



How cities live: Florence and the mobile analytics

Gianni DUGHERI[^], Laura GRASSINI^{*}, Alessandra PETRUCCI^{*}, Giorgio CECCHI^{*}

[^] Statistical Office, Comune di Firenze

* Dipartimento di Statistica, informatica, applicazioni (Università di Firenze)

PISA, 8-10 MAY 2018

WORKSHOP "SMALL AREA METHODS AND LIVING CONDITIONS INDICATORS IN EUROPEAN POVERTY STUDIES IN THE ERA OF DATA DELUGE AND BIG DATA"

FINAL EVENT OF THE JEAN MONNET CHAIR SAMPLEU

Outline of the presentation

- Mobile phone data: uses
- Evidence from a comparison between mobile phone data (Vodafone) and official statistical data (case study: Florence)
 - Resident population
 - Commuter flows to Florence
 - Temporary visitors
- Conclusions









Mobile phone tracking and city life

Many experiments

- Quantifying flows in a territory (for example a city)
- Identifying major flows
- Identifying cluster of users on the basis of calling habits
- Applications in tourism studies (14th Global Forum on Tourism Statistics, 2016)









Phone data from network probing systems

- Offer a fine temporal granulation
- Data at cell level → cell geographic areas covered by the same antennas
- Problems with area borders
- The activity of non-human-operated mobile devices must be removed









Common observational domains

	Official statistical data			
	ARCH.I.M.E.DE.	Accommodation stats (°)	Household surveys (*)	
	Residents	Tourists (at least	Tourists (at least one	
Mobile phone	Commuters	one overnight stay)	overnight stay)	
data			One-day visitors	

(°) professional accommodation facilities; (*) domestic flows









Usual environment

The concept of **usual environment** and, complementarily of **unusual environment** - typical of tourism statistics - is crucial for the use of mobile phone data (identification of temporary visitors, tourists, residents, commuters).

The usual environment of the subscriber requires time series data and interconnectivity between different data forms (i.e. same subscriber ID in domestic and outbound data, etc.).

The activities of anonymous subscribers should be followed over a long period to establish their usual environment (UN, 2017).









Definition of unusual environment in household surveys on travels and holidays

Coun	try	Distance threshold	Administrative borders	Respondents' self-evaluation	Frequency of visit	Duration (*)	
Austr	ia		Х	X	<2 per month		(*) one day visitors
Finlar	nd	30 Km			<once a<br="">week</once>	3 h	<i>Source:</i> EUROSTAT
Franc	e	100Km (*)		Х			
Germ	any		Х	Х			
Portu	igal		X(*)			3 h	
Spain			Х	Х			
Italy			Х		<once a<br="">week</once>		









Case study. ARCH.I.M.E.DE. project (ARCHivio Integrato di Microdati Economici e DEmografici)

ARCH.I.M.E.DE. project aims at providing elementary longitudinal and cross-sectional data useful for social and economic research, territorial and sector planning, for the evaluation of policies at national, regional and local level.

ARCH.I.M.E.DE. database is build through the massive use of archives produced by various public administrations in their institutional activity and integrating the information coming from the population registry.









Case study. Vodafone data (May-Sep 2016) Time detail: month/day each 6 hours

TWO DATASETS (estimated population data)

- 1- Origin destination matrix (unique subscribers; ACE census area)
- 2- Profiled data: unique presences (unique subscribers) and statistical presences (weighted with time of stay)

Italians (Italian SIM, regions of origin): residents, commuters, temporary visitors

Foreigners (Foreign SIM, country of origin)









Origin destination flows

- Vodafone records as flow, the number of Italian SIMs, that make a movement from one ACE to another. It is considered a movement only if the stay in the origin and in the destination is more than two hours.
- ARCH.I.M.E.DE. records as flow, the number of people living in a municipality, but with the work or study place in Florence (administrative archives).









Profiled data. Vodafone definitions

|--|



Residents: Florentine SIMs residing permanently in Florence in the month of survey, in the previous one and in the following one.



Commuters: SIMs resident outside Florence but who regularly come to Florence (at least 3 days a week and for 12 weeks, during the period of reference, May-September 2016).

Centro Camilo Dagum

ced Statistics for Equitable and Sustainable Development



Italian temporary visitors: Italian SIMs resident outside Florence that occasionally have been to Florence.



