W201 Assignments & Live Session Policies

Research Design and Applications for Data and Analysis (RDADA)

Spring 2019

Corey Jackson
coreybjackson@berkeley.edu

This document and the **syllabus** are your primary documents for the course. This document is for Corey Jackson's live sessions. If you are another section, please ask your instructor for their section-specific document.

Major revisions are pending until the first class. Minor revisions may be made up until a week before a scheduled class. **Check this doc every week to stay current.**

Please get started. You are required to complete the onboarding before we meet Week 1. (Don't worry--the Week 1 onboarding and discussion question will not change.)

Table of Contents

Assignments Overview

The assignments below will help

1. reinforce your comprehension of the asynchronous content,
2. refine your written and spoken communication skills and your professional style,
3. stimulate and motivate you to develop your own data science concept proposals (i.e. research designs),
4. and build a network of collaboration among your peers.

Ultimately we hope that you learn how to forge data science concepts in a way that will impress partners and add value to organizations.

Graduate school is not like undergraduate education, and the UC Berkeley Masters of Information and Data Science (MIDS) is not like most graduate programs. Expect to benefit a great deal from the impressive work and life experiences of your peers; do not hesitate to bring your own perspectives into the classroom. Rather than simply prepare your work for your instructor and peers, think bigger. Take responsibility and authority for achieving your own educational goals. In most cases this means you should chose a real stakeholder as the audience for your writing.

Think of yourself as the chef and your peers and instructor as guests at your table--eager to experience your creative vision. The assignment prompts below are guidelines, not checklists; you will not be penalized for using your judgement about how best to achieve a desirable result. Especially at the beginning, let your imagination roam freely and do not worry terribly about being wrong. We will generate a lot of ideas together, and we will help each other identify the viable ones.

The expectations for this course are high. That said, I strive for a friendly and enjoyable classroom environment. You all are very accomplished and I look forward to discussions about how the course concepts reflect your professional experiences. I anticipate we will have fun.

Course Mechanics and Onboarding

Before the first class complete a few administrative tasks. For security, copyright, and privacy purposes we limit access to course material generated in Google drive to your @ischool.berkeley.edu or @berkeley.edu account. We recommend you also adopt more secure permissions when sharing work among students. To ensure that your Berkeley Google account is set up, and to ensure you are ready for week 1 please do the following:

Completion of the on-boarding before we meet week 1 is part of your participation grade.

- Register your CalNet ID.
- Create your bConnected Google account.
    - Because it is a Berkeley organization account, docs created with your bConnected login offer increased security permissions which protect your and your peer's privacy and identity.
    - If you use multiple Google accounts, add your CalNet ID account to your Google app account switcher, OR
    - Log into your bdrive.berkeley.edu account from a private or incognito window.
- Create a portfolio. You will submit assignments and receive instructor feedback through this document. **Do not submit assignments via the i School Virtual Campus (ISVC)**.
    - Copy the RDADA portfolio template to your bDrive.
    - Open the view only copy from your bConnected account and select *File>Make a copy…*
        - The template is viewable by all @berkeley.edu accounts. If you see a page suggesting that you request permission to view the document, you are not logged in to your bConnected account. Follow Step 2 and try again.
    - Title your version of the portfolio as "W201 Spring2019 | Class Day of week, Time | Firstname Surname Portfolio"
        - (ex) "W201 Spring 19 Portfolio | Wednesday 4pm | Jane Doe"
    - Share with your instructor and the graduate reader directly as commenters coreybjackson@berkeley.edu and kborojerdi@berkeley.edu

    - Activate link sharing and set it to “Anyone at UC Berkeley with the link **can comment**”.
        - You can of course opt for even more security and manage collaborators individually; however, be sure to always add your instructor and grader (if applicable).
    - More on portfolio can be found later in this document
- Send your instructor (coreybjackson@berkeley.edu) a “Hello World!” email from your bmail.berkeley.edu account (they will not respond, but it is another proof that your account is set up correctly).
- Update your public (and present tense!) bio here. Log in on the top, then click update profile.
- Update your ISVC bio here.
- Take the **section specific** time zone survey to help coordinate office hours and group work.
    - Tues 400pm class
    - Tues 630pm class
- Activate your professional Zoom account (included with your Cal tuition). The email you receive may include language like “Zoom account invitation.”
- Activate your Slack account. See Class Correspondence (Slack) for more info.

Miscellaneous reminders for live session #1

- Use the library VPN to ensure access to the linked readings. Select the Library_VPN group.
- Breakout notes: You may find it useful to take notes during the breakout discussions in your portfolio for reference during our larger group discussion.
- Recordings: the main session is recorded. Breakout groups and open office hours after class are not recorded. More info will be provided in class.
- To join class: Click the meeting (calendar) icon on the left hand side of the I School Virtual Campus. There you’ll find a list for your upcoming meetings. You will see a "join class now" button about 15 mins before class.

Diversity and Inclusion

In an ideal world, science would be objective. However, much of science is subjective and is historically built on a small subset of privileged voices. In this class, we will make an effort to read papers from a diverse group of scientists, but limits still exist on this diversity. I acknowledge that it is possible that there may be both overt and covert biases in the material due to the lens with which it was written, even though the material is primarily of a scientific nature. Integrating a diverse set of experiences is important for a more comprehensive understanding of science. I would like to discuss issues of diversity in data science as part of the course from time to time.

Please contact me (in person or electronically) or submit anonymous feedback if you have any suggestions to improve the quality of the course materials or the discussions in live session.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities (including race, gender, class, sexuality, religion, ability, etc.) To help accomplish this:

If you have a name and/or set of pronouns that differ from those that appear in your official Berkeley records, please let me know. If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. I want to be a resource for you. Remember that you can also submit anonymous feedback (which, if necessary to address your concern, will lead me to make a general announcement to the class). If you prefer to speak with someone outside of the course, the following individuals are excellent resources: Director of Student Affairs, Associate Director of Student Affairs, Academic Director. You can find their contact info here.

