



ANTICIPARE LA CRESCITA CON LE NUOVE COMPETENZE SUI BIG DATA – EDIZIONE 2

Operazione Rif. PA 2019-11596/RER “Anticipare la crescita con le nuove competenze sui Big Data”, approvata dalla Regione Emilia-Romagna con DGR n° 789 del 20 maggio 2019 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CENTRO NAZIONALE DI RICERCA E INNOVAZIONE



Programma della lezione

- Cos'è l'analisi geospaziale
- Applicazioni
- Tipi di file
- Librerie Python disponibili
- GeoPandas
- OpenStreetMap

Cos'è l'analisi geospaziale?

- È disciplina che si occupa di acquisire, gestire, analizzare e interpretare **dati geospaziali**
- I dati geospaziali sono raccolti attraverso strumenti come
 - GPS, sensori remoti (ad esempio satelliti), rilevamenti topografici, cartografia digitale e altri mezzi di acquisizione dati geografici

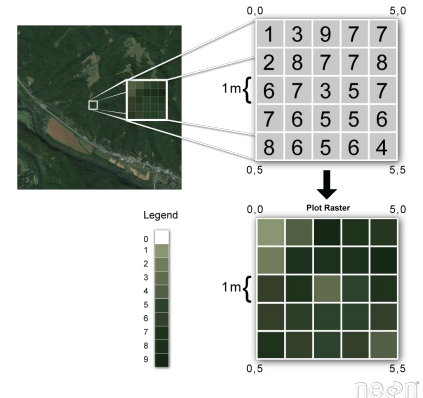
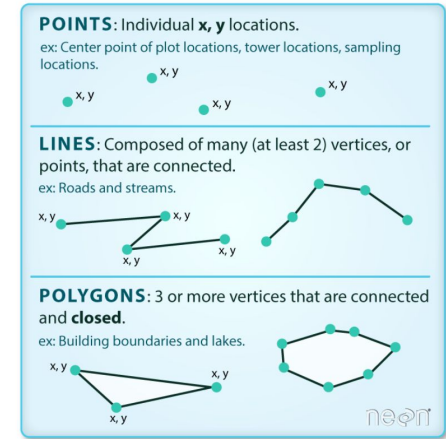


Applicazioni

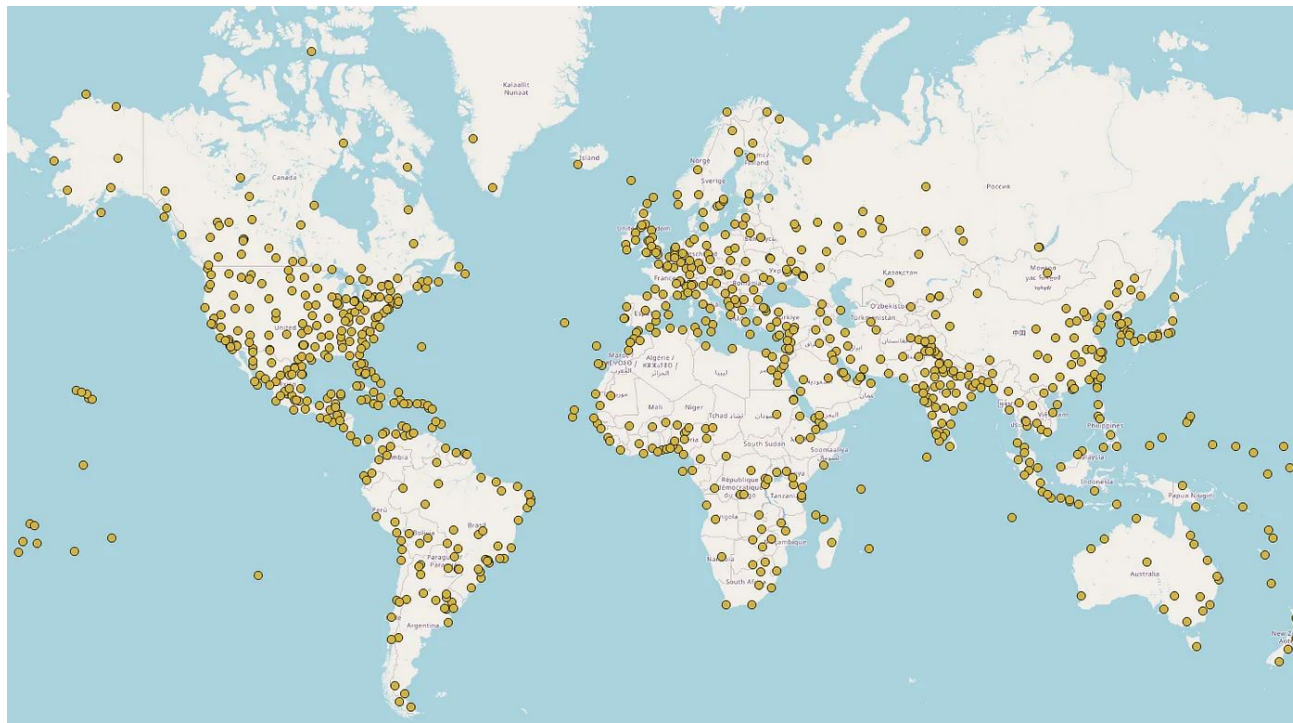
- **Analisi di mercato**
 - Esempi: individuare gruppi di clienti con caratteristiche comuni in segmenti geografici, scegliere dove aprire un nuovo negozio
- **Gestione di emergenze**
 - Esempi: individuare aree a rischio di catastrofi naturali, comunicare con il pubblico in tempo reale durante un'emergenza tramite l'analisi di immagini satellitari, analisi dei danni post-emergenza
- **Urbanistica e Pianificazione del territorio**
 - Esempi: ottimizzare le reti stradali, pianificare il trasporto pubblico, monitorare le modifiche nel tempo del territorio per valutare l'evoluzione di una città nel tempo

