

\\ 489 \\

**Analisi cluster gerarchica delle imprese
metalmeccaniche della Provincia di Modena**

Sandra Paterlini¹
Elena Pirani²
Margherita Russo³

Aprile 2005

¹ Università degli Studi di Modena e Reggio Emilia
Dipartimento di Economia Politica
Via Berengario,51
41100 Modena
e-mail: paterlini.sandra@unimore.it

²⁻³ Università degli Studi di Modena e Reggio Emilia
Dipartimento di Scienze Sociali, Cognitive e Quantitative
Via Giglioli Valle,9
42100 Reggio Emilia
e-mail: pirani.elena@unimore.it
e-mail: russo.margherita@unimore.it

1. Obiettivo dell'analisi
2. I dati: selezione del campione e dei caratteri da considerare nell'analisi
 - 2.1 La scelta dei caratteri oggetto dell'analisi
 - 2.2 La ricodifica delle variabili qualitative
 - 2.3 I gruppi di variabili e il campione delle imprese
3. L'analisi cluster: una sintesi della metodologia
 - 3.1 Il concetto di similarità e misure di distanza
 - 3.2 I metodi Gerarchici
 - 3.3 La rappresentazione grafica del risultato: il dendrogramma
 - 3.4 La determinazione del numero di cluster
4. Le fasi dell'analisi
5. I risultati
6. Conclusioni

1. Obiettivo dell'analisi

In che misura la partizione di una popolazione di imprese definita sulla base di ipotesi teoriche o osservazioni empiriche può essere irrobustita da un'analisi cluster della struttura dei dati?

Nel 2001 attraverso l'indagine empirica "Metalnet – Struttura e cambiamento delle relazioni tra le imprese metalmeccaniche della provincia di Modena"¹, è stato intervistato un campione rappresentativo delle imprese metalmeccaniche modenesi, che ha consentito la raccolta di numerose informazioni relative a questo sistema produttivo. La popolazione di imprese oggetto di studio nell'indagine "Metalnet" è quella definita nel data base Unimec² costituito da 3.900 imprese metalmeccaniche (Ateco91 da 27 a 35), attive nella provincia di Modena nel 1998 con almeno una unità produttiva, per un'occupazione totale di 49.381 addetti stimati al dicembre 1997. Da tale popolazione è stato estratto e intervistato un campione di imprese³ con più di cinque addetti, a cui è stato somministrato un questionario di 95 domande aperte e chiuse che ci consentono di descrivere in modo dettagliato aspetti salienti dei caratteri strutturali e relazionali della popolazione oggetto di indagine.

Nelle elaborazioni dei dati svolte nella prima fase dell'indagine, commentate in Pirani Russo (2001, 2003), è stata utilizzata una partizione del campione in imprese "conto proprio" e imprese "conto terzi". Tale partizione coglie un primo aspetto delle caratteristiche relazionali tra le imprese. In particolare, indichiamo imprese "in conto proprio" le imprese in cui l'attività è svolta in autonomia (anche se il disegno può essere parzialmente o anche completamente definito dal cliente); indichiamo, invece, con l'espressione imprese "in conto terzi" quelle imprese in cui l'attività produttiva è realizzata su specifiche del cliente e con progettazione dell'impresa o del cliente. Ad ogni impresa è stato quindi attribuito questo nuovo carattere confrontando il fatturato realizzato nei due tipi di attività.

Obiettivo dell'analisi che viene presentata in questo paper è definire una tipologia di imprese che - sulla base di caratteri strutturali e relazionali - ci consenta di validare la descrizione della popolazione di imprese metalmeccaniche della provincia di Modena innanzitutto rispetto alla partizione conto proprio/conto terzi. Un'altra variabile su cui normalmente si analizza la struttura di un settore produttivo è la distribuzione dei caratteri della popolazione di imprese rispetto alla dimensione in termini di addetti e di fatturato. Tale dimensione è però determinata da fattori quali il livello di integrazione verticale dell'impresa e la tecnologia in uso, che a sua volta dipende dalle caratteristiche del prodotto in termini di lunghezza della serie, numero di serie (nell'unità di tempo e cumulate nel tempo)⁴. In una prima fase della analisi abbiamo osservato in che misura i fenomeni analizzati si distribuivano per classe di dimensione, con la consapevolezza che le differenze dimensionali nascondessero altri fenomeni più difficilmente rappresentabili in modo sintetico.

Un altro obiettivo nell'analisi della struttura delle imprese metalmeccaniche è anche quello di valutare l'importanza delle relazioni tra imprese nelle fasi di produzione e commercializzazione (con fornitori, clienti, concorrenti) e nelle collaborazioni di progettazione e ricerca e sviluppo. Per quel che riguarda le imprese conto terzi il questionario conteneva delle domande che consentono di evidenziare le particolari relazioni di collaborazione con i loro committenti. Una prima elaborazione aveva consentito di evidenziare 8 tipi di imprese che, rispetto alla collaborazione nella progettazione e nella ricerca e sviluppo, sembrano evidenziare differenze nelle caratteristiche strutturali (quali la dimensione e il livello di integrazione verticale, l'intensità della concorrenza, i mercati di sbocco dei prodotti) e relazionali (durata delle relazioni con i committenti, localizzazione e dimensione dei committenti).

Per cogliere l'effettiva importanza di questi fenomeni abbiamo avviato un'analisi cluster i cui risultati vengono presentati in questo paper. Pur nel loro carattere preliminare, questa prima serie di elaborazioni ci consente di evidenziare quali variabili utilizzare nell'analisi, quali tecniche di analisi cluster siano le più appropriate e quali problemi restino ancora da affrontare

per completare la trama interpretativa della tipologia di imprese, a partire dalla struttura dei dati rilevati nell'indagine campionaria.

Il paper presenta innanzitutto i dati utilizzati nella analisi (sezione 2), viene poi descritta l'analisi cluster (sezione 3) e specificate le fasi dell'analisi (sezione 4). I principali risultati sono commentati nella sezione 5. Il paper si chiude con alcune considerazioni sulla efficacia della analisi effettuata e su quali direzioni sviluppare per ulteriori approfondimenti che rendano efficace la analisi cluster dei fenomeni indagati.

2. I dati: selezione del campione e dei caratteri da considerare nell'analisi

2.1 La scelta dei caratteri oggetto dell'analisi

Attraverso l'indagine "Metalnet – Struttura e cambiamento delle relazioni tra le imprese metalmeccaniche della provincia di Modena", è stato intervistato un campione rappresentativo delle imprese metalmeccaniche modenesi, che ha consentito la raccolta di numerose informazioni relative a questo sistema produttivo. Il questionario utilizzato in tale indagine consente una descrizione quantitativa e qualitativa della dinamica relativa alle seguenti variabili: il prodotto, il processo produttivo, il livello di integrazione verticale, la dimensione in termini di addetti, la composizione della forza lavoro, il fatturato (interno ed estero), il ricorso alla subfornitura, il grado di dipendenza dai subfornitori e dai committenti, le caratteristiche della concorrenza, i mercati di approvvigionamento degli input e dell'output, la composizione delle vendite.

Il questionario, composto da 95 domande, ha generato oltre 800 variabili, trattate fino ad ora con il programma di elaborazione statistica Spss, al fine di individuare una descrizione dell'universo delle imprese metalmeccaniche modenesi. Fra i dati raccolti, sono presenti sia variabili di carattere qualitativo – quelle rappresentative di caratteristiche relazionali – sia variabili quantitative – sostanzialmente quelle relative alle caratteristiche strutturali.

Per quanto riguarda questo ultimo aspetto, poiché la letteratura sulla cluster analysis propone diversi metodi alternativi da utilizzare nel caso in cui si stiano studiando variabili quantitative o variabili qualitative (vedi paragrafo 3.1), o entrambe, si è deciso di considerare entrambi i tipi di variabile, scegliendo una adeguata misura di distanza.

Dal punto di vista delle variabili da utilizzare nell'analisi, poiché nell'algoritmo di analisi cluster ogni variabile ha lo stesso peso, si è ritenuto opportuno effettuare un'attenta selezione delle variabili oggetto di studio, per evitare di inserire un elevato numero di variabili il cui contenuto informativo non è molto significativo. Dopo una prima scrematura che ha portato all'eliminazione di alcune variabili (sostanzialmente quelle relative alle domande di controllo inserite nel questionario), le variabili da utilizzare nell'algoritmo di analisi cluster sono state selezionate nel modo seguente:

a. Innanzitutto, nell'ottica di privilegiare le variabili relazionali, o che esprimono più direttamente caratteristiche relazionali, si è scelto di non inserire nell'analisi cluster alcune delle variabili rilevate (ICT, prodotti, macchinari utilizzati dall'impresa, componenti o lavorazioni acquistati, e variabili relative agli investimenti e all'espansione dell'impresa).

b. In secondo luogo, sono state escluse le variabili che, in base alle precedenti analisi dei dati, non presentavano sostanziali differenziazioni di risposta e che quindi, non fornendo elementi di diversificazione fra le imprese, non avrebbero aggiunto informazioni utili all'analisi.

c. In alcuni casi più variabili sono state "riassunte" in un'unica variabile (ad esempio, nel caso degli acquisti dai primi cinque fornitori strategici, anziché utilizzare cinque diverse variabili, è stata costruita una variabile sintetica che rappresenta un dato medio).

d. Infine, una annotazione particolare nel caso della domanda relativa all'integrazione verticale delle imprese. In questa domanda le imprese intervistate dovevano dichiarare, nel caso in cui svolgessero l'attività indicata, in che percentuale tale attività era svolta all'interno,

all'esterno o da altre imprese del gruppo. Poiché fra le tre modalità di risposta proposte la terza è risultata essere marginale, si è ritenuto opportuno considerare solamente le prime due; di conseguenza, questa variabile è stata assimilata ad una variabile le cui risposte possono essere raggruppate in due aggregati. A questo punto, si è potuto inserire nell'analisi cluster solamente il primo aggregato, cioè la percentuale di attività svolta all'interno, poiché la parte restante, la percentuale di attività svolta all'esterno, rappresentava semplicemente la parte restante.

Gli interventi sopra descritti hanno pertanto permesso di selezionare 33 macro-variabili, per un totale di 105 variabili, fra cui 58 numeriche, e 47 qualitative.

2.2 La ricodifica delle variabili qualitative

Per quanto riguarda le variabili qualitative sono stati effettuati alcuni interventi di ricodifica, al fine di ridurre le modalità di risposta e rendere maggiormente omogenei i dati. Nel dataset iniziale erano presenti variabili qualitative di tre tipi:

- variabili qualitative dicotomiche (sì/no)
- variabili qualitative suddivise in classi su scala ordinale, come il grado di concorrenza, l'importanza dei fattori di vantaggio competitivo, la collaborazione nella progettazione o alla R&S, le cui risposte erano ripartite su una scala ordinale da 1 a 5 che esprimeva l'intensità del fenomeno considerato,
- variabili che esprimevano caratteri qualitativi sconnessi. Queste variabili derivavano dalle risposte date a domande aperte come la localizzazione dei fornitori e dei clienti.

Mentre per le variabili dicotomiche non è stata effettuata alcuna ricodifica, per quanto riguarda le variabili qualitative ordinali, è evidente come un elevato numero di classi determini una dispersione dei dati, rendendo l'analisi meno intelligibile. Si è perciò deciso di procedere ad una riclassificazione dei dati per ridurre le modalità di risposta. Per queste variabili, che esprimevano l'intensità di un fenomeno, sono perciò state aggregate le modalità di risposta "alta" e "molto alta", o "bassa" e "media"⁵.

Nel caso delle domande sulla localizzazione dei clienti e dei fornitori, nel questionario si era lasciata alle imprese la massima libertà nell'indicazione della risposta, senza ricondurla a priori ad aggregati geografici predefiniti. Per consentire l'elaborazione dei dati, le risposte fornite dalle imprese sono state quindi codificate, aggregandole in classi. Anche in questo caso si è scelto di creare solamente tre aggregati geografici (Emilia-Romagna, Italia, Altri Paesi), sia per evitare una eccessiva dispersione dei dati, sia perché dalle analisi preliminari sui dati non sembrava rilevante utilizzare una disaggregazione maggiore. Per quanto riguarda invece le variabili relative alla ripartizione geografica del fatturato, si è ritenuto opportuno mantenere la differenziazione delle risposte inizialmente prevista nel questionario, poiché una aggregazione avrebbe fatto perdere informazioni sostanziali sui mercati di vendita e sulle differenze fra i subfornitori locali e le imprese maggiormente orientate su altri mercati.

Un ulteriore intervento preliminare sul dataset è stato effettuato con riferimento alla differenziazione delle domande previste per le imprese in conto proprio e per le imprese in conto terzi. Infatti, per la struttura stessa del questionario utilizzato per la raccolta delle informazioni, i campi da compilare riservati a imprese in conto proprio e in conto terzi erano differenti (nonostante le domande fossero le medesime, si mirava a tenere separati i due tipi di risposta). Nella fase di preparazione dei dati da elaborare attraverso l'analisi cluster, si è perciò provveduto ad integrare le due informazioni creando un'unica variabile. Per quelle imprese che, producendo sia in conto proprio che in conto terzi, hanno compilato entrambe le parti del questionario, si è privilegiata la risposta data relativamente al tipo di attività svolta in modo prevalente.

Infine, per le variabili che erano presenti solamente per il conto terzi (in particolare le domande relative alle "collaborazioni con il committente"), per le imprese conto proprio è stato inserito il valore NaN (Not a Number = non disponibile), al fine di poter comunque utilizzare queste variabili.

2.3 I gruppi di variabili e il campione delle imprese

Le variabili così selezionate e ricodificate, sono state raggruppate in insiemi omogenei, rispecchiando la struttura originaria del questionario che procedeva per temi di indagine. I gruppi di variabili, utilizzati per l'applicazione dell'algoritmo di cluster analysis sono caratteristiche strutturali, fornitori, clienti, concorrenti, collaborazioni, CT-collaborazioni, e all data che considera le variabili dei primi 5 raggruppamenti. Si noti che l'ultimo gruppo di variabili riguarda solamente le imprese che svolgono la loro attività prevalentemente in conto terzi; questo gruppo è infatti stato creato per confrontare risultati della cluster analysis con gli otto gruppi di imprese individuati in Russo e Pirani (2003) relativamente all'intensità della collaborazione nella progettazione e alla R&S delle imprese conto terzi con i loro committenti. L'Appendice 1 presenta l'elenco delle variabili utilizzate, suddivise nei "gruppi di variabili" definiti, e in base ai quali è stata svolta l'analisi cluster: accanto ad ogni variabile considerata è indicato il tipo di variabile, e per ognuna di esse viene indicato il valore minimo, medio e massimo assunto dalla variabile stessa nel campione, o l'ampiezza delle classi e le relative percentuali di risposte, nel caso di variabili qualitative.

Come detto, nell'indagine Metalnet è stato intervistato un campione statisticamente rappresentativo dell'universo delle imprese metalmeccaniche della provincia di Modena. Tuttavia, si è ritenuto opportuno escludere alcune di queste imprese dal dataset utilizzato per lo svolgimento dell'esercizio di cluster analysis, in quanto un certo numero di imprese presentavano molti dati mancanti. Si è quindi scelto di ridurre la numerosità del campione, preferendo un insieme di imprese che presentasse i dati per tutti i gruppi di variabili considerate. I risultati dell'analisi cluster fanno riferimento perciò a 134 imprese, le quali presentano dati completi per le variabili utilizzate nell'analisi.

