

3.4.1.1 Classificare i dati con i metodi degli intervalli costanti e della deviazione standard

Ipotesi Dati presenti nella cartella ECDLGIS_SYLLABUS con il seguente percorso: C:\GISeQGIS_DATI\ECDL_SYLLABUS
Carichiamo:

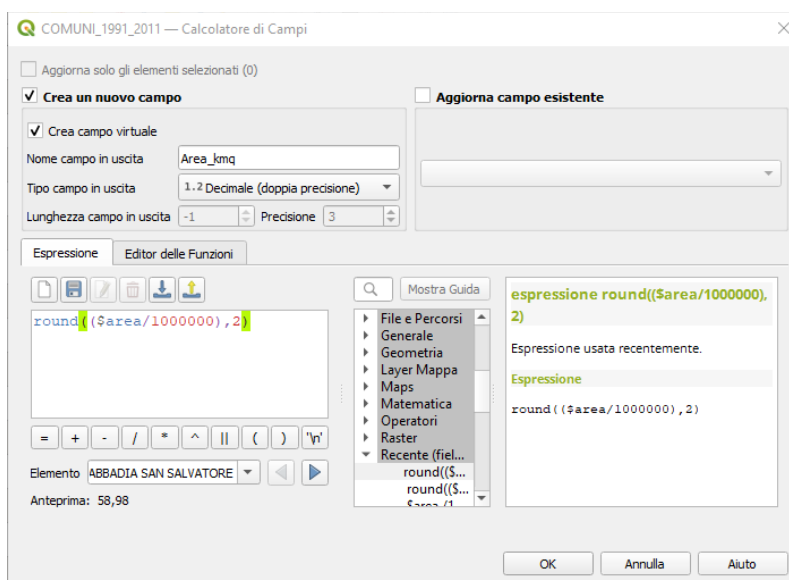
- il layer vettoriale **COMUNI_1991_2011** che ha la seguente tabella attributi:

	ISTAT1991	NOME	PRO1991	PRO2011
1	09052036	TREQUANDA	52	52
2	09052033	SINALUNGA	52	52

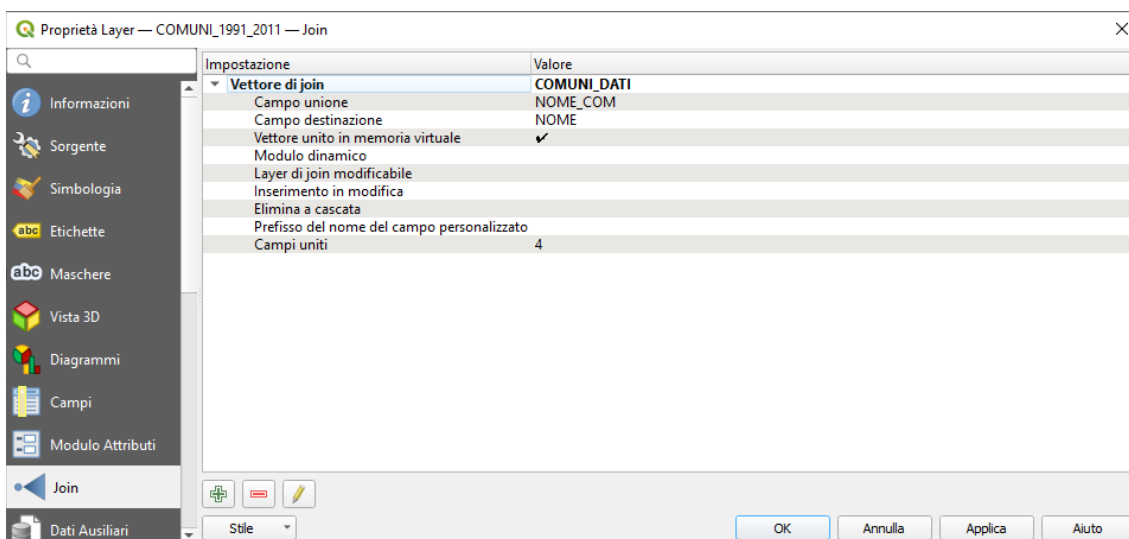
- il layer testo delimitato **COMUNI_DATI** che ha la seguente tabella attributi:

	NOME_COM	CODPR_1991	PRCOM_1991	CODPR_2011	PRCOM_2011	POP_1991	POP_2011	ALT_MIN	ALT_MAX
1	ABBADIA SAN SALVATORE	52	52001	52	52001	7243	6557	318	1725
2	ABETONE	47	47001	47	47001	758	687	821	1938
3	AGLIANA	47	47002	47	47002	13410	16792	36	50

