Argomenti di discussione

(Discussion topics)

# ANALISI DELLE DIREZIONI PROVINCIALI ATTRAVERSO IL DB GEO

*E. Carbone, A. Spingola*

N. 01/2015

Agenzia Entrate

*Argomenti di discussione* è una pubblicazione che intende divulgare contributi ed analisi su argomenti di economia, statistica, econometria e scienza delle finanze, che abbiano rilevanza per la missione dell'Agenzia delle Entrate, al fine di alimentare il dibattito scientifico sui temi di interesse strategico dell'Agenzia e favorire lo scambio di opinioni. La pubblicazione ospita contributi in lingua italiana o inglese proposti da autori sia interni sia esterni all'Agenzia. Le opinioni espresse negli articoli sono lasciate alla discrezionalità degli autori e non impegnano la responsabilità dell'Agenzia delle Entrate.

Discussion topics *is a publication aiming at disclosing contributions and analysis in economics, statistics, econometrics and public finance, which are relevant to the mission of the Italian Revenue Agency (Agenzia delle Entrate) in order to feed the scientific debate on topics of strategic interest for the Agency and to facilitate the exchange of opinions. This publication admits contributions in Italian or in English proposed by authors internal or external to the Agency. The views expressed in the articles are those of the authors and do not involve the responsibility of the Revenue Agency.*

LICENZA/LICENCE

Quest'opera è distribuita con Licenza Creative Commons Attribuzione - Non commerciale - Non opere derivate 3.0.
*This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.*

ISSN: 2420-773X

Stefano PISANI (Agenzia delle Entrate - Responsabile Scientifico/Editor in Chief)

COMITATO SCIENTIFICO/BOARD OF EDITORS
Bruno CHIARINI (Università degli Studi di Napoli "Parthenope")
Valeria DE BONIS (Università degli Studi di Roma "La Sapienza")
Antonio DI MAJO (Università degli Studi "Roma Tre")
Roberto MONDUCCI (Istituto Nazionale di Statistica)
Alessandro SANTORO (Università degli Studi "Milano-Bicocca")

REDAZIONE/MANAGING EDITORS
Nicoletta Alborino (Agenzia delle Entrate)
Marta Gallucci (Agenzia delle Entrate)

Tel. +39-06-5054-3874/5071 Fax +39-06-5054-5038
E-mail: ae.argomentididiscussione@agenziaentrate.it
Web: http://www1.agenziaentrate.gov.it/argomentididiscussione/index.htm

# Analisi delle Direzioni Provinciali attraverso il DB Geo

Ernesto Carbone, Andrea Spingola

Agenzia delle Entrate, andrea.spingola@agenziaentrate.it

**Abstract**

The aim of this paper is the analysis of the market demand faced by the Italian Revenue Agency. To perform this study, we used some variables at the provincial level from DB Geo, grouped by the following 7 thematic areas:

- size of catchment area,
- fiscal dangerousness,
- social dangerousness,
- standard of living,
- maturity of the production structure,
- use of IT services,
- availability of transport infrastructures.

This work has a twofold objective:

1. The calculation of a composite indicator for each thematic area, summarizing the characteristics of each Provincial Directorate (via a unique number for each area) in order to analyze specific aspects of the single area and, therefore, provide a more effective instrument for business decisions.

2. The identification of similarities between the Provincial Directorates, considering the 7 areas simultaneously, in order to obtain homogeneous groups (clusters) on which to apply custom management profiles; in addition, this new map of Italy allows us to make more meaningful comparisons between Provincial Directorates in the same cluster.

**Sommario**

Lo scopo del presente articolo è l'analisi della domanda di mercato incontrata dall'Agenzia delle Entrate. Per effettuare questo studio, abbiamo utilizzato alcune variabili a livello provinciale dal DB Geo, raggruppate secondo le 7 seguenti aree tematiche:

- Numerosità del bacino
- Pericolosità fiscale
- Pericolosità sociale
- Tenore di vita
- Maturità della struttura produttiva
- Livello di tecnologia e servizi

- Disponibilità di infrastrutture di trasporto

Questo lavoro ha un duplice obiettivo:

1. Il calcolo di un indicatore composito per ciascuna area tematica, che sintetizzi le caratteristiche di ciascuna Direzione Provinciale (attraverso un unico valore per ciascuna area) al fine di analizzare specifici aspetti della singola area e, di conseguenza, fornire uno strumento più efficace per le decisioni di business.

2. L'identificazione delle similarità tra le Direzioni Provinciali, considerando le 7 aree simultaneamente, in modo da ottenere gruppi omogenei (*"cluster"*) su cui applicare profili di *"custom management"*; inoltre, questa nuova mappa dell'Italia ci consente di effettuare confronti più significativi tra le Direzioni Provinciali dello stesso cluster.

**JEL classifications:** C38, C44, M31

**Key words:** Customer Relationship Management, Citizen Relationship Management, Multivariate data analysis, Factor analysis, Principal components analysis, Cluster analysis

## Indice

1. Introduzione ..... 6
2. Analisi della domanda ..... 8
3. Mappe delle DP in base agli indici sintetici di ogni area tematica ..... 10
4. Analisi di cluster delle Direzioni Provinciali ..... 14
4.1 Caratterizzazione dei cluster ..... 18
5. Conclusioni ..... 21
Appendice ..... 22
A. Risultati ACP ..... 22
B. Interpretazione delle componenti principali ..... 32
C. Indici sintetici e cluster delle Direzioni Provinciali ..... 32
Riferimenti bibliografici ..... 36
Siti internet consultati ..... 37

# 1. Introduzione

Riprendendo alcuni lavori precedenti (Mascitti et al., 2004, Carbone et al., 2010), ci si propone di analizzare il territorio e la domanda di mercato che nasce da esso per l'Agenzia delle Entrate, utilizzando alcune variabili presenti nel DB Geo provinciale.

Il DB Geo è un data base contenente dati provenienti da diverse fonti interne ed esterne che rappresentano diversi aspetti del contesto territoriale delle province italiane. Le 36 variabili utilizzate in questo studio sono raggruppate in 7 aree tematiche, come indicato in figura 1.1:

- La "Numerosità del bacino", dove è rappresentata la platea dei contribuenti che dà origine alla domanda dei servizi dell'Agenzia.
- La "Pericolosità fiscale", le cui variabili ci danno un'idea, da un lato, dell'evasione e, dall'altro, del disordine dell'economia (identificato tramite fallimenti e protesti); questo secondo aspetto può suggerire una maggiore difficoltà nel controllo da parte dell'Agenzia o essere indizio di una situazione economica non florida, che rende più difficile la *tax compliance*.
- La "Pericolosità sociale" è utile per definire il contesto di sicurezza nel quale operano sia i contribuenti sia i funzionari dell'Agenzia.
- Il "Tenore di vita" è importante sia perché rappresenta un riferimento per confronti con quanto dichiarato dai contribuenti, sia perché può influenzare le differenti tipologie di attività che interessano i singoli uffici (un territorio più ricco richiede servizi differenti rispetto a uno meno agiato).
- La "Maturità della struttura produttiva" è stata presa in considerazione partendo dal presupposto che sussistano necessità differenti (in termini di assistenza e di controllo) a seconda della specializzazione settoriale (agricoltura, industria o servizi), della dimensione delle strutture produttive presenti nel territorio e della velocità nel loro cambiamento.
- Il "Livello di tecnologia e servizi" cerca di valutare l'utilizzo di Internet e dei servizi bancari e postali per una duplice motivazione: dal lato dei servizi al contribuente, conoscere la propensione all'uso dei mezzi informatici può agevolare la calibrazione degli strumenti offerti per gli adempimenti fiscali; dal lato del controllo, la diffusione di strumenti di pagamento elettronici favorisce la *compliance* e aiuta nella strategia di contrasto.
- La "Disponibilità di infrastrutture di trasporto" rappresenta la maggiore o minore facilità per i contribuenti nel raggiungere gli uffici dell'agenzia, e per i nostri verificatori nel raggiungere i soggetti da controllare.

Sono state considerate medie triennali[1] al fine di minimizzare eventuali effetti temporanei e considerare tendenze di medio periodo.

[1] In generale è considerato il triennio 2008-2010; per le aree tematiche *Numerosità del bacino* e *Livello di tecnologia e servizi* è stato considerato il triennio 2009-2011.

**Figura 1.1**: *Elenco delle variabili inserite nel modello, catalogate per area tematica*

| Numerosità del bacino di utenti | Pericolosità fiscale | Pericolosità sociale | Tenore di vita | Maturità della struttura produttiva | Livello di tecnologia dei servizi | Disponibilità di infrastrutture di trasporto |
|---|---|---|---|---|---|---|
| Dichiarazioni 730 | Tax Gap (in %) | Delitti contro la vita ogni milione di abitanti | Depositi bancari e postali pro capite | Tasso dei lavoratori autonomi | Operazioni Bancoposta on-line ogni milione di abitanti | Km di autostrade ogni milione di abitanti |
| Dichiarazioni Unico PF | Tasso di fallimento delle imprese | Furti, rapine, estorsioni, sequestri ogni milione di abitanti | Valore aggiunto pro capite | Produzione media netta IRAP | Numero di POS attivi ogni milione di abitanti | Km di strade provinciali, regionali, statali ogni milione di abitanti |
| Dichiarazioni Unico Società e Enti non commerciali | Tasso dei protesti | Truffe e altre frodi ogni milione di abitanti | Patrimonio immobiliare pro capite | Turn-over lordo delle imprese | Tasso di Unico PF trasmessi per via telematica dai contribuenti | Km di ferrovie ogni milione di abitanti |
| Atti e dichiarazioni di successione registrati | | Attentati e associazione (mafiose, a delinquere) ogni milione di abitanti | Importo medio delle pensioni | Quota IRAP della piccola industria e agricoltura | Servizi Home Banking ogni milione di abitanti | Trasporti cittadini ogni diecimila abitanti |
| Stranieri residenti | | Reati di contrabbando ogni milione di abitanti | Consumo elettrico domestico pro capite | | | |
| | | Tutti gli altri reati ogni milione di abitanti | Produzione di rifiuti urbani pro capite | | | |
| | | | Tasso di disoccupazione | | | |
| | | | Tasso di occupazione | | | |
| | | | Premi assicurativi (ramo vita) pro capite | | | |

L'analisi effettuata sulla base di queste informazioni è composta dalle seguenti due fasi:

1. la prima ha come obiettivo il calcolo di un indicatore sintetico per ogni area tematica che riassuma le caratteristiche di ogni Direzione Provinciale (nel seguito DP), per quell'area, attraverso un numero unico;
2. nella seconda, l'intento è ripartire il territorio nazionale in gruppi di DP con caratteristiche il più possibile omogenee all'interno di ogni gruppo e significativamente differenti tra un gruppo e l'altro (ossia in cluster).

I risultati ottenuti nella prima fase consentono di ottenere una sequenza delle DP utile per analizzare particolari aspetti inerenti la singola area tematica: ad esempio se si vuole analizzare fenomeni che riguardano solo la dimensione del bacino (Spingola, 2012), oppure solo la pericolosità fiscale ecc.

I risultati ottenuti nella seconda fase, invece, permettono di identificare affinità tra le DP considerando complessivamente tutte le 7 aree tematiche. Grazie all'identificazione di queste similarità, partendo dall'assunto di base che mercati con caratteristiche similari comportano rischi ed opportunità similari, è possibile, ad esempio, ipotizzare profili di management analoghi per DP dello stesso cluster oppure creare sinergie e confronti tra di esse per affrontare le criticità nell'attività lavorativa (Carbone et al., 2010).

Gli indici sintetici ottenuti nella prima fase sono inoltre utili come analisi esplorativa preliminare alla seconda fase e consentono di identificare in maniera più semplice le caratteristiche dei cluster successivamente ottenuti.

## 2. Analisi della domanda

Il principale problema teorico di questo lavoro è l'analisi della domanda che deve essere fronteggiata da un ente pubblico.

Nel settore privato questo genere di studi è in fase avanzata e, soprattutto in alcuni settori, si è passati dai classici modelli di analisi della domanda all'utilizzo di strumenti di data mining per studiare le relazioni col cliente. Infatti, quando la domanda del bene o servizio è ampia rispetto alle capacità produttive dell'azienda, essa può permettersi di concentrarsi unicamente sul prodotto e sui processi interni, aumentando la capacità produttiva per massimizzare i profitti. Ma, nel periodo storico che stiamo attraversando, il principale problema è trovare clienti con cui instaurare relazioni di business il più possibile stabili e durature, per cui si è passati dal controllo della redditività del prodotto, tipica del mondo E.R.P., al controllo della redditività del cliente, caratteristico della strategia di C.R.M.

Il CRM è un insieme di meccanismi organizzativi e tecnologici che hanno l'obiettivo di contenere l'instabilità del mercato attraverso una migliore conoscenza delle variabili ambientali, per poter anticipare le esigenze dei clienti e rendere più stabile e programmabile l'attività di produzione (Brown, 2000).