3. L'analisi cluster: una sintesi della metodologia

L'utilizzo della analisi cluster consente di evidenziare la possibile presenza di similarità e struttura nei dati, attraverso l'identificazione di gruppi omogenei di oggetti.

L'attività tipica di raggruppamento dei dati si sviluppa in genere attraverso i seguenti passi (Jain et al. 1999):

1) Rappresentazione dei dati

Questa fase è diretta alla determinazione, in genere ad opera del ricercatore, del numero di possibili raggruppamenti e del numero, tipo e scala delle misurazioni da utilizzare nel partizionare i dati. In questo stadio si inserisce, quando presente, il processo di selezione (*feature selection*) dal dataset originale delle misurazioni più efficaci nel raggruppare i dati e di trasformazione dei caratteri di input al fine di valorizzare possibili pattern nascosti (*feature extraction*).

2) Prossimità dei dati

In questa fase viene definita una misura di similarità, generalmente sulla base di una metrica di distanza, fra gli oggetti in base alla quale procedere alla determinazione della composizione dei gruppi. La definizione della misura di similarità dipende dalla natura delle misurazioni caratterizzanti i differenti oggetti: qualitative o quantitative.

3) Clustering or Grouping

In questa fase si definisce quale algoritmo utilizzare per procedere alla determinazione della composizione dei gruppi. Gli algoritmi presenti in letteratura sono in genere suddivisi in due grandi classi: *partitivi o non gerarchici*, in cui gli oggetti sono raggruppati in modo tale da identificare la partizione che ottimizza, in genere in senso locale, un dato criterio di raggruppamento, e *gerarchici*, che producono raggruppamenti successivi ordinabili secondo livelli crescenti o decrescenti sulla base della misura di similarità.

4) Data Abstraction

Tale fase porta ad estrarre una semplice e compatta rappresentazione dell'insieme dei dati di partenza. Un esempio tipico di astrazione dei dati è la descrizione compatta di ogni cluster per mezzo di prototipi o misurazioni rappresentative quali possono essere i centroidi dei gruppi.

5) Validazione dell'output

In questa fase si vuole validare la partizione dei dati identificata. Tale analisi è condotta in genere grazie all'utilizzo di un criterio specifico di ottimalità. Una struttura di raggruppamento è valida nel caso in cui non sia identificata by chance o grazie ad opportuni artifici nell'algoritmo di raggruppamento. Nel caso in cui siano utilizzati approcci statistici per risolvere il problema di clustering, tale processo di validazione può assumere tre forme: un controllo esterno della validità che paragona la struttura identificata rispetto ad una struttura nota a priori, un controllo interno che cerca di determinare se la struttura sia appropriata dal punto di vista intrinseco rispetto ai dati ed infine un controllo di tipo relativo che paragona le due strutture e misura i loro meriti relativi.

I metodi di analisi cluster sono in genere distinti in:

- *gerarchici*, che producono raggruppamenti successivi ordinabili secondo livelli crescenti o decrescenti della distanza.
- *non gerarchici o partitivi*, che determinano un'unica suddivisione del dataset, considerata ottimale rispetto ad un dato criterio statistico.

3.1 Il concetto di similarità e misure di distanza

Al fine di svolgere una analisi cluster di un dataset è necessario definire una misura di prossimità o di similarità fra gli oggetti. Un coefficiente di similarità misura la relazione fra due oggetti considerando congiuntamente le differenti misurazioni ad essi associati. La determinazione di tale coefficiente dipende quindi dalla natura qualitativa o quantitativa delle misurazioni.

Variabili Quantitative

Sia X una matrice di misurazioni riferite a n oggetti di cui si considerano k caratteristiche, X_i il vettore k -dimensionale dell' i -esima osservazione e x_{ih} il suo elemento generico. Sia Σ^{-1} l'inversa della matrice di varianze e covarianze campionarie. Le metriche più comunemente utilizzate sono:

Distanza city-block o di Manhattan

$$d_{ij} = \sum |x_{ih} - x_{jh}|$$

Distanza euclidea

$$d_{ij} = \left[\sum (x_{ih} - x_{jh})^2 \right]^{\frac{1}{2}}$$

Distanza di Minkowsky

$$d_{ij} = \left[\sum (x_{ih} - x_{jh})^r \right]^{\frac{1}{r}}$$

Distanza di Mahalanobis

$$d_{ij} = (X_i - X_j)' \Sigma^{-1} (X_i - X_j)$$

È noto come la distanza euclidea possa essere influenzata in modo negativo dalla possibile differente scala delle k misurazioni. Una possibile soluzione a tale problema potrebbe allora essere l'utilizzo di una procedura di normalizzazione dei dati.

Una metrica di largo utilizzo è inoltre la metrica City-block, che risente in misura inferiore rispetto alla distanza euclidea degli scostamenti fra i caratteri associati alla differenti osservazioni.

La distanza Euclidea e la distanza City-block costituiscono un caso speciale della metrica di Minkowski. Tale metrica funziona bene nel caso in cui il dataset presenti cluster compatti o isolati, ma non dovrebbe essere utilizzata nel caso in cui le k misurazioni contenute nella matrice dei dati abbiano valori medi e deviazioni standard largamente differenti. In tal caso è consigliato normalizzare i dati in modo opportuno o utilizzare altre metriche ad hoc.

Una possibile fonte di distorsione nelle misure di distanza è data dalla possibile presenza di correlazione lineare fra i k caratteri. Tale distorsione è di norma evitata attraverso un processo di whitening dei dati o utilizzando la metrica di Mahalanobis. La distanza di Mahalanobis associa pesi differenti alle differenti misurazioni sulla base delle varianze e delle correlazioni lineari a coppie.

Variabili Qualitative

Nel caso in cui le misurazioni corrispondano a variabili qualitative una misura comunemente utilizzata è la distanza di Hamming, che corrisponde alla proporzione delle variabili che risultano discordanti.

$$d_{ij} = \left(\frac{\#(x_{ih} \neq x_{jh})}{n} \right)$$

$\#(\cdot)$ indica il numero di misurazioni discordanti

Variabili Miste

Nel caso in cui le misurazioni siano di natura differente (qualitative e quantitative), una possibile misura di distanza che le consideri congiuntamente può essere calcolata utilizzando l'indice di Gower (1971). Tale indice è definito come:

$$d_{ij} = 1 - \frac{\sum_{h=1}^k d_{ijh} s_{ijh}}{\sum_{h=1}^k d_{ijh}}$$

dove se la h -esima variabile è quantitativa ed $R(h)$ è il suo campo di variazione, si ha

$$s_{ijh} = 1 - \frac{|x_{ih} - x_{jh}|}{R(h)}$$

mentre se è qualitativa

$$s_{ijh} = \begin{cases} 1 & \text{se la } h\text{-esima variabile ha la stessa modalità per le osservazioni } i\text{-esima e } j\text{-esima} \\ 0 & \text{altrimenti} \end{cases}$$

ed in generale

$$d_{ijh} = \begin{cases} 1 & \text{se si conoscono i valori dell' } h\text{-esima variabile per le osservazioni } i\text{-esima e } j\text{-esima} \\ & \text{(serve quando vi sono dati mancanti)} \\ 0 & \text{nel caso contrario,} \\ & \text{oppure in caso di accordo 0/0 per variabili binarie di tipo presenza/assenza} \end{cases}$$

3.2 I metodi Gerarchici

I metodi gerarchici raggruppano in modo sequenziale gli oggetti presenti nel dataset sulla base di un'opportuna misura di similarità, in modo tale che gruppi determinati ad un livello di distanza inferiore siano contenuti in quelli determinati in corrispondenza di livelli superiori della distanza. Essi possono quindi essere ulteriormente suddivisi in:

- metodi agglomerativi: si considera inizialmente un numero di gruppi corrispondenti al numero di oggetti presenti nel dataset e si uniscono sequenzialmente i gruppi progressivamente formati fino alla determinazione di un singolo gruppo;
- metodi scissori: il processo di raggruppamento avviene in senso contrario rispetto a quello dei metodi agglomerativi; si parte da un singolo gruppo, composto dall'intero dataset, che sarà progressivamente partizionato in sottogruppi iterativamente individuati.

Un algoritmo agglomerativo prevede in genere i seguenti step:

- 1) Si calcola la matrice di prossimità contenente le distanze fra ogni coppia di oggetti. Ogni oggetto è considerato come un gruppo a sé stante.
- 2) Sulla base dell'informazione contenuta nella matrice di prossimità, le coppie di cluster che presentano un grado maggiore di similarità sono rilevate e unite in un solo gruppo. Si aggiorna la matrice di prossimità considerando tale operazione di unione.
- 3) Nel caso in cui tutti gli oggetti siano contenuti in un unico cluster, si ha l'interruzione dell'algoritmo. In caso contrario, la procedura iterativa riparte dallo step 2.

Differenti algoritmi agglomerativi possono essere implementati grazie all'utilizzo di differenti modalità di calcolo della misura di similarità e di aggiornamento della matrice di prossimità allo step 2. Possono infatti essere utilizzate differenti metriche per la determinazione e aggiornamento della matrice di prossimità.

Se indichiamo con n_r il numero di oggetti nel cluster r e n_s il numero di oggetti nel cluster s e X_{r_i} l' i -esimo oggetto nel cluster r , possiamo formalizzare i differenti metodi gerarchici agglomerativi, che si distinguono fra loro sulla base del metodo di legame (*linkage*) che utilizzano. Il metodo di legame specifica il meccanismo di aggiornamento della matrice di prossimità ad ogni iterazione.

Tra i metodi gerarchici agglomerativi i più noti sono:

METODO DEL LEGAME SINGOLO: ogni oggetto è considerato inizialmente come un gruppo a sé stante. Si costruisce la matrice di prossimità contenente le distanze di tutte le possibili coppie di oggetti nel dataset. Si costruisce una lista in ordine ascendente di tali distanze e gli oggetti che sono a distanza minima gli uni rispetto gli altri sono assegnati ad un unico gruppo. Si calcola quindi la distanza fra tale gruppo appena formato e i rimanenti oggetti come la minima distanza tra tutti gli oggetti del gruppo e gli altri oggetti. Si ripete la procedura iterativamente fino ad arrivare alla formazione di una singola partizione.

Si utilizza quindi la distanza minima fra gli oggetti nei due gruppi:

$$d(r,s)=\min(\text{dist}(X_{r_i},X_{s_j})), i ? (1,\dots,n_r), j ? (1,\dots,n_s)$$

METODO DEL LEGAME COMPLETO: tale algoritmo differisce da quello a "legame singolo" per le modalità di calcolo della distanza fra il gruppo appena formato e ognuno dei rimanenti oggetti o gruppi. In questo caso non si considera la distanza minima, ma quella massima. Il metodo del "legame completo" consente di identificare cluster più compatti rispetto a quello del "legame singolo", ma fallisce, ad esempio, nell'identificazione di cluster concentrici. Jain e Dubes (1988) hanno osservato come l'algoritmo "a legame completo" identifichi gerarchie più significative rispetto a quello a legame singolo in molti campi di applicazioni.

$$d(r,s)=\max(\text{dist}(X_{r_i},X_{s_i})), i ? (1,\dots,n_r), j ? (1,\dots,n_s)$$

METODO DEL LEGAME MEDIO: tale algoritmo differisce dai precedenti per le modalità di calcolo della distanza fra il gruppo appena formato e ognuno dei rimanenti oggetti o gruppi. Tale distanza è calcolata come la distanza tra un oggetto fittizio, avente come caratteri ad esso associati una media dei caratteri degli oggetti contenuti nel gruppo, e i rimanenti oggetti.

$$d(r, s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} dist(X_{ri}, X_{sj})$$

Altri metodi gerarchici sono quelli della mediana, della varianza minima, del legame flessibile e numerosi altri che differiscono fra loro per il meccanismo di determinazione della misura di similarità fra gli oggetti.

3.3 La rappresentazione grafica del risultato: il dendrogramma

Il raggruppamento determinato da un algoritmo gerarchico è solitamente visualizzato attraverso un dendrogramma, e quindi di un grafico a due dimensioni (n.oggetti vs misura di similarità) che schematizza i raggruppamenti identificati in corrispondenza di differenti livelli di similarità.

Ipotizziamo che l'algoritmo gerarchico abbia determinato la ripartizione di 5 oggetti secondo quanto riportato in figura 1, dove gli oggetti 1 e 3 e 5 e 4 hanno dissimilarità minima e sono raggruppati a costituire gli oggetti fittizi 6 e 7 i quali hanno similarità maggiore fra loro rispetto all'oggetto 2 e quindi costituiscono un nuovo oggetto fittizio 8.

In un dendrogramma, l'asse orizzontale rappresenta gli indici degli oggetti nel dataset originale. I collegamenti fra gli oggetti sono ottenuti con linee a forma di U capovolta. L'altezza delle U indica la distanza fra gli oggetti. Ad esempio, il legame che rappresenta il cluster che contiene gli oggetti 1 e 3 ha altezza pari a 1.

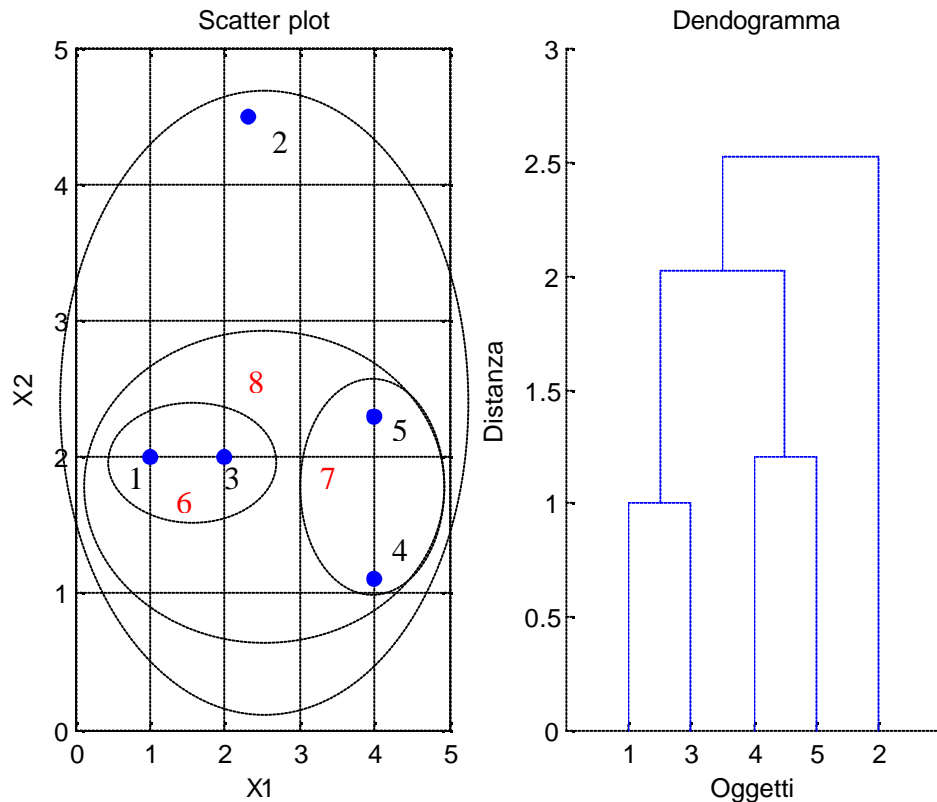


Figura 1: Raggruppamento gerarchico di 5 oggetti. Sulla sinistra, scatterplot degli oggetti nello spazio delle variabili X_1 e X_2 caratterizzanti gli oggetti. Sulla destra, dendrogramma del raggruppamento gerarchico.

