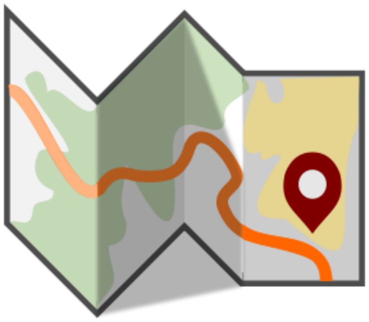
# Nearest Neighbor Analysis

## QGIS Tutorials and Tips



Author
Ujaval Gandhi
http://google.com/+UjavalGandhi

Translations by
SongHyun Choi

This work is licensed under a Creative Commons Attribution 4.0 International License.

# 최근린 분석

GIS는 객체들간의 공간관계를 분석하는데 매우 유용합니다. 그러한 분석으로 주어진 객체가 어떤 객체에 가장 가까운가를 찾는 것이 있습니다. QGIS는 **Distance Matrix**라는 하는 툴을 가지고 있는데 그러한 분석을 돕습니다. 이 예제에서는 2개의 데이터셋을 사용할 것이고 하나의 레이어의 어떤 점이 두번째 레이어의 어떤 점과 가장 가까운가를 찾을 것입니다.

## 과업 개요

이미 알려져 있는 주요 지진발생 위치가 주어지면 지진이 발생한 곳에서 가장 가깝게 인구가 밀집한 지역을 찾습니다.

## ■■ ■■

- QGIS에서 어떻게 테이블을 결합하는지 알아봅니다. (보다 자세한 내용은 :doc:`performing_table_joins`를 보기 바랍니다.)

## 데이터 획득

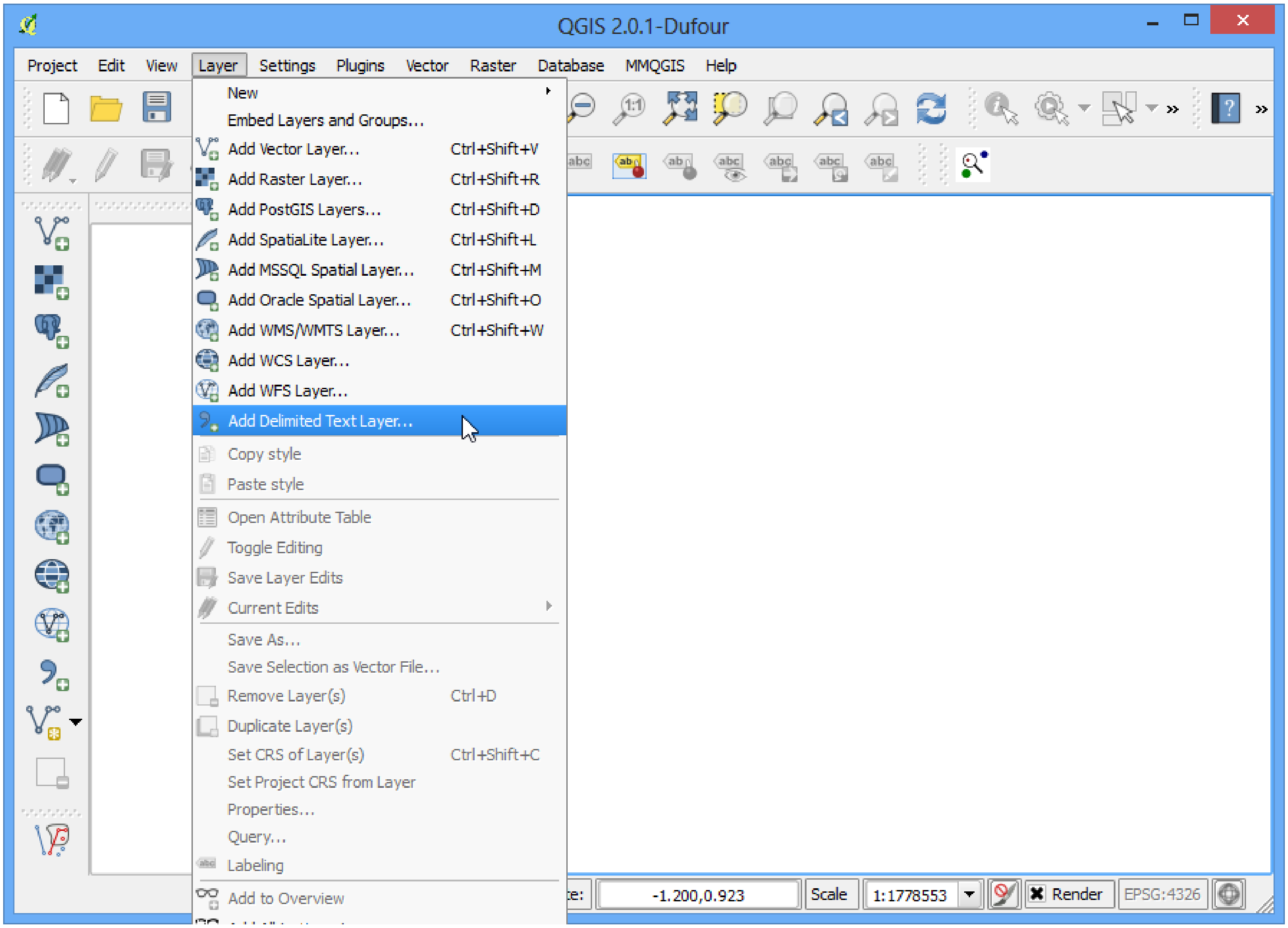
예제에서는 모든 주요 지진을 레이어로 나타내기 위해 NOAA’s National Geophysical Data Center의 중요 지진 데이터베이스 Significant Earthquake Database 를 사용할 것입니다. 탭으로 구분된 지진 데이터 `tab-delimited earthquake data <http://www.ngdc.noaa.gov/nndc/struts/results?type_0=Exact&query_0=$ID&t=101650&s=13&d=189&dfn=signif.txt>`_를 다운로드하시기 바랍니다.

Natural Earth는 훌륭한 인구밀집지역 Populated Places 데이터베이스를 가지고 있습니다. 단순 데이터셋 `simple (less columns) dataset <http://www.naturalearthdata.com/http//www.naturalearthdata.com/download/10m/cultural/ne_10m_populated_places_simple.zip>`_을 다운 받으십시오.

데이터 출처: [NGDC] [NATURALEARTH]

## 과정

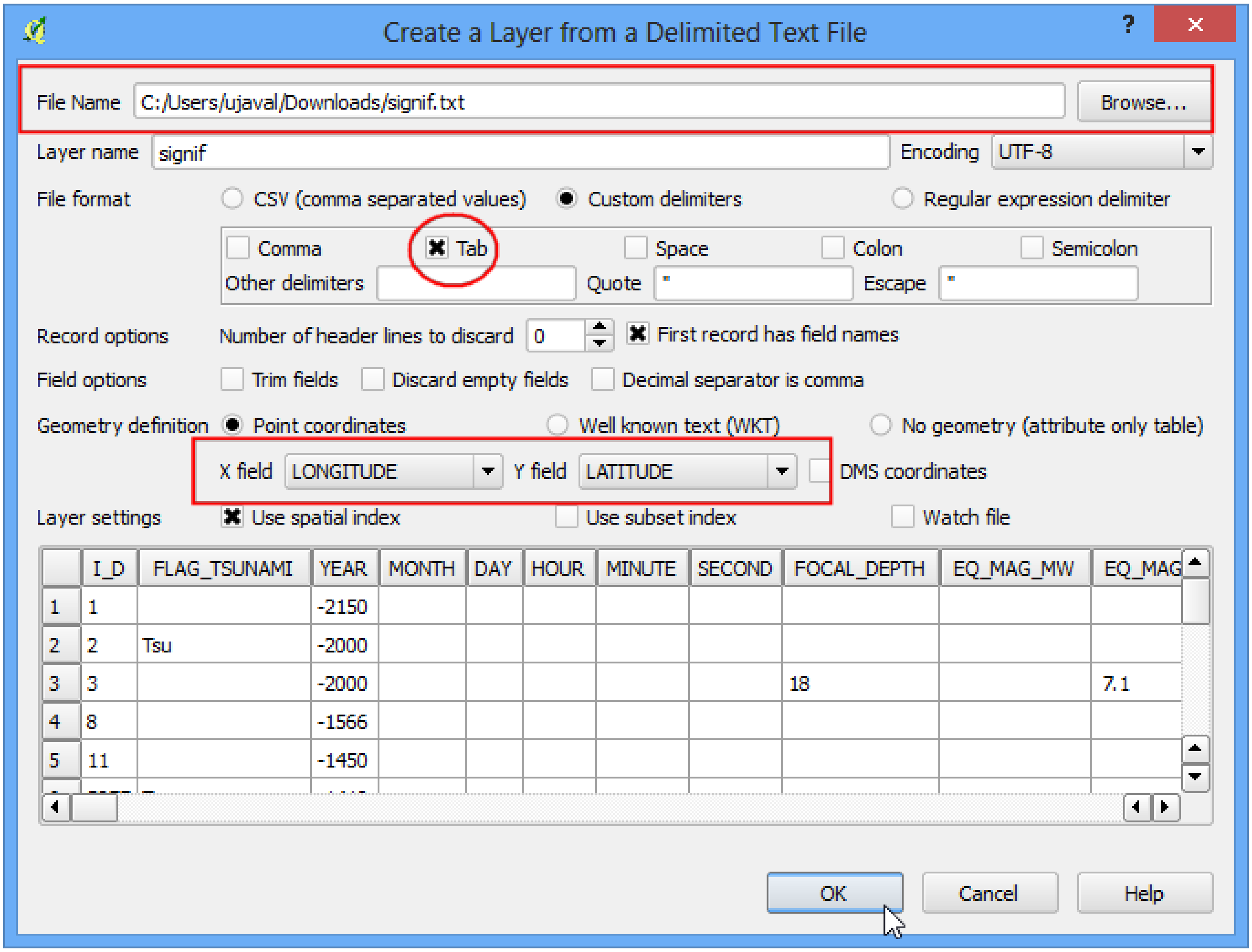
1. 메뉴: 레이어 --> 구분자로 분리된 텍스트 레이어를 추가 ...:menuselection:`Layer --> Add Delimited Text Layer`를 열고 다운로드한 ``signif.txt``파일을 찾습니다.


