Human mobility analysis and simulation with Python

Luca Pappalardo

@lucpappalard

Consiglio Nazionale delle Ricerche

Roberto Pellungrini

@pellungrobe

Università di Pisa

Gianni Barlacchi*

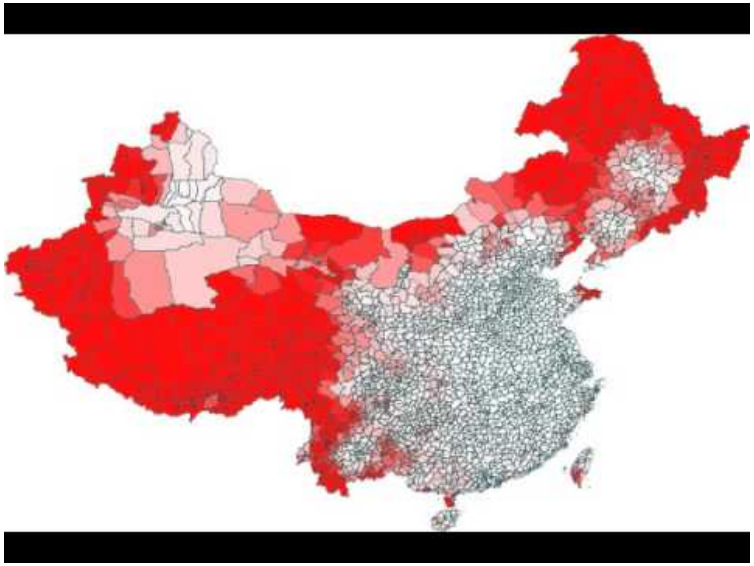
@GianniBarlacchi

Fondazione Bruno Kessler

scikit mobility

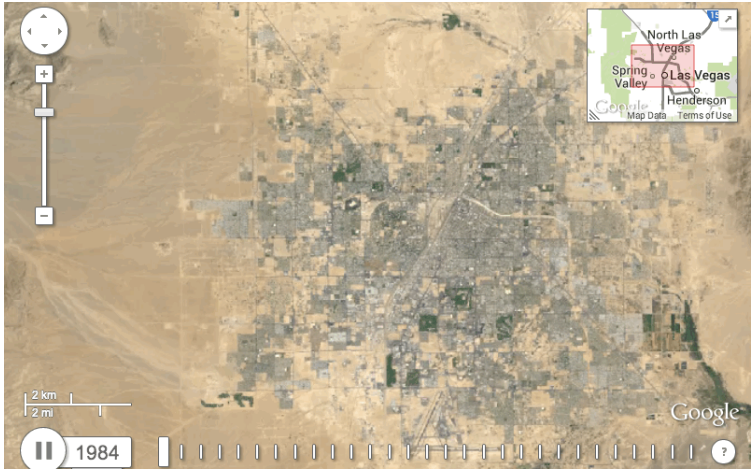
A M L D

*work done while the author was at Fondazione Bruno Kessler




Wembley
Olympic Park
Big Ben
Wimbledon
The rest of the city rises and leaves the house to begin a new day.
And our insight around driving behaviour increases second by second, minute by minute.


15:00:00


North Las Vegas
Spring Valley
Las Vegas
Henderson
Map Data
Terms of Use
2 km
2 mi
Google
1984

“*Human mobility* is the study that describes how humans move within a network or system.”

