

7. Excel táblából származó adatok behívása és hozzákapcsolása meglévő shapefile-okhoz /geokódolás/

Ehhez a feladathoz a substring vagy hasonló szövegfeldolgozó függvény ismerete szükséges.

Ebben a feladatban Excel táblából importálunk be adatokat a QGIS-be. Az importált adattáblát hozzáfűzzük már meglévő adatsorhoz. Adott egy shapefile is, ez tartalmazza Fejér megye összes települését, ennek az attribútum tábláját szeretnénk kibővíteni a lakosság korcsoportos adatsorával.

Készítendő egy olyan shape-állomány, amelynek attribútum táblázata tartalmazza az Excelből átvett adatokat.

Adatok: FEJER_TELEPULESEK.SHP, és a KSH* honlapjáról letöltött és szűkített adatsor, amely tartalmazza a férfiak és nők 5 évenkénti korcsoportok szerinti bontását. (07_4_1_2_1.csv)

*KSH:

http://www.ksh.hu/nepszamlalas/reszletes_tablak

ALUL kapott megye-->katt-->Településsoros adatok-->4.1.2.1. A népesség korcsoport és nemek szerint, 2011 Excel táblázat. Ami kell belőle:

Települések adatai az egyes korcsoportokról és a CSAK az Összesen oszlopai (férfi+nő együtt)

Ebből készítsünk egy CSV fájlt. A CSV egy olyan szöveges állomány legyen, amelyben az egyes adatokat pontosvessző választja el egymástól, egy sor egy település adatait tartalmazza. Figyeljünk arra, hogy a települések neve előtt ugyanabban a cellában kódok is szerepelnek! Ezeket célszerű az Excel függvényei segítségével levágni, illetve kitörölni a felesleges adatokat. Figyeljünk arra, hogy az Excel, beállításától függően az ezres helyi érték szerint, hármassal csoportosítja a számokat. Ne legyen szóköz a számokban! Pl. 1 024, helyesen: 1024.

Megoldás, magyarázat

A településnév kivágásához magyar nyelvű Excelben: KÖZÉP(\$A70;9;HOSSZ(\$A70))

Magyarázat: A közép függvény az SQL-ből is ismert substring függvényhez hasonlóan működik. Meg kell adni első paraméterként, hogy mely mezőben található szöveget dolgozzuk fel, majd hányadik karakter legyen a kezdő érték, és melyik legyen az utolsó. Utolsó karakter megadásához a HOSSZ függvénnyel kiszámítjuk a szöveges változó hosszát.

Megoldás:

Adjunk hozzá egy **New Delimited text file** réteget (Szöveges fájl). Itt a *Browse* gombbal hívjuk be a CSV fájlt, állítsuk be a helyes karakterkódolást. (Az excelből exportált CSV karakterkódolása windows-1250, ha konvertáltad valamely másik programmal, vedd számításba az UTF-8-at is). Elválasztó jelként válasszuk a *Custom delimiters*, ebből pedig a Semicolon (pontosvessző) tagolást. Ekkor az alsó box-ban már látható a beolvasandó fájl tagolása. Ha megtartottuk a fejléc oszlopot, a *First record has field names* legyen beikszelve. Ez az állomány nem tartalmaz koordinátákat. Ekkor a *Geometry definition*-ként a *No geometry*-t kell választani, mivel nem tartalmaz koordinátákat. A Watch file lehetőséggel a QGIS figyeli a textfile-ban időközben történt módosításokat.

Ezután az oké gombra kattintva megjelenik az attribútum tábla.

Ezután nyissuk meg a településeket tartalmazó Shapefile réteget. Nyissuk meg a település

rétegünk Properties ablakját, és válasszuk a **Joins** (Kapcsolások) fület. Ebben az ablakban alul a **zöld+** jelre kattintva adható hozzá attribútum tábla. A *Join layer* az importált CSV állományunk a *Join field* pedig az a mező, amely alapján hozzá akarjuk kapcsolni a települések rétegéhez. Ez legyen a település nevét tartalmazó oszlop. A *Target field* pedig az az oszlop a települések táblában, amelyhez hozzáfűzzük. Ez a TELNEV a településneveket tartalmazó oszlop alapján történhet. Ezen kívül kiválaszthatóak azok a mezők, amelyeket hozzá akarjuk kapcsolni (*Choose which fields are joined*). Ha ez inaktív, az összes mezőt kapcsoljuk. Megadható még a kapcsolt oszlopok előtagja is.

Ezután létrejött a kapcsolat (a Properties ablakon Apply, majd OK), ellenőrizhetjük az attribútum táblázatban.

Megjegyzés: Az előbb ismertetett folyamatot **geokódolás**nak nevezzük. Sajnos, egyik táblázatunk sem tartalmazott egyedi azonosítót (pl. KSH kódot) az egyes településeknél. Ilyenkor, ha eltérések adódnak a két állomány településnevei között (adódhat ez akár elgépelésből is), még ha ugyanarról a településről van szó is, nem jön létre a kapcsolat. Célszerű kapcsolatot létesíteni kódok, pl. KSH kód alapján. Ezek hiányában, ebben a feladatban a településneveket kellett használnom.