

Introduzione alla programmazione Python per le scienze sociali

Michele Tizzoni

Dipartimento di Sociologia e Ricerca Sociale
Via Verdi 26, Trento
Ufficio 6, 3 piano



UNIVERSITÀ
DI TRENTO

Chi sono

- Michele Tizzoni
- Email: michele.tizzoni@unitn.it (contattatemi per richieste di colloquio)
- Ufficio: n°6, piano 3, edificio di sociologia
- Aree di ricerca: computational social sciences, epidemiologia computazionale, reti complesse
- Insegnamenti AA 22/23
 - Introduzione a Python (seminari di credito, DSRS)
 - Digital health and computational epidemiology (DISI)

Seminari

- Seminari di credito per competenze informatiche divisi in due parti:
 - Parte A, 24h, introduttivo (marzo - aprile 2023)
 - Parte B, 24h, avanzato (aprile - maggio 2023)
 - La parte A è propedeutica alla B.

Seminari

- Seminari di credito per competenze informatiche divisi in due parti:
 - Parte A, 24h, introduttivo (marzo - aprile 2023)
 - Parte B, 24h, avanzato (aprile - maggio 2023)
 - La parte A è propedeutica alla B.
- **Modalità di esame:**
 - **Esame scritto in laboratorio con alcuni semplici esercizi da risolvere, simili a quelli proposti durante il corso.**
 - **Un appello da definire alla fine della parte A (metà aprile)**
 - **Due appelli da definire alla fine della parte B (giugno/luglio)**
 - **Il superamento dell'esame per la parte B convalida automaticamente anche i crediti della parte A**

Programma parte A

Data	Argomento	Orario	Laboratorio
Lunedì 27/02/2023	Introduzione al corso	15-17	13
Martedì 28/02/2023	Scripting, e uso di Jupyter notebooks.	14-17	13
Lunedì 06/03/2023	Basi di Python (variabili, sintassi)	15-17	13
Martedì 07/03/2023	Esercitazione	14-17	13
Lunedì 13/03/2023	Stringhe, liste, tuple, dizionari	15-17	13
Martedì 14/03/2023	Controllo di flusso	14-17	13
Lunedì 20/03/2023	Lettura e scrittura di file. Formati dati.	15-17	13
Martedì 21/03/2023	Funzioni, gestione errori e testing	14-17	13
Lunedì 27/03/2023	Introduzione alla visualizzazione grafica con matplotlib	15-17	13
Martedì 28/03/2023	Esercitazione	14-16	13

Programma parte B

Data	Argomento	Orario	Laboratorio
Martedì 4/04/2023	Introduzione al corso	14-16	13
Lunedì 17/04/2023	Introduzione a Pandas	15-17	13
Martedì 18/04/2023	Formato dati, parsing, lettura di file	14-16	13
Martedì 2/05/2023	Statistiche descrittive con Pandas	14-17	13
Lunedì 8/05/2023	Visualizzazione dati con matplotlib e seaborn	15-17	13
Martedì 9/05/2023	Esercitazione	14-17	13
Lunedì 15/05/2023	Introduzione alla visualizzazione di dati georeferenziati	15-17	13
Martedì 16/05/2023	Visualizzazione di dati georeferenziati con Geopandas	14-17	13
Lunedì 22/05/2023	Introduzione alla raccolta dati tramite API	15-17	13
Martedì 23/05/2023	Introduzione all'analisi di reti sociali con Networkx	14-17	13

Materiali

- Le lezioni si baseranno ampiamente sui materiali sviluppati dal Dr. David Leoni per i seminari dello scorso anno
 - SoftPython: <https://it.softpython.org/>
- Le slides saranno disponibili sul mio repository GitHub
 - https://github.com/micheletizzoni/Python_for_social_sciences

Questionario di ingresso

- A questo link trovate un breve questionario di ingresso che vi chiedo di compilare:
- <https://forms.gle/cGMF29VP9EEnrwtV9>

Grazie!

Perché programmare nelle
scienze sociali?