3.4 La determinazione del numero di cluster

Definita la misura di similarità e il metodo del legame da utilizzare, la rappresentazione grafica più comunemente utilizzata è quella del dendrogramma, che consente di raggruppare gli oggetti in un albero gerarchico. La fase finale dell'analisi è in genere diretta a dividere gli oggetti in differenti gruppi cercando di individuare un raggruppamento naturale degli oggetti, utilizzando ad esempio un criterio basato sul coefficiente di correlazione cofenetic, o determinando arbitrariamente a che livello tagliare l'albero gerarchico.

La figura sottostante mostra come sia possibile determinare due e tre raggruppamenti fissando il livello di cutoff rispettivamente in corrispondenza del valore 2,3 e 1,7. In genere si utilizza un livello di cutoff pari al 70% del valore massimo di dissimilarità.

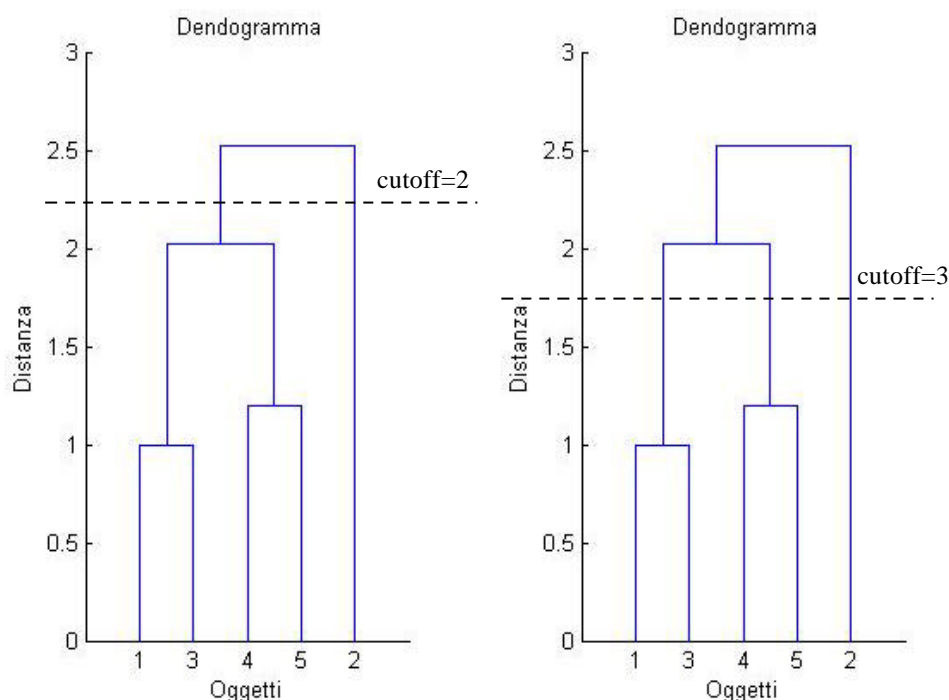


Figura 2: Determinazione del numero di cluster utilizzando tecniche di analisi cluster gerarchica. Il livello a cui si esegue il taglio ("cutoff") determina il numero di cluster in cui ripartire gli oggetti.

4. Le fasi dell'analisi

Come evidenziato nel paragrafo 3, gli algoritmi di analisi cluster possono essere distinti in due macro-classi: algoritmi di analisi cluster non gerarchica o partitivi, dove gli oggetti sono direttamente allocati in cluster e algoritmi gerarchici, che producono raggruppamenti successivi ordinabili secondo livelli crescenti o decrescenti sulla base della misura di similarità.

La nostra attenzione si è concentrata su algoritmi di analisi cluster gerarchica, in quanto interessati al raggruppamento sequenziale delle imprese e ad evidenziare il più possibile anche le singole relazioni fra imprese. Tale scelta ha richiesto la definizione del tipo di algoritmo utilizzare (se agglomerativo o scissorio), della metrica di distanza, del tipo di legame, della rappresentazione grafica e del criterio da utilizzare per l'identificazione dei cluster (vedi paragrafo 3.2).

Data la necessità di considerare congiuntamente variabili sia quantitative che qualitative, si è deciso di utilizzare l'indice di Gower (1981) quale metrica di distanza.

Indagini preliminari in cui abbiamo testato differenti metodi di legame (singolo, medio e completo) hanno supportato la scelta del metodo di legame "completo" (vedi paragrafo 3.2 per

la descrizione). Il metodo a legame “singolo” e quello a legame “completo” sono quelli più comunemente utilizzati. Quest’ultimo consente di identificare cluster più compatti (Baeza-Yates 1992) rispetto a quelli identificati dal metodo a legame singolo, che tende a produrre cluster più allungati. Jain e Dubes (1988) hanno inoltre osservato come il metodo a legame “completo” consenta di produrre gerarchie più utili rispetto al metodo a legame “singolo” in svariate applicazioni reali.

Il dendrogramma è stato quindi utilizzato per la rappresentazione grafica dei risultati ottenuti dalla cluster analysis. Indagini future saranno dirette ad esaminare possibili rappresentazioni alternative, quale l’utilizzo delle cosiddette mappe ad albero (Johnson, B., Shneiderman 1991) e il minimum spanning tree (Sedgewick, 1984).

Il livello di cutoff è stato fissato al 70% del valore massimo della dissimilarità. Tuttavia, seppur questo livello sia quello comunemente utilizzato, esso è per sua natura arbitrario. Si è pertanto deciso di determinare e salvare tutte le possibili ripartizioni in un numero di cluster da 2 a n , dove n corrisponde al numero di imprese considerate.

L’analisi cluster è stata svolta in ambiente MATLAB, utilizzando funzioni della Toolbox Statistics e implementando un programma ad hoc per il calcolo dell’indice di Gower e la rappresentazione grafica dei cluster.

5. I risultati

Per ogni gruppo di variabili usate nelle diverse elaborazioni cluster *all data*, *caratteristiche strutturali*, *fornitori*, *clienti*, *concorrenti*, *collaborazioni*, *CT-collaborazioni*, sono riportate Tabelle 2-8 (Appendice 3) i dati relativi alle variabili usate nella analisi cluster per ciascuno dei cluster individuati con un cut off del 70% e, in alcuni casi, inferiore.

Le tabelle riportano inoltre i valori di altre variabili non usate nell’analisi cluster, ma secondo noi rilevanti per interpretare i risultati, come ad esempio le informazioni che specificano se le imprese sono conto proprio o conto terzi, in base alla nostra classificazione. I valori evidenziati sono quelli che mostrano un valore significativamente più elevato o inferiore al valore medio.

Consideriamo innanzitutto i risultati della analisi cluster *alldata*. Con un cut off del 70% sono individuati due cluster, il primo dei quali, alldata1, è costituito da solo 3 grandissime imprese. Abbiamo quindi applicato un cut off del 50% che facesse emergere almeno un’altra suddivisione dell’altro cluster alldata2, in cui sono presenti rispettivamente 32 e 99 imprese. Alldata2a raggruppa imprese in prevalenza in contro proprio, di dimensione media relativamente maggiore in termini di addetti, alldata2b raggruppa invece imprese in prevalenza conto terzi, di dimensione media relativamente minore in termini di addetti.

Troviamo da questa analisi una prima conferma che le somiglianze tra le caratteristiche strutturali e relazionali delle imprese sono largamente colte dalla distinzione del tipo di impresa adottato nella nostra precedente analisi.

Cerchiamo ora di cogliere quali sono le variabili che determinano quel risultato. Questo può essere fatto osservando i risultati della analisi cluster sui sottoinsiemi di variabili che compongono l’insieme di variabili di *alldata*: una prima considerazione riguarda il grado di somiglianza che è assai difforme nelle cinque analisi cluster effettuate. Esaminiamo quindi i risultati di ciascuna di queste analisi cluster. Nell’Appendice 2 sono riportati i dendrogrammi relativi ai differenti gruppi di variabili considerate. Colori differenti evidenziano i cluster individuati quando si consideri un livello di cutoff pari al 70% della distanza massima. Si osservi come la struttura dei dendrogrammi differisca notevolmente fra i differenti gruppi di variabili considerate.

Caratteristiche strutturali: in questa cluster analysis il cutoff del 70% evidenzia nuovamente i due cluster iniziali di *alldata*, con un cutoff del 35% sono evidenziati 4 cluster, due dei quali risultano dalla scomposizione del cluster carstr2 di appena 3 imprese; gli altri due cluster

mostrano una disomogenea numerosità che ci fa propendere per un cutoff ancora minore, fino a individuare almeno un altro cluster significativo. Diventa necessario scomporre i carstr1b in almeno due cluster (carstr1b1 e carstr1b2), ma questo disaggrega anche il cluster carstr1a. Si osservi in figura 4, la forma assunta dal dendrogramma. Le tre grandi imprese risultano nettamente distinte dalle altre imprese del campione, a livelli elevati di distanza (asse y). Le altre imprese del campione risultano essere molto simili fra loro: le altezze dei rami del dendrogramma sono infatti, fatta eccezione per le tre grandi imprese, inferiori a 0.4.

Fornitori: In questa analisi cluster il gruppo di variabili coglie le caratteristiche dei fornitori (numero totale di fornitori, numero e localizzazione dei fornitori strategici) e delle relazioni con essi in termini di collaborazioni nella progettazione e nella ricerca e sviluppo. Con un cut off del 70% si evidenziano cinque cluster, forn5 è composto da 8 imprese di dimensione molto superiore alla media. Osservando la Figura 5, si osservi come i due gruppi sulla destra e il secondo gruppo da sinistra risultano essere compatti e distinti fra loro. La prevalenza di variabile qualitative non porta a separare le tre grandissime imprese, come nel caso precedente.

Clienti: tre cluster: due di imprese prevalentemente conto proprio e uno di imprese conto terzi. La Figura 6 evidenzia la presenza di due cluster fortemente distinti fra loro e compatti al loro interno: la lunghezza dei rami è infatti nella maggior parte dei legami inferiori al livello di 0.6.

Concorrenti e Collaborazioni: oltre dieci cluster. La Figura 7 evidenzia una forte eterogeneità fra le imprese del campione (lunghezza dei rami elevata) e la presenza di una molteplicità di cluster individuati nel caso in cui si considerino i dati del gruppo concorrenti. Tale eterogeneità si manifesta anche in relazione ai dati del gruppo collaborazioni, sebbene in misura inferiore (vedi Figura 8)

CT-collaborazioni: tre cluster da confrontare con gli otto gruppi individuati in Russo Pirani (2003). La Figura 9 evidenzia la presenza di quattro cluster compatti con rami aventi lunghezza inferiore al 50%.

Poiché nella analisi cluster ogni variabile ha lo stesso peso, i risultati sono fortemente influenzati da variabili che - su uno specifico fenomeno - siano particolarmente disaggregate. Dai risultati delle elaborazioni emerge che è necessario procedere con una ulteriore selezione delle variabili usate nella analisi cluster, e in alcuni casi costruire nuove variabili sintetiche.

L'elenco di variabili sarà modificato nel seguente modo.

- La composizione professionale degli addetti pesa per ben 10 variabili, ne useremo invece solo due: il numero totale degli addetti e la quota di impiegati.
- Per quel che riguarda l'integrazione verticale, selezioneremo tra le attuali 19 variabili solo quelle che corrispondono alle fasi che mostrano una maggiore variabilità tra le imprese, sia nella quota utilizzata di quella fase sia per quel che riguarda la quota realizzata all'interno dell'impresa, che è la variabile usata nella analisi cluster. Adottando questo criterio, le fasi si riducono alle seguenti nove: Progettazione e design; Costruzione stampi e modelli; Lavorazioni di carpenteria metallica; Lavorazioni su macchine utensili; Trattamenti chimici e termici; Trattamenti superficiali; Verniciatura; Assemblaggio parti meccaniche; Assemblaggio componenti elettriche e elettroniche.
- Poiché le risposte mancanti per il 3°, 4° e 5° fornitore superano in media il 30% dei casi, l'analisi cluster sarà condotta solo sui primi due fornitori.
- Per quel che riguarda la percentuale di acquisti dai primi cinque fornitori strategici, anziché utilizzare una variabile sintetica che rappresenta un dato medio, sarà utilizzata la somma delle quote dei primi 3 fornitori.
- Poiché tra i clienti è assai modesta la presenza di imprese/agenti commerciali, consumatori, committenti pubblici, si ridurranno le variabili che rappresentano la

composizione percentuale delle vendite per tipo di clienti a tre: grandi imprese; medie/piccole imprese; piccole imprese o artigiani

- Poiché tra le 13 variabili considerate ve ne sono 11 che rappresentano l'importanza delle caratteristiche di vantaggio competitivo (relative alle tre classi di risposte: "nullo", "basso- medio" e "alto-molto alto"), proponiamo di selezionare solo i seguenti fattori che hanno una importanza alta o molto alta e mostrano una significativa differenza tra le imprese (riducendo così ulteriormente il peso di questa variabile qualitativa attualmente su tre classi a una variabile quantitativa): migliore capacità di progettazione; maggiore velocità di esecuzione degli ordini; migliore qualità del prodotti; maggiore gamma dei prodotti; migliore capacità di promozione delle vendite; migliore qualità del servizio post-vendita.

Seguendo tali criteri di selezione delle variabili, sarà pertanto definito un nuovo dataset, che risulterà essere composto ancora da 6 insiemi che raggruppano 58 macro-variabili, per un totale di 71 variabili (contro le 105 fino ad ora utilizzate).

6. Conclusioni

Le analisi svolte offrono interessanti spunti di ricerca per ulteriori indagini. L'analisi di differenti gruppi di dati ha portato ad evidenziare strutture di cluster differenti in relazione all'insieme di caratteristiche considerate. Imprese simili rispetto alla natura della clientela possono essere estremamente eterogenee per caratteristiche strutturali. Le ripartizioni individuate e le strutture differenti dei dendrogrammi evidenziano questo fenomeno di eterogeneità fra imprese in relazione alle differenti caratteristiche individuate. L'analisi ha tuttavia consentito di evidenziare la presenza di cluster compatti rispetto al gruppo di variabili fornitori, clienti, CT-collaborazioni e caratteristiche strutturali. L'analisi dei gruppi concorrenti e collaborazioni evidenzia la presenza di un numero elevato di cluster con elevata dissimilarità all'interno dei cluster.