QGIS 2.0.1-Dufour
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
MMQGIS
Help
New
Embed Layers and Groups...
Add Vector Layer... Ctrl+Shift+V
Add Raster Layer... Ctrl+Shift+R
Add PostGIS Layers... Ctrl+Shift+D
Add SpatiaLite Layer... Ctrl+Shift+L
Add MSSQL Spatial Layer... Ctrl+Shift+M
Add Oracle Spatial Layer... Ctrl+Shift+O
Add WMS/WMTS Layer... Ctrl+Shift+W
Add WCS Layer...
Add WFS Layer...
Add Delimited Text Layer...
Copy style
Paste style
Open Attribute Table
Toggle Editing
Save Layer Edits
Current Edits
Save As...
Save Selection as Vector File...
Remove Layer(s) Ctrl+D
Duplicate Layer(s)
Set CRS of Layer(s) Ctrl+Shift+C
Set Project CRS from Layer
Properties...
Query...
Labeling
Add to Overview
-1.200,0.923
Scale
1:1778553
Render
EPSG:4326

2. 이 파일은 탭으로 구분된 파일 *tab-delimited file*이므로 파일포맷 :guilabel:`File format`으로 구분자정의를 체크하고 탭 :guilabel:`Tab`을 선택합니다. X필드 :guilabel:`X field`와 Y필드 :guilabel:`Y field`가 자동적으로 채워집니다. 확인 :guilabel:`OK`를 클릭합니다.

## *Note*

QGIS에 파일이 불러들여지는 동안 몇 에러메세지를 볼 수도 있습니다. 이것들은 유효한 에러이고 파일의 몇 행은 불러들여지지 않을 것입니다. 이 예제에서는 무시할 수 있습니다.



File Name: C:/Users/ujaval/Downloads/signif.txt

Layer name: signif

Encoding: UTF-8

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: 0 ☒ First record has field names

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

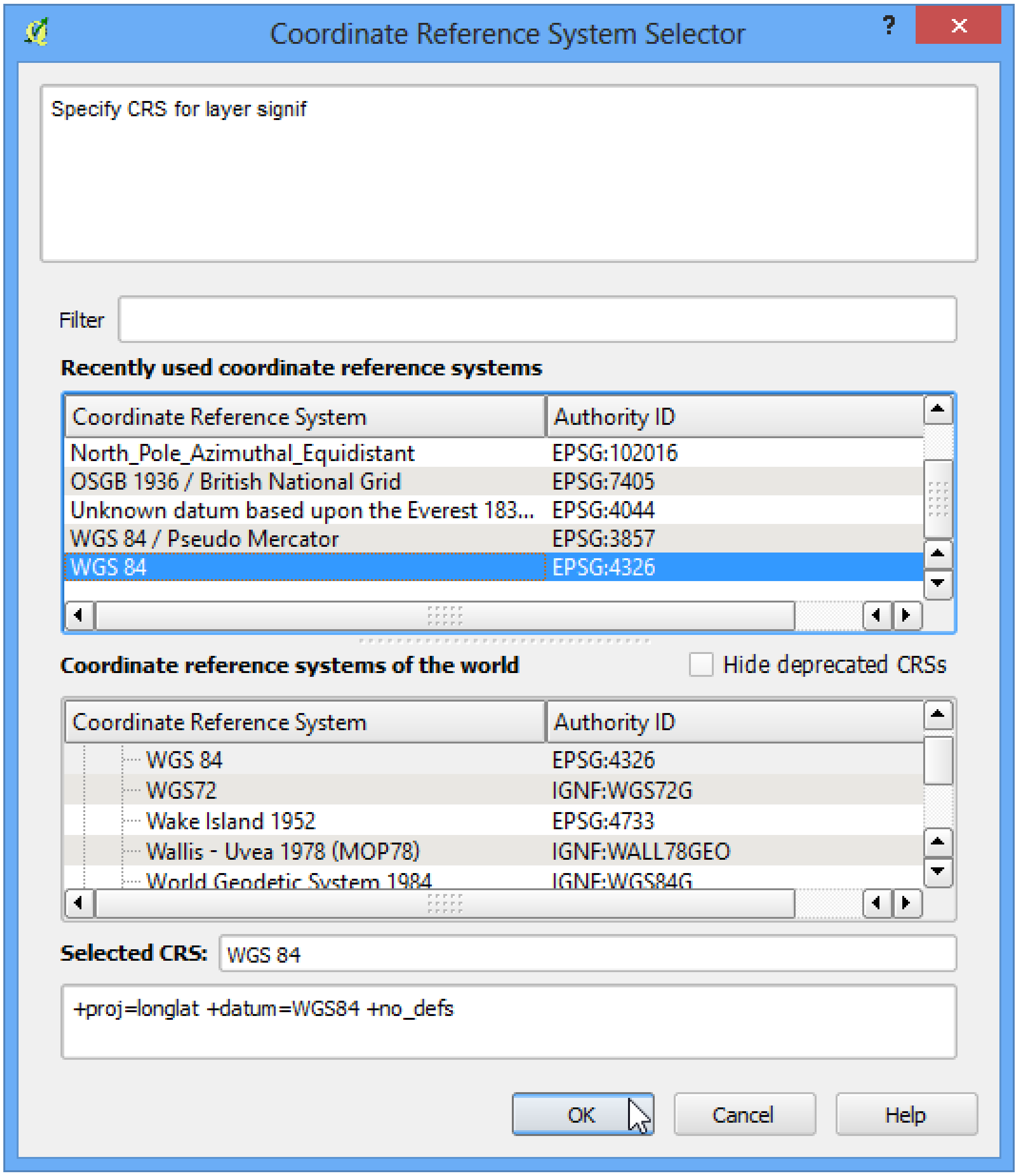
Geometry definition: ● Point coordinates ○ Well known text (WKT) ○ No geometry (attribute only table)

X field: LONGITUDE Y field: LATITUDE ☐ DMS coordinates

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. 지진 데이터셋은 위도 / 경도 좌표를 갖고 있으므로 좌표계 선택 Coordinate Reference System Selector 다이알로그에서 CRS는 WGS 84 EPSG:436 를 선택합니다.


Coordinate Reference System Selector
Filter
Hide deprecated CRSs
OK
Cancel
Help

Specify CRS for layer signif

**Recently used coordinate reference systems**

| Coordinate Reference System | Authority ID |
|---|---|
| North_Pole_Azimuthal_Equidistant | EPSG:102016 |
| OSGB 1936 / British National Grid | EPSG:7405 |
| Unknown datum based upon the Everest 183... | EPSG:4044 |
| WGS 84 / Pseudo Mercator | EPSG:3857 |
| WGS 84 | EPSG:4326 |

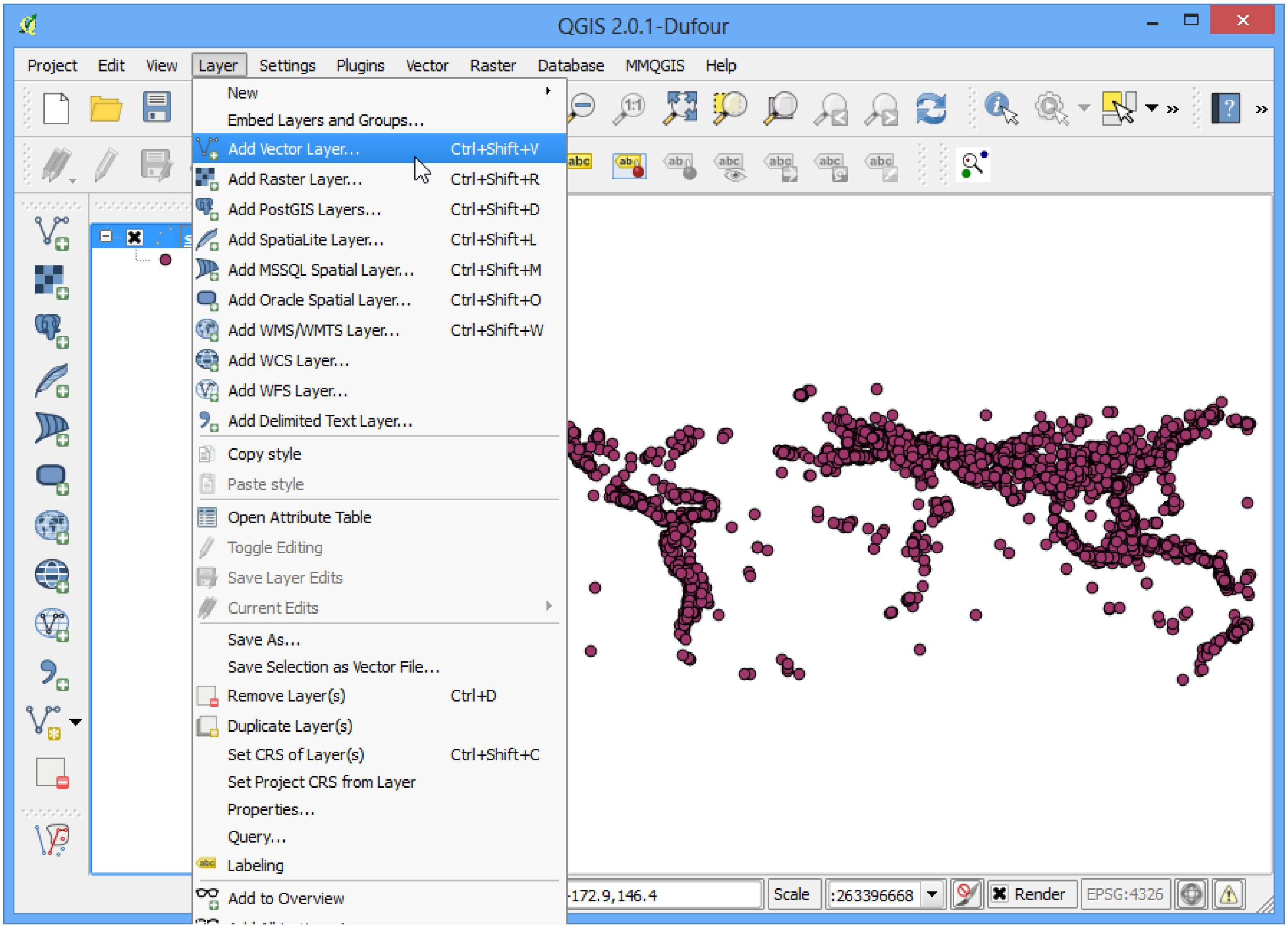
**Coordinate reference systems of the world** ☐ Hide deprecated CRSs

| Coordinate Reference System | Authority ID |
|---|---|
| WGS 84 | EPSG:4326 |
| WGS72 | IGNF:WGS72G |
| Wake Island 1952 | EPSG:4733 |
| Wallis - Uvea 1978 (MOP78) | IGNF:WALL78GEO |
| World Geodetic System 1984 | IGNF:WGS84G |

