Course Syllabus
ADTA 5240 – Harvesting, Storing, and Retrieving Data
CSCE 5200 – Information Retrieval
IPAC 4240 – Data Wrangling
COURSE INFORMATION

- ADTA 5240: Harvesting, Storing, and Retrieving Data – Sections 501 and 551 – 3 credit hours
- CSCE 5200 – Information Retrieval – Section 501 – 3 credit hours
- IPAC 4240 – Data Wrangling – Section 501 – 3 credit hours
- Monday 6:30 PM – 7:50 PM, Frisco Campus Hall Park B Room #128 and using Canvas at (UNT.Instructure.com)

Professor / Instructor Contact Information

- Instructor: Abdulrahman Habib, PhD
- Office Location: UNT New College at Frisco #145
- Office Hours: Monday 2:00 PM – 6:00 PM and online by appointment
- Email Address: habib@unt.edu

About the Professor / Instructor

Welcome to ADTA 5240 – Harvesting, Storing, and Retrieving Data. My name is Abdulrahman Habib, call me Habib. I have over 15 years of work experience in IT project management. I led projects in Smart City, e-government, e-commerce, and ERP implementation for small and medium businesses. I am passionate about how technology improves people lives and contribute to our society. I have a BS in computer science, MS in Engineering Systems, and a Ph.D. in Information Science. My current research interests include applied research on emerging technologies, innovation, and smart cities applications.

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

No required prerequisite for this course. We will work together to help you acquire all the needed skills for this class. It will be helpful if you have a background about Linux, computer network, database, using data files (text & CSV), and MS Excel.

Materials – Text, Readings, Supplementary Readings

No textbook is required for this course. The following books are for suggested reading:
Course Description
This course introduces the fundamentals of data engineering, including collecting, wrangling, storing, retrieving, and processing data. The goal of this course is to provide students with both theoretical knowledge and practical experience leading to mastery of the fundamentals of data engineering, using both small and large datasets. As these fundamentals are introduced, exemplary technologies will be employed to illustrate how storage and processing architectures can be constructed. The problems are being considered in the context of big data analytics. Exercises and examples will consider both simple and complex data structures, as well as data ranges from clean and structured to dirty and unstructured.

Course Objectives
- Develop an understanding of the fundamental concepts of the modern data management, including data science life cycle, data scaling, structuring data, and data lakes
- Develop knowledge and skills in storing, retrieving, and processing data with the Apache Hadoop framework using the cloud technology
- Develop knowledge and skills in working with the Apache Hadoop framework including Hadoop Distributed File System (HDFS), MapReduce, and Hive
- Develop knowledge and skills in working with HDFS and Spark/pySpark
- Develop knowledge and skills in cleansing/wrangling data with Google/Open Refine
- Develop knowledge and skills in collecting data using streaming technologies
- Introduce students to real-time big data using Spark Streaming

Course Topics
- Apache Hadoop framework and its major components
- Apache Hadoop framework and the cloud technology
- Storing & Retrieving data with Apache Hadoop HDFS, MapReduce, and Hive
- Storing & Retrieving data with Apache Hadoop HDFS and Spark
• Data lakes: A storage of choice for the modern data management
• Data cleansing and wrangling with Google/Open Refine
• Data collection with the streaming technologies
• Introduction to real-time big data with Spark Streaming
• Introduction to the Python API for Spark: pySpark

Course Requirements
1. The student will be responsible for checking Canvas announcements. I recommend you using Canvas mobile application and setup your course notifications to your preferred email. Read more about Canvas: https://community.canvaslms.com/docs/DOC-10701
   canvas-student-guide-table-of-contents
2. The student will access and follow all course instructions found in the syllabus, announcements, assignments, and all other class-related documents.
3. The student will complete all the class assignments in the time frame specified in the class documents including the course calendar to participate effectively in class activities.
4. The student will complete all the assessment tests and exams – if required – in the time frame specified in the class documents including the course calendar.
5. The student will complete all the projects – if required – in the time frame specified in the class documents including the course calendar.

