



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



Last time: data requirements

Describe the system in terms of **explanatory variables**

Socio-economic characteristics of potential passengers:

- total population at travel origin, car ownership
- employed/occupied population, income levels, age
- intended destinations of travel

Data source

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

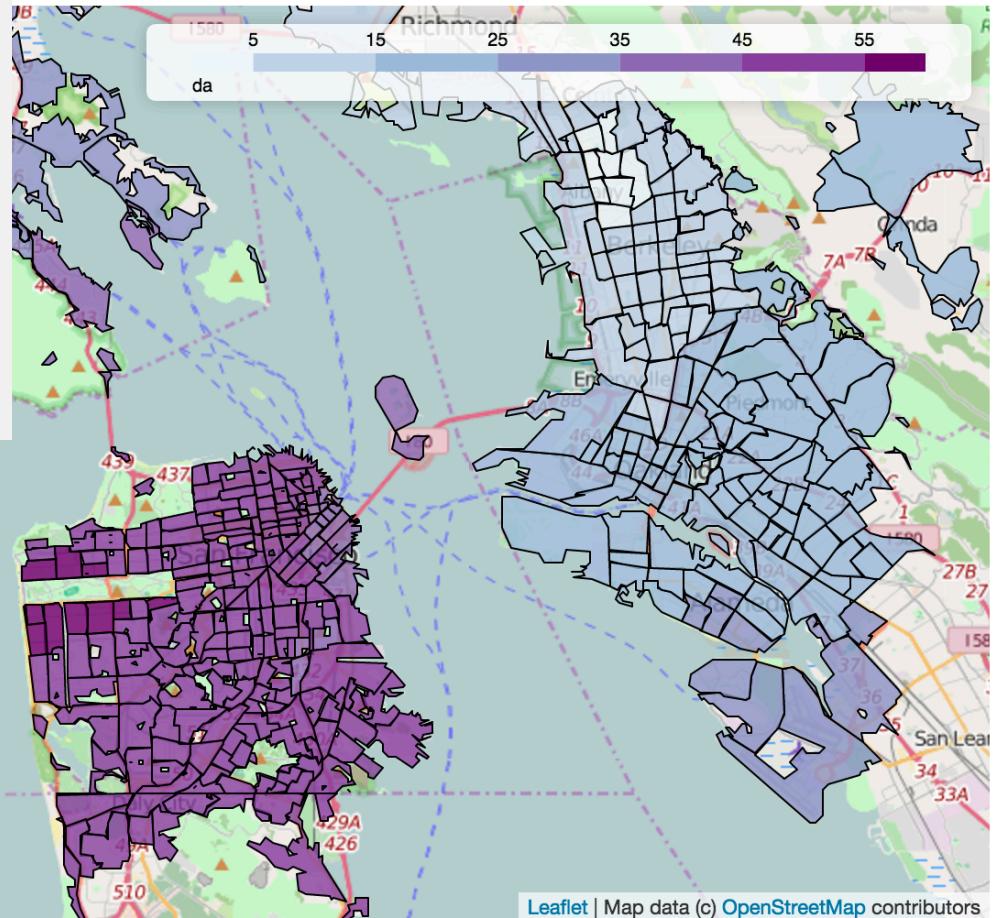
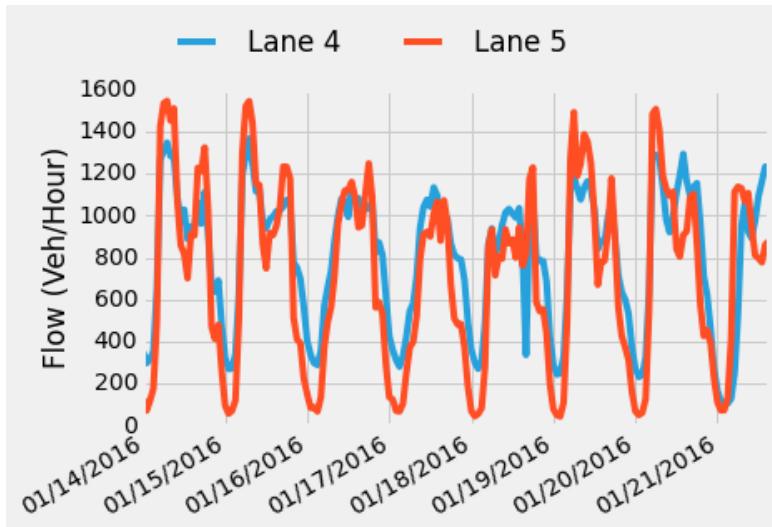
- gas prices, tolls, parking fees
- transit fare
- taxi/uber/lyft/your favorite TNC fare

Regional
transportation
agencies,
APIs, local sources

What we learned by now: labs and HWs



We have learned how to visualize and explore data in time and space!!