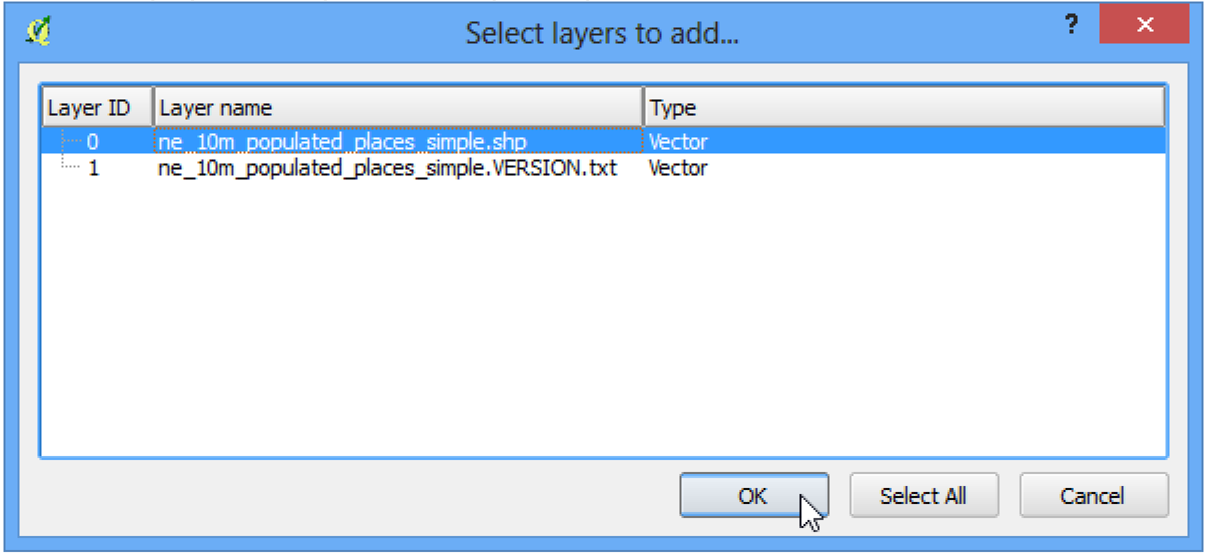
**Selected CRS:** WGS 84

```
+proj=longlat +datum=WGS84 +no_defs
```

4. 지진 포인트가 QGIS로 불러들여지고 나타납니다. 인구밀집지역 레이어도 열어보겠습니다. 메뉴 레이어 --> 레이어 추가 --> 벡터레이어 추가 :menuselection:`Layer --> Add Vector Layer`로 갑니다.

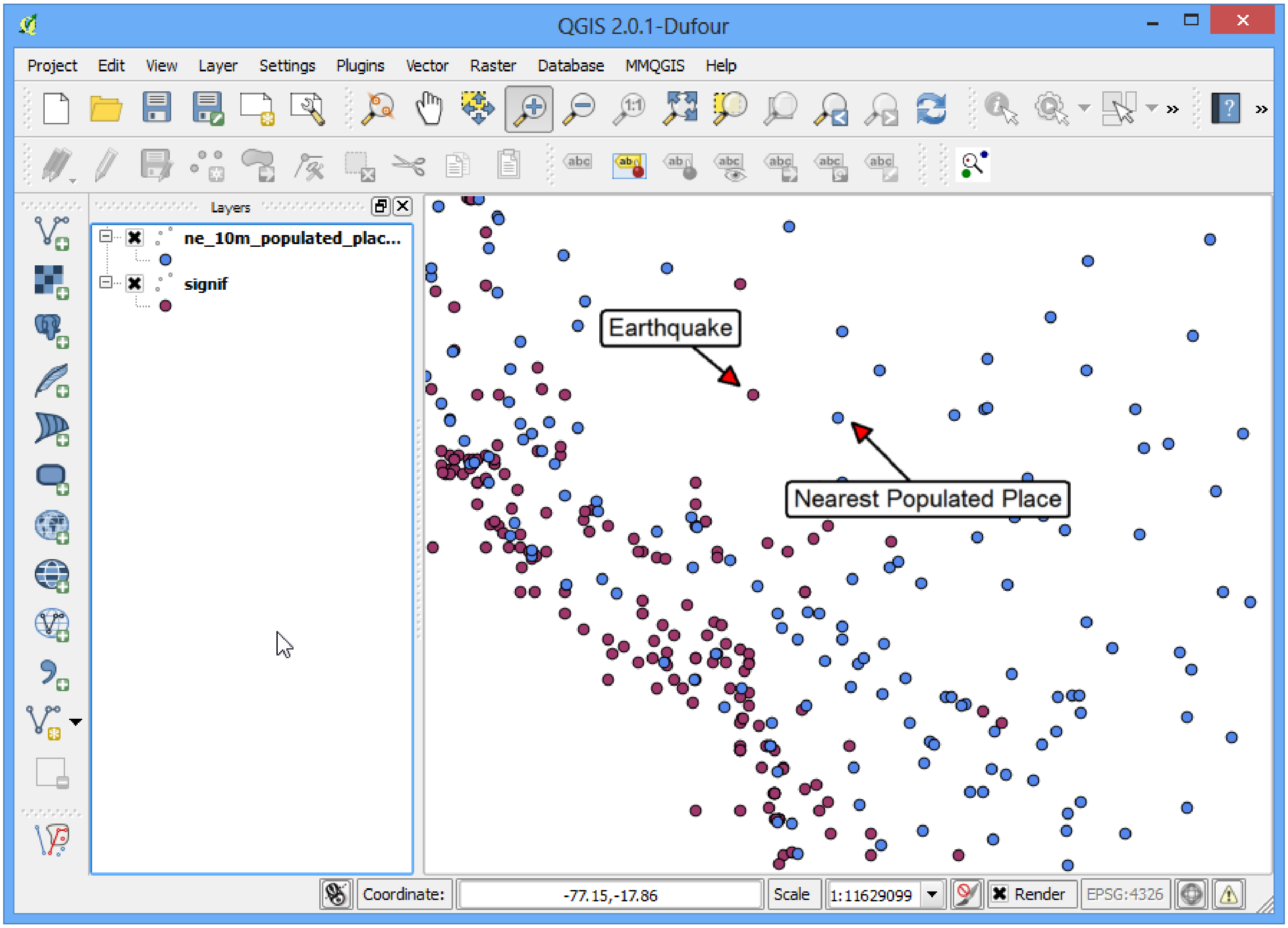

QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Help
New
Embed Layers and Groups...
Add Vector Layer... Ctrl+Shift+V
Add Raster Layer... Ctrl+Shift+R
Add PostGIS Layers... Ctrl+Shift+D
Add SpatiaLite Layer... Ctrl+Shift+L
Add MSSQL Spatial Layer... Ctrl+Shift+M
Add Oracle Spatial Layer... Ctrl+Shift+O
Add WMS/WMTS Layer... Ctrl+Shift+W
Add WCS Layer...
Add WFS Layer...
Add Delimited Text Layer...
Copy style
Paste style
Open Attribute Table
Toggle Editing
Save Layer Edits
Current Edits
Save As...
Save Selection as Vector File...
Remove Layer(s) Ctrl+D
Duplicate Layer(s)
Set CRS of Layer(s) Ctrl+Shift+C
Set Project CRS from Layer
Properties...
Query...
Labeling
Add to Overview
-172.9,146.4
Scale
:263396668
Render
EPSG:4326

5. 다운로드한 ``ne_10m_populated_places_simple.zip`` 파일을 찾고 열기 :guilabel:`Open`를 클릭합니다. 추가할 레이어 선택 :guilabel:`Select layers to add...` 다이얼로그에서 ``ne_10m_populated_places_simple.shp``를 선택합니다.

