



Harjoitustehtävä 5

# QGIS-harjoituksia: Tietokantaliitokset

Nylander, A. & Muukkonen, P. (2022)



## Johdanto

Tässä harjoituksessa jatketaan erityyppisten tiedostojen käsittelyä ja esikäsittelyä sekä myös harjoitellaan tietokantaliitoksia. Monesti paikkatieto-ohjelmaan tuotava data vaatii jonkinlaisia käsittelyä, jotta sitä voidaan hyödyntää analyyseissä. Siksi onkin tärkeä taito osata muokata dataa sellaiseksi, että se sopii tehtävään analyysiin. Tämä harjoitus sisältää kaksi tehtävää ja hyödyllisiä lisätehtäviä.

Kannattaa tallentaa tässäkin harjoituksessa käytettävät tiedostot sekä tuotettavat kuvat ym. *samaan kansioon* (luo itsellesi kutakin kurssikertaa/moduulia varten oma kansio), jotta ne löytyvät helposti. Harjoituksesta koostetaan **raportti**: jokaisesta tehtävästä raporttiin liitetään visualisointi ja mahdollisesti lyhyt pohdinta.

## Tehtävä 1: Väestöllisiä tunnuslukuja

Tässä tehtävässä käsitellään 5 x 5 km -väestöruututietokantaa. Tietokanta koostuu kahdesta tiedostosta, jotka yhdistetään yhdeksi tietokantaliitoksen (*Join*) avulla. Sitä ennen aineistoa joudutaan hieman esikäsitlemään.

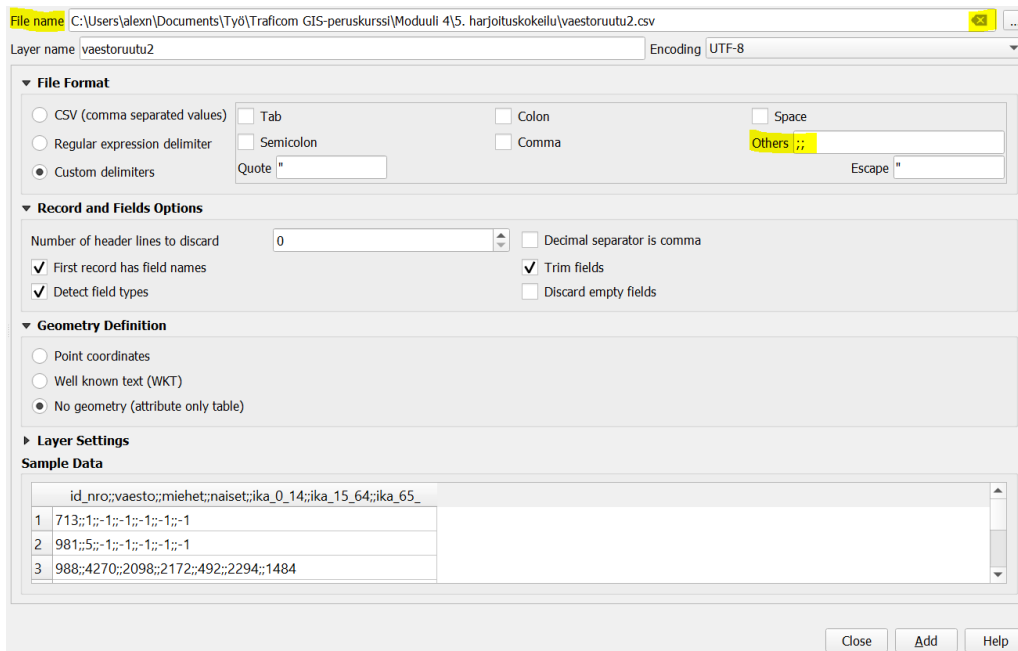
1. Aloitetaan lataamalla tämän harjoituksen (Harjoitus 5) datapaketti [**harjoitus5\_data.zip**] (<https://aoe.fi/#/materiaali/2379>). Pura paketti kurssikerran kansioosi. Voit huomata, että paketti sisältää mm. Shapefile-tiedoston (*vaestoruutu1*) ja csv-tiedoston (*vaestoruutu2*).
2. Tuo vaestoruutu1-taso QGIS:iin *Add Vector Layer (Lisää vektoritaso)* -työkalulla (). Valitse tiedosto (*vaestoruutu1.shp*) ja lisää se painamalla *Add (Lisää)*.
3. Yritetään seuraavaksi tuoda ohjelmaan csv-tiedosto. Avaa siis *Add Delimited Text Layer (Lisää erotinmerkkejä sisältävä tekstitiedosto)* -työkalu ja valitse tiedosto kurssikerran kansioosi ( -painike). Kun katsot tiedoston sisältöä *Sample Data (Esimerkkidata)* -kohdassa, voit huomata sen näyttävän kummalliselta. Niin sarakkeiden nimien kuin arvojenkin välissä on kaksi ;-merkkiä, eikä QGIS osaa lukea kyseisin luetteloerottimin erotettua tiedostoa. Näkymä säilyy samanlaisena, vaikka luetteloerottimeksi asettaisi **Others** (*Muut*) -kohtaan ”;” (kuva 1).

Laatijat:

Alex Nylander & Petteri Muukkonen\*  
 Geotieteiden ja maantieteen osasto,  
 Helsingin yliopisto & HY+

\* petteri.muukkonen@helsinki.fi

Tämä materiaali on laadittu CRITICAL-tutkimushankkeessa (2020–2023), jota on tukenut Strategisen tutkimuksen neuvosto (STN). Lisäksi materiaalin laatimista ovat tukeneet Opettajien akatemia, Helsingin yliopisto sekä HY+. Tekstiä ja kuvia saa käyttää CC BY 4.0 -lisenssillä (vapaa käyttö- ja muokkausoikeus, viittaa alkuperäiseen).

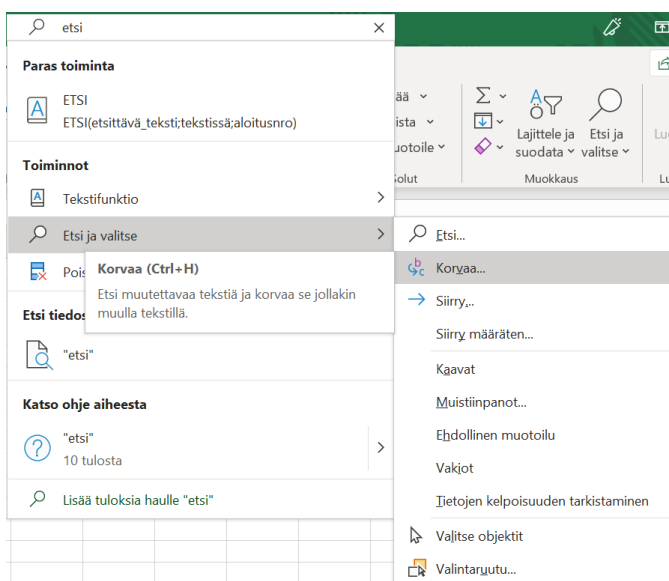


Kuva 1.

4. Klikkaa *Add Delimited Text Layer (Lisää erotinmerkkejä sisältävä tekstitiedosto)* -sivun yläreunassa olevaa -painiketta, jotta tiedosto ei jää taustalle auki ja saat avattua sen toisaalla (*File name/Tiedostonimi* -rivin tulisi siis olla tyhjä).

Eräs tapa tämän ongelman ratkaisemiseksi on avata tiedosto taulukkolaskentaohjelmassa, kuten Excelissä (seuraavien kohtien ohjeet ovat Exceliä varten).