(Example: driving times from the East Bay, mins)



Today

9:10-9:30

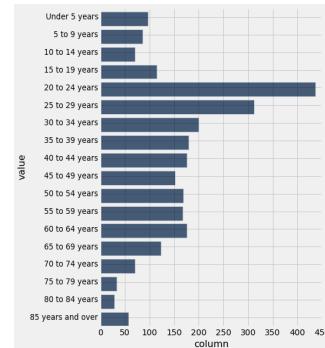
Review: urban mobility modeling

9:30-10:00

Practice simple operations on data tables
Explore socio-economic data with histograms
Mini Lab 4, Part 1



+



10:10-10:30

Impact: social and environmental. Metrics.

10:30-11:00

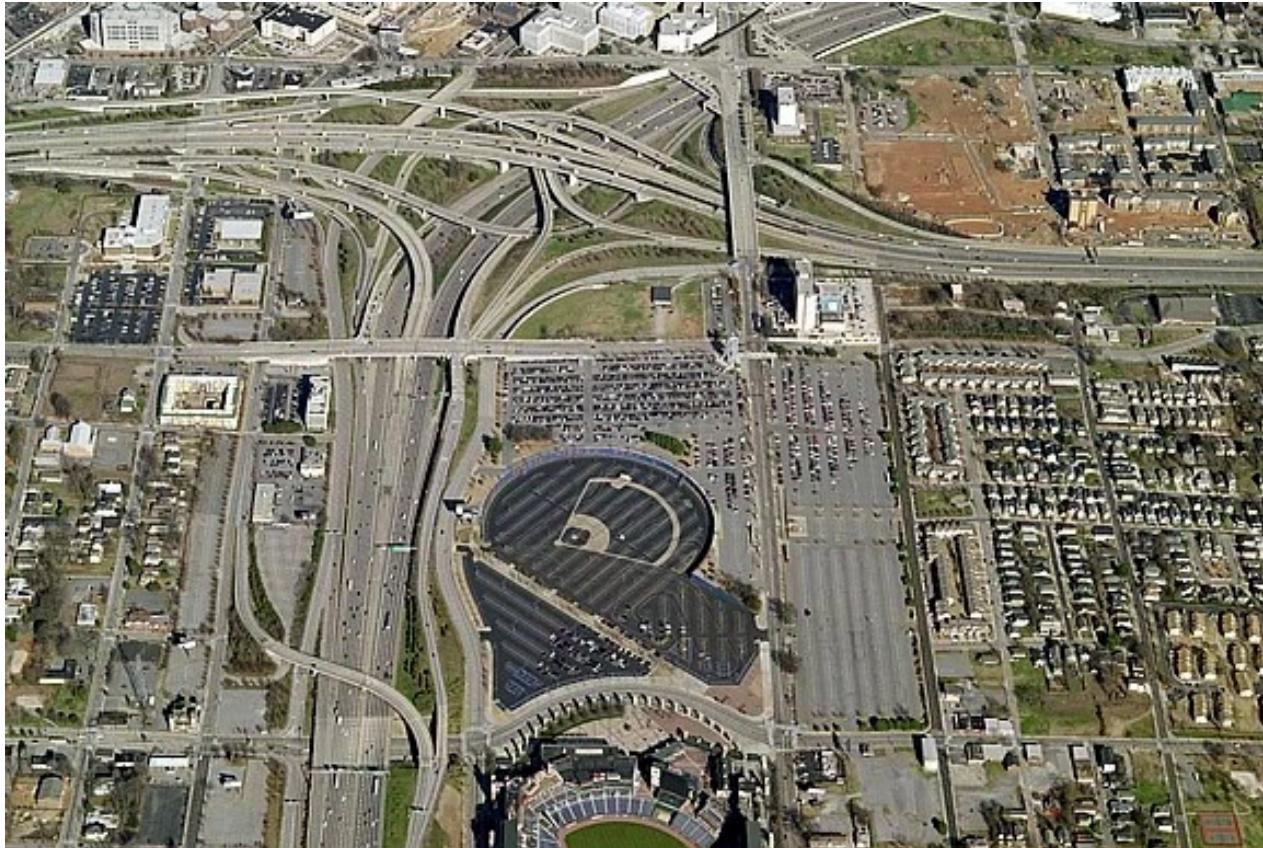
Practice advanced operations on data tables (joins)

