

GII-07.3 modulis: erdvinė statistika

1 užduotis: erdvinė autokoreliacija, vietos koeficientas ir MAUP

Atlikimo data: Šiam darbui atlikti reikia maždaug savaitės.
Užduoties įvertinimas: Maksimalus šios užduoties įvertinimas – 38 balai. Ji sudarys 6 % bendro kurso pažymio.

Aprašymas ir tikslai

Šioje užduotyje dėmesys telkiamas į erdvinės statistikos metodų taikymą atliekant įvairių nusikaltimų geografinių duomenų analizę Lietuvoje. Studentai naudos šiuos erdvinės statistikos metodus:

- 1) autokoreliacijos erdvėje skaičiavimas ir interpretavimas;
- 2) Lietuvoje įvykdomų nusikaltimų vietos koeficiento skaičiavimas, interpretavimas ir kartografavimas;
- 3) analizė pasitelkiant alternatyvias erdvines ribas, siekiant pamatyti kintamo ploto vieneto problemos poveikį.

Pateikimo reikalavimai

Atsakykite į visus klausimus atskirame dokumente („MS Word“). Būtinai įtraukite visus žemėlapius arba lenteles, sukurtas atsakant į užduoties klausimus. Visi vaizdai turi būti patalpinti į „MS Word“ dokumentą, visos lentelės turi būti kuriamos naudojant „Table“ (lentelės) funkcijas „MS Word“ aplinkoje. Visi skaičiai turi būti pateikiami suapvalinti iki trečio skaitmens po kablelio.

Pasiruošimas

Šiai užduočiai bus reikalingas šis duomenų failas: „Lithuania_2005.SHP“.

Šiame SHAPE faile yra toliau išdėstyti 2005 metais nustatyti kintamieji, kuriuos galima naudoti užduotims atlikti:

UNIQUE_ID	Unikalus kiekvienos savivaldybės identifikavimo numeris
COUNTY	Unikalus kiekvienos apskrities identifikavimo numeris
VARDAS	Savivaldybės pavadinimas
APSKRITIS	Apskrities pavadinimas
POPULATION	Savivaldybės gyventojų skaičius
CRIMES	Savivaldybės administracinėse ribose įvykdytų nusikaltimų skaičius
THEFT	Savivaldybės administracinėse ribose įvykdytų vagysčių skaičius
BNE	Įsilaužiamųjų į gyvenamąsias patalpas vagysčių skaičius savivaldybėje
AUTO	Pavogtų automobilių skaičius
JUVDEL	Paauglių padarytų nusikaltimų skaičius
CRIMRATE	Nusikalstamumo lygis savivaldybėje, skaičiuojant 10 000 gyventojų
THFTRATE	Vagysčių lygis savivaldybėje, skaičiuojant 10 000 gyventojų
BNERATE	Įsilaužiamųjų vagysčių į gyvenamąsias patalpas lygis savivaldybėje, skaičiuojant 10 000 gyventojų
AUTORATE	Automobilių vagysčių lygis savivaldybėje, skaičiuojant 10 000 gyventojų
JDELRATE	Paauglių nusikalstamumo lygis savivaldybėje, skaičiuojant 10 000 gyventojų

Visi šie duomenys gauti iš Lietuvos statistikos departamento: <http://www.stat.gov.lt/en/>

