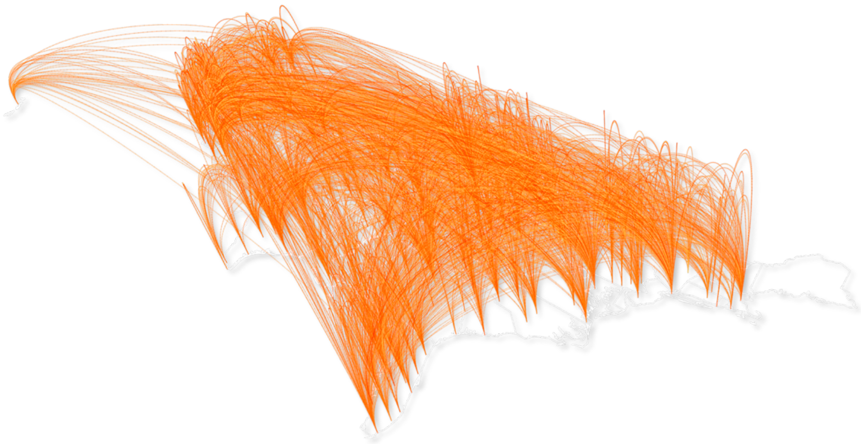
Collective models

Goal: generate/predict mobility flows between origins and destinations

Examples:

- Commuting flows

- Migration/relocation flows



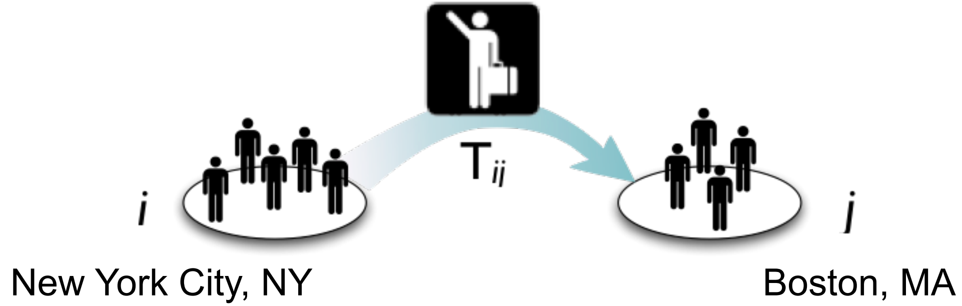
https://www.flaticon.com/authors/dave-gandy

Spatial flows and OD matrices

Mathematically, spatial flows are represented as a Origin-Destination **(OD) matrix**, **T**:

1. Define locations discretizing space, using a tessellation (e.g., counties, municipalities)

1. Element T_{ij} is the *number of trips from i to j per unit time*.



OD matrix

destination

origin

	a	b	c	d	e	f
a	-	3	27	2	1	0
b	1	-	4	0	0	5
c	8	3	-	1	13	6
d	2	1	5	-	0	2
e	11	0	6	5	-	1
f	0	3	2	2	0	-

(self-loops are usually not considered)