Klikkaa kansiossasi *vaestoruutu2.csv*-tiedostoa ja avaa se Excelissä, jossa luetteloerottimia on helppo muuttaa. Hae Excelin hakukentästä ”*Etsi*” (engl. ”*Find*”) ja valitse *Etsi ja valitse* → *Korvaa* (*Find and select* → *Replace*) (**kuva 2**). Voit myös helposti avata toiminnon painamalla **CTRL + H**.



Kuva 2. Find and select → Replace...

Avautuvassa ikkunassa aseta **kuvan 3** mukaisesti etsittävät (:) ja korvaavat arvot (,) ja paina lopuksi *Korvaa kaikki* (*Replace All*). Nyt luetteloerottimen pitäisi olla pilkku. Tallenna tiedosto ja sulje Excel.

Kuva 3. Etsi ja korvaa

5. Avaa QGIS:ssä jälleen *Add Delimited Text Layer* (*Lisää erotinmerkkejä sisältävä tekstitiedosto*) -työkalu ja hae äsken muokkaamasi tiedosto kurssikerran kansistasi. Valitse *File Format* (*Tiedostomuoto*) -kohtaan **CSV (comma separated values)**, sillä luetteloerotin on nyt tosiaan pilkku. Koska tämä aineisto ei sisällä geometriatietoja, valitaan *Geometry Definition* (*Geometrian määrittelyt*) -kohdassa **No geometry (attribute only table)** (*Ei geometriaa*).

Tarkista *Sample Data* (*Esimerkkidata*) -ruudukosta, että tiedosto näyttää siltä miltä pitää (**kuva 4**). Lisää taso kartalle painamalla *Add* (*Lisää*).

	id_nro	vaesto	miehet	naiset	ika_0_14	ika_15_64	ika_65_
1	713	1	-1	-1	-1	-1	-1
2	981	5	-1	-1	-1	-1	-1
3	988	4270	2098	2172	492	2294	1484

Kuva 4. csv-tason lisääminen.

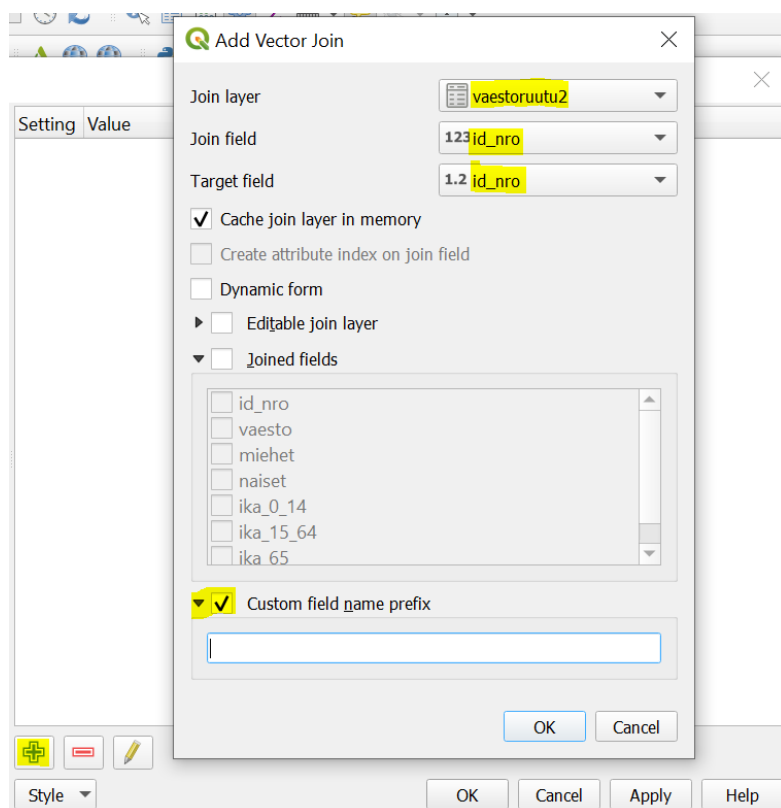
6. Tarkastele seuraavaksi kunkin tason attribuuttitaulua. Taulukkotiedosto sisältää yhtä lailla väestöruutuihin sidottua informaatiota, mutta taulukosta puuttuu kohteiden geometria. Vektoritaso (*vaestoruutu1*) sisältää taas kohteiden geometriat, mutta taulukosta puuttuu väestöön liittyviä tunnuslukuja. Näemmä *vaestoruutu1*- ja *vaestoruutu2*-tasoilla on kuitenkin eräs yhteinen nimittäjä: id-numero. Toisin sanoen jokaisella väestöruudulla on oma yksilöllinen id-numeronsa, joka löytyy kummankin tason attribuuttitaulusta. Sen ansiosta taulukkotiedoston sarakkeet voidaan liittää vektoritason attribuuttitauluun tietokantaliitoksella.

7. Klikkaa hiiren oikealla *vaestoruutu1*-tason nimeä. Valitse *Properties* → *Joins* (*Ominaisuudet* → *Liitokset*).

Paina *Joins/Liitokset*-välilehdellä -painiketta niin, että *Add Vector Join* -työkalu (**kuva 5**) avautuu. Valitse *Join Layer* -kohtaan se taso, josta tuodaan attribuuttitietoja, eli *vaestoruutu2*. *Join field* – ja *Target field* -kohtiin valitaan ne sarakkeet, jotka toimivat yhteisenä nimittäjänä kummassakin tasossa. Tällä kertaa id-sarakkeen nimi on kummassakin tasossa *id\_nro*, joten valitaan se kumpaankin kohtaan.

Rastita alemmaa kohta *Custom field name prefix* mutta kumita kenttä tyhjäksi. Tämä ”prefiksi” tarkoittaa etuliitettä, joka liitettäisiin tuotavien sarakkeiden nimien eteen. Tällä kertaa sellaista ei kuitenkaan tarvita, ja sarakkeiden nimet näyttäivät helpommin luettavilta ja informatiivisemmilta ilman sitä.

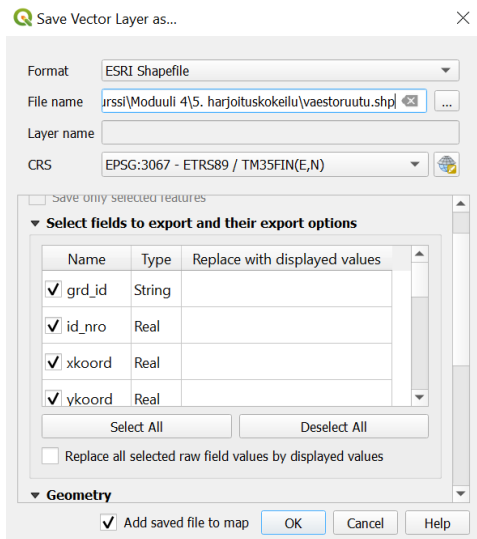
Paina *OK* ja uudelleen *OK* tallentaaksesi muutokset.



Kuva 5.

8. Tarkastele *vaestoruutu1*-tason attribuuttitaulua. Huomaat, että sarakkeita on nyt enemmän, ja mukana on toisesta taulukosta tuodut sarakkeet. Tuodut sarakkeet eivät kuitenkaan pysyvästi ole vielä attribuuttitaulussa, vaan taso täytyy tallentaa uudelleen.