Communications
Interaction with Instructor: I look forward to getting to know all of you and working with you. Contact me anytime using my UNT email (habib@unt.edu). I will check messages daily and will make every effort to respond as quickly as possible. If you would like to meet me we can do an online conference call or we can meet in person at UNT New College in Frisco #145. In both cases please send me and email to schedule the meeting. My goal is for you to gain a greater understanding of the course topics while enjoying this course. Learn how to engage in meaningful and useful online course activities and to help you achieve your course goals.

Assignments, Assessments, And Projects
1. There will be five discussions, except for the midterm take-home week and the final week. The student will respond to posted online course discussion questions each week following the instructions for discussion forums. Each student should submit his/her initial posts to respond to the discussion questions by the deadline provided on the Course Calendar. Then he/she should continue by posting responses to his/her classmates’ posts throughout the week to maximize points earned each week. Students are encouraged to enhance the conversation by providing complementary resource materials and properly referenced supplementary items.
2. There will be five technical challenges throughout the course. Students are required to submit their technical challenges on time.
3. There will be a midterm take-home challenge. Students have one week to complete it.
4. There will be a final project.
   a. The student will complete a final project following the project instructions.
   b. Students will submit the final project by the deadline.

Make-Up Policy
No make-up assessment tests or exams will be offered except for being approved in advance. Students will be required to provide necessary documentation.

Late-work Policy
All assignments are to be submitted using the UNT email. The deadline for submitting an assignment is 11:00 PM on the due date. Late submissions can still be submitted up to 24 hours after the deadline. Assignments submitted within 24 hours after the due date/time will be subject to a 25% penalty. No submissions will be accepted later than 24 hours after the deadline.
NOTES: Late work is subject to penalty described above unless previously approved by the instructor.

Grading Policy
The student’s grade (1000 points) in the course consists of the following components:

   - Five Discussions: 30 points each or 150 in total
   - Five Technical Challenges: 60 points each or 300 in total
   - Midterm Technical Challenge: 250 points
   - Final Project: 300 points

The final letter grade will be determined as follows:
- A: 90-100% (Outstanding, excellent work. The student performs well above the minimum criteria.)
- B: 80-89% (Good, impressive work. The student performs above the minimum criteria.)
- C: 70-79% (Solid, college-level work. The student meets the criteria of the assignment.)
- D: 60-69% (Below average work. The student fails to meet the minimum criteria.)
- F: 59 and below (Sub-par work. The student fails to complete the assignment.)

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
Business hours are:
- Monday-Thursday 8am-midnight
- Friday 8am-8pm
- Saturday 9am-5p
- Sunday 8am-midnight

Access & Navigation

Access and Log in Information
This course was developed and will be facilitated utilizing the University of North Texas’ Canvas. To get started with the course, please go to the website at https://canvas.unt.edu

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 the website at http://ams.unt.edu.

Technical Requirements / Assistance
The following information has been provided to assist you in preparation for the technological aspect of the course.

UIT Help Desk: http://www.unt.edu/helpdesk/index.htm
Web Browser
Word Processor
Creating and submitting files in Microsoft Office, the standard software for this course

Minimum Technical Skills Needed:
- Using the learning management system
- Using email with attachments
- Copying and pasting
- Downloading and installing software
- Using spreadsheet programs

Course Policies

Class Participation
Attendance is expected for this class. Students are required to log in regularly to the Canvas class site. Students are also required to participate in all class activities such as discussion boards, conference sessions, and group projects if assigned. To learn more about campus resources and information on how you can achieve success, go to succeed.unt.edu.
COVID-19 Impact on Attendance

While attendance is expected as outlined above, it is important for all of us to be mindful of the health and safety of everyone in our community, especially given concerns about COVID-19. Please contact me if you are unable to attend class because you are ill, or unable to attend class due to a related issue regarding COVID-19. It is important that you communicate with me prior to being absent so I may make a decision about accommodating your request to be excused from class.

