Mel and Enid Zuckerman College of Public Health
University of Arizona

SYLLABUS
Healthcare Data Science (BIOS 511)
Fall 2020

Time:  Monday and Wednesday 9:00am – 10:15am
Location:  Live Online
Instructor:  Jin Zhou Ph.D.

Instructure Office Room: Drachman Hall, Rm A242
Phone: 520-626-1393
Email: jzhou@email.arizona.edu

Office Hours:
By appointment through personal zoom room:
https://arizona.zoom.us/j/5124548764

Catalog Description:

Course Description:  This course introduces you new tools and techniques used in healthcare related data sciences. Topics include

- Basic knowledge of large clinical databases focusing on medical records
- Cohort definition and extraction
- Intro to SQL
- Linux basic
- Collaborative research using Git/GitHub
- High performance computing and cloud computing
- Shiny
- Docker
- TidyVerse
- Basic predictive modeling
- Select tool from data science for example spark (distributed analysis), TensorFlow.

Course Prerequisites:  Statistical computing, data management

Course Learning Objectives:

At the completion of the course, you will be able to:

1) communicate with healthcare stakeholders and understand whether their problem is one that can be solved by healthcare Data Science and determine how it might be best solved;
2) apply various tools and techniques to acquire, clean and store data for analysis;

3) identify various analytical problems and the appropriate modeling techniques (statistical and machine learning);

4) communicate results to healthcare stakeholders;

5) ensure analytical results reproducibility.

Program Competencies Covered:

**Analytical Skills:**

1) Be able to describe the roles biostatistics serves in the discipline of public health

2) Apply basic informatics techniques and vital statistics and public health records in the description of public health characteristics and in public health research and evaluation

3) Communicate understanding of the assumptions necessary for a given statistical and analytical procedure as well as the ability to determine if the assumptions are met for a given data set

4) Suggest preferred methodological alternatives to commonly used statistical methods when assumptions are not met

5) Recognize strengths and weaknesses of proposed approaches, including alternative designs, data sources, and analytical methods

6) Apply descriptive and inferential methodologies according to the type of study design for answering a particular research question

**Communication Skills:**

7) Communicates effectively both in writing and orally (unless a handicap precludes one of those forms of communication)

8) Interpreting and presenting accurately and effectively demographic, statistical, and scientific information for professional and lay audiences adapting and translating public health concepts to individuals and communities

9) Leading and participating in groups to address specific issues, including ability to work in teams, span organizational boundaries and cross systems

**Course Notes:** Notes will be posted online before lecture

**Required Texts/Readings:**

- *Reproducible Research with R and RStudio* by Christopher Gandrud: [Amazon](https://www.amazon.com), [GitHub repo](https://github.com).
- *Happy Git with R* by Jenny Bryan.
- *R for Data Science* by Hadley Wickham and Garrett Grolemund.
- *Advanced R* by Hadley Wickham.
- *R Packages* by Hadley Wickham.
**Course Requirements:** Successful completion of five homework, examinations, and active class participation. Homework as a programming project will be distributed every other week. Evaluation will be dependent upon commit to Git repo.

**Grading/Student Evaluation:** Due dates will be given for each assignment. Late homework will not be accepted.

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>30%</td>
</tr>
<tr>
<td>Final Project</td>
<td>30%</td>
</tr>
<tr>
<td>Student Presentation</td>
<td>30%</td>
</tr>
<tr>
<td>Attendances</td>
<td>10%</td>
</tr>
</tbody>
</table>

Final grades are based on the following point system:

- A = 90 – 100%
- B = 80 – 89%
- C = 70 – 79%
- D = 60 – 69%
- E = 59% or less

**Class Attendance/Participation:** Attendance will be not be traced, you are responsible for everything that goes on in class, including any alteration to the syllabus. If I make an announcement in class, you are responsible for it. We will have 3-5 quizzes during the semester. The quizzes will be in-class and closed book. The best way to prepare for quizzes is to attend classes, read lecture notes, and do homework.

All holidays or special events observed by organized religions will be honored for those students who show affiliation with that particular religion. Absences pre-approved by the UA Dean of Students (or Dean’s designee will be honored.)