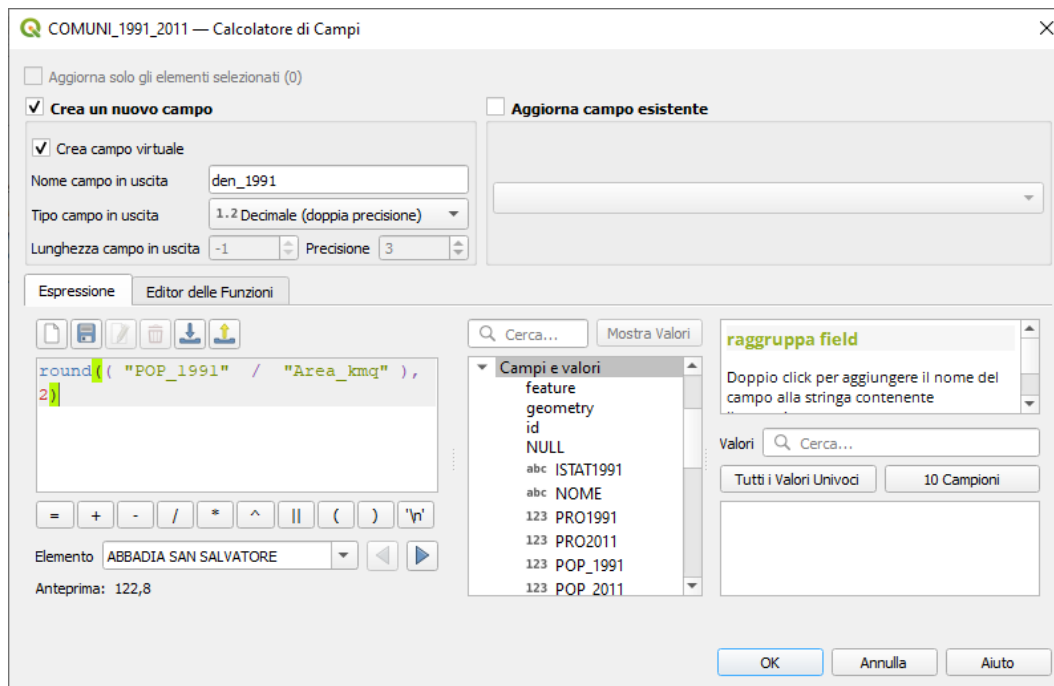
nella tabella dati di **COMUNI_1991_2011** aggiungiamo un *campo virtuale* di nome *Area_kmq* con il calcolo dell'area di ogni comune tramite il **Calcolatore di Campi** così impostato:



effettuiamo poi il join tra **COMUNI_1991_2011** campo *NOME* e **COMUNI_DATI** campo *NOME_COM* :



calcoliamo la densità di popolazione nel 1991 abitanti per kmq e memorizziamola in un ulteriore campo virtuale di nome *den_1991* della tabella attributi di **COMUNI_1991_2011** tramite il **Calcolatore di Campi** così impostato:



vediamo come, a seguito di quanto fatto, risulta modificata (solo temporaneamente nel progetto in corso) la tabella attributi di **COMUNI_1991_2011** :

Id	Nome	Alias	Tipo	Nome tipo	Lunghezza	Precisione	Commento
abc 0	ISTAT1991		Testo (stringa)	String	16	0	
abc 1	NOME		Testo (stringa)	String	50	0	
123 2	PRO1991		Intero (32 bit)	Integer	5	0	
123 3	PRO2011		Intero (32 bit)	Integer	5	0	
◀ 4	POP_1991		Intero (32 bit)	integer	0	0	
◀ 5	POP_2011		Intero (32 bit)	integer	0	0	
◀ 6	ALT_MIN		Intero (32 bit)	integer	0	0	
◀ 7	ALT_MAX		Intero (32 bit)	integer	0	0	
⊗ 8	Area_kmq		Decimale (doppia precisione)	double precision	-1	0	round((\$area /1000000),2)
⊗ 9	den_1991		Decimale (doppia precisione)	double precision	-1	0	round(("POP_1991" / "Area_kmq"),2)

su sfondo verde chiaro e celeste chiaro i campi temporanei aggiunti, in visualizzazione la tabella attributi è:

	ISTAT1991	NOME	PRO1991	PRO2011	POP_1991	POP_2011	ALT_MIN	ALT_MAX	Area_kmq	den_1991
1	09052001	ABBADIA SAN SALVATORE	52	52	7243	6557	318	1725	58,98	122,8
2	09047001	ABETONE	47	47	758	687	821	1938	31,02	24,44
3	09047002	AGLIANA	47	47	13410	16792	36	50	11,68	1148,12
4	09046001	ALTOPASCIO	46	46	9976	15072	2	79	28,61	348,69