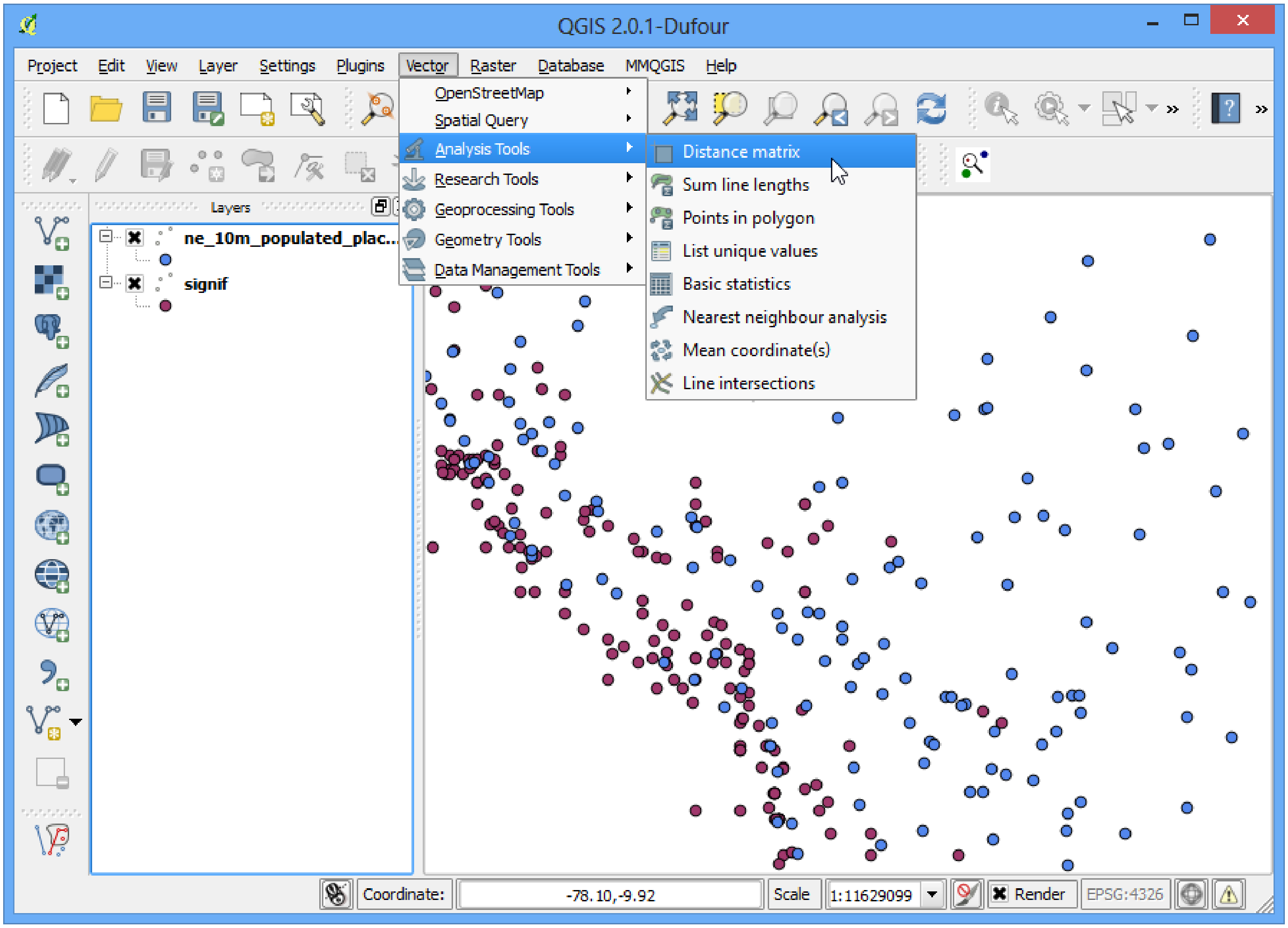

Select layers to add...
OK
Select All
Cancel

| Layer ID | Layer name | Type |
|---|---|---|
| 0 | ne_10m_populated_places_simple.shp | Vector |
| 1 | ne_10m_populated_places_simple.VERSION.txt | Vector |

6. 두 개의 데이터셋을 확대 및 탐색해 봅니다. 각각의 보라색 점은 중요한 지진이 발생한 위치를 나타내고, 각각의 파란색 점은 인구밀집지역을 나타냅니다. 지진레이어에서 각 점에 대해 인구밀집지역으로부터 가장 가까운 점을 찾는 방법이 필요합니다.


QGIS 2.0.1-Dufour
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
MMQGIS
Help
Layers
ne_10m_populated_plac...
signif
Earthquake
Nearest Populated Place
Coordinate:
-77.15,-17.86
Scale
1:11629099
Render
EPSG:4326

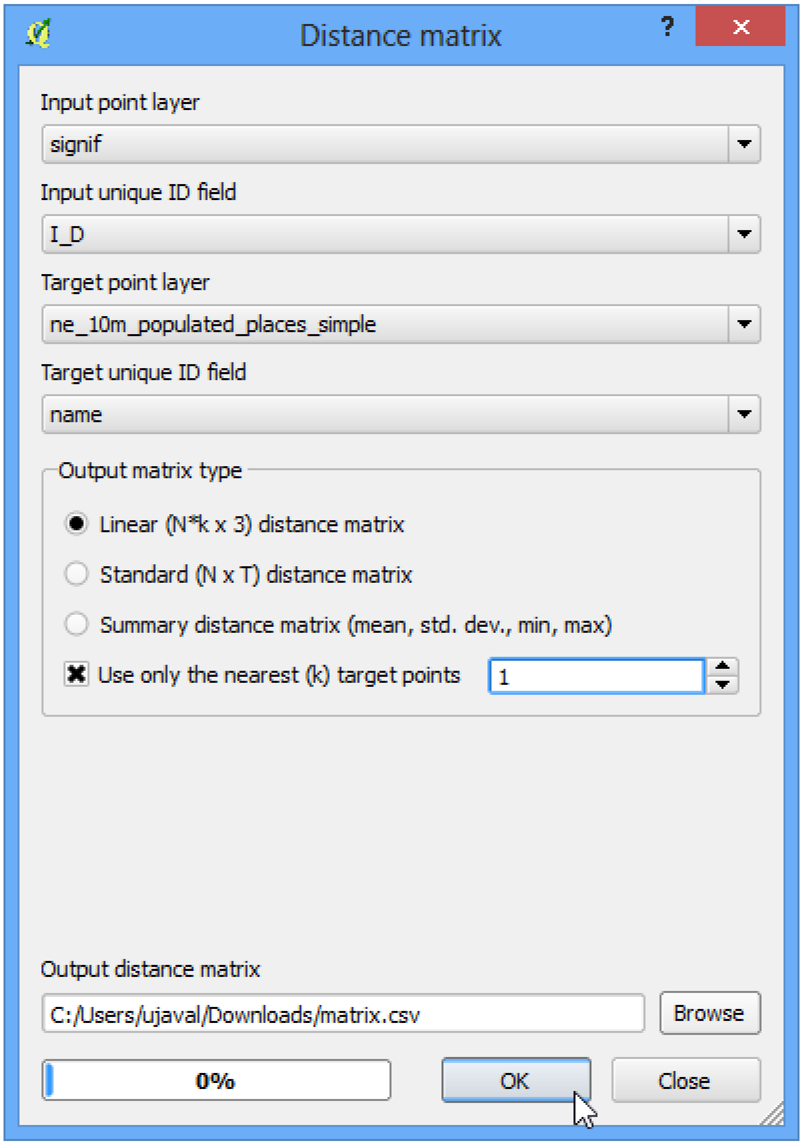
7. 메뉴 벡터 --> 분석 도구 --> 거리행렬계산하기 :menuselection:`Vector --> Analysis Tools --> Distance Matrix`로 갑니다.


QGIS 2.0.1-Dufour
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
MMQGIS
Help
OpenStreetMap
Spatial Query
Analysis Tools
Research Tools
Geoprocessing Tools
Geometry Tools
Data Management Tools
Distance matrix
Sum line lengths
Points in polygon
List unique values
Basic statistics
Nearest neighbour analysis
Mean coordinate(s)
Line intersections
Layers
ne_10m_populated_plac...
signif
Coordinate:
-78.10,-9.92
Scale
1:11629099
Render
EPSG:4326

8. 여기서 입력 점 레이어로써 지진 레이어 ***signif*** 를, 대상 점 레이어로써 인구밀집지역 ``ne_10m_populated_places_simple``를 선택합니다. 또한 이러한 각 레이어로부터 결과가 어떻게 보여질지 고유필드를 선택할 필요가 있습니다. 이번 분석에서는 가까운 점으로 **1** 만 찾으므로 최근린점만을 사용 :guilabel:`Use only the nearest(k) target points`을 체크하고 :guilabel:`1`을 입력합니다. 출력파일명으로 ``matrix.csv``을 입력하고 OK를 클릭합니다.

## Note

알아두면 편리한 것으로 오로지 1 레이어만 가지고 분석을 수행할 수도 있습니다. 입력 및 출력 두 개의 레이어를 동시에 선택합니다. 결과는 여기서 사용했던 것 처럼 다른 레이어 대신 같은 레이어로부터 최근린이 될 것입니다.



