

**Public Policy Analytics** MUSA 508 / CPLN 592

Lecture – Fridays, 10:15am–11.45am; Meyerson Hall, B3

Michael Fichman Lab (Section 402) – 12pm–1.30pm – Meyerson Hall, B3

Matt Harris Lab (Section 403) – 12pm–1.30pm – Meyerson Hall, B4

**Instructors:**

| **Name** | **Email** | **Office Hours** |
| --- | --- | --- |
| Ken Steif, Ph.D | ksteif@upenn.edu | TBA |
| Michael Fichman, MCP | mfichman@upenn.edu | Tuesday 10-12, Wed. 2-4, Calend.ly |
| Matt Harris | matthew.harris@micadatalabs.com | Monday TBA, Calend.ly |

**TAs:**

| **Name** | **Email** | **Office Hours** |
| --- | --- | --- |
| Anna Duan | annaduan@sas.upenn.edu | TBA |
| Syndey Goldstein, MCP | sydng@upenn.edu | TBA |

**Course Resources:**

Piazza page -<https://piazza.com/class/krap1pcyj696n1>

Course text -<https://urbanspatial.github.io/PublicPolicyAnalytics/> or

<https://www.routledge.com/Public-Policy-Analytics-Code-and-Context-for-Data-Science-in-Government/Steif/p/book/9780367507619>

**Course Description:** This course teaches advanced spatial analysis and an introduction to data science/machine learning in the urban planning and public policy realm. The class focuses on real-world spatial analysis applications and, in combination with introductory machine learning, provides students a modern framework for efficiently allocate limited resources across space. We will learn novel approaches for ensuring our models do not discriminate against communities/people of color as well as communities of different income levels.

The format of the class includes weekly lectures/in-class demos and labs. There are six required assignments, including two projects. Prerequisites include either CPLN503, the summer GIS course or prior experience with GIS in a formal setting. Having experience in R and the ‘tidyvserse’ is helpful.

**Grading:** The grading breakdown is as follows: 50% for homework; 20% for project 1; 20% for project 2; and 10% for participation. Your homework/project grade is dependent on your ability to motivate your analysis and communicate empirical results to a non-technical audience using maps and data visualization. Your participation grade is a function of both in-class participation and Piazza participation.

**Homework:** Homework is due on the dates indicated on the schedule below. Late homework will be accepted but penalized. Please prepare all homework as an R Markdown. You are encouraged welcome to work in groups, but **you must submit a homework assignment that is uniquely yours.** Your willingness to ask and answer questions on Piazza will count towards your participation grade. Homeworks will be graded on a scale of 1 (revise and resubmit) to 3 (superlative).

**How we will assess you:** Your grade in this course is not a function of your ability to write efficient or even ‘clean’ code. It is dependent on your ability to communicate technical Planning and public policy analytical concepts to non-technical decision-makers. This is done through clear and concise writing, and most important, through data visualization.

Some students will come in with significant experience coding in R, others, not so much. Admittedly, newcomers will have a more significant learning curve, but I urge you to form pairs and project teams with students who have diverse skillsets. Piazza is a great way to meet new people.

**Readings**: The course text, [*Public Policy Analytics*](https://urbanspatial.github.io/PublicPolicyAnalytics/)*,* includes all the code you will need for the course. A useful supplemental text on ‘tidy’ coding in R is [*R for Data Science*](https://r4ds.had.co.nz/)*.* For some great intuition on data visualization and ‘ggplot’, check out [*Data Visualization: A Practical Introduction*](https://socviz.co/index.html#preface).

**Academic Integrity:** Teamwork is essential in this class. However, please ensure that the work you turn in is uniquely your own. Your ability to copy/paste and hack together code from the readings and from the internet is essential, particularly for newcomers, but plagiarism of any kind is strictly prohibited. Do not copy and paste policy-related narrative. If you have a question about these issues, just ask one of the professors. We will adhere to the University’s [Code of Academic Integrity](https://catalog.upenn.edu/pennbook/code-of-academic-integrity/).

**Software**: This course will be taught using R and R Studio.

| **Date** | **Lecture** | **Lab** | **Readings** | **Assignment due** |
| --- | --- | --- | --- | --- |
| 3-Sep | Introduction to Public Policy Analytics | Introduction to the Tidyverse, Tidycensus, sf & ggplot | Read book [introduction](https://urbanspatial.github.io/PublicPolicyAnalytics/introduction.html) |  |
| 10-Sep | Why start with Indicators? | Collaborating on data science projects w/ Github w/ rmarkdown | Ch 1: Indicators for Transit  Oriented Development |  |
| 17-Sep | Data-driven Comprehensive Planning | Geoprocessing | Ch 2: Planning Urban Growth Area expansion |  |
| 24-Sep | Project 1 - Intro to geospatial modeling – Predicting home prices in Boulder, CO. | Intro to machine learning; regression; cross-validation. | Ch 3: Intro to Geospatial Machine Learning | TOD Assignment |
| 1-Oct | Project 1 - Intro to geospatial modeling – Predicting home prices in Boulder, CO. | Spatial autocorrelation; generalizability across space (race and class) | Ch 4: Geospatial ML; modeling the spatial process |  |
| 8-Oct | Time series prediction | Serial correlation; time series modeling in R | TBA | Predicting home prices midterm project |
| 15-Oct | NO CLASS |  |  |  |
| 22-Oct | Geospatial risk prediction – Predictive policing. See additional resources [here](https://github.com/urbanSpatial/PredictingFireRisk). | Algorithmic bias across space; Racist training data? | Ch 5: Predictive Policing |  |
| 29-Oct | Churn Prediction | Logistic regression; confusion metrics; cost/benefit analysis | Ch 6: Bounce to Work! | Risk prediction |
| 5-Nov | Data privacy, Disparate Impact, algorithmic fairness & predicting recidivism | Algorithmic bias across races; More racist training data; Memo-writing | Ch 7: Recidivism | Housing subsidy algorithm |
| 12-Nov | Space/time rideshare trip prediction | time lags; gganimate; purrr | Ch 8: Predicting ride share demand |  |
| 19-Nov | Data governance - Dennis Culhane talk/Final Project work | Final Project | Final project | Either recidivism memo OR bike share prediction |
| 26-Nov | No Class - Thanksgiving | Final Project | Final project |  |
| 3-Dec | Developing a GitHub.io portfolio / Final project work |  | Final project |  |
| 10-Dec | Final presentations - No lab |  |  | Final project |