

**University of California, Berkeley  
Master of Information and Data Science (MIDS)**

**W205 – Fundamentals of Data Engineering**

**2018 Summer, Sections 4, 5, 6, 7**

**Instructor: Kevin R. Crook**

**Instructor's Supplement to the Syllabus**

**Note:** this supplement to the syllabus is only for Kevin Crook's sections and does not apply to other sections. Information, including rubrics, in this supplement will override the main syllabus.

**Primary Syllabus**

Prior to the semester start, the primary syllabus can be found at this link:

**<https://mids-w205-fund-of-data-eng.github.io/course-content>**

After the semester starts, the syllabus will can be found in GitHub Classroom at this link (requires that you are logged into GitHub Classroom, otherwise it will not work):

**<https://github.com/mids-w205-crook/course-content>**

**Office Hours**

- Official Scheduled
  - Tuesday, 6:30 pm to 7:30 pm Berkeley Time
  - I will stay the whole time, so you can come anytime during this time period
  
- Other times I have set aside to be available
  - Please slack me to let me know you are coming. I won't stay unless someone is present or someone has told me in slack that they are coming.
  - 15 minutes before any of my classes
  - Tuesdays, 5:30 pm to 6:30 pm, Berkeley Time
  - Wednesdays, 5:30 pm to 6:10 pm, Berkeley Time
  - Thursdays, 5:30 pm to 6:10 pm, Berkeley Time

## Checklist Before Your First Synchronous Class Meeting

Here is a quick checklist of things you should complete before our first synchronous class session. More details on these items follow in subsequent sections.

- Meet prerequisites
- Slack account and joining relevant channels
- GitHub account
- Complete the Google Form to let me know your Slack username and your GitHub username
- Access to the O'Reilly textbooks using the Berkeley library or paid Safari account

## Prerequisites

In past semesters, we have had instances of students not meeting prerequisites and taking the course anyway. We have spent class time and office hour time essentially bringing students up to speed with prerequisite materials that they should have already known.

This semester we are committed to avoiding spending class time or office hour time on prerequisite materials to be fair to students who do meet the prerequisites.

### For MIDS program entry:

- A working knowledge of fundamental Computer Science concepts, including:
  - Data Structures
  - Algorithms
  - Analysis of Algorithms
  - Linear Algebra
  - Object-Oriented Programming Skills in Python or C/C++ or Java
- Students should have been given access to bridge courses or other resources to complete deficiencies in these areas with specific instructions of which ones are needed for W205

## For W205:

- Python
  - Students should have taken W200 Python Fundamentals for Data Science  
OR  
if exempted from taking W200, students should have been given access to the Python bridge course with instructions to complete it prior to taking W205
- Linux command line
- GitHub (including command line using the git utility)
- Database Management (SQL)

### Slack Account and Joining Relevant Channels

- **slack.com**
- Please use slack to communicate as it's the easiest way to reach me. Email will take longer. I try to check slack once a day, but cannot guarantee it.
- Please join the **ucbyschool** team using your **@school.berkeley.edu** email address, not your @berkeley.edu address, nor any other, as it will not be approved.
- Please join the following channels:
  - **w205**
    - all sections for all instructors for w205
  - **w205-crook-summer2018**
    - all section for Kevin Crook for Summer 2018
    - my primary means of communication with my students
    - please check daily if possible

### GitHub Account and GitHub Classroom

- **github.com**
- You will be using GitHub Classroom to submit all of your assignments this semester.
- You will need a GitHub account, which I will add to the permission group for my classroom.
- There is not a separate account for GitHub classroom.
- Assignment link – for security reasons, these links will be placed in slack. You will need to accept these links in a timely manner before they expire.
- While it is not needed for this course, GitHub has an academic discount which essentially allows you to create private repos for free.

## Please Enter Your Slack and GitHub Usernames into my Google Form

For security reasons, I'll post a link in Slack to a Google Form where you can enter your Slack and GitHub usernames.

## O'Reilly Textbooks – Berkeley Library or Safari

This course will use part of several O'Reilly Textbooks.

The Berkeley has most of these books available free:

**<http://www.lib.berkeley.edu/>**

Safari is a subscription service for online books. You may want to check with your employer to see if they provide a free subscription. They have a free trial period, but it's only a couple of weeks. They have a paid service which runs around \$45 a month.

**<https://www.safaribooksonline.com/>**

## Amazon Web Services (AWS)

Later in the semester we will be using Amazon Web Services. It's not needed for the first class. Sign up information will be posted in slack.

## Google Cloud Platform (GCP)

Later in the semester we will be using the Google Cloud Platform (GCP). It's not needed for the first class. Sign up information will be posted in slack.

## Format of the Asynchronous and Synchronous Sessions

### Reading Assignments

- Mandatory
- Must be able to read, study, and understand these on your own

### Asynchronous

- 1.5 hours of your 3 credit hours each week
- Mandatory
- Videos

- Must be able to watch, study, and understand these on your own
- Will not be directly repeated in Synchronous

### **Synchronous**

- 1.5 hours of your 3 credit hours each week
- Mandatory
- Will not be a direct repeat of the Asynchronous which you must complete on your own
- Typical activities:
  - **Cold Calling Questions** – students will be divided into groups and we will go round robin. Questions may come from the relevant prerequisites, the asynchronous material, the assigned readings, and previous synchronous sessions.
  - **Break Out Sessions** – students will work interactively in groups. In the W205 curriculum this is usually related to the assignments. It is mandatory to participate and contribute to your group.
  - **Instructor Led Activities** – instructor will lead students through an activity. It is mandatory to follow along and keep up. If you have an issue, please let the instructor know before we go on to the next step. Please do not fall several steps behind and then expect the class to stop for several minutes for you to catch up.