What you will learn today

Dealing with **mobility data**
loading, cleaning and preprocessing

Measuring mobility quantities
distances, predictability, mobility habits

Outline of the workshop

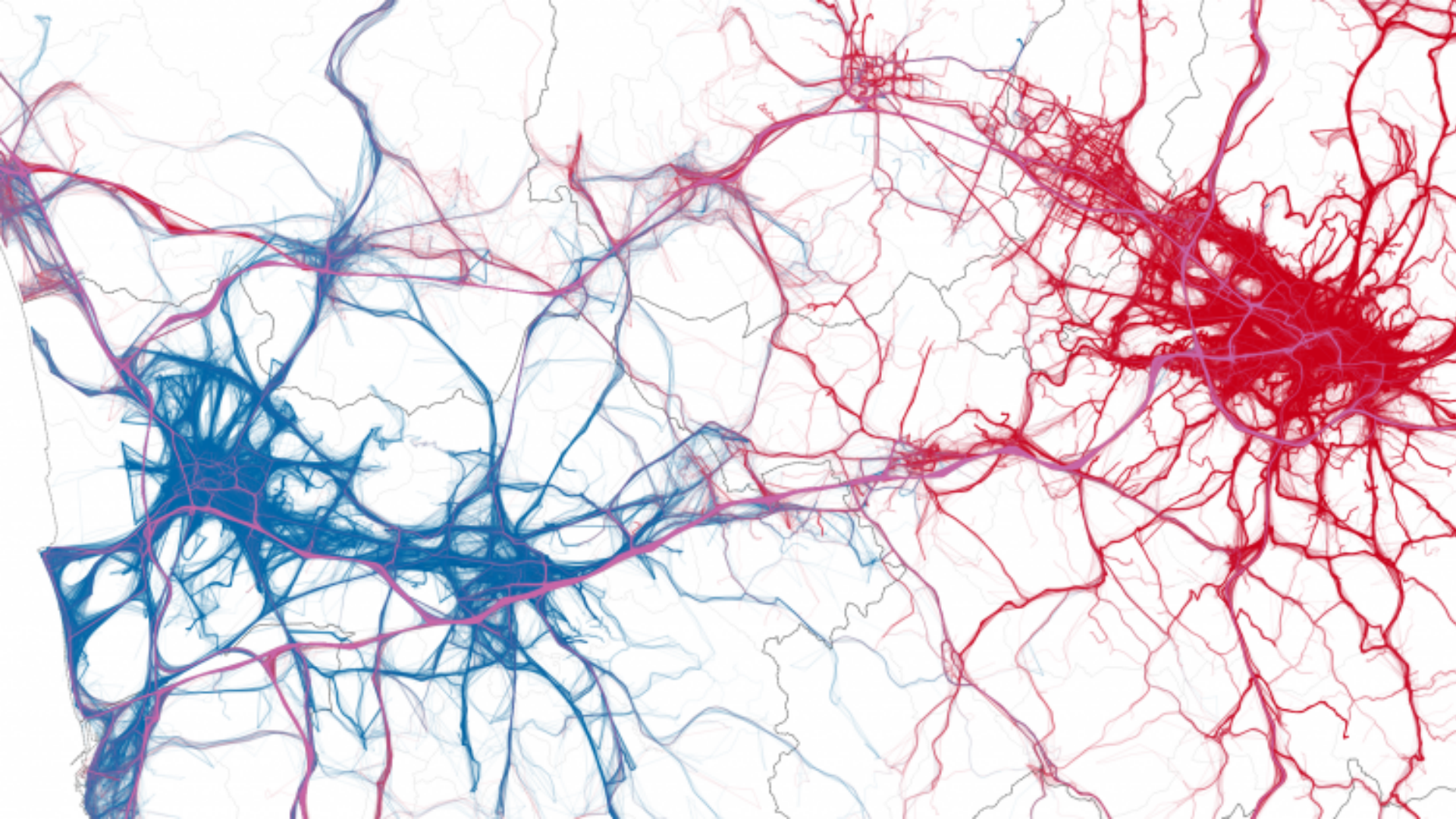
- **Digital breadcrumbs**: practical perspective on the mobility data landscape
- **Hows and whys**: examples of trajectories processing on real data
- **Under the microscope**: measuring patterns
- **A data provider's perspective**: How risky is to release mobility data?

