Course Information

ADTA 5120 – Introduction to Data Analytics / CSCE 5300 Introduction to Big Data and Data Science
Fall 2019 – August 26 – December 13, 2019
Class Meeting Time/Location – Tuesdays, 6:30-7:50 pm, FRSC 111

Instructor Contact

Denise R Philpot, PhD, MBA
FRSC 126
Office hours: Tuesdays and Wednesdays, 3:30 pm – 5:30 pm; all other times by appointment
Denise.Philpot@unt.edu

About the Professor / Instructor

Welcome to ADTA 5120 Introduction to Data Analytics / CSCE 5300 Introduction to Big Data and Data Science. I am Dr. Denise Philpot, the instructor for this course and the Advanced Data Analytics program advisor. Prior to earning my doctoral degree in Applied Technology and Performance Improvement with a minor in Management Science, I was a systems analyst/customer account manager for Xerox Corporation. While teaching high school business courses I earned my MBA in Organizational Behavior/Human Resource Management. Like most of my students, I have not followed the traditional academic path. Perhaps I am a bit biased, but I think that enables me to see both the academic and practitioner viewpoints which helps me to add meaning to textbook material that you can apply to real world experiences. The variety of career and academic experiences that we all have provide the foundation for interesting course discussions.

I am excited to have you in this course and look forward to learning more about you and your career goals. Together we will explore a variety of statistical analysis tools, learn about how and when to use them, interpret the outputs of the analysis, and describe the results in ways that will help us or others take appropriate actions to achieve the desired outcomes or goals. Together we will do great things!

Course Pre-requisites, Co-requisites, and/or Other Restrictions

This course requires that the student has successfully completed college level mathematics and a basic statistics course prior to enrollment or have relevant current work experience that will enable them to be successful in an introductory graduate-level statistics course.
Required Materials

Two books are required for this course. Other supplemental materials will be provided via a link to the UNT Willis Library website or included in the Content folders on Blackboard. Students will also need to have access to Microsoft Excel and IBM SPSS for data analysis assignments.


Technology: Students will need access to a computer with MS 2016 Office Suite, especially Excel. If you do not have this version, you can access Office 365 via your student email account. Students will use SPSS for some projects. There is access to SPSS through campus resources, via the Virtual Lab (access will be provided in about a week). If you prefer to purchase the software, you can. There is a student discount that you can take advantage of but you must make sure to purchase the correct license. The license should give you access for one year. I recommend you purchase the SPSS Statistics Premium GradPack 25 – it includes decision trees which we will use in class and ADTA 5230. Do not get the “Base” version – it has very limited capability nor the “Standard” version which does not include the decision tree analysis capability. The link below gives several options for acquiring a lease of the SPSS software.


Course Description for ADTA 5120 Introduction to Data Analytics

This course introduces fundamental concepts of data analytics, including framing business problems, data wrangling, exploratory data analysis, statistical learning models, data analysis software and programming, communicating and operationalizing analysis results, and data ethics. The course focuses on applications of data analytic methods in framing and answering strategic questions facing decision makers in a variety of business sectors. A case study approach is used to introduce key data analytic methods which are explored in more depth in other Advanced Data Analytics courses.
Course Objectives

1. By the end of the course, students should be able to:

2. Articulate key advances in contemporary data analytics, and describe the skill sets needed to be successful in a data analytics career.

3. Describe the data analytics project lifecycle and key elements of each phase.

4. Frame analysis objectives and develop an analysis plan to solve business problems.

5. Use basic tools and methods to obtain, assess and prepare data for analysis.

6. Utilize exploratory data analysis methods to understand characteristics of data sets relative to business objectives.

7. Apply appropriate survey and sampling methods, and identify limitations in survey and sample data.

8. Calculate, interpret and assess parameter estimates and associated inferential statistics.

9. Recognize situations in which multivariate data analysis methods can be applied to address business objectives, and identify which concepts and techniques are necessary to solve a specific problem.

10. Articulate best practices related to data privacy and ethics issues.

11. Effectively communicate analysis results and insights verbally and in writing, presenting descriptive statistics and models in business context and employing appropriate data visualizations.