Due tipi di dati principali

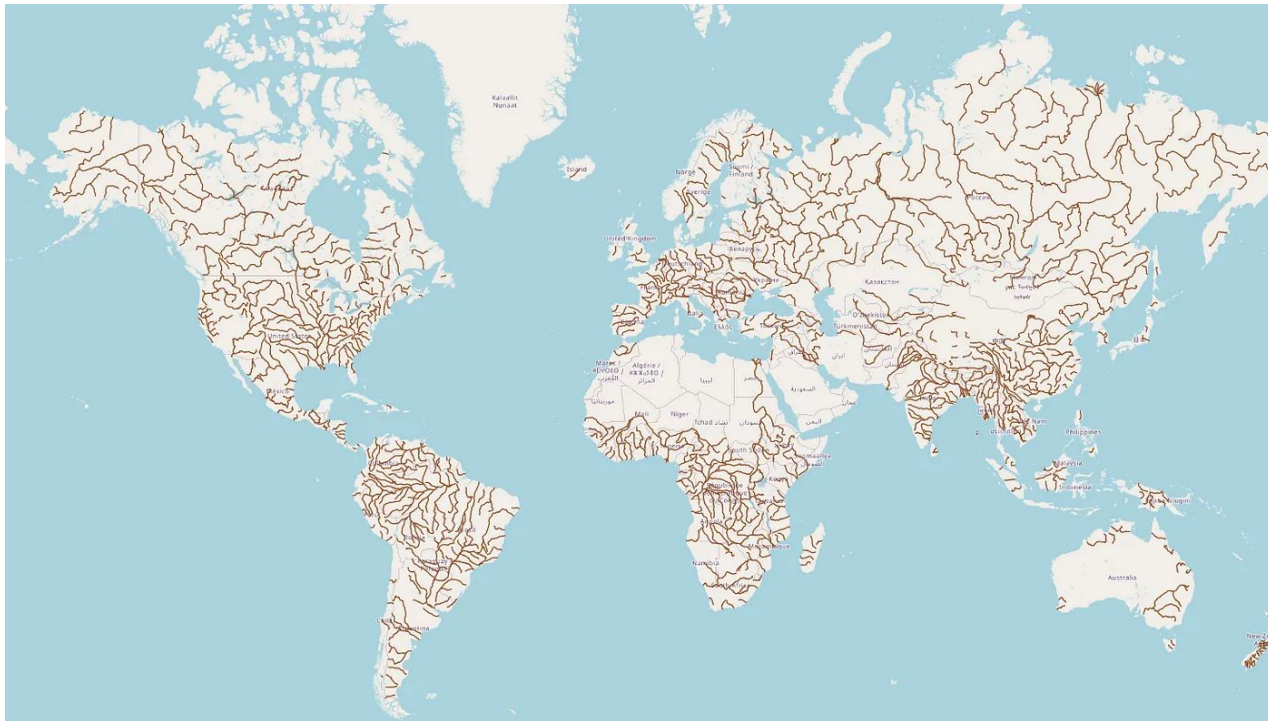
- **Dati vettoriali** = dati tabulari che contengono caratteristiche specifiche sulla superficie terrestre e assegnano attributi a tali caratteristiche
- **Dati raster** = griglia di valori visualizzati su una mappa come pixel.
 - Ogni valore di pixel rappresenta un'area sulla superficie terrestre



