course Hero, Chegg, 24HourAnswers, etc.).
- Working with another person on any assignment other than authorized group projects.
- Sharing or allowing others to access your files, whether done with permission or not.
- Reusing your own work from another semester, course, or section.
- Sharing answers with others during exams (passing notes, texting, whispering, gesturing)
- Discussing exam questions and answers with students who have not taken the exam.
- Soliciting others to complete work for you.

## Generative AI Statement

Some Use of Generative AI Permitted Under Some Circumstances or With Explicit Permission

During this class, we may use ChatGPT or other generative AI applications. You will be informed as to when, where, and how these tools are permitted to be used, along with guidance for attribution. Any use outside of this permission constitutes a violation of Ball State University's [Student Academic Ethics Policy](#).

## Statements of Inclusion

### **Accommodation:**

If you need course adaptations or accommodation because of a disability, please contact me as soon as possible. Ball State's Disability Services office coordinates services for students with disabilities. documentation of a disability needs to be on file in that office before any accommodation can be provided. Disability Services can be contacted at 765-285-5293 or [dsd@bsu.edu](mailto:dsd@bsu.edu).

If you are experiencing mental health concerns, [telehealth services are available](#).

### **FERPA and Privacy:**

As a student, your educational records are considered confidential. Under FERPA (Family Educational Rights and Privacy Act), your records are confidential and protected. Under most circumstances, your records will not be released without your written and signed consent. However, some directory information may be released to third parties without your prior consent unless a written request to restrict this is on file. You can learn more about student rights to privacy by reading [Ball State's FERPA and Privacy and Protection](#).

## Subject To Change Statement

This syllabus and schedule are subject to change in the event of extenuating circumstances.