12. Apply data analysis methods to address business problems from real world case studies.

Course Topics may include:

1. Introduction: What is data science?

2. Data Science Lifecycle

3. Ethics, Privacy and Security

4. Data Acquisition, Preparation and Management

5. Exploratory Data Analysis

6. Surveys, Sampling and Estimation

7. Data Analysis Software

8. Communicating and Operationalizing Results

9. Team Projects and Presentations

Course Description for CSCE 5300 Introduction to Big Data and Data Science

Introduction to Big Data and Data Science including an overview of the field, technical challenges, computational approaches, practical applications, structured and unstructured data processing, empirical methods in computer science, data analytics and learning, data visualization, privacy and ethics. Emphasis on Big Data and its effect on other topics within Data Science, its technical characteristics, and state-of-the-art Big Data analytics architectures and tools.
Teaching Philosophy

It is my goal to create a learning environment in which students feel respected, are engaged in the activities, and bring their questions, experiences, and ideas to the classroom. For real learning to occur, we must work together to achieve a common goal: mastery of the curriculum and the ability to apply what is learned to future activities both in and out of the classroom. In support of the learning objective, I commit to you, to be fully engaged in the classroom, to be available outside of the classroom, and to share my knowledge and experiences with you to enhance the learning process. I believe that learning should be fun (not necessarily easy or without hard work) and that I can learn from you, too. I expect each student to work at their full capacity, respect others, and participate in the classroom so that their experiences can add to the overall learning experience. Lifelong learning is the foundation of my commitment to you for ensuring that the ideas, concepts, theories, and practices I bring to the classroom are current, relevant, and of value to you.

TECHNICAL REQUIREMENTS / ASSISTANCE

Access and Log in Information

This course was developed and will be facilitated utilizing the CANVAS Learning Management System. To get started with the course, please go to: https://unt.instructure.com/login/ldap

You can access student guides on Canvas at this site. You will need your EUID and password to log in to the course. If you do not know your EUID or have forgotten your password, please go to: https://ams.unt.edu/

The Canvas Student app has a mobile version of Canvas that helps students stay current with their courses anywhere. Download the Canvas Student app on Android and iOS devices. For iOS devices, see: How do I download the Canvas Student app on my iOS device? https://community.canvaslms.com/docs/DOC-9831-18561185379

For Android devices, see: How do I download the Canvas Student app on my Android device? https://community.canvaslms.com/docs/DOC-9758-18555199445

Student Academic Support Services

Links to all these services can be found on the Online Student Resources tab within the Canvas Help function.

a. Academic Resource Center: buy textbooks and supplies, access academic catalogs and programs, register for classes, and more.

b. Center for Student Rights and Responsibilities: provides Code of Student Conduct along with other useful links.

c. Office of Disability Accommodation: ODA exist to prevent discrimination on the basis of disability and to help students reach a higher level of independence. https://disability.unt.edu/
d. Counseling and Testing Services: CTS provides counseling services to the UNT community as well as testing services; such as admissions testing, computer-based testing, career testing and other tests.  http://studentaffairs.unt.edu/counseling-and-testing-services

e. UNT Libraries: online library services http://www.library.unt.edu/services

f. Online Tutoring: chat in real time, mark up your paper using drawing tools and edit the text of your paper with the tutor’s help.

g. The Learning Center Support Programs: various program links provided to enhance the student experience.  https://learningcenter.unt.edu/

h. Supplemental Instruction: program for every student, not just for students that are struggling.

i. UNT Writing Lab: offers free writing tutoring to all UNT students, undergraduate and graduate.  http://writingcenter.unt.edu/

j. Math Tutor Lab: http://math.unt.edu/mathlab/

k. Succeed at UNT: how to be a successful student information.  https://success.unt.edu/

The following information is provided to assist you in preparation for the technological aspect of the course.

UIT Help Desk: http://it.unt.edu/help-desk-resources-students

Browser requirements: You need a browser that interfaces well with Canvas, such as Microsoft Internet Explorer or Mozilla Firefox.
https://clear.unt.edu/supported-technologies/canvas/requirements

Word Processor
Creating and submitting files in Microsoft Office, the standard software for this course.