I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to me about it. (Again, anonymous feedback is always an option. If you'd like to remain anonymous, I recommend you contact one of the resources above, make it clear that you'd like to maintain anonymity, and then communicate that you'd like your concern shared with me.)

As a participant in course discussions, you should also strive to honor the diversity of your classmates.

Class Correspondence (Slack)

Slack is our primary method of communication and should be used to contact the instructor. The instructor will do their best to respond within 24-48 hours, although during the weekend or holidays responses may take longer. Sign up for Slack and then join the appropriate channels listed below.

Log on to ucbischool.slack.com and create a Slack account. Make sure you use your Berkeley email @ischool.berkeley.edu. I strongly encourage you to use a first & last name in your user name so that you don't get messages intended for another student (e.g. mikerivera or mike.rivera is better than mike because Slack may autocomplete the username and you may send a message to an unintended recipient). If you have any issues, contact Mike Rivera, one of the iSchool Slack admins michaelrivera@ischool.berkeley.edu.

Section Specific Channel (Qs about assignments & logistics)

This channel is for specific questions, comments and announcement related to your section. Instead of sending a direct message to the instructor, I encourage you to post to the channel; oftentimes your peers can answer questions faster than the instructor can and your peers can benefit from shared insight. You are always welcome to send instructor a direct message. If you would like to send direct message to the grader, please CC the instructor so that everyone is on the same page.

Join the appropriate section channel.

Tuesday 4pm: w201-sp19-t400-jackson
Tuesday 630pm: w201-sp19-t630-jackson

Cross Section Discussion Channel (Post course-related articles & Qs about async material)

Join #discuss-jackson-sp19

This cross-section channel (i.e. all of Mike's sections) should be used to continue the discussion outside of live session. It is a place to post interesting news articles and readings, and to ask questions about the async.

Office Hours

You will find the most updated office hours info in a pinned Slack post. I'm also available after class, via Slack and by appointment.

Overview of Assignments

W201 has four types of assignments:

- Individual Assignments
- Group Presentations
- Final Group Project
- Discussion Questions (not graded)

Individual and group assignment will be graded. Discussion questions will help you digest the material and will prepare you for live session; however, they are not graded.

The breakdown of your final grade is:

- 30% individual assignments (3)
- 30% group projects (3)
- 20% final project
- 20% class engagement (This includes attendance, preparation for live session, participation in discussion, and peer review).
    - Completion of the on-boarding is also part of your participation grade.

Students can find their assignment grades on the i School Virtual Campus (ISVC). Please note that your instructor may not post final participation grades on the ISVC; if you'd like to discuss your participation in live session, please reach out to the instructor.

You will receive a letter grade at the end of the term (official grading scale). You instructor is solely responsible for assigning final grades.

Late Policy

Late assignments are accepted only under extreme circumstances. A significant late penalty will be assessed but some points > no points.

Take Risks, Try New Approaches

Make the assignments in this class as useful to you as possible. If you are grappling with a problem at work, you're welcome to address that in an assignment. If you want to pursue a project that has nothing to do with your professional domain--that's also great.

Want to expand upon a previous assignment in the latter part of the semester? Go for it. Want to get creative with the assignment and engage your "beginner brain" (week 6 reference)? Send it! Ultimately, I encourage you to take risks and try new approaches.

Similarly, don't be afraid to contribute to class discussion. We come to class with different levels of comfort speaking in public; this class will give all of us more practice. If you read/view the async material and come prepared, you will be equipped to participate. You will also realize that many questions have a range of possible answers. Even if you get the answer "wrong," think of it as a learning opportunity. I will politely redirect us toward a "more correct" answer.

Bottom line: Take risks. It's ok to fail. Learn. Have fun!

Class Engagement

In order to ensure a productive learning environment, students should prepare for and attend every live session, contribute thoughtfully to class discussion, and provide peer feedback on their peers' presentations.

Attendance and Preparation

You are expected to attend and actively participate in each live session. Come prepared. Your peers and I expect that you **read and view all readings and async lectures prior** to live session. Please let your instructor know in advance if you need to miss class for a personal or work emergency.

Recommended approach to async material

I encourage you check the weekly async info document before you start to see if the instructor has provide any additional clarification on the async material.

I recommend you watch the videos first, then read the articles; the videos provide important context for the readings. I recommend the following reading strategy. Quickly read the abstract/intro, then skim the headings, tables and charts, and then read the conclusion. After you skim the reading, give the article a *full read*. This will ensure you know the punchline before you start.

Discussion Questions & Bidirectional Learning Tool (BLT) Questions

Discussion question are an opportunity for you to synthesize the async material and apply it to your work experience and your future career goals. I highly recommend you respond to the discussion questions in your portfolio. While the instructor will not grade them directly or offer feedback, these questions will help you organize your thoughts and will help you prepare for live sessions.

Note that the discussion questions are not intended to narrow down the focus of the videos and readings for each week, but rather they provide a common focal point for the synchronous (live) sessions. As this is a graduate seminar, you are expected to come to class with your own synthesized views of the weekly materials in addition to the discussion questions that are suggested in this document. Therefore, whether or not there are discussion questions assigned for a particular week, you should come prepared to exchange your thoughts and impressions of the course materials in class.

You must come ready to participate fully in live session. The discussion questions and the BLT (Bidirectional Learning Tool) questions in the async material will help. These may serve as a starting point for your initial breakout group discussion.

Peer Review

Peer review is an essential feature of data science research. To that end, you will provide feedback on your peers' presentations. Your feedback should be critical, but constructive.

- Peer feedback will occur during live session. After each presentation there will be a brief question & answer session and an opportunity to provide concrete peer feedback via a Google form. Presenters will also self-evaluate.
    - There are two goals of peer feedback during live session. 1) Note what works in your peers' presentations and note what could be improved. 2) Reflect upon the the advice you give and compare it to your presentation strategy. Think, “what can I learn from my peers?”
    - **I encourage you to focus on the substantive aspects of the presentation in the Q & A session**. The goal is to have a meaningful back and forth conversation about the substantive topic. You are encouraged to include comments about presentation style in the written feedback.
    - In order to mimic the team environment outside of this class, you will know who provides you with specific feedback (i.e. this is not anonymous). Be honest, but professional. I know eventually you might become friends and don’t want speak critically about your classmates’ work; however, think of constructive feedback as a positive thing. It is better to hear how you can improve in-class, when the stakes are more reasonable, than to hear this feedback after a presentation for your job, where the stakes are higher.
    - For group presentations, presenter evals will be distributed to all members in one’s group. This will give you all an opportunity to reflect upon what worked and what didn't. Think of this as the postmortem analysis discussed in the async.