A dalis: Autokoreliacija erdvėje

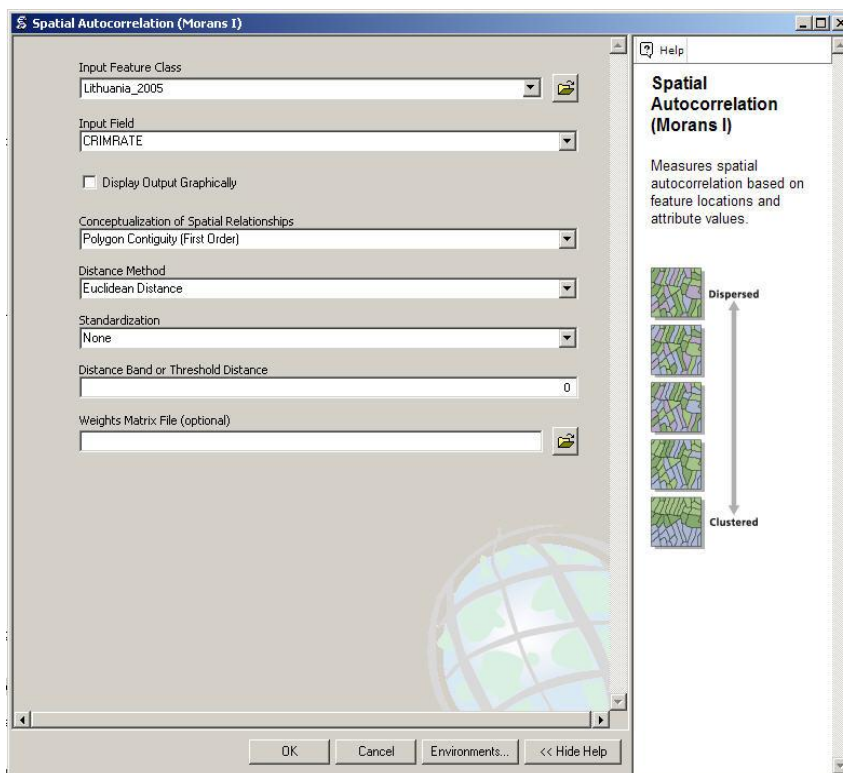
Šioje užduotyje mes naudosime Lietuvos savivaldybių ribas ir tirsime nusikalstamos veiklos erdvinio išsidėstymo ypatumus (jei tokių bus). Apskaičiuosime kelių kintamųjų autokoreliaciją erdvėje, naudodami skirtingus erdvių kaimynų nusakymo būdus.

1. Atidarykite „ArcMap“ ir pridėkite failą „Lithuania_2005.SHP“ į savo ekraną.
2. Dešiniuoju pelės klavišu spustelėkite slukšnio pavadinimą ir pasirinkite **Properties** (savybės). Tada spustelėkite kortelę **Symbolology** (simbolika).
3. Šio modulio užduotyse visuomet norėsite rodyti **Quantities** (kiekius) naudodami **Graduated colors** (graduotas spalvas).
4. **Field Value** (lauko vertei) parinkite CRIMRATE.
5. Taikykite numatytąją klasifikaciją (joje turėtų būti 5 natūralieji lūžiai), pasirinkite tinkamą spalvų paletę (naudinga naudoti spalvų paletę, kurioje nuo žalios (mažas nusikalstamumas) pereinama prie raudonos (didelis nusikalstamumas)) ir pavaizduokite kintamąjį.

1 klausimas: ar matote kokią nors aiškia verčių išsidėstymo tendenciją žemėlapyje? (2 balai)

Išvaizda gali apgauti, tad visuomet geriau apskaičiuoti statistiką ir patvirtinti arba atmesti išvadas, kurias padarome vizualaus stebėjimo metu. Žemėlapyje matoma tendencija gali būti stebima viso labo dėl naudojamos simbolikos.

1. Atidarykite „ArcToolbox“.
2. Spustelėkite **Spatial Statistics Tools** (erdvinės statistikos įrankiai) → **Analyzing Patterns** (erdvinio išsidėstymo analizė) → **Spatial Autocorrelation** (Morano I) (autokoreliacija erdvėje).
3. Atsivers toliau pateikiamas langas.



4. **Input feature class** (įvesties elementų klasė) yra „Lithuania_2005.SHP“; **Input Field** (įvesties laukas) yra CRIMRATE; o **Conceptualization of Spatial Relationships** (erdvinių ryšių konceptualizacija) – **Polygon Contiguity (First Order)** (poligono gretimumas, pirmasis lygis). Visas kitas parinktį palikite numatytąsias. Spustelėkite **OK** (gerai). Lange **Spatial Autocorrelation** (autokoreliacija erdvėje, Morano *I*) nuimkite varnelę nuo **Close this dialog when completed successfully** (sėkmingai užbaigus uždaryti šį dialogo langą) bei peržvelkite rezultatus (jei to nepadarytė anksčiau).