Foreign temporary visitors: SIMs resident outside Italy that occasionally have been to Florence.



Resident population 1/2

- Vodafone identifies as a resident each Florentine SIM residing permanently in Florence in the month of survey, in the previous one and in the following one.

- Vodafone uses the unique subscribes; this means that a person whose phone has been detected once within a cell of an ACE of the municipality of Florence enters the count.









Resident population 2/2

- With the Vodafone data it is also possible to distinguish the part of the population that spent the night in Florence from the rest.

- ISTAT data refers to the average population of May 2016.

	Vodafone*	ISTAT	
0 am-6 am	297840		•Unique subscribers.
0 am-12 pm	59619		•Working days
Total	357459	382615	
Centro Camilo Dagum	MAKIng Sustainable developmen	VELL	SMALL AREA METHODS FOR MONITORING OF POVERTY AND LIVING CONDITIONS IN EU Jean Monnet Chair Furgewar Commission

and WELL-being frameworks work for policy analysi



Commuters

- Vodafone identifies as commuters all the SIMs resident outside Florence but who regularly come to Florence (**at least 3 days a week and for 12 weeks, during the period of reference, May-September 2016**).

-ARCH.I.M.E.DE. records as commuters, the people living outside of Florence but with the work or study place in Florence (administrative archives). 2015 data.

Vodafone*	ARCHIMEDE
35479	156759

*Unique subscribers. Daily avarege, working days.









Temporary visitors

- The total number of Italian and foreign t.v. in Florence was calculated for May 2016 with Vodafone.
- Vodafone identifies as a t.v., each SIM that has been detected in a cell of a florentine ACE occasionally
- ISTAT provides data on night stays in an accommodation facility in the municipality of Florence during the reference period, May 2016.









May 2016 (total monthly data)

Vodafone*	Temporary visitors			
	Italian t.v.		Foreign t.v.	Total
0 am-6 am	1042368		969336	2011704
	Tuscanian	Non tuscanian		
	427914	614454		

* Unique subscribers.

SMALL AREA METHODS FOR MONITORING OF POVERTY AND LIVING CONDITIONS IN EU Jean Monnet Chair

	ISTAT			
		Italian	Foreign	Total
EMÆL	Night stays	182435	714189	896624
ans 1343	MATIS .	Centro Camilo Dagum Tuscan Interuniversity Centre Advanced Statistics for Equitable and Sustainable Development	MAKing Sustai and WELL-beir	KSWELL nable development ig frameworks work for policy analysis

Conclusions

- The natural target variable of mobile phone data is **mobility**.
- Integration with official domains should imply the use of similar operational definitions.
- Background information is related to the regulations of a given country
 →the applied data cleaning and related algorithms are a *black box*
- Problems in mobile phone market **selectivity** (infrastructure and individual)
- The modernization of infrastructure might introduce new problems (time comparability ?)
- Network probing system data more useful for unofficial domains: e.g. the number of unique subscribers on the site of attraction or concert (UN, 2017)









Main references

Ahas, R., Silm, S. Jarv, and E. Saluveer (2010). Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones. Journal of Urban Technology, 17(1).

EUROSTAT (2014). Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics. http://mobfs.positium.ee/data/uploads/reports/consolidated-report.pdf

Gabrielli, L., Furletti, B., Giannotti, F., Nanni, M., Rinzivillo, S. (2014). Use of Mobile Phone Data to Estimate Visitors Mobility Flows. In: Canal C., Idani A. (eds) Software Engineering and Formal Methods. SEFM 2014. Lecture Notes in Computer Science, vol 8938. Springer.

Raun, J. & Ahas, R. (2016). Defining usual environment with mobile tracking data. 14th Global Forum on Tourism Statistics, 23-25 November 2016, Venice, Italy.

Seynaeve, G. & Demunter, C. (2016). "When mobile network operators and statistical offices meet integrating mobile positioning data into the production process of tourism statistics2", 14th Global Forum on Tourism Statistics (Venice, Italy, 23-25 November 2016), <u>http://tsf2016venice.enit.it/images/articles/Papers_Forum/1.4%20Mobile%20-%20Eurostat-Proximus-StatisticsBelgium%2020111118.pdf</u>

UN (2017) Handbook on the use of Mobile Phone data for Official Statistics, UN Global Working Group on Big Data for Official Statistics, Nov. 2017.