Su questa tabella attributi faremo gli esempi:

- di questo esercizio ES_SYL_3_4_1_1 Classificare i dati con il metodo degli intervalli costanti e della deviazione standard
- del successivo ES_SYL_3_4_1_2 Classificare i dati con il metodo dei quantili
- e del successivo ES_SYL_3_4_1_3 Classificare i dati con il metodo del natural breaks

e' opportuno effettuare preventivamente una analisi dei dati sui valori della densità di popolazione nei 287 comuni, in dettaglio:

MONTEVERDI MARITTIMO	7,69	FOSCIANDORA	34,83	DICOMANO	74,16	PORTO AZZURRO	233,21
RADICONDOLI	7,78	CASTELNUOVO BERARDENGA	35,69	LUCIGNANO	74,77	AREZZO	238,22
RADICOFANI	11,01	FABBRICHE DI VALLICO	38,88	MONTESPERTOLI	75,47	LARCIANO	242,81
SAN GODENZO	11,14	VILLA COLLEMANDINA	38,90	CAMPO NELL'ELBA	76,28	PONTE BUGGIANESE	246,24
ROCCALBEGNA	11,68	GUARDISTALLO	39,71	CIVITELLA IN VAL DI CHIANA	76,41	SAN MINIATO	247,19
BADIA TEDALDA	11,70	CASTELLINA MARITTIMA	39,87	VAGLIA	77,15	ROSIGNANO MARITTIMO	248,72
PALAZZUOLO SUL SENIO	12,13	MOLAZZANA	40,10	LICCIANA NARDI	79,23	CARMIGNANO	249,39
SILLANO	12,73	PRATOVECCHIO	40,66	FOSDINOVO	81,90	PIAN DI SCO	250,63
MONTICIANO	13,19	PITEGLIO	41,16	MONTEPULCIANO	83,80	VINCI	253,63
MONTETROTTONDO MARITTIMO	13,56	CASTIGLIONE DI GARFAGNANA	41,52	PIEVE FOSCIANA	84,90	CASTELFIORENTINO	258,20
CHIUSDINO	13,57	CUTIGLIANO	41,57	VERNIO	86,20	VAIANO	259,17
MONTIERI	13,85	MULAZZO	42,08	MONTE SAN SAVINO	87,35	MONTECARLO	259,25
CAPRAIA ISOLA	13,85	PITIGLIANO	42,35	TERRICCIOLA	88,09	SANTA MARIA A MONTE	272,69
MONTECATINI VAL DI CECINA	14,06	BIBBONA	42,52	CASCIANA TERME	88,15	MASSAROSA	276,60
SAN GIOVANNI D'ASSO	14,12	ARCIDOSSO	44,35	SAN MARCELLO PISTOIESE	89,95	CAPANORI	281,15
CAMPAGNATICO	15,23	BAGNI DI LUCCA	44,63	CAMPORGIANO	90,85	PIOMBINO	281,99
MURLO	15,65	ANGHIARI	44,91	COREGLIA ANTELMINELLI	90,99	VICOPISSANO	282,04
COMANO	15,97	VICCHIO	45,17	CAPOLONA	91,39	LAMPORVECCHIO	292,02
CIVITELLA PAGANICO	16,02	STAZZEMA	45,39	PODENZANA	95,46	MONTOPOLI IN VAL D'ARNO	293,42
MAGLIANO IN TOSCANA	16,27	MINUCCIANO	46,73	PIAZZA AL SERCHIO	98,63	MONTALE	305,23
VERGEMOLI	16,92	PONTREMOLI	47,31	MARCIANO DELLA CHIANA	101,14	SAN GIULIANO TERME	306,52
SCANSANO	17,12	SANTA FIORA	47,42	CASTIGLION FIORENTINO	102,41	IMPRUNETA	308,46
CASOLE D'ELSA	17,28	BUONCONVENTO	47,86	RIO MARINA	102,51	FUCECCHIO	315,03
FIRENZUOLA	17,81	STIA	48,06	SAN PIERO A SIEVE	103,12	SERAVEZZA	321,73
SEMPRONIANO	17,91	GAVORRANO	48,49	REGGELLO	103,58	MARCIANA MARINA	330,15
GAIOLE IN CHIANTI	17,92	TRESANA	48,81	BORGO SAN LORENZO	104,46	ALTOPASCIO	348,69
CINIGIANO	18,65	ORCIANO PISANO	48,84	BORGO A MOZZANO	105,10	FIESOLE	357,81
PIENZA	18,96	MARCIANA	49,50	CAVRIGLIA	110,75	CAMAIORE	359,21
SESTINO	19,07	SAN GIMIGNANO	50,32	RIGNANO SULL'ARNO	117,45	BAGNO A RIPOLI	369,58
MANCIANO	19,22	GAMBASSI TERME	50,40	CRESPINA	119,51	PISTOIA	371,78
CASTIGLIONE D'ORCIA	20,05	VOLTERRA	50,94	TERRANUOVA BRACCIOLINI	120,98	POGGIBONSI	373,53
LAJATICO	20,29	LORO CIUFFENNA	51,50	TAVARNELLE VAL DI PESA	121,18	PORCARI	377,60
RADDA IN CHIANTI	20,31	SARTEANO	51,65	TORRITA DI SIENA	121,45	FOLLONICA	381,24