I primi studi di questo tipo sono stati compiuti sui clienti di banche ed assicurazioni. In tali settori, infatti, le tecnologie stanno conducendo sempre più a una disintermediazione

delle relazioni con il cliente e ad un utilizzo diretto dei servizi web based (Decastri et al, 2001). Questo comporta la costruzione di immense basi di dati contenenti le scelte effettuate dalla clientela, sulle quali è possibile compiere analisi finalizzate a personalizzare il contatto e la proposta di nuovi prodotti/servizi. Anche nelle società di telecomunicazioni esiste una situazione molto vivace, determinata dalle presenza di gigantesche basi di dati, costituite dalle interazioni telefoniche dei clienti. L'analisi di questi dati consente di clusterizzare i clienti stessi (Okazaki, 2006) e formulare delle offerte differenziate per i diversi cluster di appartenenza.

In altri settori privati, i sistemi CRM hanno ruoli meno centrali: simili alle telecomunicazioni per numerosità della clientela e frequenza delle interazioni con gli utenti, le aziende di distribuzione di elettricità e gas e la grande distribuzione organizzata appaiono aver seguito un percorso evolutivo meno sviluppato sul CRM analitico e direzionale (Bracchi e Motta, 2004).

Nella Pubblica Amministrazione la situazione è differente: mentre sono diversi i casi di analisi della domanda finalizzati alla sola quantificazione del fabbisogno di un servizio pubblico (acqua, trasporto pubblico), si registra un forte ritardo nell'utilizzo di tecniche volte a individuare le specificità dell'utenza, per offrire un servizio più "personalizzato". Tale attività, mutuando la definizione data per il settore privato, è indicata con il nome di Citizen Relationship Management (CiRM).

Il CiRM può essere definito come una strategia, sostenuta dalla tecnologia e orientata al cittadino, per mantenere e ottimizzare le relazioni fra PA e cittadini ed incoraggiare la partecipazione di questi ultimi (Schellong, 2007). Si tratta di una definizione operativa, perché il CiRM è un'area di ricerca emergente e rappresenta ancora un concetto gestionale, piuttosto che un argomento teorico. Nel CiRM, il cittadino non è solo un fruitore di servizi, ma una fonte di informazioni e punti di vista grazie ai quali pianificare servizi migliori e meglio erogati, o persino elaborare nuove politiche di governo (Colapinto, 2005). Oggi, infatti, anche la PA affronta nuove sfide: cittadini, aziende e realtà locali hanno richieste sempre più pressanti, a fronte di una progressiva riduzione dei budget e all'obsolescenza dei processi gestionali, in un contesto colmo di barriere culturali che limitano l'adozione di strumenti volti alla diversificazione dell'offerta (Hunter e Shine, 2001).

Nella PA italiana sono state sviluppate alcune applicazioni di CiRM operativo, finalizzate alla gestione dei contact center, anche se gran parte degli enti ha attivato dei centralini telefonici piuttosto che delle vere e proprie strutture preposte al public care; eccezioni sono, per la PA locale, i comuni di Roma e Firenze, e per la PA centrale, i casi dell'Agenzia delle Entrate e dell'INPS (Schael, 2005).

Questo lavoro costituisce un tentativo di applicare le logiche del CRM analitico alle realtà territoriali dell'Agenzia delle Entrate, con l'obiettivo di ottenere gruppi omogenei di direzioni provinciali ai quali applicare profili di management personalizzati e all'interno dei quali effettuare confronti più significativi tra DP (CRM direzionale). La missione istituzionale dell'Agenzia delle Entrate, infatti, non consiste solo nell'offrire servizi ai contribuenti per favo¬rirne l'adempimento fiscale, ma anche nell'operare la prevenzione e il contrasto dell'evasione; di conseguenza l'analisi della domanda non può basarsi solo sullo studio delle interazioni col cliente, tramite contact center o direttamente allo sportello, ma deve necessariamente fondarsi su una analisi territoriale

che tenga conto di tutte le caratteristiche dell'ambiente socio-economico della Direzione Provinciale, il suo territorio e le infrastrutture.

## 3. Mappe delle DP in base agli indici sintetici di ogni area tematica

Il primo passo è consistito in una analisi dei dati finalizzata ad individuare, tra le variabili a disposizione per ogni area tematica, quelle ritenute più significative ai fini dell'operato dell'Agenzia. Sulle stesse è stato poi condotto un esame descrittivo delle distribuzioni, soffermandosi in particolare sulle osservazioni estreme.

L'analisi bidimensionale ha riguardato il calcolo delle correlazioni fra le variabili. La presenza di una correlazione molto elevata non è stata sempre interpretata come sintomo di una informazione ridondante poiché lo studio che stiamo eseguendo non è afflitto da problemi di multicollinearità.

Infine l'analisi multidimensionale è consistita in una analisi in componenti principali. Si tratta di una tecnica statistica che definisce nuove variabili come combinazioni lineari di quelle originarie; queste grandezze "artificiali" spiegano il fenomeno in misura decrescente: le prime hanno un contenuto informativo superiore rispetto a ciascuna variabile originaria, singolarmente considerata, pertanto la loro osservazione consente di ottenere una sintesi efficace del fenomeno. È bene precisare comunque che, per le caratteristiche tipiche di questa tecnica, non è possibile assegnare a priori un significato a ciascuno degli assi appena descritti, ma tale definizione scaturisce dallo studio dei risultati conseguiti.

Si è scelto, per ogni area, un numero di componenti principali tale da rappresentare almeno il 75% della variabilità del fenomeno, e si è cercato di dare una interpretazione dei nuovi assi fattoriali.

In appendice A si riportano, per ogni area tematica, la matrice di correlazione delle variabili, con i suoi autovalori e autovettori (che rappresentano, rispettivamente, la varianza spiegata da ogni asse fattoriale e il coefficiente di ciascuna variabile nella determinazione dell'asse stesso) e i cerchi delle correlazione (dove si riproducono graficamente sia l'influenza delle variabili sugli assi sia le correlazioni fra le variabili stesse). Infine si riporta, in appendice B, uno schema con le possibili interpretazioni degli assi fattoriali scelti per rappresentare le diverse aree.

L'ultima operazione è rappresentata dal calcolo degli indici sintetici. Tra i metodi più diffusi per ottenere indici compositi, ricordiamo:

- Pesatura uniforme (Mazziotta et al., 2010)
- Metodo tassonomico di Wroclaw, (Mazziotta et al., 2010)
- Metodo delle penalità per coefficiente di variazione, (Mazziotta e Pareto, 2007)
- Analisi della correlazione (Nardo et al., 2005)
- Benefit of the Doubt (Charnes et al, 1981; Cherchye et al, 2008; Rinaldi et al, 2008)
- Tecniche fattoriali (Bolasco, 2002)

Tra tutti, si è scelto di utilizzare una tecnica fattoriale, che ha il vantaggio di calcolare direttamente i pesi delle diverse variabili (tramite la correlazione, sebbene essa, talvolta, possa non indicare il vero peso delle variabili). Nel nostro caso, per ogni DP (i) e per ogni area tematica (j), si è proceduto con il calcolo di una combinazione lineare (cl) dei valori delle componenti principali selezionate, con pesi pari agli autovalori di quei fattori. Il risultato ottenuto è stato successivamente trasformato applicando la formula 3.1:

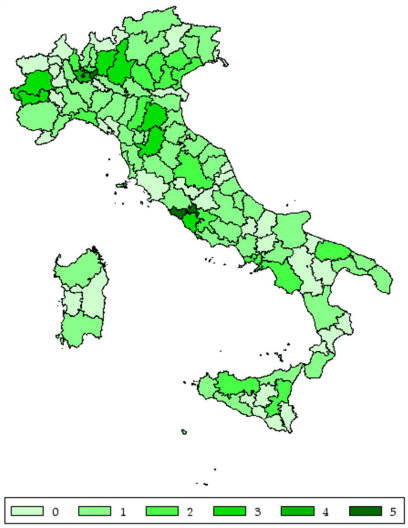
$$is_{ij} = 5 * \frac{cl_{ij} - MIN_i(cl_{ij})}{MAX_i(cl_{ij}) - MIN_i(cl_{ij})} \qquad [3.1]$$

Che consente di riportare l'indicatore sintetico (is), all'intervallo 0 – 5.

Tale livelli sono stati successivamente rappresentati con diverse gradazioni di colore in mappe che offrano una rappresentazione visiva della distribuzione dei livelli di quell'area tematica sul territorio.

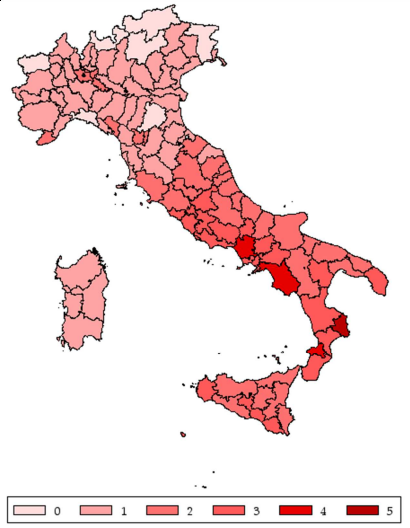
In figura 3.1 sono mostrate le mappe relative ai singoli indici, rappresentati con valori crescenti da 0 a 5.

**Figura 3.1**: *Mappe delle DP in base agli indici sintetici di ogni area tematica*



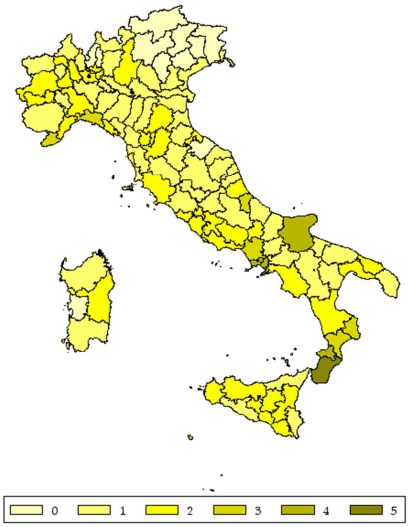

**Numerosità del bacino**

Si osserva che, a parte qualche zona popolosa del sud (Napoli, Salerno, Bari, Palermo, Catania), la platea è maggiormente concentrata al centro-nord (con le punte massime di Milano 2 e Roma 1).



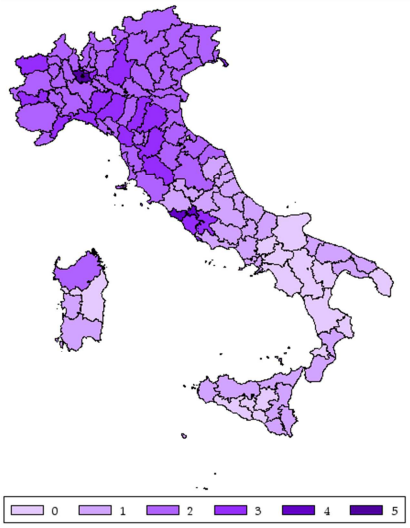

## Pericolosità fiscale

Si nota un progressivo aumento del fenomeno procedendo verso sud, dai valori minimi di Bolzano e Trento a quelli massimi di Vibo Valentia e Crotone.



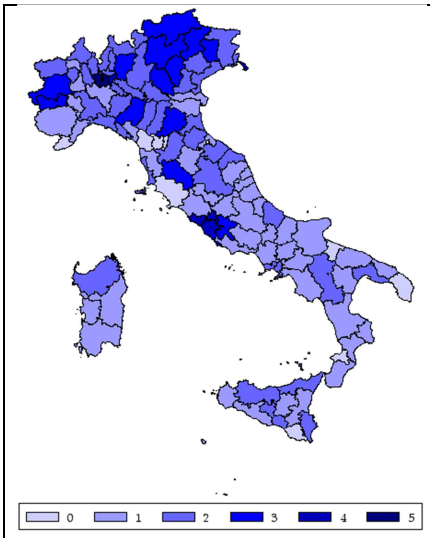

## Pericolosità sociale

Emerge una distribuzione abbastanza variegata su tutto il territorio italiano, con una bassa incidenza del fenomeno nel nord-est e picchi massimi a Reggio Calabria e Napoli.



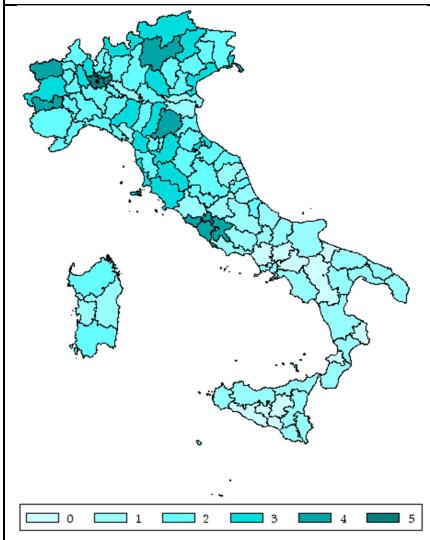

## Tenore di vita

Si evidenzia la crescita del tenore di vita in direzione sud-nord, con valori minimi ad Enna e Vibo Valentia e massimi a Roma e Milano.



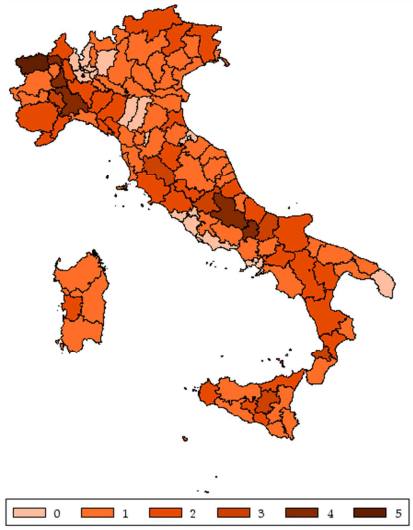

### Maturità della struttura produttiva

Si rileva una distribuzione non omogenea sul territorio nazionale: mentre, ad esempio, Siena ha un valore alto, le vicine Grosseto e Pistoia hanno i valori minimi. Inoltre, insieme alle attese Roma, Milano, Torino e Bologna, troviamo una struttura produttiva più matura a Parma, Bergamo e nel nord-est.




