CS M10R: INTRODUCTION TO R PROGRAMMING

Originator
enikjeh

College
Moorpark College

Attach Support Documentation (as needed)
Why use R.docx
2019-05-06_CS_advisory_Minutes.docx

Discipline (CB01A)
CS - Computer Science

Course Number (CB01B)
M10R

Course Title (CB02)
Introduction to R Programming

Banner/Short Title
R Programming

Credit Type
Credit

Honors
No

Start Term
Fall 2021

Catalog Course Description
Introduces computer programming and algorithm design using the R programming language. Covers an introduction to R, from installation to most of the statistical concepts, and machine learning. Includes the fundamentals of computer programming concepts: basic data types, variables, if-else, loops, functions, vectors, objects, matrices, arrays, data frames, lists, factors, basic input, data visualization, and output with files. Explains some principles of algorithm design and analysis as well as techniques for testing programs.

Taxonomy of Programs (TOP) Code (CB03)
0707.10 - *Computer Programming

Course Credit Status (CB04)
D (Credit - Degree Applicable)

Course Transfer Status (CB05) (select one only)
A (Transferable to both UC and CSU)

Course Basic Skills Status (CB08)
N - The Course is Not a Basic Skills Course

SAM Priority Code (CB09)
D - Possibly Occupational

Course Cooperative Work Experience Education Status (CB10)
N - Is Not Part of a Cooperative Work Experience Education Program

Course Classification Status (CB11)
Y - Credit Course

Educational Assistance Class Instruction (Approved Special Class) (CB13)
N - The Course is Not an Approved Special Class

Course Prior to Transfer Level (CB21)
Y - Not Applicable

Course Noncredit Category (CB22)
Y - Credit Course

Funding Agency Category (CB23)
Y - Not Applicable (Funding Not Used)

Course Program Status (CB24)
1 - Program Applicable

General Education Status (CB25)
Y - Not Applicable

Support Course Status (CB26)
N - Course is not a support course

Field trips
Will not be required

Grading method
(L) Letter Graded

Alternate grading methods
(O) Student Option- Letter/Pass
(P) Pass/No Pass Grading

Does this course require an instructional materials fee?
No

Repeatable for Credit
No

Is this course part of a family?
No

Units and Hours

Carnegie Unit Override
No

In-Class

Lecture
Minimum Contact/In-Class Lecture Hours
52.5
Maximum Contact/In-Class Lecture Hours
52.5

Activity

Laboratory

Total in-Class

Total in-Class
Total Minimum Contact/In-Class Hours
52.5
Total Maximum Contact/In-Class Hours
52.5

Outside-of-Class

Internship/Cooperative Work Experience

Paid

Unpaid

Total Outside-of-Class

Total Outside-of-Class
Minimum Outside-of-Class Hours
105
Maximum Outside-of-Class Hours
105

Total Student Learning

Total Student Learning
Total Minimum Student Learning Hours
157.5
Total Maximum Student Learning Hours
157.5

Minimum Units (CB07)
3
Maximum Units (CB06)
3

Advisories on Recommended Preparation
MATH M15 or MATH M15H and CSM10P or CSM125 or CSM10J

Requisite Justification

Requisite Type
Recommended Preparation

Requisite
MATH M15 or MATH M15H and CSM10P or CSM125 or CSM10J

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Content review

Student Learning Outcomes (CSLOs)

	Upon satisfactory completion of the course, students will be able to:
1	implement an R programming environment to develop programs for data manipulation, management, and analysis.
2	develop statistical computing solutions including statistical graphics using the R programming language.

Course Objectives

	Upon satisfactory completion of the course, students will be able to:
1	establish an efficient scientific computing environment.
2	identify and use available R packages and associated Open Source software to meet given scientific objectives.
3	design and write efficient programs using R (and similar high-level languages) to perform routine and specialized data manipulation/management and analysis tasks.
4	document, share, and collaborate on code development using a suite of Open Source standards and tools.
5	document analytical workflow using R, markdown languages, and version control.
6	compose an R program using Vectors, Matrices, or Arrays.

Course Content

Lecture/Course Content

(5%) History, overview of R, Install, and configuration of R programming environment

- History of R
- The R environment
- Installing R
- Installing R Studio
- Using R interactively
- R commands, case sensitivity, etc.
- R Script File

(5%)Basic language elements and data structures

- Variable names and assignment
- Operators
- Integers
- R Data Types
- Arithmetic

(15%) Matrices, Arrays, Vectors, and Lists

- Index matrices
- Matrix multiplication
- Arrays
- Array indexing
- The array() function
- The outer product of two arrays
- Generalized transpose of an array
- Vectors and assignment
- Vector arithmetic
- Logical vectors
- Index vectors
- Lists
- Subsetting vectors
- Subsetting matrices

(5%) Data input/output, Data storage formats, and Data Frames

- Reading data from files
- The read.table() function
- The scan() function
- Loading data from other R packages
- Data Frames

