CPLN 675 – Land Use & Environmental Modeling

Instructor: Michael Fichman

**Final Project 2022: Urban Growth Modeling –** Due May 10th at noon Eastern

In anticipation of the region’s next large-scale comprehensive planning process, the local Metropolitan Planning Organization (MPO) has asked you to forecast urban development in 2029.

Your task is to find a region of interest[[1]](#footnote-1) and gather land cover data from the USGS[[2]](#footnote-2) and land use/population/development data from federal (census), state and local open data websites[[3]](#footnote-3).

You will then construct a binary logistic regression model that predicts change in development from 2009 to 2019. You must employ the same goodness of fit metrics that we’ve used for binary models like the midterm. Spatial cross-validation is optional. Use the Houston sprawl markdown as your template for this assignment.

You will then come up with two planning scenarios – one forecasting a supply change and one, a demand change.

1. **Demand-side change:** In this scenario, you will find or approximate 2029 population projections for your study area and then ‘distribute’ the ‘new’ population among all the census tracts in the region according to a logic or plan that you devise[[4]](#footnote-4). You will use your population forecast and 2019 data to predict where new development will occur in 2029. Metropolitan planning organizations (MPOs) publish population forecasts for metro areas. If you can’t find one for your area, make some reasonable projection of your own and note that you have done so.
2. **Supply-side change:** In this scenario, you will “plan” a new large-scale development, like a new highway or public transportation line. Your model will probably use transit proximity as a predictor of land use change – what happens when you make new infrastructure and then predict the location of new development? Draw this new development in ArcGIS or the mapedit package (if it requires new infrastructure) and you will briefly describe the nature of this new development (What is it? What kind of new demand do you expect it to generate?). Please be creative in how you design/describe your intervention. After drawing the new development, come up with a scenario-based population projection for that area and create a development forecast for 2029.

Keep in mind that once you have a model that gives, for each grid cell, the ‘probability of development’, you need to come up with an **allocation procedure**. Do you bulldoze all wetlands or just a portion? Should this portion be decided by county or by neighborhood? Be clear on your allocation procedure – this can reflect your own personal values or development “cultures” you identify as reasonable or likely.

To summarize, the modeling procedure is as follows:

1. Wrangle data for time period 1 land use, infrastructure, census etc.,

2. Wrangle same data for time period 2

3. ID areas of land use change period 1 -> period 2

4. Use period 1 data to create a model ("the model") to predict land use change.

5. Feed period 2 data to the model to project for period 3 in the context of scenarios specified in the assignment

**Logistics**

You will do this project in teams of 2. You can sign up for your team here:

<https://docs.google.com/spreadsheets/d/1YcYd8Ldfe7ZBRch4hcREkS2IxNWSN8L_zLdVhT2Lp98/edit?usp=sharing>

If you cannot find a partner or want to do it alone, contact the professor.

**Deliverables:**

*Please make sure to name your documents as specified!*

1. **Pr**epare a planning memo communicating the planning rationale of the project, showing some exploratory maps/tables, presenting/discussing the data/feature engineering process, presenting model validation and results **in a 1000 word write-up with lots of data visualizations.** Please show your work (regression tables, confusion matrices /maps, etc.). You will be **judged on your ability to communicate this to planners.** The best approach is to relate ideas like development demand and allocation in a way that planners will understand. As always, bonus points for ggplot data visualizations/maps. Many of you should consider how this can be used as a portfolio item. Markdown is optional (but encouraged). Cross-validation (generalizability) tests are also optional.

Name this document:

teamMemberLastName1\_ teamMemberLastName2\_MEMO

1. Prepare an 8.5” x 11” ‘poster’ in pdf format that communicates the analysis. Use brief annotations, but try not to make it too wordy.

Name this document:

teamMemberLastName1\_ teamMemberLastName2\_POSTER.pdf

*The documents can be dropped in the following Box folder:*

https://upenn.app.box.com/f/8262a01c00fd47edbdfcaaae9fb5f07b

1. Find a region that isn’t totally built out. Make sure there is room for sprawl [↑](#footnote-ref-1)
2. Since we don’t have data on exactly 10-year intervals for land cover, and our most recent land data are 2019 - use the next best thing for the ancestral time period. You should use data from 2008 or 2011, and from the most recent survey in 2019. You can calculate change on that interval using raster calculator or map algebra in R. Find the data here: https://www.mrlc.gov/data?f%5B0%5D=category%3ALand%20Cover [↑](#footnote-ref-2)
3. There are ACS census data available from 2009 and 2019. Use the get\_acs function in tidycensus [↑](#footnote-ref-3)
4. You have to come up with an algorithm to do “allocate”, but you can find an example in the markdown, Section 5.2. [↑](#footnote-ref-4)