### Livello di tecnologia e servizi

Si constatano maggiori livelli di tecnologia nel centro-nord, in particolare ad Aosta e Bologna, che si affiancano alle metropoli di Roma e Milano.




### Disponibilità di infrastrutture di trasporto

La distribuzione del fenomeno sul territorio è piuttosto variegata. Mentre Aosta e Napoli hanno valori estremi dell'indice (basso a Napoli e alto ad Aosta) a causa, principalmente, della diversa concentrazione di popolazione sul territorio (che incide sull'intasamento dei mezzi di trasporto), sui valori di L'Aquila e Prato ha influenza la presenza o meno delle infrastrutture di trasporto (numerose all'Aquila, meno a Prato).

## 4. Analisi di cluster delle Direzioni Provinciali

La seconda fase è partita dal mettere insieme tutte le variabili che sono state selezionate nel primo stadio per ogni area tematica ed eseguire su esse una cluster analysis. Si è tentato anche di realizzare uno studio direttamente sulle componenti principali scelte e sugli indici sintetici calcolati nella prima fase, ma i risultati sono stati inferiori alle attese ai fini degli obiettivi aziendali.

Avendo già eseguito una indagine sulle singole variabili nel punto precedente, si è passati direttamente all'analisi bidimensionale, costruendo una matrice di correlazione unica, che non è stata riportata in questo documento per ragioni di spazio, ma dalla cui analisi si evidenzia che:

- tutte le variabili legate alla dimensione delle DP sono altamente correlate tra loro;
- il Tax gap in percentuale (Braiotta et al., 2012; D'agosto et al., 2013) è correlato positivamente sia con la quota della produzione IRAP derivante da aziende agricole e industrie di piccole dimensioni sia con il Tasso di disoccupazione, mentre è correlato negativamente con il Valore aggiunto Istat procapite e i POS procapite , a testimoniare che le zona più ricche sono quelle con minore intensità dell'evasione;
- l'utilizzo di servizi di Home banking è molto correlato con la Produzione media Irap e l'importo medio delle pensioni procapite, dimostrando che tenore di vita e livello di tecnologia vanno di pari passo.

Conclusa la fase di analisi esplorativa, si è proceduto alla classificazione automatica, ossia alla cluster analysis. Con questo termine si intende un insieme di tecniche che cercano di raggruppare un elevato numero di unità statistiche (nel nostro caso le Direzioni provinciali) in determinati gruppi (cluster), il più possibile omogenei al loro interno e disomogenei tra di loro, con riferimento ai valori di alcune variabili appositamente selezionate. Nel nostro caso la procedura di aggregazione è di tipo gerarchico bottom-up, (in quanto fornisce una sequenza di partizioni che va dal numero massimo di gruppi, tanti quanti sono le Direzioni, fino ad un unico gruppo che le contiene tutte) con algoritmo di base (cioè attraverso aggregazioni singole). Questa scelta è consigliata, in letteratura, quando il numero di soggetti da analizzare non è particolarmente esteso poiché consente di tenere sotto controllo il criterio di aggregazione (Bolasco, 2002).

Per realizzare tali raggruppamenti è stato scelto il criterio di Ward, attraverso il quale si ottengono dei cluster sferici con all'incirca lo stesso numero di osservazioni (Yeo, 2003), eliminando il problema della riallocazione degli outlier riscontrato nella clusterizzazione del 2010.

Il criterio di scelta del numero di cluster si è basato su alcuni indicatori statistici:

- Cubic Clustering Criterion (Sarle, 1983)
- Pseudo-F Statistic (Calinski e Harabasz, 1974)
- Pseudo-t2 Statistic (Duda e Hart, 1973)

ma ha risposto anche alla necessità di mediare tra due opposte esigenze: mantenere la maggiore quantità di informazioni possibile e avere un risultato fruibile ai fini della

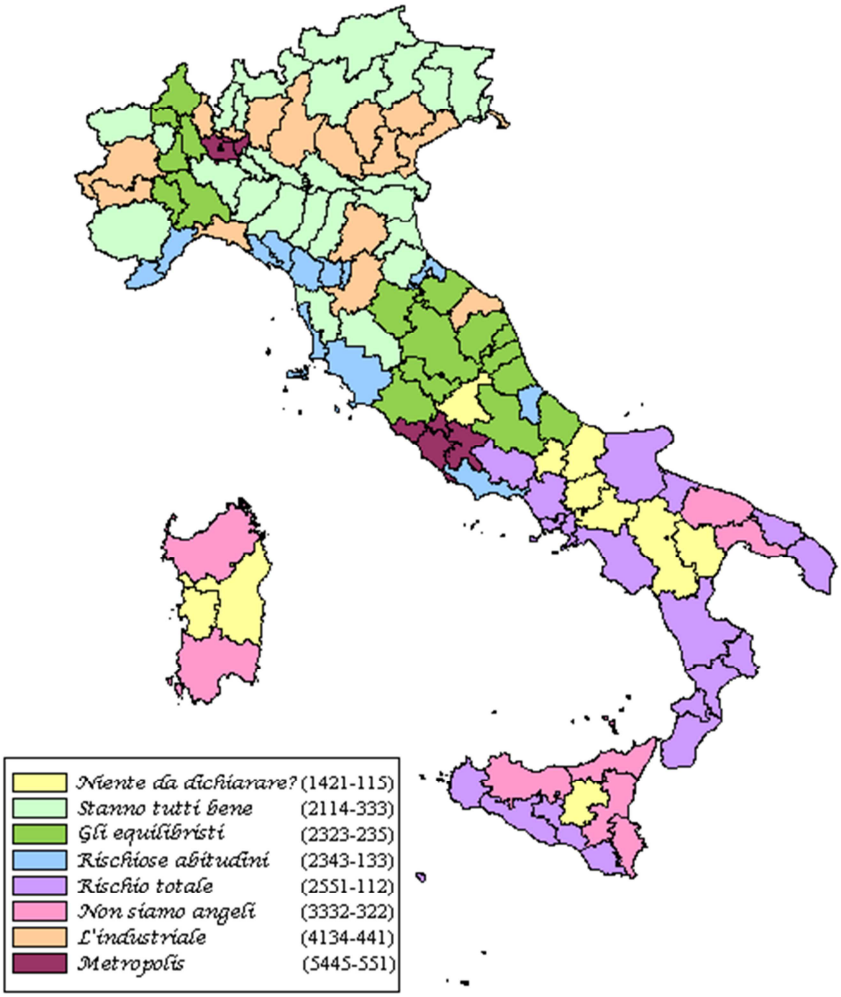
programmazione futura.

Si è ottenuta, in questo modo, una ripartizione del territorio nazionale in otto gruppi omogenei di DP con mercati di riferimento analoghi, la cui rappresentazione grafica è riportata nella mappa in Figura 4.1.

Nella legenda, è indicato il nome del gruppo e, fra parentesi, un codice a sette cifre che riassume le caratteristiche di ogni raggruppamento.

Come nome si è scelto il titolo di un film che richiamasse in modo efficace un tratto distintivo di quel cluster. Il codice esprime invece, per ogni cifra, il livello di una corrispondente area tematica: la prima cifra è legata alla numerosità del bacino, la seconda alla pericolosità fiscale e, a seguire, pericolosità sociale, tenore di vita, maturità della struttura produttiva, livello di tecnologia e servizi e, infine, disponibilità di infrastrutture di trasporto. Il valore di ogni cifra è un numero intero tra 1 e 5 e rappresenta la distanza dalla media nazionale: ad esempio, un valore uguale a 3 per la numerosità del bacino indica una dimensione media delle DP di quel cluster prossima alla media nazionale; un valore uguale a 5, invece, indica una dimensione media delle DP di quel cluster apprezzabilmente superiore alla media nazionale.

**Figura 4.1** *Mappa degli otto cluster*



Nel dettaglio, per caratterizzare le singole classi ottenute si sono utilizzati gli indici sintetici calcolati nella fase 1, confrontando la media di ogni indicatore nel singolo cluster con la media nazionale. In questo modo è stato possibile individuare quali aspetti contraddistinguono maggiormente ogni gruppo, in senso positivo o negativo.

Per calcolare i codici identificativi di ogni cluster, la matrice degli indici X (di dimensione n*p, dove n = 8 è il numero dei cluster e p = 7 è il numero degli indici sintetici calcolati) è stata trasformata in una matrice di punteggio, adottando il seguente algoritmo per ogni elemento $x_{ij}$:

**Tabella 4.1**: *Algoritmo per la definizione della matrice di punteggio*

| **Intervallo considerato** | **Score** |
|---|---|
| $\mathrm{m} \leq x_{\mathrm{ij}} < \mathrm{m} + \frac{2}{5}[\overline{X} - \mathrm{m}]$ | 1 |
| $\mathrm{m} + \frac{2}{5}[\overline{X} - \mathrm{m}] \leq x_{\mathrm{ij}} < \mathrm{m} + \frac{4}{5}[\overline{X} - \mathrm{m}]$ | 2 |
| $\mathrm{m} + \frac{4}{5}[\overline{X} - \mathrm{m}] \leq x_{\mathrm{ij}} < \overline{X} + \frac{1}{5}[\mathrm{M} - \overline{X}]$ | 3 |
| $\overline{X} + \frac{1}{5}[\mathrm{M} - \overline{X}] \leq x_{\mathrm{ij}} < \overline{X} + \frac{3}{5}[\mathrm{M} - \overline{X}]$ | 4 |
| $\overline{X} + \frac{3}{5}[\mathrm{M} - \overline{X}] \leq x_{\mathrm{ij}} < \mathrm{M}$ | 5 |

dove

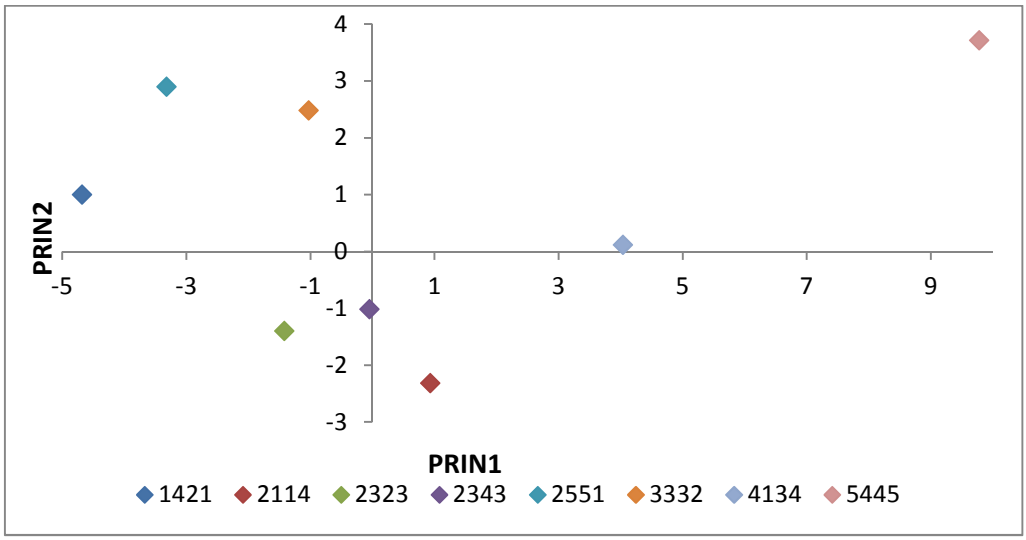
$m = \mathrm{MIN}_i\, x_{ij}$ $\qquad$ $\mathrm{M} = \mathrm{MAX}_i\, x_{ij}$ $\qquad$ $\overline{X} = Media\ nazionale\ dell'area\ tematica\ j$

In sostanza, gli intervalli fra il minimo e la media e tra la media e il massimo sono stati divisi ciascuno in cinque intervalli di uguale dimensione; raggruppando poi gli intervalli contigui a due a due, si sono ottenute 5 classi, attraverso le quali si riesce a codificare in una scala 1 - 5 la distanza della media del cluster dalla media nazionale.

Per ottenere una rappresentazione grafica significativa dei cluster ottenuti, si è eseguita una analisi in componenti principali basata su tutte le 36 variabili utilizzate nell'analisi di cluster.

L'osservazione sul piano fattoriale dei centroidi dei gruppi consente una visone sintetica pregnante dei cluster ottenuti, dato che i primi due assi fattoriali rappresentano da soli il 52% della variabilità del fenomeno (Figura 4.2). Il primo asse fattoriale rappresenta soprattutto l'aspetto dimensionale e di ricchezza, il secondo descrive in particolar modo la pericolosità fiscale e sociale.