(5%) Control structures (loops, if-else)
- if-else statements
- for loop, repeat loop, and while loop

(5%) Functions and Object-Oriented Programming
- Built-in Functions
- User Defined Functions
- String Functions
- Objects
- Classes

(5%) Graphics and visualization
- Bargraph
- Histograme
- Pie Chart
- Scatter Diagram
- Scales
- BoxPlots

(10%) Machine Learning and Grammar of data manipulation (dplyr and related tools)
- *K*-Nearest Neighbors
- Decision Tree
- Background
- Selecting Variables
- Filtering Rows
- Comparisons
- Adding New Variables
- Database Interface

(5%) Debugging/profiling
- Something is wrong
- Debugging Tools in R
- Using Traceback
- Using Recover
- Using debug

(5%) Factors, Tables, and Math
- Factors
- Generating Factor levels
- Basic tables: table() and xtable()
- Math Built-in functions

(35%) Statistical concepts
- Descriptive Statistics
- Sampling methods
- Central Tendency values
- Discrete probability distributions
  - The binomial distribution
  - The Poisson distribution
- Continuous probability distributions
    - The uniform distribution
    - Normal distribution
    - Student's t distribution
- The central limit theorem
- Confidence Interval
- The chi-squared distribution
- Single samples Hypothesis tests
  - Hypothesis test for population proportion
  - Hypothesis test for population mean
- Two Samples Hypothesis testing
- ANOVA Hypothesis testing
- Correlation Analysis

- Linear Regression
- Logistic Regression
- Time Series Analysis

Laboratory or Activity Content

Not applicable.

Methods of Evaluation

Which of these methods will students use to demonstrate proficiency in the subject matter of this course? (Check all that apply):

Problem solving exercises
Skills demonstrations
Written expression

Methods of Evaluation may include, but are not limited to, the following typical classroom assessment techniques/required assignments (check as many as are deemed appropriate):

Computational homework
Essay exams
Individual projects
Objective exams
Problem-solving exams
Quizzes
Research papers
Skills demonstrations

Instructional Methodology

Specify the methods of instruction that may be employed in this course

Computer-aided presentations
Class activities
Class discussions
Distance Education
Group discussions
Guest speakers
Instructor-guided interpretation and analysis
Internet research
Laboratory activities
Lecture

Describe specific examples of the methods the instructor will use:

Instructor will use blackboard to lecture the concepts and computer to show the programming examples like decision and loop structures to apply to real world data. Instructor will show some helpful video and power point tutorials with examples about writing R statements. Instructor will promote student discussion, group activities, use of Internet, Lab, and homework. Instructor will assess student understanding at frequent intervals throughout the learning process.

Representative Course Assignments

Writing Assignments

1. Documenting program specifics including variables, input-output, functions, and solutions of the actual code of the program.

2. Writing pseudo-code, a detailed “text-based” description of a R program, for solving a statistical problem. An example would be: Write a pseudocode to predict the price of gold as a function of time.

Critical Thinking Assignments

1. Analyze and interpret the relation between number of hours study and the test result for students at college. Collect your data from the library for different typical courses. We want to examine the relationship between amount of time spent studying for exam (X) in hours and the score that person makes on exam (Y). Use R programming to calculate coefficient of correlation, regression equation, and interpret them.

2. Collect data for a real world problem, perform data analysis, and interpret the results using the R programming language.

Reading Assignments

1. Read professional journals, white papers, and case studies about R programming in addition to the text assignments and present a summary of your reading to the class.

2. Read articles in which R programming was utilized to solve and analyze a statistical problem.Then write the conclusion and your recommendation.

Skills Demonstrations

1. Analyze the strengths and weaknesses of educational software program R.

2. Demonstrate the use of common statistical methods such as linear regression, statistical and data analysis functions in R programming.

Outside Assignments

Representative Outside Assignments

1. Read data science peer-reviewed journal articles in which R programming was utilized to solve or analyze a problem, and then write pseudocode to recreate that particular R programming code.

2. Conduct library and/or Internet research to gather information and prepare for group design projects. An example would be: Design a R program using regression methods to predict the stock market shares.

Articulation

Equivalent Courses at other CCCs

College	Course ID	Course Title	Units
Palomar College	CSIT 165	R Programming	3
College of San Mateo	CIS 140	R Programming for Big Data	4
De Anza College	CIS 64H	R Programming	4.5
Fresno City College	DS 40	Introduction to R Programming for Data Science	

District General Education

A. Natural Sciences

B. Social and Behavioral Sciences

C. Humanities

D. Language and Rationality

E. Health and Physical Education/Kinesiology

F. Ethnic Studies/Gender Studies

Course is CSU transferable

Yes

CSU Baccalaureate List effective term:

F2021

CSU GE-Breadth

Area A: English Language Communication and Critical Thinking

Area B: Scientific Inquiry and Quantitative Reasoning

Area C: Arts and Humanities

Area D: Social Sciences

Area E: Lifelong Learning and Self-Development

Area F: Ethnic Studies

CSU Graduation Requirement in U.S. History, Constitution and American Ideals:

UC TCA

UC TCA
Proposed

Date Proposed:
6/15/2021

IGETC

Area 1: English Communication

Area 2A: Mathematical Concepts & Quantitative Reasoning

Area 3: Arts and Humanities

Area 4: Social and Behavioral Sciences

Area 5: Physical and Biological Sciences

Area 6: Languages Other than English (LOTE)

Textbooks and Lab Manuals

Resource Type
Textbook

Classic Textbook
No

Description
Wickham, Hadley. *Advanced R*. 2nd ed., Chapman and Hall/CRC, Routledge, 2019.

