

Course Syllabus

Course Prefix, Number, and Title:

CIS 372-D01/DT1: Programming for Analytics

Credits:

3 credits

University Name:

Dakota State University

Academic Term/Year:

Fall 2019

Last date to Drop and receive 100% refund:

Sep 5

Last date to Withdraw and earn a grade of 'W':

Nov 8

Course Meeting Time and Location:

Tuesday/Thursday 11:00 am – 12:15 pm in East Hall RM 206 (onsite/online students)

Instructor Information:

Name:

James BOIT

Office:

East Hall, Room 317

Phone Number(s):

Office: 605-256-5074

Email Address:

james.boit@trojans.dsu.edu

Virtual Office Hours:

Communicate via chat (e.g. skype) by appointment.

Wednesdays 10:00-11:30 am.

Approved Course Description:

Catalog Description:

An introduction to programming for data analysis with an emphasis on the analysis of large datasets. A programming language common to the analytics industry will be utilized.

Prerequisite(s): CIS 123 or CIS 130 or CSC 150

Additional Course Information:

- The programming language we will use is R programming. R is an integrated suite of software facilities for data analysis, data manipulation, calculation and graphical display, and provides more advanced data visualization. It is powerful, easy to learn and fast to code.

Prerequisites:

Course Prerequisite(s):

CIS 123 Problem Solving and Programming or CIS 130 Visual Basic Programming, or CSC 150 Computer Science I

Technology Skills:

R, RStudio, D2L, Email

Course Materials:

Required Textbook(s):

Most of the content is similar to R Cookbook and R for Data Science, which are available in pdf online, though there are many optional texts available for students to refer to.

- “R Cookbook”, Paul Teetor
- “R for Data Science Import, Tidy, Transform, Visualize, and Model Data”, Hadley Wickham and Garret Golemund

Optional Materials:

- “An Introduction to R”, W.N.Venables, D.M.Smith and the R Core Team; Online copy accessible at: <http://cran.r-project.org/doc/manuals/R-intro.pdf>
- “R for Beginners”, Emmanuel Paradis; Online copy accessible at: http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf
- “Using R for Data Analysis and Graphics – Introduction, Code and Commentary”, J.H.Maindonald; Online copy accessible at: <http://cran.r-project.org/doc/contrib/usingR.pdf>
- There are many quality free on-line tutorials and resources for R such as DataCamp that students may find useful at: <https://www.datacamp.com/courses/free-introduction-to-r>

Computer/ laptop:

I recommend you bring your laptop; we will have in class exercises.

Student Support:

DSU Knowledge Base:

The DSU Knowledge Base contains links and resources to help students by providing information about the following topics: User Accounts & Passwords, Academic Tools & Resources, Software & Apps Support, WiFi & Network Access, Campus Emergency Alert System, Campus Printing, IT Security & Safe Computing, and the Support Desk (which is there to help both on and off-campus students). The Knowledge Base can be accessed through the link below:

· [DSU Knowledge Base](#)

D2L Support for Students:

The D2L Support for Students site is designed to provide DSU students a D2L support resource center that contains user guides, tutorials, and tips for using the D2L learning environment. The D2L Support for Students site can be accessed through the link below:

· [DSU D2L Support Resources for Students](#)

Course Delivery and Instructional Methods:

Instructional methods: Lecture notes, slide shows, electronic mails and programming assignments.

Class Preparation: All class materials will be posted on D2L site. Lecture notes, announcements, asynchronous discussions, etc. can be found/posted on the course D2L site <http://d2l.sdbor.edu> . The course website is located in D2L (see URL above). All students are expected to have read the chapters assigned for each week prior to the class. Students are also expected to check their Email at least once a day.

Computer software: We will use R and RStudio editor. You can download R through the following link: www.r-project.org and install RStudio via <https://www.rstudio.com/products/rstudio/> .

Class Videos: Videos of each class will be posted at: <https://videos.dsu.edu/Playlist/CIS372-FA2019>

Classroom Policies:

Attendance and Make-up Policy:

ADA Statement:

If you have a documented disability and/or anticipate needing accommodations (e.g., non-standard note taking, extended time on exams or a quiet space for taking exams) in this course, please contact the instructor. Also, please contact Dakota State University's Disabilities Office by calling 605-256-5121 or emailing Success.Center@dsu.edu as soon as possible. The DSU website contains additional information and the form to request accommodations found at <https://portal.sdbor.edu/dsu-student/student-resources/disability-services/Pages/default.aspx/>. (Students must log into the DSU portal to access this page.) You will need to provide documentation of your disability. The Disabilities Office must confirm the need for accommodations before officially authorizing them.

Academic Honesty Statement:

Cheating and other forms of academic dishonesty run contrary to the purpose of higher education and will not be tolerated in this course. Please be advised that, when the instructor suspects plagiarism, the Internet and other standard means of plagiarism detection will be used to resolve the instructor's concerns. The South Dakota Board of Regents Student Academic Misconduct Policy can be found here: [SDBOR Policy 2.33](#).