Individual & Group Assignments

There is a heavy emphasis on presentations in this class. A major takeaway of this course is that communication skills are an integral component of effective data science. Whether you present to a large audience, a board of executives, or simply present an idea to a colleague over coffee, you must be able to effectively convey your findings.

You will have opportunities to expand on particular course concepts via several individual assignments. That said, all research–whether in industry, government, or nonprofit–is collaborative. Group work is a critical part of any data science career. To that end, you will be responsible for several group presentations during the live sessions. For presentations, your grade is a function of the common group product and your individual contribution.

Portfolio

In order to submit your assignments, you will first have to create a portfolio (see Course Mechanics and Onboarding). The goal of this document is to have a central place to submit assignments and receive instructor and peer feedback.

Treat your portfolio as a living and breathing document. Feel free to edit your headings, language, ordering, exhibits, or style of presentation whenever and however you like, so long as the 14 week structure is clear. If you don't know what a portfolio is, review the onboarding section.

- Changes to assignments once submitted:
    - You can make a few minor, copyedit changes after the deadline; however, if you want to make a more substantive change, please do the following. 1) Keep the original link to your presentation in your portfolio. 2) Create a copy of the slide deck and make the changes to that updated copy. 3) Include an updated link under the original link and type "updated presentation" and include a sentence or two that indicate what changed.
        - This is not an excuse to work on the presentation after the deadline; however, I want to allow you to make adjustments if you realize a change is necessary after you've formally submitted.
- **The portfolio will house all your work. Do not submit assignments via the i School Virtual Campus (ISVC).**

Due Dates/Times

ALL assignments are due by 11:59pm PST the night before the live session of the relevant week.

- Note that group assignments also require a team process agreement. This is due earlier than the "11:59pm before live session" deadline.
- More specific dates and information can be found later in this document.

Assignment Deliverables

Individual Assignments

- Link to **annotated** google slide deck in portfolio. The slide deck template provides tips on how to create a clear and persuasive presentation. These instructions also provide useful information about the procedures and format for presentations in this class.

    - Slides must be housed in your Berkeley google drive and I must have access to the link you post.
    - The level of detail you include in the annotated slides is up to you. You do not need to include a verbatim script; however, your notes need to be intelligible and I must be able to understand the crux of your presentation by looking at your annotated slidedeck. This is important, because I will use both the video recording and the annotated slidedeck to provide feedback.

- Link to video recording in portfolio.
    - The video will help ensure you practice both the content and the timing of the presentation and will help the instructor provide feedback. In live session you will present in small groups and the instructor will see some presentations live. The videos will help the instructor provide feedback on all presentations (especially those not viewed during live session).
    - **You must use Zoom.** The recording must show your slides and include an audio voiceover. (Whether or not your face shows is up to you.) Use the cloud recording feature, go to Recordings in your Zoom profile, click Share, and copy and paste the link into your portfolio. For more info, see this link or contact 2U.
        - Make sure you have activated your professional Zoom account (included with your Cal tuition) which you need to record on the cloud. The email you receive may include language like “Zoom account invitation.”
        - Check your “User Type” in your Zoom profile. It should say something other than Basic(free). (Contact 2U student support if you have issues. If you can’t figure it out before the first assignment, simply record locally, upload the recording to Google Drive and then include that link to your portfolio.)
- Link to slidedeck in Slack. I will pin a “Post Week X Final Slides Here” post on Slack and you will comment on that post with a link to your slides. You can upload your annotated slides or a PDF without annotations. (Remember, the link you submit to your portfolio must be an annotated Google slide deck.)
    - The slides you post in Slack must be housed in your Berkeley google drive and you must grant access to your peers (i.e. Anyone at UC Berkeley with the link **can view**).

Group assignments

- Each team member should read and understand the submission guidelines
- You are required to submit the deliverables above plus…

    - Statement of Individual contribution in portfolio. This will describe the extent to which you contributed to the group product.
        - Feel free to use the “#slide=” argument to clearly specify which slides you created.
    - The Team Process Agreement Form. Instructions can be found in the document.

Clarification on group assignment deliverables

- Video recording: For group assignments, I recommend you meet at least once with you group to do a top-to-bottom practice run. That said, *each student is only responsible to record their portion* of the presentation and post a link in their portfolio.
- Slack post: Only one student from each group needs to post a link to the slidedeck to Slack; however, each student must post a link to the annotated google slide deck in their portfolio.
- Each group member will present during live session.

Define your audience

- Think carefully about your target audience before you begin. Speak to a prospective professional or consumer audience; avoid speaking to the instructor or the class.
- You must explicitly articulate the audience in the first slide of your presentation (e.g. “Our intended audience for this presentation is X”). An intended audience may be Amazon Web Services, health care insurance companies, nurse's unions, a city council, individual consumers of music, digital content providers, etc. Presentations (and writing) are more effective if you speak to *someone*. If you explicitly articulate of your audience you will narrow down the scope of your project and will make sure you keep focused.

Timing and use of scripts

- Presentations allow students to explore course concepts in greater detail and help improve one's communication skills. In general, it is best to use the most precise and concise language possible; these presentations provide practice.
- Some students use a very detailed script and strictly adhere to it. Others sketch a complete, but less thorough script and emphasize a conversational tone. There is a clear balance between 1) a tone that is overly rehearsed and sounds robotic and 2) an overly informal presentation that meanders and goes over the time limit.
- Ultimately you are free to choose what works best with your individual style. That said, I will strictly enforce the time limits described in each assignment. Detailed annotation and practice will help ensure you say within the time limit.

