Data Science for Smart Cities

CE88

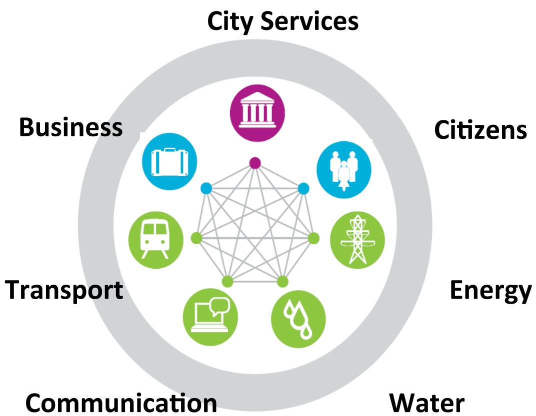
Prof: Alexei Pozdnukhov
GSI: Madeleine Sheehan

115 McLaughlin Hall

alexeip@berkeley.edu
m.sheehan@berkeley.edu
CE88 in title

Where we are in the course

Weeks 1-2. Introduction and motivation: cities as complex systems.



Lecture 1. Introduction to urban systems. Inter-dependent infrastructures with human in the loop.
Lecture 2. Modeling principles. Causality.
Lecture 3. Spatio-temporal nature of urban data.
Lecture 4. Data flows in cities

Weeks 3-5. Urban data collection, handling and processing.

Lecture 5. Data acquisition: measurement and crowd-sensing.
Lecture 6. Community surveys, population census, APIs.

Lecture 7. Demand data exploration.
Lecture 8. Supply data exploration.

Where we are in the course

Cities are complex systems: we take a holistic approach

↓

Urban data collection, handling and processing.

↓ ← **We are here!**

Data exploration and analysis. Demand, supply and impact.

↓

Modeling and forecasting.

↓

Decision making, planning and governance.

N.B. Good data science practices are:

Problem to solve ➔ Data ➔ Analysis method ➔ Solution

or

Data ➔ Analysis method ➔ Insights

… and not the other way around!