Mini Lab 4, Part 2



Impact of urbanization

Transportation and urban planning



Common features:

- Single purpose zones
- Inwardly focused
- Creating barriers and buffers
- Isolating neighborhoods

THE DEATH
AND LIFE
OF GREAT
AMERICAN
CITIES

JANE JACOBS



Impact of urbanization

Environment

Disruption of animal habitats

- Air pollution
- Water pollution
- Noise
- Aesthetics

Resource consumption

- Land consumption
- Raw materials
- Energy consumption (fossil fuels)



Government agencies can do very little to make an environmental impact report entertaining. Each study, hundreds of pages long, meticulously analyzes the possible effect of a major construction project on everything from endangered species to air quality. **When printed out, the reports weigh enough to kill a small animal if dropped.**

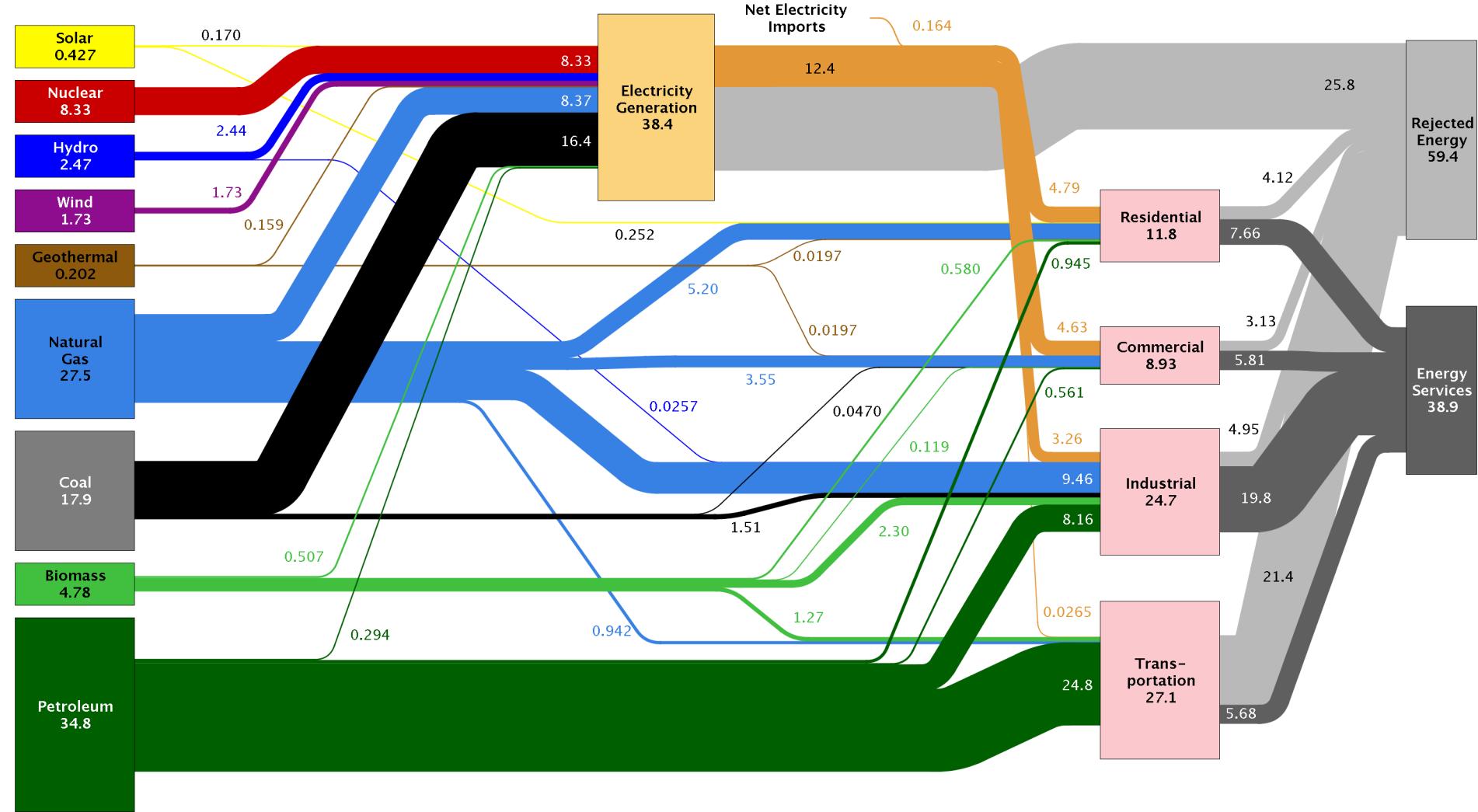
Source: <http://www.latimes.com/local/la-me-crazy-cover-20141010-story.html>