Input point layer: signif

Input unique ID field: I_D

Target point layer: ne_10m_populated_places_simple

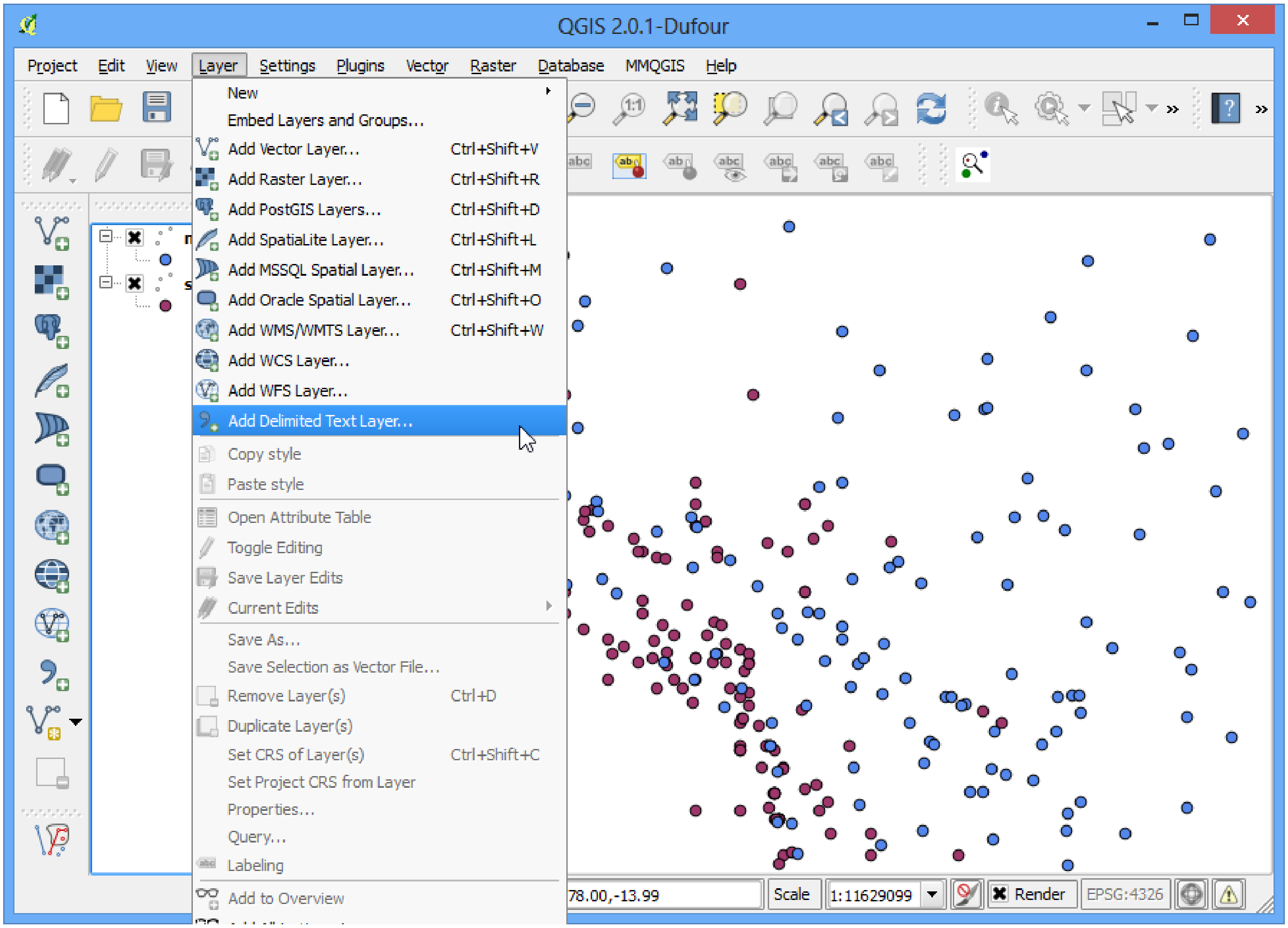
Target unique ID field: name

Output matrix type

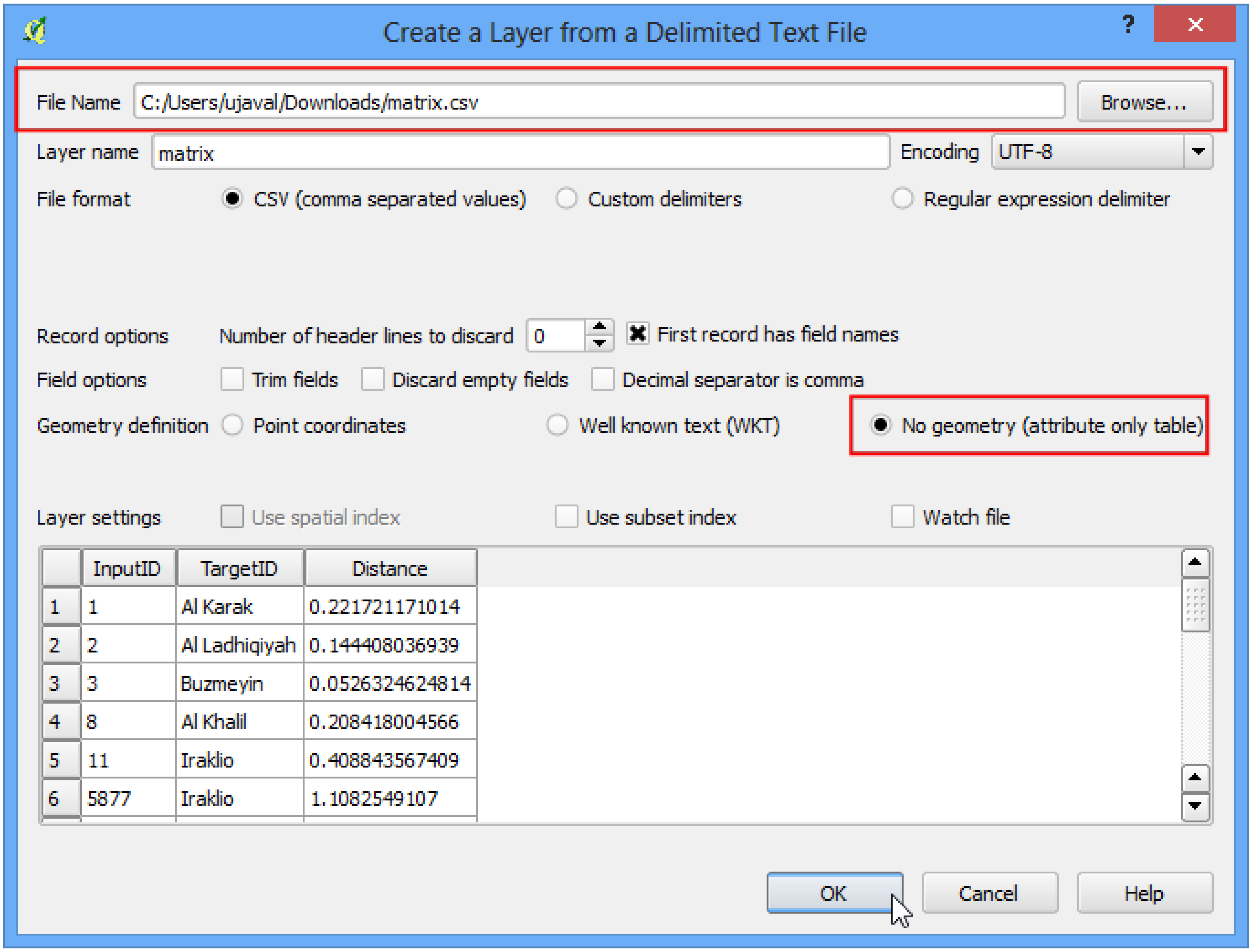
- ● Linear (N*k x 3) distance matrix
- ○ Standard (N x T) distance matrix
- ○ Summary distance matrix (mean, std. dev., min, max)
- ☒ Use only the nearest (k) target points: 1

Output distance matrix: C:/Users/ujaval/Downloads/matrix.csv

9. 일단 파일이 생성되면, 노트패드 혹은 다른 어떤 텍스트 에디터에서 그것을 볼 수 있습니다. QGIS는 CSV파일을 불러들일 수 있습니다. 그래서 그것을 QGIS에 추가하고 거기서 봅니다. 메뉴 레이어 --> 구분자로 분리된 텍스트 레이어 추가... :menuselection:`Layer --> Add Delimited Text Layer...`로 갑니다.


QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Help
New
Embed Layers and Groups...
Add Vector Layer... Ctrl+Shift+V
Add Raster Layer... Ctrl+Shift+R
Add PostGIS Layers... Ctrl+Shift+D
Add SpatiaLite Layer... Ctrl+Shift+L
Add MSSQL Spatial Layer... Ctrl+Shift+M
Add Oracle Spatial Layer... Ctrl+Shift+O
Add WMS/WMTS Layer... Ctrl+Shift+W
Add WCS Layer...
Add WFS Layer...
Add Delimited Text Layer...
Copy style
Paste style
Open Attribute Table
Toggle Editing
Save Layer Edits
Current Edits
Save As...
Save Selection as Vector File...
Remove Layer(s) Ctrl+D
Duplicate Layer(s)
Set CRS of Layer(s) Ctrl+Shift+C
Set Project CRS from Layer
Properties...
Query...
Labeling
Add to Overview
78.00,-13.99
Scale
1:11629099
Render
EPSG:4326

10. 새롭게 만들어진 ``matrix.csv`` 파일을 찾습니다. 이 파일이 텍스트 컬럼이기때문에 지오메트리 정의 :guilabel:`Geometry definition`에서 지오메트리가 아님 :guilabel:`No geometry (attribute only table)`를 선택합니다. :guilabel:`OK`를 클릭합니다.



File Name C:/Users/ujaval/Downloads/matrix.csv

Layer name matrix

Encoding UTF-8

File format ● CSV (comma separated values) ○ Custom delimiters ○ Regular expression delimiter

Record options Number of header lines to discard 0 ☒ First record has field names

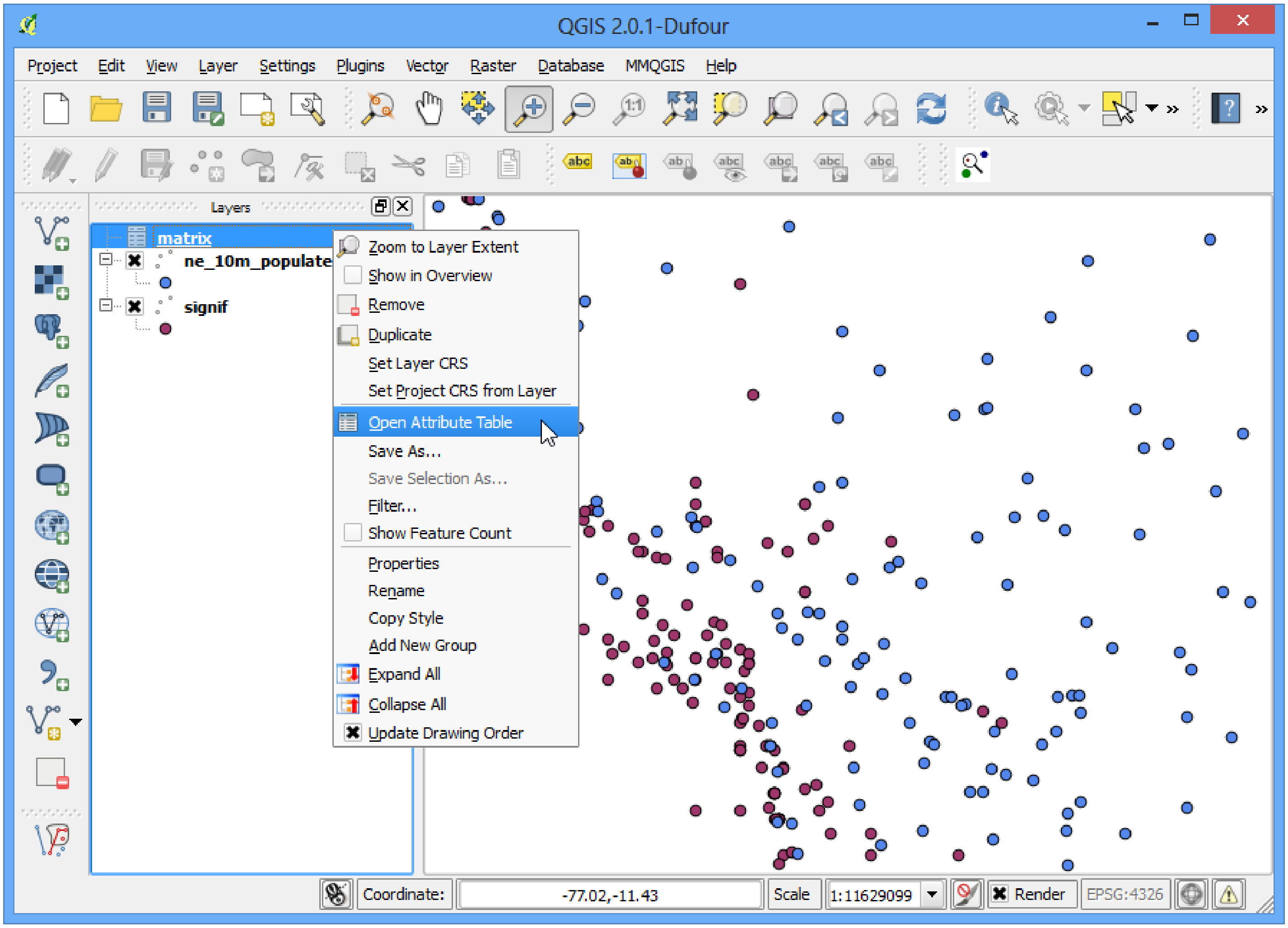
Field options ☐ Trim fields ☐ Discard empty fields ☐ Decimal separator is comma