Klikkaa *vaestoruutu1*-tasoa hiiren oikealla ja valitse *Export* → *Save features as...* (*Vie* → *Tallenna kohteet nimellä...*). Valitse [...] -painikkeella kansio, jonne haluat tason tallentaa ja tallenna se esim. nimellä *vaestoruutu* (**HUOM!** Katso, että formaatti on vektorimuotoinen, ja että tallennat tason oikeassa CRS:ssä (**kuva 6**)).



**Kuva 6.**

9. Nyt väestöruutuaineisto sisältää myös väestöön liittyviä tunnuslukuja: väestön kokonaismäärä ruuduittain, miesten määrä, naisten määrä sekä 0–14-, 15–64- ja yli 65-vuotiaiden määrä. Aineistossa on edelleen sarakkeita, joita ei voi järkevästi käyttää näihin analyyseihin: alle 10 asukkaan väestöruutujen tarkemmille tiedoille on turvallisuussyistä jätetty arvo -1.

Poistetaan nuo arvot esimerkiksi käyttämällä *Extract by Attribute* (*Irrota attribuutin perusteella*) -työkalua. Valitaan **kuva 7** mukaisesti *Input layeriksi* (*syötetasoksi*) **vaestoruutu**-taso, ja valitaan vain ne ruudut, joiden väkiluku on vähintään 10. Klikkaa *Extracted (attribute)* -kohdan oikealta puolelta **kolmea pistettä**, valitse *Save to file* (*Tallenna tiedostoon*) ja tallenna taso nimellä *vaestoruutu\_yli\_10* kurssikerran kansioosi. Klikkaa *Run* (*Suorita*).

Extract by Attribute

Parameters Log

Input layer  
vaestoruutu [EPSG:3067]

☐ Selected features only

Selection attribute  
123 vaesto

Operator  
≥

Value [optional]  
10

Extracted (attribute)  
Traficom GIS-peruskurssi/Moduuli 4/5. harjoituskokeilu/vaestoruutu\_yli\_10.shp

☒ Open output file after running algorithm

Extracted (non-matching) [optional]  
[Skip output]

0%

Run as Batch Process... Run Close Help

**Extract by attribute**

This algorithm creates a new vector layer that only contains matching features from an input layer. The criteria for adding features to the resulting layer is defined based on the values of an attribute from the input layer.

Kuva 7.

Päätä seuraavaksi, minkä edellä mainituista väestöllisistä tunnusluvuista haluaisit visualisoida kartallasi. Valitse haluamasi muuttuja, ja laske sen osuus kokonaisväestöön nähden. Avaa siis attribuuttitaulussa *Field Calculator* (*Kentän arvojen laskin*) (**kuva 8**). Anna *Output field name*ksi (*Nimi uudelle kentälle*) jokin **sopiva nimi**, muuta kenttätyyppiä **Decimal number (real)** (*Desimaalinumero (reaali)*), ja liitä seuraava lauseke **Expression**/Lauseke-kenttään (muokkaa sanan *Käyttämäsi\_muuttujan\_nimi* tilalle käyttämäsi muuttujan nimi).

**"Käyttämäsi\_muuttujan\_nimi" / "vaesto" \* 100**

☒ Create a new field ☐ Update existing field

☐ Create virtual field

Output field name: osuus\_65+

Output field type: Decimal number (real)

Output field length: 10 Precision: 2

Expression

Function Editor

"ika\_65\_" / "vaesto" \* 100

Feature: 5kmN6635E0270 Preview: 34.75409836065574

Search... Show Values

- abc grd\_id
- 1.2 id\_nro
- 1.2 xkoord
- 1.2 ykoord
- abc kunta
- 123 vaesto
- 123 miehet
- 123 naiset
- 123 ika\_0\_14
- 123 ika\_15\_64
- 123 ika\_65\_

Kuva 8. Uuden kentän luominen – esimerkkinä yli 65-vuotiaiden osuus

Lisää uusi sarake painamalla OK.

10. Muokataan seuraavaksi tason symboliikkaa (**kuva 9**). Klikkaa tason nimeä hiiren oikealla ja valitse *Properties* → *Symbolology* (*Ominaisuudet* → *Kuvaustekniikka*). Valitse sitten sivun yläreunasta esitystavaksi **Graduated** (*Porrastettu*) ja laita *Value* (*Arvo*) -kohtaan äsken luomasi tietyn ryhmän osuutta kuvaava **sarake**. Aseta mielestäsi sopiva *Color ramp* (*Liukuväri*) ja valitse *Mode* (*Tila*) -kohtaan **Natural Breaks** (*Jenks*). Paina *OK* lisätäksesi muutokset kartalle.

Symbol	Values	Legend
✓	0,00 - 20,29	0 - 20
✓	20,29 - 30,23	20 - 30
✓	30,23 - 41,18	30 - 41
✓	41,18 - 56,52	41 - 57
✓	56,52 - 100,00	57 - 100

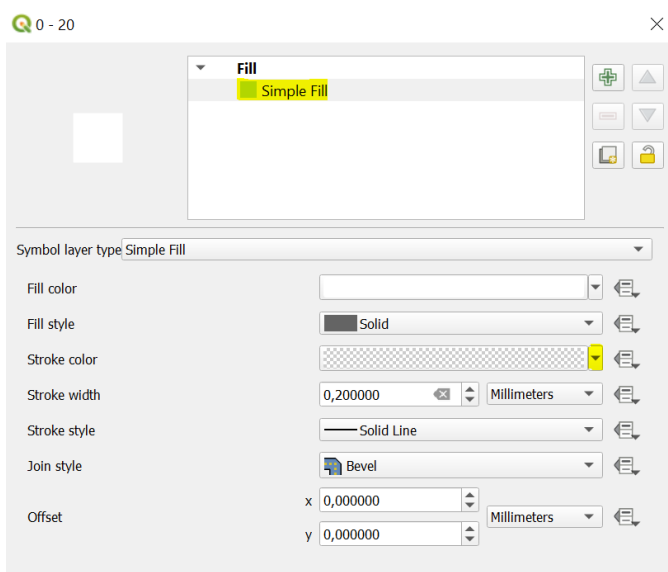
**Kuva 9.**

Voit kuitenkin huomata, että symboliikka näyttää epäselvältä, sillä alue on suuri, ja väestöruutujen mustat reunat tummentavat kokonaisuutta kohtuuttomasti. Muokataan siis edelleen *Symbolology*-sivulta tason muotoilua (**kuva 10**). Klikkaa ensin ylimmän luokan *Symbol*-kohtaa (1.) ja sitten ylempää toista *Symbol*-kenttää (2.).

Symbol	Values	Legend
✓	0,00 - 20,29	0 - 20
✓	20,29 - 30,23	20 - 30
✓	30,23 - 41,18	30 - 41
✓	41,18 - 56,52	41 - 57
✓	56,52 - 100,00	57 - 100