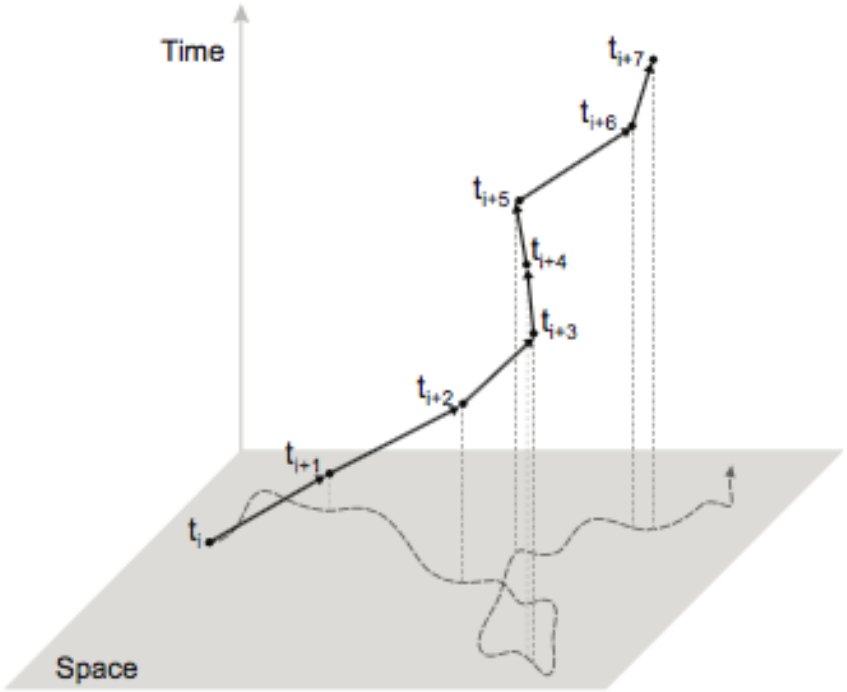
Two Applications

- Car Crash prediction
  how to exploit human mobility analysis to forecast whether a vehicle will crash

- Migration Flow prediction
  how to predict the flows of people between locations on a city/region/country

