

Allegato 5

**Non importa che sia lontano.
Il potere attrattivo dei grandi
centri commerciali attraverso
l'analisi dei Big Data**

L'area fiorentina non gode di notorietà solo per il suo immenso patrimonio culturale: secondo il report presentato a Roma da Risposte Turismo in occasione dello *Shopping Tourism Italian Monitor* lo scorso novembre, il capoluogo toscano seguirebbe subito Milano nella classifica delle città italiane preferite dai turisti dello shopping, che generano un giro di affari del valore di 261 milioni di euro.

Negli ultimi decenni sono sorti grandi centri commerciali a servire sia le zone limitrofe ai comuni su cui sorgono, sia aree più lontane. Parchi del divertimento per gli *shopping addict* di ogni tipo – che si tratti di abbigliamento, generi alimentari, arredamento. Quando è poi esploso il fenomeno degli outlet, che da quasi due decenni a questa parte ha rivoluzionato i consumi andando a creare un proficuo mercato parallelo a quello dei negozi tradizionali, l'area del fiorentino non si è trovata impreparata. In poco più di dieci anni di attività, il Barberino Designer Outlet nel Mugello e il The Mall di Reggello – il primo con la sua vasta offerta di articoli per lo sport, la casa e il vestiario, l'altro dedicato agli amanti dell'alta moda – si sono consolidati tra le realtà che propongono prodotti a prezzo molto scontato, diventando una sosta tipica di chi si reca in Toscana per affari o per piacere.

Ma se volessimo restringere il campo a coloro che risiedono in Toscana e, più nello specifico, ai comuni che fanno parte di quella che è definita "cintura metropolitana fiorentina"? Se ci volessimo domandare chi sono, come si muovono, quali sono le abitudini di questi utilizzatori? Dovremmo affidarci unicamente ai sistemi più classici, quali indagini campionarie, censimenti e altri rilevamenti statistici che riguardano i tratti socio-demografici della popolazione o esistono oggi delle alternative per lo studio dei fenomeni appena esposti? Una cosa è certa: si tratta di una moltitudine che per raggiungere ciascuno di questi colossi dello shopping si muove su vie che disegnano il reticolo stradale urbano percorso quotidianamente da numerosi veicoli, confluendo e appesantendo le arterie del sistema stradale.

Nuovi strumenti per favorire la cooperazione territoriale

Spesso le metodologie tradizionali sono molto costose, o raccolgono dati in periodi di tempo troppo lontani dalle analisi: basti pensare ai censimenti ISTAT che avvengono ogni dieci anni, o altre indagini che riguardano campioni poco numerosi. Oggi, invece, si può fare affidamento ad altre: è diventato possibile condurre studi di mobilità attraverso l'analisi di grandi quantità di dati che ci riguardano, consentendo di registrare e studiare fatti sulle abitudini delle persone con un dettaglio impensabile fino a pochi anni fa.

È in questo contesto che si inserisce una collaborazione nata tra l'ISTI-CNR di Pisa e l'IRPET di Firenze. I due centri di ricerca hanno condotto uno studio usando dati GPS, dati utili a localizzare latitudine e longitudine di oggetti e persone, in grado di rilevare continuamente i dettagli spazio-temporali e ricostruirne la posizione geografica. L'utilizzo di questi nuovi metodi ha, infatti, realizzato una sorta di "microscopio sociale" che ha stimolato le ricerche in numerosi campi con significative applicazioni in diversi ambiti, dalla sanità pubblica all'ingegneria dei trasporti, la pianificazione urbana e la progettazione delle *smart cities*.

Possiamo anche pensare alle utili ricadute che questo tipo di studi può avere in ambito di programmazione territoriale. Per fare un esempio, sappiamo che da anni si sta cercando di attuare anche in Toscana l'istituto della *perequazione territoriale* – strumento applicabile, secondo la normativa vigente, sulla base di accordi volontari tra enti amministrativi. Si tratta di favorire forme

di cooperazione fra le istituzioni nelle varie fasi che compongono l'applicazione di nuove politiche urbanistiche, al fine di mitigare la diversa distribuzione di costi e benefici tra attori economici e territori derivanti da tali scelte.

In altre parole, stiamo parlando di meccanismi di redistribuzione delle risorse economiche diretti a ripartire nel modo più equo possibile le esternalità positive (quindi le entrate finanziarie dei bilanci comunali – oneri di urbanizzazione dovuti nell'insediamento di nuove realtà, gettito ICI) e le esternalità negative (costi collettivi di vario tipo, tra cui costi ambientali, per l'eccessivo consumo del suolo e per l'aumento della mobilità privata; costi pubblici riferiti all'erogazione dei servizi connessi; costi sociali causati da una prevalenza di insediamenti monofunzionali, etc) tra territori appartenenti ad amministrazioni diverse, ma nella sostanza integrate e interdipendenti sotto gli aspetti economici e funzionali. Passare dalla concorrenzialità alla cooperazione porterebbe benefici su vari livelli, da un uso più razionale del territorio alla possibilità di sfruttare economie di scala conseguenti alla messa in comune delle risorse.

È abbastanza chiara, a questo punto, l'importanza che il coordinamento degli enti locali andrebbe a rivestire proprio nell'ambito della definizione degli assetti urbanistici delle aree produttive e commerciali, il potenziamento e l'adeguamento delle infrastrutture volte a servirle, degli ulteriori interventi di riqualificazione che si potrebbero rendere necessari. Operare nell'ottica di uno sviluppo più razionale delle risorse armonizzerebbe i vantaggi che ne deriverebbero, ammortizzando allo stesso tempo gli svantaggi. Potrebbero essere, forse, proprio i *big data* della mobilità a convincere le istituzioni a dare una svolta alle pratiche tradizionali di progettazione urbanistica per attivare, finalmente, gli schemi di collaborazione previsti dal sistema della perequazione territoriale.

I centri commerciali

Tornando allo studio, nell'osservare gli scambi che avvengono tra i comuni del territorio toscano e, in particolare, della piana fiorentina, sono state indagate le abitudini di coloro che si recano presso sette dei più rilevanti centri commerciali dell'area. La scelta di questi centri commerciali ha riguardato un duplice criterio: volendo studiare attrattori che potessero interessare il più ampio bacino di utenza in termini di popolazione e relativa provenienza, si è guardato alle dimensioni (in particolare, ai centri commerciali "grandi", con superficie di vendita superiore ai 10.000 m²) e alla tipologia (ad esempio, il The Mall sopra citato ha dimensioni più contenute rispetto agli altri centri ma una portata attrattiva interessante determinata dal tipo di offerta che propone).

Centro Commerciale	Comune	Settore Merceologico	Superficie di vendita	Anno di apertura
I Gigli	Campi Bisenzio	Alimentare/Misto	43.980 m ²	1997
Il Parco	Calenzano	Alimentare/Misto	10.000 m ²	2003
Coop	Lastra a Signa	Alimentare/Misto	12.250 m ²	1997
IKEA	Sesto Fiorentino	Non Alimentare	21.222 m ²	2002
Centro Sesto	Sesto Fiorentino	Alimentare/Misto	20.030 m ²	2003
The Mall	Reggello	Non Alimentare	5.000 m ²	2001
Barberino Designer Outlet	Barberino del Mugello	Non Alimentare	27.000 m ²	2006

Principali caratteristiche degli attrattori commerciali studiati.

Come emerge dalla tabella, si tratta di centri commerciali accomunati da alcune caratteristiche e che, invece, rimangono piuttosto eterogenei tra loro se guardiamo al settore merceologico e alle

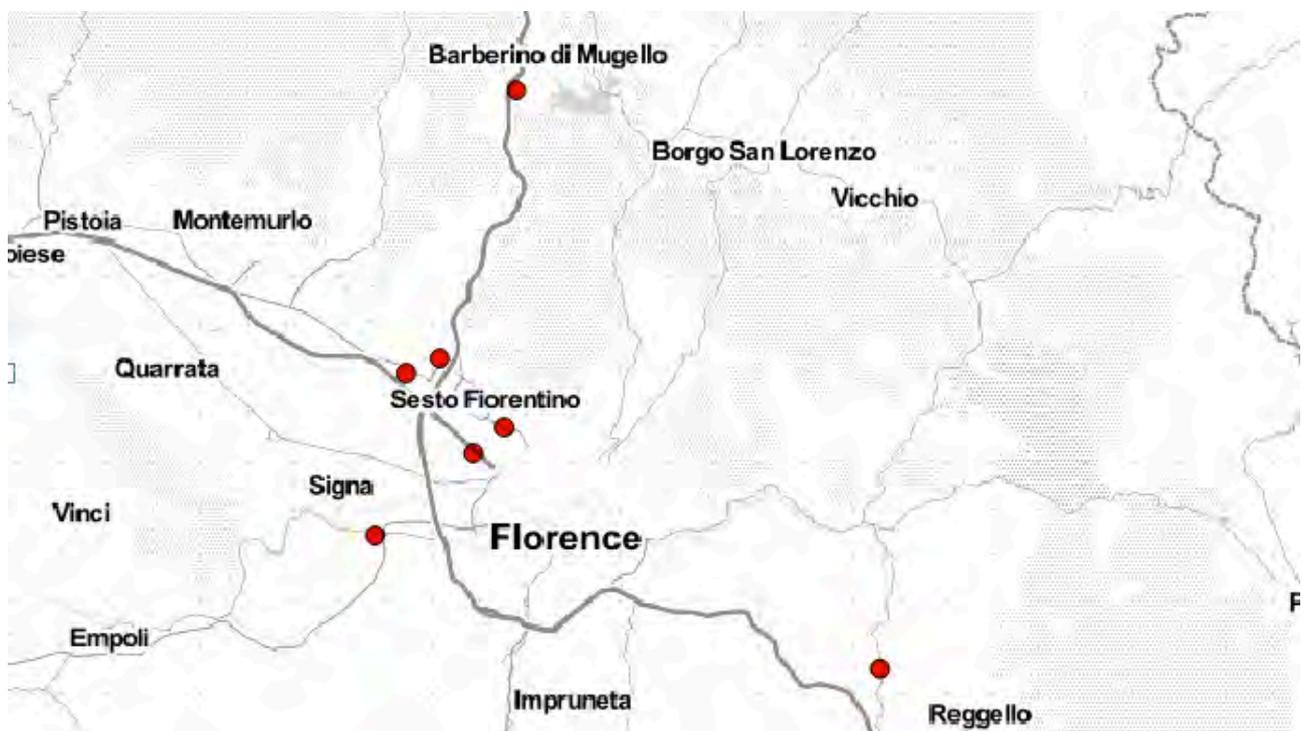
dimensioni. Da un lato i due già citati outlet di Barberino e Reggello; maggiori similitudini possono trovarsi fra tre dei grandi attrattori che compaiono nell'analisi: si tratta de Il Parco nel comune di Calenzano, la Coop di Lastra a Signa e Centro Sesto, situato a Sesto Fiorentino. In questo caso, i supermercati che ospitano questi centri sono il motivo prevalente per cui la maggior parte degli utenti si reca presso questi luoghi, supermercati che fanno comunque parte di gallerie commerciali popolate di negozi di vario genere.

Si trova nel caso opposto I Gigli, centro commerciale esteso e imponente quanto i tipici *mall* americani, che spinge la sua offerta soprattutto sul lato *retail* e della ristorazione, pur accogliendo a sua volta un grande supermercato.

Ancora, per quel che riguarda la proposta, fa caso a sé l'IKEA, senza dimenticare che il fatturato del colosso svedese dell'arredamento dipende non solo dalla vendita di mobili per la casa ma anche dal settore *food*, con una tendenza di anno in anno in forte crescita.

I dati

Per questa indagine, dunque, è stato utilizzato un dataset proveniente da una fornitura di GPS fornita da OCTO Telematics, società attiva nel mercato dell'*insurance telematics* che si occupa del tracciamento e della raccolta di dati riferiti a una flotta di veicoli tramite i dispositivi di navigazione satellitare installati a bordo delle vetture. In particolare, si tratta di mezzi privati che montano a bordo delle *black box* il cui compito è raccogliere i fatti relativi alle traiettorie degli utenti, così come previsto dai piani assicurativi a cui sono associate. Come anticipato, si tratta di dettagli spazio-temporali: possiamo dedurre la lunghezza totale della distanza percorsa tra due punti, il tempo impiegato, il giorno e l'ora in cui quel veicolo si è mosso. I dati a cui fa riferimento lo studio, inoltre, riguardano un campione di circa 250.000 veicoli che hanno transitato in Toscana tra gennaio e marzo 2014, per un arco di tempo complessivo di sette settimane.

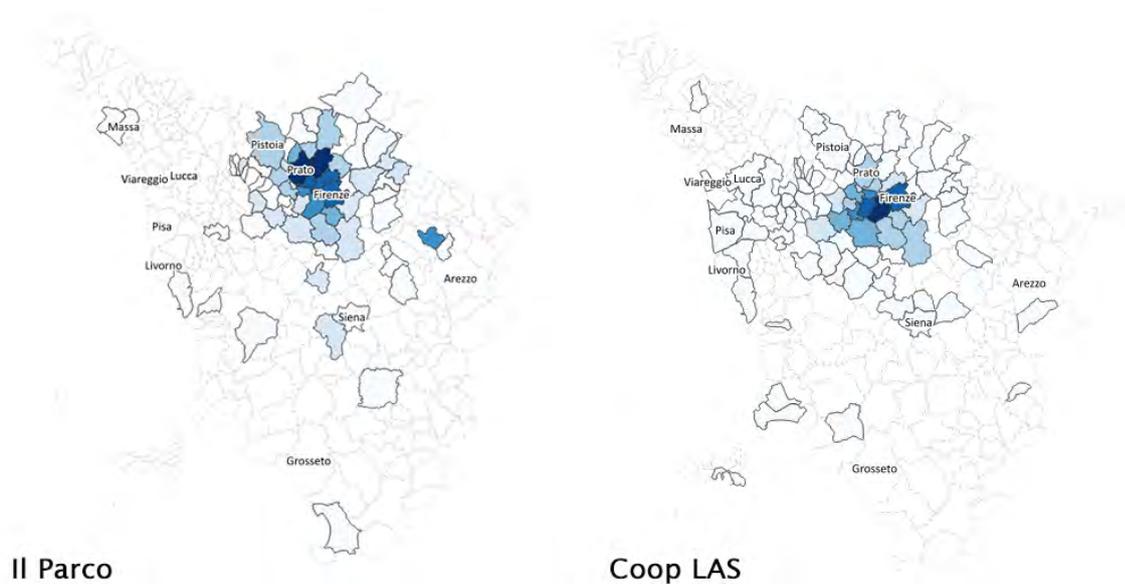


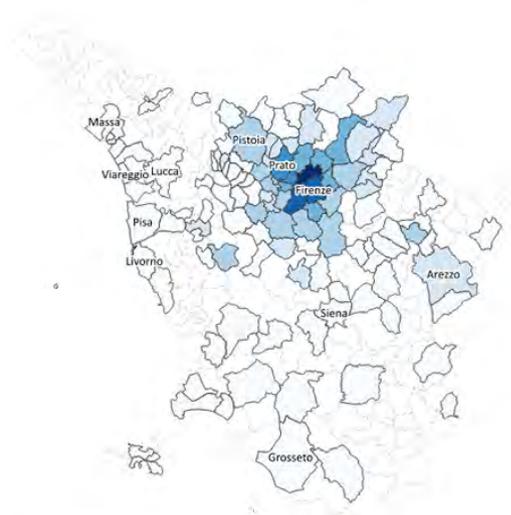
Localizzazione degli attrattori commerciali oggetto di studio

Dai dati grezzi, così come raccolti dalle *black box*, si possono però reperire altre informazioni. Raggrupparli, contarli, confrontarli: queste e altre manipolazioni hanno permesso di ottenere una nuova conoscenza sui luoghi generalmente più frequentati dai singoli utenti in modo da estrarre con buona approssimazione la zona di residenza; ancora, è stato possibile calcolare la sosta in un determinato luogo. Incrociando, poi, i dati delle scatole nere con le localizzazioni dei sette centri commerciali è stato possibile approfondire come ognuno degli utenti si serve di questi attrattori.