Come emerge chiaramente dalle considerazioni presentate in questo paper, l'analisi cluster richiede un lungo lavoro di selezione delle variabili essenziali per rappresentare i fenomeni oggetto di indagine. Questo working paper presenta lo stadio intermedio dell'analisi che ci consente di focalizzare meglio la metodologia di analisi cluster e l'ulteriore selezione delle variabili su cui calcolare gli indici di distanza.

Sulla base della nuova selezione di variabili, sarà ridefinito l'insieme di osservazioni per le quali sono disponibili i dati e sarà nuovamente effettuata l'analisi cluster sui singoli sotto insiemi di variabili (caratteristiche strutturali, fornitori, clienti, concorrenti, collaborazioni) e su alldata. Con la nuova selezione di variabili, l'analisi cluster relativa ai sottoinsiemi di variabili sarà condotta senza variabili ripetute. La presenza di variabili ripetute non consente infatti di focalizzare quali siano gli insiemi di variabili che generano coesione o eterogeneità tra le osservazioni.

La ricerca sarà inoltre diretta ad analizzare rappresentazioni alternative al fine di evidenziare cluster compatti e di evidenziare la struttura relazionale fra imprese, utilizzando legami basati su opportune metriche di distanze. Una prima analisi sarà diretta alla rappresentazione grafica dei cluster utilizzando non solo dendrogrammi, ma strutture spaziali di reti grazie all'elaborazione con software Pajek⁶. Inoltre, si valuterà se l'utilizzo di strumenti alternativi, quali il Minimum Spanning Tree (Sedgewick, 1984), consentano di isolare cluster compatti ed evidenziare imprese rappresentative dei differenti cluster.

Bibliografia

- Batagelj V., W. de Noor e A. Mrvar (2005) *Exploratory Social Network Analysis with Pajek*, Cambridge Univ. Press, Cambridge.
- Baeza-Yates R. A. (1992) Introduction to data structures and algorithms related to information retrieval. In *Information Retrieval: Data Structures and Algorithms*, W. B. Frakes and R. Baeza-Yates, Eds. Prentice-Hall.
- Brusco S. e Bigarelli X(1995) "Struttura industriale e fabbisogni formativi nei settori della maglieria e delle confezioni in Italia. Un'analisi per regione, 1993", *Rivista Italiana di Economia*, n. 0, 1995, pp. 7-47.
- Jain A. K. e Dubes, R. C. (1988) *Algorithms for Clustering Data*. Prentice-Hall advanced reference series. Prentice-Hall, Inc., Upper Saddle River, NJ.
- Jain, A.K., Murty M.N., and Flynn P.J. (1999) "Data Clustering: A Review", *ACM Computing Surveys*, Vol 31, No. 3, 264-323.
- Johnson, B., Shneiderman, B. (1991)"Treemaps: a space-filling approach to the visualization of hierarchical information structures" *Proc. of the 2nd International IEEE Visualization Conference*.
- Kaufman, L. e Rousseeuw P.J., (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley.
- Lalla M. (2001), "Struttura e cambiamento nelle relazioni tra le imprese metalmeccaniche nella provincia di Modena, II. Distribuzioni degli addetti e pesi per le stime dei parametri", *Materiali di discussione*, n. 394, Dipartimento di Economia politica, Università degli Studi di Modena e Reggio Emilia.
- Pirani E. e M. Russo (2003), "Struttura e cambiamento nelle relazioni tra le imprese metalmeccaniche nella provincia di Modena, III. Aspetti metodologici dell'indagine empirica: fase di rilevazione, controlli e statistiche preliminari", *Materiali di discussione*, n. 396, Dipartimento di Economia politica, Università degli Studi di Modena e Reggio Emilia.
- Russo M. e R. Giardino (2000), "Struttura e cambiamento nelle relazioni tra le imprese metalmeccaniche nella provincia di Modena, I. La popolazione di imprese meccaniche della provincia di Modena: procedure impiegate per integrare le informazioni amministrative del Registro Imprese e dell'Inps", *Materiali di discussione*, n. 310, Dipartimento di Economia politica, Università degli Studi di Modena e Reggio Emilia.
- Russo M. e E. Pirani (2001), "Struttura e cambiamento nelle relazioni tra le imprese metalmeccaniche nella provincia di Modena, IV. Primi risultati dell'indagine empirica", *Materiali di discussione*, n. 396, Dipartimento di Economia politica, Università degli Studi di Modena e Reggio Emilia.
- Russo M. e E. Pirani (2003), "Competition and cooperation in a metal engineering production system", *Materiali di discussione*, n. 438, Dipartimento di Economia politica, Università degli Studi di Modena e Reggio Emilia.
- Sedgwick, R., (1984), *Algorithms*. Addison-Wesley, Reading.
- Silberston A. (1972), *Economies of Scale in Theory and Practice*, *Economic Journal*, pp. 369-91.

Appendice 1

Tabella 1 Schema delle variabili utilizzate per l'analisi cluster

Caratteristiche strutturali		tipo di variabile	classi delle variabili qualitative	variabili numeriche			variabili qualitative				
				valore minimo	valore medio	valore max	distribuzione di frequenza delle risposte (% di imprese)				
1	appartenenza a gruppi	qualitativa	{1,2,3}				si, a gruppo italiano	si, a gruppo estero	no	missing	
							11,9	3,0	85,1	0,0	
2	integrazione verticale: fasi svolte all'interno (val %)	numerica									
	Progettazione e design			0,0	72,9	100,0					
	Acquisto di materie prime e componenti			0,0	78,8	100,0					
	Costruzione stampi e modelli			0,0	35,6	100,0					
	Fonderia			0,0	18,6	100,0					
	Fusioni e sinterizzazioni			0,0	9,7	100,0					
	Deformazione a caldo			0,0	18,2	100,0					
	Deformazione a freddo			0,0	55,0	100,0					
	Stampaggio parti metalliche			0,0	37,7	100,0					
	Lavorazioni di carpenteria metallica			0,0	59,1	100,0					
	Lavorazioni su macchine utensili			0,0	68,4	100,0					
	Trattamenti chimici e termici			0,0	23,7	100,0					
	Trattamenti superficiali			0,0	27,3	100,0					
	Verniciatura			0,0	42,9	100,0					
	Assemblaggio parti meccaniche			0,0	89,6	100,0					
	Assemblaggio componenti elettriche e elettroniche			0,0	60,1	100,0					
	Montaggio finale			0,0	91,9	100,0					
	Controllo e collaudo			10,0	95,6	100,0					
	Assistenza tecnica ai clienti			0,0	94,5	100,0					
	Trasporto beni prodotti			0,0	37,5	100,0					
3	numero di addetti per composizione professionale	numerica									
	Totale operai			0	59	1909					
	Operai specializzati			0	18	536					
	Totale impiegati			0	23	697					
	Impiegati tecnici e di produzione			0	13	402					
	Impiegati commerciali			0	3	120					
	Impiegati amministrativi			0	5	137					
	Informatici, CED			0	1	38					
	Totale dirigenti			0	3	98					
	Totale addetti			6	85	2704					
	Titolare e soci			0	2	11					
4	ripartizione del fatturato per aree geografiche (val. %)	numerica									
	Provincia di Modena			0,0	38,0	100,0					
	Provincia Reggio Emilia			0,0	5,5	93,0					
	Resto Emilia Romagna			0,0	17,0	100,0					
	Altre regioni italiane			0,0	23,0	100,0					
	Altri paesi Unione Europea			0,0	8,8	76,0					
	Altri paesi			0,0	7,1	100,0					
5	% lavorazioni esterne sul totale acquisti	numerica		0,0	12,5	60,2					
6	valore aggiunto/addetto (mln di lire, 2000)	numerica		22,1	126,4	435,6					
7*	quota di fatturato nella produzione di pezzi unici	numerica		0,0	21,8	100,0					
Fornitori											
8	n. fornitori abituali	numerica		0	78	600					
9	n. fornitori strategici	numerica		0	8	100					
10	localizzazione fornitori (ricodif. in 3 classi)	qualitativa	{1,2,3}				Emilia-Romagna	Italia	Estero	missing	
	Primo fornitore						42,5	28,4	6,0	23,1	
	Secondo fornitore						42,5	24,6	4,5	28,4	
	Terzo fornitore						35,1	18,7	1,5	44,8	
	Quarto fornitore						28,4	15,7	3,0	53,0	
	Quinto fornitore						20,9	15,7	3,0	60,4	
11	durata rapporto con fornitori (anni)	numerica									
	Durata rapporto con primo fornitore			1	13	41					
	Durata rapporto con secondo fornitore			1	11	45					
	Durata rapporto con terzo fornitore			1	12	41					
	Durata rapporto con quarto fornitore			1	11	40					
	Durata rapporto con quinto fornitore			1	9	25					
12	% acquisti dai 5 fornitori strategici	numerica		2,5	31,5	100,0					
13	collaboraz. nella progettazione con i fornitori (ricodif. in 3 classi)	qualitativa	{1,2,3}				mai	salturiam.	sempre	missing	
	Primo fornitore						63,4	5,2	4,5	26,9	
	Secondo fornitore						59,7	3,7	6,0	30,6	
	Terzo fornitore						43,3	6,7	4,5	45,5	
	Quarto fornitore						37,3	6,7	3,0	53,0	
	Quinto fornitore						32,1	5,2	2,2	60,4	
14	collaborazione alla R&S con i fornitori (ricodif. in 3 classi)	qualitativa	{1,2,3}				mai	salturiam.	sempre	missing	
	Primo fornitore						47,0	18,7	8,2	26,1	
	Secondo fornitore						46,3	14,9	9,0	29,9	
	Terzo fornitore						35,8	15,7	3,7	44,8	
	Quarto fornitore						30,6	12,7	4,5	52,2	
	Quinto fornitore						27,6	10,4	2,2	59,7	

	tipo di variabile	classi delle variabili qualitative	variabili numeriche			variabili qualitative			
			valore minimo	valore medio	valore max	distribuzione di frequenza delle risposte (% di imprese)			
Clienti									
15*	percentuale di fatturato per tipi di clienti	numerica							
	grandi imprese		0,0	27,5	100,0				
	medie/piccole imprese		0,0	45,1	100,0				
	piccole imprese o artigiani		0,0	19,8	100,0				
	imprese/agenti commerciali		0,0	5,8	100,0				
	consumatori		0,0	2,7	100,0				
	committente pubblico		0,0	1,0	100,0				
16*	quota fatturato principali clienti	numerica							
	quota fatturato primo cliente		1,0	38,7	100,0				
	quota fatturato secondo cliente		0,0	14,8	50,0				
	quota fatturato terzo cliente		0,0	9,5	40,0				
17*	localizzazione principali clienti (ricodif. in 3 classi)	qualitativa {1,2,3}				Emilia-Romagna	Italia	Estero	missing
	localizzazione primo cliente					57,5	22,4	14,2	6,0
	localizzazione secondo cliente					53,0	20,1	13,4	13,4
	localizzazione terzo cliente					49,3	22,4	11,9	16,4
18*	durata rapporto principale cliente (anni)	numerica	0	12	35				
Concorrenti									
19*	grado di concorrenza (ricodif. in 3 classi)	qualitativa {1,2,3}				nullo	medio-basso	alto-molto alto	missing
						9,1	35,6	55,3	0,0
20*	n. concorrenti	numerica	0	12	220				
21*	importanza delle caratteristiche di vantaggio competitivo (ricodif. in 3 classi)	qualitativa {1,2,3}				nulla	medio-bassa	alta-molto alta	missing
	migliore capacità di progettazione					37,5	28,9	33,6	0,0
	migliore velocità di esecuzione degli ordini					9,3	41,9	48,8	0,0
	migliore qualità del prodotto					6,2	23,3	70,5	0,0
	migliore gamma dei prodotti					20,2	36,4	43,4	0,0
	migliore capacità di risposta a richieste particolari					4,6	18,5	76,9	0,0
	migliore capacità di promozione delle vendite					35,2	56,3	8,6	0,0
	migliore qualità del servizio post-vendita					28,9	38,3	32,8	0,0
	prezzi più bassi					24,0	60,5	15,5	0,0
	costi più bassi					26,4	57,4	16,3	0,0
	migliori condizioni di pagamento					22,7	61,7	15,6	0,0
	tradizione di affidabilità					3,1	14,7	82,2	0,0
22*	caratteristiche dei fornitori/clienti - concorrenti	qualitativa {1,2,3}				stesso settore e area	stesso settore ma diversa area	entrambi i casi	missing
						73,5	20,6	5,9	0,0
Collaborazioni									
23*	collabora alla progettazione con il committente (ricodif. in 3 classi)	qualitativa {1,2,3}				mai	salturiam.	sempre o quasi	missing
						30,4	34,2	35,4	0,0
24*	collabora alla R&S con il committente (ricodificata in 3 classi)	qualitativa {1,2,3}				mai	salturiam.	sempre o quasi	missing
						55,7	29,1	15,2	0,0
25	attivazione accordi con altre imprese	qualitativa {1,2,3,4}				solo con imprese italiane	solo con imprese straniere	sia con imprese italiane che	nessun accordo di collaborazi
						9,0	7,5	10,4	73,1
26	attivazione di accordi di R&S	dicotomica {0,1}				sì	no		missing
						22,4	77,6		0,0
13	collaboraz. nella progettazione con i fornitori (ricodif. in 3 classi)	qualitativa {1,2,3}				mai	salturiam.	sempre	missing
	Primo fornitore					63,4	5,2	4,5	26,9
	Secondo fornitore					59,7	3,7	6,0	30,6
	Terzo fornitore					43,3	6,7	4,5	45,5
	Quarto fornitore					37,3	6,7	3,0	53,0
	Quinto fornitore					32,1	5,2	2,2	60,4
14	collaborazione alla R&S con i fornitori (ricodif. in 3 classi)	qualitativa {1,2,3}				mai	salturiam.	sempre	missing
	Primo fornitore					47,0	18,7	8,2	26,1
	Secondo fornitore					46,3	14,9	9,0	29,9
	Terzo fornitore					35,8	15,7	3,7	44,8
	Quarto fornitore					30,6	12,7	4,5	52,2
	Quinto fornitore					27,6	10,4	2,2	59,7

