

**UUSI UNISTUSI
EESTI MURDE- JA FOLKLOORIALADE
PIIRITLEMISE TEEMAL**

**Arvo Krikmann
EKM FO**

Tartu 2014

Kasutatud andmestike sisu, laad ja päritolu

Alljärgnev püüab jätkata arutelu keele- ja folkloorialade piiritlemise üle empiirilise statistilise vaatluse põhjal, mille aluseks on arvandmestikud kolmest üsna suurest sisumassiivist pärinevate üksuste geograafilise leviku kohta Eestis:

- 1) murdesõnavara levikuandmed “Väikeses murdesõnastikus” (VMS I–II);
- 2) mõistatuste levikuandmed eesti mõistatuste andmebaasis (Krikmann ja Krikmann 2012) aadressil <http://www.folklore.ee/moistatused/>;
- 3) eesti vanasõnade geograafilise leviku register “Eesti vanasõnade” IV köites (EV IV: 267–354); andmed veidi redigeeritud ja uuendatud.

Statistikaisse on võetud tüpoloogilised üksused, mida on registreeritud kahest või enamast kihelkonnast.

Geoüksuste vahelise seosetiheduse statistilisest hindamisest

Nn inimgeograafias on kasutatud ammu väga mitmekesisest statistilist arsenalist, viimase pool-sajandi jooksul ka rafineeritumaid mitmemõõtmelise analüüsi meetodeid, nagu faktoranalüüs, klasteranalüüs, nn mitmemõõtmeline skaleerimine jt. Humanitaarias on need meetodid laiemat rakendust leidnud eelkõige murdeuurimises; sh on olemas ulatuslik kirjandus murde-materjali areaalse analüüsi kvantitatiivsete meetodite kohta. Selles kirjutises ei püüta vähimalgi määral refereerida inimgeograafias kasutatud kvantitatiivseid meetodeid ega analüüsida neid teoreetilises plaanis, kuid kirjutise lõpupoolel on tehtud arglikke katseid faktor- ja klasteranalüüsi võtete tegelikuks rakendamiseks. Põhiliselt on aga siin, nagu ka mu eelmistes töedes, nähtuste seosetiheduse hinnangutena kasutatud kaht elementaarset parameetrit:

- 1) põikeid lineaarsete vm korrelatsiooniväljade normfunktsioonist (nn residuaale);
- 2) Zbigniew Pawłowski (1967: 38) nn kolligatsioonikoefitsiente e λ -koefitsiente, mis põhinevad tõenäosuste korrutamise reeglil ning leitakse valemist

$$\lambda = \frac{p(A_1, A_2)}{p(A_1) \cdot p(A_2)}$$

kus $p(A_1)$ ja $p(A_2)$ on mingite empiiriliste nähtuste summaarsed esinemissagedused ja $p(A_1, A_2)$ nende koosinemise sagedus; $\lambda = 1$ sõltumatute sündmuste korral, $\lambda > 1$ tähendab positiivset sõltuvust, $\lambda < 1$ negatiivset.

Käesolevas vaatluses ongi kihelkondade sarnasust või seosetihedust kolmes nimetatud materjalis hinnatud λ -koefitsientide põhjal. Kuna meid huvitavad just kihelkondade seosed teiste kihelkondadega, on maatrikstabelite diagonaalväärtused (s.o kihelkondade “ühisosad iseendaga”) asendatud nullidega. Koefitsiendid on edasi standardiseeritud niiviisi, et 1. ringi koefitsiendid on võetud 2. arvutusringi lähteandmeteks ning sooritatud vajalik arv iteratsioone (siin konkreetselt 20), kuni tabeli rea- ja veerusummad (ning keskmised) piisava täpsusastmeni samastuvad.

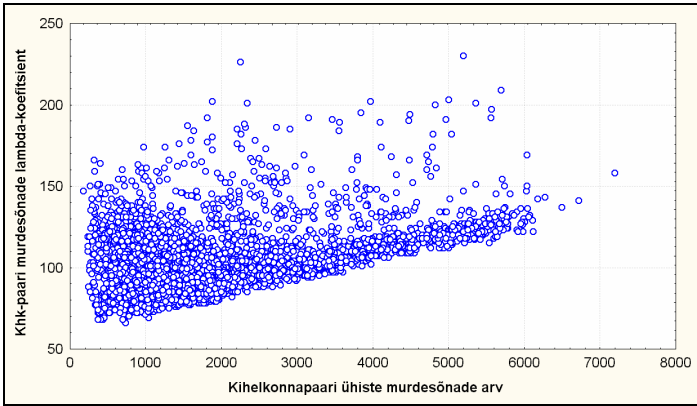
Andmete hoidmiseks ja töötlemiseks on kasutatud programmi Statistica 5.5 vahendeid. Enamik kartogramme on tekitatud “Eesti mõistatuste andmebaasi” (Krikmann ja Krikmann 2012) juurde kuuluva allmooduli, nn kaardimasina abil (<http://www.folklore.ee/moistatused/kaardimasin/index.php>). Kuna see on võimeline töötleva ainult täisarvulisi andmeid, siis on λ -koefitsiendid jm kaardistatavad sisendid tehnilistel põhjustel korrutatud või jagatud 10, 100 või 1000-ga ning ümardatud täisarvudeks.

Nähtuste sarnasuse ~ kokkukuuluvuse ~ seosetiheduse lihtsamat järku hinnangutena on λ -koefitsientidest sagedamini kasutatud muid, praegustesse statistikaprogrammidesse sisse kodeeritud mõõdikuid, nt lineaarseid korrelatsioonikordajaid ja klasteranalüüsi aluseks olevaid nn eukleidilisi kaugusi.

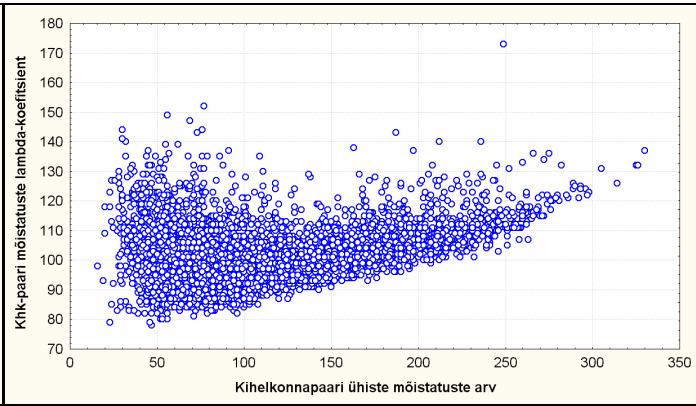
Seosemõõdikute vahetõrge geooksuste vahelise ühisosa suurusega

Lk 4 paiknevad 9 joonist kujutavad geooksuste vahelise “tүpoloogilise” ühisosa suuruse ja kolme seosemõõdiku *in spe* suhteid meie kolmes vaatlusmaterjalis. Kui ideaalis peaksid seosetugevused olema ühisosade suurusest sõltumatud, siis näeme, et tegelikult pole seda ükski neist. Standardiseeritud lambdade väljade trendid on tõusev-koonduvad, korrelatsioonikordajate väljad ülespoole või horisontaalselt koonduvad ja eukleidiliste kauguste väljad üldiselt allapoole koonduvad.

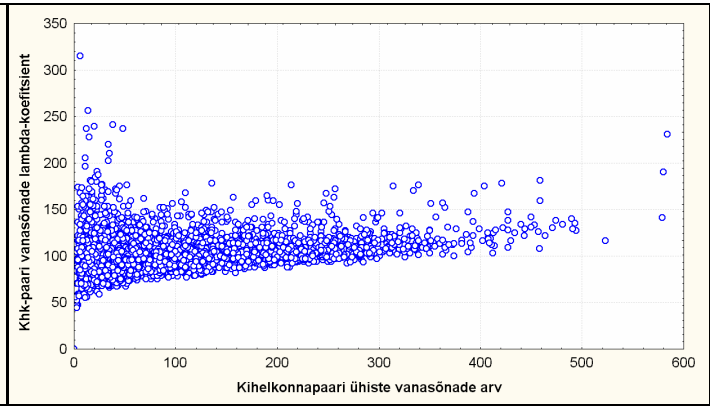
Ma ei söanda siinkohal hakata uuesti arutama küsimust sellest, kas ja millal seosekoefitsientide jm mõõdikute suhtes tuleb rakendada n-ö ühtlaseks tempereerimise tehnikat, st püüda tingimata teisendada kõverjoonelise trendiga regressiooniväljad alul lineaarseteks, seejärel normeerida tõusva või alaneva trendiga lineaarsed väljad “horisontaalseteks”, seejärel normeerida ja standardiseerida ka paisuva või koonduva tendentsiga residuaalväljad jne (vrd sellest ka Krikmann 1979: 34–36; Krikmann ja Pajusalu 2000: 147–150).



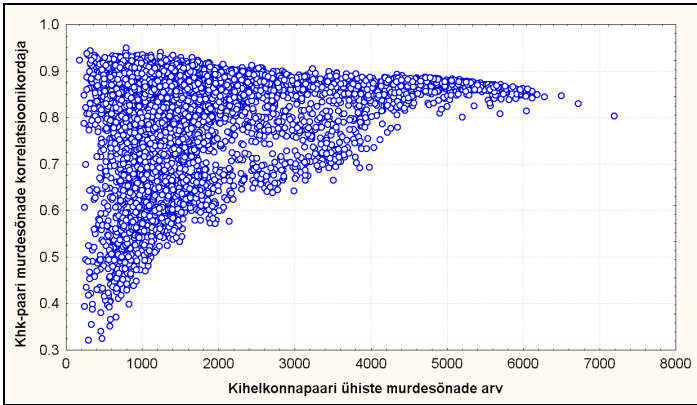
Murdesõnad: λ -koefitsiendid ($\times 100$)



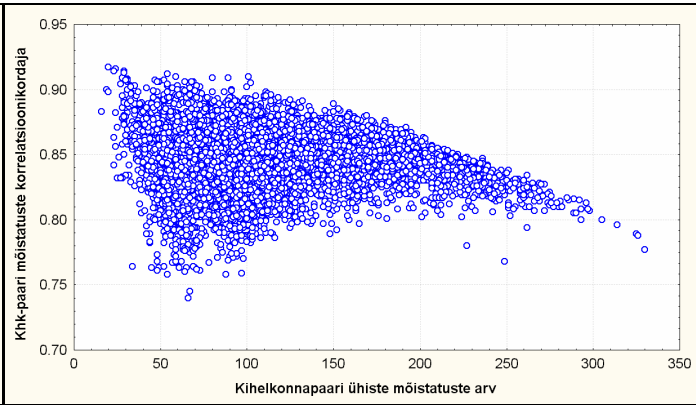
Mõistatused: λ -koefitsiendid ($\times 100$)



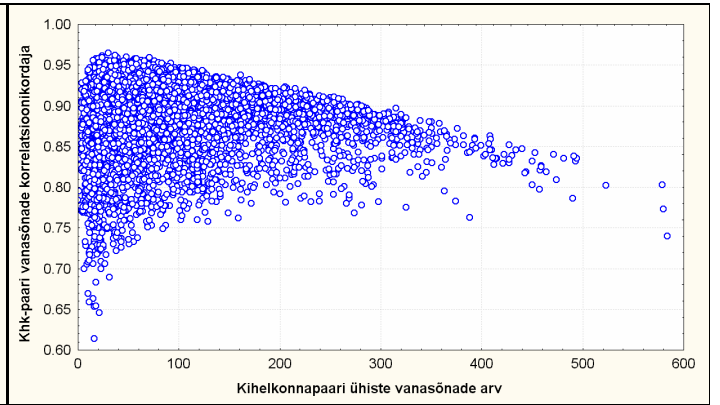
Vanasõnad: λ -koefitsiendid ($\times 100$)



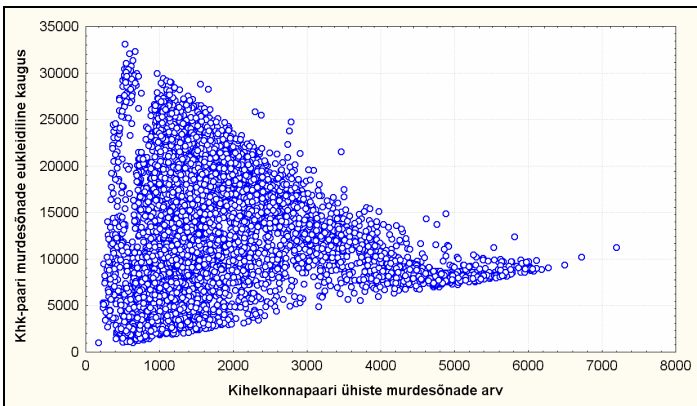
Murdesõnad: korrelatsioonikordajad



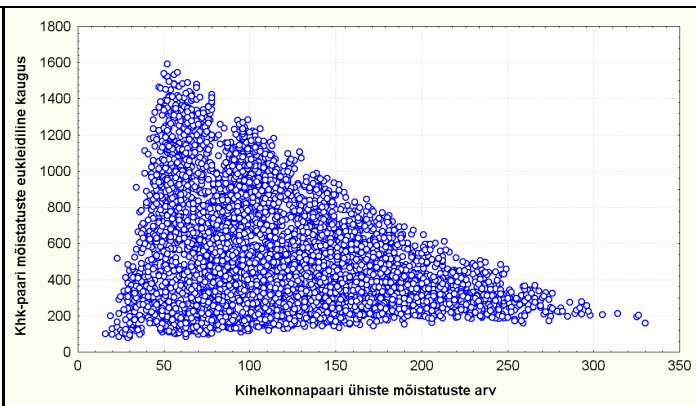
Mõistatused: korrelatsioonikordajad



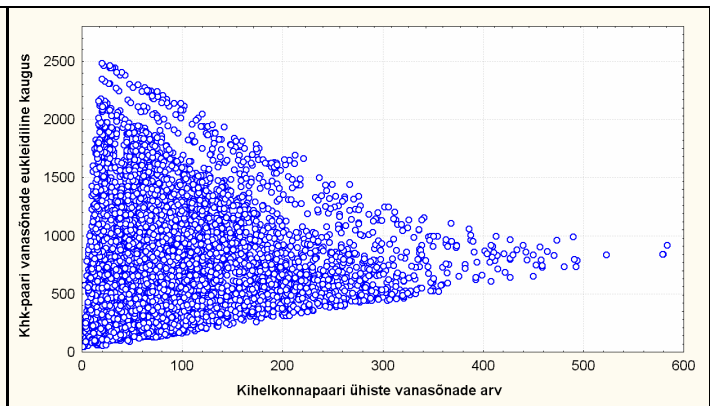
Vanasõnad: korrelatsioonikordajad



Murdesõnad: eukleidilised kaugused



Mõistatused: eukleidilised kaugused



Vanasõnad: eukleidilised kaugused

Samuti ei ole selle teksti kirjutajal piisavalt teadmisi ega praktilisi kogemusi otsustamiseks, kui sobivad on siinset laadi andmestike korral sisu- ja geouksuste vaheliste seoste mõõdikute rolli täitma nt mainitud λ -koefitsiendid, korrelatsioonikordajad või eukleidilised kaugused. Provi-soorsed katsetused on jätnud küll mulje, et λ -koefitsientide korral on tekkiva “intuitiivse müra” nivoo vähim, eukleidiliste kauguste kasutamisel – vaatamata nende arvutusvalemi äärmisele lihtsusele – kõrgeim ja korrelatsioonikordajate puhul kuskil nende kahe vahepeal.

Esimene kultuurigeograafiline triviaaltõik: mida lähemal, seda sarnasemad

Järgnevad 9 joonist (vt lk 6) näitavad selle tõiga ilmset paikapidavust meie kolmes vaatlusmaterjalis: kui sõltumatuks muutujaks võtta kihelkonnapaaride (keskmete) vaheline kaugus meil kasutatud aluskaardi tingühikutes ja sõltuvaks muutujaks nende paaride seosetihedust iseloomustavad λ -koefitsiendid, siis saame selgelt alaneva nõgusa kujuga, korrelatsioonikordajate puhul alanev-hajuva lineaarse kujuga ning eukleidiliste kauguste korral hajusa puhmakujuliselt tõusva lineaarse trendiga regressiooniväljad.

Sisu- ja geouksuste jaotussuhetest

Kui oleme oma vaatlusest eemaldanud ainult ühesainsas geouksuses esinevad sisuüksused (murdesõnad, mõistatused, vanasõnad), siis kujunevad nende summaarsed hulgad sellisteks:

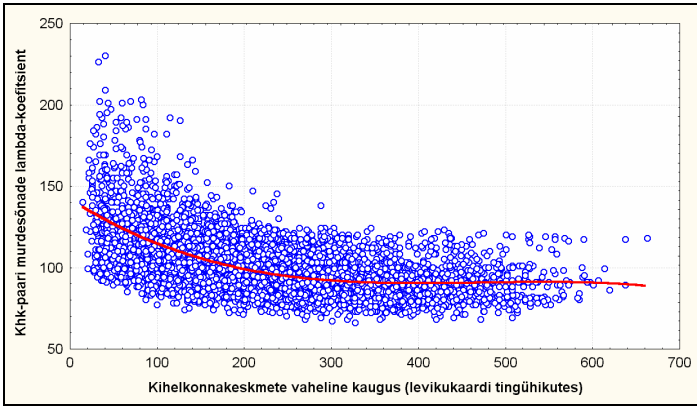
eri **murdesõnu** VMS-i andmetel **41777**;

mõistatustüüpe EM-baasi andmetel **1416**;

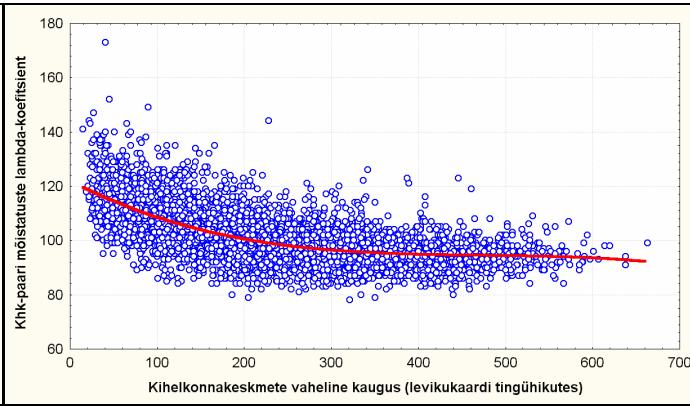
vanasõnatüüpe EV IV andmetel **4911**.

Geouksusteks on **105 Eesti kihelkonda**, sh kihelkonna õigustes esinev Setumaa. Vaatlusest on välja jäetud Vormsi, kust on kogutud vaid kaduvväike hulk eestikeelset murde- ja folklooriainest, samuti keelesaared (Lutsi, Leivu, Kraasna); vanasõnade andmestikust on eemaldatud ka paariteistkümmet kirjapanekuga esindatud ning seega statistiliselt töövõimetus Lihula ja Kirbla. Murdesõnade jagunemisest kesk-, ida- ja rannikumurde vahel mõnedes Kirde- ja Ida-Eesti kihelkondades pole hoolitud (vrd sellest ka Krikmann ja Pajusalu 2000: 135–136).

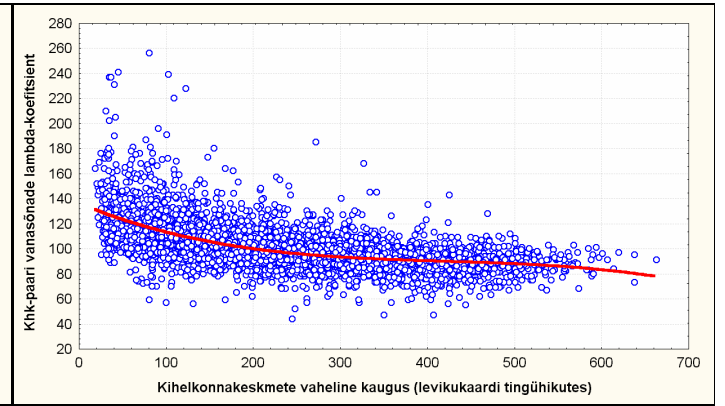
Keele- ja folklooriarhiividesse talletatud sisuainese sagedussuhted geouksustega on veidral viisil paradoksaalsed.



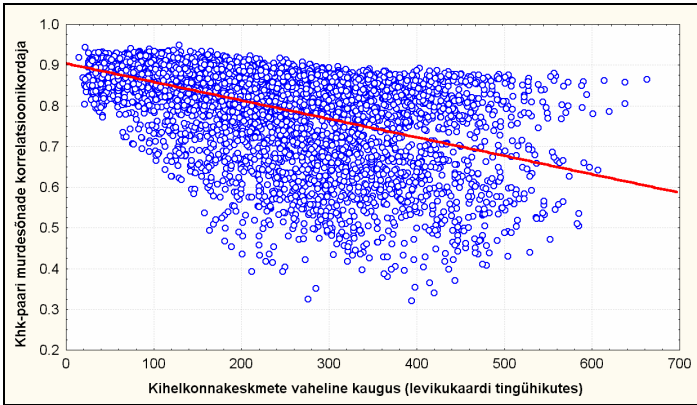
Murdesõnad: λ -koefitsiendid ($\times 100$)



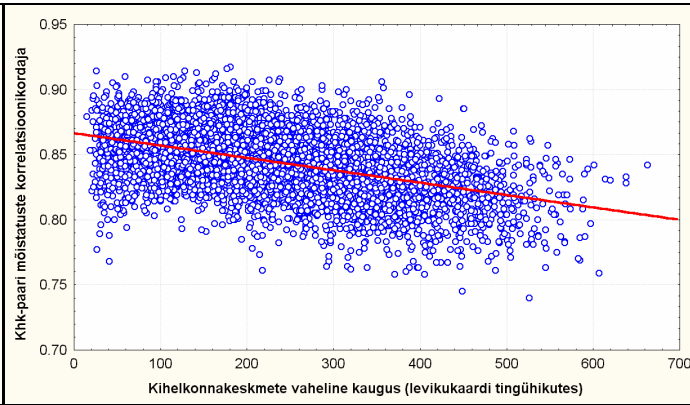
Mõistatud: λ -koefitsiendid ($\times 100$)



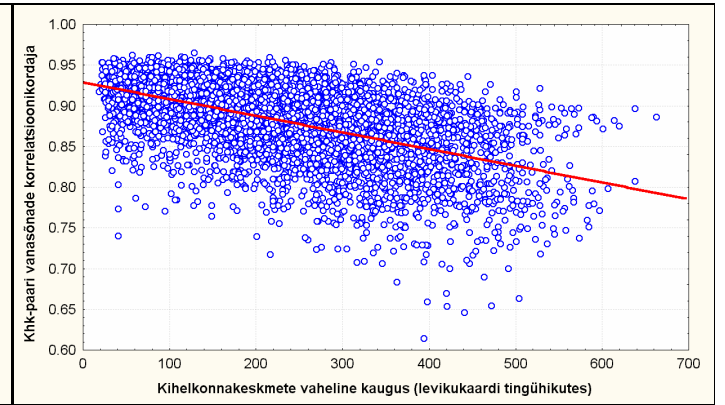
Vanasõnad: λ -koefitsiendid ($\times 100$)



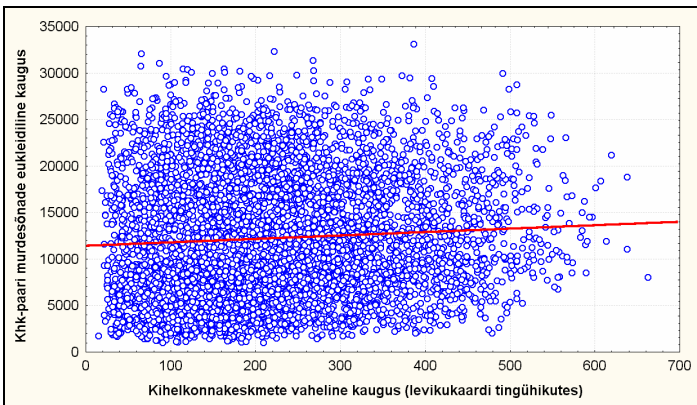
Murdesõnad: korrelatsioonikordajad



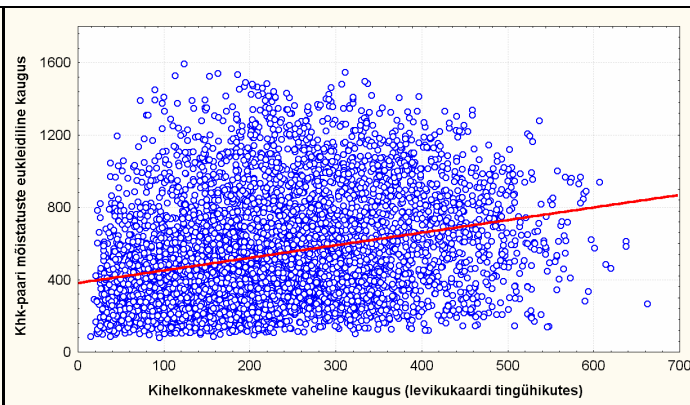
Mõistatud: korrelatsioonikordajad



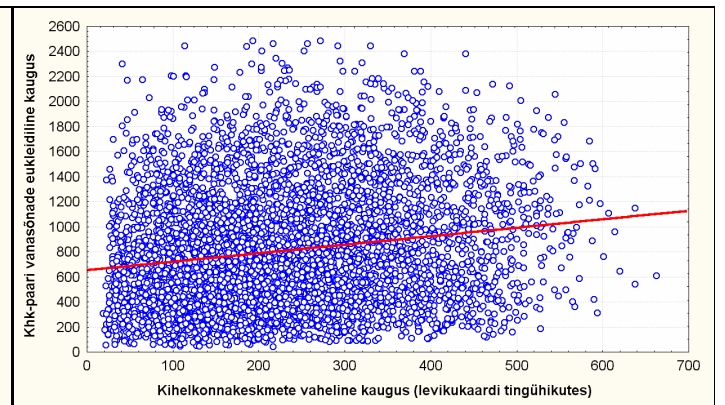
Vanasõnad: korrelatsioonikordajad



Murdesõnad: eukleidilised kaugused



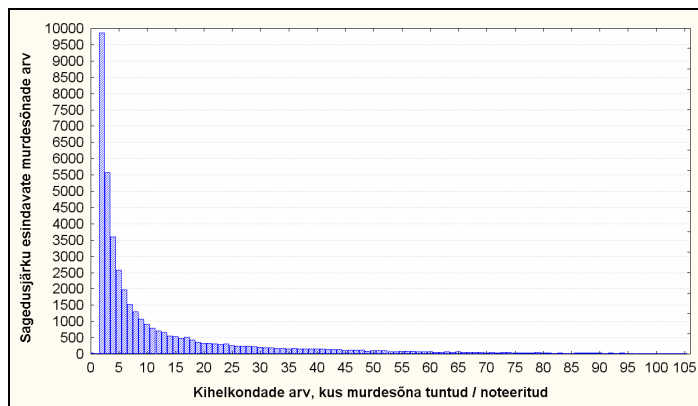
Mõistatud: eukleidilised kaugused



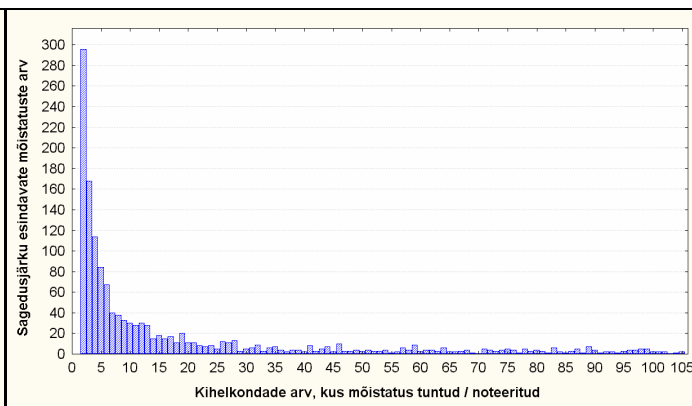
Vanasõnad: eukleidilised kaugused

Sisuüksuste jagunemisel nii tekstilise kui ka geograafilise produktiivsuse plaanis näib lasuvat universaalne “Zipfi needus”: nende sagedusjaotus mõlemas plaanis on drastiliselt ebahütlane, kusjuures kõige enam selgub ikka olevat tüpoloogiliselt või geograafiliselt “singulaarseid” üksusi, millel polegi seoseid teistega ning mida nt folklooriandmestikus justkui ei tohiks üldse olla, kuna folkloor on definitsioonikohaselt traditsioon, ja traditsiooni tuvastamiseks peaks meil olema dokumenteeritud andmeid vastavate sisuüksuste kordumise kohta. Teisalt on taoline improvisatoorsete “pürgijate” pilv kindlasti traditsiooni elujõu ja kestlikkuse sümptomiks ja garantiks. Meie vaatluse andmeid kajastab alloleva joonise ülariida.

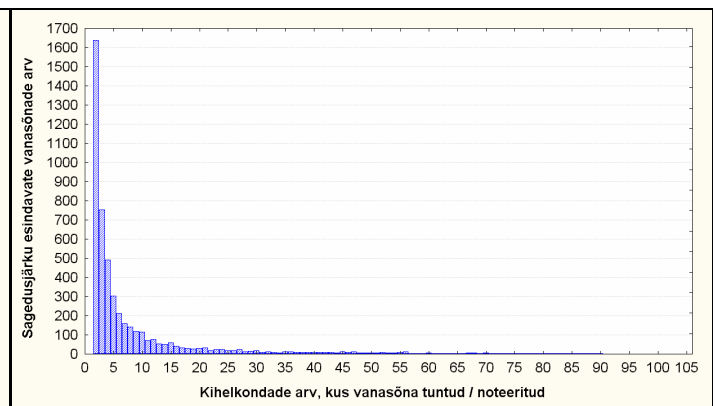
Logaritmilisel skaalal (vt joonise alarida) tavatsevad zipfilike sagedusjaotuste regressiooniväljad võtta lineaarse ilme; lineaarsete normeeringute ebatäpsused meie kolme andmestiku logaritmilistel graafikutel võivad osalt johtuda ka sellest, et histogrammidelt puuduvad ühesainsas kihelkonnas esindatud sõnade ja folklooriüksuste tohutud hulgad.



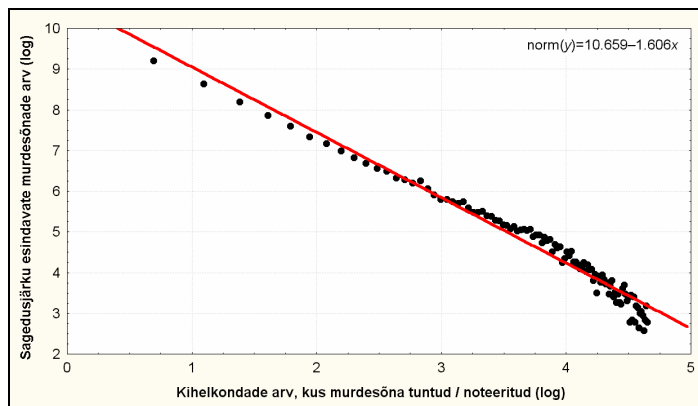
Murdesõnad



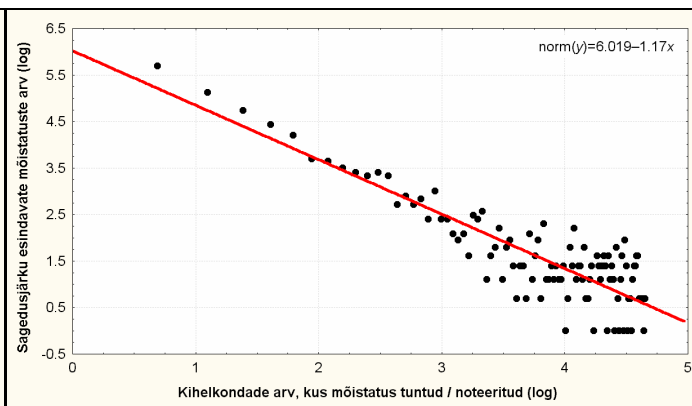
Mõistatused



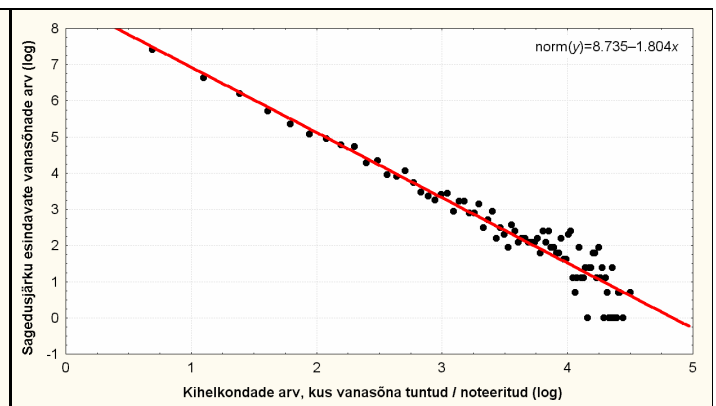
Vanasõnad



Murdesõnad (logaritmilisel skaalal)

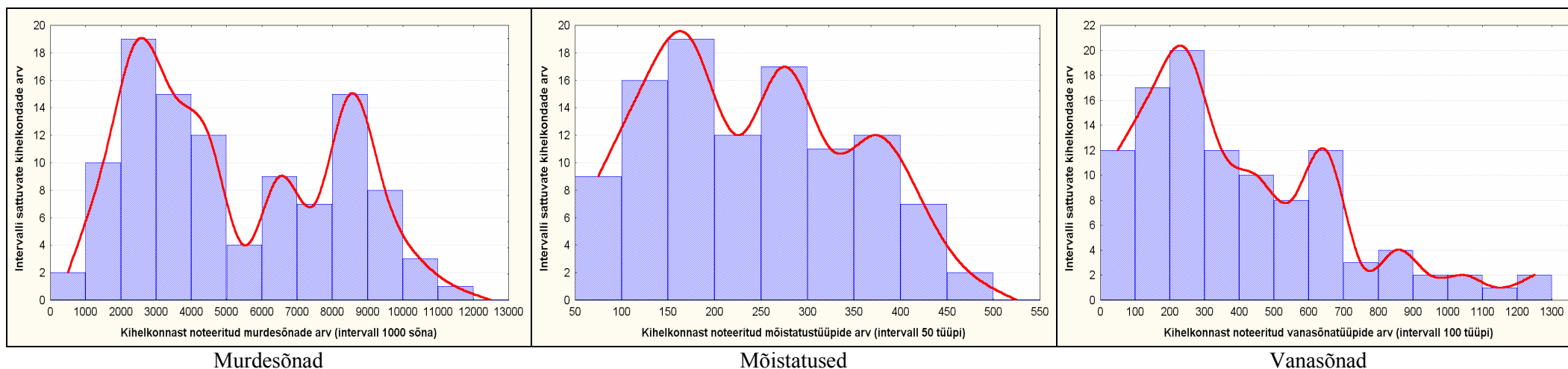


Mõistatused (logaritmilisel skaalal)



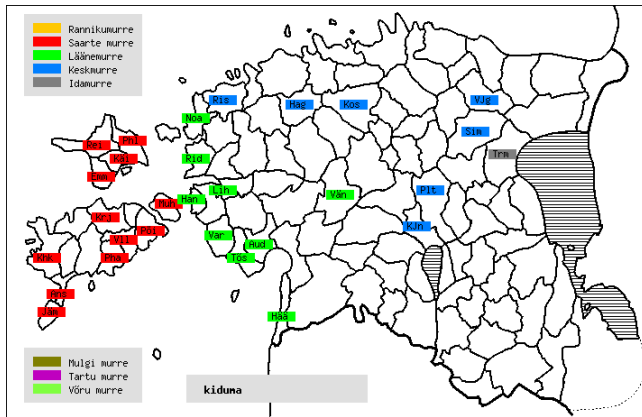
Vanasõnad (logaritmilisel skaalal)

Geoüksuste jagunemine neist registreeritud sisuüksuste hulga lõikes on hoopis teistsugune. Allolevad histogrammid meenutavad vasakule kallutatud raskuskeskmega normaaljaotust või mitme normaaljaotuse ühendit. Ka illustreerivad need graafikud kindlasti tõika, et murde- või keelearhiivist saadud sagedusandmetes on paratamatult pealistikku teave teatud sisuüksuste tegeliku tuntuse / mittetuntuse kohta teatud aladel kui ka teave neil aladel toimunud kogumistöö intensiivsuse / vähesuse kohta. Eriti sümptomaatiline on selles suhtes murdesõnade histogrammil nähtuv kõrge “parempoolne küür”: selle ainuvõimalikuks seletuseks on, et murdeuurijail tekkisid millegipoolsest esinduslikud või kõitvad n-õ lemmikkihelkonnad, mida taas ja taas väisati. Ja pole ju loomulik uskuda, et mõne kihelkonna sõnavara on tegelikult suurusjärgu võrra vaesem kui teisel – et nt Kihelkonnal on olnud tegelikult käibel 11456 mittesingulaarset murdesõna, Kuusalus 10652, Kodaveres 10547, Lüganusel 10136 jne, seevastu aga Peetris kõigest 1518, Rakveres 1289, Noarootsis 928 ja Paides 728.

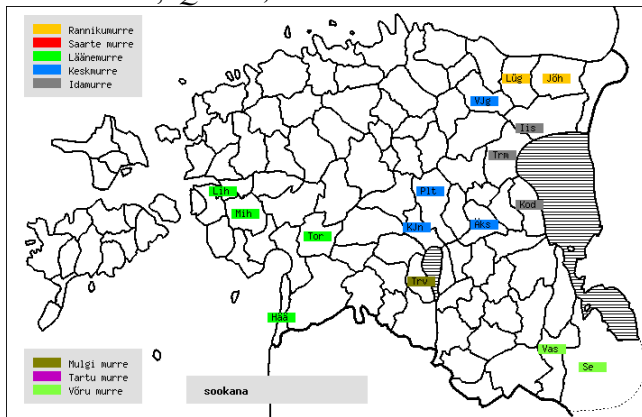


Murdeuurijad on tavatsenud murdepiire näha seal, kus ühtivad mitmete tähtsate isoglosside parmad. Meie kolmes aines seevastu on ülekaalus laiad, kuid koldelise ja katkendliku ilmega levikupildid, mis vastavad Mart Rembali (1979: 18–19) levikutüübile “Something is everywhere”.

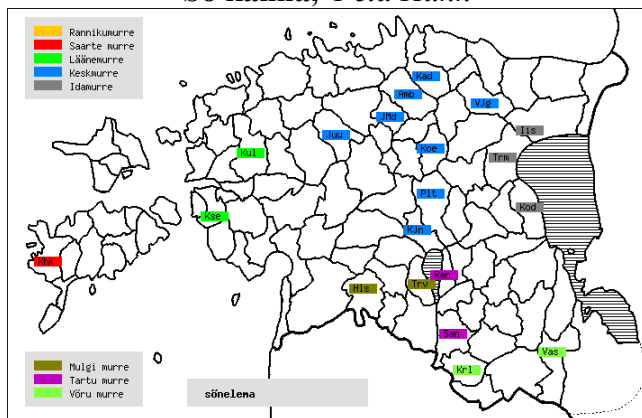
S. H. Vestringi sõnaraamatus leiduvate sõnade murdetaustast huvitunult sattusin VMS-i võrguversioonis rohketele problemaatilise ilmega kartogrammidele, näiteks:



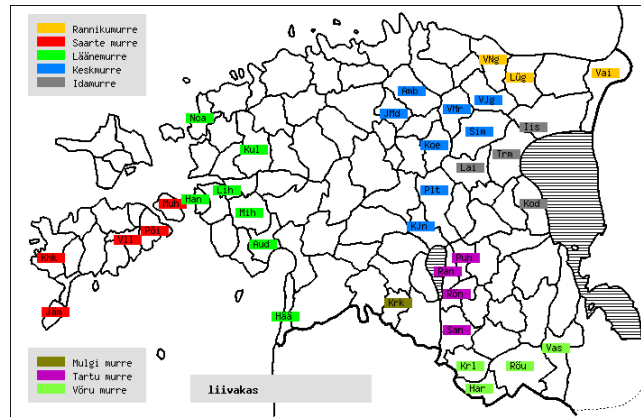
Kiddoma, Quinen, kräncken VMS 'kiratsema'



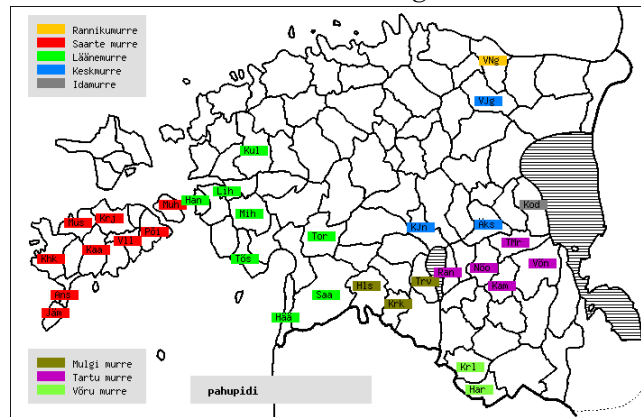
So kanna, Feld Huhn



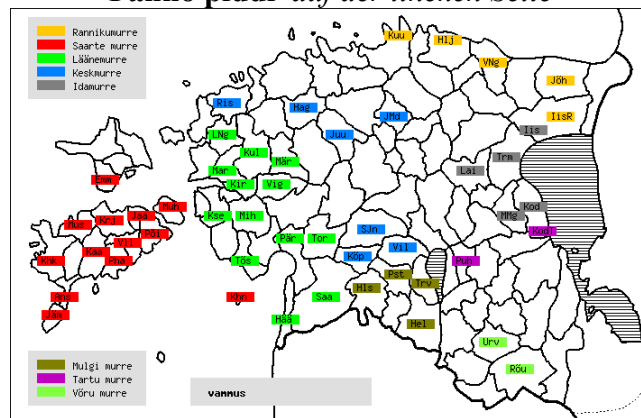
Sönnutlema | Sönnelema Worte wechseln, schelten



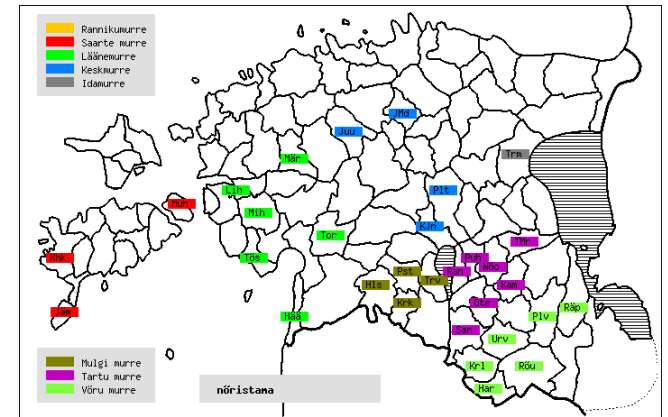
Liivakas Sandigt



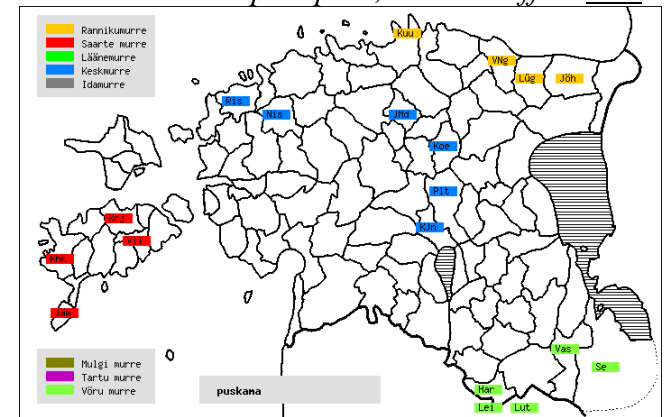
Pahho piddi auf der lincken Seite



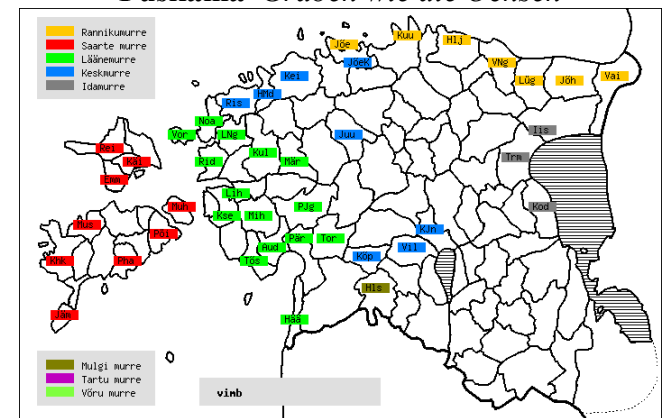
Wammus G se. Weisse baur unter kleid



Nörristama Ausplümpern, sachte lauffen. Rev.

