

# Nearest Neighbor Analysis

QGIS Tutorials and Tips



Author

Ujaval Gandhi

<http://google.com/+UjavalGandhi>

Translations by

Pino Nicolosi a.k.a Rattus

## Analisi di prossimità

Il GIS è molto utile nell'analizzare le relazioni spaziali tra le geometrie. Un'analisi di questo genere consiste nello scoprire quali geometrie sono più vicine a un dato elemento. QGIS dispone di uno strumento chiamato "Matrice di distanza" che è di grande aiuto in questo tipo di indagini. In questo tutorial useremo 2 dataset e scopriremo quali punti in un primo layer sono i più vicini a determinati punti presenti in secondo layer.

## Descrizione dell'esercizio

Data la localizzazione di tutti i più importanti terremoti conosciuti, andremo ad individuare i luoghi popolati più vicini a ciascuno dei punti in cui si sono verificati i terremoti.

### ***Altri aspetti che avremo modo di apprendere nel corso dell'esercizio.***

- Come realizzare delle operazioni di unione tabellare in QGIS. (Vedete anche [Realizzare unioni tabellari](#) qualora vi occorranو spiegazioni dettagliate).

## Ottenere i dati necessari

Useremo il [Significant Earthquake Database](#) come nostro layer di riferimento per rappresentare i principali terremoti. Scaricate il t`tab-delimited earthquake data <[http://www.ngdc.noaa.gov/nndc/struts/results?type\\_0=Exact&query\\_0=\\$ID&t=101650&s=13&d=189&dfn=signif.txt](http://www.ngdc.noaa.gov/nndc/struts/results?type_0=Exact&query_0=$ID&t=101650&s=13&d=189&dfn=signif.txt)>`\_

Natural Earth fornisce un prezioso dataset dei luoghi abitati. Scaricate il [simple \(less columns\) dataset](#)

Fonte Dati: [NGDC] [NATURALEARTH]

## Procedimento

1. Aprite *Layer* ■ *Aggiungi layer testo delimitato...* e individuate il file appena scaricato `signif.txt` .



2. Dal momento che si tratta di tab-delimited file, scegliete *Tab* alla voce *formato file*. I campi X e Y saranno compilati automaticamente. Click su *OK*.

### Note

Vedrete comparire dei messaggi di errore mentre QGIS sta importando il file. Si tratta di errori di incolonnamento e alcune righe del file non saranno importate. Per quelli che sono gli scopi di questo esercizio questi errori possono essere trascurati senza conseguenze. Chiudete.

**Create a Layer from a Delimited Text File**

File Name:

Layer name:  Encoding:

File format: ☐ CSV (comma separated values) ☒ Custom delimiters ☐ Regular expression delimiter

☐ Comma 
 ☒ Tab 
 ☐ Space 
 ☐ Colon 
 ☐ Semicolon

Other delimiters:  Quote:  Escape:

Record options: Number of header lines to discard:  ☒ First record has field names

Field options: ☐ Trim fields ☐ Discard empty fields ☐ Decimal separator is comma

Geometry definition: ☒ Point coordinates ☐ Well known text (WKT) ☐ No geometry (attribute only table)

☒ DMS coordinates  
 X field:  Y field:

Layer settings: ☒ Use spatial index ☐ Use subset index ☐ Watch file

	I_D	FLAG_TSUNAMI	YEAR	MONTH	DAY	HOUR	MINUTE	SECOND	FOCAL_DEPTH	EQ_MAG_MW	EQ_MAG
1	1		-2150								
2	2	Tsu	-2000								
3	3		-2000						18		7.1
4	8		-1566								
5	11		-1450								

