

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN ECONOMIA

CICLO XXV

**INDICATORI DI BENESSERE NELLE REGIONI
EUROPEE: ANALISI MULTIDIMENSIONALI ED
APPROCCIO BASATO SUI FUZZY SETS**

RELATORE:

CHIAR.MO PROF. SERGIO ZANI

DOTTORANDA: LARA BERZIERI

SOMMARIO

INTRODUZIONE.....	3
CAPITOLO 1 - BENESSERE: DEFINIZIONI ED APPROCCI TEORICI	7
1.1 DAL PIL PROCAPITE AL BENESSERE.....	7
1.2 SCHEMI TEORICI PER LA MISURAZIONE DEL BENESSERE	15
CAPITOLO 2 - MISURARE IL BENESSERE NELLE REGIONI EUROPEE.....	21
2.1 LE UNITA' TERRITORIALI D'ANALISI	21
2.2 TRATTAMENTO PRELIMINARE DEI DATI.....	34
2.3 CRITERI DI SELEZIONE DEGLI INDICATORI DI BENESSERE.....	37
2.3.1 VARIABILI DEMOGRAFICHE	38
2.3.2 VARIABILI LEGATE ALLA SALUTE E INCIDENTALITA'	45
2.3.3 VARIABILI RELATIVE ALL'ISTRUZIONE	47
2.3.4 VARIABILI RELATIVE AL LAVORO.....	49
2.3.5 VARIABILI LEGATE ALLA RICCHEZZA E AL TEMPO LIBERO	55
2.3.6 VARIABILI AMBIENTALI	58
CAPITOLO 3 - ANALISI DEGLI INDICATORI DI BENESSERE TRAMITE LA CARTOGRAFIA	61
CAPITOLO 4 - CREAZIONE DI UN INDICATORE COMPOSITO DI BENESSERE: ANALISI MULTIDIMENSIONALI .	93
4.1 ANALISI INDICATORI PER AMBITO DI BENESSERE	95
4.2 PRIME SINTESI PER CIASCUN AMBITO DI BENESSERE	99
4.4 SINTESI TRAMITE ANALISI DELLE COMPONENTI PRINCIPALI	119
4.4 INDIVIDUAZIONE DI GRUPPI OMOGENEI: CLUSTER ANALYSIS.....	132
CAPITOLO 5: COSTRUZIONE DEGLI INDICATORI MEDIANTE GLI INSIEMI SFOCATI	148
5.1 DEFINIZIONE DI FUZZY SETS	148
5.2 LA TEORIA DEI FUZZY SETS APPLICATA ALL'ANALISI SUL BENESSERE	150
5.2.1 POSSIBILI FUNZIONI DI APPARTENENZA.....	151
5.2.2 AGGREGAZIONE E PONDERAZIONE DEI VALORI FUZZY.....	156

CAPITOLO 6: CONFRONTI TRA I RISULTATI DEI DIVERSI METODI ED OSSERVAZIONI CONCLUSIVE.....	166
BIBLIOGRAFIA	174
INDICE DELLE FIGURE	178
APPENDICE.....	180

INTRODUZIONE

STATISTICA, DEMOCRAZIA, BENESSERE

Il perseguimento del benessere è direttamente connesso all'attribuzione e al concreto esercizio di diritti fondamentali dell'individuo che trovano nelle costituzioni democratiche un riconoscimento specifico e una garanzia
[Sen 1985]

Nella vita quotidiana si è sempre più sommersi dalle informazioni, il che genera spesso un sentimento di confusione.

Il motivo di ciò è da ricercarsi nel fatto che le informazioni, provenienti da più fonti, sono talora in disaccordo tra loro sul medesimo argomento e quindi si pone il problema di identificare il dato più attendibile e corretto.

La statistica pubblica nasce appunto per porre rimedio a tale situazione: la produzione di informazioni statistiche attendibili rientra tra i compiti che Istat, a livello nazionale, ed Eurostat, a livello europeo, si propongono di raggiungere.

L'attendibilità dell'informazione statistica pubblica è garantita da un sistema di controlli di qualità che opera a livello mondiale per il rispetto dei principi generali e dell'insieme di regole (nomenclature, definizioni, classificazioni) stabiliti e riconosciuti dai principali organismi internazionali, quali il Fondo Monetario Internazionale e la Divisione Statistica delle Nazioni Unite.

L'osservanza di tali principi nelle fasi di produzione, elaborazione e diffusione dei dati favorisce la comparabilità fra le statistiche dei vari paesi.

Pertanto la statistica, oltre allo svolgimento della funzione conoscitiva, svolge un "ruolo sociale", poiché trasforma i dati in informazione e successivamente l'informazione in conoscenza, in maniera tale che i principali decisori cioè politici, operatori economici e pure semplici cittadini, possano effettuare le loro scelte avendo a disposizione alcuni elementi conoscitivi utili a ridurre il grado di incertezza.

Solo in tal caso l'informazione è realmente un bene pubblico: più siamo informati su ciò che sta accadendo nella nostra società, migliore sarà il funzionamento della nostra democrazia.

Lo stesso termine “statistica” deriva da scienza dello stato e tale origine sottolinea il ruolo della statistica come “strumento per comprendere meglio la realtà esistente e sostenere scelte concrete per migliorare la qualità della vita dei cittadini e consentire a questi ultimi di valutare le politiche messe in atto da chi è stato a ciò deputato.” (Giovannini, 2010).

Anche i *policy-makers*, operanti ad ogni livello (locale, nazionale ed internazionale) dovrebbero garantire l'assunzione di decisioni politiche in grado di massimizzare il benessere del relativo gruppo sociale, compatibilmente con i vincoli finanziari cui devono attenersi.

Tuttavia negli ultimi anni assistiamo ad una crescente crisi della legittimità democratica i cui sintomi sono: la credibilità declinante delle istituzioni rappresentative dello Stato democratico, la crisi del sistema dei partiti, la confisca del potere politico da parte di rappresentanti senza più contatti con i loro elettori, con il conseguente acuirsi della distanza tra popolo ed élite politica. (Pharr and Putnam, 2000).

Nonostante la perdita di fiducia nelle istituzioni da parte dei cittadini da un lato (anche se in misura assai differente da paese a paese) ed i pesanti vincoli finanziari e di bilancio dall'altro, l'Unione Europea si è posta come obiettivo il miglioramento delle condizioni e della qualità della vita negli stati membri (vedi Art. 1 del Protocollo sulla Politica Sociale del Trattato di Maastricht del 1992). Pertanto, l'informazione sistematica in tal senso risulta di cruciale importanza nel processo di coesione europea ed occorre disporre di un sistema europeo di indicatori sociali ed economici (insieme di elementi strutturali e procedure di misurazione) in grado di monitorare permanentemente e analizzare la qualità della vita e i mutamenti sociali.

Andare oltre il Pil rappresenta una delle grandi sfide politiche dei nostri giorni, anche se molto probabilmente non si giungerà mai ad una sola misura alternativa; ad ogni modo una maggiore attenzione alla qualità della vita (Qdv) ed alla sua sostenibilità potrà condurre ad una crescente disponibilità di informazioni circolanti che permetteranno ai media di informare meglio i cittadini e ai cittadini di giudicare in maniera più efficace l'operato della politica.

Come osserva correttamente il Comitato delle Regioni dell'Unione Europea, “non esiste una definizione universalmente accettata di qualità della vita”; essa infatti non può essere oggettivata, “in quanto la scelta degli indicatori, così come la rappresentazione e interpretazione, implicano già di per sé un giudizio” che risulta condizionato dalla cultura e dalla prospettiva teorica adottata.

Possiamo ad ogni modo affermare che si tratta di un tema interdisciplinare (data la permeabilità tra statistica, economia, sociologia e psicologia) e multidimensionale, in quanto coinvolge una pluralità di aspetti relativi alla vita umana.

Emergono pertanto svariate difficoltà già in sede di definizione del concetto di benessere e successivamente nella scelta degli indicatori necessari per la sua misurazione. Ma, come ci dimostrano tutta una serie di studi precedenti, tale complessità non dev'essere vista come un ostacolo insormontabile, bensì come un incentivo a proseguire negli studi già intrapresi, che saranno pertanto sempre parziali e certamente migliorabili. Come sostiene Vogel nel 1997 “la ricerca sugli indicatori sociali di benessere non dovrebbe mai arrestarsi, ma adattarsi continuamente a società in perpetuo mutamento”.

SCHEMA DELLA TESI

Dopo tali premesse, la presente ricerca si propone di misurare il livello di benessere che caratterizza le regioni europee, ricorrendo ai dati rilevati da Eurostat ed utilizzando metodi statistici multivariati, quali l'analisi delle componenti principali ed in particolare un metodo basato sui *fuzzy sets*, in maniera tale da aggiungere un piccolissimo tassello all'esteso mosaico della ricerca sociale su questo tema.

Nel primo capitolo si propone una rassegna dell'ampia letteratura sul benessere: evoluzione del concetto e dei metodi utilizzati per misurarlo. (si vedano: Bruni, Porta, 2004; Chiappero Martinetti, Pareglio, 2009; ed il numero monografico di *Social Indicators Research*, volume 102 (1) 2011 “*National and personal well-being: measuring the progress of societies*”).

Si considera inoltre uno schema teorico di base per definire con chiarezza e precisione il fenomeno multidimensionale oggetto di ricerca e per identificarne le sub-componenti relative alla sfera economica, sociale ed ambientale. Si pongono in luce le analogie e le differenze di significato tra i concetti di benessere (*well-being*) e di qualità della vita (*quality of life*), spesso utilizzati come termini intercambiabili, ma che devono essere intesi come solo parzialmente sovrapponibili. (Stiglitz, Sen, Fitoussi, 2009; OECD, 2011).

Nel secondo capitolo si descrivono le unità territoriali di riferimento, ossia le regioni europee NUTS (Nomenclatura Unità Territoriali Statistiche) di livello 2, e si procede ad una ricognizione dei dati disponibili nel sito di Eurostat:

<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>.

L'analisi del benessere a livello delle regioni europee rappresenta un primo elemento di novità di questa ricerca, poiché gli studi di questo tipo considerano solitamente i Paesi (si vedano ad esempio Baliamoune-Lutz, 2006; Bérenger, Verdier-Couchane, 2007) oppure le province italiane (si veda la nota indagine condotta annualmente da Il Sole - 24 Ore). L'esame delle singole regioni europee consente un raffronto tra aree simili appartenenti a Stati differenti.

Si effettua quindi una prima scelta delle variabili che possono considerarsi come sintomi del benessere delle regioni, affrontando e trattando anche il problema dei dati mancanti.

La tappa successiva è l'analisi unidimensionale di ciascuna variabile e individuazione degli *outliers* univariati e bivariati. Si propone anche una visualizzazione cartografica dei diversi risultati ottenuti tramite il programma GVSig, mediante gli *shapes* forniti dal sito di Eurostat.

Dopo le suddette operazioni di *data cleaning* delle informazioni disponibili, si propongono diverse analisi multidimensionali, eseguite tramite il software SPSS, per svelare la struttura sottostante all'insieme degli indicatori e delle regioni, ricorrendo all'analisi delle componenti principali, con diverse scelte metodologiche (ponderazione o non ponderazione, eliminazione di variabili ridondanti, etc.). Si ottengono vari indicatori sintetici che possono essere assunti quali misure d'una variabile latente sottostante: il benessere. (Joint Research Centre of European Commission, 2008).

Il contributo metodologico di maggior rilievo di questa tesi è la proposta di utilizzare la teoria dei *fuzzy sets* (e.g. Zimmerman, 2001) per costruire indicatori sintetici che esprimano il grado di appartenenza delle singole regioni europee al sottoinsieme di quelle caratterizzate da elevato benessere. Questo approccio è stato originariamente introdotto per la misura della povertà (Cerioli, Zani, 1990; Lemmi, Betti, 2006) e successivamente esteso alla valutazione del benessere (Chiappero Martinetti, 2000; Baliamoune-Lutz, 2006) e della *customer satisfaction* (Zani, Milioli, Morlini, 2010 e 2012).

Analogamente alla povertà, il benessere è un concetto che può essere tradotto più efficacemente in termini di grado di appartenenza (d'una regione), rispetto ad una semplicistica classificazione dicotomica: con benessere, senza benessere.

Dopo aver definito, in base a ciascuna variabile, il livello di appartenenza di ogni regione europea al sottoinsieme di quelle "benestanti", si propongono diversi criteri di aggregazione per ottenere un indice composito.

I risultati conseguiti con i metodi multivariati tradizionali e quelli ricavati dall'approccio basato sui *fuzzy sets* vengono confrontati tra loro per porre in luce le analogie e le differenze.

CAPITOLO 1 - BENESSERE: DEFINIZIONI ED APPROCCI TEORICI

1.1 DAL PIL PROCAPITE AL BENESSERE

EVOLUZIONE DI CONCETTI E DI MISURE

L'espressione "qualità della vita" è divenuta nelle società moderne tanto più utilizzata quanto più indefinita: nonostante i mass-media la ripetano comunemente, la maggior parte degli utilizzatori non conosce bene il suo significato; né d'altra parte esiste un'uniformità di contenuti e definizioni tra gli stessi ricercatori sociali.

Tale concetto nel corso degli ultimi sessant'anni ha conosciuto una continua evoluzione.

Le prime riflessioni sulla qualità della vita (Qdv) sono state formulate, sebbene in maniera implicita, da Agnes Heller nell'ambito della teoria dei bisogni di derivazione marxista. La studiosa giudica il sistema capitalista colpevole di aver negato agli individui la possibilità di soddisfare i bisogni espressivi più autentici, quelli che Marx definisce "bisogni ricchi" e che attengono principalmente alla piena emancipazione dell'uomo, all'immediatezza dei rapporti sociali e non al semplice possesso della "merce feticcio".

Il termine Qdv, nelle economie avanzate, inizia a comparire negli anni cinquanta come sinonimo di ricchezza materiale ed il principale criterio di valutazione è il Pil procapite¹: maggiore è il reddito disponibile e maggiori sono le possibilità di ottenere una più elevata qualità della vita.

Nonostante l'ottima capacità di sintesi, l'indicatore Pil è stato da subito fortemente criticato per alcuni difetti intrinseci². Innanzitutto esso è un indicatore parziale delle attività che contribuiscono al benessere di una nazione, in quanto esso registra solo transazioni avvenute nei

¹ Il Pil è il valore a prezzi correnti dell'insieme di beni e servizi finali prodotti in un anno, nel territorio di una nazione. Tale misura quando è stata inventata, negli anni 30, non era stata pensata come un indice di benessere. Si era in piena recessione e il vero problema era capire quale fosse il livello della produzione, ma soprattutto come questa produzione si legasse agli investimenti in opere pubbliche, all'occupazione e così via. È soltanto nella lettura neoclassica di fine anni 50-primi anni 60 che il Pil assume il rilievo che ha oggi.

² Pur riconoscendone limiti e debolezze, ancora oggi alcuni economisti continuano a privilegiare questo indicatore per misurare la ricchezza di una nazione.

mercati formali. Il classico esempio è quello del lavoro domestico che, se svolto da una casalinga non è registrato come una transazione formale, mentre se tale servizio viene affidato ad una persona che percepisce una regolare retribuzione, esso appare nella contabilità nazionale: lo spostamento della produzione di beni e servizi dal settore informale a quello formale lascia inalterato il benessere ma aumenta il Pil, contribuendo all'allentamento del legame fra queste due grandezze.

Il secondo aspetto riguarda la qualità delle singole voci di spesa, in quanto alcune di esse servono a compensare i danni provocati da comportamenti o situazioni negative. Prendiamo il caso di un drogato in cura presso un centro di disintossicazione: la spesa per fornire metadone a tale centro è oggetto di una transazione monetaria che concorre anch'essa ad aumentare il Pil. Possiamo supporre che, in assenza di drogati, le autorità avrebbero destinato i fondi all'acquisto di nuovi macchinari medici. Ebbene, l'effetto sul prodotto interno lordo è lo stesso, ma la qualità della spesa è ben diversa.

Ciò nonostante, fino agli anni sessanta, il legame fra crescita economica e benessere è piuttosto forte ed il PIL costituisce un soddisfacente indicatore della qualità della vita; successivamente, invece, tale relazione si indebolisce³.

Dall'inizio degli anni Settanta comincia ad entrare in crisi il paradigma della crescita illimitata quando, a seguito degli shock petroliferi, si prende coscienza della scarsità delle risorse naturali e pure del problema dell'inquinamento dell'ambiente.

I ricercatori provvedono pertanto alla creazione di vari indici che consistono in elaborazioni del Pil: si cercò essenzialmente di correggerlo per il valore degli stock di capitale consumato. I seguenti esempi rappresentano le prime iniziative per estendere la misurazione del Pil oltre la ristretta nozione di capacità produttiva del sistema economico e per spostarsi verso una nozione dinamica del concetto, ossia, verso il problema della sostenibilità:

- il *Net National Product* introdotto da Lindahl, consiste nel prodotto nazionale lordo depurato dal deprezzamento degli stock di capitale fisico e umano. Esso costituisce una prima approssimazione del reddito di Hicks⁴;

³ Alcuni ricercatori (Daly, Cobb) ritengono che tale legame si sia spezzato e che stiamo ormai assistendo ad una crescita economica che riduce il benessere.

⁴ Il reddito di Hicks consiste nel massimo ammontare di beni o servizi che una persona o una nazione può consumare in un determinato periodo senza ridurre le possibilità di consumo future.

- il *Measure of Economic Welfare* di Nordhaus e Tobin, 1972, è calcolato sottraendo al PIL tutto ciò che non può essere considerato consumo corrente: il deprezzamento dei beni capitali, gli investimenti e le cosiddette spese non discrezionali (il costo di spostamento verso il lavoro e i costi che si devono sostenere per vivere nelle grandi città). Le spese per la salute e per l'educazione vengono contabilizzate come investimenti in capitale umano. Si aggiungono poi i benefici che provengono dallo stock di capitale accumulato a livello nazionale, includendo l'istruzione e la spesa sanitaria;

- l'*Index of Sustainable Economic Welfare* di Cobb e Cobb, 1994, comprende numerosi elementi contenuti negli indici precedenti con delle correzioni ed introduce aspetti innovativi quali la distribuzione del reddito, il rapporto fra capitale nazionale e capitale straniero, il deperimento delle risorse naturali, il danno ambientale, il valore del tempo libero e quello del lavoro domestico non pagato.

Sebbene tali indici introducano elementi interessanti che colgono aspetti tralasciati dall'indicatore originale, essi appaiono ancora incompleti perché trascurano una gamma di elementi rilevanti al fine di quantificare il potenziale economico di un paese: il capitale umano (inteso come la conoscenza che deriva dal cumularsi dell'esperienza oppure da investimenti mirati in istruzione e ricerca), la salute e la tecnologia.

Altri studiosi approfondiscono la relazione tra le variabili economiche e la soddisfazione/felicità individuale, dando nuovo impulso a tale filone di studi e dimostrando l'infondatezza di alcuni assunti tradizionali dell'economia, quale ad esempio la relazione positiva tra reddito e felicità a prescindere dal confronto tra paesi, tempo o persone.

Se da un lato Frey e Stutzer ci confermano che la correlazione tra reddito pro-capite e felicità in un particolare momento nel tempo dello spazio e del tempo è statisticamente significativa, dall'altro Easterlin, tramite il paradosso⁵ che da lui prende il nome, dimostra che se consideriamo le persone nel corso della vita, e nelle economie avanzate, la felicità rimane pressoché costante, o addirittura diminuisce, all'aumentare del reddito pro-capite. Lo studioso spiega che tale paradosso si verifica perché "l'aumento del reddito genera un'equivalente crescita delle aspirazioni materiali, e l'effetto negativo di questo aumento nelle aspirazioni riduce l'effetto positivo dell'aumento del reddito".

⁵ Il paradosso di Easterlin è stato riconfermato da indagini empiriche recenti condotte dallo stesso studioso, a seguito della pubblicazione nel 2008 della ricerca degli economisti Stevenson e Wolfers che ne mettevano in discussione la validità.

Un'altra possibile spiegazione al paradosso si basa sul concetto di reddito relativo: gli individui non prendono in considerazione il reddito assoluto, ma lo mettono in confronto con il reddito delle persone loro vicine. Tale ipotesi è stata formulata e testata econometricamente da Deussenberry, il quale rileva inoltre la presenza di una struttura asimmetrica di esternalità, in quanto le persone quando operano confronti guardano "in alto" e pertanto i più ricchi provocano un effetto esterno negativo sulle persone più povere e non viceversa.

In aggiunta alle precedenti osservazioni, lo psicologo Kahneman evidenzia che con l'aumentare del reddito cresce pure la domanda di giustizia sociale e quindi nei paesi più ricchi gli arricchimenti individuali possono renderci più infelici se ottenuti in parallelo con aumenti di disuguaglianza.

Scitovsky nel 1972 in *"The Joyless Economy"* spiega che nelle società opulente l'aumento dei beni e dei consumi non rende le persone più felici a causa dell'eccessivo utilizzo di beni comfort a scapito dei beni stimolo. Secondo l'autore il consumo di cose comode e non stimolanti crea infatti dipendenza e aumenta nel tempo "il costo" richiesto dal cambiare stile di vita, che quindi si auto-rafforzano, nonostante creino noia e rendano la vita scialba.

Alcuni studiosi per spiegare il paradosso di Easterlin formulano una teoria relazionale della felicità: se indichiamo con F la felicità di un individuo (considerandola una variabile misurabile in senso cardinale), con I il reddito (inteso come insieme di mezzi materiali), con R le relazioni interpersonali, (che Gui chiama "beni relazionali"), e ignoriamo altri elementi importanti, come ad esempio la salute, possiamo scrivere la "funzione di felicità" tramite la seguente:

$$F = f(I,R)$$

La felicità è pertanto espressa come una funzione del reddito individuale e dei beni relazionali. Se è vero e ragionevole supporre che l'effetto complessivo del reddito (I) contribuisce direttamente alla felicità, soprattutto per bassi livelli di reddito, bisogna anche considerare che, dopo aver superato una certa soglia, questo può diventare negativo poiché l'impegno per aumentare il reddito (assoluto o relativo) può produrre sistematicamente effetti negativi sui beni relazionali, sulla qualità e quantità delle nostre relazioni (ad esempio a causa delle risorse eccessive che impieghiamo per aumentare il reddito e che sottraiamo ai rapporti umani), e quindi, indirettamente, potrebbe smorzare, o addirittura ribaltare l'effetto totale diminuendo la felicità (è quindi una questione di derivate parziali). La funzione reddito-felicità assume pertanto una forma ad U rovesciata, simile a quella di Kutznet ⁶, in cui superata una certa soglia

⁶ Essa viene utilizzata per misurare il rapporto tra crescita ed inquinamento.

di reddito, detta “punto critico”, il segno del rapporto cambia e quegli stessi “beni” diventano “mali”.

Scitovsky sostiene che le persone superano il punto critico in modo inconsapevole, a causa di limiti cognitivi e condizionamenti sociali.

Ad ogni modo le differenze di reddito spiegano solo una parte piuttosto esigua delle differenze di felicità tra le persone, difatti esistono altri fattori ben più importanti, che vengono illustrati schematicamente da Frey e Stutzer in *Economia e felicità*, 2002:

- fattori della personalità come l'autostima, l'autocontrollo, l'ottimismo, l'estroversione e il nevroticismo;
- fattori socio-demografici come l'età, il genere, lo stato civile e l'istruzione;
- fattori economici come il reddito individuale e aggregato, la disoccupazione e l'inflazione;
- fattori contestuali e situazionali, come la salute, le condizioni di lavoro, le relazioni con il coniuge, colleghi, parenti, amici;
- fattori istituzionali come il grado di decentramento politico e i diritti alla partecipazione politica diretta dei cittadini.

Le riflessioni di tali studiosi hanno permesso di ampliare la serie di indicatori di Qdv: all'individuazione di ulteriori indicatori oggettivi di tipo sociale e relazionale si aggiungono, a partire dagli anni settanta, indicatori soggettivi sulla Qdv percepita, quali giudizi di felicità e soddisfazione.

Infatti, fin dalle prime ricerche sulla qualità della vita (USA, anni '70) si è visto che tale concetto si può “dicotomizzare in una dimensione ‘oggettiva’ - ovvero legata agli aspetti materiali, e non, del benessere individuale e/o collettivo - e in una dimensione ‘soggettiva’, basata sulla valutazione individuale delle proprie condizioni di vita”. Gli elementi oggettivi includono, seguendo tale approccio, le condizioni di vita accertabili, come gli standard di vita, le condizioni lavorative e lo stato di salute, mentre il benessere soggettivo riguarderebbe invece la valutazione specifica delle condizioni di vita individuali e le componenti affettive.

Il sociologo tedesco Wolfgang Zapf nel 1984 propose la seguente tipologia, che si basa sull'utilizzo combinato di indicatori “soggettivi” e “oggettivi” e che dà luogo a quattro diversi esiti, a seconda delle combinazioni tra condizioni oggettive di vita e soddisfazione soggettiva:

CONDIZIONI SOGGETTIVE DI VITA		
CONDIZIONI OGGETTIVE DI VITA	BUONE	CATTIVE
BUONE	WELL-BEING	DISSONANZA
CATTIVE	ADATTAMENTO	PRIVAZIONE

Figura 1. Tipologia di Zapf sulle condizioni di vita

Sono quattro i possibili atteggiamenti degli individui nei confronti della propria condizione: “adattamento” quando i soggetti si dicono soddisfatti pur non vivendo in condizioni oggettivamente favorevoli; “dissonanza” indica l’atteggiamento di chi, in buone condizioni oggettive, si dice comunque insoddisfatto; “privazione” indica la situazione in cui a condizioni oggettive sfavorevoli corrisponde effettivamente il disagio individuale; “benessere” quando a condizioni oggettive favorevoli corrisponde effettivamente un adeguato livello di soddisfazione. Cummins nel 1988 (*International Society for Quality of Life Studies*), pur segnalando che esistono almeno un centinaio di definizioni in tema, suggerisce che la seguente possa costituire un punto di accordo terminologico:

“La qualità della vita include sia l’oggettivo che il soggettivo, ed ogni asse è l’aggregazione di sette ambiti: benessere materiale, salute, produttività, affettività, sicurezza, società e benessere interiore. L’ambito oggettivo include misure rilevanti di benessere oggettivo. L’ambito soggettivo include la soddisfazione ponderata con l’importanza assegnata dall’individuo.”

In sintesi, la prima dimensione, che la letteratura sociologica chiama *having* e riguarda il possesso dei beni primari e l’attenzione ai bisogni primari del singolo individuo, mentre la seconda dimensione, chiamata *being*, riguarda gli aspetti secondari della vita delle persone contribuiscono a migliorare le condizioni di vita e comprende anche il *loving* cioè le capacità espressive e partecipative, di interazione con gli altri individui. Questa seconda dimensione si pone pertanto su una prospettiva non individualistica, poiché parte dal presupposto che la Qdv non si esaurisca nell’utilità e nelle aspirazioni individuali, ma richieda solidarietà e valori sociali condivisi da una collettività. All’attenzione per l’ambiente si aggiunge quella verso altri soggetti sociali e alle relazioni intersoggettive di partecipazione.

Negli anni novanta gli indici che conservano un approccio unidimensionale alla Qdv, vengono quindi sostituiti da una visione multidimensionale che considera gli strumenti (salute, ambiente,

istruzione, innovazione tecnologica) come fini in sé e non solo strumenti per ottenere un reddito più elevato.

La varietà delle scelte possibili che il ricercatore deve effettuare e la necessaria soggettività delle regole di aggregazione spiega la produzione nell'ultimo ventennio di una moltitudine di indicatori elaborati da centri studi, organismi internazionali o comunità locali.

L'utilizzo dell'approccio multidimensionale, se da un lato si presta ad analizzare in maniera molto più completa ed esauriente la Qdv, dall'altro introduce maggiore complessità nell'analisi ed in particolare si pone il problema se aggregare in un indice sintetico rappresentativo la pluralità di dimensioni considerate oppure presentare l'informazione disaggregata ricorrendo all'uso di vettori; ci riferiamo ai primi con il termine di indici, mentre gli altri li definiamo come raccolte di indicatori.

Un esempio del primo tipo è l'Indice di Sviluppo Umano (*Human Development Index*), pubblicato ogni anno a partire dal 1993 dal Programma delle Nazioni Unite per lo Sviluppo e ispirato alla visione di A. Sen del benessere, si basa su tre dimensioni ritenute di fondamentale importanza quali: una vita lunga ed in buona salute (utilizzando come *proxy* di tale dimensione la variabile aspettativa di vita alla nascita); il grado di educazione (misurato attraverso il tasso di alfabetizzazione degli adulti, con peso pari a due terzi, e quello di scolarizzazione primaria e secondaria, con peso pari ad un terzo); il reddito pro-capite espresso a parità di potere d'acquisto. Benché aspramente criticato da vari ricercatori, il modo di calcolare l'HDI non è stato ancora modificato.

Riguardo agli indicatori sviluppati in ambito accademico, le aree tematiche affrontate in prevalenza sono quella ambientale e quella economica; segue poi quella sociale, mentre le istituzioni sono ancora scarsamente monitorate.

Il *Well-Being Index*, creato dalla *World Conservation Union*, valuta il livello di benessere tramite due sottoindici: l'indice di benessere umano (ottenuto a sua volta aggregando indicatori che misurano la qualità della salute della popolazione, la ricchezza economica, il livello culturale e di scolarizzazione, la qualità delle relazioni civili ed il grado di equità della società) e l'indice di qualità dell'ambiente (composto da indicatori che misurano lo stato dei suoli, dell'acqua, dell'aria, delle specie e del patrimonio genetico, nonché dell'uso delle risorse naturali). Gli indici di area hanno uguale peso e la loro aggregazione avviene con una media semplice. Si dà importanza sia alla dimensione statica, ossia alla qualità della vita, sia a quella dinamica, più strettamente connessa con la sostenibilità nel tempo del benessere.

L'esperienza dell'*Environmental Sustainability Index* (ESI) di Yale e di Columbia University è un caso particolare nel quale gli autori prendono espressamente posizione in favore di una regola di aggregazione. La ricerca della sintesi, per offrire una misura immediatamente utilizzabile dai *policy makers* nell'indirizzare le loro scelte, è secondo gli autori un obiettivo importante, da perseguire anche se i criteri di scelta sono necessariamente non perfetti e universalmente condivisi.

Nel complesso notiamo un interesse crescente al controllo della sostenibilità dei processi economici, dell'organizzazione sociale e della gestione delle risorse naturali.

Negli ultimi anni si evidenzia pertanto un costante spostamento dell'attenzione da una visione statica ad una dinamica: agli indicatori di stock, che privilegiano la misura della qualità della vita in un dato istante nel tempo, si aggiungono sempre più di frequente quelli che misurano tale fenomeno nel tempo, cioè gli indicatori di flusso.

1.2 SCHEMI TEORICI PER LA MISURAZIONE DEL BENESSERE

Un interessante schema teorico di riferimento per l'analisi del benessere è l'approccio delle capacità, formulato negli anni 80 dall'economista e filosofo Amartya Sen. Lo studioso indiano precisa che il concetto di tenore di vita (*standard of living*) si limita a considerare le dimensioni del benessere che hanno a che fare con la vita del soggetto, mentre quello di benessere (*well-being*) riguarda un orizzonte più ampio, che include anche altre realizzazioni che contribuiscono ad accrescere indirettamente il benessere individuale e di cui è importante tenere conto.

Secondo il *capability approach* il benessere individuale è il risultato di un processo all'interno del quale i mezzi e le risorse acquisibili o disponibili (cioè beni e servizi pubblici), rappresentano solo uno strumento per ottenere benessere. Sen riprende pertanto la concezione aristotelica⁷ ma pone l'attenzione sulla pluralità di fattori personali e familiari e sulla varietà di contesti – sociali, ambientali, economici, istituzionali, culturali – che agiscono nella determinazione del processo di benessere individuale.



Figura 2. Il processo di determinazione del benessere

La parola benessere viene assimilata all'idea di "star bene", intesa come una condizione che include "ciò che l'individuo può fare o può essere" considerate le risorse a disposizione e le

⁷ Secondo Aristotele "è chiaro che non è la ricchezza il bene da noi cercato: essa infatti ha valore solo in quanto "utile", cioè in funzione di qualcos'altro" [Etica Nicomachea]; secondo il filosofo greco tutti i beni sono solo un mezzo per conseguire la felicità. Quest'ultima è l'unico fine non strumentale che l'essere umano persegue nella sua attività e per questo la felicità è definita "fine ultimo". Aristotele precisa però che la parola felicità (in greco *eudaimonia*) non coincide con il piacere, bensì con il concetto di "fioritura umana".

capacità delle persone di trasformare tali mezzi in risultati che intendono conseguire. Sen fa notare che a parità di risorse possedute è possibile raggiungere diversi livelli di *well-being* a seconda delle capacità di conversione di cui dispone un individuo, le quali consentono di trasformare un determinato insieme di beni in realizzazioni potenziali, cioè quello che viene chiamato lo spazio delle capacità; le capacità sono le funzioni potenziali delle persone, cioè il fare e l'essere e l'insieme delle capacità consiste nella pluralità di opzioni che il singolo individuo ha la libertà di valutare. I fattori di conversione dipendono dalle caratteristiche personali di ciascun soggetto, quali l'età, il sesso, le condizioni fisiche e psichiche, le abilità e talenti, oltre che dall'ambiente familiare, sociale, economico, naturale, culturale e politico-istituzionale circostante. Successivamente, nel passaggio dallo spazio delle capacità a quello dei funzionamenti, cioè all'insieme dei traguardi effettivamente realizzati, entrano in gioco le scelte e le responsabilità ad esse collegate di ciascuna persona. I due spazi sono a loro volta multidimensionali, poiché racchiudono una molteplicità di aspetti del benessere.

A proposito della distinzione tra spazio di capacità e spazio dei funzionamenti Sen chiarisce: “*A functioning is an achievement, whereas a capability is the ability to achieve. Functionings are, in a sense, more directly related to living conditions, since they are different aspects of living conditions. Capabilities, in contrast, are notions of freedom, in the positive sense: what real opportunities you have regarding the life you may lead*” (Sen, 1985) e sostiene che scegliere di compiere una determinata azione avendo una pluralità di opzioni a disposizione ha un valore intrinseco per il benessere dell'individuo superiore alla condizione in cui vi sia un'unica alternativa disponibile. Si conferisce pertanto valore alla maggiore libertà di scelta lasciata al soggetto, che in tal modo incrementa il proprio benessere. Ad esempio, se una donna con una certa istruzione ed esperienza lavorativa riceve un'offerta di lavoro, ma decide di non lavorare (senza nessuna pressione dalle persone intorno a lei) ha la capacità di fare quel lavoro ma non raggiunge la funzione di lavorare. Nonostante ciò il livello di benessere della donna è superiore rispetto a quello di un'altra donna alla quale nessuno offre un lavoro per mancanza di capacità.

Pur rappresentando un impianto teorico assai innovativo, la proposta di Sen presenta un notevole grado di complessità e difficoltà di applicazione legate al problema della definizione e misurazione dei suoi elementi costitutivi. Si pensi ad esempio all'elencazione dell'insieme di capacità e funzionamenti da considerare ai fini dell'analisi: Sen non ne fornisce intenzionalmente una lista, ma demanda tale scelta ai processi democratici e alle scelte sociali in

relazione al tempo, al luogo, allo schema di valori prevalente nei differenti contesti⁸; altro elemento problematico non banale riguarda la stima dei tassi di conversione.

Per questa serie di motivi risulta molto più semplice adottare lo schema concettuale formulato dall'Oecd (*Organisation for Economic Cooperation and Development*), contenuto nel “*Compendium of Oecd well-being indicators*”. In estrema sintesi la condizione di benessere umano è determinata direttamente da due componenti cioè le condizioni di vita materiali e la qualità di vita (al cui interno occorre considerare i valori medi, ma pure le differenze tra i vari gruppi), ed è influenzata dalla dimensione di sostenibilità nel tempo del benessere (che a sua volta influenza il benessere presente), così come emerge dalla Figura 3.

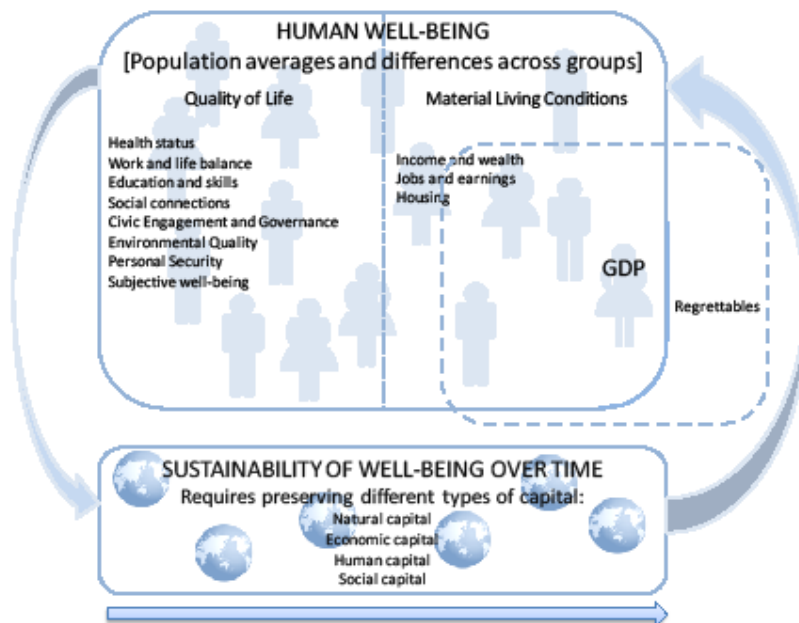


Figura 3. Struttura degli indicatori di benessere individuati dall'Oecd

⁸ Al contrario la filosofa Martha Nussbaum ritiene che le capacità possano essere elencate in una lista seppure non definita completamente: in *Giustizia Sociale e Dignità Umana* (2002) ne propone alcune, quali la possibilità di vivere una vita di normale durata, di godere di buona salute fisica, di un'alimentazione e un'abitazione adeguate, di vedere soddisfatta una serie di bisogni relativamente a integrità fisica, senso di appartenenza, istruzione, utilizzo del pensiero e dell'immaginazione, di provare affetto e riceverlo, di controllare il proprio ambiente...

La prima componente riguarda la dimensione economica del benessere, rappresenta le possibilità di consumo e di dominio sulle risorse ed è determinata da tre fattori: reddito e patrimonio, lavoro e stipendio, abitazione. Come si può osservare dalla Figura 3, il Pil non costituisce la misura numerica di tale dimensione, in quanto al Prodotto Interno Lordo occorre aggiungere le attività che non figurano nei mercati formali e sottrarre quelle che generano conseguenze negative sullo sviluppo economico, come ad esempio l'inquinamento.

La seconda componente include tutti gli altri aspetti non monetari della vita umana in grado di aumentare il benessere individuale in maniera diretta (insieme di possibilità ed opportunità di scelta) che vengono accorpati in sette dimensioni: stato di salute, equilibrio tra lavoro e tempo libero, livello d'istruzione ed abilità, partecipazione civica e politica, relazioni sociali, qualità dell'ambiente, sicurezza personale, benessere soggettivo.

Riguardo allo sviluppo sostenibile, esso viene definito dall'Oecd come "lo sviluppo nel presente che non comprometta le possibilità (in termini di risorse, ambiente e condizioni socio-economiche) delle generazioni future di soddisfare i loro bisogni". Il vincolo di sostenibilità dello sviluppo determina dunque l'esigenza di progettare sistemi economici e sociali in grado di raggiungere simultaneamente i due obiettivi, tramite la difesa dell'ambiente – inteso nella sua accezione più estesa, come ambiente naturale, artificiale e culturale – l'estensione dell'orizzonte temporale degli interventi e il perseguimento dell'equità intra e intergenerazionale" (Pearce et al. 1996). Appare chiaro, dunque, che solo uno sviluppo economico che abbia le caratteristiche sopra descritte possa garantire una condizione di reale benessere per l'intera società presente e futura. La valutazione dello sviluppo sostenibile, diversamente da quella del benessere, non prevede né il ricorso a variabili di tipo soggettivo né che si possa calcolare un indicatore aggregato.

In definitiva il concetto di benessere è ben più esteso rispetto a quello di qualità della vita: oltre a comprendere quest'ultimo, il primo considera ovviamente la dimensione materiale ed aggiunge gli aspetti legati alla sostenibilità.

La cornice teorica dell'Oecd tiene conto delle recenti osservazioni della comunità scientifica emerse dai forum sul tema organizzati negli ultimi anni (Palermo 2004, Istanbul 2007, Busan 2009) ed in particolare delle raccomandazioni contenute nel rapporto della "Commissione sulla misurazione della performance economica e del progresso sociale". Quest'ultima, detta anche Commissione Stiglitz (voluta dal presidente della Repubblica francese Nicolas Sarkozy), è

guidata dal premio Nobel per l'Economia Joseph Stiglitz e composta da venticinque economisti, tra cui altri quattro premi Nobel, quali Amartya Sen, Jean-Paul Fitoussi, Daniel Kahneman e Kenneth Arrow, ed il presidente dell'Istat Enrico Giovannini.

Gli obiettivi raggiunti dalla Commissione Stiglitz sono stati:

- individuare i limiti del PIL come indicatore della performance economica e del progresso sociale;
- considerare nuovi indicatori di progresso sociale;
- valutare la fattibilità di strumenti di misura alternativi ed individuare le modalità più adeguate per presentare le informazioni statistiche.

Gli studiosi hanno rimarcato che solo un set di indicatori può riflettere la complessità della società e la selezione di qualsiasi misura rilevante del benessere è un giudizio di valore più che una valutazione tecnica. Per tale motivo non si propone una lista precisa dei fattori di benessere, bensì l'elenco dei vari ambiti in cui il benessere può manifestarsi: condizioni materiali (reddito, consumi e patrimonio), salute, istruzione, lavoro e tempo libero, partecipazione politica, relazioni interpersonali, ambiente sociale e naturale, sicurezza economica e personale. Si precisa inoltre che tutte le dimensioni del benessere richiedono misure di dati sia oggettivi che soggettivi e che si dovrebbero valutare le disuguaglianze in modo complessivo, cioè considerare le variazioni tra gli individui, i gruppi socio-economici, tra i sessi e le generazioni, con particolare attenzione alle disuguaglianze sorte più di recente, come quelle legate all'immigrazione. Si dovrebbero inoltre valutare i nessi tra le varie dimensioni di benessere, al fine di considerarne gli effetti cumulativi.

Accanto alle dimensioni sopra elencate se ne affianca una orizzontale: la sostenibilità. Misurare e valutare la sostenibilità è una preoccupazione centrale della Commissione, poiché consente di determinare se almeno l'attuale livello di benessere può essere mantenuto per le generazioni future. Essa infatti richiede la conservazione simultanea o l'aumento di una serie di "stock" quali la quantità e qualità delle risorse naturali e delle risorse umane, del capitale sociale e fisico, nonché la qualità delle istituzioni. La valutazione della sostenibilità richiede numerose ipotesi e scelte normative ed è ulteriormente complicata dalla presenza di interazioni tra i modelli socio-economici e ambientali scelti dalle diverse nazioni. Ad ogni modo la complessa

valutazione della sostenibilità è complementare alla questione del benessere corrente e deve essere esaminata separatamente.

Gli studiosi appartenenti alla Commissione esortano i governi nazionali ad istituire delle tavole rotonde per individuare in modo condiviso gli indicatori del progresso economico e sociale, poiché il set di indicatori deve essere definito attraverso un processo democratico con i rappresentanti delle diverse componenti della società. Un ultimo punto fondamentale per la Commissione è la comunicazione, in quanto è necessario “sviluppare dei modelli e degli strumenti che possano aiutare gli istituti di statistica e le tavole rotonde dei singoli Paesi a comunicare con la gente. Strumenti semplici, che la sera a tavola una famiglia possa ascoltare al telegiornale. Solo così potrebbero contribuire a comprendere la situazione di un Paese e a cambiare i comportamenti” (Giovannini, 2010).

Come già affermato da Sen, quello che noi misuriamo configura ciò che collettivamente ci sforziamo di perseguire (e ciò che perseguiamo determina ciò che si misura) il rapporto e la sua applicazione può avere un significativo impatto sul modo in cui la nostra società guarda a se stessa e, di conseguenza, sul modo in cui le politiche sono progettate, attuate e valutate.

CAPITOLO 2 - MISURARE IL BENESSERE NELLE REGIONI EUROPEE

2.1 LE UNITA' TERRITORIALI D'ANALISI

La breve cornice teorica contenuta nel primo capitolo costituisce un punto di riferimento imprescindibile al fine di misurare i livelli di benessere nelle regioni europee.

Tuttavia, nello svolgimento della ricerca emergono fin da subito numerosi problemi metodologici, che verranno a mano a mano chiariti, tali da allontanarci dall'applicazione dei principi prima illustrati, con conseguenti ripercussioni sui risultati dell'analisi.

Le unità territoriali cui si fa riferimento nella ricerca sono le 271 regioni europee come previsto dalla classificazione Eurostat NUTS, (acronimo di Nomenclatura delle Unità Territoriali per Statistiche), di livello 2. Le regioni NUTS 2 dei 27 paesi appartenenti all'Unione Europea sono visualizzate in Figura 4.

La classificazione NUTS, benché già utilizzata a partire dal 1988, venne ufficializzata dal Parlamento Europeo solo nel 2003 e successivamente modificata nel 2004 e nel 2007, a seguito dell'ingresso dei paesi dell'Europa dell'Est nell'Unione.

Tale nomenclatura prevede tre livelli gerarchici e si basa sulle strutture amministrative già esistenti in ciascun stato membro, per ragioni pratiche relative alla disponibilità di dati (Eurostat ricorre ai dati raccolti dagli organi di statistica nazionali e locali delle singole nazioni) e pure nell'ottica del successivo riutilizzo dei dati nella realizzazione di appropriate misure di *policies*.

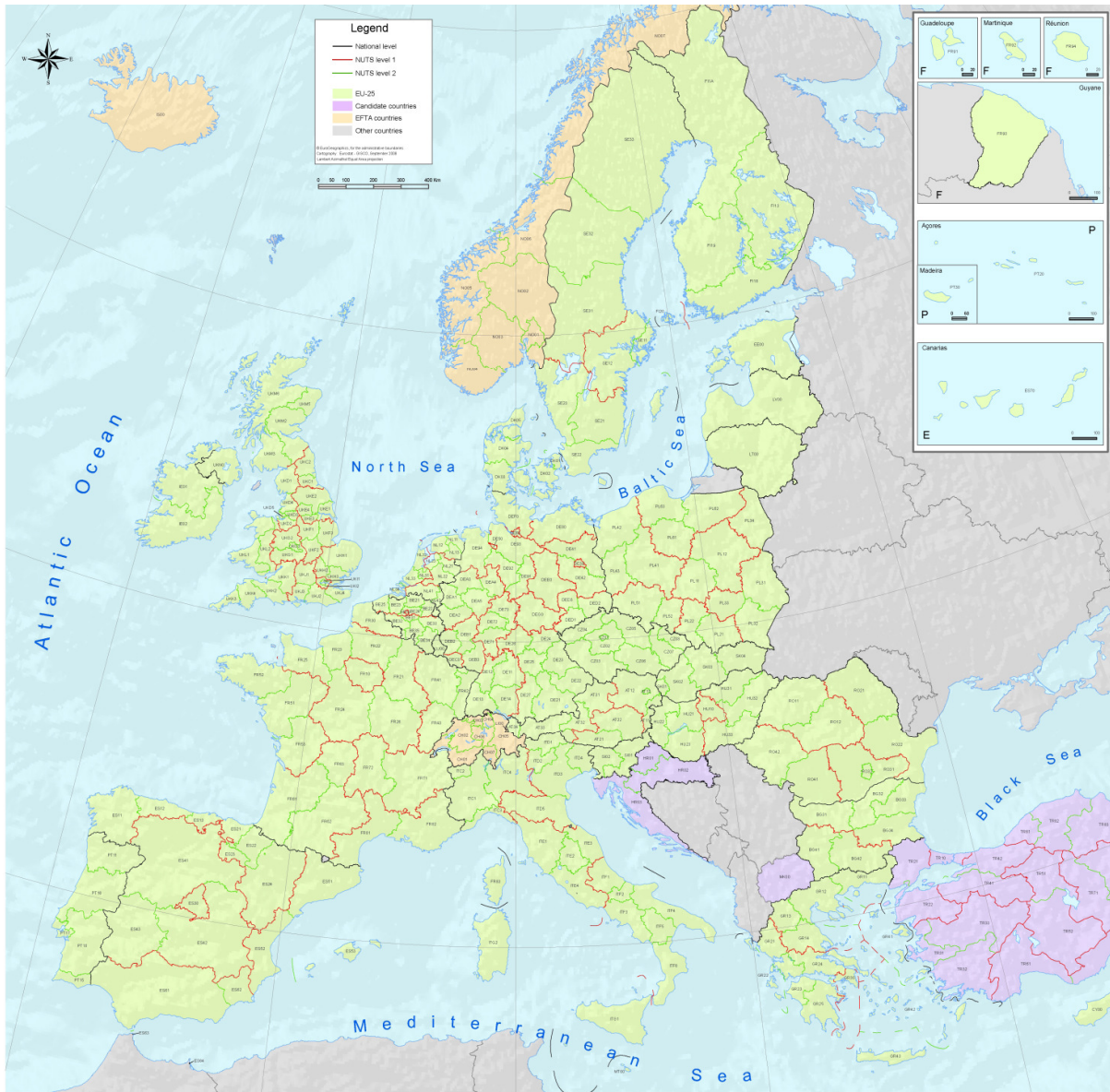


Figura 4. Regioni europee NUTS 2

La Tabella 1 mostra per ciascuna nazione la corrispondenza tra i livelli NUTS e le unità amministrative nazionali. Nel caso dell'Italia il livello NUTS 2 coincide essenzialmente con le Regioni, fatta eccezione per il Trentino Alto-Adige, che viene diviso nelle due Province autonome di Trento e Bolzano, per un totale di 21 unità territoriali.

Si sceglie di non prendere in considerazione il livello NUTS 3 (che in Italia corrisponde alle Province), in quanto si sarebbe ulteriormente aggravato il problema della disponibilità dei dati. Inoltre, com'è noto, in Italia è in programma un processo di accorpamento e di ridefinizione delle province.

	NUTS 1		NUTS 2		NUTS 3		LAU 1		LAU 2	
BE	Gewesten / Régions	3	Provincies / Provinces	11	Arrondissementen / Arrondissements	44	-		Gemeenten / Communes	589
BG	Rajoni	2	Rajoni za planirane	6	Oblasti	28	Obshtini	264	Naseleni mesta	5 329
CZ	Území	1	Oblasti	8	Kraje	14	Okresy	77	Obce	6 249
DK	-	1	Regioner	5	Landsdelel	11	Kommuner	99	Sogne	2 148
DE	Länder	16	Regierungsbezirke	39	Kreise	429	Verwaltungsgemeinschaften	1 457	Gemeinden	12 379
EE	-	1	-	1	Groups of Maakond	5	Maakond	15	Vald, linn	227
IE	-	1	Regions	2	Regional Authority Regions	8	Counties, Ctries	34	Electoral Districts	3 441
GR	Groups of development regions	4	Periferies	13	Nomoi	51	Demoi, Koinotites	1 034	Demotiko diamerisma, Koinotiko diamerisma	6 130
ES	Agrupacion de comunidades Autonomas	7	Comunidades y ciudades Autonomas	19	Provincias + islas + Ceuta, Melilla	59	-		Municipios	8 111
FR	Z.E.A.T + DOM	9	Régions + DOM	26	Départements + DOM	100	Cantons de rattachement	3 787	Communes	36 683
IT	Gruppi di regioni	5	Regioni	21	Provincia	107	-		Comuni	8 101
CY	-	1	-	1	-	1	Eparchies	6	Dimoi, koinotites	613
LV	-	1	-	1	Reģioni	6	Rajoni, republikas pilsētas	33	Pilsētas, novadi, pagasti	527
LT	-	1	-	1	Apskritis	10	Savivaldybės	60	Seniūnijos	518
LU	-	1	-	1	-	1	Cantons	13	Communes	116
HU	Statisztikai nagyrégiók	3	Tervezési-statisztikai régiók	7	Megyeék + Budapest	20	Statisztikai kistérségek	168	Települések	3 152
MT	-	1	-	1	Gzejjer	2	Distretti	6	Kunsilli	68
NL	Landsdelen	4	Provincies	12	COROP regio's	40	-		Gemeenten	443
AT	Gruppen von Bundesländern	3	Bundesländer	9	Gruppen von politischen Bezirken	35	-		Gemeinden	2 357
PL	Regiony	6	Województwa	16	Podregiony	66	Powiaty i miasta na prawach powiatu	379	Gminy	2 478
PT	Continente + Regioes autonomas	3	Comissaoes de Coordenação regional + Regioes autonomas	7	Grupos de Concelhos	30	Concelhos – Municipios	308	Freguesias	4 260
RO	Macroregiuni	4	Regiuni	8	Judet + Bucuresti	42	-		Comuni + Municipiu + Orase	3 174
SI	-	1	Kohezijske regije	2	Statistične regije	12	Upravne enote	58	Občine	210
SK	-	1	Oblasti	4	Kraje	8	Okresy	79	Obce	2 928
FI	Manner-Suomi, Ahvenanmaa / Fasta Finland, Åland	2	Suuralueet / Storområden	5	Maakunnat / Landskap	20	Seutukunnat / Ekonomiska regioner	77	Kunnat / Kommuner	416
SE	Grupper av riksområden	3	Riksområden	8	Län	21	-		Kommuner	290
UK:	Government Office Regions; Country	12	Counties (some grouped); Inner and Outer London; Groups of unitary authorities	37	Upper tier authorities or groups of lower tier authorities (unitary authorities or districts)	133	Lower tier authorities (districts) or individual unitary authorities; Individual unitary authorities or LECs (or parts thereof); Districts	443	Wards (or parts thereof)	10 664
EU-25	-	91	-	257	-	1233	-	8 134	-	113 098
EU-27	-	97	-	271	-	1303	-	8 398	-	121 601

Figura 5. Classificazione NUTS per ciascun paese membro

Di seguito si propone l'elenco delle regioni NUTS 2 con i rispettivi codici e la relativa nazione di appartenenza.