2 klausimas: kokia Morano *I* vertė ir ar ji reikšminga? Interpretuokite šią statistiką. Pastaba: jei *Z* rodiklis yra didesnis nei 1,28, vadinasi reikšmingumo lygis yra 10 %, o jei 1,645 – reikšmingumo lygis siekia 5 % (3 balai).

Dabar pakartokite šiuos veiksmus su kitais keturiais nusikalstamumo rodikliais.

3 klausimas: kokia Lietuvos kitų nusikalstamumo rodiklių Morano *I* vertė? Parenkite lentelę, kurioje būtų visos Morano *I* statistikos ir atitinkami jos *Z* rodikliai bei interpretuokite rezultatus (5 balai).

Atliekant ankstesnius autokoreliacijos erdvėje skaičiavimus, pirmojo lygio gretimumas buvo naudojamas kaip pagrindinis modulyje paminėtų kaimynų apibrėžimo metodas. Kitas metodas – atvirkščio atstumo. Tokia erdvinių svertų matrica kiekvieną erdvinį vienetą susies su visais kitais erdviniais vienetais, o ne tik su jo pirmojo lygio kaimynais. Tačiau, kadangi naudojamas atvirkščio atstumo metodas, artimesni erdviniai vienetai turės daugiau įtakos nei nutolę labiau.

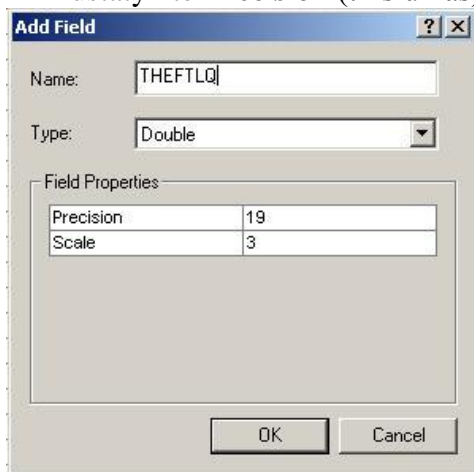
Perskaičiuokite Morano *I* statistiką, taikydami **Inverse Distance** (atvirkščio atstumo) metodą ir įrašydami jį į skiltį **Conceptualization of Spatial Relationships** (erdvinių ryšių konceptualizacija).

4 klausimas: ar, matuojant autokoreliaciją erdvėje, skirtingų erdvinių svertų matricos naudojimas (atvirkštieji atstumai prieš pirmojo lygio gretimumą) duoda kitokius rezultatus? Neįtraukite kitos lentelės: tik pateikite bendruosius pastebėjimus (2 balai).

Dabar jau turėjo būti akivaizdi esama autokoreliacijos erdvėje tendencija (jei ji buvo). Tačiau tai nereiškia, kad nėra nusikalstamumo klasterių: tiesiog nėra nusikalstamumo lygio klasterių. Toliau autokoreliacijai erdvėje analizuoti pradėsime naudoti vietos koeficientus.