LONDA	20,53	RIO NELL'ELBA	51,95	ABBADIA SAN SALVATORE	122,80	MONTETARCI	383,23
MONTEMIGNAIO	20,55	FILATIERA	52,92	GALLICIANO	126,69	LASTRA A SIGNA	405,78
TALLA	20,66	SOVICILLE	53,20	BIBBIENA	126,84	MONTETUPO FIORENTINO	407,78
MONTALCINO	20,87	LORENZANA	53,29	RUFINA	128,96	MASSA E COZZILE	411,86
SASSETTA	20,88	PESCAGLIA	53,31	PELAGO	132,32	QUARRATA	457,75
SAMBUCA PISTOIESE	21,10	BARBERINO VAL D'ELSA	53,64	LATERINA	137,80	CASCINA	461,43
ZERI	21,20	SAN ROMANO IN GARFAGNANA	53,76	COLLESALVETTI	139,68	LUCCA	468,53
PIEVE SANTO STEFANO	21,40	PECCIOLI	53,91	SINALUNGA	147,40	BUGGIANO	471,26
CAPALBIO	21,41	CHITIGNANO	54,98	SAN CASCIANO IN VAL DI PESA	148,51	SIENA	480,44
SAN CASCIANO DEI BAGNI	21,47	VILLA BASILICA	55,42	CAMPILIA MARITTIMA	150,11	UZZANO	512,24
TREQUANDA	21,49	MARLIANA	55,87	GROSSETO	150,39	PISA	532,64
CHIUSI DELLA VERNA	21,74	CETONA	56,56	BARGA	153,40	CALCI	543,46
SANTA LUCE	21,87	SAN QUIRICO D'ORCIA	56,73	VECCHIANO	153,99	CHIESINA UZZANESE	546,39
SEGGIANO	22,02	FIVIZZANO	56,76	CHIUSI	156,65	MONTIGNOSO	548,38
ORTIGNANO RAGGIOLO	22,15	SUBBIANO	57,18	VILLAFRANCA IN LUNIGIANA	161,32	MONTEMURLO	557,63
RIPARBELLA	22,39	POPPI	57,71	AULLA	167,86	MONSUMMANO TERME	558,63
SORANO	23,98	CASTAGNETO CARDUCCI	58,10	SANSEPOLCRO	172,57	PONTERERA	571,28
ABETONE	24,44	CASTEL FOCIGNANO	59,07	LARI	175,22	CECINA	572,93
CASTELLINA IN CHIANTI	25,14	PALAIA	59,91	PONTASSIEVE	178,68	PIETRASANTA	596,13
MARRADI	25,29	RAPOLANO TERME	59,93	BIENTINA	179,42	PONSACCO	616,10
CAPRESE MICHELANGELO	25,58	CAPOLIVERI	61,21	CERRETO GUIDI	181,57	PIEVE A NIEVOLE	668,95
CHIANNI	26,03	MONTERONI D'ARBIA	61,36	COLLE DI VAL D'ELSA	184,60	EMPOLI	699,49
CANTAGALLO	26,50	PIANCASTAGNAIO	63,22	FOIANO DELLA CHIANA	189,66	MASSA	708,24
ASCIANO	28,80	CASALE MARITTIMO	64,55	CALENZANO	194,32	SANTA CROCE SULL'ARNO	734,82
CASTELNUOVO DI VAL DI CECINA	30,08	CASTEL DEL PIANO	64,59	INCISA IN VAL D'ARNO	201,29	SIGNA	765,03
BAGNONE	30,39	ISOLA DEL GIGLIO	64,81	CASTELFRANCO DI SOTTO	202,66	SAN GIOVANNI VALDARNO	827,82
GIUNCUGNANO	30,66	MONTERRCHI	65,22	CHIANCIANO TERME	203,58	SCANDICCI	896,23
POMARANCE	31,26	ORBETTELLO	65,51	CAPRAIA E LIMITE	206,14	CARRARA	949,65
CAREGGINE	31,30	BARBERINO DI MUGELLO	65,53	SERRAVALLE PISTOIESE	208,08	SESTO FIORENTINO	971,83
SCARLINO	31,53	GREVE IN CHIANTI	65,76	MONTE ARGENTARIO	208,18	FORTE DEI MARMI	1071,40
CASOLA IN LUNIGIANA	31,83	CORTONA	65,94	CERTALDO	211,71	AGLIANA	1148,12
VAGLI SOTTO	32,13	BUCINE	66,54	SAN VINCENZO	215,79	MONTecatini-TERME	1168,16
MONTAIONE	32,36	MONTESCUDAIO	67,51	CAPANOLI	217,75	CAMPI BISENZIO	1197,64
CASTELLAZZARA	32,78	FAUGLIA	67,68	CALCI	218,67	POGGIO A CAIANO	1323,50
SUVERETO	32,91	CASTIGLION FIBOCCHI	67,99	FIGLINE VALDARNO	219,57	LIVORNO	1604,06
ROCCASTRADA	32,96	PERGINE VALDARNO	68,40	CASTELNUOVO DI GARFAGNANA	221,14	PRATO	1702,70
MASSA MARITTIMA	33,57	CASTELFRANCO DI SOPRA	70,68	BUTI	226,13	VIAREGGIO	1829,91
CASTEL SAN NICCOLO'	34,34	MONTERRIGIONI	71,56	PORTOFERRAIO	227,91	FIRENZE	3942,65
CASTIGLIONE DELLA PESCAIA	34,34	SCARPERIA	73,85	PESCAIA	228,94		