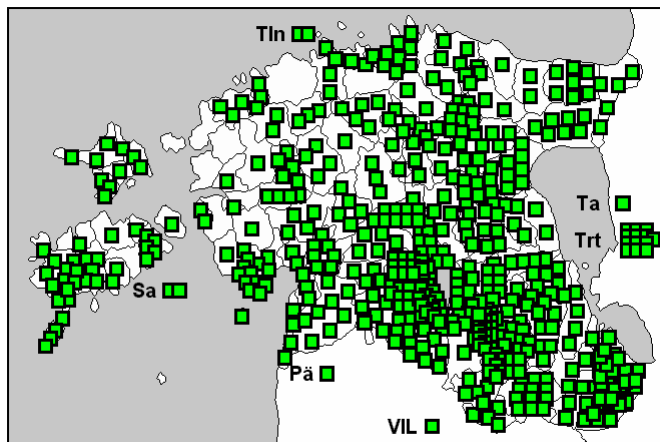


Puskama Graben wie die Ochsen

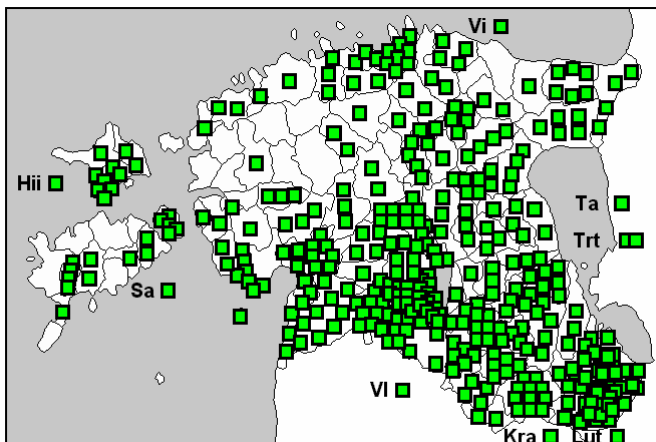


vimb Die Wemgalle

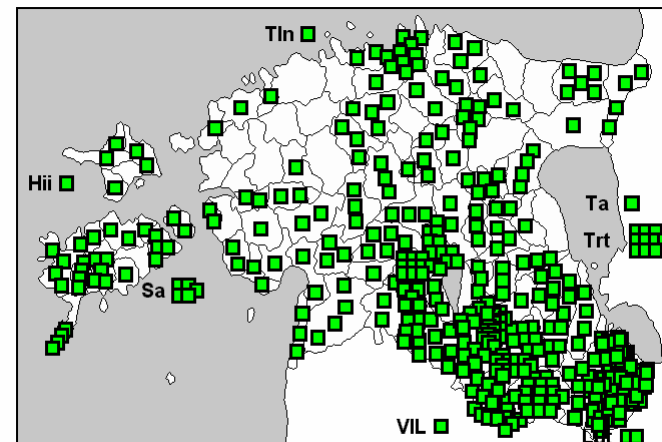
Või nt mitmed kartogrammide lõunaeesti mõistatuste kogumikus “Tere teele, tere meele...” (Krikmann 2000a: 300 jj) jätavad mulje, nagu oleks tugev kagutuul puhunud üleskirjutusi nagu lehti Kagu- vm Lõuna-Eestis asuvast hunnikust üha hõrenevalt loode suunas laiali:



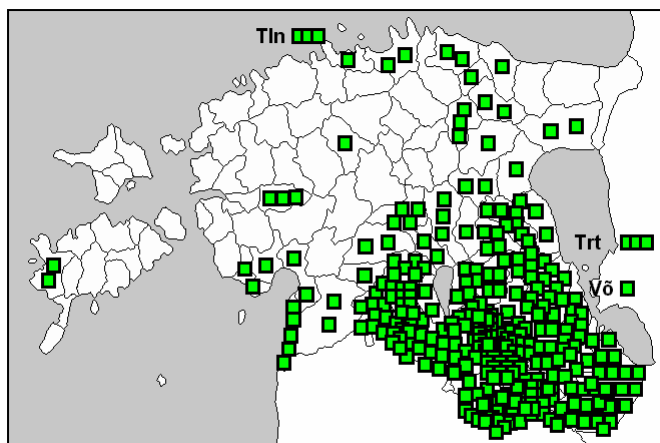
EM 2735. Üks vaat, kahesugune õlu? — Muna



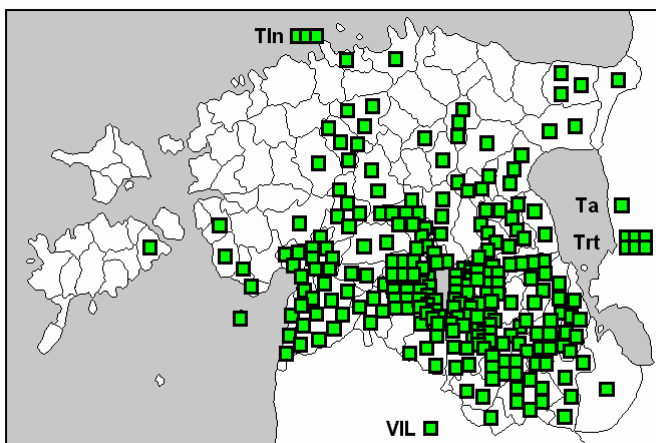
EM 1759. Puu pikkune, pilliroo jämedune? — Puu süda



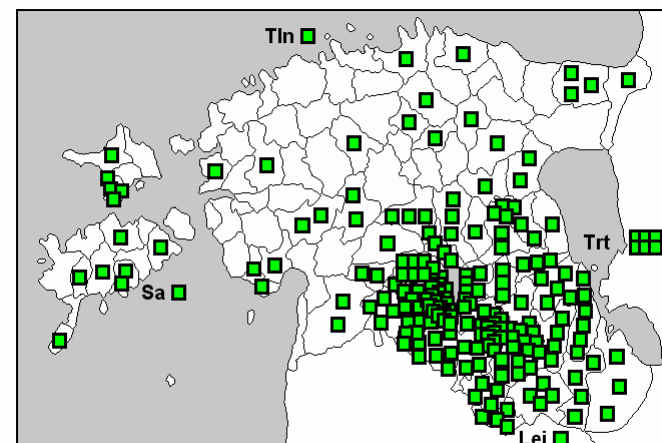
EM 533. Kats orrõtäüt valget kanno? — Hamba'



EM 566. Kahru kõnd kallast müüdä, raudsaapa jalan? — Ader



EM 1807. Päeval vorst, öösel soolikas? — Sukk



EM 359. Isa ehib, ema ehib, poeg saab poolde ilma? — Tuli, suits

Aine stereotüpsus / unikaalsus

Üldandmed geouksuste ainehulga ja stereotüpsuse kohta

Allolevas tabelis näitavad t_K -veerud kihelkondadest noteeritud murdesõnade, mõistatus- ja vanaõnatüüpide hulki, r_K -veerud neile komplektidele vastavate sisuüksuste “kohtumisi” teiste valimis esinevate geouksustega. (Neid vahekordi on selgitatud paljudes mu eelmistes töodes, nt Krikmann 1979: 28–29; 1980: 3–4; 1997: 225 jpm. Viimasest passusest pärineb ka alljärgnev illustratsioon.)

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	---	T_{4777}
K_1	×	×	×	×				×	×		×
K_2	×	×	×	×		×		×	×		×
K_3	×			×	×	×	×		×		
K_4		×	×		×	×	×		×		×
K_{110}	×			×		×			×		

Kihelkondade r_K -parameeter iseloomustab ühtlasi asjaomase materjali suhtelist stereotüpsust / unikaalsust: laiemal alal tuntud aines on stereotüpssem, kitsamal alal tuntu unikaalsem..

Khk	Murdesõnad		Mõistatused		Vanasõnad	
	t_K	r_K	t_K	r_K	t_K	r_K
Amb	3392	167174	197	11368	650	16896
Ann	2010	113065	57	4269	227	7979
Ans	2611	119287	126	8126	75	2242
Aud	4763	200822	169	10472	245	8210
Emm	6310	239227	146	8094	125	4265
Hag	2181	111296	109	7479	40	1613
Han	4206	167626	200	11789	148	4861
Har	8521	254692	280	12786	123	3562
Hel	5331	199634	418	17444	868	20918
HJn	2281	117293	190	10756	78	3016
Hlj	4161	157704	390	16920	760	17869
Hls	8097	283889	384	17179	869	20427

HMd	1543	88384	155	9513	384	11584
Hää	9086	315254	302	14851	621	13551
Iis	9465	321226	256	13850	345	10796
Jaa	2009	88639	99	5959	48	1565
JJn	2875	133386	233	13405	227	6423
JMd	7484	298596	124	8533	96	4168
Jõe	4669	182331	203	11954	444	14679
Jõh	6198	218952	236	13036	277	9300
Juu	9070	328146	326	15913	580	15905
Jür	2799	127608	232	12087	239	7699
Jäm	8960	306812	182	10579	320	9736
Kaa	4316	164600	188	11031	489	12929
Kad	5440	208392	265	13773	286	8965
Kam	4390	160325	158	8830	489	13533
Kan	3543	127909	353	15812	393	9981
Kei	3223	154839	92	6433	164	5884
Khk	11456	349539	137	8458	257	8367
Khn	4175	182489	141	8639	263	8424
Kir	1965	106408	72	4893	10	315
KJn	9170	329523	429	18127	273	8881
Kod	10547	327291	224	12503	506	14648
Koe	7628	303471	244	13320	404	12492
Köp	1801	95323	317	15419	302	9725
Kos	3683	173737	168	9714	612	15277
Krj	2877	124919	296	14734	683	17636
Krk	9568	297617	342	14871	268	7907
Krl	6971	246751	151	9186	63	2001
Kse	8831	315857	218	12560	290	10182
Ksi	3179	150110	115	7836	344	9499
Kul	3276	164456	424	17645	1208	24128
Kuu	10652	325657	197	11467	122	4194
Käi	4038	165801	139	8085	216	5632
Kär	1840	94944	162	9697	216	7355
Lai	6253	241648	381	16872	672	17815
Lih	2106	116721	73	5318	9	431

LNg	2640	132881	96	6188	108	3566
Lüg	10136	321194	135	8498	747	17794
Mar	8436	310598	117	7891	183	4588
Mih	3642	166151	175	10583	191	7619
MMg	1807	95840	276	14017	312	10395
Muh	9535	320300	236	12463	226	8182
Mus	3382	132062	152	8478	184	6099
Mär	8843	324333	195	11317	814	19292
Nis	2050	102732	78	5414	82	3227
Noa	928	52033	134	8728	52	1487
Nõo	6564	234989	329	15490	520	14022
Ote	4194	156591	434	17311	835	18786
Pai	728	43665	123	7821	106	3933
Pal	2717	143647	321	15612	707	19121
Pee	1518	83161	169	10327	99	4004
Pha	3408	141503	79	4735	228	7789
Phl	3879	156805	124	7754	116	3645
Pil	2236	107629	360	17027	302	11024
PJg	4362	184340	211	12463	188	7097
Plt	9549	338645	368	17064	564	16442
Plv	8507	260061	272	14005	662	15380
Pst	2699	115165	426	18016	543	14977
Puh	7044	264275	262	13657	406	12512
Pär	1568	86371	308	15105	328	9273
Põi	6495	237655	258	13269	526	14353
Rak	1289	77502	235	12809	681	17194
Ran	6184	224013	263	13233	401	12778
Rap	1546	88042	180	10805	243	8694
Rei	7443	279015	147	8940	124	4624
Rid	3598	167209	116	7504	109	3539
Ris	7534	293971	171	10391	193	6540
Rõn	2686	119831	341	15089	285	8236
Rõu	9326	267037	420	16663	1002	20000
Räp	6283	207172	298	14354	684	15181
Saa	3164	139284	386	17521	919	19569

San	7730	268823	295	14238	429	11831
Se	7769	219698	486	15822	926	16935
Sim	4385	185259	291	14585	273	9638
SJn	2466	121943	379	17146	679	18413
TMr	1683	93093	254	13344	658	16695
Tor	8063	312618	368	17068	310	10023
Tös	8660	318856	357	16669	607	17615
Trm	8853	321114	447	18104	1068	22396
Trv	8465	294213	327	15787	283	9595
Tür	2636	141950	206	11982	315	11521
Urv	5416	171434	331	14995	375	9578
Vai	8111	282308	274	13428	403	12741
Var	4369	178402	173	10733	32	1194
Vas	8203	243341	293	12240	1269	20438
Vig	5121	216850	298	14595	653	17296
Vil	2558	130958	467	18550	1162	23750
VJg	8137	309051	224	12428	169	6337
Vll	6966	267018	87	5649	148	5141
VMr	3874	184534	356	16375	503	14807
VNg	8995	297642	131	7261	433	11479
Vön	3481	139333	368	16599	483	13579
Vän	3181	163814	338	16356	585	15818
Äks	2870	148516	296	15137	208	7028
Σ	546862	20855840	25541	1291218	41969	1128275

Vanasõnade näidud on üsna sarnased, kuigi mitte kattuvad, meie varasemate vaatlusmaterjalide andmetega (vt nt Krikmann 1997: 227–229).

Arusaadavalt looduslikel põhjustel pole kihelkonna keele- või folklooriainese *t*- ja *r*-näitude seos lineaarne, vaid kergelt kumer: kuna sisuüksuste sagedusjaotus mistahes loomulikus ainevalimis on zipfilikult ebäühtlane (väikesi üksusi palju, suuri vähe, keskmisi keskmisel määral), siis saab pikem mittekorduvate sisuüksuste loend koosneda ainult suhteliselt unikaalsemest ainest, kuna sagedaste, stereotüüpsete üksuste varu on väike ja ammendub üsna peatselt.

Stereotüüpne südamaa / unikaalne perifeeria

Kuigi meie vaatlusest on välja lülitatud ainestike kõige unikaalsem osa, tõestavad need jätkuvalt, et peab paika teine kultuurigeograafiline põhitõde: perifeeriaste aines on unikaalne, südamaa aines stereotüüpne. Perifeeriatena tõusevad meil esile eelkõige Kagu-Eesti laias mõttes (Se, Võ, TaL) ja saared, nõrgemal määral ka lõunaeeesti kompleksi liituv Mulgi, saarte kompleksi liituv Läänemaa rannäärne mandriosa ning mõned Kirde-Ranna-Eesti kihelkonnad (eriti Kuu, Hlj, VNg).

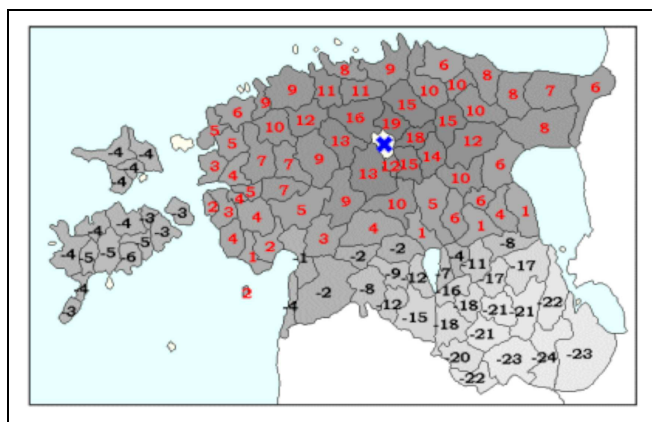
Mulle oli ammuilma teada juba üldisemaist loodusseadustest johtuv tõik, et stereotüüpse materjaliga Põhja- (ja eriti süda-Eesti) geouksuste suhtlemispilt teiste geouksustega on loid ja üheülbaline, unikaalsema materjaliga alade oma aga dünaamiline ja reljeefne. Mu eelmainitud jm varasemais töödes leidub rohkesti siinsega sarnaste hindamise meetodite alusel saadud kartogramme Eesti kihelkondadest noteeritud murdesõnavara ning mõistatus- ja vanasõnaainese stereotüüpia / unikaalsuse kohta. – näited järgnevas seerias pärinevad saitidest “Eesti kihelkondade murdesõnavaralisi suhteid” (Krikmann 1998b)

(<http://www.folklore.ee/~kriku/MURRE/Index.htm>)

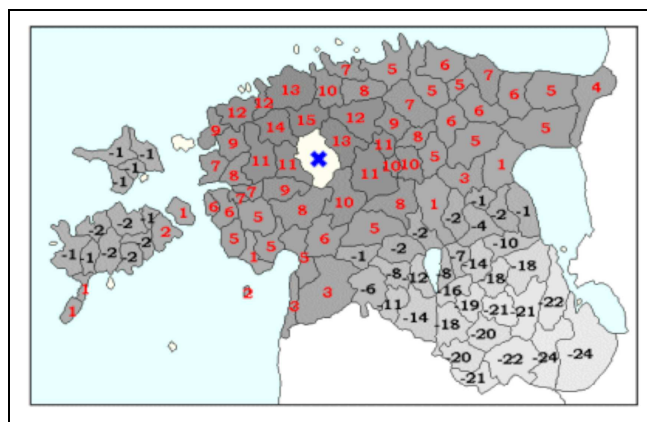
ja “Eesti kihelkondade suhtlemistihedus mõistatuses” (Krikmann 2000)

(<http://www.folklore.ee/~kriku/LEVIK/coefs.htm>)

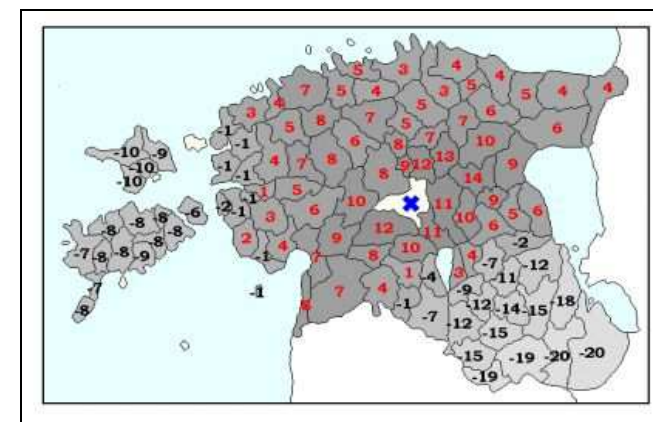
(kõik kartogrammid silutud):



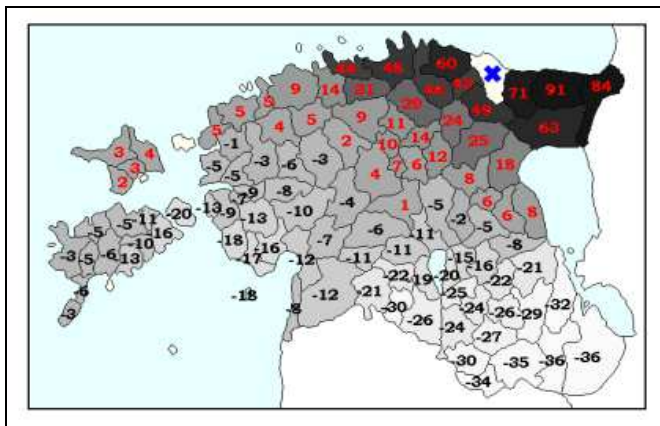
Anna murdesõnad



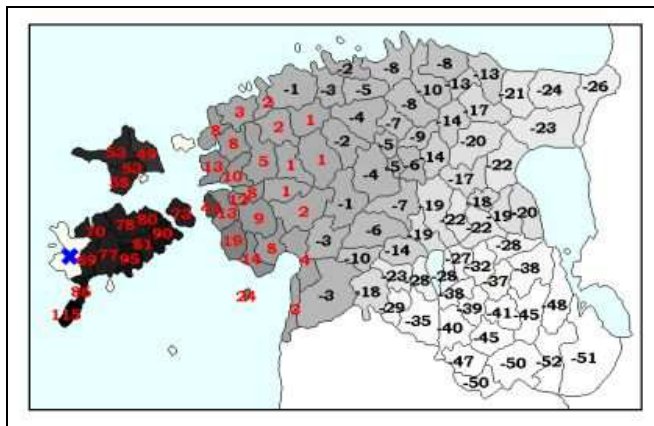
Rapla murdesõnad



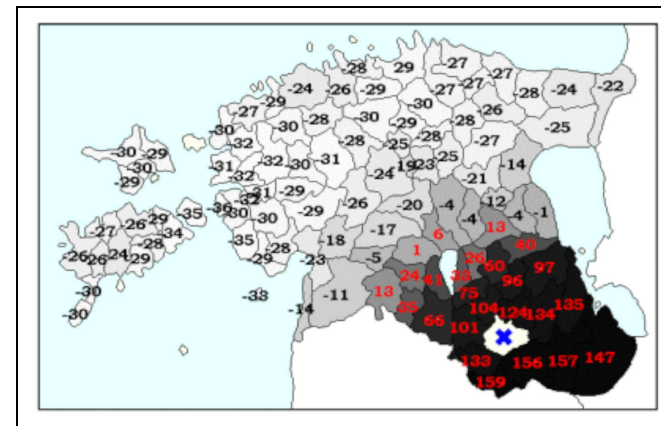
Pilistvere murdesõnad



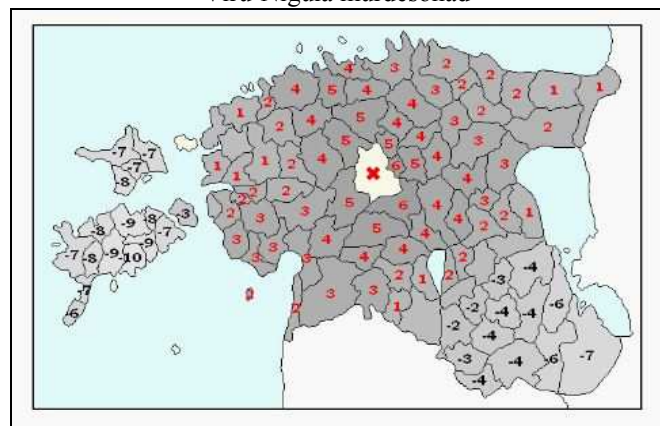
Viru-Nigula murdesõnad



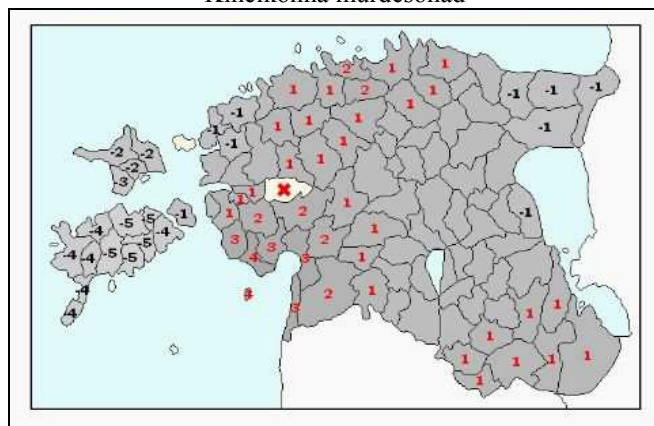
Kihelkonna murdesõnad



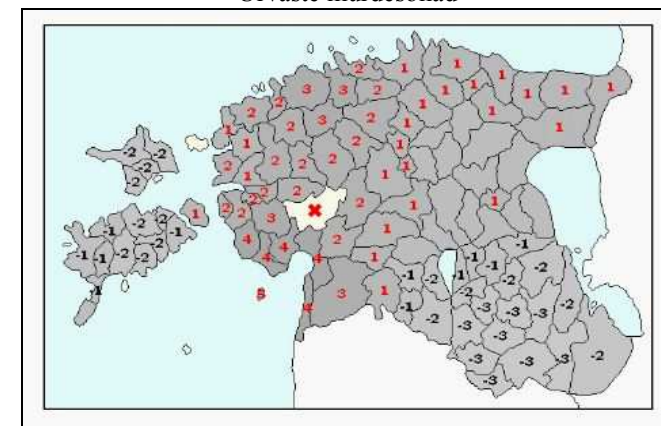
Urvaste murdesõnad



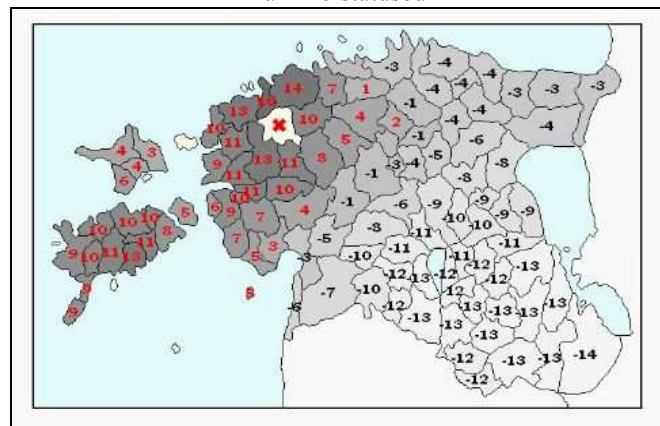
Türi mõistatused



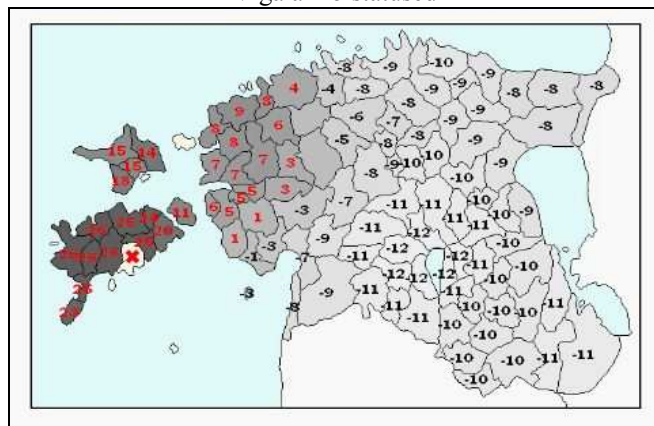
Vigala mõistatused



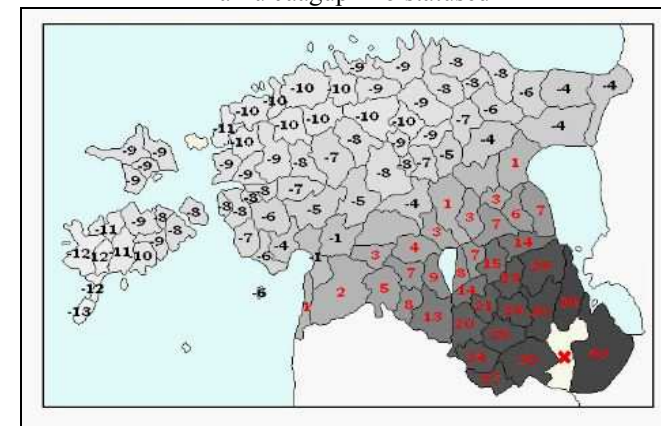
Pärnu-Jaagupi mõistatused



Nissi mõistatused



Püha mõistatused



Vastseliina mõistatused

Püüdsin kihelkondade vahelise seosetiheduse hindamiseks vanasõnade levikuandmestikus leiutada koguni omaenda meetodika, mis arvestaks mitte ainult nende ühisosa suurst, vaid ka suhtelist stereotüüpsust, olles arvanud, et küsimus pole mitte niivõrd “täpse lähendfunktsiooni leidmises, vaid sellise lähendfunktsiooni leidmises, millel oleks seesmine loogika” (Krikmann 1997: 224). Olen hiljem püüdnud testida oma “vanasõnavalemi” võimalikku üldisemat kasutatavust ja saanud kahetisi tulemusi: nt Eesti kihelkondade murdesõnavara stereotüüpsuse hindamiseks on ta sobinud päris hästi (vt Krikmann ja Pajusalu 2000: 141), H. Gösekeni, S. H. Vestringi ja A. Thor Helle vanasõnade geograafilise tausta leidmiseks aga mitte, mistõttu selleks on kasutatud teistsugust arvutusviisi (vt Krikmann 1986: 112–114).

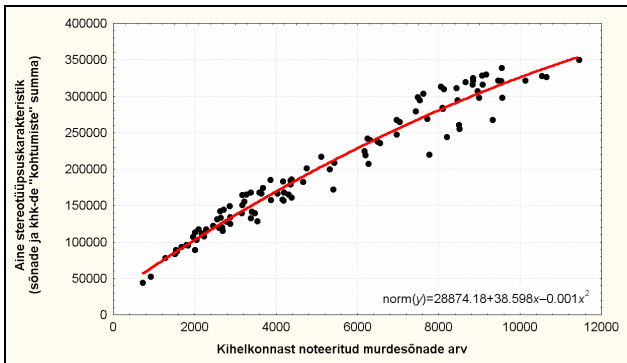
Siinses vaatluses on *t*- ja *r*-vahekordi normeeriva lähendfunktsioonina kasutatud lihtsalt 2. astme polünoomi ja kihelkondlike sisukogumite hinnangutena normipõigete (nn residuaalide) suurus (vt lk 18 ülal) – nimetagem neid siin *s*-koefitsientideks. Kolmel järgmisel kartogrammil (vt taas lk 18) on toodud *s*-koefitsientide “toored” empiirilised jagunemispildid ning kolmel viimasel silutud jagunemised. Silumised on toimunud lihtviisiliselt samamoodi nagu mu mitmetes eelmistes töödes: iga kihelkond saab igas järgmises iteratsiooniringis näiduks ta enda ja ta naabrite eelmise ringi näitude keskmised.

Näiteks:

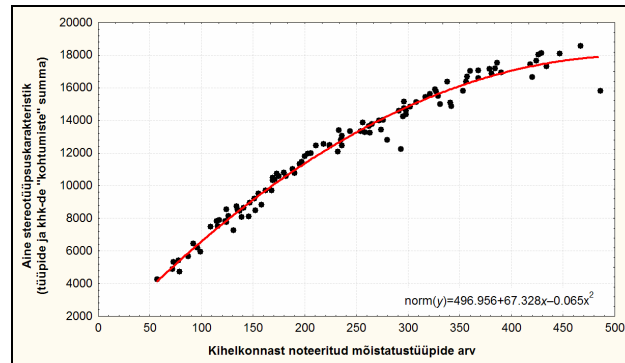
Ambla jaoks: (Amb+Kuu+Kad+VMr+JJn+JMd+Kos+HJn)/8;

Anna jaoks: (Ann+JMd+JJn+Pee+Pai+Tür+Kos)/7, jne jne.

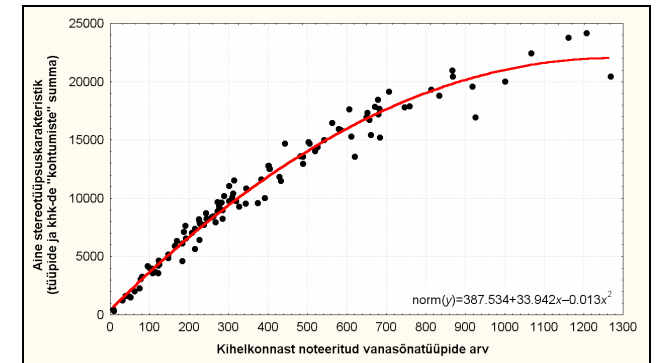
Taolisel silumisviisil on see halb kül, et kuna maa äärealade kihelkondadel on keskmiselt vähem naabreid kui mujal ja seega keskmised lamenduvad seal aeglasemalt, siis iteratsioonide tulemusel kanduvad kõrgemad näidud maa äärealadele; kolme iteratsiooni korral on see häire juba märgatav. Stereotüüpsuse / unikaalsuse hinnangute korral on “silumismüra” liiati desinformeeriv, kuna äärealade aines on ka loomuldasa suhteliselt unikaalne.



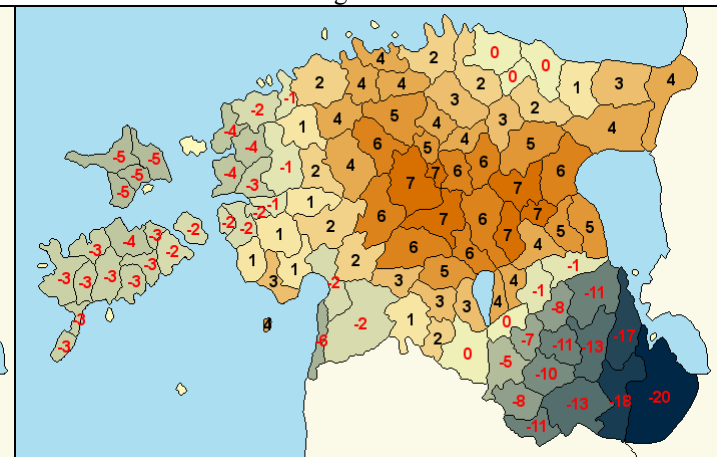
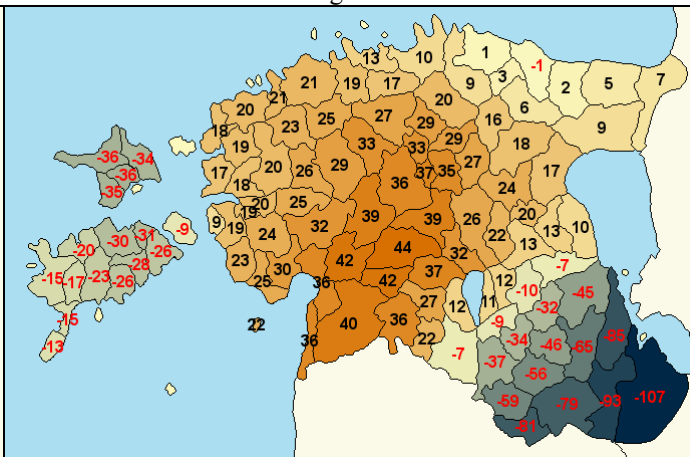
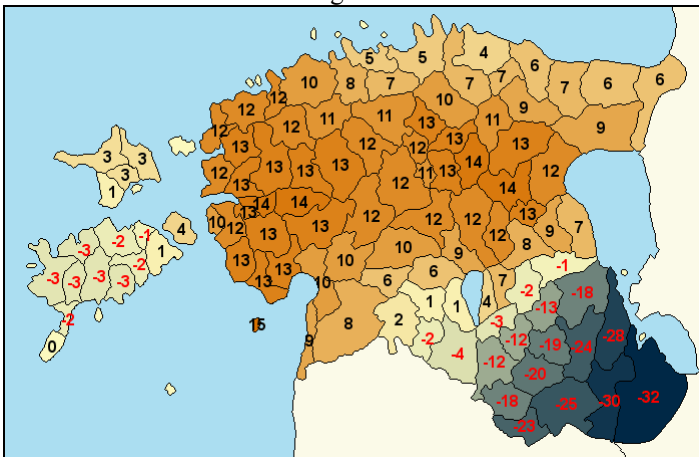
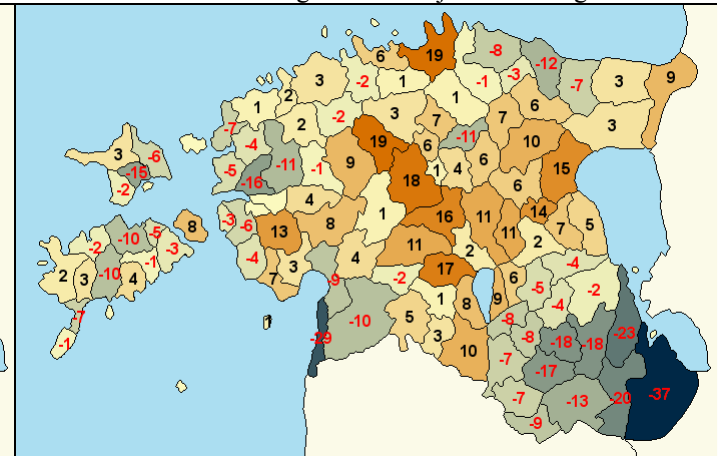
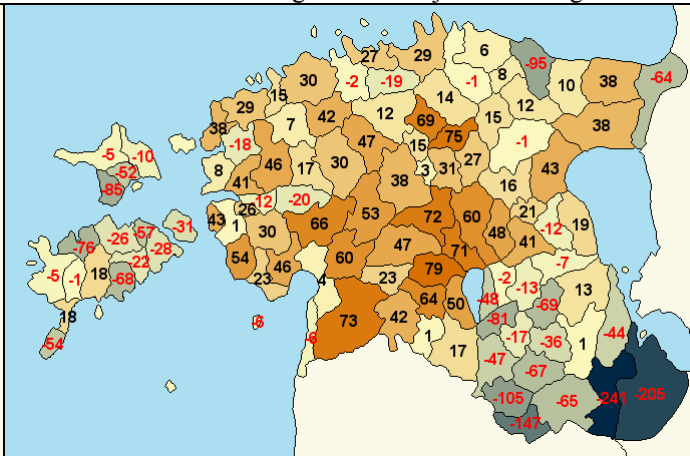
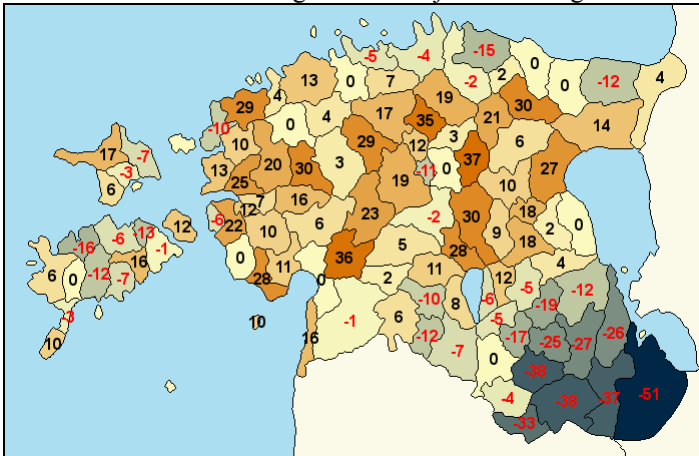
Murdesõnad: regressioonivälja normeerimine



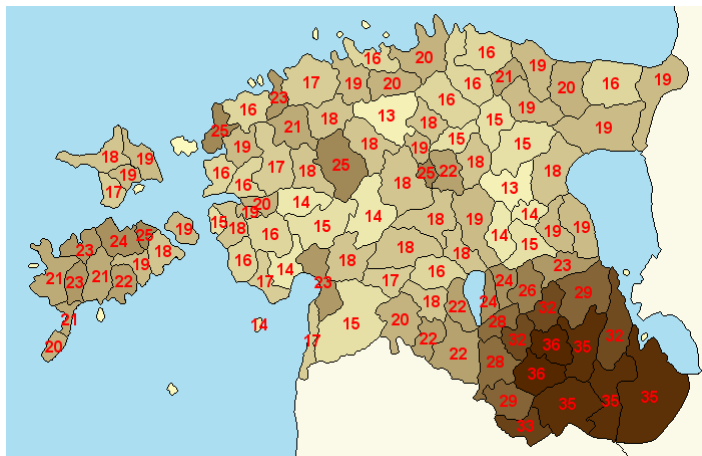
Mõistatused: regressioonivälja normeerimine



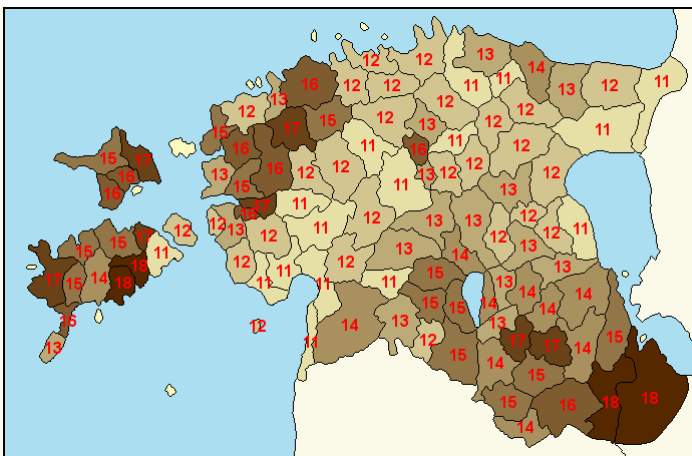
Vanasõnad: regressioonivälja normeerimine



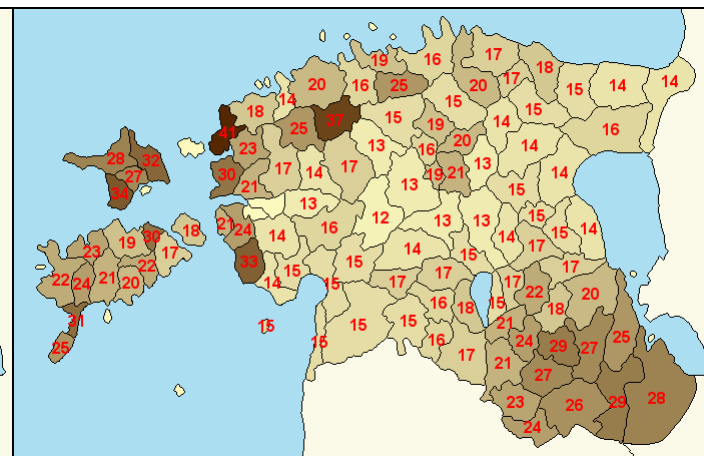
Käesolevas vaatluses on iga kihelkonna λ -koefitsientide summad itereeritud võrdseteks, kuid nad erinevad oma hälbemäärade poolest: unikaalsema ainega kihelkondadel on need eeldatavasti suuremad ning stereotüüpsema ainega kihelkondadel väiksemad. Vaatlus näitabki, et λ -koefitsientide standardhälbed sobiksid samuti päris hästi meie geouksuste stereotüüpse iseloomustajateks – vt järgmises kaardisarjas:



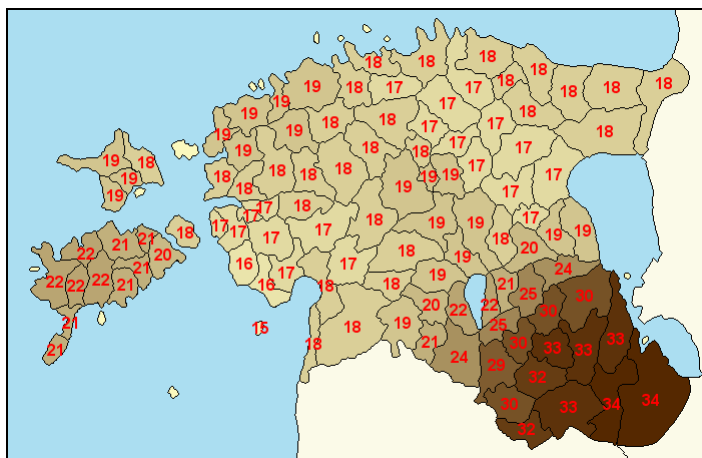
Murdesõnad: λ -koefitsientide standardhälbed



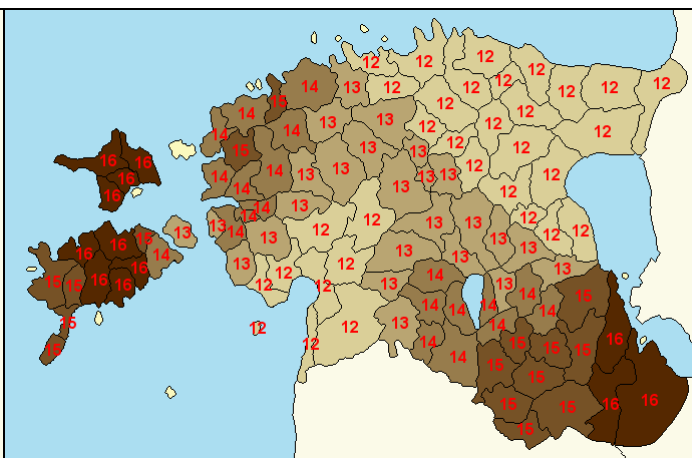
Mõistatused: λ -koefitsientide standardhälbed



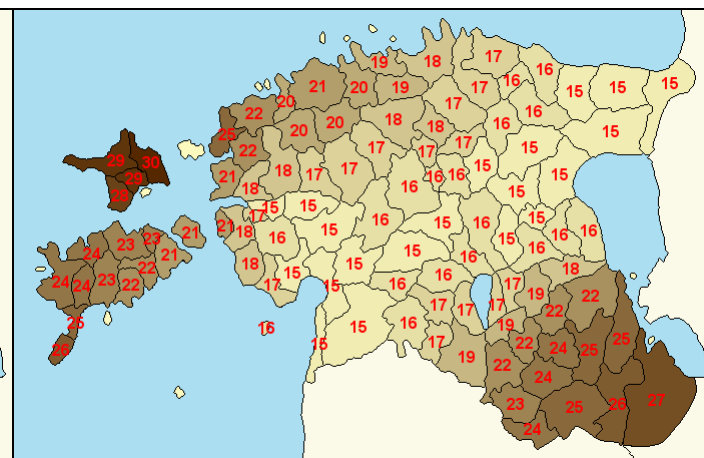
Vanasõnad: λ -koefitsientide standardhälbed



Murdesõnad: standardhälvete silumid (2 iteratsiooni)



Mõistatused: standardhälvete silumid (2 iteratsiooni)



Vanasõnad: standardhälvete silumid (2 iteratsiooni)

λ -koefitsientide pingeread ja mis nende sees toimub

Oma kolme kõnealuse andmestiku kohta koostasin samuti eelmiste voorudega analoogilised maatrikstabelid kõigi kihelkonnapaaride “kohtumiste” hulga kohta samades sisuüksustes, kõrvaldasin tabelist diagonaalväärtused (s.o kihelkondade “ühisosad iseendaga”) ja arvasin kõigi tabeliruutude kohta standardiseeritud keskmistega λ -koefitsiendid. Diagonaalne katkend mõistatuste λ -tabelist oleks nt selline:

	Amb	Ann	Ans	Aud			VNg	Võn	Vän	Äks
Amb	0	112	99	105						
Ann	112	0	121	118						
Ans	99	121	0	101						
Aud	105	118	101	0						
					0					
						0				
VNg							0	94	97	111
Võn							94	0	102	108
Vän							97	102	0	102
Äks							111	108	102	0

Edasi tuletasin iga sisukorpuse (murdesõnad, mõistatused, vanasõnad) λ -koefitsientide pooltabeli põhjal kihelkonnapaaride koefitsientide (λ -de täisarvuliste 100-kordsete) paremusjärjestuste tabeli, mille pikkuseks on murdesõnade ja mõistatuste korral $(105 \times 104) / 2 = 5460$ rida ja vanasõnade korral $(103 \times 102) / 2 = 5253$ rida.

Nt mõistatuste λ -pingerea ülaots näeb välja nõnda (K_1 ja K_2 on kihelkonnad, X_1 , Y_1 ja X_2 , Y_2 nende koordinaadid meie aluskaardi tingühikutes ning $\lambda(K_1 K_2)$ nende seosetihedushinne):

Jrk-nr	K_1	K_2	X_1	Y_1	X_2	Y_2	$\lambda(K_1K_2)$
1	Se	Vas	643	63	602	59	173
2	Emm	Phl	107	267	138	300	152
3	Ans	Vll	63	140	134	195	149
4	Jaa	Krj	139	224	113	218	147
5	Emm	Käi	107	267	121	284	144
6	Nis	Pha	286	332	118	178	144
7	Käi	Phl	121	284	138	300	143
8	Räp	Vas	608	124	602	59	143
9	Kir	Lih	256	255	243	247	141
10	Jaa	Pha	139	224	118	178	140
----	----	----	----	----	----	----	-----

Nüüd oleks loomulik huvitada küsimusest, mis toimub nende λ -pingeridade sees: millised geoüksused üksteist armastavad ja millised väldivad, millistes ruumisuhetes paarid domineerivad skaala eri osades jne.

Järgnevas on toodud väljavõtteid neist arengutest 250 sammu pikkuste etappide kaupa.

Iga tabelirea vasakpoolne kaart näitab antud vahemikus esindatud kihelkondade seostumissagedusi teistega.

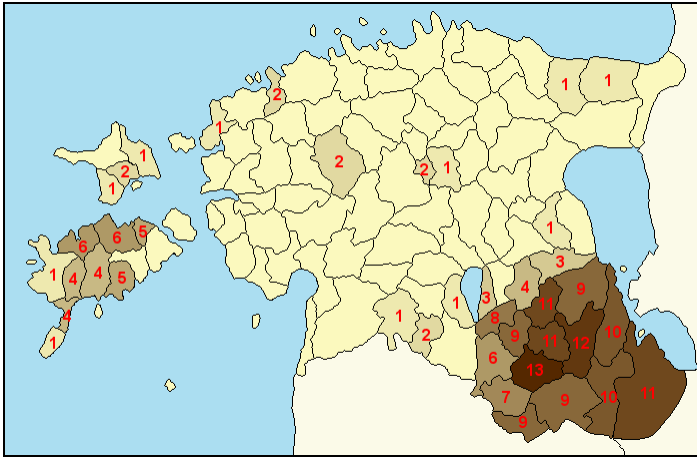
Keskmine kaart annab vasakpoolsest silutud pildi (3 iteratsiooni).

Parempoolsetel kartogrammidel ühendavad mustad sirgjooned seosesse angažeeritud kihelkondade keskmehid, sinised kastikesed näitavad nende sirgete keskpunkte ja suuremad punased tähnid siniste joonepoolitajate keskmist asukohta.