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 |

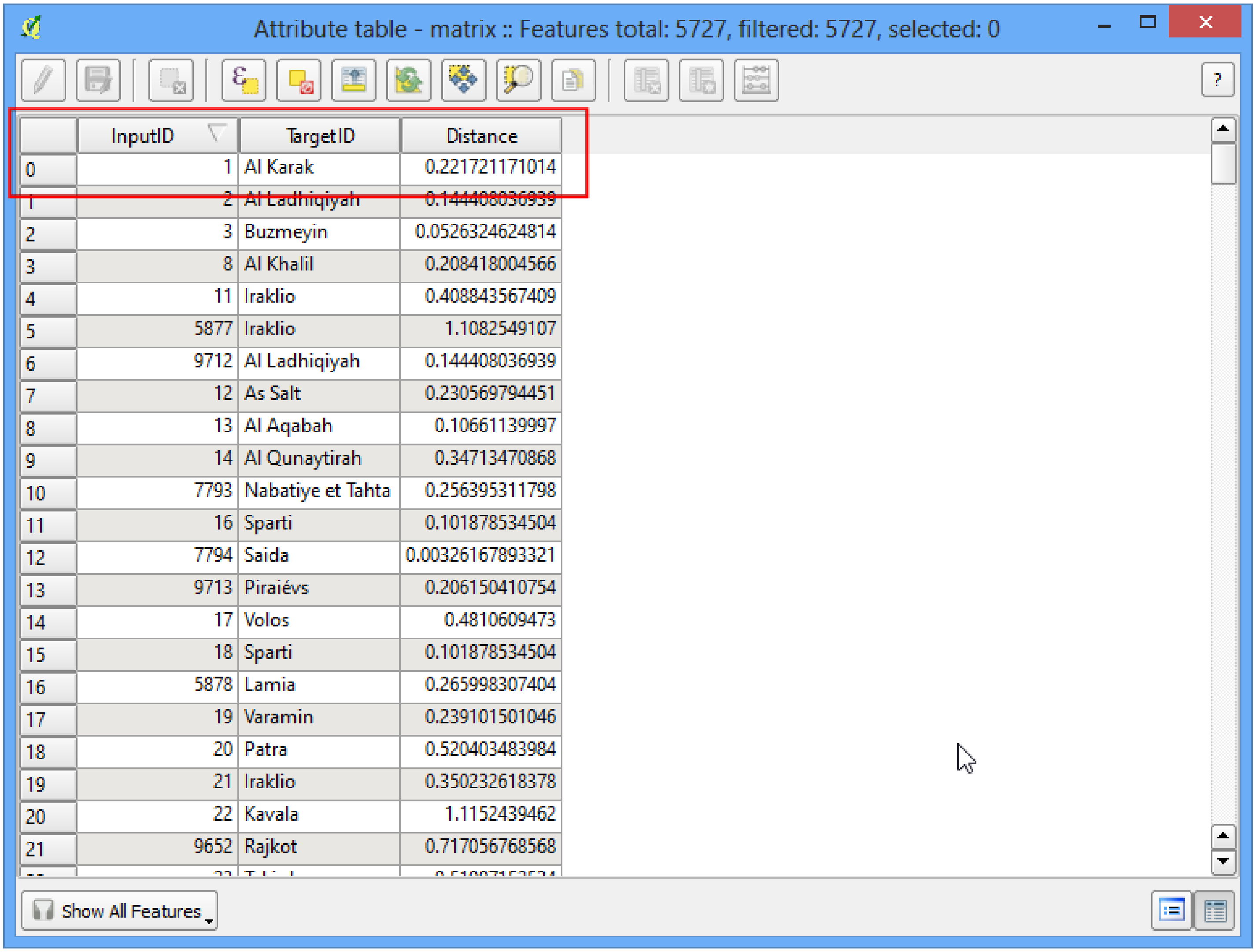
11. 테이블로 불러들여진 CSV파일을 보게 될 것입니다. 테이블 레이어를 우측클릭하고 속성테이블열기 :guilabel:`Open Attribute Table`를 선택합니다.


QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Help
Layers
matrix
ne_10m_populate
signif
Zoom to Layer Extent
Show in Overview
Remove
Duplicate
Set Layer CRS
Set Project CRS from Layer
Open Attribute Table
Save As...
Save Selection As...
Filter...
Show Feature Count
Properties
Rename
Copy Style
Add New Group
Expand All
Collapse All
Update Drawing Order
Coordinate:
-77.02,-11.43
Scale
1:11629099
Render
EPSG:4326

12. 이제 결과의 내용을 볼 수 있을 겁니다 . InputID 필드는 지진레이어의 필드명을 담고 있습니다 . TargetID 필드는 지진발생지점으로 가장 가까운 인구밀집지역으로부터의 객체명을 포함고 있습니다 . :guilabel:`Distance` 필드는 2 점간의 거리입니다 .

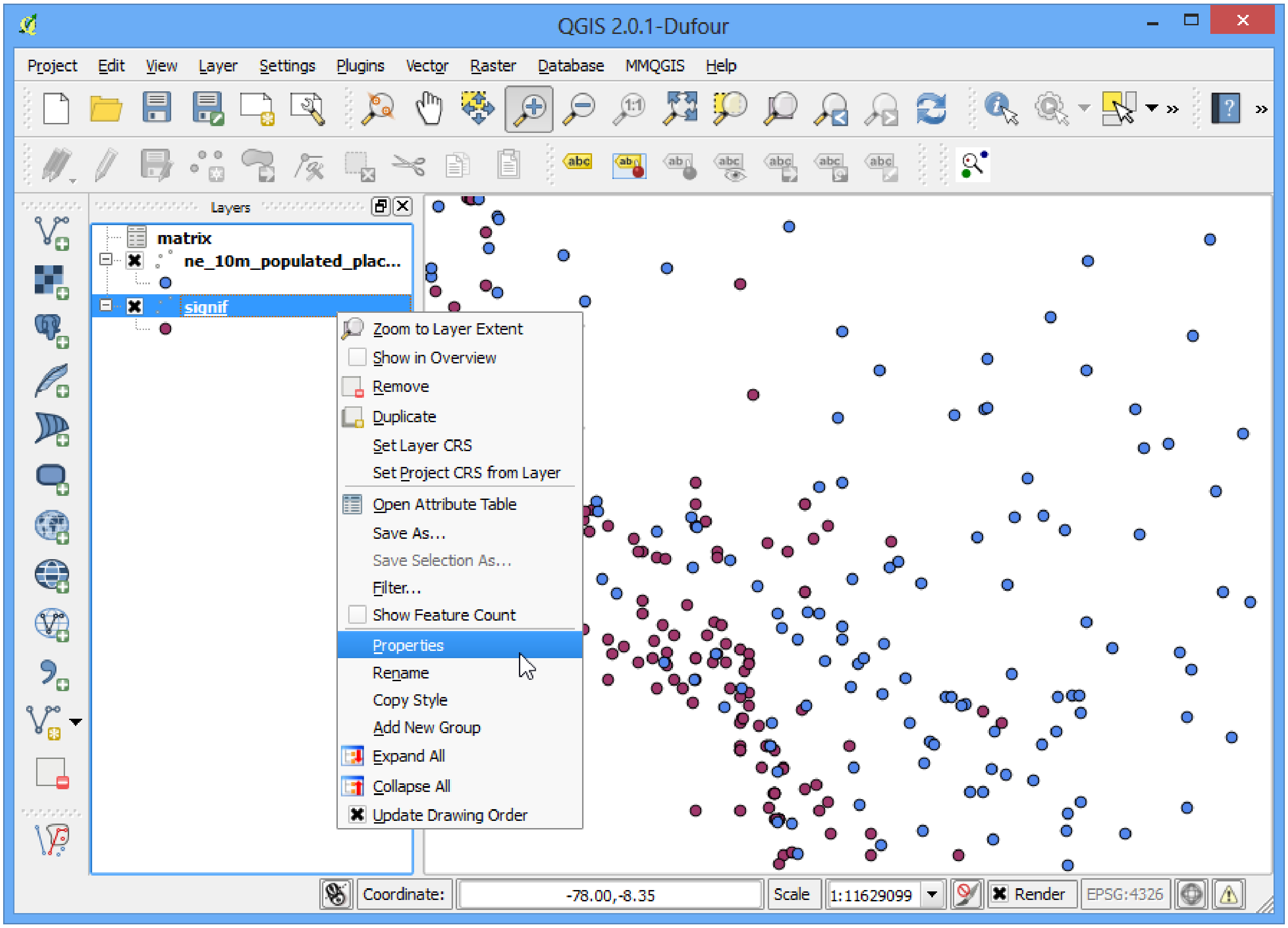
**Note**

거리 distance*의 계산은 레이어의 좌표계를 사용한 것임을 기억해야 합니다 . 여기서 거리는 원 레이어 좌표가 degree이기 때문에 *decimal degrees 단위가 될 것입니다 . 거리의 단위로 미터를 원한다면 툴을 구동시키기 전에 레이어를 재투영해야 합니다 .

