DEPARTMENT: Computer Science and Information Systems

COURSE PREFIX: CIS COURSE NUMBER:542/642CREDIT HOURS: 3

1. TITLE: Introduction to R Programming
2. COURSE DESCRIPTION AND PREREQUISITE(S):

R is increasingly becoming a *de facto* language for data analysis in many industries and academic disciplines. The main purpose of this course is to equip students with a good foundation in R, namely: R language syntax; RStudio Environment; various data science tools used in conjunction with R; and basic data manipulation, analysis, and visualization techniques implemented in R. With this foundation, students will be able to conduct basic data science projects or learn more advanced data analysis tools or packages available in R. As a part of the course requirements, students will have to complete and present their own research project in R. This course requires a basic understanding of statistics. Thus, students should have taken an introductory statistics course prior to enrolling in this course. Also, although the course covers the relevant programming concepts in R in-depth, a basic understanding or computer programming (e.g. variables, loops, controls, etc.) would be helpful

**Prerequisite(s):** CIS243 or STA135; instructor permission if you took a statistics course other than the two listed.

1. COURSE OBJECTIVES:

The student will be able to

1. Use RStudio for entering and debugging R code and working with R packages
2. Work with various data types and advanced data structures in R
3. Use the basic elements of R syntax
4. Read, organize, and visualize data in R
5. Perform basic statistical analysis involving descriptive statistics, t-tests, ANOVA, and Regression, etc.
6. Conduct Data Science projects using R and related tools
7. CONTENT OUTLINE:

The course will have the following (tentative) outline:

|  |  |  |
| --- | --- | --- |
| **Class** | **Topic** | **Readings** |
| 1 | Introduction to the Class  Installing R and RStudio | Course Syllabus  Chapter 1 |
| 2 | Introduction to R Environment | Chapter 2 |
| 3 | R Packages | Chapter 3 |
| 4 | Basics of R Syntax | Chapter 4 |
| 5 | Advanced Data Structures in R | Chapter 5 |
| 6 | Reading Data into R | Chapter 6 |
| 7 | Reading Data into R from the Web | Assigned reading |
| 8 | Visualizing Data in R: Base Graphics | Chapter 7 |
| 9 | Visualizing Data in R: ggplot2 | Chapter 7 |
| 10 | Exam 1 Review | Review Chapters 1-7 |
| 11 | Exam 1 | Review Chapters 1-7 |
| 12 | Introduction to R Markdown and ‘knitr” package | Chapter 28 |
| 13 | Using GitHub for Data Science Projects in R | Assigned Reading  Individual Project Description |
| 14 | Writing R Functions | Chapter 8 |
| 15 | Control Statements in R | Chapter 9 |
| 16 | Control Statements in R | Chapter 9 |
| 17 | Loops in R | Chapter 10 |
| 18 | Loops in R | Chapter 10 |
| 19 | Group Manipulations in R | Chapter 11-12 |
| 20 | Data Reshaping in R | Chapter 14 |
| 21 | Manipulating Stings in R | Chapter 16 |
| 22 | Exam 2 Review | Review Chapters 8-12, 14, 16, 28 |
| 23 | Exam 2 | Review Chapters 8-12, 14, 16, 28 |
| 24 | Probability Distributions in R | Chapter 17 |
| 25 | Descriptive Statistics in R | Chapter 18 |
| 26 | Correlation and Covariance in R | Chapter 18 |
| 27 | T-Tests in R | Chapter 18 |
| 28 | ANOVA in R | Chapter 18 |
| 29 | Simple Linear Regression in R | Chapter 19 (Sec 19.1 only) |
| 30 | Individual Project Presentations |  |
| 31 | Individual Project Presentations |  |
| 32 | Group Project Reports are due on the last day of the semester  Peer Evaluations are due before Exam 3 (which is also your final Exam) |  |
| 33 | Exam 3 | Review Chapters 17, 18, and 19 (Sec 19.1 only) |

1. INSTRUCTIONAL ACTIVITIES:

This course will involve the following instructional activities:

1. *Lab Work*. Virtually every class will be held in the lab. The lab lecture will involve the instructor running R code examples that illustrate a particular data analysis technique in R. Students are expected to follow along by entering and running the code demonstrated by the instructor. Some lectures will involve demonstrations of R related tools (e.g. GitHub) or technologies (e.g. R Markdown language). After covering a particular topic, the instructor will ask students to take a short, multiple choice quiz designed to assess their understanding of the topic covered. Students are allowed to take these quizzes only while in the lab. Thus, these quizzes assess both attendance and participation in each class.
2. *Independent Work*. Students are expected to read each assigned chapter to paper prior to coming to class. This will allow for a more efficient use of class time. Instead of learning the basics in class, students will have an opportunity to practice the material and clarify the concepts they do not understand. Also, students are expected to work on their individual projects outside of the class. Class time can be used to get instructor help or feedback on the individual project.
3. *Online Discussions*. We will be using the discussion forums on Canvas for all course related discussions outside of the class. This feature is especially important for those students who take this course online. Please limit the use of email only to person issues.
4. FIELD, CLINICAL, AND/OR LABORATORY EXPERIENCES:

Virtually every class will involve individual lab work.