If you are experiencing any symptoms of COVID-19 (https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html) please seek medical attention from the Student Health and Wellness Center (940-565-2333 or askSHWC@unt.edu) or your health care provider PRIOR to coming to campus. UNT also requires you to contact the UNT COVID Hotline at 844-366-5892 or COVID@unt.edu for guidance on actions to take due to symptoms, pending or positive test results, or potential exposure. While attendance is an important part of succeeding in this class, your own health, and those of others in the community, is more important.

Remote Instruction after November 28th, 2020

The UNT fall schedule requires this course to have fully remote instruction beginning November 28th. Additional remote instruction may be necessary if community health conditions change or you need to self-isolate or quarantine due to COVID-19. Information on how to be successful in a remote learning environment can be found at https://online.unt.edu/learn.

Statement on Face Covering

Face coverings are required in all UNT facilities. Students are expected to wear face coverings during this class. If you are unable to wear a face covering due to a disability, please contact the Office of Disability Access to request an accommodation. UNT face covering requirements are subject to change due to community health guidelines. Any changes will be communicated via the instructor.

ACADEMIC POLICIES

Scholarly Expectations

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.
Instructor Responsibilities and Feedback
The instructor is responsible for responding to student questions about assignments and projects, about the course material presented, and for providing additional resources to enhance understanding of course material. Timely feedback is essential for student success and the instructor is responsible for providing timely feedback to students throughout the course. The instructor will actively participate in each week’s discussion forum and will provide feedback to students each week regarding their participation. The instructor will grade submitted assignments and will post grades for students within 10 days of assignment due date.

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 as discussed at this link: http://www.albion.com/netiquette/corerules.html.

Rules of Engagement
Rules of engagement refer to the way students are expected to interact with each other and with their instructors. Here are some general guidelines:

- While the freedom to express yourself is a fundamental human right, any communication that utilizes cruel and derogatory language on the basis of race, color, national origin, religion, sex, sexual orientation, gender identity, gender expression, age, disability, genetic information, veteran status, or any other characteristic protected under applicable federal or state law will not be tolerated.
- Treat your instructor and classmates with respect in any communication online or face-to-face, even when their opinion differs from your own.
- Ask for and use the correct name and pronouns for your instructor and classmates.
- Speak from personal experiences. Use “I” statements to share thoughts and feelings. Try not to speak on behalf of groups or other individual’s experiences.
- Use your critical thinking skills to challenge other people’s ideas, instead of attacking individuals.
- Avoid using all caps while communicating digitally. This may be interpreted as “YELLING!”
- Be cautious when using humor or sarcasm in emails or discussion posts as tone can be difficult to interpret digitally.
- Avoid using “text-talk” unless explicitly permitted by your instructor.
- Proofread and fact-check your sources.
- Keep in mind that online posts can be permanent, so think first before you type.

See these Engagement Guidelines (https://clear.unt.edu/online-communication-tips) for more information.

Incompletes
This course will observe the UNT policy on incompletes, found here: http://registrar.unt.edu/grades/incompletes
Add/Drop Policy
The University of North Texas Add Drop Policy for Fall 2017 can be found at the following link: http://registrar.unt.edu/registration/fall-add-drop

Policy on Server Unavailability or Other Technical Difficulties
The University is committed to providing a reliable online course system to all users. However, in the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time-sensitive assessment activity, the instructor will extend the time windows and provide an appropriate accommodation based on the situation. Students should immediately report any problems to the instructor and also contact the UNT Student Help Desk: helpdesk@unt.edu or 940.565.2324. The instructor and the UNT Student Help Desk will work with the student to resolve any issues at the earliest possible time.

Copyright Notice
Some or all of the materials on this course Web site may be protected by copyright. Federal copyright law prohibits the reproduction, distribution, public performance, or public display of copyrighted materials without the express and written permission of the copyright owner unless fair use or another exemption under copyright law applies. Additional copyright information may be located at http://copyright.unt.edu.