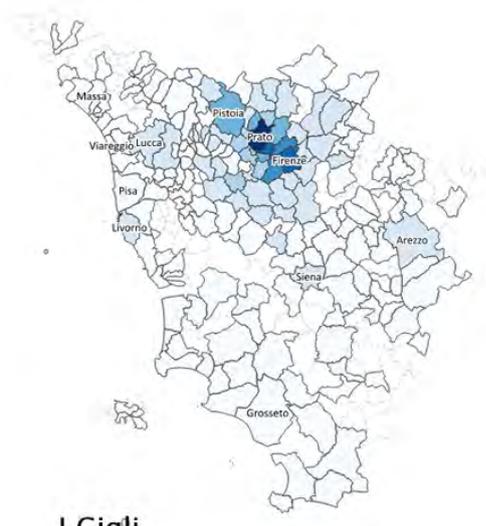
L'utenza dei grandi attrattori. Non solo clienti

È necessaria una premessa: i dati che sono stati analizzati, come anticipato, riguardano 250 mila veicoli, pari circa all'8% della popolazione veicolare circolante. Questi numeri, prima di essere rappresentati, sono stati ripesati per mezzo delle risultanze delle tabelle ACI relative alle vetture registrate in Toscana nei periodi cui si riferiscono. Partendo comunque da un campione, alcune aree – in particolar modo i comuni di dimensioni inferiori – possono risultare non rappresentate se hanno un dato di partenza pari a zero.

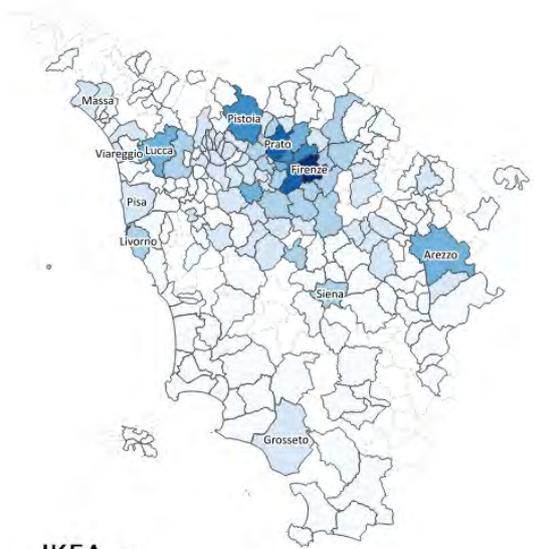




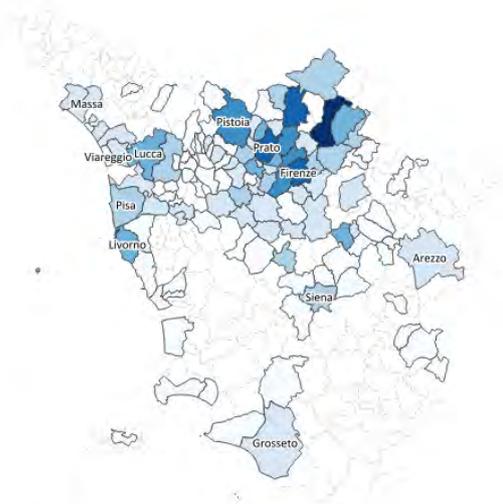
Centro Sesto



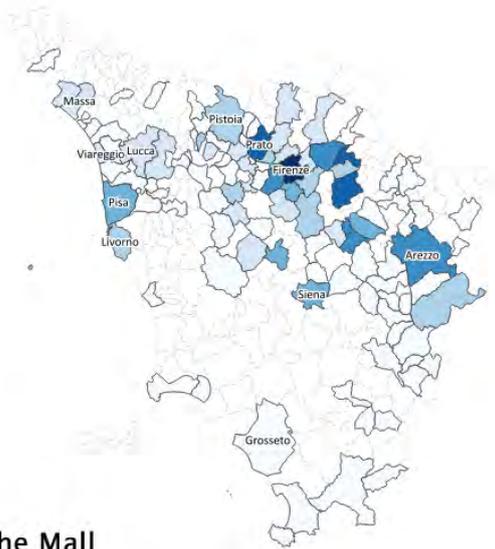
I Gigli



IKEA



Barberino Outlet



The Mall

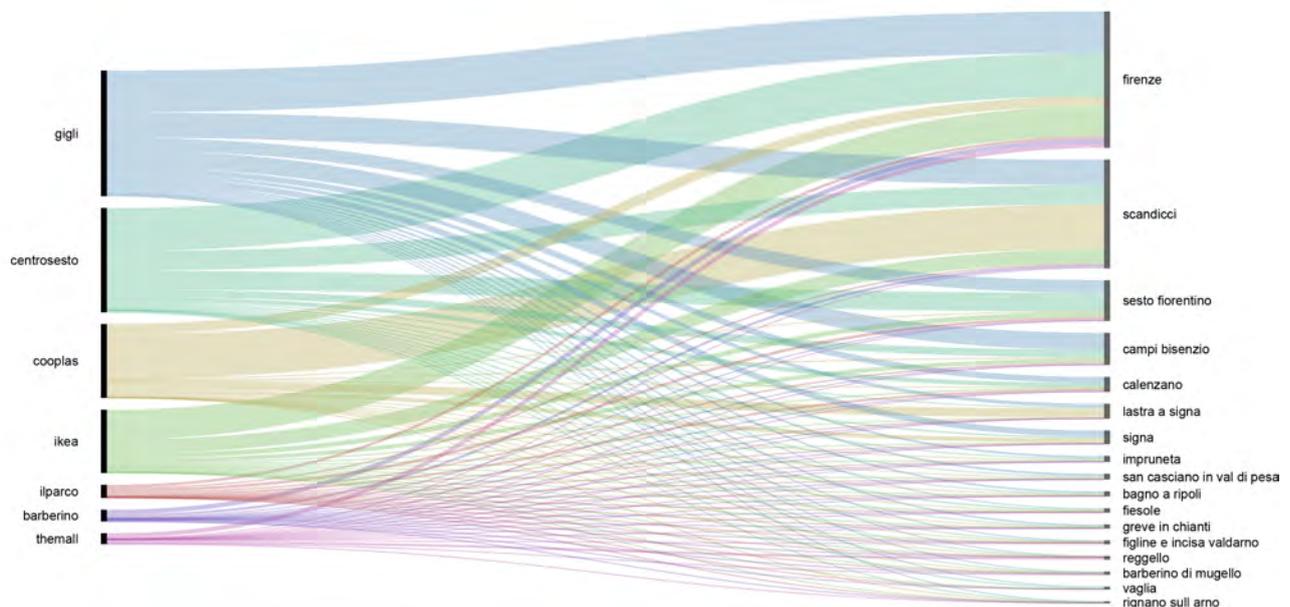
Comuni di residenza degli utenti che si sono recati presso ciascun centro commerciale nel periodo considerato.

Come ci aspetteremmo, i centri in cui la presenza del supermercato è il motivo fondamentale per visitare quello *store*, finiscono per raccogliere un bacino di utenti che in maggioranza abita nelle vicinanze di quella determinata zona. Man mano che ci spostiamo su attività che ospitano negozi di altro genere o che, in generale, hanno caratteristiche di unicità rispetto ai primi, vediamo che il raggio d'azione va a interessare luoghi sempre più lontani dai soli dintorni.

Le “mappe di calore” che vediamo mostrano aree più scure in presenza di un maggior numero di utenti, fino a diventare più chiare man mano che i visitatori scendono. E allora vediamo come Il Parco di Calenzano, il centro commerciale con la superficie di vendita inferiore tra quelli analizzati, ha visitatori concentrati nella piana fiorentina, con un'elevata affluenza proveniente dall'area dell'asse Firenze – Prato – Pistoia. Concentrazione elevata anche per la Coop di Lastra a Signa, con visitatori quasi esclusivamente dell'area sud di Firenze. Un po' più ampio, nella stessa categoria di attrattori commerciali, il raggio d'azione di Centro Sesto di Sesto Fiorentino. Legandoci a quanto introdotto prima in merito alla perequazione territoriale, è già chiaro quanto sia auspicabile il coordinamento amministrativo tra le aree su cui sorgono questi centri commerciali, dislocati in specifiche zone del territorio, strategiche dal punto di vista dell'accessibilità.

Per confermare quanto detto prima, invece, arriviamo poi al caso de I Gigli: è vero, c'è anche un *superstore* del gruppo Pam al suo interno, però coloro che si muovono alla volta dello *shopping centre* di Campi Bisenzio arrivano da numerose parti della Toscana, soprattutto quella centro-settentrionale e dei capoluoghi di provincia, e – possiamo immaginare – non soltanto per la spesa settimanale. Sicuramente sono più interessate le aree - e più frequenti, dunque, i relativi spostamenti – delle tre province che fanno parte dell'area metropolitana, ma non sono da sottovalutare le traiettorie che arrivano da tutta la regione. Fenomeno simile nel caso dell'IKEA, dove sono più intense le visite dai capoluoghi di provincia, così come nei due outlet studiati.

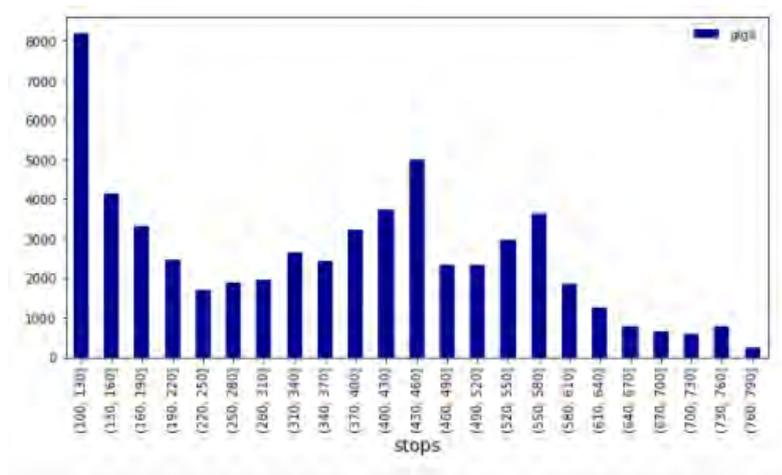
Le visualizzazioni che seguono, invece, permettono un dettaglio più focalizzato sugli scambi che avvengono tra questi grandi centri commerciali e i comuni della “cintura metropolitana fiorentina”, quell'area definita dagli enti amministrativi che circondano il capoluogo toscano.



La sankey mostra come si suddivide l'utenza di ciascun centro commerciale, con riferimento alla sola cintura metropolitana fiorentina.

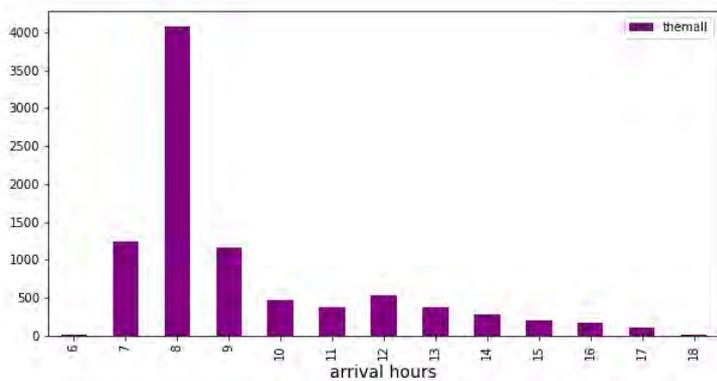
Bisogna sottolineare, a questo punto, un fatto: le informazioni che abbiamo illustrato in particolare con le sette mappe sono riferite alla totalità dei soggetti che si recano presso i centri commerciali analizzati nello studio. Come possiamo immaginare, questi si popolano non solo della loro clientela, dal momento che sono diversi e vari gli utilizzatori di queste aree che vi giungono per motivi differenti. Si tratta quindi di consumatori, ma anche di dipendenti e di corrieri che recapitano la merce ogni giorno. Se da un lato è vero che le traiettorie che si perdono sul reticolo stradale sono, in maggioranza, di coloro che si recano a fare acquisti, a un'analisi più approfondita potremmo scoprire qualcosa in più. È emerso, infatti, che ben il 26,7 % del traffico totale è generato dall'8,2 % degli utenti che utilizzano il territorio. Si tratta di coloro che compiono frequenti viaggi presso i centri commerciali, che vi sostano a lungo o, al contrario, per un breve lasso di tempo, e lo fanno in orari particolari.

Una fetta così piccola di veicoli, vuol dire, si appropria di un quarto del traffico totale. E quindi le 475 mila presenze che emergono – a vario titolo, come abbiamo capito – guardando a I Gigli durante le sette settimane analizzate non si riferiscono unicamente ai clienti delle attività commerciali presenti nel centro. Neanche gli utenti unici che vi si recano (224 mila, secondo le stime) sono solo riferite a coloro che andranno a fare acquisti nel tempo libero. Di seguito qualche statistica riferita a questi “regular users” e a come si comportano con i punti di interesse studiati.



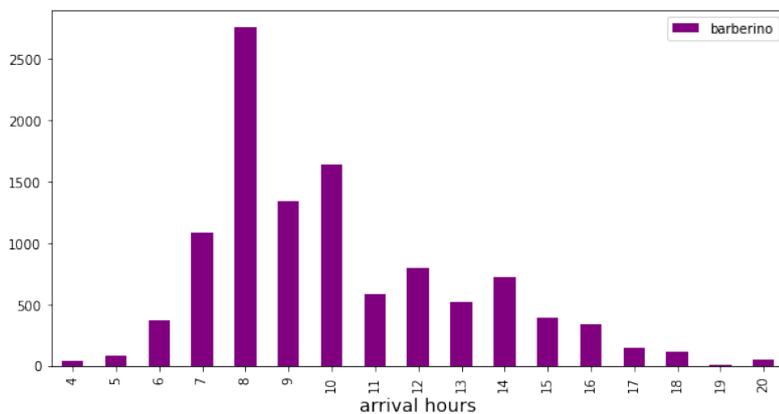
Distribuzione di frequenza sulla durata delle soste presso I Gigli espressa in intervalli di trenta minuti, riferita agli utenti più peculiari.

Nel grafico sopra rappresentato è interessante osservare il picco centrale, con molta probabilità riferito ai dipendenti (tenendo conto che 500 minuti corrispondono alle 8 ore di lavoro), mentre la coda destra individuerebbe coloro che utilizzano il parcheggio per scopi diversi dalla destinazione del polo commerciale. La coda sinistra può riguardare i corrieri e coloro che si recano più frequentemente al centro commerciale considerato, ma anche dipendenti con forme lavorative diverse rispetto al full time.



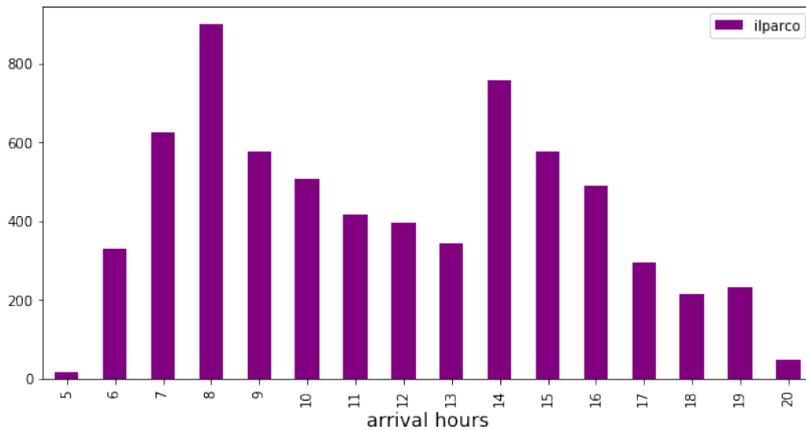
Distribuzioni di frequenza degli orari di arrivo dei regular users presso The Mall di Reggello.

È indicativo osservare l'orario di arrivo di questi utenti presso il The Mall. Qui il massimo è molto evidente: l'outlet del lusso apre al pubblico alle dieci, ma dipendenti e i corrieri che recapitano quotidianamente la merce iniziano a raggiungere la zona commerciale di Reggello dalle prime ore del mattino, con le otto che diventano l'ora di punta.



Distribuzione di frequenza degli orari di arrivo dei regular users presso l'Outlet di Barberino.

Questo comportamento emerge allo stesso modo negli altri centri commerciali, pur con delle differenze a seconda della tipologia: se anche l'outlet di Barberino denota un andamento simile, con un incremento ben definito degli arrivi presso l'area fra le sette e le dieci del mattino, negli altri poli sono più distribuiti – andamento diverso dovuto, verosimilmente, alle differenti forme contrattuali dei dipendenti che vi operano presso.