	tipo di variabile	classi delle variabili qualitative	variabili numeriche			variabili qualitative				
			valore minimo	valore medio	valore max	distribuzione di frequenza delle risposte (% di imprese)				
23*	collabora alla progettazione con il committente (ricodif.in 3 classi)	qualitativa	{1,2,3}				mai	salturiam.	sempre o quasi	missing
							30,4	34,2	35,4	0,0
24*	collabora alla R&S con il committente (ricodificata in 3 classi)	qualitativa	{1,2,3}				mai	salturiam.	sempre o quasi	missing
							55,7	29,1	15,2	0,0
27	numero di addetti	numerica		6	29	123				
25	attivazione accordi con altre imprese	qualitativa	{1,2,3,4}				solo con imprese italiane	solo con imprese straniere	sia con imp. italiane che straniere	nessun accordo di collaborazi
							6,3	3,8	2,5	87,3
26	attivazione di accordi di R&S	dicotomica	{0,1}				si	no	missing	
							11,4	88,6	0,0	
28	rapporto di subfornitura	qualitativa	{1,2,3}				solo forn. 1 livello	solo forn. 2 livello	forn. 1 e 2 livello	missing
							60,8	11,4	25,3	2,5
19	grado di concorrenza	qualitativa	{1,2,3}				nullo	medio-basso	alto-molto alto	missing
							6,4	39,7	53,9	0,0
29	previsioni dinamica addetti	qualitativa	{1,2,3}				in crescita	stabile	in diminuz.	missing
							55,7	40,5	3,8	0,0
30	previsioni dinamica fatturato	qualitativa	{1,2,3}				in crescita	stabile	in diminuz.	missing
							81,0	16,5	2,5	0,0
4	ripartizione del fatturato per aree geografiche (val. %)	numerica								
	Provincia di Modena			0,0	48,4	100,0				
	Provincia Reggio Emilia			0,0	6,9	93,0				
	Resto Emilia Romagna			0,0	21,7	100,0				
	Altre regioni italiane			0,0	18,2	100,0				
	Altri paesi Unione Europea			0,0	3,0	50,0				
	Altri paesi			0,0	1,7	60,0				
7	quota di fatturato nella produzione di pezzi unici	numerica		0,0	20,4	100,0				
16	percentuale di fatturato per tipo di clienti	numerica								
	grandi imprese			0,0	24,2	100,0				
	medie/piccole imprese			0,0	50,6	100,0				
	piccole imprese o artigiani			0,0	23,3	100,0				
	imprese/agenti commerciali			0,0	0,4	20,0				
	consumatori			0,0	0,1	5,0				
	committente pubblico			0,0	0,0	0,0				
	altro			0,0	1,3	100,0				
18*	durata rapporto principale cliente (anni)	numerica		1	13	30				
31	modifiche ai prodotti negli ultimi due anni	dicotomica	{0,1}				si	no	missing	
							59,5	39,2	1,3	
32	tipo di modifiche	dicotomica	{0,1}				si	no	missing	
	introduzione di nuovi prodotti						35,4	64,6	0,0	
	innovazioni sostanziali ai prodotti esistenti						24,1	75,9	0,0	
	piccole modifiche ai prodotti esistenti						35,4	64,6	0,0	
33	iniziativa dell'introduzione di nuovi prodotti	dicotomica	{0,1}				si	no	missing	
	iniziativa e progetto propri						34,2	65,8	0,0	
	iniziativa propria, ma progettazione esterna						5,1	94,9	0,0	
	iniziativa del cliente, ma progettazione propria						26,6	73,4	0,0	
	iniziativa e progetto esterni						26,6	73,4	0,0	

* = variabili che erano separate per il conto proprio/conto terzi

Appendice 2

Rappresentazione grafica dei dendrogrammi in relazione ai differenti insiemi di dati considerati: *all data, Caratteristiche Strutturali, Fornitori, Clienti, Concorrenti, Collaboratori, CT- collaborazioni*

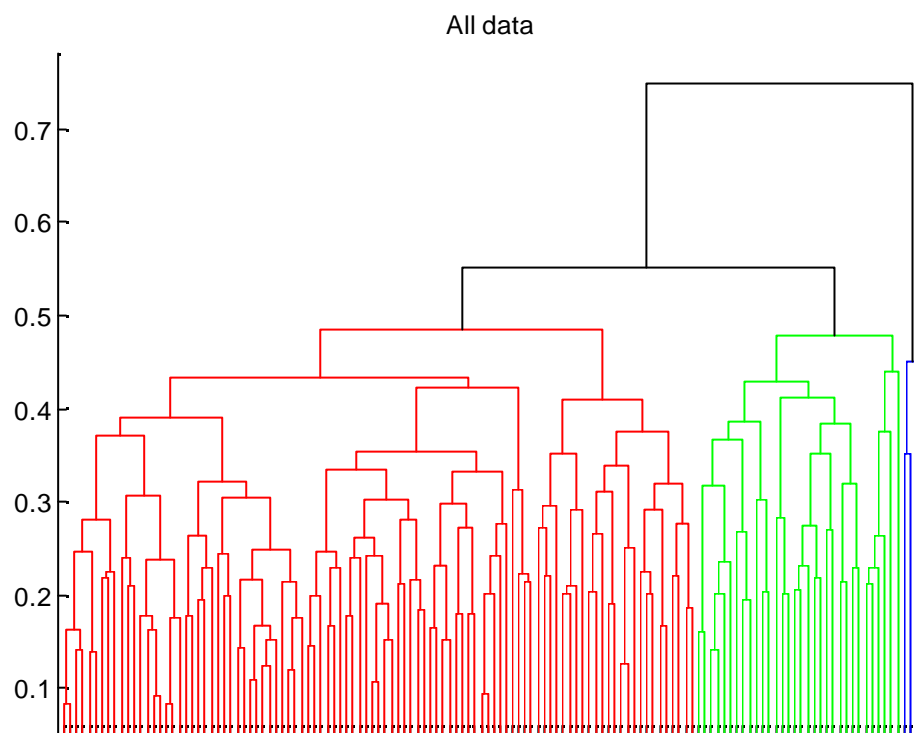


Figura 3: All data – Dendrogramma. Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

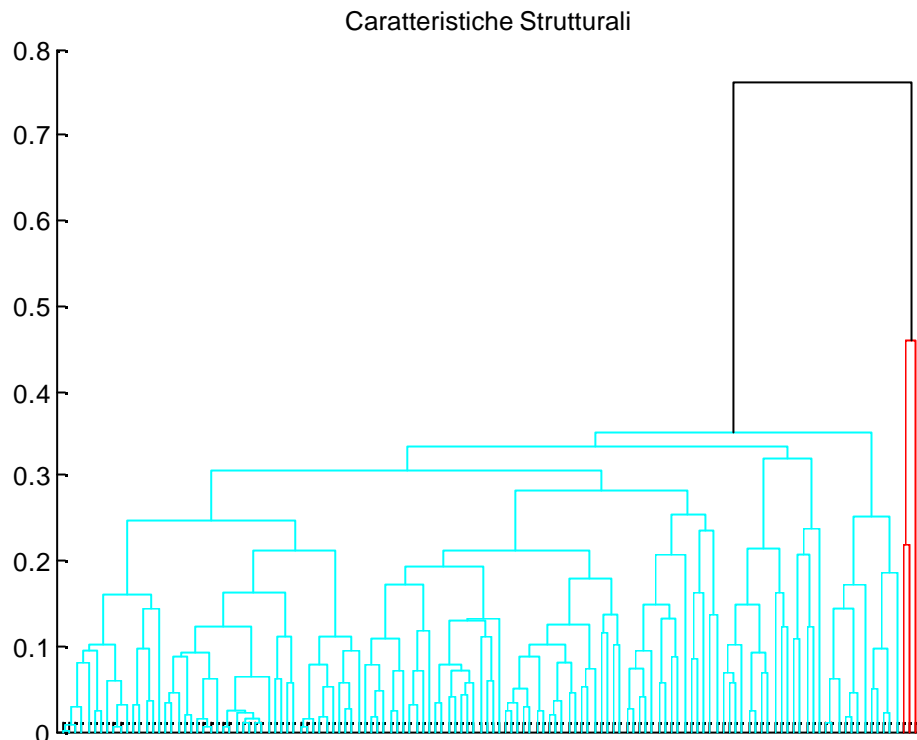


Figura 4: Caratteristiche Strutturali – Dendrogramma Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

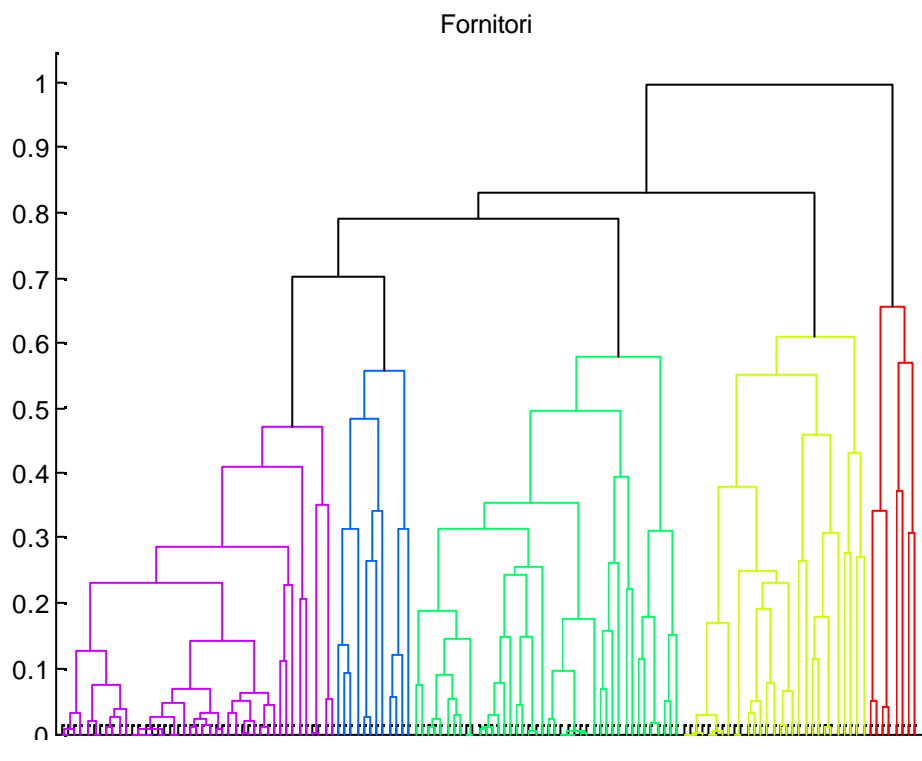


Figura 5: Fornitori – Dendrogramma Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

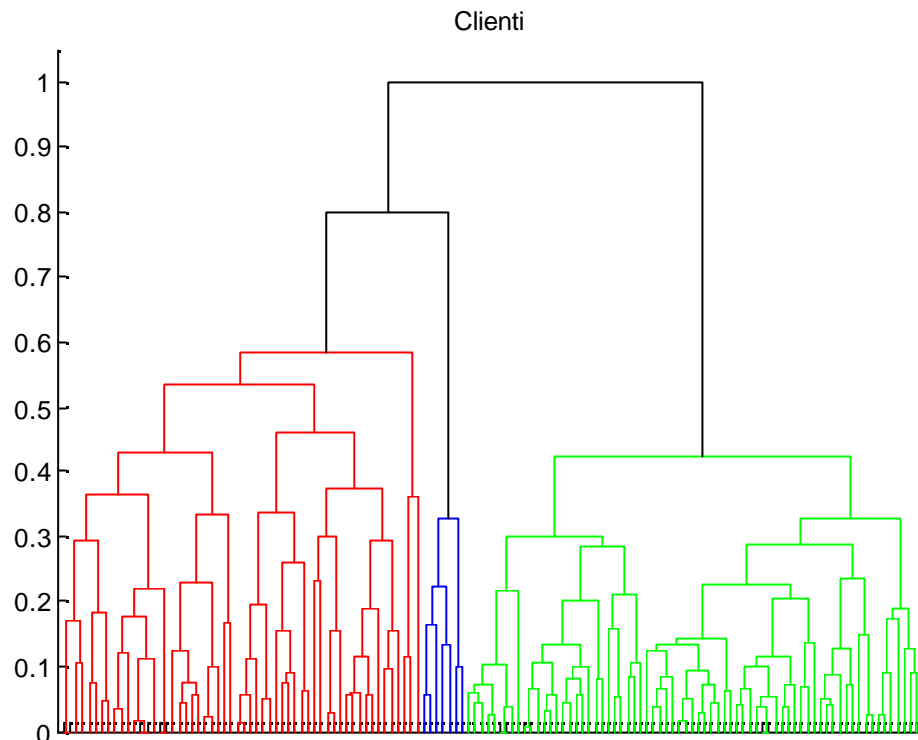


Figura 6: Clienti – Dendrogramma. Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

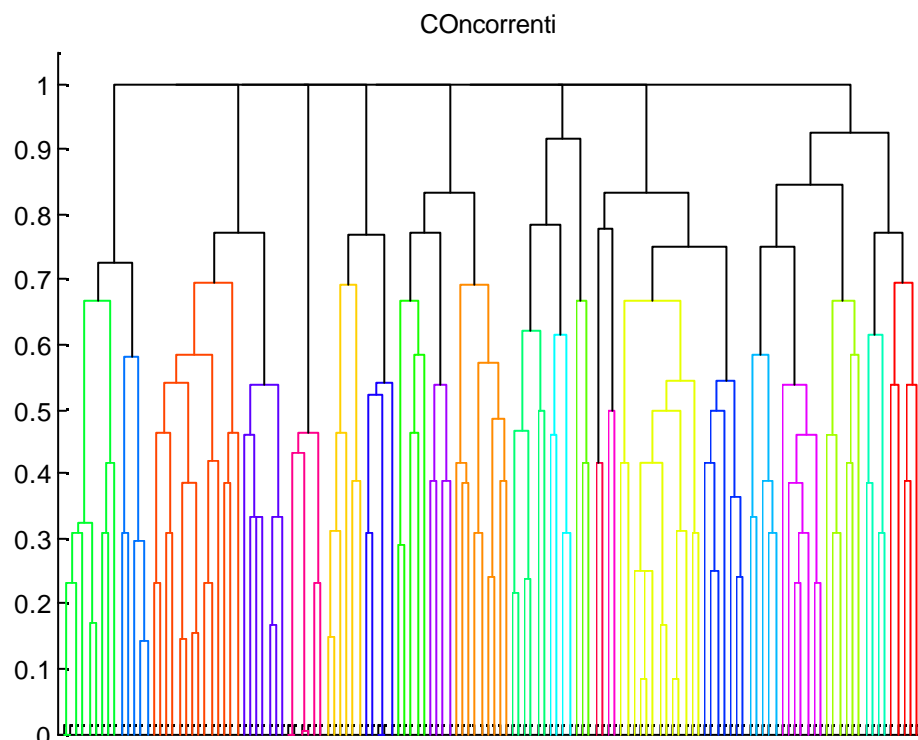


Figura 7: Concorrenti – Dendrogramma. Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

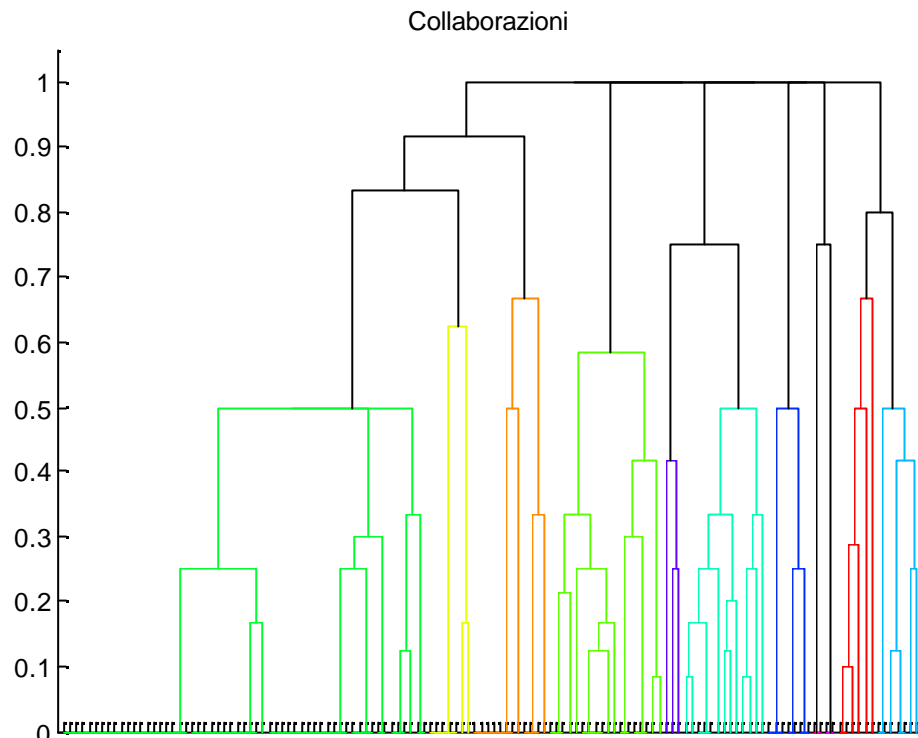


Figura 8: Collaborazioni – Dendrogramma Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo

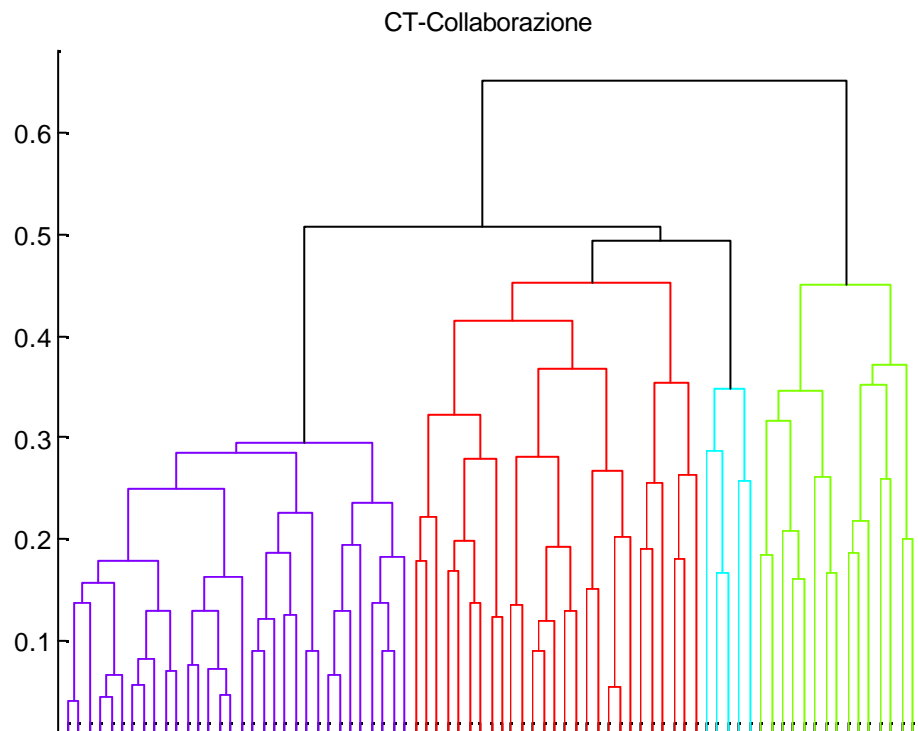


Figura 9: CT-Collaborazione – Dendrogramma Colori differenti evidenziano i cluster individuati nel caso in cui si consideri un cutoff a livello del 70% del valore massimo.

Appendice 3

Dati relativi a ciascuno dei cluster individuati

Tabella 2 *all data*
variabili: 1-28

	N. imprese	val. ass.	Cut off al 70% del livello di dissimilarità: 2		Cut off al 50% del livello di dissimilarità: 3 gruppi		
			1	2	alldata1	alldata2a	alldata2b
			3	131	3	32	99
	conto proprio/ conto terzi	Conto proprio	100,0	39,7	100,0	62,5	32,3
		Conto terzi		60,3		37,5	67,7
		6-9		6,9			9,1
		10-19		29,0		9,4	35,4
	classi di addetti	20-49		40,5		40,6	40,4
		50-99		11,5		18,8	9,1
		Oltre 99	100,0	12,2	100,0	31,3	6,1
		0-1 mld		5,6		3,1	6,5
	classi di fatturato	1-5 mld		44,8		21,9	52,7
		5-20 mld		35,2		37,5	34,4
		Oltre 20 mld	100,0	14,4	100,0	37,5	6,5
1	Appartenenza a un gruppo	Sì, a un gruppo italiano			66,7	18,8	8,1
		Sì, a un gruppo estero			33,3	9,4	
		No				71,9	91,9
2		Progettazione e design			76,7	78,8	69,2
		Acquisto di materie prime e componenti			66,7	81,9	78,1
		Costruzione stampi e modelli			46,7	9,8	49,3
		Fonderia			42,5	20,0	14,1
		Fusioni e sinterizzazioni			0,0	7,7	11,8
		Deformazione a caldo				0,0	28,6
		Deformazione a freddo			100,0	30,0	62,0
		Stampaggio parti metalliche			40,0	19,1	45,4
	Livello di integrazione verticale: fasi svolte all'interno	Lavorazioni di carpenteria metallica			35,0	36,2	68,4
		Lavorazioni su macchine utensili			85,0	47,5	75,6
		Trattamenti chimici e termici			45,0	10,5	29,3
		Trattamenti superficiali			50,0	20,0	29,5
		Verniciatura			95,0	41,8	41,1
		Assemblaggio parti meccaniche			100,0	91,9	88,1
		Assemblaggio componenti elettriche e elettroniche			85,0	53,7	62,6
		Montaggio finale			100,0	93,3	90,9
		Controllo e collaudo			100,0	98,2	94,6
		Assistenza tecnica ai clienti			100,0	94,8	94,2
		Trasporto beni prodotti			0,0	18,2	45,8
3		Totale operai			1074	60	27
		Operai specializzati			245	26	8
		Totale impiegati			354	37	9
	N. medio addetti per composizione professionale	Impiegati tecnici e di produzione			198	24	3
		Impiegati commerciali			52	4	1
		Impiegati amministrativi			87	6	2
		Informatici, CED			17	1	0
		Totale dirigenti			42	3	1
		Titolare e soci			0	2	2
3tot	Numero medio di addetti (val. ass.)		1469	53	1469	101	38
4		Provincia di Modena	1,5	38,6	1,5	20,0	44,8
		Provincia di Reggio Emilia		5,6		3,0	6,5
	ripartizione del fatturato per aree geografiche (val. %)	Resto dell'Emilia-Romagna	15,5	17,1	15,5	6,8	20,5
		Altre regioni italiane		23,3		34,2	19,7
		Altri paesi UE	28,5	8,4	28,5	20,1	4,5
		Altri paesi	27,5	6,7	27,5	16,0	3,7
5	% acquisti di lavorazioni esterne				5,8	12,1	12,8
6	Valore aggiunto per addetto (mln di lire)				148,6	123,3	127,3
7	Quota di fatturato nella produzione di pezzi unici			22,3		22,1	22,4
8	N. fornitori abituali	val. ass.	366,0	71,6	366	104	61
9	Numero fornitori strategici	val. ass.	30,7	7,9	31	7	8

N. imprese	val. ass.	Cut off al 70% del livello di dissimilarità: 2		Cut off al 50% del livello di dissimilarità: 3 gruppi			
		1	2	alldata1	alldata2a	alldata2b	
		3	131	3	32	99	
10_1°	localizzazione primo fornitore - ricodificata	Emilia-Romagna			44,0	61,3	
		Italia		100,0	36,0	34,7	
		Estero			20,0	4,0	
10_2°	localizzazione secondo fornitore - ricodificata	Emilia-Romagna		33,3	58,3	60,9	
		Italia		66,7	41,7	30,4	
		Estero				8,7	
10_3°	localizzazione terzo fornitore - ricodificata	Emilia-Romagna			77,8	62,3	
		Italia		100,0	22,2	34,0	
		Estero				3,8	
10_4°	localizzazione quarto fornitore - ricodificata	Emilia-Romagna			62,5	63,6	
		Italia		66,7	31,3	31,8	
		Estero		33,3	6,3	4,5	
10_5°	localizzazione quinto fornitore - ricodificata	Emilia-Romagna		100,0	21,4	61,1	
		Italia			64,3	33,3	
		Estero			14,3	5,6	
11	Durata del rapporto con i 5 principali fornitori strategici (anni)	primo fornitore			21	13	
		secondo fornitore			13	13	
		terzo fornitore			6	14	
		quarto fornitore			4	11	
		quinto fornitore			10	9	
12	% acquisti dai fornitori strategici			16,7	14,2	28,5	
13_1°	collaborazione progettazione primo fornitore	mai	33,3	88,4	33,3	75,0	93,0
		saltuar. o per alcuni comp/lavoraz.	33,3	6,3	33,3	12,5	4,2
		per gran parte/tutti i compon/lavoraz.	33,3	5,3	33,3	12,5	2,8
13_2°	collaborazione progettazione secondo fornitore	mai	33,3	87,8	33,3	73,9	92,5
		saltuar. o per alcuni comp/lavoraz.	66,7	3,3	66,7	4,3	3,0
		per gran parte/tutti i compon/lavoraz.		8,9		21,7	4,5
13_3°	collaborazione progettazione terzo fornitore	mai	66,7	80,0	66,7	50,0	90,4
		saltuar. o per alcuni comp/lavoraz.	33,3	11,4	33,3	33,3	3,8
		per gran parte/tutti i compon/lavoraz.		8,6		16,7	5,8
13_4°	collaborazione progettazione quarto fornitore	mai	100,0	78,3	100,0	50,0	88,6
		saltuar. o per alcuni comp/lavoraz.		15,0		43,8	4,5
		per gran parte/tutti i compon/lavoraz.		6,7		6,3	6,8
13_5°	collaborazione progettazione quinto fornitore	mai	33,3	84,0	33,3	57,1	94,4
		saltuar. o per alcuni comp/lavoraz.	33,3	12,0	33,3	28,6	5,6
		per gran parte/tutti i compon/lavoraz.	33,3	4,0	33,3	14,3	
14_1°	collaborazione R&S primo fornitore	mai		65,6		33,3	76,4
		saltuar. o per alcuni comp/lavoraz.	66,7	24,0	66,7	37,5	19,4
		per gran parte/tutti i compon/lavoraz.	33,3	10,4	33,3	29,2	4,2
14_2°	collaborazione R&S secondo fornitore	mai		68,1		43,5	76,5
		saltuar. o per alcuni comp/lavoraz.	100,0	18,7	100,0	26,1	16,2
		per gran parte/tutti i compon/lavoraz.		13,2		30,4	7,4
14_3°	collaborazione R&S terzo fornitore	mai	33,3	66,2	33,3	38,9	75,5
		saltuar. o per alcuni comp/lavoraz.	66,7	26,8	66,7	44,4	20,8
		per gran parte/tutti i compon/lavoraz.		7,0		16,7	3,8
14_4°	collaborazione R&S quarto fornitore	mai	66,7	63,9	66,7	25,0	77,8
		saltuar. o per alcuni comp/lavoraz.	33,3	26,2	33,3	56,3	15,6
		per gran parte/tutti i compon/lavoraz.		9,8		18,8	6,7
14_5°	collaborazione R&S quinto fornitore	mai		72,5		35,7	86,5
		saltuar. o per alcuni comp/lavoraz.	100,0	21,6	100,0	50,0	10,8
		per gran parte/tutti i compon/lavoraz.		5,9		14,3	2,7
15	percentuale di fatturato per tipi di clienti	grandi imprese			41,7	25,4	23,5
		medie/piccole imprese			6,7	28,1	51,9
		piccole imprese o artigiani			5,0	21,7	19,6
		imprese/agenti commerciali			13,3	14,5	2,6
		consumatori			33,3	3,5	1,5
		committente pubblico			0,0	3,2	0,3
16	quota di fatturato dei principali clienti	quota fatturato primo cliente	55,0	38,4	55,0	41,3	37,6
		quota fatturato secondo cliente	4,0	15,0	4,0	12,4	15,8
		quota fatturato terzo cliente	2,5	9,6	2,5	8,0	10,0

N. imprese	val. ass.	Cut off al 70% del livello di dissimilarità: 2		Cut off al 50% del livello di dissimilarità: 3 gruppi		
		1	2	alldata1	alldata2a	alldata2b
		3	131	3	32	99
17_1°	Emilia-Romagna		62,1		34,5	70,5
	localizzazione primo cliente		Italia		31,0	22,1
		100,0	13,7	100,0	34,5	7,4
17_2°	Emilia-Romagna		61,7		26,1	70,7
	localizzazione secondo cliente		Italia		26,1	22,8
		100,0	14,8	100,0	47,8	6,5
17_3°	Emilia-Romagna		59,5		36,4	65,2
	localizzazione terzo cliente		Italia		18,2	29,2
		100,0	13,5	100,0	45,5	5,6
18	durata relazione col principale cliente			14,0	12,8	11,9
19	nullo		9,3		22,6	5,1
	grado di concorrenza	33,3	35,7	33,3	29,0	37,8
		66,7	55,0	66,7	48,4	57,1
20	Numero medio di concorrenti			7	8	13
21	nulla				21,9	44,1
	migliore capacità di progettazione			33,3	34,4	26,9
				66,7	43,8	29,0
	maggior velocità di esecuzione degli ordini				18,8	6,4
				33,3	43,8	41,5
				66,7	37,5	52,1
	migliore qualità del prodotto				6,3	6,4
				66,7	21,9	22,3
				33,3	71,9	71,3
	maggior gamma dei prodotti			33,3	21,9	19,1
					46,9	34,0
				66,7	31,3	46,8
	migliore capacità di risposta a richieste particolari			66,7	9,4	1,1
					9,4	22,1
				33,3	81,3	76,8
	migliore capacità di promozione delle vendite			33,3	18,8	40,9
				66,7	62,5	53,8
					18,8	5,4
	migliore qualità del servizio post-vendita				12,5	35,5
				66,7	53,1	32,3
				33,3	34,4	32,3
				33,3	31,3	21,3
	prezzi più bassi			33,3	53,1	63,8
				33,3	15,6	14,9
				66,7	21,9	26,6
	costi più bassi				59,4	58,5
				33,3	18,8	14,9
				33,3	28,1	20,4
	migliori condizioni di pagamento			33,3	62,5	62,4
				33,3	9,4	17,2
					9,4	1,1
	tradizione di affidabilità				21,9	12,8
				100,0	68,8	86,2
22	a: stesso settore e area geografica				75,0	73,1
	b: stesso settore ma diversa area geografica				12,5	23,1
	entrambi i casi				12,5	3,8
23	Mai		31,4		7,7	35,6
	collabora nella progettazione (solo per il CT)		32,6		7,7	37,0
			36,0		84,6	27,4
24	Mai		55,8		7,7	64,4
	partecipa alla ricerca e sviluppo (solo per il CT)		26,7		46,2	23,3
			17,4		46,2	12,3
25	Solo con imprese italiane				9,4	9,1
	attivazione di accordi con altre imprese			33,3	18,8	3,0
				33,3	18,8	7,1
				33,3	53,1	80,8
26	Sì			100,0	43,8	13,1
	Collaborazioni di ricerca e sviluppo				56,3	86,9