Visualizzando il **Pannello Statistiche** sul campo **den_1991** risulta:

Statistica	Valore
Numero	287
Somma	57591,1
Media	200,666
Mediana	73,85
Deviazione standard (popolazione)	352,897
Deviazione standard (campione)	353,514
Minimo	7,69
Massimo	3942,65
Intervallo	3934,96
Minoranza	7,69
Maggioranza	13,85
Varietà	285
Q1	34,585
Q3	228,425
Scarto interquartile (IQR)	193,84
Valori mancanti (nulli)	0

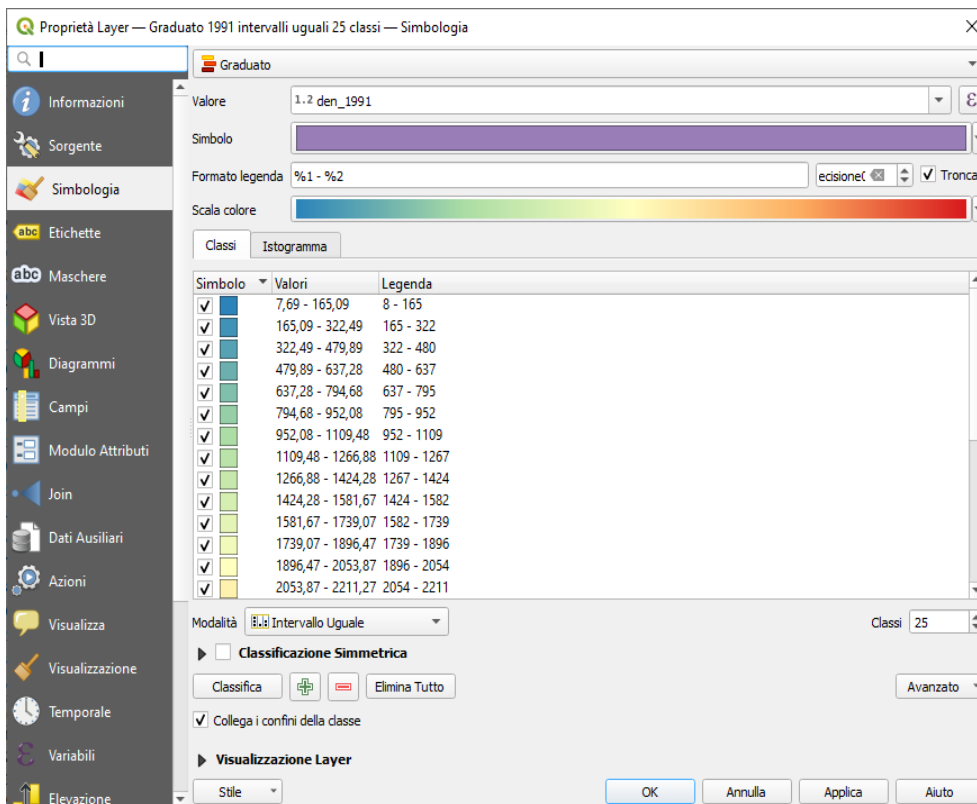
Vogliamo rappresentare i dati in una carta tematica utilizzando la simbologia Graduato: **Layer (Barra dei Menu) ► Proprietà... ► Simbologia ► Graduato**, questa scelta ci permette di suddividere i dati di una data colonna in un certo numero di classi e quindi scegliere uno stile differente per ciascuna classe.