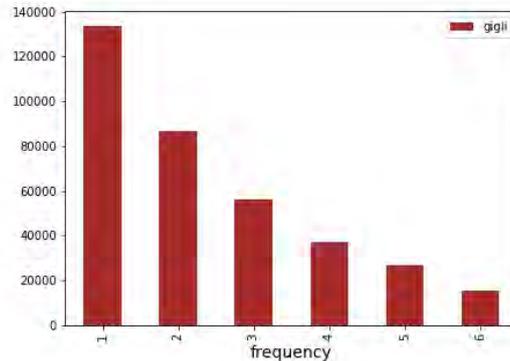
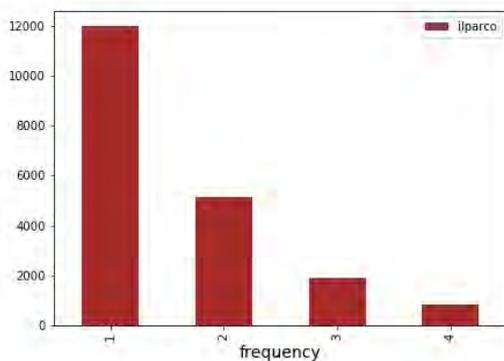
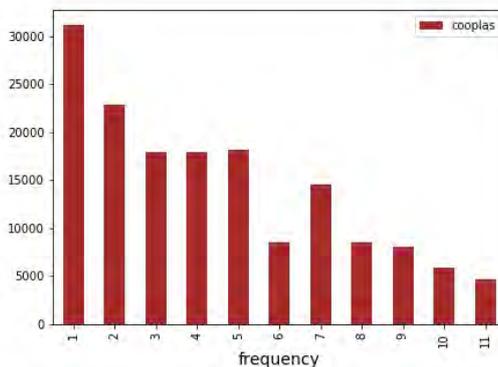
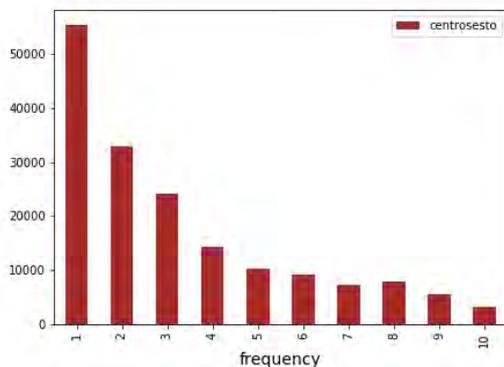


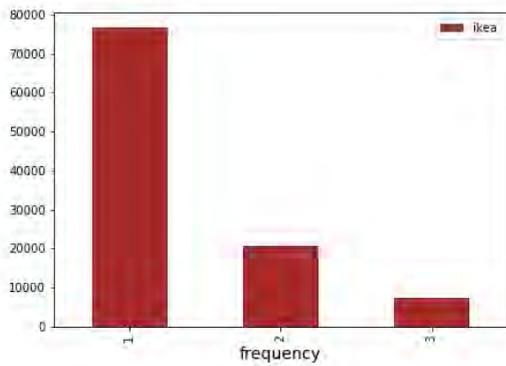
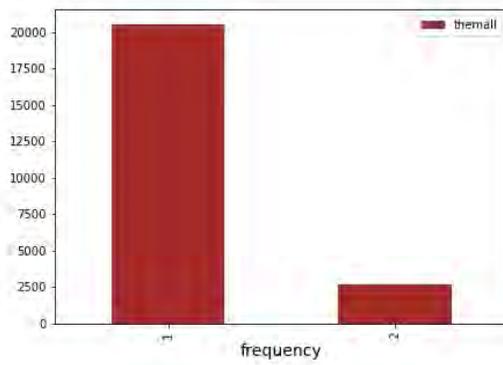
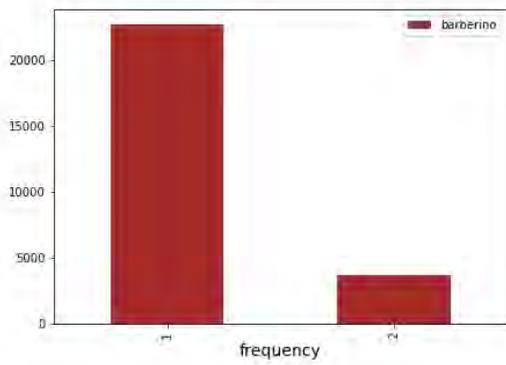
Distribuzione di frequenza degli orari di arrivo dei regular users presso Il Parco di Calenzano.

Nell'esempio illustrato è abbastanza evidente come gli orari di arrivo di questi utenti al centro commerciale Il Parco abbiano due momenti di massimo, corrispondenti ai due turni lavorativi. La proiezione dei dati ha, inoltre, fatto emergere una curiosità: non è affatto infrequente che i parcheggi adiacenti ai centri commerciali siano utilizzati come parcheggi scambiatori.

Questi che abbiamo tracciato finora sono aspetti legati ai veicoli che utilizzano il territorio per determinati motivi, e che generano un discreto ammontare di traffico. Sono utenti che condividono comportamenti omogenei; a partire dalla frequenza dei viaggi verso ciascuna destinazione sono stati poi studiati i loro orari di arrivo e il periodo di sosta. Ma per quel che riguarda i consumatori?

Partiamo proprio dalle distribuzioni di frequenza. Se i dipendenti e i corrieri sopra citati sono caratterizzati da frequenti viaggi verso le relative destinazioni nell'arco temporale definito, si parla di numeri diversi quando guardiamo a chi è interessato a fare acquisti.

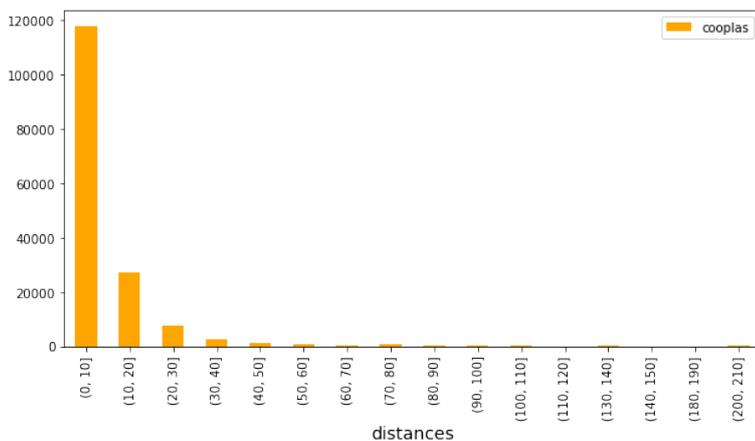


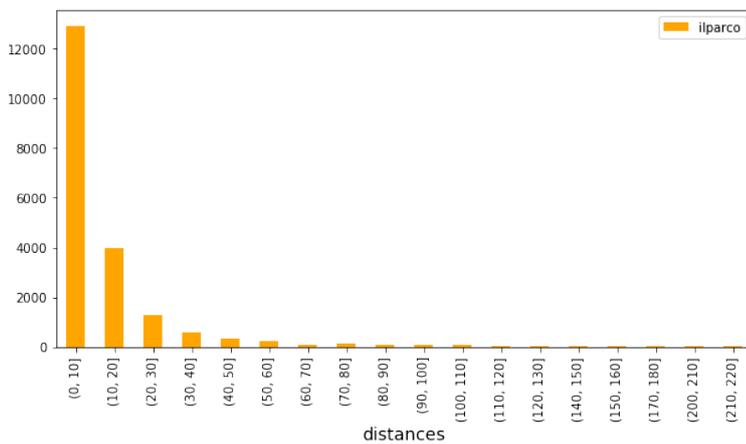
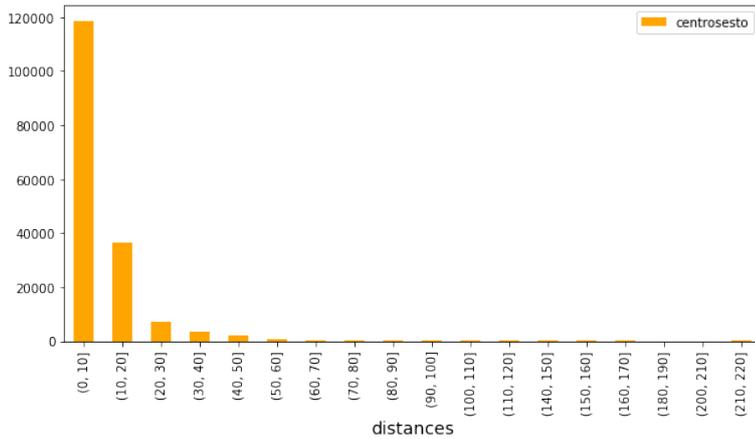


Distribuzione di frequenza delle visite presso ciascun attrattore commerciale

Come si evince chiaramente dalle distribuzioni in figura, i centri commerciali che ospitano un supermercato sono visitati più di frequente rispetto alle altre categorie. Gli outlet, infatti, vedono una netta prevalenza di persone che si recano a fare acquisti una volta sola nel mese e mezzo considerato, con qualche eccezione – ma, in ogni caso, tornando al massimo un'altra volta per andare a caccia di affari; discorso simile per chi è interessato ai mobili di IKEA.

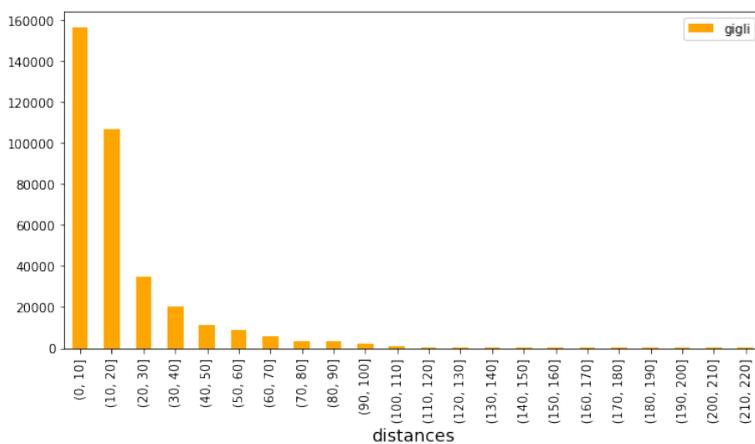
Qual è, invece, la distanza percorsa mediamente dai clienti di ogni centro commerciale? È interessante scoprire che ci sono differenze anche significative a seconda delle caratteristiche merceologiche di ciascun polo attrattivo.

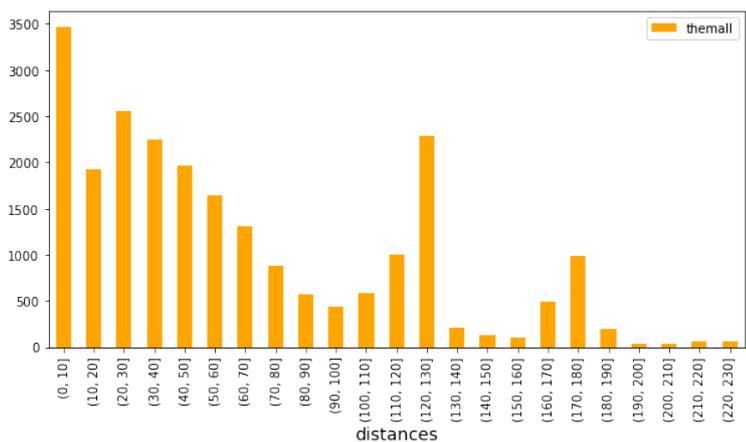
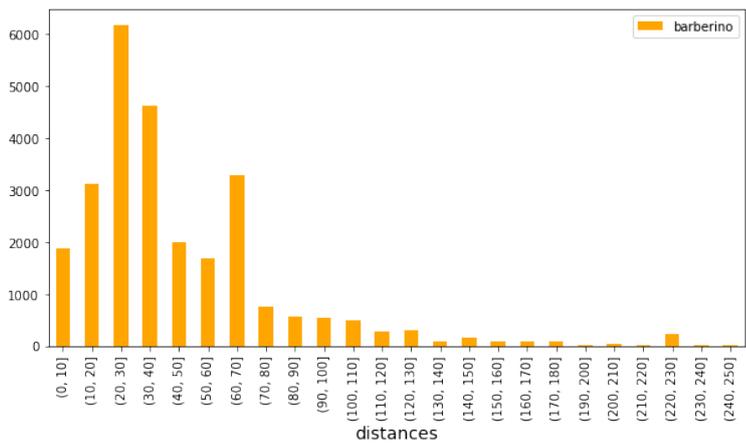
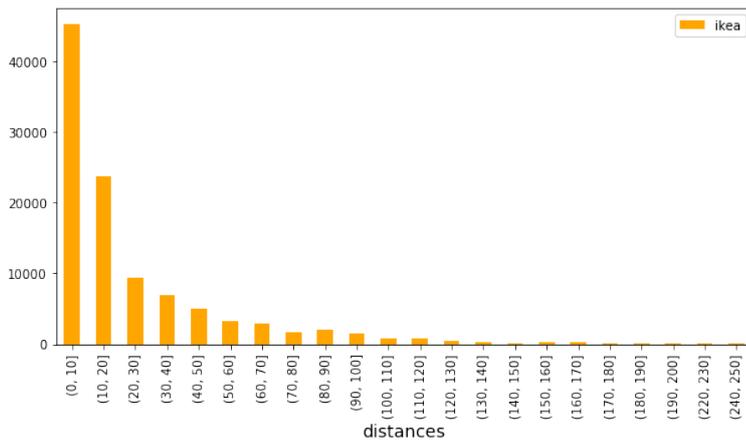




Coop di Lastra a Signa, Centro Sesto, Il Parco: Distribuzione di frequenza delle distanze percorse km per raggiungere i centri commerciali (km)

I luoghi dei supermercati, come al solito, mantengono un comportamento simile: i clienti risiedono nelle vicinanze, la maggioranza di loro infatti si muove entro i dieci chilometri dal comune di residenza o poco più.





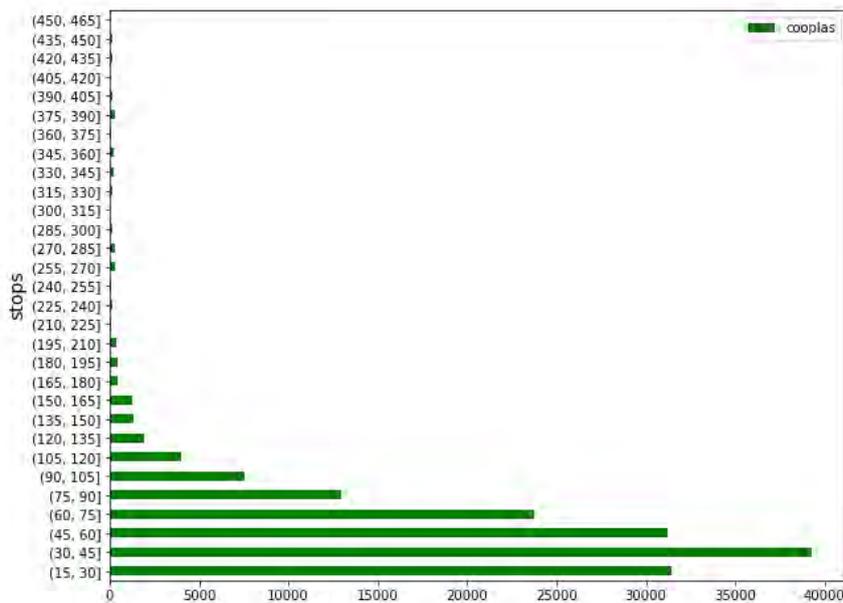
I Gigli, IKEA, Barberino Outlet e The Mall: distribuzioni di frequenza delle distanze percorse per raggiungere i centri commerciali (km)

La situazione inizia a cambiare nel caso di IKEA e I Gigli, con valori quasi sovrapponibili; le specialità che contraddistinguono i due centri commerciali avvicinano i clienti più lontani, fino ad arrivare in alcuni casi a cento chilometri di distanza. Distanza che aumenta visibilmente nel caso dei due outlet della piana fiorentina: non è difficile pensare che alcune di queste traiettorie siano riferite proprio a clienti che arrivano da fuori regione. Inoltre, la dislocazione di questi store è strategica perché situata proprio presso le uscite autostradali.