Impact of urbanization: energy use



Estimated U.S. Energy Use in 2014: ~98.3 Quads

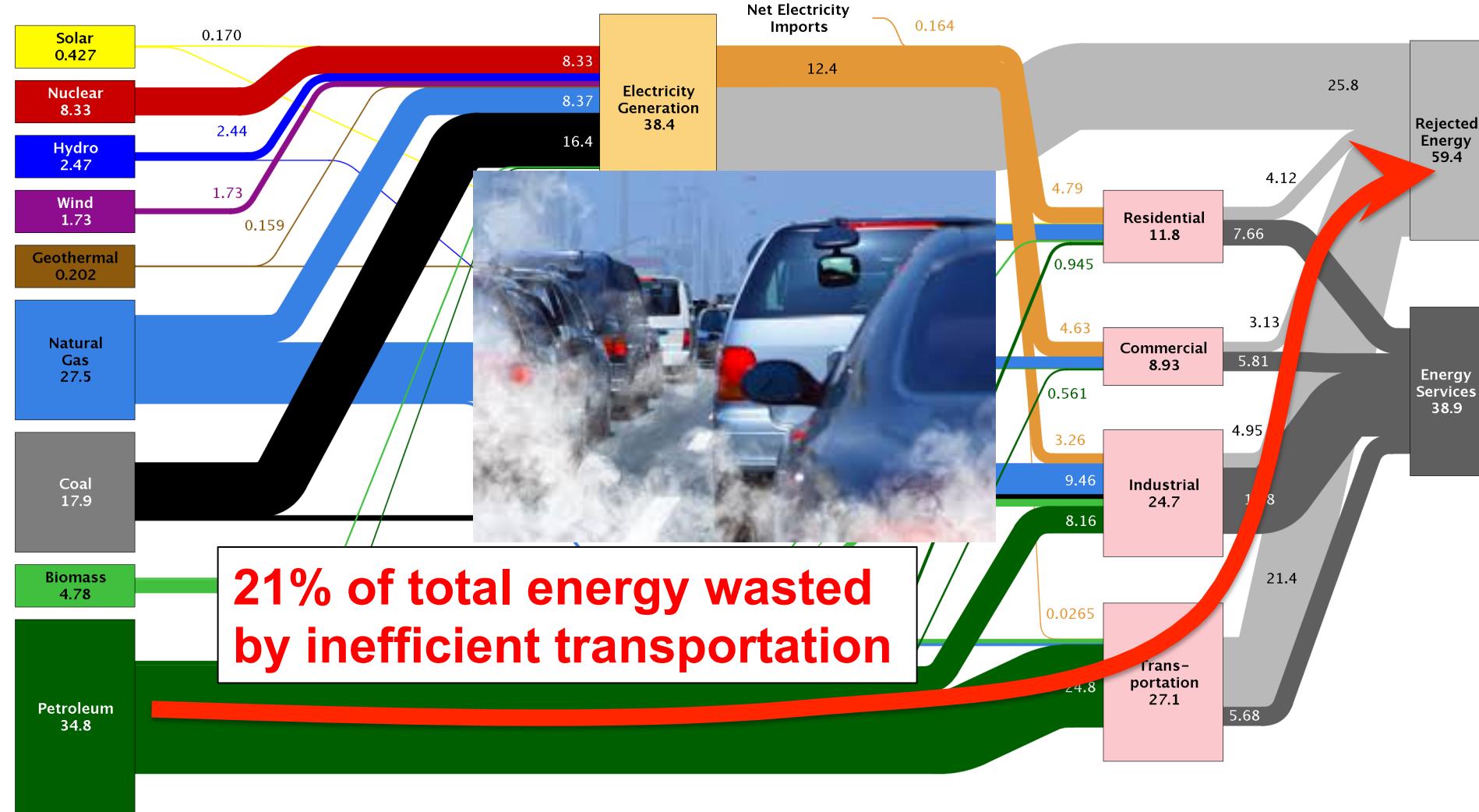
 Lawrence Livermore
National Laboratory



Source: LLNL 2015. Data is based on DOE/EIA-0035(2015-03), March, 2014. If this information or a reproduction of it is used, credit must be given to the Lawrence Livermore National Laboratory and the Department of Energy, under whose auspices the work was performed. Distributed electricity represents only retail electricity sales and does not include self-generation. EIA reports consumption of renewable resources (i.e., hydro, wind, geothermal and solar) for electricity in BTU-equivalent values by assuming a typical fossil fuel plant "heat rate." The efficiency of electricity production is calculated as the total retail electricity delivered divided by the primary energy input into electricity generation. End use efficiency is estimated as 65% for the residential and commercial sectors 80% for the industrial sector, and 21% for the transportation sector. Totals may not equal sum of components due to independent rounding. LLNL-MI-410527

Impact of urbanization: energy use

Estimated U.S. Energy Use in 2014: ~98.3 Quads



Source: LLNL 2015. Data is based on DOE/EIA-0035(2015-03), March, 2014. If this information or a reproduction of it is used, credit must be given to the Lawrence Livermore National Laboratory and the Department of Energy, under whose auspices the work was performed. Distributed electricity represents only retail electricity sales and does not include self-generation. EIA reports consumption of renewable resources (i.e., hydro, wind, geothermal and solar) for electricity in BTU-equivalent values by assuming a typical fossil fuel plant "heat rate." The efficiency of electricity production is calculated as the total retail electricity delivered divided by the primary energy input into electricity generation. End use efficiency is estimated as 65% for the residential and commercial sectors 80% for the industrial sector, and 21% for the transportation sector. Totals may not equal sum of components due to independent rounding. LLNL-MI-410527