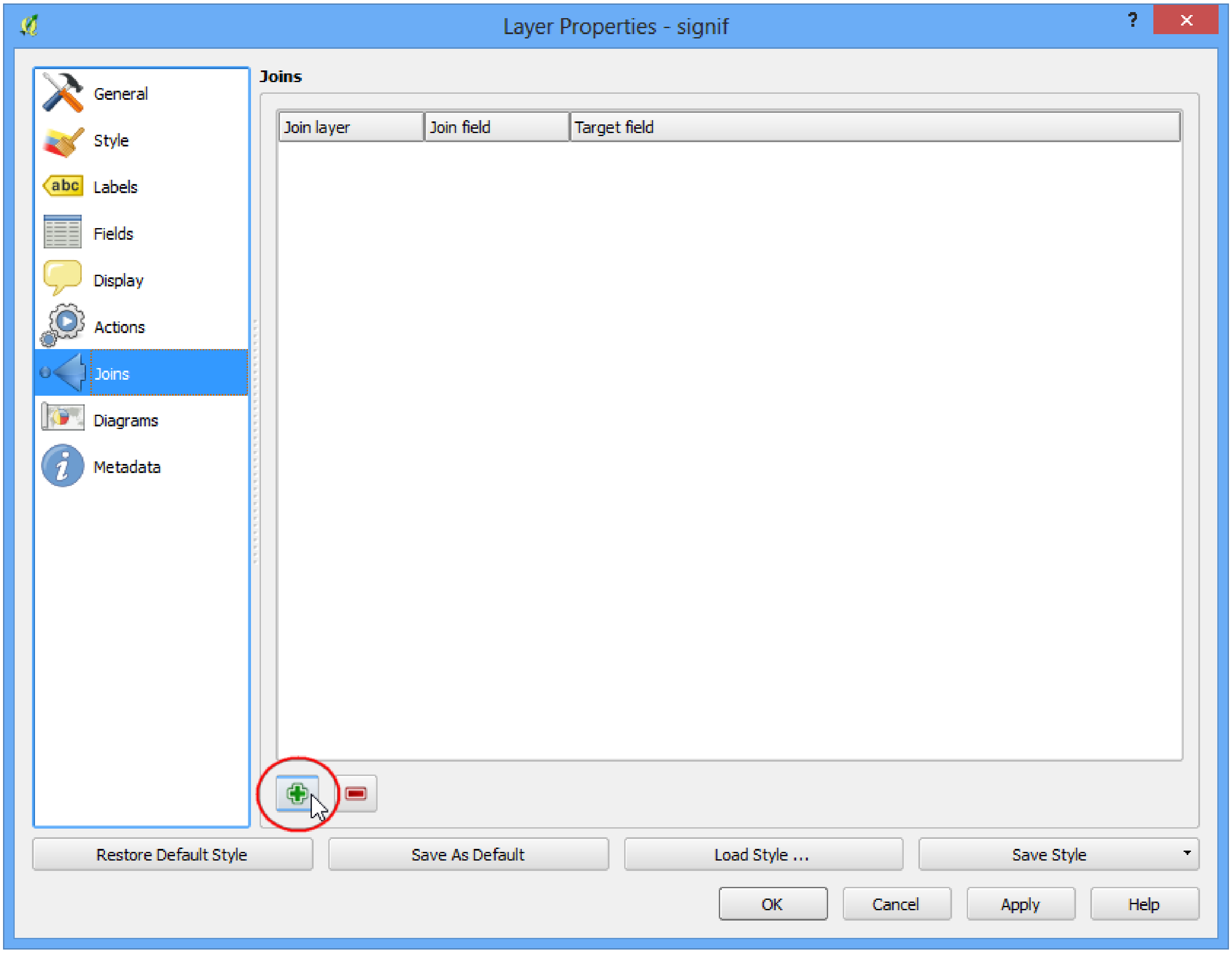

Attribute table - matrix :: Features total: 5727, filtered: 5727, selected: 0
?
Show All Features

|  | 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 |

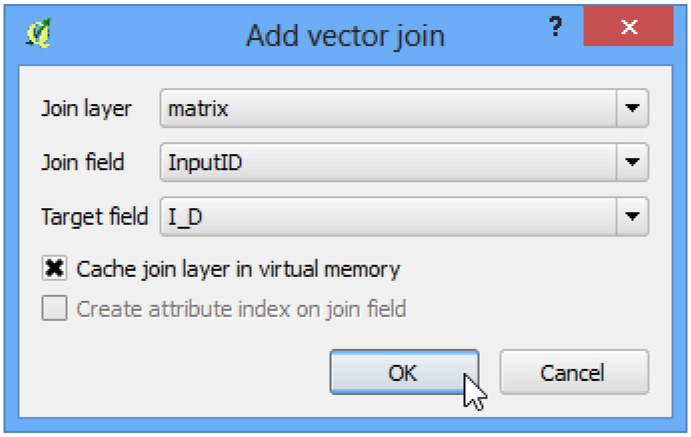
13. 이것이 원래 찾던 결과와 가장 가까운 것입니다 . 몇 사용자들에게 이 테이블은 충분할 것입니다 . 그러나 이 결과를 테이블 결합 **Table Join**을 사용해서 원 지진레이어에 통합시킬 수 있습니다 . 지진 레이어를 우측클릭하고 속성 :guilabel:`Properties` 을 선택합니다 .


QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Help
Layers
matrix
ne_10m_populated_plac...
signif
Zoom to Layer Extent
Show in Overview
Remove
Duplicate
Set Layer CRS
Set Project CRS from Layer
Open Attribute Table
Save As...
Save Selection As...
Filter...
Show Feature Count
Properties
Rename
Copy Style
Add New Group
Expand All
Collapse All
Update Drawing Order
Coordinate:
-78.00,-8.35
Scale
1:11629099
Render
EPSG:4326

14. 결합 Joins 탭으로 가서 + 버튼을 클릭하십시오 .

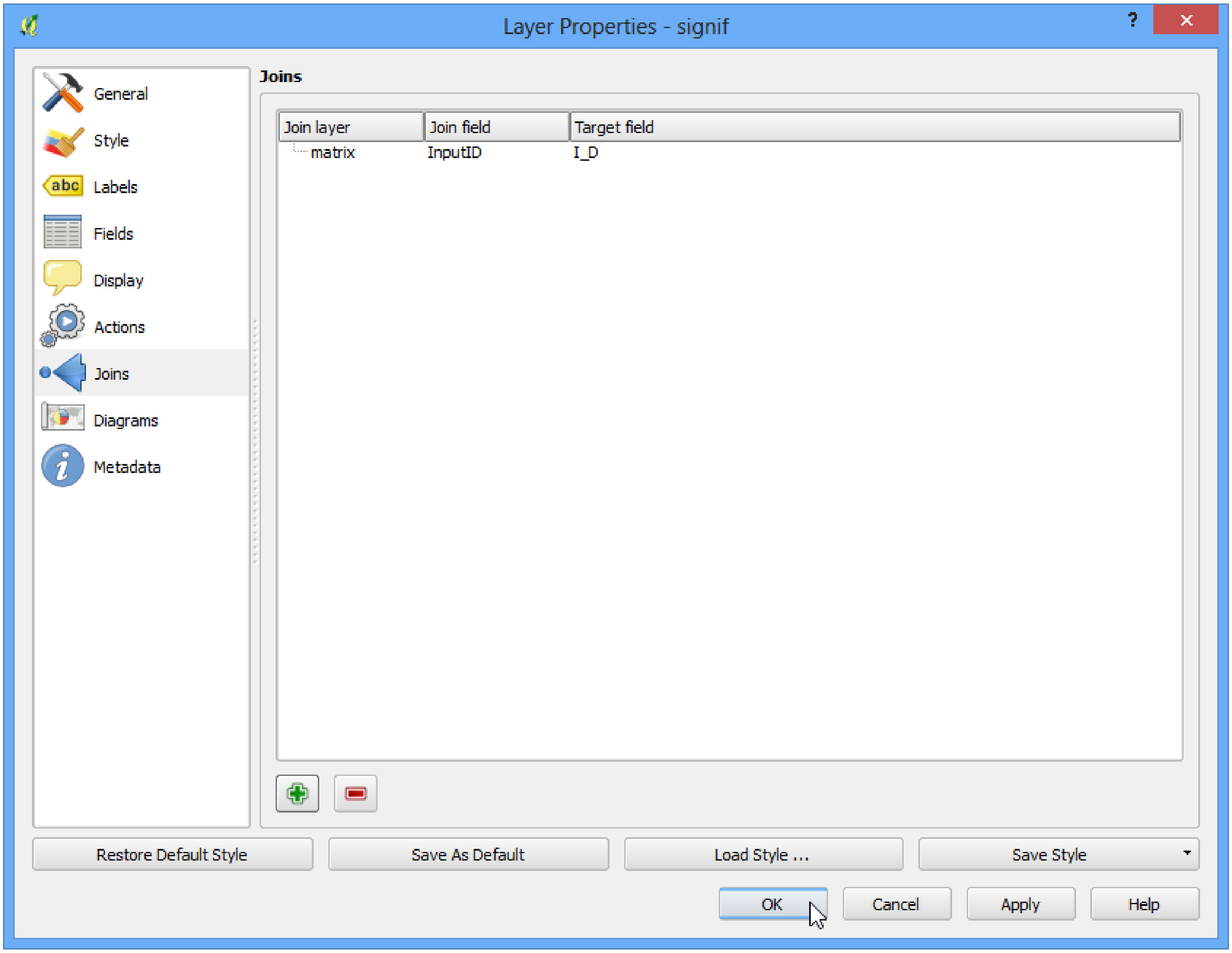

Layer Properties - signif
General
Style
Labels
Fields
Display
Actions
Joins
Diagrams
Metadata
Joins
Join layer
Join field
Target field
Restore Default Style
Save As Default
Load Style ...
Save Style
OK
Cancel
Apply
Help

15. 분석결과 (***matrix.csv***)로 부터 이 레이어를 결합하고자 합니다 . 같은 값을 갖는 각 레이어로부터 필드를 선택할 필요가 있습니다 . 아래와 같이 필드를 선택합니다 .