Total out-flow from i

$$\sum_j T_{ij} = O_i$$

Total in-flow to j

$$\sum_i T_{ij} = D_j$$

Total flow

$$\sum_{ij} T_{ij} = N$$

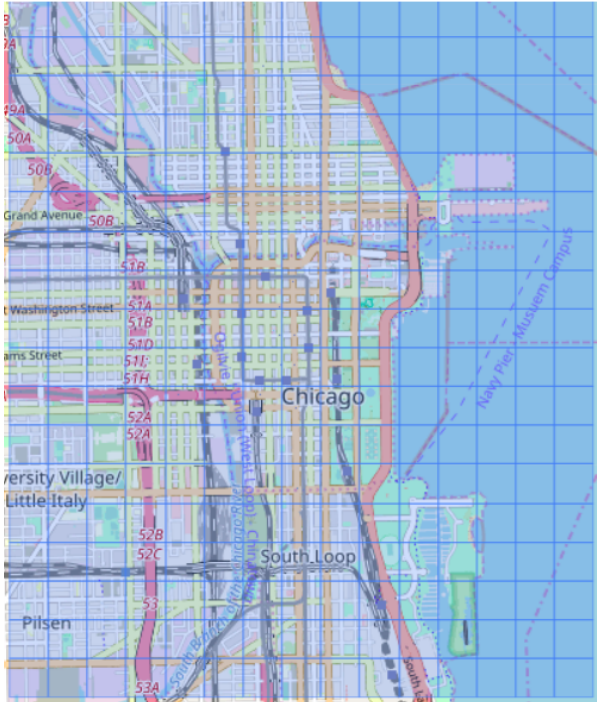
Probabilistic models of spatial flows

- The model assigns a **probability** to each possible OD-matrix **T**

- Methods to **fit** the model's parameters
    - <u>maximizing the likelihood of observed **T***</u>
    - minimizing the distance from observed **T***

Probabilistic models of spatial flows

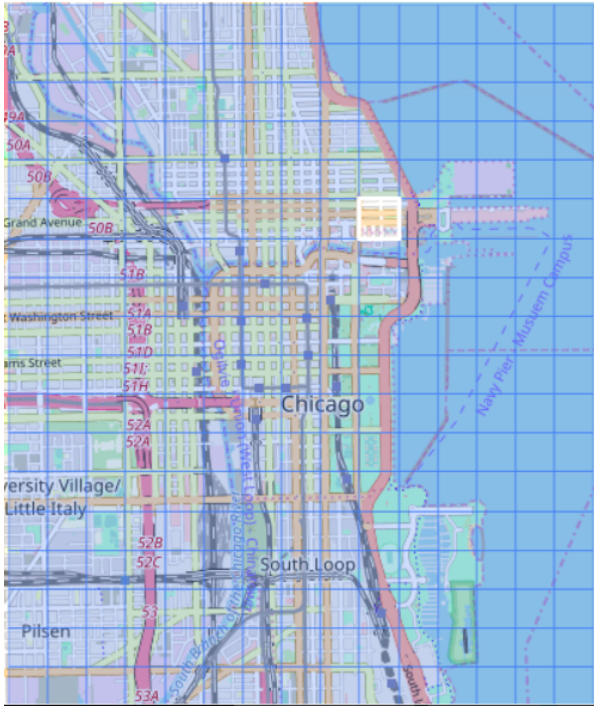
Idea: Interpret the problem as a classification task



classes = locations

Probabilistic models of spatial flows

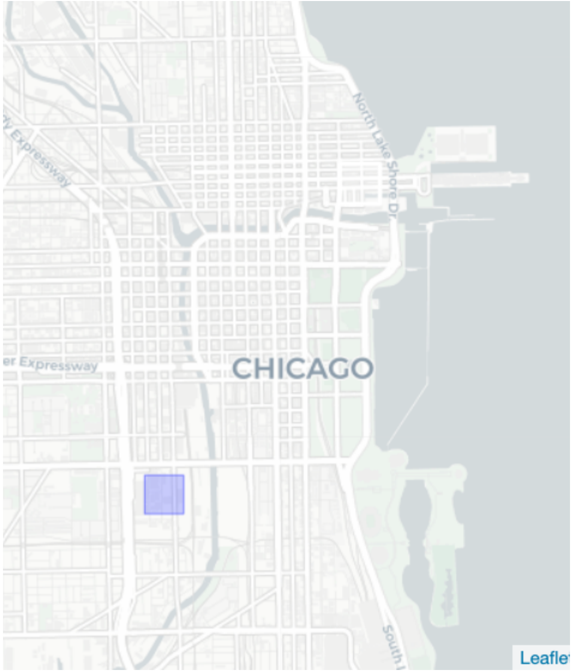
Idea: Interpret the problem as a classification task