**Kuva 10.**

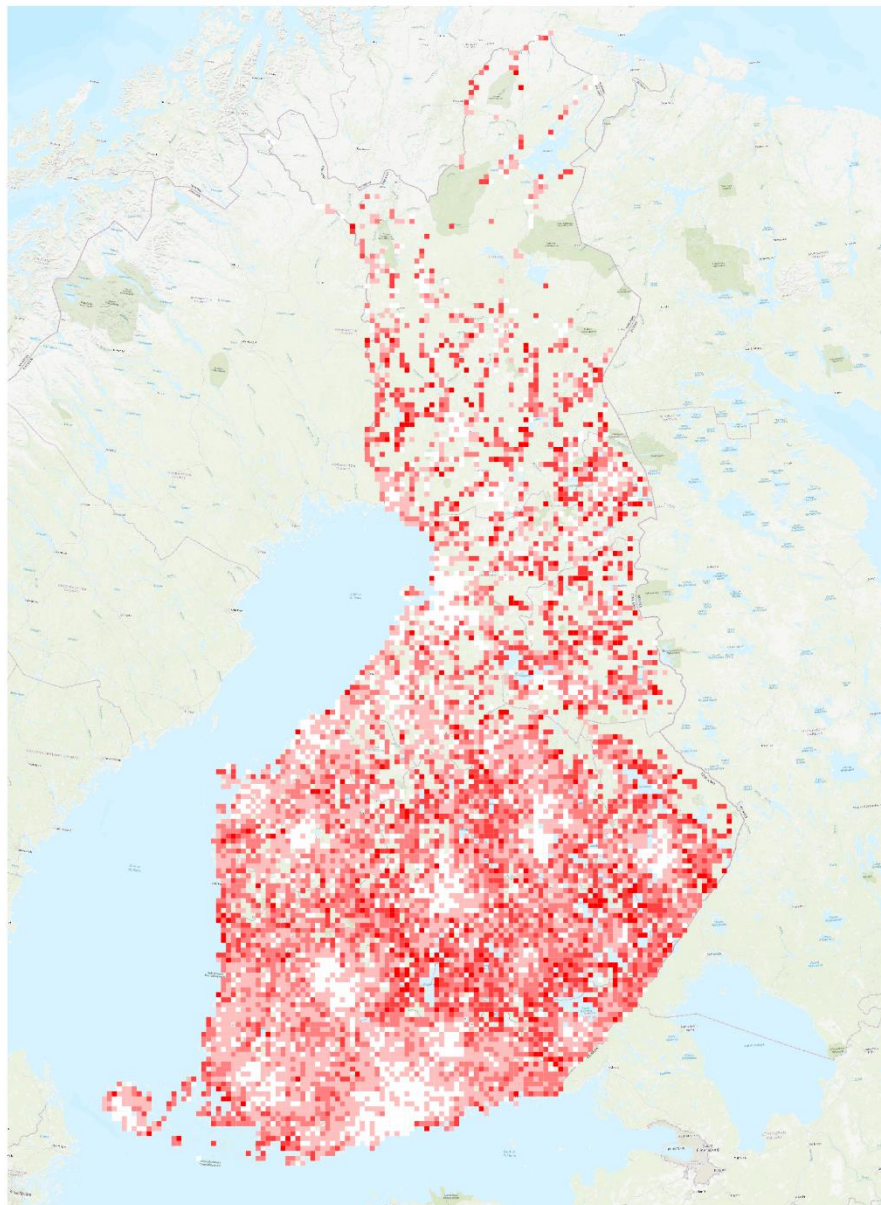
Avautuvasta ikkunasta (**kuva 11**) klikkaa *Simple Fill* (*Yksinkertainen täyttö*) -kohtaa ja avaa pienestä nuolinäppäimestä valikko kohdassa *Stroke color* (*Viivan väri*). Valitse valikon yläosasta *Transparent Stroke* (*Läpikuultava viiva*). Paina OK tallentaaksesi muutokset. Toista seuraavaksi samat muutokset muille neljälle luokalle!



**Kuva 11.**

Nyt alkaa olla valmista: kartan pitäisi näyttää huomattavasti selkeämmältä. Luodaan vielä lopullinen kartta. Klikkaa siis *Project → New Print Layout...* (*Projekti → Uusi taitto...*). Saat vaihdettua taiton pystyasentoiseksi (Suomen karttaan sopivammaksi) klikkaamalla taittonäkymää hiiren oikealla ja painamalla *Item Properties → Orientation → Portrait* (*Sivun ominaisuudet → Suunta → Pysty*). Seuraavaksi lisää kartta, mittakaava, pohjoisnuoli ja legenda. Kun olet karttaan tyytyväinen, tallenna se kurssikerran kansioosi painamalla *Layout → Export as Image* (*Taitto → Vie kuvaksi*). **Liitä kartta raporttiisi!** Karttasi voi näyttää esim. tältä (**kuva 12**):





0 75 150 km



Yli 65-vuotiaiden osuus 5x5km väestöruuduittain, %

0 - 20

20 - 30

30 - 41

41 - 57


57 - 100

Taustakartta: ESRI Topo

**Kuva 12.** Yli 65-vuotiaiden osuus (%) 5x5km väestöruuduittain. Kartta kattaa ruudut, joiden väkiluku on vähintään 10 asukasta.

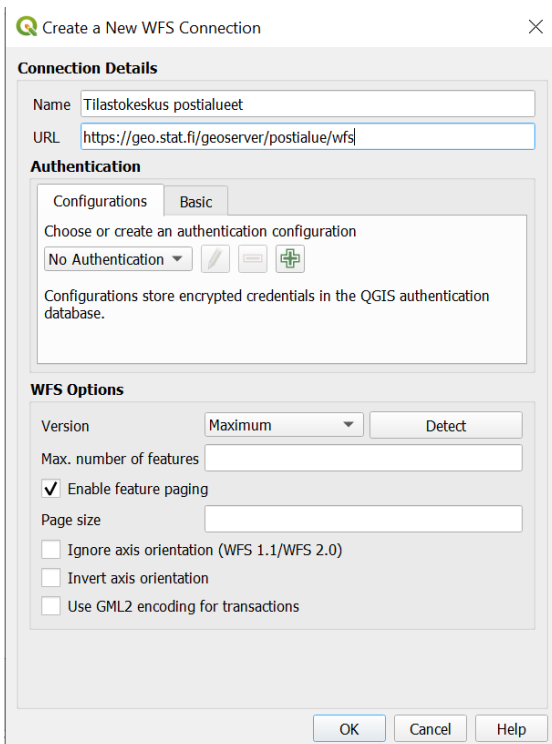
## Tehtävä 2: Sosioekonominen summaindeksi

Tässä tehtävässä harjoitellaan edelleen datan esikäsitlemistä, muokkaamista ja visualisointia. Tehtävänä on visualisoida esim. tätä [Helsingin kaupungin raporttia](#) mukaillen laskettava ”huono-osaisuuden summaindeksi”. Indeksä on em. raporttiin nähden hieman yksinkertaistettu sosioekonomista segregatiota kuvaava indeksi: korkeampi arvo kuvaa alueita, jonne on keskittynyt huono-osaisuutta.

1. Aloitetaan tuomalla postialueittain jaettu tilastoaineisto Tilastokeskuksen WFS-rajapinnan kautta. Klikkaa *Add WFS Layer* (*Lisää WFS-taso*, ) -painiketta. Paina *Server Connections* (*Palvelinyhteydet*) -valikon alta *New* (*Uusi*) lisätäksesi uuden WFS-yhteyden (**kuva 13**). Lisää yhteys lisäämällä nimi ja URL-osoite ja paina sitten OK.

Name: Tilastokeskus postialueet

URL: <https://geo.stat.fi/geoserver/postialue/wfs>



Kuva 13.

Klikkaa *Connect* (*Yhdistä*) yhdistääksesi WFS-rajapintaan. Valitse sitten *Paavo-tilastoaineisto ja -postinumeroalueet 2021* -niminen taso (pitäisi olla listan ylinnä) ja lisää se painamalla *Add* (*Lisää*). Voit myös lisätä taustakartan esimerkiksi tässä vaiheessa.

2. Tarkastellaan attribuuttitaulua. Se näyttää sisältävän monta kymmentä saraketta, jotka sisältävät tietoa eri ryhmiin kuuluvien ihmisten määrästä. Nimet ovat epäinformatiivisia, mutta aineistokuvaus ja selitteet löytyvät tästä [täältä](#)!