BELGIO

BE10 Région de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest
BE21 Prov. Antwerpen
BE22 Prov. Limburg (BE)
BE23 Prov. Oost-Vlaanderen
BE24 Prov. Vlaams-Brabant
BE25 Prov. West-Vlaanderen
BE31 Prov. Brabant Wallon
BE32 Prov. Hainaut
BE33 Prov. Liège
BE34 Prov. Luxembourg (BE)
BE35 Prov. Namur

BULGARIA

BG31 Северозападен / Severozapaden
BG32 Северен централен / Severen tsentralen
BG33 Североизточен / Severoiztochen
BG34 Югоизточен / Yugoiztochen
BG41 Югозападен / Yugozapaden
BG42 Южен централен / Yuzhen tsentralen

REPUBBLICA CECA

CZ01 Praha
CZ02 Střední Čechy
CZ03 Jihozápad
CZ04 Severozápad
CZ05 Severovýchod
CZ06 Jihovýchod
CZ07 Střední Morava
CZ08 Moravskoslezsko

DANIMARCA

DK01 Hovedstaden
DK02 Sjælland
DK03 Syddanmark
DK04 Midtjylland
DK05 Nordjylland

GERMANIA

DE11	Stuttgart
DE12	Karlsruhe
DE13	Freiburg
DE14	Tübingen
DE21	Oberbayern
DE22	Niederbayern
DE23	Oberpfalz
DE24	Oberfranken
DE25	Mittelfranken
DE26	Unterfranken
DE27	Schwaben
DE30	Berlin
DE40	Brandenburg
DE50	Bremen
DE60	Hamburg
DE71	Darmstadt
DE72	Gießen
DE73	Kassel
DE80	Mecklenburg-Vorpommern
DE91	Braunschweig
DE92	Hannover
DE93	Lüneburg
DE94	Weser-Ems
DEA1	Düsseldorf
DEA2	Köln
DEA3	Münster
DEA4	Detmold
DEA5	Arnsberg
DEB1	Koblenz
DEB2	Trier
DEB3	Rheinessen-Pfalz
DEC0	Saarland
DED2	Dresden
DED1	Chemnitz
DED3	Leipzig
DEE0	Sachsen-Anhalt
DEF0	Schleswig-Holstein
DEG0	Thüringen
ESTONIA	
EE00	Eesti
IRLANDA	
IE01	Border, Midland and Western
IE02	Southern and Eastern

GRECIA

EL11	Ανατολική Μακεδονία, Θράκη / Anatoliki Makedonia, Thraki
EL12	Κεντρική Μακεδονία / Kentriki Makedonia
EL13	Δυτική Μακεδονία / Dytiki Makedonia
EL14	Θεσσαλία / Thessalia
EL21	Ήπειρος / Ipeiros
EL22	Ιόνια Νησιά / Ionia Nisia
EL23	Δυτική Ελλάδα / Dytiki Ellada
EL24	Στερεά Ελλάδα / Sterea Ellada
EL25	Πελοπόννησος / Peloponnisos
EL30	Αττική / Attiki
EL41	Βόρειο Αιγαίο / Voreio Aigaio
EL42	Νότιο Αιγαίο / Notio Aigaio
EL43	Κρήτη / Kritis

SPAGNA

ES11	Galicia
ES12	Principado de Asturias
ES13	Cantabria
ES21	País Vasco
ES22	Comunidad Foral de Navarra
ES23	La Rioja
ES24	Aragón
ES30	Comunidad de Madrid
ES41	Castilla y León
ES42	Castilla-La Mancha
ES43	Extremadura
ES51	Cataluña
ES52	Comunidad Valenciana
ES53	Illes Balears
ES61	Andalucía
ES62	Región de Murcia
ES63	Ciudad Autónoma de Ceuta
ES64	Ciudad Autónoma de Melilla
ES70	Canarias

FRANCIA

FR10 Île de France
FR21 Champagne-Ardenne
FR22 Picardie
FR23 Haute-Normandie
FR24 Centre
FR25 Basse-Normandie
FR26 Bourgogne
FR30 Nord - Pas-de-Calais
FR41 Lorraine
FR42 Alsace
FR43 Franche-Comté
FR51 Pays de la Loire
FR52 Bretagne
FR53 Poitou-Charentes
FR61 Aquitaine
FR62 Midi-Pyrénées
FR63 Limousin
FR71 Rhône-Alpes
FR72 Auvergne
FR81 Languedoc-Roussillon
FR82 Provence-Alpes-Côte d'Azur
FR83 Corse
FR91 Guadeloupe
FR92 Martinique
FR93 Guyane
FR94 Réunion

ITALIA

ITC1 Piemonte
ITC2 Valle d'Aosta/Vallée d'Aoste
ITC3 Liguria
ITC4 Lombardia
ITF1 Abruzzo
ITF2 Molise
ITF3 Campania
ITF4 Puglia
ITF5 Basilicata
ITF6 Calabria
ITG1 Sicilia
ITG2 Sardegna
ITD1 Provincia Autonoma di Bolzano/Bozen
ITD2 Provincia Autonoma di Trento
ITD3 Veneto
ITD4 Friuli-Venezia Giulia
ITD5 Emilia-Romagna
ITE1 Toscana
ITE2 Umbria
ITE3 Marche
ITE4 Lazio

CIPRO

CY00 Κύπρος / Kýpros

LITUANIA

LV00 Latvija

LETTONIA

LT00 Lietuva

LUSSEMBURGO

LU00 Luxembourg

UNGHERIA

HU10 Közép-Magyarország

HU21 Közép-Dunántúl

HU22 Nyugat-Dunántúl

HU23 Dél-Dunántúl

HU31 Észak-Magyarország

HU32 Észak-Alföld

HU33 Dél-Alföld

MALTA

MT00 Malta

OLANDA

NL11 Groningen

NL12 Friesland (NL)

NL13 Drenthe

NL21 Overijssel

NL22 Gelderland

NL23 Flevoland

NL31 Utrecht

NL32 Noord-Holland

NL33 Zuid-Holland

NL34 Zeeland

NL41 Noord-Brabant

NL42 Limburg (NL)

AUSTRIA

AT11 Burgenland (AT)

AT12 Niederösterreich

AT13 Wien

AT21 Kärnten

AT22 Steiermark

AT31 Oberösterreich

AT32 Salzburg

AT33 Tirol

AT34 Vorarlberg

POLONIA

PL11 Łódzkie

PL12 Mazowieckie

PL21 Małopolskie

PL22 Śląskie

PL31 Lubelskie

PL32 Podkarpackie

PL33 Świętokrzyskie

PL34 Podlaskie

PL41 Wielkopolskie

PL42 Zachodniopomorskie

PL43 Lubuskie

PL51 Dolnośląskie

PL52 Opolskie

PL61 Kujawsko-Pomorskie

PL62 Warmińsko-Mazurskie

PL63 Pomorskie

PORTOGALLO

PT11 Norte
PT15 Algarve
PT16 Centro (PT)
PT17 Lisboa
PT18 Alentejo
PT20 Região Autónoma dos Açores
PT30 Região Autónoma da Madeira

ROMANIA

RO11 Nord-Vest
RO12 Centru
RO21 Nord-Est
RO22 Sud-Est
RO31 Sud - Muntenia
RO32 București - Ilfov
RO41 Sud-Vest Oltenia
RO42 Vest

SLOVENIA

SI01 Vzhodna Slovenija
SI02 Zahodna Slovenija

SLOVACCHIA

SK01 Bratislavský kraj
SK02 Západné Slovensko
SK03 Stredné Slovensko
SK04 Východné Slovensko

FINLANDIA

FI19 Länsi-Suomi
FI1B Helsinki-Uusimaa
FI18 Etelä-Suomi
FI1A Pohjois- ja Itä-Suomi
FI20 Åland

SVEZIA

SE11 Stockholm
SE12 Östra Mellansverige
SE21 Småland med öarna
SE22 Sydsverige
SE23 Västsverige
SE31 Norra Mellansverige
SE32 Mellersta Norrland
SE33 Övre Norrland

GRAN BRETAGNA

UKC1 Tees Valley and Durham
UKC2 Northumberland and Tyne and Wear
UKD1 Cumbria
UKD3 Greater Manchester
UKD4 Lancashire
UKD6 Cheshire
UKD7 Merseyside
UKE1 East Yorkshire and Northern Lincolnshire
UKE2 North Yorkshire
UKE3 South Yorkshire
UKE4 West Yorkshire
UKF1 Derbyshire and Nottinghamshire
UKF2 Leicestershire, Rutland and Northamptonshire
UKF3 Lincolnshire
UKG1 Herefordshire, Worcestershire and Warwickshire
UKG2 Shropshire and Staffordshire
UKG3 West Midlands
UKH1 East Anglia
UKH2 Bedfordshire and Hertfordshire
UKH3 Essex
UKI1 Inner London
UKI2 Outer London
UKJ1 Berkshire, Buckinghamshire and Oxfordshire
UKJ2 Surrey, East and West Sussex
UKJ3 Hampshire and Isle of Wight
UKJ4 Kent
UKK1 Gloucestershire, Wiltshire and Bristol/Bath area
UKK2 Dorset and Somerset
UKK3 Cornwall and Isles of Scilly
UKK4 Devon
UKL1 West Wales and The Valleys
UKL2 East Wales
UKM2 Eastern Scotland
UKM3 South Western Scotland
UKM5 North Eastern Scotland
UKM6 Highlands and Islands
UKN0 Northern Ireland

La fonte dei dati è costituita dal database on-line predisposto da Eurostat all'indirizzo http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Eurostat elabora un sistema di indicatori per analizzare lo stato della sostenibilità e della qualità della vita a livello dei singoli stati membri dell'Unione Europea.

Tale database costituisce uno strumento di attuazione e monitoraggio della Strategia di sviluppo sostenibile, finalizzata a conciliare lo sviluppo economico con la coesione sociale e la protezione per l'ambiente. La dimensione nazionale di analisi non pare però sufficiente a soddisfare i fabbisogni informativi: un'analisi più dettagliata basata quindi sul livello regionale è in grado di evidenziare meglio i fenomeni oggetto di studio. Osservando l'intera lista di indicatori, si nota la presenza pressoché esclusiva di dati quantitativi ed in particolare una sovrabbondanza di dati riguardanti l'ambito economico-finanziario riferiti all'intero territorio nazionale. Cliccando sul menù ad espansione "*Regional statistic*" si nota subito che l'albero di navigazione (riportato nella sua interezza in appendice) è piuttosto articolato; tuttavia, osservandolo con maggior dovizia di particolari, ci si rende conto della limitata disponibilità di indicatori utilizzabili per realizzare l'analisi sulla qualità della vita. Com'era logico aspettarsi, la quantità di dati a disposizione è sempre minore via via che l'unità territoriale considerata è più ristretta. A dimostrazione di ciò si ravvisa la scarsa attenzione ad indagare gli ambiti non economici del benessere, quali *in primis* la soddisfazione individuale e pure le connessioni sociali e l'impegno civico: indicatori riferiti a tali temi sono pressoché assenti se prendiamo in considerazione il livello regionale. Anche le due dimensioni trasversali, cioè l'equità e la sostenibilità, costituiscono due aspetti pressoché tralasciati all'interno delle regioni.

Tenuto conto di tale limite, tra gli indicatori a disposizione vengono individuati quelli che meglio sembrano descrivere gli ambiti di benessere definiti dall'Oecd. (Come precisato nel capitolo primo, si è consapevoli del fatto che il processo di selezione degli indicatori, per essere realmente condiviso e legittimato, dovrebbe coinvolgere tutti i rappresentanti delle diverse componenti della società.)

Al fine di rimanere per quanto possibile fedeli allo schema Oecd, vengono dapprima proposti gli indicatori relativi al benessere materiale.

REDDITO, RICCHEZZA, LAVORO E CONDIZIONI ABITATIVE

- Pil a prezzi correnti per abitante;
- Reddito familiare disponibile;
- Reddito familiare primario;
- Numero medio di ore lavorate;
- Numero veicoli privati posseduti;
- Titolo di occupazione della casa (% proprietari case private);
- Tasso deprivazione (casa affollata con almeno un aspetto sintomo di deprivazione);
- Grado di indipendenza (completamente indipendente, semi indipendente, appartamento).

Relativamente al benessere non materiale si esaminano separatamente gli aspetti relativi ai bisogni primari dell'individuo e quelli descrittivi dei bisogni secondari, legati ai rapporti interpersonali e all'ambiente circostante.

BISOGNI PRIMARI

SALUTE e SANITA'

- Aspettativa di vita alla nascita;
- Tasso fertilità;
- Tasso mortalità infantile;
- Indice di vecchiaia;
- Indice di dipendenza;
- Personale medico;
- Numero letti ospedalieri;
- Numero vittime di incidenti stradali

OCCUPAZIONE

- Tasso popolazione attiva
- Tasso di occupazione 15-64;
- Tasso di disoccupazione;
- Tasso di disoccupazione di lungo periodo;

ISTRUZIONE

- % popolazione con istruzione secondaria;
- % popolazione con istruzione terziaria;
- Numero brevetti applicati
- Adulti che partecipano a corsi di formazione;
- Famiglie con accesso ad Internet da casa;
- Famiglie in possesso di banda larga;
- Persone che usano regolarmente Internet;
- Persone che non hanno mai usato un Pc

BISOGNI SECONDARI

AMBIENTE

- Ricchezza del paesaggio;
- % suolo occupato da prati;
- % suolo occupato da foreste;
- % suolo occupato artificialmente;
- Saldo bilancio energetico
- Risorse acqua dolce;
- Smaltimento rifiuti urbani;

EQUITA' E SOSTENIBILITA'

- Disparità nella distribuzione del reddito;
- Disparità nel mercato del lavoro (differenza tasso occupazione per genere)
- % di energie rinnovabili sul totale dei consumi

2.2 TRATTAMENTO PRELIMINARE DEI DATI

Il possesso di informazioni affidabili è la condizione necessaria per l'ottenimento di analisi altrettanto affidabili.

La qualità dei dati Eurostat è garantita dal rispetto delle raccomandazioni impartite da FMI e OECD: tali organismi valutano la qualità osservando 6 dimensioni: pertinenza, precisione, tempestività, accessibilità, comprensibilità, coerenza.

Ipotizziamo pertanto di possedere dati di buona qualità, condizione strettamente necessaria per un corretto svolgimento dell'analisi dei dati.

Per tenere conto della diversa unità di misura in cui i singoli indicatori sono espressi e del loro diverso grado di variabilità territoriale, è necessario eliminare tale fattore di incomparabilità mediante il calcolo dei rispettivi valori standardizzati. (Si potrebbero comunque utilizzare altre tecniche di normalizzazione);

Si segnala una pluralità di dati mancanti, cioè diversi casi in cui l'unità statistica, la regione, presenta un valore che non è possibile conoscere.

Varie sono le strategie adottabili per ovviare a tale problema e ciascuna di esse comporta pregi e difetti, difatti non esiste una metodologia migliore in assoluto per il trattamento di *missing data* (Rubin, 1987). I metodi perseguibili sono essenzialmente di due tipi: deterministici e stocastici. Optando per la prima tipologia, si tralascia l'ipotesi di *Complete Case Analysis* (tecnica basata sull'eliminazione di qualsiasi unità sia stata solo parzialmente osservata) poiché comporterebbe l'impensabile cancellazione delle regioni per le quali non si dispone dei dati di tutti gli indicatori.

Laddove è possibile si ricorre al metodo deduttivo, in base al quale il valore imputato è dedotto da informazioni o relazioni note: in pratica si riempie il *blank* con il valore riferito all'anno precedente.

Qualora tale espediente non sia possibile, si procede con l'assegnazione di un opportuno valore ai dati mancanti; in particolare si sceglie di imputare il valore medio, cioè a ciascuno dei dati mancanti d'una variabile si sostituisce la media calcolata su tutte le regioni per le quali si conoscono i valori.

Occorre però tenere presente che tale scelta "... è sia seducente sia pericolosa perché può cullare l'utente in uno stato di piacevole sensazione credendo che i dati, dopo tutto, siano completi" (Dempster, Rubin, 1983).

Ad ogni modo se il numero di dati mancanti per ciascun indicatore si mantiene al di sotto di una soglia ragionevole (quale può essere il 15% delle unità), si opta per l'imputazione del valor medio: si sostituiscono tutte le mancate risposte relative ad una determinata variabile con un unico valore, cioè la media calcolata sul totale delle regioni.

Si sceglie di eliminare la variabile che presenta un numero di valori mancanti superiore ai 30.

Ci si trova pertanto nella condizione di dover eliminare tutte le variabili relative all'utilizzo del Personal Computer ed Internet, in quanto presentano una percentuale di valori mancanti che si avvicina al 40%; riguardo all'ambito della salute si escludono i caratteri "numero letti in ospedale" e "numero di dottori per 100.000 residenti" (cui corrispondono rispettivamente il 27 ed il 33% di valori mancanti); nella categoria ambiente si tralasciano le variabili riguardanti il trattamento delle acque reflue, dei rifiuti e dell'energia (con 57, 46 e 37% di *missing data*).

Statistiche univariate

	N	Media	Deviazione std.	Mancante		N. di estremi ^a	
				Conteggio	Percentuale	Min	Max
pop_tot	266	1863086,82	1555936,559	0	,0	0	16
area	266	16097,4082	19843,36458	0	,0	0	19
densità	266	382,763	953,1302	0	,0	0	26
under15	266	291332,015	257929,5270	0	,0	0	14
tra15_64	266	1248509,049	1047992,240	0	,0	0	16
over65	266	323245,192	267792,0548	0	,0	0	23
ind_vecc	266	116,3176	35,27814	0	,0	1	7
ind_dip	266	49,9257	5,52600	0	,0	0	0
under10	260	10,3835	1,81358	6	2,3	0	5
over74	266	8,0803	2,23356	0	,0	7	1
fertilità	266	1,6254	,32846	0	,0	0	1
t_mov_nat	266	,7587	3,61571	0	,0	1	11
t_mov_mig	266	2,7803	4,25372	0	,0	0	2
agricoltura	236	48,5278	17,74250	30	11,3	2	0
foreste	236	22,6296	15,99501	30	11,3	0	0
caccia_pesca	236	1,3804	2,51706	30	11,3	0	18
alto_impatto_amb	236	5,0897	4,27254	30	11,3	0	18
servizi_resid	236	13,4318	11,52379	30	11,3	0	15
TRAT_ACQUE_2008	114	79,5489	24,82322	152	57,1	7	0
TRATT_RIFIUTI_2008	144	94,4531	12,39451	122	45,9	27	0
Rif_pers_kg	266	378,7969	261,75140	0	,0	0	0
Energyrec_recycling	197	26,8500	27,50110	69	25,9	0	0
Activity_rate_1564	266	71,2933	6,38340	0	,0	4	0
Employ_rate15_64	266	64,634	7,7082	0	,0	1	0
E_over64	266	5,030	3,9450	0	,0	0	8
employ_Males	266	70,533	6,9527	0	,0	2	0
employ_Females	266	58,714	9,3887	0	,0	4	0
Employ_MF	266	11,819	6,0106	0	,0	1	20
unemploy_rate	266	9,388	4,7701	0	,0	0	13
long_unempl	266	3,8356	2,85938	0	,0	0	10
unemp_sex_Males	266	9,247	4,6331	0	,0	0	16
unemp_sex_Females	266	9,652	5,4421	0	,0	0	14
UNEMPL_FM	266	,405	2,8702	0	,0	4	16
unempl_15_24	266	21,979	11,0428	0	,0	0	6
DIFF_DIS_GIOV_AD	266	14,006	7,7831	0	,0	0	2
GDP	266	96,699	46,3612	0	,0	0	7
PPT	266	96,387	36,2889	0	,0	0	9
reddito_primario_netto	253	99,5577	31,32037	13	4,9	0	1
reddito_disponibile_netto	253	100,0552	26,86994	13	4,9	0	1
vehicle1000	249	1029,5582	904,72781	17	6,4	0	19
t_non_lavoro	266	77,6015	1,62949	0	,0	1	10
asp_vita	264	79,614	2,5940	2	,8	22	0
infant_mort	266	4,0756	1,96076	0	,0	1	18
letti_osp	195	526,230	184,2390	71	26,7	0	0
dottori	179	334,044	116,9672	87	32,7	0	9
vittime_inc	266	77,9624	38,61721	0	,0	0	5
lower_secondary	256	27,220	14,7333	10	3,8	0	7
upper_sec	256	47,515	14,7788	10	3,8	0	0
terz_edu	256	25,265	8,6028	10	3,8	0	1
no_use_pc	167	21,425	14,2872	99	37,2	0	0
bandalarga	135	62,904	15,9593	131	49,2	1	0
internet_acc	166	67,831	15,8442	100	37,6	0	0
online_order	164	37,927	23,7113	102	38,3	0	0
int_reg_use	166	65,301	17,0541	100	37,6	0	0
R_d_gdp	248	1,6039	1,32492	18	6,8	0	12
formaz_ad	249	10,087	7,1282	17	6,4	0	5
patent_app	241	42,9224	53,88503	25	9,4	0	11

a. Numero di casi non compresi nell'intervallo (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

2.3 CRITERI DI SELEZIONE DEGLI INDICATORI DI BENESSERE

Si ribadisce che nella scelta degli indicatori ha influito notevolmente la disponibilità dei dati: laddove possibile si ricorre ad identificare variabili *proxy*; diversamente si descrive quale sarebbe l'indicatore più opportuno per colmare l'assenza informativa.

Si prendono in considerazione solo indicatori relativi che tengano conto della diversa dimensione delle regioni: le misure sono rapportate alla popolazione residente o un altro aggregato di riferimento più appropriato, oppure calcolate in percentuale.

L'anno di riferimento è quello più recente a disposizione, in prevalenza gli anni 2009 e 2010.

Consapevoli dell'ineliminabile arbitrarietà nella scelta degli indicatori elementari di benessere, si illustrano le metodologie di analisi dei dati che consentono di spiegare il fenomeno oggetto di analisi, cioè il grado di benessere che caratterizza le regioni europee.

Per soddisfare l'esigenza di sintesi occorre effettuare un'ulteriore selezione degli indicatori, verificando che ciascuno di essi presenti valore discriminante tra i territori soggetti a confronto, cioè la variabilità del singolo indicatore dovrà essere sufficientemente ampia da rendere significative le differenze tra una regione e l'altra: a tal scopo si calcola il coefficiente di variazione.

Si procede con il calcolo di matrici di correlazione per studiare le relazioni esistenti tra gli indicatori appartenenti a ciascun ambito di benessere ed eliminare le variabili ridondanti, cioè scartare quegli indicatori che misurano lo stesso aspetto del fenomeno analizzato.

L'obiettivo è pertanto quello di ridurre il numero di indicatori per facilitare l'interpretazione dei risultati, minimizzando l'inevitabile perdita di informazioni.

2.3.1 VARIABILI DEMOGRAFICHE

Qualsiasi tipo di ricerca sociale non può prescindere dalla considerazione della situazione demografica che caratterizza il relativo ambito di analisi.

Di seguito si esaminano gli indicatori estratti dal database predisposto da Eurostat. Si tratta della categoria più numerosa, aggiornata e completa, proprio a sottolinearne l'importanza ai fini della programmazione delle strutture sanitarie, previdenziali, scolastiche etc.

Si ritiene inoltre che la qualità dei dati demografici sia particolarmente elevata.

ETICHETTA	INDICATORE DEMOGRAFICO
ind_vecc	Indice di vecchiaia: $\text{vecchiaia} > 64 / (0-14) * 100$
ind_dip	Indice di dipendenza $(<15+>64) / (15-64) * 100$
%<10	Percentuale popolazione di età inferiore a 10 anni
%>74	Percentuale popolazione di età superiore a 74 anni
fert_rate	Tasso di fertilità totale: numero medio di figli per donna in età fertile
nat_change_M 2006_2010	Tasso netto movimento naturale popolazione (media anni 2006-2010)
nat_migration_M 2006_2010	Tasso netto movimento migratorio popolazione (media anni 2006-2010)

Possiamo suddividerle in variabili di stato e caratteri dinamici.

Esaminando la struttura per età della popolazione, si prendono in considerazione la fascia giovanile, cioè coloro che ancora non hanno compiuto 10 anni di età, i cosiddetti “grandi anziani”, ossia quanti abbiano superato la soglia dei 74 anni, ed inoltre i classici indici di vecchiaia e di dipendenza. Il primo misura lo squilibrio della struttura per età della popolazione; indica infatti il numero di persone di 65 anni e oltre, su 100 persone di età fino a 14 anni; il secondo è un indicatore di carico

sociale, poiché rapporta la popolazione non attiva a quella in età lavorativa, infatti è calcolato come rapporto tra le persone di età inferiore a 15 anni o superiore a 64, sul totale delle persone aventi un'età compresa tra i 15 ed i 64 anni, in percentuale.

Relativamente al tema del ricambio della popolazione, si analizzano:

- il tasso di fertilità totale, cioè il numero medio di figli per donna. Esso si calcola come somma dei quozienti specifici di fecondità calcolati rapportando, per ogni età feconda (15-49 anni) il numero di nati vivi all'ammontare medio annuo della popolazione femminile;

- i tassi legati al movimento naturale e migratorio, di cui si calcola una media semplice degli ultimi 5 anni in modo da eliminare l'effetto delle oscillazioni annuali (rispettivamente la differenza tra nati e morti, quella tra emigrati ed immigrati, entrambe calcolate sulla popolazione residente per mille).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
ind_vecc	266	11,73	241,11	116,32	35,28	30,33
ind_dip	266	34,93	64,69	49,93	5,53	11,07
tra 0 e 9	266	7,20	24,37	10,38	1,80	17,32
75 years or over	266	1,62	13,85	8,23	1,88	22,89
tasso fertilità totale	266	1,03	3,56	1,63	0,33	20,21
tasso movimento naturale pop	266	-10,18	25,45	0,76	3,62	476,59
tasso movimento migratorio	266	-7,25	15,86	2,78	4,25	152,99
Validi (listwise)	266					

Le misure descrittive forniscono una misura di sintesi di ciascun fenomeno, rappresentata dalla media, e due misure di variabilità assoluta e relativa, cioè rispettivamente lo scostamento quadratico medio ed il coefficiente di variazione.

La matrice di correlazione riporta il coefficiente di correlazione di Pearson calcolato per ciascuna delle possibili coppie di indicatori: com'è noto tale coefficiente misura la relazione lineare tra due variabili ed assume valori nell'intervallo chiuso -1 e +1.

Correlazioni

		ind_vecc	ind_dip	tra 0 e 9	75 years or over	tasso fertilità totale	tasso movimento naturale pop	tasso movimento migratorio
ind_vecc	Correlazione di Pearson	1	,360**	-,803**	,795**	-,508**	-,758**	-,015
	Sig. (2-code)		,000	,000	,000	,000	,000	,806
	N	266	266	266	266	266	266	266
ind_dip	Correlazione di Pearson	,360**	1	,135*	,674**	,487**	-,033	,104
	Sig. (2-code)	,000		,028	,000	,000	,593	,092
	N	266	266	266	266	266	266	266
tra 0 e 9	Correlazione di Pearson	-,803**	,135*	1	-,510**	,833**	,893**	,114
	Sig. (2-code)	,000	,028		,000	,000	,000	,064
	N	266	266	266	266	266	266	266
75 years or over	Correlazione di Pearson	,795**	,674**	-,510**	1	-,154*	-,523**	,181**
	Sig. (2-code)	,000	,000	,000		,012	,000	,003
	N	266	266	266	266	266	266	266
tasso fertilità totale	Correlazione di Pearson	-,508**	,487**	,833**	-,154*	1	,695**	,086
	Sig. (2-code)	,000	,000	,000	,012		,000	,161
	N	266	266	266	266	266	266	266
tasso movimento naturale pop	Correlazione di Pearson	-,758**	-,033	,893**	-,523**	,695**	1	,218**
	Sig. (2-code)	,000	,593	,000	,000	,000		,000
	N	266	266	266	266	266	266	266
tasso movimento migratorio	Correlazione di Pearson	-,015	,104	,114	,181**	,086	,218**	1
	Sig. (2-code)	,806	,092	,064	,003	,161	,000	
	N	266	266	266	266	266	266	266

** La correlazione è significativa al livello 0,01 (2-code).
* La correlazione è significativa al livello 0,05 (2-code).

La variabile “percentuale di residenti aventi un’età inferiore ai 10 anni” presenta una correlazione diretta molto elevata con il tasso di movimento naturale ($r=0,893$) ed il tasso di fertilità totale ($r=833$); per contro la relazione tra la percentuale di over 10 e l’indice di vecchiaia è ovviamente di tipo inverso ($r=-0,803$).

L’esistenza di una pluralità di apprezzabili relazioni lineari tra le varie coppie di indicatori fa supporre l’opportunità dell’analisi in componenti principali (ACP). Tale metodologia consente di

sintetizzare i sette indicatori sopra descritti in poche variabili, dette componenti principali, con una perdita limitata di informazioni.

L'ACP riesce a dare una prima interpretazione di come la correlazione tra le variabili considerate possa essere spiegata da fattori comuni sottostanti; ogni componente principale è espressa come combinazione lineare degli scostamenti standardizzati degli indicatori originari.

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	3,659	52,278	52,278	3,659	52,278	52,278
2	1,943	27,758	80,037	1,943	27,758	80,037
3	,985	14,065	94,102			
4	,176	2,521	96,622			
5	,111	1,589	98,211			
6	,091	1,299	99,509			
7	,034	,491	100,000			

Metodo di estrazione: Analisi componenti principali.

Comunalità

	Iniziale	Estrazione
Z(ind_vecc)	1,000	,916
Z(ind_dip)	1,000	,909
Z: tra 0 e 9	1,000	,964
Z: 75 years or over	1,000	,931
Z: tasso fertilità totale	1,000	,889
Z: tasso movimento naturale pop	1,000	,868
Z: tasso movimento migratorio	1,000	,125

Metodo di estrazione: Analisi componenti principali.

La variabile standardizzata tasso movimento migratorio è spiegata solo per il 12,5% della sua varianza e pertanto è ragionevole eliminarla dall'analisi delle componenti principali.

Se ne ricava un netto miglioramento della percentuale di varianza spiegata dalle due prime componenti principali, che in questo modo sale al 92,3% (valore ben superiore a quello soglia, pari al 73,5%)

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	3,653	60,890	60,890	3,653	60,890	60,890
2	1,884	31,394	92,284	1,884	31,394	92,284
3	,218	3,631	95,915			
4	,119	1,983	97,898			
5	,092	1,528	99,426			
6	,034	,574	100,000			

Comunalità

	Iniziale	Estrazione
Z(ind_vecc)	1,000	,916
Z(ind_dip)	1,000	,957
Z: tra 0 e 9	1,000	,968
Z: 75 years or over	1,000	,915
Z: tasso fertilità totale	1,000	,924
Z: tasso movimento naturale pop	1,000	,857

Metodo di estrazione: Analisi componenti principali.

Ora tutte le variabili presentano una percentuale di varianza spiegata che va oltre l'85%

La prima componente può essere interpretata come un indicatore di giovinezza della popolazione residente, presenta infatti una forte relazione diretta con la percentuale di residenti aventi fino a 9 anni ($r=0,948$), il tasso relativo al movimento naturale ($r=0,917$) ed inversa con l'indice di vecchiaia ($-0,922$); la seconda è strettamente legata all'indice di dipendenza ($r=0,969$) ed in misura minore alla percentuale di residenti aventi 75 anni o più.

Matrice di componenti^a

	Componente	
	1	2
Z(ind_vecc)	-,922	,257
Z(ind_dip)	-,136	,969
Z: tra 0 e 9	,948	,265
Z: 75 years or over	-,716	,634
Z: tasso fertilità totale	,731	,625
Z: tasso movimento naturale pop	,917	,127

2.3.2 VARIABILI LEGATE ALLA SALUTE E INCIDENTALITA'

Esaminando la dimensione relativa alla salute, il database permette di estrarre unicamente variabili strumentali, dette anche indicatori di mezzo; per contro si risente della mancanza di variabili finali, cioè di indicatori di risultato, ma se ne comprendono le difficoltà insite nella rilevazione.

<i>ETICHETTA</i>	<i>INDICATORE SALUTE</i>
asp_vita_2009	Speranza di vita alla nascita: numero medio di anni di vita alla nascita
infant_mort	Tasso di mortalità infantile: morti nel primo anno di vita su 1.000 nati vivi
vittime_inc_2009	Morti in incidenti stradali sul colpo o entro 30 giorni su 100.000 residenti

La speranza di vita alla nascita differisce a seconda del genere, a vantaggio della componente femminile, e dipende da una serie di fattori tra cui le differenze a livello di stile di vita, clima e alimentazione, ma anche di istruzione, reddito e accesso all'assistenza sanitaria e ad altri servizi sociali che incidono sulla salute delle persone (in realtà sarebbe meglio considerare gli anni di vita vissuti in buona salute psico-fisica).

Come variabile *proxy* della qualità del servizio sanitario prestato ai cittadini (anche se altri fattori sono importanti) abbiamo a disposizione unicamente il tasso di mortalità infantile⁹ su 1000 nati vivi.

Si prende in esame, inoltre, il numero di persone decedute a seguito di sinistro stradale, avvenuta entro 30 giorni dall'evento (in realtà sarebbe opportuno tener conto anche degli individui rimasti invalidi o che hanno riportato danni permanenti in conseguenza dell'incidente stradale)

⁹ Ovviamente anche altri fattori, quali lo stato di salute e l'età della madre, sono determinanti.

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
SMEAN(asp_vita)	266	72,90	83,30	79,61	2,58	3080,78
tasso mortalità infantile	266	,0	15,8	4,08	1,96	207,86
vittime incidenti stradali	266	14,0	238,0	77,96	38,62	201,89
Validi (listwise)	266					

Correlazioni

		SMEAN(asp_vita)	tasso mortalità infantile	vittime incidenti stradali
SMEAN(asp_vita)	Correlazione di Sig. (2-code)	1	-,600**	-,413**
	N	266	266	266
tasso mortalità infantile	Correlazione di Sig. (2-code)	-,600**	1	,254**
	N	266	266	266
vittime incidenti stradali	Correlazione di Sig. (2-code)	-,413**	,254**	1
	N	266	266	266

** . La correlazione è significativa al livello 0,01 (2-code).

2.3.3 VARIABILI RELATIVE ALL'ISTRUZIONE

<i>ETICHETTA</i>	<i>INDICATORE ISTRUZIONE</i>
lower_secondary	Basso livello di educazione, livelli 0, 1, 2
upper_sec	Medio livello di educazione, livelli 3 e 4
terz_edu	Alto livello di educazione, livelli 5 e 6
formaz_ad	Life long learning
r_d	Investimenti in ricerca e sviluppo in % rispetto al Pil
patent_app	Brevetti riconosciuti dall'European Patent Office per 1.000.000 residenti

Riguardo al livello d'istruzione, Eurostat adotta la classificazione internazionale ISCED: *International Standard Classification of Education*. Essa è basata su 6 livelli, a loro volta raggruppabili in tre classi: basso grado di educazione, corrispondente ai livelli 0, 1 e 2 (scuola d'infanzia, primaria e secondaria inferiore); medio grado di educazione, coincidente con i livelli 3 e 4 (scuola secondaria superiore); alto grado di educazione, comprendente i livelli 5 e 6 (istruzione post-secondaria non universitaria, istruzione universitaria).

Gli altri indicatori estratti hanno lo scopo di misurare il grado di crescita culturale delle persone adulte nell'arco dell'intera vita, cioè l'apprendimento permanente, in inglese *Life Long Learning*: a tal scopo viene presa in esame la percentuale di partecipanti ad un corso di formazione (svoltosi nelle quattro settimane che precedono lo svolgimento dell'intervista) finalizzato a migliorare conoscenze, competenze e abilità all'interno dell'ambito lavorativo, sociale o personale; l'importanza attribuita al settore Ricerca e sviluppo possiamo desumerla sia dall'entità degli investimenti ad esso assegnato (calcolati in percentuale rispetto al Pil), che dai risultati concretamente realizzati, quale il numero di brevetti riconosciuti e registrati, che testimonia il grado di ingegno e creatività intellettuale raggiunto dalla popolazione.

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
SMEAN(lower_secondary)	266	3,3	78,4	27,17	14,82	183,35
SMEAN(upper_sec)	266	11,8	78,5	47,46	14,70	322,96
SMEAN(terz_edu)	266	9,0	53,1	25,37	8,55	296,78
SMEAN(R_d_gdp)	266	,10	7,93	1,60	1,28	125,39
SMEAN(formaz_ad)	266	,8	36,1	10,09	6,90	146,28
SMEAN(patent_app)	266	,14	278,45	42,92	51,28	83,70
Validi (listwise)	266					

Correlazioni

		SMEAN(lower_se condary)	SMEAN(upper_s ec)	SMEAN(terz_edu)	SMEAN(R_d_gdp)	SMEAN(formaz_ ad)	SMEAN(patent_a pp)
SMEAN(lower_secondary)	Correlazi Sig. (2- N	1	-,832** ,000	-,303** ,000	-,220** ,000	-,125** ,041	-,211** ,001
SMEAN(upper_sec)	Correlazi Sig. (2- N	-,832** ,000	1	-,277** ,000	-,040 ,518	-,204** ,001	,163** ,008
SMEAN(terz_edu)	Correlazi Sig. (2- N	-,303** ,000	-,277** ,000	1	,451** ,000	,568** ,000	,086 ,162
SMEAN(R_d_gdp)	Correlazi Sig. (2- N	-,220** ,000	-,040 ,518	,451** ,000	1	,334** ,000	,429** ,000
SMEAN(formaz_ad)	Correlazi Sig. (2- N	-,125** ,041	-,204** ,001	,568** ,000	,334** ,000	1	,042 ,497
SMEAN(patent_app)	Correlazi Sig. (2- N	-,211** ,001	,163** ,008	,086 ,162	,429** ,000	,042 ,497	1
		266	266	266	266	266	266

** . La correlazione è significativa al livello 0,01 (2-code).* . La correlazione è significativa al livello 0,05 (2-code).

Ad eccezione della coppia basso e medio grado di educazione ($r=-0,832$), le rimanenti coppie di variabili non presentano valori dell'indice di correlazione molto elevati in valore assoluto.

Per evitare problemi di multicollinearità, in quanto la somma delle prime tre variabili è pari a 100, si sceglie di eliminare la variabile relativa al livello culturale più basso, conservando quelle riguardanti un grado di educazione medio ed elevato, poiché registrano coefficienti di variazione maggiori.

2.3.4 VARIABILI RELATIVE AL LAVORO

<i>ETICHETTA</i>	<i>INDICATORE LAVORO</i>
Activity_rate_15_64	tasso di attività (Popolazione attiva 15_64/Popolazione 15_64*100)
E15_64	tasso di occupazione (Occupati 15_64/Popolazione 15_64*100)
employ_m	tasso di occupazione maschile
employ_f	tasso di occupazione femminile
employ_M-F	differenza tra tasso occupazione maschile e quello femminile
disoccup_totale	tasso disoccupazione totale (disoccupati ed in cerca primo lavoro su popolazione attiva)
long_unempl	tasso disoccupazione di lungo periodo (superiore ai 12 mesi) su popolazione attiva
unemp_M	tasso disoccupazione maschile
unemp_F	tasso disoccupazione femminile
unemp_F-M	differenza tra tasso disoccupazione femminile e quello maschile
disoccupaz_giov	tasso disoccupazione giovanile
diff_dis_giov-ad	differenza tra tasso disoccupazione giovanile e quello riferito agli over 24
weekly_worked_hours	numero ore settimanali lavorate

Gli indicatori estratti descrivono le caratteristiche quantitative dell'occupazione e della disoccupazione, considerandone le differenze esistenti in termini di genere ed età.

In base alla definizione Eurostat, il tasso di attività e quello di occupazione vengono calcolati sul totale dei residenti di età compresa tra i 15 ed i 64 anni.

Il tasso di attività misura la propensione al lavoro di una determinata popolazione. Si precisa che la popolazione attiva comprende gli occupati, i disoccupati o quelli in cerca di prima occupazione che abbiano compiuto 15 anni.

Il tasso di disoccupazione misura le difficoltà legate all'inserimento nel mercato del lavoro ed è calcolato come rapporto tra le persone che cercano attivamente un'occupazione e la popolazione attiva. In particolare si tiene conto della disoccupazione che perdura da oltre 12 mesi, in quanto la persistenza del fenomeno nel lungo periodo comporta l'acuirsi degli effetti negativi monetari e non monetari ad esso correlati, determinando l'aggravarsi della situazione di disagio sociale.

Sempre riguardo alla disoccupazione si osservano inoltre le disparità di genere e quelle intergenerazionali, calcolando rispettivamente la differenza tra il tasso di disoccupazione femminile e quello maschile e lo scarto tra il tasso di disoccupazione giovanile, relativo cioè agli individui tra i 15 ed i 24 anni, ed il medesimo tasso riferito agli adulti (persone di età compresa tra i 25 ed i 64 anni).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
Activity_rate_15 64	266	46,44	81,87	71,2933	6,38	8,95
tasso occupazione tra 15 e 64	266	39,9	78,0	64,634	7,71	11,93
tasso occupazione maschile	266	49,8	83,8	70,533	6,95	9,86
tasso occupazione femminile	266	25,7	75,7	58,714	9,39	15,99
differenza occupazione m-f	266	-1,9	36,1	11,819	6,01	50,85
tasso disoccupazione	266	2,7	28,9	9,388	4,77	50,81
tasso disoccup lunga durata	266	,19	18,77	3,8356	2,86	74,55
disocc_maschile	266	2,3	29,2	9,247	4,63	50,10
disocc_femm	266	2,5	31,9	9,652	5,44	56,38
differenza disocc f-m	266	-7,6	13,7	,405	2,87	708,88
disoccupazione giovanile	266	4,2	60,2	21,979	11,04	50,24
differenza disoc giov-ad	266	,1	41,6	14,006	7,78	55,57
Validi (listwise)	266					

Le variabili esaminate presentano elevati coefficienti di correlazione in modulo, tali da rendere superfluo l'utilizzo di alcuni indicatori: banalmente il tasso di attività registra una relazione lineare

molto alta con il tasso di occupazione ($r=0,915$); si conserva quest'ultimo, poiché rappresenta un aspetto più specifico del mercato del lavoro.

Calcolando il coefficiente di correlazione tra il tasso di occupazione totale e quelli specifici per ciascun genere, notiamo l'elevato valore del coefficiente tra il primo indicatore e gli altri due: per soddisfare un criterio di parsimonia si prosegue l'analisi considerando solo il tasso globale, ma tenendo conto delle differenze esistenti tra i tassi di occupazione maschile e quello femminile.

Correlazioni

		tasso occupazione tra 15 e 64	tasso occupazione maschile	tasso occupazione femminile	differenza occupazione m-f
tasso occupazione tra 15 e 64	Correlazione di Pearson	1	,923**	,955**	-,424**
	Sig. (2-code)		,000	,000	,000
	N	266	266	266	266
tasso occupazione maschile	Correlazione di Pearson	,923**	1	,769**	-,044
	Sig. (2-code)	,000		,000	,474
	N	266	266	266	266
tasso occupazione femminile	Correlazione di Pearson	,955**	,769**	1	-,673**
	Sig. (2-code)	,000	,000		,000
	N	266	266	266	266
differenza occupazione m-f	Correlazione di Pearson	-,424**	-,044	-,673**	1
	Sig. (2-code)	,000	,474	,000	
	N	266	266	266	266

** La correlazione è significativa al livello 0,01 (2-code).

Riguardo al tasso di disoccupazione totale e quello specifico per genere si procede nello stesso modo: dato il valore del coefficiente di correlazione prossimo ad 1, si conserva la variabile riferita al totale.

Tra il tasso di disoccupazione giovanile e la variabile “differenza tra il tasso di disoccupazione giovanile e quello riferito agli adulti” si salva quest'ultima.

Correlazioni

		tasso occupazione tra 15 e 64	differenza occupazione m-f	tasso disoccupazione	tasso disoccup lunga durata	differenza disocc f-m	differenza disoc giov-ad
tasso occupazione tra 15 e 64	Correlazione di Pearson	1	-,424**	-,731**	-,744**	-,364**	-,748**
	Sig. (2-code)		,000	,000	,000	,000	,000
	N	266	266	266	266	266	266
differenza occupazione m-f	Correlazione di Pearson	-,424**	1	,121*	,158**	,749**	,341**
	Sig. (2-code)	,000		,049	,010	,000	,000
	N	266	266	266	266	266	266
tasso disoccupazione	Correlazione di Pearson	-,731**	,121*	1	,893**	,254**	,712**
	Sig. (2-code)	,000	,049		,000	,000	,000
	N	266	266	266	266	266	266
tasso disoccup lunga durata	Correlazione di Pearson	-,744**	,158**	,893**	1	,283**	,635**
	Sig. (2-code)	,000	,010	,000		,000	,000
	N	266	266	266	266	266	266
differenza disocc f-m	Correlazione di Pearson	-,364**	,749**	,254**	,283**	1	,418**
	Sig. (2-code)	,000	,000	,000	,000		,000
	N	266	266	266	266	266	266
differenza disoc giov-ad	Correlazione di Pearson	-,748**	,341**	,712**	,635**	,418**	1
	Sig. (2-code)	,000	,000	,000	,000	,000	
	N	266	266	266	266	266	266

** La correlazione è significativa al livello 0,01 (2-code).

* La correlazione è significativa al livello 0,05 (2-code).

Se ne ottiene pertanto una matrice di dimensioni ridotte su cui andare ad implementare l'analisi delle componenti principali.

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	3,611	60,190	60,190	3,611	60,190	60,190
2	1,445	24,086	84,276	1,445	24,086	84,276
3	,384	6,399	90,674			
4	,323	5,382	96,056			
5	,146	2,431	98,487			
6	,091	1,513	100,000			

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	3,611	60,190	60,190	3,611	60,190	60,190
2	1,445	24,086	84,276	1,445	24,086	84,276
3	,384	6,399	90,674			
4	,323	5,382	96,056			
5	,146	2,431	98,487			
6	,091	1,513	100,000			

Metodo di estrazione: Analisi componenti principali.

Comunalità

	Iniziale	Estrazione
Z: t occupaz tot	1,000	,810
Z: diff occ m-f	1,000	,893
Z: tasso disoc	1,000	,904
Z: tasso disoc LP	1,000	,863
Z: diff disoc f-m	1,000	,842
Z: diff disoc giov-ad	1,000	,744

L'analisi risulta valida: infatti la percentuale di varianza totale spiegata è pari all'84,3%, contro un valore soglia di 73,5%, e tutte le variabili sono spiegate in misura soddisfacente.

Matrice di componenti^a

	Componente	
	1	2
Z: t occupaz tot	-,895	,097
Z: diff occ m-f	,517	,791
Z: tasso disoc	,857	-,411
Z: tasso disoc LP	,852	-,370
Z: diff disoc f-m	,586	,706
Z: diff disoc giov-ad	,859	-,079

Metodo estrazione: analisi componenti principali.

La prima componente spiega il fenomeno della disoccupazione, compresa quella di lungo periodo, tenendo conto anche delle disuguaglianze intergenerazionali, mentre la seconda componente spiega la disuguaglianza tra i sessi sia riguardo al tasso di occupazione sia a quello di disoccupazione.

2.3.5 VARIABILI LEGATE ALLA RICCHEZZA E AL TEMPO LIBERO

L'importanza rivestita dal benessere economico nel determinare almeno in parte il livello della qualità di vita delle persone è fuori discussione.

<i>ETICHETTA</i>	<i>INDICATORE RICCHEZZA</i>
GDP	Prodotto interno lordo a prezzi correnti di mercato (100 = valore medio)
PPT	Pil a parità di potere d'acquisto (100= valore medio)
Primary income	reddito primario netto delle famiglie
Disposable income	reddito disponibile netto delle famiglie
Vehicles	Veicoli posseduti per 1000 abitanti (eccetto rimorchi e motocicli)
Ore non lavoro	Percentuale ore settimanali non dedicate al lavoro retribuito

Nonostante i limiti descritti in precedenza, non si può non considerare nell'analisi il Prodotto interno lordo ai prezzi di mercato, che si definisce come il risultato finale dell'attività di produzione delle unità produttrici residenti, corrispondente alla produzione totale di beni e servizi dell'economia, diminuita dei consumi intermedi e aumentata dell'Iva gravante e delle imposte indirette sulle importazioni. Esso viene utilizzato tradizionalmente per misurare e comparare il grado di attività economica relativo ad un'area territoriale. Il PIL procapite è il valore complessivo dei beni e servizi prodotti all'interno di un territorio in un certo intervallo di tempo e destinati ad usi finali diviso per il numero di abitanti. Il PIL è un indicatore del livello della ricchezza nazionale, mentre nei confronti internazionali viene utilizzato il PIL a parità di potere di acquisto (PPS), al fine di eliminare le distorsioni indotte dai differenti livelli dei prezzi: esso è calcolato come la media pesata dei prezzi di un paniere di beni e servizi che sono omogenei, comparabili e rappresentativi di ogni paese membro dell'Unione Europea. Pertanto il Pil calcolato a PPT tiene conto del costo della vita e dell'inflazione.

In entrambi i casi il Pil è procapite e posto uguale a 100 il valore medio riferito all'intera Unione Europea, si calcolano i valori di ciascuna regione in riferimento a questa base.

Adottando invece le famiglie come unità di riferimento per indagare la relativa situazione economica, ricorriamo ai seguenti aggregati:

- il reddito primario, che rappresenta la capacità delle famiglie di produrre reddito con l'impiego del proprio lavoro e del proprio patrimonio; si compone di reddito da lavoro dipendente, reddito netto derivante dalla proprietà di abitazioni e capitali e reddito risultante dall'attività imprenditoriale svolta dalle famiglie;
- reddito disponibile, che comprende il reddito primario e ad esso aggiunge prestazioni sociali ed altri trasferimenti privati e pubblici, togliendo imposte correnti e contributi sociali.

Per valutare il grado di benessere economico sarebbe necessario quantificare anche il patrimonio familiare, verificando l'eventuale proprietà dell'abitazione (ed alcune caratteristiche salienti come la superficie in rapporto al numero di residenti, la presenza di due o più bagni etc.) e l'ammontare dei depositi in banca. Purtroppo tali dati non sono però fruibili; si dispone invece di dati relativi al numero di automobili possedute per 1.000 residenti, che costituisce un indicatore indiretto di benessere economico.

Sarebbe auspicabile disporre di dati relativi alla quota di persone a rischio di povertà e alla concentrazione nella distribuzione del reddito, come ad esempio l'indice di Gini.

Nel database Eurostat, non avendo a disposizione indicatori relativi alla conciliazione tra lavoro e tempo libero, si estrae la variabile "numero di ore lavorate a settimana", ricavando pertanto per differenza il numero di ore settimanali dedicate al tempo libero. In questo contesto con l'espressione "tempo libero" s'intende tutto il tempo non finalizzato al lavoro retribuito, di cui fanno parte anche tempi destinati a soddisfare bisogni primari (dormire, mangiare, ecc.), tempi di spostamento, di cura (dai lavori domestici all'assistenza ai bambini).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
prodotto interno lordo pc (%media)	266	12	323	96,70	46,36	208,58
pil ppt pc (% media)	266	27	332	96,39	36,29	265,61
SMEAN(reddito_primario_netto)	266	24,10	216,21	99,5577	30,54	325,96
SMEAN(reddito_disponibile_netto)	266	33,40	180,00	100,0552	26,20	381,85
SMEAN(vehicle1000)	266	23	6480	1029,56	875,23	117,63
%ore non lavorate	266	73,0	82,6	77,601	1,63	4762,32
Validi (listwise)	266					

Correlazioni

		prodotto interno lordo pc (%media)	pil ppt pc (% media)	SMEAN(reddito_ primario_netto)	SMEAN(reddito_dis ponibile_netto)	SMEAN(vehicle 1000)	%ore non lavorate
prodotto interno lordo pc (%media)	Correlazione di Sig. (2-code) N	1 266	,948** 266	,798** 266	,719** 266	,114 266	,530** 266
pil ppt pc (% media)	Correlazione di Sig. (2-code) N	,948** 266	1 266	,766** 266	,683** 266	,121 266	,391** 266
SMEAN(reddito_primario_netto)	Correlazione di Sig. (2-code) N	,798** 266	,766** 266	1 266	,954** 266	,165** 266	,505** 266
SMEAN(reddito_disponibile_netto)	Correlazione di Sig. (2-code) N	,719** 266	,683** 266	,954** 266	1 266	,178** 266	,419** 266
SMEAN(vehicle1000)	Correlazione di Sig. (2-code) N	,114 266	,121 266	,165** 266	,178** 266	1 266	,062 266
%ore non lavorate	Correlazione di Sig. (2-code) N	,530** 266	,391** 266	,505** 266	,419** 266	,062 266	1 266

** . La correlazione è significativa al livello 0,01 (2-code).* . La correlazione è significativa al livello 0,05 (2-code).

Per evitare variabili ridondanti nelle analisi successive si escludono il “Prodotto interno lordo” ed il “Reddito primario netto”.

L’indice di correlazione tra la variabile Pil a parità di potere d’acquisto ed il reddito disponibile netto è pari a 0,68 e pertanto la relazione tra i due fenomeni è diretta, ma non eccessivamente elevata.

2.3.6 VARIABILI AMBIENTALI

<i>ETICHETTA</i>	<i>INDICATORE AMBIENTALE</i>
agricoltura	% suolo utilizzato per l'agricoltura
foreste	% suolo coperto da foreste e arbusteti
caccia_pesca	% suolo adibito ad attività di caccia e pesca
alto impatto_ambientale	% suolo su cui vengono svolte attività ad alto impatto ambientale
servizi_residenze	% zone residenziali, commerciali ed industriali
rifiuti_proc_kg	rifiuti urbani prodotti per persona in kg
dens	Densità abitativa per km

Ai fini della misurazione del benessere ambientale si evidenzia una situazione di grave carenza informativa: mancano indici di biodiversità, di ricchezza del paesaggio, della presenza di risorse di acqua fresca, dell'assenza di inquinamento dell'aria, del suolo e delle acque che siano riferiti al territorio regionale.

Interessante sarebbe pure disporre dell' "impronta ecologica", in grado di valutare l'impatto ecologico e la sostenibilità ambientale: esso si presenta come un indicatore aggregato e sintetico relativo allo stato di pressione umana sui sistemi naturali concettualmente abbastanza semplice e ad elevato contenuto comunicativo. L'impronta ecologica misura, infatti, il consumo alimentare, materiale ed energetico della popolazione umana sulla superficie terrestre o marina necessaria per produrre le risorse naturali o, nel caso dell'energia, sulla superficie terrestre necessaria ad assorbire le emissioni di anidride carbonica.¹⁰

¹⁰ L'impronta ecologica di una persona è data dalla somma di 6 differenti componenti:

- la superficie di terra coltivata necessaria per produrre gli alimenti
- l'area di pascolo necessaria per produrre i prodotti animali
- la superficie di foresta necessaria per produrre legname e carta
- la superficie marina necessaria per produrre pesci e frutti di mare

I principali dati ambientali a disposizione provengono dall'indagine svolta nel 2009 denominata LUCAS, acronimo di *Land Use and Cover Area frame Survey*. La classificazione utilizzata è caratterizzata da una netta separazione tra uso del suolo e copertura del suolo.

Per “*Land Cover*” si intende la copertura fisica della superficie terrestre, mentre il termine “*Land Use*” descrive le stesse aree in base alla funzione socio-economica del territorio.

Dati gli scopi della ricerca, si preferisce analizzare la dimensione funzionale del territorio, la cui struttura gerarchica si basa su 15 categorie, a loro volta raggruppate da Eurostat in 5 ambiti:

Agricoltura

Foreste

Caccia e pesca

Aree residenziali, per commercio e servizi (commercio, finanza, affari, tempo libero, sport)

Aree con attività ad alto impatto ambientale (produzione energia, industrie, reti di trasporto e comunicazione, costruzioni, acque e trattamento acque reflue).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	CV
dens	266	3,0	9622,0	382,76	953,13	249,01
SM(agricoltura)	266	,00	83,52	48,53	16,71	34,43
SM(foreste)	266	,00	67,68	22,63	15,06	66,56
SM(caccia_pesca)	266	,00	19,00	1,38	2,37	171,71
SM(alti_impatto_amb)	266	1,12	33,82	5,09	4,02	79,05
SM(servizi_resid)	266	2,12	90,00	13,43	10,85	80,79
rifiuti procapite kg	266	,00	1076,60	378,80	261,75	69,10
Validi (listwise)	266					

-
- la superficie di terra necessaria per ospitare infrastrutture edilizie
 - la superficie forestale necessaria per assorbire le emissioni di anidride carbonica risultanti dal consumo energetico dell'individuo considerato.

La densità abitativa, calcolata come rapporto tra la popolazione compresa in un'area e l'area stessa, è espressa in chilometri quadrati. L'addensamento demografico attorno a poli gravitazionali dipende essenzialmente da motivazioni di ordine socio-economico, come la disponibilità di abitazioni, le opportunità di lavoro e la vicinanza a determinati servizi (Egidi, Giovannini 1998). A dimostrazione di ciò la relazione tra densità e aree residenziali e commerciali è diretta e abbastanza forte ($r=0,728$).

Correlazioni

		dens	SM (agricoltura)	SM(foreste)	SM(caccia_pesca)	SM (servizi_resid)	SM (alti_impatto_amb)	rifiuti procapite kg
dens	Correlazione di Pearson	1	-,330**	-,206**	-,092	,728**	,468**	,028
	Sig. (2-code)		,000	,001	,134	,000	,000	,653
	N	266	266	266	266	266	266	266
SM(agricoltura)	Correlazione di Pearson	-,330**	1	-,576**	-,059	-,329**	-,198**	,000
	Sig. (2-code)	,000		,000	,338	,000	,001	,997
	N	266	266	266	266	266	266	266
SM(foreste)	Correlazione di Pearson	-,206**	-,576**	1	-,093	-,412**	-,336**	,029
	Sig. (2-code)	,001	,000		,129	,000	,000	,638
	N	266	266	266	266	266	266	266
SM(caccia_pesca)	Correlazione di Pearson	-,092	-,059	-,093	1	-,097	-,108	,031
	Sig. (2-code)	,134	,338	,129		,113	,080	,618
	N	266	266	266	266	266	266	266
SM(servizi_resid)	Correlazione di Pearson	,728**	-,329**	-,412**	-,097	1	,545**	-,091
	Sig. (2-code)	,000	,000	,000	,113		,000	,140
	N	266	266	266	266	266	266	266
SM(alti_impatto_amb)	Correlazione di Pearson	,468**	-,198**	-,336**	-,108	,545**	1	,158**
	Sig. (2-code)	,000	,001	,000	,080	,000		,010
	N	266	266	266	266	266	266	266
rifiuti procapite kg	Correlazione di Pearson	,028	,000	,029	,031	-,091	,158**	1
	Sig. (2-code)	,653	,997	,638	,618	,140	,010	
	N	266	266	266	266	266	266	266

** La correlazione è significativa al livello 0,01 (2-code).

Per evitare multicollinearità, si elimina variabile “Caccia e pesca” (motivo: valori coefficienti più bassi).

CAPITOLO 3 - ANALISI DEGLI INDICATORI DI BENESSERE TRAMITE LA CARTOGRAFIA

“In all situation we need to have words, numbers and graphic working together” (Tuftte, 2001)

Per effettuare le rappresentazioni cartografiche si utilizza il programma GvSig (si tratta della versione gratuita disponibile on-line di GIS, *Geographic Information System*).

Dopo aver scaricato gli *Shapefile* (formato vettoriale di registrazione di identità geometriche e delle informazioni associate) dal sito di Eurostat, si provvede ad unire la tabella degli attributi contenente i codici delle regioni con quella in cui si riportano i dati relativi agli indicatori di benessere descritti in precedenza.

Con lo scopo di fornire un'immediata comprensione di ciascun fenomeno tramite la mera osservazione della cartina, si è provveduto a trasformare i dati quantitativi in ordinali utilizzando i quintili. In tal modo si risolve anche il problema dell'eventuale presenza di valori anomali, in quanto trattasi di misure robuste. Pertanto si dividono le 267 regioni in 5 gruppi di uguale numerosità ed il rango uno corrisponde al valore più piccolo, mentre il rango cinque a quello più grande.