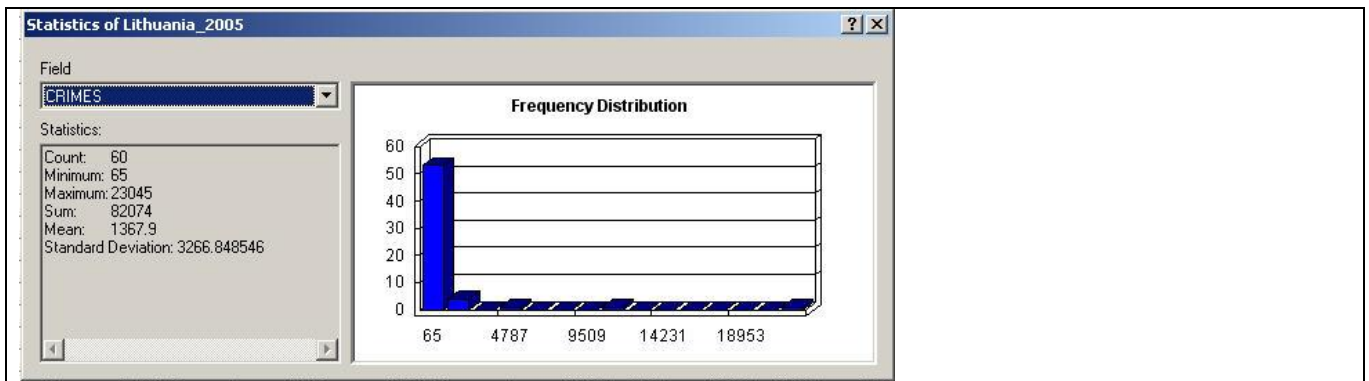
B dalis: vietos koeficientas

1. Dešiniuuoju pelės klavišu spustelėkite „Lithuania_2005“ ir atidarykite atributų lentelę.
2. Iš mygtuko **Options** (parinktys) spustelėkite **Add Field...** (įtraukti lauką).
3. Skiltyje **Name** (pavadinimas) įveskite THEFTLQ; skiltyje **Type** (tipas) įveskite **Double**; nustatykite **Precision** (tikslumas) į 19, o **Scale** (mastelis) – į 3.



Dabar apskaičiuosime keturių skirtingų specifinių nusikalstamumo lygių vietos koeficientus: bendrajam nusikalstamumo lygiui vietos koeficiento apskaičiuoti negalima. Šiam skaičiavimui paaiškinti reikia išstudijuoti atitinkamą kurso teorinę medžiagą. Atminkite, kad vietos koeficientas matuoja konkretaus nusikaltimo procentinį lygį savivaldybėje lyginant su to paties nusikaltimo lygiu Lietuvoje. Taigi, bendras vieno nusikaltimo vietos koeficiento vardiklis yra toks pat visoms savivaldybėms – jis kinta tik skaičiuojant kitą vietos koeficientą. Mūsų pirmasis uždavinys – nustatyti vagysčių, įsilaužiamųjų į gyvenamąsias patalpas vagysčių, automobilių vagysčių ir paauglių nusikalstamumo procentinį lygį lyginant su nusikalstamumo lygiu visoje Lietuvoje.

1. Atidarykite atributų lentelę „Lithuania_2005“.
2. Dešiniuuoju pelės klavišu spustelėkite CRIMES ir tada spustelėkite **Statistics...** (statistika).
3. Gausite toliau iliustruojamą informaciją:



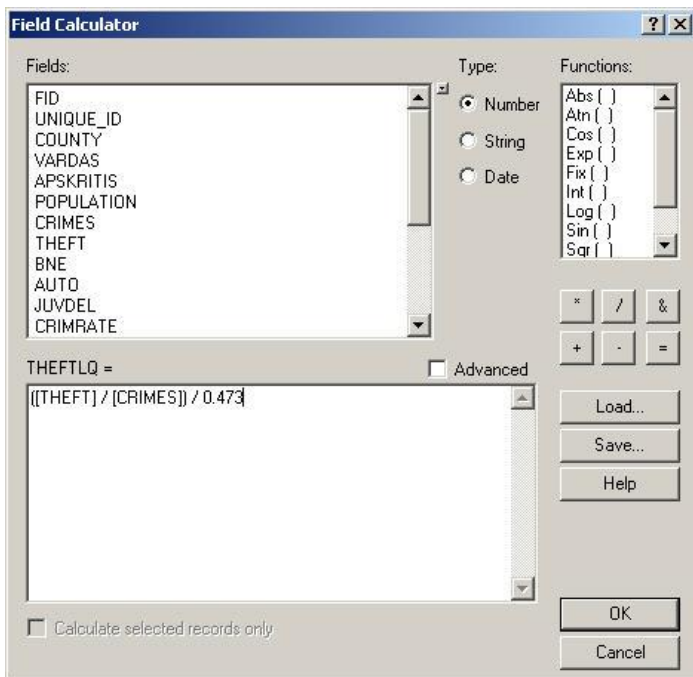
4. Kaip matote, iš viso 2005 metais įvykdyti 82 074 nusikaltimai. Tai – nedidelis nusikaltimų skaičius valstybei (arba bet kokiai teritorijai), kurioje gyvena 3,5 mln. žmonių.
5. Iš šio lango galima pakeisti lauką (**Field**) ir gauti keturių specifinių nusikaltimų kategorijų statistiką.
6. Gaukite sumas šioms nusikaltimų rūšims: CRIMES, THEFT, BNE, AUTO ir JUVDEL.

