A Clarification on Ethical and Professional Commitments

We are making a change to our grading and guidance about ethical commitments, to accept a broader range of them. In Weeks 1-3 the grading team has applied a narrower, careful definition of “ethical commitments”; the video lectures and some of the samples employed a broader notion that also encompasses what might be called “professional commitments”. We have **decided to accept the broader definition** and will be retroactively going back through the grading of ethical commitments from weeks 1-3 to assign extra points where warranted.

I think it is helpful for data scientists to articulate some ethical and professional commitments that remind them of how they want to behave, in the face of pressures that may arise to do otherwise. In some cases, these commitments may be ethical in nature, but in some cases they may just be commitments to following best professional practices (e.g., evaluate on held-out data, not on data used for estimating the model). Not following these practices can lead to bad data science outcomes, which will, of course, have bad outcomes for the people and organizations affected by the decisions the data science project supports. For our grading, we will now accept explanations of importance of an ethical or professional commitment that give an example of what could go wrong within the confines of the project, and do not describe who will be harmed when the faulty data science project is put into practice

In some professions, there are articulated universal commitments (e.g., in medicine, the Hippocratic oath of “first, do no harm”). In professions with less history, there may not be such universal commitments. One of my former mentors, the late Prof. Michael Cohen, used to teach about the history of professionalization and say that “real professions have ethics”. In 503, you will have an assignment of adapting an existing oath, pledge, or checklist from another field for use in data science.

So, if you’ve been getting points off for providing ethical commitments that were professional in nature rather than ethical, and you’ve articulated harms only within the context of the data science project, we are declaring that it is **good enough for the manifestos.** (Of course you are still required to explain the commitment using project details; the rubric requirements remain the same.) **You do not need to replace or reframe these in more directly ethical terms for your final project.**

Week 3 Grading/Feedback notes

The most important thing we can tell you right now is to PLEASE review the feedback you’ve been given (as well as the feedback in this document, the exemplars, and Slack conversations) as you prepare your final manifesto. We expect that you will integrate the feedback you have received in Weeks 1-3 into your final manifesto in the sections Application in Domain of Interest, QMECs (weeks 1-3), interview reflection, and tweet (Note that you do NOT need to revise your reading responses or plans for knowledge acquisition (weeks 1-3) for this assignment.)

General feedback

- Submitting your PDF: when you are asked to select pages per question, please note that you can and should select multiple pages for a single question when your answer is split between two pages.

QMECs

- Suggestion: if you're struggling to create a QMEC, don't forget about your reading response insights! The insights you extract from the readings are often great material for the QMECs.
- **Ethical Commitments** should have an actionable element that you can commit to do or not do. It is not the same as a professional good practice. (We've said this in previous weeks). There should be negative implications for some identifiable set of stakeholders if you do not follow through on the commitment.
    - Here is some advice: Ethics is more than consistently applying good data science practices (should go without saying that you'll apply techniques properly) or not doing something wrong, bad, or illegal. An ethical commitment takes into account your actions on others, in particular those with less power than you. For this assignment, select a commitment that addresses how your actions as a data scientist affect others. Discuss what benefits or harms might result, who might be affected by your actions (stakeholders in ethics terms), and how following the commitment will help you avoid those harms or provide those benefits. In addition, make sure your chosen ethical commitment pertains to the stage we are covering in a given week.
    - **Specific to the Week 3 stage, Analysis & Modeling**: p-hacking, multiple testing and assuming causation instead of correlation are not self-evidently ethical issues, not even HARKing. In your explanations you need to actively address the ethical component. Ask yourself: what possible harms might happen as a result of not following this commitment? Which stakeholders are likely to be affected, and how? In other words, go beyond the outcome of someone being misinformed to identify actions that might be taken that have a negative impact on someone (investors, customers, etc.)
- Look at the QMEC Grading rubric section from Week 1 feedback: some of you are still misunderstanding the "Importance" section as "'why the project is important for the domain' as opposed to 'why the QMEC is important to the project'.

Reading Responses

- a maxim from the 501 grading team: when in doubt, choose "expertise over maxim"
    - Reminder: a maxim is a short, pithy statement that expresses some widely applicable point of wisdom. Expertise is something that may be gained from some combination of education and (professional) experience. Expertise may certainly be expressed in a maxim, but is not required to be.