- To ensure a smooth transition between speakers in group presentations please 1) choose one person to share their screen and advance the slides, and 2) have each group member test their audio *before* you begin the presentation.
- To stay within the time limit I recommend you use your own timer (e.g. cell phone or Google slides timer. To access the timer in Google slides: view > present > settings > open speaker notes).

General Presentation Tips

- **While it is not required, I strongly urge students to meet with the instructor outside of class to discuss both individual and group assignments.** Our meeting give students an opportunity to chat informally about their ideas. If you'd like to meet to discuss a group assignment, each group member is not required to attend. In the past, a couple of teammates met with me and then reported back to the group. A meeting with the instructor and a report back to the group can be part of a student's responsibility for the presentation (ie. Team Process Agreement).
- If you're not actively speaking during presentations please mute your audio. This will eliminate background noise.
- You will engage the audience the most if you look into the camera when presenting. I recommend you view your presentation notes on your computer, or hold your notes just below the camera.
- During live session you will use the share screen feature in Zoom. I recommend you practice the share screen functionality before class. Contact 2U if you have any questions.
- You may use any source material you would be comfortable using in the professional setting. For example, you may use news articles, but you may hesitate to use wikipedia pages.
- The best presentations have a recognizable conclusion, which could be as simple as "thanks for listening!"
- If you use Google slides presentation mode, share the window with the presentation (not your screen) so that the audience can't see your notes. You can also "hide" your notes on a second screen.
- Group specific notes:
    - For group presentations there will be four student groups. Two groups will present in breakout and I will be view the entire presentations live. Two groups will present in the main discussion area and will be recorded in Zoom. I will view the recording and will provide feedback after class.

- Pace your workload: I encourage you to decide on the division of labor well before the Team Process Agreement Form is formally due.
- Communication method: I recommend you create either a Slack channel or a group direct message. If you create a Slack channel and want to make it private, please contact Mike Rivera, one of the iSchool Slack admins @mikerivera
- Decide on a topic early: Students say this is one of the more challenging aspects of the assignment. Some students pin a google doc in their Slack channel and then ask each student to post ideas by a given date, and then they either vote on a topic via Slack or vote and then meet via Zoom to make the final decision.

Assignment feedback

Generally I will evaluate students based on:

- argument/logic/strategy
- writing/narrative/storytelling
- format/presentation

Students can expect to receive several tiers of feedback.

- The instructor will provide individual student-level / group-level feedback via the portfolio.
- The instructor will discuss class-level feedback in live session. They will discuss what the students mastered and what most students need to work on.
- Students will receive feedback from their peers.
- Students will self-evaluate.
- Students who want additional feedback from the instructor can set up an appointment, come to office hours, or chat with the instructor after live session.

Key Dates

Below you can the key dates for this course.

Calendar week that begins on Monday	Corresponding Syllabus & Assignments Week	More Info
1/7	1	
1/14	2	

1/21	3	**Individual Assignment #1** Tues. classes: Present during live session (1/22), Deliverables due 11:59pm PST the night before live session.
1/28	4	
2/4	5	
2/11	6	
2/18	7	
2/25	8	**Group Project #1** Tues. classes: Present during live session (2/27), Deliverables due 11:59pm PST the night before live session.
3/4	9	**Individual Assignment #2** Tues. classes: Present during live session (3/5), Deliverables due 11:59pm PST the night before live session.
3/11	10	
3/18	11	**Group Project #2** Tues classes: Present during live session (3/19), Deliverables due 11:59pm PST the night before live session.
3/25	Spring Break	No Jackson Live Session
4/1	12	**Individual Assignment #3** Wed classes: Present during live session (4/2), Deliverables due 11:59pm PST the night before live session.
4/8	13	**Group Project #3** Tues. classes: Present during live session (4/9), Deliverables due 11:59pm PST the night before live session.
4/15	14	**Final Group Project** Tues. classes: Present during live session (4/16), Deliverables due 11:59pm PST the night before live session.

Note: Please note that the syllabus week across courses may not sync up entirely Do not assume that because we are on week x, your other instructors are also on week x (and visa versa). Spring 2019.

Assignment List and Details

Below you can find specific information about each assignment. As you prepare your assignments and view your peers' presentations, think about topics and ideas that interest you. This will help inform your choice of topic for future assignments.

I. Introduction and Overview

Week 1: Prospective Biography (Discussion Question)

Please review both the syllabus and assignment documents thoroughly. Complete the tasks in the Course Mechanics and Onboarding section before we meet week 1. This is part of your class engagement (participation) grade.

To prepare for the live session, use the Week 1 discussion section of your portfolio and consider the following scenario:

Think of yourself three years from now in the data science world. What are you doing? What does that entail? Feel free to be a bit inventive, so long as it helps you imagine and reflect. Try to be specific; for instance, you might name a real or imagined firm or organization, a title for yourself or team description, or a project you see yourself engaged in.

Remember, you should have set up a portfolio by now. If you don't know what a portfolio is, you'll need to re-read the preceding portions of *this* document.

II. Decision Making

Week 2: Tactical and Strategic Decisions (Discussion Question)

Identify an organization that could be a client for a future data science project; it can be your current organization or a new one you want to learn about. In this organization, identify two decisions—one tactical, the other strategic—where data science is or can be used to significant effect. Think about the data and the organizational requirements of applying data science in both of those settings (tactical and strategic). What might be some significant roadblocks to success?

Hint: It may help if you decide on a criterion or working definition for “tactical” and “strategic” decisions.

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes.

Week 3: Impact of Data Science and Policy Implications (Individual Assignment #1)

Tues. classes: Present during live session (1/22)
Deliverables due 11:59pm PST the night before live session

Objective: To articulate, in concrete terms, how data science could improve a domain of your choice.

Do a preliminary brainstorm about topics, domains, or industries that interest you. This may be a current (or upcoming) public or corporate issue where you believe the application of data science tools and methods could make a substantive difference--in practice and/or in how debates and decisions around the issue develop. For this assignment, you will use a “policy brief” or “business brief” style of communication. Think of this as a brief you would present to a decision maker. In other words, don't get bogged down in the technical aspects of your idea; articulate the data science gap and your solution to fill it.