5 klausimas: kokios yra visų penkių nusikaltimų kategorijų sumos ir keturių specifinių nusikaltimų kategorijų procentiniai lygiai (THEFT, BNE, AUTO ir JUVDEL)? (2 balai).

6 klausimas: ką atitinka bendras nusikalstamumo procentinis lygis, kuris nesiekia vieneto (arba 100 % – priklausomai nuo jūsų naudotų procentų stiliaus)? (1 balas).

Dabar apskaičiuokime vietos koeficientus.

1. Dešiniuoju pelės klavišu spustelėkite lauko pavadinimą: THEFTLQ
2. Spustelėkite **Field Calculator...** (lauko skaičiuotuvas).
3. Ignoruokite perspėjimą dėl nebuvimo redagavimo seanse, kadangi tai – naujas kintamasis, tad, padarius klaidą, jį galima nesunkiai pašalinti.
4. Laukų skaičiuotuvo lange įrašykite atitinkamą formulę.
5. Ji turi atrodyti panaši į iliustraciją:



6. Pakartokite šiuos veiksmus ir sukurkite: BNELQ, AUTOLQ ir JUVDELLQ.

Naudodamiesi pirmiau aprašytais nuostatomis (5 natūraliaisiais lūžiais ir žalios-raudonos spalvų palete), sukurkite du žemėlapius: vienas jų bus skirtas vagysčių rodikliui, kitas – vagysčių vietos koeficientui.

7 klausimas: aprašykite ir įtraukite į darbo ataskaitą šių dviejų žemėlapių ekranų nuotraukas. Ar tarp jų yra pastebimų skirtumų? Interpretuokite šiuos skirtumus (7 balai).

Nepateikite jų šioje užduotyje: tik peržvelkite kitų rodiklių ir atitinkamų vietos koeficientų žemėlapius bei pamatykite, kaip kinta jūsų požiūris į nusikaltimus naudojant vietos koeficientą. Savivaldybės, kuriose nusikalstamumo lygis yra žemas, vis tiek gali turėti problemų su tam tikros rūšies nusikaltimais.

Dabar, naudodamiesi pirmiau išdėstytais metodais, apskaičiuokite šių keturių vietos koeficientų autokoreliaciją erdvėje. Visais atvejais taikykite pirmojo lygio gretimumą.