The big picture

- ◆ L'immagine digitale del mondo riproduce una copia sempre più fedele del mondo reale, fisico.
- ◆ Questo consente di usare metodi computazionali per misurare delle ricorrenze e definire dei nessi causali, attraverso metodi di statistica, machine learning, data mining
- ◆ Possiamo mettere in relazione le tracce digitali lasciate dall'attività umana di ogni giorno con misure empiriche ottenute con metodi tradizionali (indagini a campione, osservazioni sul campo)
- ◆ Possiamo sviluppare modelli predittivi del comportamento umano su larga scala e testarli quasi in tempo reale

New York City

1.1 miliardi di chiamate taxi
6 anni
350 Gb di dati





facebook

P. Butler
December 2010

tracce digitali del comportamento umano

[slide by C. Cattuto]

Italia:

97% della popolazione possiede uno smartphone

84% della popolazione accede a internet

6+ ore al giorno online

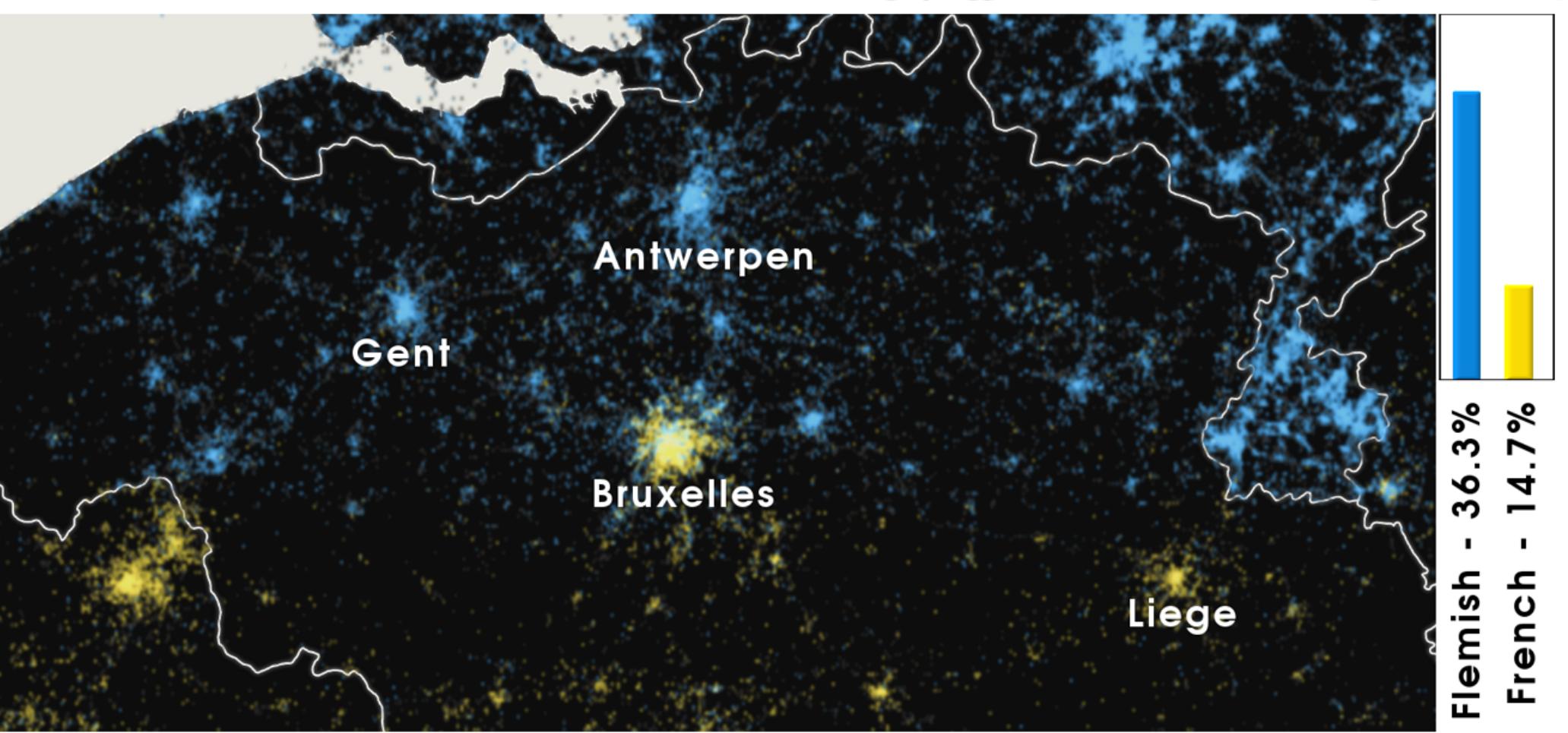
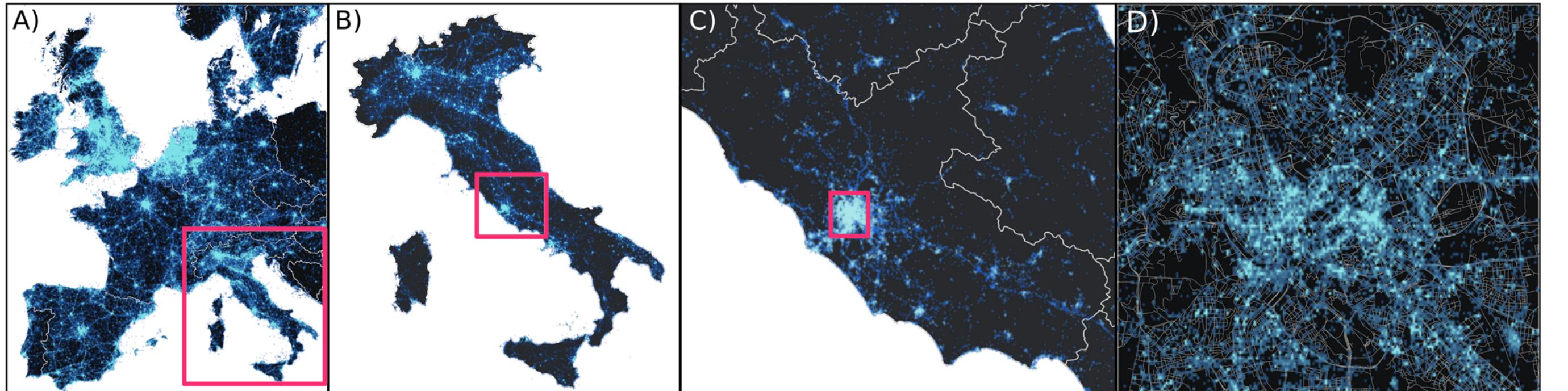
2 ore al giorno su social media