Figura 6. DEMOGRAFIA: INDICE DI DIPENDENZA

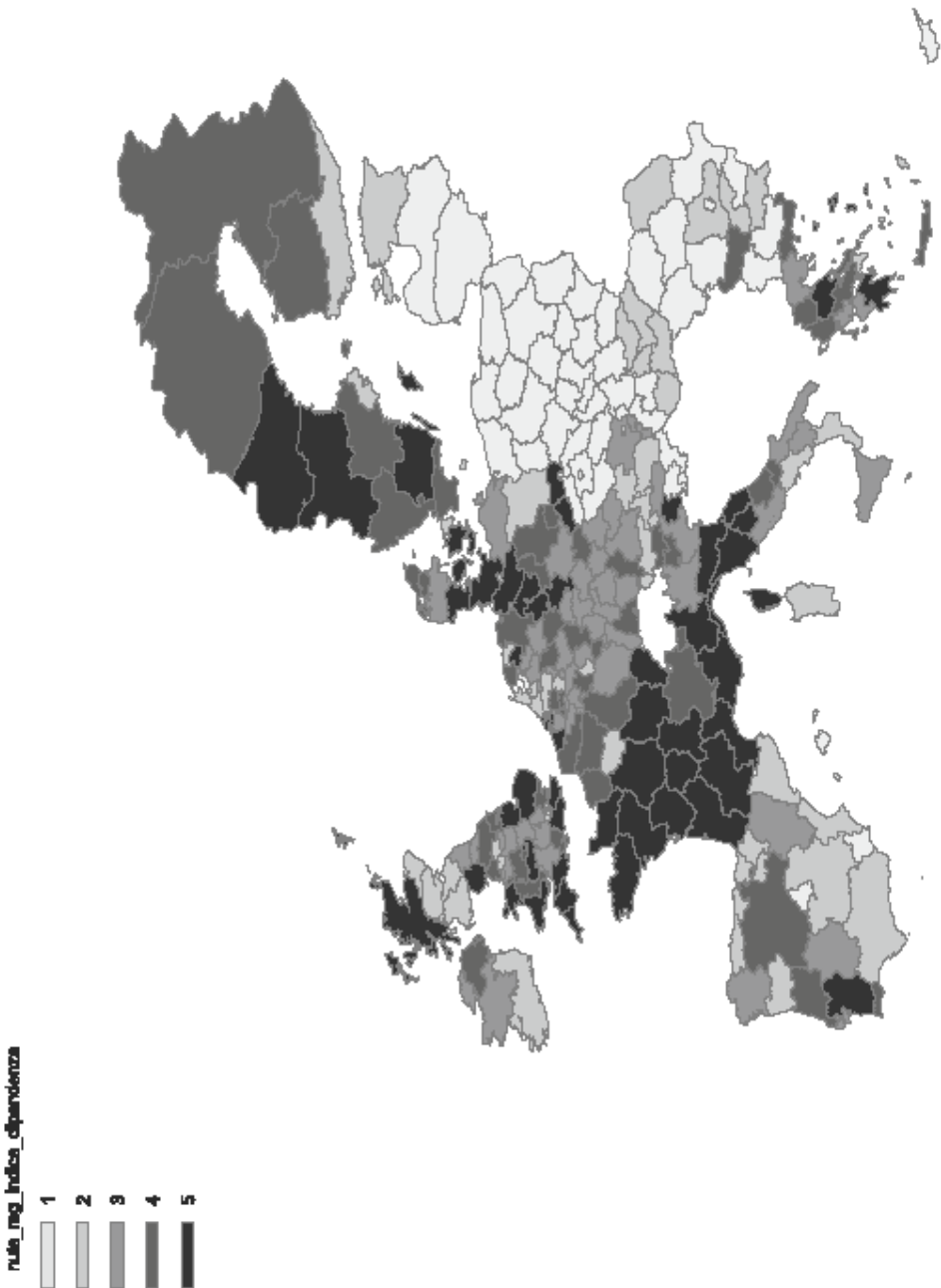


Figura 7. DEMOGRAFIA: INDICE DI VECCHIAIA

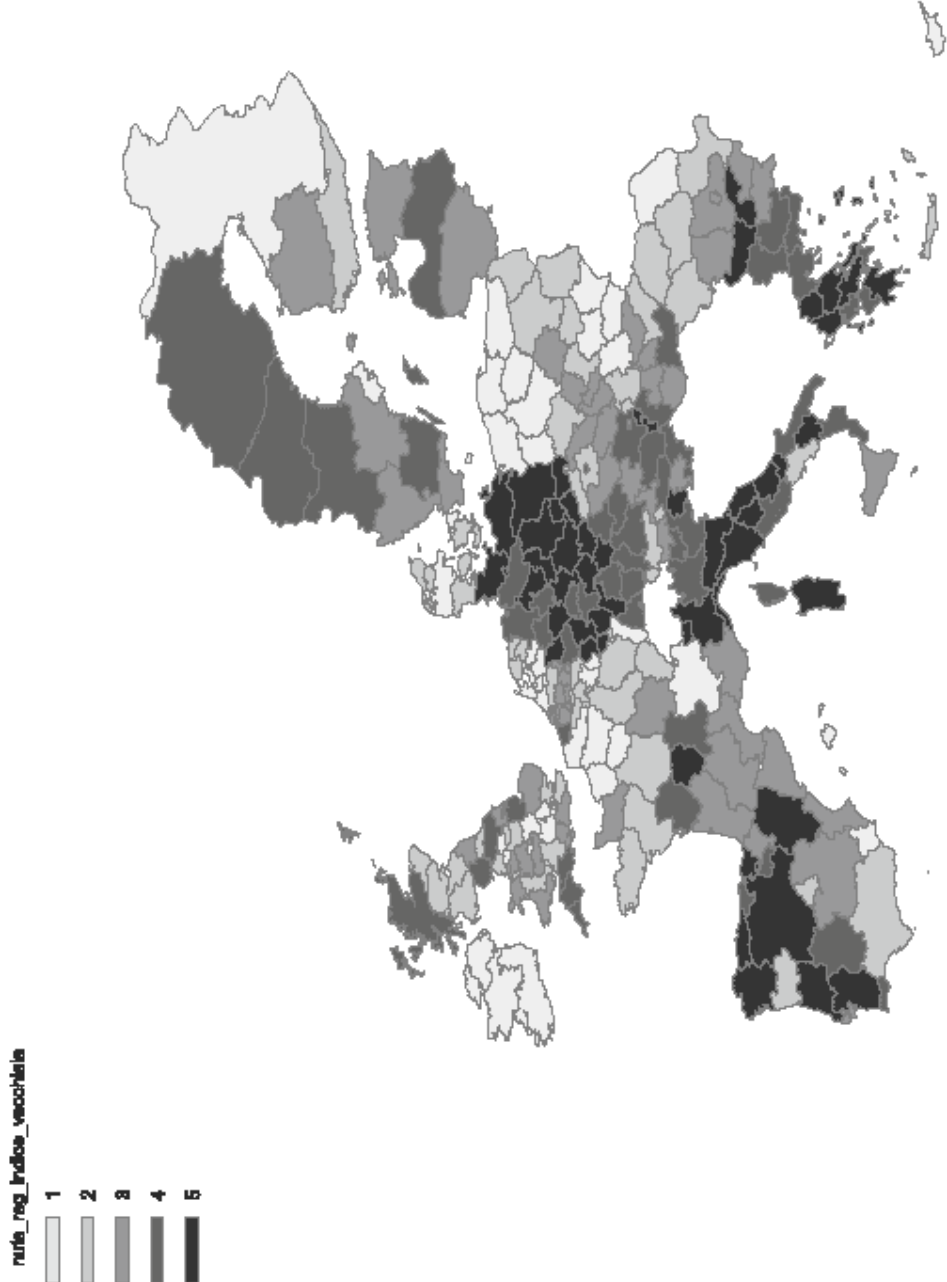


Figura 8. DEMOGRAFIA: PERCENTUALE POPOLAZIONE OVER 74

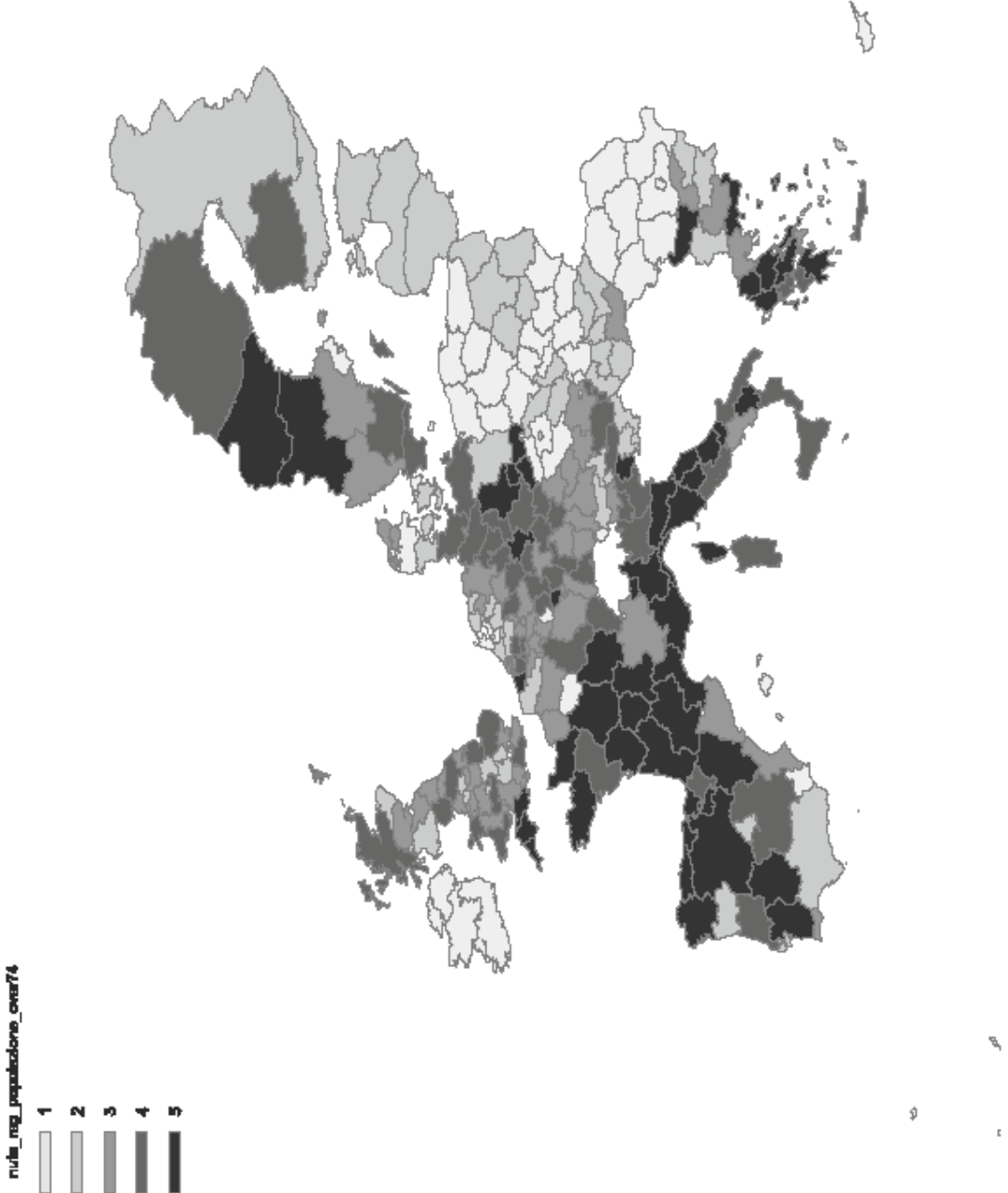


Figura 9. DEMOGRAFIA: PERCENTUALE POPOLAZIONE UNDER 10

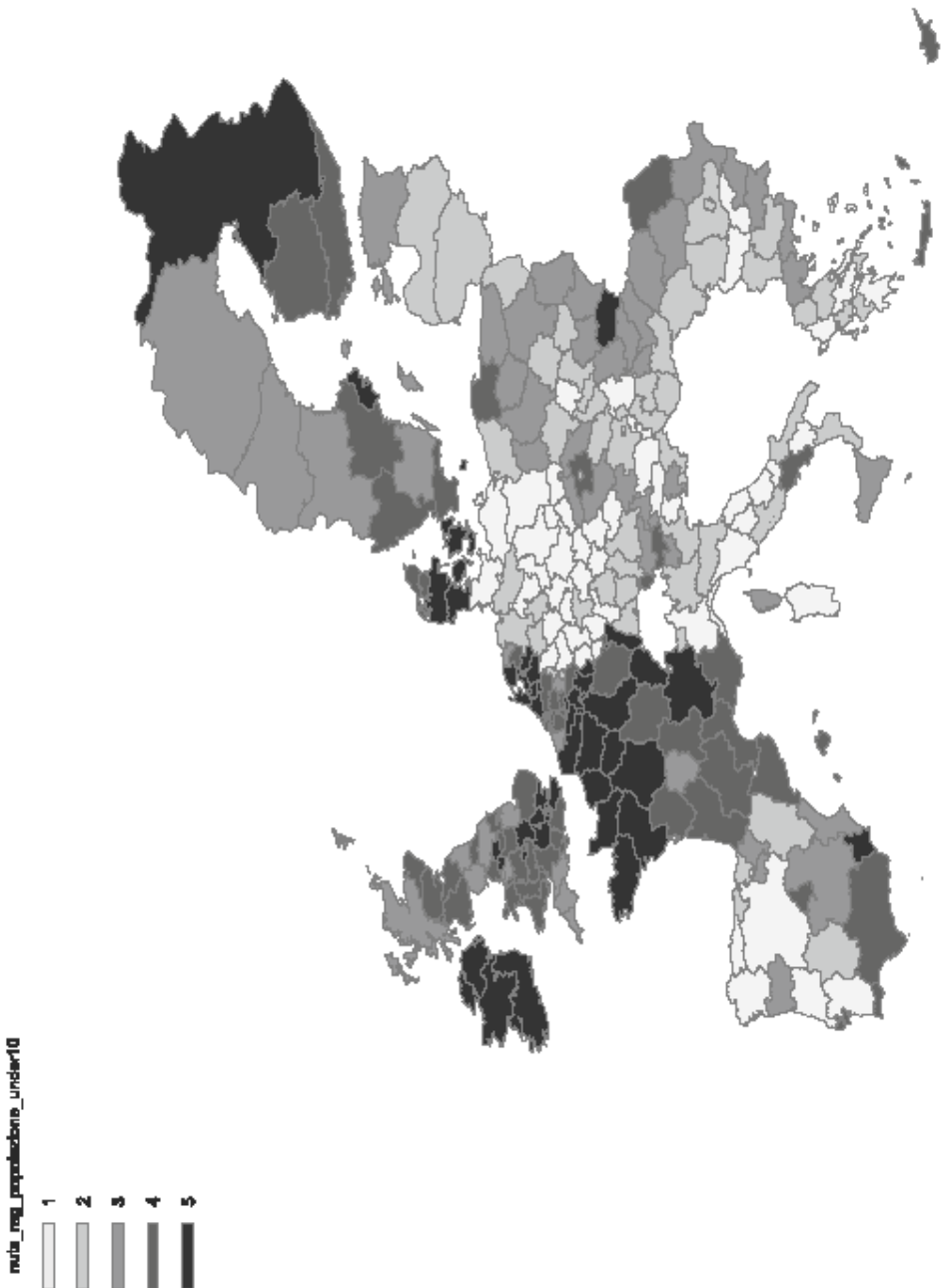


Figura 10. DEMOGRAFIA: TASSO DI FERTILITA'