<b>Grading Rubrics</b>
------------------------

### **Professional Conduct**

In the past, the vast majority of students have conducted themselves professionally, however, we have had some issues with students in the following areas:

- Taking personal responsibility for meeting prerequisites
- Taking personal responsibility for scheduling and managing the coursework
- Taking personal responsibility for coming to class prepared
- Treating others in a respectful manner
- Accepting things you cannot change, working respectfully within the system to change things you can, and choosing your battles wisely

### **Attendance including Active & Enthusiastic Participation**

- High quality connectivity: high speed wired internet, wired headset, quiet environment, noise cancellation technology, muted when not speaking, etc.

- Prepares for class by watching and studying the asynchronous materials and readings.
- Entering the online classroom a few minutes early to ensure you are ready to go when class starts. Remains engaged until the end of class.
- Video is on and active during class
- Answers “cold calling” questions frequently
- Actively participates and frequently takes a leadership role in break outs

### **Making Up Missed Classes**

- Whether excused or unexcused absence, missed classes must be made up in order to earn a grade of A- or higher for the semester
- To make up a class, students must watch the synchronous video, work any in-class exercises on their own, write up an executive summary, and submit the executive summary to the instructor using a direct message on slack before the next class meeting start time
- Executive Summary should be a PDF file, 1 to 1.5 pages in length, professionally formatted, 12 point Calibri font, 1 inch margins, 1.08 line spacing

### **Assignments**

- 12 Assignments
- 10 of 12 will count
  - 2 lowest will be dropped
  - However, to be dropped, the assignment must be turned in with a reasonable attempt
- Each assignment is graded in whole numbers between 0 and 10 points
  - Subjective = 2 of 10 points
    - Most students will score a 1 on subjective on most assignments
    - A score of 2 on subjective indicates exceptional performance and typically only 3 to 5 students in a section will make this
  - Objective = 8 of 10 points
- Late Submissions
  - Assignments 1 through 11
    - When grading, a zero will be put in as a “placeholder” in ISVC for missing submissions and changed when they are submitted
    - 1 of 10 points penalty for 1 second late
    - Additional 1 point penalty for each additional 24 hours late
    - After 10 days late, the assignment will be worth zero points

- Assignment 12
  - Due to end of semester time constraints, late submissions for assignment 12 will have the following late penalties
    - 5 of 10 points for 1 second late
    - Zero for the assignment if 24 hours late

Semester Letter Grade	Rubrics
<b>A +</b>	<ul style="list-style-type: none"> <li>• Professional Conduct</li> <li>• Attendance (or make ups) in all 14 synchronous sessions</li> <li>• Frequently answers “cold calling” questions</li> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 9.5000 without rounding or higher</li> <li>• Completed the course (non-emergency incompletes cannot receive higher than B+)</li> </ul>
<b>A</b>	<ul style="list-style-type: none"> <li>• Professional Conduct</li> <li>• Attendance (or make ups) in all 14 synchronous sessions</li> <li>• Frequently answers “cold calling” questions</li> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 8.5000 without rounding or higher</li> <li>• Completed the course (non-emergency incompletes cannot receive higher than B+)</li> </ul>
<b>A -</b>	<ul style="list-style-type: none"> <li>• Professional Conduct</li> <li>• Attendance (or make ups) in all 14 synchronous sessions</li> <li>• Frequently answers “cold calling” questions</li> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 8.0000 without rounding or higher</li> <li>• Completed the course (non-emergency incompletes cannot receive higher than B+)</li> </ul>
<b>B +</b>	<ul style="list-style-type: none"> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 7.5000 without rounding or higher</li> </ul>
<b>B</b>	<ul style="list-style-type: none"> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 7.0000 without rounding or higher</li> </ul>
<b>B -</b>	<ul style="list-style-type: none"> <li>• All 12 assignments submitted with a reasonable attempt</li> <li>• The average of the highest 10 of the 12 assignments should be 6.5000 without rounding or higher</li> </ul>

<b>Below B -</b>	<ul style="list-style-type: none"> <li>• Continue in 0.5000 increments</li> <li>• Assignments not completed will be scored as zero</li> <li>• Average of highest 10 assignments</li> </ul>
------------------	--

<b>Due Dates for Assignments</b>
----------------------------------

Assignment Number	Due Date
00	Does not count. We will work on it in class. Instructor will mock grade 00 so students can get used to how GitHub classroom works
01	Monday, May 21 <sup>st</sup> , 6 am Berkeley Time (instructor will be flexible with due date and not count late penalties up to 1 week late)
02	Monday, May 28 <sup>th</sup> , 6 am Berkeley Time (instructor will be flexible with due date and not count late penalties up to 1 week late)
03	Monday, June 4 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
04	Monday, June 11 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
05	Monday, June 18 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
06	Monday, June 25 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
07	Monday, July 2 <sup>nd</sup> , 6 am Berkeley Time (late penalties apply)
08	Monday, July 9 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
09	Monday, July 16 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
10	Monday, July 23 <sup>rd</sup> , 6 am Berkeley Time (late penalties apply)
11	Monday, July 30 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)
12	Monday, August 6 <sup>th</sup> , 6 am Berkeley Time (late penalties apply)