Peer review (review of W2 work)

- The directions for this week were copied from Week 2 (review of W1 content). For this reason, we only graded your feedback on Week 2 QMEC content.

Interview reflections

Please make sure there's a clear mapping between the interview QMEC and the project stage. Make sure it's clear to an external reader why the stage you've chosen fits.

Week 2 Grading/Feedback notes

The first piece of feedback is that many of you would benefit from (re)reading the Week 1 feedback notes. We saw a lot of improved work, but also a lot of the same mistakes carried over from Week 1.

Please review your PDFs before submitting to ensure that you have accurately mapped your answers to the questions in Gradescope. This will make the lives of your hardworking graders much easier.

Information for Week 3 assignments

- In the Week 3 manifesto assignment,there are two additional "questions" to map pages to: your two problem descriptions. These are **unscored/ungraded**, and we will not be providing feedback on them. We are asking you for this mapping just to create a quick reference point for the instructional team in case they need to look at your problem descriptions while grading.
- **Pro tip: Look at the grading rubric detail!** When grading your weekly manifesto sections, the instructional team uses a detailed rubric. This is very close to the same rubric we will use for the final manifesto. As you work on Week 3, take a look not only at your Week 2 grades and feedback, but also the specific point distribution of the rubrics. These will help you to improve your responses, as we have told you here what we are looking for.
- **Informational Interview Reflection:**
  The reflection will be worth 8 points. You will be able to revise/edit your content as part of the final manifesto submission. You can choose the format (paragraph, bullets, etc.) that suits you. We will grade your required content as follows:
    - Three insights relevant to course content: one question, one maxim, one ethical commitment. Insights should be related to course content, and described in adequate detail. *[1 point each for question, maxim, ethical commitment]*
    - Map each insight to a single stage in the data science project stages framework: project stage mapping should be correct and clear to an external

reader (i.e. if it's not obvious why you chose this stage, explain it!) *[1 point each for question, maxim, ethical commitment]*
        - Note: this may be any of the four stages; it should just be clear why you chose the stage you did *[1/25 addition in response to student question]*
    - Three follow up questions: you need to include three, and they should be specific, related to the interview (e.g. the person's work as a data scientist, other content from the interview, etc.) and seem important to ask (e.g. you're not asking about hobbies, etc.). *[1 point]*
    - Length/detail: your reflection should be within the recommended word count range, neither too short nor too long; in other words, it contains a sufficient amount of detail to demonstrate what has been learned from the interview. *[1 point]*

Questions, Maxims, Ethical Commitments (QMEC) (Week 2)

- **The QMEC should be a general principle.** Each QMEC you choose should be something general you will always ask at the project stage covered in the week's course content. It should be able to be applied to many different projects or scenarios you encounter in your work as a data scientist rather than be specific to a project. You will then use the details of your chosen project to illustrate or expand your reader's understanding of the general principle (the QMEC) and what it would mean to address it in the "meaning in context" and "importance" sections (Rubric points #3 & #4). NOTE: *If you submit project-specific QMECs for your final manifesto, you will not receive full credit for these answers.* (This advice is repeated from last week.)
- **The QMEC should be applicable to the project stage covered in class this week** (i.e. data collection and cleaning for Week 2)**.** If you receive feedback that one of your QMECs did not fit the project stage, consider relocating (and editing) that one to the appropriate stage for the final manifesto and choose something more appropriate for the other week. (This advice is repeated from last week.)
- **Project descriptions (rubric point #2)**: The project descriptions for QMECs should be short but descriptive. A helpful tip would be to make sure the input and output is clear from the one-sentence description.
- **Ethical commitments:**
    - When considering ethical commitments, you should be able to identify specific potential harms that could happen as a result of not following the commitment. You should think of an ethical commitment as something directed toward your clients, users, company, or society at large. Many of the ethical commitments we see each week are variations on themes of 1) protecting yourself from liability or culpability, 2) not breaking the law, or 3) a version of "I will do my job well." This is not sufficient: committing to avoid something that is just poor data science practice (or is illegal) is not in and of

itself an ethical commitment. Another example: Statistical bias is not the same thing as prejudice, and is not automatically an ethical issue.

Tweet

- If you are not an active Twitter user or are unfamiliar with the tweet format, please take a look at the Rogati article linked in the assignment directions (in the manifesto template). They will help you to understand what we are looking for.
- It is perfectly acceptable to talk about this (fictional) project in the past tense as if you have completed it; doing this will help you to be specific about the project outcomes (rubric point #2).
- Unless your tweet is one of the two projects in your Application in Domain of Interest section (and it does not have to be!), please do not use the tweet as your sample project in your Week 2 QMEC responses.