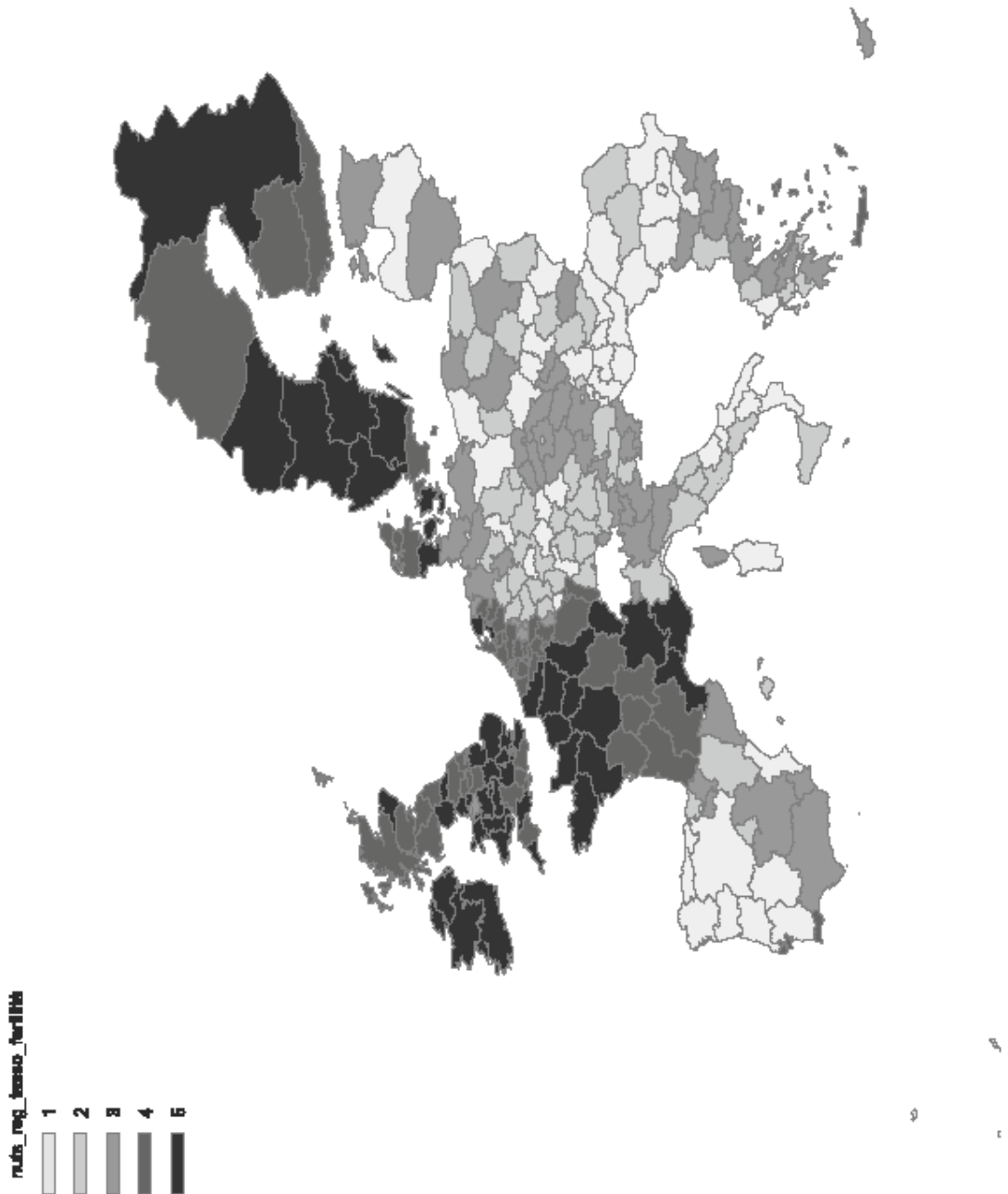


Figura 11. DEMOGRAFIA: TASSO NETTO MOVIMENTO MIGRATORIO

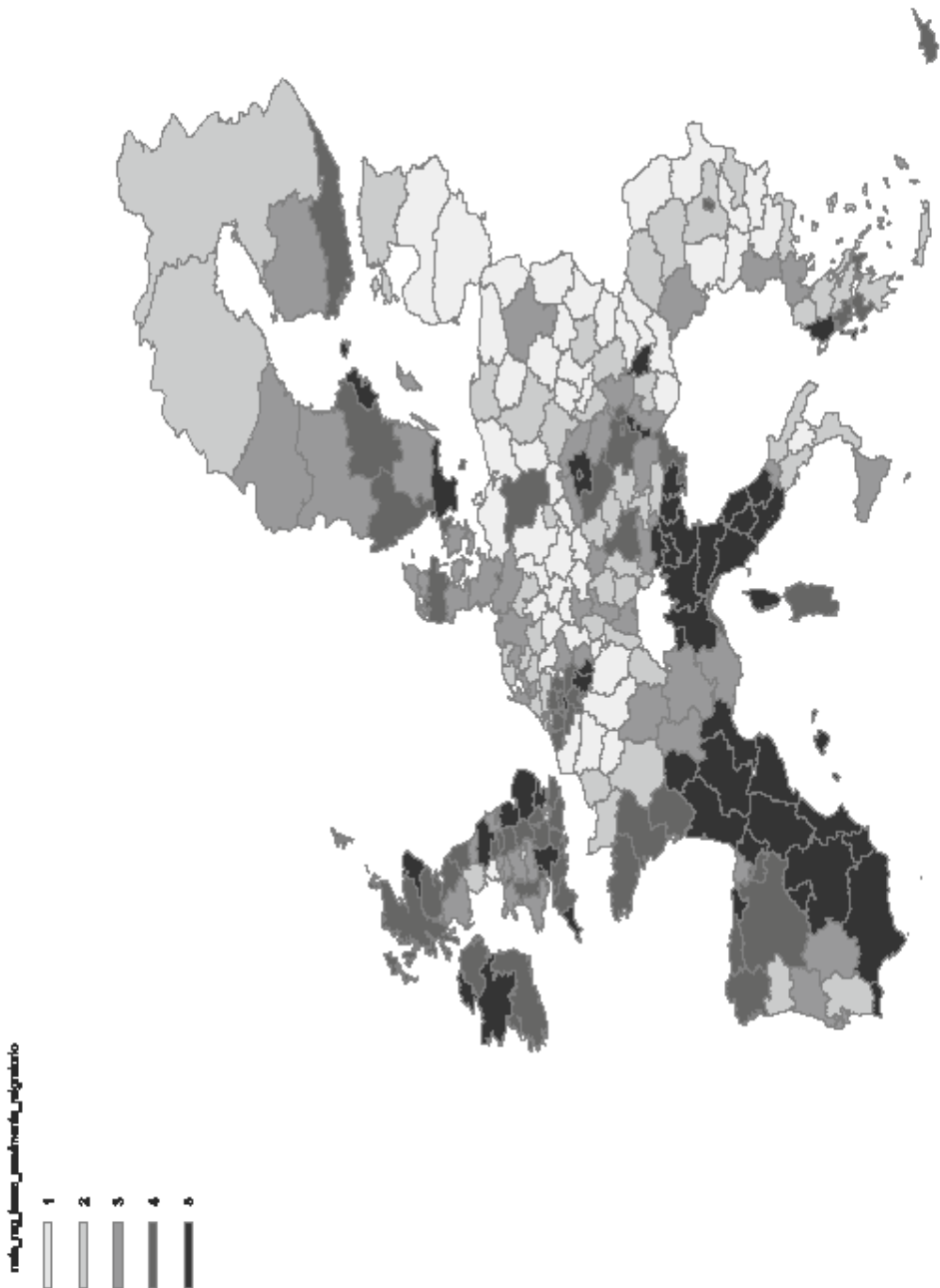


Figura 12. DEMOGRAFIA: TASSO NETTO MOVIMENTO NATURALE

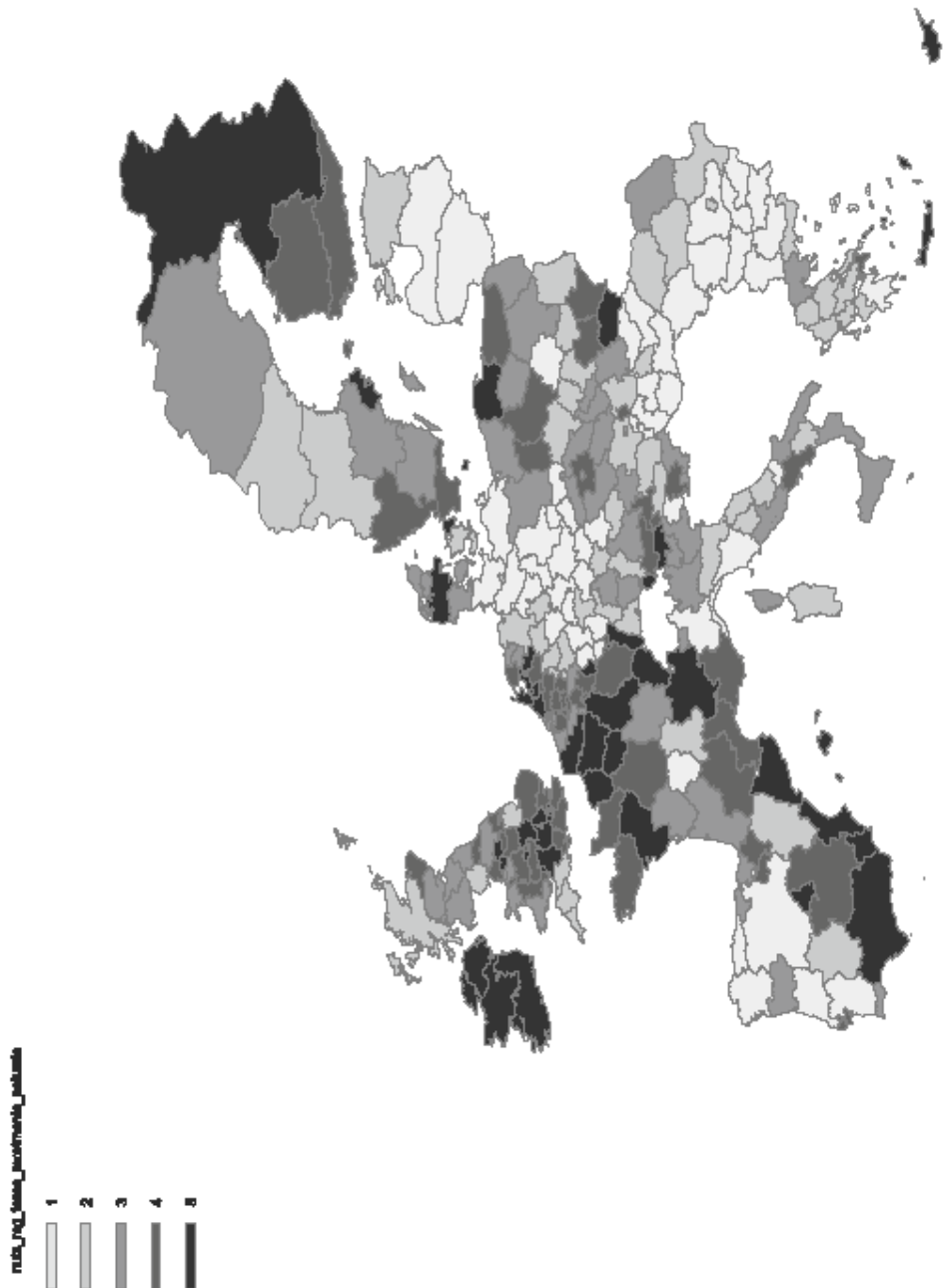


Figura 13. SALUTE: ASPETTATIVA DI VITA

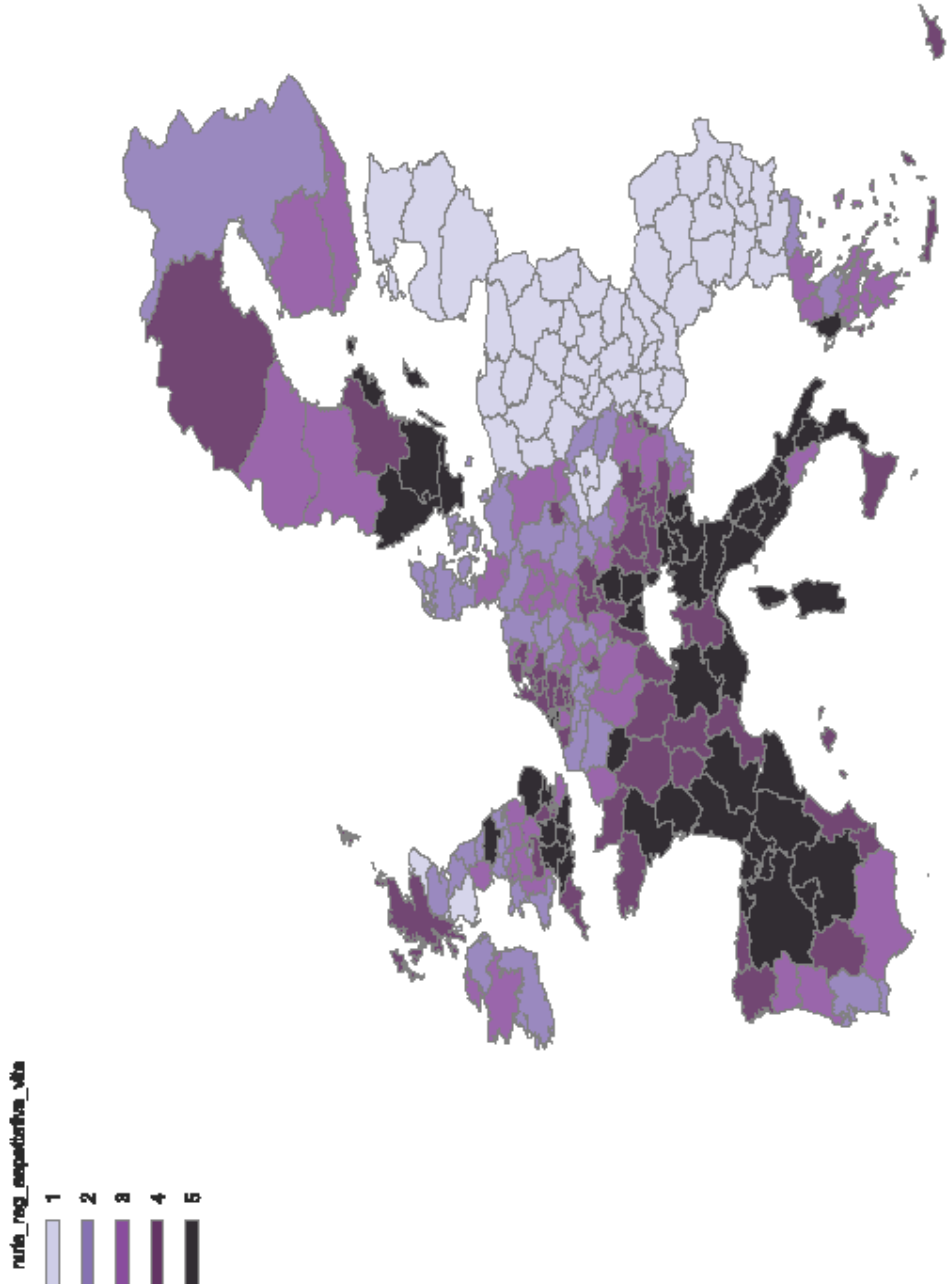


Figura 14. SALUTE: TASSO DI MORTALITA' INFANTILE

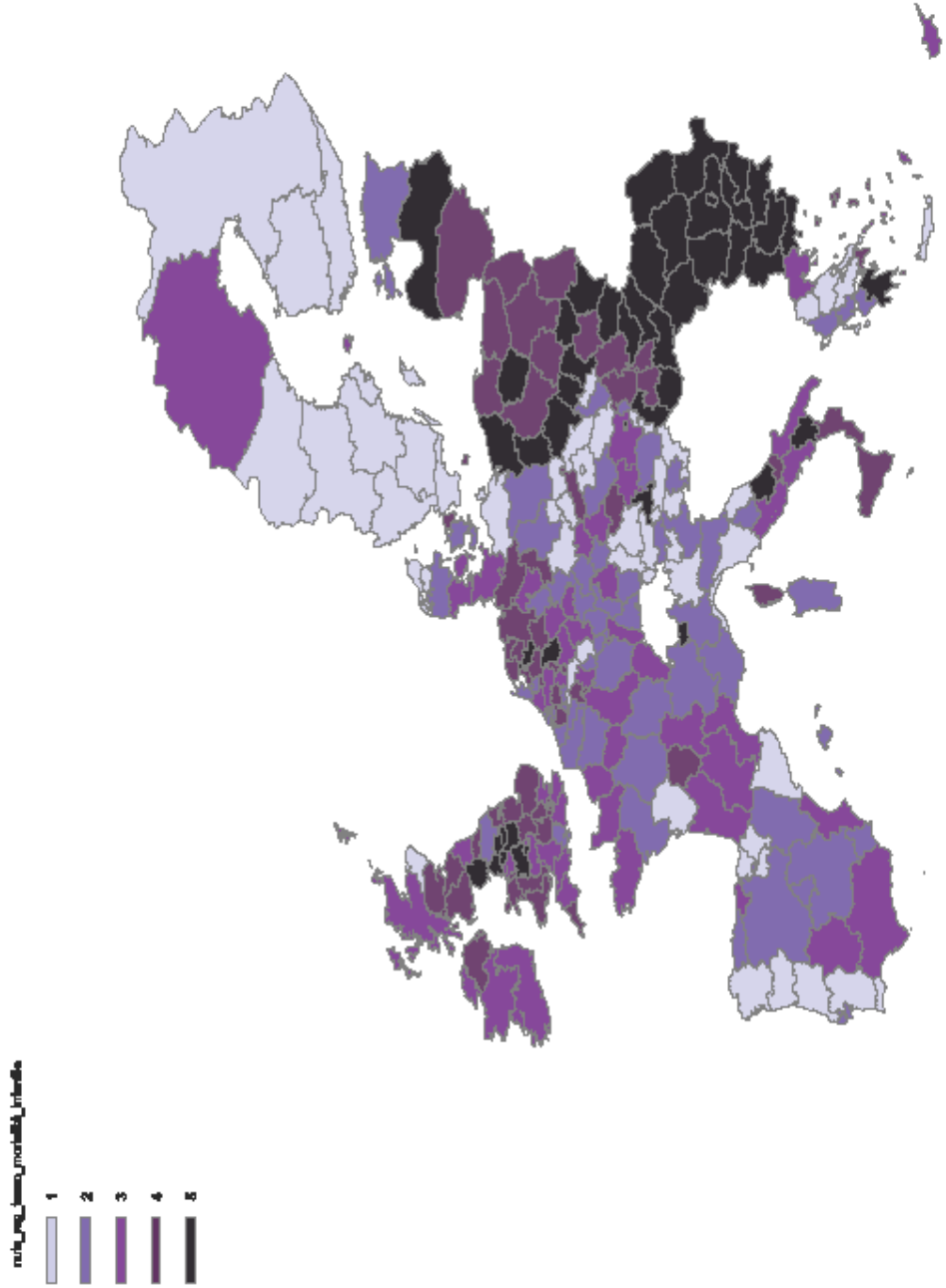


Figura 15. SALUTE: TASSO MORTALITA' INCIDENTI STRADALI

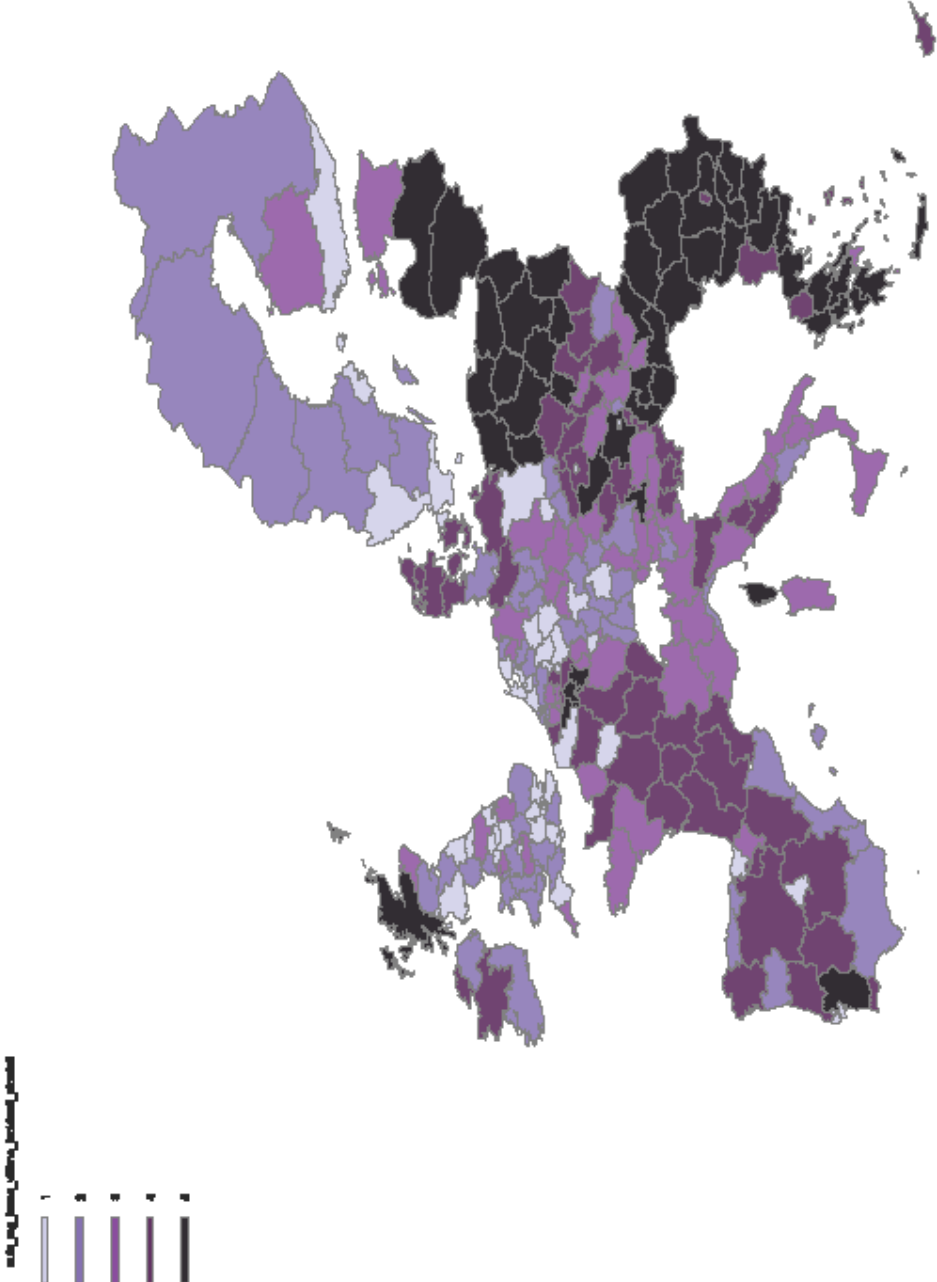


Figura 16. RICCHEZZA E TEMPO LIBERO: ORE DI NON LAVORO

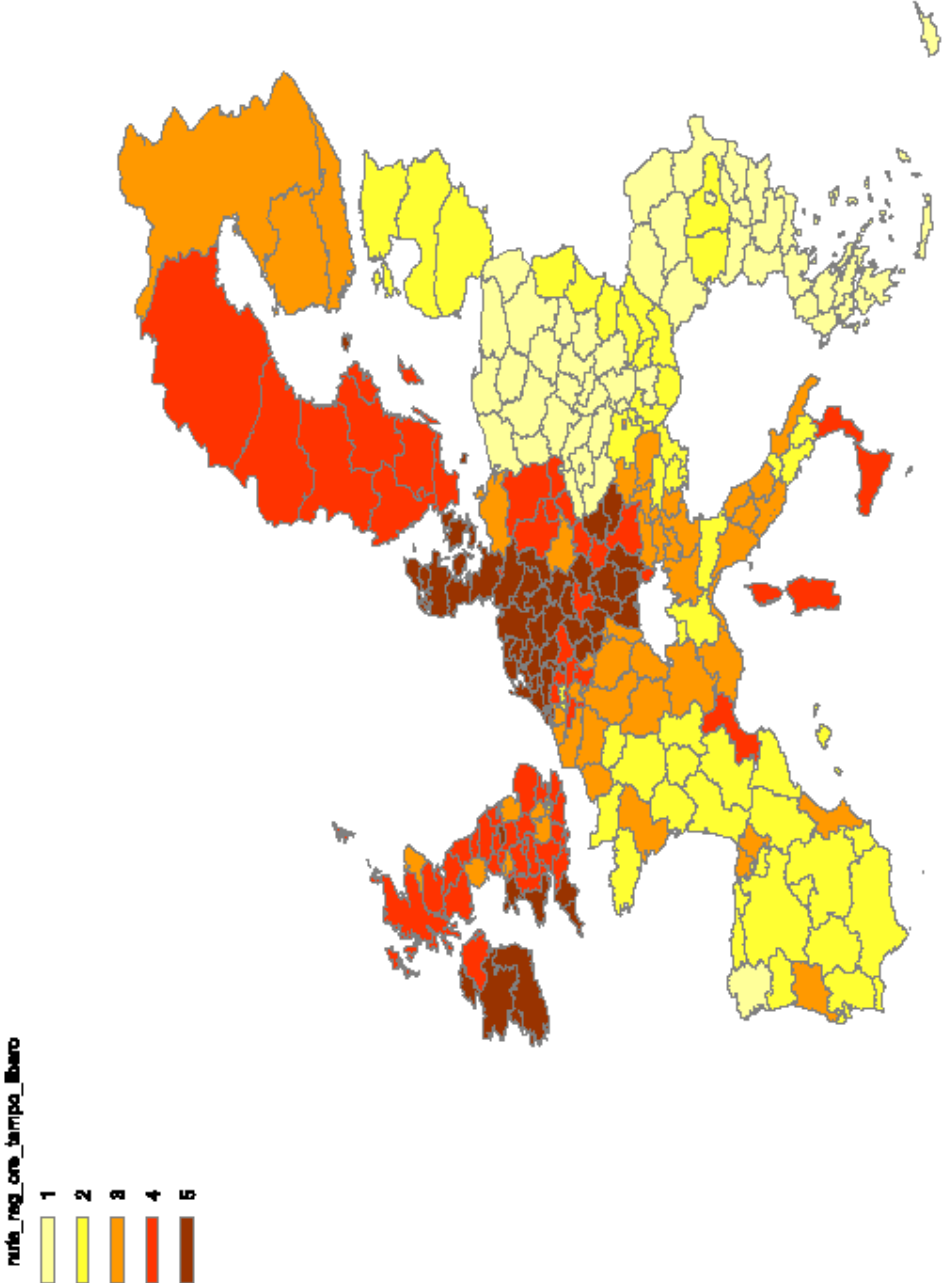


Figura 17. RICCHEZZA E TEMPO LIBERO: PIL A PARITA' DI POTERE D'ACQUISTO

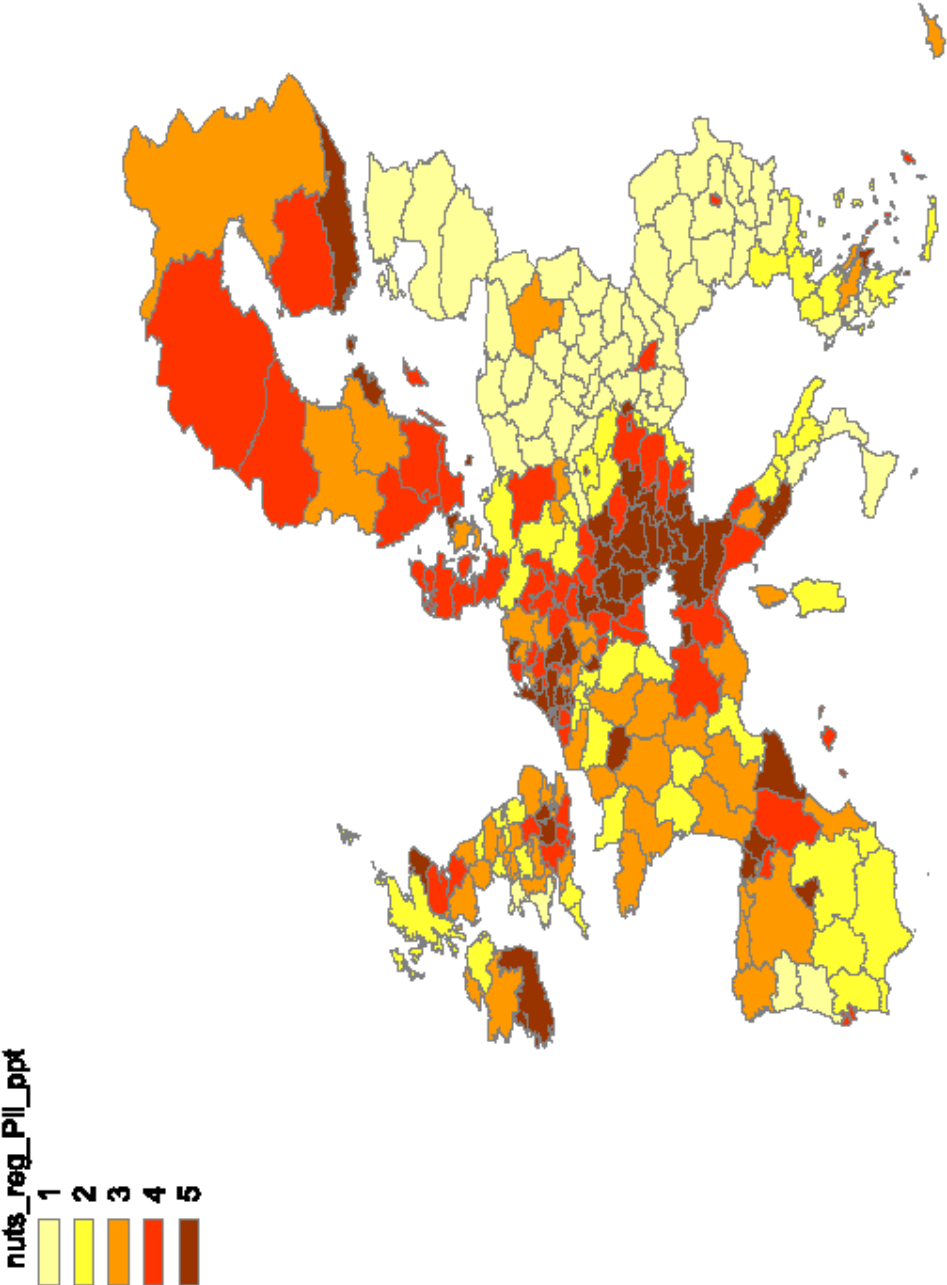


Figura 18. RICCHEZZA E TEMPO LIBERO: VEICOLI POSSEDUTI PER 1000 ABITANTI

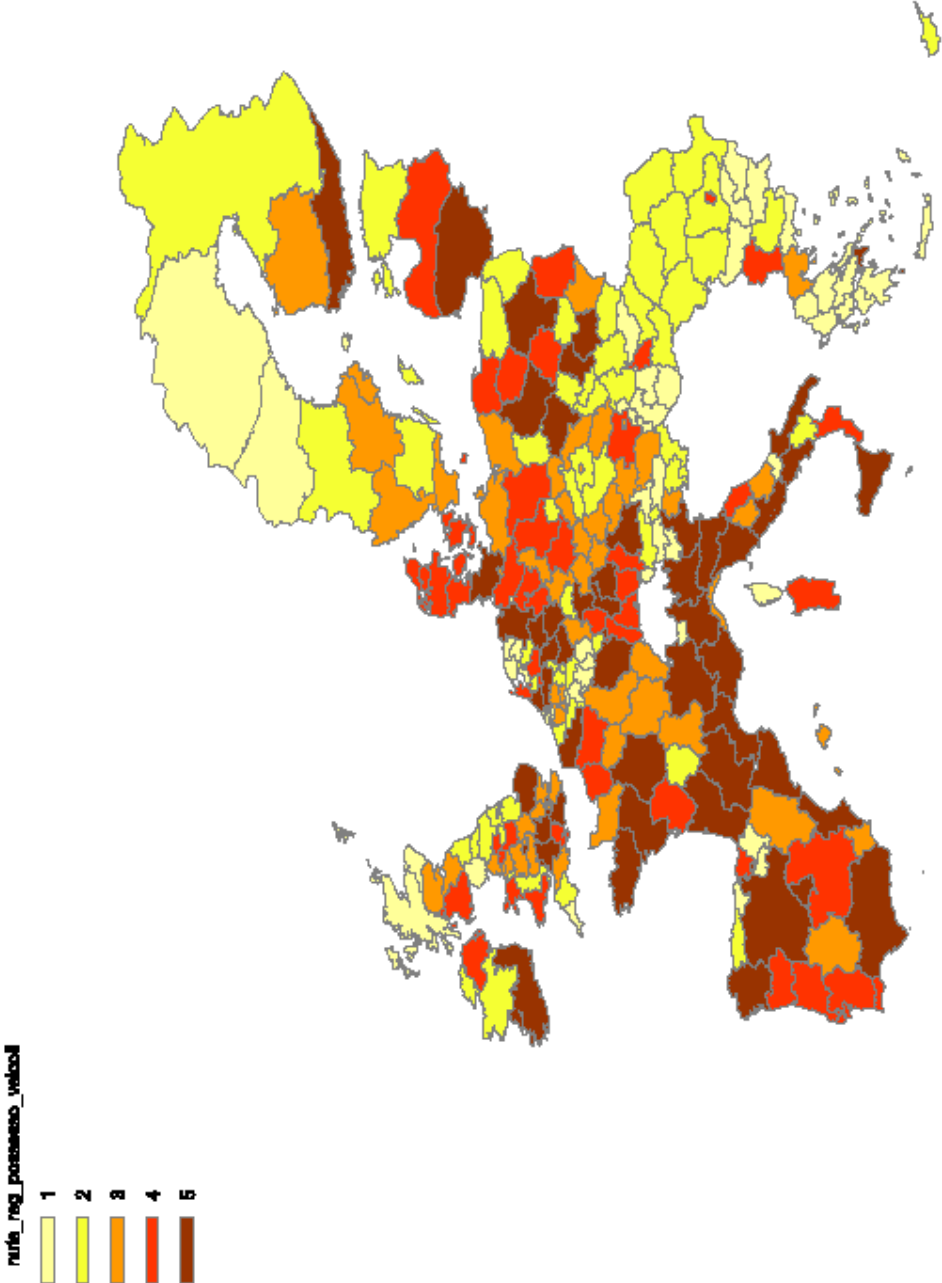


Figura 19. RICCHEZZA E TEMPO LIBERO: REDDITO FAMILIARE NETTO DISPONIBILE

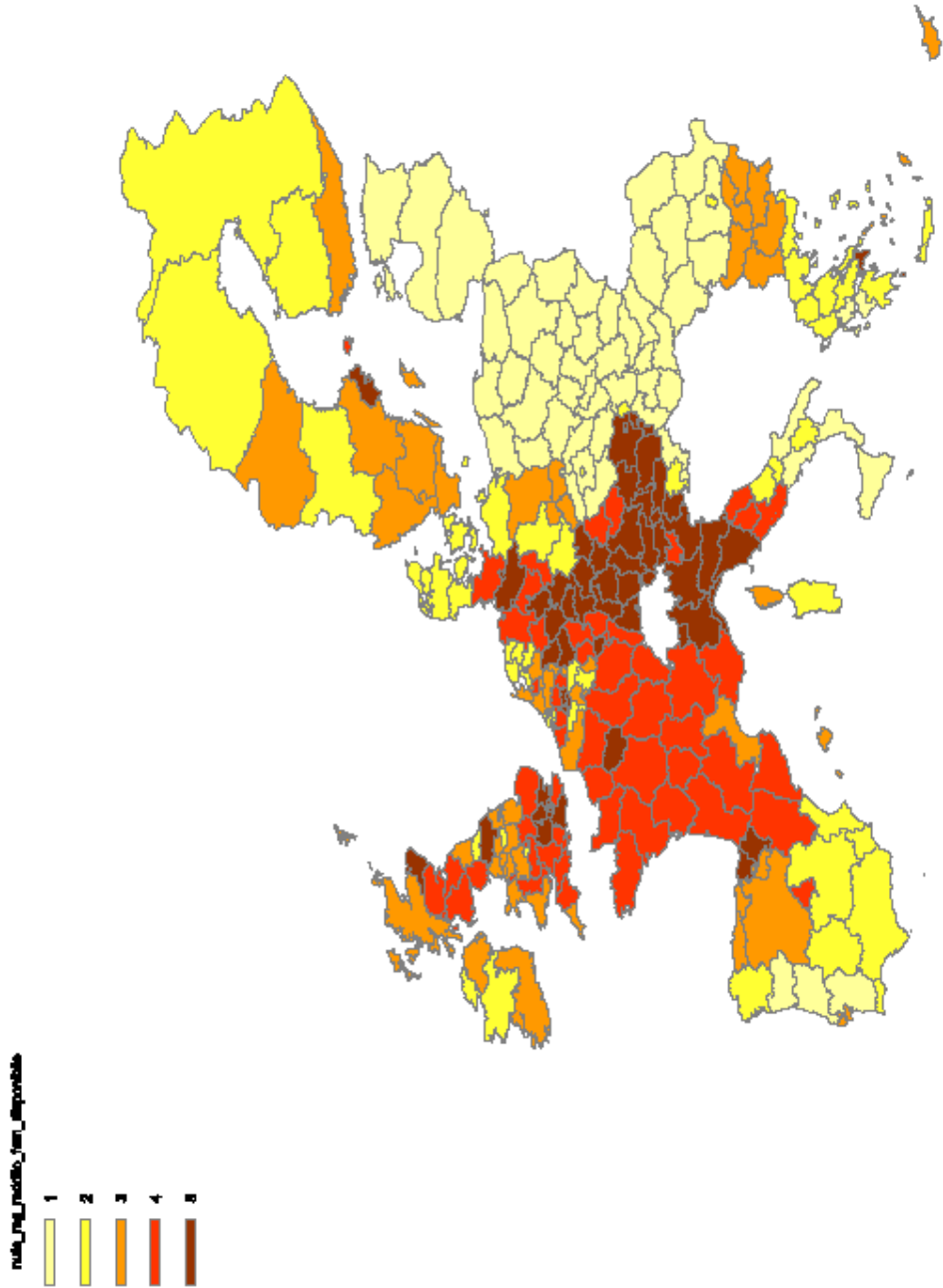


Figura 20. ISTRUZIONE: PERCENTUALE ISTRUITI DI MEDIO LIVELLO

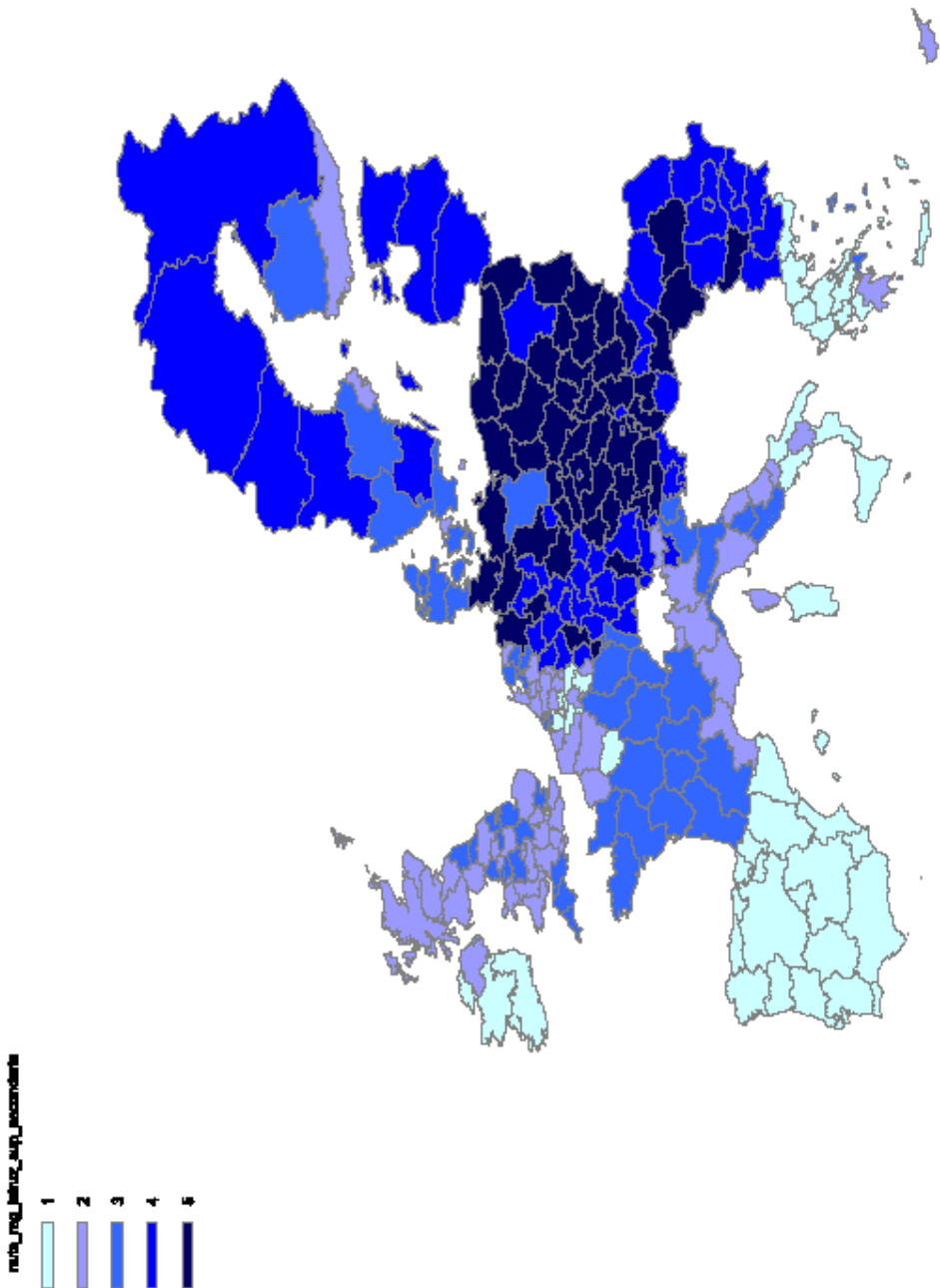


Figura 21. ISTRUZIONE: PERCENTUALE ISTRUITI DI ALTO LIVELLO

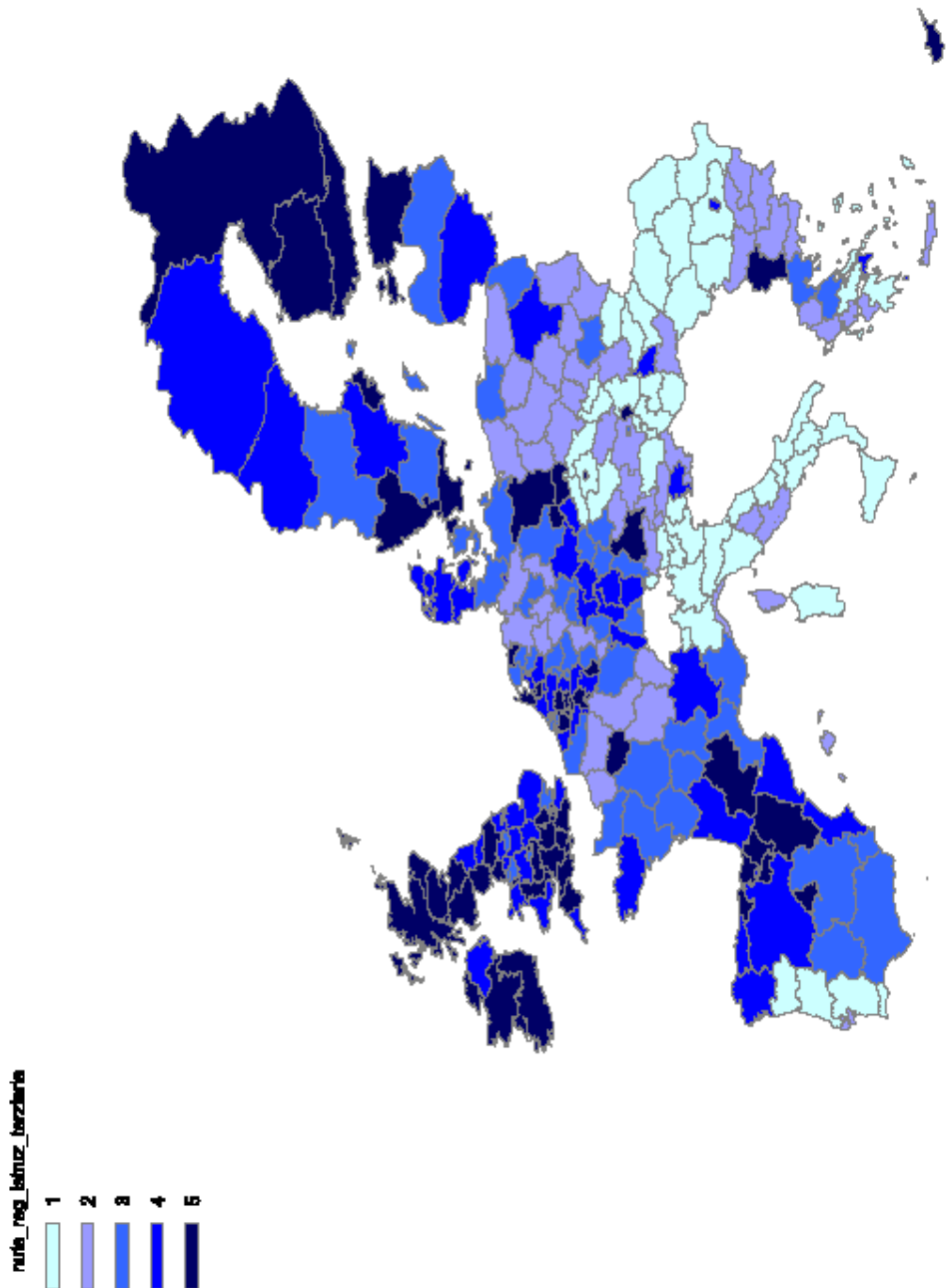


Figura 22. ISTRUZIONE: PERCENTUALE FORMAZIONE PERMANENTE ADULTI

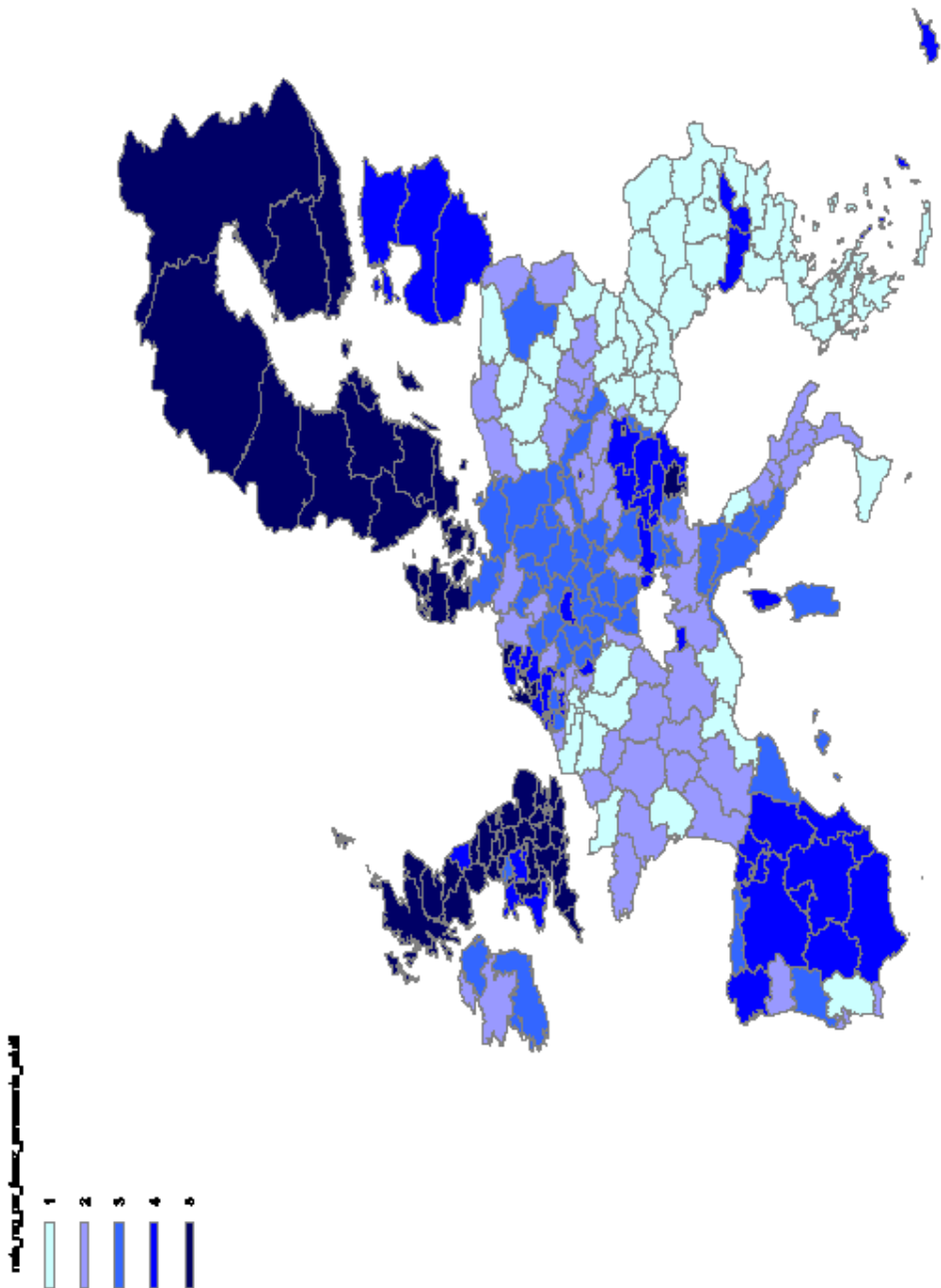


Figura 23. ISTRUZIONE: PERCENTUALE SPESA RICERCA E SVILUPPO

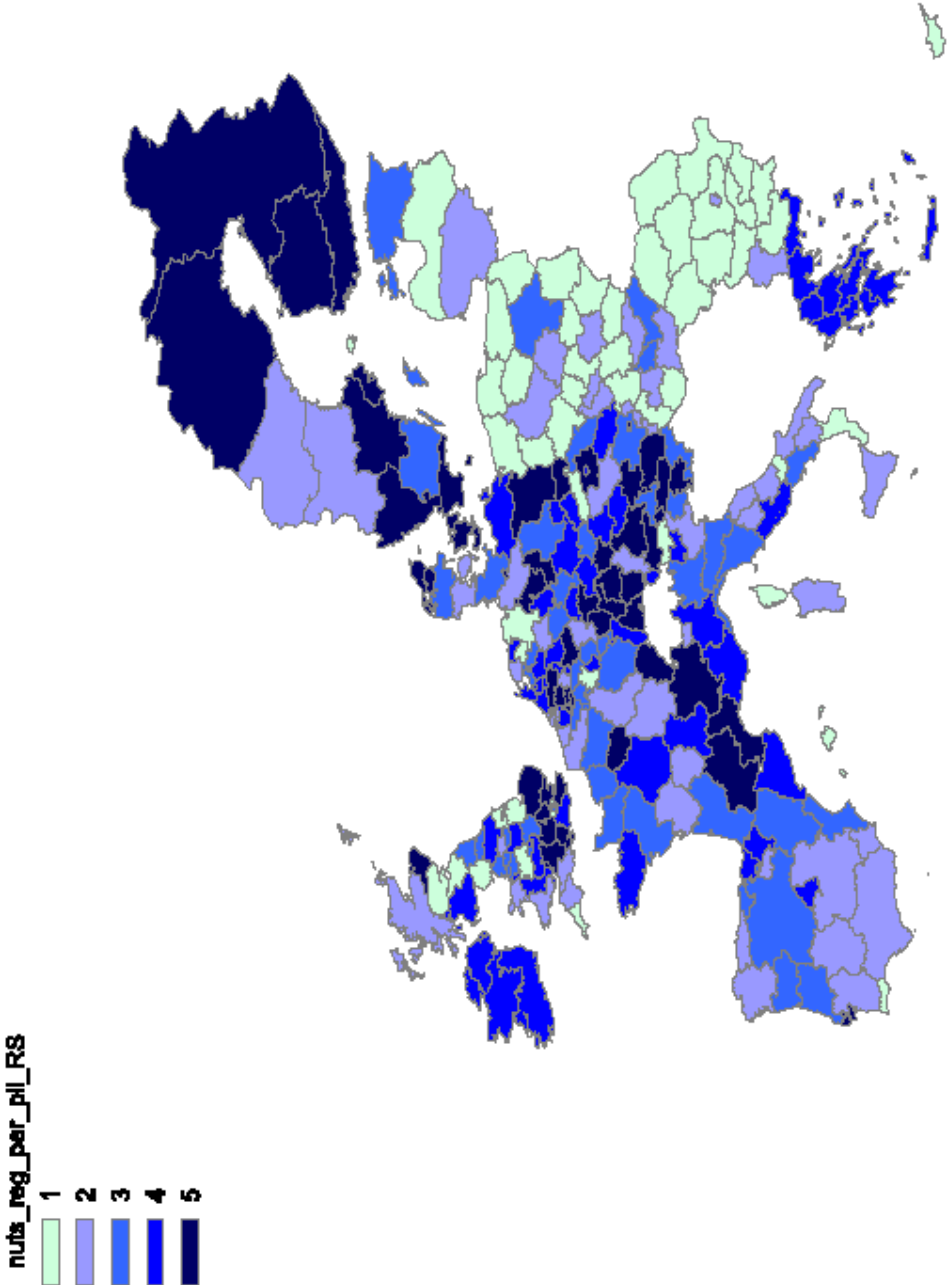


Figura 24. ISTRUZIONE: PERCENTUALE BREVETTI ACCETTATI

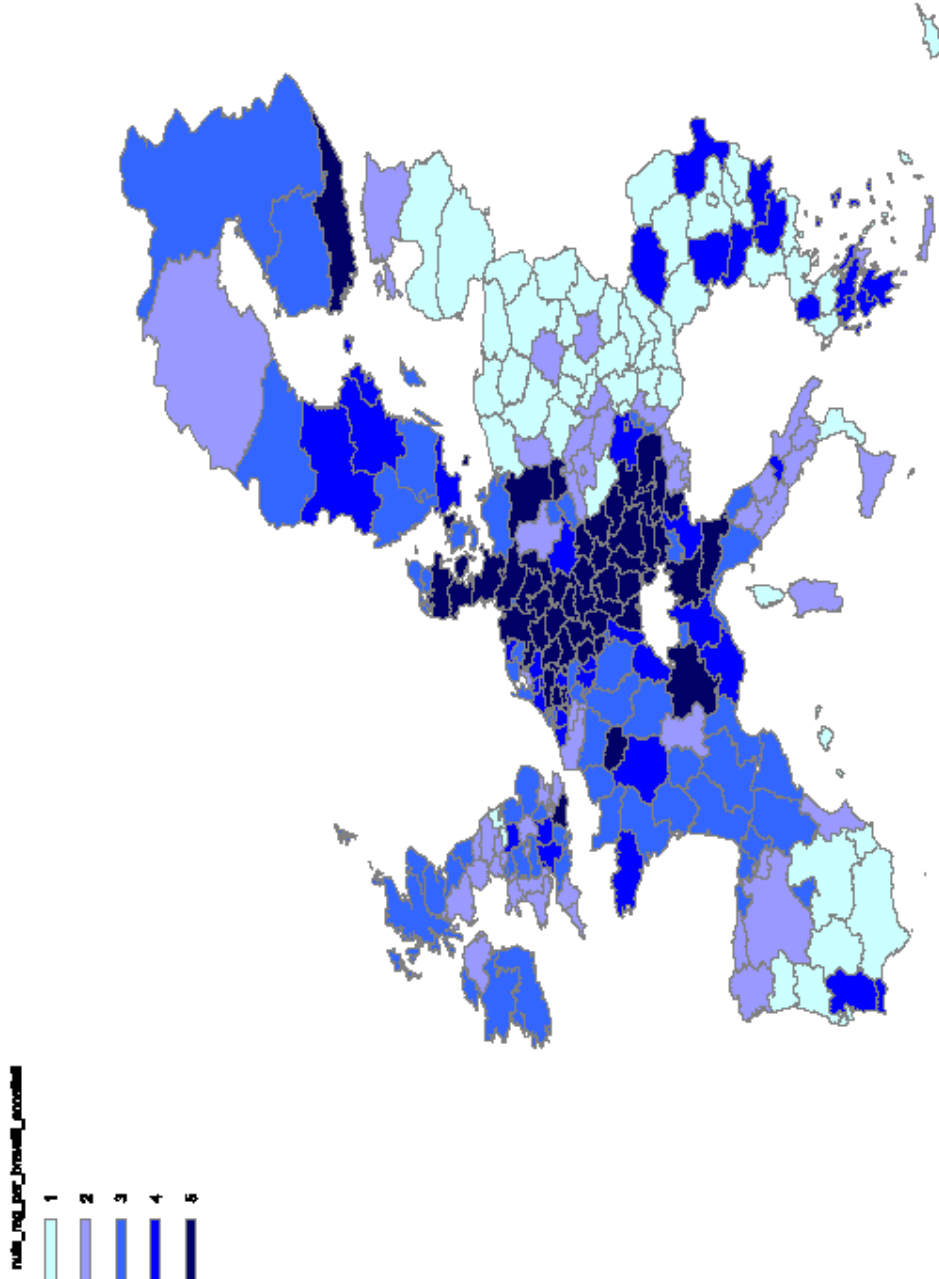


Figura 25. LAVORO: TASSO DI OCCUPAZIONE

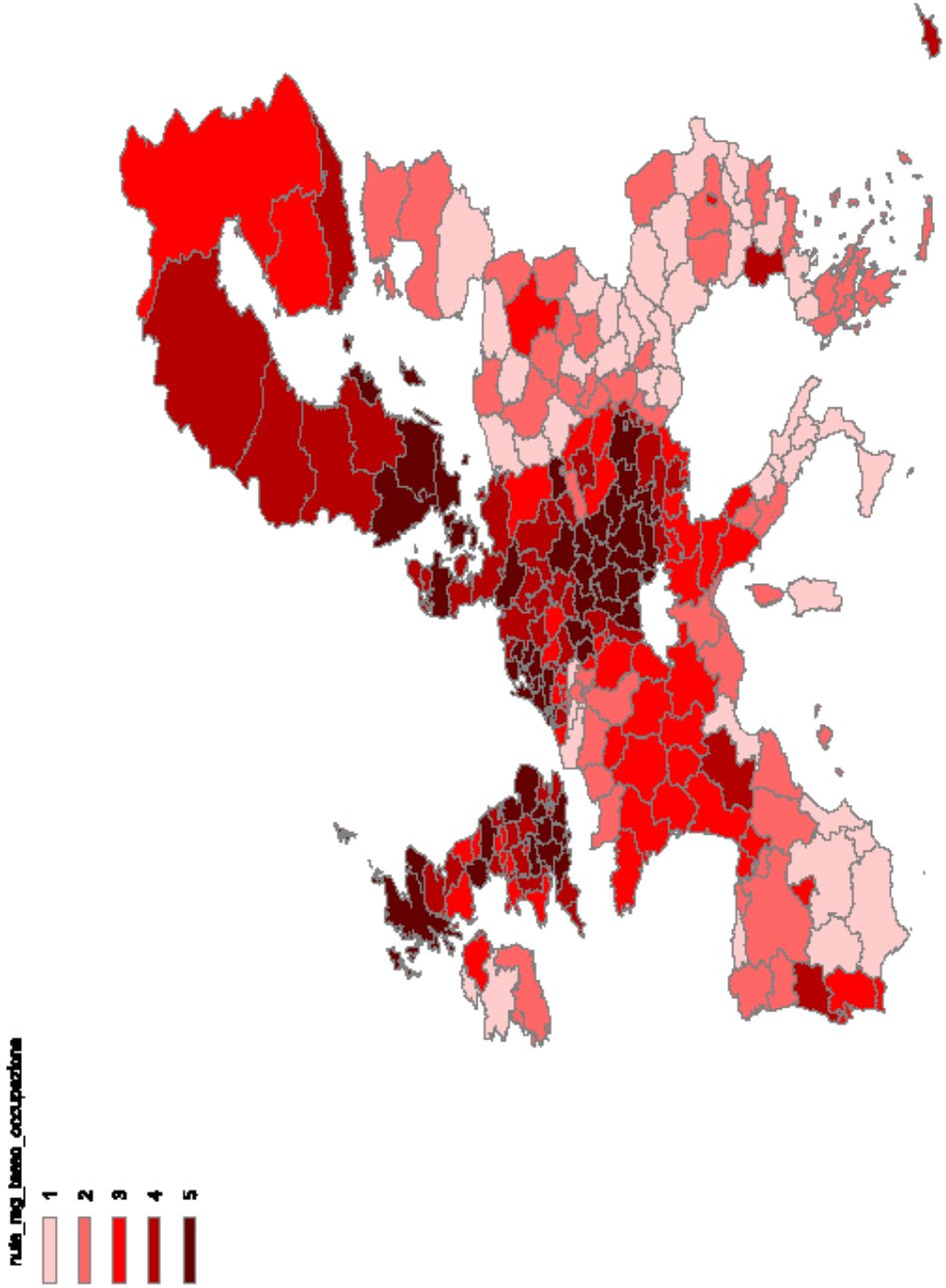


Figura 26. LAVORO: DIFFERENZA OCCUPAZIONE MASCHI-FEMMINE

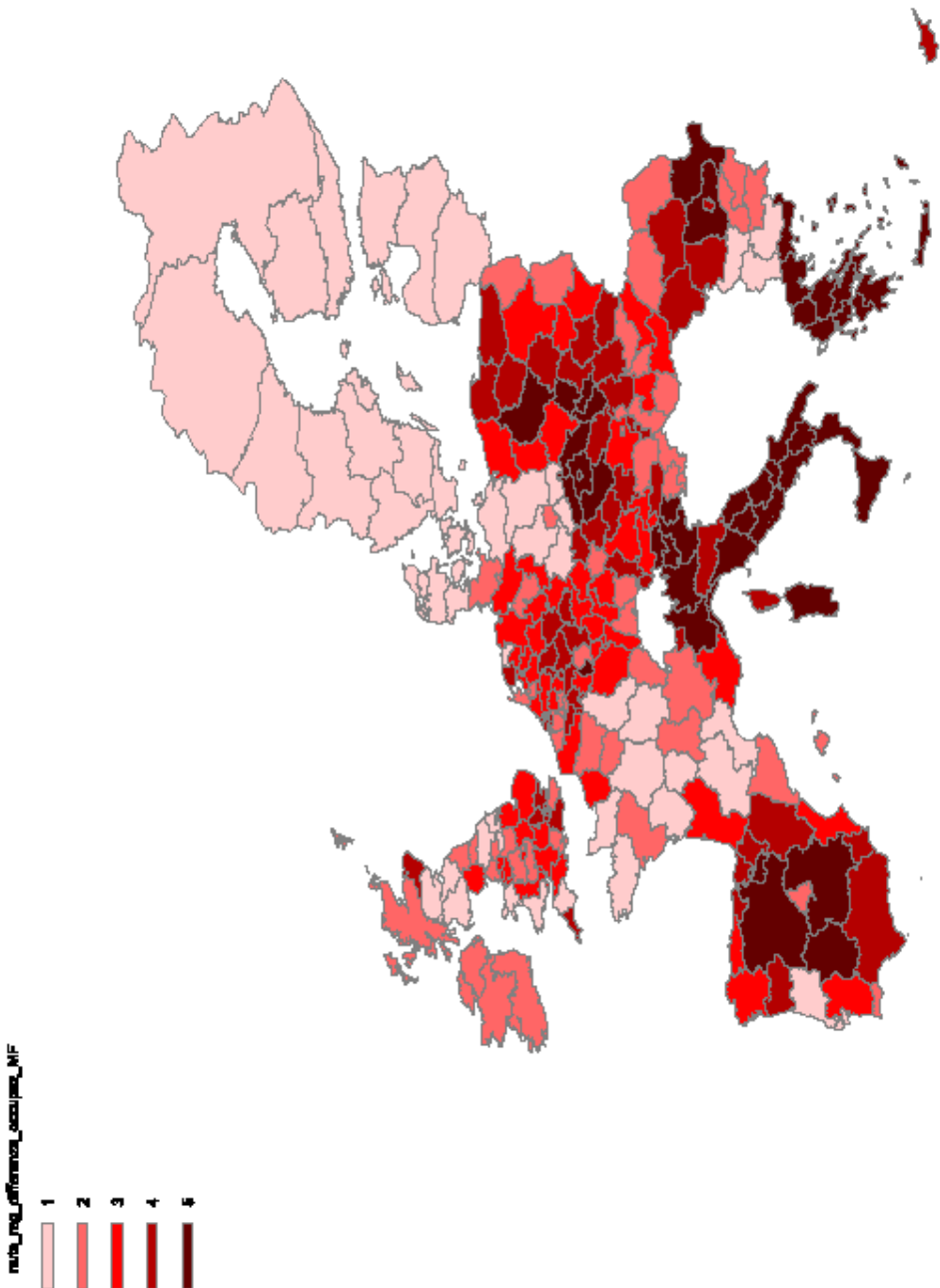


Figura 27. LAVORO: TASSO DI DISOCCUPAZIONE

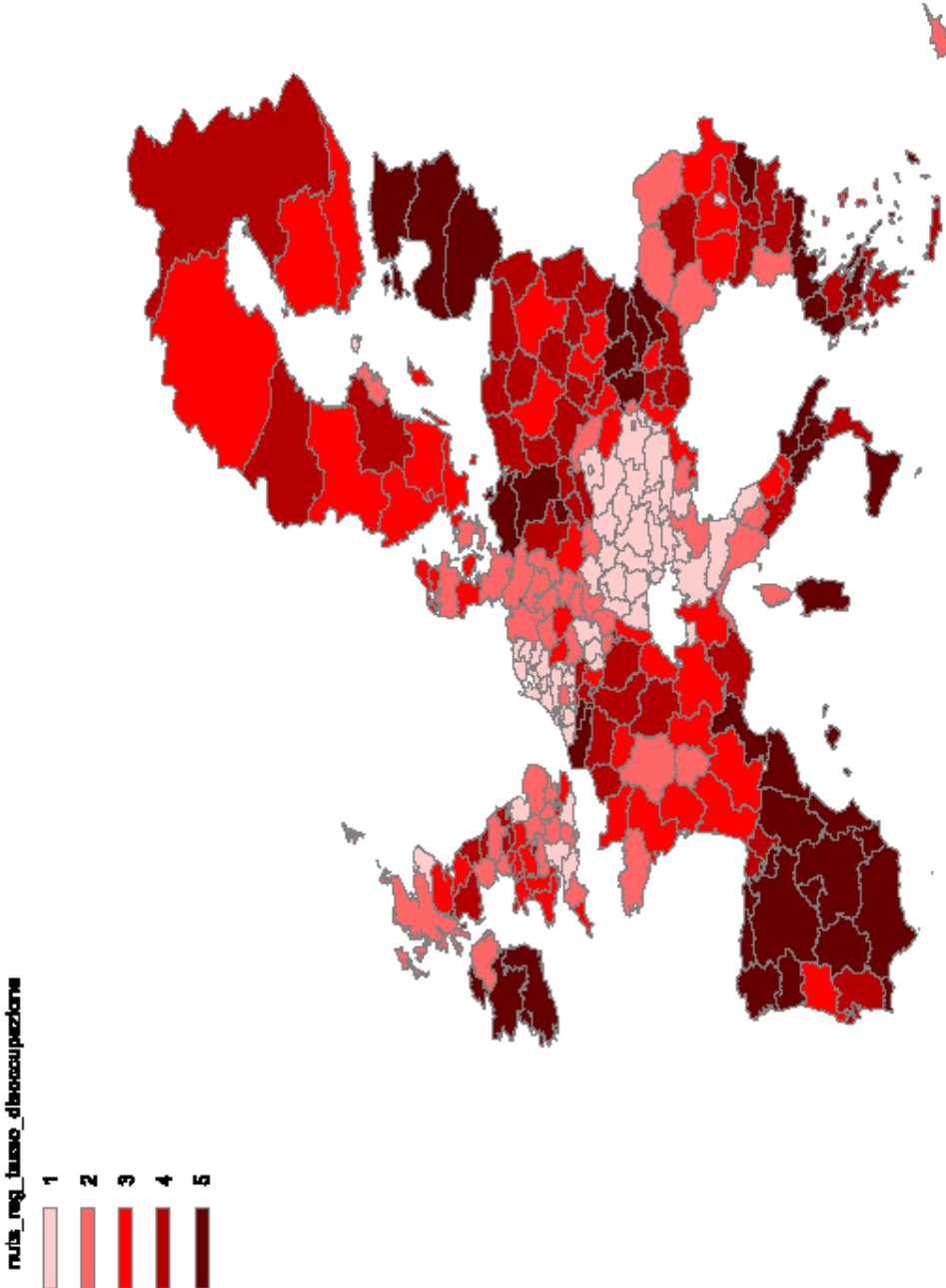


Figura 28. LAVORO: DIFFERENZA DISOCCUPAZIONE FEMMINE-MASCHI

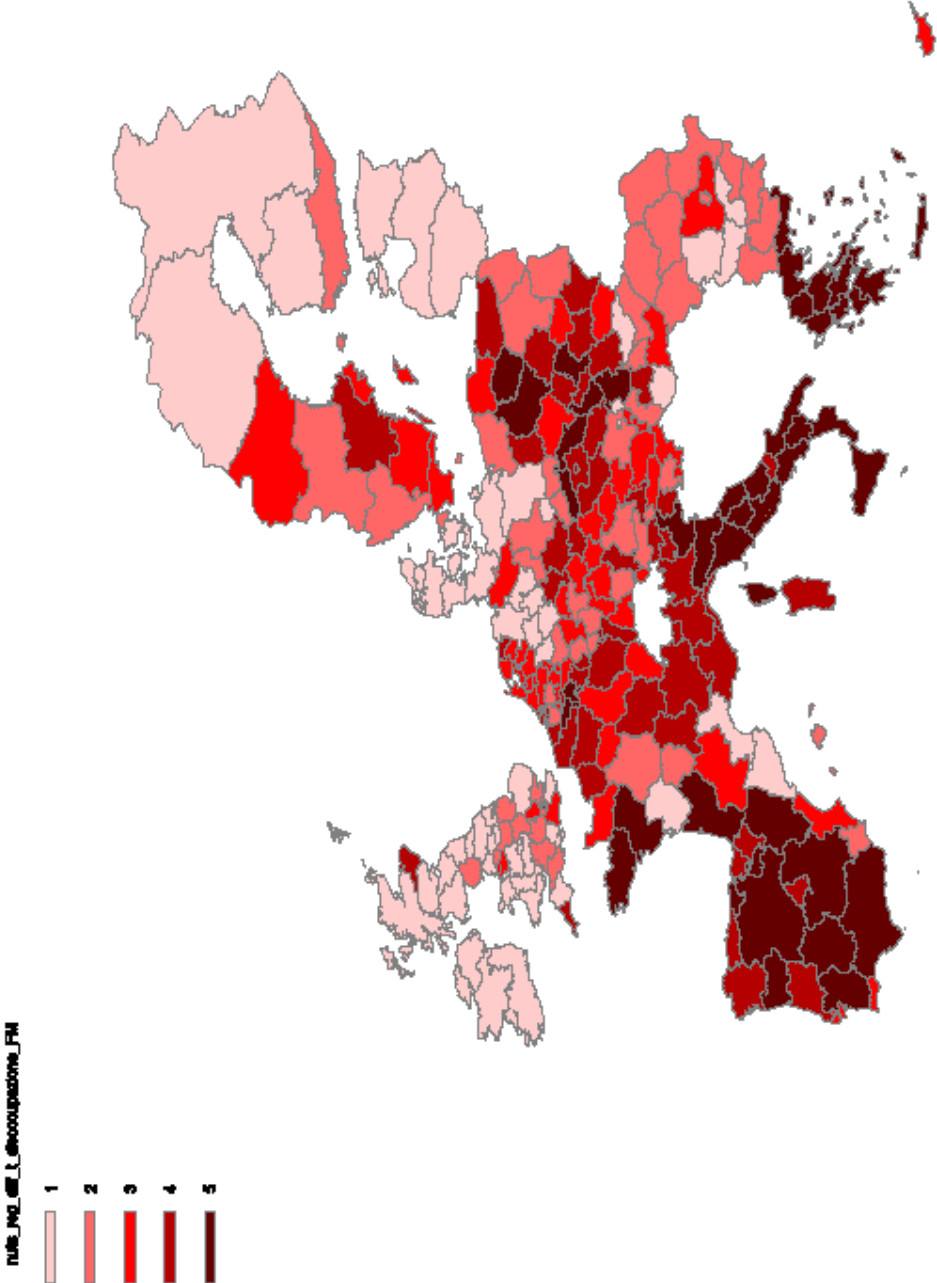


Figura 29. LAVORO: TASSO DI DISOCCUPAZIONE DI LUNGO PERIODO

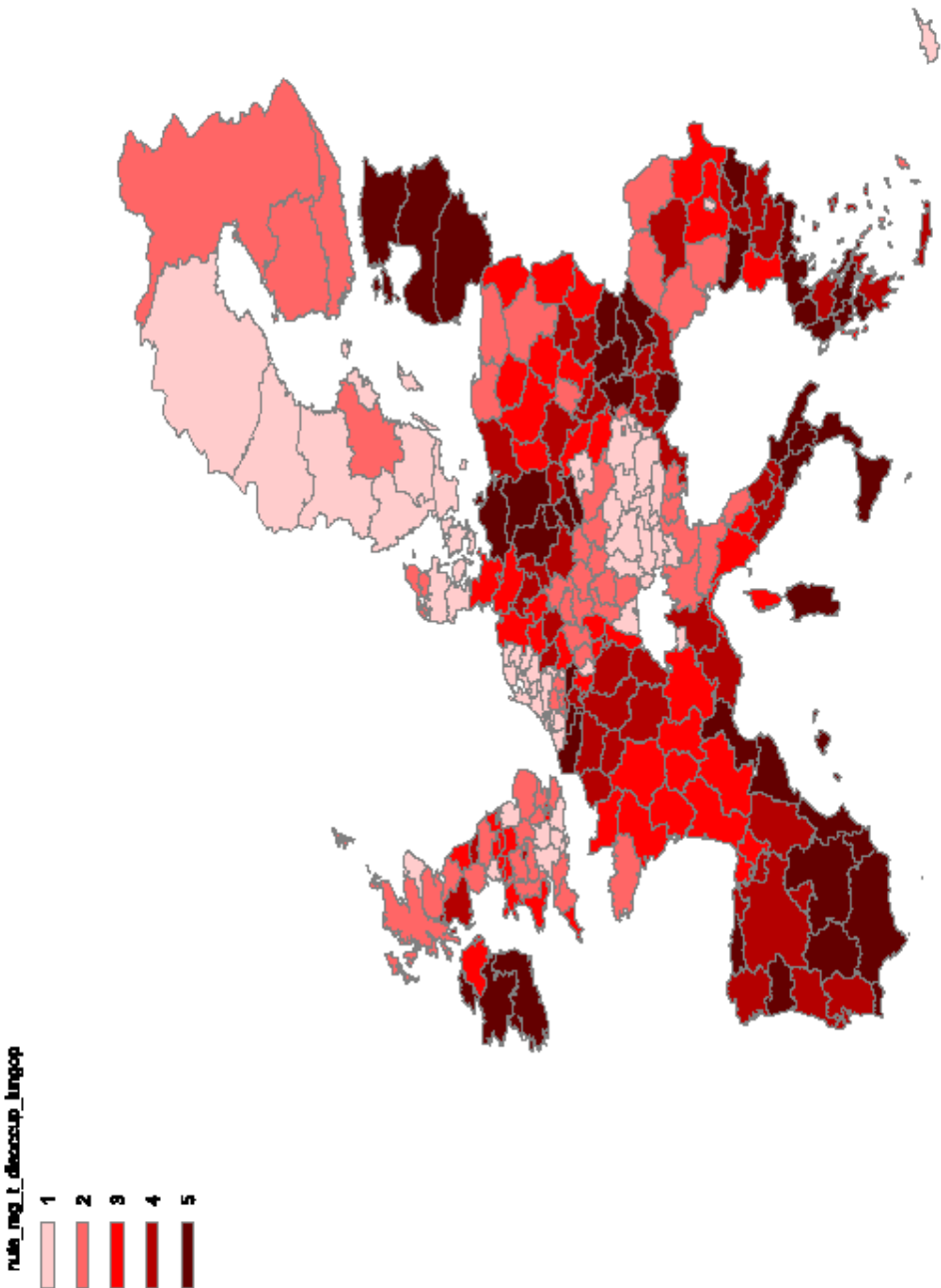


Figura 30. LAVORO: DIFFERENZA DISOCCUPAZIONE GIOVANI-ADULTI

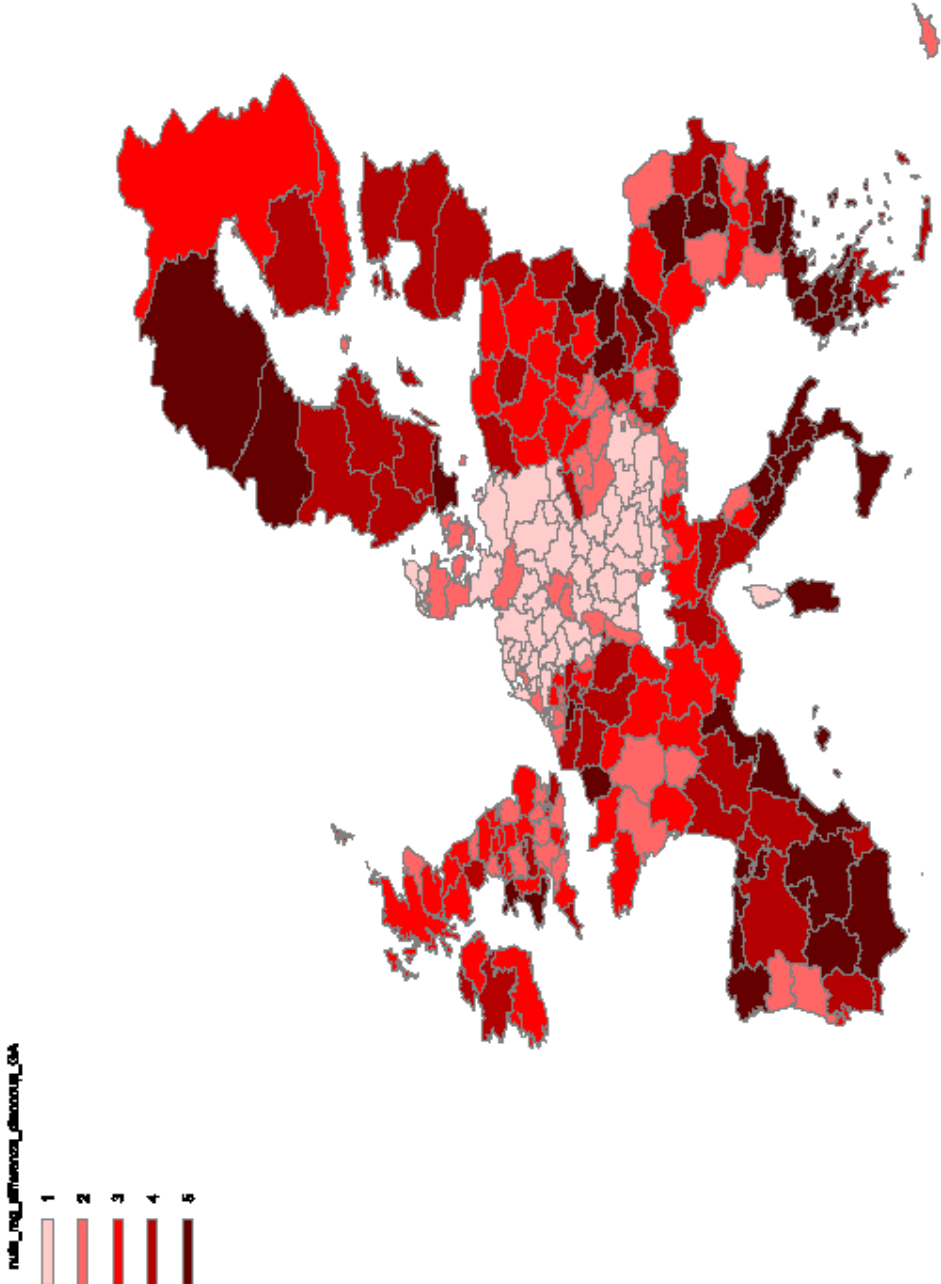


Figura 31. AMBIENTE: PERCENTUALE TERRENI AGRICOLI

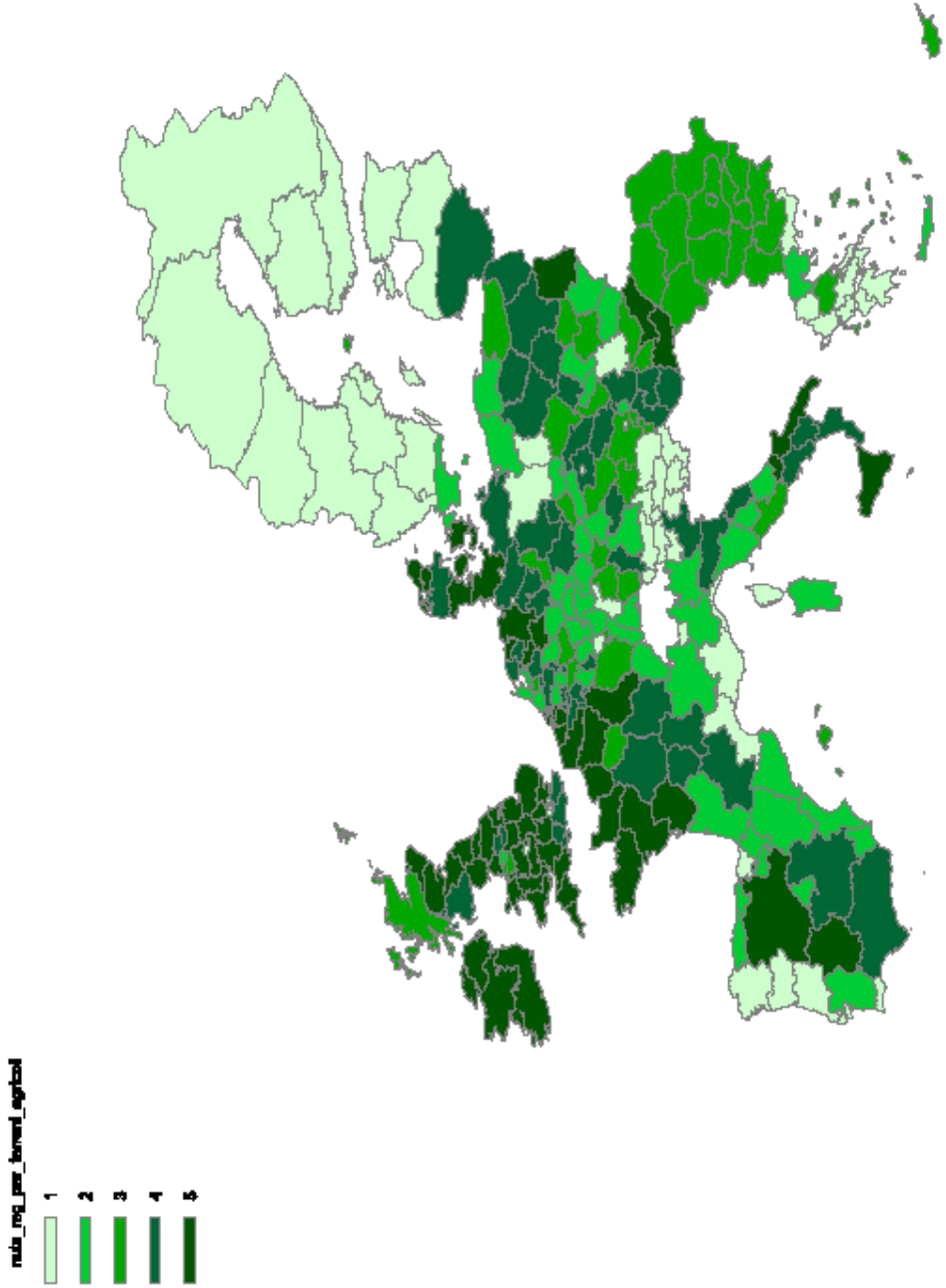


Figura 32. AMBIENTE: PERCENTUALE TERRENI COPERTI DA FORESTE

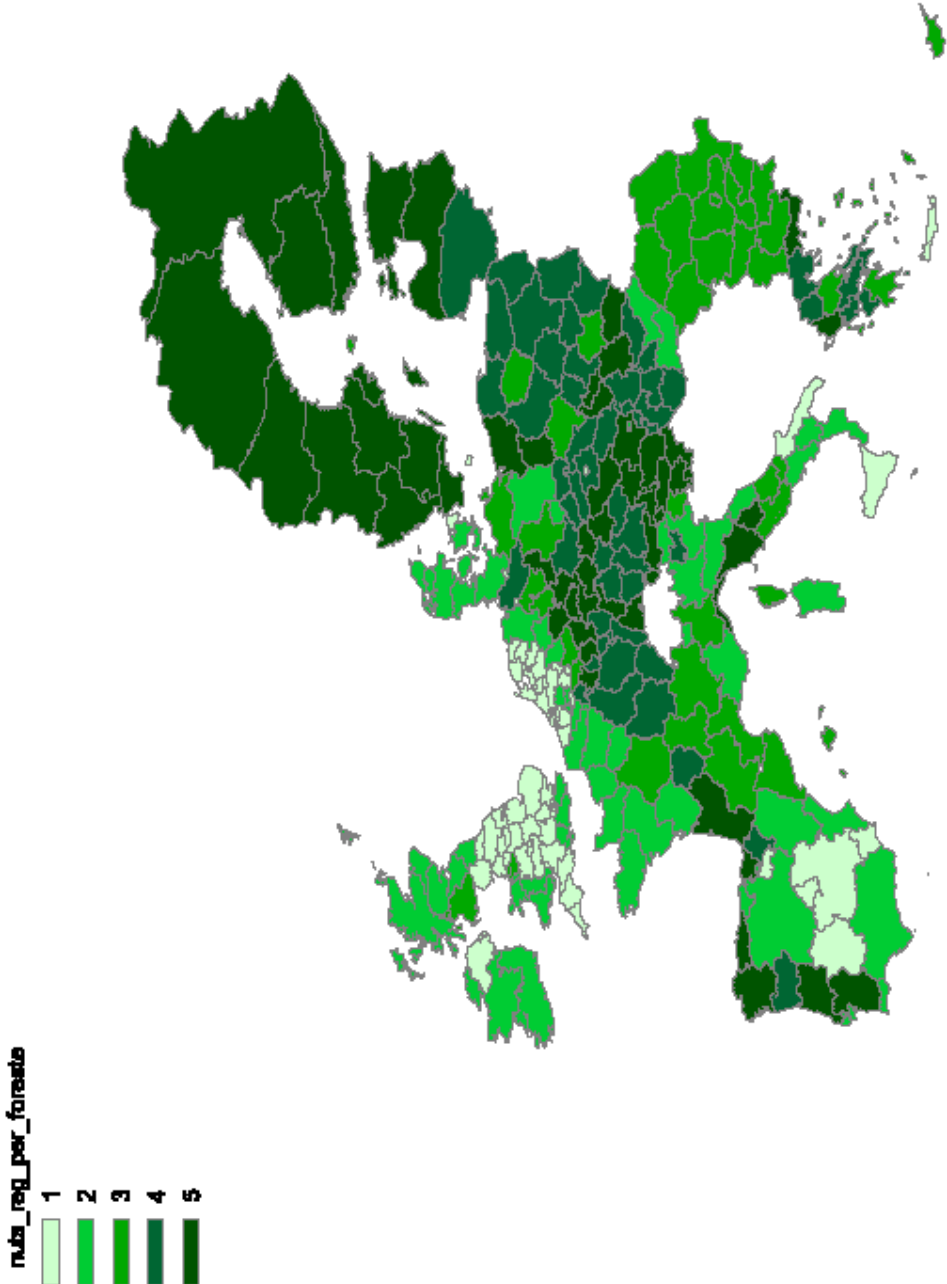


Figura 33. AMBIENTE: PERCENTUALE ZONE RESIDENZIALI, COMMERCIALI, INDUSTRIALI

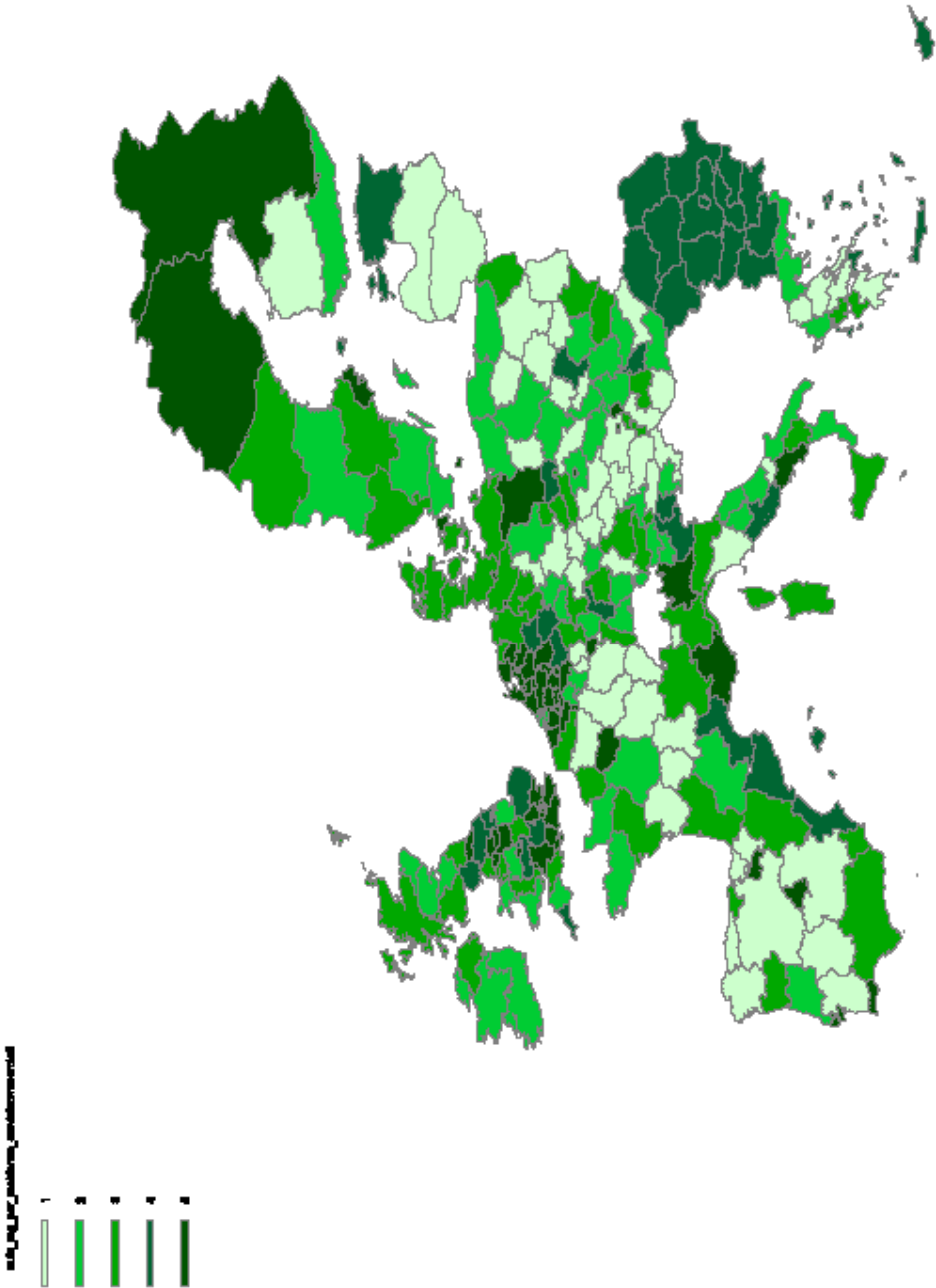


Figura 34. AMBIENTE: PERCENTUALE TERRENI SU CUI VENGONO SVOLTE ATTIVITA' AD ALTO IMPATTO AMBIENTALE

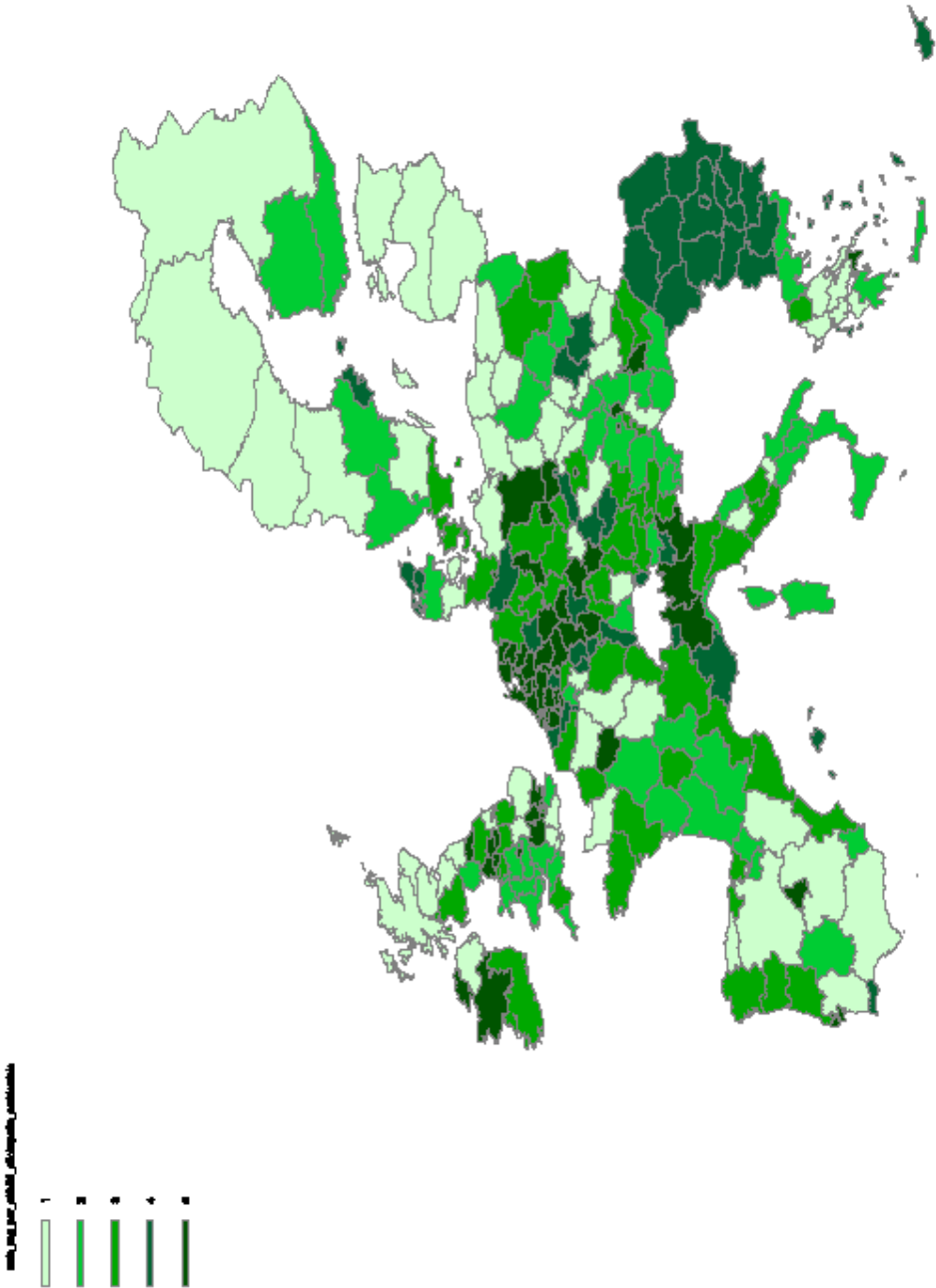


Figura 35. AMBIENTE: RIFIUTI PROCAPITE

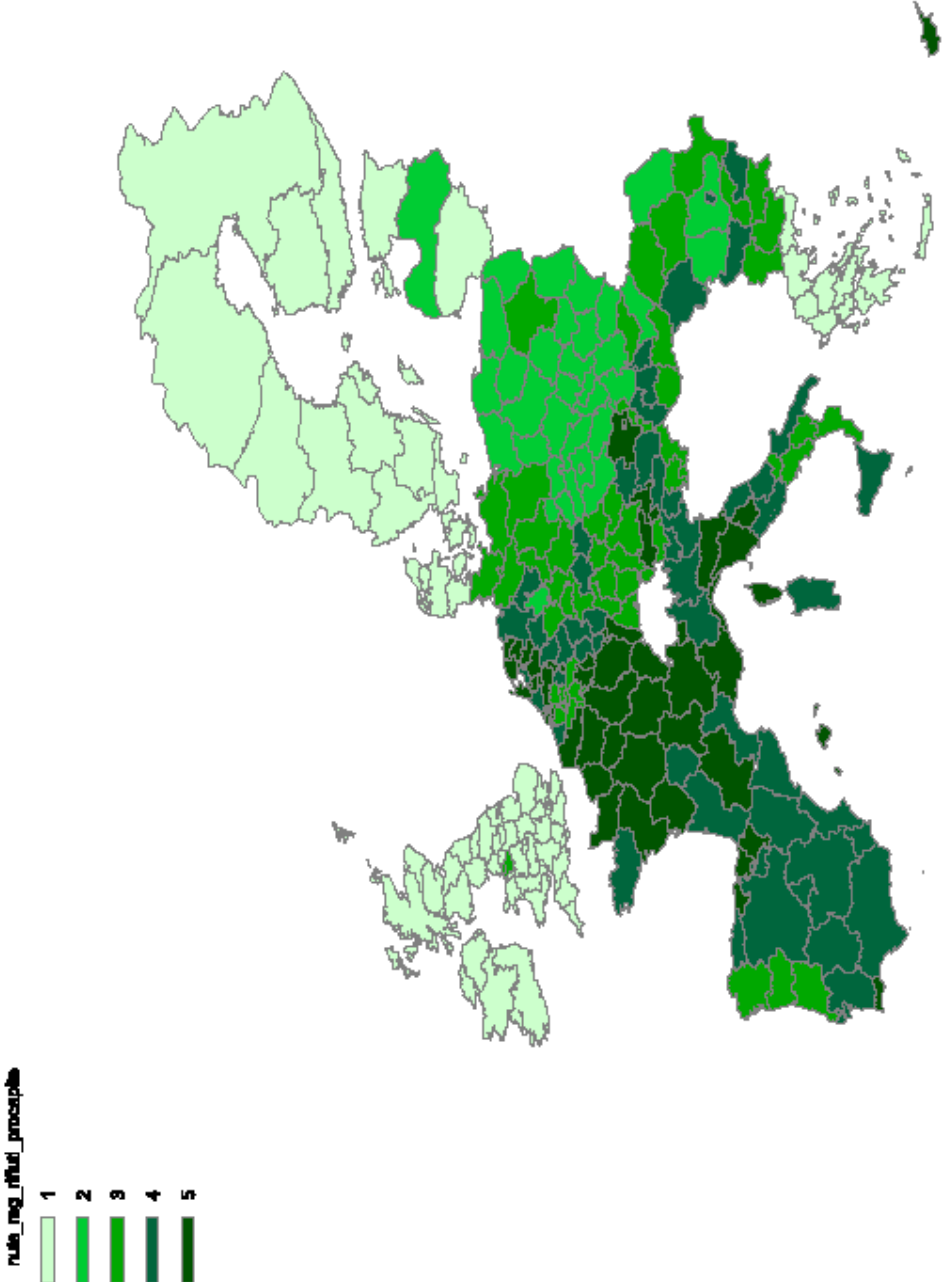
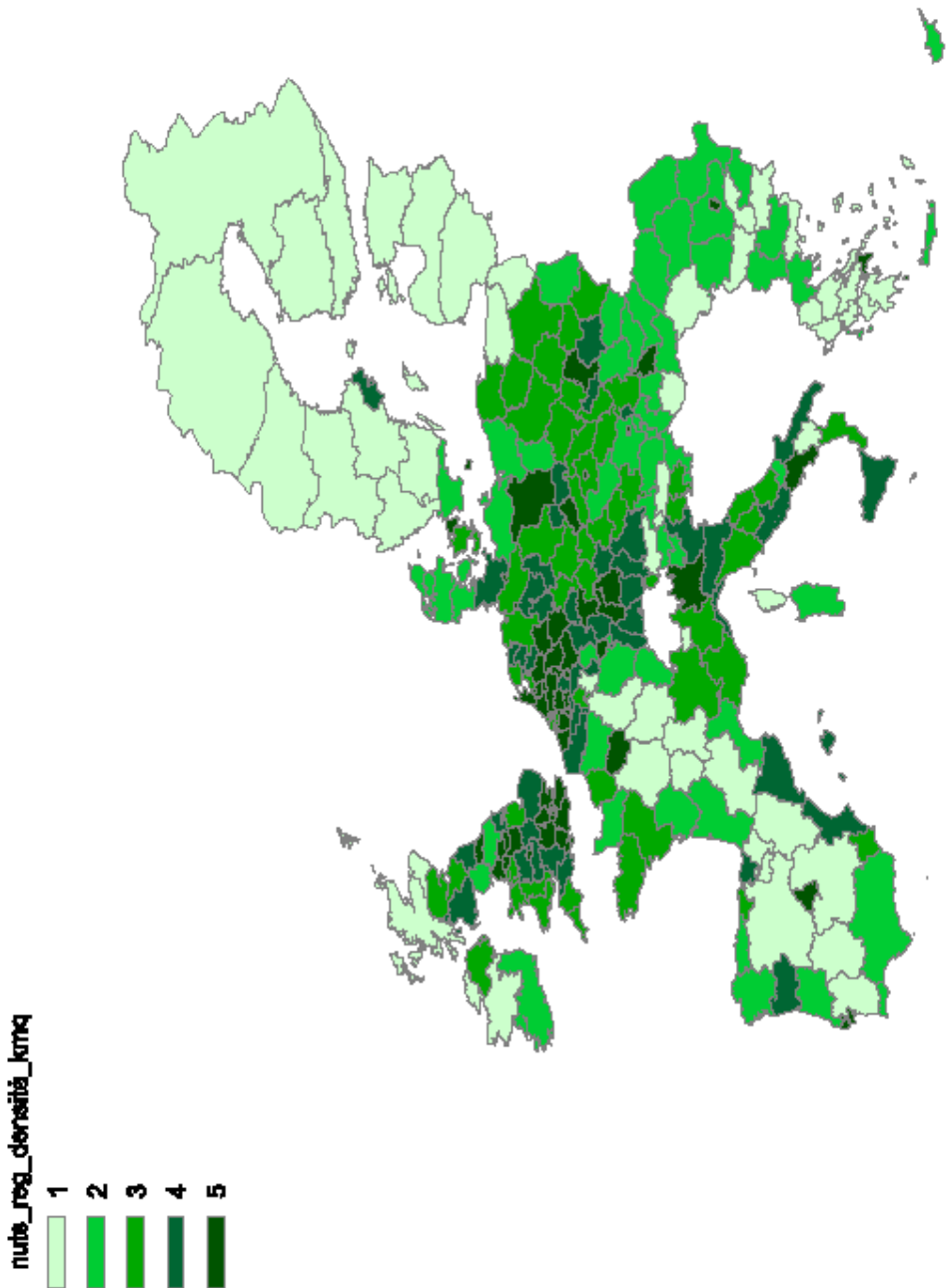


Figura 36. AMBIENTE: DENSITA' ABITATIVA IN KMQ



CAPITOLO 4 - CREAZIONE DI UN INDICATORE COMPOSITO DI BENESSERE: ANALISI MULTIDIMENSIONALI

Si presenta l'annoso ed irrisolto problema sull'opportunità della creazione di un indice composito di benessere.¹¹

Da un lato molti studiosi contestano l'uso di indici sintetici che portino alla determinazione di un unico valore per ciascuna area geografica, prediligendo il cosiddetto *dashboard* ossia un 'cruscotto' nel quale sia possibile identificare diverse dimensioni del fenomeno, tutte rilevanti, senza che esse siano ulteriormente aggregate. Dal punto di vista statistico si tratta di una scelta inoppugnabile e vari sono gli studiosi che si schierano contro l'utilizzo degli indici composti, adducendo le seguenti motivazioni: possono condurre a messaggi politici fuorvianti se mal costruiti o mal interpretati; la selezione degli indicatori ed i pesi ad essi assegnati possono costituire oggetto di dispute politiche ed indurre a conclusioni di *policy* semplicistiche.