Impact of transportation: metrics

Environmental impacts

Energy consumption

Air pollution

Noise



Vehicles-Miles Travelled (VMT)

Economic impacts

Time wasted



Vehicles-Hours Travelled (VHT)

What do we need to know in order to compute impact?

Travel demand

- traveling population, mode choice
- destination choice

Feedback effects

- delays due to congestion



Demand forecasting: mode choice

Set of explanatory variables (EVs) for the total population

Population data
(geo-demographics)

Spatial data
(system layout)

System performance data
(travel times)

Individual level data
on travel decisions

Mode choice model
development

Set of EVs **and** the
mode choices for a
sample of population

Prediction of modal
split for the total
population

Travel demand for
the new line (and all
other modes)



Destination choice

Set of explanatory variables (EVs) for the total population

Population data
(geo-demographics)

Spatial data
(system layout)

System performance data
(travel times)

Individual level data
on travel
destinations

Set of EVs **and**
destinations for a
sample of population

Destination choice
model

?

Origin-Destination
Flows

Aggregated models of destination choice

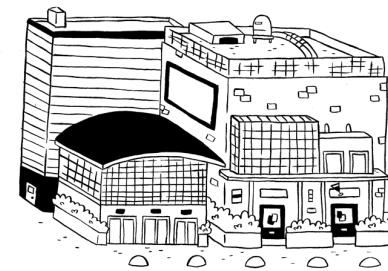
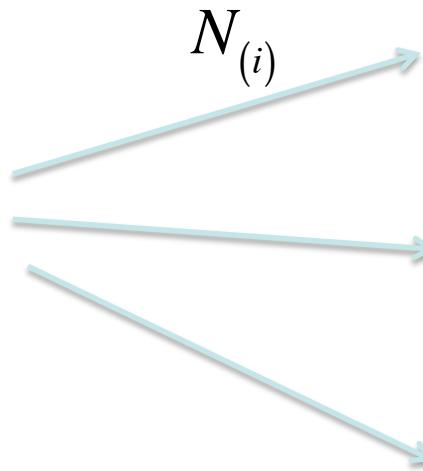


Since destination choice decisions are more difficult to infer and generalize to the total population, it is common to use simplified aggregated models to compute number of trips between origins and destinations.



$$N_{(i)} = \gamma P A_i f(t_i)$$

$$f(t_i) = t_i^\alpha e^{\beta t_i}$$



A_i



Today

9:10-9:30

Review: urban mobility modeling

9:30-10:00

Practice simple operations on data tables
Explore data with histograms
Mini Lab 4, Part 1.

10:10-10:30

Mode, destination choice and Impact

10:30-11:00

Practice advanced operations on data tables (joins)

Mini Lab 4, Part 2:

VHT & VMT computation

HW4:

**VHT & VMT with mode choice
and heterogeneous population**