Tehdään tämä analyysi pääkaupunkiseudun mittakaavassa. Valitaan siis seuraavaksi vain ne postinumerot, jotka (täysin tai pääosin) sijaitsevat Helsingin, Espoon, Vantaan tai Kauniaisten kaupunkien alueella. Se onnistuu esimerkiksi kuntakoodien (sarake ”*kunta*”) avulla.

Pääkaupunkiseudun kuntien kuntakoodit ovat 049 (Espoo), 091 (Helsinki), 092 (Vantaa) ja 235 (Kauniainen).

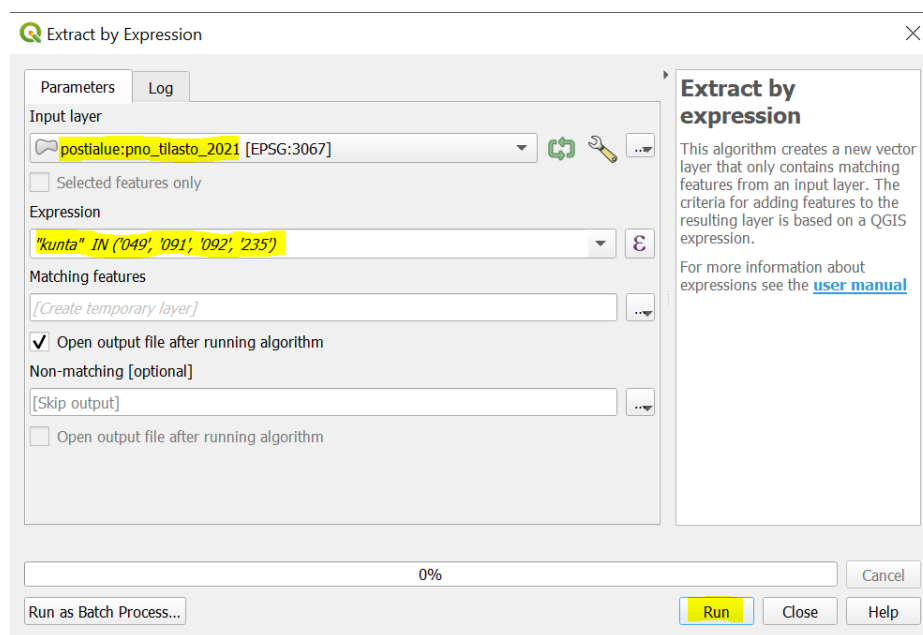
Muun muassa *Extract by Expression* (Irrota käyttäen lauseketta) -työkalua käyttämällä voidaan kuntakoodien avulla valita vain tiettyjen kuntien postinumeroalueet.

3. Avaa esim. *Processing Toolboxin* (Prosessointityökalut) kautta *Extract by Expression* -työkalu (kuva 14).

Lisää *Input layer* (syötetaso) -kohtaan postialueet sisältävä **taso**. *Expression* (Lauseke) -kohtaan asetetaan lauseke, jolla haetaan halutut objektit tiettyjä arvoja apuna käyttäen. Eräs tapa on käyttää ns. **SQL IN** -lauseketta:

Kirjoita siis (tai kopioi tästä) seuraava lauseke *Expression* (Lauseke) -kenttään:

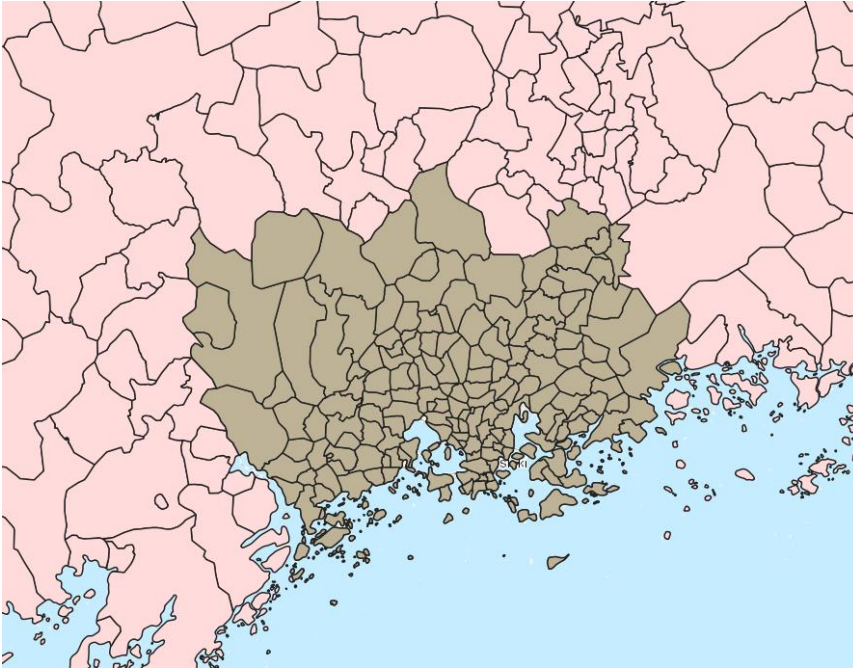
**"Kunta" IN ('049', '091', '092', '235')**



Kuva 14.


Paina **Run** (Suorita).

4. Nyt tasovalikkoon pitäisi olla ilmestynyt *Matching features* -niminen väliaikainen taso, jonka tulisi näyttää suurin piirtein tältä (kuva 15):



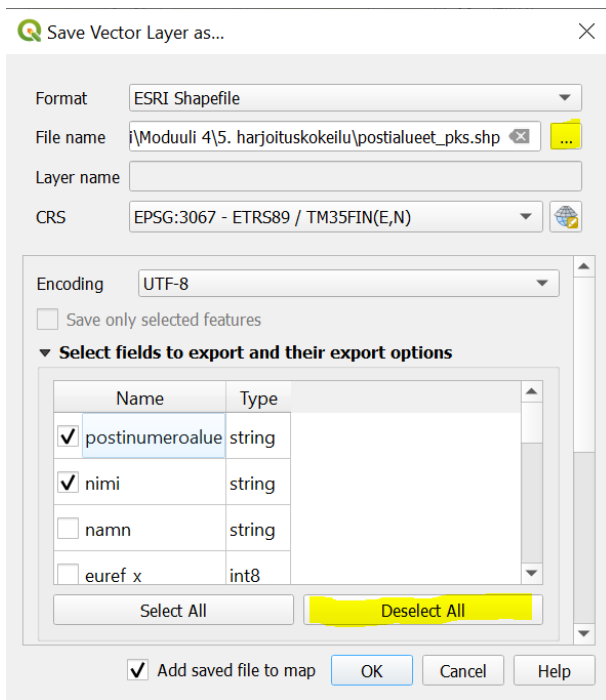
Kuva 15.

Tallennetaan se seuraavaksi pysyväksi ja valitaan samalla vain ne sarakkeet, joiden tietoja tarvitsemme analyysin suorittamisessa. Klikkaa siis *Matching features* -tasoa hiiren oikealla, ja paina *Export* → *Save features as...* (Vie → Tallenna kohteet nimellä...).