tracce digitali

disponibili come effetto collaterale di attività ordinarie
alto livello di copertura, accesso alle grandi scale
possibilità di elaborazione automatica

prospettiva storica
orizzonte temporale limitato
riproducibilità limitata
contesto limitato
privacy e protezione dei dati

Demografia digitale



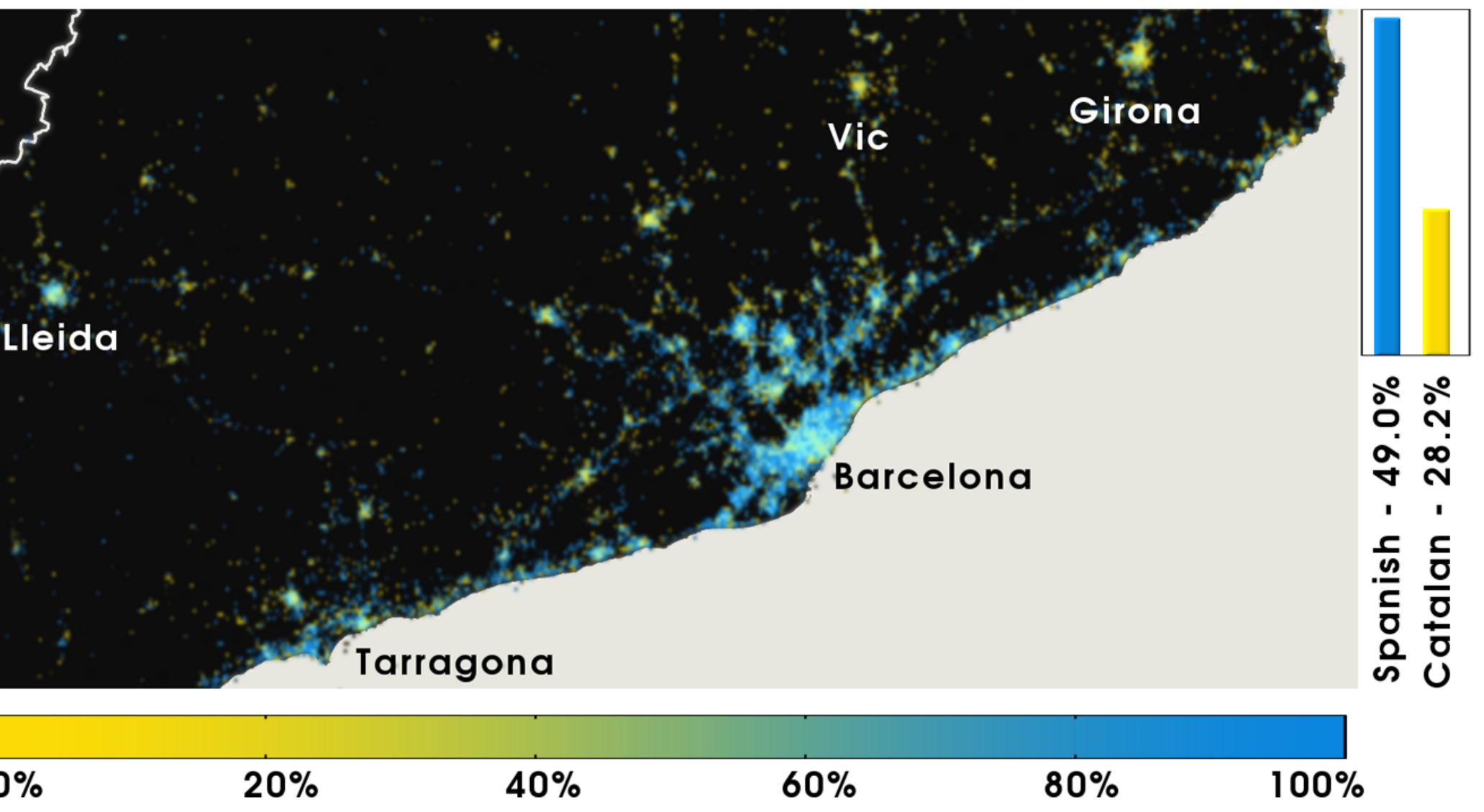
PLOS ONE

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

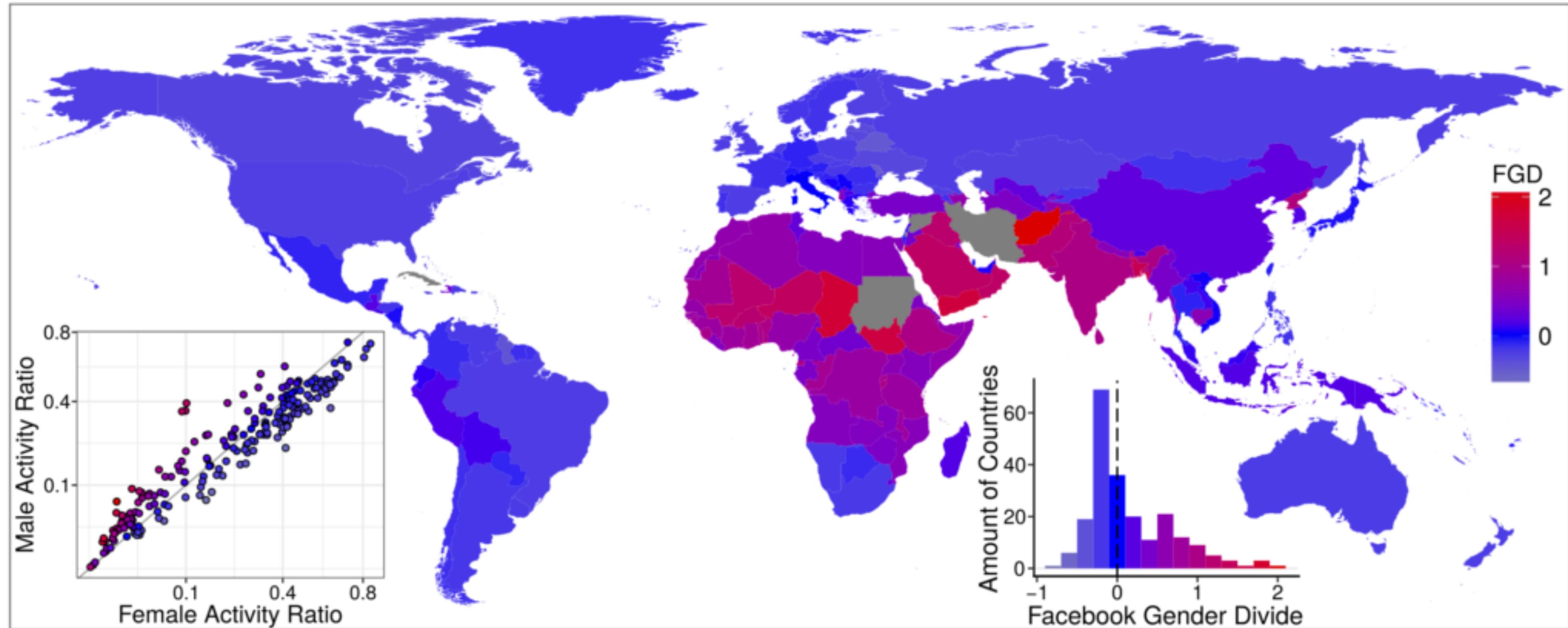
The Twitter of Babel: Mapping World Languages through Microblogging Platforms