Se si vanno a calcolare le distanze medie si ha un'ulteriore conferma di queste evidenze. Fra i nove e i dodici chilometri per i supermercati, con un tempo di permanenza medio poco inferiore a un'ora;

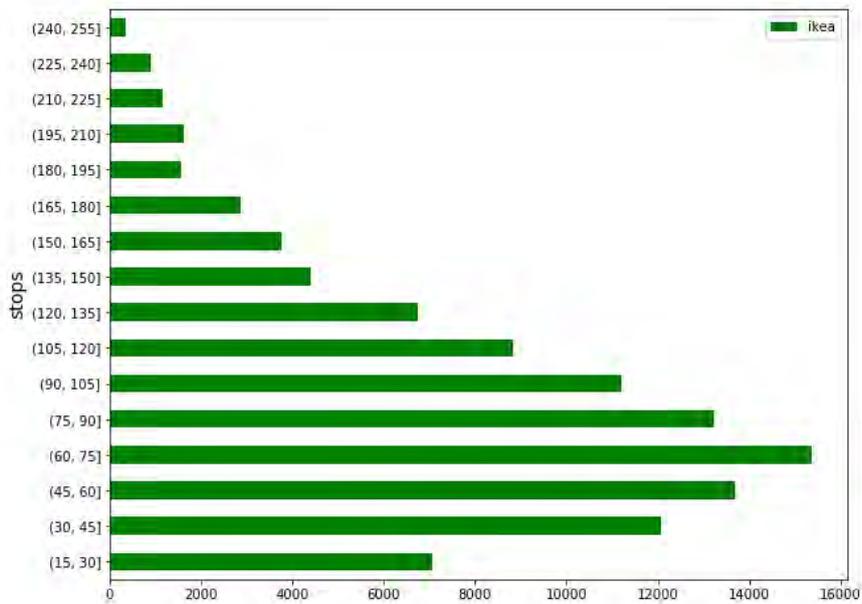
tra i diciotto e i ventiquattro chilometri per Gigli e IKEA che sembrano, così, delineare delle similitudini, soprattutto pensando che i clienti di entrambe mediamente sostano un'ora e mezzo; si passa, infine, ai quarantasei chilometri medi di Barberino Outlet e i sessantadue del The Mall, dove si stimano circa due ore di shopping per ciascun visitatore. La nicchia dell'alta moda, dunque, ripaga con una clientela disposta a percorrere più strada per raggiungere la meta.

A proposito di soste, siamo in grado di ottenere un dettaglio più specifico sui singoli centri commerciali. Ancora una volta, prendiamo – uno per tutti – l'esempio della Coop di Lastra a Signa: Così come negli altri super, la maggioranza riesce a concludere gli acquisti rimanendo ampiamente entro l'ora di tempo.



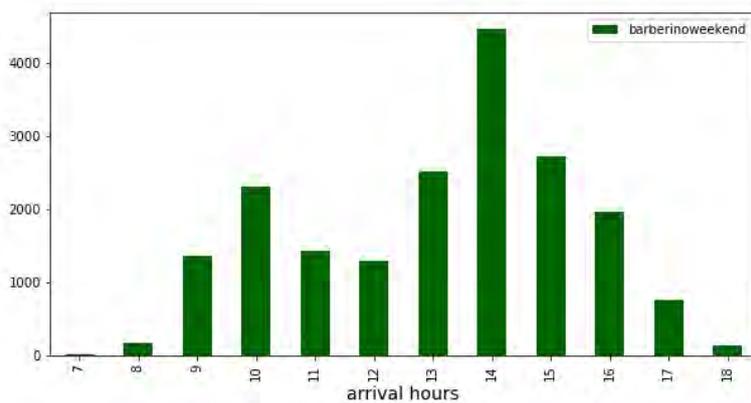
Distribuzione di frequenza delle soste presso la Coop di Lastra a Signa, raggruppate in intervalli di quindici minuti

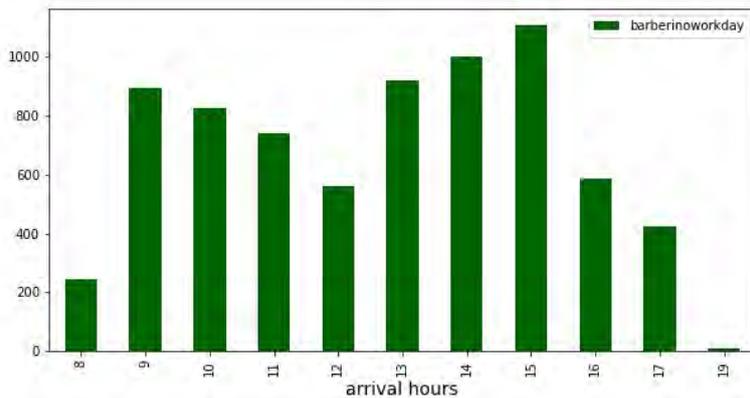
Distribuiti in modo molto differente, invece, i tempi di permanenza all'IKEA. Scegliere l'arredamento della nuova casa può portare via molte ore della giornata, senza dimenticare che chi ha la possibilità di raggiungere lo store in tempi rapidi sceglie anche di visitarlo per il suo ristorante o per acquisti che richiedono meno tempo.



Distribuzione di frequenza delle soste presso IKEA, raggruppate in intervalli di quindici minuti

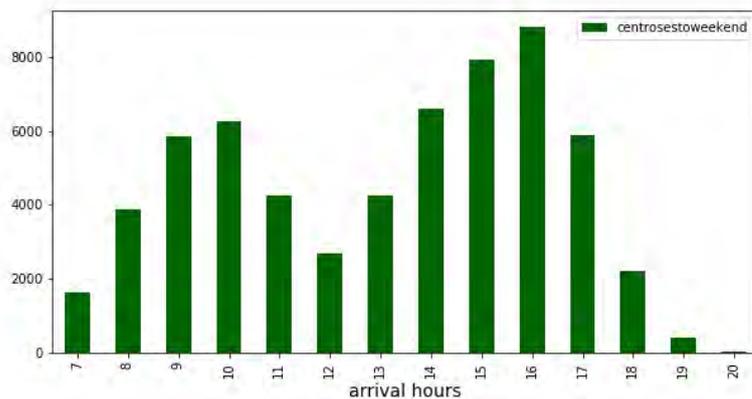
E per quel che riguarda gli orari di maggiore affluenza? Vediamo che i risultati mostrano una situazione abbastanza tipica in tutti i casi considerati.





Distribuzione di frequenza degli orari di arrivo presso l'Outlet di Barberino, a seconda che si tratti di giorni lavorativi o weekend

In generale, gli arrivi si distribuiscono in maniera sostanzialmente omogenea durante i giorni normali, mentre nel fine settimana coloro che si alzano presto per recarsi ai centri dello shopping sono in numero molto inferiore rispetto alle presenze che si registrano dopo l'ora di pranzo; l'esempio sopra riportato di Barberino riassume in maniera fedele quello che accade un po' ovunque. Se, poi, il nostro orario di lavoro ci tiene impegnati le classiche otto ore durante la settimana, con ritmi serrati dalle 9 alle 18, possiamo decidere qual è il momento ottimale in termini di affollamento per andare al supermercato nei giorni in cui siamo liberi. Basta dare uno sguardo al grafico seguente: vogliamo evitare file alla cassa? Secondo i risultati di Centro Sesto, a Sesto Fiorentino, conviene sbrigare la spesa settimanale il sabato o la domenica a mezzogiorno.



Distribuzione di frequenza degli orari di arrivo del weekend presso Centro Sesto a Sesto Fiorentino

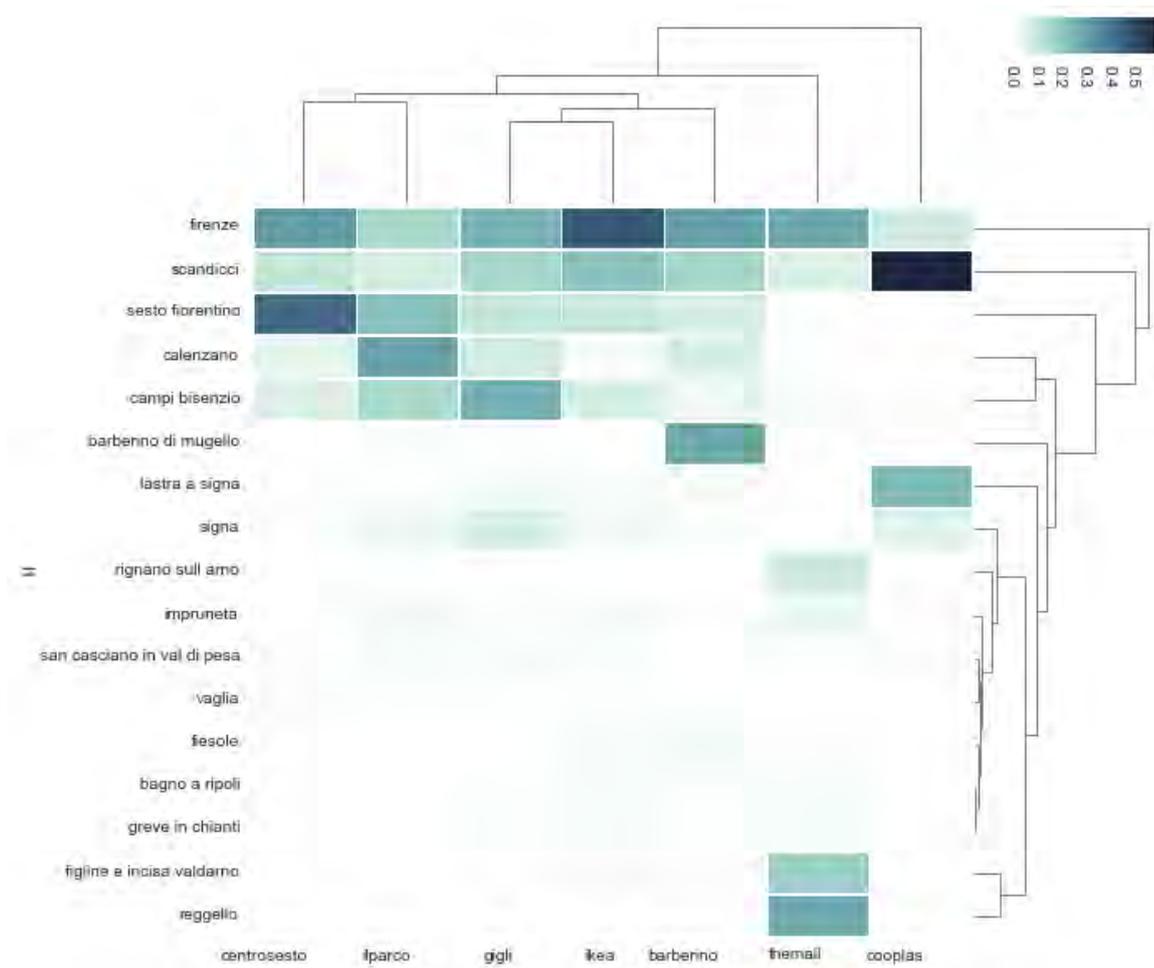
Machine learning e mobilità

Finora sono stati presentati i risultati ottenuti manipolando in vario modo i *big data* della mobilità che abbiamo incontrato all'inizio; il focus si è concentrato sulle traiettorie dei veicoli a disposizione, ed è emersa la caratterizzazione di due tipi principali di utenza e un utilizzo diverso delle strutture commerciali che si differenzia, più che per la localizzazione sul territorio, per la loro offerta prevalente.

Ma questa grande mole di dati può essere esplorata anche in altro modo. È possibile, infatti, applicare metodi di *clustering* – approcci di *machine learning* che definiscono, per mezzo di algoritmi

predefiniti, come raggruppare i dati in modo da creare la maggiore similarità all'interno dei gruppi così ottenuti e, allo stesso tempo, più eterogeneità tra gli stessi. Nel nostro caso, abbiamo provato a vedere cosa succede adottando uno di questi sistemi di apprendimento automatico per scoprire se esistono relazioni significative utilizzando mezzi alternativi, e se questi validano o smentiscono i risultati ad ora presentati.

Il grafico proposto, in questo caso, a partire dai poli commerciali oggetto di studio considera le origini dei veicoli con un dettaglio sui comuni della cintura metropolitana fiorentina. Ad esempio, a partire dai numeri delle visite totali di IKEA, è stata calcolata la percentuale di utenti che visita il negozio abitando a Fiesole o a Scandicci.



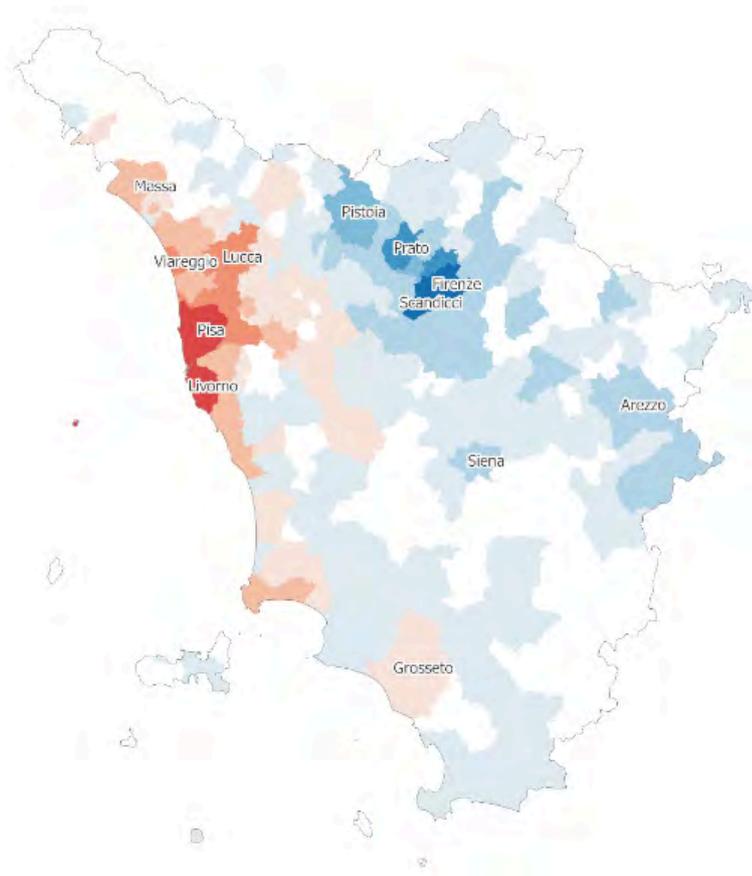
Un esempio di clustering gerarchico applicato ai comuni e agli attrattori.

Seguendo i diagrammi ad albero tracciati a margine, scopriamo che in prima istanza l'algoritmo considera parte dello stesso gruppo lo store di Centro Sesto e Il Parco, i Gigli e IKEA: questo risultato in qualche modo conferma le evidenze incontrate finora sul simile utilizzo dei due centri commerciali, finendo per considerare insieme, da una parte, i supermercati e, dall'altra, i due poli attrattori più particolari; ulteriori elaborazioni finiscono per unire l'outlet di Barberino ai secondi, e così via. In modo simile, sull'altro lato si possono leggere i comuni che, sempre secondo l'algoritmo, denotano un comportamento analogo. La heatmap che accompagna il grafico, invece, indica il volume di traffico originato da ciascun comune e diretto verso i centri commerciali; più scura è la cella, maggiori sono gli spostamenti degli abitanti di un comune verso quell'attrattore.