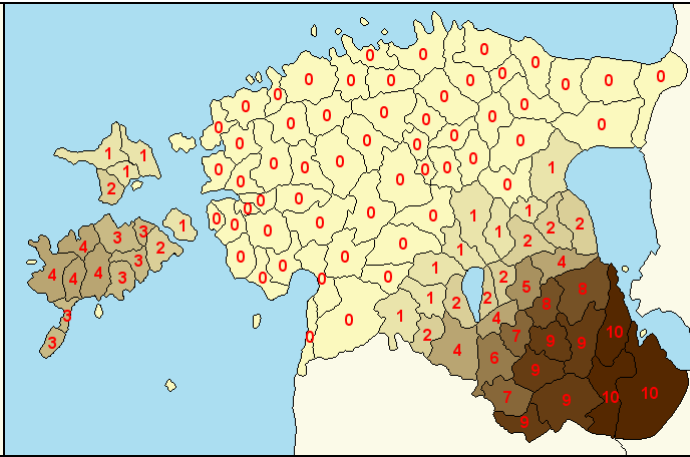
Vajalike arvutuste tegemiseks ja andmete parempoolsete must-valge põhjaga kartogrammide koostamiseks kasutasin möödunud sajandi lõpuaastail koostatud *offline*-režiimil töötavat kaardistajat, mis põhineb samuti Statistica 5.5 vahenditel; arvandmete ja graafika vahendamiseks oli vaja kirjutada mõningaid uusi skripte. Sini-kollase põhjaga kartogrammid on tehtud eelnimetatud *online*-kaardimasina abil, mis laseb hõlpsasti teha ka silutud kaarte, kuid ei võimalda kanda kaartidele joongraafilisi objekte.

Piiranguid tulenes ka sellest, et Statistica 5.5 ei luba oma .stg-laiendiga joonistele kanda rohkem kui 901 lisamärki; selle maksimumi ületamisega kaasnevad häälekad protestid.

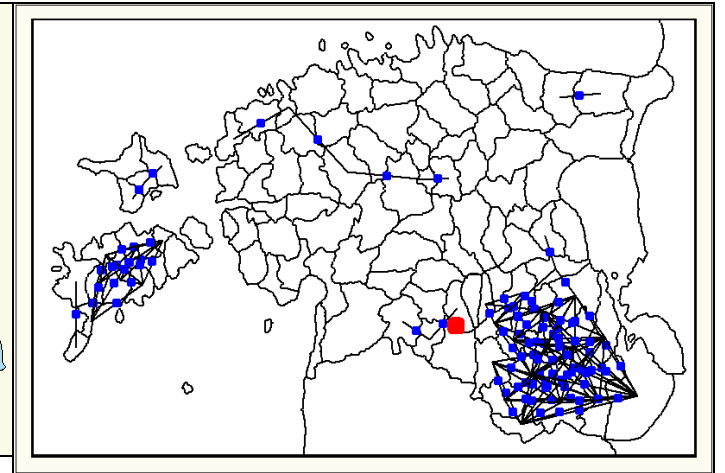
Murdesõnad



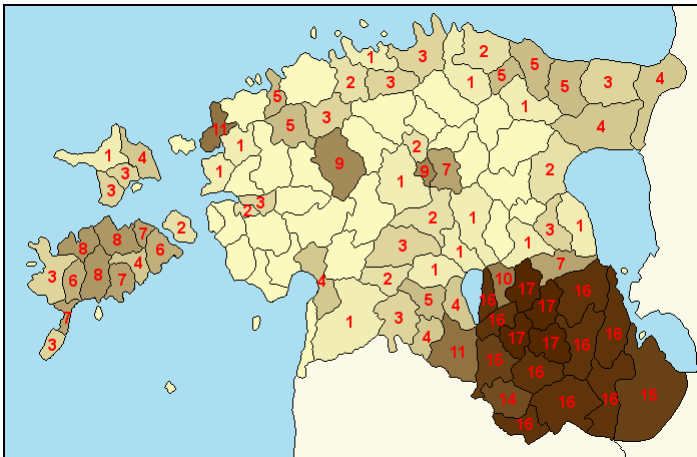
Murdesõnad: 1–100 toorkoefitsiendid



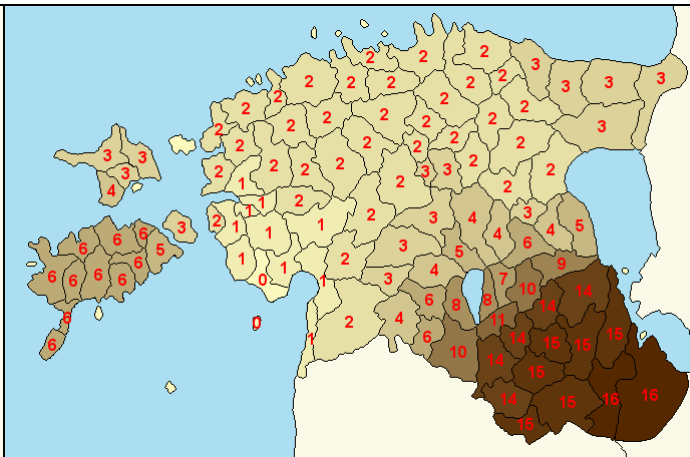
Murdesõnad: 1–100 silutud koefitsiendid



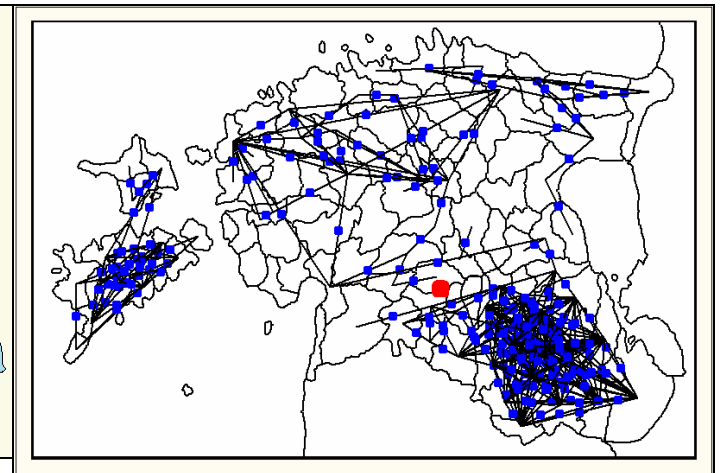
Murdesõnad: 1–100 seoste graaf



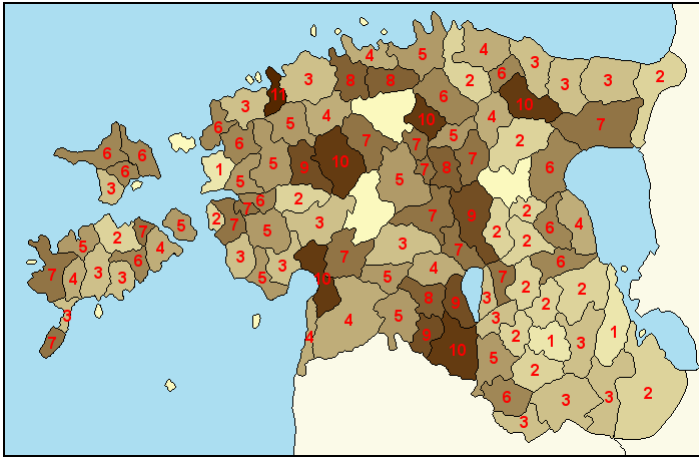
Murdesõnad: 1–250 toorkoefitsiendid



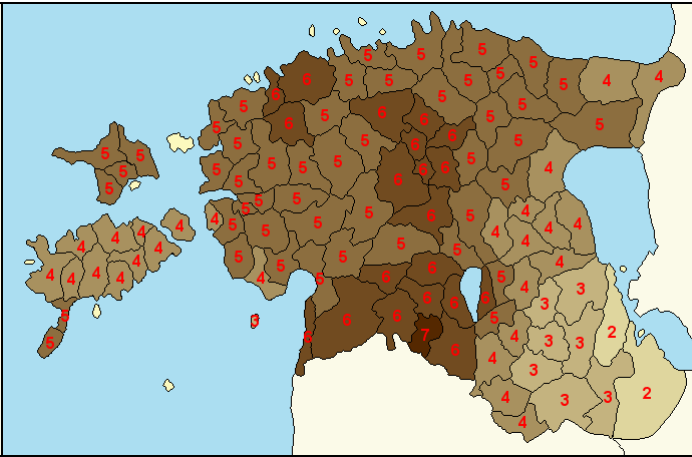
Murdesõnad: 1–250 silutud koefitsiendid



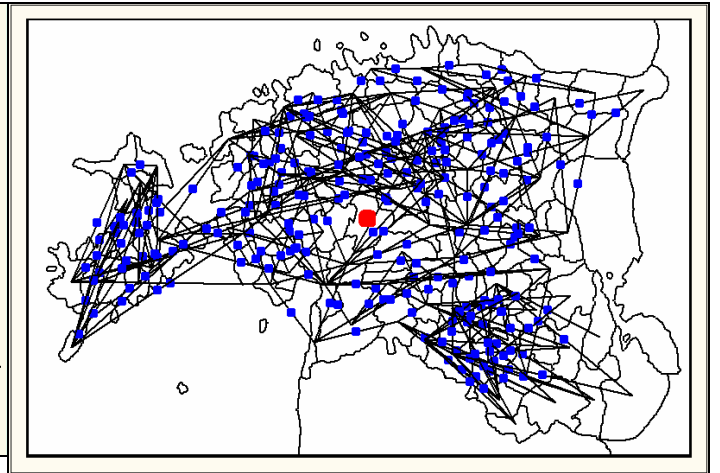
Murdesõnad: 1–250 seoste graaf



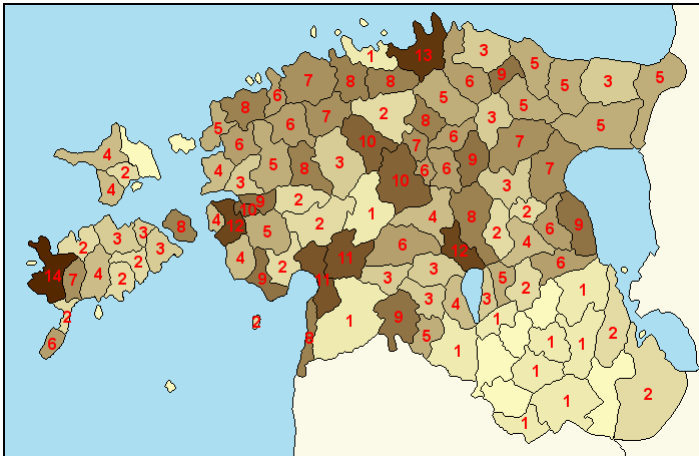
Murdesõnad: 251–500 toorkoefitsiendid



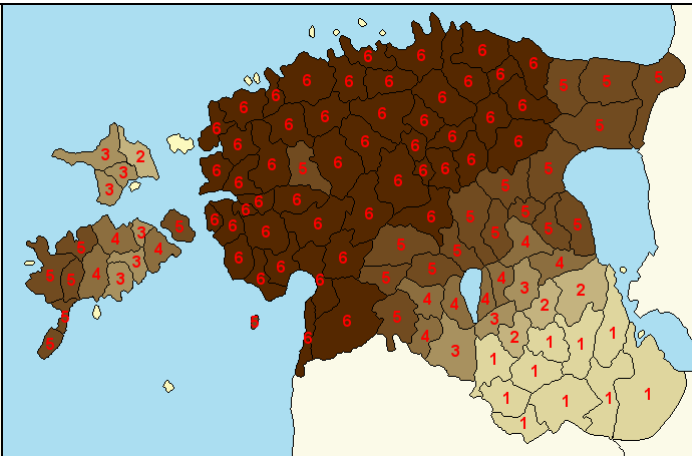
Murdesõnad: 251–500 silutud koefitsiendid



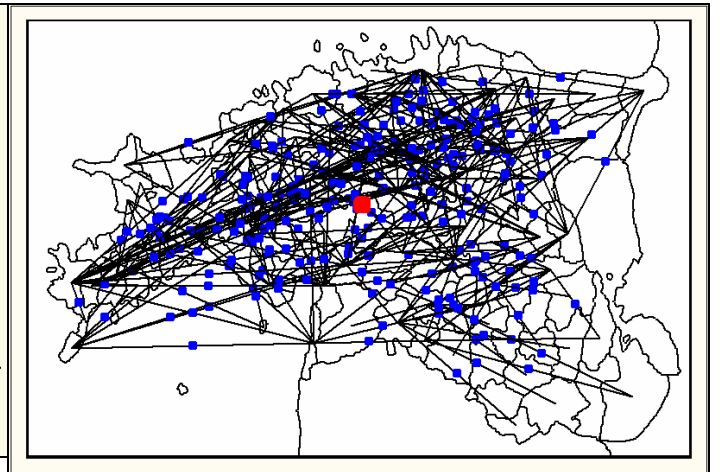
Murdesõnad: 251–500 seoste graaf



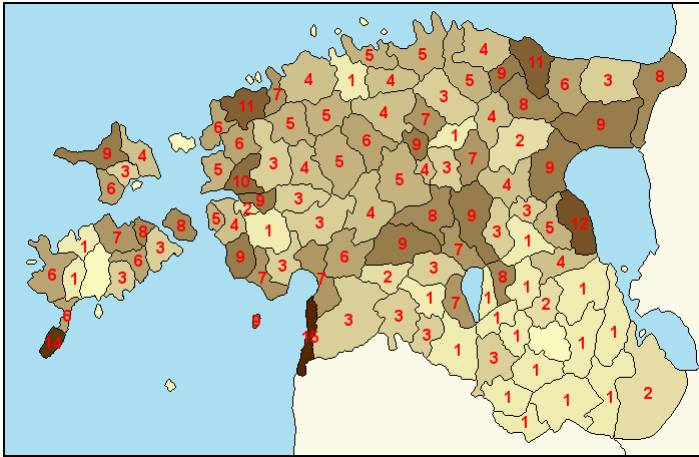
Murdesõnad: 501–750 toorkoefitsiendid



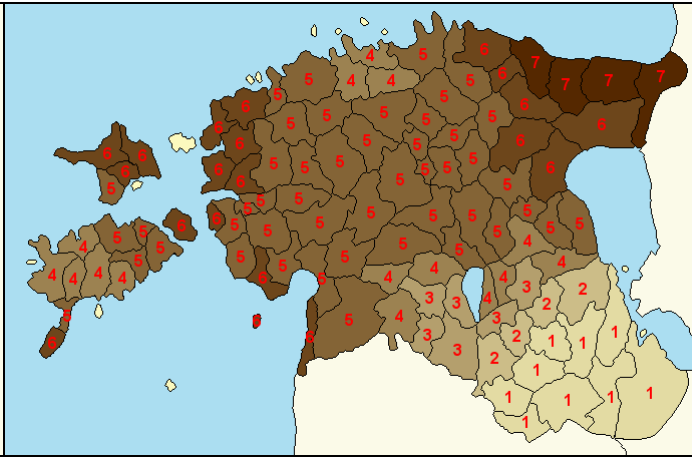
Murdesõnad: 501–750 silutud koefitsiendid



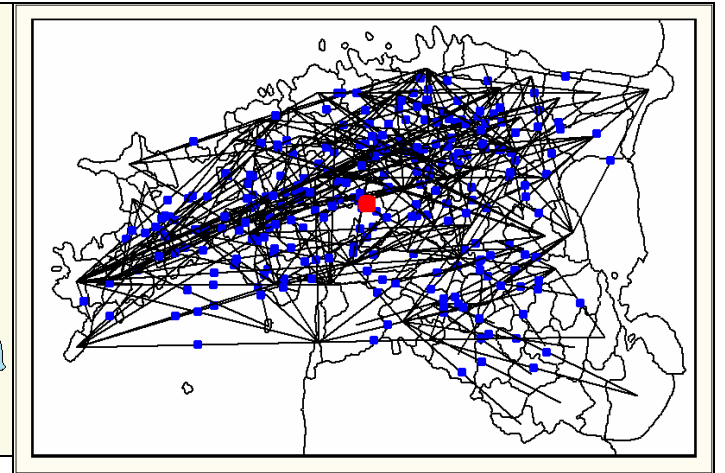
Murdesõnad: 501–750 seoste graaf



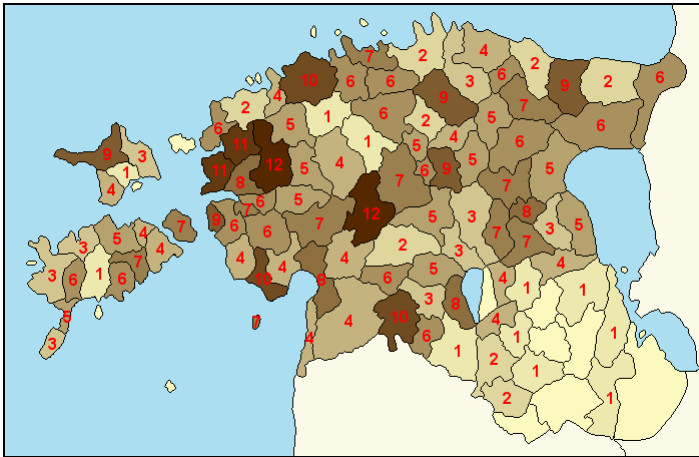
Murdesõnad: 751–1000 toorkoefitsiendid



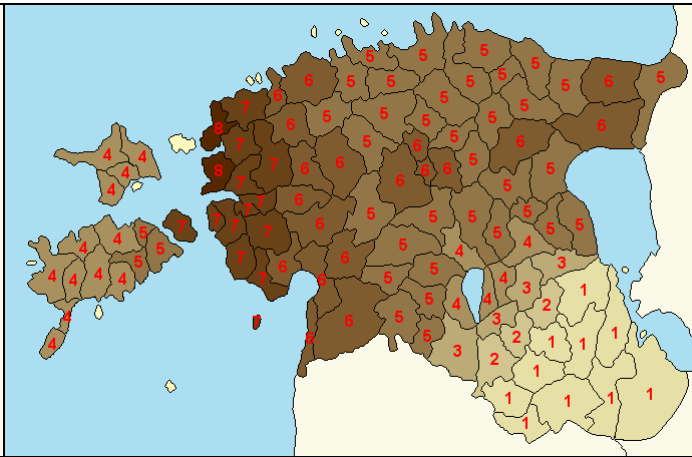
Murdesõnad: 751–1000 silutud koefitsiendid



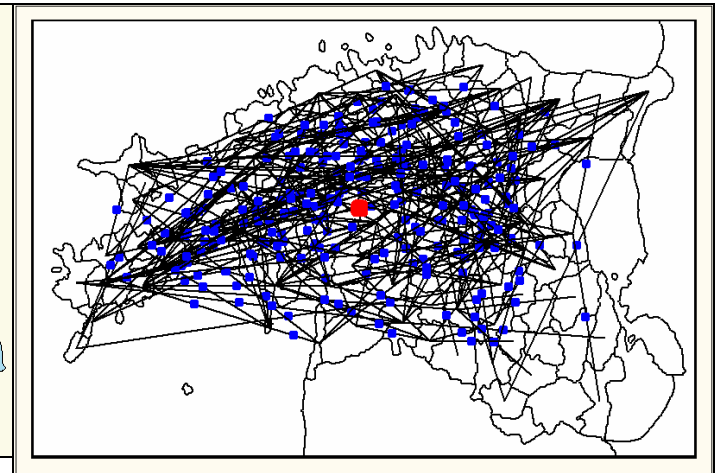
Murdesõnad: 751–1000 seoste graaf



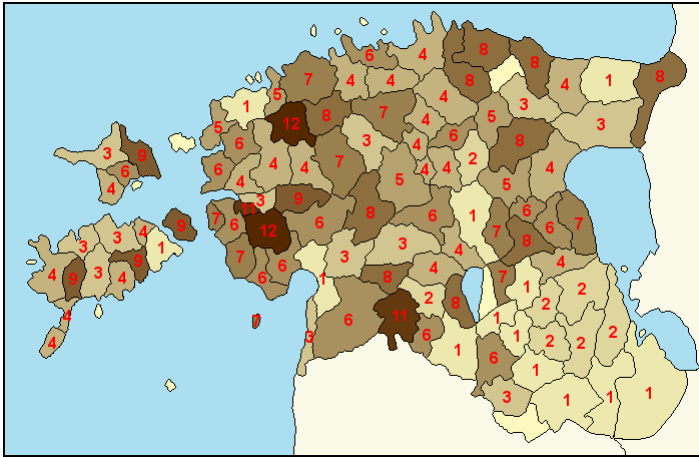
Murdesõnad: 1001–1250 toorkoefitsiendid



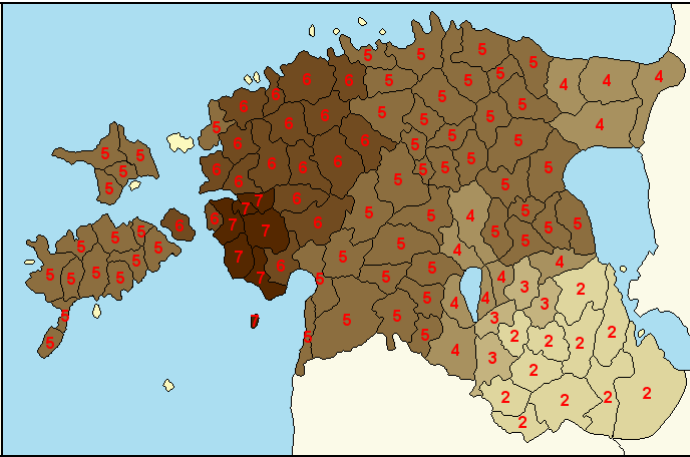
Murdesõnad: 1001–1250 silutud koefitsiendid



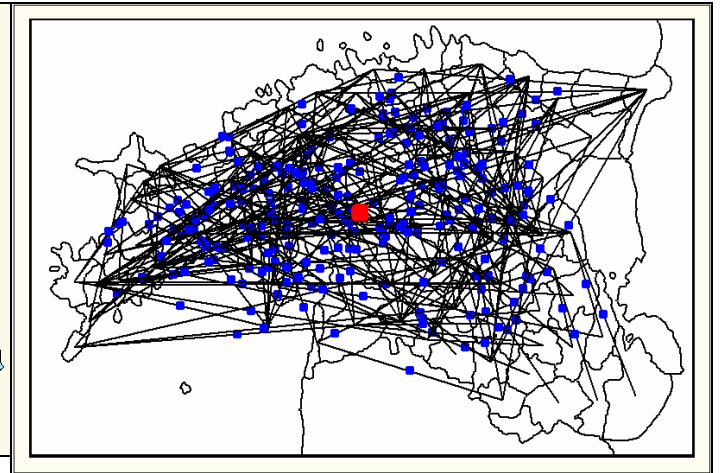
Murdesõnad: 1001–1250 seoste graaf



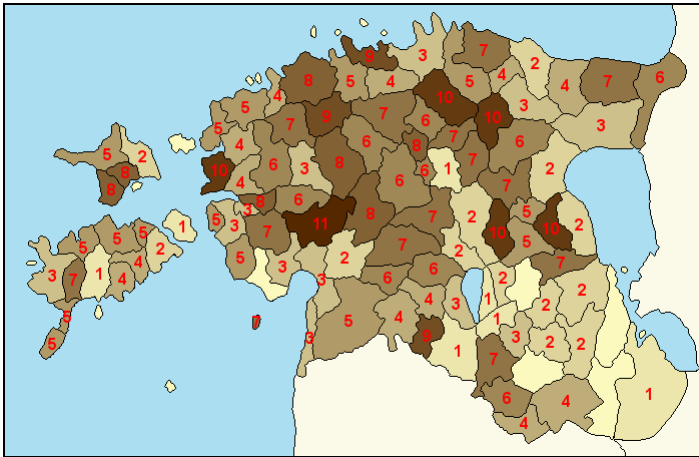
Murdesõnad: 1251–1500 toorkoefitsiendid



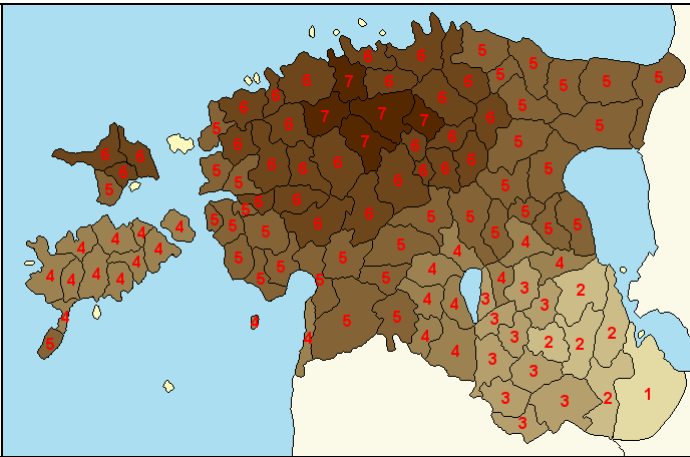
Murdesõnad: 1251–1500 silutud koefitsiendid



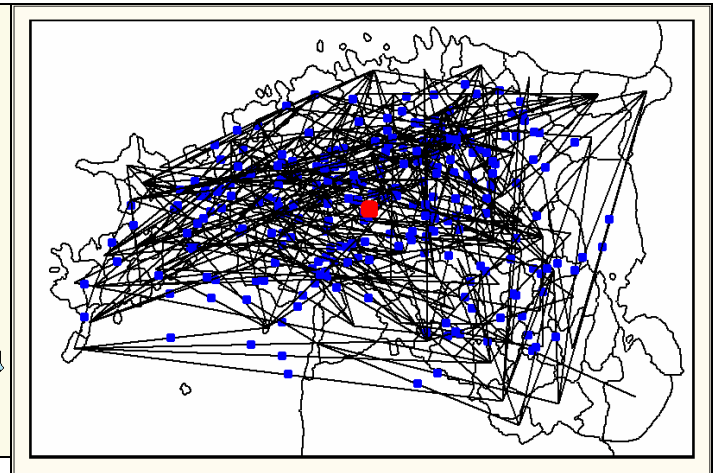
Murdesõnad: 1251–1500 seoste graaf



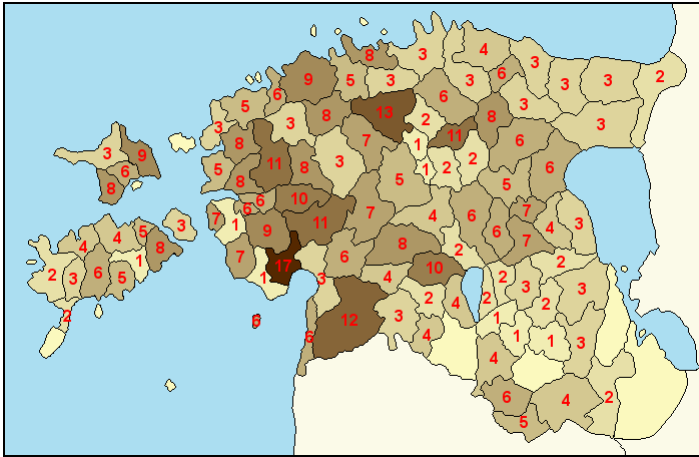
Murdesõnad: 1501–1750 toorkoefitsiendid



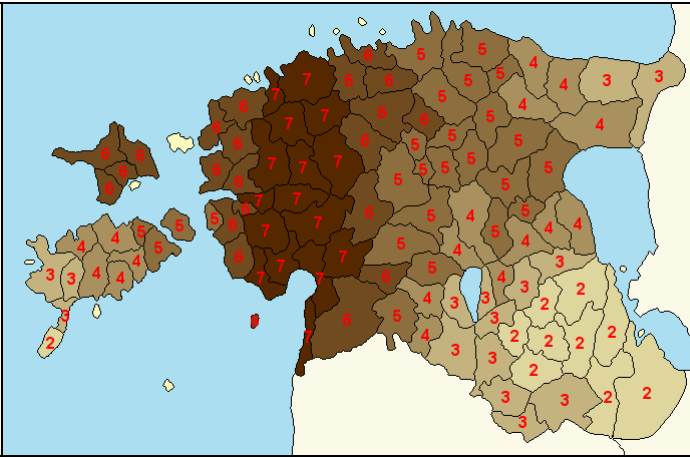
Murdesõnad: 1501–1750 silutud koefitsiendid



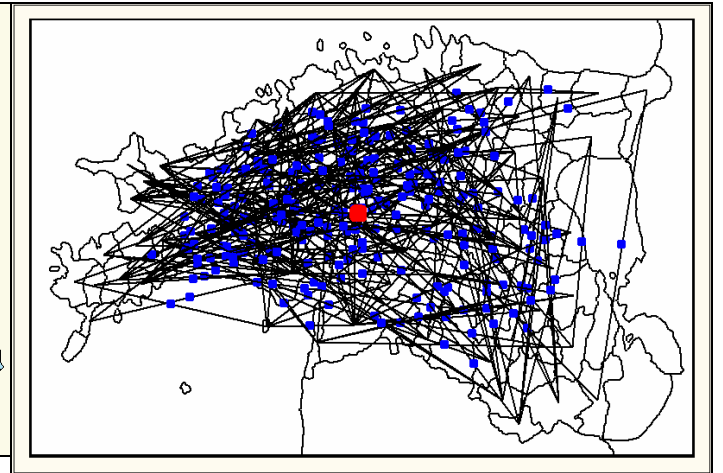
Murdesõnad: 1501–1750 seoste graaf



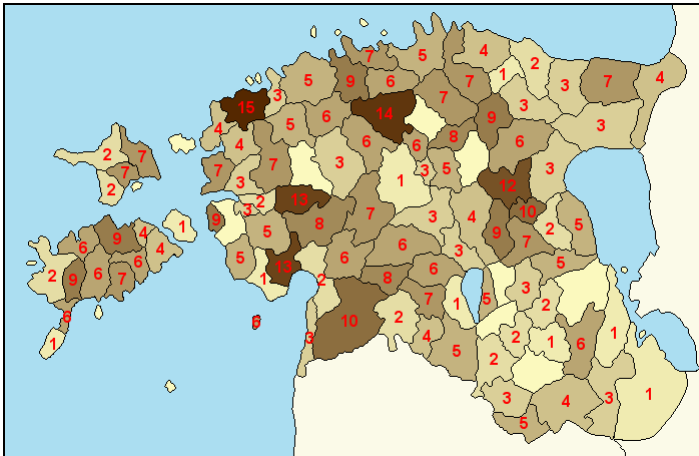
Murdesõnad: 1751–2000 toorkoefitsiendid



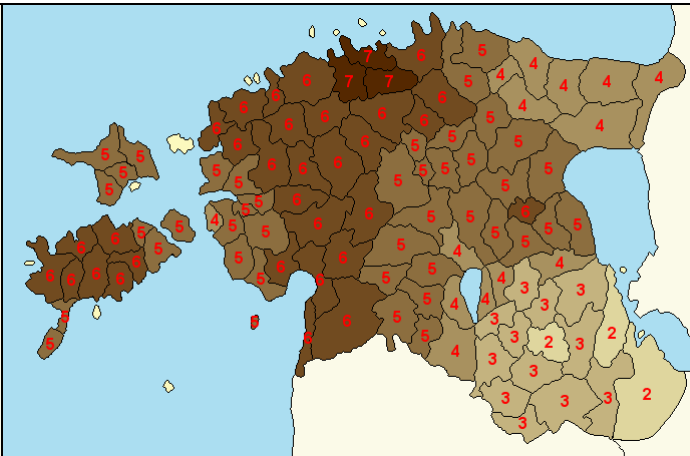
Murdesõnad: 1751–2000 silutud koefitsiendid



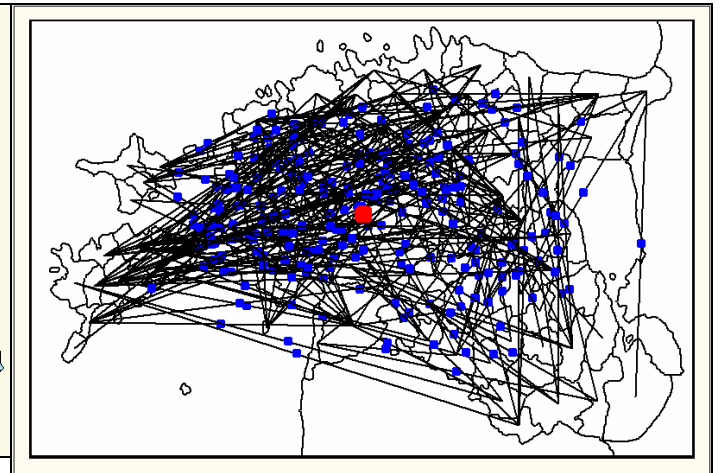
Murdesõnad: 1751–2000 seoste graaf



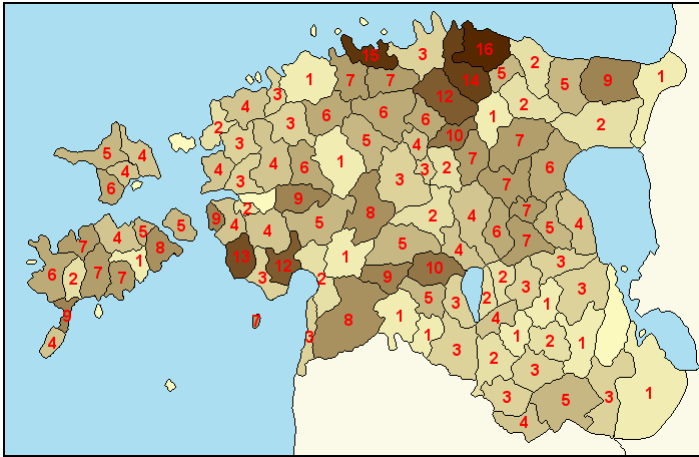
Murdesõnad: 2001–2250 toorkoefitsiendid



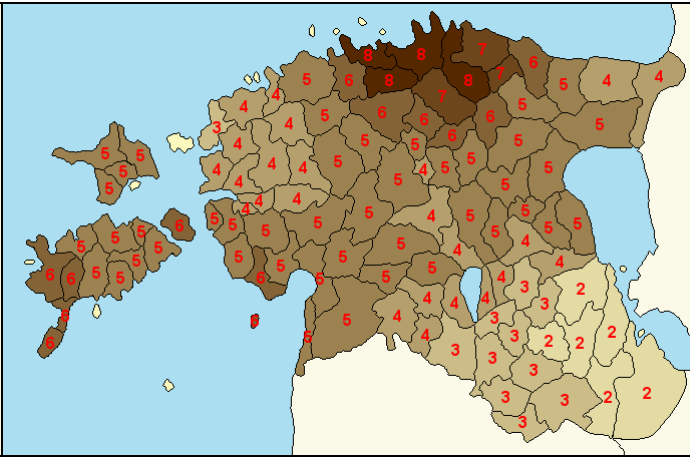
Murdesõnad: 2001–2250 silutud koefitsiendid



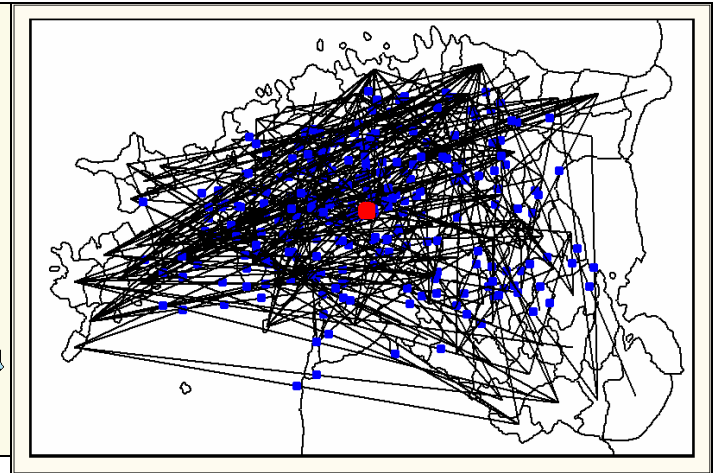
Murdesõnad: 2001–2250 seoste graaf



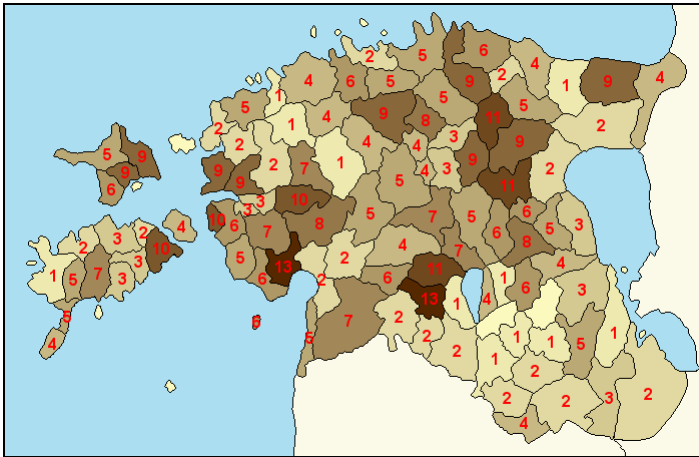
Murdesõnad: 2251–2500 toorkoefitsiendid



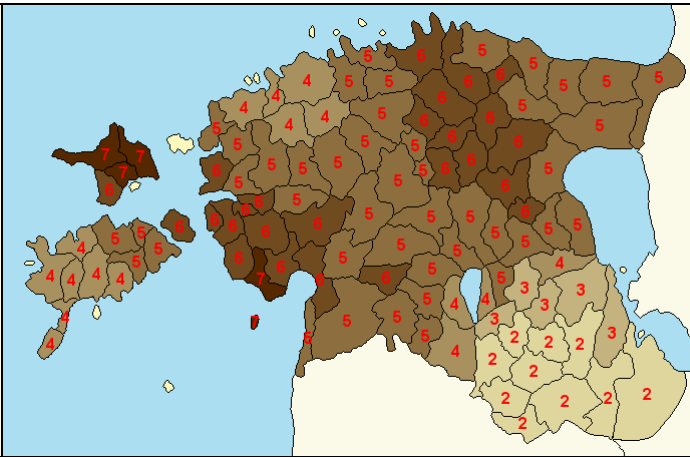
Murdesõnad: 2251–2500 silutud koefitsiendid



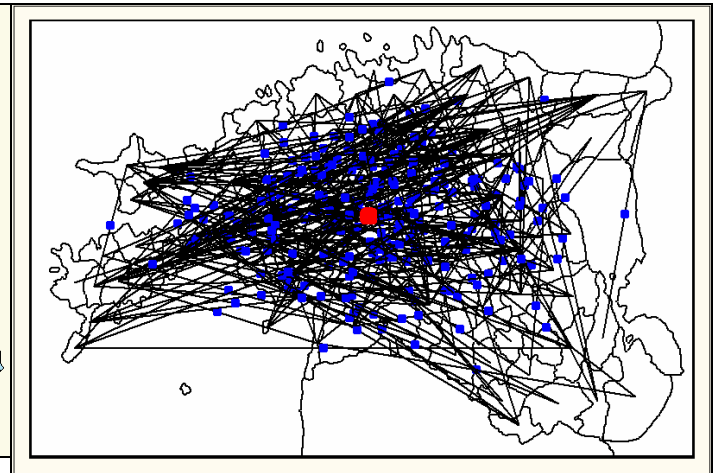
Murdesõnad: 2251–2500 seoste graaf



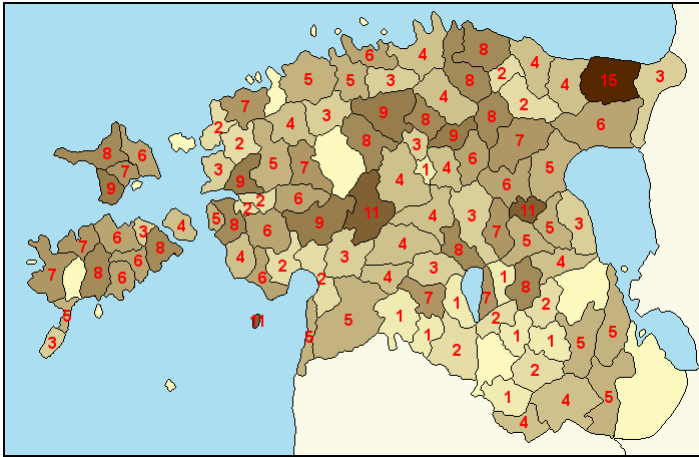
Murdesõnad: 2501–2750 toorkoefitsiendid



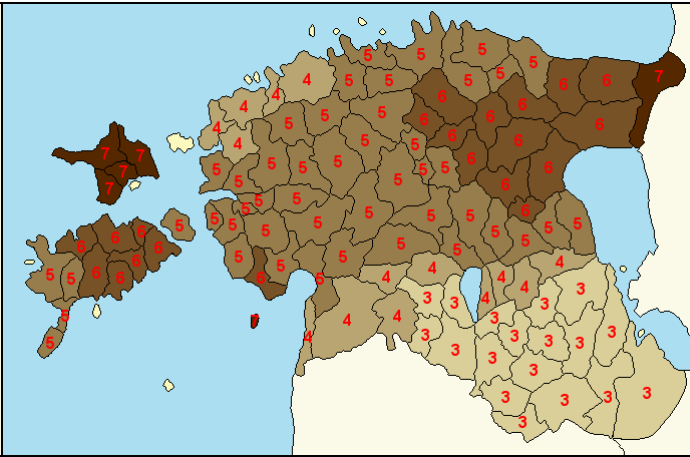
Murdesõnad: 2501–2750 silutud koefitsiendid



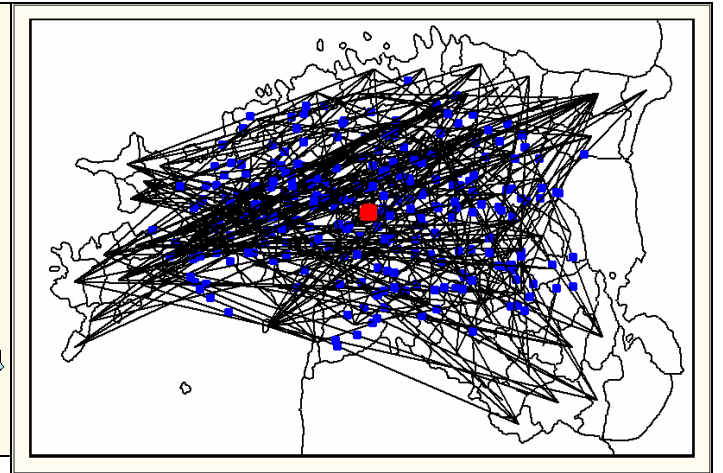
Murdesõnad: 2501–2750 seoste graaf



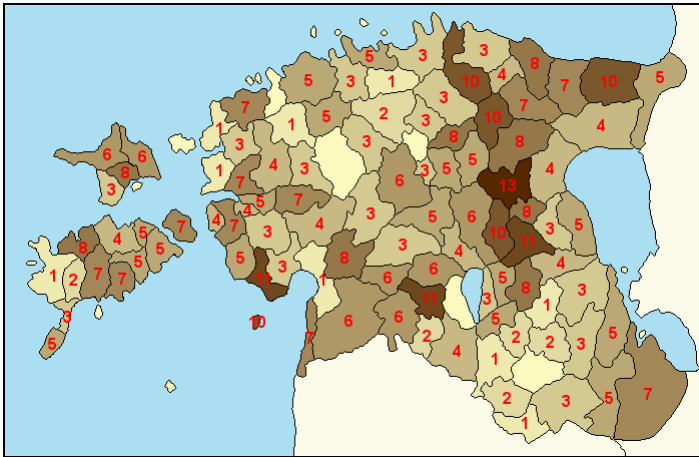
Murdesõnad: 2751–3000 toorkoefitsiendid



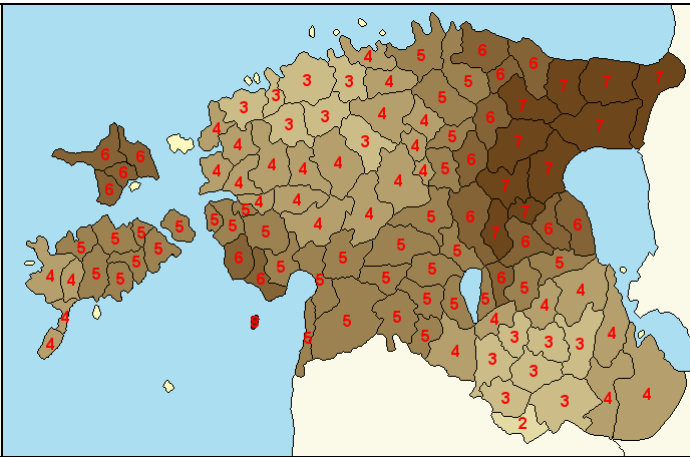
Murdesõnad: 2751–3000 silutud koefitsiendid



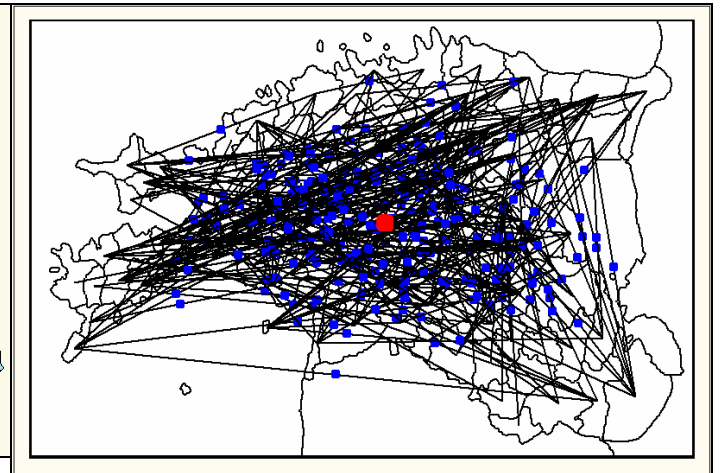
Murdesõnad: 2751–3000 seoste graaf



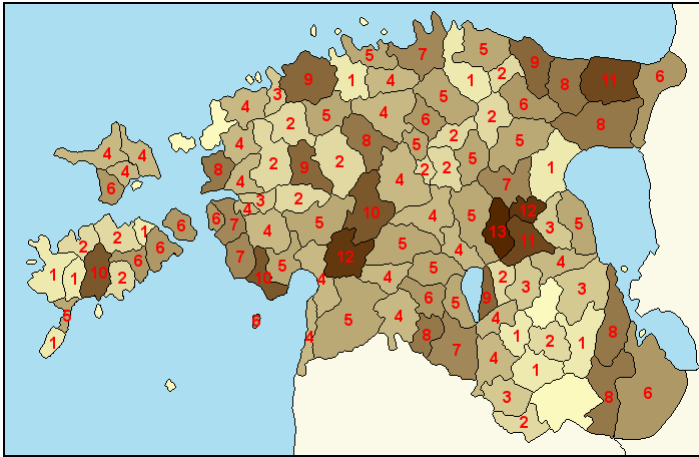
Murdesõnad: 3001–3250 toorkoefitsiendid



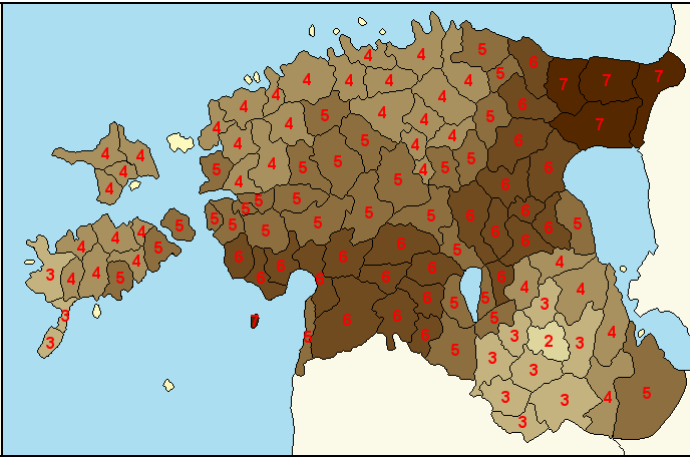
Murdesõnad: 3001–3250 silutud koefitsiendid