Dati puntuali



Dati di tipo linea



Dati di tipo poligono



Dati raster



Tipi di file

Il file più comune è lo Shapefile. Quando lo scarichi, hai un file zip che contiene i seguenti file che sono collegati tra di loro:

- **.shp** è il file più importante che fornisce la geometria
- **.dbf** contiene i campi non geografici che permettono di capire il contesto dei dati geospaziali
- **.shx** fornisce l'indice della geometria
- **.prj** contiene informazioni sulla proiezione cartografica utilizzata per rappresentare i dati geografici

Proiezione geografica



(lon, lat)



(x, y)

Librerie Python

- **geopandas** per lavorare con dati vettoriali
- **rasterio** per lavorare con dati raster
- **Shapely** per manipolare dati vettoriali
- **folium** per visualizzare mappe interattive

3 fonti di dati

1. Natural Earth Data
2. USGS Earth Explorer
3. OpenStreetMap

<https://gisgeography.com/best-free-gis-data-sources-raster-vector/>



GeoPandas 0.14.0

GeoPandas is an open source project to make working with geospatial data in python easier. GeoPandas extends the datatypes used by [pandas](#) to allow spatial operations on geometric types. Geometric operations are performed by [shapely](#). Geopandas further depends on [fiona](#) for file access and [matplotlib](#) for plotting.

Description

The goal of GeoPandas is to make working with geospatial data in python easier. It combines the capabilities of pandas and shapely, providing geospatial operations in pandas and a high-level interface to multiple geometries to shapely. GeoPandas enables you to easily do operations in python that would otherwise require a spatial database such as PostGIS.

[Getting started](#)

[Documentation](#)

[About GeoPandas](#)

[Community](#)

☰ On this page

[Description](#)

[Useful links](#)

[Supported by](#)

[Indices and tables](#)

 [Show Source](#)

Dataframe Vs Series

Series

	apples
0	3
1	2
2	0
3	1

+

Series

	oranges
0	0
1	3
2	7
3	2

=

DataFrame

	apples	oranges
0	3	0
1	2	3
2	0	7
3	1	2

GeoDataFrame



Dati del laboratorio

- I file geografici delle ripartizioni geografiche, regioni, province, città metropolitane e comuni dal sito dell'istat: <https://www.istat.it/it/archivio/222527>
(Versione generalizzata, 2023, Comune)
- tavole di dati Ambiente urbano: <https://www.istat.it/it/archivio/281184>

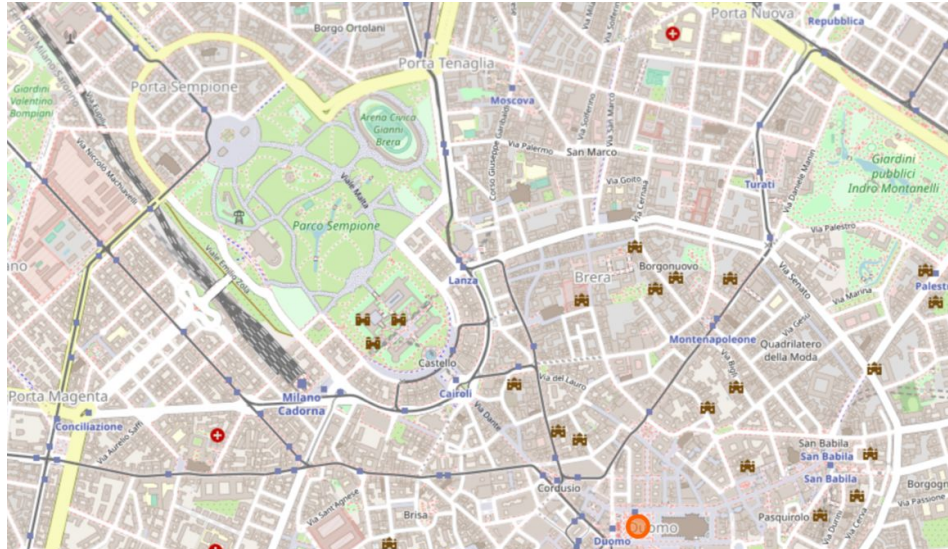
Trovate i file necessari nella mia repository GitHub:

