CYPLAN 255: Urban informatics and visualization, Spring 2022

Final Project

Due May 9, 2022

Past students have described the final project as the most fun and rewarding part of this class. The project gives you an opportunity to choose an empirical research topic you're excited about, explore available data, and perform an in-depth analysis using the tools and concepts from class. Have fun with it and good luck!

Submission criteria

- Projects must be submitted in the form of a webpage hosted on your own GitHub account. The contents of your submission may be spread across multiple internal pages, but they should all be accessible from a main landing page using hyperlinks.
- To submit your project, upload the URL of your webpage to bCourses by the due date: 11:59 pm on May 9
- At a minimum, your project site should consist of the following:
    1. a write-up in the form of an essay or longform blog post which describes
        a. your research question(s)
        b. methodology
        c. key findings
    2. multiple visual elements which highlight key points from the text, and may include tables, charts, graphs, maps, and illustrations
    3. references as appropriate
- Your methodology need not be limited to what we've done in class. You can use whatever programming languages or software you like to complete your analyses. A successful project should, however, demonstrate mastery of at least *some* of the skills we've covered this semester.
- You will likely use some combination of Markdown and/or HTML to render your webpage. You can, of course, use one of the many GitHub Pages templates out there for structuring your site, and then populate the template with your own text and figures. If you want more control over the way your site looks, you might consider starting from scratch. Alternatively, you could choose to style your entire project, analysis and all, from within a Jupyter Notebook, and use a tool like nbconvert to convert your notebook to HTML. These are all perfectly acceptable options.
- You will also be expected to present your work in class during the final week of instruction. These presentations can be delivered using slides and/or your webpage as visual aids. The duration of the presentations will depend on how many individual presentations there are, but should be no more than 5-10 minutes.

Grading scheme

The final project is worth a total of 75 points (75% of your total grade for the class). Those points are distributed as follows

1. Project proposal + initial analysis, AKA Assignment 4 (10 pts)
2. Final presentation (20 pts)
3. GitHub Pages site (45 pts)

Project evaluation criteria

Each of you is coming into this class with a different level of experience, so there are no technical benchmarks that I expect you to achieve beyond what is mentioned above in the submission criteria. In general, the things I'm looking for are:

- Effort in applying skills from the class
- Creativity, both in your research design and in your problem solving
- Thoughtfulness and clarity in how you communicate your ideas, analyses, and findings
- Insight that you gain about either the research topic or the analytical tools you've implemented
- A website that looks clean/professional; doesn't need to be fancy, but it should be visually pleasing – e.g. consistent formatting, legible text (esp. figure axes labels), good use of color, etc.

Suggested work plan

1. Choose one or more research questions that can guide your investigation. Examples: What's the relationship between transit service and residential demographics in the East Bay? How much housing could plausibly be added to Los Angeles under current zoning rules, and where? What are the similarities and differences between public spaces in São Paulo and Mexico City?

2. Search for data sources that might be able to shed light on your questions. Hopefully some of the data will be readily analyzable (like a pre-prepared shapefile from a public agency), but you can get creative too: web scraping, .pdf parsing, etc. NOTE: you might need to iterate between steps 1 and 2 a few times before settling on a topic.

3. As a first pass, you can explore the data in a spreadsheet program or GIS software. You'll also write code to load data into Jupyter notebooks, and then more code to explore it, subset it, calculate descriptive statistics, join it with other data sources, generate charts and maps, and so on.

4. After a lot of exploratory analysis, you will write up your findings, including a description of your research question and methodology, and produce a set of data visualizations that support your key points (think longform blog post or essay). You will probably end up answering some of your original questions, but abandoning others, and going in some new directions as well.

Teams

You are encouraged to work in teams of two. Individual projects are also welcome, and larger teams will probably be permitted if you check with me first. More people does not mean less work per person.

Picking a topic

You're welcome to work on any topic that interests you. Some of your research questions may be more academic, while some of you may be more interested in using the tools we've learned to build *other* tools. Either is fine. If you're thinking about things that might fall outside the general scope of this class, come talk to me first so we can be sure we'll be able to provide you with appropriate support.

Also, don't worry if your topic is similar to someone else's, or a study that's already been done. You might be surprised by what you find when you try to recreate or extend an existing analysis. Along these lines, try searching Google Scholar to see what other work has been done on your topic.

Check out the example projects page on the class wiki for some inspiration of what you should be striving for. You should also peruse the list of public datasets that we've compiled on the class wiki.

Some example topics from previous years:

- Inequality and health outcomes
- BART and suburban job density
- Access to greenspace
- Disadvantaged communities
- Transit and equity
- Access to bicycle infrastructure in six Latin American cities
- Travel mode choice in San Francisco
- Rent burdens
- Mapping architectural typologies
- Identifying micro-geographies of gentrification in the Bay Area
- Effects of transit schedules on accessibility
- Replicating Google Maps "areas of interest" with open data
- Public housing in Rio de Janeiro
- Public transport in Cordoba, Argentina
- Vacant parcels in the Bay Area
- Measuring affordable housing
- Asthma rates near refineries
- Gerrymandering
- Lyme disease in California
- Bicycle infrastructure and business turnover
- Land use and flooding
- Incorporating hills into bicycle accessibility metrics
- Mapping iconic sights on Instagram
- Comparing travel behaviors across neighborhoods

Strategic advice

- Big data is not always good data. Small data can be just as impactful.
- Negative results are still results!
- Be realistic. Allot yourself a time budget and target an end product that you think you can hit.