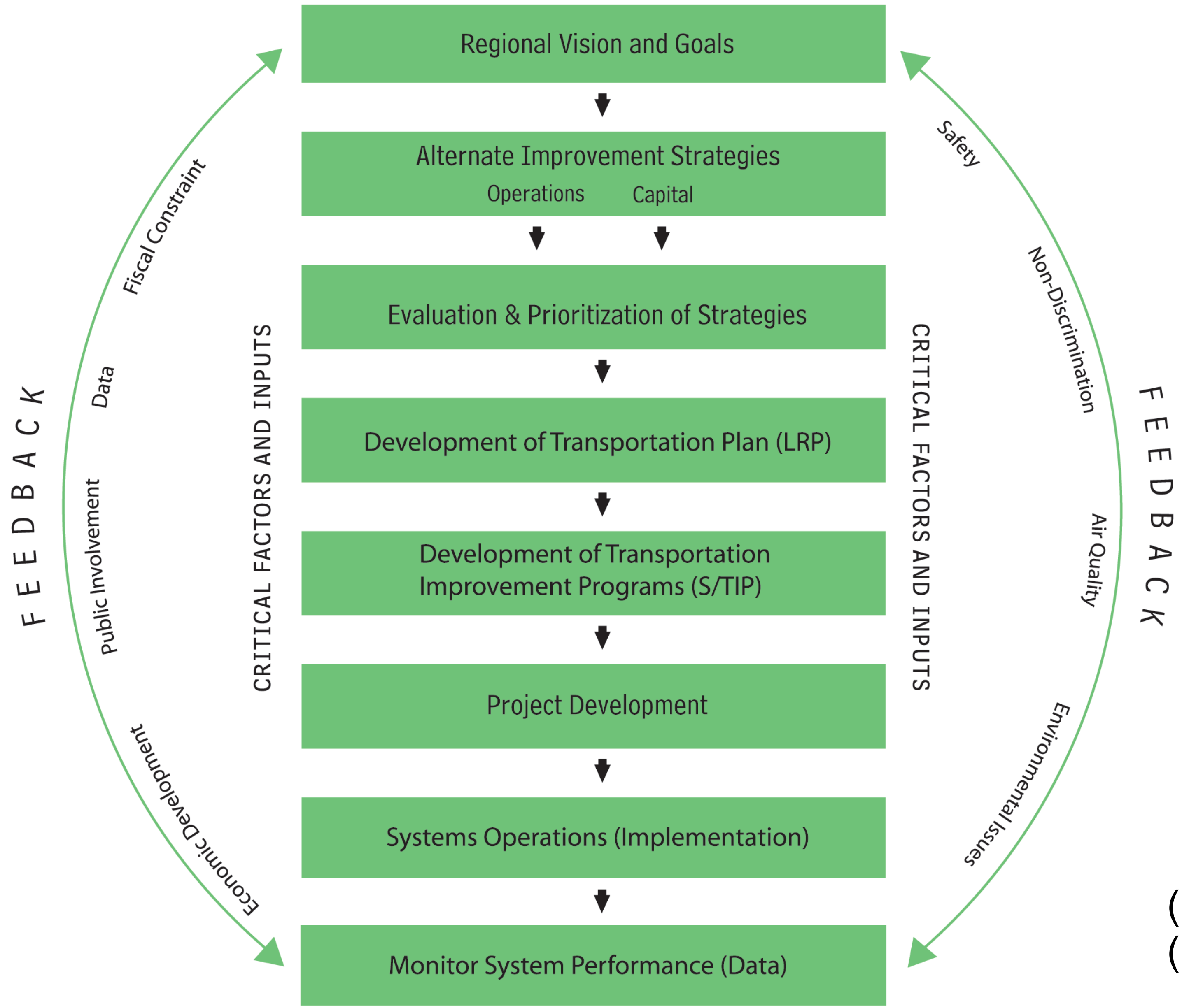
Transportation and urban planning

Data science methods provide tools, inform decisions and support models:

- Better planning tools to help urban transportation planners understand the impact of their decisions on the transportation network and the natural and human environment.
- Better decision support tools for communities to help them tackle land use, community development, economic development, and environmental protection challenges.
- Decision support and visualization to assist planners with conveying information to all the stakeholders to encourage successful community design and informed decision-making.

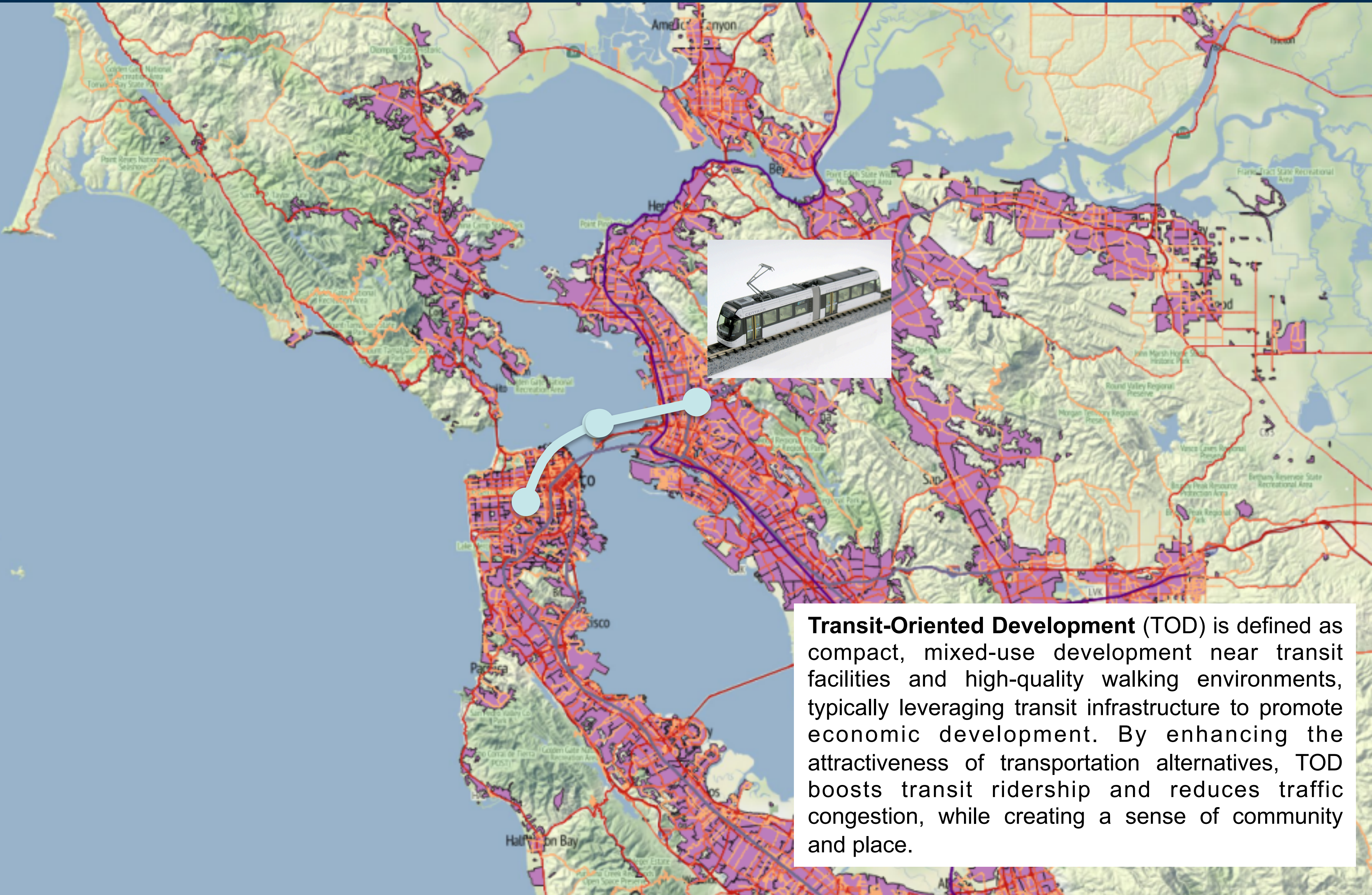
Examples of data-rich planning tools include transportation models, land use models, GIS, interactive maps and satellite imagery, as well as various specialized scenario planning and evaluation models.