Delia Mocanu, Andrea Baronchelli, Nicola Perra, Bruno Gonçalves, Qian Zhang, Alessandro Vespignani

Published: April 18, 2013 • <https://doi.org/10.1371/journal.pone.0061981>

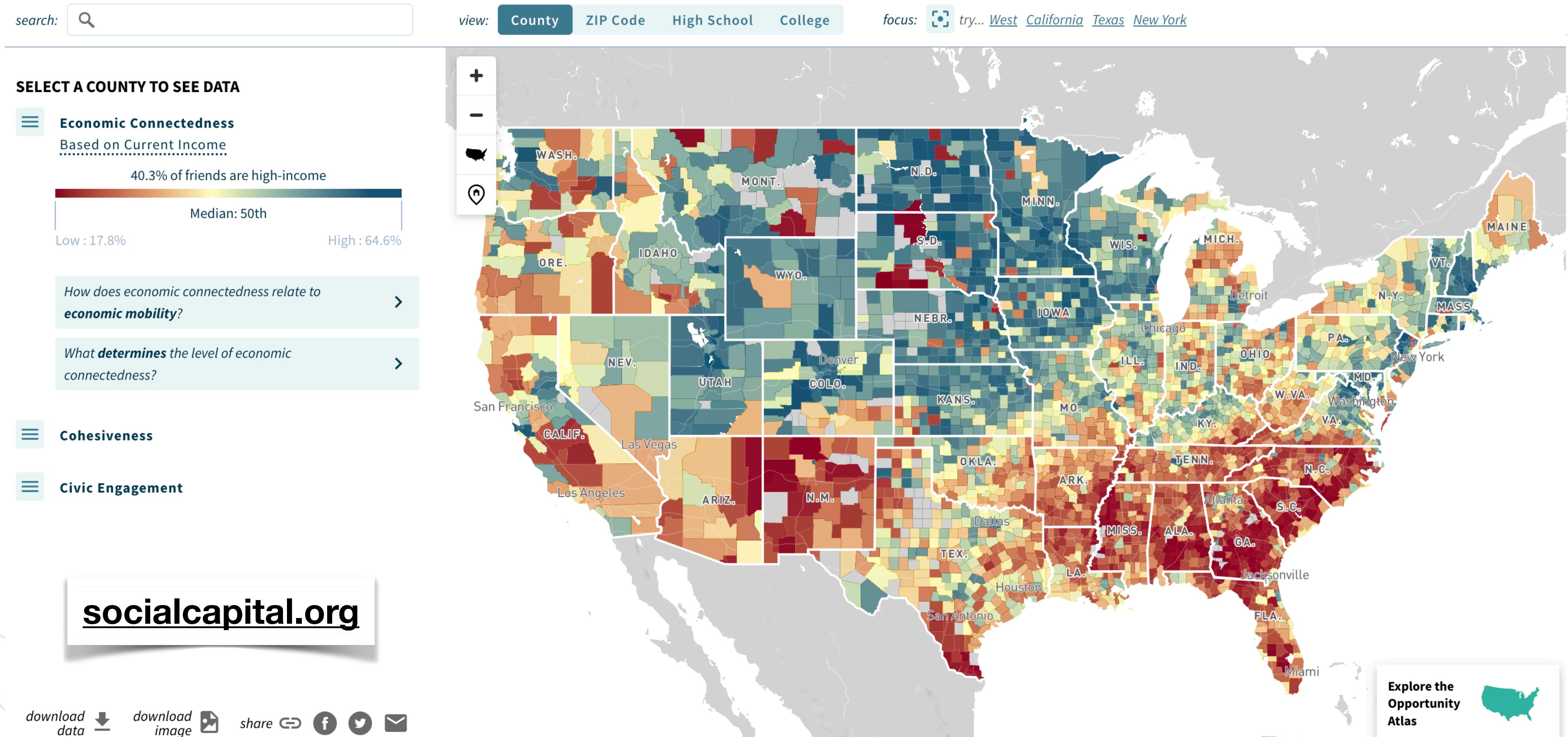


Disuguaglianze di genere



Garcia et al. PNAS 2018

Capitale sociale



Dati nel settore umanitario



Search Datasets

DATA

LOCATIONS

ORGANISATIONS

DATAVIZ ▾

ADD DATA

data.humdata.org

The Humanitarian Data Exchange

Find, share and use humanitarian data all in one place

[LEARN MORE](#)

FIND DATA

Search Datasets



20,538

DATASETS

254

LOCATIONS

1,856

SOURCES

ADD DATA



Make your dataset available
on HDX

[UPLOAD FILE](#)



HDX Connect: let others request
your data

[ADD METADATA](#)

Learn how the HDX team supports [responsible data sharing](#).