given a trip's origin location, predict the destination

Probabilistic models of spatial flows

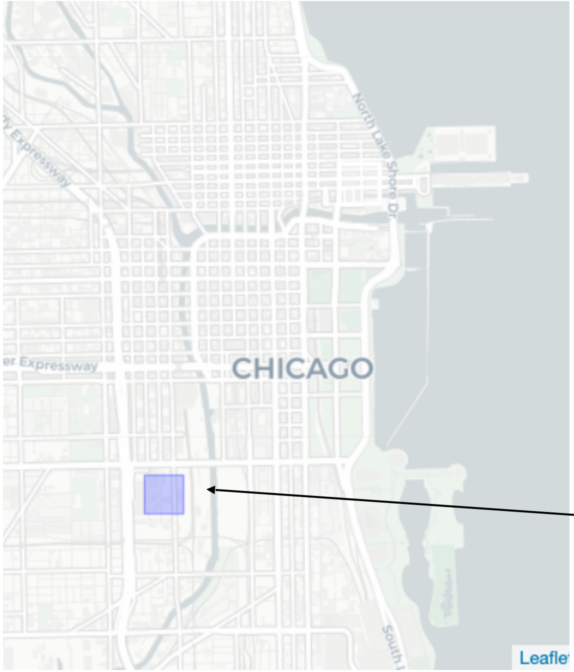
Goal: find the correct class (= location of destination)

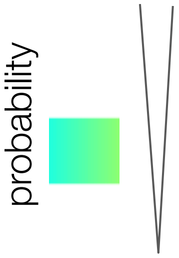


Each location has some probability to be the destination

How do we estimate these probabilities?

Gravity model



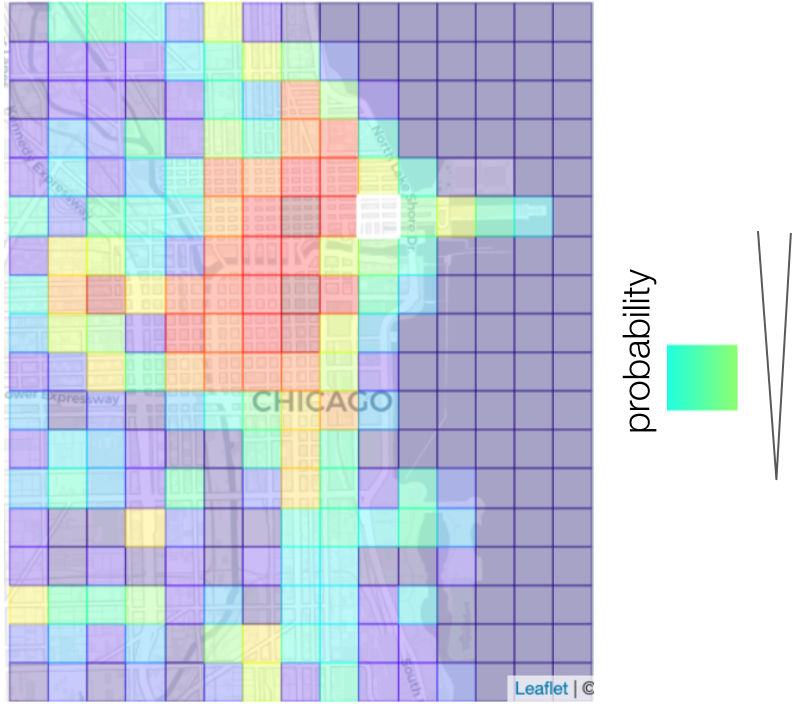


- Flows **decay** with **distance**
- Flows **grow** with **population**

The probability to observe a trip between *i* and *j* is large

$$p_{ij} \propto \frac{pop_j^{\beta}}{r_{ij}^{\gamma}}$$

Gravity model



Parameters β and y can be estimated using **maximum likelihood**

$$p_{ij} \propto \frac{pop_j^{\beta}}{r_{ij}^{\gamma}}$$

Validation of collective models

Comprehensive survey on distance/similarity measures between probability density functions. (Cha, S. H., 2007, City, 1(2))

Common metrics to compare OD matrices

- Sorensen-Dice similarity
  (Common part of commuters)

$$\frac{\sum_{ij} \min(T^e_{ij}, T^m_{ij})}{\sum_{ij} T^e_{ij}}$$

- Root Mean Squared Error

$$\sqrt{\frac{\sum_{ij} (T^e_{ij} - T^m_{ij})^2}{n^2}}$$

- More (cosine similarity, correlation, …)