Regional development context



(c) FHWA
(c) FTA

New transit line: an example



Transit-Oriented Development (TOD) is defined as compact, mixed-use development near transit facilities and high-quality walking environments, typically leveraging transit infrastructure to promote economic development. By enhancing the attractiveness of transportation alternatives, TOD boosts transit ridership and reduces traffic congestion, while creating a sense of community and place.

New transit line: step 1

Let us be mindful of the regional and local context, and find out: what are the current characteristics of the population and existing transportation options?

Step 1. Describe the system in terms of **explanatory variables**:	Data source
Socio-economic characteristics of potential passengers: - total population at travel origin, car ownership - employed/occupied population, income levels, age - intended destinations of travel	Census, surveys (traditional) APIs & crowd-sensing (emerging)
Level-of-service variables - accessibility of transit, travel times - driving times (including delays due to congestion)	Routing services, online maps, APIs
System parameters and policy variables - gas prices, tolls, parking fees - transit fare - taxi/uber/lyft/your favorite TNC fare	Regional transportation agencies, APIs, local sources

New transit line: step 2

Step 2. Connect explanatory variables with travel behaviors.

Travelers make particular choices on how to get to their desired destinations. Consider mode choice in a daily commute to work: will one drive (or carpool), bike, take transit, call uber? The outcome of such decision process depends on a particular context within which such decisions are made. We have described decision makers and the context with a set of explanatory variables.