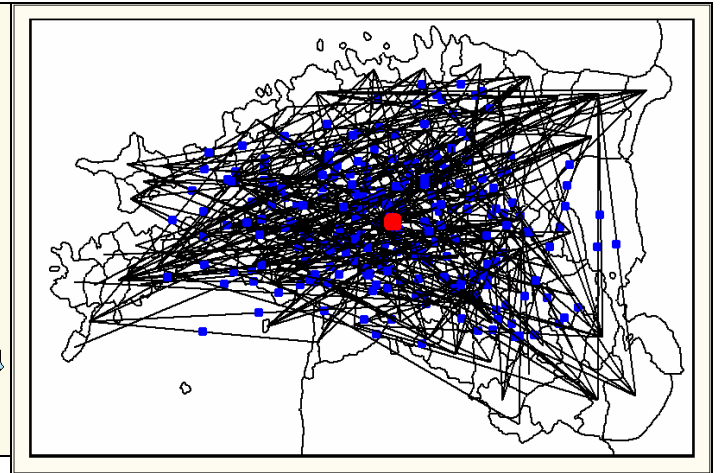
Murdesõnad: 3001–3250 seoste graaf



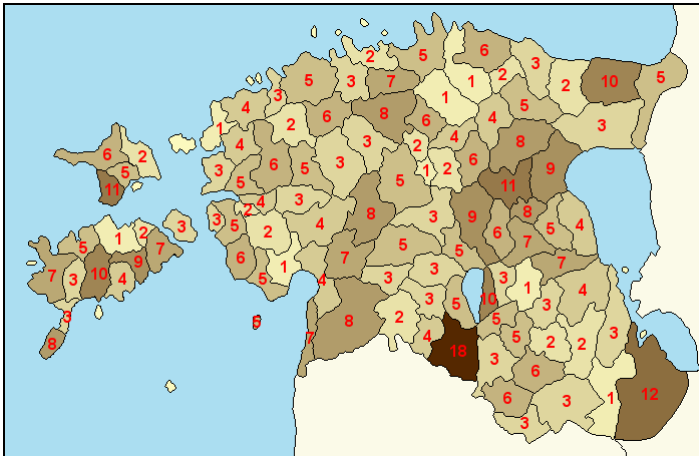
Murdesõnad: 3251–3500 toorkoefitsiidid



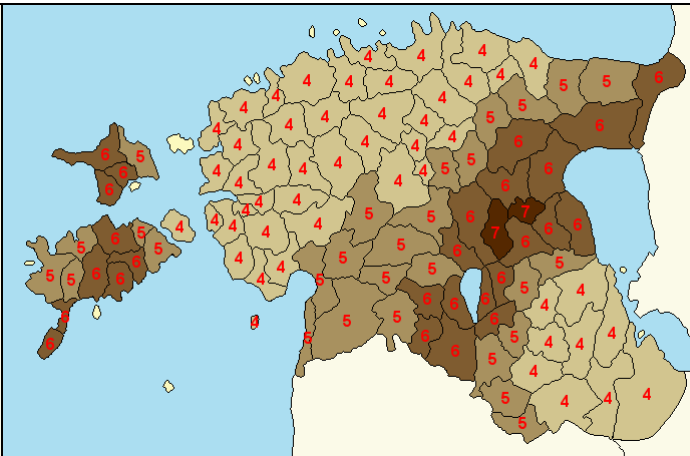
Murdesõnad: 3251–3500 silutud koefitsiidid



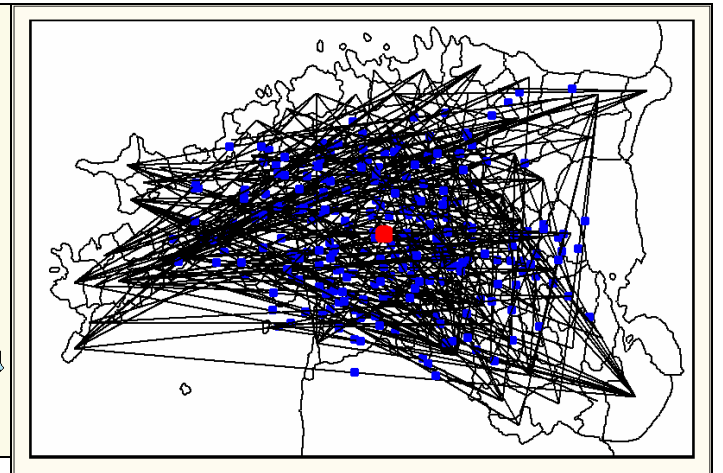
Murdesõnad: 3251–3500 seoste graaf



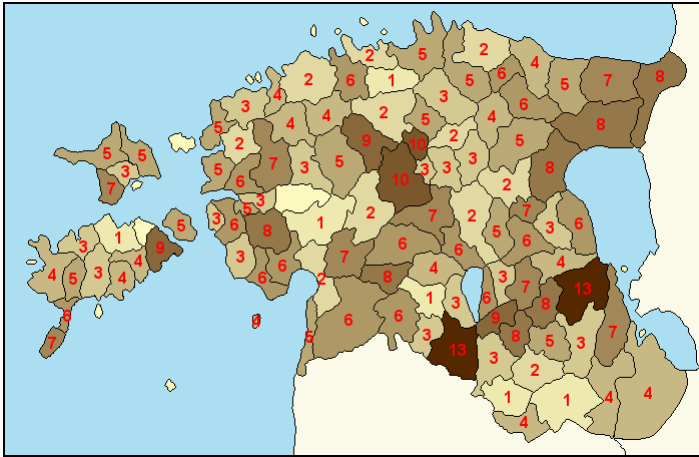
Murdesõnad: 3501–3750 toorkoefitsiidid



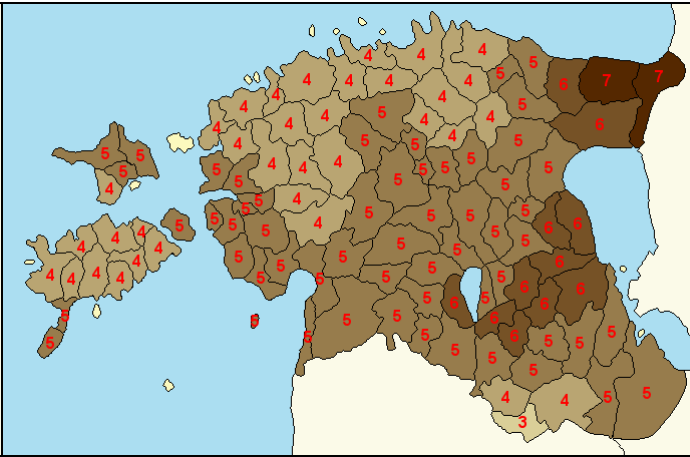
Murdesõnad: 3501–3750 silutud koefitsiidid



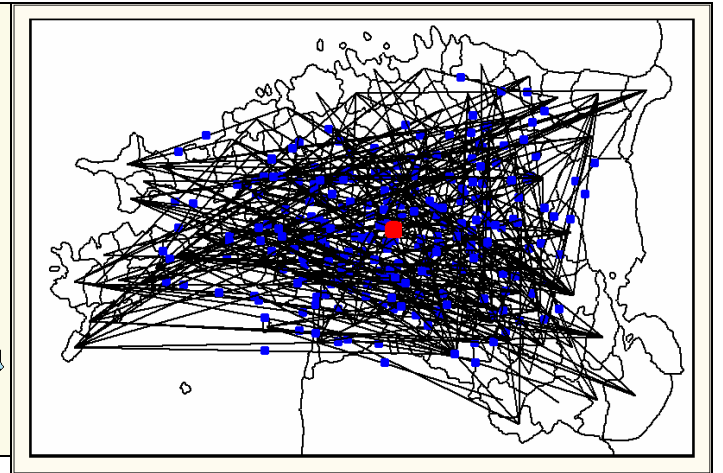
Murdesõnad: 3501–3750 seoste graaf



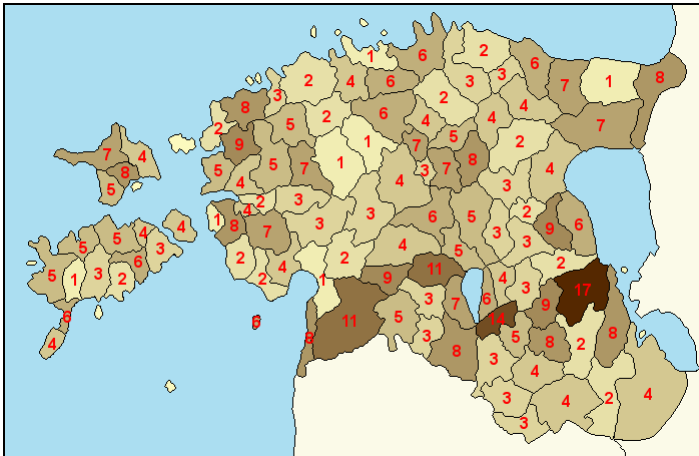
Murdesõnad: 3751–4000 toorkoefitsiidid



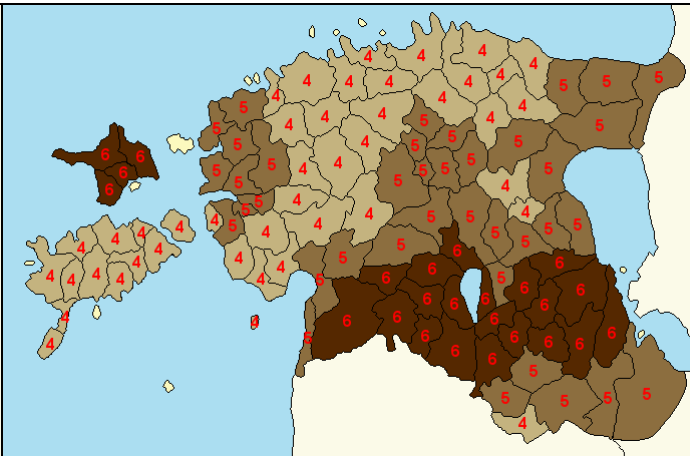
Murdesõnad: 3751–4000 silutud koefitsiidid



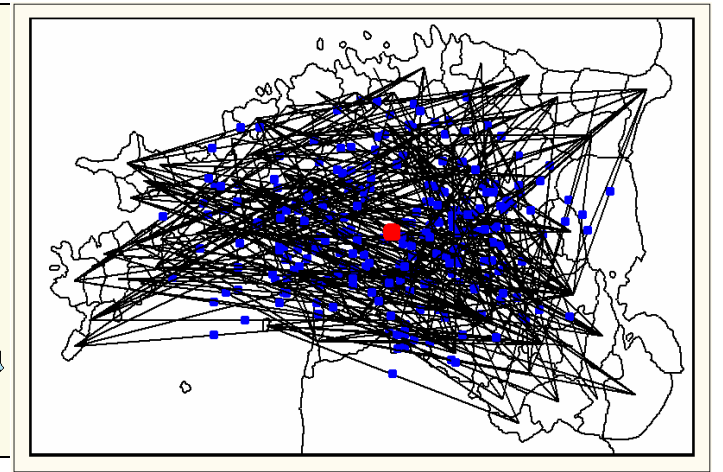
Murdesõnad: 3751–4000 seoste graaf



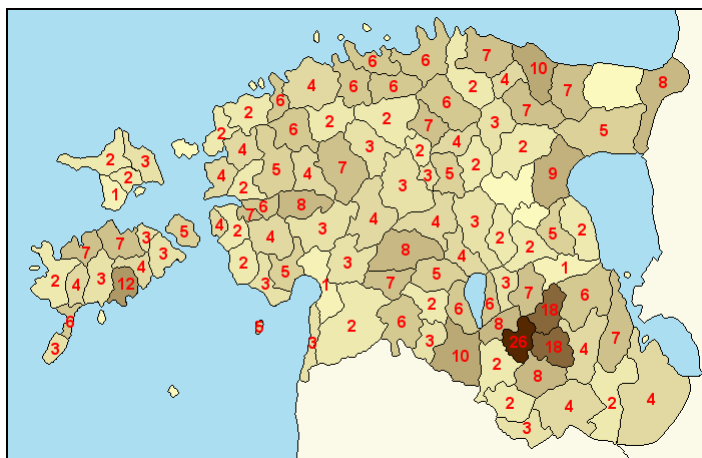
Murdesõnad: 4001–4250 toorkoefitsiidid



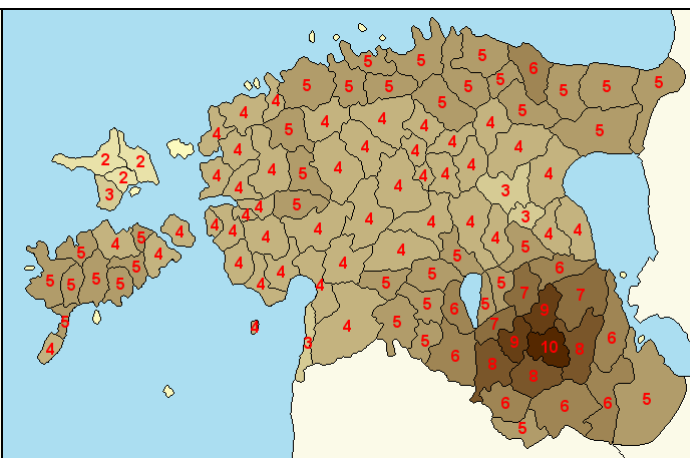
Murdesõnad: 4001–4250 silutud koefitsiidid



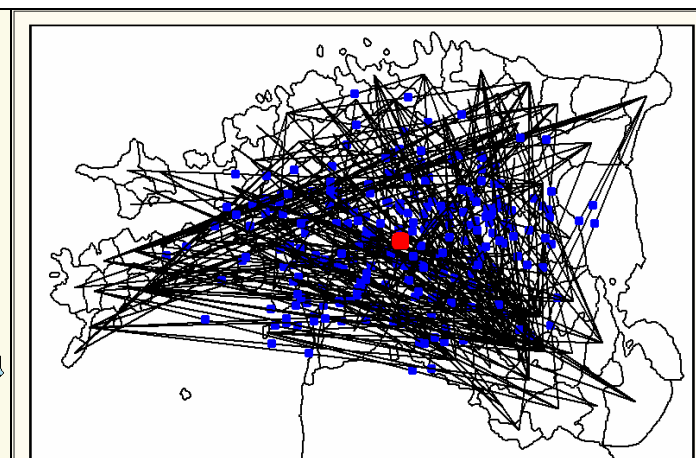
Murdesõnad: 4001–4250 seoste graaf



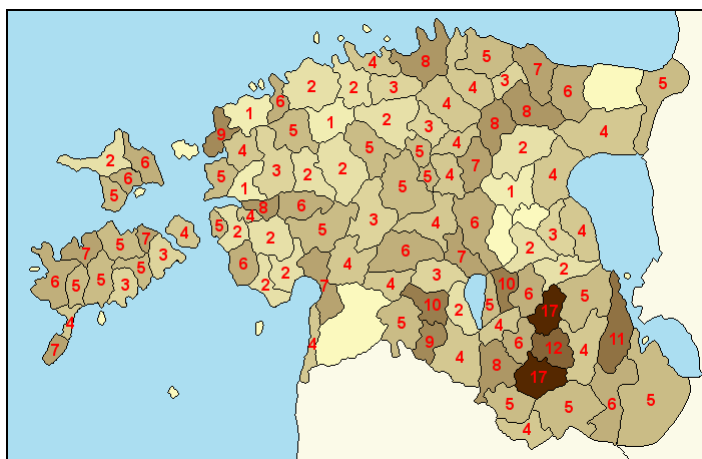
Murdesõnad: 4251–4500 toorkoefitsiidid



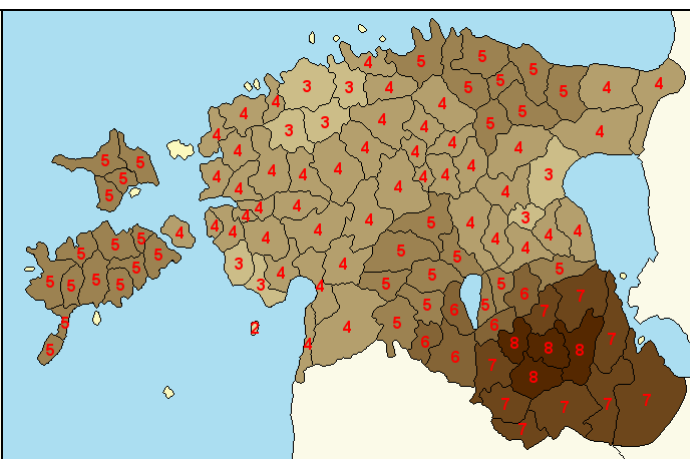
Murdesõnad: 4251–4500 silutud koefitsiidid



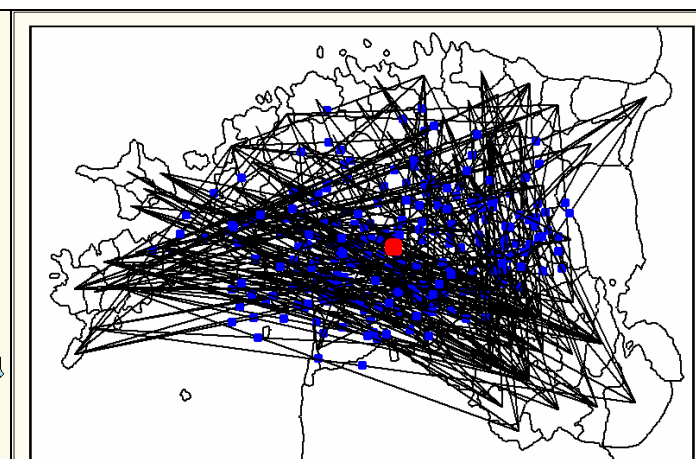
Murdesõnad: 4251–4500 seoste graaf



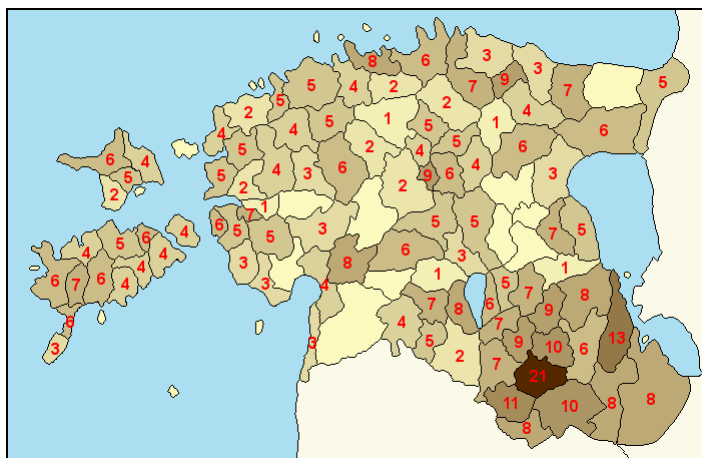
Murdesõnad: 4501–4750 toorkoefitsiidid



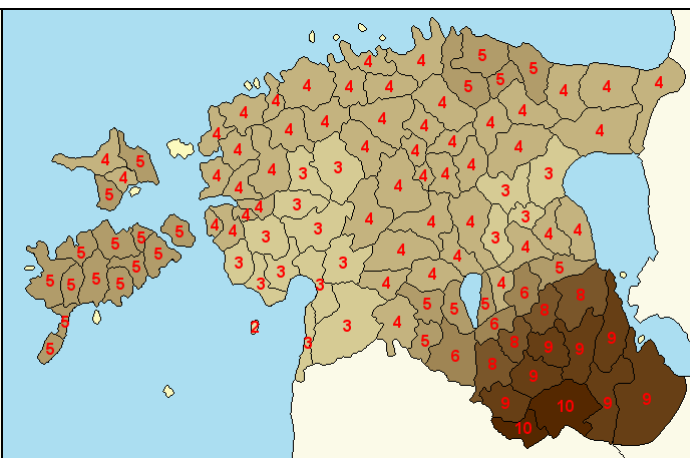
Murdesõnad: 4501–4750 silutud koefitsiidid



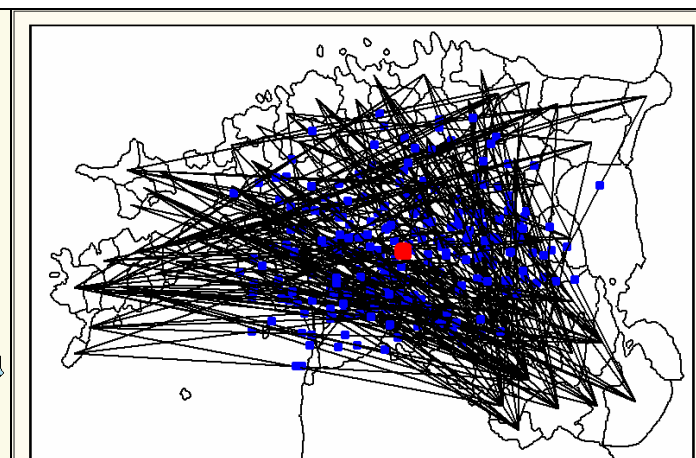
Murdesõnad: 4501–4750 seoste graaf



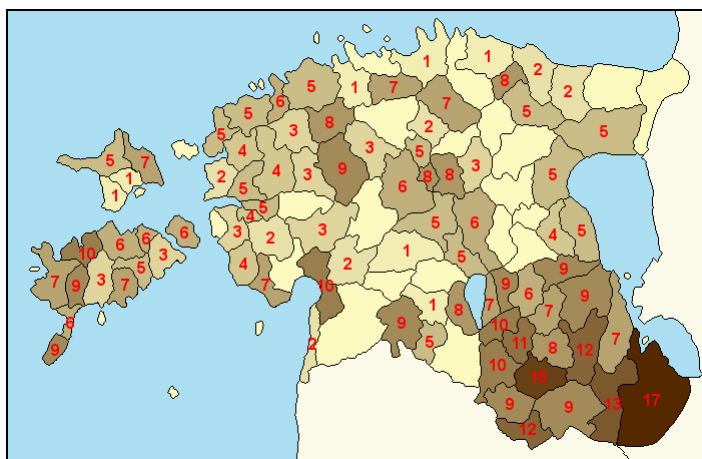
Murdesõnad: 4751–5000 toorkoefitsiidid



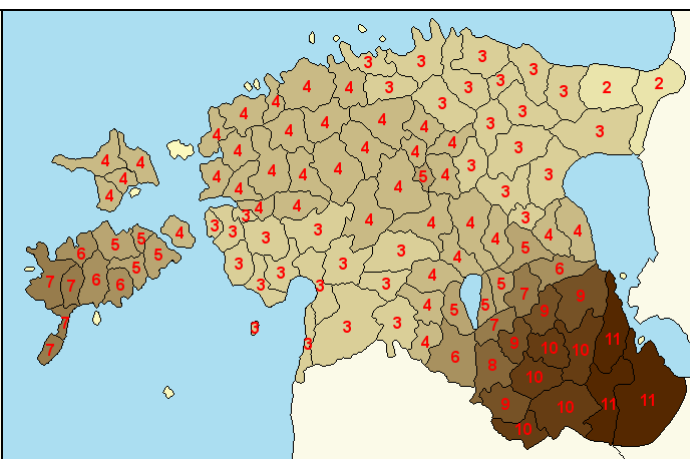
Murdesõnad: 4751–5000 silutud koefitsiidid



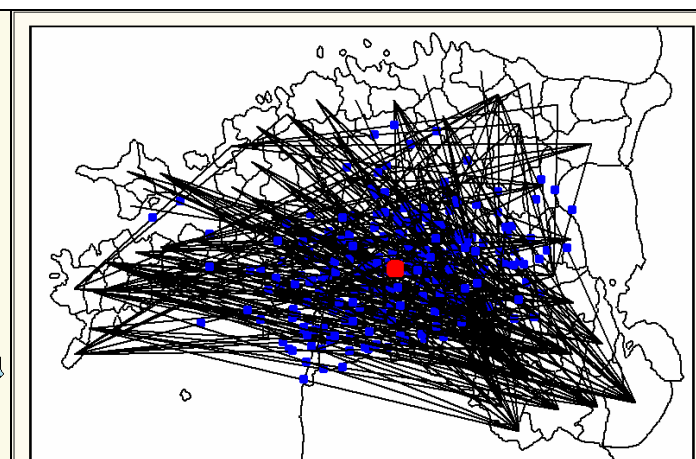
Murdesõnad: 4751–5000 seoste graaf



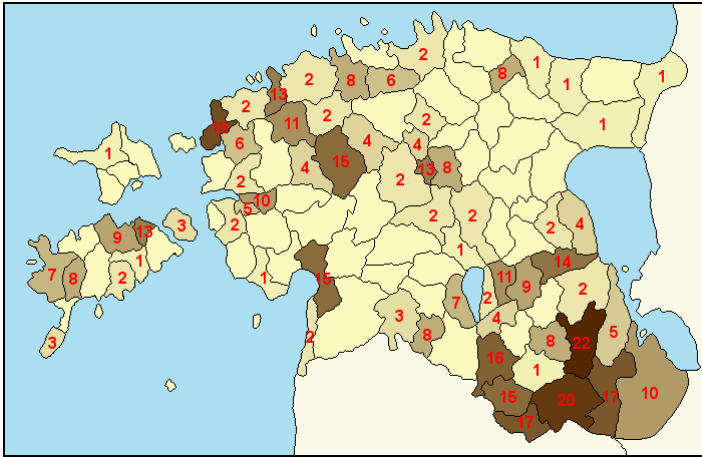
Murdesõnad: 5001–5250 toorkoefitsiidid



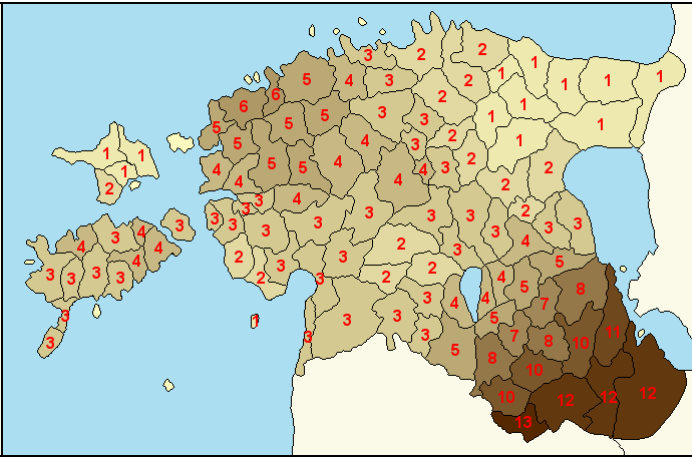
Murdesõnad: 5001–5250 silutud koefitsiidid



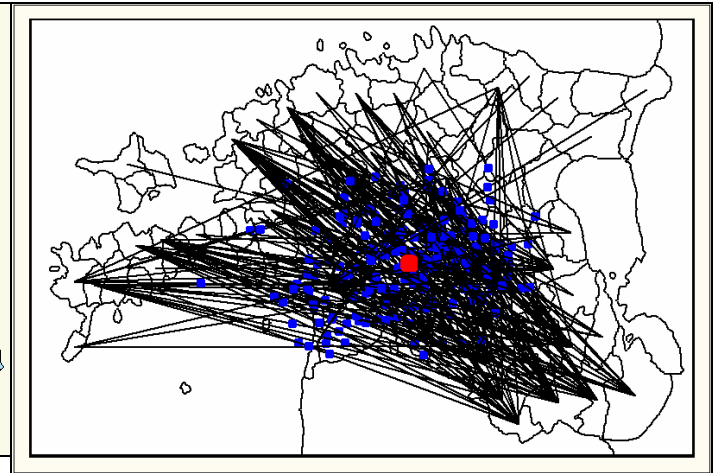
Murdesõnad: 5001–5250 seoste graaf



Murdesõnad: 5251–5460 toorkoefitsiendid

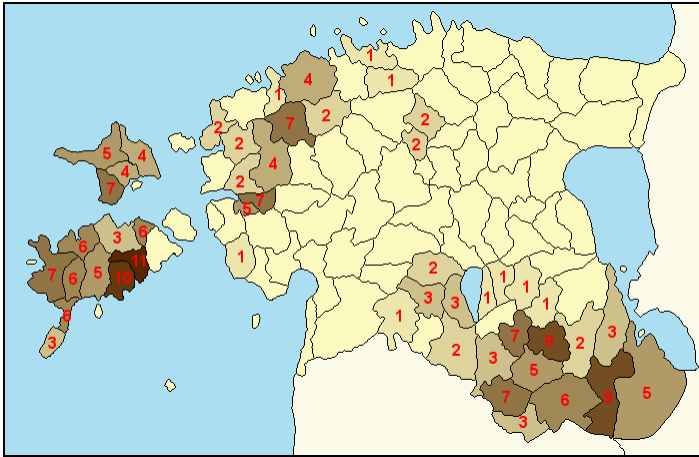


Murdesõnad: 5251–5460 silutud koefitsiendid

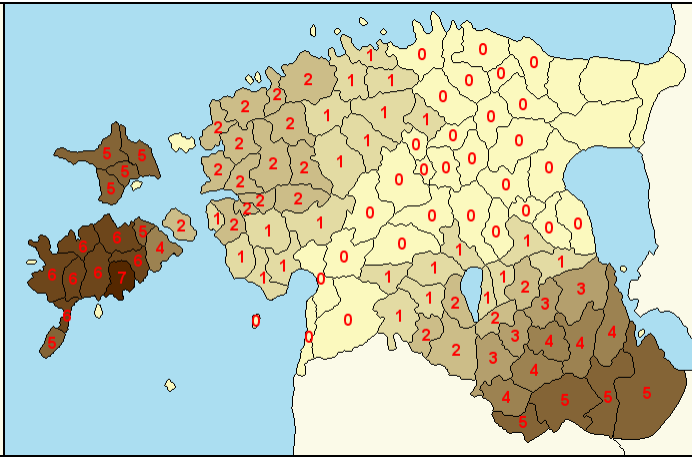


Murdesõnad: 5251–5460 seoste graaf

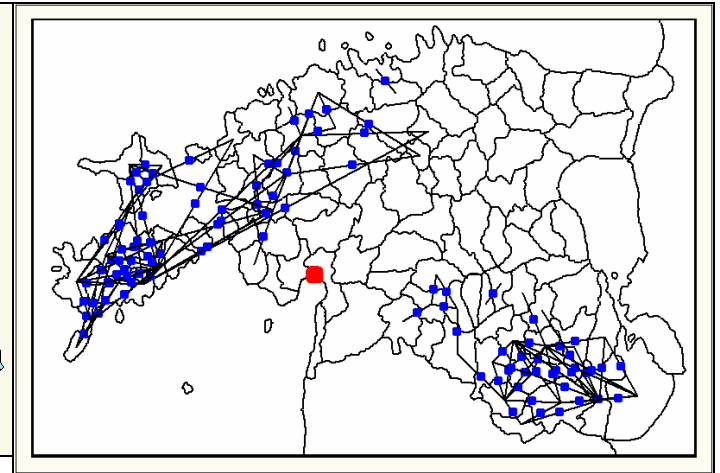
Mõistatud



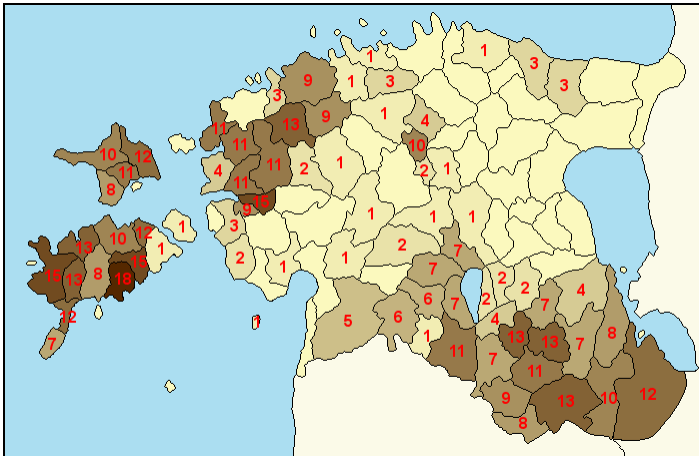
Mõistatud: 1–100 toorkoefitsiendid



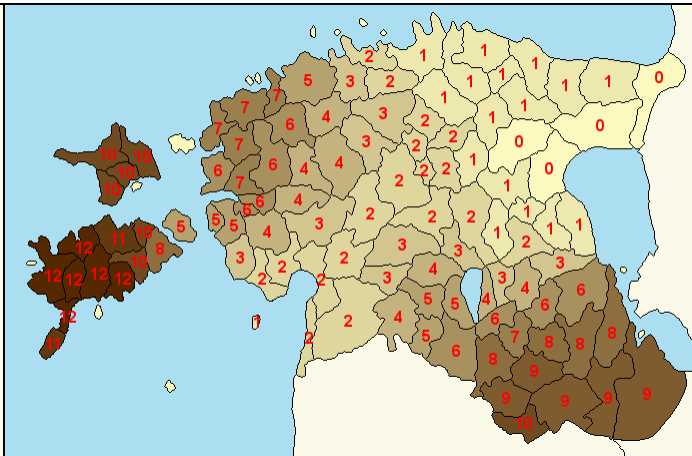
Mõistatud: 1–100 silutud koefitsiendid



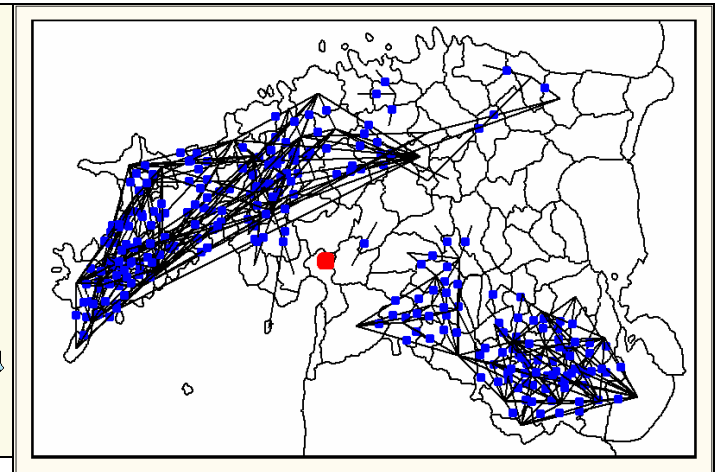
Mõistatud: 1–100 seoste graaf



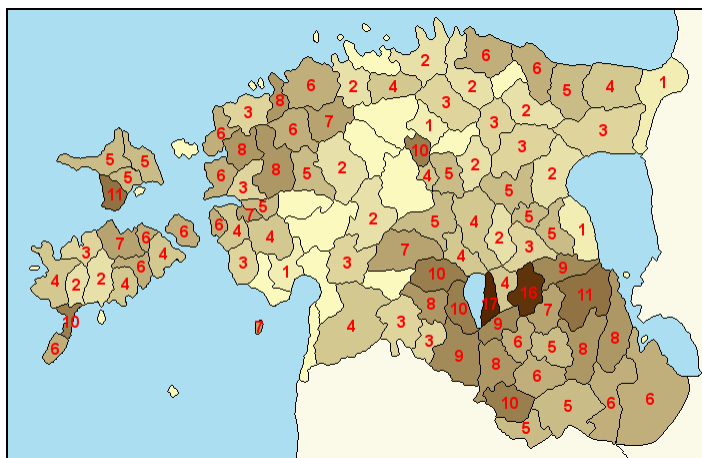
Mõistatud: 1–250 toorkoefitsiendid



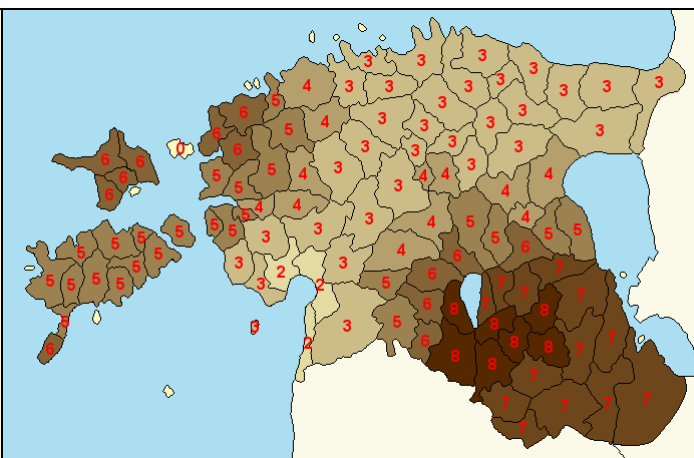
Mõistatud: 1–250 silutud koefitsiendid



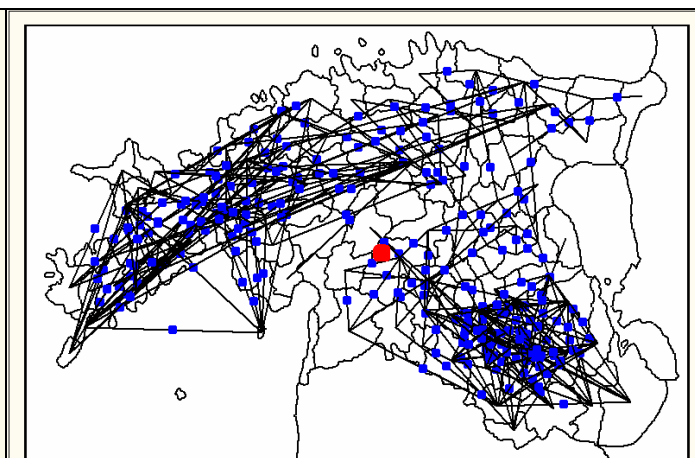
Mõistatud: 1–250 seoste graaf



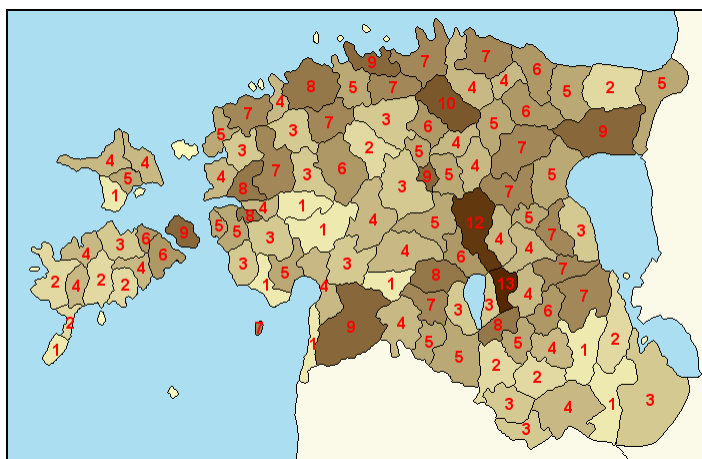
Mõistatused: 251–500 toorkoefitsiendid



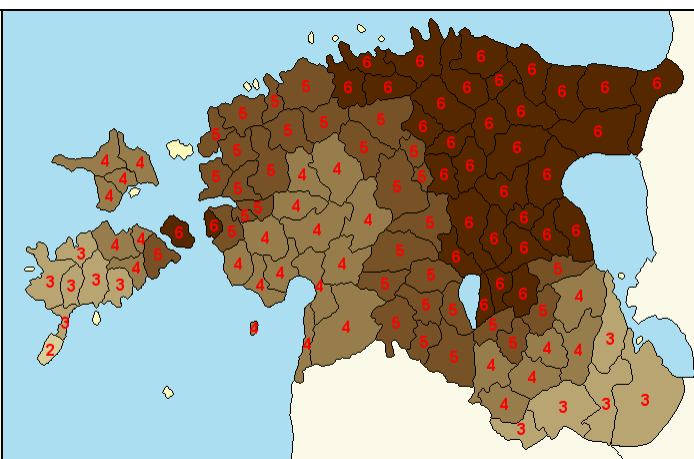
Mõistatused: 251–500 silutud koefitsiendid



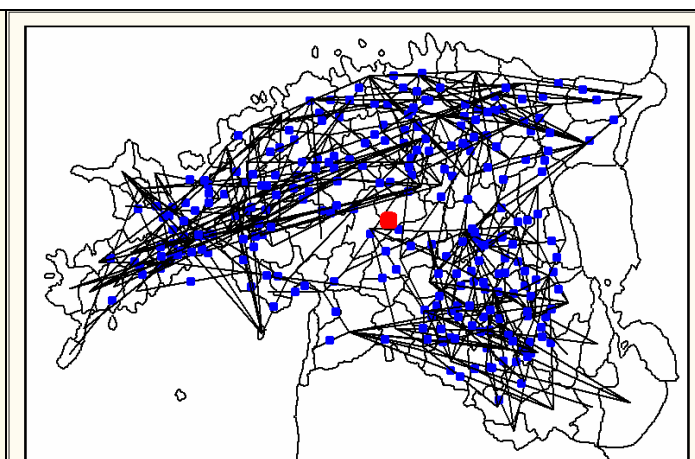
Mõistatused: 251–500 seoste graaf



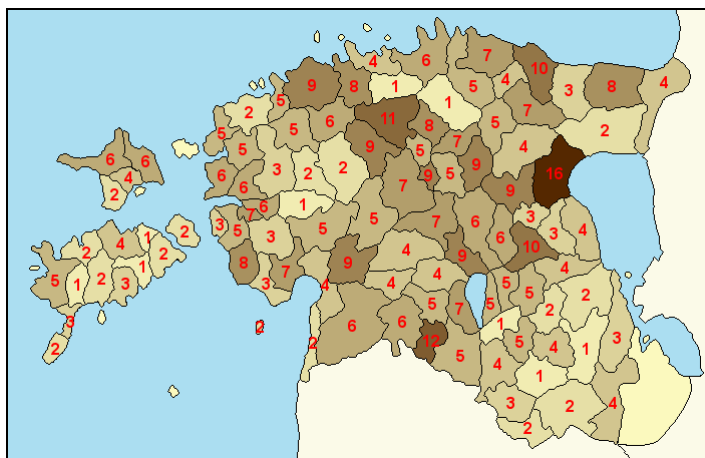
Mõistatused: 501–750 toorkoefitsiendid



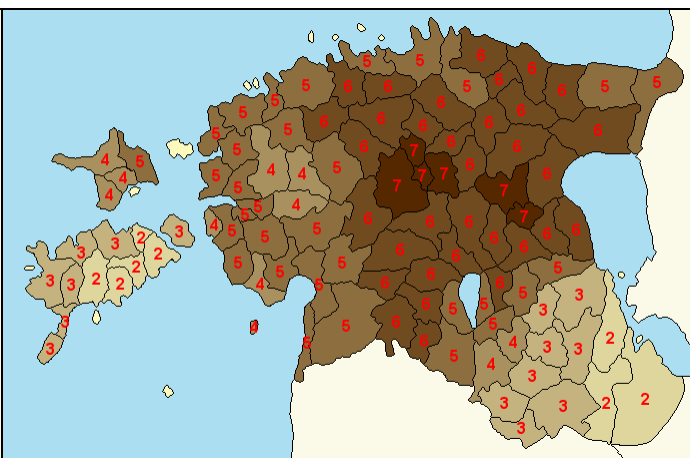
Mõistatused: 501–750 silutud koefitsiendid



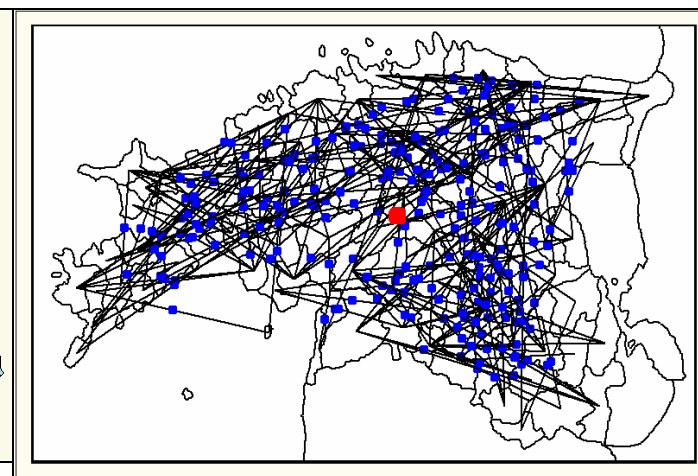
Mõistatused: 501–750 seoste graaf



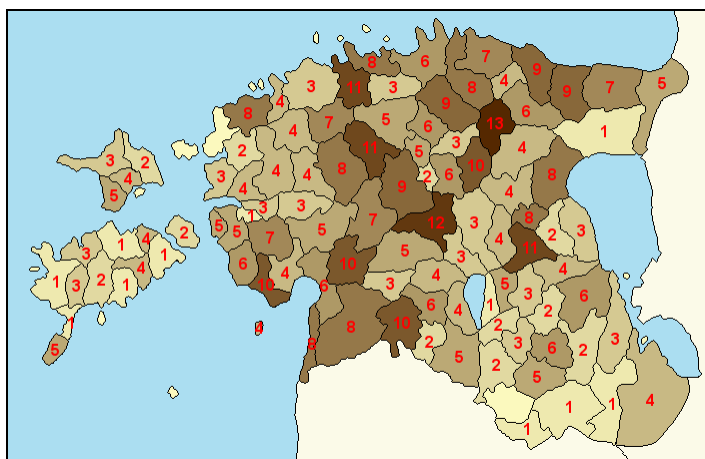
Mõistatud: 751–1000 toorkoefitsiendid



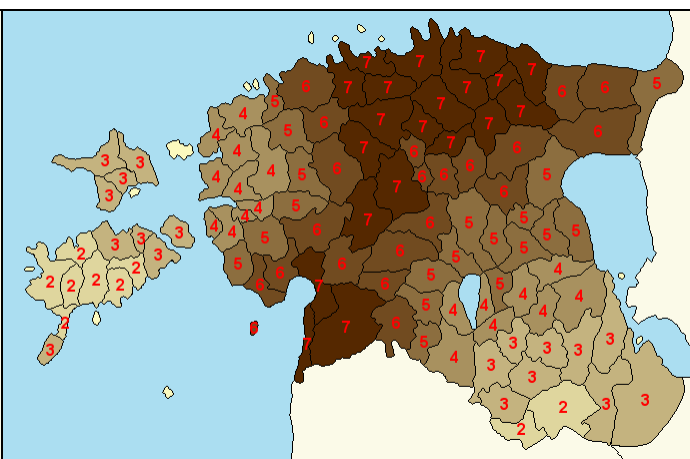
Mõistatud: 751–1000 silutud koefitsiendid