Peer feedback

- You do not need to provide feedback on the skills inventory, reading insights, or other sections not included in the directions.
- Please submit your marked up copy of your partner's assignment rather than a paragraph summary of your feedback. This allows the instructional team to look at your feedback in the context of your partner's assignment.
- In order to get full credit for your feedback, you must submit a PDF with the text of your feedback visible in Gradescope. Please check your submission after uploading. There were multiple threads about mark up strategies on Slack in Week 2; please search these if you have received feedback that your comments were not accessible to instructors.
- If your grader informs you that you have provided feedback to your partner that is wrong (e.g.you have inadvertently encouraged your partner to make changes that will make their submission worse), PLEASE reach out to your partner to let them know. You can also tell your partner to reach out to one of the lecturers (Melissa, Bobby, Anthea) if they would like to discuss it.
- Your feedback needs to be specific, preferably actionable - it should provide your partner with insight into how they can improve their submissions for the final manifesto (or build on strengths already in the drafts). It needs to go beyond a summary of what your partner wrote or enthusiasm (e.g. "great work!"). For ideas of what to look for, you could use 1) the grading rubric 2) feedback you received from the graders on your own work 3) feedback provided in this feedback document.
    - In particular, for the QMEC: look carefully at how well your partner has used the details of their chosen project to illustrate or expand understanding of the general principle identified in the QMEC. How might this be improved?

Reading responses

- The instructions require that you identify *and* explain the insight. You can use a somewhat general quote as your insight, but you need to explain its meaning and how it connected to the specific reading.

Week 1 Grading/Feedback Notes

You may have received a lower grade than you are accustomed to receiving. Don't panic. The course grading is structured around giving you feedback on the first assignment without it affecting your course grade very much. (The first week counts for only 5% of the total course points).

Here are some resources for making sense of the grades and feedback, and figuring out what to do next.

1. Week two has several sections that have the same structure and expectations as week 1: questions, maxims, and ethical commitments; plan for knowledge acquisition; reading responses. Make sure you look at the grades and feedback from week 1 and adjust what you write accordingly.
2. Your week one material will be resubmitted and graded anew at the end of week four. That will be higher stakes (80% of your grade). You may choose to edit your week one material now, or you can edit everything at the end of week four.
3. You may choose to change the two projects with problem type classification from what you submitted in week one. You can make this change at any time throughout the four weeks. However, at the end of the four weeks, you will need to have a coherent set of maxims, questions, and ethical commitments that are all applied to the same two projects. So if you change the projects, that may require rewriting some maxims, questions, or ethical commitments from previous weeks.
4. **Getting Advice.**
   If you got feedback that you are having trouble understanding, or perhaps you disagree with it, and would like some substantive advice on how to improve what you wrote, please follow these steps:
   a. Write up as specific a question as you can, including enough context for the instructor team to understand it. It may help to include a screenshot of what you wrote and what the feedback was.
      i. Be sure to include your best interpretation of what the feedback is telling you. If you have multiple possible interpretations, feel free to describe them and the instructional team can then help you decide which was intended.
      ii. If you have a specific question about a possible improvement and you'd like advice about whether that would be better, please include that.
   b. Send a slack DM to the whole lecturer team (@Prof. Resnick, @Melissa Chalmers, @Bobby Madamanchi, @Anthea Josias) with your question or

concern (written up as an attachment or inline within slack). One of us will get back to you in the slack thread, usually within 24 hours.

5. **Requesting a Regrade.**
   If you think the grade is incorrect (it happens sometimes!), and would like to request a regrade, please follow these steps:
   a. Submit a regrade request through Gradescope (instructions)
   b. Per the regrade policy as stated in the syllabus: "Regrade requests will be handled by a different grader. The entire question will be regraded, applying all elements of the grading rubric. Your grade could go up or down."

Some General Feedback from Week 1 Graders

You may find it helpful to spend time looking at the samples and exemplar manifestos we have provided for you in the Resources section of the course on Coursera:
https://www.coursera.org/learn/siads501/resources/8wQqe

Project Descriptions and Types feedback

Make sure you are trying to solve a data science problem.