Add vector join
OK
Cancel

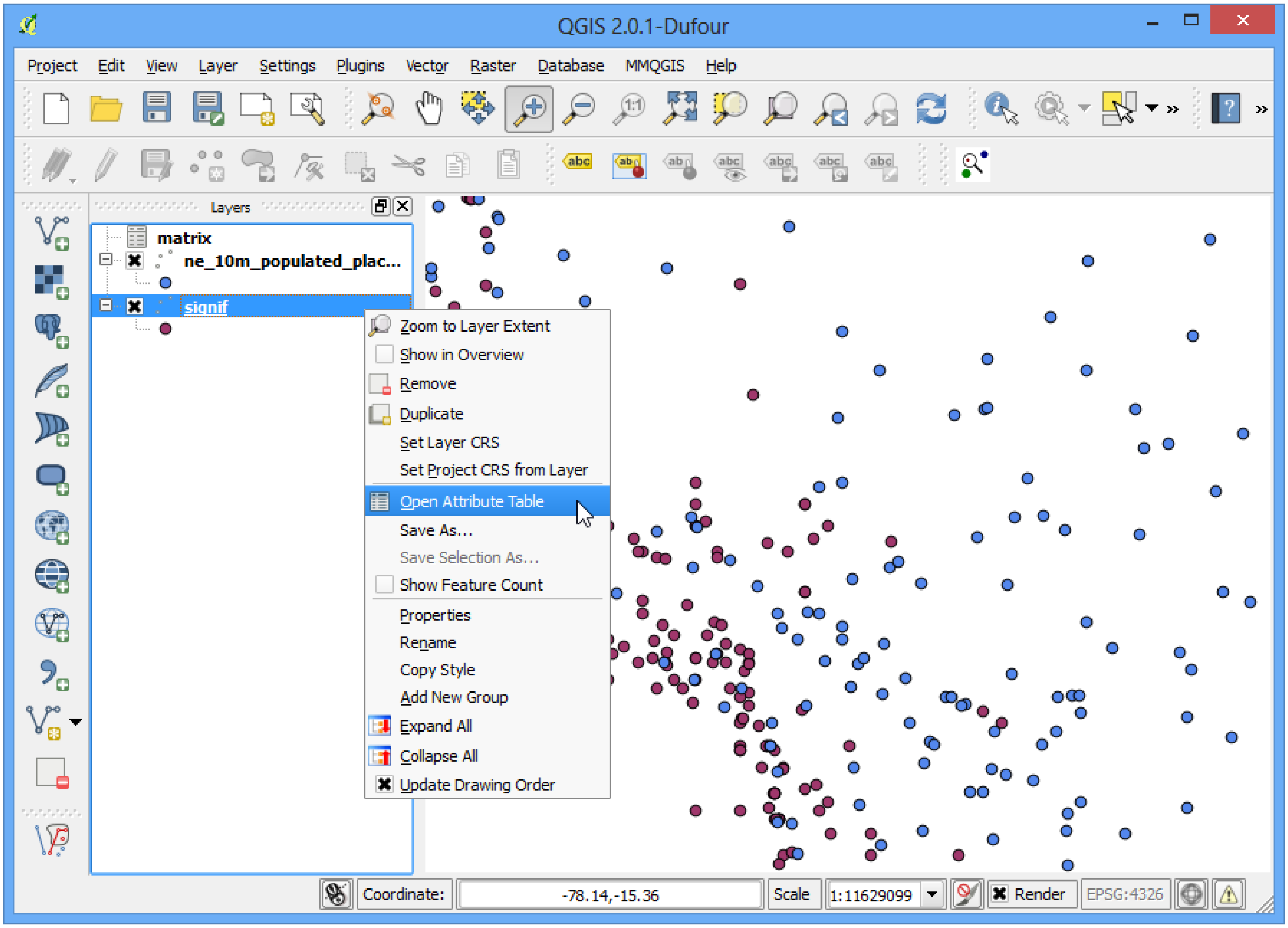
Join layer: matrix
Join field: InputID
Target field: I_D
☒ Cache join layer in virtual memory
☐ Create attribute index on join field

16. 결합 :guilabel:`Joins` 탭에서 결합이 나타날 것입니다 . :guilabel:`OK`를 클릭하십시오 .

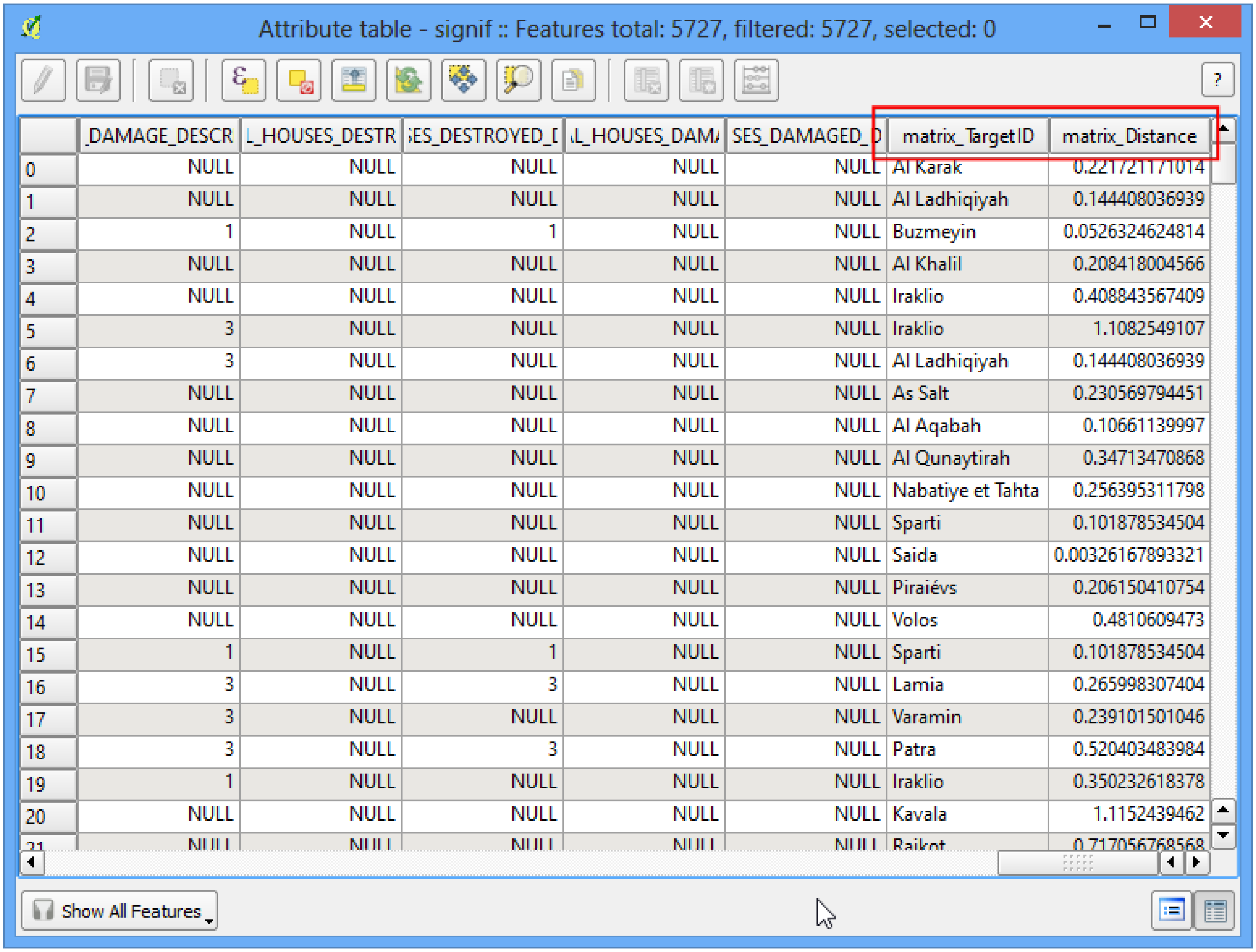

Layer Properties - signif
General
Style
Labels
Fields
Display
Actions
Joins
Diagrams
Metadata
Joins
Restore Default Style
Save As Default
Load Style ...
Save Style
OK
Cancel
Apply
Help

| Join layer | Join field | Target field |
|---|---|---|
| matrix | InputID | I_D |

17. 이제 우측클릭을 통해 지진 레이어의 속성테이블을 열고 속성테이블 열기를 선택합니다 .


QGIS 2.0.1-Dufour
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
MMQGIS
Help
Layers
matrix
ne_10m_populated_plac...
signif
Zoom to Layer Extent
Show in Overview
Remove
Duplicate
Set Layer CRS
Set Project CRS from Layer
Open Attribute Table
Save As...
Save Selection As...
Show Feature Count
Properties
Rename
Copy Style
Add New Group
Expand All
Collapse All
Update Drawing Order
Coordinate:
-78.14,-15.36
Scale
1:11629099
Render
EPSG:4326

18. 모든 지진 객체에 대해 최근린 (인구밀집지역에 가장 가까운 ) 그리고 최근린 거리의 속성을 가지는 것을 볼 수 있을 것입니다.



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

| | _DAMAGE_DESCR | L_HOUSES_DESTR | ES_DESTROYED_ | 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 |