Tabella 3 *Caratteristiche Strutturali*
variabili:1-7

N. imprese	val. ass.	livello di dissimilarità: 2 gruppi		Cut off al 35% del livello di dissimilarità: 4 gruppi			
		carstr1	carstr2	carstr1a	carstr1b	carstr2a	carstr2b
		131	3	12	119	1	2
conto proprio/ conto terzi	Conto proprio	39,7	100,0	33,3	40,3	100,0	100,0
	Conto terzi	60,3		66,7	59,7		
classi di addetti	6-9	6,9			7,6		
	10-19	29,0		33,3	28,6		
	20-49	40,5		41,7	40,3		
	50-99	11,5			12,6		
	Oltre 99	12,2	100,0	25,0	10,9	100,0	100,0
3tot N. medio addetti	val. ass.	54	1436	49	55	2704	802
classi di fatturato	0-1 mld	5,6			6,2		
	1-5 mld	44,8		41,7	45,1		
	5-20 mld	35,2		41,7	34,5		
	Oltre 20 mld	14,4	100,0	16,7	14,2		100,0
7 produzione di pezzi unici		22,3		3,4	24,3		
5 % acquisti di lavorazioni esterne		12,7	3,3	10,7	12,9	0,8	4,5
6 Valore aggiunto per addetto (mln di lire)		127,1	100,6	118,8	128,1		100,6
2 Livello di integrazione verticale: fasi svolte all'interno	Progettazione e design	71,9	100,0	56,7	73,1	100,0	100,0
	Acquisto di materie prime e compo	79,9	33,3	41,7	84,2	100,0	
	Costruzione stampi e modelli	35,5	40,0	38,6	35,2		80,0
	Fonderia	16,6	42,5	66,0	9,5		85,0
	Fusioni e sinterizzazioni	10,0		25,0	7,7		
	Deformazione a caldo	18,2			22,2		
	Deformazione a freddo	52,9	100,0		58,4		100,0
	Stampaggio parti metalliche	37,6	40,0		38,6		80,0
	Lavorazioni di carpenteria metallic	59,8	35,0	18,0	63,4	50,0	20,0
	Lavorazioni su macchine utensili	68,0	85,0	96,0	64,9	90,0	80,0
	Trattamenti chimici e termici	22,9	45,0	16,7	23,7		90,0
	Trattamenti superficiali	26,6	50,0	38,3	25,4		100,0
	Verniciatura	41,4	95,0	24,7	43,0	90,0	100,0
	Assemblaggio parti meccaniche	89,4	100,0	48,0	92,2	100,0	100,0
	Assemblaggio componenti elettrici	60,9	35,0	25,0	63,4	70,0	
	Montaggio finale	91,6	100,0	52,0	94,2	100,0	100,0
Controllo e collaudo	95,5	100,0	93,9	95,7	100,0	100,0	
Assistenza tecnica ai clienti	94,4	100,0	100,0	94,0	100,0	100,0	
Trasporto beni prodotti	38,5		43,4	38,1			
16 quota di fatturato principali clienti	quota fatturato primo cliente	39,0	10,0	50,8	37,6		10,0
	quota fatturato secondo cliente	14,9	8,0	11,8	15,2		8,0
	quota fatturato terzo cliente	9,5	5,0	6,5	9,8		5,0
4 ripartizione del fatturato per aree geografiche (val. %)	Provincia di Modena	38,6	0,5	63,9	35,9		0,5
	Provincia di Reggio Emilia	5,6		3,4	5,9		
	Resto dell'Emilia-Romagna	17,3	0,5	16,7	17,4		0,5
	Altre regioni italiane	23,3		10,8	24,7		
	Altri paesi UE	8,6	21,0	5,3	8,9		21,0
Altri paesi	5,9	77,5		6,6		77,5	

Tabella 4 *Fornitori*
variabili: 8-14

		Cut off al 90% del livello di dissimilarità: 2 gruppi		Cut off al 70% del livello di dissimilarità: 5 gruppi				
		1	2	forn1	forn2	forn3	forn4	forn5
N. imprese	val.ass.	8	126	43	12	42	29	8
conto proprio/ conto terzi	Conto proprio	87,5	38,1	34,9	58,3	38,1	34,5	87,5
	Conto terzi	12,5	61,9	65,1	41,7	61,9	65,5	12,5
	6-9		7,1	4,7	8,3	9,5	6,9	
	10-19		30,2	39,5	25,0	35,7	10,3	
classi di addetti	20-49	25	40,5	39,5	50,0	31,0	51,7	25,0
	50-99	12,5	11,1	11,6	8,3	11,9	10,3	12,5
	Oltre 99	62,5	11,1	4,7	8,3	11,9	20,7	62,5
3tot	N. medio addetti	494	59	48	67	49	88	494
	0-1 mld	14,3	5,0	4,8		7,5	3,8	14,3
classi di fatturato	1-5 mld	14,3	45,8	59,5	41,7	40,0	34,6	14,3
	5-20 mld		36,7	26,2	41,7	40,0	46,2	
	Oltre 20 mld	71,4	12,5	9,5	16,7	12,5	15,4	71,4
6	Valore aggiunto per addetto	146,0	125,9	123,5	119,4	126,2	131,8	146,0
5	% acquisti di lavorazioni esterne	8,2	12,7	13,9	14,7	12,6	10,4	8,2
2	Progettazione e design	72,5	72,9	68,5	72,7	77,5	75,3	72,5
	Acquisto di materie prime e componenti	75,0	79,0	90,1	83,3	69,3	75,2	75,0
	Costruzione stampi e modelli	41,4	35,0	43,0	61,3	10,6	35,3	41,4
	Fonderia		22,0	16,7		9,8	71,3	
	Fusioni e sinterizzazioni		10,7			11,1	28,6	
	Deformazione a caldo		22,2		25,0		50,0	
	Deformazione a freddo	40,0	56,5	64,4	37,5	66,7	50,0	40,0
	Stampaggio parti metalliche	3,3	40,6	55,4	37,5	25,8	40,0	3,3
Livello di integrazione verticale: fasi svolte all'interno	Lavorazioni di carpenteria metallica	34,0	61,1	69,4	51,7	60,3	52,3	34,0
	Lavorazioni su macchine utensili	38,0	70,4	72,9	60,9	74,0	66,7	38,0
	Trattamenti chimici e termici		25,9	18,8		28,5	45,4	
	Trattamenti superficiali		29,3	18,6	14,3	26,3	57,5	
	Verniciatura	36,8	43,4	35,2	34,4	48,6	52,9	36,8
	Assemblaggio parti meccaniche	89,3	89,7	94,3	95,0	84,6	87,7	89,3
	Assemblaggio componenti elettriche e elettroniche	50,0	61,3	60,5	42,5	68,1	65,5	50,0
	Montaggio finale	91,3	92,0	91,9	100,0	88,0	93,0	91,3
	Controllo e collaudo	99,4	95,4	95,7	99,6	93,9	95,0	99,4
	Assistenza tecnica ai clienti	100,0	94,1	94,3	100,0	89,2	98,3	100,0
	Trasporto beni prodotti	32,5	37,9	44,2	51,8	37,5	22,7	32,5
8	N. fornitori abituali	320	65	50	115	73	54	320
9	Numero fornitori strategici	18	8	5	5	13	5	18
10_1°	localizzazione primo fornitore - ricodificata	Emilia-Romagna	58,8	100,0	54,5	40,0	33,3	
		Italia	66,7	35,1	18,2	54,3	61,9	66,7
		Estero	33,3	6,2	27,3	5,7	4,8	33,3
10_2°	localizzazione secondo fornitore - ricodificata	Emilia-Romagna	33,3	61,1	96,3	72,7	48,5	26,3
		Italia	33,3	34,4	3,7	27,3	42,4	68,4
		Estero	33,3	4,4			9,1	5,3
10_3°	localizzazione terzo fornitore - ricodificata	Emilia-Romagna	60,0	63,8	94,4	87,5	50,0	41,2
		Italia	40,0	33,3		12,5	46,2	58,8
		Estero		2,9	5,6		3,8	
10_4°	localizzazione quarto fornitore - ricodificata	Emilia-Romagna	40,0	62,1	83,3	87,5	54,5	43,8
		Italia	20,0	34,5	8,3	12,5	40,9	56,3
		Estero	40,0	3,4	8,3		4,5	40,0
10_5°	localizzazione quinto fornitore - ricodificata	Emilia-Romagna	80,0	50,0	80,0	20,0	50,0	40,0
		Italia		43,8	20,0	40,0	50,0	53,3
		Estero	20,0	6,3		40,0	6,7	20,0

N. imprese	val.ass.	Cut off al 90% del livello di dissimilarità: 2 gruppi		Cut off al 70% del livello di dissimilarità: 5 gruppi					
		1	2	forn1	forn2	forn3	forn4	forn5	
		8	126	43	12	42	29	8	
11	primo fornitore	16	13	12	16	11	13	16	
Durata del rapporto con i 5 principali fornitori (anni)	secondo fornitore	16	11	11	13	12	10	16	
	terzo fornitore	15	11	11	13	10	11	15	
	quarto fornitore	13	11	14	13	8	9	13	
	quinto fornitore	9	9	11	15	8	7	9	
12	% acquisti dai fornitori strategici	9,0	25,2	32,3	12,9	21,4	26,0	9,0	
13_1°	collaborazione	mai	33,3	90,2	92,9	60,0	100,0	85,0	33,3
progettazione primo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	16,7	6,5	7,1	10,0		15,0	16,7	
		50,0	3,3		30,0			50,0	
13_2°	collaborazione	mai	33,3	89,7	96,0	50,0	97,1	88,9	33,3
progettazione secondo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	33,3	3,4			2,9	11,1	33,3	
		33,3	6,9	4,0	50,0			33,3	
13_3°	collaborazione	mai	40,0	82,4	83,3	57,1	100,0	64,7	40,0
progettazione terzo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	20,0	11,8	5,6	28,6		29,4	20,0	
		40,0	5,9	11,1	14,3		5,9	40,0	
13_4°	collaborazione	mai	100,0	77,6	75,0	42,9	100,0	62,5	100,0
progettazione quarto fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.		15,5		42,9		37,5		
			6,9	25,0	14,3				
13_5°	collaborazione	mai	60,0	83,3	90,0	40,0	100,0	73,3	60,0
progettazione quinto fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	20,0	12,5		40,0		26,7	20,0	
		20,0	4,2	10,0	20,0			20,0	
14_1°	collaborazione R&S primo fornitore	mai	16,7	66,7	92,9		94,3	15,0	16,7
progettazione primo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	66,7	22,6	7,1		5,7	85,0	66,7	
		16,7	10,8		100,0			16,7	
14_2°	collaborazione R&S secondo fornitore	mai	33,3	68,2	96,0		91,4	22,2	33,3
progettazione secondo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	66,7	18,2			8,6	72,2	66,7	
			13,6	4,0	100,0			5,6	
14_3°	collaborazione R&S terzo fornitore	mai	60,0	65,2	94,4		92,6	17,6	60,0
progettazione terzo fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	40,0	27,5	5,6	28,6	7,4	82,4	40,0	
			7,2		71,4				
14_4°	collaborazione R&S quarto fornitore	mai	100,0	61,0	83,3		95,8	18,8	100,0
progettazione quarto fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.		28,8	8,3	28,6	4,2	81,3		
			10,2	8,3	71,4				
14_5°	collaborazione R&S quinto fornitore	mai	40,0	71,4	90,0	20,0	100,0	40,0	40,0
progettazione quinto fornitore	saltuariam. o solo per alcuni compon/lavora per gran parte/tutti i compon/lavoraz.	60,0	22,4	10,0	20,0		60,0	60,0	
			6,1		60,0				
23*	collabora nella progettazione (solo per il CT)	Mai	31,8	35,5	20,0	39,3	19,0		
partecipa alla ricerca e sviluppo (solo per il CT)	Saltuariamente o solo per alcuni clienti	100,0	31,8	32,3	40,0	35,7	23,8	100,0	
	Sempre o quasi sempre		36,5	32,3	40,0	25,0	57,1		
24*	partecipa alla ricerca e sviluppo (solo per il CT)	Mai	56,5	58,1	40,0	75,0	33,3		
partecipa alla ricerca e sviluppo (solo per il CT)	Saltuariamente o solo per alcuni clienti	100,0	25,9	32,3	20,0	10,7	38,1	100,0	
	Sempre o quasi sempre		17,6	9,7	40,0	14,3	28,6		

Tabella 5 *Clienti*

variabili 15-18

N. imprese	val. ass.	Cut off al 70% del livello di dissimilarità: 3 gruppi			Cut off al 50% del livello di dissimilarità: 5 gruppi				
		clien1	clien2	clien3	clien1	clien2a	clien2b	clien2c	clien3
		7	56	71	7	27	26	3	71
conto proprio/ conto terzi	Conto proprio	85,7	66,1	16,9	85,7	55,6	73,1	100,0	16,9
	Conto terzi	14,3	33,9	83,1	14,3	44,4	26,9		83,1
classi di addetti	6-9		1,8	11,3				33,3	11,3
	10-19	14,3	19,6	36,6	14,3	22,2	15,4	33,3	36,6
	20-49	14,3	33,9	46,5	14,3	37,0	30,8	33,3	46,5
	50-99	42,9	17,9	2,8	42,9	18,5	19,2		2,8
	Oltre 99	28,6	26,8	2,8	28,6	22,2	34,6		2,8
3tot	N. medio addetti	100	159	25	100	209	125	16	25
classi di fatturato	0-1 mld		1,9	8,8				33,3	8,8
	1-5 mld	14,3	32,7	55,9	14,3	29,2	32,0	66,7	55,9
	5-20 mld	28,6	36,5	33,8	28,6	41,7	36,0		33,8
	Oltre 20 mld	57,1	28,8	1,5	57,1	29,2	32,0		1,5
7	produzione di pezzi unici	30,7	21,3	21,2	30,7	15,8	21,0	100,0	21,2
15*	percentuale di fatturato per tipi di clienti								
	grandi imprese	18,3	31,8	18,9	18,3	35,5	27,8	33,0	18,9
	medie/piccole imprese	33,3	34,8	54,4	33,3	42,4	31,0	0,3	54,4
	piccole imprese o artigiani	13,3	15,1	24,2	13,3	5,0	27,5		24,2
	imprese/agenti commerciali	28,3	8,1	1,9	28,3	4,2	11,5	13,3	1,9
consumatori	6,7	5,3	0,2	6,7	3,7	1,5	53,3	0,2	
committente pubblico		1,9	0,4		3,7	0,3		0,4	
16*	quota fatturato principali clienti								
	quota fatturato primo cliente	21,0	41,5	38,7	21,0	56,0	28,0	29,3	38,7
	quota fatturato secondo cliente	8,9	13,8	16,1	8,9	10,5	14,4	34,3	16,1
	quota fatturato terzo cliente	5,0	8,3	10,7	5,0	5,0	8,0	34,3	10,7
18*	durata relazione col principale cliente	2	10	15	2	11	8	20	15
19*	grado di concorrenza								
	nullo	14,3	12,7	5,7	14,3	19,2	7,7		5,7
	medio-basso	14,3	27,3	44,3	14,3	19,2	34,6	33,3	44,3
	alto-molto alto	71,4	60,0	50,0	71,4	61,5	57,7	66,7	50,0
17*_1°	localizzazione primo cliente								
	Emilia-Romagna		19,6	98,5		25,0	4,2	100,0	98,5
	Italia		56,9	1,5		66,7	54,2		1,5
	Estero	100,0	23,5		100,0	8,3	41,7		
17*_2°	localizzazione secondo cliente								
	Emilia-Romagna	16,7	25,0	89,4	16,7		39,1	66,7	89,4
	Italia	33,3	40,9	10,6	33,3	100,0			10,6
	Estero	50,0	34,1		50,0		60,9	33,3	
17*_3°	localizzazione terzo cliente								
	Emilia-Romagna		23,3	87,5		11,8	26,1	66,7	87,5
	Italia	60,0	46,5	10,9	60,0	76,5	30,4		10,9
	Estero	40,0	30,2	1,6	40,0	11,8	43,5	33,3	1,6
23*	collabora nella progettazione (solo per il CT)								
	Mai		42,9	29,0		38,5	50,0		29,0
	Saltuariamente o solo per alc	33,3	19,0	37,1	33,3	23,1	12,5		37,1
	Sempre o quasi sempre	66,7	38,1	33,9	66,7	38,5	37,5		33,9
24*	partecipa alla ricerca e sviluppo (solo per il CT)								
	Mai		47,6	61,3		46,2	50,0		61,3
	Saltuariamente o solo per alc	33,3	38,1	22,6	33,3	30,8	50,0		22,6
	Sempre o quasi sempre	66,7	14,3	16,1	66,7	23,1			16,1