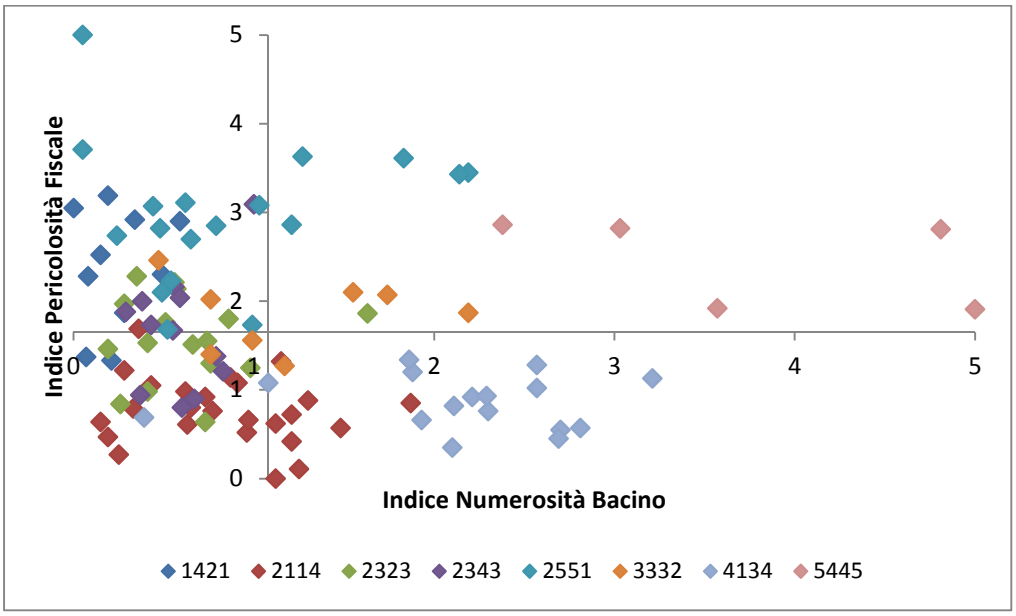
**Figura 4.2** *Centroidi dei cluster sul piano fattoriale*




Dato che il significato attribuito ai fattori è però piuttosto complesso, si può ottenere una rappresentazione altrettanto significativa dei cluster ottenuti considerando la distribuzione delle direzioni provinciali in base all'indice di numerosità del bacino e dell'indice di pericolosità fiscale, fissando l'origine degli assi sulla media nazionale (Figura 4.3).

Si osserva una addensarsi della nuvola su valori inferiori alla media nazionale lungo l'asse della numerosità del bacino, mentre lungo l'asse della pericolosità fiscale i dati appaiono più distribuiti e si staccano le DP appartenenti al cluster 2551 sul semiasse positivo, e quelli appartenenti al cluster 2114 sul semiasse negativo.

**Figura 4.3** *Distribuzione delle DP in base agli indici di numerosità del bacino e pericolosità fiscale*



## 4.1 Caratterizzazione dei cluster

Di seguito sono presentati i cluster in ordine crescente di numerosità del bacino di utenti e di pericolosità.

Il *cluster* "**Niente da dichiarare?**" (**1421-115** - colorato in rosa scuro nella Figura 4.1) si compone di dieci DP e raggruppa le Direzioni dell'entroterra del sud insieme a Rieti, Nuoro ed Oristano[2].

Risulta avere una piccola numerosità del bacino e bassi livelli di tenore di vita, maturità della struttura produttiva e di tecnologia e servizi; mentre la pericolosità sociale è medio - bassa, quella fiscale è abbastanza rilevante; infine, raggiunge i valori massimi di disponibilità di infrastrutture di trasporto.

All'interno del raggruppamento, risaltano alcuni valori che si discostano dalla media del cluster, ovvero:

- Isernia, con una pericolosità sociale medio - alta
- Rieti, con un tenore di vita medio
- Avellino, Enna e Potenza, con una maturità della struttura produttiva media
- Nuoro, con una disponibilità di infrastrutture di trasporto medio - bassa

Il *cluster* "**Stanno tutti bene**" (**2114-333** - colorato in verde chiaro nella Figura 4.1) è costituito da ventisei Direzioni[3] site nel centro-nord, che si caratterizzano per dimensioni medio - piccole ed un tenore di vita abbastanza elevato; non risultano pericolose né fiscalmente, né socialmente e si posizionano nella media per quanto riguarda maturità della struttura produttiva, livello di tecnologia e servizi e disponibilità di infrastrutture di trasporto.

Tra le DP che divergono dalla media del cluster, si distinguono:

- Modena, con una numerosità del bacino medio – alta
- Lodi, con una pericolosità fiscale media
- Biella e Sondrio, con una pericolosità sociale media
- Como, Lecco, Modena e Reggio Emilia, con una disponibilità di infrastrutture di trasporto bassa
- Aosta, con una pericolosità sociale media ed alti livelli di tecnologia e servizi e disponibilità di infrastrutture di trasporto

[2] Avellino, Benevento, Campobasso, Enna, Isernia, Matera, Nuoro, Oristano, Potenza, Rieti.

[3] Aosta, Belluno, Biella, Bolzano, Como, Cremona, Cuneo, Ferrara, Forlì - Cesena, Gorizia, Lecco, Lodi, Mantova, Modena, Parma, Pavia, Piacenza, Pisa, Pordenone, Ravenna, Reggio Emilia, Rovigo, Siena, Sondrio, Trento, Udine.

Il *cluster* "**Gli equilibristi**" (**2323-235** - colorato in verde scuro nella Figura 4.1**Errore. L'origine riferimento non è stata trovata.**) è formato da sedici Direzioni del centro-nord[4].

Accanto ad una disponibilità di infrastrutture di trasporto elevata, si osservano valori medi di pericolosità fiscale, tenore di vita e tecnologia e servizi e valori medio – bassi per le restanti aree tematiche.

Rispetto alla media del cluster, si differenziano le seguenti DP:

- Novara, con una pericolosità fiscale bassa ed una maturità della struttura produttiva medio -alta
- Fermo, Perugia e Pesaro - Urbino, con una medio - bassa disponibilità di infrastrutture di trasporto

Il *cluster* "**Rischiose abitudini**" (**2343-133** - colorato in azzurro nella Figura 4.1) è composto da dodici Direzioni[5] e riunisce alcune DP liguri e toscane a Rimini, Pescara e Latina.

In generale, sono caratterizzate da valori medi in tutte le aree tematiche, eccetto che nella numerosità del bacino (medio - bassa), nella pericolosità sociale (medio - alta) e nella maturità della struttura produttiva (bassa).

Rispetto alla media del cluster, si differenziano le seguenti DP:

- Lucca e Massa - Carrara, con una pericolosità sociale medio - bassa
- La Spezia, Livorno e Pescara, con una media maturità della struttura produttiva
- Prato, Rimini e Latina, con una disponibilità di infrastrutture di trasporto bassa

Nel *cluster* "**Rischio totale**" (**2551-112** - colorato in lilla nella Figura 4.1) abbiamo diciotto DP del sud[6]. Si sottolinea che le due DP di Napoli, inizialmente formanti un gruppo a sè, sono state accorpate in questo insieme per maggiore analogia relativamente alle pericolosità ed al tenore di vita.

Le caratteristiche di questa serie di DP sono: valori massimi di pericolosità fiscale e sociale, bassi livelli di tenore di vita, maturità della struttura produttiva e tecnologia e servizi e grado medio - basso di disponibilità di infrastrutture di trasporto e numerosità del bacino.

Le DP che non si allineano alle medie del cluster sono:

---

[4] Alessandria, Arezzo, Ascoli, Asti, Chieti, Fermo, L'Aquila, Macerata, Novara, Perugia, Pesaro - Urbino, Teramo, Terni, Verbano-Cusio-Ossola, Vercelli, Viterbo.

[5] Grosseto, Imperia, La Spezia, Latina, Livorno, Lucca, Massa - Carrara, Pescara, Pistoia, Prato, Rimini, Savona.

[6] Agrigento, Barletta Andria e Trani, Brindisi, Caltanissetta, Caserta, Catanzaro,Cosenza, Crotone, Foggia, Frosinone, Lecce, Napoli I, Napoli II, Ragusa, Reggio Calabria, Salerno, Trapani, Vibo Valentia.

- Napoli I e II, con una numerosità del bacino medio - alta ed una maturità della struttura produttiva media
- Foggia, con una media pericolosità fiscale e medio - alta disponibilità di infrastrutture di trasporto
- Agrigento, BAT e Brindisi, con una media pericolosità fiscale
- Lecce, con una medio - bassa pericolosità sociale
- Caltanissetta e Reggio Calabria, con una maturità della struttura produttiva media
- Vibo Valentia, con medio - alta disponibilità di infrastrutture di trasporto

Il *cluster* "**Non siamo angeli**" (**3332-322** - colorato in rosa nella Figura 4.1**Errore. L'origine riferimento non è stata trovata.**) raccoglie otto Direzioni[7] del sud e isole e presenta valori in linea con la media nazionale, con l'esclusione di tenore di vita, tecnologia e servizi e disponibilità di infrastrutture di trasporto, che hanno numeri leggermente inferiori.

Non ci sono significativi spostamenti delle DP dalle medie del cluster.

Il *cluster* "**L'industriale**" (**4134-441** - colorato in salmone nella Figura 4.1) riunisce sedici DP[8] del centro-nord.

Mentre la pericolosità fiscale e la disponibilità di infrastrutture di trasporto sono basse, la pericolosità sociale si accosta alla media nazionale, che è invece superata in tutte le restanti aree tematiche.

Casi particolari difformi dalle medie del cluster sono:

- Trieste, con una bassa numerosità del bacino
- Monza e Brianza, con una media pericolosità fiscale
- Treviso e Vicenza, con una bassa pericolosità sociale
- Ancona e Bologna, con una media disponibilità di infrastrutture di trasporto

Il *cluster* "**Metropolis**" (**5445-551** - colorato in bordeaux nella Figura 4.1) è formato dall'unione di due cluster a sé stanti, ovvero quello contenente le tre DP di Roma e quello con le due DP di Milano.

È un cluster con valori massimi per tutte le aree tematiche, con l'eccezione della disponibilità di infrastrutture di trasporto (bassa) e delle pericolosità fiscale e sociale, che si attestano su livelli medio - alti.

---

[7] Bari, Cagliari, Catania, Messina, Palermo, Sassari, Siracusa, Taranto.

[8] Ancona, Bergamo, Bologna, Brescia, Firenze, Genova, Monza e Brianza, Padova, Torino I, Torino II, Treviso, Trieste, Varese, Venezia, Verona, Vicenza.

In appendice C si riporta una tabella riassuntiva contenente il cluster di appartenenza di ogni Direzione Provinciale, insieme ai valori degli indici sintetici di ogni area tematica.

## 5. Conclusioni

In questo lavoro è stata analizzata la domanda di mercato nei confronti dell'Agenzia delle Entrate, utilizzando alcune variabili presenti nel DB Geo provinciale.

Tramite gli indici sintetici sono state riassunte (attraverso un numero unico per ogni area tematica) le caratteristiche di ogni Direzione Provinciale, al fine di poter analizzare particolari aspetti inerenti la singola area e, quindi, fornire uno strumento più efficace per le decisioni aziendali.

Sono stati osservati livelli maggiori di numerosità del bacino, tenore di vita e tecnologia e servizi al centro-nord, mentre la pericolosità fiscale e, in buona parte, anche quella sociale sono maggiormente concentrati al sud. I livelli di maturità della struttura produttiva e disponibilità di infrastrutture di trasporto presentano invece distribuzioni più variegate all'interno del territorio nazionale.

Nella seconda fase, si sono identificate le affinità tra le DP considerando complessivamente tutte le 7 aree tematiche, con lo scopo di ottenere gruppi omogenei sui quali applicare profili di management personalizzati; inoltre, questa suddivisione dell'Italia consente di effettuare confronti più significativi tra DP dello stesso cluster.

Sono stati ottenuti otto raggruppamenti. Nello specifico, si evidenzia che le zone maggiormente rischiose a livello fiscale e sociale sono raggruppate nel cluster "Rischio totale", presente nel sud Italia e nelle isole maggiori, mentre una certa pericolosità fiscale si riscontra anche in zone meno pericolose socialmente, appartenenti soprattutto all'entroterra del sud (cluster "Niente da dichiarare?") e alle metropoli di Roma e Milano (cluster "Metropolis"). Le DP meno pericolose fiscalmente e socialmente appartengono, invece, al cluster "Stanno tutti bene", situato nel centro-nord e caratterizzato da un tenore di vita abbastanza elevato.