STUDENT TECHNICAL SUPPORT

The University of North Texas UIT Student Helpdesk provides student technical support in the use of Canvas and supported resources. The student help desk may be reached at:
Email: helpdesk@unt.edu Phone: 940.565-2324
In Person: Sage Hall, Room 130
Our hours are:

- Monday-Thursday 8am-midnight
- Friday 8am-8pm
- Saturday 9am-5pm
- Sunday 8am-midnight

Technical Skill Requirements

Students should be able to upload and download files, perform data analysis using Microsoft Excel, and access the Internet for course support materials. Effective navigation of Blackboard
is necessary as course assignments and support materials will be made available through this application. Email will be used to communicate to students via the UNT provided student email accounts.

Netiquette

Student behavior that interferes with an instructor’s ability to conduct a class or other students’ opportunity to learn is unacceptable and disruptive and will not be tolerated in any instructional forum at UNT. This includes but is not limited to comments made on discussion boards or other unacceptable communications between students in an online or blended learning environment. Inappropriate behaviors will be handled based upon the UNT Student Conduct and Discipline Policy which can be found at deanofstudents.unt.edu/conduct.

For those students that are new to online learning or assignments on a Learning Management System like Canvas, you may find these guidelines developed by Albion and Seth T. Ross to be very helpful.


Additional tips for communicating via email and discussion boards

When emailing:

- Use clear and concise language.
- Remember that all college level communication should have correct spelling and grammar. Avoid slang and texting abbreviations. Limit the use of emoticons.
- Avoid using the caps lock – AS IT CAN BE INTERPRETTED AS YELLING.
- Be cautious when using humor or sarcasm. Tone is sometimes lost in an email or discussion post and your message may be taken seriously.
- Use a professor’s proper title – Dr. or Prof. – or if you are in doubt use Mr. or Ms. Don’t refer to faculty by their first name unless specifically invited.
- Use a descriptive subject line.
- Avoid attachments unless you are sure your recipients can open them.
- Sign your message with your name.
- Think before you send an e-mail to more than one person. Does everyone really need to see your message? Likewise, be sure you really want everyone to receive your response when you click, “reply all.”
- Be sure that a message’s author intended for the information to be distributed before you click the “forward” button.
- Be careful with personal information (both yours and other’s).

When posting on discussion boards:

- Make posts that are on topic and within the scope of the course material.
- Take your posts seriously. Review and edit your posts before sending.
- Be as brief as possible while still making a thorough comment.
- Always give proper credit when referencing or quoting another source.
- Be sure to read all messages in a thread before replying.
• Avoid short, generic replies such as, “I agree.” You should include why you agree or add to the previous point.
• Always be respectful of others’ opinions even when they differ from your own. Express differing opinions in a respectful, non-critical way.
• Do not make personal or insulting remarks.
Course Requirements

This course is organized around real-world case studies, which are used to introduce data analytic methods and build professional skills. There will be a directed assignment, an analysis report and/or a presentation associated with each case study.

Your final grade will be determined based on weekly assignments, class participation via discussion boards, journals, and a final team project. Class participation 20%; weekly analysis assignments 40%; and, the final research project/presentation 40%. The total number of points received will be divided by the total possible number of points to determine your final grade.

<table>
<thead>
<tr>
<th>Assignments</th>
<th>Points Possible</th>
<th>Percentage of Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class Participation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 3 discussion board assignments @ 25 points each</td>
<td>200 points</td>
<td>20%</td>
</tr>
<tr>
<td>• 3 journal assignments @ 25 points each</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Peer Review associated with Team Project @ 50 points</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bi-Weekly Analysis Assignments</strong></td>
<td>400 points</td>
<td>40%</td>
</tr>
<tr>
<td>• See Case Study table with case studies and deliverables</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Team Project and Presentation</strong></td>
<td>400 points</td>
<td>40%</td>
</tr>
<tr>
<td>• See Case Study table for deliverables</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Points Possible</strong></td>
<td>1000 points</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Deliverable</th>
<th>Possible Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm-up Exercises</td>
<td>Written solutions to exercises.</td>
<td>100</td>
</tr>
<tr>
<td>Retail Location Analysis (data file subject to change)</td>
<td>Analysis report detailing results of data cleaning, EDA, variable selection for optimal retail location, description of multi-variate regression model with location recommendation.</td>
<td>150</td>
</tr>
<tr>
<td>Market Research (Factor Analysis)</td>
<td>Research report on survey design and how it facilitates factor analysis. Include discussion on validity and reliability, defining constructs/dimensions/ and sampling design.</td>
<td>150</td>
</tr>
<tr>
<td>Automated Medical Diagnosis (data file subject to change)</td>
<td>Analysis report detailing results of EDA, variable selection for automated medical diagnosis (decision tree) model. Presentation of findings.</td>
<td>300</td>
</tr>
</tbody>
</table>