Representing mobility data



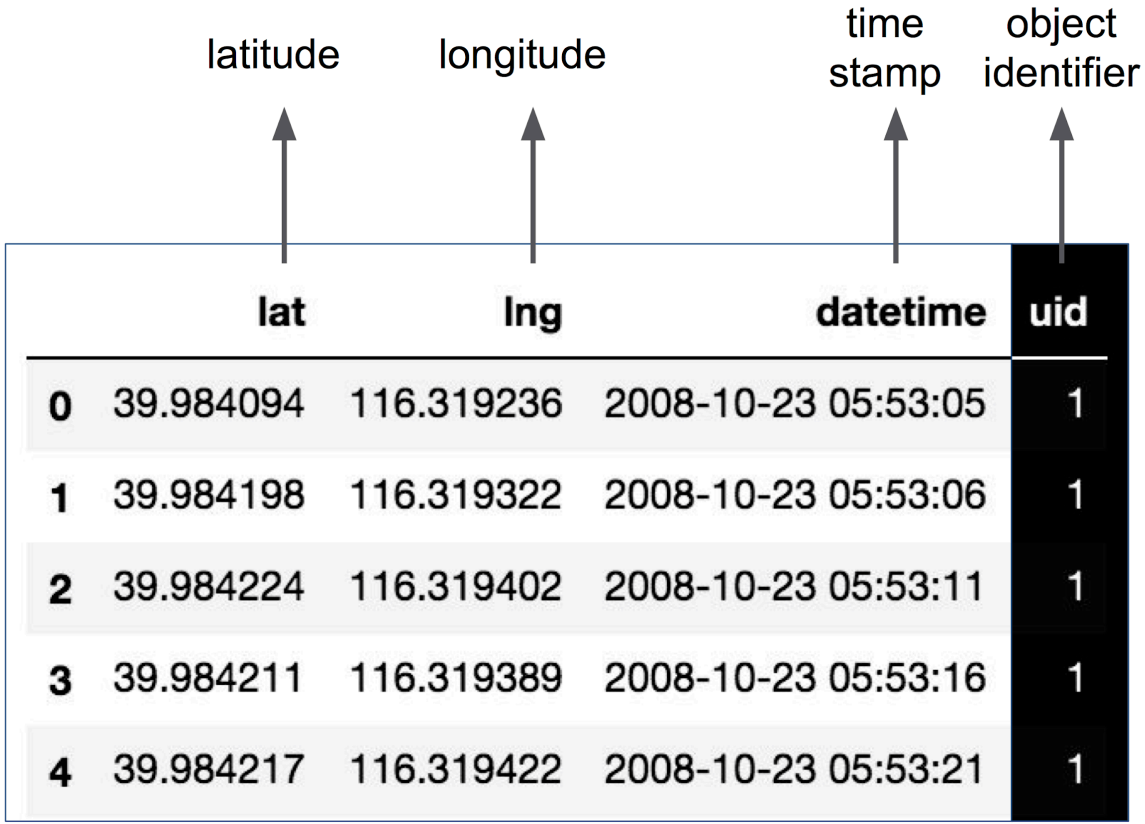


Trajectory. The trajectory of an object *u* is a temporally ordered sequence of tuples

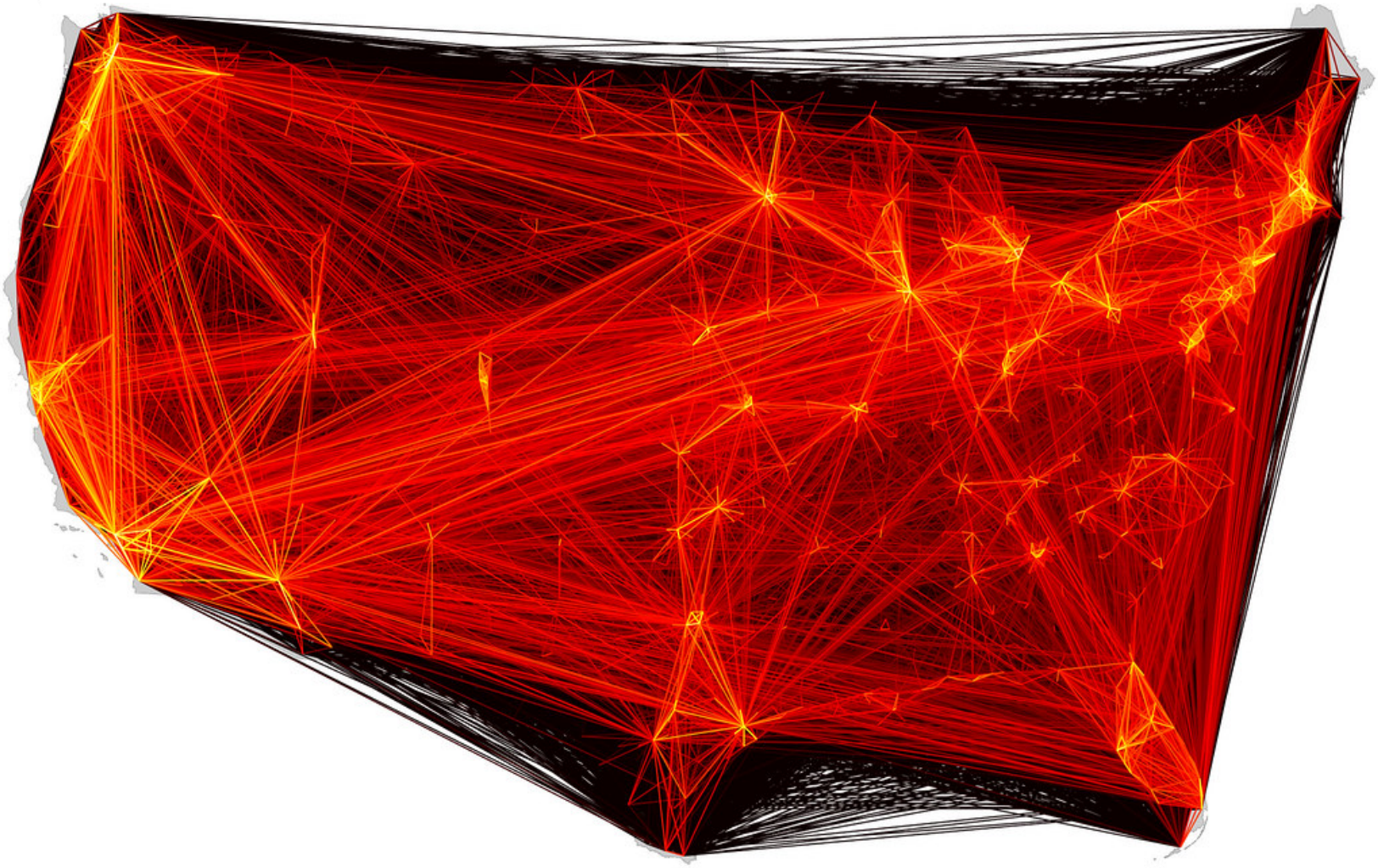
$$T_u = \langle (l_1, t_1), (l_2, t_2), \ldots, (l_n, t_n) \rangle$$