Valitse formaatiksi esimerkiksi .shp ja valitse  -painikkeen avulla kansio, jonne taso tallennetaan. Anna tasolle nimeksi esimerkiksi *postalueet\_pks*. Koska vain osa sarakkeista halutaan viedä uuteen tiedostoon, klikataan ensin **Deselect All** (*Poista valinnat*, **kuva 16**). Sitten rastitetaan listasta ne sarakkeet, joita tarvitaan sosioekonomisen summaindeksin laskemisessa:

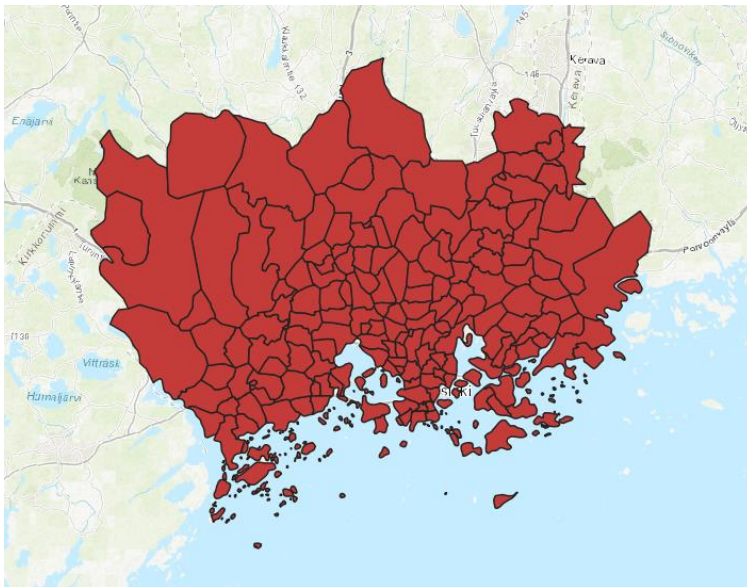
- **postinumeroalue** (postinumero)
- **nimi** (postinumeroalueen nimi)
- **ko\_ika18y** (täysi-ikäisten ihmisten määrä)
- **ko\_perus** (vain peruskoulutuksen omaavien määrä)
- **hr\_pi\_tul** (pienimpään tuloviidennekseen kuuluvien määrä)
- **pt\_tyott** (työttömänä vuoden lopussa olleiden 18-64v määrä)





**Kuva 16.**

Klikkaa *OK* tuodaksesi tason kartalle. Voit poistaa väliaikaisen *Matching features* -tason ja alkuperäisen postinumerotason kartalta. Näkymäsi pitäisi näyttää suurin piirtein tältä (**kuva 17**).

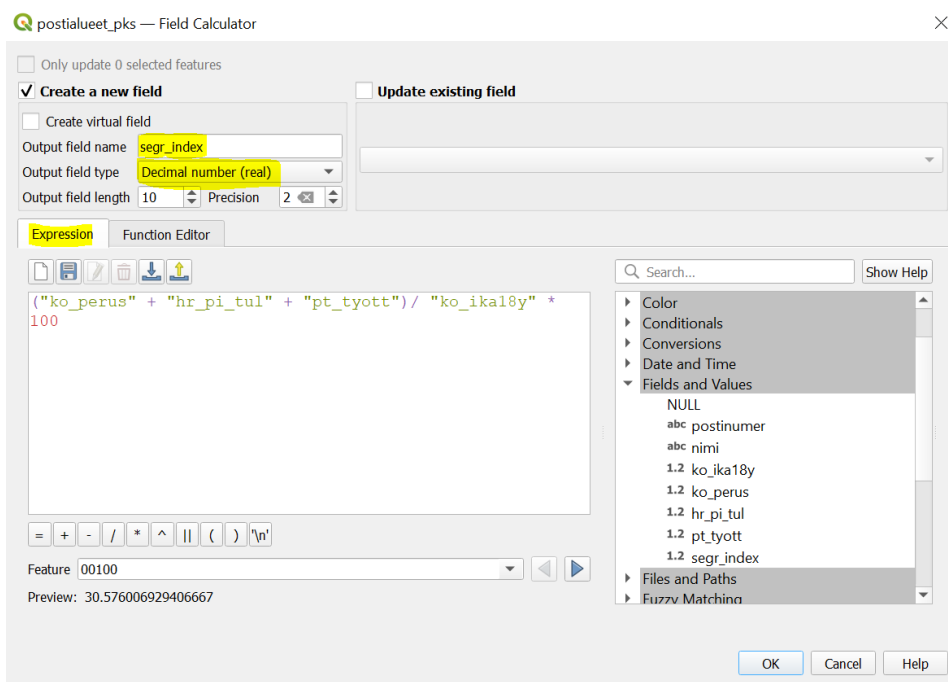


*Kuva 17.*

**5.** Kun avaat uuden tason attribuuttitaulun, siellä pitäisi nyt olla ainoastaan äsken valitut kuusi (6) saraketta. Lasketaan seuraavaksi niiden avulla summaindeksi kullekin postinumeroalueelle. Avaa jälleen *Field Calculator (Kentän arvojen laskin)* (**kuva 18**), jolla luodaan uusi sarake. Anna *Output field nameksi* (*Nimi uudelle kentälle*) esimerkiksi *segr\_index*, ja määritä kentän tyyppiä *Decimal number (real)*. Kirjoita vielä *Expression (Lauseke)* -kohtaan seuraava lauseke:

$$("ko\_perus" + "hr\_pi\_tul" + "pt\_tyott") / "ko\_ika18y" * 100$$

jossa matalasti koulutettujen, alhaisimpaan tuloluokkaan kuuluvien ja työttömien työikäisten määrä jaetaan kaikkien täysi-ikäisten kesken ja kerrotaan sadalla.

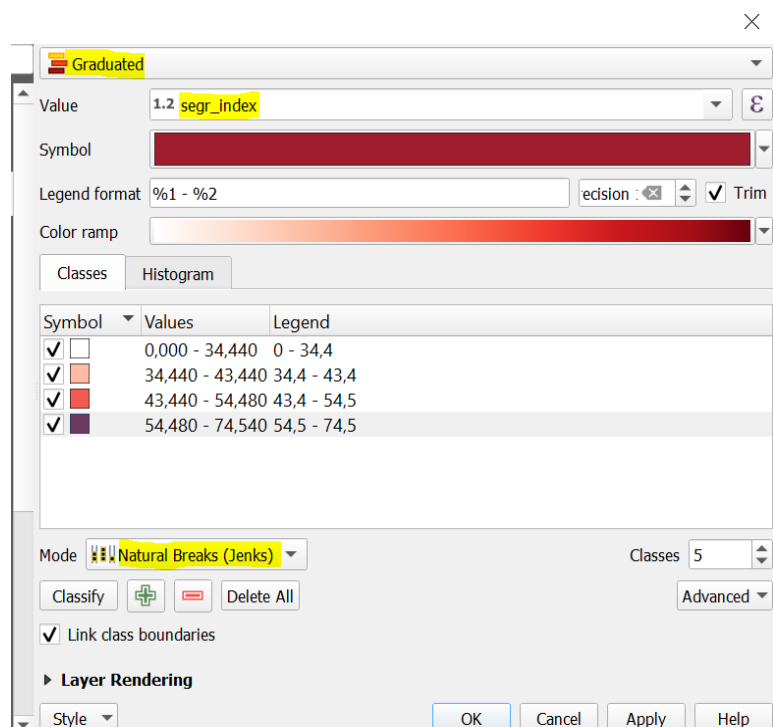


Kuva 18.

Klikkaa **OK** lisätäksesi uuden sarakkeen. Lopeta attribuuttitaulun muokkaaminen ja tallenna taso.

Alkaa olla valmista! Muutetaan vielä karttatason symboliikkaa klikkaamalla tason nimeä hiiren oikealla ja valitsemalla *Properties* → *Symbolology* (*Ominaisuudet* → *Kuvaustekniikka*) (**kuva 19**).