<https://github.com/eugeniarling/machine-learning-course-eugenia/tree/main/data>

Introduzione a OpenStreetMap

OpenStreetMap è il più grande progetto di Informazione Geografica Volontaria

- Iniziato da Steve Coast in UK nel 2004

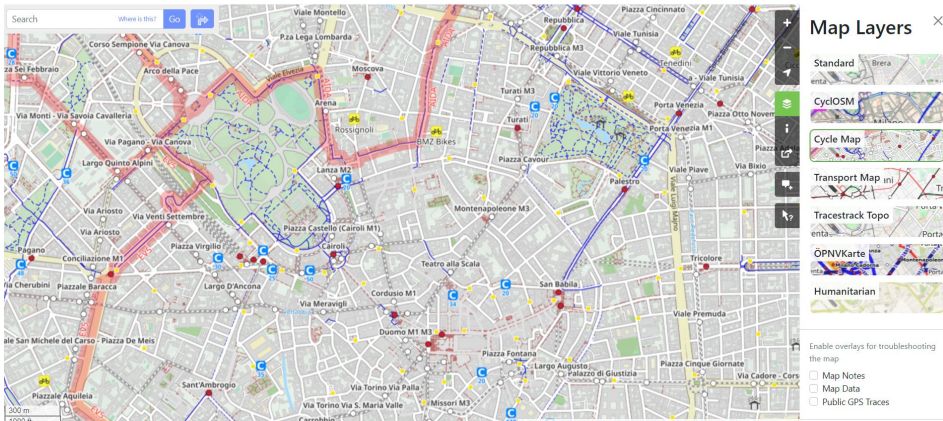


Un database, tante mappe

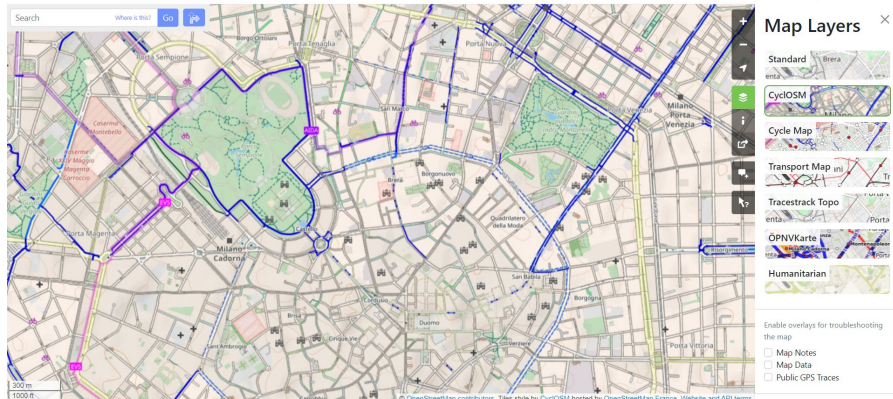
OpenStreetMap Edit History Export GPS Traces User Diaries Communities Copyright Help About Log In Sign Up



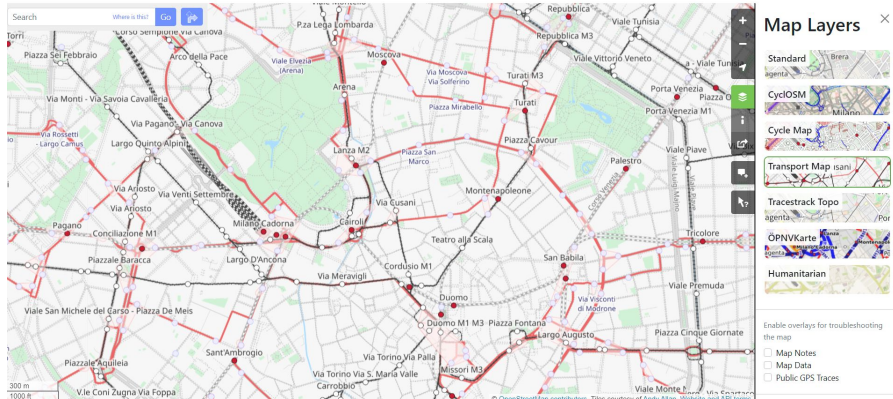
OpenStreetMap Edit History Export GPS Traces User Diaries Communities Copyright Help About Log In Sign Up