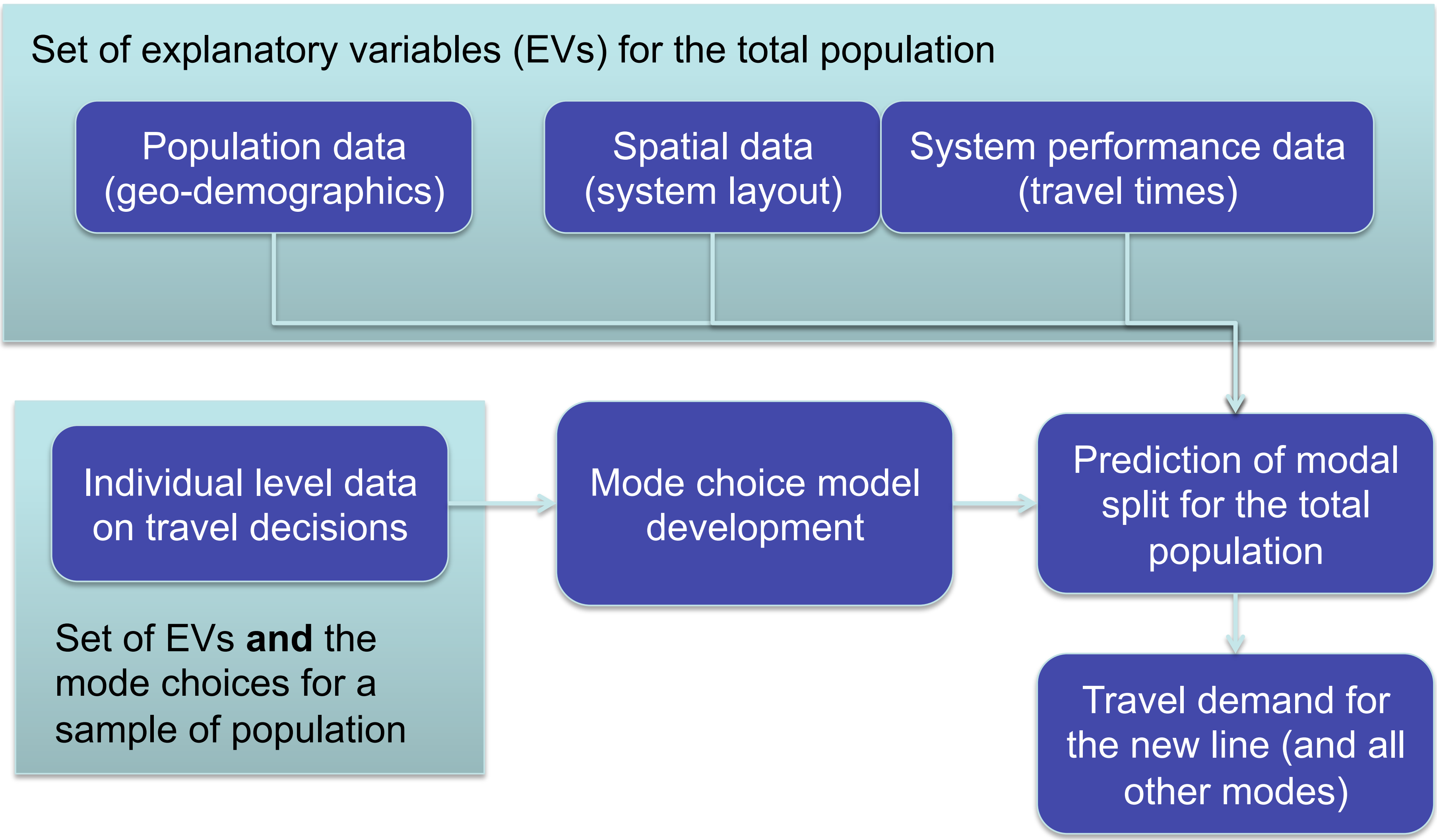
Since we would like to understand and predict demand on a new transit line, we have to be able to predict choices of individual travelers. If we had an algorithm (a model) that accurately predicts these choices for most (better yet, all) travelers in all the variety of context situations, we could apply it to the total population to get the desired demand forecast.

To build such an algorithm, it would require micro-level data, i.e. an empirical evidence of individual choices made by a large and diverse group of travellers.

Such micro-level data are difficult to collect. It requires extensive travel surveys, either manually conducted or based on the crowd-sensing approaches via locational data collected from wearable sensors / smartphones.

Demand forecasting

When we accomplish steps 1 and 2, the modeling process would look like:

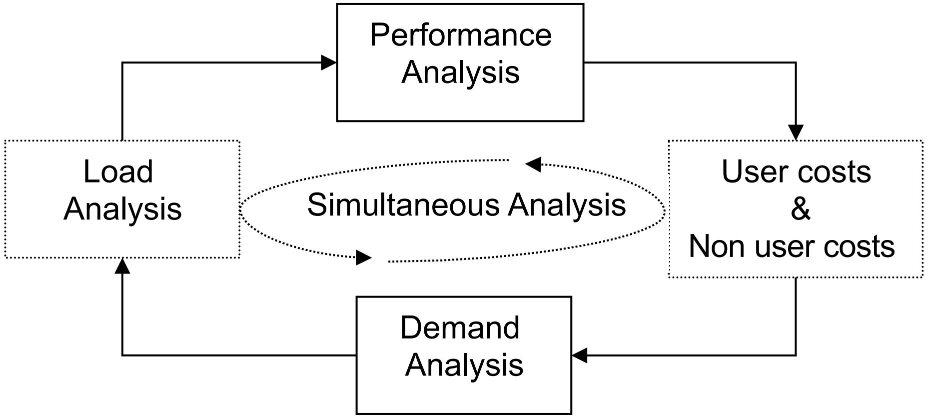


New transit line: step 3

One could stop after steps 1 and 2 above, however… Could the changes in demand have impact on the supply side of the system – leading to changes in performance variables and, consequently, have a feedback effect on demand?