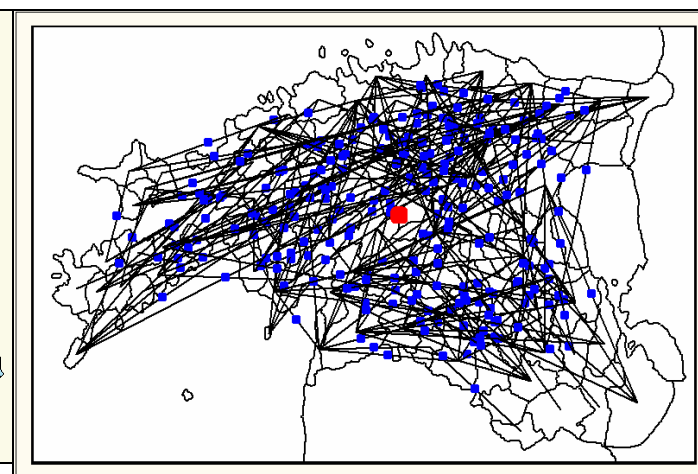
Mõistatud: 751–1000 seoste graaf



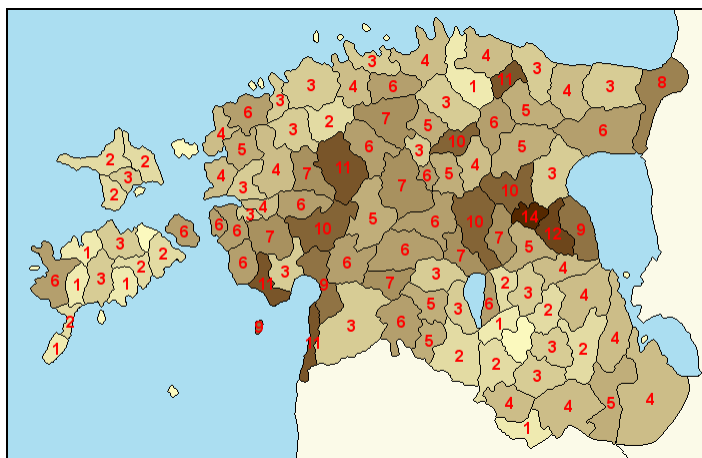
Mõistatud: 1001–1250 toorkoefitsiendid



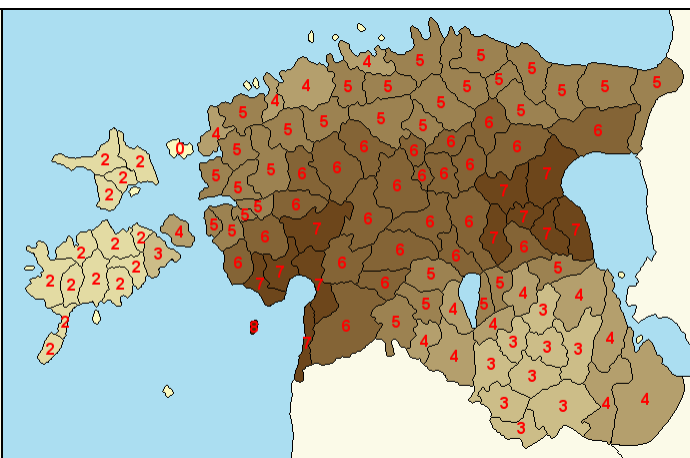
Mõistatud: 1001–1250 silutud koefitsiendid



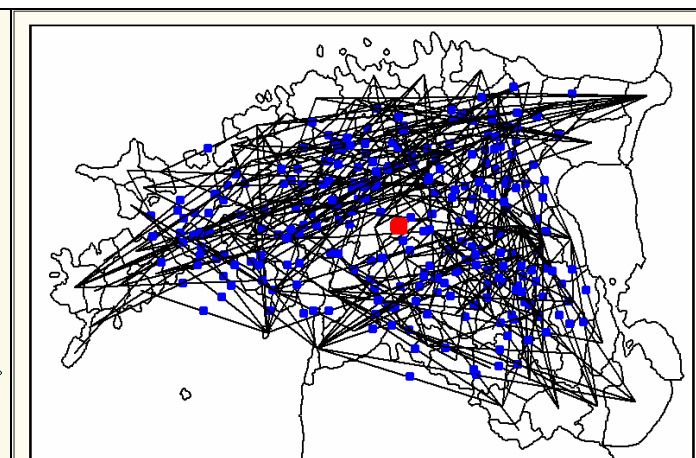
Mõistatud: 1001–1250 seoste graaf



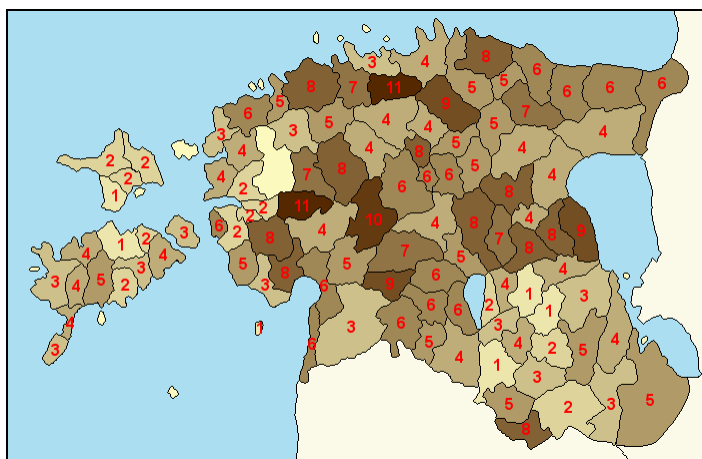
Mõistatud: 1251–1500 toorkoefitsiendid



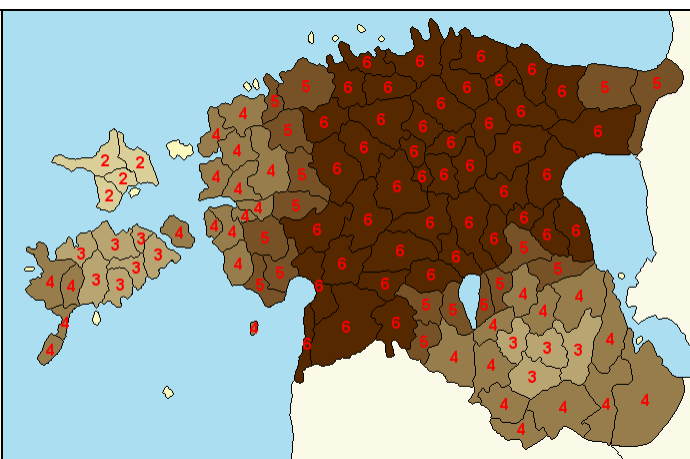
Mõistatud: 1251–1500 silutud koefitsiendid



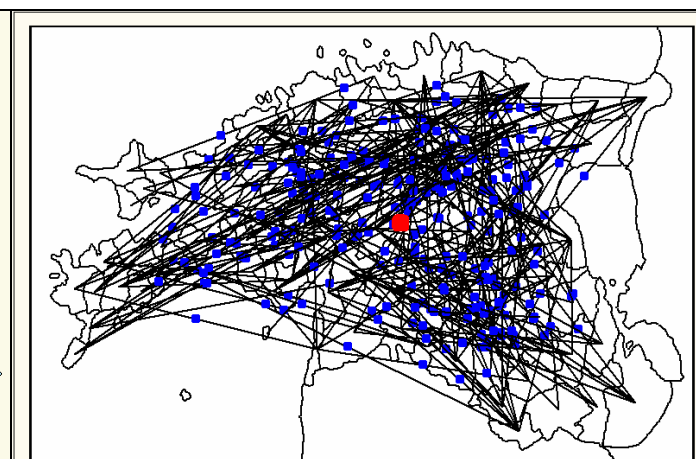
Mõistatud: 1251–1500 seoste graaf



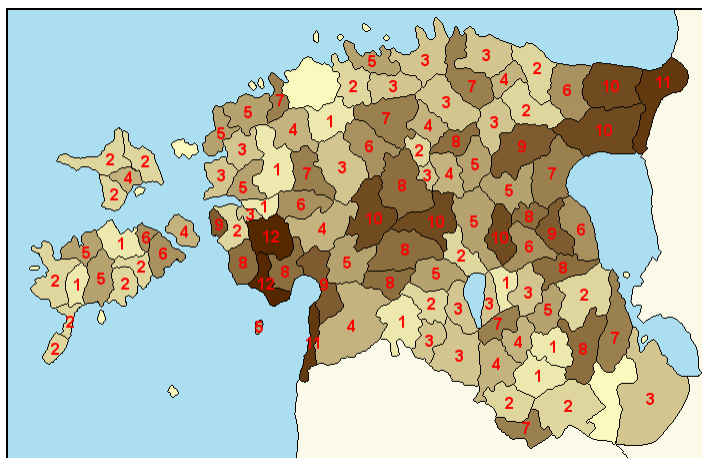
Mõistatud: 1501–1750 toorkoefitsiendid



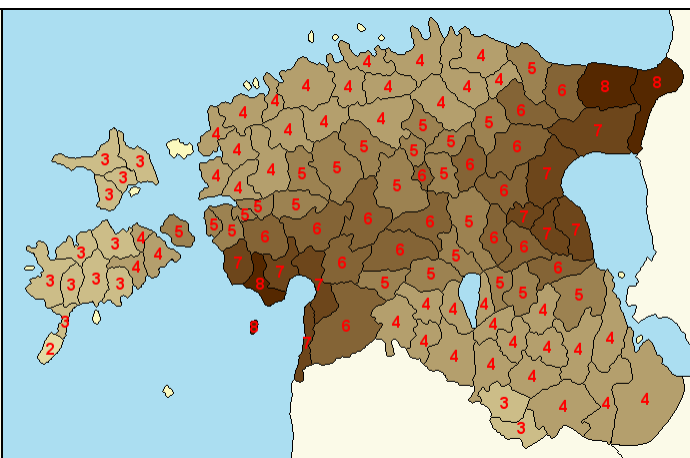
Mõistatud: 1501–1750 silutud koefitsiendid



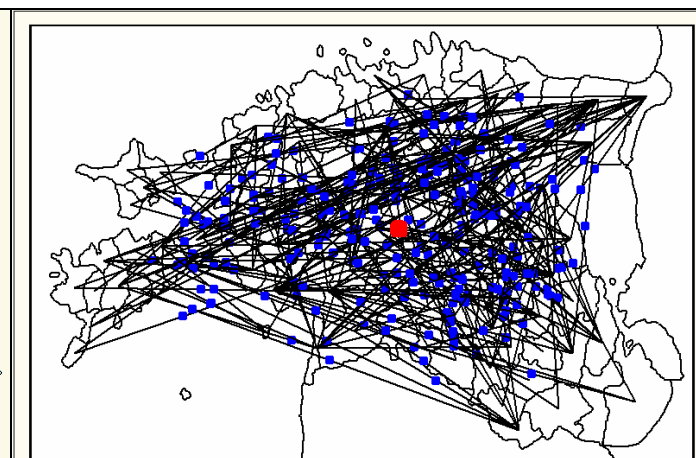
Mõistatud: 1501–1750 seoste graaf



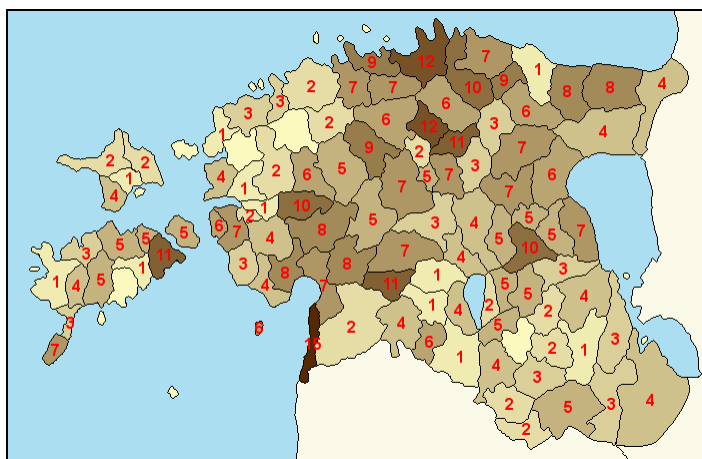
Mõistatused: 1751–2000 toorkoefitsiendid



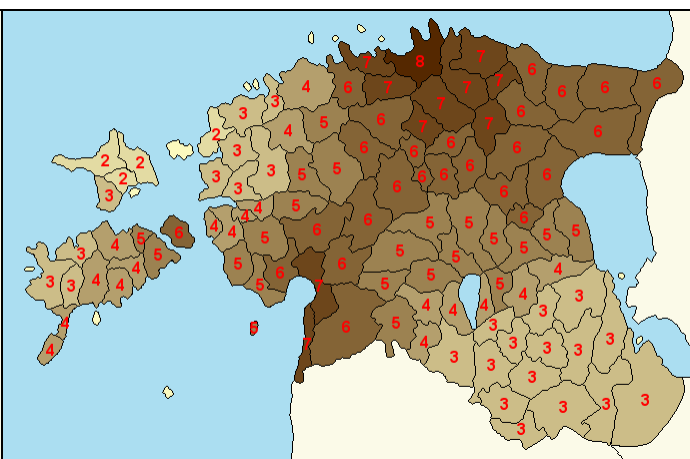
Mõistatused: 1751–2000 silutud koefitsiendid



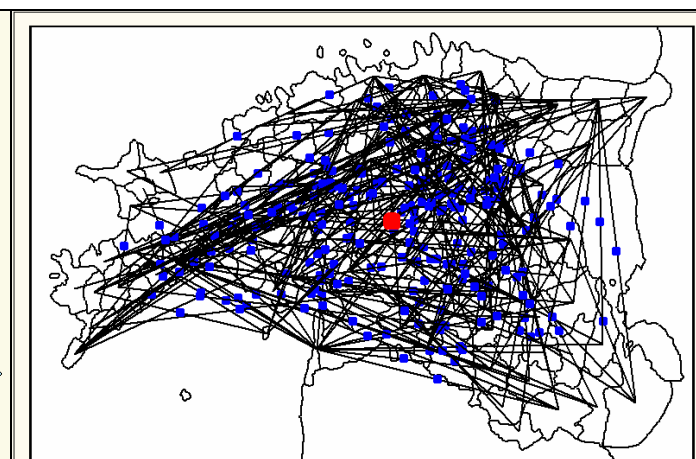
Mõistatused: 1751–2000 seoste graaf



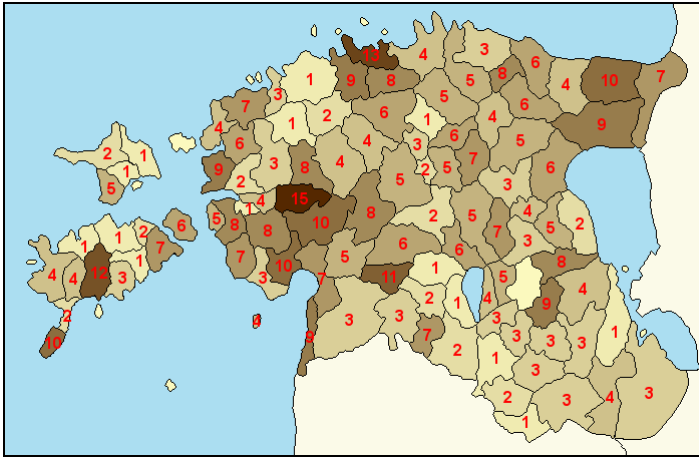
Mõistatused: 2001–2250 toorkoefitsiendid



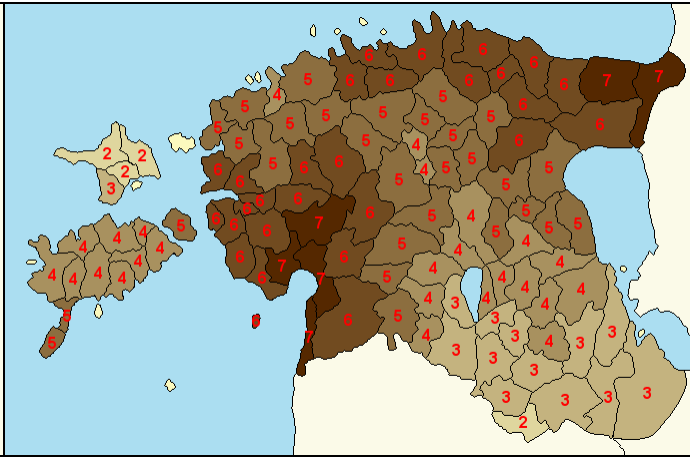
Mõistatused: 2001–2250 silutud koefitsiendid



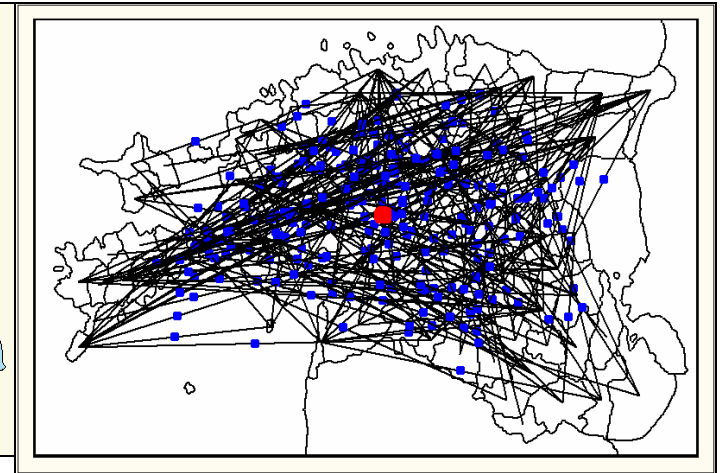
Mõistatused: 2001–2250 seoste graaf



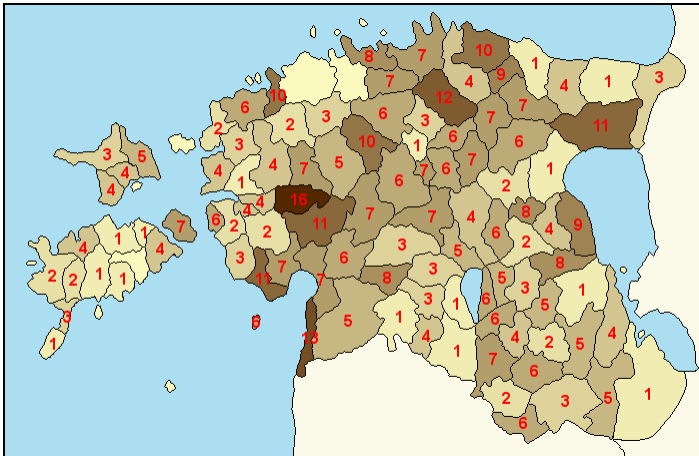
Mõistatused: 2251–2500 toorkoefitsiendid



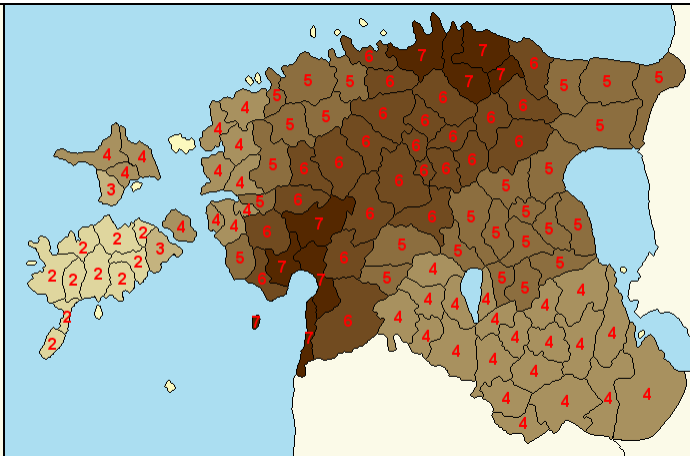
Mõistatused: 2251–2500 silutud koefitsiendid



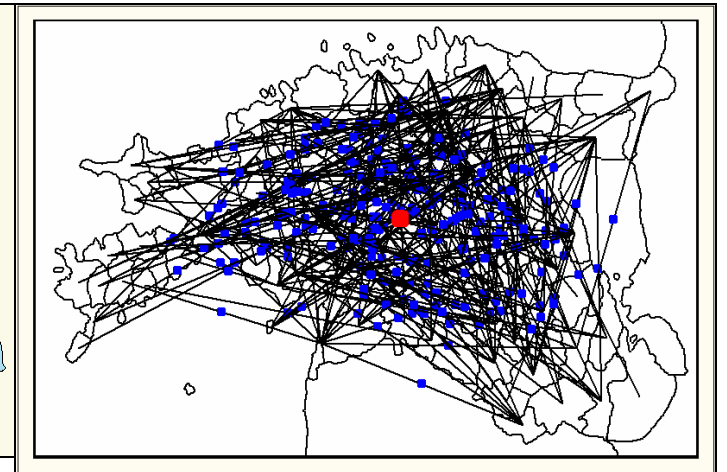
Mõistatused: 2251–2500 seoste graaf



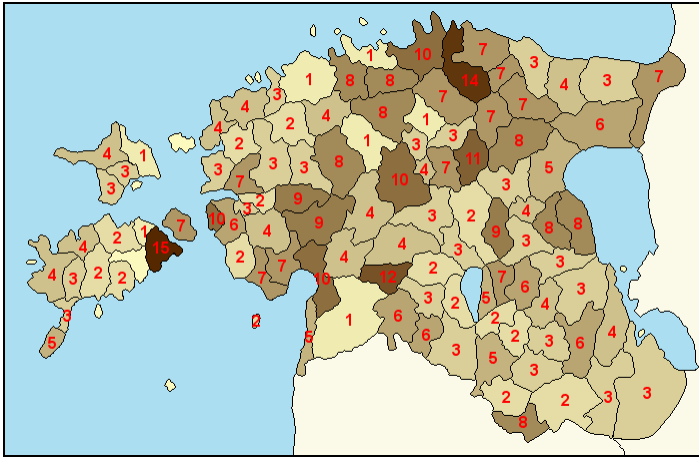
Mõistatused: 2501–2750 toorkoefitsiendid



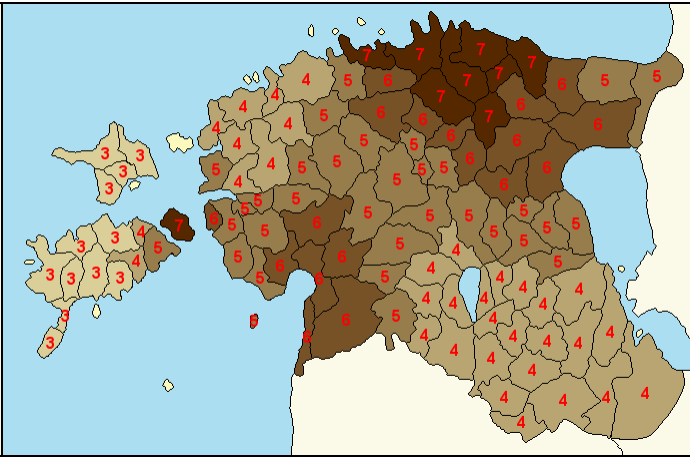
Mõistatused: 2501–2750 silutud koefitsiendid



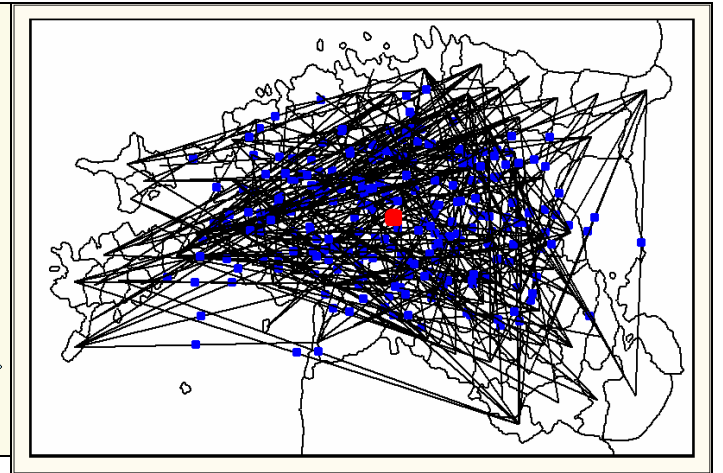
Mõistatused: 2501–2750 seoste graaf



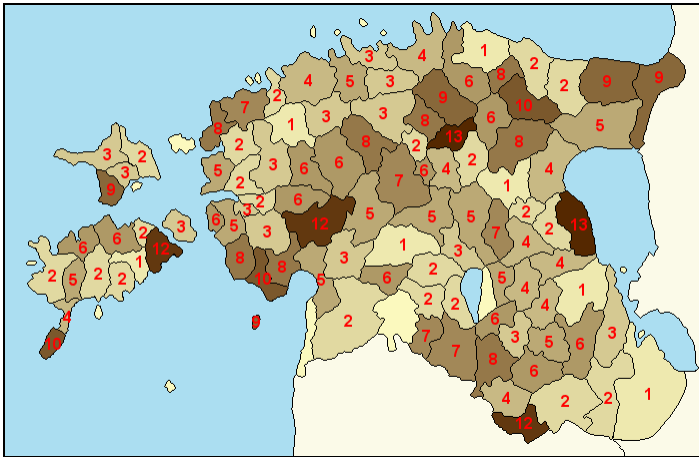
Mõistatud: 2751–3000 toorkoefitsiendid



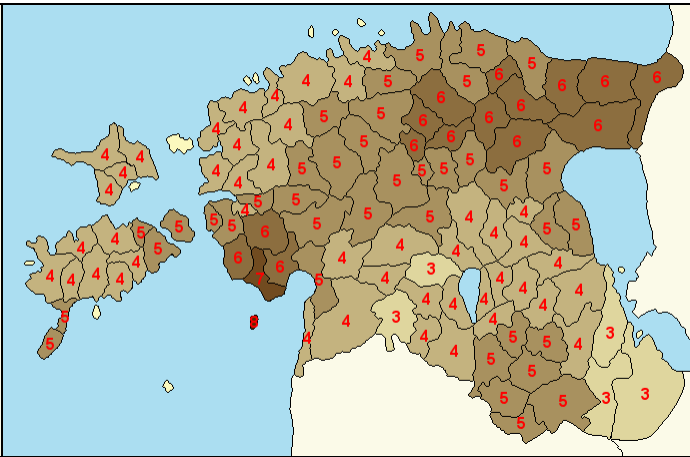
Mõistatud: 2751–3000 silutud koefitsiendid



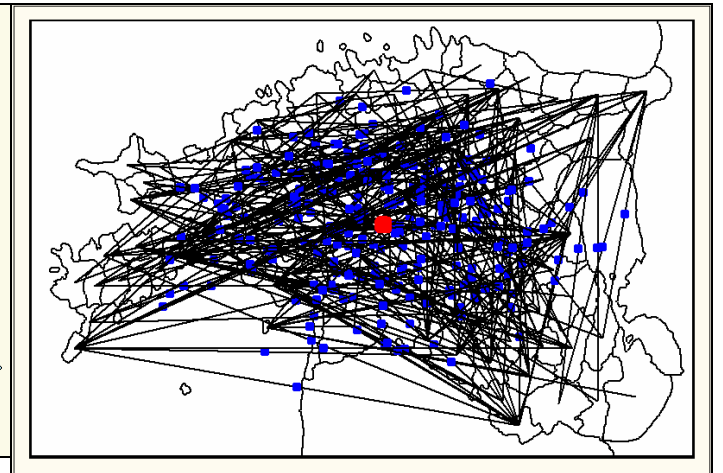
Mõistatud: 2751–3000 seoste graaf



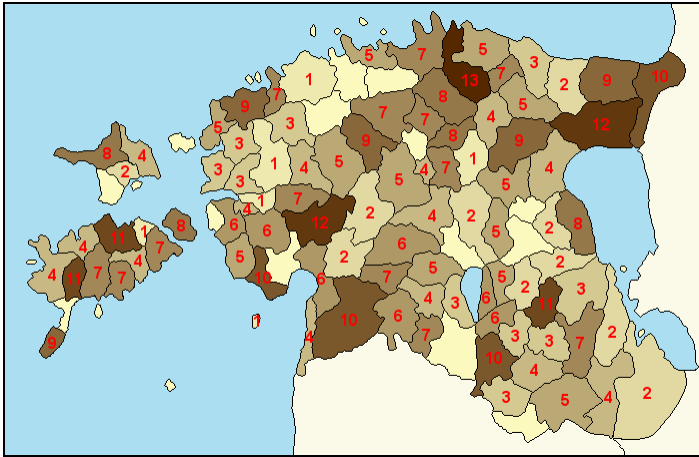
Mõistatud: 3001–3250 toorkoefitsiendid



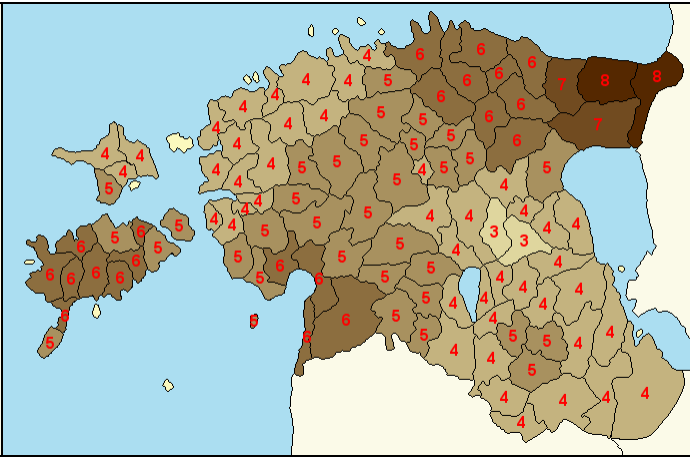
Mõistatud: 3001–3250 silutud koefitsiendid



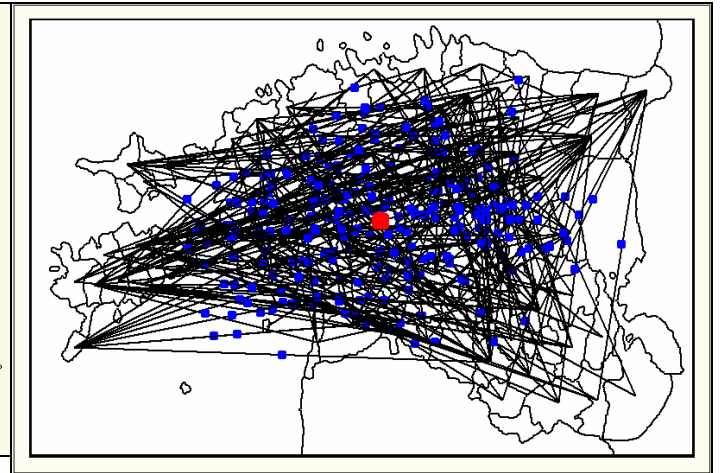
Mõistatud: 3001–3250 seoste graaf



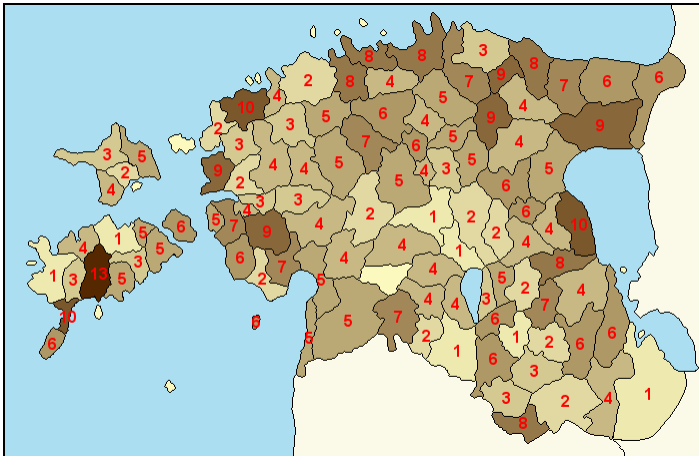
Mõistatud: 3251–3500 toorkoefitsiendid



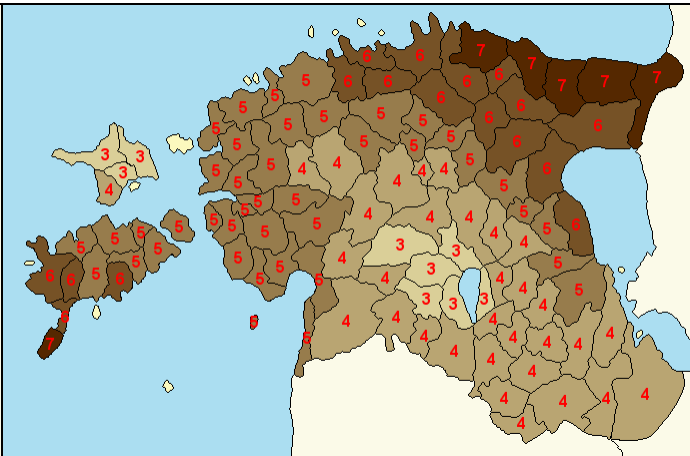
Mõistatud: 3251–3500 silutud koefitsiendid



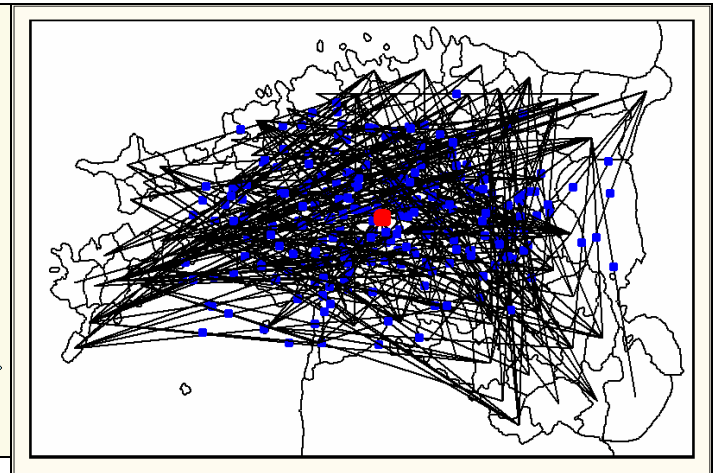
Mõistatud: 3251–3500 seoste graaf



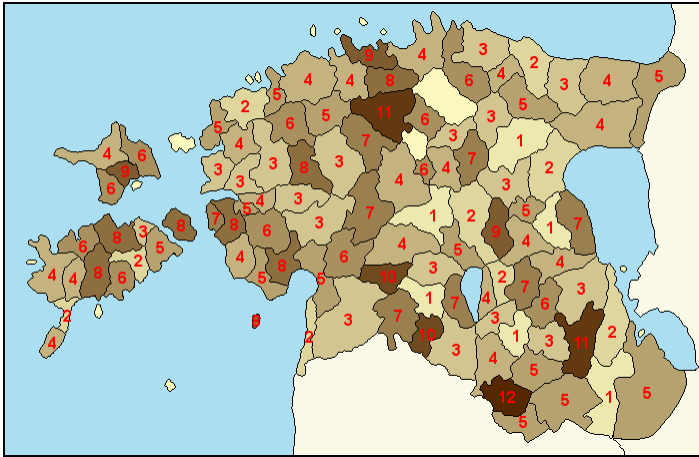
Mõistatud: 3501–3750 toorkoefitsiendid



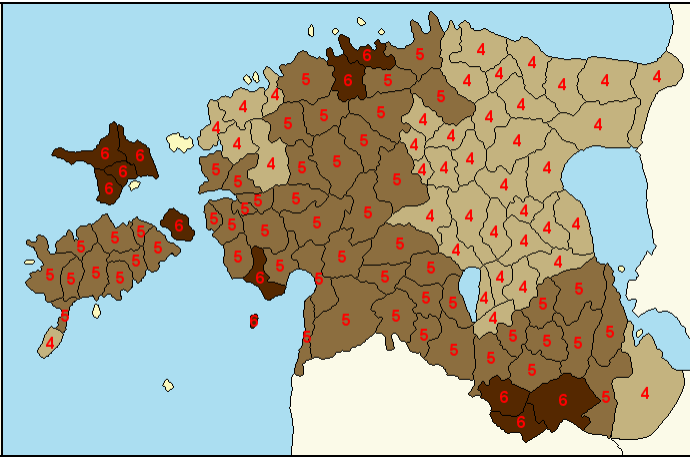
Mõistatud: 3501–3750 silutud koefitsiendid



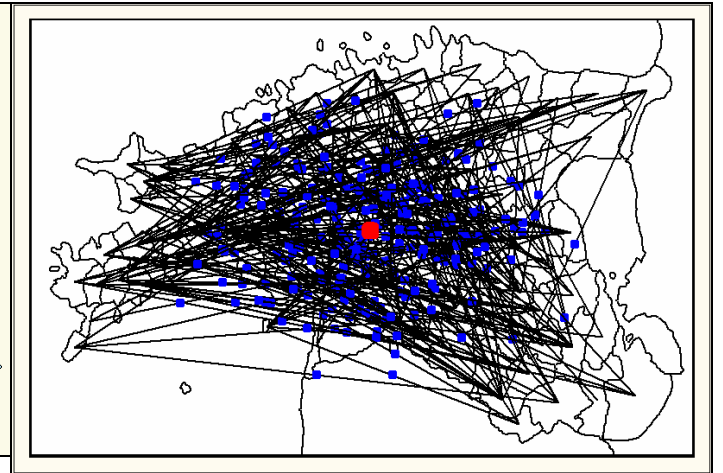
Mõistatud: 3501–3750 seoste graaf



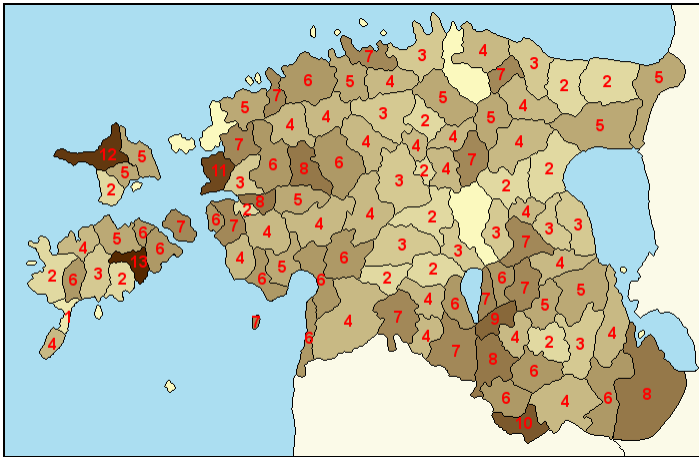
Mõistatud: 3751–4000 toorkoefitsiendid



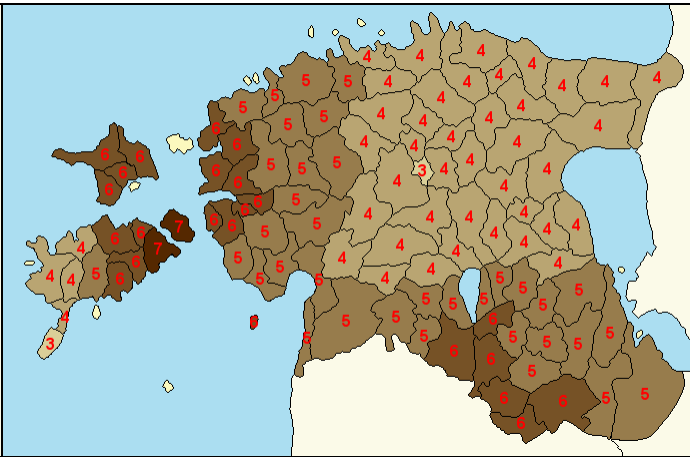
Mõistatud: 3751–4000 silutud koefitsiendid



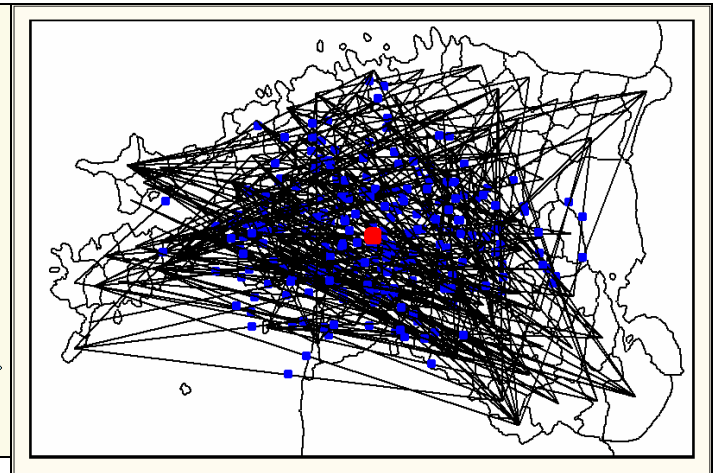
Mõistatud: 3751–4000 seoste graaf



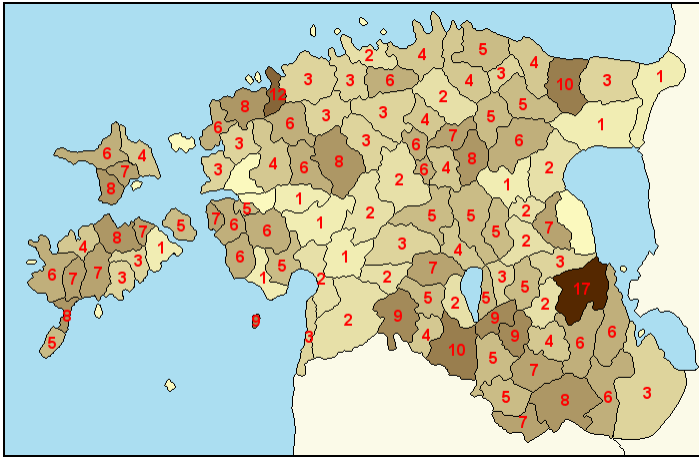
Mõistatud: 4001–4250 toorkoefitsiendid



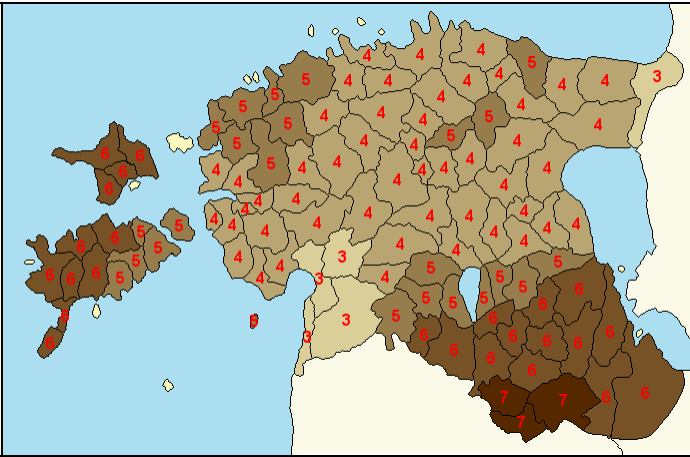
Mõistatud: 4001–4250 silutud koefitsiendid



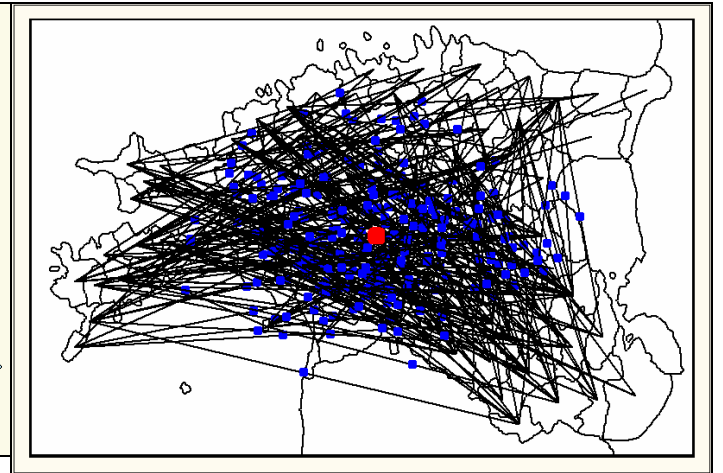
Mõistatud: 4001–4250 seoste graaf



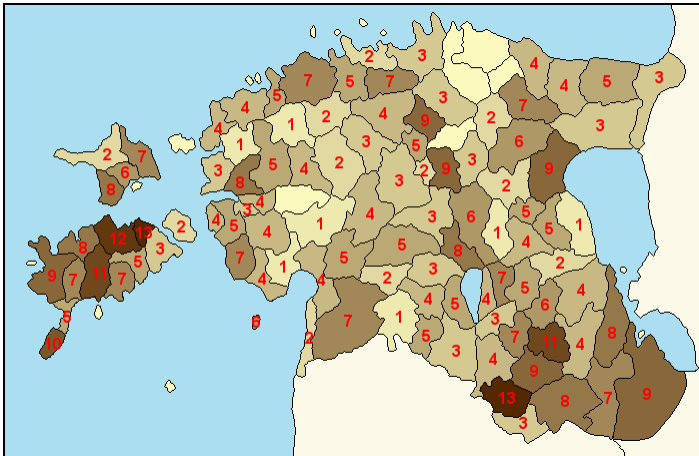
Mõistatused: 4251–4500 toorkoefitsiendid



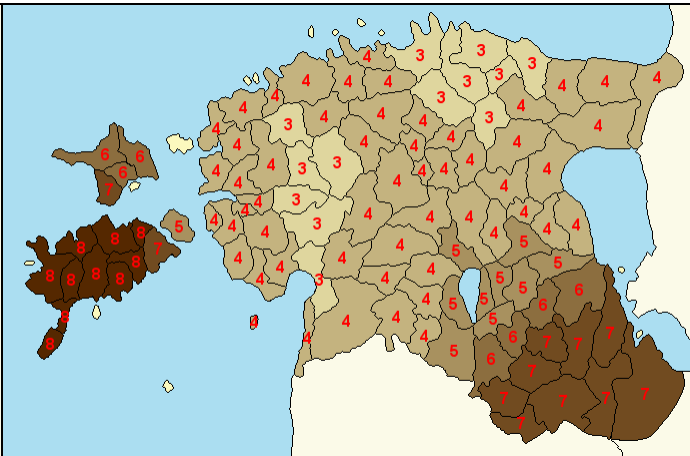
Mõistatused: 4251–4500 silutud koefitsiendid