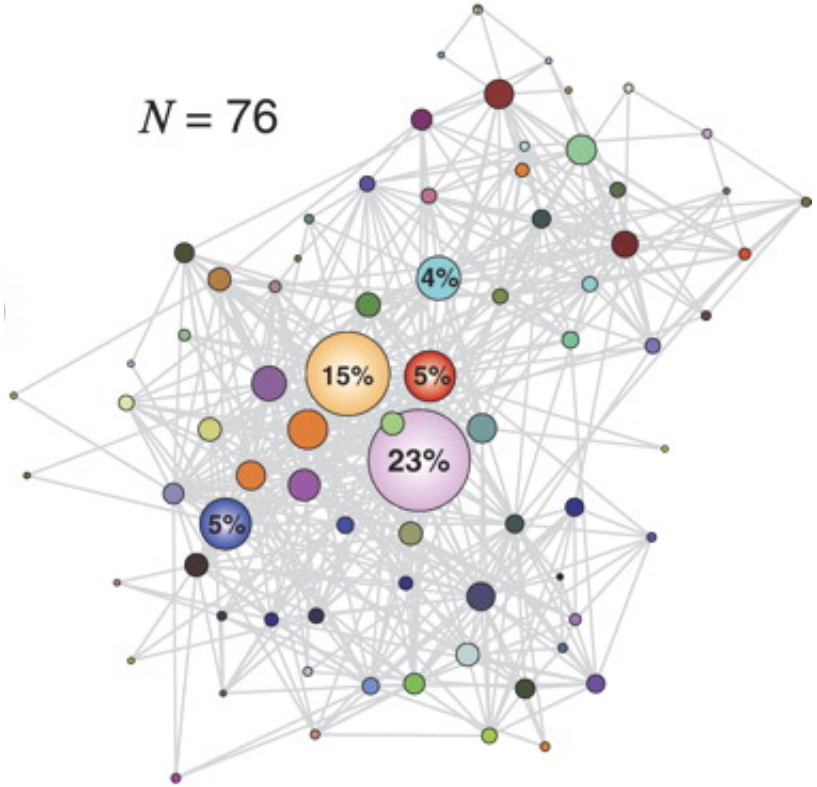
where $l_i = (x_i, y_i)$ is a location, x_i and y_i are the coordinates of the location, and t_i is the corresponding timestamp, with $t_i < t_j$ if $i < j$.

scikit mobility



	lat	lng	datetime	uid
0	39.984094	116.319236	2008-10-23 05:53:05	1
1	39.984198	116.319322	2008-10-23 05:53:06	1
2	39.984224	116.319402	2008-10-23 05:53:11	1
3	39.984211	116.319389	2008-10-23 05:53:16	1
4	39.984217	116.319422	2008-10-23 05:53:21	1





Flows. An Origin-Destination matrix T is a $n \times m$ matrix where:

- n is the number of distinct “origin” locations
- m is the number of distinct “destination” locations
- T_{ij} is the number of objects traveling from i to j.

scikit mobility

	flow	origin	destination
0	121606	36001	36001
1	5	36001	36005
2	29	36001	36007
3	11	36001	36017
4	30	36001	36019

	tile_id	population	geometry
51	36001	304564	POLYGON ((-73.933672 42.76071, -73.809603 42.7...
23	36003	48787	POLYGON ((-78.308611 42.086675, -78.3088390000...
18	36005	1397366	POLYGON ((-73.783519 40.881033, -73.74806 40.8...
57	36007	199346	POLYGON ((-75.850388 42.327731, -75.8437919999...
20	36009	79819	POLYGON ((-79.06070800000001 42.347917, -79.06...