Graduate Online Course Attendance Policy
Students are expected to participate actively each week and to meet all deadlines for course assignments as detailed in the Course Calendar. Information about the University of Texas' Attendance Policy may be found at http://policy.unt.edu/policy/15-2-5

Administrative Withdrawal
This course will observe the UNT policy on academic withdrawal found here: https://deanofstudents.unt.edu/withdrawals

Syllabus Change Policy
Changes to the course syllabus or due dates are not anticipated but should they be necessary, the instructor will provide ample notification to students to allow them to complete assignments in a timely manner without penalty.

UNT GENERAL POLICIES

Student Conduct and Discipline: Student Handbook. You are encouraged to become familiar with the University’s Policy of Academic dishonesty found in the Student Handbook. The content of the Handbook applies to this course. If you are in doubt regarding the requirements, please consult with me before you complete any requirements of the course.

The UNT Code of Student Conduct can be found here:
The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking 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 an accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. ... Note that students must obtain a new letter of 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 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://disability.unt.edu/. You may also contact them by phone at 940.565.4323.

The University of North Texas (UNT) prohibits discrimination and harassment because of race, color, national origin, religion, sex, sexual orientation, gender identity, gender expression, age, disability, genetic information, veteran status, or any other characteristic protected under applicable federal or state law in its application and admission processes; educational programs and activities; employment policies, procedures, and processes; and university facilities. The University takes active measures to prevent such conduct and investigates and takes remedial action when appropriate.

UNT uses a system called Eagle Alert to quickly notify students with critical information in the event of an emergency (i.e., severe weather, campus closing, and health and public safety emergencies like chemical spills, fires, or violence). In the event of a university closure, please refer to Canvas for contingency plans for covering course materials.

Student records pertaining to this course are maintained in a secure location by the instructor of record. All records such as exams, answer sheets (with keys), and written papers submitted during the duration of the course are kept for at least one calendar year after course completion. Course work completed via the Canvas online system, including grading information and comments, is also stored in a safe electronic environment for one year. Students have the right to view their individual record; however, information about student’s records will not be divulged to other individuals without proper written consent. Students are encouraged to review the Public Information Policy and the Family Educational Rights and
Privacy Act (FERPA) laws and the University’s policy. See UNT Policy 10.10, Records Management and Retention for additional information.

**Acceptable Student Behavior**

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. Students engaging in unacceptable behavior will be directed to leave the classroom and the instructor may refer the student to the Dean of Students to consider whether the student's conduct violated the Code of Student Conduct. The University's expectations for student conduct apply to all instructional forums, including University and electronic classroom, labs, discussion groups, field trips, etc. Visit UNT’s [Code of Student Conduct](https://deanofstudents.unt.edu/conduct) to learn more.

**Access to Information - Eagle Connect**

Students’ access point for business and academic services at UNT is located at: [my.unt.edu](https://my.unt.edu). All official communication from the University will be delivered to a student’s Eagle Connect account. For more information, please visit the website that explains Eagle Connect and how to forward e-mail [Eagle Connect](https://it.unt.edu/eagleconnect).

**Student Evaluation Administration Dates**

Student feedback is important and an essential part of participation in this course. The student evaluation of instruction is a requirement for all organized classes at UNT. The survey will be made available during weeks 13, 14 and 15 [insert administration dates] of the long semesters to provide students with an opportunity to evaluate how this course is taught. Students will receive an email from "UNT SPOT Course Evaluations via IASystem Notification" ([no-reply@iasystem.org](mailto:no-reply@iasystem.org)) with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. Once students complete the survey they will receive a confirmation email that the survey has been submitted. For additional information, please visit the [SPOT website](http://spot.unt.edu/) or email [spot@unt.edu](mailto:spot@unt.edu).

**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**


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.

**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.
Student Verification
UNT takes measures to protect the integrity of educational credentials awarded to students enrolled in distance education courses by verifying student identity, protecting student privacy, and notifying students of any special meeting times/locations or additional charges associated with student identity verification in distance education courses. See UNT Policy 07-002 Student Identity Verification, Privacy, and Notification and Distance Education Courses (https://policy.unt.edu/policy/07-002).