- Find a recent magazine, newspaper, or academic article relevant to your domain of choice. Try to pick one that discusses how data and/or analysis could fill an information gap and make an impact for organizations in that domain. If the article is not about data science per se, infer what the data science obstacle and opportunity is. Feel free to use any sources you would use in your professional setting.
- You should use the one article as the focal point of your presentation. You are not required to--but are welcome to--bring in ideas from additional articles, podcasts, etc. (cite when necessary).
- Start with the slide deck template and guide; these instructions provide useful information about the procedures and format for presentations in this class. In this

individual presentation, skip the group-specific instructions in the template. Provide a bullet-point style brief of:

- (Slide 1) Always have a title slide! How will you pique our interest? Have a hook.
- (Slide 2) Where/when/why the article was written. Include the original author's 1) most important and 2) most problematic claims. This is an opportunity to articulate any grievances you have about the article (i.e. What it got right and what you find problematic).
- (Slide 3) What is the gap to be filled by data science. Is it a data gap, an analysis gap, or both?
- (Slide 4) How the gap(s) might be filled.
- (Slide 5) Who would be motivated to 1) invest in the data science and/or 2) use the data science findings. Why would they invest in the science or use the findings? You may identify multiple stakeholders but focus on one.
- (Slide 6) What are the most important opportunities and obstacles to 1) acquire the data and make it available and/or 2) analyze the data and communicate the findings.
- (Slide 7) Include a final bibliography slide with full references and link(s) in any citation style. Do not include full references in your main slides.
- You are welcome to have more slides if you want, but be mindful of the time limit.

You have up to 8 minutes to present in-class in small breakout groups. The time limit will be strictly enforced. An objective of this class is to help you improve your communication skills. This includes using the most precise and concise language possible. Practice! You will be cut off and will explicitly lose points if you go over 8 minutes.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.

Week 4: Bias (Discussion Question)

Prepare a scene that describes bias in action. Think of an organization where decision bias is visible and contributes significantly to what people do or where decision bias affects an outcome. You may use your Week 2 discussion question as a starting point or you can use a new client that interests you. Be prepared to explain the story and demonstrate the bias. Be sure to:

- Give the scene a substantive title.
- Identify the most important decision makers (not necessarily a boss or CEO). Describe the scene or organizational context in which they are making the decision (e.g. resources, stressors, constraints, etc.).
- Identify and specify the bias present in the situation. Think of the Kahneman text. You may also describe a new kind of bias you see.
- Demonstrate the bias in action (provide an illustrative example).
- Imagine yourself in a professional context giving advice or designing products or services. Propose a means to compensate for or counteract the bias.

Try to empathize to the decision-maker what it would be like to resist or alleviate the bias. What are the benefits of an attempt to counteract this bias? Hint: To claim you will fully prevent the bias may be unrealistic; instead, think about how to “move toward the truth.”

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes.

III. Research Design

Week 5: Kuhn’s Data Science Revolution? (Discussion Question)

Consider the emergence of data science and think of its impact on a field or industry that you are in or would like to be involved in.

Consider Kuhn’s framework of analysis and the specific components of his account of science. Discuss how that field or industry does and does not conform to Kuhn’s typology. Explain and analyze how that field is in transition/transformation as a result of data science. Does it support Kuhn’s ideas of scientific change? Is it a “revolution” in Kuhn’s terms? How useful do you find the idea of “revolution”?

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes.

Week 6: Research Design (Discussion Question)

Pick a contemporary business or trade magazine (e.g., Business Week, Harvard Business Review, Wall Street Journal, The Economist, etc.). Find an article that addresses an important question, but falls short. In other words, find an article that poorly frames or articulates an important question.

Explain why the question is poorly framed.

- Consider why the framers made the mistakes they did.
- How would you reframe the question as a research question for better results?
- Explain why that reframe would enable better, more actionable research outcomes.
- Reframe the question clearly and show the before-and-after versions of the question.

Hints:

- Think about the difference between a research question and how people usually ask questions when they are lazy about research design (Use Creswell as your guide).
- You may have sub-questions, but a research question is typically a single sentence. You may include additional sentences to provide context, but you should focus on how to craft a single, clear research question sentence.
- Research is an iterative process. You will rarely formulate the research question correctly the first time. Overall, you will get more insight out of a research question that is almost right than one that is terrible. It is valuable to identify the “right” research question, but it is also valuable to find smart people who missed the mark.

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes.

Week 7: Fallacies (Discussion Question)

Find an article that makes a meaningful fallacy of argumentation. Describe the error of logic and present your argument for why it is fallacious. Don’t worry if you can’t precisely label the fallacy. Rather than fix it, tell us why you think the author made the suspect argument and describe why the fallacy is sticky and difficult to remedy. Think of answers to the following questions:

- What is the fallacy?
- Why is the fallacy so sticky?
- Why is the fallacy difficult to fix?

- In your experience, how prevalent is the fallacy in business decision making? Share an example.
- How impactful is the fallacy on a scale of 1 (low) to 10 (high)? Why?

Extra hint: Don't look at too many articles. Almost every article in an opinion (or news) section of a major publication will contain some fallacious reasoning.

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes.

Week 8: Case Studies--Research Design (Group Project #1)

Tues. classes: Present during live session (2/26)
Deliverables due 11:59pm PST the night before live session

This assignment will take some real work. It is the most substantial project you will do this semester apart from the final assignment.

Objective: To apply research design concepts to identify a business problem and research question, and to describe how data science can answer that question.

Develop a case study of how data science could be used by an organization to inform an important decision. Go into more detail than the brief case studies in the async videos. Write in the present tense and utilize **at least three** concepts we've discussed over the past few weeks. For instance, try to name a logical fallacy or a cognitive bias in the decision-making process; weave them into the storyline so that they flow nicely and are not blatantly listed.