1. TEXT(S) AND RESOURCES:

Textbook:

* Jared Lander (2017). R for Everyone: Advanced Analytics and Graphics (2nd Edition). Addison-Wesley Data & Analytics Series. ISBN-13: 978-0134546926
  + Either a hard copy or an e-book – whatever works best for you.

Software:

* The latest version of R (can be installed from <https://www.r-project.org/>)
* The latest version of RStudio Desktop (can be installed from <https://www.rstudio.com/products/rstudio/download/>)
  + You don’t have to buy the software: it is “open source”. Also, it will be preinstalled for you in the lab):

1. EVALUATION AND GRADING PROCEDURES:

Course performance will be evaluated through the following activities:

|  |  |  |
| --- | --- | --- |
| **Assessment Activity** | **Weight** | **Description** |
| Quizzes | 20% | At the end of virtually every class there will be a quiz based on the material covered in class. These quizzes can be taken only in class. Thus, this grade component assesses not only how well you master the material for each class, but also your attendance, and, to a lesser degree, participation in each course. Attendance of each and every class is essential for your success in this course. There will be approximately 23 quizzes in this course (one for every lecture). I will drop the three lowest quiz grades at the end of the semester. |
| Exam 1 | 20% | Multiple-choice exam over Chapters 1-7 |
| Exam 2 | 20% | Multiple-choice exam over Chapters 8-12, 14, 16, 28 and assigned readings on sourcing data from the Web, R Markdown, ‘knitr”, and GitHub |
| Exam 3 | 20% | Multiple-choice exam over Chapters 17, 18, and 19 (Sec 19.1 only) |
| Individual Research Project | 20% | You will be required to conduct and present your own research project using R, R Markdown, ‘knitr’, and GitHub. You will have to document every step of your research project: from data input to report generation. All the code files should be posted to GitHub and a URL to your GitHub repository should be submitted via Canvas. Both undergraduate and graduate students are required describe how their data was sources and organized, generate descriptive statistics, produce data visualizations, and provide interpretations of results.  While undergraduate students can use any analytical technique for their research project (including simple descriptives), graduate students are expected to use more advanced analysis techniques, such as t-test, ANOVA, regression, etc. Moreover, graduate research projects should include a separate write-up that follows the format of a typical research paper. The write-up should contain an introduction, literature review, theory with hypothesis, discussion of results, implications, and a conclusion. R code (generated with the help of R Markdown language and ‘knitr’ package) should be included as an appendix. Undergraduate students are only required to submit well-commented R code produced using R Markdown and ‘knitr”.  You grade will be subdivided as follows:   * 10% is devoted to the report itself that you will generate using ‘knitr”. The report will be graded by the instructor using a rubric * 5% is devoted to a short presentation covering your research project that you will make in class. The presentation will be graded by the instructor using a rubric * 5% is devoted to peer evaluations. Your report will be evaluated by several of your peers using a rubric. If you fail to submit your own peer evaluations for others by the date of the final exam or your peer evaluations are of low quality, up to 2% will be subtracted from your 5% grade component. Thus means you will lose a substantial share of this grade component. |

The final grade will be based on the following scale:

1. 90 and above
2. 80 to 89
3. 70 to 79
4. 60 to 69
5. 59 and below

Final grades will not be rounded up

1. ATTENDANCE POLICY:

Students are expected to adhere to the MSU Attendance Policy outlined in the current *MSU Bulletin*.

1. ACADEMIC HONESTY POLICY:

Murray State University takes seriously its moral and educational obligation to maintain high standards of academic honesty and ethical behavior. Instructors are expected to evaluate students’ academic achievements accurately, as well as ascertain that work submitted by students is authentic and the result of their own efforts, and consistent with established academic standards. Students are obligated to respect and abide by the basic standards of personal and professional integrity.

**Violations of Academic Honesty include:**

**Cheating** - Intentionally using or attempting to use unauthorized information such as books, notes, study aids, or other electronic, online, or digital devices in any academic exercise; as well as unauthorized com­munication of information by any means to or from others during any academic exercise.

**Fabrication and Falsification** - Intentional alteration or invention of any information or citation in an academic exercise. Falsification involves changing information whereas fabrication involves inventing or counterfeiting information.

**Multiple Submission** - The submission of substantial portions of the same academic work, including oral reports, for credit more than once without authorization from the instructor.

**Plagiarism** - Intentionally or knowingly representing the words, ideas, creative work, or data of someone else as one’s own in any academic exercise, without due and proper acknowledgement.

Instructors should outline their expectations that may go beyond the scope of this policy at the beginning of each course and identify such expectations and restrictions in the course syllabus. When an instructor receives evidence, either directly or indirectly, of academic dishonesty, he or she should investigate the instance. The faculty member should then take ap­propriate disciplinary action.