Grading

Course grades will be assigned based on this percentage with a standard 10-point grading scale (100% — 90%, A; 89% — 80%, B; 79% — 70%, C; 69% — 60%, D; 59% — 0%, F).
Introduction to Data Analytics

Course Assignment, Examination, and or Project Policies

Team Data Analytics Research Project
The final project and presentation are due at the end of the course. Each person will submit a research paper that includes an introduction, brief review of statistical technique being used and how it might improve outcomes, problem statement/hypothesis, methods/analysis section (what you did/how you did it), results, and discussion. Also part of the final project is a brief presentation which should include visual aids such as a PowerPoint presentation. Total points for the final project/presentation will be 400 points. It is expected that the paper be free from grammatical errors and appropriately use APA style for citations and reference list. The minimum requirement for the paper will be 10 pages of content, double-spaced, 1-inch margins, using Arial or Times Roman 12 point font. The submitted research paper should also include a separate cover page that includes your names and the title of your paper as well as a reference list formatted using the current APA style guide. You are not required to include an abstract for this paper. A rubric for the project will be provided. The paper and presentation files are due on May 8th, at 10:00 pm CST. Late papers will not be accepted. The paper will be submitted for grading via software that checks for plagiarism. Plagiarism is a violation of the Student Code of Conduct and will be handled per university policy.

Discussion Boards (25 points each)
There will be three discussion board assignments. Each discussion board forum will focus on a question related to the textbook reading or supplemental readings that will be posted to Blackboard. To earn full points on discussion boards, students must be actively engaged in the group discussion and provide input to each of the assigned questions. As graduate students, it is expected that your responses be thoughtful, respectful, grammatically correct, and show your understanding of the topic being discussed.

Journals (25 points each)
There will be three journal assignments. These are reflective in nature and are designed for you to share your thoughts and experiences related to the topic presented. There will be pre-reading assigned with each journal assignment that will be provided by your professor or come from the textbook. As graduate students, it is expected that your responses be thoughtful, grammatically correct, and show your understanding of the topic being discussed. Journal assignments are not seen by your peers and do not require responses to their entries.

Warm Up and Case Study Assignments (see table for points)
There will be three (3) class assignments given during the course that are related to data and specific industry problems presented in the course. You will be required to conduct data analysis and interpret the results. Written responses are expected to be free of grammatical errors. Data analysis should include a brief discussion of the steps you used to complete the analysis.

Course Expectations
It is my goal to create a learning environment in which students feel respected, are engaged in the activities, and bring their questions, experiences, and ideas to the classroom. For real learning to occur, we must work together to achieve a common goal: mastery of the curriculum and the ability to apply what is learned to future activities both in and out of the classroom. In
support of the learning objective, I commit to you, to be fully engaged in the classroom, to be available outside of the classroom, and to share my knowledge and experiences with you to enhance the learning process. I believe that learning should be fun (not necessarily easy or without hard work) and that I can learn from you, too. I expect each student to work at their full capacity, respect others, and participate in the classroom so that their experiences can add to the overall learning experience. Lifelong learning is the foundation of my commitment to you for ensuring that the ideas, concepts, theories and practices I bring to the classroom are current, relevant, and of value to you.

As a student in this course you have responsibilities that must be met to ensure your success as listed below.