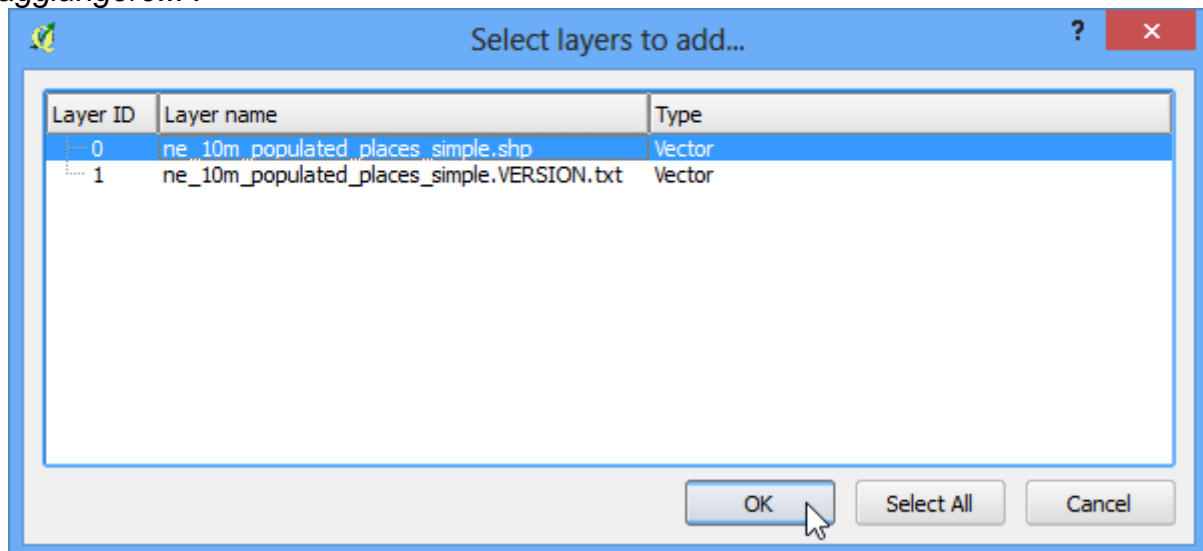
3. As the earthquake dataset has Latitude/Longitude coordinates, choose *WGS 84 EPSG:436* as the CRS in the *Coordinate Reference System Selector* dialog.



4. Adesso i punti del layer dei terremoti devono essere caricati e visualizzati in QGIS. Apriamo allora anche il file dei luoghi abitati. Andate su *Layer* ■ *Aggiungi vettore*.



5. Individuate il file appena scaricato `ne_10m_populated_places_simple.zip` e fate click su *Open*. Selezionate quindi il file *Open* nella finestra di dialogo *Seleziona layer da aggiungere...*



6. Fate degli zoom ed esplorate entrambi i dataset. Ciascun punto viola indica il luogo in cui si è verificato un importante terremoto e ciascun punto blu rappresenta un luogo fortemente popolato. Noi cerchiamo un modo per individuare il punto appartenente al layer dei luoghi abitati che sia il più vicino a uno dei punti del layer dei terremoti.



7. Go to *Vettore* ■ *Strumenti di analisi* ■ *Matrice di distanza*.



8. Selezioniamo il layer `signif` come vettore di punti in input e il layer `ne_10m_populated_places_simple` come vettore di punti in output. Dovrete anche selezionare un unico campo da ciascuno di questi layer che indicherà come i vostri risultati verranno visualizzati (?). In questa analisi a noi interessa ottenere **1** punto più vicino, quindi spuntate la casella guilabel: *Usa solo i punti di destinazione più vicini*, e inserite il valore **1**. Chiamate il file di output `matrix.csv` e fate click su OK.