- In several examples, students tried to solve non data science problems and called it causal modeling. For things that need to be simulated because there are known (or hypothesized) physical or biological processes, causal modeling isn't sufficient. However, you can apply data science to interpretation and fitting of simulations up the data. For example optimization can be used to test different model parameters or to evaluate model fit.

Make sure you explain how data science is adding something to the available data. Sometimes students propose to do something like use a classifier to find the highest or lowest performers in a population. If you already have a measure of performance in your data set, then ask yourself: what does data science bring to the table? Data science can help to predict performance in other populations or can be used to identify the factors associated with performance. Please explain how data science is going to be applied.

Grading rubric: The grading rubric for this section will remain the same for the final manifesto (Week 4). Here are some tips for improving your final manifesto:

- "Explanation for classification is correct and specific" (+1)
  - The assignment instructions did not specify that a separate explanation for problem type is required; you had to guess from the examples we provided. Here's a clarification on what we're looking for with this rubric element: does your problem explanation contain sufficient information for a reader to determine that your problem type is correct?
- "It's clear what kinds of data would be used." (+1)
  - Other than the grade, this is an important step for you in developing your projects. If you can clearly state what the inputs to your model are, even

examples of possible variables, it would help you see 1) if the project is realistic, and 2) if the type is correct.

Questions, Maxims, Ethical Commitments (QMEC) general feedback

Each QMEC you choose should be something general you will always ask at the project stage covered in the week's course content. It should be able to be applied to many different projects or scenarios you encounter in your work as a data scientist rather than be specific to a project.

- Be strategic: If you got Week 1 feedback that one of your QMECs did not fit the project stage, consider relocating (and editing) that one to the appropriate stage for the final manifesto and choose something more appropriate for problem formulation for Week 1.

Defining MAXIMS

A Maxim is a short, pithy statement that expresses some general truth in a memorable way. If it takes several sentences to express your maxim, it is probably not a maxim. If it only applies to the particular project, it is not a general maxim.

QMEC Grading rubric

The grading rubric for QMEC sections will remain the same for Weeks 1-4. With this in mind, here are some general things to remember for your QMECs:

- *Your QMEC is *one* sentence, not multiple. (+1 / -0.5).*
    - Importantly, your sentence contains one primary idea rather than multiple.
- *Provide a *one* sentence summary of which of your two projects your question pertains to*. It is meant to remind us. If we need more info, we can look back. (+1)
    - Here are some examples. Bad = "This Q applies to Project 1". Good: "Project 1 trains a machine learning model to recognize misinformation on Twitter to flag for human review."
- *You explain how the general QMEC you chose applies in the context of your Project at the relevant project stage.* (+1)
    - Make sure you have provided an explanation of the general principle AND how it relates directly to your project. This is your chance to use the details of your chosen project to illustrate or expand your reader's understanding of the QMEC and what it would mean to address it.
- *You explain why it is valuable to ask your QMEC at the relevant project stage, and suggest how it would make your project better*. (+1)
    - Some questions you could address here: how does addressing this QMEC shape the project at a specific stage (and moving forward)? What would the consequences be for the project if you did not address this MQEC?
    - **From the template - "Importance" section:** Think of this as being labeled "Importance *for this stage of the project*." This is not a place to explain why the project as a whole is important to undertake; it's a place to explain why the QMEC you have chosen is important to be taken up at this stage of the project.

Additional general QMEC tips

- Be specific and concrete. Specific sentences include words about people, places, and things. Concrete sentences are about particular things that happen at particular times.
- Use complete sentences. Introduce bullet points or lists with a sentence so we know why you are enumerating items.
- Use details from the projects you have chosen in your explanations! Examples are great ways to illustrate and clarify your points.
- You don't have to use them, but here are some sentence snippets to get your thinking started:
    - "It is important to ask [this Q/M/E/C] at the [relevant project stage] because…[specific, concrete reasons]".
    - "In the context of [my Project], it is valuable to ask [my QMEC], because [specific, concrete reasons]. For example, [example]."

Reading Responses

A note on labeling: if it's plausible for the chosen insight to be applicable to multiple project stages, make sure that your description makes it clear why you've chosen this particular stage.

Plan for Knowledge Acquisition

To add specificity to your plans for each inventory option, you may find it helpful to consult the list of MADS courses: https://www.si.umich.edu/programs/master-applied-data-science-online/curriculum/mads-courses