OpenStreetMap Edit History Export GPS Traces User Diaries Communities Copyright Help About Log In Sign Up

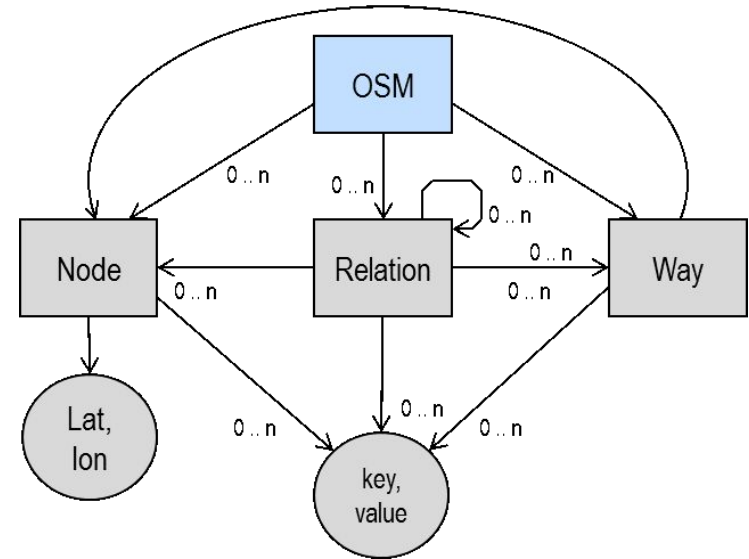


OpenStreetMap Edit History Export GPS Traces User Diaries Communities Copyright Help About Log In Sign Up



Struttura dei dati

- **Nodo** = singolo punto, identificato da (lat,lon)
 - Oggetti puntuali: ristorante, bar, scuole
- **Way** = insieme ordinato da 2 a 2000 nodi
 - Oggetti lineari: strada, fiume
 - Oggetti areali: cattedrali, parchi molto grandi
- **Relazione** = insieme ordinato di oltre 2000 nodi
 - Esempi: linee di trasporto
- **Tag** = descrizione degli elementi geografici, che consiste di due componenti, chiave e valore
 - Esempio: building=supermarket, building = hotel



Nodo

Search Where is this? Go

Node: Unicredit (4275242130)

Version #2

Created 20 gift shops, 19 clothes shops, and 45 other objects; Updated a bank, a cafe, and 3 other objects