Valitse ylävalikosta *Graduated* (*Porrastettu*), aseta *Value/Arvo*-kohtaan summaindeksiä kuvaava sarakke (*segr\_index*) ja valitse mieleisesi *Color ramp* (*Liukuväri*). Aseta *Modeksi* *Natural Breaks* (*Jenks*). Klikkaa ylimpänä olevan luokan nimeä (luokka joka sisältää arvot 0 - 0) ja poista se -painikkeen avulla. Kun väriskaala on mielestäsi hyvä, tallenna muutokset painamalla **OK**.

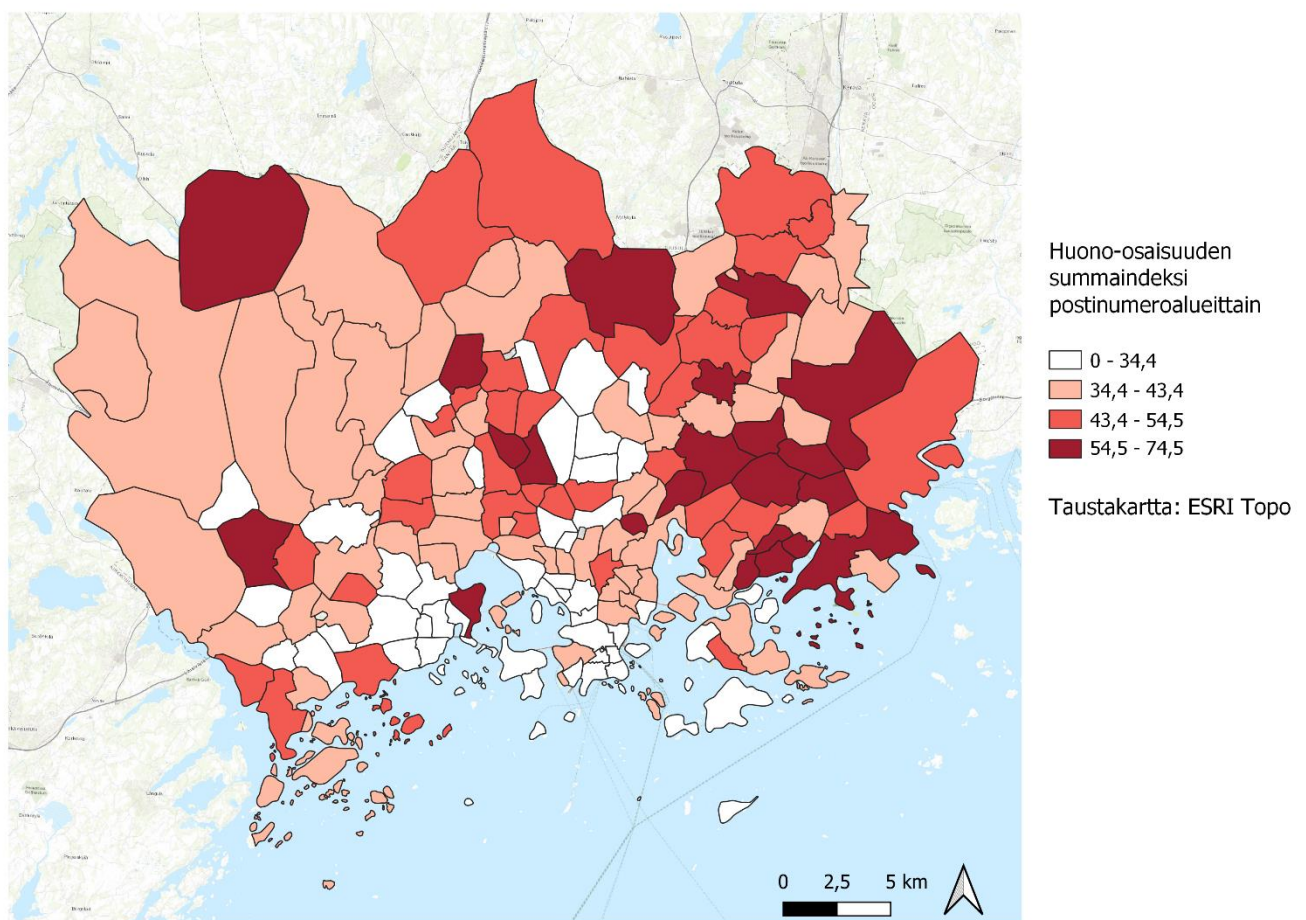


Kuva 19.

6. Lisää mielestäsi sopiva taustakartta, jos et ole tähän mennessä lisännyt sellaista. Viedään karttataso Layout-näkymään klikkaamalla Project → New Print Layout... (*Projekti → Uusi taitto*).

Lisää karttataso *Layout*-näkymään ja liitä pohjoisnuoli ja mittakaava. Lisää seuraavaksi legenda ja muokkaa merkkienselitteitä mahdollisimman kuvaaviksi (paina *Auto update* pois päältä muokataksesi legendan kohteiden määrää ja tyyliä). Tallenna lopuksi karttaesitys klikkaamalla *Layout → Export as Image* (*Taitto → Vie kuvaksi*) ja tallenna kuva kurssikerran kansioosi.

**Liitä kartta raporttiisi ja pohdi muutamalla lauseella**, mitä kartta kertoo, ja mitä heikkouksia sillä ehkä on. Karttasi voi näyttää esimerkiksi tältä (**kuva 20**). Huomioi, että kun visualisoidaan harvan tuntemaa indeksiä, on hyvä hieman avata sen merkitystä kuvatekstissä.



**Kuva 20.** Mukautettu huono-osaisuuden summaindeksi, jota voidaan käyttää eräänlaisena segregaaion tunnuslukuna. Indeksii kertoo postinumeralueittain matalasti palkattujen, työttömien ja matalasti koulutettujen yhteenlasketun osuuden (%) alueen koko väestöön nähden. Korkea indeksi kuvaa alueellisesti kasaantunutta huono-osaisuutta. Data: Tilastokeskus/2021.

## Lisätehtävä 1: Peruskartan georeferointi

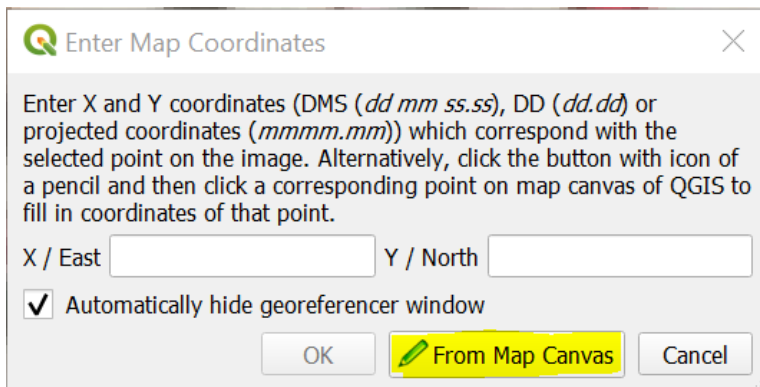
**HUOM!** Tämä on hyödyllinen tehtävä, jonka tekeminen kannattaa, vaikka se onkin lisätehtävä.

Mikäli tulee vastaan tilanne, että haluaa esimerkiksi käyttää vanhaa paperikarttaa tai vaikkapa ilmakuvaa paikkatietanalyysissä, on georeferointi hyödyllinen taito. Georeferoinnilla tarkoitetaan kartan kiinnittämistä paikkaan; toisin sanoen sitä, että jokaisella kartan kohteella on koordinaatit.