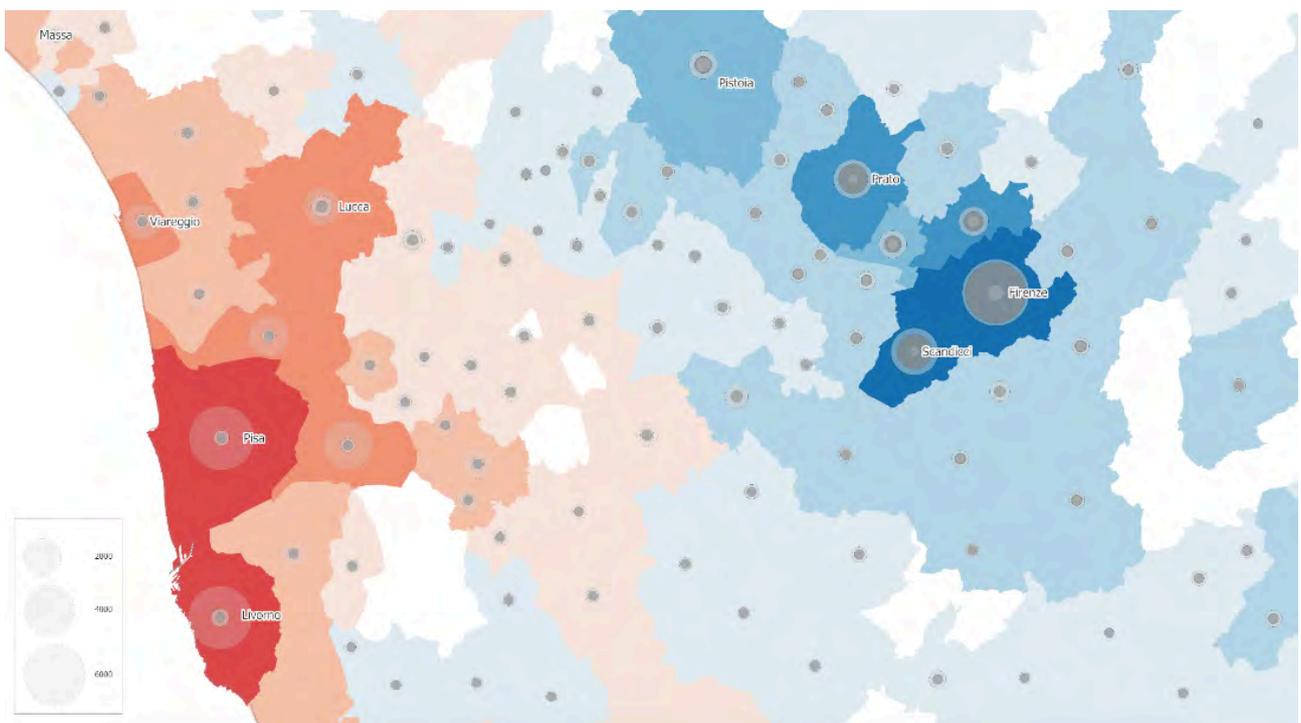
Un modello per i confronti: il caso Ikea

Le analisi finora presentate hanno dimostrato risultati interessanti, qualche volta devianti dal fatto che i centri commerciali studiati, in ogni caso, hanno caratteristiche tra loro differenti. Prevedere le preferenze degli utenti di un luogo rispetto a un altro, però, non è sempre agevole, dal momento che sono necessari dati di partenza il più possibile confrontabili. Eppure, da qualche anno, due realtà perfettamente identiche per offerta esistono sul territorio regionale. Stiamo parlando delle due IKEA toscane: a inizio 2014, infatti, ha aperto al pubblico anche lo *store* di Pisa. Entrambe dislocate in zona industriale, in prossimità dei rispettivi aeroporti, vicine a svincoli autostradali; una ad accogliere la clientela proveniente dalla costa e l'altra quella dall'area più a est della regione. Un certo effetto cannibalismo si è inevitabilmente verificato, portando gli interessati che abitano nelle zone limitrofe a ciascun attrattore a preferire, ovviamente, quello più vicino. Meno immediato, invece, il comportamento degli utenti provenienti da comuni equidistanti dai due capoluoghi. Sono diverse le domande che possiamo porci. La scelta di un negozio rispetto a un altro dipende solo da motivi di accessibilità? Qual è il volume di clientela che la nuova apertura ha sottratto al centro commerciale di Firenze? C'è da tenere in conto che per questo esperimento i dati a disposizione hanno riguardato un periodo limitato, circa due settimane osservate il mese precedente all'inaugurazione dell'IKEA pisana e le due settimane seguenti, e che in certi casi i volumi di traffico sono stati influenzati proprio da un "effetto novità" che si è generato. Il metodo presentato rimane comunque un modello di riferimento per ulteriori studi ed esplorazioni.

Nelle mappe si possono trovare altre informazioni. Innanzitutto, blu e rosso distinguono la propensione degli abitanti di ciascuna area a recarsi presso un negozio o l'altro. Maggiore è l'intensità del colore, più è netta la predilezione degli utenti di ciascun comune verso quella destinazione. È interessante vedere come i confini delle aree amministrative territoriali siano, in qualche modo, ridisegnati con lo studio delle traiettorie, andando a creare nuovi bacini funzionali.



Mappa di preferenza e fidelizzazione della clientela proveniente da ciascun comune verso gli store (rosso per l'IKEA pisana, blu per quella fiorentina).



Volumi di traffico che originano da ciascun comune, con focus sull'area centro-settentrionale della regione. I cerchi grigio scuri individuano gli utenti che si sono recati al negozio di Sesto Fiorentino nell'arco di tempo considerato

prima dell'apertura di Pisa, i cerchi grigio chiari sono riferiti al periodo post apertura, mentre in trasparenza il volume di traffico riferito allo store pisano.

Il colpo d'occhio della visualizzazione favorisce l'evidenza del giro di affari che Firenze ha perso a vantaggio del centro commerciale pisano. In realtà, osservando le mappe si potrebbero evincere altri fatti interessanti. Da una parte, infatti, è come se IKEA avesse creato un bisogno in precedenza inespreso o piuttosto debole per le località della costa Toscana. In altre parole, quei comuni che prima poco alimentavano la clientela fiorentina, adesso si sono trasformati in un'importante fetta di clientela del centro commerciale di Pisa; ne è un chiaro esempio il caso delle aree di Pontedera, Livorno, Lucca e Pisa stessa.

Dall'altra parte, gli abitanti dei comuni che prima si recavano frequentemente a Sesto Fiorentino non sono stati particolarmente interessati dall'apertura pisana, anche se si possono notare gli effetti suscitati dalla novità, appunto, come ad esempio nei casi di Pistoia e Prato, che registrano qualche viaggio anche in direzione della costa. In ogni caso, rimangono di particolare interesse le aree più chiare sulla mappa, che non prendono particolare posizione nei confronti di uno o dell'altro centro commerciale rimanendo contese dai due store. Oltre a criteri di accessibilità, la differenza potrebbe essere fatta dall'offerta data dal territorio, fuori dai quattro muri dei negozi.

In conclusione, è indubbio che studi di questo genere – pur con i relativi limiti – potrebbero apportare notevoli vantaggi durante i processi decisionali degli enti amministrativi, con un diverso punto di vista che potrebbe risolvere aspetti critici della pianificazione territoriale. Inoltre, questo lavoro potrà essere sviluppato anche in ulteriori direzioni, automatizzando il processo di analisi di nuovi dati e generalizzando il modello ottenuto per renderlo applicabile alle nuove situazioni, per evidenziare i punti di interesse che hanno un impatto negativo sul traffico locale.

Allegato 6

**Le presenze turistiche
nell'Area metropolitana di Firenze,
Prato e Pistoia attraverso i dati
di telefonia mobile**

Indice

INTRODUZIONE	3
DATI	3
RISULTATI	4
ANALISI PRESENZA PER SINGOLA ANTENNA	5
FIRENZE	5
PRATO	7
PISTOIA	8
ANALISI ANTENNE AGGREGATE	9
FIRENZE	9
PRATO	10
PISTOIA	11
CO-PRESENZA DEI TURISTI NELL'AREA DI FIRENZE	12
CONCLUSIONI	16

Introduzione

«Il primo di luglio prossimo andrò all'estero, in Italia, dove resterò un po' di tempo», così annunciava Dickens nel maggio del 1844 all'amico Macvey Napier.

Era l'Europa dei Grand Tour e lo scrittore inglese si apprestava a compiere il suo viaggio alla scoperta dell'Italia. La corrispondenza epistolare avrebbe dato vita all'opera 'Lettere dall'Italia', raccolta di scritti in cui Dickens documentava il proprio viaggio.

Oggi, maggio 2018, lo scrittore al suo arrivo alla stazione di Santa Maria Novella a Firenze non dovrebbe far altro che, impugnare il proprio smartphone per comunicare a Macvey Napier il proprio arrivo nella città di Dante.

Se da un lato questo ci avrebbe privato di uno dei capolavori letterari del 1800, l'illustre scrittore, così come migliaia di altri turisti in arrivo nella culla del Rinascimento, collegandosi all'antenna GSM avrebbero lasciato l'impronta digitale del proprio passaggio.

L'utilizzo delle moderne tecnologie di comunicazione ci consente di disporre di un enorme mole di dati da raccogliere, analizzare e interpretare per comprendere meglio le dinamiche sociali, culturali ed economiche.

Se nel 1800 solo nobili e letterati potevano permettersi di affrontare viaggi per l'Europa a fine ludici e culturali, oggi in un mondo sempre più interconnesso milioni di persone ogni giorno si spostano per piacere, lavoro o necessità. Diventa pertanto di fondamentale importanza per istituzioni, aziende e policy maker approfondire la conoscenza delle dinamiche di tali flussi e degli effetti che essi hanno sul territorio urbano.

Il nostro lavoro ha come obiettivo quello di interpretare i dati GSM, raccolti nell'area metropolitana di Firenze, Prato e Pistoia, per identificare i comportamenti di coloro che, più avanti definiremo 'visitatori' con l'intento di contribuire al patrimonio informativo a disposizione delle istituzioni preposte al governo del territorio.

Dati

Negli anni recenti la diffusione dei dispositivi mobili ha portato ad un incremento del volume dei dati relativi a "dove" e "quando" le persone si spostano.

L'analisi è stata compiuta su un campione di dati GSM, ovvero dati di telefonia mobile raccolti dall'operatore telefonico durante l'uso del telefono cellulare con chiamate, SMS o traffico dati. Il fornitore dei dati è uno dei maggiori operatori nel settore delle telecomunicazioni in Italia. Per compiere le analisi di nostro interesse sono stati utilizzati i Call Data Record (CDR), ovvero record registrati durante le attività di chiamata e che contengono l'informazione temporale, un identificatore anonimo dell'utente, la cella telefonica e la durata della chiamata. Pertanto, essendo le attività dei CDR riferite a chiamate telefoniche, le analisi sono state compiute escludendo le attività di dati inerenti alla connessione Internet.

La nostra analisi vuole concentrarsi su flussi e spostamenti nell'area metropolitana di interesse da parte di soggetti che non sono ivi residenti. Per individuare nei nostri dati tali soggetti, è stato interpellato il KDDLab (ISTI – CNR, Pisa), il quale ha messo a disposizione

lo strumento del Sociometro, strumento di *data mining* che, attraverso l'analisi delle abitudini di chiamata degli utenti classifica gli stessi in categorie comportamentali.

La nostra analisi si concentra su coloro che, dal Sociometro sono classificati come "Visitor", ovvero coloro che non vivono e non lavorano/studiano nell'area di interesse, ma che occasionalmente la visitano.

Il campione risultante è composto da circa 9 milioni di record, nel range temporale che va dal 05 Ottobre al 30 novembre 2015.

User_id	Data	Ora	Antenna
---------	------	-----	---------

Per ciascuna antenna presa in esame disponiamo delle coordinate geografiche (latitudine e longitudine) che abbiamo utilizzato per le rappresentazioni su mappa.

Risultati

In questa sezione presentiamo i risultati ottenuti dai dati a nostra disposizione, cercando di rispondere alle domande: come si muovono i visitatori nell'area metropolitana di Firenze, Prato e Pistoia? Esistono differenze fra centro e periferia? Come cambia la fruizione del territorio a seconda del giorno o dell'orario?

In questo lavoro, la presenza sul territorio di un visitatore in uno specifico luogo è determinato dal fatto, che questi si colleghi ad una certa antenna telefonica.

Il limite di tali dati è quello di non fornire una precisa collocazione geografica, ma di indicarci la presenza in un'area, non consentendo alla nostra analisi di scendere a un livello di maggior dettaglio.

Nei seguenti paragrafi si propongono analisi che aiutino a comprendere la mobilità di tali utenti nell'area metropolitana di Firenze, Prato e Pistoia, quali:

1. analisi per ogni antenna della presenza media oraria nei diversi giorni della settimana, al fine di identificare differenze di comportamento dei visitatori e la presenza di eventi eccezionali;
2. Abbiamo ripetuto l'analisi aggregando le antenne sulla base di caratteristiche qualitative comuni, quali la vicinanza geografica, la vicinanza ad arterie di comunicazione, la presenza di attrattori, seguendo criteri suggeriti da esperti di dominio;
3. Ci siamo infine concentrati sulla sola area di Firenze, per rappresentare su mappa le aree 'toccate' da ciascun visitatore durante la propria permanenza nella città (grafo di co-occorrenza).

Analisi presenza per singola antenna

L'obiettivo è quello di fornire una panoramica a livello statistico delle presenze medie orarie e giornaliere degli utenti identificati dal sociometro come *visitatori*.

L'analisi è stata compiuta, attraverso lo strumento 'R', un linguaggio di programmazione e un ambiente di sviluppo specifico per l'analisi statistica dei dati. Sono state oggetto di analisi solo le antenne con il maggior numero di presenze, ovvero il maggior numero di visitatori che si sono agganciati a quell'antenna.

Di seguito riportiamo i grafici denominati '*Call intensity*', che rappresentano la media oraria delle chiamate, divise per giorno della settimana. Tale rappresentazione consente di verificare le differenze nell'utilizzo del territorio da parte dei visitatori della città.

Per esemplificare abbiamo riportato per ciascuna area in esame il grafico relativo alle antenne in cui le differenze sono maggiormente evidenti.

Firenze

Per l'area di Firenze abbiamo selezionato le antenne:

- FI001 centro storico della città
- FI023 area stadio comunale Artemio Franchi
- FI033 area di Novoli
- FI037 area Cascine