All forms of academic dishonesty will result in an F for that assignment or quiz (minimum) to an F for the course grade and possible notification of the Academic Integrity Committee.

Communication and Feedback:

Requirements for Course interaction:

Preferred contact method is through email. Email messages should be sent to me at james.boit@trojans.dsu.edu rather than through D2L email. When emailing me, please place “CIS 372 – your name” at the beginning of the subject line. This will aid me in responding to your email quicker.

In addition, I can communicate with you via chat (e.g. skype) by appointment.

Email Response Time:

I generally respond to email messages within 24 hours. Response time may be a little longer on weekends and holidays.

Feedback on Assignments:

Feedback for assignments is usually provided within 2 weeks of the assignment due date. I'll notify the class if delay is expected.

Student Learning Outcomes:

Upon completing the course, the student will be able to:

- Get general information about R programming
- Learn R basic functions
- Perform data manipulations in R
- Use R to perform data preprocessing and explanatory data analysis.
- Use R to perform regression analysis and advanced statistical analysis

Evaluation Procedures:

Assessments:

The course grade will be calculated as follows and it is subject to change.

Work item	Weight
Assignments	80%
Project	20%

Assignments: All assignments are individual assignments. Your assignments must be typed. Please keep in mind that hand-written reports will NOT be accepted. It is expected that you complete assignments on time. I will not accept assignments that are submitted later than one week after the due date. Assignments submitted after the due day will lose 50%. We have four assignments, which account for 80% of your final grade (20 points for each assignment).

Project: You will work on one project, which accounts for 20% of your final grade. The project is an individual one. The project grade will be based on the quality of each component of your work. Evaluation of the reports is based on the following criteria: technical soundness, organization, and clarity.

Final Examination:

There is no final exam.

Performance Standards and Grading Policy:

The final letter grade will be based on the following scale:

90% - 100%.....	"A"
80 - 89.99%.....	"B"
70 - 79.99%.....	"C"
60 - 69.99%.....	"D"
Less than 60%	"F"

Student Verification Statement and Proctoring Policy:

Federal law requires that universities verify the identity of students when course materials and/or course assessment activities are conducted either partially or entirely online. A student's Desire2Learn (D2L) login and password are intended to provide the student with secure access to course materials and are also intended to help the university meet this federal mandate. Some DSU Faculty also require the use of a proctor for exams in distance-delivered (Internet) courses and this requirement provides a second level of student identity verification. Students are responsible for any proctoring fees, if applicable. Finally, an instructor who uses web conferencing technology may require students to use a webcam during exams, as another means of student identity verification through voice and visual recognition.

Tentative Course Outline and Schedule:

Week	Date	Topics	Details/Reading documents	Assignment/Project, Dues
1	Aug 26-29	Introduction, First Session in R	Course overview, objectives, Intro to Analytics Prog., installation, some basics R cookbook: CH 1-2	
2	Sep 3-5	Data Structures, Navigating the software	Vectors, matrix, list, data frames, arrays, packages, built-in data sets R cookbook: CH 3-5	
3	Sep 10-12	Data Structures, Navigating the software (con't)	Vectors, matrix, list, data frames, arrays, packages, built-in data sets R cookbook: CH 3-5	Assignment 1 out
4	Sep 17-19	Basic statistics, Functions, Control statements	Computing basic statistics, defining functions, if, while, for loops R Data Sci: CH 15	
5	Sep 24-26	Getting data in and out of R	Input/Output, Read from/write to files R cookbook: CH 4	Assignment 1 due

Week	Date	Topics	Details/Reading documents	Assignment/Project, Dues
6	Oct 1-3	Data transformation, Strings and dates	Cleaning data, tidy data with 'tidyR', 'lubridate' package for dates R Data Sci: CH 9 R cookbook: CH 6-7	
7	Oct 8- 10	Data transformation (con't)	Tidy data with 'dplyr' R Data Sci: CH 9	Assignment 2 out
8	Oct 15-17	Data visualization	Visualization basics, plotting R cookbook: CH 10	
9	Oct 22-24	Data visualization (con't)	Visualization with ggplot2 R Data Sci: CH 1	Assignment 2 due
10	Oct 29-31	Data visualization (con't)	Visualization with ggplot2 R Data Sci: CH 1	Assignment 3 out
11	Nov 5-7	Linear regression	Performing simple and multiple linear regression, regression statistics	Project out
12	Nov 12- 14	Linear regression (con't)	Performing simple and multiple linear regression, regression statistics R cookbook: CH 11	Assignment 3 due
13	Nov 19-21	Explore advanced statistical techniques	Cluster analysis, Principal component analysis, Factor analysis R cookbook: CH 13	Assignment 4 out
14	Nov 26	Explore advanced statistical techniques (con't)	Cluster analysis, Principal component analysis, Factor analysis R cookbook: CH 13	
15	Dec 3-5	Case study - Review		Assignment 4 due
16	Dec 10-12	Project Submission		Project due

Nov 27-Dec1 (Thanksgiving holiday)

Dec 12-18 (Exam period)