1. Students will regularly access Blackboard, attend virtual class meetings, and participate in online discussions.
2. Students will be responsible for checking course announcements in Blackboard and checking student email daily.
3. Students will complete weekly readings and assigned work by stated deadlines.
4. Students will be responsible for downloading data used for projects as directed.
5. Students will be responsible for obtaining software required for completing assigned work as directed.

Policies
Assignment Policy / Late Work
All work for this course is due no later than 11:59 pm CST on the designated due date (Sundays, throughout the semester, unless specifically noted). Please note that the final project and presentation are due on a Wednesday, May 8th, because the semester ends on Friday, May 10th. Any assignment submitted after that time will receive a highest possible score of 50%. Additional points may be deducted when the assignment is graded based on the quality of the work submitted. Work submitted more than 48 hours after the due date will not be accepted, and the student will receive a zero for that assignment. Please don’t lose valuable points this semester by turning in work late.

**Late work is subject to penalty described above unless previously approved by the instructor**

Instructor Responsibilities and Feedback
- As the instructor, it is my responsibility to help students grow and learn; provide clear instructions for projects and assessments, answer questions about assignments, identify additional resources as necessary, provide rubrics, and continually review and update course content based upon learning outcomes and changes in the field of study.
- Feedback on assignments will be provided in a timely manner. Students can expect responses to emails within 24 hours. Grades for weekly assignments will be posted the following week. Project grades will be posted as they are completed.

Turnitin Notice
All works submitted for credit must be original works created by the scholar uniquely for the class. It is considered inappropriate and unethical, particularly at the graduate level, to make
duplicate submissions of a single work for credit in multiple classes, unless specifically requested by the instructor. Work submitted at the graduate level is expected to demonstrate higher-order thinking skills and be of significantly higher quality than work produced at the undergraduate level. Turnitin is used as a tool to assist students in their scholarly writing to address plagiarism issues. It is recommended that students use this resource to ensure their work is free of copyright issues prior to final submission of their projects.

**Class Participation**
Students are required to login regularly to the online class site. The instructor will use the tracking feature in Blackboard to monitor student activity. Students are also required to participate in all class activities such as discussion board, chat or conference sessions and group projects.

**Virtual Classroom Citizenship**
The same guidelines that apply to traditional classes should be observed in the virtual classroom environment. Please use proper netiquette when interacting with class members and the professor.

**Incompletes**
Incompletes will only be given per university policy.

http://registrar.unt.edu/grades/incompletes

**UNT POLICIES**

**Student Conduct and Discipline:**
You are encouraged to become familiar with the University’s Code of Student Conduct and the Policy of Academic Integrity found on the Dean of Students website. The policies contained on this website apply to this course. If you have questions regarding any of the information presented regarding academic integrity, please feel free to contact me. I will be happy to review any of your work prior to final submission for grading.

The UNT Code of Student Conduct can be found here: http://deanofstudents.unt.edu/conduct

The UNT policy regarding Academic Integrity can be found here: http://policy.unt.edu/policy/06-003

**ADA Policy**
The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking reasonable accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide you with a reasonable accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You may request reasonable accommodations at any time, however, ODA notices of reasonable accommodation should be provided as early as possible in the semester to avoid any delay in implementation. Note that students must obtain a new letter of reasonable accommodation for every semester and must meet with each faculty member prior to implementation in each class. Students are strongly encouraged to deliver letters of reasonable accommodation during faculty office hours or by
appointment. Faculty members have the authority to ask students to discuss such letters during their designated office hours to protect the privacy of the student. For additional information see the Office of Disability Accommodation website at http://www.unt.edu/oda. You may also contact them by phone at 940.565.4323.

Sexual Assault Prevention

UNT is committed to providing a safe learning environment free of all forms of sexual misconduct, including sexual harassment sexual assault, domestic violence, dating violence, and stalking. Federal laws (Title IX and the Violence Against Women Act) and UNT policies prohibit discrimination on the basis of sex, and therefore prohibit sexual misconduct. If you or someone you know is experiencing sexual harassment, relationship violence, stalking, and/or sexual assault, there are campus resources available to provide support and assistance. UNT’s Survivor Advocates can assist a student who has been impacted by violence by filing protective orders, completing crime victim’s compensation applications, contacting professors for absences related to an assault, working with housing to facilitate a room change where appropriate, and connecting students to other resources available both on and off campus. The Survivor Advocates can be reached at SurvivorAdvocate@unt.edu or by calling the Dean of Students Office at 940-565-2648. Additionally, alleged sexual misconduct can be non-confidentially reported to the Title IX Coordinator at oeo@unt.edu or at (940) 565 2759.