## Appendice

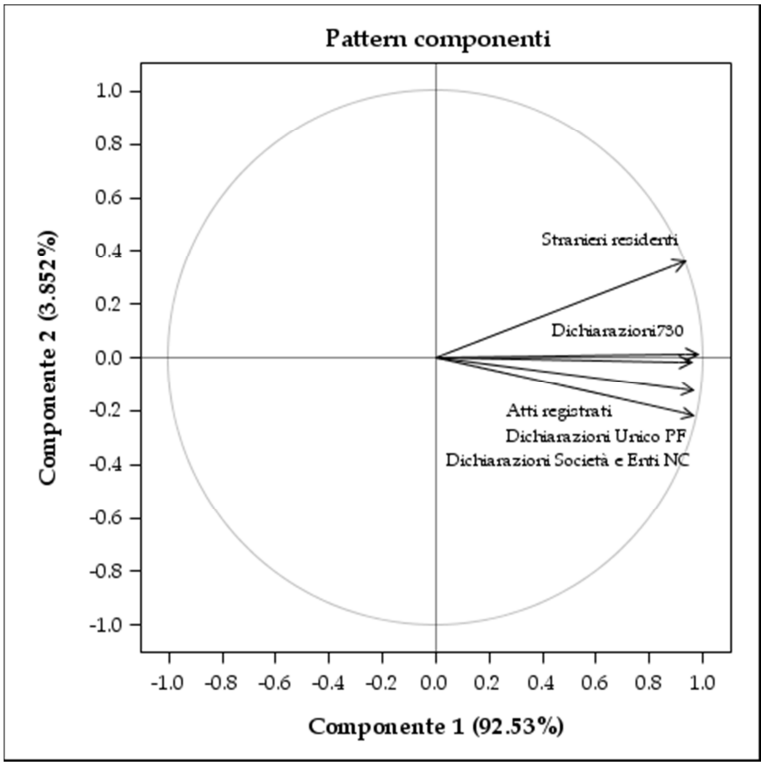
### A. Risultati ACP

***Numerosità del Bacino***

| *Matrice di correlazione* | | | | | |
|---|---|---|---|---|---|
| | *Dichiarazioni 730* | *Dichiarazioni Unico PF* | *Dichiarazioni Società e Enti NC* | *Atti registrati* | *Stranieri residenti* |
| Dichiarazioni730 | 1.0000 | 0.9452 | 0.9051 | 0.8812 | 0.8854 |
| Dichiarazioni Unico PF | 0.9452 | 1.0000 | 0.9371 | 0.9391 | 0.8259 |
| Dichiarazioni Società e Enti NC | 0.9051 | 0.9371 | 1.0000 | 0.9683 | 0.9177 |
| Atti registrati | 0.8812 | 0.9391 | 0.9683 | 1.0000 | 0.8578 |
| Stranieri residenti | 0.8854 | 0.8259 | 0.9177 | 0.8578 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 4.62650540 | 0.9253 |
| 2 | 0.19258324 | 0.9638 |
| 3 | 0.13890254 | 0.9916 |
| 4 | 0.02544964 | 0.9967 |
| 5 | 0.01655917 | 1.0000 |

| *Autovettori* | |
|---|---|
| | *PRIN1* |
| Dichiarazioni730 | 0.446269 |
| Dichiarazioni Unico PF | 0.449522 |
| Dichiarazioni Società e Enti NC | 0.457193 |
| Atti registrati | 0.449422 |
| Stranieri residenti | 0.433324 |

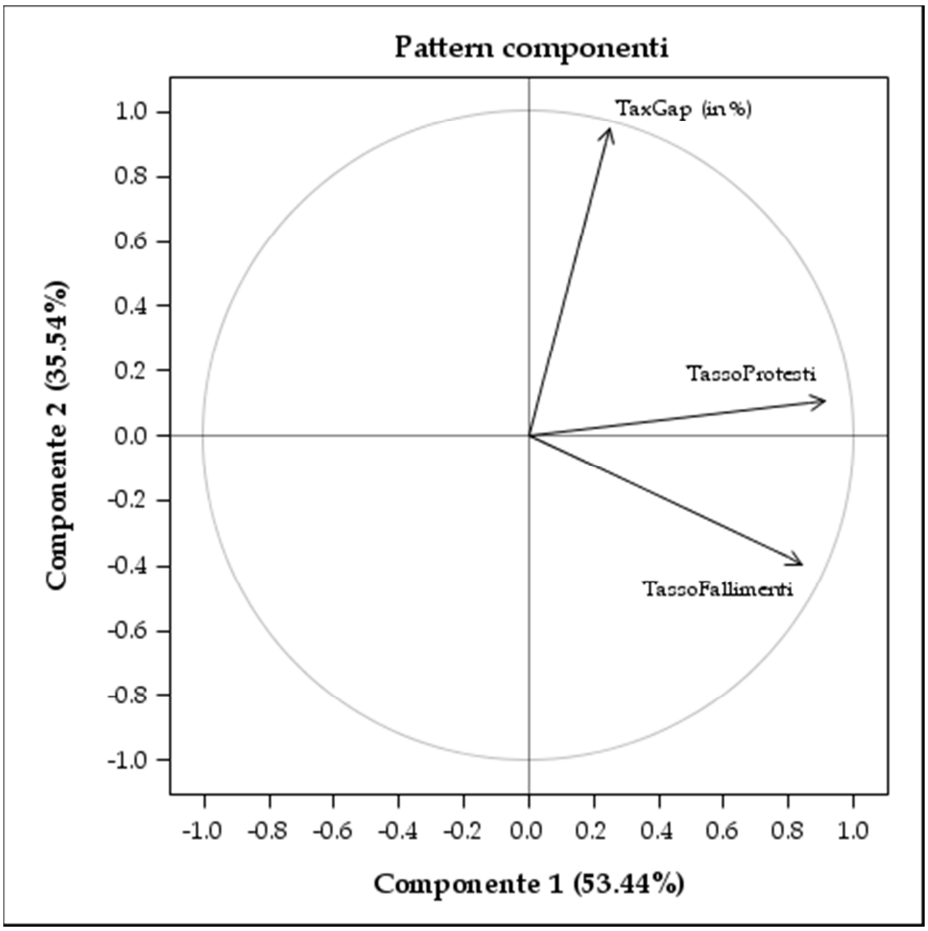

Pattern componenti
Stranieri residenti
Dichiarazioni730
Atti registrati
Dichiarazioni Unico PF
Dichiarazioni Società e Enti NC
Componente 2 (3.852%)
Componente 1 (92.53%)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

## *Pericolosità fiscale*

| *Matrice di correlazione* | | | |
|---|---|---|---|
| | *TaxGap (in %)* | *TassoFallimenti* | *TassoProtesti* |
| TaxGap (in %) | 1.0000 | -.0923 | 0.2487 |
| TassoFallimenti | -.0923 | 1.0000 | 0.5812 |
| TassoProtesti | 0.2487 | 0.5812 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 1.60323078 | 0.5344 |
| 2 | 1.06607603 | 0.8898 |
| 3 | 0.33069319 | 1.0000 |

| *Autovettori* | | |
|---|---|---|
| | *PRIN1* | *PRIN2* |
| TaxGap (in %) | 0.195557 | 0.918679 |
| TassoFallimenti | 0.664732 | -.381469 |
| TassoProtesti | 0.721033 | 0.102521 |

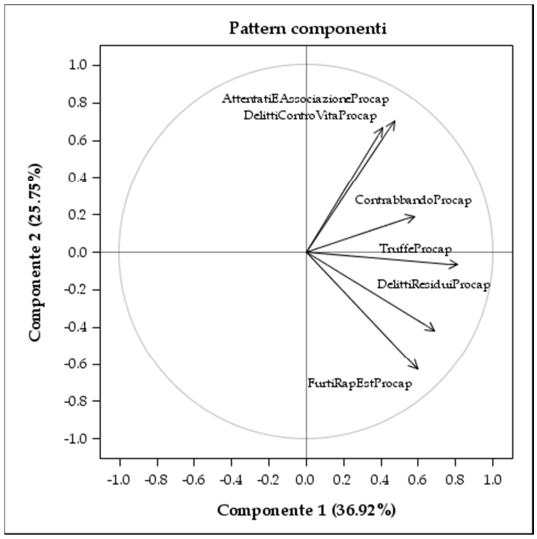


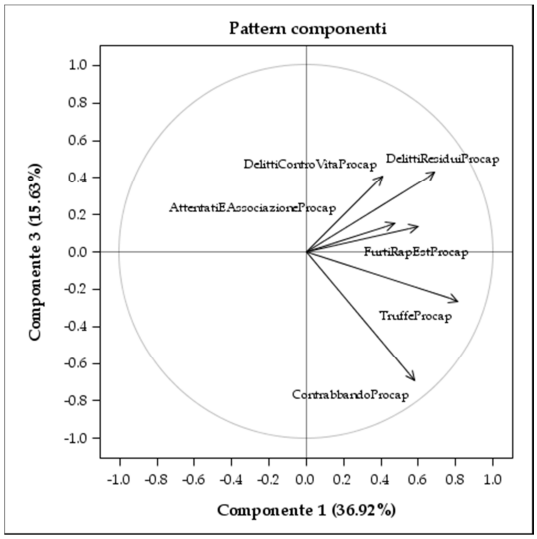
## *Pericolosità sociale*

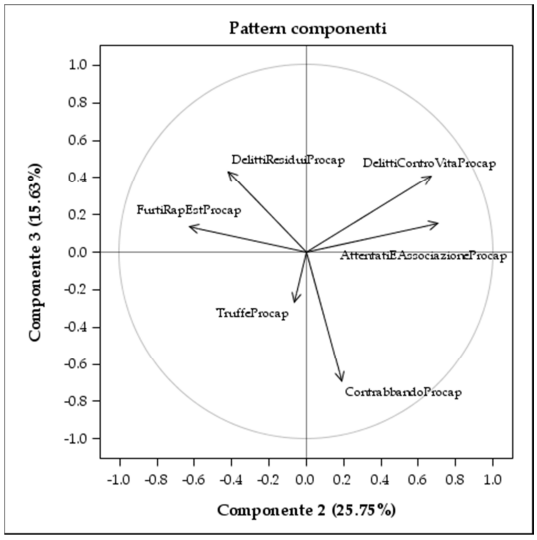
| *Matrice di correlazione* | | | | | | |
|---|---|---|---|---|---|---|
| | *Delitti ControVita Procap* | *FurtiRapEst Procap* | *Truffe Procap* | *Attentati e Associazione Procap* | *Contrabbando Procap* | *Delitti Residui Procap* |
| DelittiControVitaProcap | 1.0000 | -.0702 | 0.1162 | 0.5027 | 0.1798 | 0.1451 |
| FurtiRapEstProcap | -.0702 | 1.0000 | 0.3754 | -.0881 | 0.1648 | 0.5722 |
| TruffeProcap | 0.1162 | 0.3754 | 1.0000 | 0.3092 | 0.4605 | 0.4398 |
| AttentatiEAssociazioneProcap | 0.5027 | -.0881 | 0.3092 | 1.0000 | 0.2165 | 0.0663 |
| ContrabbandoProcap | 0.1798 | 0.1648 | 0.4605 | 0.2165 | 1.0000 | 0.0829 |
| DelittiResiduiProcap | 0.1451 | 0.5722 | 0.4398 | 0.0663 | 0.0829 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 2.21521613 | 0.3692 |
| 2 | 1.54527680 | 0.6267 |
| 3 | 0.93795527 | 0.7831 |
| 4 | 0.57341174 | 0.8786 |
| 5 | 0.41596416 | 0.9480 |
| 6 | 0.31217591 | 1.0000 |

| *Autovettori* | | | |
|---|---|---|---|
| | *PRIN1* | *PRIN2* | *PRIN3* |
| DelittiControVitaProcap | 0.273372 | 0.536149 | 0.420488 |
| FurtiRapEstProcap | 0.400535 | -.503261 | 0.139191 |
| TruffeProcap | 0.545851 | -.052978 | -.271768 |
| AttentatiEAssociazioneProcap | 0.317623 | 0.565632 | 0.157105 |
| ContrabbandoProcap | 0.389731 | 0.153439 | -.712741 |
| DelittiResiduiProcap | 0.462722 | -.336128 | 0.444157 |


Pattern componenti
AttentatiEAssociazioneProcap
DelittiControVitaProcap
ContrabbandoProcap
TruffeProcap
DelittiResiduiProcap
FurtiRapEstProcap
Componente 2 (25.75%)
Componente 1 (36.92%)


Pattern componenti
DelittiControVitaProcap
DelittiResiduiProcap
AttentatiEAssociazioneProcap
FurtiRapEstProcap
TruffeProcap
ContrabbandoProcap
Componente 3 (15.63%)
Componente 1 (36.92%)

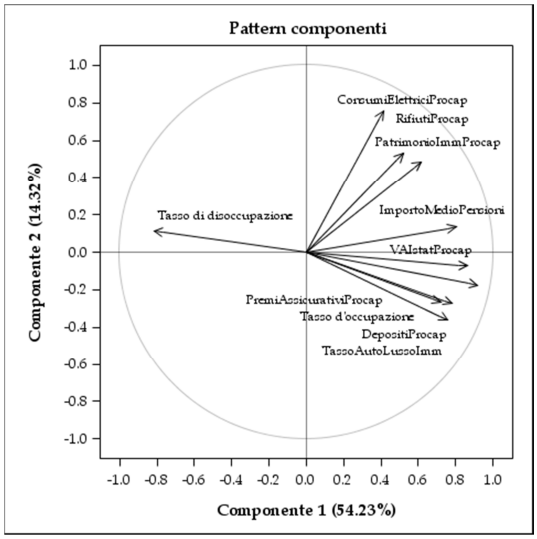

Pattern componenti
DelittiResiduiProcap
DelittiControVitaProcap
FurtiRapEstProcap
AttentatiEAssociazioneProcap
TruffeProcap
ContrabbandoProcap
Componente 3 (15.63%)
Componente 2 (25.75%)