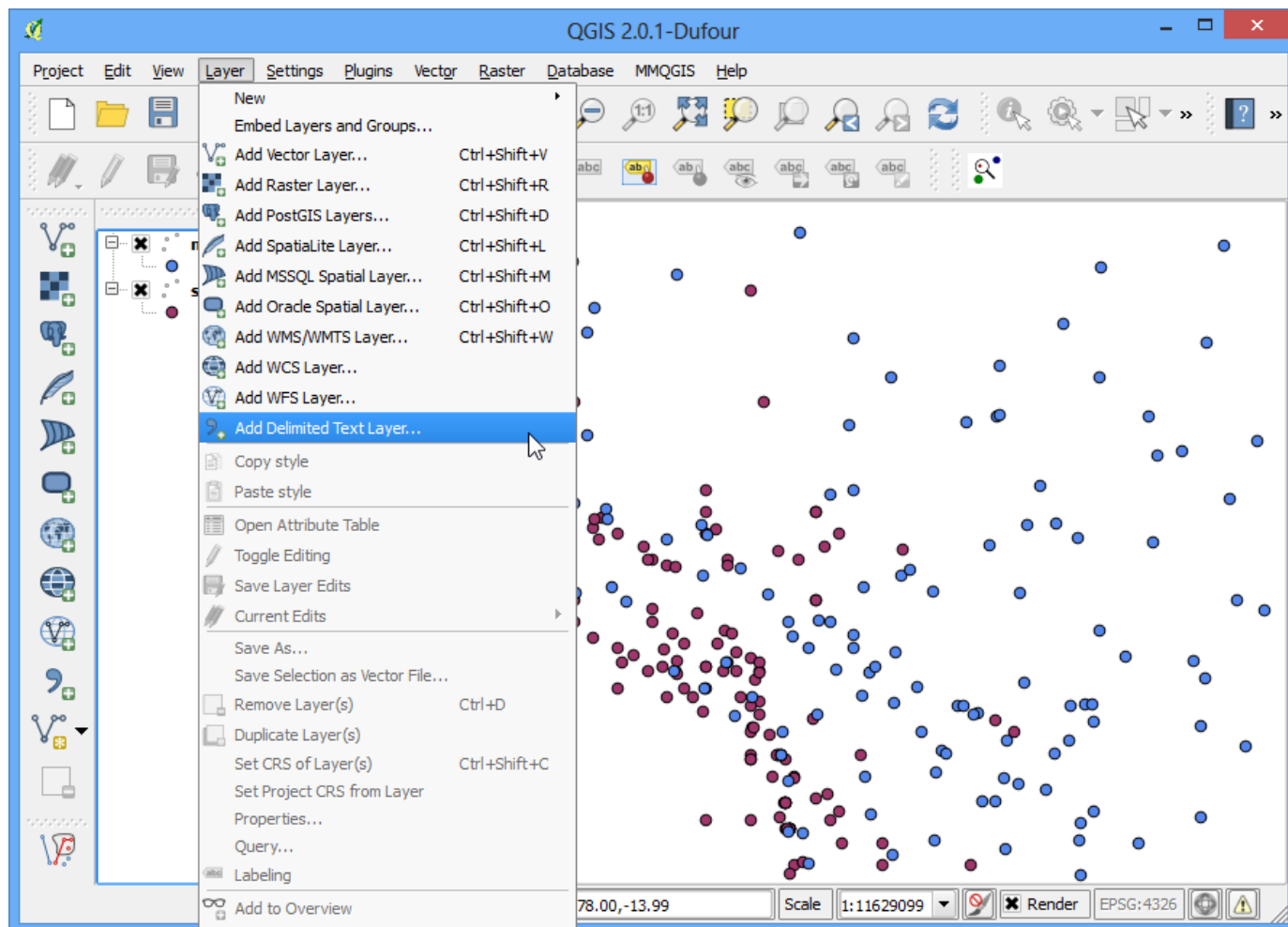
### Note

Un'osservazione di una certa utilità è quella che eventualmente potreste anche realizzare questa analisi con un solo layer. Selezionando lo stesso layer come vettore di punti in input e vettore di punti in output. Il risultato saranno i punti più vicini nello stesso layer invece dei punti più vicini in due layer diversi.