Dall'altro lato gli indicatori composti sono sempre più riconosciuti come strumento utile per l'analisi socio-economica e per la comunicazione pubblica. Essi sintetizzano concetti multidimensionali che non possono essere spiegati da un singolo indice ed in particolare consentono semplici confronti tra differenti aree territoriali e sono di facile interpretazione da parte dei *policy maker* e del pubblico.

La costruzione di un indicatore composito, per garantire coerenza e trasparenza a tutto il processo, deve procedere di pari passo alla presentazione della documentazione di riferimento, sia per ciò che concerne le metodologie utilizzate, sia per ciò che riguarda i dati di partenza.

Dopo un'accorta selezione delle variabili da trattare, nel caso in cui si decida per l'aggregazione degli indici, il ricercatore si trova a dover operare due scelte fondamentali, che inevitabilmente andranno ad incidere sui risultati della ricerca: i pesi da assegnare ai vari sottoindicatori ed il metodo di aggregazione.

¹¹ "È difficile immaginare che il dibattito sull'uso di indicatori composti possa mai essere risolto."(Saisana, 2005)

Riguardo ai criteri di ponderazione, tecnicamente la procedura consiste nell'individuazione di pesi da assegnare ad ogni indicatore elementare, dopodiché ogni peso verrà moltiplicato per il corrispondente valore individuale dell'indicatore elementare.

Evitando valutazioni soggettive, e allo scopo di procedere alla difficile scelta tra i differenti approcci per la definizione dei pesi, il ricercatore deve tenere conto (Nardo, 2005):

- della struttura teorica e razionale sul quale la variabile latente e, conseguentemente l'indicatore sintetico, sono basati;
- del contributo di ogni indicatore elementare all'aggregazione;
- della qualità dei dati e l'adeguatezza degli indicatori elementari

Pertanto, i pesi rappresentano un giudizio sul valore di ciascun singolo indicatore e proprio per evitare tale valutazione la maggior parte degli indicatori composti suppone *equal weighting*, cioè che le variabili abbiano tutte lo stesso peso.

Nel caso si adotti invece una procedura "*different weighting*", esistono diverse tecniche, alcune derivate da modelli statistici come l'analisi delle componenti principali, l'analisi fattoriale e gli *unobserved components models* (UCM), alcune derivate da metodi partecipativi come *budget allocation processes* (BAP), *analytic hierarchy process* (AHP) e *conjoint analysis* (CA), altre che utilizzano algoritmi di programmazione lineare per ricavare la frontiera efficiente.

Passando ad esaminare il metodo di aggregazione, vari sono quelli disponibili, aventi ognuno specifici vantaggi e svantaggi. I più usati sono i metodi additivi che vanno dalla somma dei ranghi per ciascun indicatore alla media ponderata di valori trasformati degli indicatori originari. Tuttavia, i metodi additivi si basano su requisiti e proprietà che spesso sono non desiderabili o difficili da soddisfare: si assume ad esempio una completa sostituibilità tra le componenti dell'indice, cioè che valori bassi registrati da alcuni indicatori siano perfettamente bilanciati da valori alti relativi ad altri indicatori.

Al contrario, con i metodi di aggregazione moltiplicativi vengono risolti i problemi legati alla compensazione dei caratteri, ma se ne introducono altri. Ad esempio la media geometrica assegna un peso maggiore ai valori più bassi e soprattutto non può essere calcolata in presenza di valori negativi o nulli; in aggiunta a ciò essa presuppone l'esistenza di un legame moltiplicativo tra le variabili.

Di seguito si attuano le procedure necessarie per la costruzione dell'indicatore sintetico di benessere, cercando di operare con la massima trasparenza possibile.

4.1 ANALISI INDICATORI PER AMBITO DI BENESSERE

CALCOLO DELL'INDICE DI AFFIDABILITA' ALPHA DI CRONBACH

Si analizzano le relazioni infragruppo, cioè quelle esistenti tra gli indicatori appartenenti al medesimo ambito di benessere, tramite il calcolo dell'Alpha di Cronbach (Cronbach, 1951). Esso è in grado di valutare come un insieme di variabili, nel nostro caso i sotto-indicatori, riescano o meno a misurare il costrutto, cioè il benessere relativo a ciascun ambito analizzato.

Tale coefficiente misura la porzione di variabilità del set di dati spiegata dalla correlazione tra i sottoindicatori e varia nell'intervallo chiuso [0, 1]: se non esiste correlazione tra i sottoindicatori Alpha è uguale a zero, mentre se i sottoindicatori sono perfettamente correlati il coefficiente è uguale ad 1. Pertanto esso misura la consistenza interna, ovvero quanto le variabili considerate costituiscono una misura soddisfacente del fenomeno latente sottostante, cioè nel nostro caso il benessere in ciascun ambito.

La formula per calcolare tale indice è la seguente:

$$\alpha = (p \cdot \bar{r}) / 1 + \bar{r} (p-1)$$

dove p è il numero di variabili ed \bar{r} è la media delle correlazioni esistenti tra ogni coppia delle p variabili.

Empiricamente si ritengono accettabili analisi con valori di α superiori a 0,70, anche se una consistenza interna più bassa può essere accettabile, se la teoria suggerisce che il costrutto in questione è multidimensionale.

AMBITO DI BENESSERE	ALPHA CRONBACH
ISTRUZIONE	0,479
SALUTE E INCIDENTALITA'	0,687
RICCHEZZA E TEMPO LIBERO	0,641
AMBIENTE	0,543
DEMOGRAFIA	0,780
LAVORO	0,860
BENESSERE TOTALE	0,735

Nella nostra analisi il valore di Alpha è elevato per quanto riguarda gli ambiti lavoro e demografia, mentre è più basso se riferito al settore dell'istruzione e dell'ambiente.

Nel caso vengano considerate tutte le 31 variabili, si segnala un valore dell'indice soddisfacente, pari a 0,735. Pertanto, l'insieme delle 31 variabili può ragionevolmente ritenersi espressione di una sola variabile latente, il benessere.

“BIPARTIZIONE” DEI SOTTOINDICATORI

Ponendosi come obiettivo quello di stilare una classifica del benessere nelle regioni europee, occorre operare una bipartizione dei sottoindicatori in maniera da distinguere i fattori di benessere da quelli di malessere: se per la maggior parte degli indici si tratta di un'operazione rapida e scontata e la relazione è diretta o inversa, come nel caso del tasso di mortalità infantile, del tasso di occupazione, del livello di istruzione, ecc., per taluni di essi la soluzione ultima rimane controversa, poiché la variabile in questione può condurre ad effetti sul benessere di segno opposto, a seconda

della diversa dimensione del benessere considerata (relazione non lineare con il benessere medesimo) .

Alcuni esempi chiariscono il concetto:

- il possesso dell'automobile se da un lato è un indice di disponibilità economica, dall'altro produce un'esternalità negativa, in quanto costituisce una fonte di inquinamento;
- un numero di ore lavorate in aumento comporta solitamente un duplice effetto: un incremento nel reddito da lavoro, ma pure una diminuzione del tempo a disposizione per il tempo libero ed i rapporti sociali;
- le zone urbanizzate comprendono edifici ed infrastrutture necessari per esigenze abitative, commerciali e di circolazione; tuttavia l'eccessiva antropizzazione del territorio comporta innegabili ripercussioni negative da un punto di vista ambientale (erosione del suolo, dissesto idrogeologico, congestione del traffico, inquinamento).

La soluzione adottata è quella di valutare l'impatto del singolo indicatore, considerandone l'effetto prevalente all'interno della dimensione di benessere nel quale viene inserito. Nella Tabella seguente si riporta il segno della componente: + se la variabile esprime un aspetto di benessere, - se l'indicatore esprime un aspetto di malessere.

Di conseguenza nel database si procede ad invertire il segno delle componenti di malessere, in maniera tale che tutti gli indicatori elementari presentino correlazione positiva con la variabile latente, cioè il benessere globale.

<i>ETICHETTA</i>	<i>INDICATORE BENESSERE</i>	<i>COMPONENTE</i>
	DEMOGRAFICI	
ind_vecc	Indice di vecchiaia: $\text{vecchiaia} > 64 / (0-14) * 100$	-
ind_dip	Indice di dipendenza $(<15 + >64) / (15-64) * 100$	-
%<10	Percentuale popolazione di età inferiore a 10 anni	+
%>74	Percentuale popolazione di età superiore a 74 anni	-
fert_rate	Tasso di fertilità totale	+
nat_change	Tasso netto movimento naturale popolazione (media anni 2006-2010)	+
nat_migration	Tasso netto movimento migratorio popolazione (media anni 2006-2010)	+
	RICCHEZZA E TEMPO LIBERO	
GDP PPT	Pil a parità di potere d'acquisto (100= valore medio)	+
Disposable income	Reddito disponibile netto delle famiglie (100= valore medio)	+
Vehicles	Veicoli posseduti per 1000 abitanti (eccetto rimorchi e motocicli)	+
Ore tempo libero	% ore settimanali non dedicate a lavoro retribuito	+
	LAVORO	
E15_64	tasso di occupazione (Occupati 15_64/Popolazione 15_64*100)	+
employ_M-F	differenza tra tasso occupazione maschile e quello femminile	-
disoccup_totale	tasso disoccupazione totale (disoccupati ed in cerca primo lavoro su popolazione att	-
long_unempl	tasso disoccupazione di lungo periodo (superiore ai 12 mesi) su popolazione attiva	-
unemp_F-M	differenza tra tasso disoccupazione femminile e quello maschile	-
diff_dis_giov-ad	differenza tra tasso disoccupazione giovanile e quello riferito agli over 24	-
	SALUTE E INCIDENTALITA'	
asp_vita	Speranza di vita alla nascita: numero medio di anni di vita alla nascita	+
infant_mort	Tasso di mortalità infantile: morti nel primo anno di vita su 1000 nati vivi	-
vittime_inc	Morti in incidenti stradali sul colpo o entro 30 giorni su 100.000 residenti	-
	ISTRUZIONE	
upper_sec	Medio livello di educazione, livelli 3 e 4	+
terz_edu	Alto livello di educazione, livelli 5 e 6	+
formaz_ad	Life long learning	+
r_d	Investimenti in ricerca e sviluppo in % rispetto al Pil	+
patent_app	Brevetti riconosciuti dall'European Patent Office per 1.000.000 residenti	+
	AMBIENTE	
foreste	% suolo coperto da foreste e arbusteti	+
alto impatto_amb	% suolo su cui vengono svolte attività ad alto impatto ambientale	-
servizi_residenze	% zone residenziali, commerciali ed industriali	-
rifiuti_proc_kg	rifiuti urbani prodotti per persona in kg	-
dens	Densità abitativa in kmq	-

Figura 37. Elenco dei 31 indicatori di benessere divisi per ambito

4.2 PRIME SINTESI PER CIASCUN AMBITO DI BENESSERE

Per quanto riguarda l'analisi sul benessere, è ragionevole assumere una completa sostituibilità tra gli indicatori appartenenti a ciascun ambito di benessere, quindi si aggregano gli indicatori all'interno di ciascuna area mediante la media aritmetica (*equal weighting*).

Pertanto, si calcola la media semplice tra gli indicatori di benessere appartenenti al medesimo settore di benessere. Prima di effettuare tale calcolo si provvede a trasformare le variabili su scala ordinale, in maniera tale da eliminare l'influenza di valori anomali.

Ancora una volta ricorriamo agli output di GvSig divisi per ciascun ambito di interesse e basati sulla divisione in quintili delle regioni europee (assegnando il rango 1 al valore più piccolo).

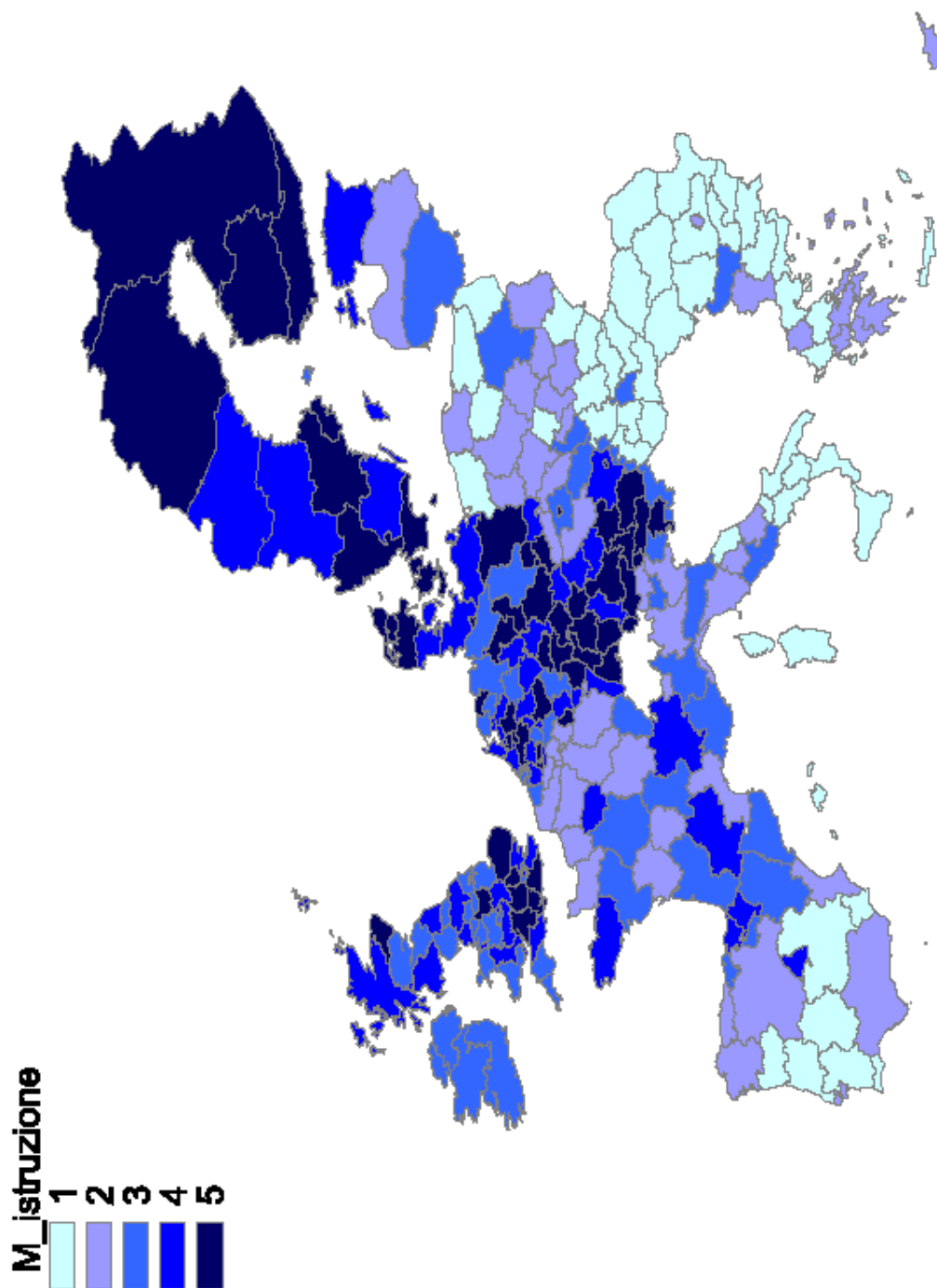


Figura 38. Media benessere ambito istruzione

La cartografia relativa all'indicatore istruzione ci mostra il primato delle regioni nord europee: i valori medi più elevati si registrano nelle regioni svedesi, finlandesi e danesi, seguite dalle regioni tedesche, austriache, olandesi e la parte meridionale dell'Inghilterra. Tali regioni investono ingenti somme in ricerca e sviluppo e puntano sulla formazione continua dei lavoratori. Le regioni tedesche si contraddistinguono per il maggior numero di brevetti riconosciuti dal preposto organo europeo.

Al contrario Romania, Bulgaria, Portogallo e le regioni meridionali dell'Italia registrano i valori più bassi in riferimento alla percentuale di laureati e pure agli investimenti in ricerca e formazione.

Osservando l'immagine riguardante l'indicatore salute, notiamo una concordanza di risultati in riferimento alle regioni del nord Europa: Svezia, Germania, Austria e Olanda si differenziano per i tassi più bassi di mortalità infantile, sia a quella dovuta sia agli incidenti stradali. Il primato della longevità spetta alle regioni italiane e ad alcune regioni francesi (Ile de France, Midi-Pyrénées e Rhône-Alpes) e della Spagna(Comunidad Foral de Navarra e Madrid, La Rioja, Castilla y Leon).

Sono invece caratterizzate da valori elevati di mortalità e nel contempo bassa longevità le regioni dell'est europeo, cioè quelle appartenenti a Romania, Bulgaria, Ungheria, Slovacchia, Polonia.

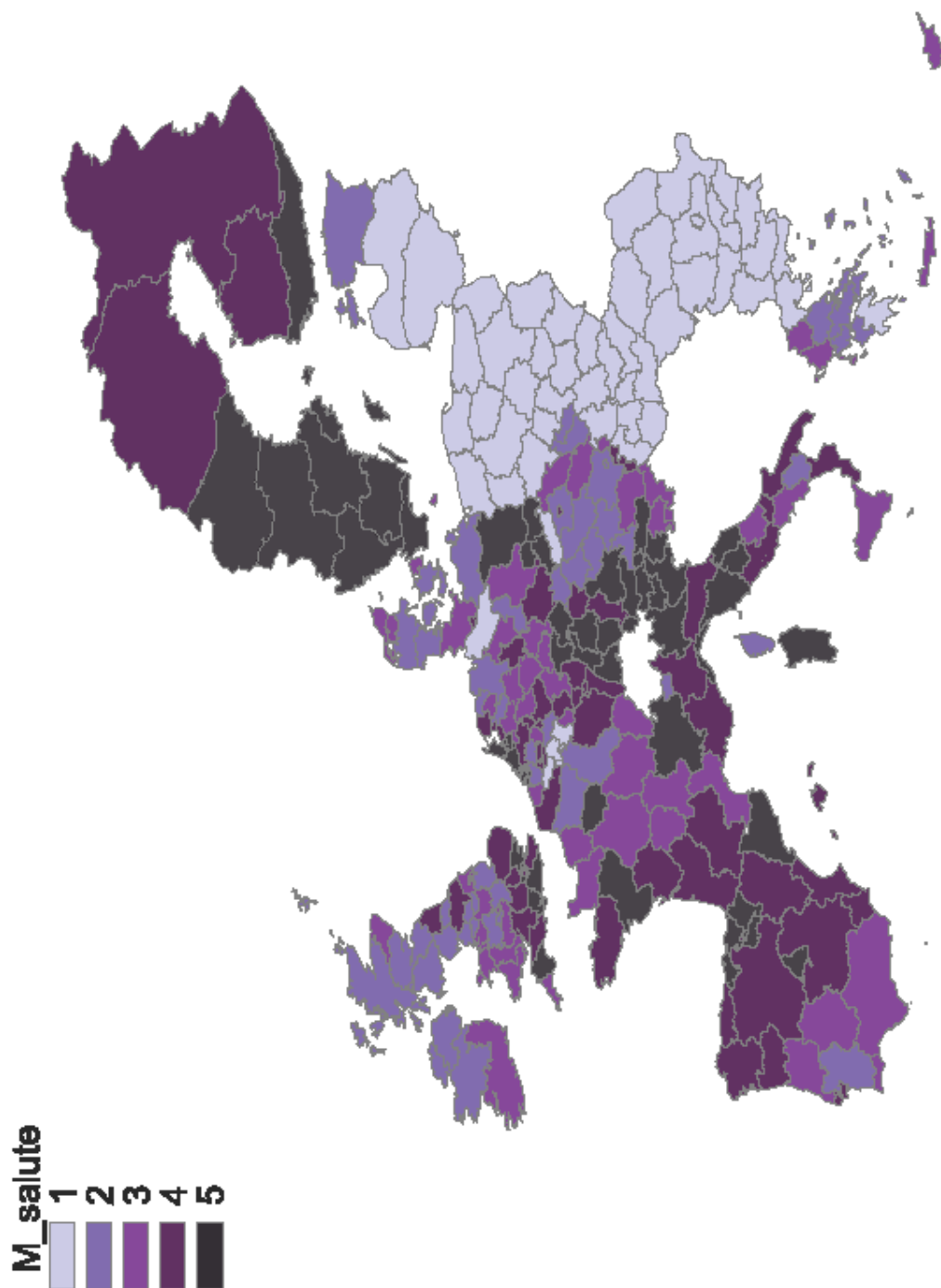


Figura 39. Media benessere ambito salute

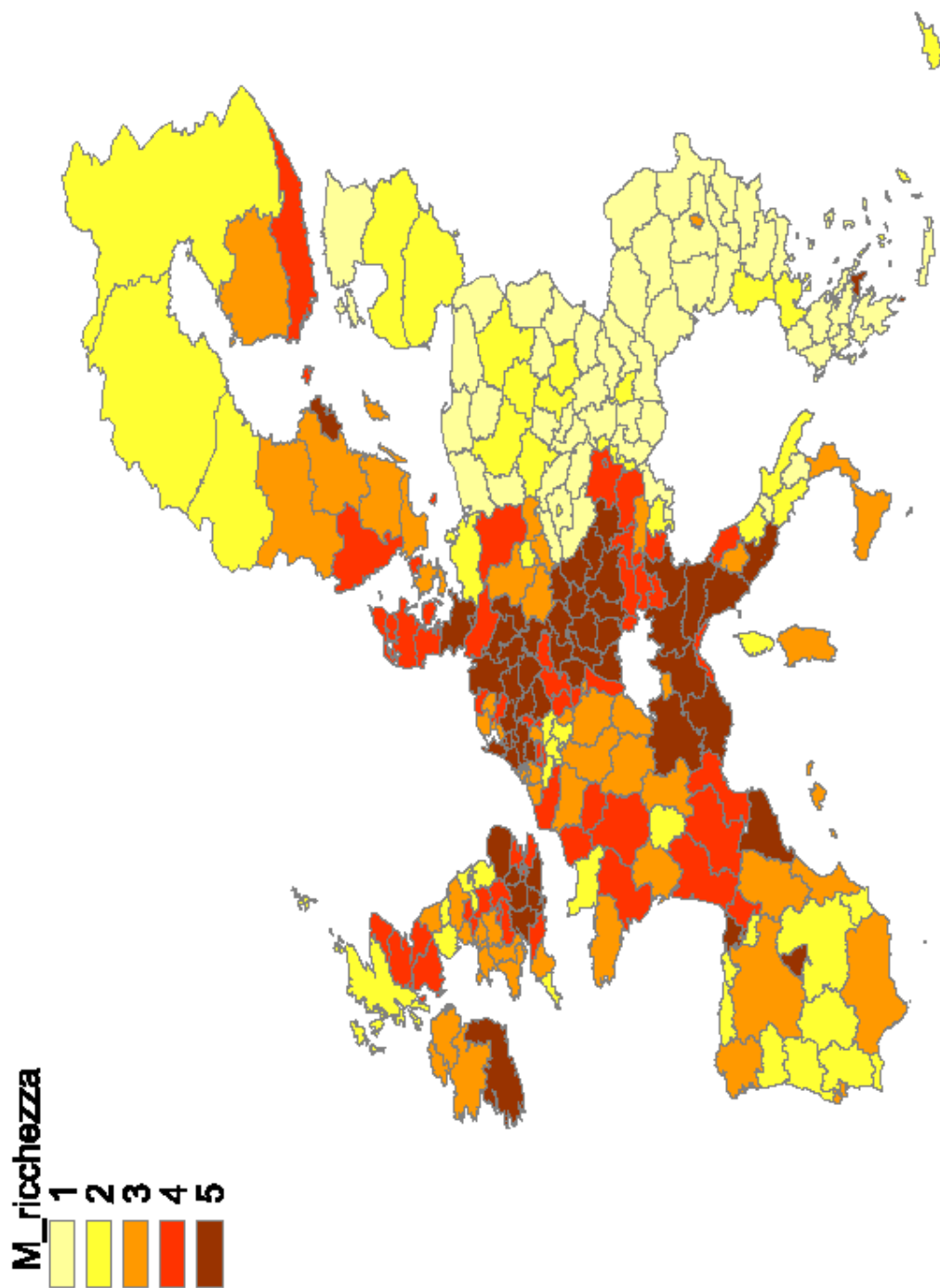


Figura 40. Media benessere ambito ricchezza

L'indicatore di ricchezza, di risorse economiche e tempo libero mostra il primato delle regioni tedesche riguardo a tutti aspetti esaminati in tale ambito; seguono poi alcune regioni olandesi (Zuid-Holland, Noord-Holland, Noord-Brabant) inglesi (Berkshire, Surrey, Gloucestershire) e italiane (Lombardia, Emilia-Romagna, Lazio, Veneto e Toscana). Queste ultime spiccano per l'elevato numero di autoveicoli posseduti.

Così come per l'ambito salute, anche in riferimento alla ricchezza le regioni dell'est europeo costituiscono il fanalino di coda dell'Europa; a queste si aggiungono le regioni della Grecia (ad eccezione di Atene).

Passando ad esaminare gli indicatori relativi al lavoro e alle disuguaglianze esistenti al suo interno, si nota che le regioni di eccellenza corrispondono circa alle medesime descritte nell'ambito della ricchezza: a dimostrazione di ciò le regioni a Sud della Germania (Trier, Oberbayern e Freiburg) si collocano entro le prime dieci posizioni. Il mercato del lavoro di tali regioni si contraddistingue inoltre per i valori più bassi di disuguaglianza esistenti tra i generi e tra generazioni. Accanto a queste spicca la regione finlandese Aland, quelle austriache (Tirol, Niederösterreich), danesi (Sjælland, Midtjylland) e olandesi (Utrecht, Noord-Holland).

Si registrano invece alti livelli di disoccupazione, unitamente a marcate disuguaglianze, nelle regioni spagnole (Ciudad Autónoma de Melilla e de Ceuta, Extremadura), dell'Italia meridionale ed insulare (Campania, Sicilia, Basilicata) e della Grecia (Dytiki Makedonia, Thraky e Sterea Ellada).

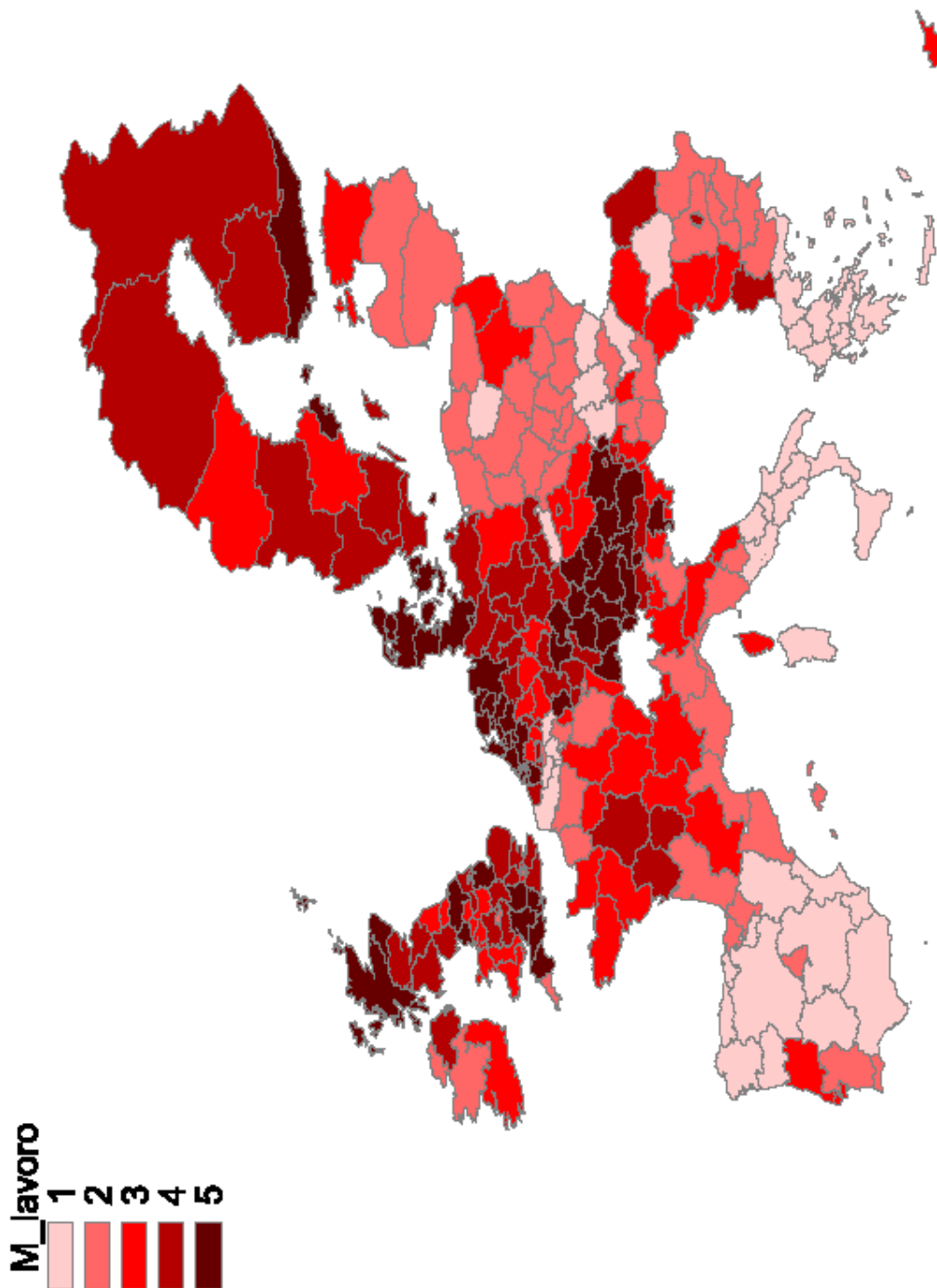


Figura 41. Media benessere ambito lavoro

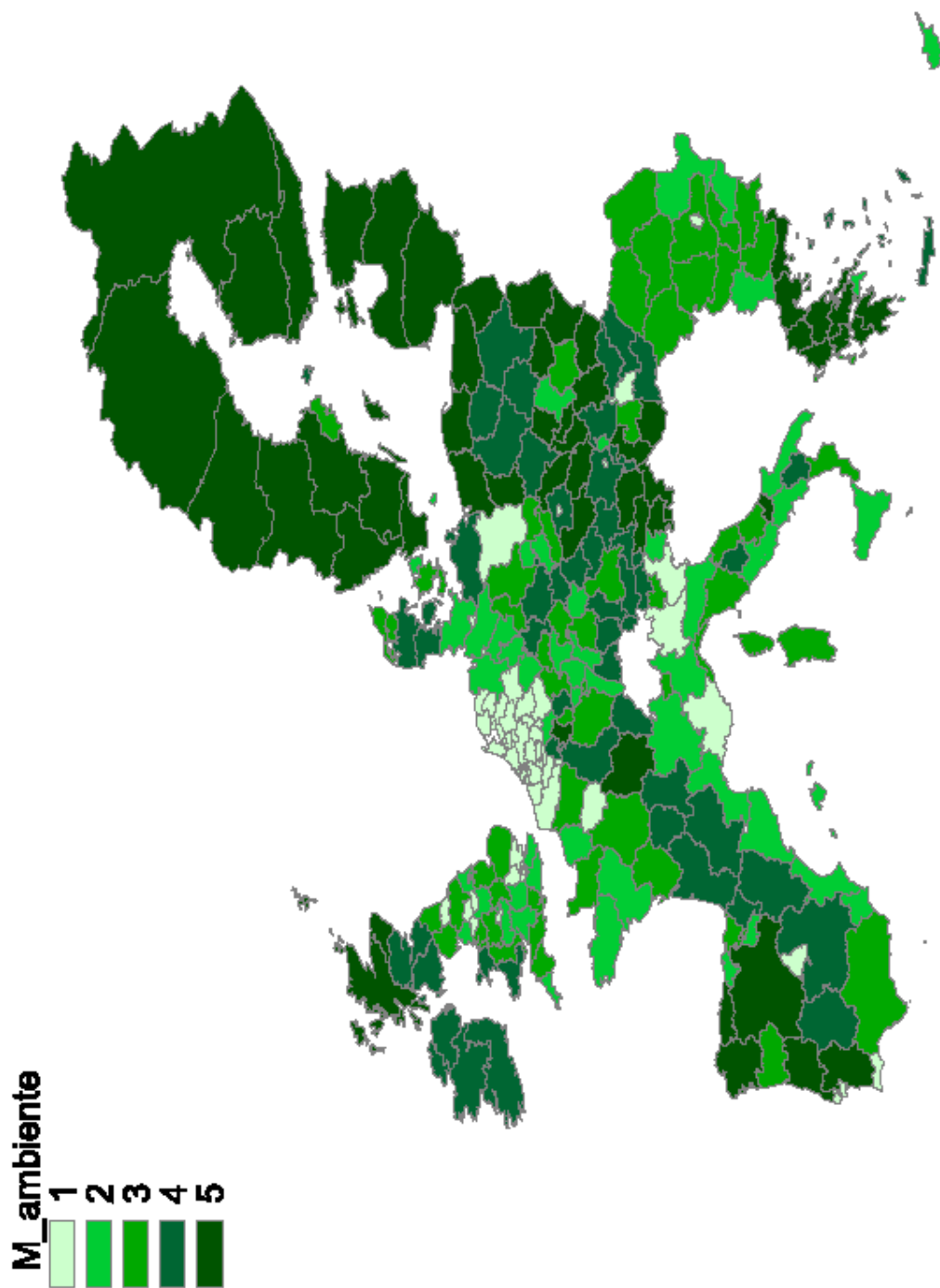


Figura 42. Media benessere ambito ambiente

Osservando gli indicatori ambientali, vediamo che le regioni di Svezia e Finlandia occupano i primi posti grazie a: bassa densità abitativa, presenza di estese foreste e presenza di attività a basso impatto ambientale. In aggiunta a queste le regioni elleniche spiccano per una densità abitativa limitata e quelle austriache e portoghesi per l'elevata percentuale di suolo coperto da foreste.

All'estremo opposto le regioni olandesi e danesi spiccano per opposte ragioni: alta densità abitativa, minore presenza di foreste ed elevata concentrazione di attività ad alto impatto ambientale.

Affrontando infine l'ambito demografico, occorre ragionare in ottica di singola regione per scoprire le situazioni migliori: le regioni olandesi Flevoland, Utrecht e quelle inglesi Northern Ireland, Berkshire e Londra registrano valori bassi dell'indice di vecchiaia accanto ad un'elevata percentuale di under 10; le regioni spagnole Murcia, Ceuta e le isole Baleari si caratterizzano per l'elevato tasso migratorio; quelle francesi Reunion, Nord - Pas-de-Calais e Picardie per l'alto tasso di fertilità e di movimento naturale.

Alti tassi di vecchiaia e una bassa percentuale di bambini al di sotto dei 10 anni caratterizzano le regioni tedesche ed italiane (Liguria, Molise e Basilicata in particolare).

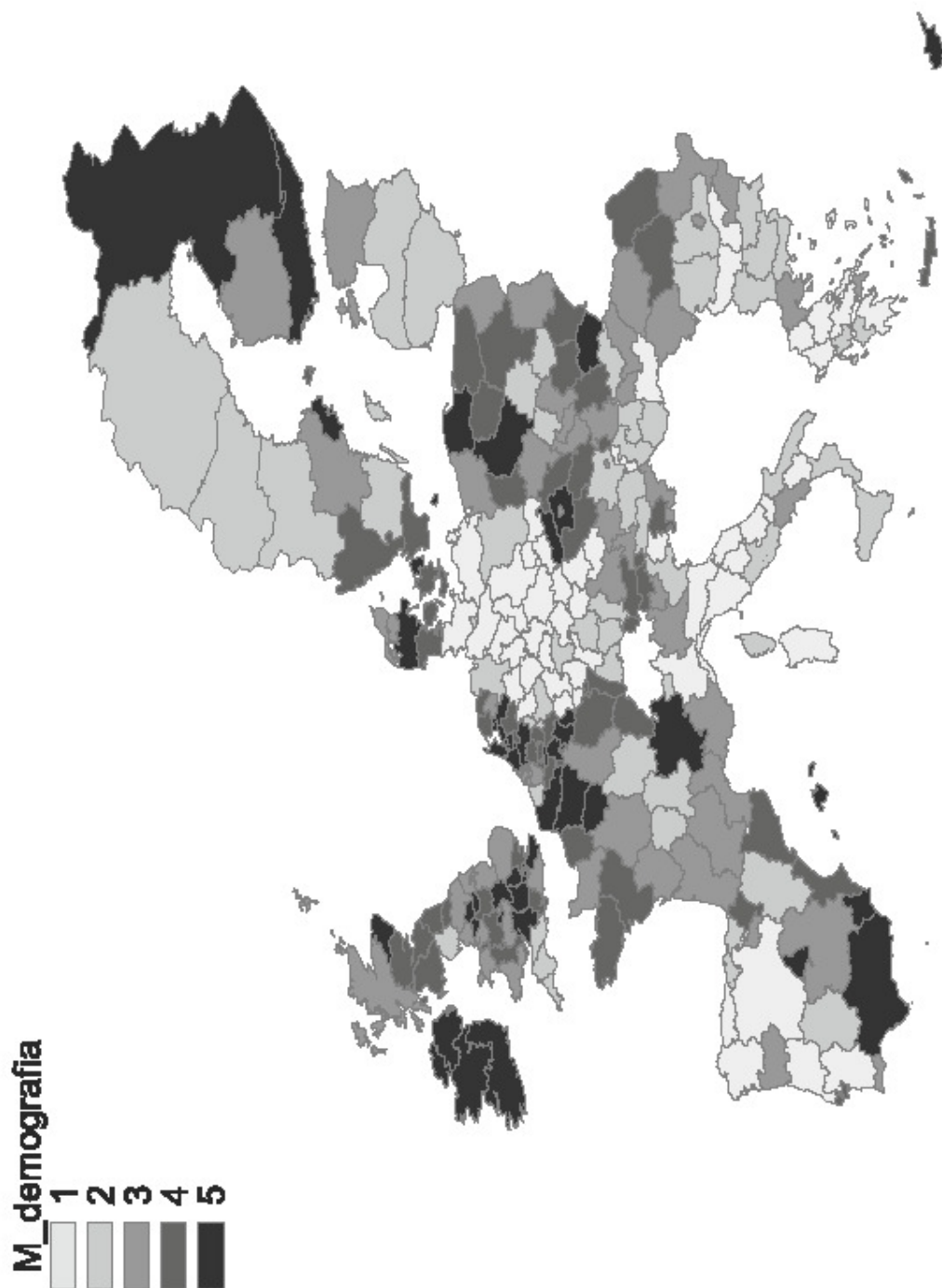


Figura 43. Media benessere ambito demografia

AGGREGAZIONE TRAMITE MEDIA PONDERATA

Si aggregano gli indicatori per ciascun ambito di benessere ricorrendo alla media ponderata, ricorrendo all'Alpha di Cronbach per definire i pesi. Tale scelta è giustificata dal fatto che la disponibilità di dati e nel nostro caso di indicatori atti a misurare tutti gli aspetti del benessere cambia a seconda dell'ambito preso in considerazione: gli aspetti demografici, così come quelli relativi al lavoro sono indagati con particolare accuratezza, al contrario l'aspetto ambientale a livello regionale viene esaminato in maniera più marginale.

AMBITO DI BENESSERE	ALPHA CRONBACH	W
ISTRUZIONE	0,479	0,120
SALUTE E INCIDENTALITA'	0,687	0,172
RICCHEZZA E TEMPO LIBERO	0,641	0,161
AMBIENTE	0,543	0,136
DEMOGRAFIA	0,78	0,195
LAVORO	0,86	0,216
TOTALE	3,99	1

Si riporta la classifica che ne scaturisce, confrontandola con quella che si otterrebbe tramite il calcolo della media aritmetica semplice.

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
SE11	Stockholm	SE	1	1
FI18	Etelä-Suomi	FI	2	2
SE23	Västsverige	SE	3	3
DE21	Oberbayern	DE	4	4
DE14	Tübingen	DE	5	8
SE22	Sydsverige	SE	6	9
UKJ3	Hampshire and Isle of Wight	UK	7	6
UKJ1	Berkshire, Buckinghamshire and Oxfordshire	UK	8	5
UKM5	North Eastern Scotland	UK	9	12
FI20	Åland	FI	10	7
AT33	Tirol	AT	11	11
DE13	Freiburg	DE	12	15
UKK1	Gloucestershire, Wiltshire and Bristol/Bath area	UK	13	13
SE12	Östra Mellansverige	SE	14	21
DK04	Midtjylland	DK	15	14
FI19	Länsi-Suomi	FI	16	23
DE11	Stuttgart	DE	17	20
UKJ2	Surrey, East and West Sussex	UK	18	18
SE21	Småland med öarna	SE	19	22
NL31	Utrecht	NL	20	10

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
AT34	Vorarlberg	AT	21	19
DE71	Darmstadt	DE	22	24
FI1A	Pohjois- ja Itä-Suomi	FI	23	25
DK01	Hovedstaden	DK	24	17
NL32	Noord-Holland	NL	25	16
DE12	Karlsruhe	DE	26	32
IE02	Southern and Eastern	IE	27	28
AT31	Oberösterreich	AT	28	30
UKH2	Bedfordshire and Hertfordshire	UK	29	27
UKH1	East Anglia	UK	30	31
AT22	Steiermark	AT	31	35
DE27	Schwaben	DE	32	34
DK05	Nordjylland	DK	33	33
NL33	Zuid-Holland	NL	34	26
DE25	Mittelfranken	DE	35	40
SE31	Norra Mellansverige	SE	36	53
NL41	Noord-Brabant	NL	37	29
FR71	Rhône-Alpes	FR	38	37
ES22	Comunidad Foral de Navarra	ES	39	57
AT21	Kärnten	AT	40	52
UKM2	Eastern Scotland	UK	41	38
UKF2	Leicestershire, Rutland and Northamptonshire	UK	42	39
DE26	Unterfranken	DE	43	58
UKE2	North Yorkshire	UK	44	42
SE33	Övre Norrland	SE	45	59
SI02	Zahodna Slovenija	SI	46	48
LU00	Luxembourg	LU	47	41
DK02	Sjælland	DK	48	44
DEB2	Trier	DE	49	49
ITD1	Provincia Autonoma di Bolzano/Bozen	IT	50	36
DK03	Syddanmark	DK	51	46
ITD2	Provincia Autonoma di Trento	IT	52	50
UKJ4	Kent	UK	53	45
SE32	Mellersta Norrland	SE	54	67
UKL2	East Wales	UK	55	56
AT12	Niederösterreich	AT	56	62
UKH3	Essex	UK	57	51
NL22	Gelderland	NL	58	47
UKK2	Dorset and Somerset	UK	59	60

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
FR10	Île de France	FR	60	55
AT32	Salzburg	AT	61	64
FR62	Midi-Pyrénées	FR	62	66
DEB3	Rhein Hessen-Pfalz	DE	63	69
UKI2	Outer London	UK	64	54
ES21	País Vasco	ES	65	83
BE24	Prov. Vlaams-Brabant	BE	66	63
DEA4	Detmold	DE	67	70
NL23	Flevoland	NL	68	43
UKK4	Devon	UK	69	65
DE23	Oberpfalz	DE	70	76
UKN0	Northern Ireland	UK	71	61
UKM3	South Western Scotland	UK	72	75
FR42	Alsace	FR	73	71
DE60	Hamburg	DE	74	68
ES51	Cataluña	ES	75	77
UKG1	Herefordshire, Worcestershire and Warwickshire	UK	76	72
FR24	Centre	FR	77	74
DE30	Berlin	DE	78	85
UKC2	Northumberland and Tyne and Wear	UK	79	81
DEA2	Köln	DE	80	86
UKF1	Derbyshire and Nottinghamshire	UK	81	80
DEF0	Schleswig-Holstein	DE	82	84
ES30	Comunidad de Madrid	ES	83	79
DE22	Niederbayern	DE	84	88
FR51	Pays de la Loire	FR	85	78
FR52	Bretagne	FR	86	87
UKM6	Highlands and Islands	UK	87	89
DED2	Dresden	DE	88	100
FR43	Franche-Comté	FR	89	93
NL21	Overijssel	NL	90	73
NL11	Groningen	NL	91	82
DE72	Gießen	DE	92	105
UKI1	Inner London	UK	93	92
DE73	Kassel	DE	94	110
IE01	Border, Midland and Western	IE	95	97
UKG2	Shropshire and Staffordshire	UK	96	91
FR53	Poitou-Charentes	FR	97	96
AT13	Wien	AT	98	95
DE92	Hannover	DE	99	108
DE24	Oberfranken	DE	100	122
DEG0	Thüringen	DE	101	121
BE21	Prov. Antwerpen	BE	102	98
NL34	Zeeland	NL	103	90

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
FR61	Aquitaine	FR	104	111
CZ01	Praha	CZ	105	99
DE94	Weser-Ems	DE	106	104
UKE4	West Yorkshire	UK	107	94
DE91	Braunschweig	DE	108	132
UKD4	Lancashire	UK	109	102
UKL1	West Wales and The Valleys	UK	110	117
DED3	Leipzig	DE	111	125
BE31	Prov. Brabant Wallon	BE	112	112
UKD3	Greater Manchester	UK	113	103
AT11	Burgenland (AT)	AT	114	115
ITC4	Lombardia	IT	115	109
FR41	Lorraine	FR	116	123
DEA5	Arnsberg	DE	117	130
SI01	Vzhodna Slovenija	SI	118	129
NL12	Friesland (NL)	NL	119	101
DEA3	Münster	DE	120	118
UKF3	Lincolnshire	UK	121	113
FR26	Bourgogne	FR	122	134
CY00	Kýpros	CY	123	107
SK01	Bratislavský kraj	SK	124	106
DEB1	Koblenz	DE	125	133
CZ02	Střední Čechy	CZ	126	126
UKE1	East Yorkshire and Northern Lincolnshire	UK	127	119
NL42	Limburg (NL)	NL	128	116
BE23	Prov. Oost-Vlaanderen	BE	129	114
DE80	Mecklenburg-Vorpommern	DE	130	145
ITD3	Veneto	IT	131	131
DEA1	Düsseldorf	DE	132	137
FR72	Auvergne	FR	133	144
BE22	Prov. Limburg (BE)	BE	134	120
CZ06	Jihovýchod	CZ	135	141
ITD5	Emilia-Romagna	IT	136	138
DED1	Chemnitz	DE	137	151
BE10	Région de Bruxelles-Capitale	BE	138	140
ITE1	Toscana	IT	139	149
DECO	Saarland	DE	140	146
ITC3	Liguria	IT	141	156
FR21	Champagne-Ardenne	FR	142	139
ITD4	Friuli-Venezia Giulia	IT	143	153
FR82	Provence-Alpes-Côte d'Azur	FR	144	143
UKE3	South Yorkshire	UK	145	128
PL12	Mazowieckie	PL	146	148
NL13	Drenthe	NL	147	127

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
ES70	Canarias	ES	148	142
EE00	Eesti	EE	149	168
UKC1	Tees Valley and Durham	UK	150	135
DE93	Lüneburg	DE	151	154
PT30	Região Autónoma da Madeira	PT	152	124
FR25	Basse-Normandie	FR	153	147
ES24	Aragón	ES	154	171
CZ03	Jihozápad	CZ	155	158
ITE3	Marche	IT	156	157
FR63	Limousin	FR	157	155
ES53	Illes Balears	ES	158	136
BE34	Prov. Luxembourg (BE)	BE	159	159
CZ05	Severovýchod	CZ	160	164
DE50	Bremen	DE	161	152
FR81	Languedoc-Roussillon	FR	162	160
UKG3	West Midlands	UK	163	150
DEE0	Sachsen-Anhalt	DE	164	170
ITE2	Umbria	IT	165	173
UKD1	Cumbria	UK	166	165
ITC2	Valle d'Aosta/Vallée d'Aoste	IT	167	162
FR22	Picardie	FR	168	166
ES62	Región de Murcia	ES	169	163
ITC1	Piemonte	IT	170	175
FR23	Haute-Normandie	FR	171	169
ES52	Comunidad Valenciana	ES	172	172
BE25	Prov. West-Vlaanderen	BE	173	161
ITE4	Lazio	IT	174	177
LT00	Lietuva	LT	175	179
UKK3	Cornwall and Isles of Scilly	UK	176	176
PT17	Lisboa	PT	177	167
FR92	Martinique	FR	178	178
FR94	Réunion	FR	179	181
FR30	Nord - Pas-de-Calais	FR	180	174
ES41	Castilla y León	ES	181	196
ES61	Andalucía	ES	182	182
ES23	La Rioja	ES	183	184
ES11	Galicia	ES	184	195
PL63	Pomorskie	PL	185	180
BE33	Prov. Liège	BE	186	187
PT16	Centro (PT)	PT	187	183
FR93	Guyane	FR	188	204
EL30	Attiki	EL	189	192
ES42	Castilla-La Mancha	ES	190	197
FR91	Guadeloupe	FR	191	201

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
ES13	Cantabria	ES	192	199
CZ08	Moravskoslezsko	CZ	193	191
ES12	Principado de Asturias	ES	194	206
CZ07	Střední Morava	CZ	195	194
ES64	Ciudad Autónoma de Melilla	ES	196	186
PL41	Wielkopolskie	PL	197	193
FR83	Corse	FR	198	188
BE35	Prov. Namur	BE	199	202
BG41	Yugozapaden	BG	200	185
EL42	Notio Aigaio	EL	201	203
HU10	Közép-Magyarország	HU	202	198
PT11	Norte	PT	203	200
PT20	Região Autónoma dos Açores	PT	204	189
RO32	București - Ilfov	RO	205	190
PL21	Małopolskie	PL	206	205
LV00	Latvija	LV	207	212
PL43	Lubuskie	PL	208	210
EL43	Kriti	EL	209	209
ITG2	Sardegna	IT	210	214
PL31	Lubelskie	PL	211	211
SK04	Východné Slovensko	SK	212	213
PL11	Łódzkie	PL	213	215
PT15	Algarve	PT	214	207
PL51	Dolnośląskie	PL	215	218
PL32	Podkarpackie	PL	216	219
PL62	Warmińsko-Mazurskie	PL	217	217
ITF2	Molise	IT	218	227
PL42	Zachodniopomorskie	PL	219	220
ES63	Ciudad Autónoma de Ceuta	ES	220	216
MT00	Malta	MT	221	208
ITF1	Abruzzo	IT	222	225
SK03	Stredné Slovensko	SK	223	226
PL34	Podlaskie	PL	224	222
EL12	Kentriki Makedonia	EL	225	233
PL61	Kujawsko-Pomorskie	PL	226	224
EL24	Sterea Ellada	EL	227	240
ITF6	Calabria	IT	228	229
PT18	Alentejo	PT	229	234
EL14	Thessalia	EL	230	239
ITF3	Campania	IT	231	228
EL21	Ipeiros	EL	232	243
ES43	Extremadura	ES	233	241
HU22	Nyugat-Dunántúl	HU	234	232
EL13	Dytiki Makedonia	EL	235	245

CODE	LABEL	COUNTRY	BEN_NONPOND	BEN_POND
CZ04	Severozápad	CZ	236	231
BE32	Prov. Hainaut	BE	237	230
EL23	Dytiki Ellada	EL	238	244
SK02	Západné Slovensko	SK	239	235
RO11	Nord-Vest	RO	240	223
PL52	Opolskie	PL	241	236
PL22	Łódzkie	PL	242	237
ITF4	Puglia	IT	243	238
RO21	Nord-Est	RO	244	221
PL33	Świętokrzyskie	PL	245	242
EL22	Ionia Nisia	EL	246	247
ITG1	Sicilia	IT	247	248
HU31	Észak-Magyarország	HU	248	249
EL41	Voreio Aigaio	EL	249	256
RO41	Sud-Vest Oltenia	RO	250	246
HU23	Dél-Dunántúl	HU	251	252
HU21	Közép-Dunántúl	HU	252	251
EL25	Peloponnisos	EL	253	262
BG31	Severozapaden	BG	254	257
HU33	Dél-Alföld	HU	255	259
HU32	Észak-Alföld	HU	256	258
BG34	Yugoiztochen	BG	257	254
RO42	Vest	RO	258	250
BG42	Yuzhen tsentralen	BG	259	255
ITF5	Basilicata	IT	260	263
BG33	Severoiztochen	BG	261	253
RO12	Centru	RO	262	260
RO22	Sud-Est	RO	263	261
EL11	Anatoliki Makedonia, Thraki	EL	264	266
BG32	Severen tsentralen	BG	265	264
RO31	Sud - Muntenia	RO	266	265

Dividendo entrambe le distribuzioni in quintili, possiamo osservarne la sostanziale coincidenza:

Tavola di contingenza Percentile Group of M_BEN_NONP * Percentile Group of M_BEN_POND

Conteggio

		Percentile Group of M_BEN_POND					Totale
		1	2	3	4	5	
Percentile Group of M_BEN_NONP	1	51	2	0	0	0	53
	2	2	48	3	0	0	53
	3	0	3	46	5	0	54
	4	0	0	5	45	3	53
	5	0	0	0	3	50	53
Totale		53	53	54	53	53	266

Misure simmetriche

		Valore	E.S. asint. ^a	T appross. ^b	Sig. appross.
Ordinale per ordinale	Tau-b di Kendall	,952	,009	105,579	,000
N. di casi validi		266			

a. Senza assumere l'ipotesi nulla.

b. Viene usato l'errore standard asintotico in base all'assunzione dell'ipotesi nulla.

