ECON-GA-4005: Project Guide

Revised: April 7, 2022

Overview

One key outcome of this course is for you to produce a piece of work you can use to demonstrate your computational skills to potential employers. That work will come in the form of a Jupyter notebook. As you know, these notebooks allow you to combine code, text, and graphics in one user-friendly document. We will add them to our GitHub repository (with your permission) so that you can use a link to show others what you've done. You will also be able to see what your classmates have done.

The relatively loose structure in self-directed projects like this makes them more challenging than most things you do in school. We think that also makes them more interesting. They give you a chance to indulge your curiosity and show off your creativity.

We have divided the project into components to keep you on track. The intent is to make the project easier by breaking it down into a number of small, manageable sub-projects. The early steps are graded only on whether you do them: you get full points if you do them, none if you don't. We start with individual work. Partway along, we encourage you to form groups of two or three, but if you'd prefer to do this on your own, that's ok, too.

Project outline

The components of the project are

Assignment	Format	Individual/Group	Points	Due
Three Project Ideas	Electronic (quiz on brightspace)	Individual	0	Apr 7
Project Proposal	Written (professional)	Group	5	Apr 23
Data Report	Written (professional)	Group	10	May 7
Final Project	Jupyter notebook	Group	85	May 21

Dates are firm and not open for negotiation.

Each part of the project must be uploaded to Brightspace before the deadline.

The components consist of:

- **Three project ideas.**

  Write down three project ideas. One or two sentences each is enough. Use your imagination. Be creative. Speak to others. Write down things that interest you. This will not be graded, but it will give you something to work with later on.

  Over the coming weeks, we recommend you keep your ears open for possible **project ideas** — and **data sources**.

- **Project proposal.**

  Form a group of between one and three — no more — and choose a single project from those you submitted — or perhaps some other idea if you get a sudden flash of inspiration. Flesh out the project in more detail, **including the data source, ideas for a model** and two figures you plan to produce. Total length should be no more than two pages.

  On the off-chance you missed the point: you should **be clear about the data and type of model you plan to use**. If you can't get the data, your whole project can be derailed.

  Did we mention data? Or models?

- **Data and Model report.**

  Describe your data and how you accessed it in enough detail that someone else could do it. Yes, that's "accessed," past tense. You should have done this already to make sure you have it. Include your Python code.

  Internet sources are preferred, because it allows others to use your code and follow up on your work.

  Also include notes about the type of model you plan to fit. You to not necessarily need to have fit the model at this stage, but you should show us what you are thinking. Now that you have the data, being explicit and clear (through writing) about how you plan to use the data is extremely helpful in spotting potential roadblocks.

- **Final project.**

  You should submit your **notebook** to Brightspace by the due date listed above. The submission should include any data or supplementary files needed to execute your project. The final writeup should be a Jupyter notebook and the file name should be your last names separated by dashes and a short title — something like `Jones-Smith-Zhang.ipynb`.

  Your project should include:

  – Description cell. Put a Markdown cell at the top of your notebook that includes the title of your project, a list of authors, and a short summary of what you do. Think of the last one as an advertisement to potential readers.

  – Data. Description of data sources and the code to read the data into Python and reformat it as needed. This should be done in enough detail that someone else can reproduce what you've done.

  – Model. Description of the model(s) used to analyze the data and the code to fit the model. This should be done in enough detail that someone else can reproduce what you've done.

  – Graphics. A series of figures that tell us something interesting. Three or four would be enough, but do what you think works best for your project. Think about the narrative: What story do you want to tell?

  – Conclusion. What was the main outcome or result of your project? What did you learn? What should readers take away? Were you surprised by anything? Is there futher work you or someone else could to do extend your analysis?

Free advice

Some things to keep in mind:

- **Keep it simple.**

  Most project ideas turn out to be too big. You're generally well-advised to carve out a manageable subset of what you think you can do. There's no reason to worry about this at the idea generation stage, but as you develop your project you may find that you need to focus more narrowly on a part of it.

- **Find data.** Make sure you can get the data you need. One way to assure this is to start with data and ask what you can do with it. Ideally you want the intersection (picture a Venn diagram) of an interesting idea and good data. You can start with either one, but ideas are often easier to find than data. Or start with an existing project that uses data you can access yourself and extend it in some way.

- **Ask for help.** We have years of experience with this kind of thing. If you're stuck, let us know and we'll try to help. You can also post questions on the class discussion board.

Grading

Projects will be graded on their overall quality. This includes, but is not restricted to, these categories:

- Quality of the idea. Is the question clearly articulated? Is it interesting? Does it have general appeal?

- Quality of the data. Does the data support the idea? Is it the best data for this question?

- Quality of the model. Did you choose an appropriate model for your data and question? Where you able to fit the model effectively?

- Degree of difficulty. Some ideas are harder than others to implement. As in Olympic diving, you get credit for taking on a challenge.

- Professional look. Does your project look professional? Are the graphs easy to understand? Are they clearly labeled?

- Demonstration of skills. Does your project showcase what you've learned in this class? Do you use some of the theory or algorithms we've worked on? How does your application of these tools impact your final writeup?

CAP-6318: Project Grade Sheet

Revised: April 7, 2022

Overall

- Clear idea and message?
- Appropriate data?
- Documentation of sources?
- Effective graphics?
- Professional appearance?

Idea

Data

Model

Graphics

Difficulty

Professional look