Edited about 6 years ago by Humus015

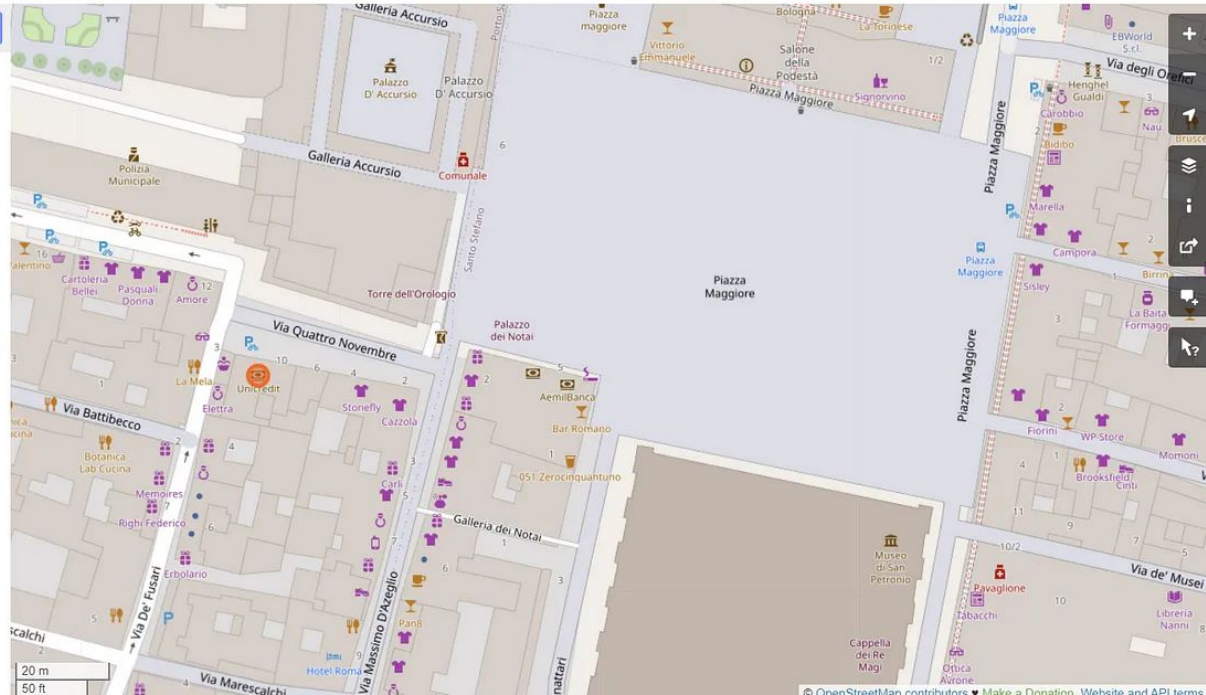
Changeset #47997954

Location: 44.4934888, 11.3417943

Tags

amenity	bank
atm	yes
name	Unicredit

[Download XML](#) · [View History](#)



Way

OpenStreetMap

Edit

History

Export

[GPS Traces](#) [User Diaries](#) [Communities](#) [Copyright](#) [Help](#) [About](#)

Way: Via Castiglione (95408276)

Version #9

Precisazione: inutile mettere i marciapiedi di fianco a tutte le strade; conviene metterli solo quando sono in sede propria, altrimenti si crea solo confusione

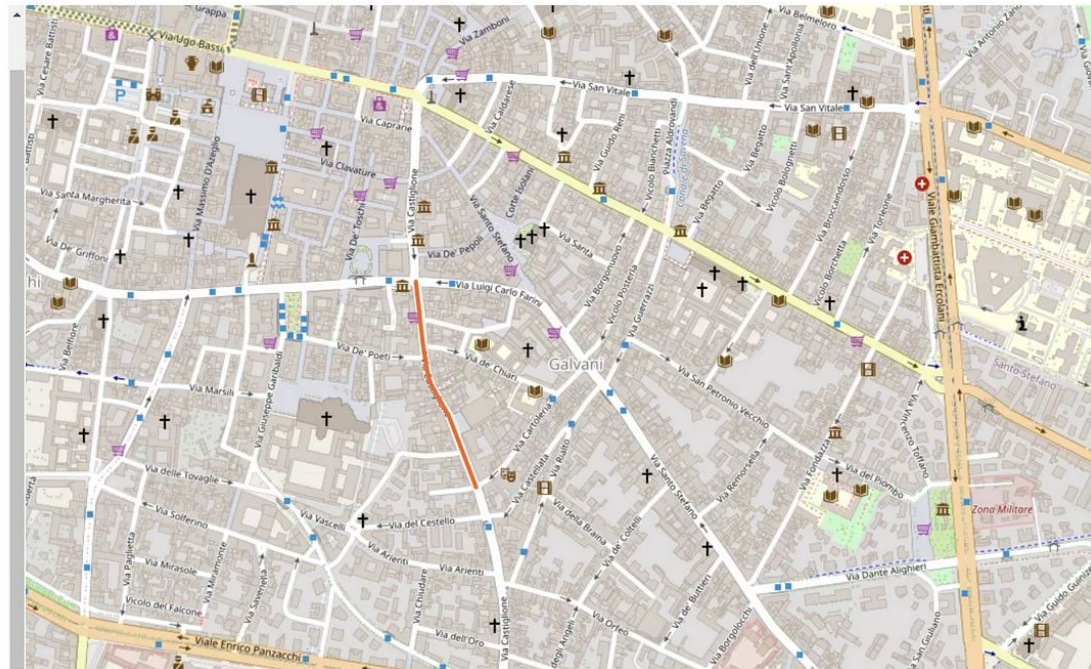
Edited almost 2 years ago by [Aldeyde](#)
Changeset #105146221

Tags

highway	tertiary
maxspeed	30
name	Via Castiglione
oneway	yes
surface	asphalt
zone:maxspeed	IT:30

Nodes

► 12 nodes



Relation



[GPS Traces](#) [User Diaries](#) [Communities](#) [Copyright](#) [Help](#) [About](#)

Search

Where is this?