Important Notice for F-1 Students taking Distance Education Courses

Federal Regulation

To read detailed Immigration and Customs Enforcement regulations for F-1 students taking online courses, please go to the Electronic Code of Federal Regulations website at http://ecfr.gpoaccess.gov. The specific portion concerning distance education courses is located at "Title 8 CFR 214.2 Paragraph (f)(6)(i)(G)" and can be found buried within this document: http://frwebgate.access.gpo.gov/cgi-bin/get-cfr.cgi?TITLE=8&PART=214&SECTION=2&TYPETEXT

The paragraph reads:

(G) For F–1 students enrolled in classes for credit or classroom hours, no more than the equivalent of one class or three credits per session, term, semester, trimester, or quarter may be counted toward the full course of study requirement if the class is taken on-line or through distance education and does not require the student's physical attendance for classes, examination or other purposes integral to completion of the class. An on-line or distance education course is a course that is offered principally through the use of television, audio, or computer transmission including open broadcast, closed circuit, cable, microwave, or satellite, audio conferencing, or computer conferencing. If the F–1 student's course of study is in a language study program, no on-line or distance education classes may be considered to count toward a student's full course of study requirement.

University of North Texas Compliance

To comply with immigration regulations, an F-1 visa holder within the United States may need to engage in an on-campus experiential component for this course. This component (which must be approved in advance by the instructor) can include activities such as taking an on-campus exam, participating in an on-campus lecture or lab activity, or other on-campus experience.
integral to the completion of this course.

If such an on-campus activity is required, it is the student’s responsibility to do the following:

(1) Submit a written request to the instructor for an on-campus experiential component within one week of the start of the course.
(2) Ensure that the activity on campus takes place and the instructor documents it in writing with a notice sent to the International Student and Scholar Services Office. ISSS has a form available that you may use for this purpose.

Because the decision may have serious immigration consequences, if an F-1 student is unsure about his or her need to participate in an on-campus experiential component for this course, s/he should contact the UNT International Student and Scholar Services Office (telephone 940-565-2195 or email internationaladvising@unt.edu) to get clarification before the one-week deadline.
## Course Calendar – Fall 2019