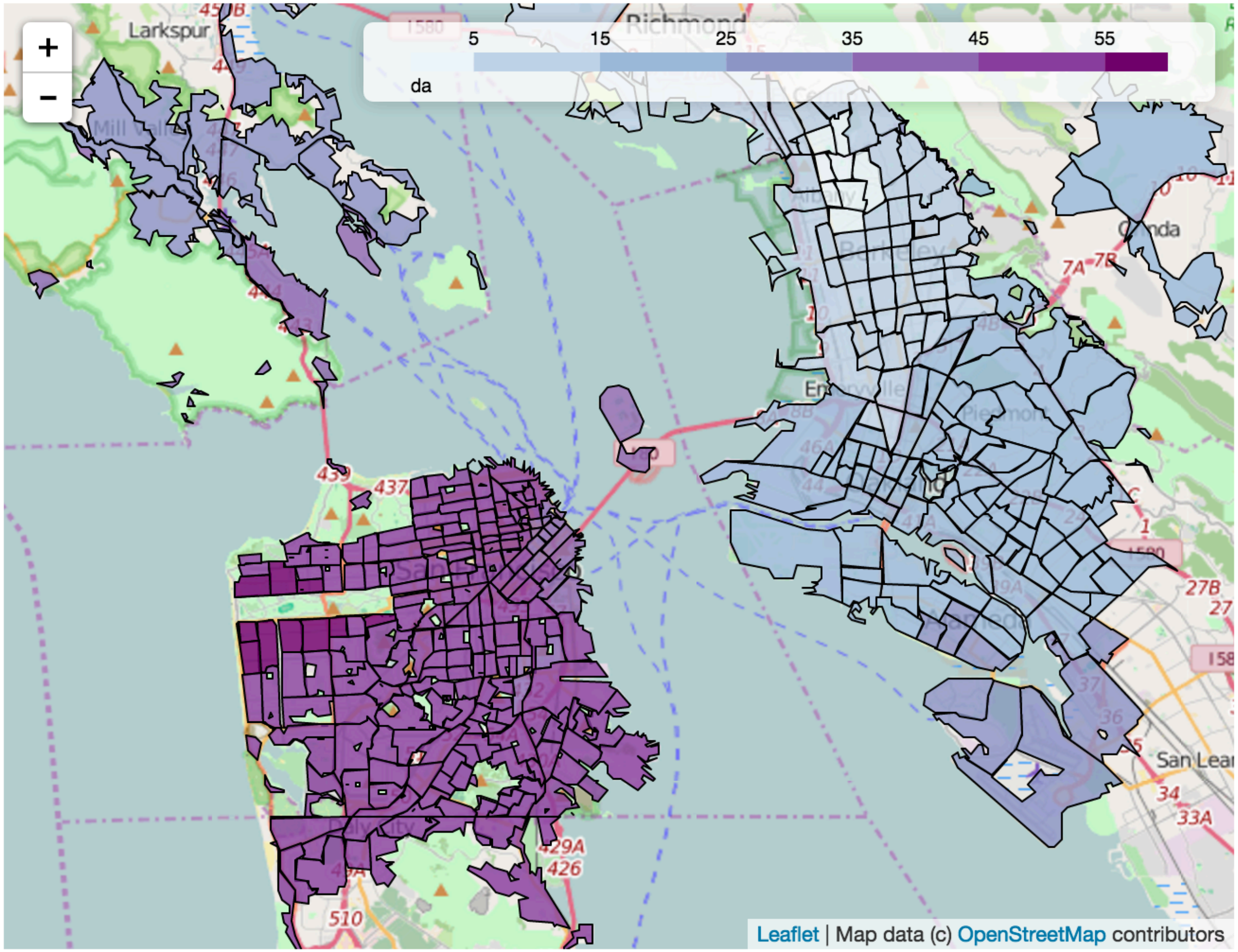
Step 3. Account for supply-demand feedback effects.

1. Short-term: shifting commute trips to a new transit line might decrease congestion and reduce travel delays, making driving more attractive, hence influencing modal split.
2. Long-term: citizens (as well as businesses) might relocate to more accessible locations. Accessibility of transit might also cause changes in housing prices that will change the demographics and, in turn, spatial pattern of travel demand.



Lab 3

Visualize and explore key socio-economic and level-of-service variables.



()