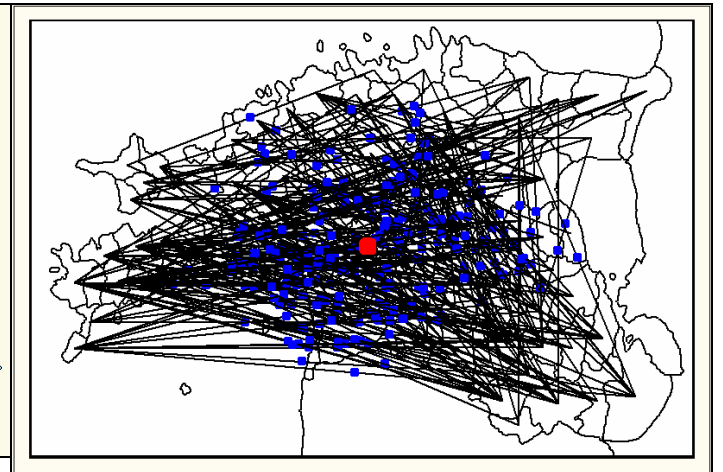
Mõistatused: 4251–4500 seoste graaf



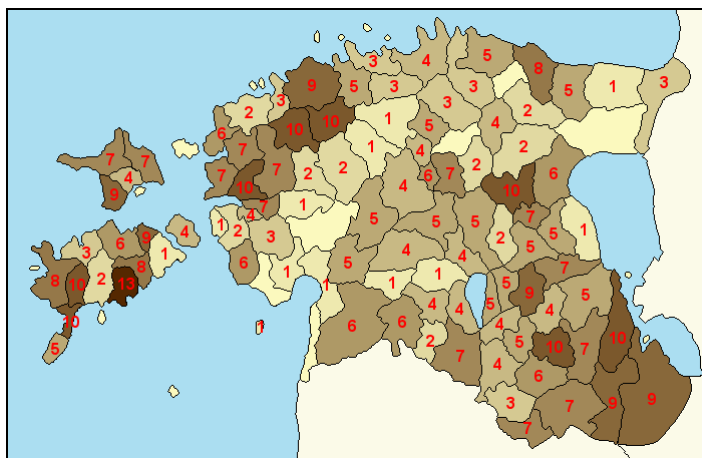
Mõistatused: 4501–4750 toorkoefitsiendid



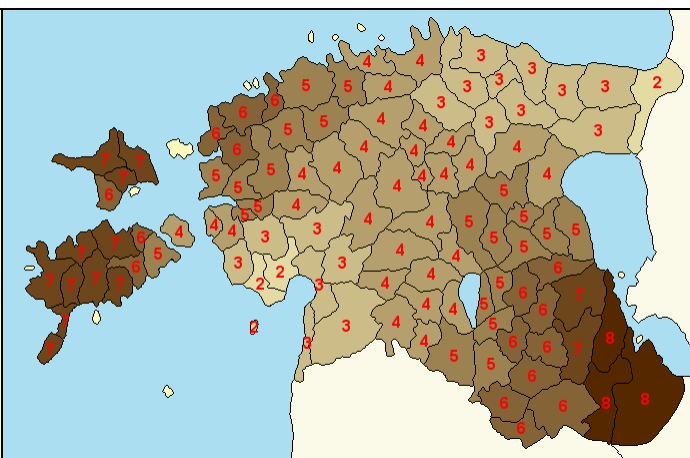
Mõistatused: 4501–4750 silutud koefitsiendid



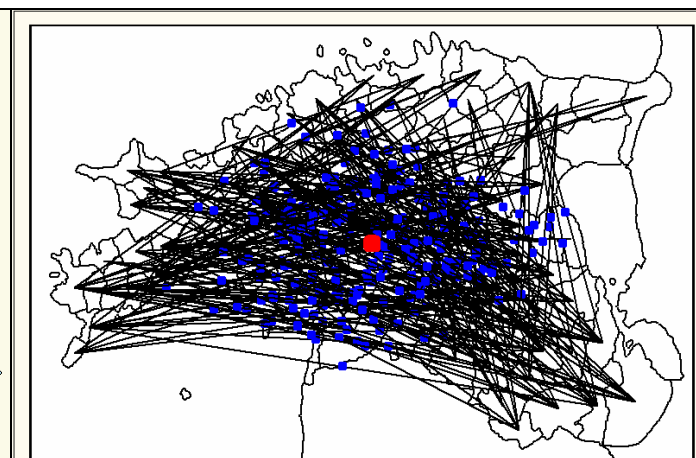
Mõistatused: 4501–4750 seoste graaf



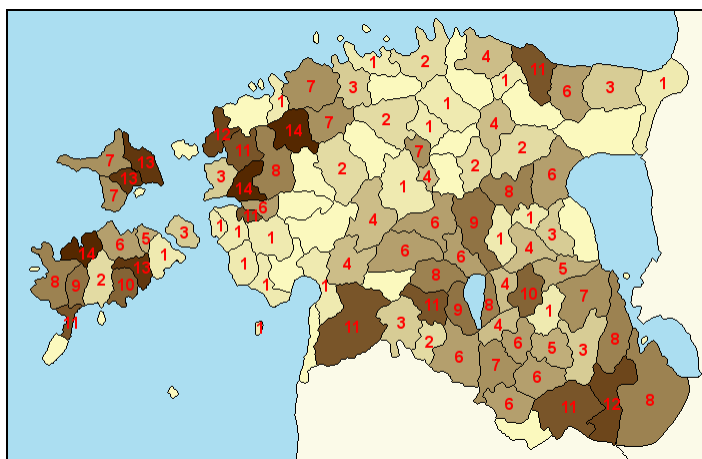
Mõistatud: 4751–5000 toorkoefitsiendid



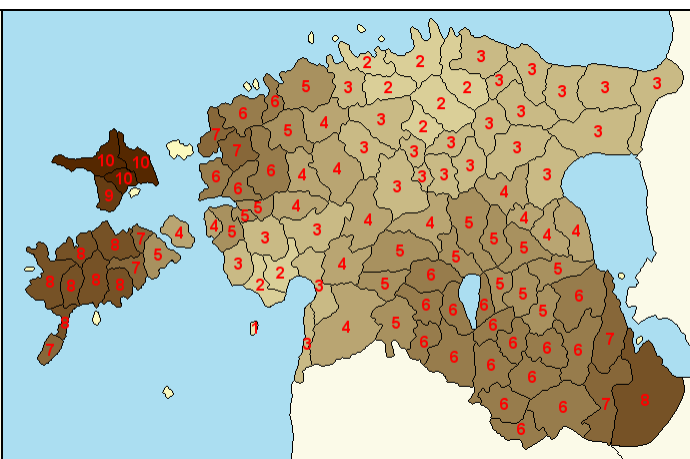
Mõistatud: 4751–5000 silutud koefitsiendid



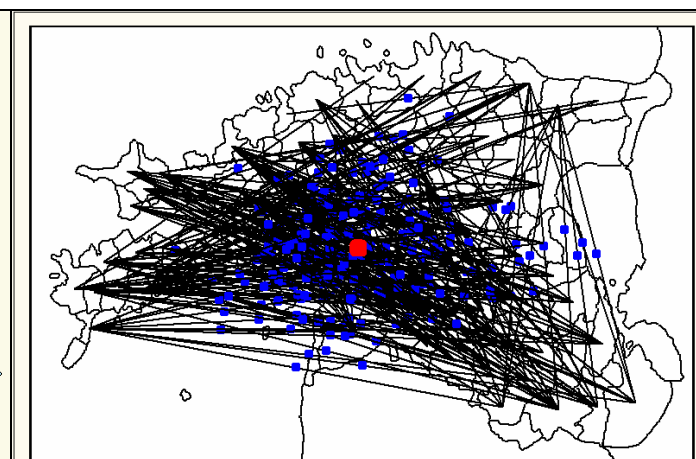
Mõistatud: 4751–5000 seoste graaf



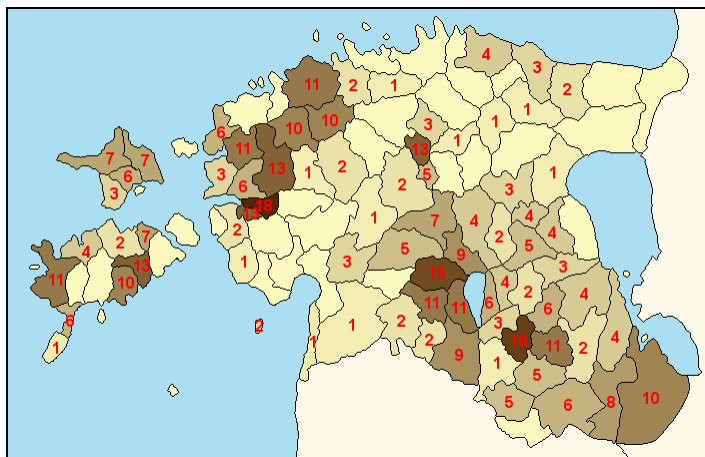
Mõistatud: 5001–5250 toorkoefitsiendid



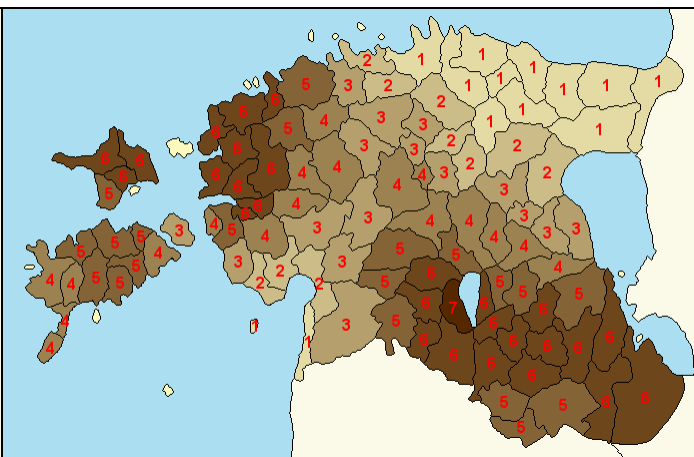
Mõistatud: 5001–5250 silutud koefitsiendid



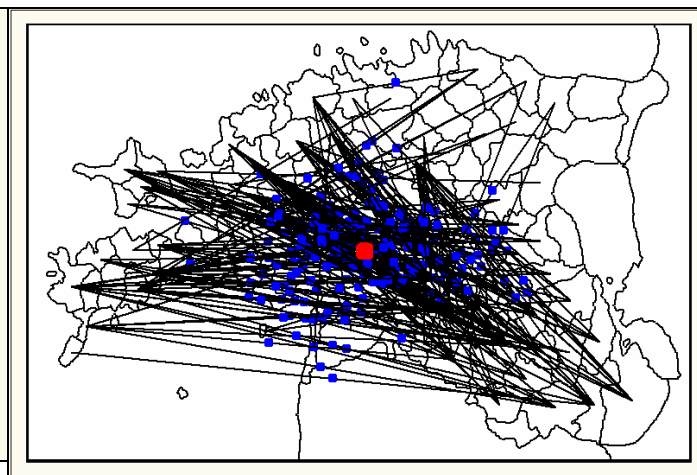
Mõistatud: 5001–5250 seoste graaf



Mõistatud: 5251–5460 toorkoefitsiendid

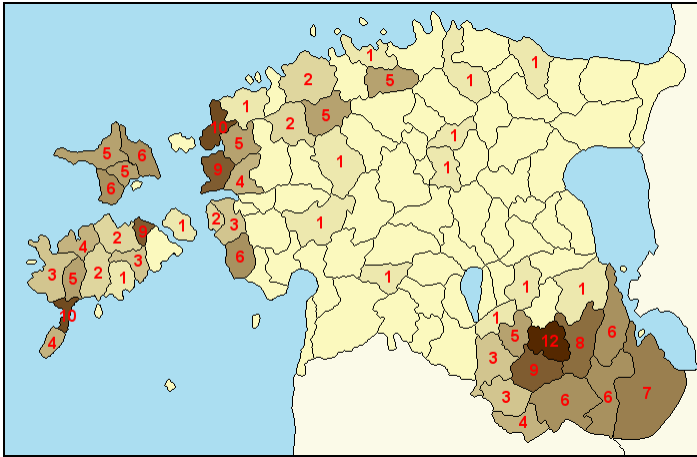


Mõistatud: 5251–5460 silutud koefitsiendid

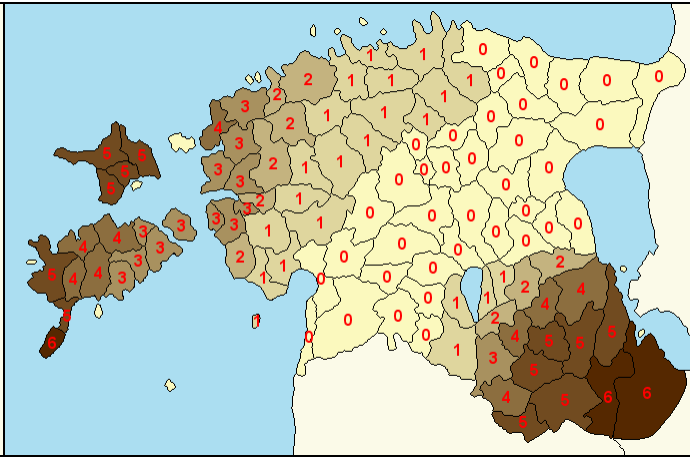


Mõistatud: 5251–5460 seoste graaf

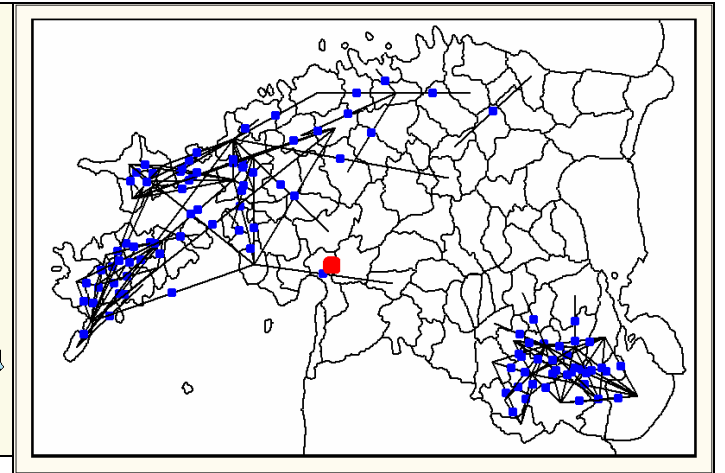
Vanasõnad



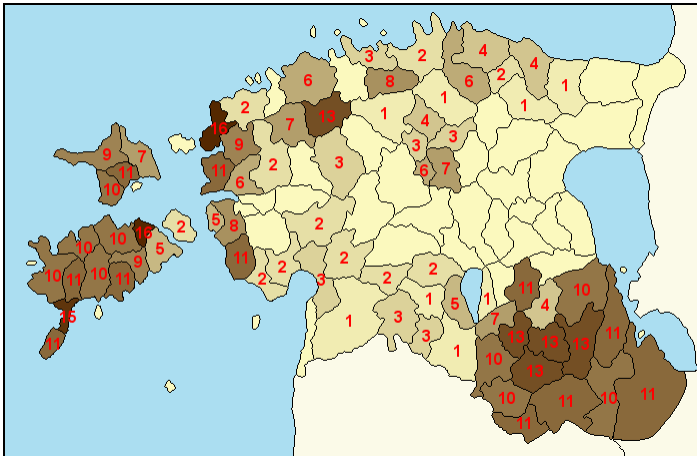
Vanasõnad: 1–100 toorkoefitsiendid



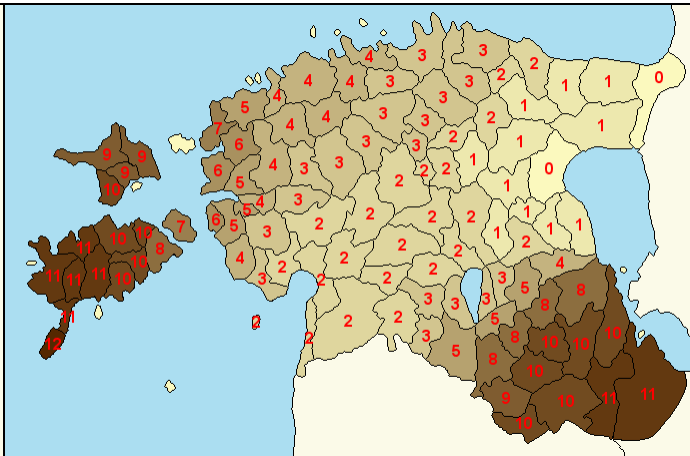
Vanasõnad: 1–100 silutud koefitsiendid



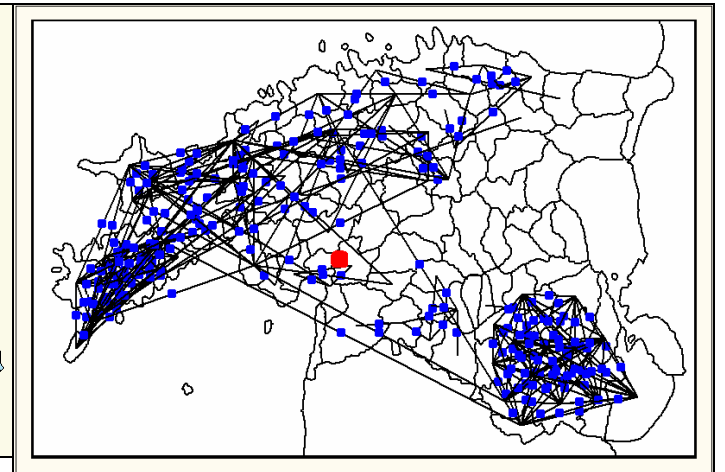
Vanasõnad: 1–100 seoste graaf



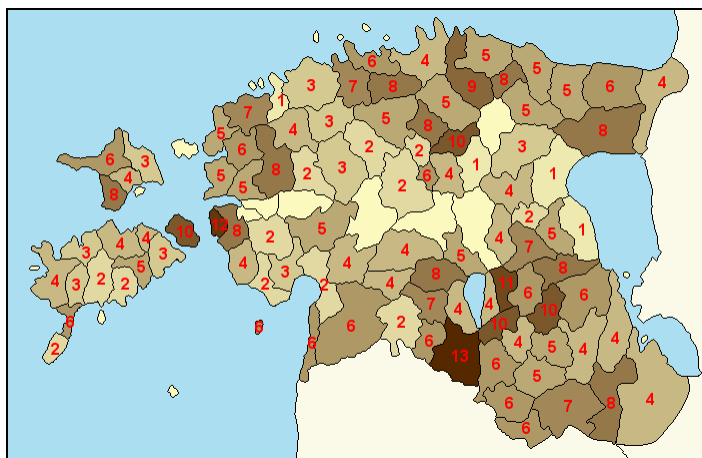
Vanasõnad: 1–250 toorkoefitsiendid



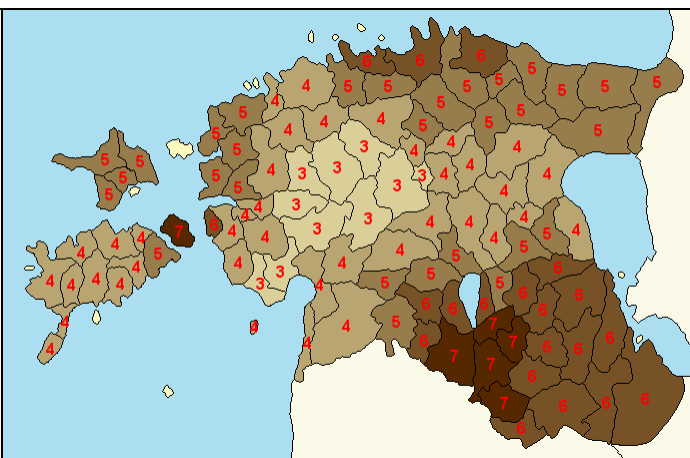
Vanasõnad: 1–250 silutud koefitsiendid



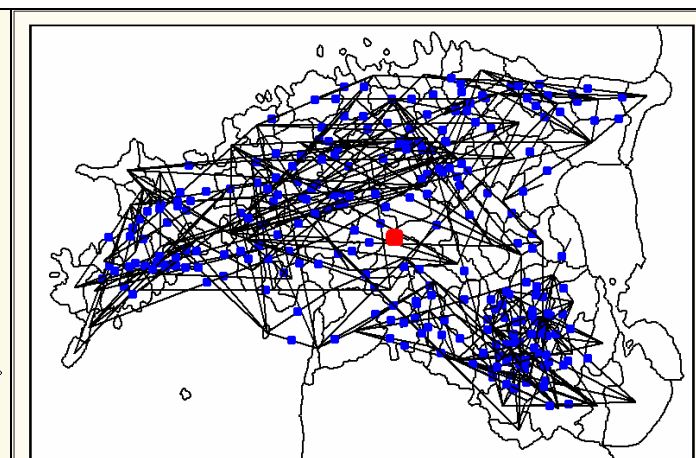
Vanasõnad: 1–250 seoste graaf



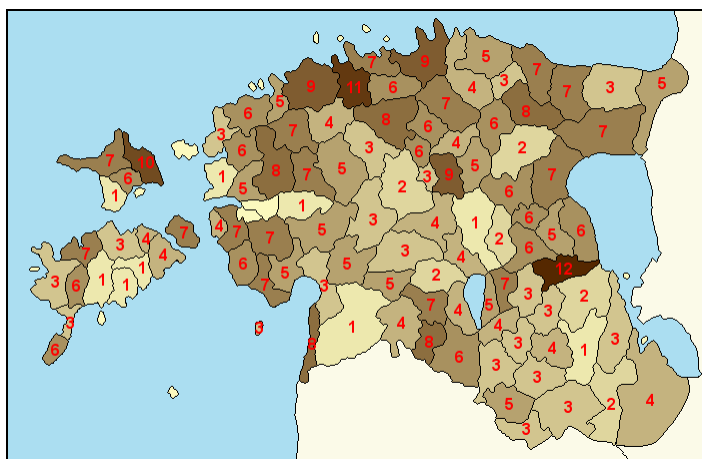
Vanasõnad: 251–500 toorkoefitsiendid



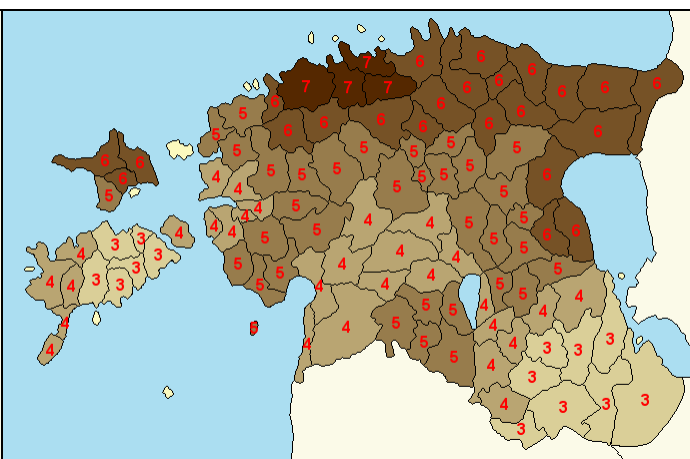
Vanasõnad: 251–500 silutud koefitsiendid



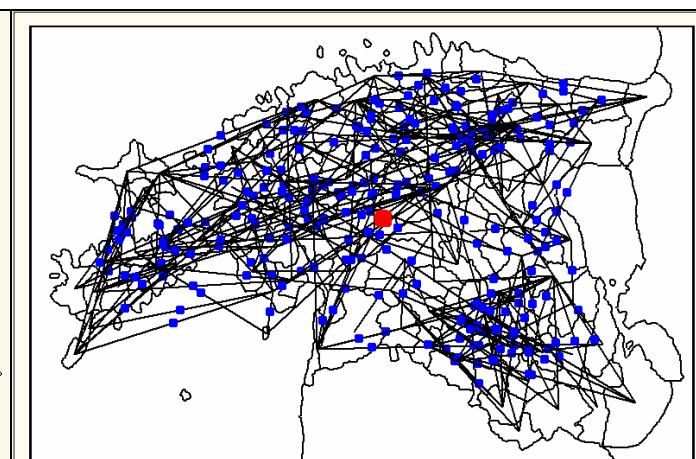
Vanasõnad: 251–500 seoste graaf



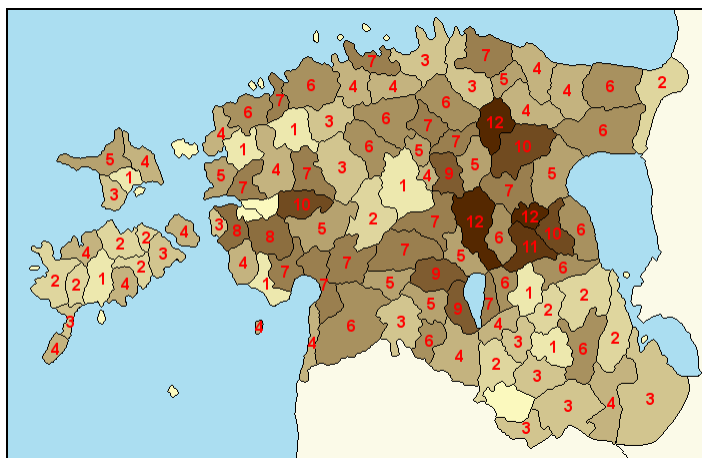
Vanasõnad: 501–750 toorkoefitsiendid



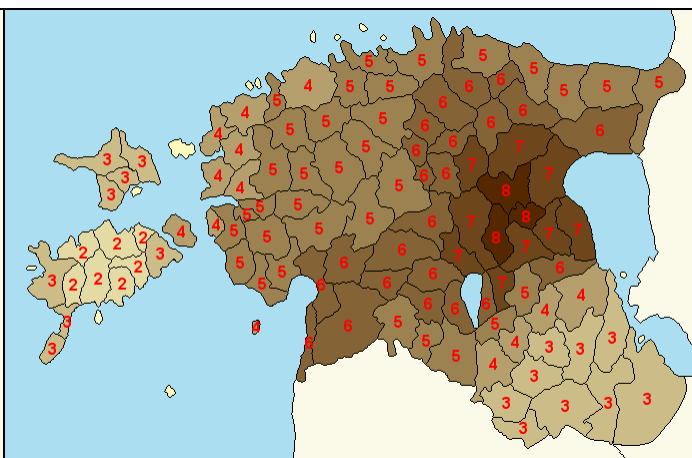
Vanasõnad: 501–750 silutud koefitsiendid



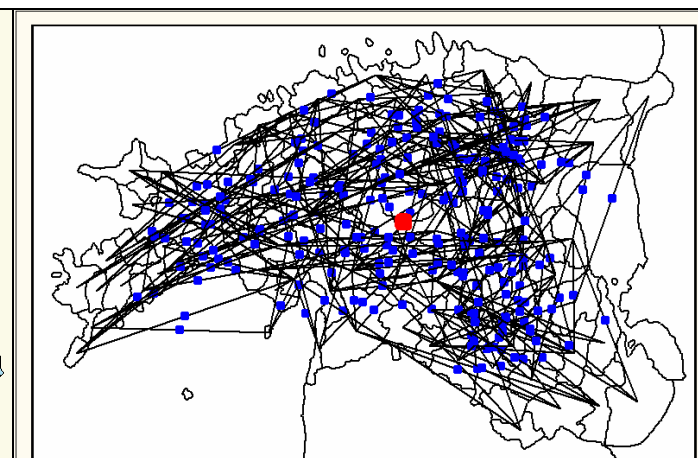
Vanasõnad: 501–750 seoste graaf



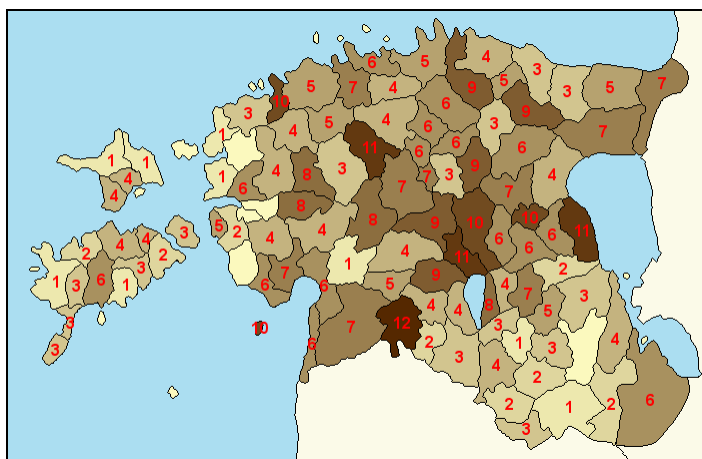
Vanasõnad: 751–1000 toorkoefitsiendid



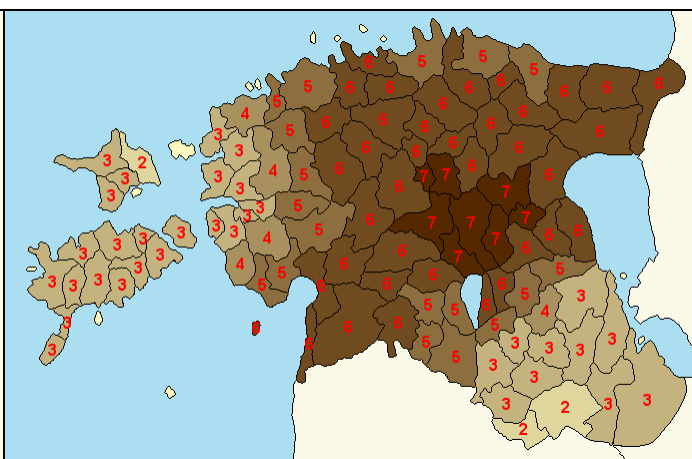
Vanasõnad: 751–1000 silutud koefitsiendid



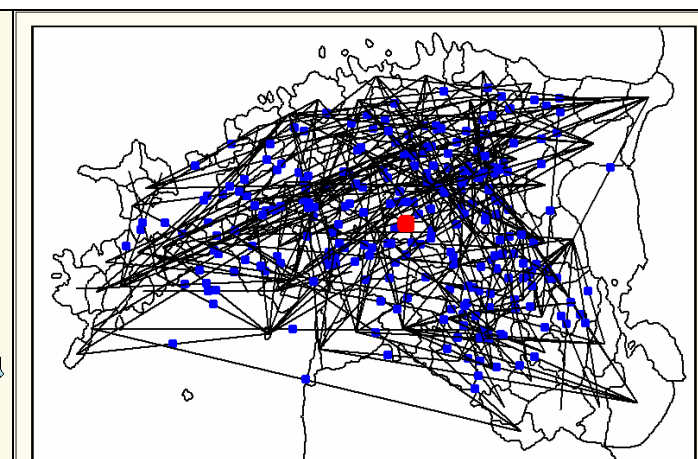
Vanasõnad: 751–1000 seoste graaf



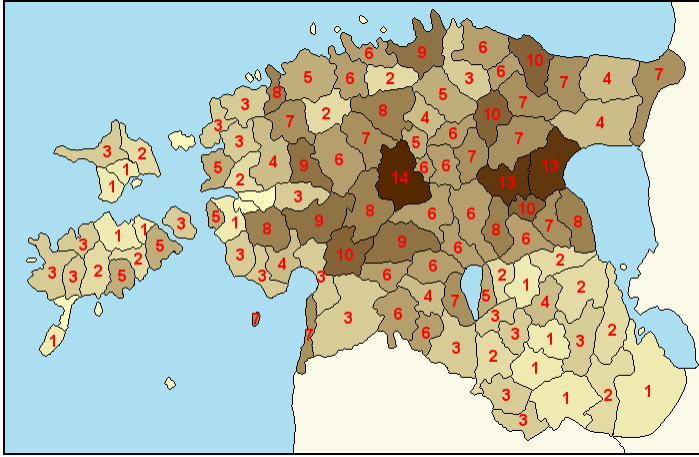
Vanasõnad: 1001–1250 toorkoefitsiendid



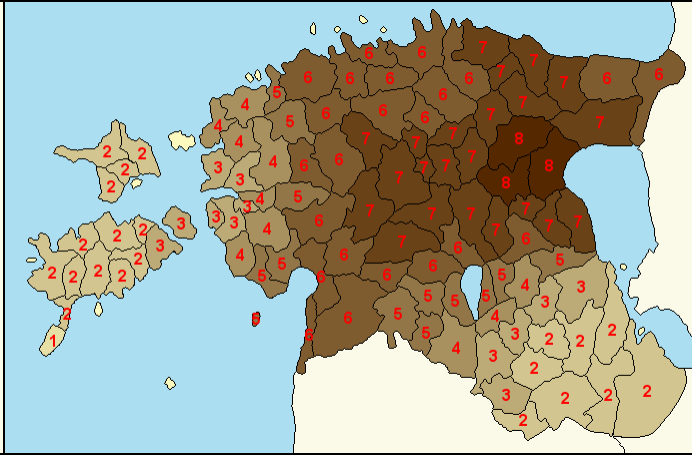
Vanasõnad: 1001–1250 silutud koefitsiendid



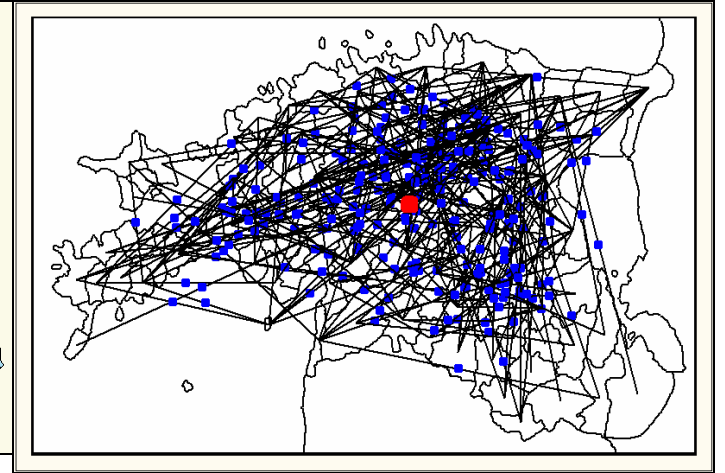
Vanasõnad: 1001–1250 seoste graaf



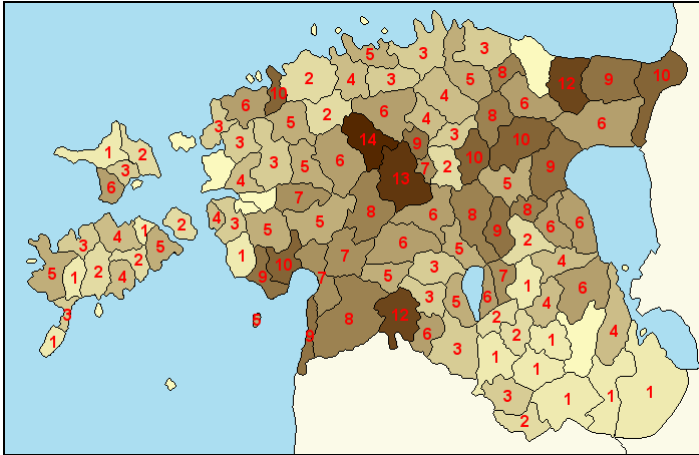
Vanasõnad: 1251–1500 toorkoefitsiendid



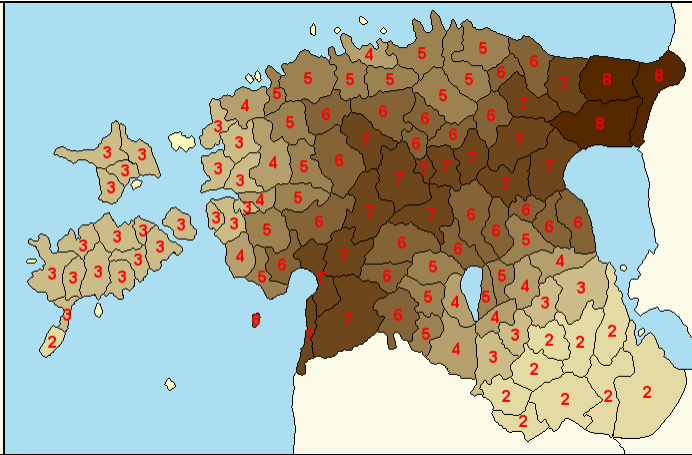
Vanasõnad: 1251–1500 silutud koefitsiendid



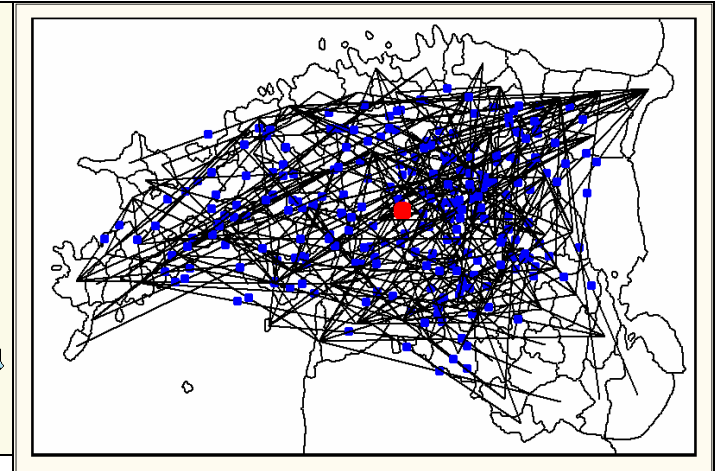
Vanasõnad: 1251–1500 seoste graaf



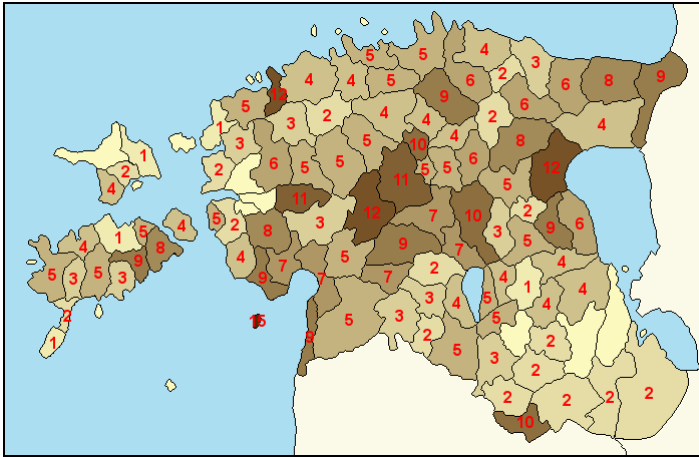
Vanasõnad: 1501–1750 toorkoefitsiendid



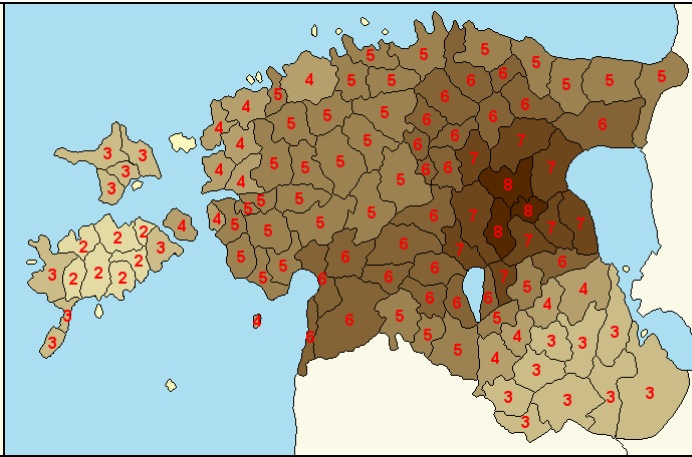
Vanasõnad: 1501–1750 silutud koefitsiendid



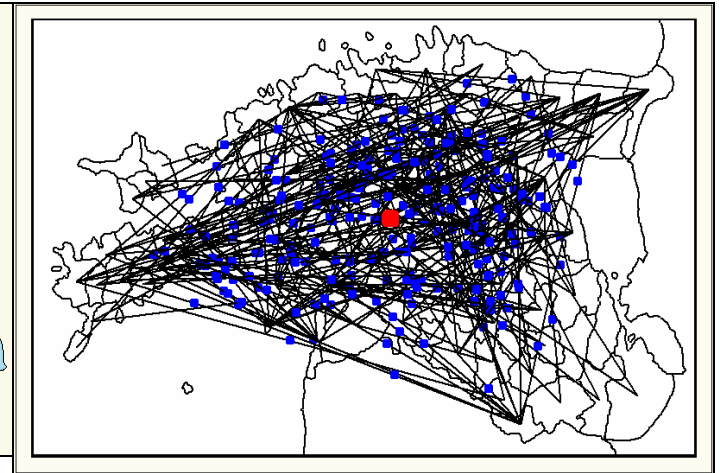
Vanasõnad: 1501–1750 seoste graaf



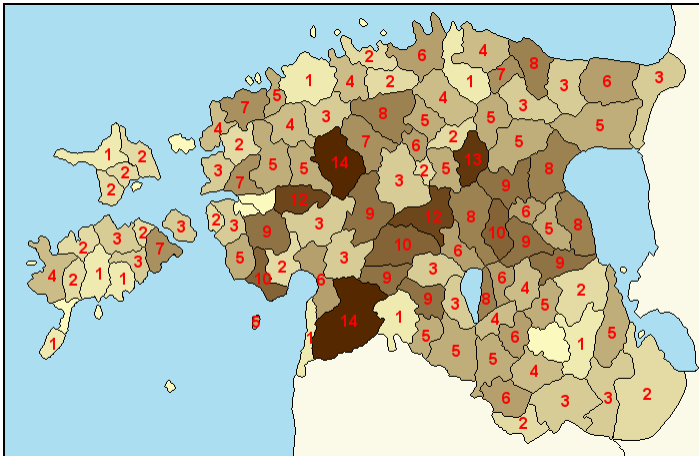
Vanasõnad: 1751–2000 toorkoefitsiendid



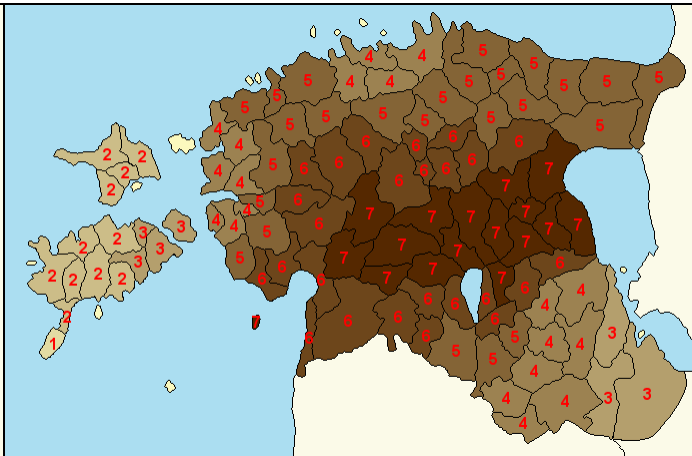
Vanasõnad: 1751–2000 silutud koefitsiendid



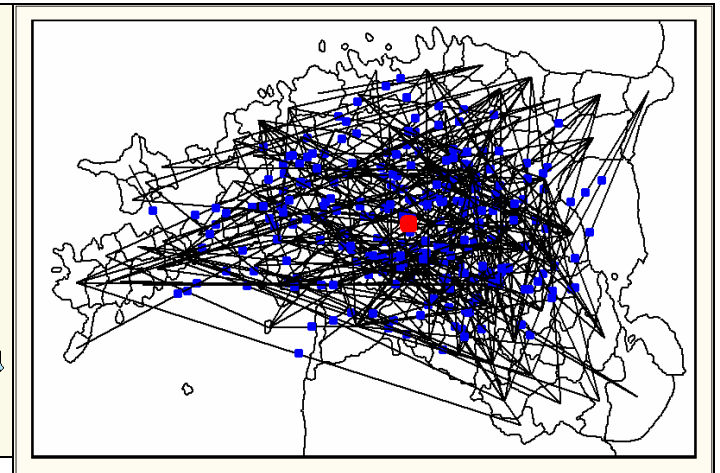
Vanasõnad: 1751–2000 seoste graaf



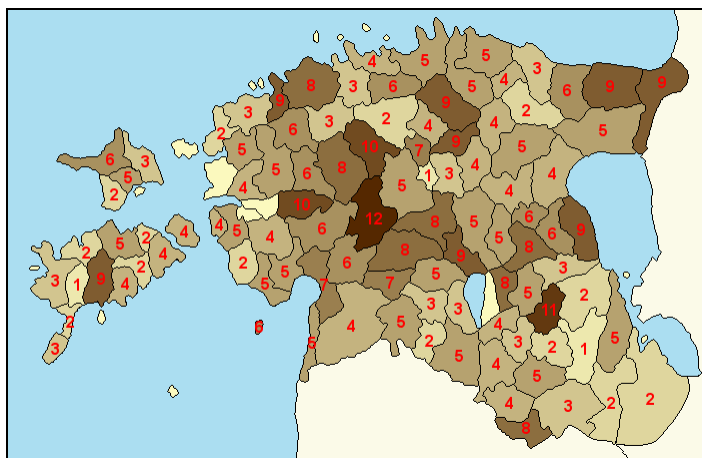
Vanasõnad: 2001–2250 toorkoefitsiendid



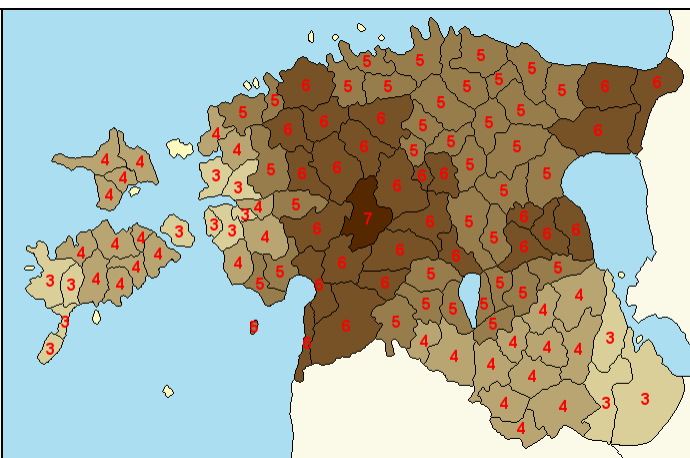
Vanasõnad: 2001–2250 silutud koefitsiendid



