



SYLLABUS   
College of Computing and Software Engineering

School of Data Science and Analytics

DATA 5020: Essential Softwares for Data Science and Analytics   
Fall 2024

# Course Information

****

Class meeting time: *Asynchronous Online*

Modality and Location: *Online course*  
*Syllabus is posted in D2L*

# Instructor Information

****

Name: Austin Brown

Email: abrow708@kennesaw.edu  
Office Location: Atrium Building 347

Office phone: 470-578-7827

Office Hours: Tuesdays & Thursdays 12:30 – 1:45 PM or by appointment!  
Preferred method of communication: KSU email – abrow708@kennesaw.edu

# Course Description

****

This course introduces students to modern, in-demand programming languages and environments used by data scientists, statisticians, research methodologists, and analysts in a variety of industries. Students will learn core, foundational programming knowledge, such as data importation, dataset/variable transformation, summary analyses, and exportation. The course will emphasize practical applications through a series of hands-on exercises and real-world examples.

Prerequisites: Admission to MSDSA Program

# Course Materials

****

Recommended Texts:

* *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data 2nd Edition* – Hadley Wickham, Mine Çetinkaya-Rundel, & Garrett Grolemund; ISBN: 978-1492097402 (<https://r4ds.hadley.nz/>)
* *Python for Data Analysis 3rd Edition* – Wes McKinney; ISBN: 978-1098104030 (<https://wesmckinney.com/book/>)

Technology requirements: Consistent and reliable access to a computer, the internet, the course D2L website, RStudio, SAS Studio, and choice of Python IDE.

# Learning Outcomes

****

Upon completion of this course, students should:

1. Import data of various types into various integrated development environments (IDEs) using multiple and appropriate programming languages.
2. Extract and understand basic metadata from columns using multiple programming languages.
3. Perform basic data transformations and data queries using multiple programming languages.
4. Perform basic data analysis using multiple programming languages.

# Course Requirements and Assignments

****

***Assignments and Point Totals:***

|  |  |
| --- | --- |
| **Assignment** | **Points** |
| Lab 1: Data Importation | 100 |
| Lab 2: Data Structure and Column Contents | 100 |
| Lab 3: Data Transformations | 100 |
| Lab 4: Data Analysis | 100 |
| Total Points | 400 |

***Grading Policy:***

|  |  |
| --- | --- |
| A | ≥ 360 points |
| B | 320 - 359 points |
| C | 280 - 319 points |
| D | 240 - 279 points |
| F | < 240 points |

**Labs:** During the semester, each lab assignment will involve a real-world situation in which you will be asked to take a dataset and perform some set of data science tasks using multiple programming languages. For example, you may be asked to import a CSV and determine how many columns contains categorical data and how many contain quantitative data using multiple programming languages.

# Course Policies

****

***Attendance and Participation:***

* While this course is asynchronous, students are expected to engage with course content as well as the instructor on a regular basis to ensure success in the course. Students are also encouraged to regularly communicate with the instructor to address any questions or concerns.

***Feedback Expectations & Response Time:***

* All assignments will be evaluated within a week’s time, unless otherwise noted.
* Emails will generally be responded to within 24 hours during the week. Emails sent over the weekend will be responded to on Monday.

***Instructor Expectations:***

* Each student is responsible for maintaining the pace of the course and turning in assignments on time. Students should utilize the course schedule, calendar, and task list to keep up with the course lectures and assignments.
* All work turned in is expected to be complete and professional. All assignments should be well organized and easy to follow. All written work should observe proper grammar, spelling, and sentence structure. If applicable, all visualizations should be well-formatted with descriptive titles and labels. Points will be deducted for sloppy unprofessional work.
* For content related questions, students should use the “Discussion” tab in D2L. Students must include the subject in the forum title, or the post may be deleted. In addition to the instructor, students are encouraged to respond to other student’s questions, if they can be of assistance. Please check the discussion titles, to make sure your question has not already been answered before starting a new thread. Please be as descriptive as possible with your questions and include pictures, references, screenshots, etc., when applicable.
* For personal questions, students should reach out to the professor directly at [abrow708@kennesaw.edu](mailto:abrow708@kennesaw.edu) and correspond through email or set up a meeting.
* The instructor will use email AND D2L to communicate important information to both individual students and the class as a whole. Therefore, the instructor expects students to check their D2L and email accounts on a **daily** basis.
* Students are highly encouraged to retain copies of all graded materials that are returned for the duration of the semester. If there is a discrepancy between the instructor's records and those of the student, the student will have to provide the appropriate documentation. The burden of proof is on the student.

Please note: all work submitted to the instructor must be an authentic product of each individual student. If evidence arises that a student has submitted someone else's work as their own, cheated on an assignment, or has committed some other academic violation of the University Code of Conduct, the student will receive an F in the class and be referred to the University Judiciary Program for formal charges.

# Department or College Policies

****

***Academic Integrity Concerning GroupMe, Group Chats, Group Texts, and Other Technology-Related Group Conversations***

Kennesaw State University encourages students to use technology to help them learn. However, it is important for students to understand the difference between appropriate collaboration and inappropriate uses of technology for plagiarism and cheating. Students who participate in group texts or other group conversations through mobile apps such as GroupMe or WhatsApp are subject to consequences if any member of that group is found to plagiarize material or facilitate cheating. By virtue of membership in the conversation or participation in the group, any student who is part of a group conversation where cheating or plagiarism occurs may receive the same penalty as students who actively cheat within the group.  Additionally, any students who are found to purchase, sell, or otherwise distribute or collect existing course material are also subject to academic dishonesty hearings. This includes use of Quizlet, Hero, and student organization test files.

# Institutional Policies

****

[Federal, BOR, & KSU Required Syllabus Policies](https://cia.kennesaw.edu/instructional-resources/syllabus-policy.php)

# KSU Student Resources

****

This link contains information on help and resources available to students: [KSU Student Syllabus Resources](https://cia.kennesaw.edu/instructional-resources/syllabus-resources.php)

# Course Schedule

****

|  |  |
| --- | --- |
| **Weeks** | **Concepts Covered** |
| 1 | **Course Introduction and Importing Data of Various File Types** |
| 2 | **Understanding Data Structure and Column Contents** |
| 3 | **Data Transformations and Queries** |
| 4 | **Basic Data Analysis** |