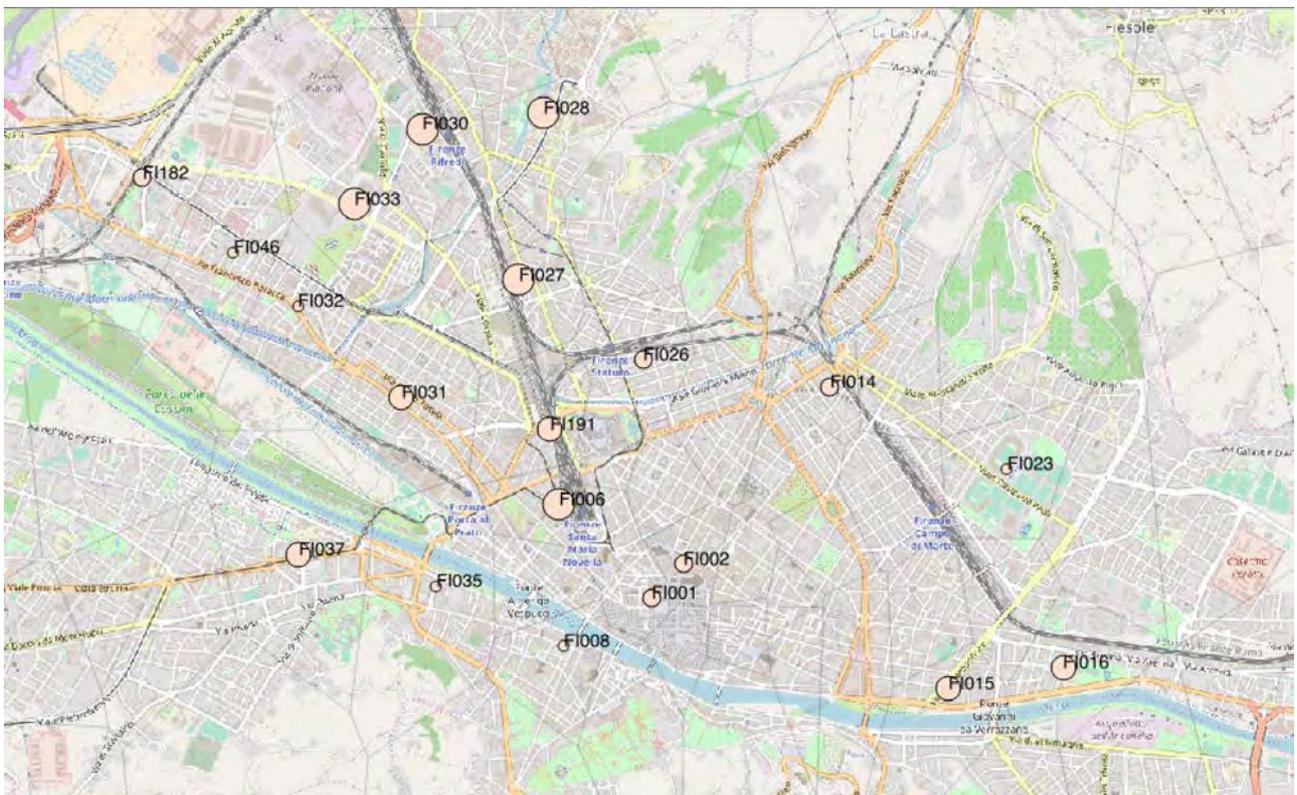


Figura 1 Top antenne di Firenze

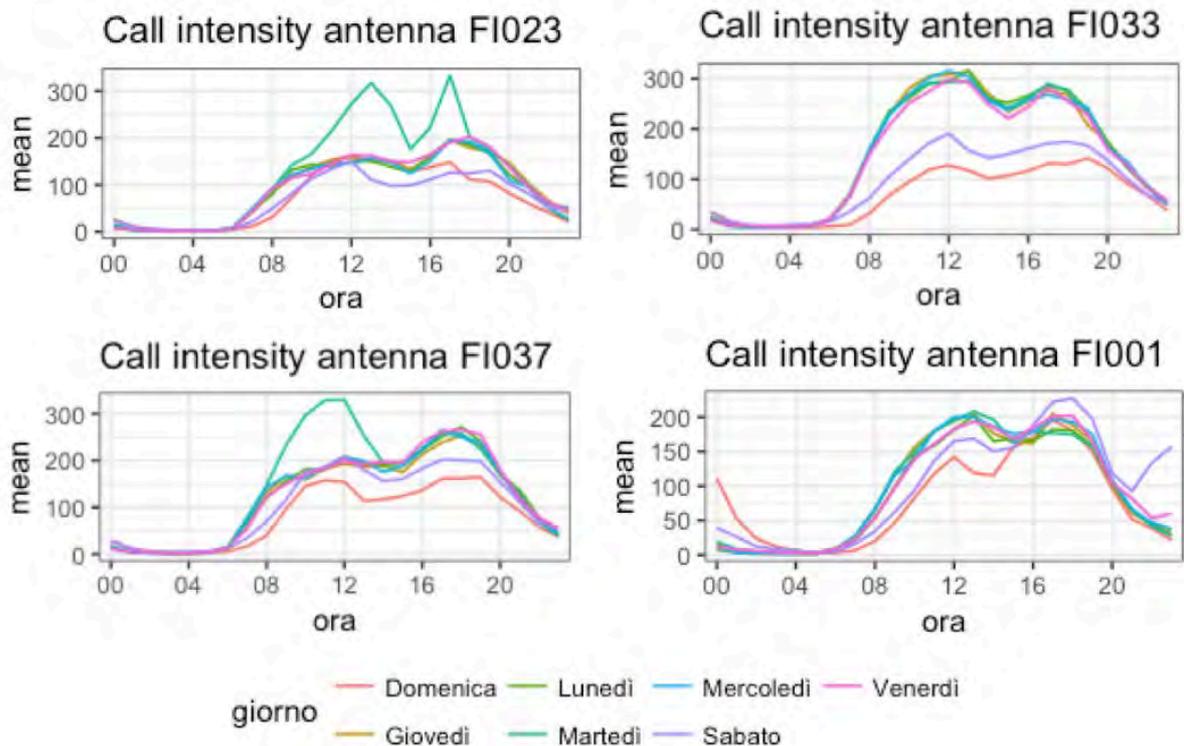


Figura 2 Call Intensity - Focus antenne Firenze

I quattro grafici mostrano differenze rispetto alle presenze di visitatori, sia per quanto riguarda i giorni della settimana che, in relazione all'orario considerato.

Nella zona del centro storico (FI001) non sembrano esserci differenze rilevanti nella presenza di visitatori fra i giorni lavorativi e week-end. È interessante notare che nel fine settimana, a differenza di altre aree, si registrano presenze di visitatori in tarda serata/notte, probabilmente trattasi di residenti al di fuori dell'area metropolitana che, raggiungono i locali del centro di Firenze per trascorrere le serate del fine settimana.

L'area di Novoli (FI033), sede sia del polo universitario che di uffici, mostra un'importante differenza fra le presenze nei giorni lavorativi e il fine settimana.

Per l'area delle Cascine (FI037) il grafico evidenzia un andamento che differisce rispetto alle altre aree nella mattinata del martedì. L'area infatti è sede del più grande mercato all'aperto di Firenze che si tiene tutti i martedì mattina.

L'antenna relativa alla zona dello stadio (FI023) riporta un andamento inatteso nel giorno di martedì. Ad una più attenta analisi tale scostamento è da imputarsi alla visita pastorale del Santo Padre tenutasi martedì 10 novembre 2015, che ha rappresentato un evento eccezionale, richiamando per l'occasione 50 mila persone durante la Santa Messa.

Prato

Per l'area di Prato abbiamo selezionato le antenne:

- PO007 centro storico della città
- PO011 SS719

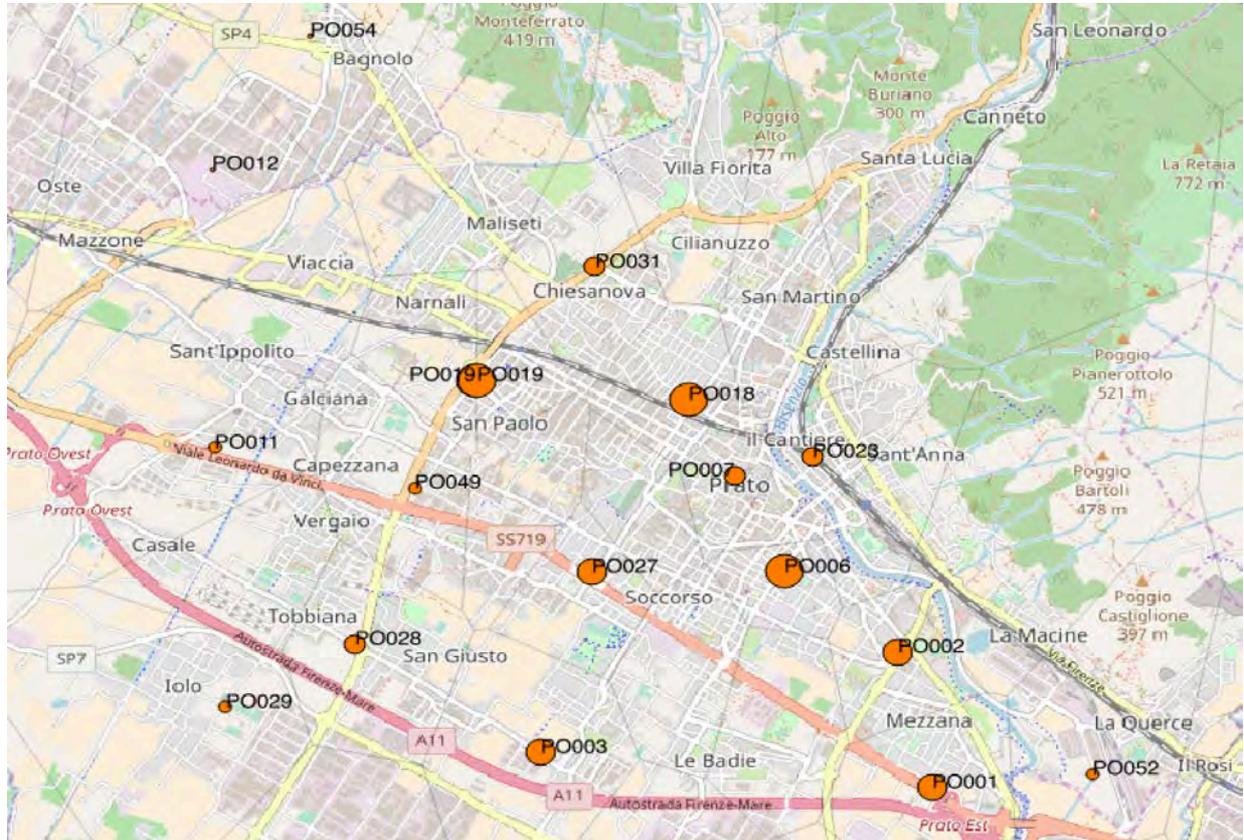


Figura 3 Top antenne di Prato

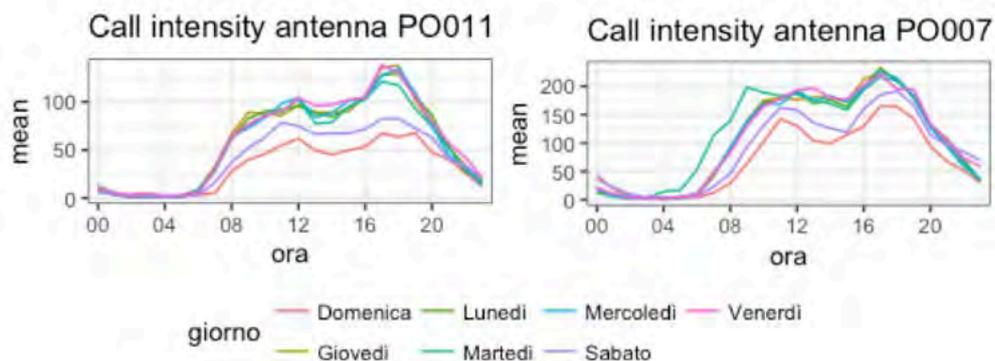


Figura 4 Call Intensity – Focus antenne Prato

La zona del centro storico di Prato (PO007) è anch'essa stata interessata dalla visita Pastorale del Santo Padre, nella primissima mattinata di martedì 10 novembre 2015. Dai dati riportati dal comune di Prato si stima fossero presenti 30 mila persone, 5 mila solo in Piazza Duomo.

Dall'area di grande comunicazione (PO011) si può osservare nei giorni lavorativi il picco di presenze dalle 16 alle 19 circa, presumibilmente l'orario di maggior congestione del traffico, specie per immettersi sull'autostrada A11.

Pistoia

Per l'area di Pistoia abbiamo selezionato le antenne:

- PT005 centro storico della città di Pistoia
- PT015 centro storico Montecatini Terme

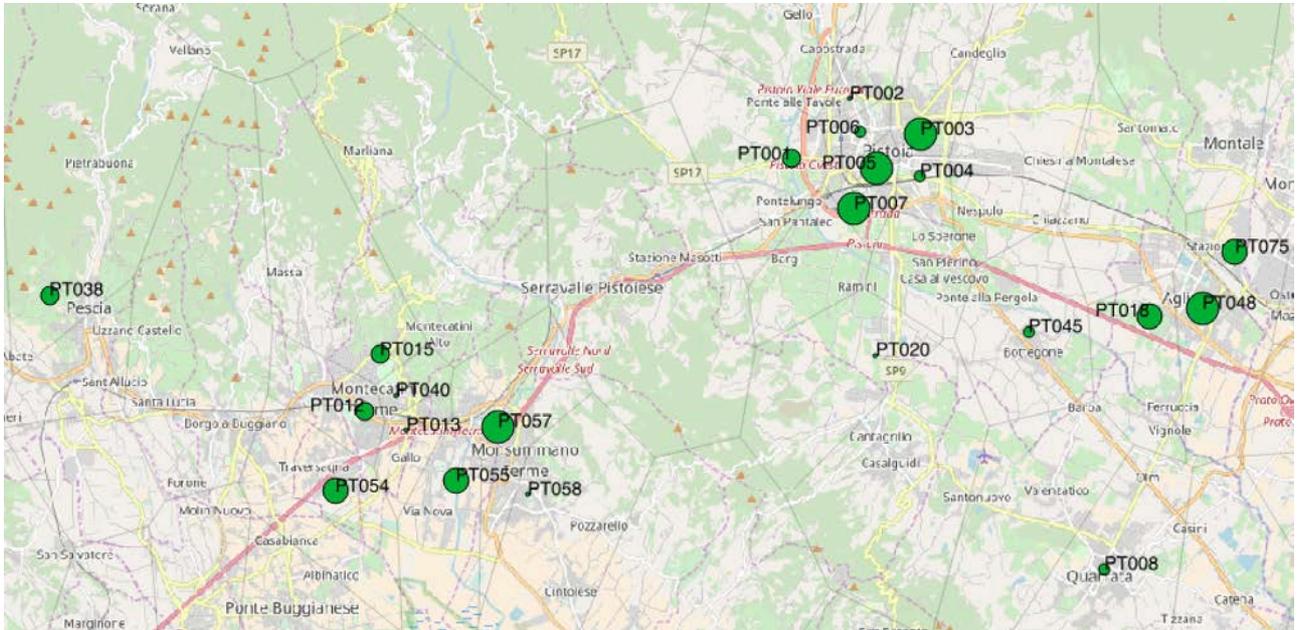


Figura 5 Top antenne Pistoia

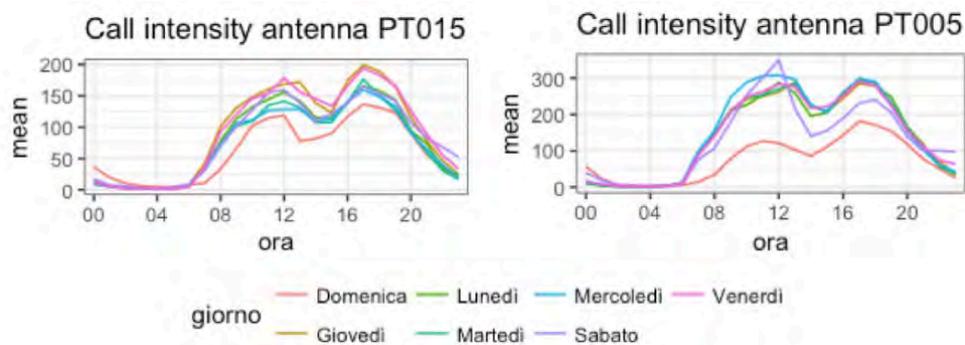


Figura 6 Call Intensity - Focus antenne Pistoia

Il centro storico di Pistoia (PT005) mostra un picco di presenze il sabato, specie nella mattina, in coincidenza del mercato settimanale.

Di particolare interesse, l'area del centro di Montecatini Terme (PT015), località a forte vocazione turistica, in cui contrariamente alle altre aree non si osservano comportamenti diversi fra i giorni lavorativi e il fine settimana, possibile evidenza che i visitatori della città si identificano con la categoria dei turisti che fanno base a Montecatini Terme per visitare la Toscana.

Analisi antenne aggregate

In considerazione del gran numero di antenne analizzate, si è scelto di aggregare le antenne sulla base di caratteristiche qualitative, quali la vicinanza geografica e l'appartenenza ad aree in cui si collocano specifiche attività socio-economiche (i dettagli dell'aggregazione sono specificati in appendice).

Abbiamo poi rappresentato tramite *heatmap* la 'Call Intensity' media oraria per giorno della settimana.