Each case study should be partly descriptive (of a real domain and organization), and partly hypothetical. Draw on published news stories and analysis to establish the situation, organizational goals, and even the decision. (You are welcome to draw from experiences in your own organization or one you are familiar with, but you cannot rely entirely on internal knowledge for this presentation). You can suggest a decision the domain or org *should* consider even if they aren't. It is your job to frame the decision well, to frame the research question(s) appropriately, and to set up the broader research design.

Do not simply summarize previous research designs; you need to come up with a creative design. This is **not** a literature review; build on the organization's existing situation, but question some of their assumptions. While your case does not need to involve big data in an extreme sense, try to at least consider one of the 3Vs of big data (volume, velocity, or variety, [veracity]).

Make sure the research design is crafted to influence one organizational decision. Good research design seeks to answer specific questions. Therefore, to create an algorithm that gives recommendations or answers to users is fine, but the research design **must** try to answer a question faced by a meaningful decision maker (i.e. a client or customer).

Make sure you address each of the following points. You may make one slide per bullet point.

- The bulk of your work should:
    - Describe the domain, organization, and stakeholders.
    - Identify the core business problem(s) and research question(s).
    - Identify the decision that must be made. Who will make the decision? The decision maker may or may not be from the c-suite.
    - Imagine and describe a creative data collection and analysis effort to address the research question.
    - Describe how the decision is made (i.e. provide context on the decision making process.) How would decision makers utilize or implement your findings?
- To conclude:
    - What impact do you expect your data science science innovation to have? How is the organization likely to respond to the data science solution, findings, or product?
    - Name a likely obstacle that might lead to unintended consequences. How might you overcome that obstacle?

This assignment can be as many slides as you would like, but the total time of presentation should not exceed 13 minutes. The time limit will be strictly enforced. An objective of this class is to help you improve your communication skills. This includes using the most precise and concise language possible. Practice! You will be cut off if you go over 13 minutes. (Keep in mind that future group assignment will be capped at 10 minutes. Use this presentation to practice the use of precise and concise language.)

Many hands make for light work, so feel free to divide up the labor in any way you see fit. That said, each student must speak during the presentation. Be sure to leave time to review the whole product and make sure the pieces flow, fit, and reinforce each other.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.

IV. Conveying Findings

Week 9: Visualization as Story (Individual Assignment #2)

Tues. classes: Present during live session (3/5)
Deliverables due 11:59pm PST the night before live session

Objective: To improve your data science narrative/storytelling skills.

Find an example from a (business) magazine source or an (academic) article where an argument is made in an extremely data-centric or data-driven fashion with intensive use of visualizations (e.g. Economist, BusinessWeek, Wall Street Journal, an academic article in your domain of choice, etc.). You are welcome to find a new article, use one you discussed in a previous assignment or discussion question, or use an article you saw on a Slack channel.

Recast the argument as a story that features **the minimum necessary amount of data** to communicate the story. Use the absolute minimum amount of text on the slide. You should have a title for each slide, but **you should not have** bullet points as you had in previous presentations. In other words, create a visual storyboard that follows the story structure in the asynchronous material:

- Once upon a time…
- And every day…
- Until one day…
- And because of this… (repeat)
- Until finally…
- And ever since that day…

Create a visually-rich presentation. You may use visualizations from the original article, from other sources (cite them) and/or you may construct your own. In this assignment you will bring the *invisible Ink* framework to the surface. In the past, the best presentations used few, if any,

graphs and tables. Instead, those presentations **explicitly** used the *Invisible Ink* narrative framework and pictures. Past student examples can be found here: example 1, example 2.

The goal is to demonstrate that a strong narrative structure does not need to rely on tables, graphics, or data. Use the absolute minimum amount of text on the slide. You should have a title for each slide, but are not required to have a series of bullet points as you had in previous presentations. Remember, this precise and concise narrative structure does not mean you need to tell an overly simplistic story.

In future presentations the *Invisible Ink* framework will likely below the surface; however, if we bring the structure front-and-center, we can better uncover its usefulness.

Define your audience on the first slide (See Define Your Audience earlier in this document). For this assignment it may be difficult to think about an audience beyond this class. Perhaps this could be a lunchtime talk at company X. Or perhaps this is a talk at a particular data conference on panel Y. Pick a specific audience outside of this course.

You will have up to 6 minutes to present. The time limit will be strictly enforced. An objective of this class is to help you improve your communication skills. This includes using the most precise and concise language possible. Practice! You will be cut off by the panel chair you go over 6 minutes.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.

Week 10: Visualization (Discussion Question)

Find two publicly available data visualizations: one that you find distinctly powerful and one that is weak or distracting. What makes them effective and poorly executed?

Consider whether Tufte's notions of chartjunk and data-ink maximization help explain your judgment of what is good or bad (they may not). If you have your own quality criteria, name and explain them (e.g. interactive visualizations).

Be prepared to share and discuss your thoughts in class. Use the discussion section of your portfolio to write your notes. Be prepared to have these images readily available either as a URL

or a PDF file, as you may share them during the live sessions.

Week 11: Persuasion and the Future of Data Science (Group Project #2)

Tues. classes: Present during live session (3/19),
Deliverables due 11:59pm PST the night before live session.

Objective: To explore a controversial debate in data science that will likely affect a domain of your choice.

Identify a specific example of a highly controversial and heated debate that is or will take place in the public eye, and whose outcome will significantly affect data science in the future. For example, debates over the National Security Agency (NSA) surveillance are likely to have an effect on how government agencies may collect and/or access certain data.
Keep the following points in mind as your prepare for the presentation:

1. The context of the issue or debate.
2. Pro side's argument and who's making it (e.g. What's the argument in favor of NSA surveillance and who is making that argument?)
3. Con side's argument and who's making it.
4. Who you think is likely to win the argument and why. (ie. what is the most likely outcome and why?)
5. Suggest a "third way" that applies data science to either transcend the debate or solve the business problem ethically.

When you consider pros and cons, be aware of organizational and individual biases, as well as fallacies of argumentation.