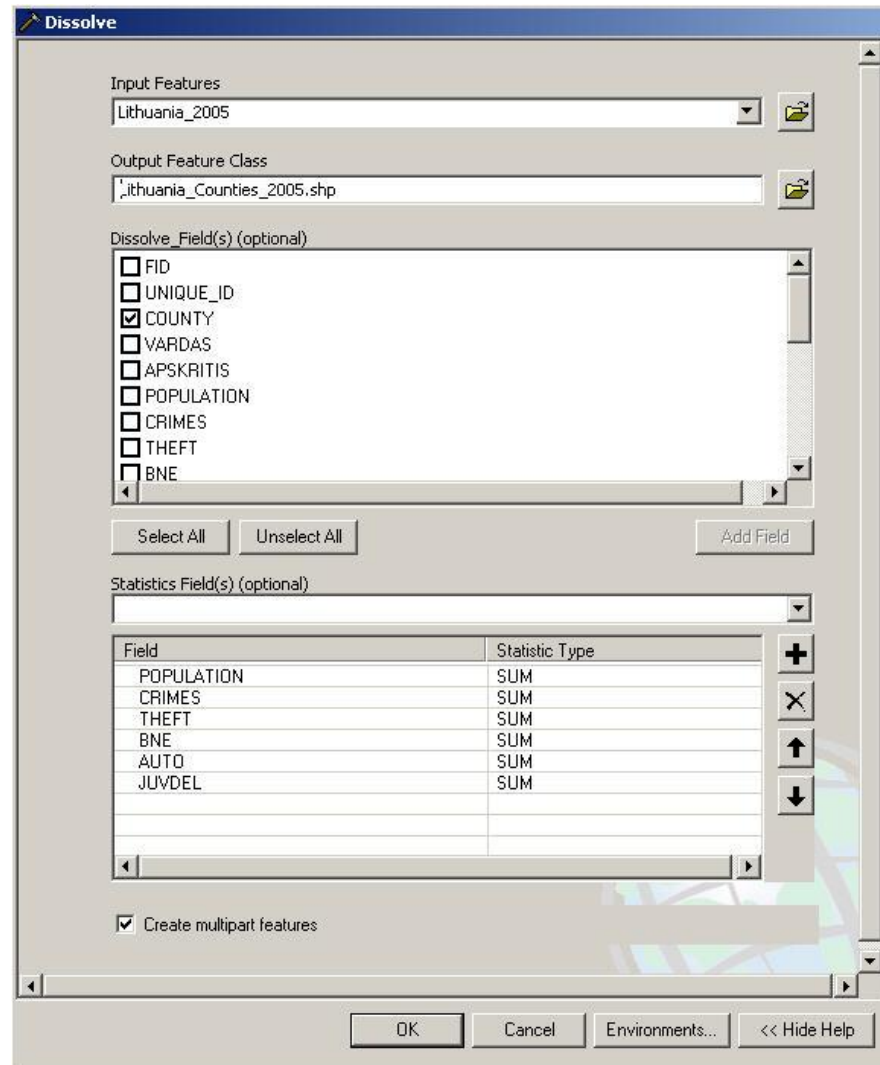
8 klausimas: kokia šių vietos koeficientų Morano *I* statistika ir *Z* rodikliai? Pateikite juos lentelėje. Interpretuokite šių statistikų prasmę (8 balai).

C dalis: kintamo plotinio vieneto problema ir autokoreliacija erdvėje

Iki šiol visa mūsų analizė apsiribojo Lietuvos savivaldybėmis. Tačiau, kaip užsiminta šio modulio teorinėje dalyje, tai gali būti tik naudojamos analizės erdvinių vienetų rezultatas. Tikėtina, kad apskrityse ne tik aukštesnis autokoreliacijos erdvėje lygis, bet ir autokoreliacija erdvėje yra reikšmingesnė. Dabar sukursime apskričių sluoksnį ir pamatysime, kokie analizėje įvyksta pokyčiai.