Resource Type
Textbook

Classic Textbook
No

Description
Kabacoff, Robert. *R in Action, Data Analysis and Graphics with R.* 2nd ed., Manning Publications, 2015.

Resource Type
Textbook

Description
Matloff, Norman. *The Art of R Programming: Tour of Statistical Software Designs*. William Pollock/No Starch Press, 2011.

Resource Type
Textbook

Classic Textbook
No

Description
Venables, William, David M. Smith, and the R Development Core Team. *An Introduction to R*. 3.6.3 ed., Network Theory Ltd, 2020.

Resource Type
Software

Description
R Programming. R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS, 3.5.2 ed., Open Source, free online. https://www.r-project.org/

Library Resources

Assignments requiring library resources
Use of the Library's print and online resources to locate current articles in newspapers, magazines and scholarly journals in preparation for a report on aspects of R programming practices such as functions, R data frame, and Correlation in R.

Sufficient Library Resources exist
Yes

Example of Assignments Requiring Library Resources
Write a paper that compares and contrasts two popular programming languages that have been used for data science (R and Python).

Distance Education Addendum

Definitions

Distance Education Modalities
Hybrid (51%–99% online)
Hybrid (1%–50% online)
100% online

Faculty Certifications

Faculty assigned to teach Hybrid or Fully Online sections of this course will receive training in how to satisfy the Federal and state regulations governing regular effective/substantive contact for distance education. The training will include common elements in the district-supported learning management system (LMS), online teaching methods, regular effective/substantive contact, and best practices.
Yes

Faculty assigned to teach Hybrid or Fully Online sections of this course will meet with the EAC Alternate Media Specialist to ensure that the course content meets the required Federal and state accessibility standards for access by students with disabilities. Common areas for discussion include accessibility of PDF files, images, captioning of videos, Power Point presentations, math and scientific notation, and ensuring the use of style mark-up in Word documents.
Yes

Regular Effective/Substantive Contact

Hybrid (1%–50% online) Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
E-mail	Instructor will email students with announcements about the course or an upcoming event. Students in turn may email the instructor with their questions or concerns.
Other DE (e.g., recorded lectures)	Instructor may recorded the lectures and post them for students to view within a specified time frame to be ready for the accompanying assignment. Students will upload their assignment to the course webpage.

Hybrid (51%–99% online) Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
Synchronous Dialog (e.g., online chat)	Instructor may be available on certain day(s) with certain time frames to help students and answer questions via online chat.
E-mail	Instructor will email students with announcements about the course or an upcoming event. Students in turn may email the instructor with their questions or concerns.
Other DE (e.g., recorded lectures)	Instructor may recorded the lectures and post them for students to view within a specified time frame to be ready for the accompanying assignment. Students will upload their assignment to the course webpage.
Telephone	Instructor may provide his/her phone number to students where they can leave a voicemail and expect a call back with in 24 hour.

100% online Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
Asynchronous Dialog (e.g., discussion board)	Instructor will post a question, students will respond to the question.
E-mail	•Instructor will email students with announcements about the course or an upcoming event. Students in turn may email the instructor with their questions or concerns. •Students will email their projects/assignments to the instructor.
Other DE (e.g., recorded lectures)	Instructor may record the lectures and post them for students to view within a specified time frame.
Synchronous Dialog (e.g., online chat)	Instructor may be available on certain day(s) with certain time frames to help students and answer questions via online chat.
Telephone	Instructor may provide his/her phone number to students where they can leave a voicemail and expect a call back with in 24 hour.

Examinations

Hybrid (1%–50% online) Modality

Online
On campus

Hybrid (51%–99% online) Modality

Online
On campus

Primary Minimum Qualification

COMPUTER SCIENCE

Additional local certifications required

Recent experience in teaching Statistics. Statistics is about 35% of this course and this knowledge is essential to the effective teaching of this course.

Review and Approval Dates

Department Chair
03/19/2021

Dean
03/19/2021

Technical Review
03/25/2021

Curriculum Committee
4/6/2021

DTRW-I
04/08/2021

Curriculum Committee
MM/DD/YYYY

Board
05/11/2021

CCCCO
MM/DD/YYYY

DOE/accreditation approval date
MM/DD/YYYY