Your group will have up to 10 minutes to present during the live session. Each student must speak during the presentation. The time limit will be strictly enforced. An objective of this class is to help you improve your communication skills. This includes using the most precise and concise language possible. Practice! You will be cut off if you go over 10 minutes.

Note: In week 8 I wanted you to explicitly identify three course concepts in your presentation. In this and subsequent assignments, this should happen organically.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.

The following will help you digest the async material for this week. We will talk briefly about the aync in live session.

Identify one of the most difficult persuasion challenges you've faced in the past year. This could be from business or any other part of your life.

Keep the following questions in mind:

- What was the persuasion challenge?
- Why was it so difficult?
- What tools do you wish you had (not limited to data) that would have made the persuasion more effective?

You're welcome to use the discussion section of your portfolio to write your notes to prepare for live session.

(Week of 3/25--No Jackson live session because of spring break)

V. Future of Data Science

Weeks 12-14 Overview

Write the week 12 abstract as if the research project **was successfully completed.** The project in week 13-14 is a research design that articulates **how you would execute the idea you choose from week 12**. You can also think of the week 13-14 assignment as something similar to a grant proposal. Before someone gives you a ton of money, convince them that your work matters and that you have a realistic idea of how to complete the project.

Curious how your final project might relate to the Capstone course at the end of the MIDS program? Watch this short video of a Capstone instructor presentation during a Fall 2017 w201 live session. (You will have to download Adobe Connect, our old live session platform.)

Week 12: Final Project Abstract (Individual Assignment #3)

Tues. classes: Present during live session (4/2),
Deliverables due 11:59pm PST the night before live session.

Objective: To propose an idea for a final group project that could answer an important data science question. To practice the delivering an elevator speech.

Pitch your final project proposal to the class. Use the abstract style of a published journal or magazine article. In 300-500 words, include the following as if your final project idea has been implemented:

- Title – The catchier the better but make sure it's informative; sometimes you can pack your whole argument in the title!
- Abstract – A paragraph that summarizes the most important features of your idea. Draw from the beginning (problem), middle (research), and end (findings) of your vision; try to give the reader a taste of the total project. Use precise and concise language.
- Keywords – If someone searched online, what search terms would help them find your project?
- About the Author – This is your project; describe yourself in-context after the project is completed. This should include a description of the role that allowed you to execute this project, but it can also refer to your other affiliations, accomplishments, or irons-in-the-fire. *You may want to start with your biography from week 1 or the onboarding.* (To be clear, this can include hypothetical details.)

To give you a sense of the format, below you can find an example from a previous semester. Your style and content should be different, but the format and relative lengths should be similar.

> **Author:** Max Power, Ph.D.
>
> **Title:** Gordon™—A Virtual Cooking Coach in the Internet of Things
>
> **Abstract:** Home appliances provide information on demand and can act as limited agents on our behalf; until now, they have lacked the capacity to observe our actions and teach us how to do anything. I describe an AI platform (Gordon) that integrates sensors within kitchen appliances and a typical smartphone to teach users how to cook. First, Gordon queries users for their dietary preferences/restrictions and nutritional requirements to inform a user's meal plan. Next, Gordon

walks users through a series of tutorials to gauge their cooking skill; Gordon does this by making simple observations such as ascertaining if a user is employing the right pan, checking if s/he has set the stove top to the correct temperature, employing visual recognition to evaluate the form and size of diced ingredients (ChopCheck™), and checking basic aspects of food chemistry (e.g., acidity, salinity) with our custom line of SmartSaucier™ pans.

Gordon employs these data to develop a menu that is within the user's ability to execute and is also customizable to both available ingredients and the amount of effort a user wishes to put forth in preparing a given meal. Gordon provides instructions and advice (e.g., "Don't forget to start caramelizing your onions now; you should also preheat your oven to 375°F.") to users in real time. Gordon also notes areas for improvement and can suggest additional instructional resources (e.g., video tutorials) to address a user's specific deficiencies. Users can also provide Gordon with input about the cooking experience ("I had more trouble making a roux than I expected.") and resulting food ("That jambalaya isn't spicy enough!") to inform future meals.

Gordon's target audience is working professionals (ages 25–55). We present evidence to show that (*i*) users in this group enjoy working with Gordon and keep doing so for a sustained period, (*ii*) their cooking skills improve over time, and (*iii*) aspects of user health (e.g., weight, cholesterol level, blood pressure) improve as they learn to cook more meals at home from scratch. The data we gleaned from Gordon has also helped us improve user experiences for AI platforms in other consumer arenas.

Keywords: Gordon, smart kitchen, cooking, internet of things, virtual assistant

About the Author: A scientist by trade, Max is also a passionate home cook, lifelong teacher/learner, and founder/CEO of MP Technologies. Prior to completing a postdoctoral fellowship in genetics at the National Institutes of Health, Max earned his Ph.D. in cellular biology at Yale University and a B.S. in biology at the University of Montana. He also has a master's degree in information and data science from the University of California, Berkeley.

For examples of other data science bios and abstracts take a peek at the DataEdge 2016 speaker schedules (here is a good example).

Deliverables for week 12 are unique. Follow the instructions below.

- Instead of a slidedeck, you will create an abstract only document.
    - Create a separate google doc (example) that includes your title, abstract, keywords, and about the author. Slides must be housed in your Berkeley google drive. Make sure you set the appropriate document permissions.
- Post abstract document to Slack.

    - I will pin a “Week 12 Final Abstract Here” post on Slack and you will comment on that post with a link to your separate abstract google document. Make sure you set the appropriate document permissions.
- Link to video recording in portfolio.
    - Only audio is required. You are not required to show your face.

Week 12 Presentation Tips

- During the live session you will have up to 1 minute to present an elevator speech.
    - You will present your one minute elevator speech that highlights the crux of your abstract. You must stick to the 1 minute time limit. (Most elevator rides are shorter than this.) Choose your words carefully. Practice! I will cut you off after 1 minute and say “Unfortunately our elevator ride has ended. Goodbye.”
    - There are no slide for your elevator speech; it is especially important to practice your speech; the audience will not have any visual cues.
    - Don’t forget to identify your intended audience in your abstract and in your speech. In the concluding sentences of your abstract I would recommend something like "this innovation is likely to be used by xyz" or "this innovation will be especially applicable for those in xyz industry." In the elevator speech, you could say something like “this innovation is likely to be used by xyz” or you could say something like “hello, Mary, CEO of company X.”
        - You’re welcome to organically communicate the audience in your elevator speech. If instead you want to set up the audience at the beginning, the setup must be within your 1 minute time limit.