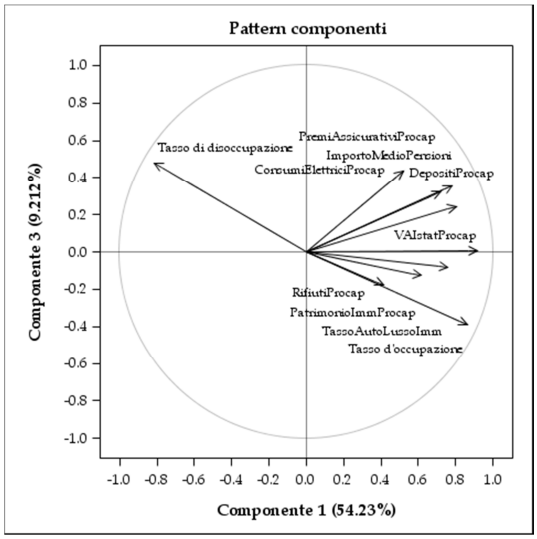
## *Tenore di vita*

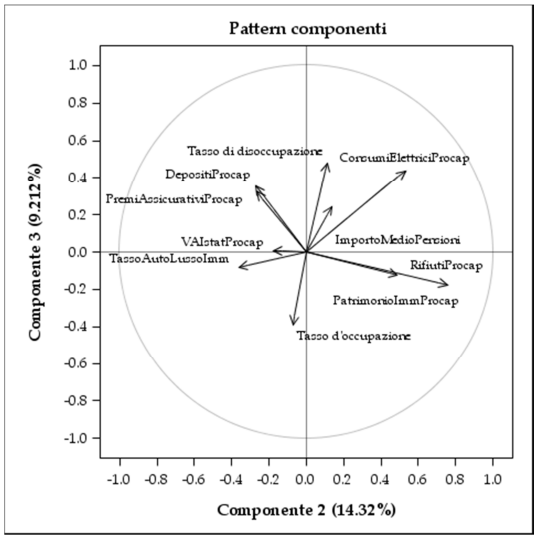
| *Matrice di correlazione* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Depositi Procap* | *VAIstat Procap* | *Patrimonio Imm Procap* | *Importo Medio Pensioni* | *Tasso Auto Lusso Imm* | *Consumi Elettrici Procap* | *Rifiuti Procap* | *Tasso d'occup* | *Tasso di disoccup* | *Premi Assicurativi Procap* |
| DepositiProcap | 1.0000 | 0.7953 | 0.3884 | 0.6318 | 0.6220 | 0.3240 | 0.0885 | 0.5309 | -.4600 | 0.6596 |
| VAIstatProcap | 0.7953 | 1.0000 | 0.4753 | 0.6419 | 0.7280 | 0.3708 | 0.2827 | 0.7858 | -.7281 | 0.6894 |
| PatrimonioImmProcap | 0.3884 | 0.4753 | 1.0000 | 0.5221 | 0.2620 | 0.3749 | 0.5067 | 0.4617 | -.4561 | 0.2678 |
| ImportoMedioPensioni | 0.6318 | 0.6419 | 0.5221 | 1.0000 | 0.4970 | 0.5352 | 0.3136 | 0.6020 | -.5482 | 0.5486 |
| TassoAutoLussoImm | 0.6220 | 0.7280 | 0.2620 | 0.4970 | 1.0000 | 0.2434 | 0.1008 | 0.6745 | -.6272 | 0.5193 |
| ConsumiElettriciProcap | 0.3240 | 0.3708 | 0.3749 | 0.5352 | 0.2434 | 1.0000 | 0.4543 | 0.3340 | -.2115 | 0.2960 |
| RifiutiProcap | 0.0885 | 0.2827 | 0.5067 | 0.3136 | 0.1008 | 0.4543 | 1.0000 | 0.3239 | -.2963 | 0.1383 |
| Tasso d'occupazione | 0.5309 | 0.7858 | 0.4617 | 0.6020 | 0.6745 | 0.3340 | 0.3239 | 1.0000 | -.9188 | 0.4796 |
| Tasso di disoccupazione | -.4600 | -.7281 | -.4561 | -.5482 | -.6272 | -.2115 | -.2963 | -.9188 | 1.0000 | -.5006 |
| PremiAssicurativiProcap | 0.6596 | 0.6894 | 0.2678 | 0.5486 | 0.5193 | 0.2960 | 0.1383 | 0.4796 | -.5006 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 5.42316920 | 0.5423 |
| 2 | 1.43175434 | 0.6855 |
| 3 | 0.92120129 | 0.7776 |
| 4 | 0.58849633 | 0.8365 |

| *Autovettori* | | | |
|---|---|---|---|
| | *PRIN1* | *PRIN2* | *PRIN3* |
| DepositiProcap | 0.336022 | -.228535 | 0.369540 |
| VAIstatProcap | 0.394179 | -.146733 | 0.006602 |
| PatrimonioImmProcap | 0.264600 | 0.407485 | -.130199 |
| ImportoMedioPensioni | 0.344783 | 0.112237 | 0.255744 |
| TassoAutoLussoImm | 0.325361 | -.302446 | -.085338 |
| ConsumiElettriciProcap | 0.223338 | 0.442939 | 0.457838 |
| RifiutiProcap | 0.177315 | 0.630731 | -.183491 |
| Tasso d'occupazione | 0.370666 | -.060505 | -.406790 |
| Tasso di disoccupazione | -.350489 | 0.094325 | 0.501447 |
| PremiAssicurativiProcap | 0.308881 | -.222606 | 0.336967 |


Pattern componenti
ConsumiElettriciProcap
RifiutiProcap
PatrimonioImmProcap
ImportoMedioPensioni
Tasso di disoccupazione
VAIstatProcap
PremiAssicurativiProcap
Tasso d'occupazione
DepositiProcap
TassoAutoLussoImm
Componente 2 (14.32%)
Componente 1 (54.23%)


Pattern componenti
PremiAssicurativiProcap
Tasso di disoccupazione
ImportoMedioPensioni
ConsumiElettriciProcap
DepositiProcap
VAIstatProcap
RifiutiProcap
PatrimonioImmProcap
TassoAutoLussoImm
Tasso d'occupazione
Componente 3 (9.212%)
Componente 1 (54.23%)

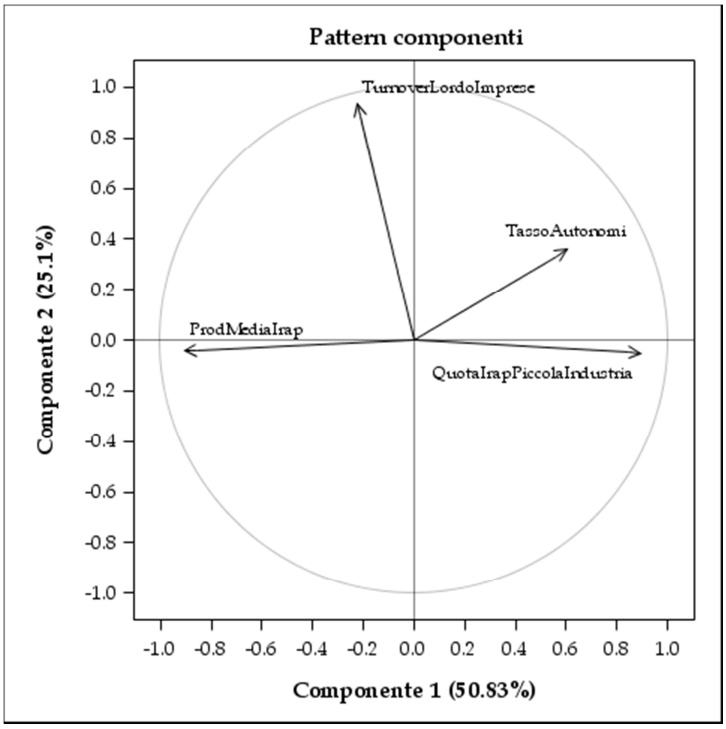

Pattern componenti
Tasso di disoccupazione
ConsumiElettriciProcap
DepositiProcap
PremiAssicurativiProcap
VAIstatProcap
ImportoMedioPensioni
TassoAutoLussoImm
RifiutiProcap
PatrimonioImmProcap
Tasso d'occupazione
Componente 3 (9.212%)
Componente 2 (14.32%)

## *Maturità della struttura produttiva*

| *Matrice di correlazione* | | | | |
|---|---|---|---|---|
| | *TassoAutonomi* | *ProdMediaIrap* | *TurnoverLordoImprese* | *QuotaIrapPiccolaIndustria* |
| TassoAutonomi | 1.0000 | -.3755 | 0.0011 | 0.3154 |
| ProdMediaIrap | -.3755 | 1.0000 | 0.1025 | -.7698 |
| TurnoverLordoImprese | 0.0011 | 0.1025 | 1.0000 | -.1571 |
| QuotaIrapPiccolaIndustria | 0.3154 | -.7698 | -.1571 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 2.03329554 | 0.5083 |
| 2 | 1.00412488 | 0.7594 |
| 3 | 0.73691839 | 0.9436 |
| 4 | 0.22566120 | 1.0000 |

| *Autovettori* | | |
|---|---|---|
| | *PRIN1* | *PRIN2* |
| TassoAutonomi | 0.422119 | 0.357635 |
| ProdMediaIrap | -.635848 | -.043340 |
| TurnoverLordoImprese | -.157886 | 0.931504 |
| QuotaIrapPiccolaIndustria | 0.626565 | -.050197 |



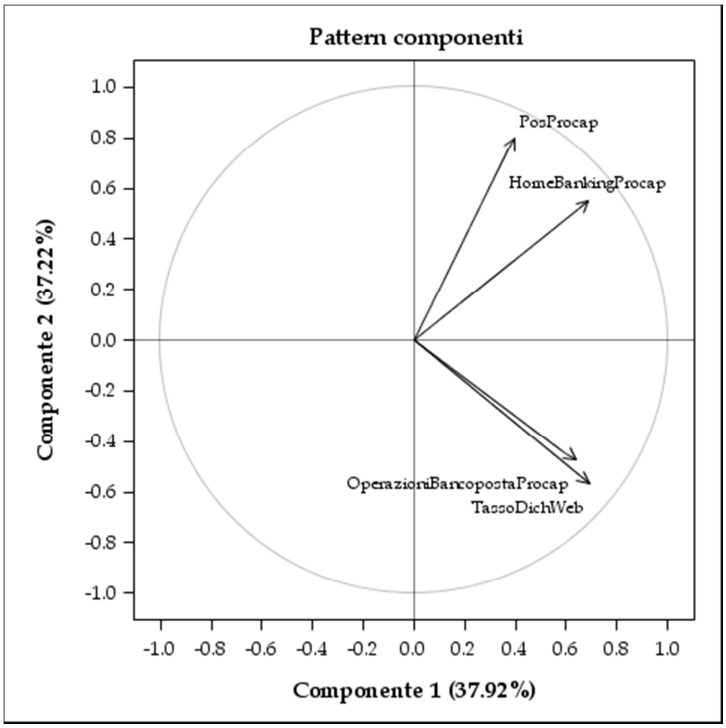

La classifica ottenuta è stata poi invertita per mostrare il grado di maturità di ogni DP.

## *Livello di tecnologia e servizi*

| *Matrice di correlazione* | | | | |
|---|---|---|---|---|
| | *Operazioni Bancoposta Procap* | *Pos Procap* | *Tasso DichWeb* | *Home Banking Procap* |
| OperazioniBancopostaProcap | 1.0000 | 0.0383 | 0.4583 | -.0050 |
| PosProcap | 0.0383 | 1.0000 | -.2346 | 0.4993 |
| TassoDichWeb | 0.4583 | -.2346 | 1.0000 | 0.2317 |
| HomeBankingProcap | -.0050 | 0.4993 | 0.2317 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 1.51699241 | 0.3792 |
| 2 | 1.48878543 | 0.7514 |
| 3 | 0.74929821 | 0.9388 |
| 4 | 0.24492395 | 1.0000 |

| *Autovettori* | | |
|---|---|---|
| | *PRIN1* | *PRIN2* |
| OperazioniBancopostaProcap | 0.517373 | -.389384 |
| PosProcap | 0.322217 | 0.654239 |
| TassoDichWeb | 0.562732 | -.464979 |
| HomeBankingProcap | 0.558421 | 0.451825 |