Go

Relation: 3573271

Version #7

fix inner edificio, ripristinato tag infopoint

Edited over 1 year ago by [ivanbranco](#)

Changeset #115371273

Tags

landuse	residential
ref.corine_land_cover	1111
source	http://dati.emilia-romagna.it/dato/ite_m/11-11-uso-del-suolo-2008.html
type	multipolygon

Members

▼ 9 members

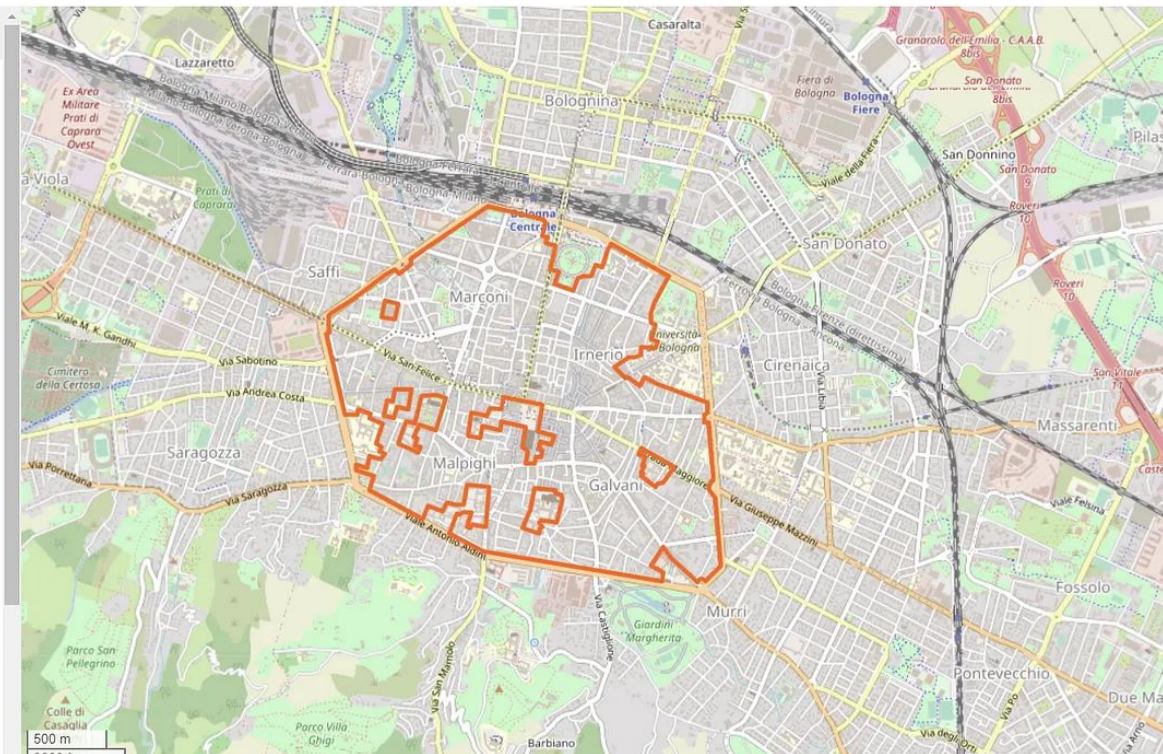
Way 265933367 as outer

Way 265933370 as inner

Way 265933372 as inner

Way 619538803 as inner

Way 967089902 as inner



Modi per scaricare i dati da OpenStreetMap

- Manualmente dal sito OpenStreetMap per dati “piccoli”, per esempio una città
- [Geofabrik](#) per scaricare dati di stati/continenti
- Pyrosm è una libreria Python che permette di scaricare, leggere e manipolare dati da OpenStreetMap

Risorse utili

- Documentazione di [GeoPandas](#)
- [Corso di geospatial analysis di Trento](#)
- [Serie di tutorial della geospatial data analysis](#)