Use of Student Work
A student owns the copyright for all work (e.g. software, photographs, reports, presentations, and email postings) he or she creates within a class and the University is not entitled to use any student work without the student’s permission unless all of the following criteria are met:

- The work is used only once.
- The work is not used in its entirety.
- Use of the work does not affect any potential profits from the work.
- The student is not identified.
- The work is identified as student work.

If the use of the work does not meet all of the above criteria, then the University office or department using the work must obtain the student’s written permission.

Download the UNT System Permission, Waiver and Release Form

Transmission and Recording of Student Images in Electronically-Delivered Courses

1. No permission is needed from a student for his or her image or voice to be transmitted live via videoconference or streaming media, but all students should be informed when courses are to be conducted using either method of delivery.

2. In the event an instructor records student presentations, he or she must obtain permission from the student using a signed release in order to use the recording for future classes in accordance with the Use of Student-Created Work guidelines above.

3. Instructors who video-record their class lectures with the intention of re-using some or all of recordings for future class offerings must notify students on the course syllabus if students' images may appear on video. Instructors are also advised to provide accommodation for students who do not wish to appear in class recordings.
Example: This course employs lecture capture technology to record class sessions. Students may occasionally appear on video. The lecture recordings will be available to you for study purposes and may also be reused in future course offerings.

No notification is needed if only audio and slide capture is used or if the video only records the instructor’s image. However, the instructor is encouraged to let students know the recordings will be available to them for study purposes.

Class Recordings & Student Likenesses

All the video recordings in this course, including lecture videos and live class activity recordings, are for students enrolled in the class section to refer to throughout the semester. Class recordings are the intellectual property of the university or instructor and are reserved for use only by students in this class and only for educational purposes. Students may not post or otherwise share the recordings outside the class, or outside the Canvas Learning Management System, in any form. Failing to follow this restriction is a violation of the UNT Code of Student Conduct and could lead to disciplinary action.

ACADEMIC SUPPORT & STUDENT SERVICES

Student Support Services

Mental Health

UNT provides mental health resources to students to help ensure there are numerous outlets to turn to that wholeheartedly care for and are there for students in need, regardless of the nature of an issue or its severity. Listed below are several resources on campus that can support your academic success and mental well-being:

- [Student Health and Wellness Center](https://studentaffairs.unt.edu/student-health-and-wellness-center)
- [Counseling and Testing Services](https://studentaffairs.unt.edu/counseling-and-testing-services)
- [UNT Care Team](https://studentaffairs.unt.edu/care)
- [UNT Psychiatric Services](https://studentaffairs.unt.edu/student-health-and-wellness-center/services/psychiatry)
- [Individual Counseling](https://studentaffairs.unt.edu/counseling-and-testing-services/services/individual-counseling)
Chosen Names
A chosen name is a name that a person goes by that may or may not match their legal name. If you have a chosen name that is different from your legal name and would like that to be used in class, please let the instructor know. Below is a list of resources for updating your chosen name at UNT.

- UNT Records
- UNT ID Card
- UNT Email Address
- Legal Name

*UNT eulDs cannot be changed at this time. The collaborating offices are working on a process to make this option accessible to UNT community members.*

Pronouns
Pronouns (she/her, they/them, he/him, etc.) are a public way for people to address you, much like your name, and can be shared with a name when making an introduction, both virtually and in-person. Just as we ask and don’t assume someone’s name, we should also ask and not assume someone’s pronouns. You can add your pronouns to your Canvas account so that they follow your name when posting to discussion boards, submitting assignments, etc. Below is a list of additional resources regarding pronouns and their usage:

- What are pronouns and why are they important?
- How do I use pronouns?
- How do I share my pronouns?
- How do I ask for another person’s pronouns?
- How do I correct myself or others when the wrong pronoun is used?