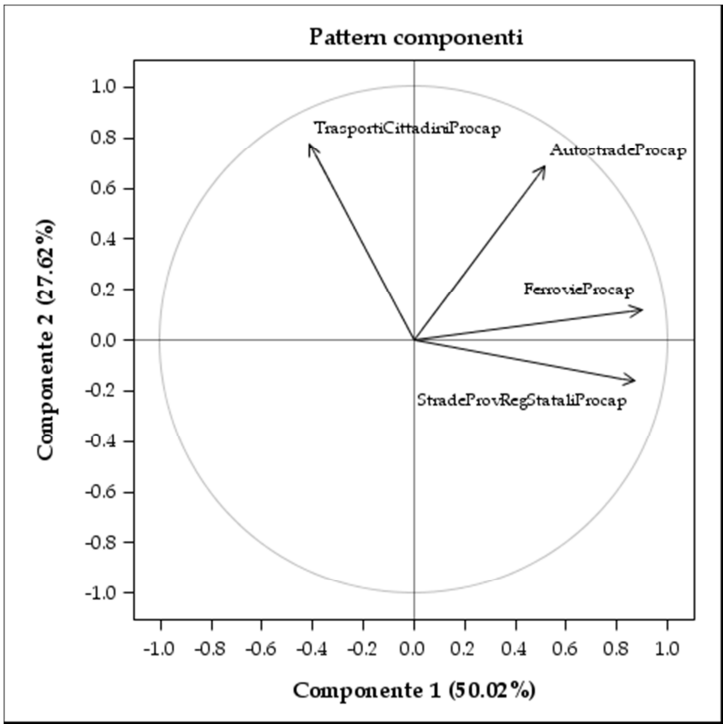


## *Disponibilità di infrastrutture di trasporto*

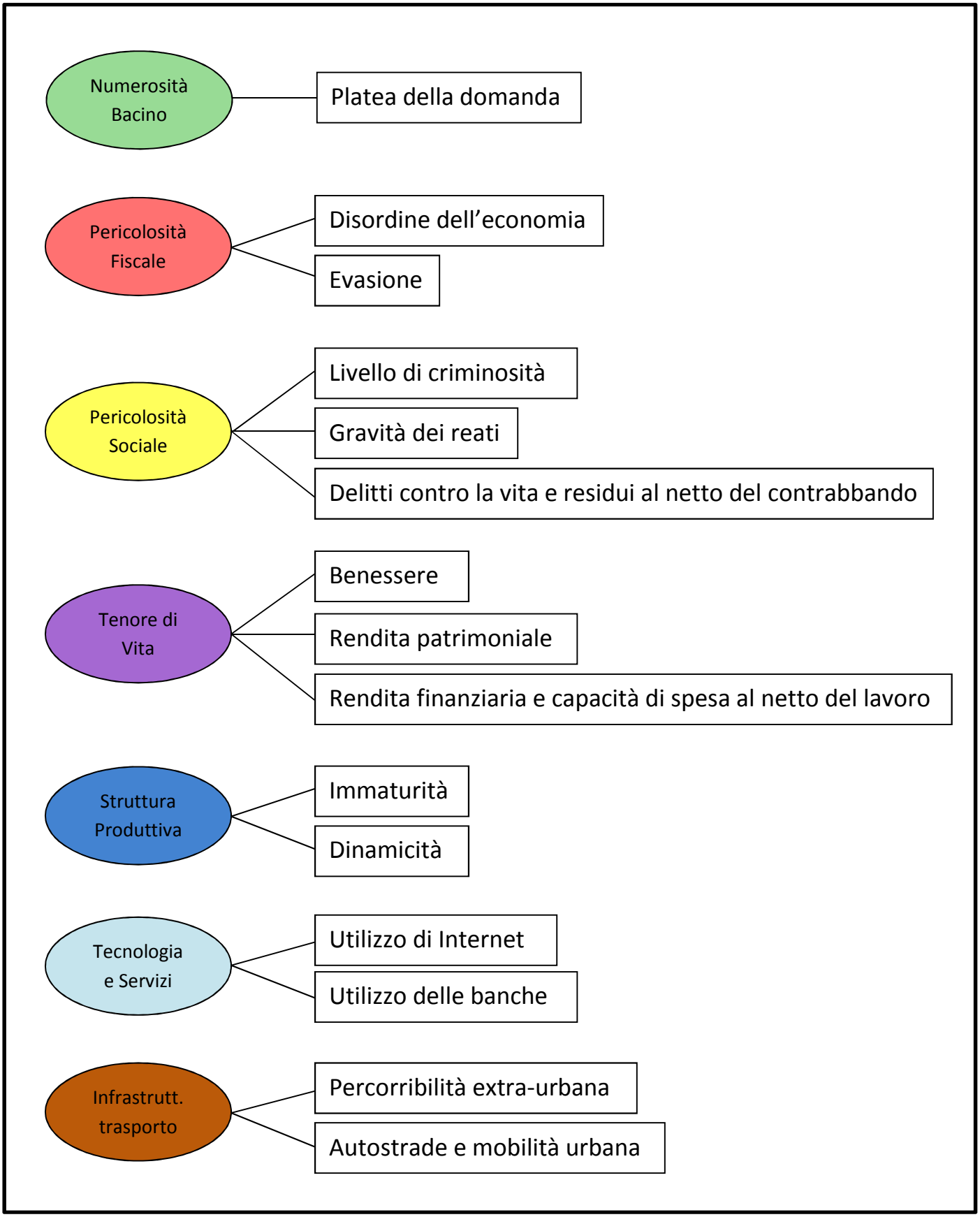
| *Matrice di correlazione* | | | | |
|---|---|---|---|---|
| | *Autostrade Procap* | *Strade ProvRegStatali Procap* | *Ferrovie Procap* | *Trasporti Cittadini Procap* |
| AutostradeProcap | 1.0000 | 0.2160 | 0.4010 | 0.0770 |
| StradeProvRegStataliProcap | 0.2160 | 1.0000 | 0.7006 | -.3193 |
| FerrovieProcap | 0.4010 | 0.7006 | 1.0000 | -.1950 |
| TrasportiCittadiniProcap | 0.0770 | -.3193 | -.1950 | 1.0000 |

| *Autovalori della matrice di correlazione* | | |
|---|---|---|
| | *Autovalore* | *Cumulativa* |
| 1 | 2.00070478 | 0.5002 |
| 2 | 1.10479237 | 0.7764 |
| 3 | 0.62739761 | 0.9332 |
| 4 | 0.26710524 | 1.0000 |

| *Autovettori* | | |
|---|---|---|
| | *PRIN1* | *PRIN2* |
| AutostradeProcap | 0.364686 | 0.651714 |
| StradeProvRegStataliProcap | 0.615957 | -.151943 |
| FerrovieProcap | 0.634274 | 0.111154 |
| TrasportiCittadiniProcap | -.292057 | 0.734730 |



**B. Interpretazione delle componenti principali**



## C. Indici sintetici e cluster delle Direzioni Provinciali

| Direzione Provinciale | Cluster | Numerosità bacino | Pericolosità fiscale | Pericolosità sociale | Tenore di vita | Struttura produttiva | Tecnologia e servizi | Infrastrutture di trasporto |
|---|---|---|---|---|---|---|---|---|
| Agrigento | 2551-112 | 0,54 | 2,23 | 1,33 | 0,21 | 0,87 | 0,11 | 0,72 |
| Alessandria | 2323-235 | 0,98 | 1,25 | 2,08 | 1,95 | 1,57 | 2,26 | 3,62 |
| Ancona | 4134-441 | 1,08 | 1,08 | 1,13 | 1,69 | 2,34 | 2,40 | 1,27 |
| Aosta | 2114-333 | 0,19 | 0,47 | 1,44 | 2,88 | 2,05 | 4,24 | 5,00 |
| Arezzo | 2323-235 | 0,76 | 1,30 | 1,24 | 1,78 | 1,36 | 1,78 | 1,69 |
| Ascoli Piceno | 2323-235 | 0,35 | 2,28 | 0,92 | 1,25 | 0,94 | 2,04 | 1,35 |
| Asti | 2323-235 | 0,41 | 0,98 | 1,30 | 1,61 | 1,10 | 1,63 | 3,09 |
| Avellino | 1421-115 | 0,59 | 2,90 | 1,25 | 0,21 | 1,36 | 0,34 | 1,85 |
| Bari | 3332-322 | 2,19 | 1,87 | 1,48 | 0,98 | 1,46 | 1,20 | 0,50 |
| Barletta Andria e Trani | 2551-112 | 0,49 | 2,10 | 1,49 | 0,57 | 0,32 | 0,13 | 0,64 |
| Belluno | 2114-333 | 0,33 | 0,79 | 0,01 | 1,97 | 2,78 | 2,62 | 1,82 |
| Benevento | 1421-115 | 0,34 | 2,92 | 1,12 | 0,29 | 0,59 | 0,34 | 1,34 |
| Bergamo | 4134-441 | 2,57 | 1,28 | 0,78 | 2,02 | 2,66 | 2,34 | 0,33 |
| Biella | 2114-333 | 0,28 | 1,22 | 1,19 | 2,06 | 1,90 | 2,33 | 0,58 |
| Bologna | 4134-441 | 2,69 | 0,45 | 2,17 | 3,03 | 2,79 | 3,67 | 1,11 |
| Bolzano | 2114-333 | 1,12 | 0,00 | 0,47 | 2,37 | 2,59 | 2,60 | 1,98 |
| Brescia | 4134-441 | 3,21 | 1,13 | 1,73 | 2,63 | 2,34 | 2,33 | 0,50 |
| Brindisi | 2551-112 | 0,52 | 1,68 | 2,14 | 0,68 | 1,28 | 0,59 | 0,61 |
| Cagliari | 3332-322 | 1,17 | 1,27 | 0,89 | 1,46 | 1,47 | 2,00 | 0,66 |
| Caltanissetta | 2551-112 | 0,24 | 2,74 | 2,25 | 0,39 | 1,70 | 0,16 | 1,67 |
| Campobasso | 1421-115 | 0,28 | 1,87 | 1,14 | 0,53 | 0,67 | 0,91 | 2,98 |
| Caserta | 2551-112 | 1,27 | 3,63 | 2,80 | 0,59 | 0,73 | 0,40 | 0,73 |
| Catania | 3332-322 | 1,55 | 2,10 | 1,84 | 0,87 | 1,38 | 1,30 | 0,53 |
| Catanzaro | 2551-112 | 0,44 | 3,07 | 3,14 | 0,61 | 1,08 | 0,90 | 1,70 |
| Chieti | 2323-235 | 0,66 | 1,51 | 0,74 | 0,89 | 1,96 | 1,24 | 1,63 |
| Como | 2114-333 | 1,21 | 0,72 | 0,44 | 1,94 | 2,05 | 2,26 | 0,17 |
| Cosenza | 2551-112 | 1,03 | 3,08 | 1,54 | 0,35 | 0,78 | 0,64 | 1,96 |
| Cremona | 2114-333 | 0,73 | 0,92 | 0,89 | 1,96 | 1,60 | 1,31 | 0,99 |
| Crotone | 2551-112 | 0,05 | 5,00 | 3,14 | 0,19 | 0,69 | 0,00 | 1,43 |
| Cuneo | 2114-333 | 1,48 | 0,57 | 0,76 | 2,33 | 1,22 | 2,07 | 2,41 |
| Enna | 1421-115 | 0,08 | 2,28 | 1,68 | 0,00 | 1,34 | 0,28 | 2,67 |
| Fermo | 2323-235 | 0,28 | 1,97 | 0,98 | 1,19 | 1,09 | 0,87 | 1,00 |
| Ferrara | 2114-333 | 0,77 | 0,76 | 0,75 | 2,13 | 1,13 | 1,29 | 0,63 |
| Firenze | 4134-441 | 2,57 | 1,02 | 1,81 | 2,66 | 2,18 | 3,37 | 0,99 |
| Foggia | 2551-112 | 0,99 | 1,73 | 3,74 | 0,43 | 0,68 | 0,53 | 2,11 |
| Forli'-Cesena | 2114-333 | 0,97 | 0,66 | 0,89 | 2,19 | 1,55 | 1,97 | 0,73 |
| Frosinone | 2551-112 | 0,79 | 2,85 | 1,89 | 0,91 | 1,29 | 0,90 | 1,44 |
| Genova | 4134-441 | 2,10 | 0,35 | 2,71 | 2,54 | 2,36 | 2,43 | 1,04 |