La pandemia COVID-19

- ◆ Per la prima volta nella storia, durante una pandemia, è stato possibile osservare e misurare il comportamento delle persone su larga scala attraverso l'analisi tracce digitali
- ◆ Analisi della mobilità umana con dati di telefonia mobile, geolocalizzati
- ◆ Analisi della diffusione di informazione e mis-informazione su social media
- ◆ Misura dei cambiamenti nei comportamenti quotidiani (sonno/veglia, socialità) attraverso dati da piattaforme digitali



Mobilità



See how your community is moving around differently due to COVID-19

As global communities respond to COVID-19, we've heard from public health officials that the same type of aggregated, anonymized insights we use in products such as Google Maps could be helpful as they make critical decisions to combat COVID-19.

These Community Mobility Reports aim to provide insights into what has changed in response to policies aimed at combating COVID-19. The reports chart movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential.



Apple Maps Mobility Trends Reports

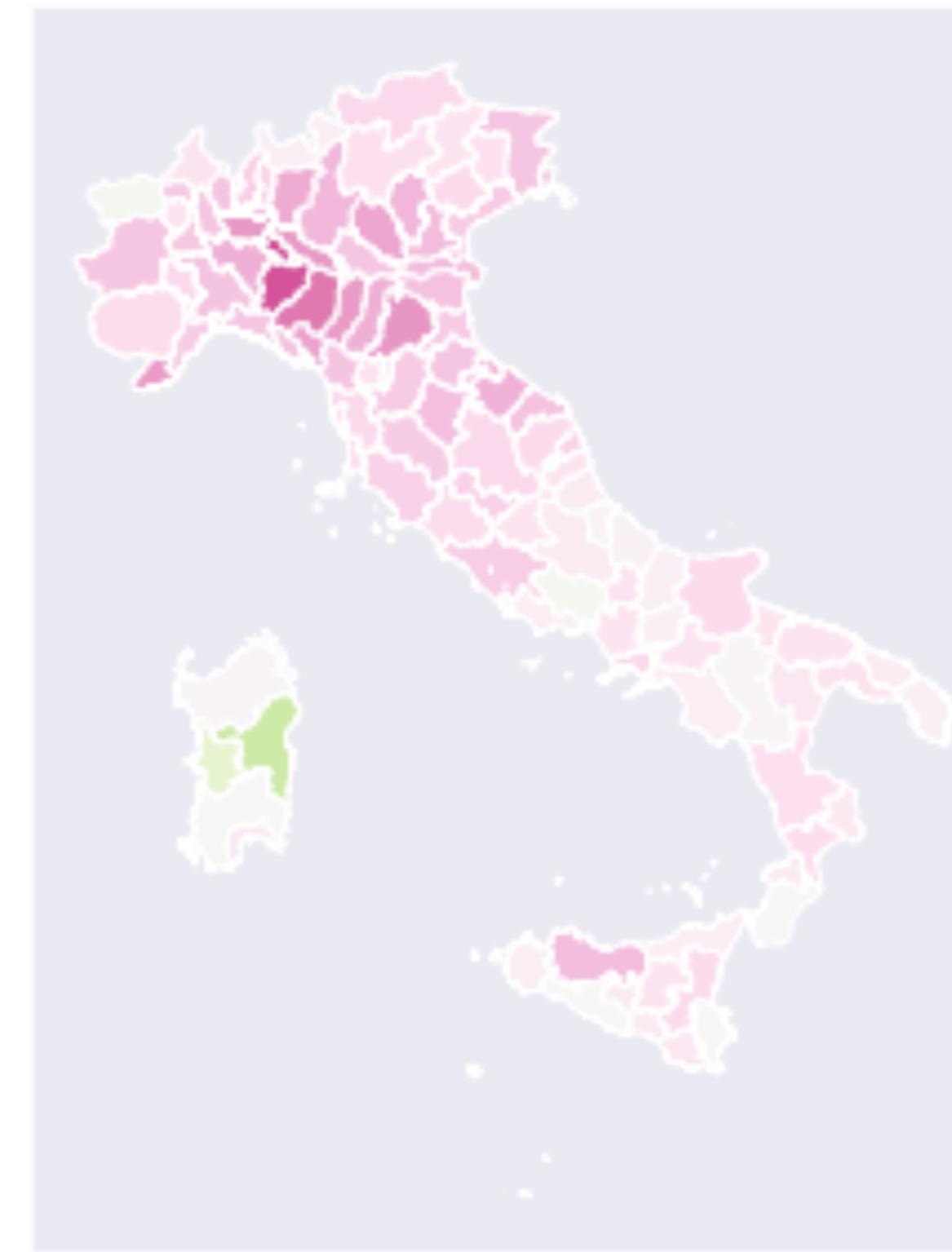
Learn about COVID-19 mobility trends. Reports are published daily and reflect requests for directions in Apple Maps. Privacy is one of our core values, so Maps doesn't associate your data with your Apple ID, and Apple doesn't keep a history of where you've been.