L'elevato valore del tau-b di Kendall ci segnala che la corrispondenza tra le due graduatorie è quasi perfetta.

Anche le immagini prodotte con GvSig confermano la sostanziale uguaglianza di risultati.

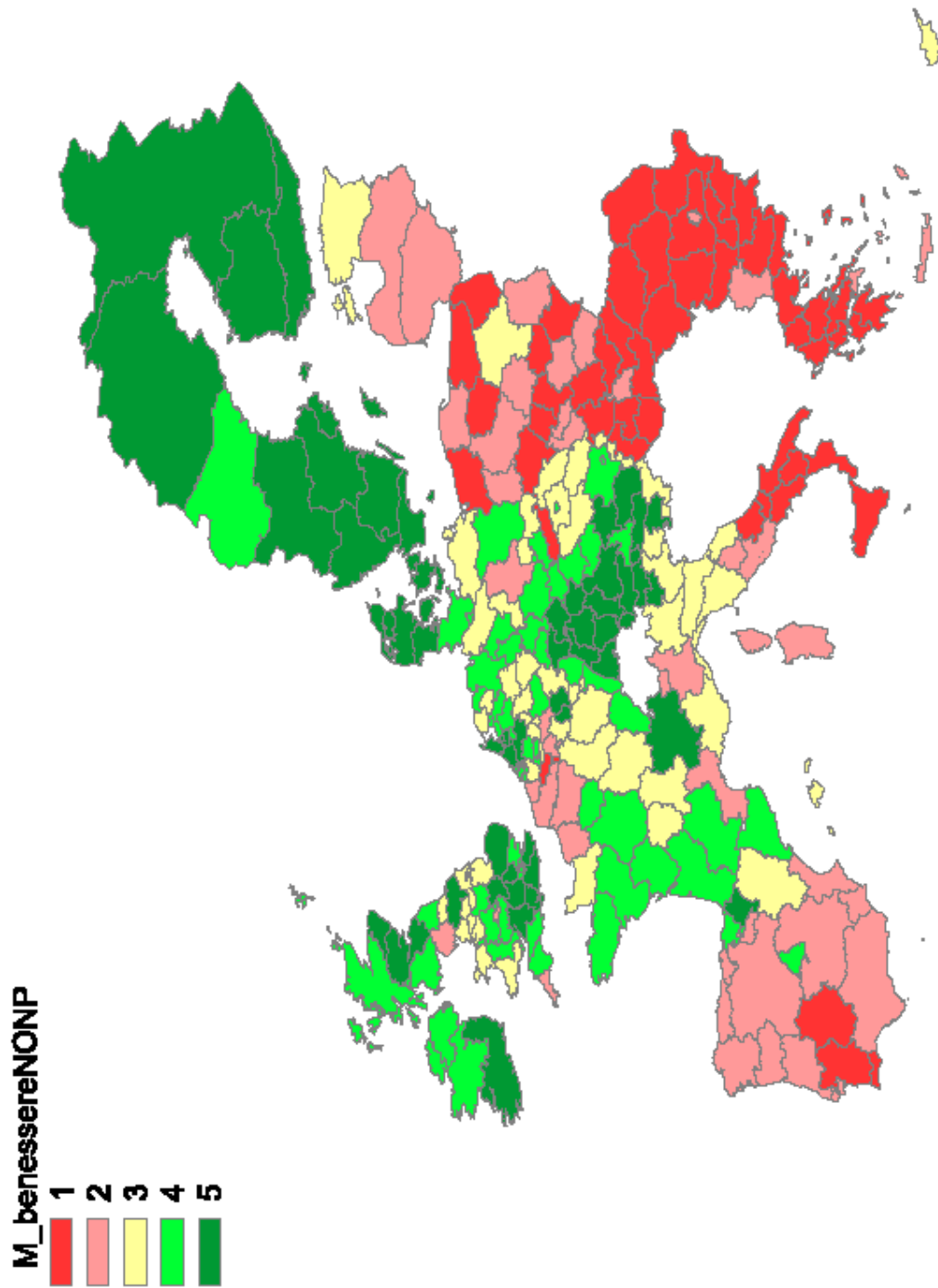


Figura 44. Media non ponderata benessere globale

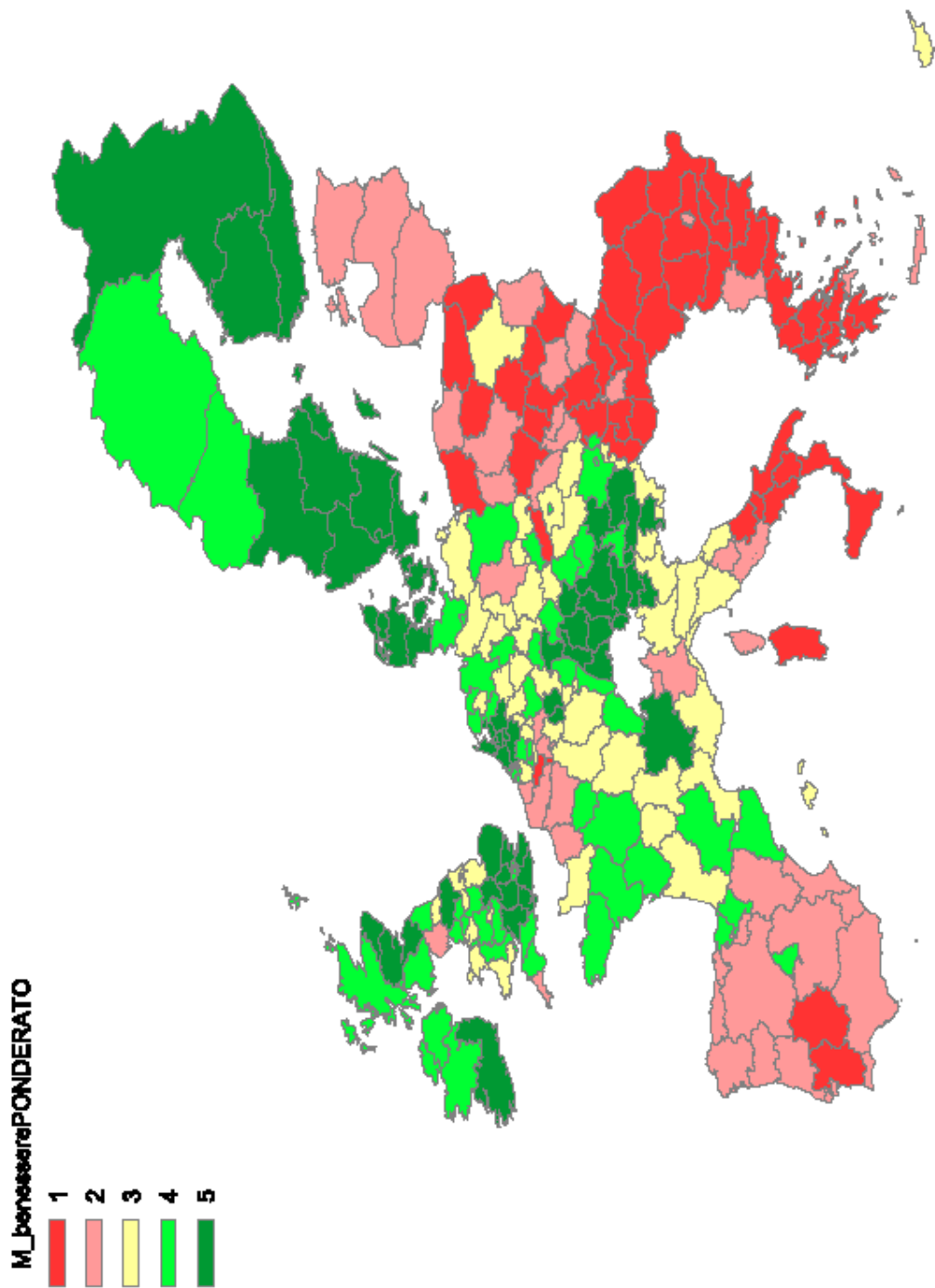


Figura 45. Media ponderata benessere globale

4.4 SINTESI TRAMITE ANALISI DELLE COMPONENTI PRINCIPALI

In secondo luogo, per sintetizzare ulteriormente il fenomeno in oggetto, calcoliamo la matrice di correlazione tra le medie di ciascun settore di benessere ed esaminiamo la possibilità di ricorrere all'Analisi delle Componenti Principali.

Correlazioni

		M2ISTR	M2SALUTE	M2RICC	M2DEMO	M2LAV	M2AMB
M2ISTR	Correlazione di Pearson	1	,511**	,707**	,082	,696**	-,183**
	Sig. (2-code)		,000	,000	,180	,000	,003
	N	266	266	266	266	266	266
M2SALUTE	Correlazione di Pearson	,511**	1	,662**	-,009	,270**	-,240**
	Sig. (2-code)	,000		,000	,882	,000	,000
	N	266	266	266	266	266	266
M2RICC	Correlazione di Pearson	,707**	,662**	1	,028	,531**	-,463**
	Sig. (2-code)	,000	,000		,653	,000	,000
	N	266	266	266	266	266	266
M2DEMO	Correlazione di Pearson	,082	-,009	,028	1	,005	-,283**
	Sig. (2-code)	,180	,882	,653		,934	,000
	N	266	266	266	266	266	266
M2LAV	Correlazione di Pearson	,696**	,270**	,531**	,005	1	-,179**
	Sig. (2-code)	,000	,000	,000	,934		,003
	N	266	266	266	266	266	266
M2AMB	Correlazione di Pearson	-,183**	-,240**	-,463**	-,283**	-,179**	1
	Sig. (2-code)	,003	,000	,000	,000	,003	
	N	266	266	266	266	266	266

** La correlazione è significativa al livello 0,01 (2-code).

Evidenziamo l'esistenza di elevate relazioni lineari dirette tra le coppie “media istruzione-media ricchezza” (indice di correlazione pari a 0,71), “media istruzione e media lavoro” ($r=0,69$) e pure “media salute e media ricchezza” ($r=0,66$).

Nonostante ciò non esistono legami di natura lineare particolarmente forti tra le rimanenti coppie di variabili e pertanto l'analisi delle componenti principali eseguita con queste modalità non risulta sufficientemente valida: come ci mostra l'output di Spss la quota di varianza spiegata dalle prime due componenti estratte, pari al 67,9% è inferiore al valore soglia, uguale a 73,5%.

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2,876	47,938	47,938	2,876	47,938	47,938
2	1,200	19,992	67,930	1,200	19,992	67,930
3	,836	13,941	81,871			
4	,646	10,764	92,634			
5	,260	4,337	96,971			
6	,182	3,029	100,000			

Metodo di estrazione: Analisi componenti principali.

Matrice di componenti^a

	Componente	
	1	2
Punteggio Z(M2ISTR)	,862	,187
Punteggio Z(M2SALUTE)	,723	,103
Punteggio Z(M2RICC)	,909	,014
Punteggio Z(M2DEMO)	,123	-,830
Punteggio Z(M2LAV)	,728	,231
Punteggio Z(M2AMB)	-,489	,642

Metodo estrazione: analisi componenti principali.

a. 2 componenti estratti

Si ripete pertanto l'Analisi delle Componenti Principali inserendo la totalità degli indicatori di benessere, cioè i trenta indicatori originari standardizzati .

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	6,905	23,018	23,018	6,905	23,018	23,018
2	4,915	16,383	39,402	4,915	16,383	39,402
3	3,511	11,702	51,104	3,511	11,702	51,104
4	2,358	7,859	58,963	2,358	7,859	58,963
5	1,623	5,409	64,371	1,623	5,409	64,371
6	1,589	5,295	69,667	1,589	5,295	69,667
7	1,430	4,766	74,433	1,430	4,766	74,433
8	1,224	4,081	78,514	1,224	4,081	78,514
9	,901	3,004	81,518			
10	,753	2,509	84,027			
11	,630	2,098	86,126			
12	,576	1,919	88,045			
13	,556	1,852	89,897			
14	,433	1,442	91,339			
15	,388	1,293	92,632			
16	,362	1,207	93,840			
17	,299	,996	94,836			
18	,259	,862	95,698			
19	,238	,792	96,490			
20	,211	,703	97,193			
21	,191	,636	97,829			
22	,143	,478	98,307			
23	,135	,451	98,758			
24	,111	,369	99,127			
25	,076	,253	99,380			
26	,066	,218	99,599			
27	,049	,163	99,762			
28	,037	,122	99,884			
29	,027	,090	99,974			
30	,008	,026	100,000			

Metodo di estrazione: Analisi componenti principali.

Le componenti Principali aventi autovalore maggiore di uno sono otto, ma in tal caso il criterio di sintesi non sarebbe soddisfatto e pertanto si esegue l'analisi inserendo l'opzione di estrazione delle prime due Componenti Principali, in maniera tale da considerare lo spazio bidimensionale.

Osservando la percentuale di varianza spiegata dalle prime due CP notiamo che esse superano il valore soglia pari a 21,46. Tuttavia dalla Tabella Comunalità apprendiamo che la percentuale di varianza spiegata in riferimento ad una pluralità di indicatori è molto bassa; in particolare non risultano pressoché spiegate le variabili “veicoli posseduti”, “rifiuti procapite prodotti”, “tasso movimento migratorio”.

Comunalità

	Iniziale	Estrazione
Z: SM(upper_sec)	1,000	,220
Z: SM(terz_edu)	1,000	,540
Z: SM(R_d_gdp)	1,000	,331
Z: SM(formaz_ad)	1,000	,493
Z: SM(patent_app)	1,000	,358
Z: SM(asp_vita)	1,000	,425
Z_MORT_INF	1,000	,253
Z_MORTI_INC	1,000	,484
Z: pil ppt pc (% media)	1,000	,587
Z: SM (reddito_disponibile_ netto)	1,000	,610
Z: SM(vehicle1000)	1,000	,017
Z: %ore non lavorate	1,000	,495
Z_ALTOIMP_AMB	1,000	,185
Z_SERV_RES	1,000	,371
Z_RIF_PERS	1,000	,012
Z_DENS	1,000	,269
Z: SM(foreste)	1,000	,209
Z_IND_VECC	1,000	,615
Z_IND_DIP	1,000	,204
Z_OVER74	1,000	,371
Z: tra 0 e 9	1,000	,758
Z: tasso fertilità totale	1,000	,544
Z: tasso movimento naturale pop	1,000	,794
Z: tasso movimento migratorio	1,000	,095
Z: t occupaz tot	1,000	,752
Z_DIFF_OCC_MF	1,000	,148
Z_UNEMP_R	1,000	,526
Z_LONG_UNEMP_R	1,000	,489
Z_DIFF_UNEMP_FM	1,000	,119
Z_DIFF_UNEMP_GA	1,000	,547

Metodo di estrazione: Analisi componenti principali.

Pertanto si riesegue l'analisi conservando solo gli indicatori la cui quota di varianza spiegata supera il 40%. Si escludono quindi le variabili: "istruzione secondaria", "spesa in ricerca e sviluppo", "brevetti riconosciuti", "mortalità infantile", "veicoli posseduti", "indice di dipendenza", "maggiori di 74 anni", "movimento migratorio", "differenze occupazione tra i generi", "differenze disoccupazione tra i generi". A queste si aggiungono tutte le variabili ambientali: densità abitativa, suolo coperto da foreste, suolo con attività ad alto impatto ambientale, suolo occupato da residenze, attività commerciali e industriali, rifiuti prodotti.

Nella costruzione di un indicatore composito è raccomandata la considerazione di tutti gli aspetti del fenomeno multidimensionale analizzato (ciascun aspetto viene misurato da diversi gruppi di variabili) e pertanto affinché tale indicatore possa considerare tutte le varie facce del fenomeno, esso dovrà contenere almeno una variabile per ogni gruppo. Alla luce di tali considerazioni risulta impensabile l'esclusione dell'intero ambito ambientale e pertanto conserviamo la variabile ambientale avente la maggiore quota di varianza spiegata, che risulta essere la "percentuale di suolo occupato da residenze e attività commerciali ed industriali" (che può essere considerata una variabile *proxy* del grado di cementificazione).

Dopo questa serie di operazioni il database rimane composto da 16 variabili; eseguendo l'analisi di affidabilità con tali indicatori otteniamo un miglioramento dell'Alpha di Cronbach, che raggiunge il valore di 0,844.

Ritornando all'ACP si procede con l'estrazione di tutte le componenti il cui autovalore è maggiore di uno.

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	6,000	37,497	37,497	6,000	37,497	37,497
2	3,230	20,187	57,684	3,230	20,187	57,684
3	1,810	11,311	68,995	1,810	11,311	68,995
4	1,062	6,635	75,631	1,062	6,635	75,631
5	,909	5,681	81,312			
6	,697	4,358	85,670			
7	,513	3,208	88,878			
8	,497	3,107	91,985			
9	,337	2,105	94,090			
10	,294	1,836	95,926			
11	,230	1,435	97,362			
12	,164	1,028	98,389			
13	,115	,719	99,108			
14	,068	,426	99,534			
15	,040	,252	99,786			
16	,034	,214	100,000			

Metodo di estrazione: Analisi componenti principali.

La percentuale di varianza totale spiegata dalle quattro componenti estratte è pari al 75,6%, e pure la quota di varianza spiegata dalle tre CP per ciascuna variabile risulta sempre superiore allo 0,5.

Comunalità

	Iniziale	Estrazione
Z: SM(terz_edu)	1,000	,569
Z: SM(formaz_ad)	1,000	,543
Z: SM(asp_vita)	1,000	,873
Z_MORTI_INC	1,000	,694
Z: pil ppt pc (% media)	1,000	,660
Z: SM (reddito_disponibile_ netto)	1,000	,814
Z: %ore non lavorate	1,000	,587
Z_SERV_RES	1,000	,640
Z_IND_VECC	1,000	,879
Z: tra 0 e 9	1,000	,934
Z: tasso fertilità totale	1,000	,742
Z: tasso movimento naturale pop	1,000	,843
Z: t occupaz tot	1,000	,843
Z_UNEMP_R	1,000	,857
Z_LONG_UNEMP_R	1,000	,854
Z_DIFF_UNEMP_GA	1,000	,769

Metodo di estrazione: Analisi componenti principali.

Matrice di componenti^a

	Componente			
	1	2	3	4
Z: SM(terz_edu)	,677	,235	-,222	-,083
Z: SM(formaz_ad)	,719	,113	,027	-,111
Z: SM(asp_vita)	,586	-,108	-,520	,498
Z_MORTI_INC	,667	-,021	-,348	-,356
Z: pil ppt pc (% media)	,735	-,081	-,335	-,011
Z: SM (reddito_disponibile_ netto)	,697	-,297	-,415	,260
Z: %ore non lavorate	,709	-,162	-,019	-,240
Z_SERV_RES	-,456	-,273	,212	,559
Z_IND_VECC	,234	,772	,465	-,104
Z: tra 0 e 9	,475	,793	,256	,115
Z: tasso fertilità totale	,634	,469	,198	,285
Z: tasso movimento naturale pop	,543	,729	,016	,132
Z: t occupaz tot	,780	-,427	,225	,039
Z_UNEMP_R	,540	-,530	,524	,093
Z_LONG_UNEMP_R	,618	-,395	,539	,162
Z_DIFF_UNEMP_GA	,491	-,598	,318	-,264

Metodo estrazione: analisi componenti principali.

a. 4 componenti estratti

Tuttavia se considerassimo solo le prime due componenti la percentuale di varianza totale spiegata sarebbe comunque sufficiente (dato che il valore soglia è pari al 44%) e consentirebbe la rappresentazione dei risultati nel piano cartesiano. Privilegiando il criterio di sintesi, si sceglie di continuare l'analisi considerando lo spazio bidimensionale.

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	6,000	37,497	37,497	6,000	37,497	37,497
2	3,230	20,187	57,684	3,230	20,187	57,684
3	1,810	11,311	68,995			
4	1,062	6,635	75,631			
5	,909	5,681	81,312			
6	,697	4,358	85,670			
7	,513	3,208	88,878			
8	,497	3,107	91,985			
9	,337	2,105	94,090			
10	,294	1,836	95,926			
11	,230	1,435	97,362			
12	,164	1,028	98,389			
13	,115	,719	99,108			
14	,068	,426	99,534			
15	,040	,252	99,786			
16	,034	,214	100,000			

Metodo di estrazione: Analisi componenti principali.

Come ci mostra la tabella Comunalità, le variabili spiegate peggio rimangono quella ambientale “% suolo occupato da residenze, attività commerciali e industriali” e quelle relative all’ambito salute, cioè “aspettativa di vita” e “tasso di mortalità per incidenti”, la cui percentuale di varianza spiegata non raggiunge il 50, ma si arrestano rispettivamente al 28,2%, 35,5% e 44,5%; all’estremo opposto troviamo le variabili “percentuale di under 10” e “tasso di movimento naturale della popolazione” e le cui quote superano lo 0,8.

Comunalità

	Iniziale	Estrazione
Z: SM(terz_edu)	1,000	,513
Z: SM(formaz_ad)	1,000	,530
Z: SM(asp_vita)	1,000	,355
Z_MORTI_INC	1,000	,445
Z: pil ppt pc (% media)	1,000	,547
Z: SM (reddito_disponibile_ netto)	1,000	,574
Z: %ore non lavorate	1,000	,529
Z_SERV_RES	1,000	,282
Z_IND_VECC	1,000	,652
Z: tra 0 e 9	1,000	,855
Z: tasso fertilità totale	1,000	,622
Z: tasso movimento naturale pop	1,000	,826
Z: t occupaz tot	1,000	,791
Z_UNEMP_R	1,000	,573
Z_LONG_UNEMP_R	1,000	,537
Z_DIFF_UNEMP_GA	1,000	,599

Metodo di estrazione: Analisi componenti principali.

Dalla Matrice di componenti possiamo desumere l'interpretazione delle due Componenti principali estratte: la prima costituisce un indicatore sintetico diretto di occupazione (coefficiente di correlazione lineare pari a 0,78), ricchezza economica e tempo libero (l'indice r è di circa 0,7 con le variabili Pil pro-capite, reddito netto familiare e percentuale ore non dedicate al lavoro), di istruzione (percentuale di laureati e di partecipanti a formazione continua) e longevità (aspettativa di vita, inverso di mortalità causa incidente);

la seconda componente costituisce un indicatore sintetico della situazione demografica ed in particolare di giovinezza della stessa (% under 10, inverso dell'indice di vecchiaia) e di rinnovamento (tasso di movimento naturale).

Matrice di componenti^a

	Componente	
	1	2
Z: SM(terz_edu)	,677	,235
Z: SM(formaz_ad)	,719	,113
Z: SM(asp_vita)	,586	-,108
Z_MORTI_INC	,667	-,021
Z: pil ppt pc (% media)	,735	-,081
Z: SM (reddito_disponibile_ netto)	,697	-,297
Z: %ore non lavorate	,709	-,162
Z_SERV_RES	-,456	-,273
Z_IND_VECC	,234	,772
Z: tra 0 e 9	,475	,793
Z: tasso fertilità totale	,634	,469
Z: tasso movimento naturale pop	,543	,729
Z: t occupaz tot	,780	-,427
Z_UNEMP_R	,540	-,530
Z_LONG_UNEMP_R	,618	-,395
Z_DIFF_UNEMP_GA	,491	-,598

Metodo estrazione: analisi componenti principali.

a. 2 componenti estratti

Discorso a parte meritano le ultime tre variabili, tutte invertite e relative alla condizione di disoccupazione: tasso di disoccupazione, tasso di disoccupazione di lungo periodo e differenza nei tassi di disoccupazione tra giovani e adulti; esse vengono spiegate in parte dalla prima componente, con cui stabiliscono una relazione lineare diretta, ed in parte dalla seconda, instaurando una relazione lineare inversa.

Il Grafico componenti (detto grafico dei pesi fattoriali in Spss) rappresenta i punti variabile e visualizza in un unico grafico tutte le informazioni contenute nelle tabelle “Matrice dei componenti” e “Comunalità”: la lunghezza di ciascun vettore-variabile è proporzionale alla rispettiva quota di varianza spiegata complessivamente dalle prime due Componenti Principali; l’angolo tra ciascun vettore ed ognuno dei due assi mostra la correlazione tra la variabile in questione e rispettivamente la prima e la seconda CP.

In aggiunta a ciò, l'angolo tra due vettori segnala la correlazione esistente tra due generiche variabili: l'angolo esistente tra il vettore riferito al "Pil a parità di potere d'acquisto" e quello relativo alla "Formazione adulta permanente" è molto piccolo e quindi indica l'esistenza di un'elevata correlazione diretta tra le due variabili; l'angolo formato dai vettori "% under 10" e "Reddito familiare disponibile" essendo prossimo a 90° segnala invece l'esistenza di incorrelazione tra le variabili.

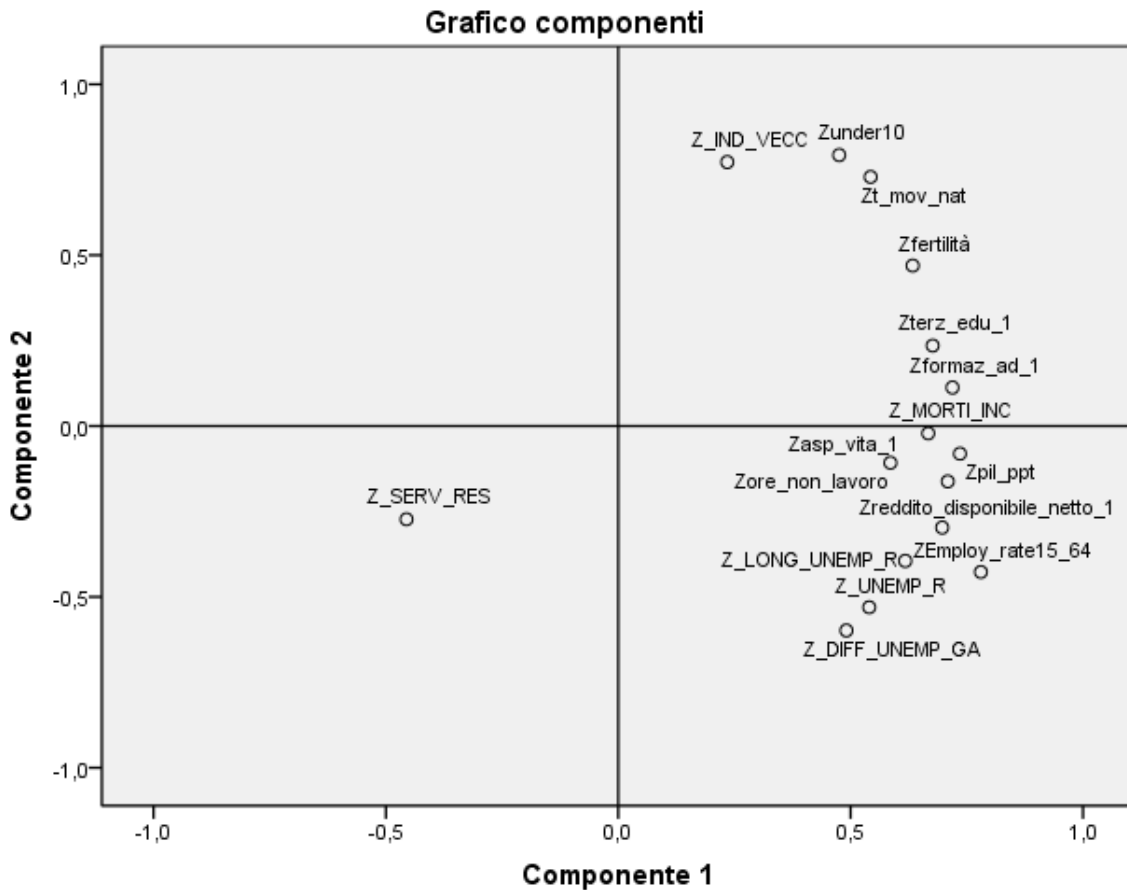


Figura 46. Grafico dei pesi fattoriali delle prime due componenti principali

Il diagramma di dispersione degli *scores* delle prime due Componenti principali, rappresenta sull'asse cartesiano i punti-unità: punti vicini all'origine corrispondono a regioni che presentano valori dei 16 indicatori di benessere prossimi alle rispettive medie; al contrario punti molto lontani dal centroide costituiscono dei valori anomali. Sono outlier relativamente alla seconda componente le isole di Guyane (unità 139), Reunion (140), Martinique (138) e Guadalupe (137) appartenenti alla Francia e le città autonome spagnole (ma situate sulla costa marocchina) di Melilla (109) e Ceuta (108).

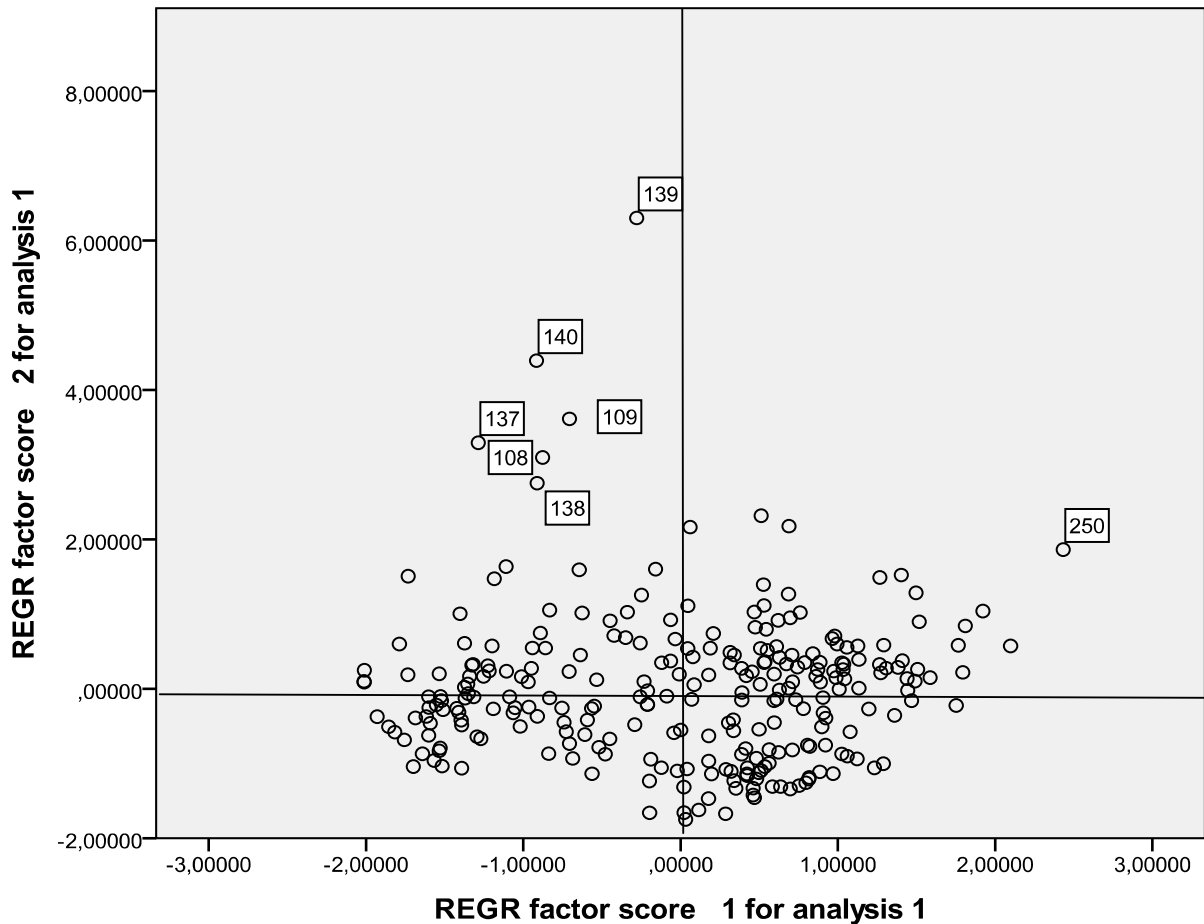


Figura 47. Diagramma di dispersione degli scores con 266 regioni

Cancellate le unità anomale otteniamo un diagramma di dispersione più chiaramente leggibile.

Onde evitare distorsioni nelle analisi successive, scegliamo di eliminare tali regioni non appartenenti al continente europeo e pertanto i metodi statistici di classificazione delle unità che vengono presentati nei prossimi paragrafi (analisi dei gruppi e teoria degli insiemi sfocati) si basano sulla nuova matrice dei dati avente dimensioni 260 per 16, in quanto 260 sono le regioni e 16 le variabili di benessere considerate.

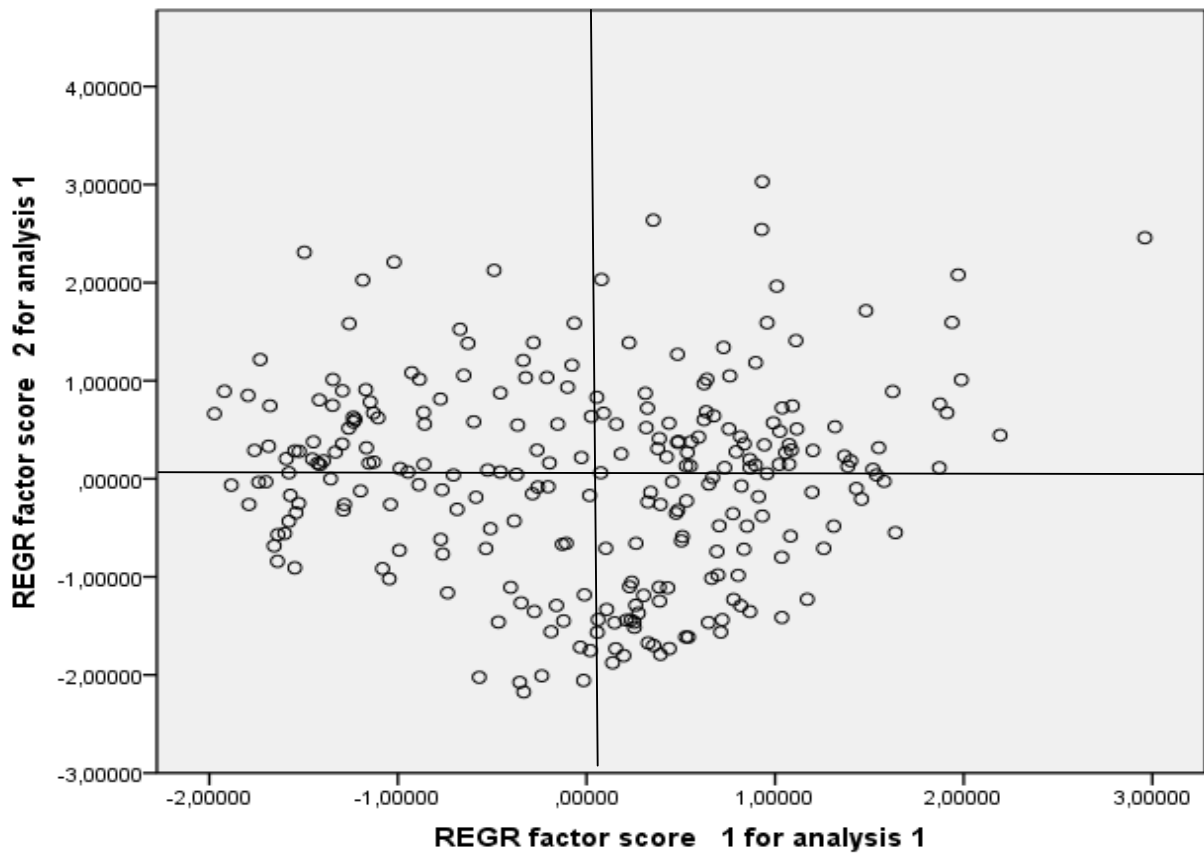


Figura 48. Diagramma di dispersione degli *scores* con 260 regioni

4.4 INDIVIDUAZIONE DI GRUPPI OMOGENEI: CLUSTER ANALYSIS

L'obiettivo della *Cluster Analysis* è quello di classificare le unità in gruppi omogenei, consentendo in tal modo una più semplice interpretazione dei dati.

L'analisi dei gruppi rientra tra i metodi esplorativi; pertanto l'effettiva esistenza di tali gruppi nel *data set* va verificata empiricamente e i cluster devono apparire con naturalezza nelle osservazioni. Tali gruppi di unità dovranno possedere contemporaneamente caratteristiche di coesione interna e separazione esterna. Contemporaneamente dovrà essere soddisfatto pure il criterio di sintesi soddisfacente delle n unità di partenza in g gruppi ($g \ll n$).

Dallo scopo che il ricercatore si propone di raggiungere mediante la classificazione delle unità derivano le scelte da effettuare per condurre l'analisi.

Prima tra tutte, la scelta delle variabili è quella che, lasciando ampi margini di soggettività, condiziona fortemente i risultati dell'analisi.

Nel caso degli indicatori sul benessere si considerano le 16 variabili standardizzate sopra descritte e si ritiene opportuno utilizzare il metodo di formazione dei gruppi non gerarchico, in particolare quello delle k -medie, poiché il numero di regioni, cioè le unità statistiche da classificare, è elevato. Infatti, la lettura del dendrogramma, cioè dell'output grafico che scaturisce dalla cluster applicando il metodo gerarchico di formazione dei gruppi, risulterebbe di difficile lettura.

La classificazione non gerarchica si sostanzia in un problema di ottimizzazione di tipo iterativo, in quanto si ricerca la partizione, con numero di gruppi g prefissato, cui corrisponde la massima coesione all'interno dei gruppi, definita come devianza nei gruppi. La metrica utilizzata è quella euclidea, in quanto garantisce la convergenza della procedura iterativa.

I due principali inconvenienti legati alla scelta del metodo non gerarchico riguardano la necessità di dover definire a priori il numero di gruppi che si vuole ottenere dall'analisi ed i poli della partizione iniziale. Riguardo al primo aspetto il numero ottimo di gruppi deve essere tale da produrre una classificazione che sia contemporaneamente sintetica e contenente gruppi di unità dotati di coesione interna. Tale numero g viene individuato per tentativi, iterando la procedura per diversi valori di g , valutando ogni volta la bontà della partizione ottenuta tramite l'indice R^2 e scegliendo la classificazione più soddisfacente.

Tale indice è basato sulla scomposizione della devianza totale delle p variabili:

$$R^2 = B/T = 1 - W/T$$

esprime il rapporto tra la devianza fra i gruppi (B) e la devianza totale (T) e varia nell'intervallo [0,1]. La devianza totale è data dalla somma tra la devianza fra i gruppi (B) e la devianza entro i gruppi (W). Esso pertanto misura la quota di variabilità totale nella matrice dei dati che può essere spiegata dalla partizione considerata. Ovviamente il valore di R^2 cresce all'aumentare di g (dato che i gruppi sono più omogenei), quindi occorre osservare la misura dell'incremento del nostro indice: se l'aumento di R^2 passando da g a $(g+1)$ è modesto, si sceglie la partizione più sintetica formata da g gruppi.

Procediamo con il calcolo dell'indice di R^2 applicando la procedura k -medie (indicando però con g il numero di gruppi¹²) al nostro database avente dimensioni 260 per 16 con diversi valori di g .

g	R²
2	0,275
3	0,367
4	0,429
5	0,506
6	0,528

Considerando la partizione con 3 o 4 gruppi l'aumento dell'indice R^2 è tale da giustificare il passaggio a partizioni maggiormente disaggregate; al contrario alla classificazione con 6 gruppi corrisponde un miglioramento dell'indice non rilevante e pertanto appare ragionevole scegliere la partizione con 5 cluster, anche se spiega poco più del 50% della variabilità totale della matrice di partenza (soglia minima per considerare valida l'analisi), soprattutto per poter confrontare tale classificazione con quelle degli indicatori compositi prima ottenuti con l'analisi delle componenti principali, i cui valori sono stati suddivisi in base ai quintili.

La Tabella ci mostra la numerosità che caratterizza ciascun gruppo. Osserviamo immediatamente che il quarto gruppo è formato solamente da 3 unità e pertanto è chiaro che tali regioni costituiscono delle unità anomale, che è necessario rimuovere per evitare distorsioni nei risultati.

¹² Tale algoritmo dovrebbe pertanto essere chiamato più coerentemente come "metodo delle g -medie".

Numero di casi in ogni cluster

Cluster	1	52,000
	2	59,000
	3	81,000
	4	3,000
	5	65,000
Validi		260,000
Mancanti		,000

Come si evince dalla Tabella seguente, in termini di scostamenti standardizzati, le regioni anomale (cluster 4) coincidono con le capitali Bruxelles (unità 10) e Londra (Inner London unità 244, Outer London unità 245), esse presentano valori eccezionalmente elevati riguardo agli indicatori economici (Pil procapite e Reddito familiare), relativi all'istruzione (% laureati e adulti in formazione continua), alla % di suolo cementificato e alla giovinezza della popolazione.

Centri dei cluster finali

	Cluster				
	1	2	3	4	5
Z: SM(terz_edu)	-,26041	-,76183	,86748	2,33088	-,27121
Z: SM(formaz_ad)	-,51943	-,79177	1,07495	1,33119	-,27194
Z: SM(asp_vita)	,49540	-1,53164	,42740	,47183	,44226
Z_MORTI_INC	-,1304	-1,1377	,5802	1,2679	,3512
Z: pil ppt pc (% media)	-,21258	-1,03307	,39792	3,32368	,51757
Z: SM (reddito_disponibile_netto)	-,07219	-1,37730	,32033	1,56121	,83442
Z: %ore non lavorate	-,42564	-1,07798	,72431	-,07250	,36580
Z_SERV_RES	,2320	,3859	-,3182	-5,3451	,1073
Z_IND_VECC	-,0662	,2508	,4631	1,4546	-,9912
Z: tra 0 e 9	,04656	-,34385	,57600	1,59632	-,81072
Z: tasso fertilità totale	-,18803	-,71157	,87546	1,06941	-,59635
Z: tasso movimento naturale pop	,17059	-,54780	,45579	2,38072	-,60849
Z: t occupaz tot	-,91555	-,67618	,72206	-,22498	,67056
Z_UNEMP_R	-,9593	-,1941	,4687	-,5197	,6615
Z_LONG_UNEMP_R	-,7664	-,1689	,6056	-,5215	,4064
Z_DIFF_UNEMP_GA	-,9047	-,2953	,2139	-,7187	,9963

Dal grafico degli *scores* possiamo osservare come tali unità si distinguono nettamente dalle rimanenti.

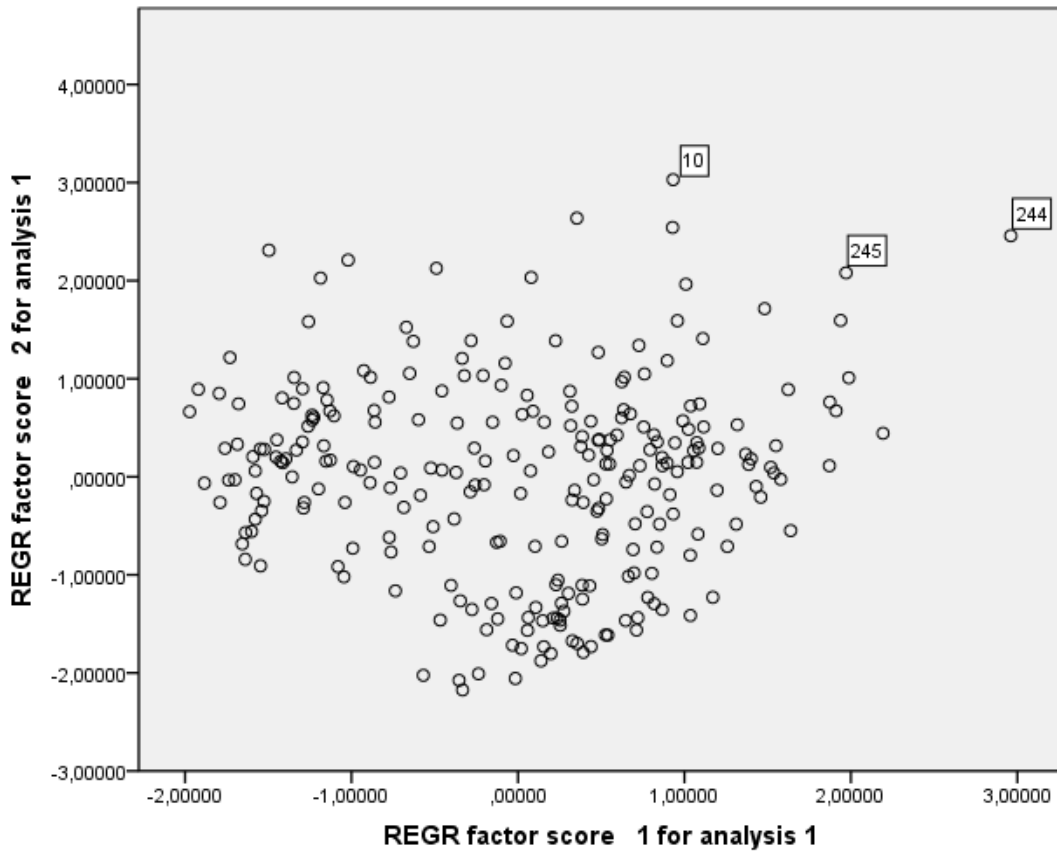


Figura 49. Diagramma di dispersione degli scores con tre outliers

Ripetendo l'analisi delle k-medie sul database 257x16 e fissando il numero di gruppi pari a cinque, otteniamo un discreto miglioramento dell'indice R^2 , che raggiunge quota 0,53. Non dobbiamo però dimenticare che la variabilità all'interno dei gruppi (l'errore) è pari al 47% della devianza totale.

La tavola di analisi della varianza fornisce i dati necessari per il calcolo dell'indice R^2 per ciascuna variabile: moltiplicando le varianze per i rispettivi gradi di libertà ottengo le devianze.

ANOVA

	Cluster		Errore		F	Sig.
	Media dei quadrati	df	Media dei quadrati	df		
Z: SM(terz_edu)	25,822	4	,569	252	45,389	,000
Z: SM(formaz_ad)	49,083	4	,237	252	207,122	,000
Z: SM(asp_vita)	45,690	4	,317	252	143,966	,000
Z_MORTI_INC	26,602	4	,577	252	46,104	,000
Z: pil ppt pc (% media)	23,655	4	,447	252	52,941	,000
Z: SM(reddito_disponibile_netto)	43,264	4	,320	252	135,282	,000
Z: %ore non lavorate	35,429	4	,475	252	74,556	,000
Z_SERV_RES	5,230	4	,601	252	8,706	,000
Z_IND_VECC	24,567	4	,540	252	45,517	,000
Z: tra 0 e 9	24,698	4	,265	252	93,332	,000
Z: tasso fertilità totale	37,972	4	,196	252	193,919	,000
Z: tasso movimento naturale pop	16,050	4	,406	252	39,532	,000
Z: t occupaz tot	37,501	4	,316	252	118,745	,000
Z_UNEMP_R	25,265	4	,410	252	61,558	,000
Z_LONG_UNEMP_R	17,424	4	,347	252	50,189	,000
Z_DIFF_UNEMP_GA	31,236	4	,370	252	84,357	,000

INDICATORI DI BENESSERE	B	W	T	R^2
Z: SM(terz_edu)	103,288	143,363	246,652	0,419
Z: SM(formaz_ad)	196,331	59,718	256,048	0,767
Z: SM(asp_vita)	182,762	79,977	262,739	0,696
Z_MORTI_INC	106,409	145,405	251,814	0,423
Z: pil ppt pc (% media)	94,619	112,597	207,216	0,457
Z: SM(reddito_disponibile_netto)	173,057	80,591	253,648	0,682
Z: %ore non lavorate	141,715	119,749	261,463	0,542
Z_SERV_RES	20,920	151,386	172,306	0,121
Z_IND_VECC	98,268	136,013	234,280	0,419
Z: tra 0 e 9	98,792	66,686	165,477	0,597
Z: tasso fertilità totale	151,888	49,345	201,233	0,755
Z: tasso movimento naturale pop	64,198	102,309	166,507	0,386
Z: t occupaz tot	150,004	79,584	229,588	0,653
Z_UNEMP_R	101,060	103,428	204,488	0,494
Z_LONG_UNEMP_R	69,696	87,487	157,184	0,443
Z_DIFF_UNEMP_GA	124,944	93,311	218,255	0,572

I 5 gruppi della partizione risultano quindi più omogenei nel loro interno con riferimento alla formazione continua degli adulti, al tasso di fertilità ed all'aspettativa di vita, mentre i cluster sono meno omogenei in relazione alla % di suolo cementificato ed al tasso di movimento naturale.

Si riportano i poli della partizione iniziale (che vengono scelti automaticamente dal programma) e le successive modifiche ai centroidi di gruppo, sino ad arrivare al nono passo in cui non si

apportano più correzioni, in quanto si è raggiunto un “ottimo” locale. Pertanto ogni cluster è associato con un centroide ed ogni unità viene assegnata al cluster il cui centroide risulta più vicino.

Centri dei cluster iniziali

	Cluster				
	1	2	3	4	5
Z: SM(terz_edu)	-,18393	,93896	-1,38870	,26054	1,20798
Z: SM(formaz_ad)	,21939	-,65071	-1,34680	2,46716	-,02714
Z: SM(asp_vita)	,72981	,18806	-2,44330	,92329	,26545
Z_MORTI_INC	1,1125	-,4671	-1,4511	1,0866	1,6563
Z: pil ppt pc (% media)	-,25868	-,20357	-1,66407	1,86318	,26490
Z: SM(reddito_disponibile_netto)	-,48166	-,27883	-2,33897	,56925	,10850
Z: %ore non lavorate	-,46117	1,10374	-,82938	,77849	,63120
Z_SERV_RES	,0000	,5733	,0000	,0000	-4,4921
Z_IND_VECC	,9065	1,7346	,0582	,2796	-1,1752
Z: tra 0 e 9	,02763	2,53642	-,71778	,15001	-1,13499
Z: tasso fertilità totale	-1,63009	1,50579	-1,38653	,62287	-,86896
Z: tasso movimento naturale pop	,69180	2,18943	-1,41567	,17737	-,21536
Z: t occupaz tot	-1,75581	-,92553	-,70499	1,73396	,12529
Z_UNEMP_R	-4,0487	-,9669	,3957	1,3181	-,7992
Z_LONG_UNEMP_R	-2,9218	-,9983	,5160	1,2750	-1,2816
Z_DIFF_UNEMP_GA	-1,4510	-,3975	,2963	,8360	1,2985

Cronologia iterazioni^a

Iterazione	Modifiche ai centri dei cluster				
	1	2	3	4	5
1	3,526	3,177	2,607	2,870	3,614
2	,401	,211	,246	,203	1,198
3	,192	,107	,085	,485	,932
4	,114	,189	,086	,546	,472
5	,000	,238	,000	,300	,238
6	,000	,000	,000	,165	,166
7	,000	,104	,000	,033	,050
8	,000	,136	,000	,078	,000
9	,000	,000	,000	,000	,000

Dalle tabelle “Centri dei cluster finali” e “Numero di casi in ogni cluster” apprendiamo gli elementi necessari per la descrivere le caratteristiche e la numerosità dei cluster individuati, con riferimento a 257 regioni, escludendo le tre sopra identificate come anomale.

Centri dei cluster finali

	Cluster				
	1	2	3	4	5
Z: SM(terz_edu)	-,30453	,42430	-,81479	,85996	-,29100
Z: SM(formaz_ad)	-,38068	-,59124	-,79595	1,44755	-,26082
Z: SM(asp_vita)	,71513	,48403	-1,50570	,36011	,43929
Z_MORTI_INC	-,0403	-,2165	-1,0364	,6843	,3667
Z: pil ppt pc (% media)	-,22447	,09881	-1,03298	,40353	,51545
Z: SM(reddito_disponibile_netto)	-,18906	,41772	-1,38725	,24930	,82483
Z: %ore non lavorate	-,55407	-,08698	-1,08230	,89943	,36240
Z_SERV_RES	,3455	,0814	,3597	-,3869	,0955
Z_IND_VECC	-,5563	,6042	,2546	,4038	-,9780
Z: tra 0 e 9	-,48568	,86745	-,34043	,49258	-,81800
Z: tasso fertilità totale	-,83956	,83106	-,71573	,88892	-,61743
Z: tasso movimento naturale pop	-,11050	,71871	-,52675	,33189	-,59864
Z: t occupaz tot	-1,24628	-,19027	-,67415	,86875	,65240
Z_UNEMP_R	-1,3465	-,0746	-,2504	,5170	,6631
Z_LONG_UNEMP_R	-,9825	-,0297	-,2477	,6774	,4050
Z_DIFF_UNEMP_GA	-1,2867	-,2357	-,3054	,2574	1,0083

Numero di casi in ogni cluster

Cluster	1	29,000
	2	37,000
	3	61,000
	4	65,000
	5	65,000
Validi		257,000
Mancanti		,000

Il primo gruppo, formato da sole 29 regioni, presenta valori inferiori rispetto alla media relativamente agli indicatori economici, di istruzione ed in particolare gli indicatori sull'ambito lavorativo (occupazione, disoccupazione nel breve e lungo periodo, disuguaglianza tra adulti e giovani) registrano i valori peggiori.

Il secondo gruppo è costituito da 37 unità e registra valori superiori alla media per quanto riguarda la percentuale di laureati, l'aspettativa di vita ed il reddito familiare ed inoltre rappresenta il gruppo avente la miglior situazione demografica: popolazione giovane, elevati tassi di fertilità e movimento naturale; al contrario l'indice relativo alla formazione continua degli adulti è inferiore alla media.

Il terzo gruppo, di 61 unità, presenta valori di tutte le variabili inferiori alla media e pertanto unisce regioni con situazioni di malessere in tutti gli ambiti esaminati, in particolare rispetto agli indicatori economici, di aspettativa di vita e istruzione.

Nel quarto gruppo vi sono 65 regioni che al contrario presentano una condizione di benessere generalizzato a tutti gli ambiti, con scostamenti rispetto alla media più accentuati riguardo all'istruzione, al tasso di fertilità e di occupazione.

Anche il quinto gruppo unisce 65 regioni aventi valori superiori alla media relativamente all'ambito economico e lavorativo, ma valori molto inferiori alla media riguardo l'aspetto demografico (popolazione più vecchia e bassa fertilità).

Ancora una volta il diagramma degli *scores* è particolarmente utile ad esprimere graficamente quanto contenuto negli output di Spss: alle unità appartenenti ad uno stesso gruppo viene associato il medesimo simbolo.