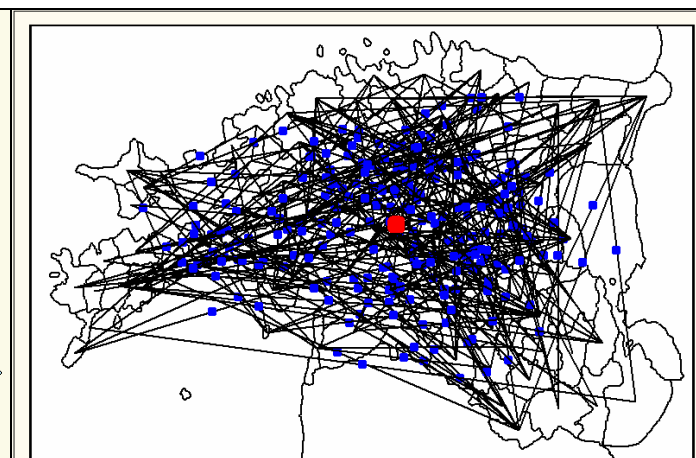
Vanasõnad: 2001–2250 seoste graaf



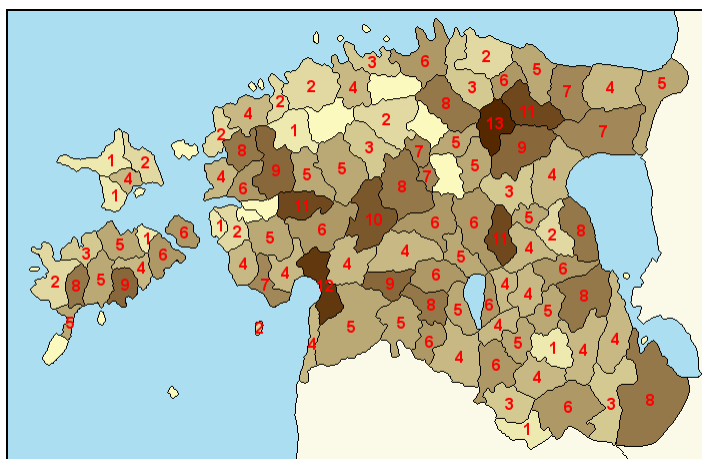
Vanasõnad: 2251–2500 toorkoefitsiendid



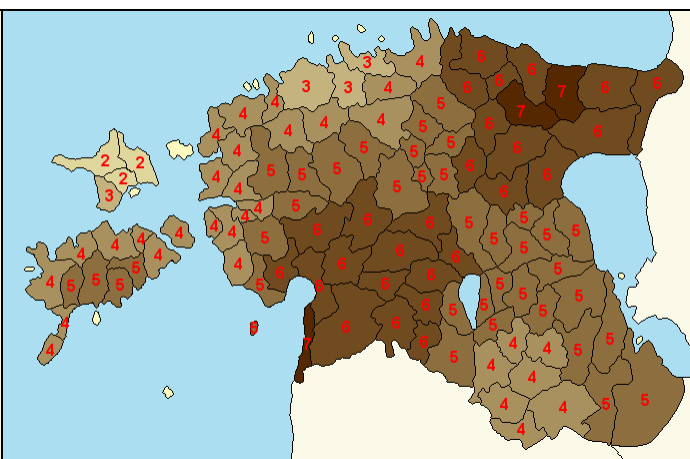
Vanasõnad: 2251–2500 silutud koefitsiendid



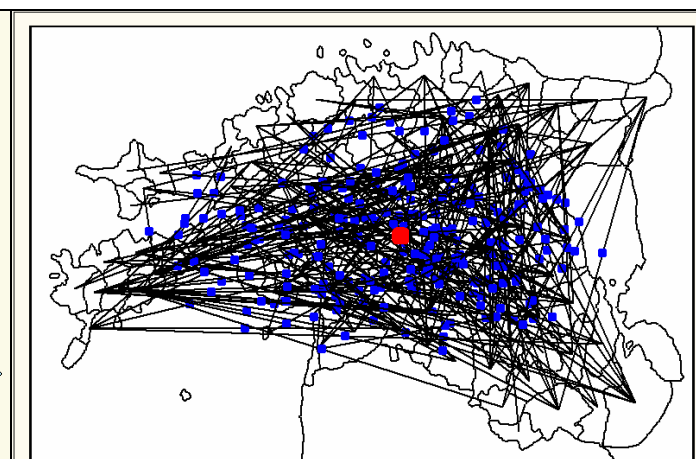
Vanasõnad: 2251–2500 seoste graaf



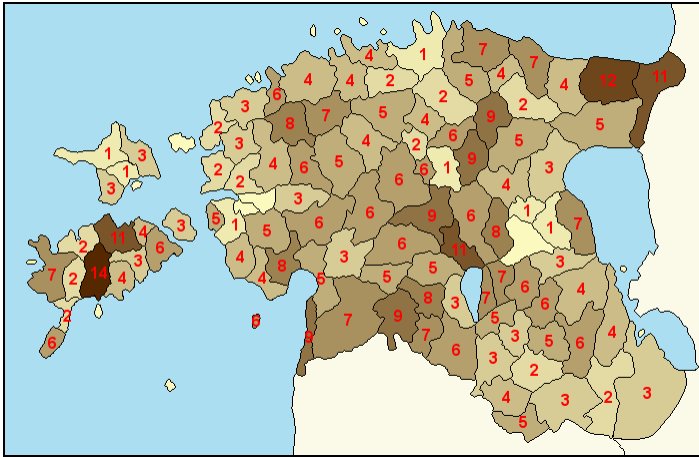
Vanasõnad: 2501–2750 toorkoefitsiendid



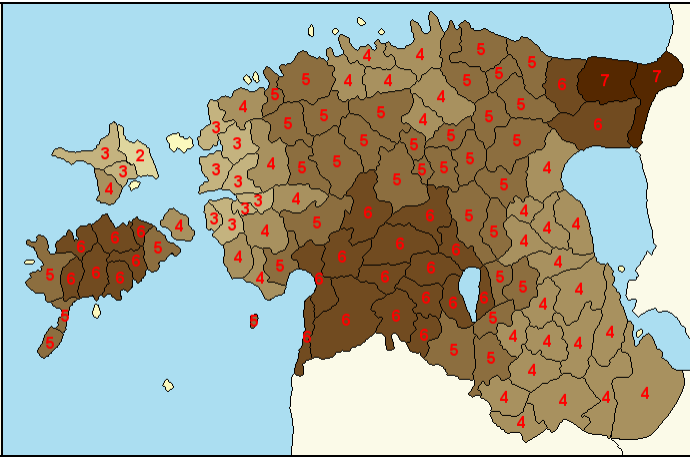
Vanasõnad: 2501–2750 silutud koefitsiendid



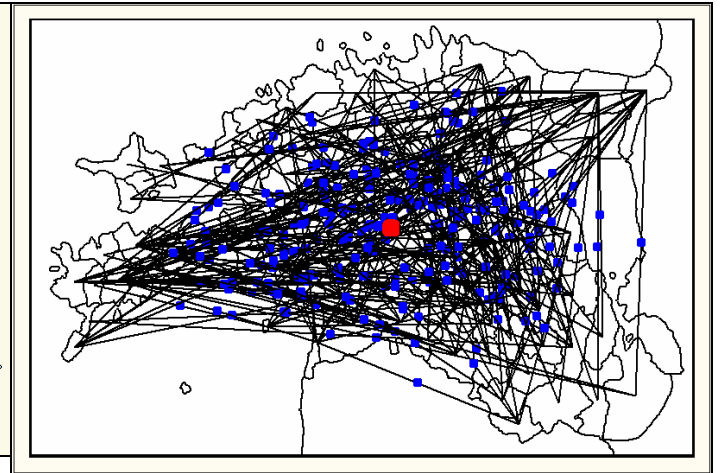
Vanasõnad: 2501–2750 seoste graaf



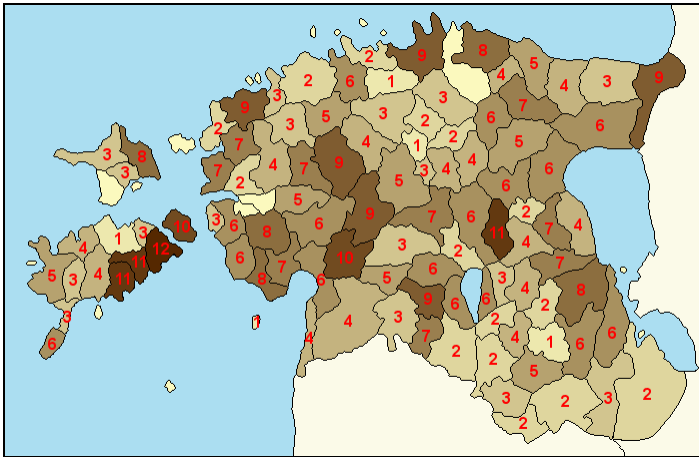
Vanasõnad: 2751–3000 toorkoefitsiendid



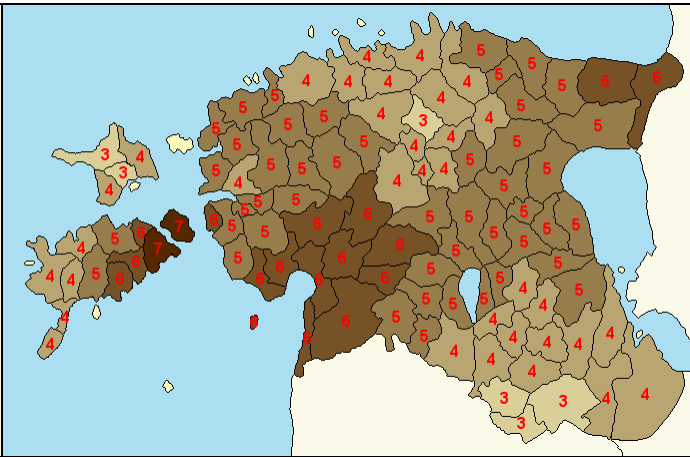
Vanasõnad: 2751–3000 silatud koefitsiendid



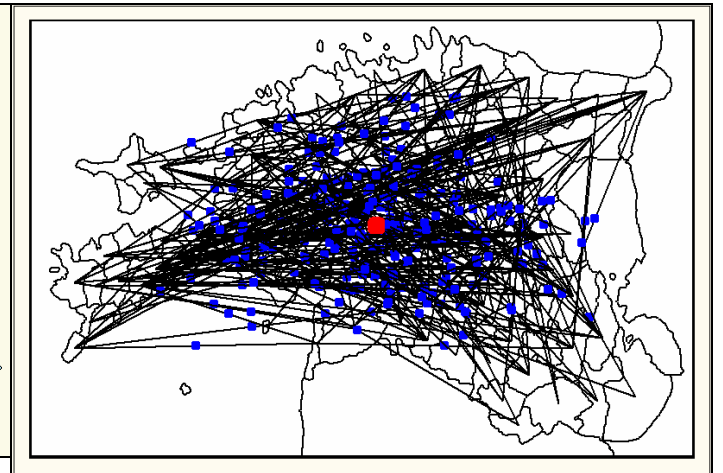
Vanasõnad: 2751–3000 seoste graaf



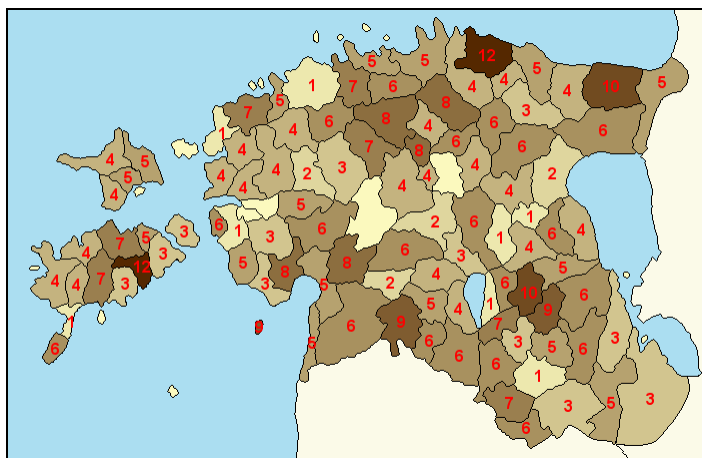
Vanasõnad: 3001–3250 toorkoefitsiendid



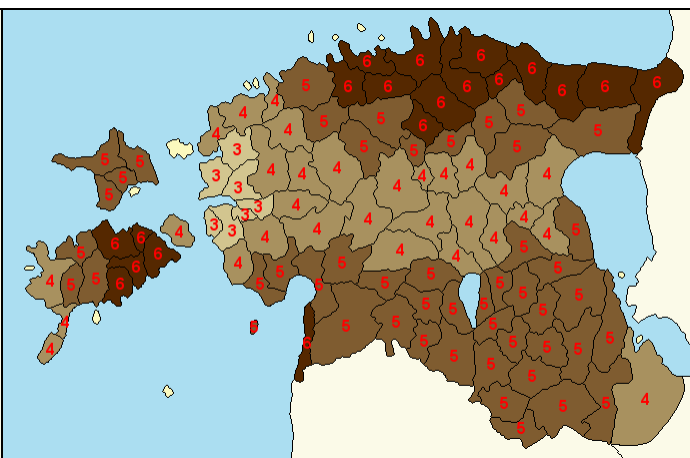
Vanasõnad: 3001–3250 silatud koefitsiendid



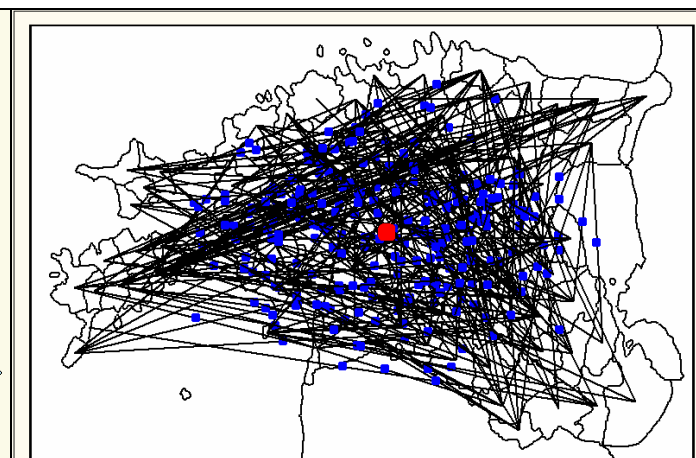
Vanasõnad: 3001–3250 seoste graaf



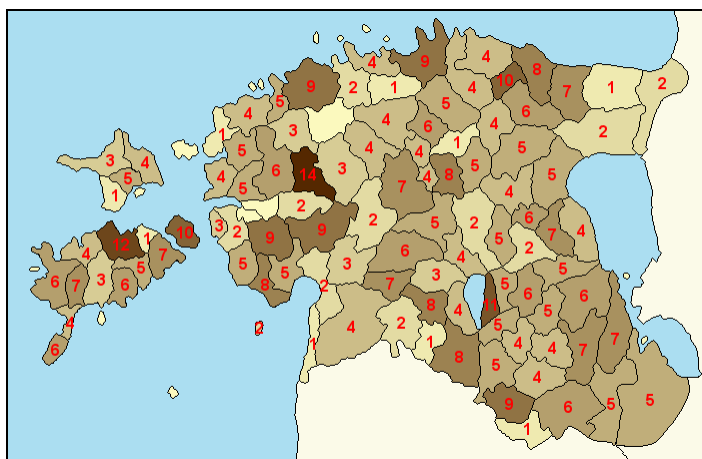
Vanasõnad: 3251–3500 toorkoefitsiendid



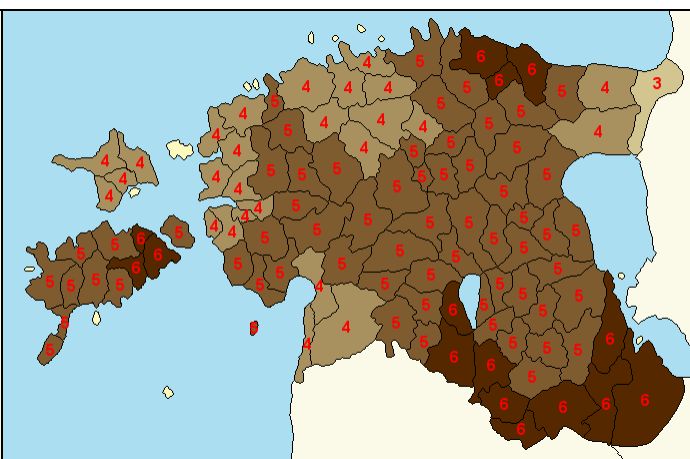
Vanasõnad: 3251–3500 silutud koefitsiendid



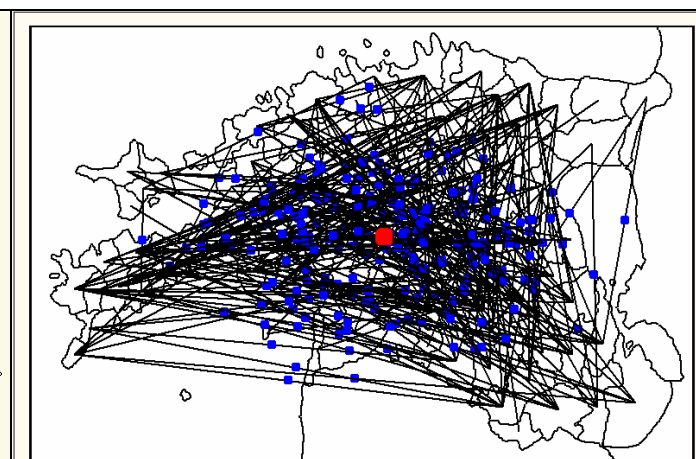
Vanasõnad: 3251–3500 seoste graaf



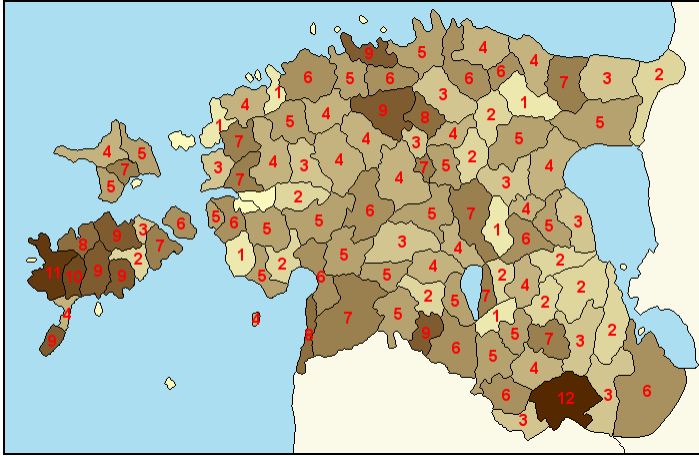
Vanasõnad: 3501–3750 toorkoefitsiendid



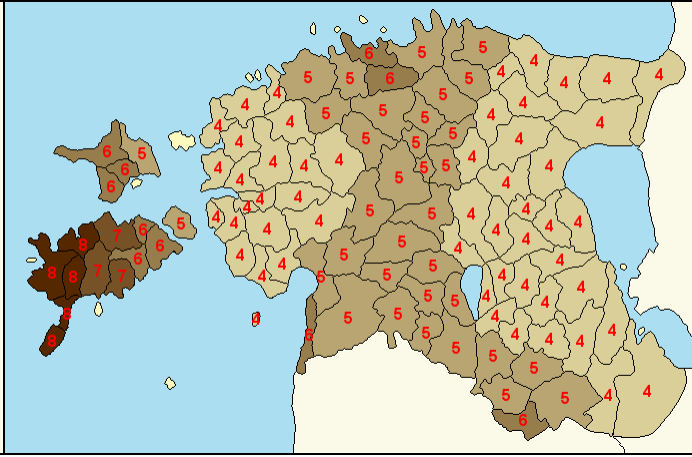
Vanasõnad: 3501–3750 silutud koefitsiendid



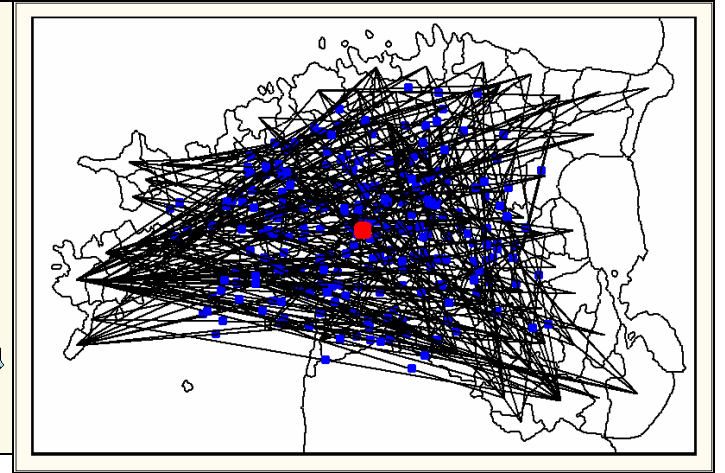
Vanasõnad: 3501–3750 seoste graaf



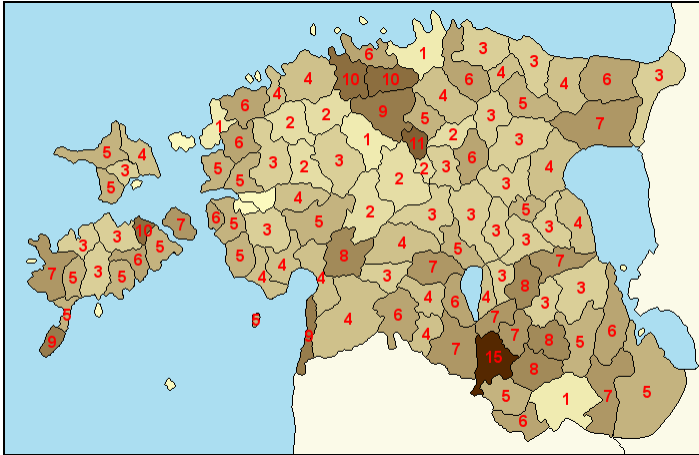
Vanasõnad: 3751–4000 toorkoefitsiendid



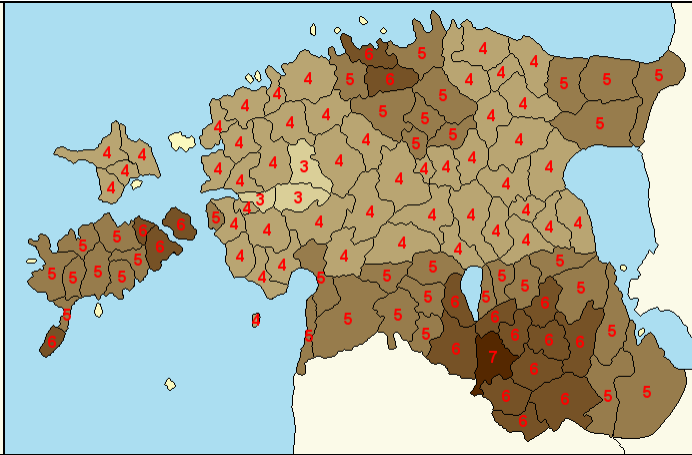
Vanasõnad: 3751–4000 silutud koefitsiendid



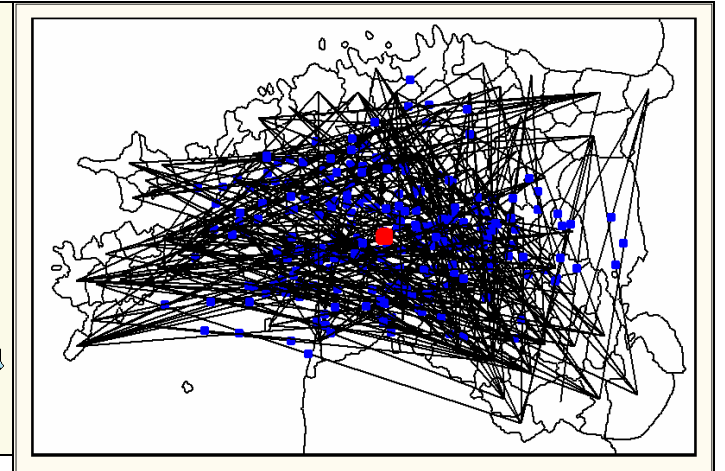
Vanasõnad: 3751–4000 seoste graaf



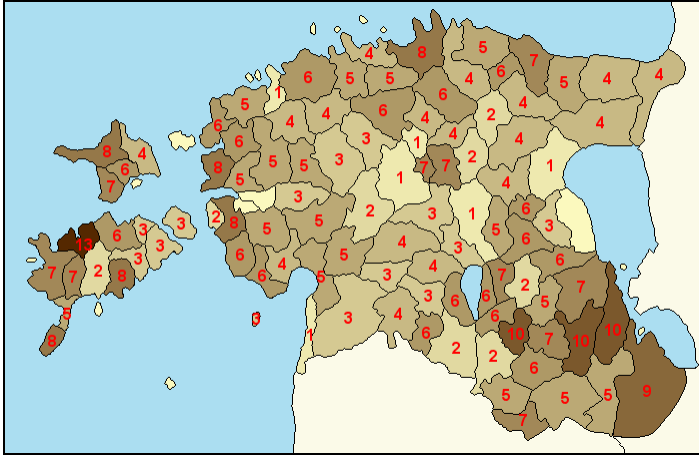
Vanasõnad: 4001–4250 toorkoefitsiendid



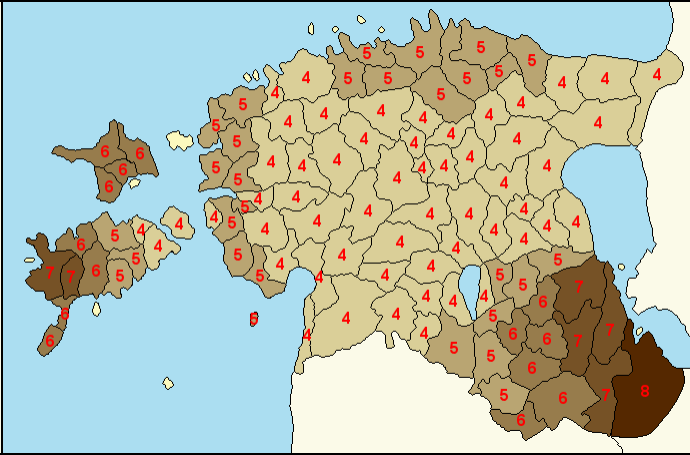
Vanasõnad: 4001–4250 silutud koefitsiendid



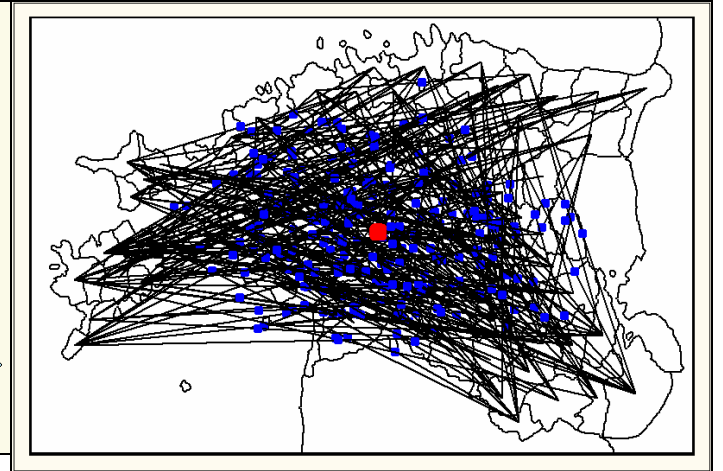
Vanasõnad: 4001–4250 seoste graaf



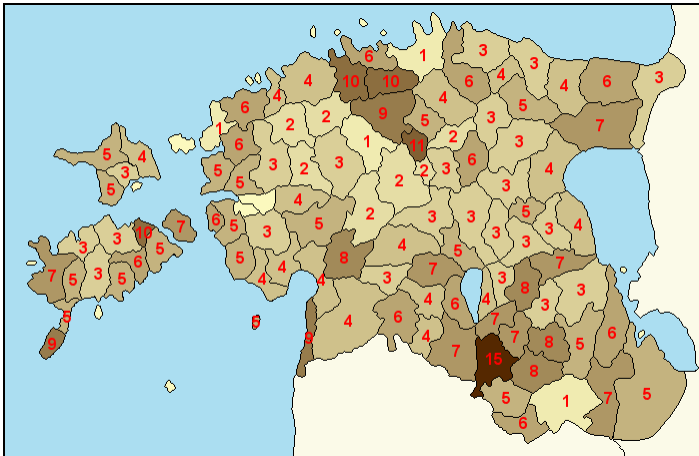
Vanasõnad: 4251–4500 toorkoefitsiendid



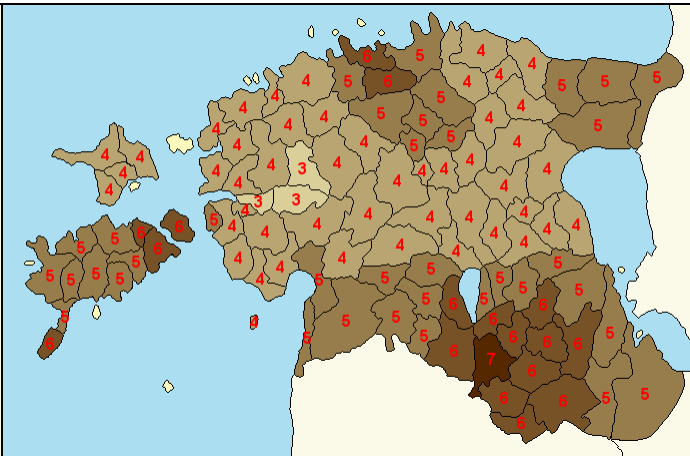
Vanasõnad: 4251–4500 silutud koefitsiendid



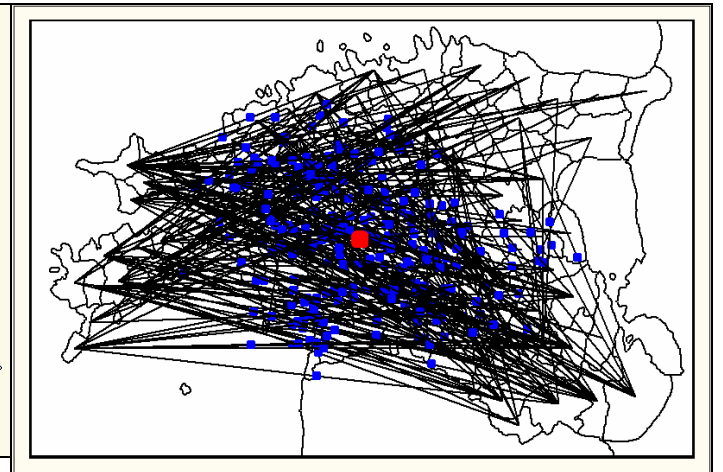
Vanasõnad: 4251–4500 seoste graaf



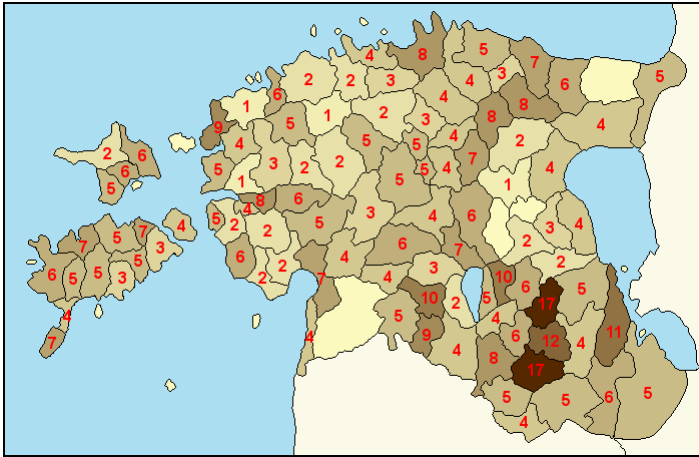
Vanasõnad: 4501–4750 toorkoefitsiendid



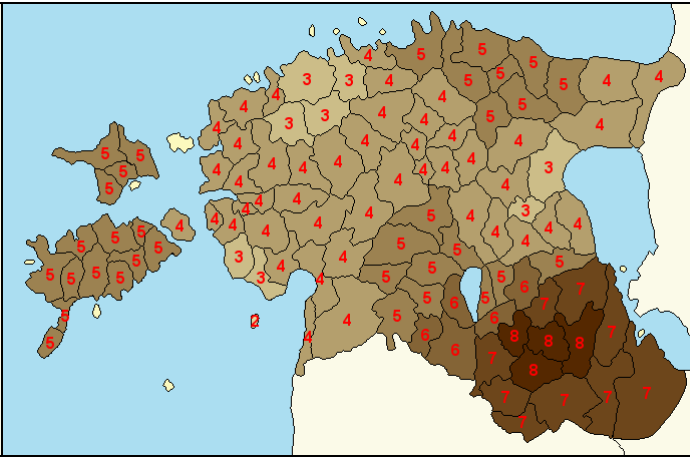
Vanasõnad: 4501–4750 silutud koefitsiendid



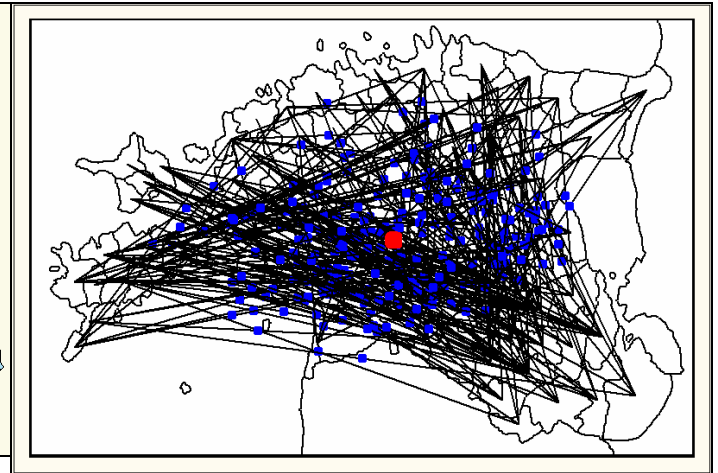
Vanasõnad: 4501–4750 seoste graaf



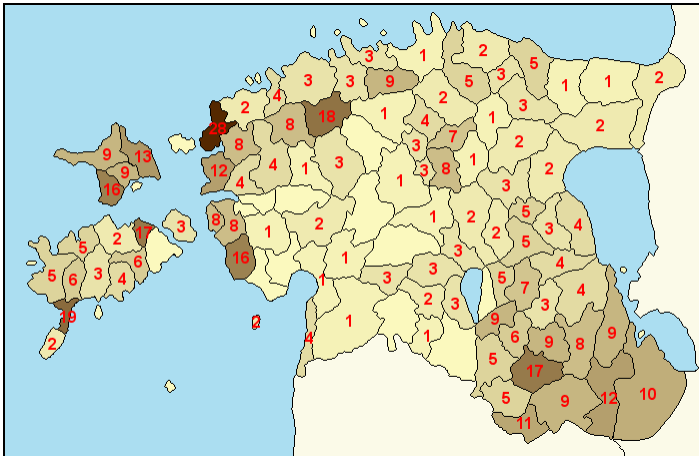
Vanasõnad: 4751–5000 toorkoefitsiendid



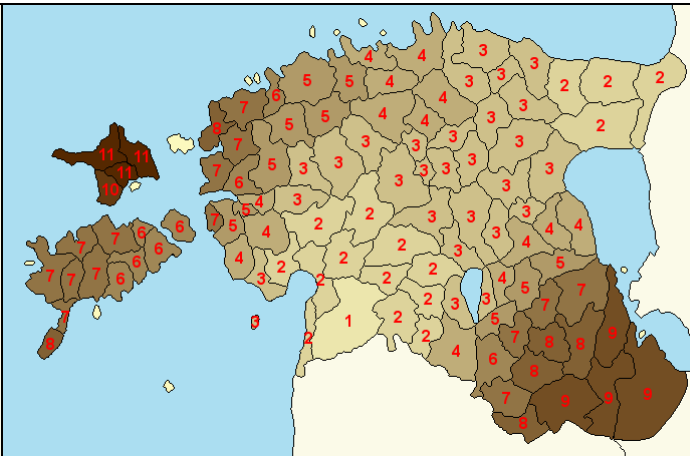
Vanasõnad: 4751–5000 silutud koefitsiendid



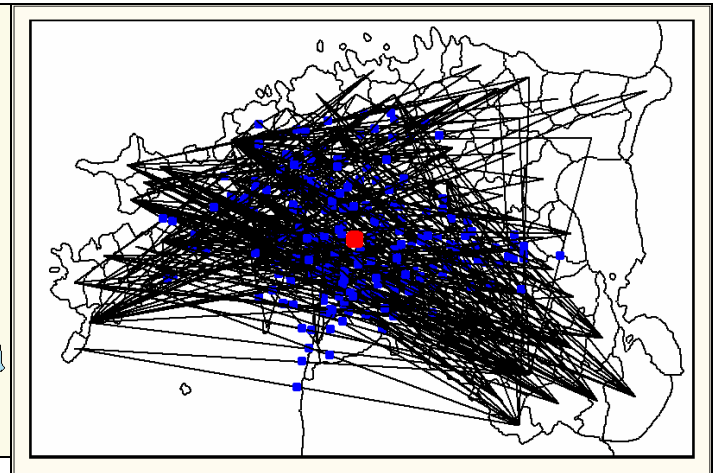
Vanasõnad: 4751–5000 seoste graaf



Vanasõnad: 5001–5253 toorkoefitsiendid



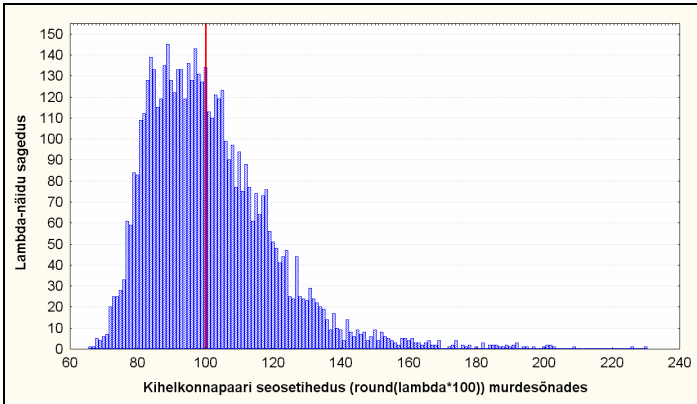
Vanasõnad: 5001–5253 silutud koefitsiendid



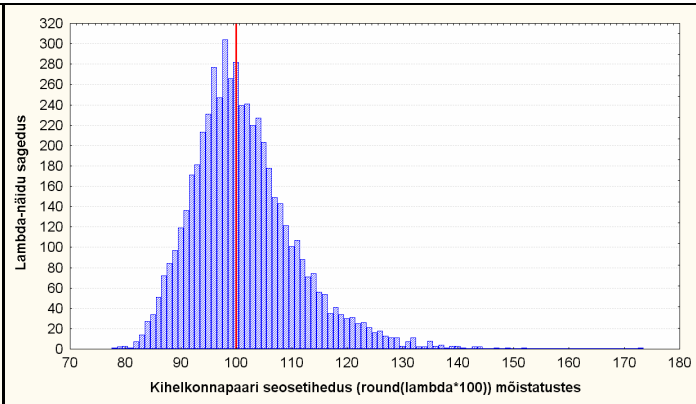
Vanasõnad: 5001–5253 seoste graaf

Mind valdas hetkeline nukrus, kui taipasin, et mul pole mingit võimalust enamikule kaardiridadele juhuslikesse kohtadesse tekkivaist värvilaikudest ja praktiliselt kogu maad katvaist joonetihnikutest leida seaduspärasusi või tendentse, mis võiksid kehtida meie λ -skaalade põhiosas. Järgnevas lk 58 on toodud kõrvutatult λ -sageduste histogrammid ja neile vastavad “pinnalaotuslikud” variatsiooniread; nende vastavused teineteisele n-õ vasakult paremale on pöördpidised ja suhete neutraalsushetke näitavad punased kontrolljooned asuvad eri telgedel. Neid vaadates sugenes lohutav oletus: siit pole võib-olla põhjust mingeid seaduspärasusi otsidagi. Kõik tuvastatavad murde- või folkloorialad konstitueeritakse juba λ -astmike tajutavalt diskreetsete astakutega “paremikes”, esimese u 500 astme jooksul, kus nidusad võrgustikud tekivad kõigepealt Kagu-Eestisse ja/või saartele ning hargnevad sealt hajusamate ja hõredamate metastaasidena vastavalt Lõuna-Tartumaale ja Mulgisse ning Loode- ja Lääne-Eesti mandriossa, ning ühtlasi tekivad mõned uued, kuid nõrgemad ja hõredamad seostekolded Ranna-Eestisse, Järvamaale ja mujale. Suur osa kõigist seostumistest toimub aga järgneval pikkamisi lineaarselt alaneval kaldteel, mis katab λ -skaalade koguulatusest ainult kitsa, kõige neutraalsemaid suhteid hõlmava vahemiku (u 1,1 ... 0,9; meil tehniliselt: 110 ... 90). Põhiprotsessid toimuvad siin Põhja-Eestis, kuid põhimõtteliselt võivad igalt alalt pärit geooksused seostuda mistahes muult alalt pärinevatega. Ning skaala lõpuosa, kus koefitsientide kahanemine taas kiireneb, hõlmavad pms unikaalsete alade esindajate (nõrgimad) seosed teistelt unikaalsetelt aladelt pärinevate geopartneritega.

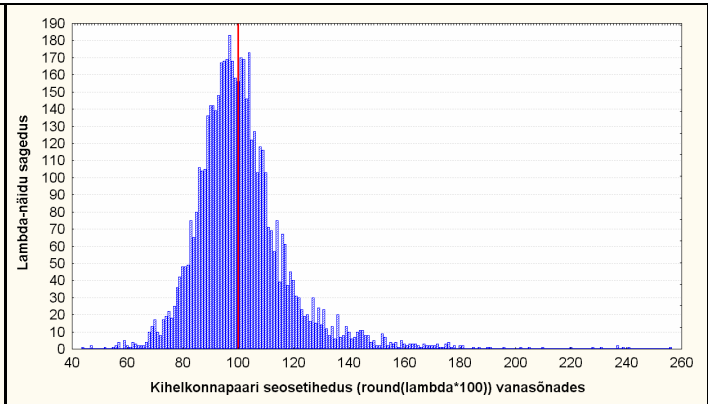
Lk 58 graafikolmikute viimane rida näitab “siniste kastikeste” (st kihelkonnakeskmeid ühendavate sirgete keskpunktide) köiepuntra taolisi teekondi läbi kõigi kolme λ -pingerea (näidud silutud libiseva keskmise meetodil: 3 iteratsiooni, silumisvahemik – pingerea 50 astet). Pärast esialgseid keskmistumisi Lõuna-Eesti ja saarte seosekolde vahel suundub murdesõnade köis loodesse, kus saarte nidum kiiresti Loode-Eestisse laieneb, mõistatuste ja vanasõnade oma aga kirde suunas, kuna neis žanrides kanduvad seosed sügavamale Põhja-Eestisse. Pärast kaootilist keerlemist naasevad köite lõpuotsad taas alguspunktide suunas – st vastavalt kagusse ja edelasse, sest siin hakkavad domineerima Lõuna-Eesti kihelkondade kõige negatiivsemad seosed kaugeimate Põhja-Eesti partneritega.