Additional Student Support Services

- Registrar (https://registrar.unt.edu/registration)
- Financial Aid (https://financialaid.unt.edu/)
- Student Legal Services (https://studentaffairs.unt.edu/student-legal-services)
- Career Center (https://studentaffairs.unt.edu/career-center)
- Multicultural Center (https://edo.unt.edu/multicultural-center)
- Counseling and Testing Services (https://studentaffairs.unt.edu/counseling-and-testing-services)
- Pride Alliance (https://edo.unt.edu/pridealliance)
- UNT Food Pantry (https://deanofstudents.unt.edu/resources/food-pantry)
Academic Support Services

- Academic Resource Center (https://clear.unt.edu/canvas/student-resources)
- Academic Success Center (https://success.unt.edu/asc)
- UNT Libraries (https://library.unt.edu/)
- Writing Lab (http://writingcenter.unt.edu/)
### Class Schedule
The following is a tentative schedule. Should any change become necessary, it will be announced via the Canvas announcement. It is the student’s responsibility to check for changes in the schedule.

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topics / Required Reading</th>
<th>Assignment Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Aug 31 – Sep 6</td>
<td>Cleansing and Wrangling Data with Google/open Refine and Excel</td>
<td>Technical Challenge #1</td>
</tr>
<tr>
<td>3</td>
<td>Sep 7 – Sep 13</td>
<td>* Monday Sep 7th is Labor day, No Class&lt;br&gt;Introduction to Apache Hadoop Framework&lt;br&gt;Hadoop The Definitive Guide: Ch1</td>
<td>Discussion #1</td>
</tr>
<tr>
<td>4</td>
<td>Sep 14 – Sep 20</td>
<td>Introduction to Google Cloud&lt;br&gt;Apache Hadoop Framework in the cloud&lt;br&gt;Hadoop The Definitive Guide: Ch2</td>
<td>Technical Challenge #2</td>
</tr>
<tr>
<td>5</td>
<td>Sep 21 – Sep 27</td>
<td>Exploring Apache Hadoop System&lt;br&gt;Hadoop The Definitive Guide: Ch3</td>
<td>Discussion #2</td>
</tr>
<tr>
<td>6</td>
<td>Sep 28 – Oct 4</td>
<td>Introduction to Apache Hive&lt;br&gt;Hadoop The Definitive Guide: Ch4</td>
<td>Technical Challenge #3</td>
</tr>
<tr>
<td>7</td>
<td>Oct 5 – Oct 11</td>
<td>Storing &amp; Retrieving Data with HDFS, MapReduce, &amp; Hive&lt;br&gt;Hadoop The Definitive Guide: Ch 5, 7, &amp; 17</td>
<td>Discussion #3</td>
</tr>
<tr>
<td>8</td>
<td>Oct 12 – Oct 18</td>
<td>Midterm Exam</td>
<td>Midterm Exam</td>
</tr>
<tr>
<td>9</td>
<td>Oct 19 – Oct 25</td>
<td>Introduction to Data Lakes&lt;br&gt;Introduction to Apache Spark</td>
<td>Discussion #4</td>
</tr>
<tr>
<td>10</td>
<td>Oct 26 – Nov 1</td>
<td>Apache Spark in the Apache Hadoop Ecosystem&lt;br&gt;High Performance Spark: Ch 1</td>
<td>Technical Challenge #4</td>
</tr>
<tr>
<td>11</td>
<td>Nov 2 – Nov 8</td>
<td>Storing and Querying Data with HDFS &amp; Spark API’s&lt;br&gt;High Performance Spark: Ch 2</td>
<td>Discussion #5</td>
</tr>
<tr>
<td>12</td>
<td>Nov 9 – Nov 15</td>
<td>Introduction to Spark Streaming&lt;br&gt;Collecting data with Streaming Technologies</td>
<td>Technical Challenge #5</td>
</tr>
<tr>
<td>13</td>
<td>Nov 16 – Nov 22</td>
<td>Final Project Presentations &amp; Peer review</td>
<td>Final Project</td>
</tr>
<tr>
<td>14</td>
<td>Nov 23 – Nov 29</td>
<td>* Thanksgiving break Nov 26 – 27, No Class</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Nov 30 – Dec 6</td>
<td>Grading</td>
<td></td>
</tr>
</tbody>
</table>