| Direzione Provinciale | Cluster | Numerosità bacino | Pericolosità fiscale | Pericolosità sociale | Tenore di vita | Struttura produttiva | Tecnologia e servizi | Infrastrutture di trasporto |
|---|---|---|---|---|---|---|---|---|
| Gorizia | 2114-333 | 0,15 | 0,64 | 0,77 | 1,69 | 2,14 | 2,59 | 0,92 |
| Grosseto | 2343-133 | 0,43 | 1,73 | 1,95 | 2,23 | 0,00 | 2,70 | 2,01 |
| Imperia | 2343-133 | 0,38 | 2,00 | 3,21 | 2,07 | 0,30 | 2,33 | 1,90 |
| Isernia | 1421-115 | 0,00 | 3,05 | 2,17 | 0,59 | 1,22 | 0,84 | 3,67 |
| La Spezia | 2343-133 | 0,37 | 0,94 | 1,85 | 1,97 | 1,63 | 2,35 | 2,00 |
| L'Aquila | 2323-235 | 0,51 | 1,76 | 0,86 | 0,96 | 0,95 | 1,27 | 4,31 |
| Latina | 2343-133 | 1,00 | 3,09 | 1,74 | 1,43 | 1,29 | 1,39 | 0,39 |
| Lecce | 2551-112 | 1,21 | 2,86 | 0,73 | 0,44 | 0,41 | 0,69 | 0,14 |
| Lecco | 2114-333 | 0,62 | 0,98 | 0,38 | 2,08 | 2,11 | 2,48 | 0,51 |
| Livorno | 2343-133 | 0,67 | 0,90 | 1,36 | 2,23 | 1,66 | 2,61 | 1,12 |
| Lodi | 2114-333 | 0,36 | 1,69 | 0,79 | 1,80 | 2,18 | 1,90 | 1,18 |
| Lucca | 2343-133 | 0,83 | 1,21 | 0,94 | 2,57 | 0,47 | 2,51 | 1,29 |
| Macerata | 2323-235 | 0,74 | 1,55 | 1,04 | 1,35 | 1,16 | 1,49 | 1,18 |
| Mantova | 2114-333 | 0,91 | 1,08 | 0,83 | 2,14 | 1,82 | 1,51 | 0,99 |
| Massa-Carrara | 2343-133 | 0,29 | 1,88 | 0,81 | 1,67 | 0,59 | 1,42 | 1,93 |
| Matera | 1421-115 | 0,19 | 3,19 | 0,84 | 0,38 | 1,09 | 0,69 | 2,00 |
| Messina | 3332-322 | 0,99 | 1,56 | 1,17 | 0,85 | 1,65 | 1,11 | 2,02 |
| Milano I | 5445-551 | 3,57 | 1,92 | 2,05 | 5,00 | 5,00 | 5,00 | 0,43 |
| Milano II | 5445-551 | 5,00 | 1,91 | 2,03 | 3,35 | 3,80 | 4,86 | 0,22 |
| Modena | 2114-333 | 1,87 | 0,85 | 0,95 | 2,52 | 2,08 | 2,64 | 0,41 |
| Monza e Brianza | 4134-441 | 1,86 | 1,34 | 2,03 | 2,17 | 2,08 | 2,99 | 0,09 |
| Napoli I | 2551-112 | 2,19 | 3,45 | 3,84 | 0,92 | 1,89 | 1,30 | 0,00 |
| Napoli II | 2551-112 | 2,14 | 3,43 | 3,86 | 0,57 | 1,82 | 1,29 | 0,09 |
| Novara | 2323-235 | 0,73 | 0,64 | 2,11 | 1,98 | 2,34 | 3,01 | 2,31 |
| Nuoro | 1421-115 | 0,21 | 1,33 | 1,71 | 0,15 | 0,53 | 0,53 | 1,08 |
| Oristano | 1421-115 | 0,07 | 1,37 | 0,39 | 0,79 | 0,64 | 0,64 | 1,84 |
| Padova | 4134-441 | 2,29 | 0,93 | 1,10 | 2,26 | 2,38 | 2,27 | 0,64 |
| Palermo | 3332-322 | 1,74 | 2,07 | 1,59 | 0,85 | 2,14 | 1,29 | 0,87 |
| Parma | 2114-333 | 1,12 | 0,62 | 1,11 | 2,72 | 2,69 | 2,51 | 1,82 |
| Pavia | 2114-333 | 1,15 | 1,32 | 0,81 | 1,97 | 1,28 | 2,10 | 1,89 |
| Perugia | 2323-235 | 1,63 | 1,86 | 1,04 | 1,55 | 1,85 | 2,08 | 1,01 |
| Pesaro-Urbino | 2323-235 | 0,86 | 1,80 | 0,37 | 1,65 | 1,67 | 1,59 | 1,01 |
| Pescara | 2343-133 | 0,55 | 1,67 | 3,26 | 1,12 | 1,34 | 2,09 | 1,10 |
| Piacenza | 2114-333 | 0,65 | 0,80 | 0,57 | 2,50 | 2,21 | 1,95 | 2,08 |
| Pisa | 2114-333 | 0,87 | 1,14 | 0,92 | 2,03 | 1,44 | 2,00 | 1,11 |
| Pistoia | 2343-133 | 0,56 | 2,16 | 1,55 | 1,95 | 0,30 | 2,24 | 0,88 |
| Pordenone | 2114-333 | 0,65 | 0,89 | 0,08 | 1,87 | 2,73 | 1,93 | 1,09 |
| Potenza | 1421-115 | 0,49 | 2,30 | 0,95 | 0,15 | 1,66 | 0,30 | 2,31 |
| Prato | 2343-133 | 0,59 | 2,04 | 1,77 | 2,24 | 0,49 | 1,64 | 0,01 |
| Ragusa | 2551-112 | 0,48 | 2,82 | 2,25 | 0,92 | 0,48 | 1,00 | 0,85 |

| Direzione Provinciale | Cluster | Numerosità bacino | Pericolosità fiscale | Pericolosità sociale | Tenore di vita | Struttura produttiva | Tecnologia e servizi | Infrastrutture di trasporto |
|---|---|---|---|---|---|---|---|---|
| Ravenna | 2114-333 | 0,96 | 0,52 | 1,06 | 2,42 | 1,95 | 2,34 | 1,53 |
| Reggio Calabria | 2551-112 | 0,65 | 2,70 | 5,00 | 0,59 | 1,44 | 0,74 | 1,39 |
| Reggio Emilia | 2114-333 | 1,30 | 0,88 | 0,86 | 2,48 | 2,19 | 2,04 | 0,47 |
| Rieti | 1421-115 | 0,15 | 2,52 | 0,53 | 1,33 | 0,76 | 0,91 | 3,19 |
| Rimini | 2343-133 | 0,79 | 1,38 | 1,57 | 2,22 | 1,25 | 3,29 | 0,47 |
| Roma I | 5445-551 | 4,81 | 2,81 | 1,68 | 3,59 | 3,51 | 3,66 | 0,36 |
| Roma II | 5445-551 | 3,03 | 2,82 | 1,69 | 3,35 | 3,64 | 3,82 | 0,19 |
| Roma III | 5445-551 | 2,38 | 2,86 | 1,68 | 2,50 | 3,08 | 3,65 | 0,78 |
| Rovigo | 2114-333 | 0,43 | 1,05 | 0,71 | 1,64 | 1,07 | 1,24 | 1,48 |
| Salerno | 2551-112 | 1,83 | 3,61 | 1,66 | 0,46 | 0,91 | 0,79 | 1,13 |
| Sassari | 3332-322 | 0,76 | 1,40 | 1,43 | 1,81 | 1,82 | 2,25 | 1,38 |
| Savona | 2343-133 | 0,60 | 0,80 | 2,37 | 2,64 | 0,63 | 2,42 | 2,49 |
| Siena | 2114-333 | 0,63 | 0,61 | 0,73 | 2,52 | 2,89 | 3,12 | 3,02 |
| Siracusa | 3332-322 | 0,47 | 2,46 | 1,30 | 0,88 | 1,81 | 1,14 | 0,87 |
| Sondrio | 2114-333 | 0,25 | 0,27 | 1,22 | 2,11 | 2,36 | 3,49 | 1,10 |
| Taranto | 3332-322 | 0,76 | 2,02 | 1,50 | 1,03 | 1,65 | 0,91 | 0,65 |
| Teramo | 2323-235 | 0,56 | 2,21 | 1,56 | 0,98 | 0,89 | 1,20 | 2,04 |
| Terni | 2323-235 | 0,41 | 1,53 | 1,10 | 1,48 | 1,42 | 1,54 | 2,43 |
| Torino I | 4134-441 | 2,70 | 0,55 | 2,12 | 2,57 | 3,04 | 3,60 | 0,55 |
| Torino II | 4134-441 | 2,81 | 0,57 | 2,14 | 1,85 | 2,67 | 3,46 | 0,96 |
| Trapani | 2551-112 | 0,62 | 3,11 | 1,72 | 0,68 | 1,17 | 0,90 | 1,94 |
| Trento | 2114-333 | 1,25 | 0,11 | 0,43 | 2,42 | 2,78 | 3,51 | 1,40 |
| Treviso | 4134-441 | 2,21 | 0,92 | 0,00 | 1,91 | 2,40 | 1,80 | 0,66 |
| Trieste | 4134-441 | 0,39 | 0,69 | 0,98 | 3,31 | 3,36 | 2,97 | 0,65 |
| Udine | 2114-333 | 1,21 | 0,42 | 0,35 | 1,91 | 2,03 | 2,29 | 2,22 |
| Varese | 4134-441 | 1,88 | 1,20 | 1,19 | 2,01 | 2,14 | 3,11 | 0,29 |
| Venezia | 4134-441 | 1,93 | 0,66 | 1,11 | 2,18 | 2,39 | 3,04 | 0,75 |
| Verbano-Cusio-Ossola | 2323-235 | 0,19 | 1,46 | 0,93 | 1,67 | 1,57 | 2,05 | 1,89 |
| Vercelli | 2323-235 | 0,26 | 0,84 | 1,68 | 1,95 | 1,37 | 2,22 | 4,29 |
| Verona | 4134-441 | 2,30 | 0,76 | 1,10 | 2,24 | 2,62 | 2,95 | 0,91 |
| Vibo Valentia | 2551-112 | 0,05 | 3,71 | 3,72 | 0,06 | 0,42 | 0,22 | 2,55 |
| Vicenza | 4134-441 | 2,11 | 0,82 | 0,28 | 1,91 | 2,54 | 2,22 | 0,59 |
| Viterbo | 2323-235 | 0,57 | 2,14 | 1,30 | 1,24 | 1,01 | 0,97 | 2,01 |

## Riferimenti bibliografici


Bolasco S. (2002), "Analisi Multidimensionale dei Dati", Carocci, Roma.

Braiotta A., Pisani S., Pisano E. (2012) "Evasione fiscale e distribuzione primaria del reddito" in Rivista di diritto finanziario e scienza delle finanze, n.2.

Bracchi G., Motta G. (2004), " Le strategie di Customer Relationship Management", Sviluppo & Organizzazione N. 202.

Brown S. (2000), "Customer Relationship Management: Linking People, Process, and Technology", John Wiley & Sons, New York.

Calinski, T. and Harabasz, J. (1974), "A Dendrite Method for Cluster Analysis," Communications in Statistics, 3, 1–27.

Carbone E., Gentili F., Marino O., Spingola A. (2010), "Analisi Georeferenziata delle Direzioni Provinciali", Documento Interno Agenzia delle Entrate, Roma.

Carbone E., Gentili F., Marino O., Spingola A. (2010), "Analisi Georeferenziata delle Direzioni Provinciali: Applicazioni", Documento Interno Agenzia delle Entrate, Roma.

Charnes, A., W.W. Cooper e E. Rhodes (1981), "Measuring the Efficiency of Decision Making Units", European Journal of Operational Research.

Cherchye, L., Moesen, W., Rogge, N., van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., Tarantola, S. (2008), "Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index", Journal of the Operational Research Society, 59, 239-251.

Colapinto C. (2005), "Gestire il cliente con il CRM e la Business Intelligence: tra storia e dati", Business Intelligence CRM e Data Mining, Utet, Torino.

D'Agosto E., Marigliani M., Pisani S. (2013), "Asymmetries in the territorial VAT gap", articolo presentato alla conferenza The Shadow Economy, Tax Evasion and Governance, 25-28 luglio, Munster, Germania.

Decastri M., De Marco M., Rajola F. (2001), "Processi decisionali e flessibilità: il Customer Relationship Management in Banca", 2° Workshop dei Docenti e Ricercatori di Organizzazione Aziendale, Università degli Studi di Padova.

Duda, R.O. and Hart, P.E. (1973), Pattern Classification and Scene Analysis, New York: John Wiley & Sons, Inc.

Hunter D., Shine S. (2001), "Customer Relationship Management: A Blueprint for Government", The Government Executive Series, Accenture.

Mascitti M., Pisani S., Scialpi G.S. (2004), "Un'Ipotesi di Classificazione degli Uffici Locali tramite l'Applicazione di Criteri di Marketing", Documento Interno Agenzia delle Entrate, Roma.

Mazziotta C., Mazziotta M., Pareto A., Vidoli F. (2010), "La sintesi di indicatori territoriali di dotazione infrastrutturale: metodi di costruzione e procedure di ponderazione a confronto", Rivista di Economia e Statistica del Territorio N.1 gennaio-aprile 2010, Franco Angeli.

Mazziotta C., Pareto A., Talucci V. (2010), "La costruzione di indicatori di disuguaglianza sociale: il caso delle regioni italiane", XXXI Conferenza italiana di scienze regionali.

Mazziotta M., Pareto A. (2007), "Un indicatore sintetico di dotazione infrastrutturale: il metodo delle penalità per coefficiente di variazione", in Lo sviluppo regionale nell'Unione Europea - Obiettivi, strategie, politiche, Atti della XXVIII Conferenza Italiana di Scienze Regionali, Bolzano.

Nardo M., M. Saisana, A. Saltelli and S. Tarantola (EC/JRC), A. Hoffman and E. Giovannini (OECD) (2005), "Handbook on Constructing Composite Indicators: Methodology and Userguide", OECD, Statistics Working Paper.

Okazaki S. (2006), "What do we know about mobile Internet adopters? A cluster analysis", Information & Management, Volume 43, Issue 2.

Rinaldi A., Vidoli F., Zelli R. (2008), "Costruzione di indicatori di competitività a livello territoriale", AISLO, Napoli.

Sarle, W.S. (1983), Cubic Clustering Criterion, SAS Technical Report A-108, Cary, NC: SAS Institute Inc.

Schael (2004), Report Benchmarking 2004 – Strategie di CRM e trend del customer care in Italia, Customer Management Forum, irso, Milano.

Schellong (2007), "Citizen Relationship Management", Encyclopedia of Digital Government (Volume I), Hershey (USA).

Spingola A. (2012), "The Optimal Allocation of Human Resources using Multivariate Data Analysis", Stochastic Modeling Techniques and Data Analysis International Conference, Chania (Crete).

Yeo D. (2003), "Applied Clustering Techniques Course Notes", Sas Institute Inc., Cary (Usa).

## Siti internet consultati

http://www.aci.it/

http://www.bancaditalia.it/

http://www.istat.it/

http://www.terna.it

http://www.isprambiente.gov.it

http://www.infocamere.it/movimprese

http://www.ivass.it