**Course Schedule:** Refer to excel table in D2L.

**Required Statements:**

**Communications:** You are responsible for reading emails sent to your UA account from your professor and the announcements that are placed on the course web site. Information about readings, news events, your grades, assignments and other course related topics will be communicated to you with these electronic methods. The official policy can be found at: [http://www.registrar.arizona.edu/emailpolicy.htm](http://www.registrar.arizona.edu/emailpolicy.htm)

**Disability Accommodation:** If you anticipate issues related to the format or requirements of this course, please meet with me. I would like us to discuss ways to ensure your full participation in the course. If you determine that formal, disability-related accommodations are necessary, it is very important that you be registered with Disability Resources (621-3268; drc.arizona.edu) and notify me of your eligibility for reasonable accommodations. We can then plan how best to coordinate your accommodations. The official policy can be found at: [http://catalog.arizona.edu/2013%2D14/policies/disability.htm](http://catalog.arizona.edu/2013%2D14/policies/disability.htm)

**Academic Integrity:** All UA students are responsible for upholding the University of Arizona Code of Academic Integrity, available through the office of the Dean of Students and online: The official policy found at: [http://deanofstudents.arizona.edu/codeofacademicintegrity](http://deanofstudents.arizona.edu/codeofacademicintegrity)

**Classroom Behavior:** (Statement of expected behavior and respectful exchange of ideas)
The Dean of Students has set up expected standards for student behaviors and has defined and identified what is disruptive and threatening behavior. This information is available at: http://deanofstudents.arizona.edu/disruptiveandthreateningstudentguidelines

Students are expected to be familiar with the UA Policy on Disruptive and Threatening Student Behavior in an Instructional Setting found at: http://policy.arizona.edu/disruptive-behavior-instructional and the Policy on Threatening Behavior by Students found at: http://deanofstudents.arizona.edu/sites/deanofstudents.arizona.edu/files/Disruptive_threat_bklt_2012.pdf

Grievance Policy: Should a student feel he or she has been treated unfairly, there are a number of resources available. With few exceptions, students should first attempt to resolve difficulties informally by bringing those concerns directly to the person responsible for the action, or with the student's graduate advisor, Assistant Dean for Student and Alumni Affairs, department head, or the immediate supervisor of the person responsible for the action. If the problem cannot be resolved informally, the student may file a formal grievance using the Graduate College Grievance Policy found at: http://grad.arizona.edu/academics/policies/academic-policies/grievance-policy

Grade Appeal Policy: http://catalog.arizona.edu/2013-14/policies/gradappeal.htm

Syllabus Changes: Information contained in the course syllabus, other than the grade and absence policies, may be subject to change with reasonable advance notice, as deemed appropriate.

Plagiarism: What counts as plagiarism?
• Copying and pasting information from a web site or another source, and then revising it so that it sounds like your original idea.
• Doing an assignment/essay/take home test with a friend and then handing in separate assignments that contain the same ideas, language, phrases, etc.
• Quoting a passage without quotation marks or citations, so that it looks like your own.
• Paraphrasing a passage without citing it, so that it looks like your own.
• Hiring another person to do your work for you, or purchasing a paper through any of the on- or off-line sources.

Live Online Modality:

• This class is scheduled to be taught in the LIVE ONLINE modality.
• Meeting Times: The class will meet Mon/Wed 9am-10:15am via Zoom. Our synchronous meetings will give us the opportunity to go through course material together, troubleshooting the commuting techniques together, and Q&As in real time.
• Staying current: You are required to complete three homeworks and one final project on your own time to accomplish learning objectives listed above.
• Class attendance:
  o If you feel sick or may have been in contact with someone who is infectious, stay home. Except for seeking medical care, avoid contact with others and do not travel.
  o Notify your instructors if you will be missing an in person or online course.
  o Campus Health is testing for COVID-20. Please call (520) 621-9202 before you visit in person.
  o Visit the UArizona COVID-20 page for regular updates.
• **Academic advising:** If you have questions about your academic progress this semester, or your chosen degree program, please note that advisors at the Advising Resource Center can guide you toward university resources to help you succeed.

• **Life challenges:** If you are experiencing unexpected barriers to your success in your courses, please note the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office can be reached at 520-621-2057 or DOS-deanofstudents@email.arizona.edu.