(Nota bene: la **Varietà** indica quanti valori sono differenti, nel caso specifico sono duplicati i valori 13,85 e 34,34; in **Maggioranza** viene indicato il

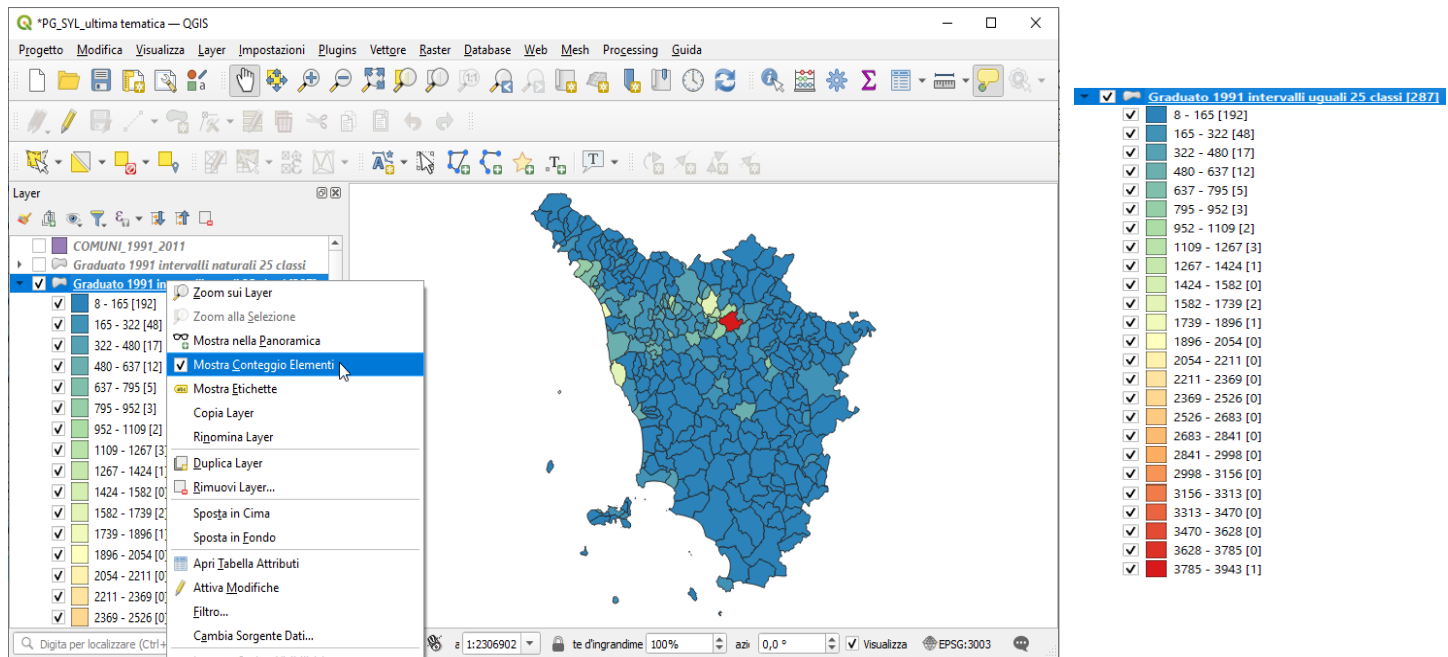
primo dei valori duplicati).

Classificazione dei dati con il metodo degli intervalli costanti

Scegliamo **Valore den_1991** Modalità **Intervallo Uguale** **Classi 25** e **Scala colore Spectral invertita**. La modalità Intervallo Uguale crea classi della stessa misura ottenute dividendo l'intervallo dei valori presenti per il numero delle classi scelto, quindi $3934,96 / 25 = 157,3984$. Risulta:



nella legenda i valori vengono approssimati, scegliamo di mostrare accanto alle classi il numero di comuni la cui densità ricade nella classe:



Si nota che ci sono molte classi con nessun comune nella classe, i comuni hanno valori prevalentemente nelle prime due classi poi c'è il comune di Firenze che ha un valore molto elevato a se stante.