9. Una volta che il file è stato prodotto, potrete vederlo nel Notepad o in qualsiasi editor di testo. Visto che anche QGIS può importare file di testo CSV, lo aggiungeremo a QGIS e lo vedremo da lì. Andate su **Layer ■ Aggiungi layer testo delimitato...**



10. Individuate il file appena creato `matrix.csv`. Dal momento che si tratta di un file composto da colonne di solo testo, spuntate la casella *Nessuna geometria (solo tabella degli attributi)* nella riga *Definizione geometria*. Click su **OK**.

**Create a Layer from a Delimited Text File**

File Name:

Layer name:  Encoding:

File format: ☒ CSV (comma separated values) ☐ Custom delimiters ☐ Regular expression delimiter

Record options: Number of header lines to discard:  ☒ First record has field names

Field options: ☐ Trim fields ☐ Discard empty fields ☐ Decimal separator is comma

Geometry definition: ☐ Point coordinates ☐ Well known text (WKT) ☒ No geometry (attribute only table)

Layer settings: ☐ Use spatial index ☐ Use subset index ☐ Watch file

	InputID	TargetID	Distance
1	1	Al Karak	0.221721171014
2	2	Al Ladhiqiyah	0.144408036939
3	3	Buzmeyin	0.0526324624814
4	8	Al Khalil	0.208418004566
5	11	Iraklio	0.408843567409
6	5877	Iraklio	1.1082549107

11. Vedrete il file CSV caricato in forma di tabella. Click sul tasto destro del layer tabellare alla voce *Apri Tabella degli Attributi*.



12. A questo punto dovrete essere in grado di esaminare il contenuto dei vostri risultati. Il campo *InputID* contiene il campo nome del layer dei terremoti. Il campo *TargetID* contiene il nome degli elementi provenienti dal layer dei luoghi abitati che sono più vicini ai punti che indicano i terremoti. Il campo *Distance* contiene la distanza tra i 2 punti.

### Note

Ricordate che il calcolo della *distanza* viene ottenuto usando il Sistema di Riferimento dei layer. Qui la distanza è misurata in unità di *gradi decimali* perché le coordinate del nostro layer sorgente sono in gradi. Qualora desideraste avere le distanze misurate in metri dovrete riproiettare i layer prima di utilizzare lo strumento di analisi.

Attribute table - matrix :: Features total: 5727, filtered: 5727, selected: 0

	InputID	TargetID	Distance
0	1	Al Karak	0.221721171014
1	2	Al Ladhiqiyah	0.144408036939
2	3	Buzmeyin	0.0526324624814
3	8	Al Khalil	0.208418004566
4	11	Iraklio	0.408843567409
5	5877	Iraklio	1.1082549107
6	9712	Al Ladhiqiyah	0.144408036939
7	12	As Salt	0.230569794451
8	13	Al Aqabah	0.10661139997
9	14	Al Qunaytirah	0.34713470868
10	7793	Nabatiye et Tahta	0.256395311798
11	16	Sparti	0.101878534504
12	7794	Saida	0.00326167893321
13	9713	Piraiévs	0.206150410754
14	17	Volos	0.4810609473
15	18	Sparti	0.101878534504
16	5878	Lamia	0.265998307404
17	19	Varamin	0.239101501046
18	20	Patra	0.520403483984
19	21	Iraklio	0.350232618378
20	22	Kavala	1.1152439462
21	9652	Rajkot	0.717056768568