Disciplinary action may include, but is not limited to the following:

1) Requiring the student(s) to repeat the exercise or do additional related exercise(s).

2) Lowering the grade or failing the student(s) on the particular exercise(s) involved.

3) Lowering the grade or failing the student(s) in the course.

**If the disciplinary action results in the awarding of a grade of *E* in the course, the student(s) may not drop the course.**

Faculty reserve the right to invalidate any exercise or other evaluative measures if substantial evidence exists that the integ­rity of the exercise has been compromised. Faculty also reserve the right to document in the course syllabi further academic honesty policy elements related to the individual disciplines.

A student may appeal the decision of the faculty member with the department chair in writing within five working days. Note: If, at any point in this process, the student alleges that actions have taken place that may be in violation of the Murray State University Non-Discrimination Statement, this process must be suspended and the matter be directed to the Office of Institutional Diversity, Equity and Access. Any appeal will be forwarded to the appropriate university committee as determined by the Provost.

1. NON-DISCRIMINATION POLICY AND STUDENTS WITH DISABILITIES:

**Policy Statement**

Murray State University endorses the intent of all federal and state laws created to prohibit discrimination. Murray State University does not discriminate on the basis of race, color, national origin, sex, gender identity, sexual orientation, religion, age, veteran status, or disability in employment or application for employment, admissions, or the provision of services and provides, upon request, reasonable accommodation including auxiliary aids and services necessary to afford individuals with disabilities equal access to participate in all programs and activities.

In particular and without limiting the preceding and pursuant to and consistent with the requirements of Title VI of the Civil Rights Act of 1964 and its regulations 34 CFR 100 et seq.; Section 504 of the Rehabilitation Act of 1973 and its regulations 34 CFR 104; Title IX of the Education Amendments of 1972, 20 USC 1681 et seq., and its regulations 34 CFR 106 et seq; and the Age Discrimination Act of 1975 and its regulations 34 CFR 110, Murray State University does not discriminate on the basis of race, color, national origin, sex, handicap, or age in its educational programs and activities. This non-discrimination in education programs and activities extends to employment and admissions and to recruitment, financial aid, academic programs, student services, athletics, and housing. Murray State is required by Title IX and 34 CFR part 106 not to discriminate on the basis of sex and the prohibition against sex discrimination specifically includes a prohibition of sexual harassment and sexual violence.

For information regarding nondiscrimination policies contact the Executive Director of IDEA/Title IX Coordinator, Camisha Duffy, Office of Institutional Diversity, Equity and Access, 103 Wells Hall, Murray, KY 42071. Telephone: 270-809-3155 (Voice) 270-809- 3361 (TDD).

**Students with Disabilities**

Students requiring special assistance due to a disability (temporary or permanent) should visit the Office of Student Disability Services immediately for assistance with accommodations. For more information, students with disabilities should contact the Office of Student Disability Services, Ken Ashlock, 423 Wells Hall, Murray, KY 42071. Telephone: 270-809-2018 (Voice) 270-809- 5889 (TDD).

1. OTHER REQUIRED DEPARTMENTAL OR COLLEGIATE COMMITTEE INFORMATION

**Email Policy**

Please use your official MSU email account to contact the instructor. I check email at least once a day during the work week. I will do my best to answer your email the same day I receive it. Please limit email content to personal issues. All course related questions should be submitted via Canvas.

**Canvas Policy:**

We will be using the discussion forums on Canvas for all course related questions and answers. This policy is intended to make sure all students get the benefits of your questions and my answers. I will check Canvas postings and respond to them at least once a day during the work week.

**Assignment Submission Policy:**

All assignments should be submitted via Canvas (no email or hard copy submissions please). Each assignment must be completed by the end of the day (no later than 11:59pm of the date announcement as the deadline). Late submissions will be penalized by 10% for each day late. No submissions will be accepted if they are more than 3 days late. There will be a bonus assignment given out at the end of the semester for those wishing to earn a bonus or make-up for an assignment they failed to submit.

**Make-Up Exams Policy:**

The lab exams and in-class exams will be administered during the scheduled class time. Missing an exam will result in 0 points for the exam. If you have a valid excuse for missing an exam (e.g. illness, school event related travel or unexpected travel plans, etc.), you must inform the professor before the scheduled exam time. You must provide appropriate supporting document to qualify for a make-up exam. Make-up exams will be scheduled by the instructor towards the end of the semester.

**AACSB’s 25% Rule:**

This course counts toward undergraduate business hours.

Business majors/areas include: Accounting, Business Administration, Computer Information Systems, Finance, Logistics and Supply Chain Management, Management, and Marketing. All other majors/areas are considered non-business.  Non-business majors are not permitted to complete more than 30 hours of business courses from the prefixes ACC, BUS, CIS, FIN, MGT, MKT, RES or the courses LST 240 and LST 540 without completing the Business Core Requirements.

If you plan to complete more than 30 hours, must contact the Center for Undergraduate Business Advising in the College of Business for special permission.