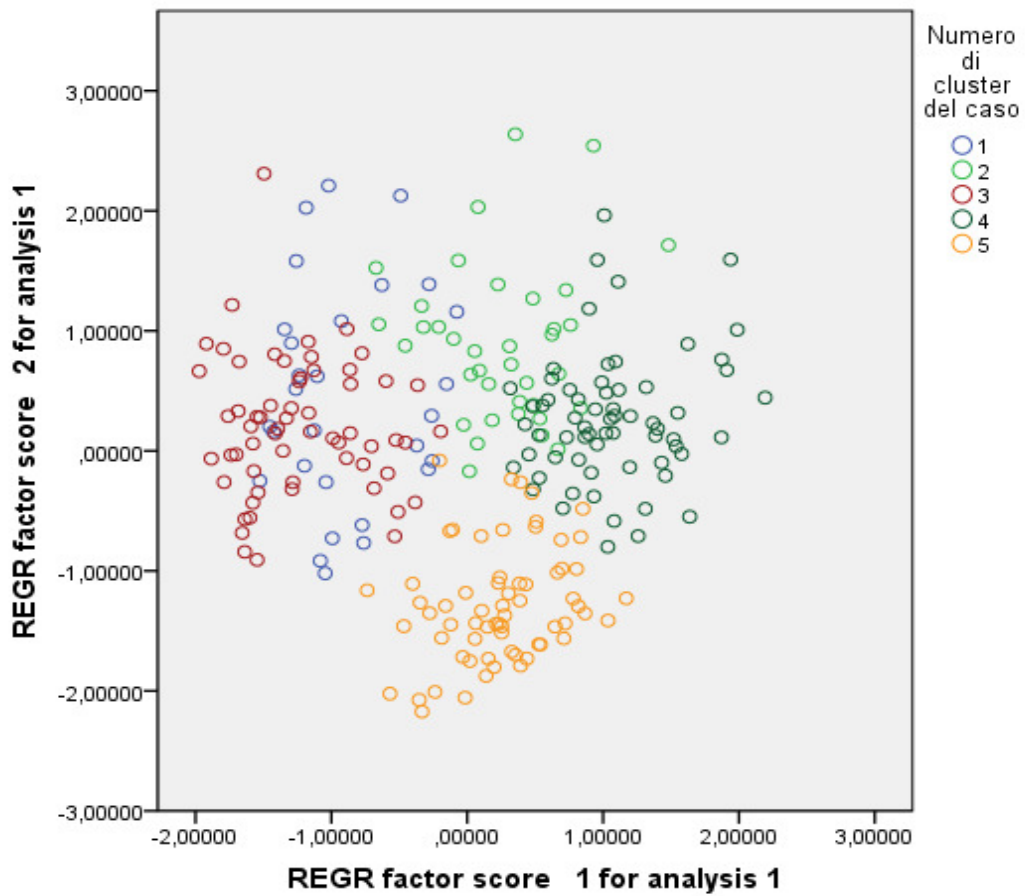


Figura 50. Diagramma di dispersione degli scores divisi per cluster di appartenenza

Per completezza si calcolano le medie dei sedici indicatori di benessere per ciascuno dei cinque gruppi individuati.

CLUSTER INDICE	1			2			3		
	Media	SD	CV	Media	SD	CV	Media	SD	CV
Z: SM(terz_edu)	-0,305	0,892	2,929	0,424	0,853	2,011	-0,815	0,691	0,848
Z: SM(formaz_ad)	-0,381	0,477	1,253	-0,591	0,334	0,565	-0,796	0,454	0,570
Z: SM(asp_vita)	0,715	0,333	0,466	0,484	0,517	1,069	-1,506	0,850	0,564
Z_MORTI_INC	-0,040	0,692	17,181	-0,217	0,943	4,356	-1,036	0,988	0,953
Z: pil ppt pc (% media)	-0,224	0,489	2,177	0,099	0,581	5,875	-1,033	0,532	0,515
SM(reddito_disponib ile_netto)	-0,189	0,538	2,843	0,418	0,371	0,888	-1,387	0,708	0,511
Z: %ore non lavorate	-0,554	0,682	1,231	-0,087	0,685	7,876	-1,082	0,495	0,458
Z_SERV_RES	0,346	0,414	1,198	0,081	0,669	8,225	0,360	0,337	0,938
Z_IND_VECC	-0,556	0,948	1,705	0,604	0,519	0,859	0,255	0,715	2,806
Z: tra 0 e 9	-0,486	0,649	1,336	0,867	0,619	0,714	-0,340	0,449	1,320
Z: tasso fertilità totale	-0,840	0,481	0,573	0,831	0,541	0,650	-0,716	0,440	0,615
Z: tasso movimento naturale pop	-0,110	0,689	6,232	0,719	0,672	0,935	-0,527	0,709	1,346
Z: t occupaz tot	-1,246	0,864	0,693	-0,190	0,483	2,539	-0,674	0,540	0,801
Z_UNEMP_R	-1,346	1,127	0,837	-0,075	0,577	7,739	-0,250	0,687	2,745
Z_LONG_UNEMP_ R	-0,982	0,644	0,656	-0,030	0,589	19,814	-0,248	0,813	3,281
Z_DIFF_UNEMP_G A	-1,287	0,430	0,334	-0,236	0,466	1,976	-0,305	0,616	2,016

CLUSTER INDICE	4			5		
	Media	SD	CV	Media	SD	CV
Z: SM(terz_edu)	0,860	0,608	<i>0,707</i>	-0,291	0,816	<i>2,803</i>
Z: SM(formaz_ad)	1,448	0,686	<i>0,474</i>	-0,261	0,327	<i>1,254</i>
Z: SM(asp_vita)	0,360	0,434	<i>1,204</i>	0,439	0,431	<i>0,981</i>
Z_MORTI_INC	0,684	0,490	<i>0,717</i>	0,367	0,637	<i>1,738</i>
Z: pil ppt pc (% media)	0,404	0,862	<i>2,137</i>	0,515	0,676	<i>1,311</i>
Z: SM(reddito_disponibile_netto)	0,249	0,545	<i>2,186</i>	0,825	0,537	<i>0,651</i>
Z: %ore non lavorate	0,899	0,910	<i>1,011</i>	0,362	0,588	<i>1,623</i>
Z_SERV_RES	-0,387	0,982	<i>2,537</i>	0,096	0,984	<i>10,302</i>
Z_IND_VECC	0,404	0,521	<i>1,291</i>	-0,978	0,911	<i>0,932</i>
Z: tra 0 e 9	0,493	0,517	<i>1,049</i>	-0,818	0,431	<i>0,527</i>
Z: tasso fertilità totale	0,889	0,454	<i>0,511</i>	-0,617	0,343	<i>0,556</i>
Z: tasso movimento naturale pop	0,332	0,587	<i>1,770</i>	-0,599	0,566	<i>0,946</i>
Z: t occupaz tot	0,869	0,467	<i>0,538</i>	0,652	0,543	<i>0,832</i>
Z_UNEMP_R	0,517	0,420	<i>0,812</i>	0,663	0,504	<i>0,759</i>
Z_LONG_UNEMP_R	0,677	0,287	<i>0,423</i>	0,405	0,537	<i>1,327</i>
Z_DIFF_UNEMP_GA	0,257	0,680	<i>2,642</i>	1,008	0,661	<i>0,656</i>

Si propone l'elenco delle regioni in base alla classificazione così ottenuta e si fornisce la distanza di ciascuna unità amministrativa dal rispettivo centroide del gruppo di appartenenza. Com'è noto minore è tale distanza e tanto più l'unità considerata presenta caratteristiche vicine a quelle medie del cluster.

ELENCO REGIONI APPARTENENTI AL GRUPPO 1

CODICE	REGIONE	NAZIONE	DISTANZA_CENTRO
EL13	Dytiki Makedonia	EL	1,916
EL12	Kentriki Makedonia	EL	2,083
EL14	Thessalia	EL	2,443
EL30	Attiki	EL	2,692
EL21	Ipeiros	EL	2,955
EL41	Voreio Aigaio	EL	3,184
ES42	Castilla-La Mancha	ES	1,698
ES43	Extremadura	ES	1,925
ES13	Cantabria	ES	2,038
ES24	Aragón	ES	2,237
ES52	Comunidad Valenciana	ES	2,513
ES51	Cataluña	ES	2,548
ES11	Galicia	ES	2,561
ES23	La Rioja	ES	2,574
ES53	Illes Balears	ES	2,671
ES41	Castilla y León	ES	2,809
ES12	Principado de Asturias	ES	3,326
ES62	Región de Murcia	ES	3,528
ES61	Andalucía	ES	3,616
ES70	Canarias	ES	4,116
ITG2	Sardegna	IT	2,093
ITF5	Basilicata	IT	2,212
ITF4	Puglia	IT	2,225
ITF1	Abruzzo	IT	2,265
ITE4	Lazio	IT	2,341
ITG1	Sicilia	IT	2,621
ITF6	Calabria	IT	2,651
ITF2	Molise	IT	2,873
ITF3	Campania	IT	3,229

ELENCO REGIONI APPARTENENTI AL GRUPPO 2

CODICE	REGIONE	NAZIONE	DISTANZA_CENTRO
BE23	Prov. Oost-Vlaanderen	BE	2,047
BE33	Prov. Liège	BE	2,191
BE21	Prov. Antwerpen	BE	2,225
BE35	Prov. Namur	BE	2,355
BE31	Prov. Brabant Wallon	BE	2,804
BE22	Prov. Limburg (BE)	BE	2,873
BE32	Prov. Hainaut	BE	3,214
BE34	Prov. Luxembourg (BE)	BE	3,473
CY00	Kýpros	CY	2,635
EL43	Kriti	EL	3,375
EL42	Notio Aigaió	EL	3,779
ES22	Comunidad Foral de Navarra	ES	2,963
ES30	Comunidad de Madrid	ES	3,569
FR25	Basse-Normandie	FR	1,105
FR41	Lorraine	FR	1,209
FR61	Aquitaine	FR	1,313
FR21	Champagne-Ardenne	FR	1,313
FR43	Franche-Comté	FR	1,371
FR52	Bretagne	FR	1,425
FR82	Provence-Alpes-Côte d'Azur	FR	1,467
FR42	Alsace	FR	1,473
FR71	Rhône-Alpes	FR	1,494
FR51	Pays de la Loire	FR	1,497
FR24	Centre	FR	1,501
FR23	Haute-Normandie	FR	1,557
FR62	Midi-Pyrénées	FR	1,569
FR26	Bourgogne	FR	1,664
FR53	Poitou-Charentes	FR	1,729
FR72	Auvergne	FR	1,924
FR22	Picardie	FR	2,026
FR81	Languedoc-Roussillon	FR	2,162
FR30	Nord - Pas-de-Calais	FR	2,455
FR10	Île de France	FR	4,030
IE01	Border, Midland and Western	IE	3,411
IE02	Southern and Eastern	IE	3,745
PT17	Lisboa	PT	2,247
UKNO	Northern Ireland	UK	1,905

ELENCO REGIONI APPARTENENTI AL GRUPPO 3

CODICE	REGIONE	NAZIONE	DISTANZA_CENTRO
BG42	Yuzhen tsentralen	BG	2,294
BG33	Severozitochen	BG	2,442
BG34	Yugoiztochen	BG	2,457
BG32	Severen tsentralen	BG	3,133
BG41	Yugozapaden	BG	3,135
BG31	Severozapaden	BG	3,850
CZ08	Moravskoslezsko	CZ	1,478
CZ07	Strední Morava	CZ	1,650
CZ04	Severozápad	CZ	1,725
CZ05	Severovýchod	CZ	1,996
CZ06	Jihovýchod	CZ	2,192
CZ03	Jihozápad	CZ	2,200
CZ02	Strední Cechy	CZ	2,527
EE00	Eesti	EE	3,196
EL22	Ionia Nisia	EL	3,193
EL11	Anatoliki Makedonia, Thraki	EL	3,442
EL23	Dytiki Ellada	EL	3,532
EL25	Peloponnisos	EL	4,513
EL24	Sterea Ellada	EL	4,683
HU21	Közép-Dunántúl	HU	1,428
HU22	Nyugat-Dunántúl	HU	1,484
HU33	Dél-Alföld	HU	1,960
HU23	Dél-Dunántúl	HU	1,962
HU32	Észak-Alföld	HU	2,418
HU10	Közép-Magyarország	HU	2,485
HU31	Észak-Magyarország	HU	3,066
LT00	Lietuva	LT	3,027
LV00	Latvija	LV	3,062
PL51	Dolnoslaskie	PL	1,032
PL33	Swietokrzyskie	PL	1,190
PL61	Kujawsko-Pomorskie	PL	1,278
PL22	Slaskie	PL	1,383
PL11	Lódzkie	PL	1,391
PL42	Zachodniopomorskie	PL	1,419
PL34	Podlaskie	PL	1,511
PL31	Lubelskie	PL	1,516
PL43	Lubuskie	PL	1,565
PL52	Opolskie	PL	1,637
PL21	Małopolskie	PL	1,672
PL41	Wielkopolskie	PL	1,717
PL62	Warmińsko-Mazurskie	PL	1,816
PL63	Pomorskie	PL	2,096
PL32	Podkarpackie	PL	2,269
PL12	Mazowieckie	PL	2,729
PT15	Algarve	PT	2,779
PT11	Norte	PT	2,903
PT20	Região Autónoma dos Açores	PT	3,174
PT30	Região Autónoma da Madeira	PT	3,346
PT18	Alentejo	PT	3,549
RO22	Sud-Est	RO	2,103
RO11	Nord-Vest	RO	2,163
RO12	Centru	RO	2,227
RO42	Vest	RO	2,363
RO41	Sud-Vest Oltenia	RO	2,443
RO31	Sud - Muntenia	RO	2,549
RO21	Nord-Est	RO	2,890
RO32	Bucuresti - Ilfov	RO	3,019
SI01	Vzhodna Slovenija	SI	2,759
SK02	Západné Slovensko	SK	2,182
SK03	Stredné Slovensko	SK	2,855
SK04	Východné Slovensko	SK	4,465

ELENCO REGIONI APPARTENENTI AL GRUPPO 4

<i>CODICE</i>	<i>REGIONE</i>	<i>NAZIONE</i>	<i>DISTANZA_CENTRO</i>
AT13	Wien	AT	3,549
BE24	Prov. Vlaams-Brabant	BE	2,705
DK03	Syddanmark	DK	2,367
DK05	Nordjylland	DK	2,525
DK04	Midtjylland	DK	2,703
DK02	Sjælland	DK	2,786
DK01	Hovedstaden	DK	3,905
FI19	Länsi-Suomi	FI	1,881
FI18	Etelä-Suomi	FI	1,898
FI20	Åland	FI	2,566
FI1A	Pohjois- ja Itä-Suomi	FI	2,598
LU00	Luxembourg	LU	5,027
NL33	Zuid-Holland	NL	2,074
NL41	Noord-Brabant	NL	2,449
NL12	Friesland (NL)	NL	2,505
NL22	Gelderland	NL	2,624
NL34	Zeeland	NL	2,720
NL13	Drenthe	NL	2,725
NL21	Overijssel	NL	2,777
NL42	Limburg (NL)	NL	2,955
NL32	Noord-Holland	NL	2,961
NL11	Groningen	NL	3,254
NL31	Utrecht	NL	3,472
NL23	Flevoland	NL	4,397
SE12	Östra Mellansverige	SE	1,724
SE23	Västsverige	SE	1,768
SE21	Småland med öarna	SE	1,821
SE33	Övre Norrland	SE	1,872
SE22	Sydsverige	SE	1,910
SE31	Norra Mellansverige	SE	2,327
SE32	Mellersta Norrland	SE	2,716
SE11	Stockholm	SE	3,158
SI02	Zahodna Slovenija	SI	2,591
UKH1	East Anglia	UK	0,929
UKF1	Derbyshire and Nottinghamshire	UK	1,053
UKK1	Gloucestershire, Wiltshire and Bristol/Bath area	UK	1,182
UKL2	East Wales	UK	1,333
UKF2	Leicestershire, Rutland and Northamptonshire	UK	1,392
UKD4	Lancashire	UK	1,410
UKG2	Shropshire and Staffordshire	UK	1,431
UKJ4	Kent	UK	1,449
UKJ3	Hampshire and Isle of Wight	UK	1,456
UKE1	East Yorkshire and Northern Lincolnshire	UK	1,477
UKH3	Essex	UK	1,499
UKG1	Herefordshire, Worcestershire and Warwickshire	UK	1,532
UKM2	Eastern Scotland	UK	1,553
UKC1	Tees Valley and Durham	UK	1,649
UKE2	North Yorkshire	UK	1,670
UKE3	South Yorkshire	UK	1,685
UKK4	Devon	UK	1,796
UKC2	Northumberland and Tyne and Wear	UK	1,910
UKE4	West Yorkshire	UK	1,911
UKK2	Dorset and Somerset	UK	1,998
UKD1	Cumbria	UK	2,052
UKM3	South Western Scotland	UK	2,116
UKJ2	Surrey, East and West Sussex	UK	2,143
UKK3	Cornwall and Isles of Scilly	UK	2,169
UKH2	Bedfordshire and Hertfordshire	UK	2,197
UKF3	Lincolnshire	UK	2,282
UKM6	Highlands and Islands	UK	2,358
UKL1	West Wales and The Valleys	UK	2,403
UKJ1	Berkshire, Buckinghamshire and Oxfordshire	UK	2,827

ELENCO REGIONI APPARTENENTI AL GRUPPO 5

CODICE	REGIONE	NAZIONE	DISTANZA_CENTRO
AT22	Steiermark	AT	1,652
AT21	Kärnten	AT	1,717
AT11	Burgenland (AT)	AT	1,895
AT12	Niederösterreich	AT	2,306
AT31	Oberösterreich	AT	2,319
AT33	Tirol	AT	2,426
AT32	Salzburg	AT	2,830
AT34	Vorarlberg	AT	3,021
BE25	Prov. West-Vlaanderen	BE	2,369
CZ01	Praha	CZ	3,867
DE92	Hannover	DE	0,917
DEB3	Rheinhesse n-Pfalz	DE	0,992
DEF0	Schleswig-Holstein	DE	1,011
DE72	Gießen	DE	1,059
DE73	Kassel	DE	1,128
DEA2	Köln	DE	1,222
DE25	Mittelfranken	DE	1,262
DEA4	Detmold	DE	1,296
DEB1	Koblenz	DE	1,310
DE24	Oberfranken	DE	1,313
DE12	Karlsruhe	DE	1,323
DE94	Weser-Ems	DE	1,342
DEA3	Münster	DE	1,371
DE26	Unterfranken	DE	1,372
DE93	Lüneburg	DE	1,478
DE23	Oberpfalz	DE	1,511
DE91	Braunschweig	DE	1,537
DE22	Niederbayern	DE	1,638
DE13	Freiburg	DE	1,662
DEB2	Trier	DE	1,665
DE27	Schwaben	DE	1,669
DEA5	Arnsberg	DE	1,683
DE71	Darmstadt	DE	1,752
DEA1	Düsseldorf	DE	1,776
DE11	Stuttgart	DE	1,795
DE14	Tübingen	DE	1,797
DEC0	Saarland	DE	2,083
DE21	Oberbayern	DE	2,499
DEG0	Thüringen	DE	2,768
DED2	Dresden	DE	2,822
DE80	Mecklenburg-Vorpommern	DE	3,009
DED3	Leipzig	DE	3,195
DEE0	Sachsen-Anhalt	DE	3,505
DED1	Chemnitz	DE	3,794
DE50	Bremen	DE	4,466
DE60	Hamburg	DE	4,510
DE30	Berlin	DE	5,602
ES21	País Vasco	ES	3,696
FR63	Limousin	FR	2,360
FR83	Corse	FR	3,030
ITE3	Marche	IT	1,830
ITD3	Veneto	IT	2,051
ITC4	Lombardia	IT	2,093
ITD2	Provincia Autonoma di Trento	IT	2,182
ITE2	Umbria	IT	2,225
ITD4	Friuli-Venezia Giulia	IT	2,265
ITC2	Valle d'Aosta/Vallée d'Aoste	IT	2,268
ITD5	Emilia-Romagna	IT	2,337
ITE1	Toscana	IT	2,372
ITC1	Piemonte	IT	2,593
ITD1	Provincia Autonoma di Bolzano/Bozen	IT	3,059
ITC3	Liguria	IT	3,220

CAPITOLO 5: COSTRUZIONE DEGLI INDICATORI MEDIANTE GLI INSIEMI SFOCATI

5.1 DEFINIZIONE DI FUZZY SETS

« Finché le leggi della matematica si riferiscono alla realtà, non sono certe, e finché sono certe, non si riferiscono alla realtà. » (Albert Einstein da *Sidelights on Relativity*, Dover)

La teoria degli insiemi sfocati, in inglese *fuzzy sets*, trova applicazione nelle situazioni in cui l'informazione è imperfetta a causa dell'imprecisione. Tale ambiguità è dovuta alla vaghezza dei dati del problema.

Tale teoria, introdotta dal matematico iraniano Lotfi Asker Zadeh nel 1965, costituisce un'estensione della teoria classica degli insiemi, in quanto alla rigida partizione matematica degli elementi dell'universo X in appartenenti o non appartenenti ad un certo insieme, si sostituisce il concetto di grado di appartenenza: si postula l'esistenza di una funzione di appartenenza che associa a ogni elemento di X un numero reale scelto in un conveniente insieme, abitualmente $[0,1]$. In tal modo si traducono in termini matematici tutte le situazioni reali proprie delle scienze umane (economia, sociologia, psicologia, ecc.) nelle quali non è possibile passare bruscamente dal possedere al non possedere una certa caratteristica (si vedano ad esempio i volumi di Klir, Folger, 1988; Zinnerman, 2001).

L'utilizzo di tale metodo è quindi particolarmente indicato in tutte le situazioni in cui si voglia quantificare l'imprecisione insita nei concetti umani, poiché "getta un ponte tra la rigida precisione della matematica classica e l'imprecisione con cui ci si presenta il mondo della realtà fenomenica"

Di seguito si propongono le relative definizioni matematiche

Nella logica classica il concetto di appartenenza ad un insieme si riferisce unicamente alla condizione VERO-FALSO e definiamo "*crisp*" l'insieme A la cui funzione di appartenenza è una funzione booleana che assume solo due valori: uno nel caso di appartenenza, zero nel caso di non appartenenza

$\mu_A(x) : X \rightarrow \{0,1\}$ per ogni $x \in X$

Come già anticipato, il concetto di insieme *fuzzy* estende il concetto di insieme *crisp* mediante l'introduzione di un grado di appartenenza.

Dato un insieme non vuoto X e detto x il suo generico elemento, dicesi (sotto)insieme sfocato di X e lo si indica con A , l'insieme ordinato delle coppie $[x, \mu_A(x)]$, essendo $\mu_A(x)$ una funzione detta di appartenenza, di dominio X ed avente per codominio l'intervallo chiuso $[0, 1]$

$\mu_A(x) : X \rightarrow [0,1]$ per ogni $x \in X$

Pertanto, X è l'insieme delle unità di riferimento, A un corrispondente insieme sfocato e l'appartenenza ad A da parte del singolo elemento x è tanto maggiore quanto più il valore della funzione di appartenenza si avvicina ad 1.

In particolare se $\mu_A(x)=0$ significa che l'elemento x non appartiene per nulla all'insieme A , mentre se $\mu_A(x)=1$ tale elemento appartiene completamente ad A .

Il numero reale $\mu_A(x)$ è detto grado di appartenenza dell'elemento $x \in X$ all'insieme sfocato A .

Sugli insiemi *fuzzy* continuano a valere gli operatori insiemistici di unione, intersezione, complemento e le leggi di De Morgan; al contrario non valgono il principio del terzo escluso (un elemento appartiene ad un insieme oppure al suo insieme complementare e mai ad entrambi) e il principio di non-contraddizione (l'intersezione di un insieme con il suo complementare è un insieme vuoto).

Com'è stato evidenziato più volte, anche il concetto di benessere è vago ed impreciso e quindi possiamo utilizzare proficuamente la teoria degli insiemi sfocati, considerando il benessere come un problema di grado: ogni regione può appartenere "in parte" a quelle "benestanti" e può essere assegnata simultaneamente a due (o più) gruppi, seppur con un parziali e diversi gradi di appartenenza, ma tale che la loro somma sia uguale a 1: due regioni saranno tanto più simili quanto più i valori della funzione di appartenenza ai medesimi gruppi sono prossimi tra loro.

5.2 LA TEORIA DEI FUZZY SETS APPLICATA ALL'ANALISI SUL BENESSERE

Consideriamo un insieme di n regioni ed un sottoinsieme A di queste che presenta una condizione di elevato benessere (che per brevità chiameremo “*regioni benestanti*”).

Adottando l'approccio *fuzzy* il numero reale

$$\mu_A(i) \quad (i = 1, 2, \dots, n)$$

indica, per ciascuna regione i -esima, il grado di appartenenza al sottoinsieme di regioni benestanti.

Se $\mu_A(i) = 0$ la regione i -esima non è certamente in uno stato di benessere;

Se $\mu_A(i) = 1$ la regione i -esima appartiene completamente al sottoinsieme di elevato benessere;

Se $0 < \mu_A(i) < 1$ la regione i -esima mostra un'appartenenza parziale al sottoinsieme di regioni benestanti.

Il problema principale dell'approccio *fuzzy* è trovare una forma funzionale adatta al fenomeno in esame, cioè un'appropriata funzione di appartenenza.

Esistono una pluralità di criteri per trasformare gli indicatori di benessere in numeri *fuzzy* e pertanto le funzioni di appartenenza possono assumere forme diverse.

Ipotizziamo che il sottoinsieme sfocato A definisca la posizione di ciascuna unità rispetto al raggiungimento del costrutto latente, cioè il grado di appartenenza di ciascuna regione all'insieme di quelle benestanti.

5.2.1 POSSIBILI FUNZIONI DI APPARTENENZA

Per ciascuna variabile X occorre scegliere una soglia inferiore, l (*lower*) ed una superiore u (*upper*). In tal modo la funzione di appartenenza è pari a 0 nel caso in cui la variabile risulti uguale o inferiore alla soglia *lower*, (indicando quindi assenza del fenomeno), 1 nell'ipotesi in cui la variabile sia pari o superiore alla soglia *upper* (segnalando completa presenza), mentre nei casi rimanenti, cioè x compreso tra le due soglie, la funzione di appartenenza deve essere una funzione crescente che per semplicità, in prima istanza, sceglieremo lineare.

$$\left\{ \begin{array}{ll} \mu_A(x_i)=0 & x_i \leq l \\ \mu_A(x_i) = \frac{x_i - l}{u - l} & l < x_i < u \\ \mu_A(x_i)=1, & x_i \geq u \end{array} \right.$$

Trattandosi di caratteri quantitativi è inevitabile un elemento di soggettività dato dal criterio scelto per la suddivisione delle classi. Di seguito si ricorre ai percentili per definire tali soglie.

Fissiamo la soglia inferiore l al quinto decile, e quella superiore u , al nono decile, per $i=1, \dots, 260$ e $j=1, \dots, 16$. Questa scelta ci appare razionale, poiché se una regione presenta un valore d'una variabile (misura d'un aspetto del benessere) minore della mediana non appartiene certamente a quelle "benestanti" (da tale punto di vista), mentre se si colloca nella fascia del 10% migliore viene ritenuta completamente appartenente ad A.

Pertanto, la distribuzione di ogni indicatore di benessere j viene divisa in tre classi ed alla funzione di appartenenza alle regioni benestanti si assegna:

- valore pari a 1 nel caso in cui la regione appartenga al 10% delle regioni con i valori migliori della variabile considerata;
- valore uguale a 0 se la regione appartiene al 50% delle regioni con i valori peggiori dell'indicatore in oggetto;
- nelle regioni rimanenti, il valore *fuzzy* di ciascuna variabile è dato da una funzione lineare tra il valore delle due soglie prima stabilite.

Si rappresenta graficamente la funzione di appartenenza che ne risulta.

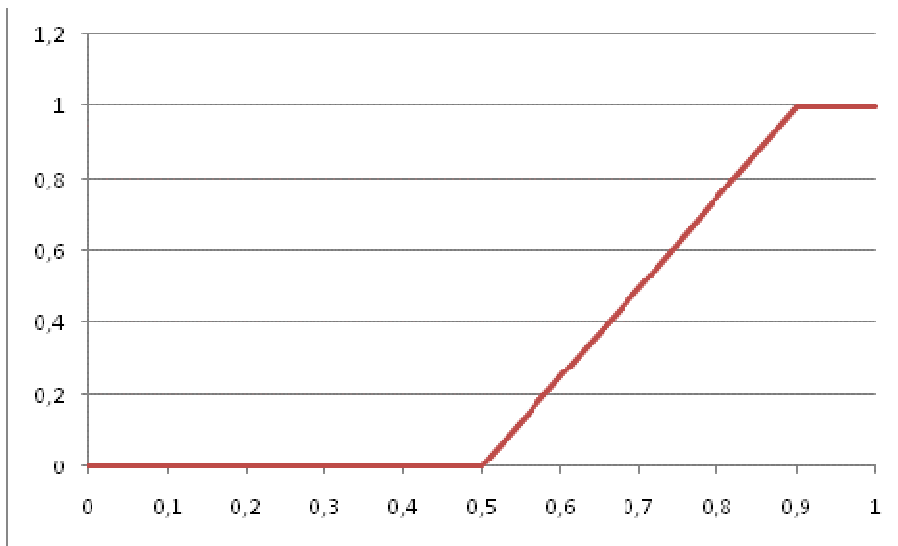


Figura 51. Funzione di appartenenza lineare

Ovviamente, si potrebbero scegliere altri valori per i limiti *lower* e *upper*.

Per effettuare i calcoli necessari all'analisi ci si avvale del programma Matlab. Si riportano i comandi matlab per calcolare gli indicatori *fuzzy* per ogni variabile per ciascuna regione, indipendentemente dai valori delle soglie.

```
soglie = xlsread('benessere_fuzzy260_15.xls','db','D2:R3');  
A = xlsread('benessere_fuzzy260_15.xls','db','D5:R264');  
  
[n,k] = size(A);  
B = zeros(n,k);  
  
for i = 1:n  
    for j = 1:k  
        if (soglie(1,j)>A(i,j) || abs(soglie(1,j)-A(i,j))<10^(-16))
```

```

        B(i,j) = 0;
    else if (A(i,j)>soglie(2,j) || abs(A(i,j)-soglie(2,j))<10^(-16))
        B(i,j) = 1;
    else
        B(i,j) = (A(i,j)-soglie(1,j))/(soglie(2,j)-soglie(1,j));
    end
end
end
end

xlswrite('benessere_fuzzy260_15.xls',B,'db fuzzy','D2:R261')

```

Le prime due righe servono per importare i dati relativi ai valori di soglia e agli indicatori rispettivamente nelle matrici “**soglie**” e “**A**”. Dati k indicatori di benessere ed n unità (regioni), “**soglie**” è una matrice di 2 righe e k colonne, in cui l'elemento $(1, j)$ con $j=1, \dots, k$ è la soglia inferiore relativa al j -esimo indicatore, mentre l'elemento $(2, j)$ $j=1, \dots, k$ è il valore della soglia superiore relativa al j -esimo indicatore. In altri termini, sulla prima riga ci sono i valori delle soglie inferiori, mentre nella seconda ci sono quelle superiori.

La matrice **A** è di n righe e k colonne, in cui l'elemento (i, j) $i=1, \dots, n$ e $j=1, \dots, k$ è il valore assoluto relativo al j -esimo indicatore per la i -esima regione. In pratica, ogni riga corrisponde ad una regione diversa e contiene il valore degli indicatori di quella regione. La matrice **B** ha la stessa struttura della matrice **A**, ma contiene i valori degli indicatori in seguito alla trasformazione *fuzzy*. L'ultima riga serve per esportare nel medesimo file Excel, nel foglio denominato “db fuzzy” la matrice “**B**” che contiene gli indicatori *fuzzy* che sono stati calcolati. Siccome i valori delle soglie hanno 15 cifre decimali, mentre i valori assoluti degli indicatori possiedono meno cifre decimali, per determinare se un valore è uguale al valore di soglia, si verifica che il valore assoluto della differenza fra questi due valori sia inferiore a $10^{(-16)}$.

Successivamente, tenendo fisse le due soglie prima utilizzate si procede introducendo altre funzioni di appartenenza aventi forma non lineare, come ad esempio quella esponenziale:

$$\mu_A(x_i) = \frac{1}{1 + e^{-a(x-b)}}$$

Tale funzione è introdotta da Zimmerman (2001) con l'obiettivo di studiare la relazione tra misure fisiche ed entità percepita delle medesime e viene spesso utilizzata nelle indagini di soddisfazione della clientela. Tale funzione prevede che il parametro a fornisca una misura della vaghezza del costrutto latente, mentre il parametro b è il punto in cui l'attitudine del soggetto da positiva si trasforma in negativa.

Anche lo studio di Balamone-Lutz (2006), basato sull'approccio di Sen, ricorre ad essa per misurare il benessere in maniera *fuzzy*.

Per la scelta del valore da attribuire al parametro a , che rappresenta il livello di sfocatura, non esistono basi teoriche cui affidarsi, ma convenzionalmente si attribuisce valore 2, frutto di studi empirici. Riguardo al parametro b , si fa riferimento alla mediana. I comandi Matlab necessari per misurare i gradi di appartenenza con la funzione esponenziale:

```
B = zeros(n,k);
for i = 1:n
    for j = 1:k
        B(i,j) = 1./(1+exp(-2*(A(i,j)-soglie(1,j))));
    end
end
xlswrite('ben_fuzzy260_15_5_9.xls',B,'db fuzzy zimm','D2:R261')
```

In aggiunta a ciò si propone come ulteriore funzione di appartenenza, quella cubica.

$$\left\{ \begin{array}{ll} \mu_A(x_i) = 0 & x_i \leq l \\ \mu_A(x_i) = \frac{\sqrt[3]{\left(2 \frac{x-l}{u-l} - 1\right)} + 1}{2} & l < x_i < u \\ \mu_A(x_i) = 1, & x_i \geq u \end{array} \right.$$

Come nel caso precedente u è l'estremo superiore dell'intervallo e coincide con il nono decile, ed l quello inferiore, dato dalla mediana.

```

B = zeros(n,k);
for i = 1:n
    for j = 1:k
        if (soglie(1,j)>A(i,j) || abs(soglie(1,j)-A(i,j))<10^(-16))
            B(i,j) = 0;
        else if (A(i,j)>soglie(2,j) || abs(A(i,j)-soglie(2,j))<10^(-16))
            B(i,j) = 1;
        else
            B(i,j) = (rq(2.*(A(i,j)-soglie(1,j))./(soglie(2,j)-...
                soglie(1,j))-1,3)+1)./2;
        end
    end
end
end

xlswrite('ben_fuzzy260_15_5_9.xls',B,'db fuzzy rad cubica','D2:R261')

```

La funzione “rq” che compare nei calcoli con la radice cubica è una funzione che serve per calcolare la radice cubica ed è composta dalle seguenti righe di codice:

```

function [y] = rq(x,i)

y = x./abs(x.^((i-1)/i));

```

Questa funzione è necessaria per calcolare radici con indici maggiori di 2, perché in Matlab non esistono funzioni specifiche per tale scopo e non si può ovviare al problema elevando semplicemente il numero ad una potenza frazionaria (per esempio a 1/3 per calcolare la radice cubica).

5.2.2 AGGREGAZIONE E PONDERAZIONE DEI VALORI FUZZY

L'analisi viene condotta sulle sedici variabili già utilizzate nel capitolo precedente per la costruzione d'un indicatore sintetico mediante le componenti principali.

Una volta calcolati i valori *fuzzy* di ciascuna di tali variabili per le 260 regioni con le tre diverse funzioni di appartenenza, occorre aggregare tali valori. Ricordiamo che per la costruzione dell'indice composito di benessere le operazioni di aggregazione e assegnazione dei pesi alle diverse componenti costituiscono un passaggio cruciale per il raggiungimento di risultati affidabili (Munda, Nardo, 2005).

Riguardo alla funzione di aggregazione, tra i tanti criteri cui possiamo ricorrere, se ne scelgono due: quello più semplice che prevede di assegnare pesi uguali a ciascuna variabile; il secondo si basa sui pesi normalizzati ottenuti dall'Analisi delle componenti principali, in quanto la prima componente principale spiega il 37% della varianza totale delle variabili. In tal modo il calcolo dei pesi è basato su un criterio non soggettivo, bensì derivato dalle relazioni tra i singoli indicatori.

Applicando il primo metodo, il grado di appartenenza della *i*-esima regione all'insieme delle regioni con elevato benessere è dato dalla proporzione :

$$\mu_A(x_i) = \frac{1}{k} \sum_{j=1}^k \mu(x_{ij}) \quad (i=1,2,\dots,n; j=1,2,\dots,k)$$

Il livello di benessere globale relativo ad ogni regione viene così calcolato attraverso i *k* indicatori osservati X_j ($j=1,2,\dots,k$) in modo che il grado di appartenenza globale rappresenti dunque una sorta

di frequenza relativa *fuzzy* di indicatori di benessere che denotano per la regione *i*-esima un elevato livello di qualità della vita.

Nel caso si voglia invece assegnare importanza diversa ai singoli indicatori di benessere, per determinare il grado di appartenenza si ricorre alla seguente

$$\mu_A(x_i) = \frac{\sum_{j=1}^k x_{ij} w_j}{\sum_{j=1}^k w_j} \quad (i= 1,2,\dots,n; j= 1,2,\dots,k)$$

I pesi w_j corrispondono ai *factor loadings*, cioè ai coefficienti delle espressioni dei punteggi normalizzati in funzione lineare degli scostamenti standardizzati delle variabili e pertanto il peso (*loading*) di ciascun indicatore è proporzionale alla rispettiva correlazione con la prima componente (*factor*). Si riporta la matrice dei coefficienti di punteggio dei componenti ed i pesi relativi, la cui somma è pari all'unità.

Poiché la relazione tra la variabile latente e l'unico indicatore ambientale (“Percentuale di suolo occupato da attività residenziali e commerciali”) è negativa, si è provveduto ad eliminare la suddetta variabile per poter procedere correttamente al calcolo dei pesi.

Matrice dei coefficienti di punteggio dei componenti

	Componente	
	1	2
Z: SM(terz_edu)	,114	,080
Z: SM(formaz_ad)	,123	,045
Z: SM(asp_vita)	,103	-,025
Z_MORTI_INC	,112	-,005
Z: pil ppt pc (% media)	,124	-,023
Z: SM(reddito_disponibile_netto)	,121	-,088
Z: %ore non lavorate	,123	-,043
Z_IND_VECC	,036	,249
Z: tra 0 e 9	,078	,259
Z: tasso fertilità totale	,108	,161
Z: tasso movimento naturale pop	,089	,236
Z: t occupaz tot	,138	-,121
Z_UNEMP_R	,097	-,157
Z_LONG_UNEMP_R	,111	-,111
Z_DIFF_UNEMP_GA	,088	-,183

INDICATORI	PESI
Z: SM(terz_edu)	0,07278866
Z: SM(formaz_ad)	0,07866984
Z: SM(asp_vita)	0,06566662
Z_MORTI_INC	0,07192207
Z: pil ppt pc (% media)	0,07956638
Z: SM(reddito_disponibile_netto)	0,07733986
Z: %ore non lavorate	0,07841683
Z_IND_VECC	0,02310943
Z: tra 0 e 9	0,04971817
Z: tasso fertilità totale	0,06893293
Z: tasso movimento naturale pop	0,05659262
Zoccupaz	0,08829520
Z_UNEMP_R	0,06210602
Z_LONG_UNEMP_R	0,07073078
Z_DIFF_UNEMP_GA	0,05614462

Pertanto si procede a determinare i gradi di appartenenza di ciascuna regione al gruppo delle regioni ad elevato benessere calcolando la media semplice e quella ponderata degli indicatori *fuzzy*, calcolati con la funzione lineare, esponenziale e cubica, ricavando altrettante classifiche. Per verificare il grado di somiglianza tra le classifiche così ottenute calcoliamo la matrice dei coefficienti di correlazione sui gradi di appartenenza ottenuti con le diverse funzioni con e senza ponderazione delle variabili.

Correlazioni

		F5_9NP	F5_9P	FCUB_NP	FCUB_P	F_ZIM_NP	F_ZIM_P
F5_9NP	Correlazione di Pearson	1	,992**	,995**	,986**	,953**	,944**
	Sig. (2-code)		,000	,000	,000	,000	,000
	N	260	260	260	260	260	260
F5_9P	Correlazione di Pearson	,992**	1	,987**	,995**	,949**	,954**
	Sig. (2-code)	,000		,000	,000	,000	,000
	N	260	260	260	260	260	260
FCUB_NP	Correlazione di Pearson	,995**	,987**	1	,991**	,939**	,931**
	Sig. (2-code)	,000	,000		,000	,000	,000
	N	260	260	260	260	260	260
FCUB_P	Correlazione di Pearson	,986**	,995**	,991**	1	,936**	,941**
	Sig. (2-code)	,000	,000	,000		,000	,000
	N	260	260	260	260	260	260
F_ZIM_NP	Correlazione di Pearson	,953**	,949**	,939**	,936**	1	,993**
	Sig. (2-code)	,000	,000	,000	,000		,000
	N	260	260	260	260	260	260
F_ZIM_P	Correlazione di Pearson	,944**	,954**	,931**	,941**	,993**	1
	Sig. (2-code)	,000	,000	,000	,000	,000	
	N	260	260	260	260	260	260

** La correlazione è significativa al livello 0,01 (2-code).

I coefficienti di correlazione calcolati tra le varie coppie mostrano livelli molto elevati, in particolare le classifiche ottenute tramite la funzione lineare e quella cubica sono somiglianti al 99,5%, indipendentemente dal fatto che le variabili siano o meno ponderate.

La stabilità dei risultati ottenuti con metodi diversi è indice della validità della classificazione e quindi garanzia dell'esistenza di una reale struttura sottostante ai dati. Tale grande rassomiglianza negli esiti delle diverse analisi segnala anche che la scelta della funzione che traduce i gradi di appartenenza tra 0 e 1 ha scarsa influenza sui risultati, nel caso in esame. In un'indagine ulteriore, si potrebbe inoltre valutare la "sensitività" dell'analisi al variare delle soglie inferiore e superiore.

Si riporta anche la matrice di cograduazione ottenendo valori dell'indice Rho di Spearman molto simili a quelli del coefficiente di correlazione.

Correlazioni

			F5_9NP	F5_9P	FCUB_NP	FCUB_P	F_ZIM_NP	F_ZIM_P
Rho di Spearman	F5_9NP	Coefficiente di correlazione	1,000	,991**	,995**	,987**	,967**	,967**
		Sig. (2-code)	.	,000	,000	,000	,000	,000
		N	257	257	257	257	257	257
	F5_9P	Coefficiente di correlazione	,991**	1,000	,987**	,995**	,961**	,973**
		Sig. (2-code)	,000	.	,000	,000	,000	,000
		N	257	257	257	257	257	257
	FCUB_NP	Coefficiente di correlazione	,995**	,987**	1,000	,991**	,955**	,956**
		Sig. (2-code)	,000	,000	.	,000	,000	,000
		N	257	257	257	257	257	257
	FCUB_P	Coefficiente di correlazione	,987**	,995**	,991**	1,000	,950**	,964**
		Sig. (2-code)	,000	,000	,000	.	,000	,000
		N	257	257	257	257	257	257
	F_ZIM_NP	Coefficiente di correlazione	,967**	,961**	,955**	,950**	1,000	,992**
		Sig. (2-code)	,000	,000	,000	,000	.	,000
		N	257	257	257	257	257	257
	F_ZIM_P	Coefficiente di correlazione	,967**	,973**	,956**	,964**	,992**	1,000
		Sig. (2-code)	,000	,000	,000	,000	,000	.
		N	257	257	257	257	257	257

** La correlazione è significativa al livello 0,01 (2-code).

Di seguito si riportano i gradi di appartenenza per ciascuna regione in base alle tre funzioni di appartenenza senza ponderazione delle variabili e con ordinamento decrescente in base alla funzione lineare.

CODE	LABEL	COUNTRY	F_lineare	F_cubica	F_esponenziale
NL31	Utrecht	NL	0,88821	0,91079	0,89034
UKJ1	Berkshire, Buckinghamshire and Oxfordshire	UK	0,86386	0,87453	0,86086
SE11	Stockholm	SE	0,80783	0,80413	0,83519
NL32	Noord-Holland	NL	0,80096	0,80525	0,84695
UKH2	Bedfordshire and Hertfordshire	UK	0,77518	0,81415	0,82220
NL33	Zuid-Holland	NL	0,74166	0,80304	0,82168
NL23	Flevoland	NL	0,71998	0,72540	0,80792
UKK1	Gloucestershire, Wiltshire and Bristol/Bath area	UK	0,70080	0,76834	0,80150
NL21	Overijssel	NL	0,69084	0,70091	0,78933
DK01	Hovedstaden	DK	0,68957	0,72001	0,77391
UKJ2	Surrey, East and West Sussex	UK	0,68613	0,68308	0,79688
FI20	Åland	FI	0,68240	0,67261	0,80315
NL22	Gelderland	NL	0,67269	0,73500	0,79897
NL41	Noord-Brabant	NL	0,66347	0,71474	0,79272
UKJ3	Hampshire and Isle of Wight	UK	0,65433	0,67120	0,77816
UKM5	North Eastern Scotland	UK	0,64101	0,61008	0,76574
NL12	Friesland (NL)	NL	0,63914	0,65790	0,76725
DK04	Midtjylland	DK	0,63153	0,68054	0,73360
FR10	Île de France	FR	0,59964	0,59986	0,73005
AT34	Vorarlberg	AT	0,59786	0,60604	0,71565
UKK2	Dorset and Somerset	UK	0,58154	0,65668	0,70804
NL34	Zeeland	NL	0,58099	0,58993	0,75557
UKH3	Essex	UK	0,56580	0,59029	0,74047
DE11	Stuttgart	DE	0,56103	0,58708	0,67319
NL11	Groningen	NL	0,55749	0,58288	0,73420
DE21	Oberbayern	DE	0,55303	0,54806	0,68295
SE23	Västsverige	SE	0,54271	0,54319	0,70835
ITD1	Provincia Autonoma di Bolzano/Bozen	IT	0,54062	0,53734	0,69697
SE22	Sydsverige	SE	0,53619	0,54232	0,69671
DE14	Tübingen	DE	0,53247	0,54147	0,67586
UKF2	Leicestershire, Rutland and Northamptonshire	UK	0,52858	0,48906	0,73589
AT33	Tirol	AT	0,51909	0,49627	0,68203
UKD4	Lancashire	UK	0,51900	0,53343	0,70104
IE02	Southern and Eastern	IE	0,51352	0,52000	0,64588
UKJ4	Kent	UK	0,51235	0,55763	0,71404
LU00	Luxembourg	LU	0,51088	0,48558	0,71337
AT32	Salzburg	AT	0,50186	0,48683	0,65000
DE13	Freiburg	DE	0,50077	0,48473	0,63813
BE24	Prov. Vlaams-Brabant	BE	0,50062	0,45875	0,71606
DK03	Syddanmark	DK	0,49926	0,51431	0,67684
NL13	Drenthe	NL	0,49898	0,51720	0,71246
UKH1	East Anglia	UK	0,49743	0,48549	0,72076
FI18	Etelä-Suomi	FI	0,49131	0,48620	0,71234
UKD3	Greater Manchester	UK	0,48098	0,49446	0,67807
DE12	Karlsruhe	DE	0,48041	0,47959	0,61857
DE71	Darmstadt	DE	0,47297	0,47861	0,64562
DK02	Sjælland	DK	0,46921	0,43983	0,62590
CZ01	Praha	CZ	0,46254	0,46727	0,59482
DK05	Nordjylland	DK	0,46238	0,46236	0,65735

CODE	LABEL	COUNTRY	F_lineare	F_cubica	F_esponenziale
ES30	Comunidad de Madrid	ES	0,45985	0,46117	0,59659
UKL2	East Wales	UK	0,45685	0,45306	0,70519
AT31	Oberösterreich	AT	0,44689	0,42646	0,64994
DE25	Mittelfranken	DE	0,44588	0,48503	0,59472
UKE4	West Yorkshire	UK	0,44487	0,42997	0,67749
UKM2	Eastern Scotland	UK	0,44406	0,44631	0,68956
NL42	Limburg (NL)	NL	0,44360	0,47433	0,66686
DE26	Unterfranken	DE	0,44329	0,44973	0,58763
FI1A	Pohjois- ja Itä-Suomi	FI	0,44158	0,42594	0,63894
UKK4	Devon	UK	0,43905	0,43640	0,65173
UKE2	North Yorkshire	UK	0,43446	0,44804	0,66662
UKG1	Herefordshire, Worcestershire and Warwickshire	UK	0,43414	0,42580	0,67290
DE27	Schwaben	DE	0,43160	0,45080	0,59960
SE21	Småland med öarna	SE	0,42806	0,41159	0,64490
UKG3	West Midlands	UK	0,42017	0,43810	0,61943
CY00	Kýpros	CY	0,41958	0,39670	0,62731
FR71	Rhône-Alpes	FR	0,41917	0,40264	0,67222
FR42	Alsace	FR	0,41166	0,41199	0,68222
BE21	Prov. Antwerpen	BE	0,40861	0,41137	0,67301
ITD2	Provincia Autonoma di Trento	IT	0,40653	0,39848	0,62748
SE12	Östra Mellansverige	SE	0,40132	0,40357	0,64857
DE60	Hamburg	DE	0,39993	0,36700	0,61731
UKF3	Lincolnshire	UK	0,39643	0,39753	0,63714
DE23	Oberpfalz	DE	0,39529	0,43338	0,55576
UKG2	Shropshire and Staffordshire	UK	0,39494	0,36409	0,67632
DE22	Niederbayern	DE	0,39272	0,42723	0,54700
IE01	Border, Midland and Western	IE	0,38364	0,39328	0,53192
UKN0	Northern Ireland	UK	0,38102	0,34852	0,65421
FR51	Pays de la Loire	FR	0,37851	0,35183	0,63631
DEB2	Trier	DE	0,37340	0,38332	0,56452
SE33	Övre Norrland	SE	0,36980	0,38246	0,60836
ES21	País Vasco	ES	0,36247	0,34940	0,55391
ES22	Comunidad Foral de Navarra	ES	0,36238	0,34677	0,60783
ITC2	Valle d'Aosta/Vallée d'Aoste	IT	0,36138	0,34398	0,56004
BE31	Prov. Brabant Wallon	BE	0,36030	0,38020	0,59981
AT13	Wien	AT	0,35979	0,31022	0,62735
AT12	Niederösterreich	AT	0,35677	0,34489	0,54622
SE32	Mellersta Norrland	SE	0,35383	0,31790	0,57226
UKC1	Tees Valley and Durham	UK	0,35223	0,31468	0,62376
AT22	Steiermark	AT	0,35078	0,34610	0,54781
DE50	Bremen	DE	0,34647	0,33793	0,53306
BE23	Prov. Oost-Vlaanderen	BE	0,34213	0,33415	0,65166
UKF1	Derbyshire and Nottinghamshire	UK	0,34188	0,33796	0,65583
UKM6	Highlands and Islands	UK	0,34037	0,33181	0,59918
UKD1	Cumbria	UK	0,33987	0,30831	0,59258
DEA1	Düsseldorf	DE	0,33794	0,33487	0,53193
UKE3	South Yorkshire	UK	0,33282	0,36985	0,62185
UKM3	South Western Scotland	UK	0,33199	0,31560	0,60381
DEA2	Köln	DE	0,33059	0,32914	0,57102

CODE	LABEL	COUNTRY	F_lineare	F_cubica	F_esponenziale
AT21	Kärnten	AT	0,32971	0,31845	0,54026
DEB1	Koblenz	DE	0,32508	0,33394	0,51812
BE25	Prov. West-Vlaanderen	BE	0,32335	0,29744	0,59092
FR52	Bretagne	FR	0,31968	0,25649	0,60467
FR30	Nord - Pas-de-Calais	FR	0,31861	0,32688	0,50089
SK01	Bratislavský kraj	SK	0,31861	0,29198	0,54656
DEA4	Detmold	DE	0,31730	0,29391	0,56184
DE24	Oberfranken	DE	0,31670	0,32782	0,51919
FR24	Centre	FR	0,31197	0,28242	0,59395
AT11	Burgenland (AT)	AT	0,31091	0,32324	0,47210
SE31	Norra Mellansverige	SE	0,30932	0,29167	0,56671
DE73	Kassel	DE	0,30586	0,33633	0,52431
FR23	Haute-Normandie	FR	0,30282	0,28183	0,53622
DEF0	Schleswig-Holstein	DE	0,30219	0,31861	0,53413
BE34	Prov. Luxembourg (BE)	BE	0,30190	0,31462	0,54250
DEA3	Münster	DE	0,29903	0,28440	0,53993
ITC4	Lombardia	IT	0,29583	0,27991	0,53215
FR43	Franche-Comté	FR	0,29574	0,31243	0,56980
ES51	Cataluña	ES	0,29561	0,27243	0,52301
DEC0	Saarland	DE	0,29521	0,28260	0,47684
DEB3	Rheinhessen-Pfalz	DE	0,29443	0,34451	0,54635
DE72	Gießen	DE	0,29405	0,30068	0,55164
UKE1	East Yorkshire and Northern Lincolnshire	UK	0,29103	0,25177	0,60792
FI19	Länsi-Suomi	FI	0,29081	0,24387	0,61037
ITD5	Emilia-Romagna	IT	0,29020	0,31166	0,46702
FR62	Midi-Pyrénées	FR	0,28835	0,25341	0,58986
DEA5	Arnsberg	DE	0,28614	0,27485	0,47840
BE22	Prov. Limburg (BE)	BE	0,28553	0,27315	0,60418
FR22	Picardie	FR	0,28224	0,27298	0,46685
DE94	Weser-Ems	DE	0,27356	0,26598	0,53595
DE30	Berlin	DE	0,27286	0,26630	0,49286
UKK3	Cornwall and Isles of Scilly	UK	0,27209	0,23893	0,55640
ES53	Illes Balears	ES	0,27025	0,26221	0,48082
DE93	Lüneburg	DE	0,26990	0,30631	0,49357
ES62	Región de Murcia	ES	0,26832	0,23063	0,46118
EL42	Notio Aigaio	EL	0,26794	0,26971	0,47306
UKL1	West Wales and The Valleys	UK	0,26651	0,30378	0,55405
DE92	Hannover	DE	0,26171	0,25357	0,51773
FR82	Provence-Alpes-Côte d'Azur	FR	0,26114	0,25664	0,54550
DED3	Leipzig	DE	0,26010	0,26342	0,46106
DED2	Dresden	DE	0,25693	0,26302	0,48887
SI02	Zahodna Slovenija	SI	0,25563	0,22307	0,57689
DE91	Braunschweig	DE	0,25373	0,21772	0,49116
ES70	Canarias	ES	0,24757	0,25766	0,39964
ITD4	Friuli-Venezia Giulia	IT	0,24004	0,25722	0,44072
ITD3	Veneto	IT	0,23903	0,25962	0,50795
PT20	Região Autónoma dos Açores	PT	0,23730	0,18179	0,46547
ITC3	Liguria	IT	0,23096	0,22223	0,43887
CZ02	Střední Čechy	CZ	0,22833	0,20492	0,43742
UKC2	Northumberland and Tyne and Wear	UK	0,22817	0,21923	0,58639

CODE	LABEL	COUNTRY	F_lineare	F_cubica	F_esponenziale
FR21	Champagne-Ardenne	FR	0,22668	0,24703	0,51079
RO21	Nord-Est	RO	0,22033	0,26424	0,35491
ES23	La Rioja	ES	0,21393	0,17544	0,44718
FR25	Basse-Normandie	FR	0,21234	0,17619	0,53394
SK04	Východné Slovensko	SK	0,21176	0,20917	0,31147
FR83	Corse	FR	0,21116	0,19921	0,49254
BE35	Prov. Namur	BE	0,20984	0,24131	0,45324
DED1	Chemnitz	DE	0,20400	0,20255	0,42291
FR41	Lorraine	FR	0,20389	0,22708	0,53184
PT30	Região Autónoma da Madeira	PT	0,20032	0,15971	0,49386
FR61	Aquitaine	FR	0,19958	0,16348	0,53998
ITE3	Marche	IT	0,19858	0,16508	0,45160
ES13	Cantabria	ES	0,19220	0,19084	0,39676
PT17	Lisboa	PT	0,19197	0,16240	0,49136
ES24	Aragón	ES	0,19176	0,16738	0,41967
BE32	Prov. Hainaut	BE	0,19116	0,22661	0,37986
ES61	Andalucía	ES	0,18821	0,19448	0,40736
ITE1	Toscana	IT	0,18765	0,20948	0,42423
FR81	Languedoc-Roussillon	FR	0,18675	0,19389	0,42843
EL43	Kriti	EL	0,18408	0,19312	0,39707
MT00	Malta	MT	0,18153	0,18441	0,47938
ES52	Comunidad Valenciana	ES	0,17823	0,17675	0,43131
RO32	București - Ilfov	RO	0,17773	0,15608	0,38063
BE33	Prov. Liège	BE	0,17567	0,18633	0,44340
PL63	Pomorskie	PL	0,16685	0,15013	0,38285
FR53	Poitou-Charentes	FR	0,16551	0,15611	0,51075
EL30	Attiki	EL	0,16398	0,14920	0,41355
DEG0	Thüringen	DE	0,16259	0,14502	0,43424
PL12	Mazowieckie	PL	0,15370	0,15401	0,43674
FR26	Bourgogne	FR	0,15198	0,14610	0,48710
ITE4	Lazio	IT	0,14753	0,17199	0,40128
FR63	Limousin	FR	0,14473	0,13263	0,46061
ITC1	Piemonte	IT	0,14469	0,14124	0,38438
ES41	Castilla y León	ES	0,14226	0,13442	0,32806
CZ03	Jihozápad	CZ	0,13290	0,13114	0,37631
BG41	Yugozapaden	BG	0,13226	0,10176	0,38339
DEE0	Sachsen-Anhalt	DE	0,13086	0,13151	0,37106
ITE2	Umbria	IT	0,12811	0,09639	0,40480
DE80	Mecklenburg-Vorpommern	DE	0,12651	0,12794	0,36843
PL62	Warmińsko-Mazurskie	PL	0,12640	0,09529	0,31196
FR72	Auvergne	FR	0,12625	0,09702	0,49913
ES12	Principado de Asturias	ES	0,12286	0,10114	0,29976
ITF3	Campania	IT	0,12260	0,09933	0,32459
PL41	Wielkopolskie	PL	0,11953	0,11805	0,35095
ES42	Castilla-La Mancha	ES	0,10824	0,08611	0,33854
RO42	Vest	RO	0,10641	0,11907	0,23923
RO11	Nord-Vest	RO	0,10415	0,08566	0,27660
EE00	Eesti	EE	0,09666	0,07850	0,33979
PL43	Lubuskie	PL	0,09410	0,07769	0,28865
PL21	Małopolskie	PL	0,08867	0,07754	0,33921
ES11	Galicia	ES	0,08796	0,07902	0,27930
ITF6	Calabria	IT	0,08667	0,07927	0,27725

CODE	LABEL	COUNTRY	F_lineare	F_cubica	F_esponenziale
PL32	Podkarpackie	PL	0,08624	0,07556	0,27878
SI01	Vzhodna Slovenija	SI	0,08048	0,10051	0,39861
ITG2	Sardegna	IT	0,07970	0,07320	0,25320
PL61	Kujawsko-Pomorskie	PL	0,07782	0,07084	0,27500
PT16	Centro (PT)	PT	0,07703	0,06757	0,32521
ITF4	Puglia	IT	0,07370	0,06920	0,27514
LT00	Lietuva	LT	0,07268	0,06729	0,27172
PT11	Norte	PT	0,07211	0,03663	0,32881
PL42	Zachodniopomorskie	PL	0,06772	0,06705	0,23081
ITF1	Abruzzo	IT	0,06766	0,06700	0,30371
EL21	Ipeiros	EL	0,06667	0,06667	0,18176
SK03	Stredné Slovensko	SK	0,06667	0,06667	0,20546
ITF2	Molise	IT	0,06549	0,06615	0,27072
CZ04	Severozápad	CZ	0,06265	0,06234	0,29183
ITF5	Basilicata	IT	0,06263	0,06526	0,21410
PT15	Algarve	PT	0,05836	0,03233	0,33595
CZ05	Severovýchod	CZ	0,05613	0,02297	0,36396
HU10	Közép-Magyarország	HU	0,05569	0,02705	0,33172
RO12	Centru	RO	0,05337	0,05943	0,19947
RO41	Sud-Vest Oltenia	RO	0,05185	0,02267	0,21873
HU32	Észak-Alföld	HU	0,05169	0,05710	0,17424
ITG1	Sicilia	IT	0,04799	0,02167	0,28105
CZ06	Jihovýchod	CZ	0,04755	0,02135	0,37831
ES43	Extremadura	ES	0,04375	0,02362	0,26920
PL31	Lubelskie	PL	0,03966	0,05249	0,27086
PL52	Opolskie	PL	0,03669	0,01503	0,24186
CZ08	Moravskoslezsko	CZ	0,03577	0,01524	0,30528
EL22	Ionia Nisia	EL	0,03434	0,04373	0,21403
PL51	Dolnośląskie	PL	0,03337	0,03668	0,26042
SK02	Západné Slovensko	SK	0,03259	0,02396	0,19508
PL34	Podlaskie	PL	0,02889	0,01630	0,25564
BG34	Yugoiztochen	BG	0,02765	0,01158	0,24692
RO22	Sud-Est	RO	0,02597	0,01318	0,19736
PL11	Łódzkie	PL	0,02326	0,00855	0,25378
EL41	Voreio Aigaio	EL	0,02264	0,01052	0,25827
LV00	Latvija	LV	0,02176	0,00828	0,19956
PL33	Świętokrzyskie	PL	0,02037	0,00900	0,19993
PL22	Śląskie	PL	0,02017	0,00872	0,27731
CZ07	Střední Morava	CZ	0,01790	0,00667	0,31857
BG33	Severoiztochen	BG	0,01707	0,00613	0,23800
BG31	Severozapaden	BG	0,01704	0,00642	0,21110
EL11	Anatoliki Makedonia, Thraki	EL	0,01509	0,00607	0,19309
HU21	Közép-Dunántúl	HU	0,01416	0,00532	0,19450
BG32	Severen tsentralen	BG	0,01284	0,00499	0,18665
EL12	Kentriki Makedonia	EL	0,01207	0,00464	0,26738
EL14	Thessalia	EL	0,01132	0,00431	0,23920
EL25	Peloponnisos	EL	0,01132	0,00431	0,21962
EL24	Stereia Ellada	EL	0,00732	0,00258	0,19629
PT18	Alentejo	PT	0,00513	0,00181	0,18151
HU31	Észak-Magyarország	HU	0,00466	0,00163	0,18008
EL23	Dytiki Ellada	EL	0,00202	0,00069	0,19157
BG42	Yuzhen tsentralen	BG	0,00000	0,00000	0,17413
EL13	Dytiki Makedonia	EL	0,00000	0,00000	0,19314
HU22	Nyugat-Dunántúl	HU	0,00000	0,00000	0,17390
HU23	Dél-Dunántúl	HU	0,00000	0,00000	0,13875
HU33	Dél-Alföld	HU	0,00000	0,00000	0,14915
RO31	Sud - Muntenia	RO	0,00000	0,00000	0,19941

CAPITOLO 6: CONFRONTI TRA I RISULTATI DEI DIVERSI METODI ED OSSERVAZIONI CONCLUSIVE

In questo capitolo conclusivo effettuiamo alcuni confronti tra i principali risultati ottenuti con i diversi metodi, allo scopo di farne emergere le analogie e le differenze.