1. Avaa uusi QGIS-projekti ja tuo sinne ensimmäisenä OpenStreetMap (OSM) Standard - taustakartta (*Web/Verkko* → *QuickMapServices* → *OSM* → *OSM Standard*). Katso, että projektin CRS (QGIS:n oikea alakulma) on *EPSG:3857*.

2. Klikkaa *Raster* → *Georeferencer (Rasteri* → *Georeferoija)*, niin georeferointityökalu aukeaa. Paina seuraavaksi -painiketta tuodaksesi harjoituksen datakansioista [**harjoitus5\_data**] *203403\_1979*-niminen tiedosto. Kyseessä on kuva vuoden 1979 1:20 000 peruskartasta, joka kattaa osan itäistä Espoota ja läntistä Helsinkiä ja josta puuttuvat geometriatiedot. *Georeferencer*-työkalulla kartta voidaan sitoa koordinaatteihin.

3. Seuraavaksi etsitään vastinpisteitä: koita löytää kartasta kohteita (teiden risteykset, rakennusten kulmat jne.), jotka eivät ole muuttuneet neljänkymmenen vuoden aikana. Klikkaa sellaista kohdetta peruskartasta, jolloin aukeaa *Enter Map Coordinates* -ikkuna (**kuva 21**). Paina **From Map Canvas** -painiketta ja etsi sitten *OpenStreetMap*-kartasta sama piste ja klikkaa sitä. Paina *OK* lisätäksesi vastinpisteen *GCP Table (Muunnospistetaulu)* -listaan.



Kuva 21.

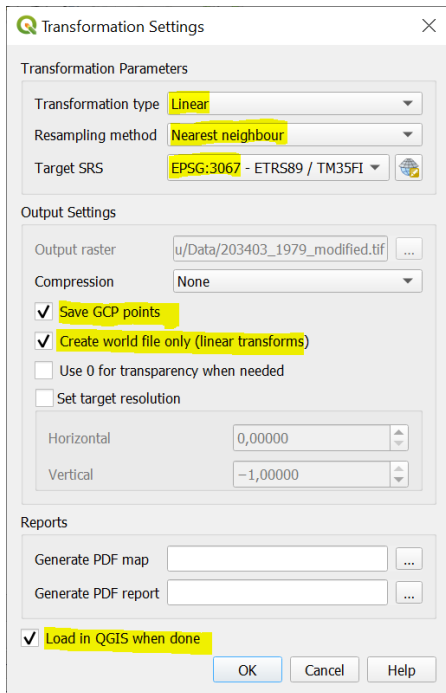
4. Etsi seuraavaksi muutama vastinpiste lisää ympäri karttaa – mitä enemmän pisteitä on, sitä paremmin georeferointi onnistuu. Noin 6–10 pistettä riittää.

5. Kun olet valinnut riittävästi pisteitä, klikkaa *Settings* → *Transformation Settings (Asetukset* → *Muunnoksen asetukset*). Valitse seuraavasti:

Transformation type/Muunnoksen tyyppi:	<b>Linear/Lineaarinen</b>
Resampling method/Uudelleenotannan menetelmä:	<b>Nearest neighbour</b>
Target SRS/Tuloskoordinaattijärjestelmä:	<b>EPSG:3067 – ETRS89 / TM35FIN</b>

Klikkaa rastit myös kohtiin **Save GCP points** (*Tallenna muunnospisteet*) ja **Create world file only (linear transforms)** (*Luo vain world-tiedosto*) sekä **Load in QGIS when done** (*Avaa tulorasteri QGIS:ssä kun valmis*).

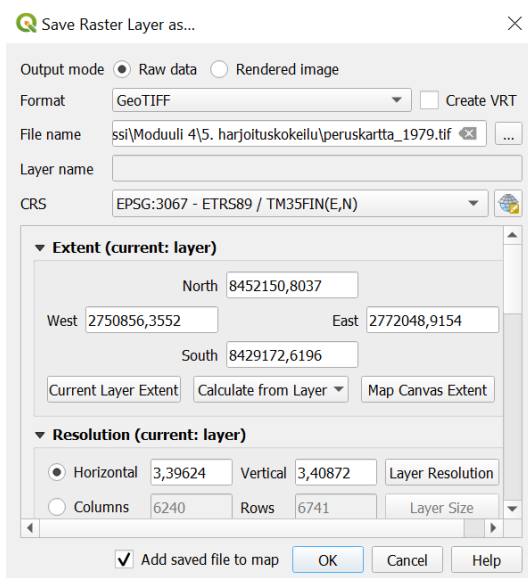




Kuva 22.

Määritettyäsi asetukset, voit aloittaa itse georeferoinnin painamalla -painiketta. Kun georeferointi on valmis, voit huomata, että taso on ilmestynyt kartalle. Anna sille projektin CRS:ää vastaava koordinaattijärjestelmä (EPSG:3857) klikkaamalla hiiren oikealla tason nimeä ja valitsemalla *Layer CRS* → *Set to EPSG:3857* (*Tason koordinaattijärjestelmä* → *Aseta EPSG:3857*).

**6.** Klikkaa seuraavaksi karttatason nimeä hiiren oikealla ja valitse *Export* → *Save to file...* (*Vie* → *Tallenna tiedostoon*) (**kuva 23**). Tallenna sitten kartta kurssikerran kansioosi esimerkiksi nimellä *peruskartta\_1979*. Anna sitten tasolle CRS:ksi EPSG:3067 ja paina *OK*.



Kuva 23.

7. Alkaa olla valmista. Vaihda nyt vielä projektin koordinaattijärjestelmä (QGIS:n oikea alakulma) vastaamaan äsken tallennetun tason koordinaattijärjestelmää (EPSG:3067). Nyt peruskartan pitäisi näkyä oikeassa kohdassa kartalla. Klikkaa peruskartta\_1979-tason nimeä hiiren oikealla, valitse *Properties* → *Transparency* (*Ominaisuudet* → *Läpinäkyvyys*) ja muokkaa hieman tason läpinäkyvyyttä (esim. 70 %:iin), jotta näet taustakartan samaan aikaan.

Kuinka tarkkaan kartta on georeferoitu? Liitä **kuvankaappaus kartasta ja parin lauseen pohdinta** raporttiisi.

## Lisätehtävä 2: Geokoodaus

---

1. Lataa QGIS:n *Manage and Install Plugins* (*Hallitse ja asenna lisäosia*) -valikosta itsellesi Digitransit.fi Geocoder -niminen lisäosa. Kun olet ladannut lisäosan, hae *Processing Toolbox* (*Prosessointityökalut*) -hausta *Geocode Addresses in a CSV file* (*Geokoodaa CSV-tiedoston sijainteja*) -niminen toiminto ja klikkaa sitä.

2. Tuo Input CSV file (*Geokoodattava CSV-tiedosto*) -kohtaan jokin osoitteita sisältävä CSV-tiedosto (voit käyttää kurssikerran kansioista löytyvää ruokakaupat-tiedostoa tai voit luoda/etsiä oman tiedoston). Aseta *Column Separator* (*Sarake-erotin*) -kohtaan luetteloerotinmerkki, joka CSV-tiedostossa on käytössä.

3. Address field name(s) as a comma separated list (*Osoitekenttien nimet pilkuin eroteltuna listana*) -kohtaan kirjoita niiden sarakkeiden nimi (tai nimet) jotka muodostavat osoitteet. Määritä loput asetukset ja paina OK. **Liitä raporttiisi kuvankaappaus** osoitepisteistä. Tarkastele, miten hyvin geokoodaus toimi: ovatko osoitteet oikeilla paikoillaan?