• **Physical and mental-health challenges:** If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call (520-621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.

• **Exams:** We do not have exams. We will have one final project that requires presentation during last class.

• **Class Recordings:**
  - All the course video’s will be posted after class. If course recordings are being made, I will notify the class by email. If you do not wish to be identified by name, please let me know and we can discuss possible ways to address it.
  - For lecture recordings, which are used at the discretion of the instructor, students must access content in D2L only. Students may not modify content or re-use content for any purpose other than personal educational reasons. All recordings are subject to government and university regulations. Therefore, students accessing unauthorized recordings or using them in a manner inconsistent with UArizona values and educational policies are subject to suspension or civil action.
<table>
<thead>
<tr>
<th>Week</th>
<th>Day</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mon</td>
<td>8/24/20</td>
<td>Introduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Basic Programming Skills</strong></td>
</tr>
<tr>
<td>2</td>
<td>Wed</td>
<td>8/26/20</td>
<td>Linux basics</td>
</tr>
<tr>
<td>3</td>
<td>Mon</td>
<td>8/31/20</td>
<td>Linux basics</td>
</tr>
<tr>
<td>4</td>
<td>Wed</td>
<td>9/2/20</td>
<td>Reproducible research, Git/GitHub</td>
</tr>
<tr>
<td>5</td>
<td>Mon</td>
<td>9/7/20</td>
<td>Holiday</td>
</tr>
</tbody>
</table>
| 6    | Wed   | 9/9/20   | Markdown Tutorial (watch video and study the content on: [https://rmarkdown.rstudio.com/lesson-1.html](https://rmarkdown.rstudio.com/lesson-1.html))  
No need to come to class but be sure to study it as we are using rmarkdown for homeworks |
| 7    | Mon   | 9/14/20  | Statistical computing using R                                         |
|      |       |          | **Advance R**                                                         |
| 8    | Wed   | 9/16/20  | R programming (benchmark, debug, profile), parallel computing, R package |
| 9    | Mon   | 9/21/20  | Data Visualization, ggplot2                                           |
| 10   | Wed   | 9/23/20  | dplyr                                                                |
| 11   | Mon   | 9/28/20  | tidy data                                                            |
| 12   | Wed   | 9/30/20  | stringr                                                              |
| 13   | Mon   | 10/5/20  | Shiny for interactive documentation                                   |
|      |       |          | **Advanced (e.g., cluster/cloud) computing**                        |
| 14   | Wed   | 10/7/20  | Web scraping                                                          |
| 15   | Mon   | 10/12/20 | Web scraping                                                          |
| 16   | Wed   | 10/14/20 | HPC at UofA (guest lecture)                                          |
| 17   | Mon   | 10/19/20 | Cluster computing at UofA                                            |
| 18   | Wed   | 10/21/20 | Cloud computing with Google Cloud Platform                           |
| 20   | Mon   | 10/26/20 | Docker                                                               |
| 21   | Mon   | 11/2/20  | Neural network and deep learning (intro)                             |
| 22   | Wed   | 11/4/20  | Neural network and deep learning (examples)                          |
|      |       |          | **Database and healthcare informatics**                             |
| 23   | Mon   | 11/9/20  | Databases and EMR                                                   |
| 24   | Wed   | 11/11/20 | Introduction to databases                                            |
| 25   | Mon   | 11/16/20 | Publicly available recourse for collaborative medicine, MIMIC-III clinical database |
| 26   | Wed   | 11/23/20 | Obs data analysis: study design, assessing causal relationship; methods for causal Inference, IPW, PS |
| 27   | Mon   | 11/25/20 | No class                                                             |
| 28   | Wed   | 11/29/20 | Case Studies - outcome prediction                                    |
| 29   | Mon   | 12/2/20  | Case Studies - comparative effectiveness research                    |
|      |       |          | **Conclusions**                                                      |
| 30   | Wed   | 12/4/20  | Guest lecture (DS experience in Facebook)                            |
| 31   | Mon   | 12/9/20  | Concluding remarks (e.g., recourse in DS, topics not covered but important) |
| 32   | Wed   | 12/11/20 | PresentationHW                                                       |

HW1 due

Student

Student

Student

HW2

HW2 due/HW3

Final Report

Final Report Due