Firenze

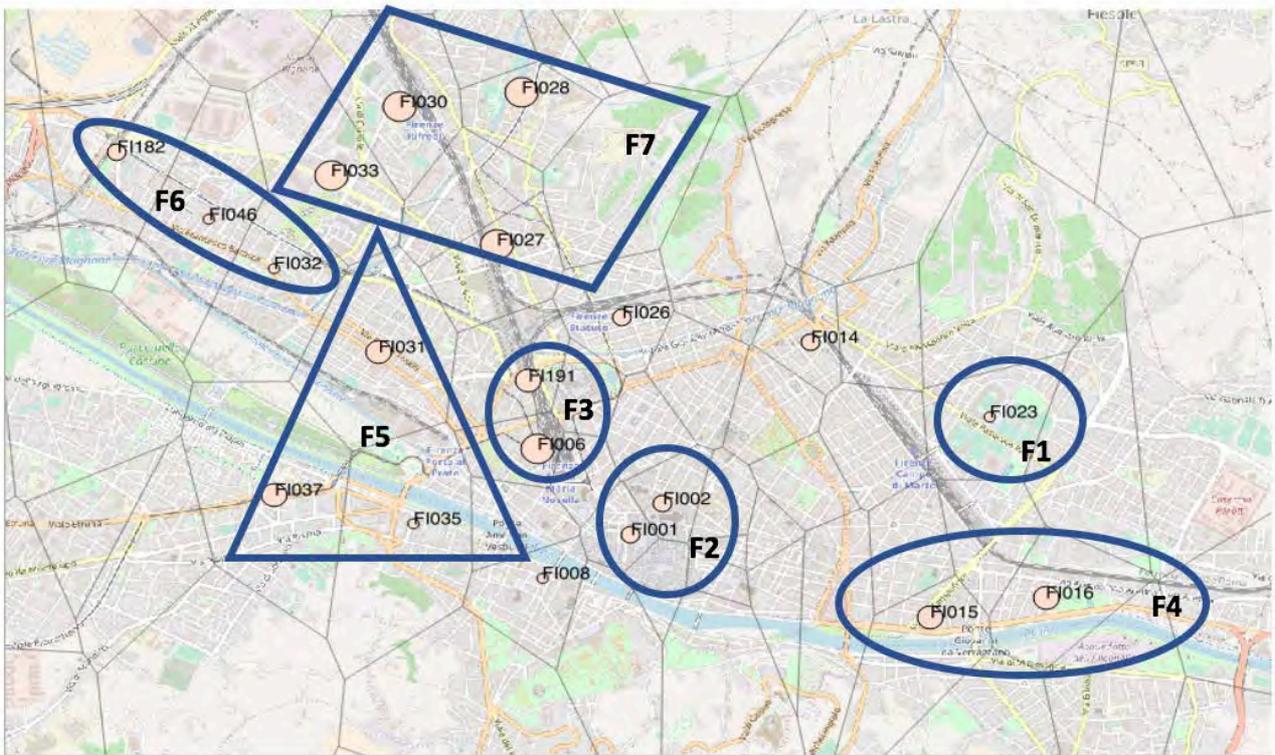


Figura 7 Antenne aggregate a Firenze

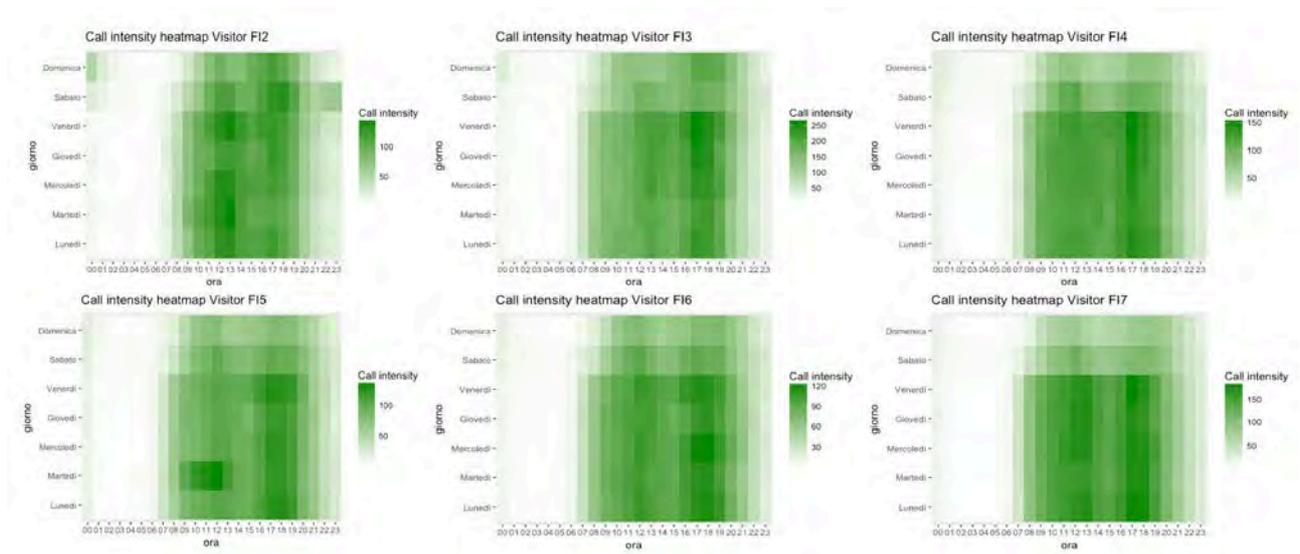


Figura 8 Heatmap Visitor Firenze

Prato

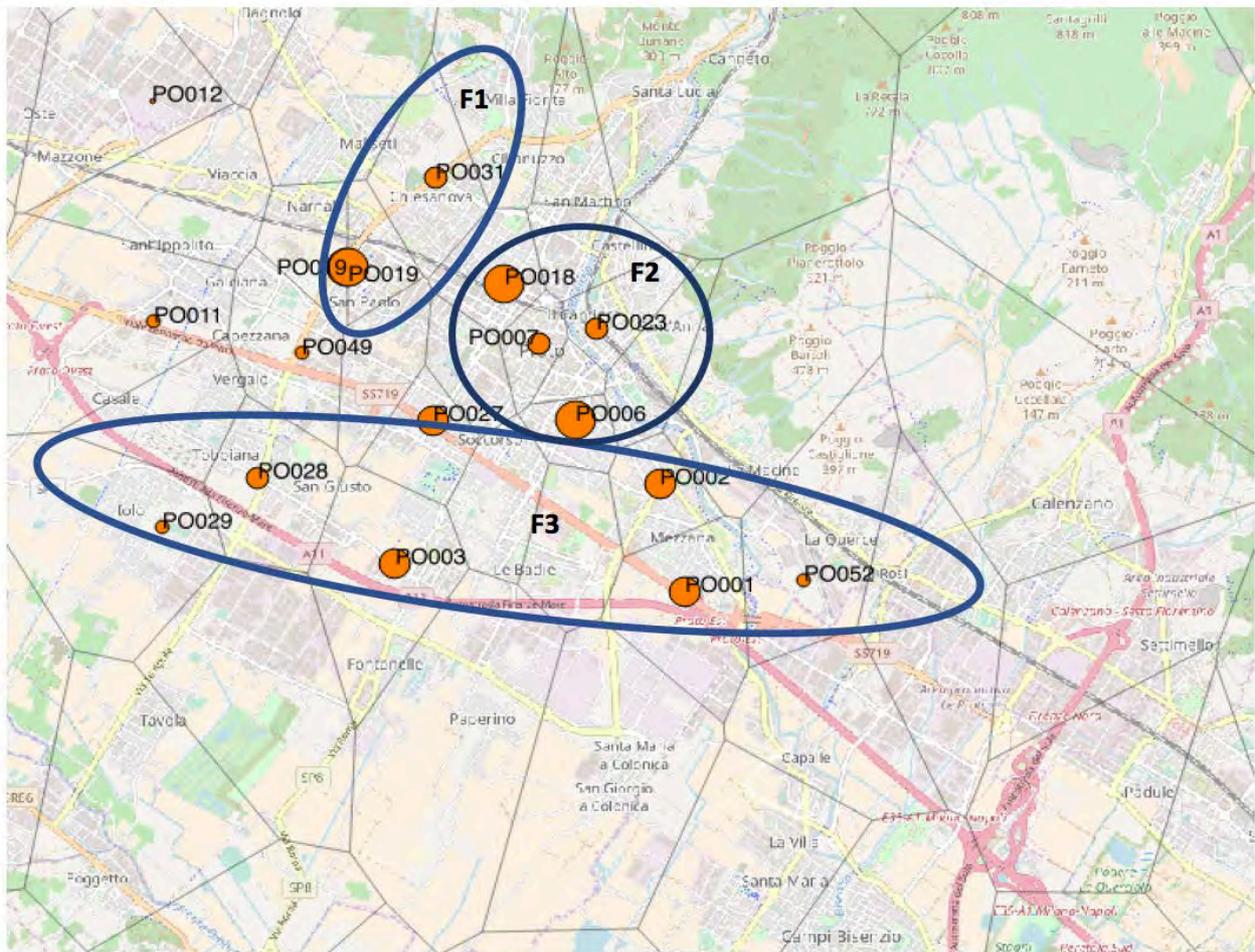


Figura 9 Antenne aggregate Prato

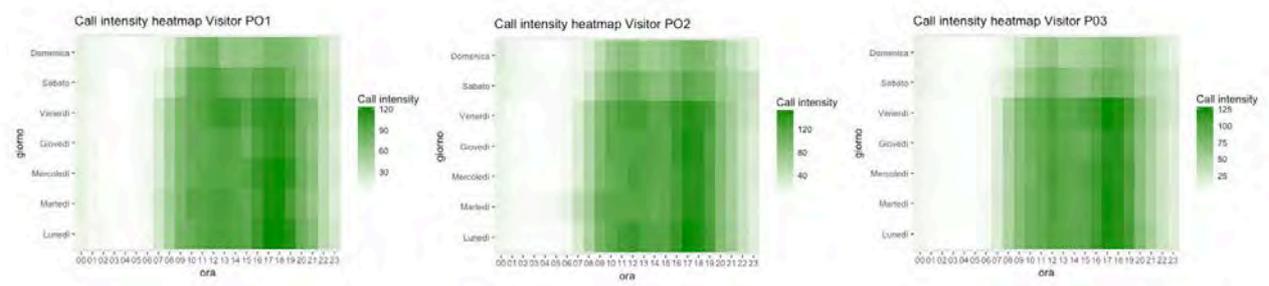


Figura 10 Heatmap Visitor Prato

Pistoia

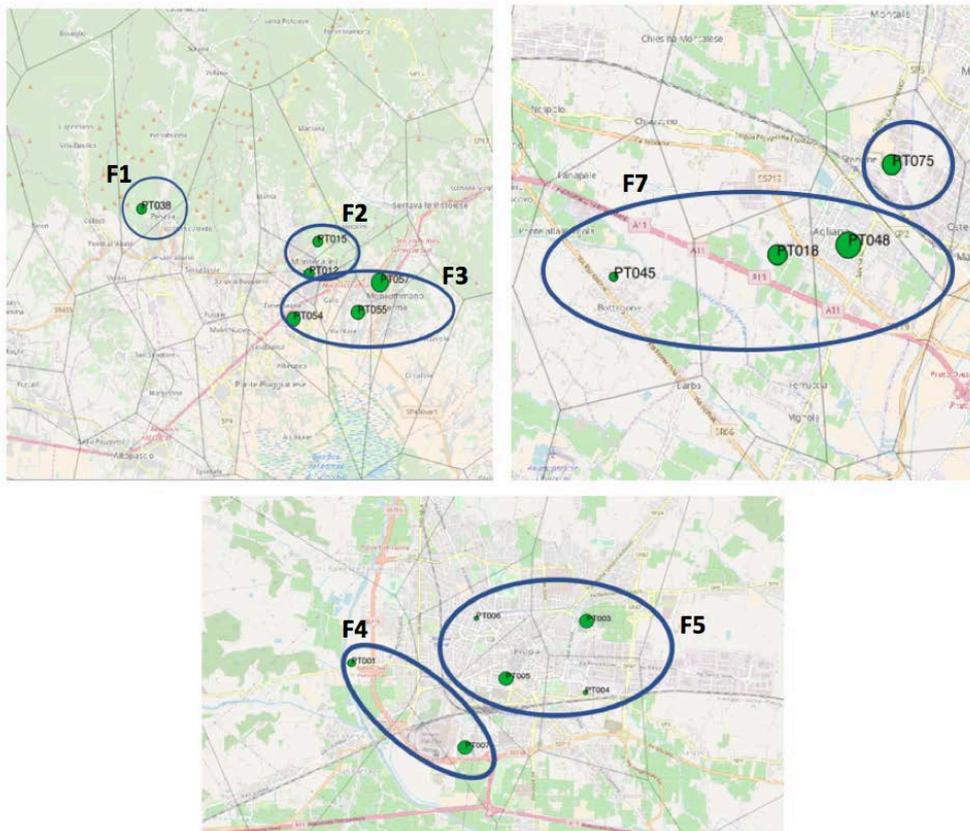


Figura 11 Antenne aggregate Pistoia

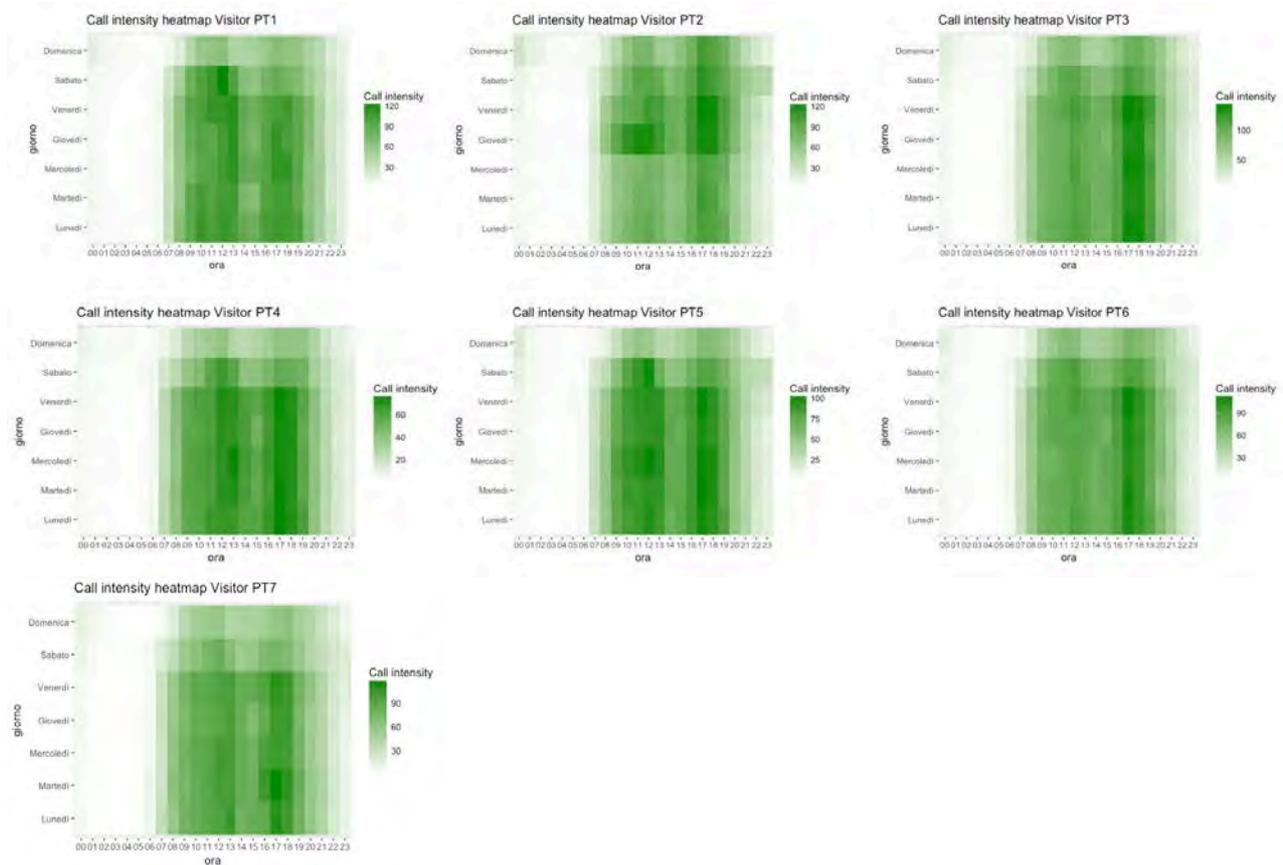


Figura 12 Heatmap visitor Prato

Co-presenza dei turisti nell'area di Firenze

Per comprendere il comportamento degli utenti classificati come *visitatori*, nell'area di Firenze, è stata compiuta un'analisi per individuare le aree della città 'toccate' dai visitatori durante la loro permanenza.

Sono stati utilizzati i dati relativi al mese di Ottobre 2015 e per ciascun utente sono state individuate le antenne a cui si sono collegati durante il loro soggiorno.

Le osservazioni sono state rappresentate tramite un grafo di co-occorrenza, i cui nodi sono le antenne unite fra di loro, sulla base del fatto che un determinato utente, nel proprio periodo di permanenza, si sia collegato ad esse. Il peso degli archi del grafo è rappresentato dal numero di utenti che si sono agganciati ad entrambe le antenne durante il proprio soggiorno a Firenze.