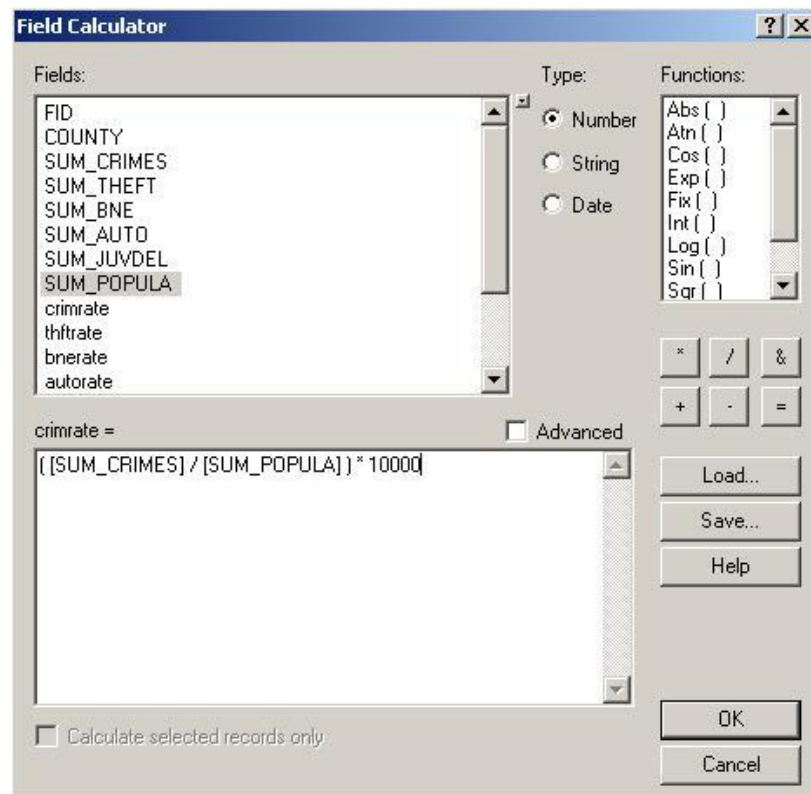
1. Atidarykite „ArcToolbox“.

2. Eikite į: **Data Management Tools** (duomenų valdymo įrankiai) → **Generalization** (apibendrinimas) → **Dissolve** (sulieti).
3. Atsiveria suliejimo funkcijos langas.
4. **Input Feature** (įvesties elementas) yra „Lithuania_2005.SHP“.
5. **Output Feature Class** (išvesties elementų klasės) pavadinimas: „Lithuania_Counties_2005.SHP“.
6. **Dissolve Field** (suliejimo laukas) yra: COUNTY
7. Parinktini **Statistical Fields** (statistikos laukai) yra: POPULATION, CRIMES, THEFT, BNE, AUTO ir JUVDEL.
8. Visų jų **Statistic Type** (statistikos tipas) yra: SUM



Naudojant šiuos naujus kintamuosius (patikrinkite naujus pavadinimus savo atributų lentelėje) ir siekiant, kad šis naujas duomenų rinkinys galėtų būti ekvivalentiškas savivaldybės duomenų rinkiniui, reikia apskaičiuoti apskričių nusikalstamumo lygius ir vietos koeficientus. Naujų laukų kūrimo procedūra buvo apibūdinta pirmiau, kartu su vietos koeficientų kūrimu. Tačiau nusikaltimų lygiai nepateikiami. Visi lygiai turi būti apskaičiuoti 10 000 gyventojų, kaip parodyta tolesnėje iliustracijoje. Kaip paminėta pirmiau, visi naujieji laukai turi būti šio **Type** (tipo): **Double**, turėti **Precision** (tikslumo lygi) 19 ir **Scale** (mastelį) 3.

Būtinai išsaugokite šį SHAPE failą, kadangi jo reikės atliekant kitą užduotį.



Dabar, turėdami naują kintamųjų rinkinį, pavaizduokite juos ir ieškokite išsidėstymo erdvėje tendencijų. Turėtumėte pastebėti akivaizdžius skirtumus tarp nusikalstamumo lygių ir atitinkamų vietos koeficientų.

Naudojant šiuos naujus atributų laukus, reikia kiekvienam rodikliui apskaičiuoti autokoreliaciją erdvėje. Erdviniams ryšiams nustatyti naudokitės pirmojo lygio gretimumu. Ignoruokite erdvinių vienetų skaičiaus perspėjimą.

9 klausimas: kokia kiekvieno iš šių apskričių nusikalstamumo rodiklių Morano I statistika ir Z rodikliai? Įtraukite juos į lentelę ir interpretuokite (5 balai).

10 klausimas: kokį poveikį kintamo plotino vieneto problema turi ką tik atliktai analizei? (3 balai).