Mobilità



February 21-28

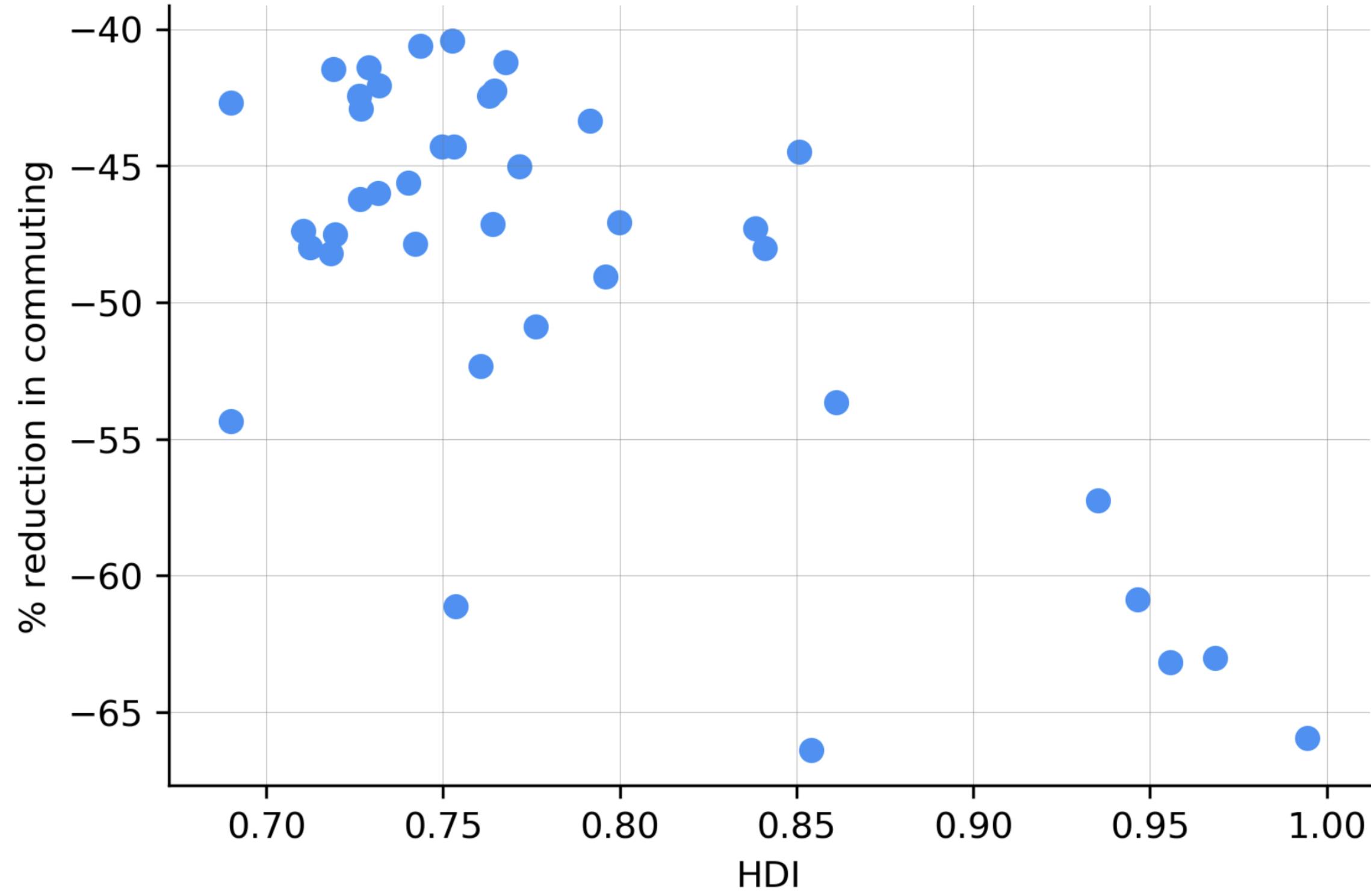
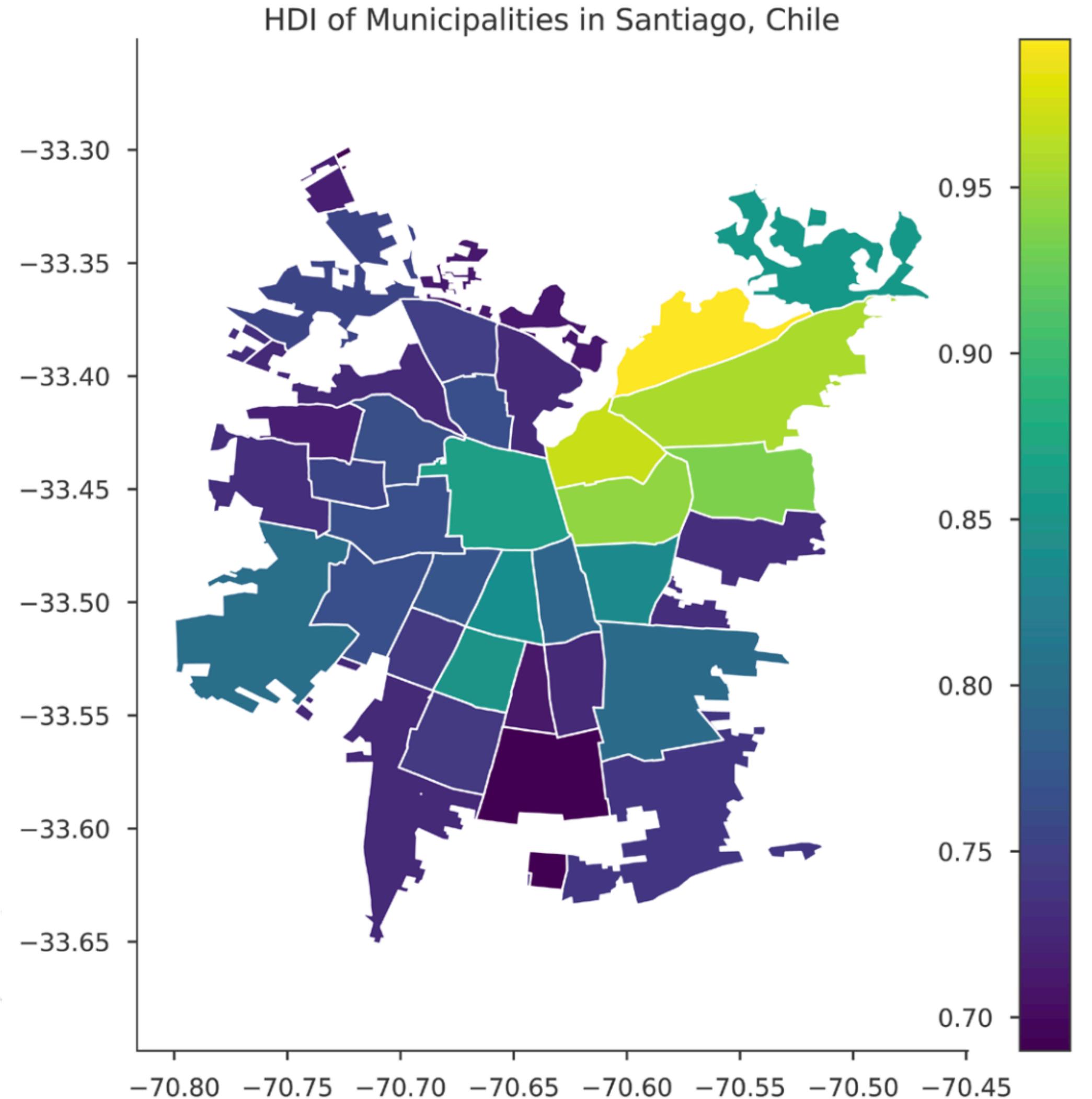


February 29 -
March 6



March 7 - 13

Effetti socio-economici



Breve intro a Python