È stato deciso di verificare se il comportamento dei visitatori e l'utilizzo dello spazio urbano di questi, cambi all'aumentare del periodo di permanenza e se all'interno della città si possono individuare aree all'interno delle quali il comportamento del visitatore risulti omogeneo. Per indagare tali aspetti è stata effettuata una partizione del grafo massimizzando la *modularity (Louvain Method)*, che ci ha permesso di misurare come la rete possa essere scomposta in comunità, che consentono di raggruppare le antenne nella città Firenze.

Ogni community è stata rappresentata con un colore diverso, i colori attribuiti su ciascun grafo rappresentano la numerosità della comunità stessa (ad esempio il colore blu rappresenta sempre la comunità con la dimensione maggiore).

Di seguito rappresentiamo i grafi ottenuti, considerando il periodo di permanenza di almeno 1 giorno e almeno 5 giorni.

¹ Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, in *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10), P1000.

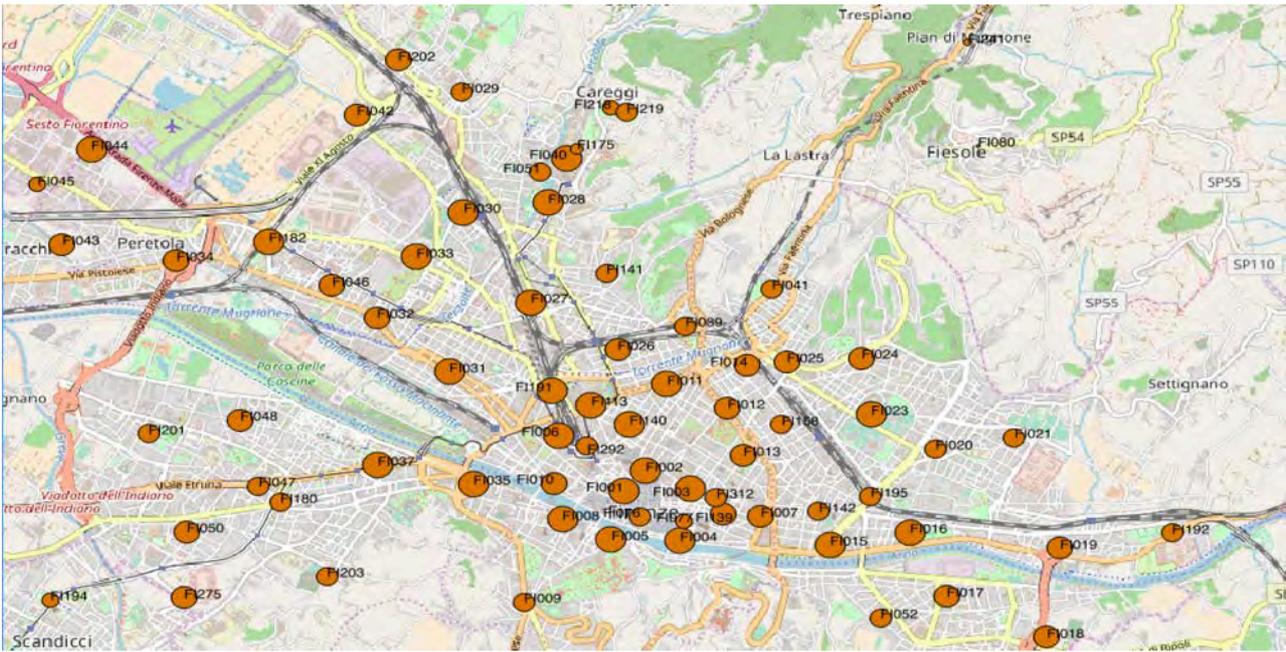


Figura 13 Count presenza antenne 1 giorni

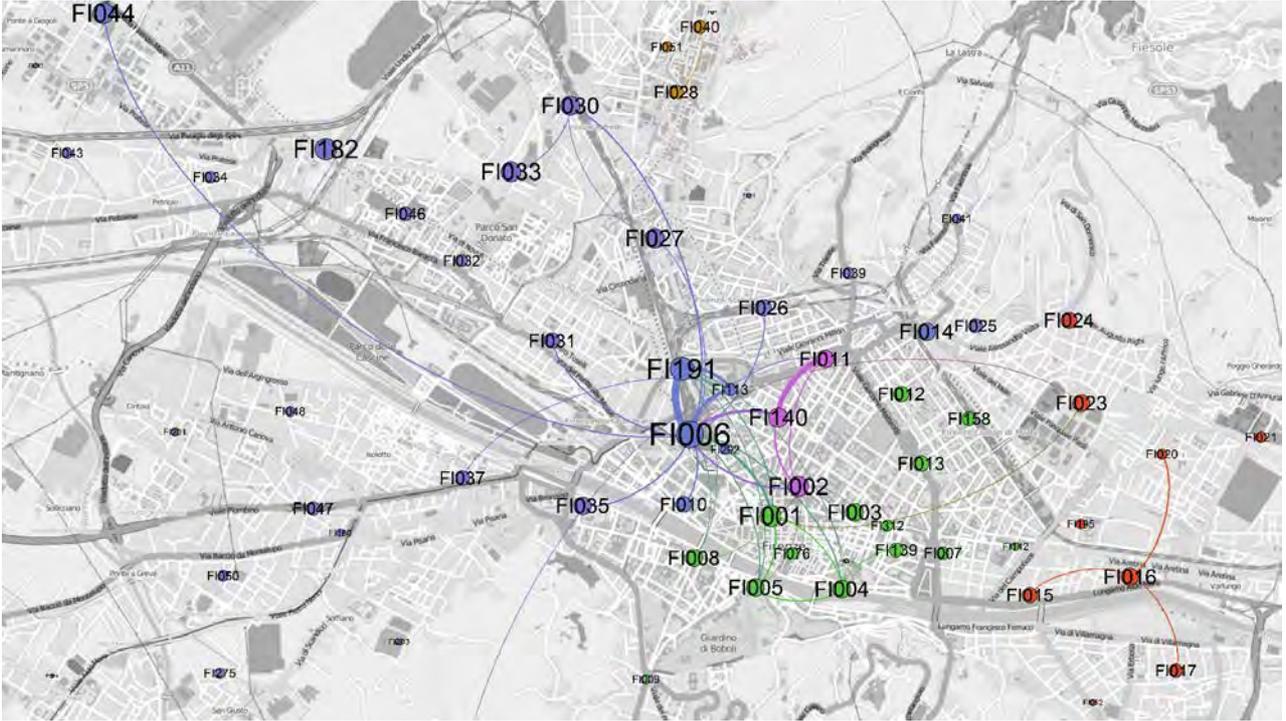


Figura 14 Grafo visitor 1 giorno

Nodes: 139
 Edges: 1712
 Undirected graph
 Modularity: 0,253
 Modularity with resolution: 0,253
 Number of Communities: 15

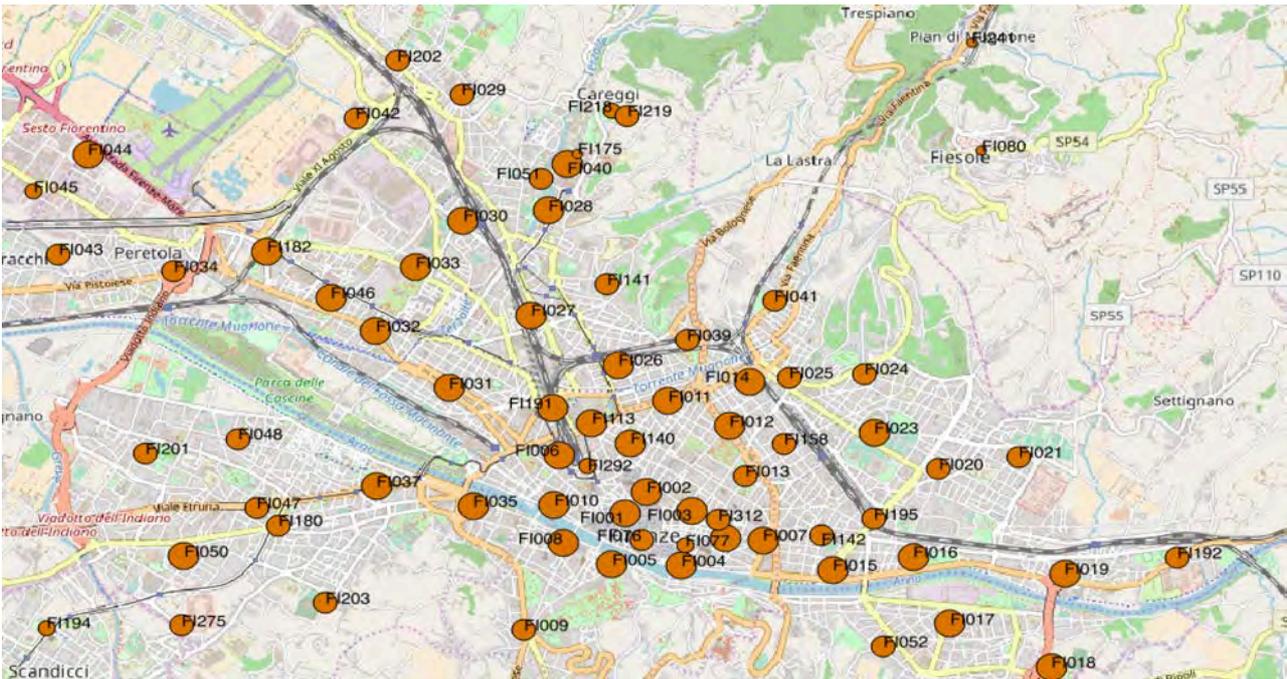


Figura 15 Count presenza antenna 5 giorni

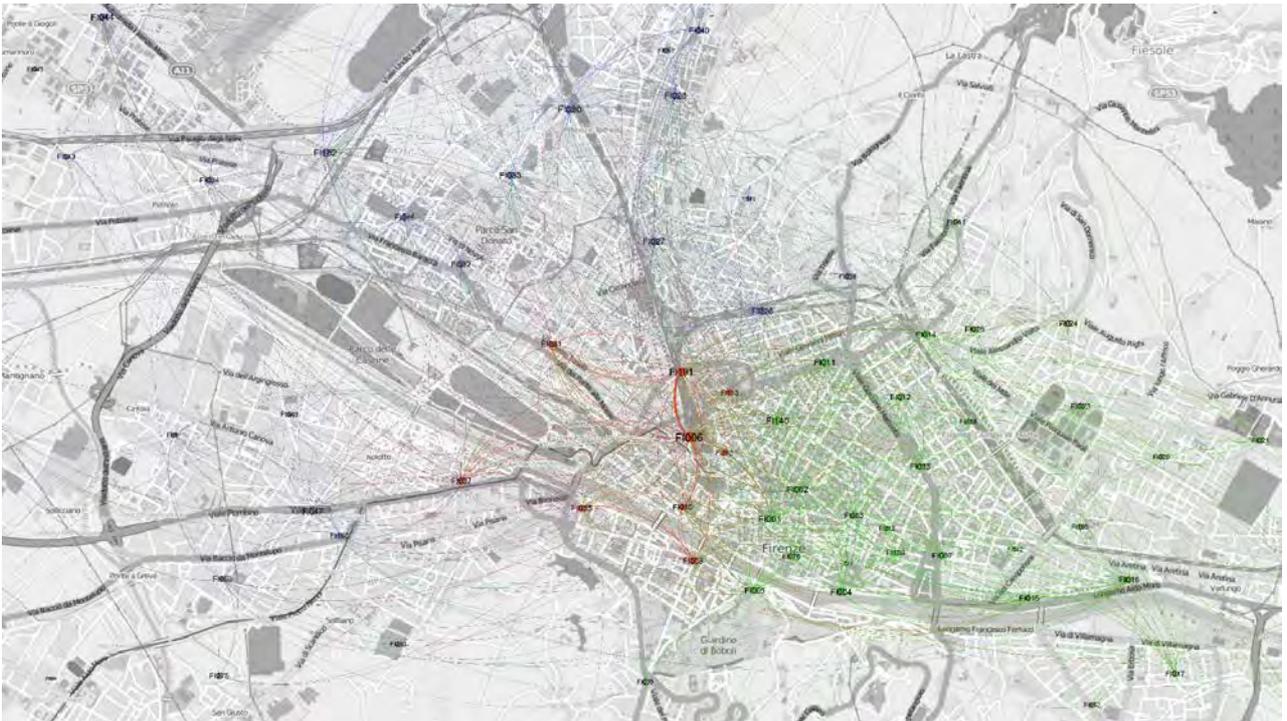


Figura 16 Grafo visitor 5 giorni

Nodes: 155
 Edges: 3814
 Undirected graph
 Modularity: 0,232
 Modularity with resolution: 0,232
 Number of Communities: 6

La scelta di ripetere l'osservazione su periodi di permanenza incrementali (da 1 a 5 giorni) è funzionale a verificare se al prolungarsi della permanenza in città, il comportamento dei visitatori cambi e vengano coinvolte aree urbane diverse.

I grafi ottenuti sembrano confermare l'ipotesi per cui, al crescere della durata del soggiorno, cresca di conseguenza l'estensione urbana coinvolta dal flusso dei visitatori.

Pare che i visitatori che soggiornano per un periodo più breve, concentrino i propri spostamenti nel centro storico, nei dintorni della stazione e delle principali attrazioni turistiche. Al crescere della permanenza aumentano gli spostamenti che coinvolgono zone più periferiche di Firenze.

Come si può osservare dai grafi, un punto nevralgico della città è la stazione di Santa Maria Novella (antenna FI006) a testimonianza del fatto che questa ricopre un ruolo di porta di accesso per molti visitatori. La stazione risulta il nodo con il grado maggiore, ma allo stesso tempo si può notare che, al crescere della durata della permanenza, l'antenna della stazione non risulta più appartenente alla community di dimensioni maggiori.

Questo a parziale conferma, che al crescere della durata di permanenza, vengono coinvolte dal flusso dei visitatori aree diverse.

L'analisi di *community detection* pare individuare aree della città in cui i visitatori potrebbero appartenere a categorie diverse. L'area del centro storico, fortemente collegata alla stazione centrale, per la quale i principali fruitori sono visitatori che si soffermano e risultano interessanti alle principali attrazioni turistiche della città.

L'area di Firenze sud/Campo di Marte in cui si collocano i principali luoghi destinati ad eventi, come ad esempio lo stadio Artemio Franchi, l'OBI Hall e il Mandela Forum, in cui i visitatori probabilmente appartengono alla categoria di coloro che si recano a Firenze in occasione di specifici eventi.

L'area di Firenze Nord in cui si collocano sia la principale struttura ospedaliera di Careggi, che i poli Universitari, il Nuovo Palazzo di Giustizia e il principale accesso autostradale – A11, che porta a ipotizzare che i visitatori appartengono a categorie che si recano a Firenze principalmente per motivi di lavoro/studio o motivi connessi alla presenza del polo ospedaliero.

Conclusioni

Il lavoro svolto si propone di dimostrare come i dati generati attraverso l'utilizzo dei dispositivi mobili, possa rappresentare un valido strumento per la comprensione del comportamento dei visitatori, come essi si spostino sul territorio nei diversi giorni della settimana e in fasce orarie diverse, come il comportamento possa variare in funzione delle aree coinvolte e delle caratteristiche socio-economiche di queste.

I dati GSM rappresentano solo un esempio, essi potrebbero essere integrati, con quelli provenienti da dispositivi dotati di GPS, che possono fornire un livello di granularità molto fine (ad esempio a livello di singola attrazione turistica). Dobbiamo comunque considerare che i dati GSM raccolgono una popolazione diversa rispetto ai GPS, permettendo di vedere l'uso della città a prescindere dai vincoli della rete stradale.

Anche i dati pubblici provenienti dall'utilizzo dei Social Network possono rappresentare sia una valida alternativa, in quanto meno costosi da reperire, sia un'integrazione in grado di fornire informazioni aggiuntive, ad esempio relative alle caratteristiche del visitatore (paese di provenienza, genere).

In aree urbane, che sempre più rappresentano poli di attrazione interessati da flussi di persone che si spostano per motivi di piacere, lavoro o necessità, diventa indispensabile approfondire la conoscenza e le dinamiche di tali fenomeni per poter elaborare le corrette policy di gestione, monitoraggio e governo del territorio.