Show All Features

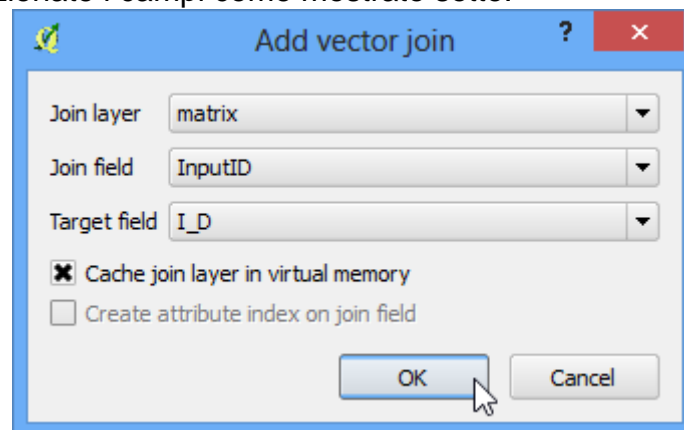
13. Questo dato si avvicina abbastanza al risultato che stavamo cercando. Per alcuni utenti questa tabella può essere considerata sufficiente. Tuttavia, noi intendiamo anche integrare i risultati nel nostro layer dei terremoti usando una **Unione Tabellare**. Click con il tasto destro del layer dei terremoti e dal menu selezionate *Proprietà*.



14. Spostatevi sulla scheda *Join* e fate click sul pulsante + .



15. Vogliamo unire i dati provenienti dalla nostra analisi di prossimità (`matrix.csv`) a questo layer. Abbiamo bisogno di selezionare due campi uno da ciascuno dei layer che abbiano gli stessi valori. Selezionate i campi come mostrato sotto.



16. Vedrete apparire la scheda *Join*. Click su *OK*.



17. Adesso, aprite la tabella degli attributi del layer dei terremoti facendo click sul tasto destro e selezionando *Apri Tabella degli Attributi*.





18. Vedrete che, per ogni punto dei terremoti, abbiamo adesso un attributo che è il punto vicino più prossimo (cioè il posto abitato più vicino) e la distanza da questo punto.

Attribute table - signif :: Features total: 5727, filtered: 5727, selected: 0								
	DAMAGE_DESCR	L_HOUSES_DESTR	SES_DESTROYED_I	L_HOUSES_DAM	SES_DAMAGED_D	matrix_TargetID	matrix_Distance	
0	NULL	NULL	NULL	NULL	NULL	Al Karak	0.221721171014	
1	NULL	NULL	NULL	NULL	NULL	Al Ladhiqiyah	0.144408036939	
2	1	NULL	1	NULL	NULL	Buzmeyin	0.0526324624814	
3	NULL	NULL	NULL	NULL	NULL	Al Khalil	0.208418004566	
4	NULL	NULL	NULL	NULL	NULL	Iraklio	0.408843567409	
5	3	NULL	NULL	NULL	NULL	Iraklio	1.1082549107	
6	3	NULL	NULL	NULL	NULL	Al Ladhiqiyah	0.144408036939	
7	NULL	NULL	NULL	NULL	NULL	As Salt	0.230569794451	
8	NULL	NULL	NULL	NULL	NULL	Al Aqabah	0.10661139997	
9	NULL	NULL	NULL	NULL	NULL	Al Qunaytirah	0.34713470868	
10	NULL	NULL	NULL	NULL	NULL	Nabatiye et Tahta	0.256395311798	
11	NULL	NULL	NULL	NULL	NULL	Sparti	0.101878534504	
12	NULL	NULL	NULL	NULL	NULL	Saida	0.00326167893321	
13	NULL	NULL	NULL	NULL	NULL	Piraiévs	0.206150410754	
14	NULL	NULL	NULL	NULL	NULL	Volos	0.4810609473	
15	1	NULL	1	NULL	NULL	Sparti	0.101878534504	
16	3	NULL	3	NULL	NULL	Lamia	0.265998307404	
17	3	NULL	NULL	NULL	NULL	Varamin	0.239101501046	
18	3	NULL	3	NULL	NULL	Patra	0.520403483984	
19	1	NULL	NULL	NULL	NULL	Iraklio	0.350232618378	
20	NULL	NULL	NULL	NULL	NULL	Kavala	1.1152439462	
21	NULL	NULL	NULL	NULL	NULL	Raikot	0.717056768568	