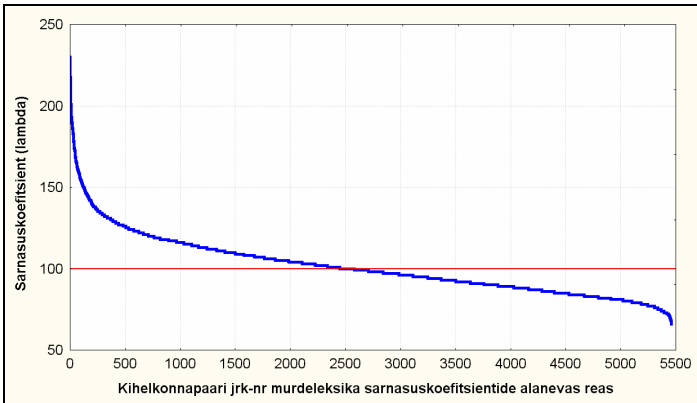
Murdesõnad



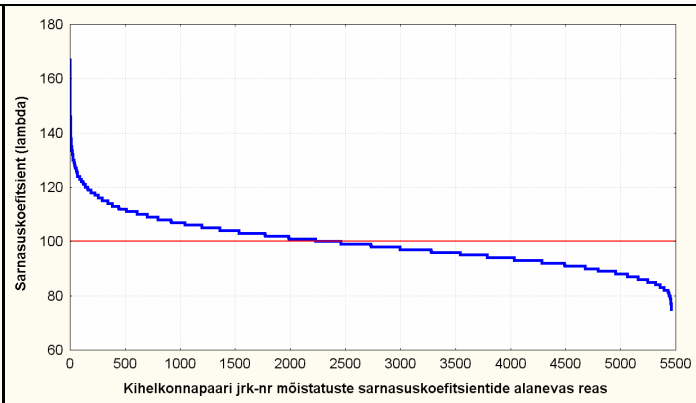
Mõistatused



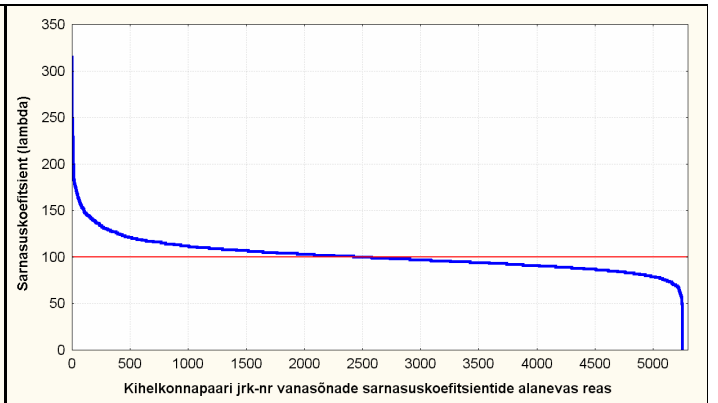
Vanasõnad



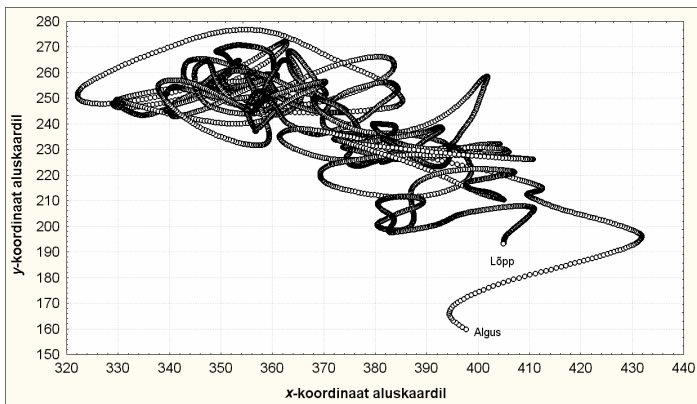
Murdesõnad



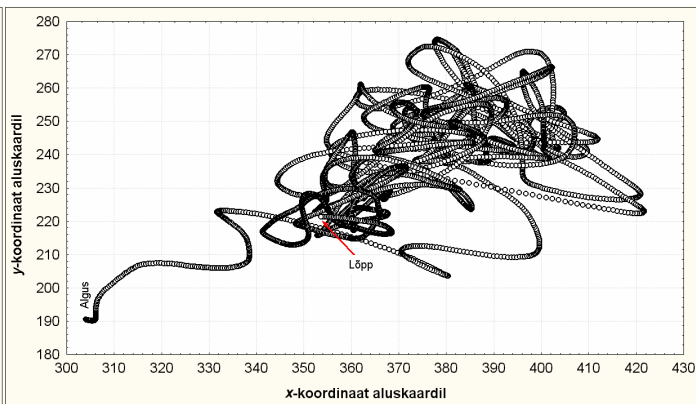
Mõistatused



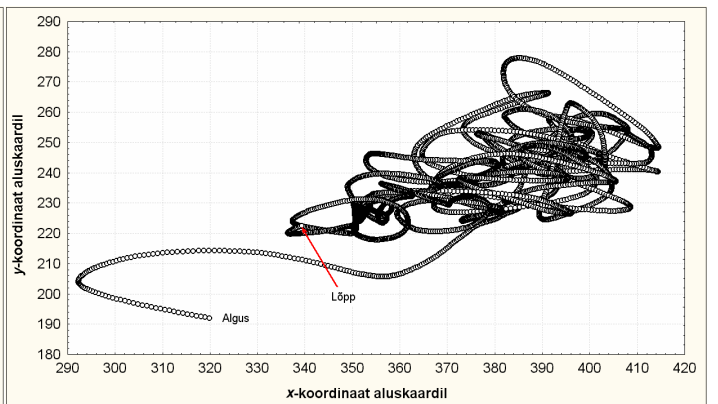
Vanasõnad



Murdesõnad

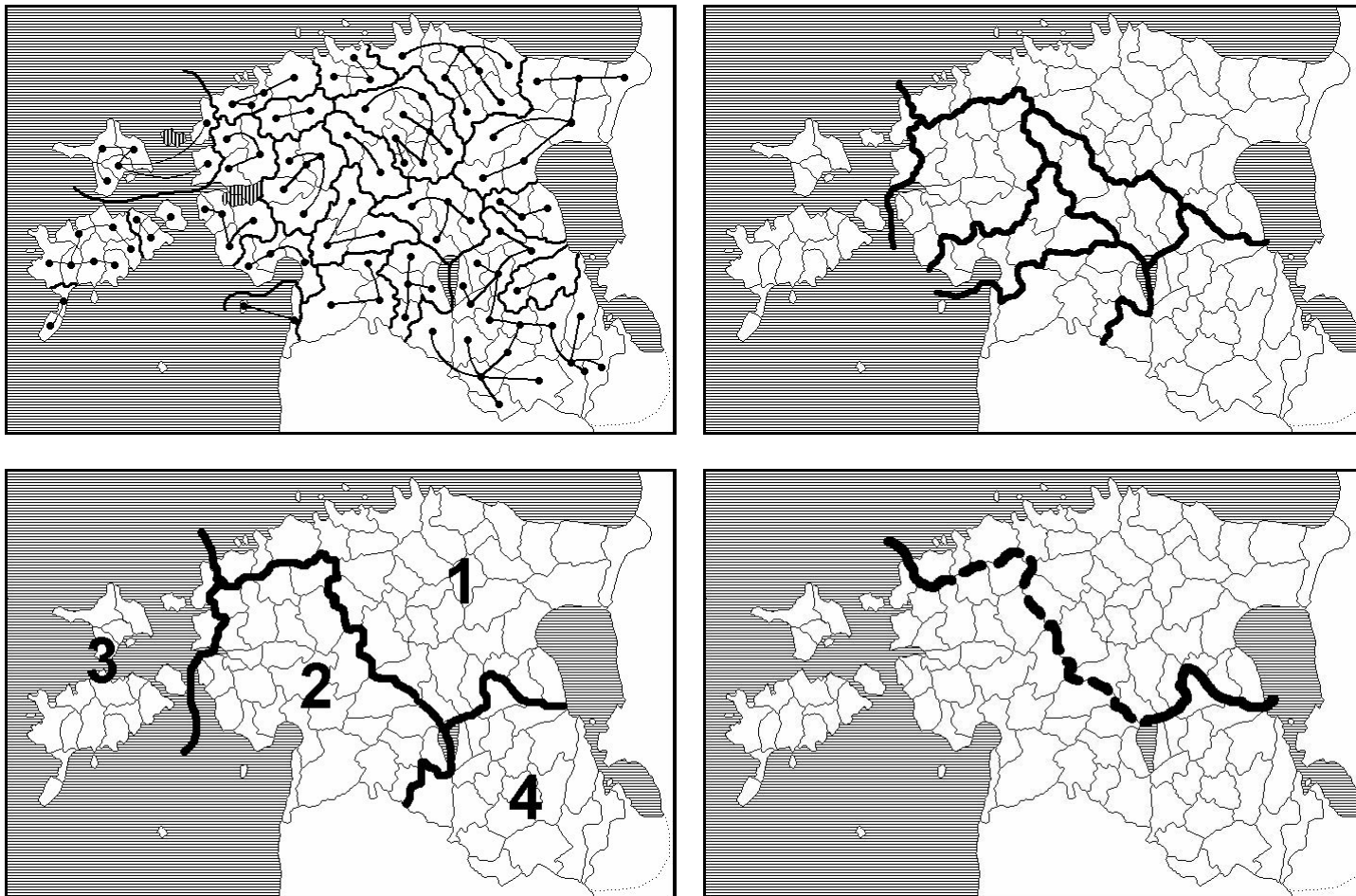


Mõistatused



Vanasõnad

Taipasin ühtlasi, et mu ammune katse “Eesti vanasõnaraajoonide” leidmiseks (vt Krikmann 1980) oli olnud meetoodiliselt väär, kuna kahepoolsed 1. astme rajoonid tekitati nii, et tugevaimatest seostest alates hakati maad täitma n a a b e r kihelkondi ühendavate graafidega ja jätkati seda seni, kuni iga kihelkond oli haaratud mõnesse graafi ning “veelahkmeid” graafide vahel käsitati 1. astme rajoonide piiridena. Edasi hinnati 1. astme rajoonide vahelisi seose- tihedusi ja sarnaseima sisuga rajoonid sulatati üheks suuremaks rajooniks ning jätkati seda seni, kuni maad läbis üksainus “viimne piir”, nii nagu näitavad neli järgnevat kartogrammi:



Mu järgnevad kogemused, sh siinse kolme andmestikuga, on siiski näidanud, et kuigi tugevaimad seosed tekivad tõesti just perifeersete naaberkihelkondade vahele, siiski üsna pea

(ja eriti Põhja-Eestis) sugeneb rohkesti positiivseid suhteid ka mitte-naaberkihelkondade vahel, mis muudavad sellise rajoneerimisviisi küsitavaks.

Kõik varasemad ja siinsed vaatlused jätavad igatahes mulje, et geoüksuste perifeersus / südamaisuus, aine unikaalsus / stereotüüpsus ja geoüksustest moodustuvate gruppide koherents / hajusus on sisuliselt ühe ja sama asja erinevad mõõtmed.

Vaadeldud andmestikes nähtuvad kaks markantset ja üks vähem eenduv keele- ja traditsiooniala:

- 1) Kagu-Eesti laias mõttes (Se + Võ + TaL) koos nende külge haakuva hajusama Mulgiga;**
- 2) saared koos nende külge haakuva hajusama Lääne- ja Loode-Eestiga;**
- 3) Eesti põhja- ja kirderannik.**

Ka Põhja-Eesti keskosas kehtib üldiselt reegel “Mida ligemal, seda sarnasemad”, kuid mingeid suuremaid selgemalt koonduvaid georühmi ei nähtu.

Faktor- ja klasteranalüüsi katsed

Kontrollimaks neid oletusi suuremal proovitükil tegin järgmise lisavaatluse.

Summeerisin samades positsioonides paiknevad λ -näidud läbi kõigi kolme maatriksi (murdesõnad + mõistatused + vanasõnad), käsitasin neid summasid algandmetena ja arvutasin nende põhjal lahtritele uued standardiseeritud λ -koefitsiendid.

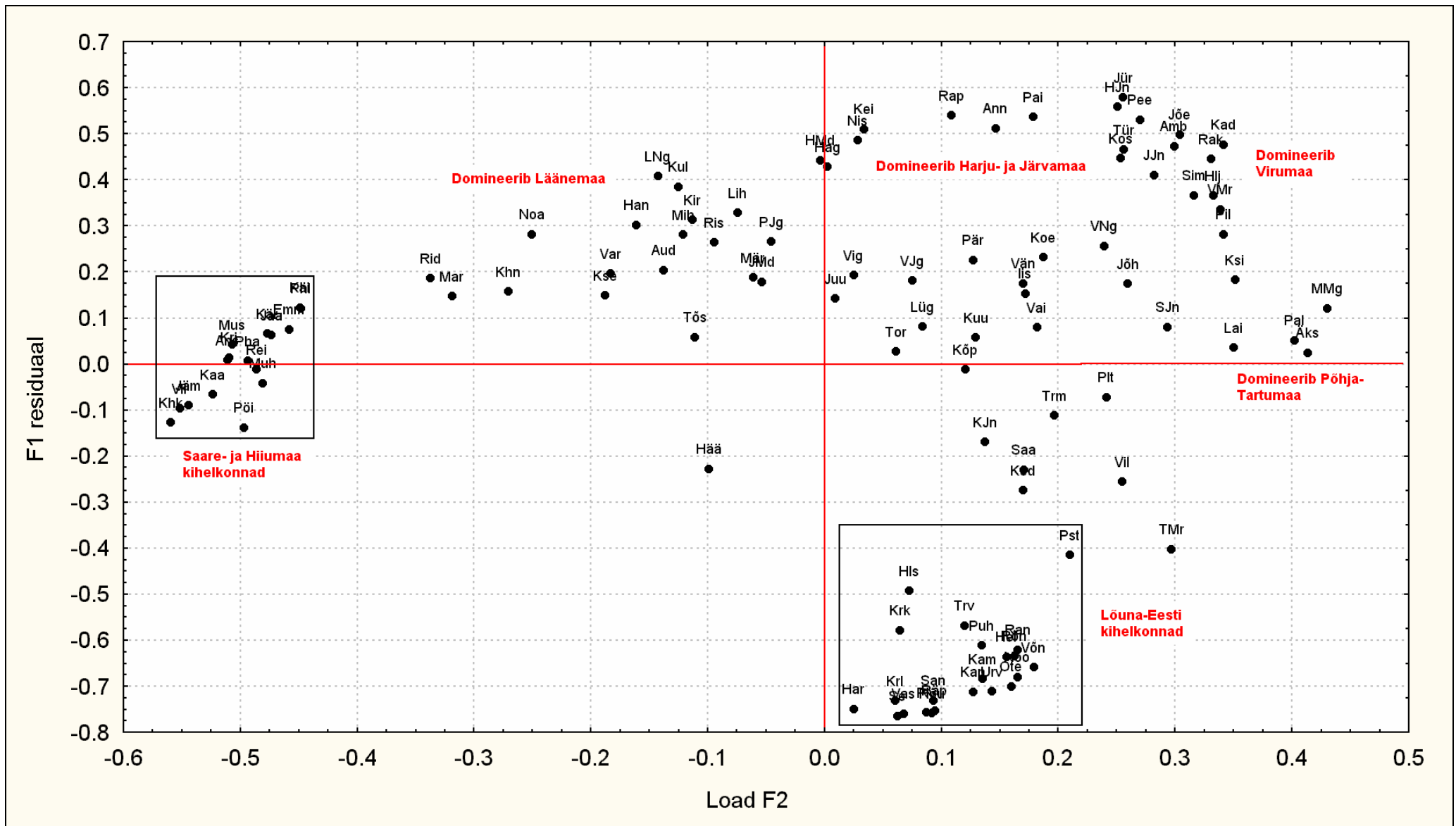
Valisin Statistica 5.5 faktoranalüüsi moodulist optsioonide jada:

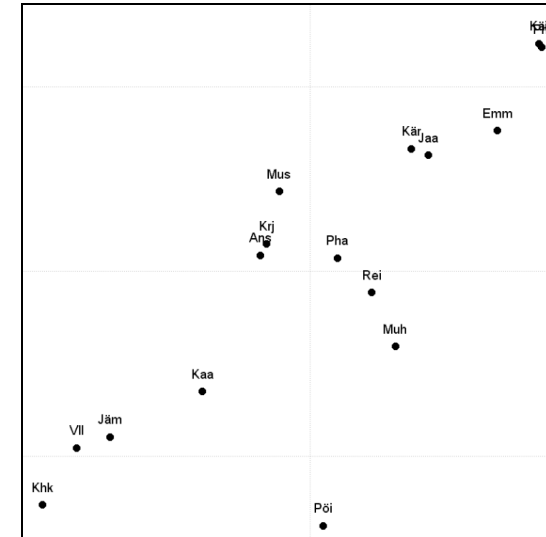
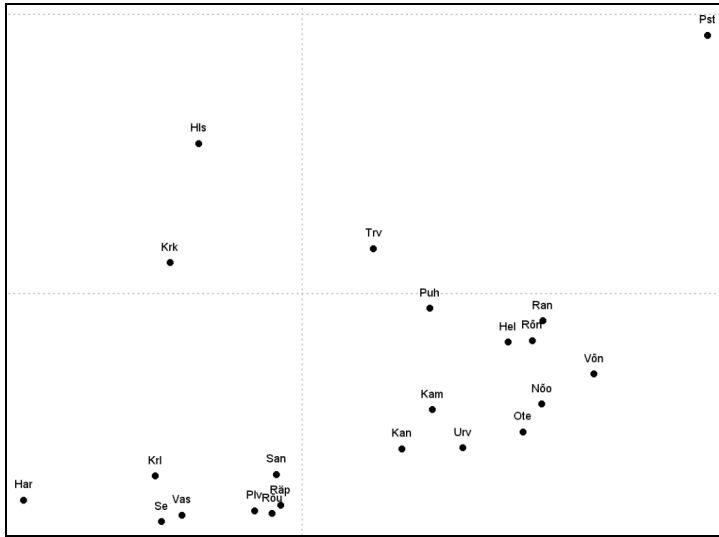
Input file: Raw data. MD deletion: Casewise → Variables: Select All → Extraction method: Principal components. Maximum no. of factors: 3. Minimum eigenvalue: 1.000 → Factor rotation: Varimax raw → Plot of loadings: 2D → Factor 2 versus Factor 1.

Selle tulemusel tekkis graaf, mis meenutas äratuntavalt Eesti kihelkondade geograafilist konfiguratiooni ning kihelkondade koonduvust ilmakaarte ja südamaisuuse / perifeersuse lõikes. Faktorite 1 ja 2 vahel selgus olevat kerge positiivne lineaarne suhe; pärast selle väljanormeerimist suurenes sarnasus Eesti kaardiga veelgi (vt graaf lk 61).

Graafi paremas allservas näeme Lõuna-Eesti klastrit tiheda Kagu-Eesti ja hajusama Mulgi komponendiga. Vasakul on teine ilmne tihend – saarte rühm. Sellest paremale eristub suhteliselt selgesti pms Lääne- ja Loode-Eesti kihelkondadest moodustuv hajusam jada, mis edasi alla paremale liikudes teiseneb pms Harju- ja Järvamaa, siis Virumaa ja edasi Põhja-Tartumaa kihelkondadest koosnevaks kaareks. Selle kaare ja Lõuna-Eesti klastri vahele graafi paremkeskmisse ossa jääb kõige hajusam piirkond, kus domineerivad Põhja-Pärnumaa ja Põhja-Viljandimaa kihelkonnad. On huvitav mainida, et Häädemeeste kihelkond, mille aines

koosneb valdavalt legendaarse murde- ja rahvaluulekoguja Marta Mäesalu kirjapanekuist, paikneb uhkes üksinduses telgede ristumispunkti lähedal.

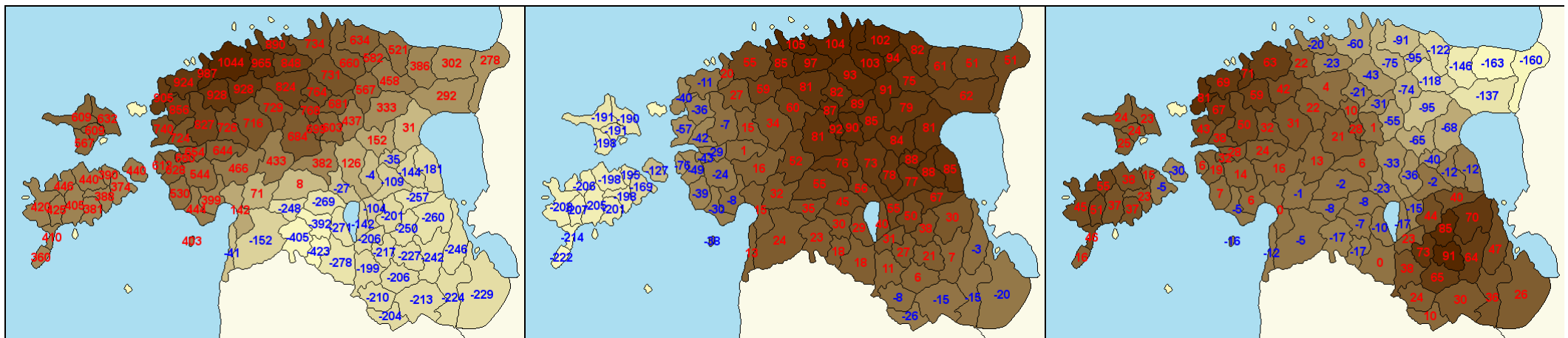




Lõuna-Eesti sülempi suurenemine näitab selgesti oma Võru-Setu, Lõuna-Tartumaa ja Mulgi allosade lähedusi ja tihedusi

Saarte sülemis pole selgemaid allrühmi näha – nt Sõrve kihelkonnad asuvad teineteisest üsna kaugel; Käina ja Pühalepa on väga lähedistiku, kuid Emmaste ja Reigi neist tükk maad eemal jne.

Kui kihelkondade faktoriskoore (tehniliselt: nende täisarvulisi korrutisi 100-ga 2 korda silutuna) mõtestada mingite neid iseloomustavate arvparameetritena, siis saaksime sellised kartograafilised interpretatsioonid:



Faktor 1 tähendab loodepoolsust

Faktor 2 tähendab “põhja-kirdepoolset südamaisust”?

Faktor 3 tähendab “loode-kagusuunalist perifeersust”?

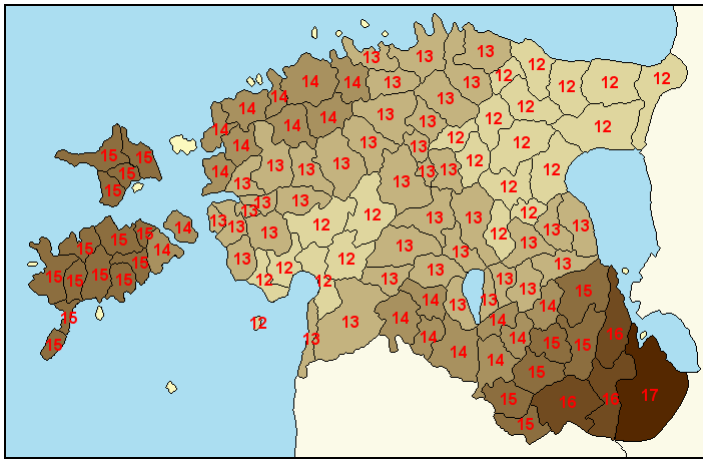
Programmil Statistica 5.5 on ka mõningad klasteranalüüsi võimed.

Valisin selle klasteranalüüsi moodulist opsioonide jada

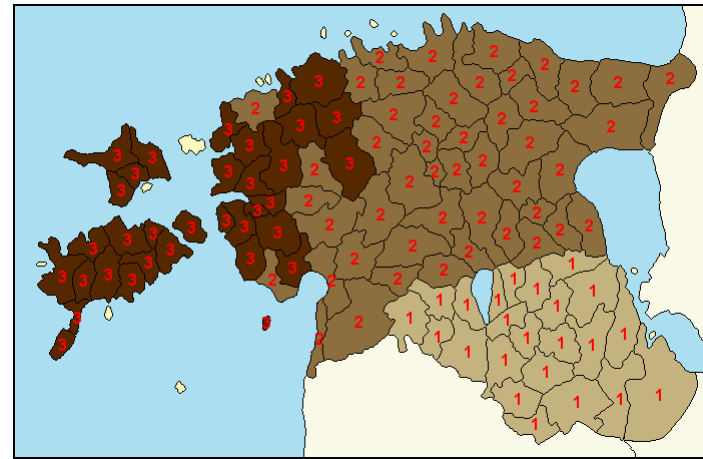
K-means clustering → Variables: Select All → Number of clusters: 3. Number of iterations: 10.

Missing data: Casewise deleted. Choose observations to maximise initial between-cluster distances.

Kaks järgnevat kaarti toovad tulemuste geograafilised interpretatsioonid:



Klastrite vahelised distantid (2 ringi silutuna)

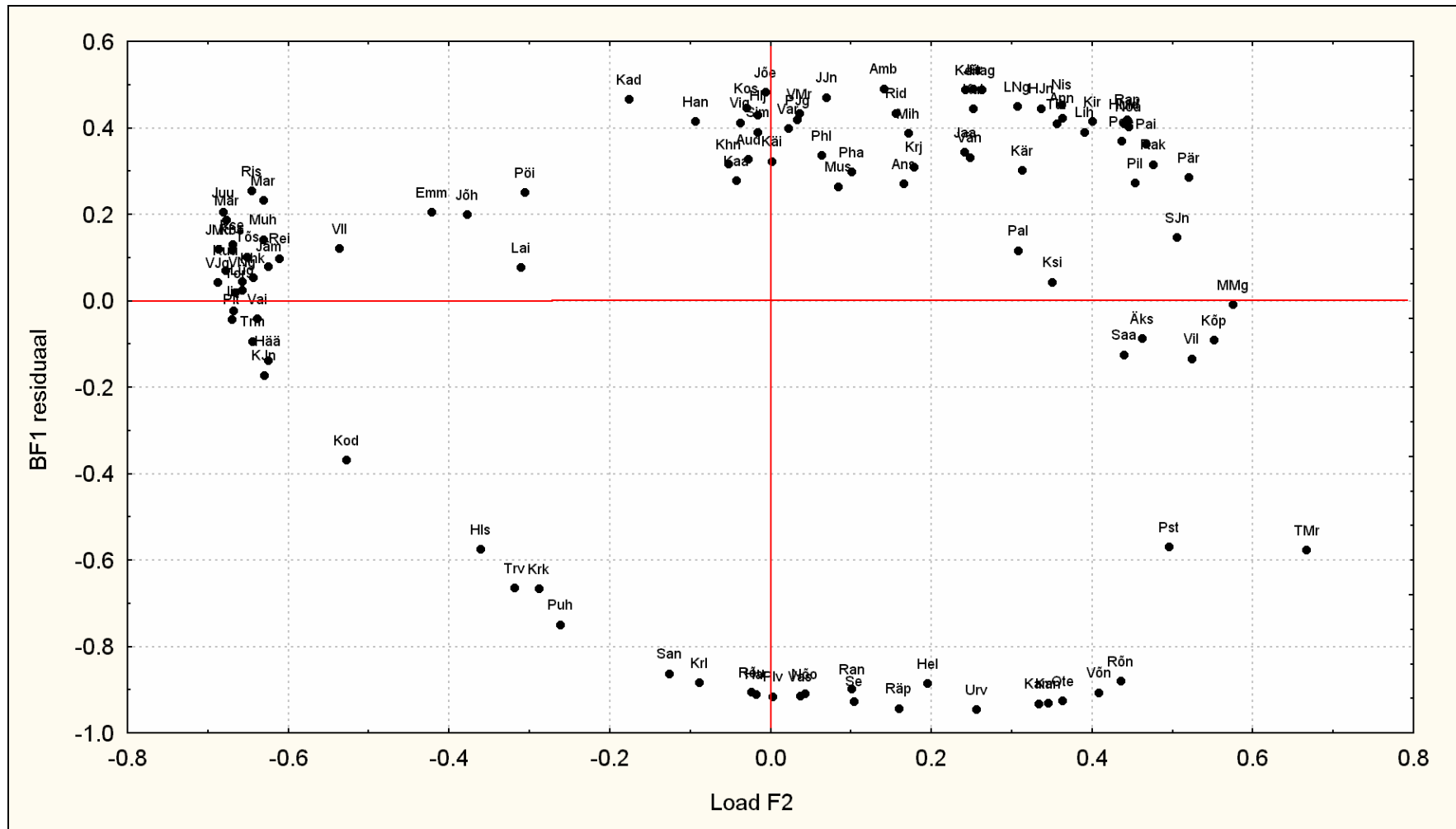


Klastrid (näidud tähendavad jrk-numbreid)

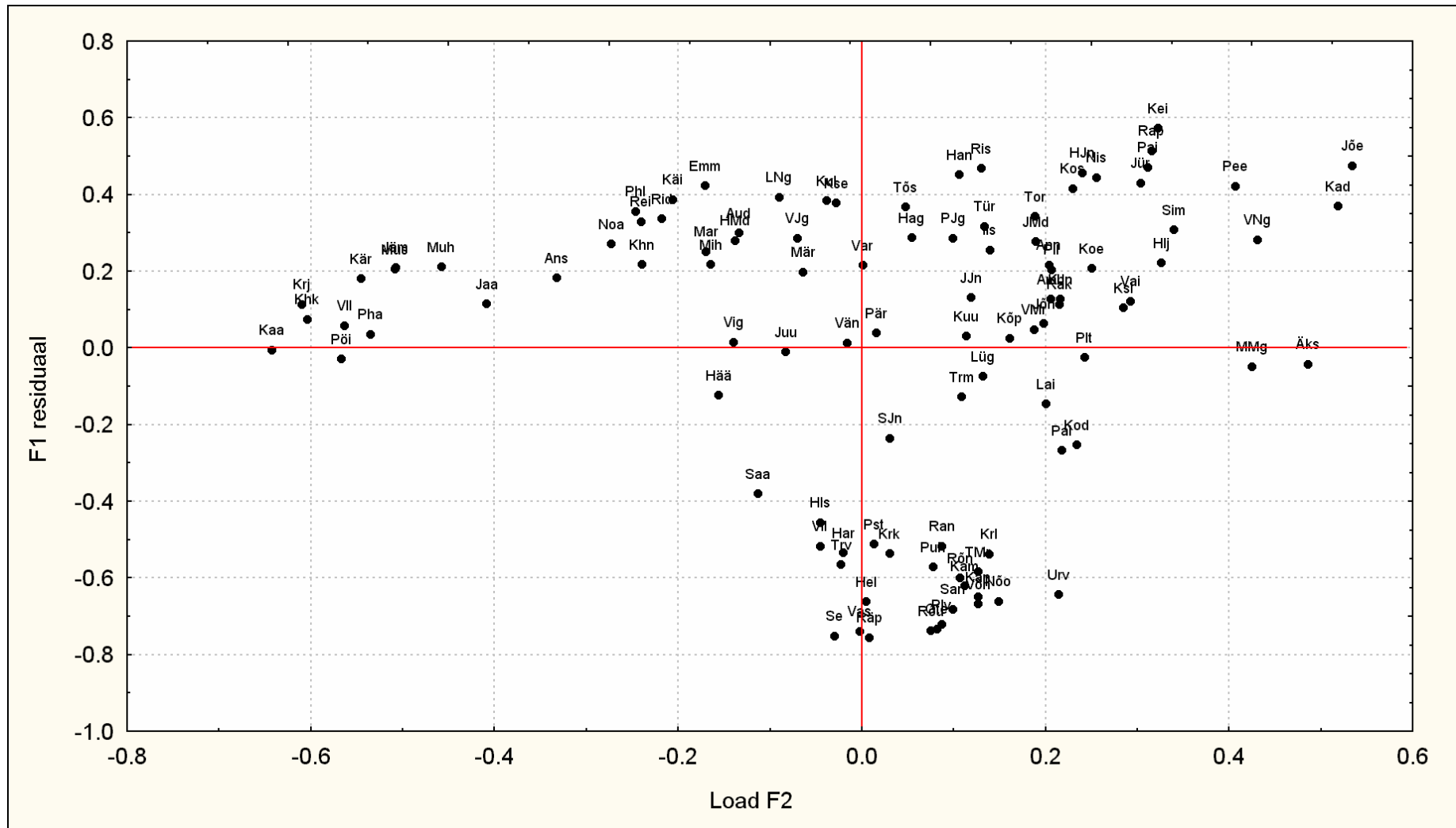
Näeme, et ka see kolmikjaotus vastab üldjoontes eelöeldud kujutelmadele: maa jaguneb Lõuna-Eestiks, saare-lääne-loodeosaks ning põhjaeestiliseks “kõigeks ülejäänuks”.

Katsetasin faktoranalüüsi võimalusi ka murdesõnade, mõistatuste ja vanasõnade λ -maatriksitel eraldi. Nende graafilised tulemused olid mitmekesisema väljanägemisega ega olnud nii heas vastavuses Eesti geooksuste tegeliku konfiguratsiooniga.

Alloleval murdesõnade graafil moodustub ümber telgede lõikepunkti faktoranalüüsi tulemustes sagedane sõõritaoline formatsioon, kuid murdealalises plaanis eristub siin selgesti vaid alumine “Lõuna-Eesti kaar”, kogu ülejäänud (ülalpoolses) Põhja-Eesti osas paikneb kaootiliselt läbisegi nii lääne-, põhja- ja idapoolseid kui ka keskeetilisi geooksusi.



Vanasõnade menetlusvariant oli täpselt samane sellega, mida katsetati koondmaterjalil (vt eespool, lk 60), ning ka tulem sellega kõige sarnasem. Lõuna-Eesti sülem eristub selgesti, kuid sisaldab ka mõningaid lähikihelkondi (Vil, TMr). Hiiumaa kihelkonnad on saarte sülemi küljest lahti veninud ning sulanud kokku Loode- ja Lääne-Eesti kihelkondadega. Hajumi põhja- ja idapoolsed servad on täidetud pms Harju-, Viru- ja Põhja-Tartumaa kihelkondadega, stereotüüpseimad Kesk-Eesti alad jäävad taas kõigi ülejäänute keskele, kõige lähemale telgede lõikepunktile.



*

Need lisavaatlused kinnitavad seega üldiselt siin lk 60 tehtud oletusi, kuid kirde-rannaeestilise levikuala olemasolu ei tõuse faktor- ega klasteranalüüsi testides üldsegi esile.

Epiloogi asemel

14 aastat tagasi, kui kirjutasime koos Karl Pajusaluga artiklit keskmurde keskpunkti kohta (Krikmann ja Pajusalu 2000), tegin ühtlasi hulganisti mitmesuguseid regressiooni- ja λ -arvutusi kombineerivaid vaatlusi keskmisest unikaalsemate murdesõnadega, mis olid tuntud 2–33 kihelkonnas (vt Krikmann 2000b). Vaatasin need tabelid uuesti läbi ja katsetasin neid faktoranalüüsi objektidena, enamasti endises, lk 60 näidatud režiimis:

Input file: Raw data. MD deletion: Casewise → Variables: Select All → Extraction method: Principal components. Maximum no. of factors: 3. Minimum eigenvalue: 1.000 → Factor rotation: Varimax raw → Plot of loadings: 2D → Factor 2 versus Factor 1.

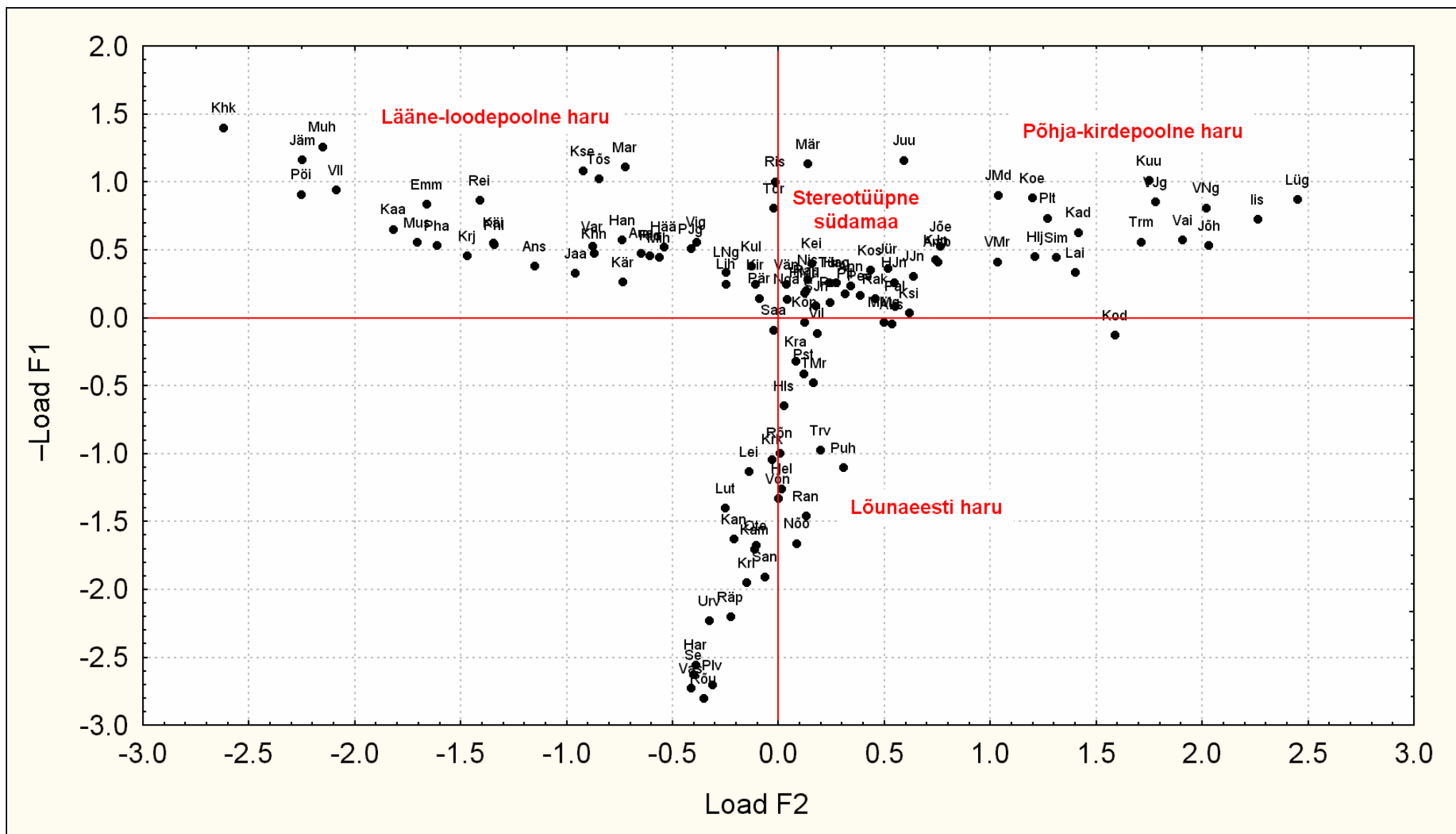
Kuna vanasõnade graafil (lk 66) eendus kirdepoolne “Harju-Viru nukk” juba päris selgesti, tegin samasuguse katse ka 2–33 kihelkonnast noteeritud mõistatuste ja vanasõnadega.

Meie kolmes vaatlusaines olid 2–33 kihelkonnas esindatud sisuüksuste (sõnade, tüüpide) hulgad järgmised:

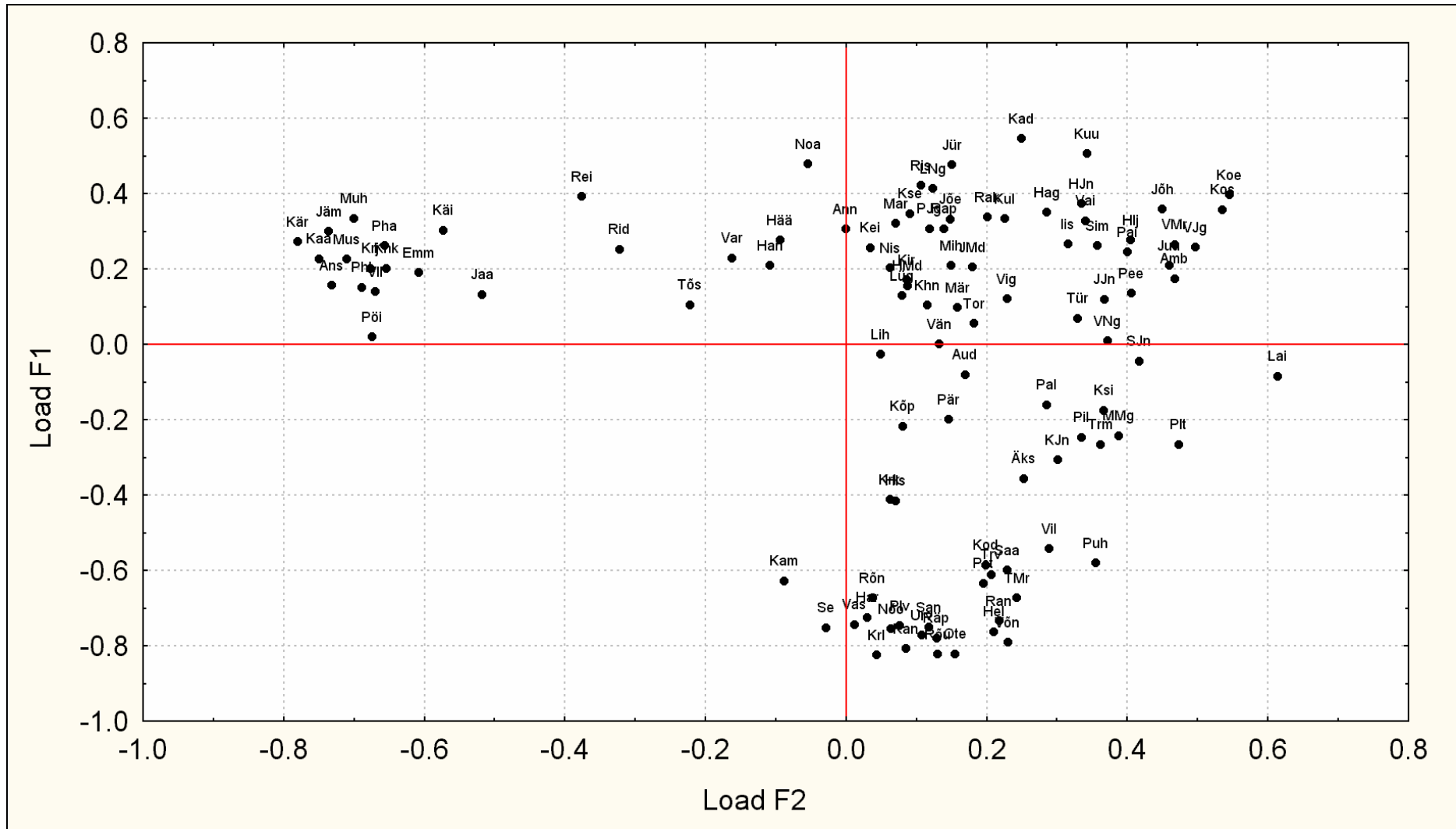
murdesõnu – 36978 (koos keelesaartega 37359);
mõistatusi – 1164;
vanasõnu – 4635.

Unikaalsemate murdesõnade faktorigraafidest jättis kõige sügavama mulje lk 68 toodu, mille aluseks olevad arvandmed on saamisviisilt lähedased parem-keskmistele kartogrammidele kihelkondade murdesõnalisi suhtluspilte näitavates kuuikutes saidis Krikmann 2000b.

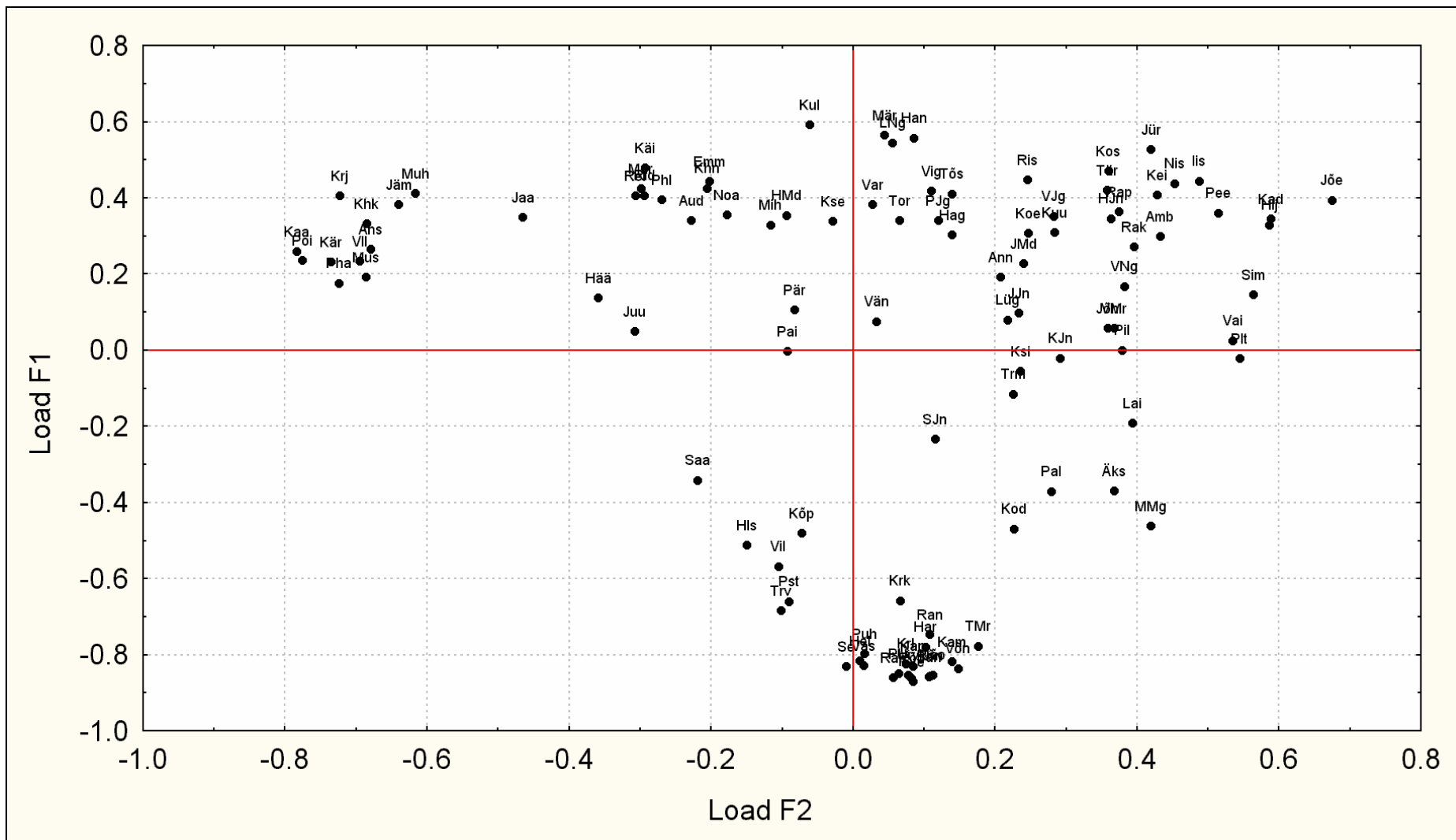
Siin sirutuvad selgesti ja ilmakaareliselt enam-vähem õigetes suundades kolm unikaalsemat poolsaart – lääne-loodepoolne, põhja-kirdepoolne ja lõunaeestiline, ning nende ristumiskohta jääb stereotüüpne südamaa.



Unikaalsemate mõistatuste graaf, mille lähteandmed on saadud siin kasutatud tavalise λ -protseduuriga, eristab selgesti kaks saart – Saaremaast ja mõnedest Hiiumaa kihelkondadest koosneva ning Lõuna-Eesti. Pms Läänemaa ja Põhja-Tartumaa kihelkondadest koosnevad riburajad ühendavad neid parem-ülapoolse hajusa Põhja-Eesti platsdarmiga, mille kirdenurka hakkab tekkima unikaalsematest kihelkondadest koosnev tipp.



Unikaalsemate vanasõnade graaf (lähteandmed samuti siinse tavalise λ -protseduuri tulem) näitab kaht eraldiseisvat tihedakoelist saart – Saaremaad (v.a Jaani) ja Kagu-Eestit, ning kaht hajusamat – Mulgi ja Põhja-Tartumaa sülemit. Geograafiliselt segase koostisega Põhja-Eesti hajumist graafi parem-ülalpooles osas küünitub lääne suunas kitsas Pärnu-Lääne-Hiiumaa kihelkondadest koosnev poolsaar.



Seega näib, et faktoranalüüsi võimed Eesti murde- ja folkloorialade kaardistamiseks küünivad kaugele üle metafoorsuse piiride. Hulganisti Eesti kihelkondade murde- ja folklooriainese sarnasust mõõtvaid λ -hindeid, mis ei kätkenud mingit geograafilist informatsiooni, sisestati analüsaatorisse, mis toimis samuti vaid puhtabstraktsete ruumide tasandil, ja saadi vastuseks hulganisti graafilisi väljundeid, mis meenutasid Eesti geograafilist kaarti ning andsid ühtlasi kvaasigeograafilises mõõdustikus teavet sellel paiknevate objektide lähedussuhete kohta.

Kirjandus

- EV IV = Eesti vanasõnad IV. Lisad. Toim A. Krikmann ja I. Sarv. Tallinn: Eesti Raamat, 1988.
- Krikmann, Jaak, Krikmann, Arvo (koost) 2012. Eesti mõistatused. Andmebaas. Tartu: EKM FO. – <http://www.folklore.ee/moistatused/> (8. VI 2014)
- Krikmann, Arvo 1979. Some aspects of proverb distribution. – Symposium “Mathematical processing of graphic data”. Tallinn, 28–44.
- Krikmann, Arvo 1980. Towards the typology of Estonian folkore regions. Preprint KKI-16. Tallinn, ENSV TA.
- Krikmann, Arvo 1986. Fraseoloogiline aines eesti vanimais grammatikates ja sõnastikes. Tallinn: Valgus.
- Krikmann, Arvo 1997. Sisesevaateid folkloori lühivormidesse I: Põhimõisteid, žanrisuhteid, üldprobleeme. Tartu: TÜ Kirjastus.
- Krikmann, Arvo [1998]. Eesti kihelkondade suhtlemistihedus mõistatuses. Tartu: EKM FO. – <http://www.folklore.ee/~kriku/LEVIK/coefs.htm> (8. VI 2014)
- Krikmann, Arvo (koost) 2000a. Tere teele, tere meele, tere egalõ talolõ. Valik lõunaeesti mõistatusi. Tartu.
- Krikmann, Arvo [2000b]. Eesti kihelkondade murdesõnavaralisi suhteid. Tartu: EKM FO. – <http://www.folklore.ee/~kriku/MURRE/Index.htm> (8. VI 2014)
- Krikmann, Arvo ja Pajusalu, Karl 2000. Kus on keskmurde keskpunkt. – Inter dialectos nominaque. Pühendusteos Mari Mustale 11. novembril 2000. Koost ja toim Jüri Viikberg. EKI Toimetised 7. Tallinn: EKSA, lk 131–172.
- Rańłowski, Zbigniew = Павловский, Збигнев 1967. Введение в математическую статистику. Перевод с польского В. Д. Меникера; под ред. Ф. Д. Лившица. Москва : Статистика.
- Rommel, Mart 1979. Large data bases in quantitative history: Some handling techniques. Preprint KKI-12. Tallinn, ENSV TA.
- VMS = Väike murdesõnastik I–II. Toim Valdek Pall. Tallinn: Valgus, 1982–1989.