Approximate format of live session

- 1 minutes elevator pitch
- After each presentation, the audience will take 2 minutes to provide peer feedback via Google form to identify the impact of each project and one potential challenge.
    - Note that I will post presenter and audience feedback to the section slack channel so that you and your group can process the feedback as soon as possible to help you select a topic.The feedback will be public since you are welcome to pick topics that originate outside of your group. You will receive brief instructor feedback via your individual portfolio.
    - The thoroughness of the feedback you provide in live session will be different from previous sessions. In earlier weeks, students had time to provide detailed feedback. This week, you will provide very brief, high-level feedback on the potential impact of the project and a challenge you foresee.

- After individual elevator speeches, you will meet with your group to rank the top three ideas you are likely to work on in weeks 13 & 14.
    - During the group break out you may view the abstracts on slack (see the “Post week 12 final abstracts here” pinned post). Think about the potential impact of each project (e.g. Business impact/monetary value; public policy implications, etc). Also think about a challenge you foresee and how you might overcome it. Your top three can include any of the ideas presented in class.
    - If you pursue an idea that originated outside of your group, please give your peer credit on the first slide of your final presentation.
    - You can view the abstracts in the Slack post.

VI. Wrap-up

The last two weeks build toward a larger group project. For Weeks 13 and 14 work on the project you selected from Week 12 and explore the current state of data science in that domain. You are welcome to take the idea in a different direction than initially presented. If you pursue an idea that originated outside of your group, please give your peer credit on the first slide of your final presentation.

Week 13: Final Project Proposal (Group Project #3)

Wed classes: Present during live session (4/9),
Deliverables due 11:59pm PST the night before live session.

Objective: To receive feedback from instructor on your final group project.

Choose one topic from those presented in week 12 live session. For Week 13, produce at least three slides of your final presentation to get feedback from the instructor. I encourage you to look ahead and check the requirements for the final project under Week 14.

Your slides will include:

1) Research question and intended audience
    - Before you meet with your instructor, your group must select one research question within a domain of your choice (you may have sub-questions). As usual, it will help to think about a specific decision and a specific decision-maker/client. To help focus your project, take the perspective of your decision-maker. What questions do they ask, or what unsolved problems do they currently perceive?
2) The data you might need and proposed analysis
    - Begin to think about the data and analysis you would need to address the question. The analysis should help someone choose between viable options (include the option to maintain the status quo). Clearly articulate how your findings will be actionable by your decision maker (i.e. show that your analysis will impact how decisions are made). It is important that the insight will help to make a meaningful decision--not just predict or optimize.
3) How you would convey the findings / final product or service delivered to client
    - Consider how you would 1) use analytical methods to maximize the likelihood of gaining actionable insights from the data and 2) convey your findings effectively to stakeholders. In other words, what is the final product or service that will be delivered to your client(s)?

Week 13's Live Session will be primarily group-instructor meetings. This is an opportunity to meet with the instructor to receive feedback before the final presentation. You must post the slides to be discussed prior to live session just like other assignments. The quality of the discussion and feedback will be a function of how developed the slides are. Group meetings with the instructor will be approximately 15 mins. While your group waits to meet with the instructor, use the time to review the week 14 assignment, fill out the week 14 Team Process Agreement, prepare questions for the instructor, brainstorm, work, plan, etc.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.

- For week 13 you do not need to record a video. That said, you must annotate your slides so that I can come to our meeting with prepared remarks.
- The meeting with the instructor will be more workshop than formal presentation.
- I will provide verbal student feedback during live session; there will be no written feedback via your portfolio.

Week 14: Final Project (Group Project)

Wed classes: Present during live session (4/16),
Deliverables due 11:59pm PST the night before live session.

Objective: To produce a research design that could answer an important data science question.

Use the feedback from Week 13 and create detailed research design in the form of a pitch-style presentation (i.e. imagine you are in front of an audience that may fund your research). Convince us that your work is valuable and compelling! Integrate course concepts when appropriate and do so organically. (e.g. Rather than say “and here we see recency bias,” you can say “we should be mindful of how we present our findings because the industry has had strong recent earning, but we should not forget the challenges we have had in recent years.)

You have up to 10 mins to present. The time limit will be strictly enforced. An objective of this class is to help you improve your communication skills. This includes using the most precise and concise language possible. Practice! You will be cut off if you go over 10 minutes.

The final presentation should include the following:

- The domain and state of data science in that domain (i.e. What types of decisions confront leaders in that domain? What data do they use to inform those decisions?),
- your research question and why it matters (i.e. the “so what?”),
- the data and methods needed to address the question,
- the final product or service that would be delivered to your clients,
- your plan to convey findings and persuade decision makers (See note below); this should include how you will manage issues, risks and contingencies (e.g. potential legal or ethical issues when you acquire or use the data, issues in how you use the results, etc.),
- and further questions and projects that could come from this study’s findings.

Clarification on "your plan to convey findings and persuade decision makers."
Think of this as your strategy to persuade decision makers. More specifically, do you think there will be pushback/dissent in a particular area? If so, how might you counteract any pushback/dissent? In your presentation you may say something like "you all may be concerned

about X, but let me assure you that Y & Z." (e.g. cost, political pushback, concern you will be unable to get the data, concern the idea isn't in line with the company's vision, etc).

Think of this as an opportunity to address broad arguments against your idea and then develop counter arguments. Here, you should address more broad conceptual issues. This is distinct from "how you will manage issues, risks and contingencies" where you will address specific legal/ethical issues.

Review the Individual & Group Assignments section above for general reminders about presentation best practices and for details on the expected assignment deliverables.