Il coefficiente di correlazione calcolato tra l'indicatore sintetico *fuzzy* dei gradi di appartenenza all'insieme delle regioni benestanti, ottenuti tramite funzione lineare, ed i punteggi relativi alla prima componente principale risulta pari a 0,93: ciò dimostra una forte somiglianza tra risultati ottenuti con questi metodi, cioè l'analisi condotta con l'approccio dei *fuzzy sets* e l'analisi delle componenti principali. Il calcolo dell'indice Rho di Spearman, uguale a 0,95, conferma quanto sopra affermato: le graduatorie delle regioni europee, dalla migliore alla peggiore, sono molto simili.

Si può quindi ritenere che entrambe le suddette indagini – con differenti procedure metodologiche – siano state in grado, nel loro complesso, di porre in luce la reale “struttura” delle situazioni di benessere nelle regioni europee, misurabile attraverso un opportuno insieme di variabili.

Correlazioni

		F_L16NP	REGR factor score 1 for analysis 1
F_L16NP	Correlazione di Pearson	1	,931**
	Sig. (2-code)		,000
	N	257	257
REGR factor score 1 for analysis 1	Correlazione di Pearson	,931**	1
	Sig. (2-code)	,000	
	N	257	257

** La correlazione è significativa al livello 0,01 (2-code).

Un altro interessante confronto è quello che scaturisce dal calcolo del coefficiente di correlazione tra l'indicatore dei gradi di appartenenza all'insieme delle regioni benestanti, ottenuti tramite funzione lineare, e le seguenti variabili (in termini di scostamenti standardizzati):

- Pil procapite a parità di potere d'acquisto;
- reddito familiare disponibile netto.

Nel primo caso l'indice risulta pari allo 0,7, mentre nel secondo caso sfiora quota 0,6, confermando quanto emerge dagli studi precedenti, cioè che la considerazione unica di tali indicatori non è sufficiente a spiegare il complesso ed articolato fenomeno del benessere.

		Correlazioni		
		F_L16NP	Z: pil ppt pc (% media)	Z: SM(reddito_disponibile_netto)
F_L16NP	Correlazione di Pearson	1	,712**	,596**
	Sig. (2-code)		,000	,000
	N	257	257	257
Z: pil ppt pc (% media)	Correlazione di Pearson	,712**	1	,698**
	Sig. (2-code)	,000		,000
	N	257	257	257
Z: SM(reddito_disponibile_netto)	Correlazione di Pearson	,596**	,698**	1
	Sig. (2-code)	,000	,000	
	N	257	257	257

*. La correlazione è significativa al livello 0,01 (2-code).

Anche in questo frangente si provvede al calcolo dell'indice Rho di Spearman sulle medesime coppie di variabili e si osservano sostanzialmente risultati analoghi ai precedenti: le graduatorie delle regioni europee calcolate con i tre criteri sopra indicati presentano tra loro una relazione diretta, ma di entità non particolarmente elevata.

		Correlazioni		
		F_L16NP	Z: pil ppt pc (% media)	Z: SM(reddito_disponibile_netto)
Rho di Spearman	F_L16NP	Coefficiente di correlazione	1,000	,758**
		Sig. (2-code)		,000
		N	257	257
Z: pil ppt pc (% media)		Coefficiente di correlazione	,758**	1,000
		Sig. (2-code)	,000	,000
		N	257	257
Z: SM(reddito_disponibile_netto)		Coefficiente di correlazione	,621**	,742**
		Sig. (2-code)	,000	,000
		N	257	257

** La correlazione è significativa al livello 0,01 (2-code).

Di seguito si mostra la tabella di contingenza tra i quintili dell'indicatore *fuzzy* dei gradi di appartenenza al gruppo delle regioni "benestanti" ed i quintili relativi alle variabili "Pil procapite a parità di potere d'acquisto" e "Reddito familiare disponibile netto".

Tavola di contingenza Percentile Group of F_L16NP * Percentile Group of pil_ppt

Conteggio

		Percentile Group of pil_ppt					Totale
		1	2	3	4	5	
Percentile Group of F_L16NP	1	35	14	1	1	0	51
	2	15	17	11	7	2	52
	3	1	13	14	15	8	51
	4	0	7	16	16	13	52
	5	0	1	9	13	28	51
Totale		51	52	51	52	51	257

Misure simmetriche

		Valore	E.S. asint. ^a	T appross. ^b	Sig. appross.
Ordinale per ordinale	Tau-b di Kendall	,620	,031	20,218	,000
N. di casi validi		257			

a. Senza assumere l'ipotesi nulla.

b. Viene usato l'errore standard asintotico in base all'assunzione dell'ipotesi nulla.

Tavola di contingenza Percentile Group of F_L16NP * Percentile Group of reddito familiare_disponibile_netto_1

Conteggio

		Percentile Group of reddito_disponibile_netto_1					Totale
		1	2	3	4	5	
Percentile Group of F_L16NP	1	33	13	5	0	0	51
	2	17	13	12	8	2	52
	3	1	6	13	17	14	51
	4	0	10	10	18	14	52
	5	0	10	11	9	21	51
Totale		51	52	51	52	51	257

Misure simmetriche

	Valore	E.S. asint. ^a	T appross. ^b	Sig. appross.
Ordinale per ordinale Tau-b di Kendall	,498	,038	13,007	,000
N. di casi validi	257			

a. Senza assumere l'ipotesi nulla.

b. Viene usato l'errore standard asintotico in base all'assunzione dell'ipotesi nulla.

Il calcolo dell'indice Tau di Kendall conferma i risultati precedentemente ottenuti tramite coefficiente di correlazione: l'associazione ordinale nel primo caso è uguale a 0,620 e nel secondo 0,498. La suddivisione in classi in base all'indicatore *fuzzy* ed ai singoli indici di ricchezza presenta analogie, ma non è certamente sovrapponibile.

Un'ulteriore considerazione: calcolando la media del grado di appartenenza all'insieme di regioni ad elevato benessere per ciascuno dei cinque cluster individuati tramite l'analisi *k*-medie, possiamo apprezzare la coerenza nei risultati ottenuti mediante il metodo *fuzzy sets* e l'analisi dei gruppi *k*-medie. Osservando le medie così calcolate e la Figura 50, cioè il diagramma di dispersione degli *scores* divisi per cluster di appartenenza, vediamo che il gruppo 4 si conferma quello che raccoglie le regioni con massimo benessere; i gruppi 5 e 2 corrispondono a regioni caratterizzate da benessere superiore alla media; il gruppo 1 è formato da regioni aventi benessere inferiore alla media ed infine il gruppo 3 corrisponde a regioni con il minimo grado di benessere.

CLUSTER	n	M_GA
1	29	0,143
2	37	0,309
3	61	0,089
4	65	0,491
5	65	0,331

Si è inoltre calcolata la tabella di contingenza considerando i quintili dell'indicatore *fuzzy* dei gradi di appartenenza delle regioni al sottoinsieme delle regioni benestanti ed i cinque gruppi omogenei ottenuti con il metodo *k*-medie. Questa classificazione congiunta conferma le caratteristiche di ciascun cluster ottenuto.

Tavola di contingenza Percentile Group of F_L16NP * Numero di cluster del caso

Conteggio

		Numero di cluster del caso					Totale
		1	2	3	4	5	
Percentile Group of F_L16NP	1	10	0	41	0	0	51
	2	15	8	18	0	11	52
	3	4	12	2	7	26	51
	4	0	15	0	22	15	52
	5	0	2	0	36	13	51
Totale		29	37	61	65	65	257

OSSERVAZIONI CONCLUSIVE

Un recentissimo articolo sulla rivista “*Social Indicators Research*” s’intitola “*Beyond GDP: Classifying Alternative Measures for Progress*” (Bleys, 2012). La nostra ricerca s’inserisce nel solco della vastissima letteratura economica e statistica sugli approcci integrativi del semplice PIL pro capite quali misure dei concetti di *economic welfare*, *well-being* e *quality of life*. (Si veda, a titolo d’esempio, l’ampia ed aggiornata rassegna riportata nell’articolo citato).

Segnaliamo inoltre l’articolo di Gualaccini e Bruni (2012) che presenta una ricerca congiunta CNEL-ISTAT ancora in corso per la misurazione del benessere equo e sostenibile con approfondimenti fino a livello regionale.

Le analisi qui condotte con riferimento alle regioni europee, a livello NUTS 2, secondo la classificazione Eurostat, hanno confermato che l’approccio multidimensionale è senza dubbio più soddisfacente e ricco d’informazioni per la valutazione complessiva del benessere. Nel presente lavoro si è cercato anche di ovviare alle principali obiezioni che vengono mosse alla suddetta procedura:

- la scelta soggettiva delle variabili osservate che vengono assunte quali “sintomi” della variabile latente sottostante, cioè il benessere;
- i criteri di sintesi e ponderazione di tali variabili per ottenere uno o più indicatori sintetici, o compositi.

Con riferimento al primo aspetto, si sono quindi considerate le diverse “facce” del concetto poliedrico di benessere, desumibili dalla letteratura ed in particolare dallo schema contenuto nel

“*Compendium of OECD well-being indicators*”. Le categorie di variabili utilizzate sono state le seguenti:

- variabili demografiche;
- salute e incidentalità;
- istruzione
- lavoro;
- ricchezza e tempo libero;
- variabili ambientali.

Per ciascuno di tali aspetti si è condotta un’accurata analisi di tutta la documentazione statistica fornita da Eurostat a livello regionale NUTS 2, effettuando operazioni di *data cleaning* ed eliminando in primo luogo le variabili con troppi valori mancanti. Si è quindi effettuato uno studio di tipo unidimensionale per la singole variabili restanti, proponendone anche una rappresentazione cartografica.

Si sono studiate le relazioni tra le variabili, mediante il calcolo delle matrici di correlazione, allo scopo anche di eliminare le variabili ridondanti, cioè con coefficiente di correlazione molto prossimo all’unità, per cui forniscono sostanzialmente le medesime informazioni. Si è così pervenuti ad un insieme di 31 variabili, articolate nella maniera seguente: variabili demografiche 7; salute e incidentalità 3; istruzione 5; lavoro 6; ricchezza e tempo libero 4; variabili ambientali 6.

Nell’intento di rispondere all’obiezione riguardante i criteri di sintesi e ponderazione per ottenere indicatori compositi, si sono considerate e confrontate diverse procedure:

- analisi delle componenti principali per ciascuna delle categorie del benessere sopra riportate, per ricavare un indicatore sintetico di ogni singolo ambito;
- sintesi di tali indicatori parziali mediante media semplice e media ponderata con il valore dell’indice Alpha di Cronbach di ciascun ambito.
- analisi delle componenti principali applicata all’insieme delle 31 variabili, riferite a tutti gli ambiti, allo scopo di individuare le dimensioni effettivamente sottostanti al *data set* utilizzato;

- definizione d'un sottoinsieme di 16 variabili, in base ai coefficienti di correlazione delle stesse con le prime componenti principali sopra calcolate, che presenta un valore di Alpha di Cronbach, uguale a 0,844, e sintesi, mediante le componenti principali, di tali variabili in due dimensioni, la prima interpretabile come indicatore composito dell'occupazione, della ricchezza, dell'aspettativa di vita e dell'istruzione, la seconda come sintesi degli aspetti demografici ed in particolare della giovinezza della popolazione.

Il confronto i valori degli indicatori sintetici del benessere ottenuti con la varie procedure – effettuato anche mediante la rappresentazione cartografiche degli stessi, con classi corrispondenti ai quintili – ha posto in luce forti analogie: le aree con più elevato benessere (di colore verde scuro) sono molto simili nelle varie mappe ed anche le zone peggiori (di colore rosso) presentano forti rassomiglianze nelle cartine ottenute con i diversi indicatori compositi. I criteri di aggregazione delle variabili osservate e di ponderazione delle medesime non incidono pertanto in maniera significativa sui risultati, che riflettono in ogni caso la reale struttura dei livelli di benessere nelle regioni europee.

La *cluster analysis* con il metodo delle *k*-medie, applicata al sottoinsieme delle 16 variabili più rappresentative con un numero di gruppi pari a 5 (dopo aver eliminato tre regioni metropolitane che risultano anomale: Bruxelles, Inner London e Outer London), ha fornito una sostanziale conferma dei risultati precedenti, pur con un diverso angolo di visuale. Il gruppo migliore (il quarto) è costituito da 65 regioni che presentano una condizione di benessere generalizzato a tutti gli ambiti, con scostamenti rispetto alla media più accentuati per l'istruzione, il tasso di fertilità e l'occupazione; il gruppo peggiore (il terzo), formato da 61 unità, presenta valori di tutte le variabili inferiori alla media e pertanto unisce regioni con situazioni di malessere in tutti gli ambiti esaminati, in particolare rispetto agli indicatori economici, all'aspettativa di vita ed all'istruzione.

Per ciascuno dei cinque gruppi individuati si sono rappresentati nel piano cartesiano gli *scores* delle prime due componenti principali calcolate sul sottoinsieme delle 16 variabili (si veda la figura 50): essa pone in luce una distinzione abbastanza soddisfacente tra i 5 gruppi (le zone di sovrapposizione di punti appartenenti a cluster diversi sono limitate). Inoltre, i punti in verde scuro del gruppo delle 65 regioni migliori si collocano nella parte destra del grafico (elevati *scores* della prima componente) ed i punti delle 61 regioni peggiori (in rosso) si situano chiaramente in una posizione molto a sinistra, mostrando punteggi assai inferiori alla media. Inoltre, i gruppi n. 2 (in verde chiaro) e n. 5 (in giallo) presentano valori degli *scores* della prima componente

approssimativamente nel medesimo *range*, ma si differenziano tra loro per i punteggi della seconda componente, che sintetizza gli aspetti demografici.

Il capitolo 5 presenta il contributo più originale del presente studio: l'applicazione dei *fuzzy sets* per la misura del benessere a livello regionale. Lavori recenti su questo tema sono l'articolo di Bérenger and Verdier-Chouchane (2007) che considera un approccio *fuzzy*, ma a livello di Paesi, e l'articolo di Okulicz-Kozaryn (2012), che fa riferimento ad un sottoinsieme di regioni europee NUTS 2, ma utilizza un modello *multilevel*.

Impiegando ancora le 16 variabili precedenti, si definisce per ciascuna di esse il grado di appartenenza di ognuna delle regioni europee al sottoinsieme *fuzzy* delle regioni con elevato benessere - per brevità "regioni benestanti" - utilizzando diverse funzioni. L'aggregazione di tali funzioni di appartenenza riferite alle singole variabili in un unico indicatore *fuzzy* di benessere avviene sia con il criterio della media semplice, sia con quello delle media ponderata.

Come si è mostrato in questo stesso capitolo, l'approccio *fuzzy* fornisce risultati abbastanza simili, ma non coincidenti, con quelli desunti dalle analisi multidimensionali applicate al medesimo sottoinsieme di variabili, mentre si differenzia in maniera piuttosto sensibile dalla valutazione del benessere economico basata semplicemente sul PIL pro capite o sul reddito familiare disponibile.

Tutto questo consente di riaffermare la superiorità metodologica e la maggiore ricchezza di informazioni dell'insieme di tecniche di analisi dei dati multidimensionali utilizzate in questa tesi – componenti principali, *cluster analysis*, indicatori compositi *fuzzy* – per la misura del benessere a livello territoriale.

BIBLIOGRAFIA

- Abbott, P., Wallace, C. (2011). Social quality: a way to measure the quality of society. *Social Indicators Research*, 108, 153-167.
- Anand, P. Hunter, G. Smith, R. (2005). Evidence based on the Sen-Nussbaum approach to welfare. *Social Indicators Research*, 74,(1), 9-55.
- Andelman, R., Board, R., Cummis, R. et al. (1988). *Quality of life definition and terminology: A discussion document from the International Society of Quality of Life Studies*. Blacksburg: ISQLS.
- Atkinson, A.B., Marlier, E. (2010). *Income and living conditions in Europe*. Luxembourg: Publications office of the European Union.
- Baliamoune-Lutz, M. (2006). On the measurement of human well-being: fuzzy set theory and Sen's capability approach. In M. McGillivray and M. Clarke (eds). *Understanding Human Well-being*. New York: United Nation University Press.
- Bérenger, V., Verdier-Couchane, A. (2007), Multidimensional Measures of Well-Being: Standard of Living and Quality of Life Across Countries, *World Development*, 35, (7), 1259-1276.
- Bleys, B. (2012), Beyond GDP: Classifying Alternative Measures for Progress, *Social Indicators Research*, 109, 355-376.
- Bruni, L., Porta, P. (2004). *Felicità ed economia*. Guerini e Associati.
- Bruni, L. (2004). *L'economia, la felicità e gli altri*. Città nuova.
- Cerioli, A., Zani, S. (1990). A fuzzy approach to the measurement of poverty. In C. Dagum and M. Zenga (eds). *Income and Wealth Distribution, Inequality and Poverty*. Berlin: Springer, 272-284.
- Cheli, B., Lemmi, A. (1995). A totally fuzzy and relative approach to the multidimensional analysis of poverty. *Economic Notes*, 24(1), 115-134.
- Chiappero Martinetti, E. (2000). A multidimensional assessment of well-being based on Sen's functioning approach. *Rivista internazionale di Scienze Sociali*, 108, (2), 207-239.
- Chiappero Martinetti, E., Pareglio, S. (2009). *Sviluppo umano sostenibile e qualità della vita*. Roma: Carocci.
- Cobb, C.W., Cobb, J.B. (1994). *The green national product: a proposed index of sustainable economic welfare*. Lanham: University Press of America.
- Cronbach, L. J. (1951). *Coefficient alpha and the internal structure of tests*. *Psychometrika*, 16(3), 297-334.

- Delbono, F., Lanzi, D. (2007). *Povert  di che cosa? Risorse, opportunit , capacit *. Bologna: Il Mulino.
- Dempster, A.P., Rubin, D.B. (1983). *Incomplete data in sample surveys*. New York: Academic Press.
- De Muro, P., Mazziotta, M., Pareto, A. (2011). Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, 104, (1), 1-18.
- Easterlin, R. (2010). *Happiness, Growth and the life cycle*. New York: Oxford University Press.
- Egidi, V., Giovannini, E. (1998). *Sistemi informativi integrati per l'analisi di fenomeni complessi e multidimensionali*. Roma: Atti IV Conferenza Nazionale di Statistica, 11-13 Novembre 1998.
- Eurostat Statistical Books (2010). *Eurostat Regional Yearbook 2010*. Luxembourg: Publications office of the European Union.
- Eurostat Statistical Books (2011). *European Economic Statistics*. Luxembourg: Publications office of the European Union.
- European Methodologies and working papers (2007). *Regions in the European Union, Nomenclature of territorial units for statistics*. Publications office of the European Commission.
- Frey, B., Stutzer, A. (2006). *Economia e felicit *. Il Sole24ore.
- Gualaccini, G.P., Bruni, S. (2013) Dal Pil al Bes: il significato e la misurazione del benessere equo e sostenibile. *Statistica e Societ *, 1, (3), 36-38.
- Giovannini, E. (2010). *Le statistiche economiche*. Bologna: Il Mulino.
- Kahneman, D. (2007). *Economia della felicit *. Il Sole24ore.
- Klir, G.J., Folger, T.A. (1988). *Fuzzy sets, Uncertainty and Information*. London: Prentice Hall International.
- Joint Research Centre of European Commission, Oecd (2008). *Handbook on constructing composite indicators*. Oecd Publishing.
- Lemmi, A., Betti, G. (2006). *Fuzzy set approach to multidimensional Poverty measurement*. New York: Springer.
- Mazziotta, M., Pareto, A. (2010). Measuring quality of life: an approach based on the non-substitutability of indicators. *Statistica e Applicazioni*, 8, (2), 169-180.
- Munda, G. (2011). Choosing Aggregation Rules for Composite Indicators. *Social Indicators Research*, 109, (3), 337-354.
- Munda, G., Nardo, M. (2005). *Constructing Consistent Composite Indicators: the Issue of Weights*. Institute for the protection and security of the citizen, Joint Research Centre.

- Nordhaus, W.D., Tobin, J. (1972). *Is growth obsolete?* New York: Columbia University Press.
- Oecd (2011). *Compendium of Oecd well-being indicators*. Oecd Publishing.
- Okulicz-Kozaryn, A. (2011). Income and Well-being Across European Provinces. *Social Indicators Research*, 106, 371-392.
- Osberg, L., Sharpe, A. (2002). An index of economic well-being for selected Oecd countries. *Review of Income and Wealth*.
- Pacheco, J., Casado, S., Porras, S. (2012). Exact methods for variable selection in principal component analysis: guide function and pre-selection. *Computational Statistics and Data Analysis* 57, 95-111.
- Pearce, D.W., Hamilton, K., Atkinson, G. (1996). Measuring sustainable development: progress on indicator. *Environment and Development Economics*.
- Pharr, S., Putnam, R. (2000). *Disaffected democracies*. Princeton: University Press.
- Piccolo, D. (2010). *Statistica per le decisioni*. Bologna: Il Mulino.
- Plebe, A. (1957). *Aristotele. Etica nicomachea*. Bari: Laterza.
- Qizilbash, M., Clark, D.A. (2005). The capability approach and fuzzy poverty measures: an application to the south African context. *Social Indicators Research*, 74, 103-139.
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Saltelli, A., Jesinghaus, J., Munda, G. (2007). *Well Being Stories*. European Commission, Joint Research Centre. Brussels: Contribution to the Expert Workshop at the Beyond GDP Conference.
- Scitovsky, P. (1976). *The joyless economy. The Joyless Economy: An inquiry into human satisfaction and consumer dissatisfaction*. Oxford: University Press.
- Segre, E., Rondinella, T., Mascherini, M. (2011). Well-being in italian regions. Measures, Civil society Consultation and Evidence. *Social Indicators Research*, 102, 47-69.
- Sen, A.K. (1985). Capability and Well-being. Nussbaum, M. Sen, A.K. (eds). *The Quality of Life*. London: Clarendon Press.
- Stiglitz, J., Sen, A., Fitoussi, J-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*.
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.
- United Nations Environment Programme, International Institute of Sustainable Development (2004). *Exploring the Links. Human well-being, poverty and ecosystems services*. Nairobi: UNEP-IISD.

- Vendramin, L., Campello, R.J.G.B., Hruschka, E.R. (2010). Relative clustering validity criteria: A comparative overview. *Statistical Analysis and Data Mining*, 3, 4, 209-235.
- Zadeh, L.A. (1965). *Fuzzy sets. Information and Control*, 8, 338-353.
- Zani, S., Cerioli, A. (2007). *Analisi dei dati e data mining per le decisioni aziendali*. Giuffrè editore, Milano.
- Zani, S., Milioli, M. A., Morlini, I. (2010). Fuzzy composite indicators: An application for measuring customer satisfaction. *Proceedings of the Meeting of Italian Statistical Society*. Padova.
- Zani, S., Milioli, M. A., Morlini, I. (2012). Fuzzy methods and satisfaction indices. *Modern Analysis of Customer Surveys: with applications using R.*, 339-456. New York: Wiley.
- Zapf, W., Noll, H. (1984.) *Social Indicators Research: societal monitoring and social reporting. Trends and perspective in empirical social research*. Berlin.
- Zhou, P., Ang, B.W., Zhou, D. Q. (2010). Weighting and aggregation in composite indicator construction: a multiplicative Optimization Approach. *Social Indicators Research*, 96, 1, 169-181.
- Zimmerman, H.J. (2001). *Fuzzy set theory and its application*, 4th edn. Boston: Kluwer.

INDICE DELLE FIGURE

Figura 1. Tipologia di Zapf sulle condizioni di vita	12
Figura 2. Il processo di determinazione del benessere	15
Figura 3. Struttura degli indicatori di benessere individuati dall'Oecd	17
Figura 4. Regioni europee NUTS 2	22
Figura 5. Classificazione NUTS per ciascun paese membro.....	23
Figura 6. DEMOGRAFIA: INDICE DI DIPENDENZA.....	62
Figura 7. DEMOGRAFIA: INDICE DI VECCHIAIA.....	63
Figura 8. DEMOGRAFIA: PERCENTUALE POPOLAZIONE OVER 74.....	64
Figura 9. DEMOGRAFIA: PERCENTUALE POPOLAZIONE UNDER 10	65
Figura 10. DEMOGRAFIA: TASSO DI FERTILITA'	66
Figura 11. DEMOGRAFIA: TASSO NETTO MOVIMENTO MIGRATORIO.....	67
Figura 12. DEMOGRAFIA: TASSO NETTO MOVIMENTO NATURALE.....	68
Figura 13. SALUTE: ASPETTATIVA DI VITA	69
Figura 14. SALUTE: TASSO DI MORTALITA' INFANTILE.....	70
Figura 15. SALUTE: TASSO MORTALITA' INCIDENTI STRADALI	71
Figura 16. RICCHEZZA E TEMPO LIBERO: ORE DI NON LAVORO.....	72
Figura 17. RICCHEZZA E TEMPO LIBERO: PIL A PARITA' DI POTERE D'ACQUISTO	73
Figura 18. RICCHEZZA E TEMPO LIBERO: VEICOLI POSSEDUTI PER 1000 ABITANTI.....	74
Figura 19. RICCHEZZA E TEMPO LIBERO: REDDITO FAMILIARE NETTO DISPONIBILE	75
Figura 20. ISTRUZIONE: PERCENTUALE ISTRUITI DI MEDIO LIVELLO	76
Figura 21. ISTRUZIONE: PERCENTUALE ISTRUITI DI ALTO LIVELLO	77
Figura 22. ISTRUZIONE: PERCENTUALE FORMAZIONE PERMANENTE ADULTI	78
Figura 23. ISTRUZIONE: PERCENTUALE SPESA RICERCA E SVILUPPO.....	79
Figura 24. ISTRUZIONE: PERCENTUALE BREVETTI ACCETTATI.....	80
Figura 25. LAVORO: TASSO DI OCCUPAZIONE	81

Figura 26. LAVORO: DIFFERENZA OCCUPAZIONE MASCHI-FEMMINE.....	82
Figura 27. LAVORO: TASSO DI DISOCCUPAZIONE	83
Figura 28. LAVORO: DIFFERENZA DISOCCUPAZIONE FEMMINE-MASCHI	84
Figura 29. LAVORO: TASSO DI DISOCCUPAZIONE DI LUNGO PERIODO.....	85
Figura 30. LAVORO: DIFFERENZA DISOCCUPAZIONE GIOVANI-ADULTI	86
Figura 31. AMBIENTE: PERCENTUALE TERRENI AGRICOLI	87
Figura 32. AMBIENTE: PERCENTUALE TERRENI COPERTI DA FORESTE.....	88
Figura 33. AMBIENTE: PERCENTUALE ZONE RESIDENZIALI, COMMERCIALI, INDUSTRIALI....	89
Figura 34. AMBIENTE: PERCENTUALE TERRENI SU CUI VENGONO SVOLTE ATTIVITA' AD ALTO IMPATTO AMBIENTALE.....	90
Figura 35. AMBIENTE: RIFIUTI PROCAPITE.....	91
Figura 36. AMBIENTE: DENSITA' ABITATIVA IN KMQ.....	92
Figura 37. Elenco dei 31 indicatori di benessere divisi per ambito	98
Figura 38. Media benessere ambito istruzione	100
Figura 39. Media benessere ambito salute	102
Figura 40. Media benessere ambito ricchezza	103
Figura 41. Media benessere ambito lavoro.....	105
Figura 42. Media benessere ambito ambiente	106
Figura 43. Media benessere ambito demografia	108
Figura 44. Media non ponderata benessere globale.....	117
Figura 45. Media ponderata benessere globale.....	118
Figura 46. Grafico dei pesi fattoriali delle prime due componenti principali.....	129
Figura 47. Diagramma di dispersione degli <i>scores</i> con 266 regioni	130
Figura 48. Diagramma di dispersione degli <i>scores</i> con 260 regioni	131
Figura 49. Diagramma di dispersione degli <i>scores</i> con tre outliers	135
Figura 50. Diagramma di dispersione degli <i>scores</i> divisi per cluster di appartenenza.....	140
Figura 51. Funzione di appartenenza lineare	152

APPENDICE

Elenco completo indicatori presenti sul sito Eurostat

- Regional statistics (reg)
 - Regional agriculture statistics (reg_agr)
 - Agricultural accounts according to EAA 97 Rev.1.1 (agr_r_accts)
 - Animal populations (December) (agr_r_animal)
 - Areas harvested, yields, production (agr_r_crops)
 - Land use (agr_r_landuse)
 - Production of cows's milk on farms (1 000 t) (agr_r_milkpr)
 - Structure of agricultural holdings by NUTS, main indicators (ef_r_nuts)
 - Regional demographic statistics (reg_dem)
 - Population and area (reg_dempoar)
 - Population at 1st January by sex and age from 1990 onwards (demo_r_d2jan)
 - Population at 1st January by sex and age, at Territorial level 2 (source: OECD) (demo_r_d2janoecd)
 - table Population at 1st January by sex and age, at Territorial level 3 (1 000) (source: OECD) (demo_r_d3janoecd)

- 📄 Annual average population by sex (1 000) (demo_r_d3avg)
- 📄 Area of the regions (demo_r_d3area)
- 📄 Area of the regions, at Territorial level 2 (source: OECD) (demo_r_d2aroeed)
- 📄 Area of the regions, at Territorial level 3 (source: OECD) (demo_r_d3aroeed)
- 📄 Population density (demo_r_d3dens)
- 📄 Population density, at Territorial level 2 (source: OECD) (demo_r_d2deoeed)
- 📄 Population density, at Territorial level 3 (source: OECD) (demo_r_d3deoeed)
- 📄 Population by sex and age groups on 1 January - NUTS level 3 regions (demo_r_pjanaggr3)
- 📄 Demographic balance and crude rates - NUTS level 2 and 3 regions (demo_r_gind3)
- 📁 Population change (reg_dempch)
 - 📄 Births and deaths (1 000) (demo_r_d3natmo)
 - 📄 Births by age of the mother (demo_r_d2natag)
 - 📄 Deaths by sex and age (demo_r_d2morag)
 - 📄 Infant mortality (demo_r_d2infmo)
 - 📄 Fertility rates by age - NUTS level 2 regions (demo_r_frate2)
- 📁 Census: Regional level census 2001 round (reg_democens)
 - 📁 Population structure (cens_rstr)
 - 📄 Population by sex, age group, marital and cohabitational status (cens_rsmarcoh)
 - 📄 Population by sex, age group and selected social indicator (cens_rssocind)
 - 📄 Population by sex, country of citizenship and indicator of birth (cens_rscztz)
 - 📁 Active population (cens_ract)
 - 📄 Population by sex, group of age, economical status (cens_rapop)
 - 📄 Total population and active population by sex, age and indicator of internal or international migration (cens_ramigr)
 - 📄 Employed persons aged 15 and over by sex, major branch of economic activity, indicator of citizenship and status of(...) (cens_ractz)
 - 📁 Educational level (cens_redu)
 - 📄 Population by sex, age group, highest educational attainment and occupation (cens_reisco)
 - 📄 Population by sex, age group, highest educational attainment, current economical activity (cens_rews)
 - 📁 Households (cens_rhou)
 - 📄 Population by sex, group of age, type of household and household status (cens_rhtype)
 - 📄 Population by sex, age group, size of household (cens_rhsize)
 - 📄 Private households by type and number of member (cens_rheco)
 - 📄 Private households by type and age group of children (cens_rhagchi)
 - 📄 Private households by type, adults by age group and economic activity (cens_rhact)
 - 📁 Dwellings (cens_rdws)
 - 📄 Dwellings by indicator of conventional character, occupancy status and type of buildings (cens_rdh)
 - 📄 Dwellings by number of rooms, of persons, type of building (cens_rdbuild)
- 📁 Life table - NUTS level 2 regions (reg_mlifetable)
 - 📄 Age specific death rate (Mx) (demo_r_mdthrt)
 - 📄 Probability of dying between exact ages (qx) (demo_r_mpbdt)
 - 📄 Probability of surviving between exact ages (px) (demo_r_mpbsurv)
 - 📄 Number left alive at given exact age (lx) (demo_r_msurv)
 - 📄 Number dying between exact ages (dx) (demo_r_mdie)
 - 📄 Person-years lived between exact age (Lx) (demo_r_mpyliv)
 - 📄 Total person-years lived above given exact age (Tx) (demo_r_mtotpyliv)
 - 📄 Life expectancy at given exact age (ex) (demo_r_mlifexp)
- 📁 Regional economic accounts - ESA95 (reg_eco)

- Gross domestic product indicators - ESA95 (reg_ecogdp)
 - Gross domestic product (GDP) at current market prices at NUTS level 2 (nama_r_e2gdp)
 - Gross domestic product (GDP) at current market prices at NUTS level 3 (nama_r_e3gdp)
 - Gross domestic product (GDP) at current market prices, at Territorial level 2 (source: OECD) (nama_r_e2gdpoecd)
 - Gross domestic product (GDP) at current market prices, at Territorial level 3 (source: OECD) (nama_r_e3gdpoecd)
 - Real growth rate of regional gross value added (GVA) at basic prices at NUTS level 2 - percentage change on previou(...) (nama_r_e2grgdp)
 - Dispersion of regional GDP at Nuts level 2 and 3 (%) (nama_r_e0digdp)
 - Branch accounts - ESA95 (reg_ecobrhc)
 - Gross fixed capital formation at NUTS level 2 (nama_r_e2gfcf)
 - Compensation of employees at NUTS level 2 (nama_r_e2rem)
 - Employment (in hours worked) at NUTS level 2 (nama_r_e2emp95hw)
 - Gross value added at basic prices at NUTS level 3 (nama_r_e3vabp95)
 - Employment (in persons) at NUTS level 3 (nama_r_e3empl95)
 - Household accounts - ESA95 (reg_ecohh)
 - Allocation of primary income account of households at NUTS level 2 (nama_r_ehh2p)
 - Secondary distribution of income account of households at NUTS level 2 (nama_r_ehh2s)
 - Income of households at NUTS level 2 (nama_r_ehh2inc)
- Regional education statistics (reg_educ)
 - Number of students by level of education, orientation, sex and region (educ_renrfrg1)
 - Number of students by age, sex and region (educ_renrfrg3)
 - Regional indicators (educ_regind)
 - Participation of adults aged 25-64 in education and training, at NUTS levels 1 and 2 (from 2008) - % (trng_ifse_04)
 - Persons aged 25-64 with lower secondary education attainment, by sex and NUTS 2 level (%) (from 2008) (edat_ifse_09)
 - Persons aged 25-64 with upper secondary education attainment, by sex and NUTS 2 level (%) (from 2008) (edat_ifse_10)
 - Persons aged 25-64 with tertiary education attainment by sex and NUTS 2 level (%) (from 2008) (edat_ifse_11)
 - Persons aged 30-34 with tertiary education attainment, by sex and NUTS 1 level (%) (from 2008) (edat_ifse_12)
 - Persons aged 25-64 and 20-24 with upper secondary or tertiary education attainment, by sex and NUTS 2 level (%) (fr(...)) (edat_ifse_13)
 - Early leavers from education and training by sex and NUTS level 1 (from 2008) (edat_ifse_16)
- Regional science and technology statistics (reg_sct)
 - R&D expenditure and personnel (reg_rd)
 - Total intramural R&D expenditure (GERD) by sectors of performance and region (rd_e_gerdreg)
 - Total R&D personnel and researchers by sectors of performance, region and sex (rd_p_persreg)
 - Human Resources in Science and Technology (HRST) (reg_hrst)
 - Annual data on HRST and sub-groups (NUTS level 0, 1 and 2) (hrst_st_rcat)
 - Annual data on HRST and sub-groups by gender (NUTS level 0 and 1) (hrst_st_rsex)
 - Annual data on HRST and sub-groups by age (NUTS level 0 and 1) (hrst_st_rage)
 - Annual data on HRST and sub-groups, employed, by sector of economic activity (NUTS level 0 and 1) (1994-2008, NACE (...)) (hrst_st_rsec)
 - Annual data on HRST and sub-groups, employed, by sector of economic activity (NUTS level 0 and 1), (from 2008, NAC(...)) (hrst_st_rsec2)
 - Employment in high technology sectors (reg_htec)

- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by gender (1994-2008(...)) (htec_emp_reg)
- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by gender (from 200(...)) (htec_emp_reg2)
- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by type of occupatio(...) (htec_emp_risco)
- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by type of occupatio(...) (htec_emp_risco2)
- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by level of educatio(...) (htec_emp_riscoed)
- 📄 Annual data on employment in technology and knowledge-intensive sectors at the regional level, by level of educatio(...) (htec_emp_riscoed2)
- 📁 European patent applications to EPO (reg_pat)
 - 📄 Patent applications to the EPO by priority year at the regional level (pat_ep_rtot)
 - 📄 Patent applications to the EPO by priority year at the regional level by IPC sections and classes (pat_ep_ripc)
 - 📄 High-tech patent applications to the EPO by priority year at the regional level (pat_ep_rtec)
 - 📄 ICT patent applications to the EPO by priority year at the regional level (pat_ep_rict)
 - 📄 Biotechnology patent applications to the EPO by priority year at the regional level (pat_ep_rbio)
- 📁 Regional structural business statistics (reg_sbs)
 - 📄 Regional data (NACE Rev.2) (sbs_r_nuts06_r2)
 - 📄 Regional data (NUTS 06) (sbs_r_nuts03)
 - 📄 Number of local units, persons employed and Wages and salaries by region (sbs_cre_reg)
 - 📄 Multi yearly regional statistics (NUTS 06) (sbs_r_3f_my)
- 📁 Regional health statistics (reg_hlth)
 - 📁 Causes of death (reg_hlth_odeath)
 - 📄 Causes of death by region - Crude death rate (per 100,000 inhabitants) (Annual data) (hlth_cd_acdr)
 - 📄 Causes of death by region- Absolute Number (3 years average) - Total (hlth_cd_ynrt)
 - 📄 Causes of death by region- Absolute Number (3 years average) - Males (hlth_cd_ynrm)
 - 📄 Causes of death by region- Absolute Number (3 years average) - Females (hlth_cd_ynrf)
 - 📄 Causes of death by region - Crude death rate (per 100,000 inhabitants - 3 years average) - Total (hlth_cd_ycdrt)
 - 📄 Causes of death by region - Crude death rate (per 100,000 inhabitants - 3 years average) - Males (hlth_cd_ycdrm)
 - 📄 Causes of death by region - Crude death rate (per 100,000 inhabitants - 3 years average) - Females (hlth_cd_ycdrf)
 - 📄 Causes of death by region - Standardised death rate (per 100,000 inhabitants - 3 years average) (hlth_cd_ysdr1)
 - 📁 Health care: resources and patients (non-expenditure data) (reg_hlth_care)
 - 📄 Health personnel by NUTS 2 region (hlth_rs_prsrg)
 - 📄 Hospital beds by NUTS 2 region (hlth_rs_bdsrg)
 - 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, in-patients, total number - Total (hlth_co_disch1t)
 - 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, in-patients, total number - Males (hlth_co_disch1m)
 - 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, in-patients, total number - Females (hlth_co_disch1f)
 - 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, in-patients, per 100,000 inhabitants - Females (hlth_co_disch2f)
 - 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, in-patients, per 100,000 inhabitants - Males (hlth_co_disch2m)

- 📄 Hospital discharges by diagnosis (ISHMT) and region, in-patients, per 100,000 inhabitants - Total (hlth_co_disch2t)
- 📄 In-patient average length of stay (ISHMT, in days) by NUTS 2 region - Females (hlth_co_inpstf)
- 📄 In-patient average length of stay (ISHMT, in days) by NUTS 2 region - Males (hlth_co_inpstm)
- 📄 In-patient average length of stay (ISHMT, in days) by NUTS 2 region - Total (hlth_co_inpstt)
- 📄 Hospital days of in-patients (ISHMT) by NUTS 2 region - Females (hlth_co_hosdayf)
- 📄 Hospital days of in-patients (ISHMT) by NUTS 2 region - Males (hlth_co_hosdaym)
- 📄 Hospital days of in-patients (ISHMT) by NUTS 2 region - Total (hlth_co_hosdayt)
- 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, day cases, total number - Females (hlth_co_disch3f)
- 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, day cases, total number - Males (hlth_co_disch3m)
- 📄 Hospital discharges by diagnosis (ISHMT) and NUTS 2 region, day cases, total number - Total (hlth_co_disch3t)
- 📁 Regional tourism statistics (reg_tour)
 - 📁 Occupancy in collective accommodation establishments : domestic and inbound tourism (reg_tour_occ)
 - 📄 Arrivals in tourist accommodation establishments - regional - annual data (tour_occ_arm2)
 - 📄 Nights spent in tourist accommodation establishments - regional - annual data (tour_occ_nin2)
 - 📁 Capacity of collective tourist accommodation : establishments, bedrooms and bedplaces (reg_tour_cap)
 - 📄 Number of establishments, bedrooms and bedplaces - regional - annual data (tour_cap_nuts3)
 - 📁 Indicators on internal tourism (reg_tour_int)
 - 📄 Share of nights spent by non-residents - regional data - annual data (tour_int_shn2)
 - 📄 Tourism intensity - regional data - annual data (tour_int_tin2)
 - 📄 Growth rate - regional data - annual data (tour_int_grm2)
- 📁 Regional transport statistics (reg_tran)
 - 📄 Road, rail and navigable inland waterways networks at regional level (tran_r_net)
 - 📄 Stock of vehicles by category at regional level (tran_r_vehst)
 - 📄 Road transport of goods - Journeys made by vehicles at regional level (tran_r_veh_jour)
 - 📄 Victims in road accidents at regional level (tran_r_acci)
 - 📄 Maritime transport of passengers at regional level (tran_r_mapa_nm)
 - 📄 Maritime transport of freight at regional level (tran_r_mago_nm)
 - 📄 Air transport of passengers at regional level (tran_r_avpa_nm)
 - 📄 Air transport of freight at regional level (tran_r_avgo_nm)
 - 📄 Maritime transport of passengers at regional level (questionnaire) (tran_r_mapa_om)
 - 📄 Maritime transport of freight at regional level (questionnaire) (tran_r_mago_om)
 - 📄 Air transport of passengers at regional level (questionnaire) (tran_r_avpa_om)
 - 📄 Air transport of freight at regional level (questionnaire) (tran_r_avgo_om)
 - 📄 Annual road freight transport by region of loading (1 000 t, Mio Tkm, 1 000 Jmys) (road_go_ta_rl)
 - 📄 Annual road freight transport by region of unloading (1 000 t, Mio Tkm, 1 000 Jmys) (road_go_ta_ru)
 - 📄 Annual national and international railway goods transport by region of loading and region of unloading (tran_r_rago)
 - 📄 Railway transport - Annual national and international railway passenger transport by region of embarkation and regi(...) (tran_r_rapa)
- 📁 Regional labour market statistics (reg_lmkt)
 - 📁 Regional economically active population - LFS series and LFS adjusted series (lfst_r_lfpop)
 - 📄 Economically active population by sex and age, at NUTS levels 1, 2 and 3 (1 000) (lfst_r_lfp3pop)
 - 📄 Economically active population by sex and age, at NUTS levels 1 and 2 (1 000) (lfst_r_lfp2act)
 - 📄 Economically active population, at Territorial level 2 (1 000) (source:OECD) (lfst_r_lfp2aoecd)
 - 📄 Economically active population, at Territorial level 3 (1 000) (source:OECD) (lfst_r_lfp3aoecd)
 - 📄 Economic activity rates by sex and age, at NUTS levels 1 and 2 (%) (lfst_r_lfp2actrt)

- ☞ Economically active population by sex, age and highest level of education attained, at NUTS levels 1 and 2 (1 000) (lfst_r_lfp2acedu)
- ☞ Regional employment - LFS series (lfst_r_lfemp)
 - ☞ Employment by sex and age, at NUTS levels 1 and 2 (1 000) (lfst_r_lfe2emp)
 - ☞ Employment by economic activity, at NUTS levels 1 and 2 (1 000) (NACE Rev.2) (lfst_r_lfe2en2)
 - ☞ Employment by economic activity, at NUTS levels 1 and 2 (1 000) (1999-2009, NACE Rev.1.1) (lfst_r_lfe2en1)
 - ☞ Employment, at Territorial level 2 (1 000) (source: OECD) (lfst_r_lfe2emoecd)
 - ☞ Employment, at Territorial level 3 (1 000) (source: OECD) (lfst_r_lfe3emoecd)
 - ☞ Employment by professional status, at NUTS levels 1 and 2 (1 000) (lfst_r_lfe2estat)
 - ☞ Employment by full-time/part-time and sex, at NUTS levels 1 and 2 (1 000) (lfst_r_lfe2eftpt)
 - ☞ Employment by sex, age and highest level of education attained, at NUTS levels 1 and 2 (1 000) (lfst_r_lfe2eedu)
 - ☞ Employment and commuting among NUTS level 2 regions (1 000) (lfst_r_lfe2ecomm)
 - ☞ Employment rates by sex and age, at NUTS levels 1 and 2 (%) (lfst_r_lfe2emprr)
 - ☞ Average number of usual weekly hours of work in main job, at NUTS levels 1 and 2 (hours) (lfst_r_lfe2ehour)
- ☞ Regional unemployment - LFS adjusted series (lfst_r_lfu)
 - ☞ Unemployment by sex and age, at NUTS levels 1, 2 and 3 (1 000) (lfst_r_lfu3pers)
 - ☞ Unemployment, at Territorial level 2 (1 000) (source: OECD) (lfst_r_lfu2peoecd)
 - ☞ Unemployment, at Territorial level 3 (1 000) (source: OECD) (lfst_r_lfu3peoecd)
 - ☞ Unemployment rates, at Territorial level 2 (%) (source: OECD) (lfst_r_lfu2rtoecd)
 - ☞ Unemployment rates, at Territorial level 3 (%) (source: OECD) (lfst_r_lfu3rtoecd)
 - ☞ Unemployment rates by sex and age, at NUTS levels 1, 2 and 3 (%) (lfst_r_lfu3rt)
 - ☞ Long-term unemployment (12 months and more), at NUTS levels 1 and 2 (1 000; %) (lfst_r_lfu2ltu)
- ☞ Regional socio-demographic labour force statistics - LFS series (lfst_r_lfsd)
 - ☞ Number of households by degree of urbanisation of residence, at NUTS levels 1 and 2 (1 000) (lfst_r_lfsd2hh)
 - ☞ Population aged 15 and over by sex and age, at NUTS levels 1 and 2 (1 000) (lfst_r_lfsd2pop)
- ☞ Regional labour market disparities - LFS series and LFS adjusted series (lfst_r_lmd)
 - ☞ Dispersion of regional (NUTS level 2 and 3) employment rates of age group 15-64 (%) (lfst_r_lmdr)
 - ☞ Dispersion of regional (NUTS levels 2 and 3) unemployment rates (%) (lfst_r_lmdur)
 - ☞ Underperforming regions regarding employment (NUTS level 2 and 3) (lfst_r_lmdue)
 - ☞ Underperforming regions regarding unemployment (NUTS level 2 and 3) (lfst_r_lmduu)
- ☞ Regional labour market data based on pre-2003 methodology (data up to 2001) - LFS adjusted series (reg_lfh)
 - ☞ Economically active population by sex and age, at NUTS levels 1, 2 and 3 (1 000) (reg_lfh3wpop)
 - ☞ Economically active population by sex and age, at NUTS levels 1 and 2 (1 000) (reg_lfh2act)
 - ☞ Economic activity rates by sex and age, at NUTS levels 1 and 2 (%) (reg_lfh2actrr)
 - ☞ Employment by sex and age, at NUTS levels 1 and 2 (1 000) (reg_lfh2emp)
 - ☞ Employment by economic activity, full-time/part-time and sex, at NUTS levels 1 and 2 (1 000) (reg_lfh2empn)
 - ☞ Employment rates of age group 15-64 by sex, at NUTS levels 1 and 2 (%) (reg_lfh2emprr)
 - ☞ Dispersion of regional (NUTS level 2) employment rates of age group 15-64 (%) (reg_lfh2cvert)
 - ☞ Unemployment by sex and age, at NUTS levels 1, 2 and 3 (1 000) (reg_lfh2pers)
 - ☞ Unemployment rates by sex and age, at NUTS levels 1, 2 and 3 (%) (reg_lfh3unrr)
 - ☞ Dispersion of regional (NUTS levels 2 and 3) unemployment rates (%) (reg_lfh3stdv)
 - ☞ Long-term unemployment (12 months and more), at NUTS levels 1 and 2 (1 000; %) (reg_lfh2ltu)
 - ☞ Number of households by degree of urbanisation of residence, at NUTS levels 1 and 2 (1 000) (reg_lfh2hh)

- 🔗 Austria - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-reg(...) (migr_r_2at)
- 🔗 Poland - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-regi(...) (migr_r_2pl)
- 🔗 Romania - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-reg(...) (migr_r_2ro)
- 🔗 Slovenia - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-re(...) (migr_r_2si)
- 🔗 Slovakia - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-re(...) (migr_r_2sk)
- 🔗 Finland - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-reg(...) (migr_r_2fi)
- 🔗 Sweden - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-regi(...) (migr_r_2se)
- 🔗 Croatia - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-reg(...) (migr_r_2hr)
- 🔗 Norway - Internal regional migration by sex, region of origin (GEO) and destination (PARTNER), excluding intra-regi(...) (migr_r_2no)
- 📁 Regional environment statistics (reg_env)
 - 📁 Regional waste statistics (reg_env_was)
 - 🔗 Regional generation and treatment of municipal waste (1 000 t) (env_nwas_gen)
 - 🔗 Regional coverage rate of municipal waste collection (%) (env_nwas_cov)
 - 📁 Regional water statistics (reg_env_wat)
 - 🔗 Freshwater resources (NUTS2) - mio m³ (env_n2_fr)
 - 🔗 Water abstraction (NUTS2) - mio m³ (env_n2_wa)
 - 🔗 Water use (NUTS2) - mio m³ (env_n2_wu)
 - 🔗 Population connected to public water supply (NUTS2) - % (env_n2_pws)
 - 🔗 Population connected to wastewater treatment (NUTS2) - % (env_n2_pww)
 - 🔗 Treatment capacity of wastewater treatment plants (NUTS2) (env_n2_cwp)
 - 🔗 Wastewater generation and discharge (NUTS2) (env_n2_wwd)
 - 📁 Other regional environment statistics (reg_env_oth)
 - 🔗 Specific land use parameters (NUTS2) (env_rlu)
 - 🔗 Energy : primary production and final consumption (NUTS2) - 1 000 toe (env_rpep)
 - 🔗 Share of Renewable Energy in Final Energy Consumption (NUTS2) - % (env_rfec)
 - 🔗 Specific transport parameters (NUTS2) (env_rtr)
- 📁 Urban audit (urb)
 - 🔗 Key indicators for core cities (urb_ikey)
 - 🔗 Derived indicators for core city (urb_icity)
 - 🔗 Derived indicators for larger urban zones (urb_iluz)
 - 🔗 Derived indicators for sub-city districts (urb_iscd)
 - 🔗 Reduced set of derived indicators for 570 cities (urb_ilca)
 - 🔗 Data collected for core city (urb_vcity)
 - 🔗 Data collected for larger urban zones (urb_vluz)
 - 🔗 Reduced set of data collected for 570 cities (urb_vlca)
 - 🔗 Perception survey results (urb_percep)
- 📁 Land cover and land use, soil, landscape (lan)
 - 📁 Land Cover (lan_lcv)
 - 🔗 Land cover overview, by NUTS 2 regions (lan_lcv_oww)
 - 🔗 Land covered by artificial land, by NUTS 1 regions (lan_lcv_art)
 - 🔗 Land covered by grassland, by NUTS 2 regions (lan_lcv_grs)

- 📄 Land covered by shrubland, by NUTS 2 regions (lan_lcv_shr)
- 📄 Land covered by woodland, by NUTS 2 regions (lan_lcv_woo)
- 📁 Land Use (lan_lu)
 - 📄 Land use overview , by NUTS 2 regions (lan_lu_oww)
 - 📄 Land use in agriculture, by NUTS 2 regions (lan_lu_agr)
 - 📄 Land use with heavy environmental impact, by NUTS 2 regions (lan_lu_heav)
 - 📄 Land use in services and residential, by NUTS 2 regions (lan_lu_inf)
- 📁 Landscape indicators (lan_lcs)
 - 📄 Landscape richness, % (lan_lcs_ric)
 - 📄 Landscape structure, % (lan_lcs_str)
 - 📄 Landscape dissection, % (lan_lcs_diss)
 - 📄 Landscape Shannon Diversity Index, % (lan_lcs_sdi)
 - 📄 Landscape Shannon Evenness Index, % (lan_lcs_sei)