Tabella 7 *Collaborazioni*

variabili: 13-14, 23*-26

Cut off al 90% del livello di dissimilarità: 7 gruppi

		coll1+2	coll3+4	coll5+6	coll7+8	coll9	coll10	coll11	
N. imprese	val.ass.	4	16	13	64	12	7	18	
conto proprio/ conto terzi	Conto proprio	50,0	43,8	46,2	26,6	50,0	71,4	66,7	
	Conto terzi	50,0	56,3	53,8	73,4	50,0	28,6	33,3	
classi di addetti	6-9				10,9	8,3		5,6	
	10-19	25,0	12,5	38,5	39,1	16,7	14,3	11,1	
	20-49	50,0	43,8	38,5	34,4	58,3	14,3	50,0	
	50-99	25,0	6,3	7,7	7,8		28,6	27,8	
	Oltre 99		37,5	15,4	7,8	16,7	42,9	5,6	
3tot	N. medio addetti	43	173	68	44	271	83	51	
classi di fatturato	0-1 mld	66,7			6,3			6,3	
	1-5 mld		35,7	53,8	55,6	45,5	14,3	18,8	
	5-20 mld		35,7	30,8	30,2	45,5	42,9	50,0	
	Oltre 20 mld	33,3	28,6	15,4	7,9	9,1	42,9	25,0	
23*	collabora nella progettazione (solo per il CT)	Mai	50,0		48,0			28,6	
		Saltuariamente o solo per alcuni cli	50,0	55,6	12,5	32,0	28,6	42,9	
		Sempre o quasi sempre		44,4	87,5	20,0	71,4	100,0	28,6
24*	partecipa alla ricerca e sviluppo (solo per il CT)	Mai	100,0	22,2	12,5	78,0		57,1	
		Saltuariamente o solo per alcuni cli		55,6	25,0	20,0	42,9	66,7	14,3
		Sempre o quasi sempre		22,2	62,5	2,0	57,1	33,3	28,6
13_1°	collaborazione progettazione primo fornitore	mai	50,0	81,3	91,7	95,2		100,0	100,0
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	18,8	8,3	4,8	16,7		83,3
13_2°	collaborazione progettazione secondo fornitore	mai		85,7	91,7	95,0		100,0	100,0
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	100,0	14,3	8,3	2,5	33,3		66,7
13_3°	collaborazione progettazione terzo fornitore	mai	50,0	84,6	33,3	96,7		50,0	92,9
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	15,4	33,3	3,3	100,0		7,1
13_4°	collaborazione progettazione quarto fornitore	mai	50,0	75,0	33,3	95,5	66,7	100,0	92,3
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	25,0	55,6	11,1	4,5	33,3	
13_5°	collaborazione progettazione quinto fornitore	mai	50,0	72,7	57,1	100,0	33,3	100,0	90,9
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	27,3	42,9		33,3		9,1
14_1°	collaborazione R&S primo fornitore	mai			41,7	92,9		33,3	100,0
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	100,0	8,3	7,1	33,3	66,7	
			50,0		50,0		66,7		
14_2°	collaborazione R&S secondo fornitore	mai			50,0	95,0		100,0	88,9
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	92,9		2,5	50,0		11,1
			50,0	7,1	50,0	2,5	50,0		
14_3°	collaborazione R&S terzo fornitore	mai	50,0	15,4		100,0		50,0	93,3
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	84,6	55,6		100,0	50,0	6,7
			50,0		44,4				
14_4°	collaborazione R&S quarto fornitore	mai	50,0	25,0		95,5	33,3	100,0	92,9
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora	50,0	75,0	44,4	4,5	66,7		
			50,0		55,6				7,1
14_5°	collaborazione R&S quinto fornitore	mai	100,0	27,3	28,6	100,0	33,3		91,7
		saltuariam. o solo per alcuni compc per gran parte/tutti i compon/lavora		72,7	28,6		66,7	100,0	8,3
				42,9					
25	attivazione di accordi con altre imprese	Solo con imprese italiane	50,0	6,3	7,7				44,4
		Solo con imprese straniere		6,3	15,4			100,0	
		Sia con imprese italiane che imprese straniere	50,0	25,0					44,4
		Nessun accordo di collaborazione		62,5	76,9	100,0	100,0		11,1
26	Collaborazioni di ricerca e sviluppo	Sì		43,8	46,2	9,4	8,3	42,9	38,9
		No	100,0	56,3	53,8	90,6	91,7	57,1	61,1

Tabella 8 CT- collaborazioni

variabili: 29-44

			Cut off al 70% del livello di dissimilarità: 3 gruppi			Cut off al 45% del livello di dissimilarità: 4 gruppi			
			Ctcoll1	Ctcoll2	Ctcoll3	CTcoll1a	CTcoll1b	CTcoll2	CTcoll4
N. imprese			32	32	15	5	27	32	15
Conto terzi			100,0	100,0	100,0	100,0	100,0	100,0	100,0
	6-9		9,4	9,4		20,0	7,4	9,4	
	10-19		37,5	46,9			44,4	46,9	
	20-49		40,6	40,6	73,3	60,0	37,0	40,6	73,3
	50-99		9,4		20,0		11,1		20,0
	Oltre 99		3,1	3,1	6,7	20,0		3,1	6,7
3tot	N. medio addetti		27	23	44	46	24	23	44
	0-1 mld		3,2	13,3		25,0		13,3	
	1-5 mld		54,8	70,0	20,0	25,0	59,3	70,0	20,0
	5-20 mld		38,7	16,7	73,3	25,0	40,7	16,7	73,3
	Oltre 20 mld		3,2		6,7	25,0			6,7
23*	collabora nella	Mai	25,0	50,0			29,6	50,0	
	progettazione (solo per il CT)	Saltuariamente o solo per alcuni clienti	37,5	31,3	33,3	60,0	33,3	31,3	33,3
		Sempre o quasi sempre	37,5	18,8	66,7	40,0	37,0	18,8	66,7
24*	partecipa alla ricerca e sviluppo (solo per il CT)	Mai	53,1	78,1	13,3	60,0	51,9	78,1	13,3
		Saltuariamente o solo per alcuni clienti	34,4	15,6	46,7	40,0	33,3	15,6	46,7
		Sempre o quasi sempre	12,5	6,3	40,0		14,8	6,3	40,0
7	fatt. nella produzione di pezzi unici		17,1	21,0	26,1	2,0	20,0	21,0	26,1
4	ripartizione del fatturato per aree geografiche (val. %)	Provincia di Modena	51,8	55,3	27,6	51,8	51,8	55,3	27,6
		Provincia di Reggio Emilia	6,6	7,6	6,1	1,0	7,4	7,6	6,1
		Resto dell'Emilia-Romagna	19,9	21,7	25,3	23,5	19,4	21,7	25,3
		Altre regioni italiane	18,5	13,7	26,4	19,8	18,4	13,7	26,4
		Altri paesi UE	2,8	0,8	8,0	3,5	2,7	0,8	8,0
		Altri paesi	0,3	0,8	6,5	0,5	0,3	0,8	6,5
15	percentuale di fatturato per tipi di clienti	grandi imprese	25,2	26,5	17,7	20,0	26,1	26,5	17,7
		medie/piccole imprese	49,0	53,1	48,9	54,0	48,1	53,1	48,9
		piccole imprese o artigiani	22,6	20,4	30,9	26,0	21,9	20,4	30,9
		imprese/agenti commerciali	0,2	0,0	1,9		0,2	0,0	1,9
		consumatori			0,7				0,7
		committente pubblico							
18	durata relazione col principale cliente (anni)		14	14	12	16	13	14	12
13_1°	collaborazione	mai	95,0	100,0	78,6	100,0	94,4	100,0	78,6
	progettazione primo fornitore	saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.	5,0		21,4		5,6		21,4
13_2°	collaborazione	mai	100,0	100,0	85,7	100,0	100,0	100,0	85,7
	progettazione secondo fornitore	saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.			14,3				14,3
13_3°	collaborazione	mai	92,3	100,0	66,7	100,0	91,7	100,0	66,7
	progettazione terzo fornitore	saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.	7,7		33,3		8,3		33,3
13_4°	collaborazione	mai	100,0	100,0	66,7	100,0	100,0	100,0	66,7
	progettazione quarto fornitore	saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.			33,3				33,3
13_5°	collaborazione	mai	100,0	100,0	81,8	100,0	100,0	100,0	81,8
	progettazione quinto fornitore	saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.			18,2				18,2
14_1°	collaborazione R&S primo fornitore	mai	95,0	71,4	50,0	100,0	94,4	71,4	50,0
		saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.	5,0	28,6	28,6		5,6	28,6	28,6
					21,4				21,4
14_2°	collaborazione R&S secondo fornitore	mai	100,0	78,9	42,9	100,0	100,0	78,9	42,9
		saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.		21,1	35,7			21,1	35,7
					21,4				21,4
14_3°	collaborazione R&S terzo fornitore	mai	100,0	76,9	33,3	100,0	100,0	76,9	33,3
		saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.		23,1	50,0			23,1	50,0
					16,7				16,7
14_4°	collaborazione R&S quarto fornitore	mai	100,0	77,8	33,3	100,0	100,0	77,8	33,3
		saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.		22,2	41,7			22,2	41,7
					25,0				25,0
14_5°	collaborazione R&S quinto fornitore	mai	100,0	85,7	45,5	100,0	100,0	85,7	45,5
		saltuariam. o solo per alcuni compon/lavo per gran parte/tutti i compon/lavoraz.		14,3	36,4			14,3	36,4
					18,2				18,2

25	N. imprese		Cut off al 70% del livello di dissimilarità: 3 gruppi			Cut off al 45% del livello di dissimilarità: 4 gruppi			
			Ctcoll1	Ctcoll2	Ctcoll3	CTcoll1a	CTcoll1b	CTcoll2	CTcoll4
			32	32	15	5	27	32	15
		Solo con imprese italiane	3,1	6,3	13,3	20,0		6,3	13,3
26	attivazione di accordi con altre imprese	Solo con imprese straniere	6,3		6,7	40,0			6,7
		Sia con imprese italiane che imprese straniere	6,3			40,0			
28		Nessun accordo di collaborazione	84,4	93,8	80,0		100,0	93,8	80,0
	Collaborazioni di ricerca e sviluppo	Sì	6,3	3,1	40,0		7,4	3,1	40,0
		No	93,8	96,9	60,0	100,0	92,6	96,9	60,0
29		Solo fornitore di 1° livello	68,8	58,1	57,1	60,0	70,4	58,1	57,1
	rapporto di subfornitura	Solo fornitore di 2° livello	6,3	22,6		40,0		22,6	
		Fornitore sia di 1° che di 2° livello	25,0	19,4	42,9		29,6	19,4	42,9
30		In crescita	40,6	65,6	66,7	80,0	33,3	65,6	66,7
	Previsioni sulla dinamica degli addetti	Stabile	56,3	28,1	33,3	20,0	63,0	28,1	33,3
		In diminuzione	3,1	6,3			3,7	6,3	
31		In crescita	68,8	90,6	86,7	100,0	63,0	90,6	86,7
32		Stabile	28,1	6,3	13,3		33,3	6,3	13,3
	Previsioni sulal dinamica del fatturato	In diminuzione	3,1	3,1			3,7	3,1	
	Introduzione di modifiche ai prodotti negli ultimi due anni		96,9	3,2	100,0	100,0	96,3	3,2	100,0
33		Introduzione di nuovi prodotti	59,4		60,0	80,0	55,6		60,0
	Tipo di modifiche introdotte	Innovazioni sostanziali ai prodotti esistenti	18,8	3,1	80,0	20,0	18,5	3,1	80,0
		Piccole modifiche di prodotti esistenti	53,1		73,3	20,0	59,3		73,3
		Iniziativa e progetto propri	40,6		93,3	80,0	33,3		93,3
19		iniziativa propria, ma progettazione esterna	3,1		20,0		3,7		20,0
	Iniziativa dell'introduzione di nuovi prodotti	Iniziativa del cliente, ma progettazione propria	34,4		66,7	80,0	25,9		66,7
		Iniziativa e progetto esterni	59,4		13,3	20,0	66,7		13,3
		nullo	9,7	6,3			11,5	6,3	
	grado di concorrenza	medio-basso	48,4	40,6	20,0	100,0	38,5	40,6	20,0
		alto-molto alto	41,9	53,1	80,0		50,0	53,1	80,0

¹ Il progetto di ricerca, coordinato da Margherita Russo, è stato finanziato dall'Università di Modena e Reggio Emilia, dal Murst, dalla Camera di Commercio, dalla Provincia e dal Comune di Modena.

² Russo e Giardino (2000)

³ Per quanto riguarda la determinazione del campione, è stata adottata la tecnica del campionamento stratificato con allocazione ottimale (di Neyman) per strato, determinando quindi la dimensione campionaria strato per strato, in modo da poter controllare la precisione delle stime all'interno di ciascuno di essi. Per un'analisi più dettagliata si veda Lalla (2001).

⁴ Silberston (1972), Brusco e Bigarelli (1995)

⁵ In base a precedenti analisi dei dati, è emerso che le sfumature presenti fra le modalità di risposta "alto" e "molto alto", e "basso" e "medio" non configurano differenze sostanziali nella risposta, e l'aggregazione non riduce il contenuto e il significato delle risposte. Al contrario, per questo motivo, si è invece lasciata in evidenza la modalità di risposta "nullo".

⁶ Pajek è un software per l'analisi di reti di relazioni sociali di grandi dimensioni, progettato e sviluppato da Vlado Batagelj e Andrej Mrvar dell'Università di Ljubljana. Può essere scaricato dalla rete all'indirizzo: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.