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic / Required Reading</th>
<th>Assignments</th>
</tr>
</thead>
</table>
| Week 1 | Course overview and Syllabus review  
*Introduction to Data Analytics*  
Course and ADA program introduction, data analysis fundamentals, data analysis process  
Due Sept 1 @ 10:59 pm  
Read Chapters  
Review Warm Up Problems |
| Week 2 | *Introduction to Data Analytics*  
Course and ADA program introduction, data analysis fundamentals, data analysis process  
Reading – Practical Statistics, Chap 1: Exploratory Data Analysis | Review Warm Up Problems |
| Week 3 | *Exploratory Data Analysis*  
EDA mind-set, summary statistics, distributions, visualizations  
Reading – Doing Data Science, Chap 5: Logistic Regression And Chap 7: Extracting Meaning from Data | Work on Warm Up Problems  
Due Sept 24 @ 10:59 pm  
Complete Journal #1  
Due Sept 17 |
| Week 4 | *Exploratory Data Analysis*  
EDA mind-set, summary statistics, distributions, visualizations  
Reading – Practical Statistics, Chap 2: Data and Sampling Distributions | Work on Warm Up Problems  
Due Sept 24 @ 10:59 pm |
| Week 5 | *Introduction to Retail Location Analysis*  
Retail location optimization, data cleaning  
Reading – Practical Statistics, Chap 3: Statistical Experiments and Significance Testing | Review Retail Location Project  
Read Chapters |
<table>
<thead>
<tr>
<th>Week</th>
<th>Topic / Required Reading</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 6</td>
<td><strong>Introduction to Retail Location Analysis</strong>&lt;br&gt;Retail location optimization, data cleaning&lt;br&gt;Reading – Doing Data Science, Chap 9: Data Visualization and Fraud Detection</td>
<td>Review Retail Location Project Read Chapters</td>
</tr>
<tr>
<td>Oct 1</td>
<td><strong>Introduction to Retail Location Analysis – Multivariate Regression</strong>&lt;br&gt;Overview of Multivariate Regression&lt;br&gt;Reading – Practical Statistics, Chap 4: Regression and Prediction</td>
<td>Retail Location Project Due Oct 22 @ 10:59 pm Read Chapters Complete Journal #2 Due Oct 15</td>
</tr>
<tr>
<td>Week 8</td>
<td><strong>Introduction to Retail Location Analysis – Multivariate Regression</strong>&lt;br&gt;Overview of Multivariate Regression&lt;br&gt;Reading – Doing Data Science, Chap 9: Data Visualization and Fraud Detection&lt;br&gt;Practical Statistics, Chap 4: Regression and Prediction</td>
<td>Retail Location Project Due Oct 22 @ 10:59 pm Read Chapters</td>
</tr>
<tr>
<td>Oct 15</td>
<td><strong>Introduction to Market Research Case Study</strong>&lt;br&gt;Survey Design and Sampling&lt;br&gt;Reading – Doing Data Science, Chap 10: Social Networks and Data Journalism and Chap 11: Causality</td>
<td>Review Survey Design and Sampling Research paper requirements Read Chapters Virtual Class meeting</td>
</tr>
<tr>
<td>Week 9</td>
<td><strong>Introduction to Market Research Case Study Continued</strong>&lt;br&gt;Survey Design and Sampling&lt;br&gt;Reading – Doing Data Science, Chap 10: Social Networks and Data Journalism and Chap 11: Causality</td>
<td>Review Survey Design and Sampling Research paper requirements Due Nov 12 Read Chapters Complete Journal #3 Due Nov 5</td>
</tr>
<tr>
<td>Oct 22</td>
<td><strong>Introduction to Decision Trees</strong>&lt;br&gt;Reading – Doing Data Science, Chap 13: Lessons Learned from Data Competitions: Data Leakage and Model Evaluation and Chap 14: Data Engineering: MapReduce, Pregel, and Hadoop</td>
<td>Submit research paper on survey design – Due Nov 12</td>
</tr>
<tr>
<td>Week 11</td>
<td><strong>Introduction to Decision Trees</strong>&lt;br&gt;Reading – Doing Data Science, Chap 13: Lessons Learned from Data Competitions: Data Leakage and Model Evaluation and Chap 14: Data Engineering: MapReduce, Pregel, and Hadoop</td>
<td></td>
</tr>
<tr>
<td>Week</td>
<td>Topic / Required Reading</td>
<td>Assignments</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>Week 12</td>
<td>Continue instruction on Decision Trees</td>
<td>Review Project – Medical Diagnosis with Decision Trees</td>
</tr>
<tr>
<td>Nov 12</td>
<td>Reading – Practical Statistics, Chap 6: Statistical Machine Learning</td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td>Reading and Journal Assignment No physical class meeting</td>
<td>Complete Journal #4 Due Nov 25</td>
</tr>
<tr>
<td>Nov 19</td>
<td>Decision Trees</td>
<td>Medical Diagnosis Decision Tree Project Due Dec 3 @ 5:00 pm</td>
</tr>
<tr>
<td></td>
<td>Reading – Doing Data Science, Chap 16: Next Generation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data Scientists, Hubris, and Ethics</td>
<td></td>
</tr>
<tr>
<td>Week 14</td>
<td>Class Presentations of Final Project</td>
<td></td>
</tr>
<tr>
<td>Nov 26</td>
<td>Final Exam Week</td>
<td>Complete Peer Review Dec 12 @ 11:00 pm</td>
</tr>
<tr>
<td></td>
<td>Class Presentations of Final Project</td>
<td></td>
</tr>
<tr>
<td>Week 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec 10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>