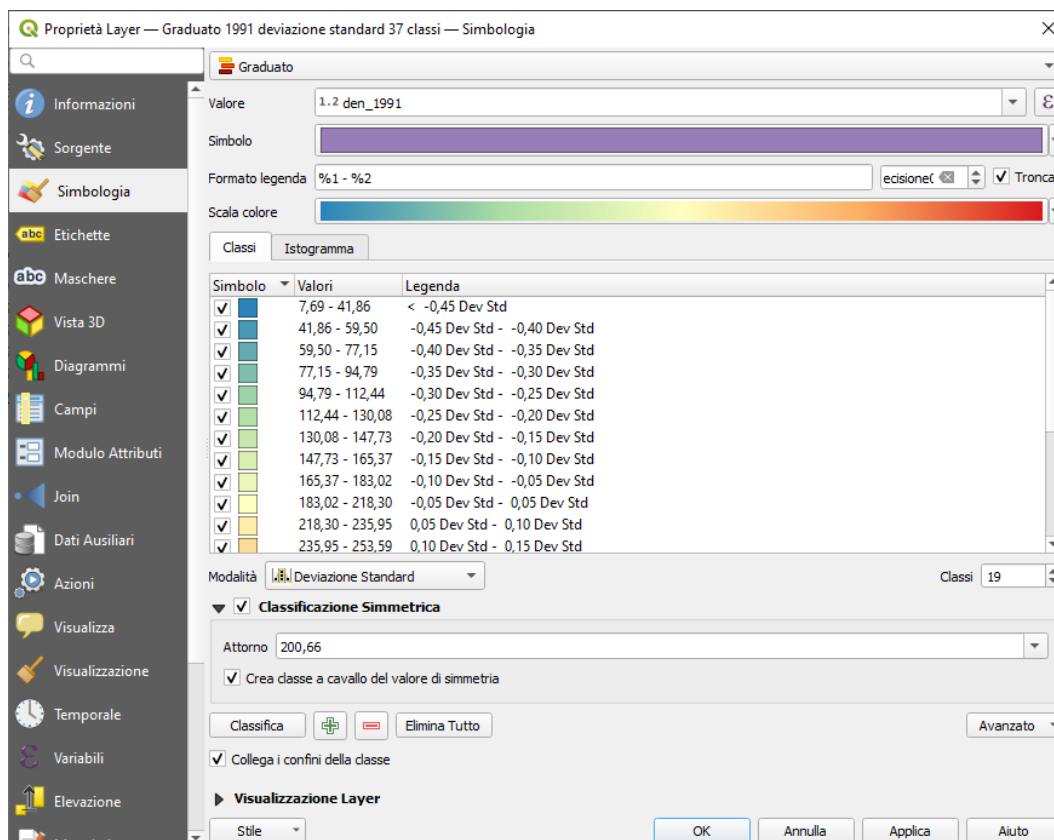
In conclusione questo tipo di rappresentazione Intervalli Uguali non è adatta alla distribuzione dei valori della den_1991.

Classificazione dei dati con il metodo della deviazione standard

Questo metodo calcola la media dei dati e crea le classi sulla base della deviazione standard dalla media. Abbiamo visto precedentemente che la *media dei valori di den_1991* nei comuni è **200,666** e che la *deviazione standard su tutto l'insieme dei dati* è **352,897**.

Notiamo che la distribuzione dei valori non è adatta a questo tipo di classificazione in quanto la deviazione standard ha significato per distribuzione di valori che siano in prevalenza attorno al valore medio e con valori uniformemente in aumento e in diminuzione rispetto al valore medio con numerosità fortemente crescente e decrescente rispetto al valore medio (distribuzione a campana rispetto al valore medio), in tal caso la deviazione standard assume un valore basso rispetto al valore medio, nel nostro caso così non è !!

Applichiamo comunque il metodo: scegliamo *Valore den_1991 Modalità Deviazione Standard, Classi 19*, *Scala colore Spectral invertita* e *Classificazione Simmetrica Attorno* il valore **220,66** della media. Risultata:



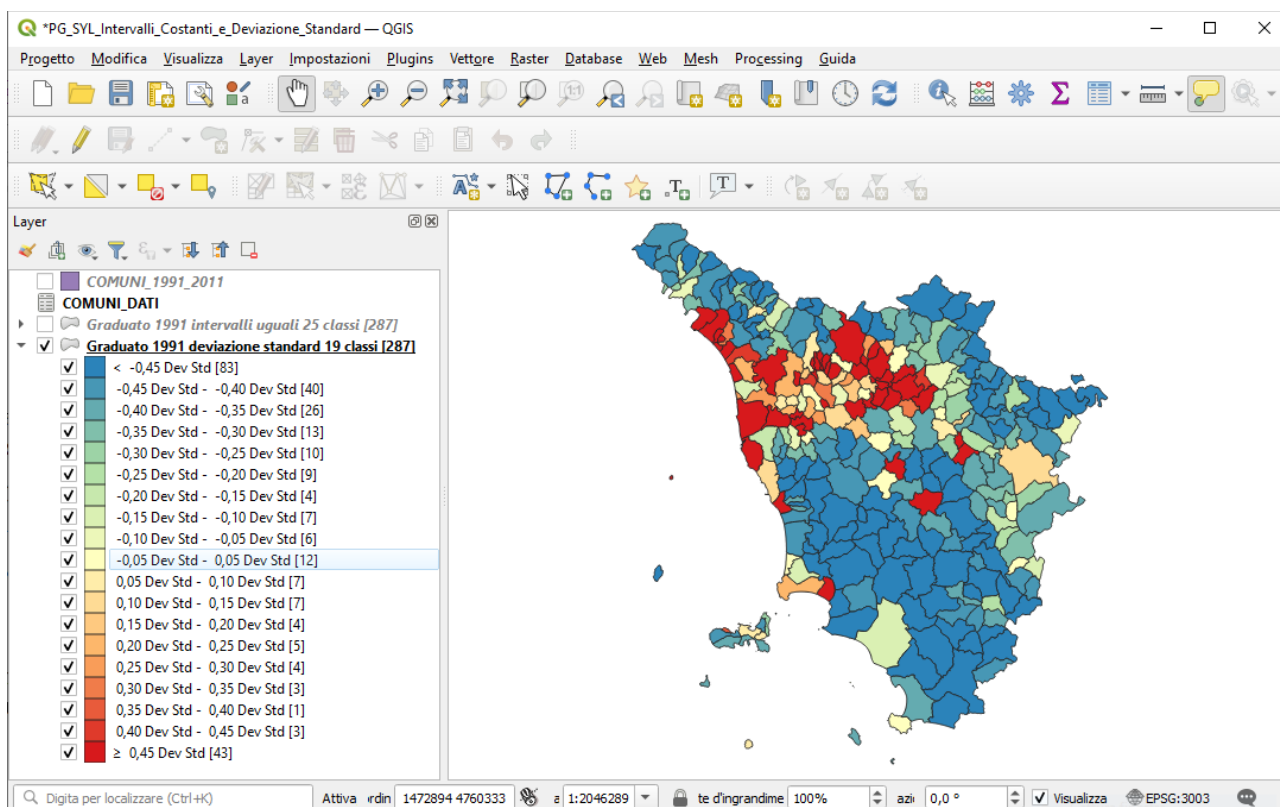
nella legenda i valori vengono approssimati, così facendo abbiamo definito:

- 1 classe centrale attorno al valore medio pari a $2 \times 352,897 / 20 = 35,2897$
- e 9 classi a decrescere + 9 classi a crescere di ampiezza $352,897 / 20 = 17,644850$

quindi sono una particolare divisione in intervalli uguali come frazione della deviazione standard su un intervallo a cavallo del valore medio di ampiezza doppia:

classe	descrizione	ampiezza
1	meno di 41,86235	
2	Tra 41,86235 e 59,50720	17,64485
3	Tra 59,50720 e 77,15205	17,64485
4	Tra 77,15205 e 94,79690	17,64485
5	Tra 94,79690 e 112,44175	17,64485
6	Tra 112,44175 e 130,08660	17,64485
7	Tra 130,08660 e 147,73145	17,64485
8	Tra 147,73145 e 165,37630	17,64485
9	Tra 165,37630 e 183,02115	17,64485
10	Tra 183,02115 e 218,31085	35,2897
11	Tra 218,31085 e 235,95570	17,64485
12	Tra 235,95570 e 253,60055	17,64485
13	Tra 253,60055 e 271,24540	17,64485
14	Tra 271,24540 e 288,89025	17,64485
15	Tra 288,89025 e 306,53510	17,64485
16	Tra 306,53510 e 324,17995	17,64485
17	Tra 324,17995 e 341,82480	17,64485
18	Tra 341,82480 e 359,46965	17,64485
19	più di 359,46965	

scegliamo di mostrare accanto alle classi il numero di comuni la cui densità ricade nella classe, risulta:



la distribuzione nelle classi è l'opposto di una distribuzione adatta alla rappresentazione tramite il metodo della deviazione standard essendo la numerosità delle densità di popolazione grande nelle classi estreme rispetto al valore medio.

Comunque questo tipo di rappresentazione nel caso specifico rispetto a quella precedente ottenuta con il metodo degli intervalli uguali ha il vantaggio di discretizzare meglio le differenze di densità di popolazione.

Salviamo le elaborazioni fatte

- nel progetto **PG_SYL_Intervalli_Costanti_e_Deviazione_Standard.qgz**
- nella cartella **C:\GISeQGIS_PROGETTI\PG_ECDLGIS_SYLLABUS**

in questo modo salviamo i campi provvisori creati nella tabella attributi di **COMUNI_1991_2011** che potremmo utilizzare in successivi esercizi.