

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

Laurea in Statistica e Tecnologie Informatiche



Relazione finale

Customer Satisfaction e analisi dei dati
Clienti Triveneto della **STEFANI-IDM**

Relatore:

Ch.mo Prof. Adelchi Azzalini

Laureando:

Fabio Bosco 491836

A.A. 2004-05

INDICE

1 Introduzione	3
1.1 SCM Group spa	3
1.2 I marchi	6
1.3 Profilo aziendale STEFANI	8
1.3.1 I prodotti	9
1.4 Profilo aziendale IDM	9
1.4.1 STEFANI – IDM	10
1.5 Definizione delle mansioni aziendali	12
2 Disegno complessivo di ricerca	22
2.1 Lo scopo	23
2.2 Popolazione, campionamento e qualità dei dati	23
2.2.1 Verifica per provincia	26
2.2.2 Verifica per tipo cliente	28
2.3 Formulazione del questionario	30
2.4 Modalità di somministrazione	32
2.4.1 Risposte totali ottenute	37
2.5 Fattori che possono compromettere la qualità dei dati	41
2.5.1 Cause della non risposta al questionario	42
3 Analisi	43
3.1 Tabulazione, pulitura e codifica dati	43
3.2 Modello per la previsione del numero di dipendenti	49
3.2.1 Richiami sul modello lineare	50
3.2.2 Stima del modello e previsione dei dati mancanti	51
3.3 Domanda 1	56
3.3.1 Classificazione mediante il modello di regressione logistico	57
3.3.1.a Richiami sulla regressione logistica	57
3.3.1.b Determinazione del modello logistico	58

3.3.2	Albero di classificazione	60
3.3.2.a	Richiami sull'albero di classificazione	60
3.3.2.b	Determinazione dell'albero di classificazione	61
3.3.3	Analisi discriminante lineare	63
3.3.3.a	Richiami sull'analisi discriminante lineare	63
3.3.3.b	Determinazione del modello mediante analisi discriminante lineare	65
3.3.4	Considerazioni domanda 1	67
3.4	Domanda 2	70
3.4.1	Prima parte – acquisto prodotti non STEFANI-IDM	70
3.4.1.a	Classificazione mediante il modello di regressione logistico	70
3.4.1.b	Albero di classificazione	74
3.4.1.c	Analisi discriminante lineare	76
3.4.1.d	Prime considerazioni prima parte	78
3.4.2	Seconda parte – acquisto prodotti STEFANI-IDM	81
3.4.2.a	Classificazione mediante il modello di regressione logistico	81
3.4.2.b	Albero di classificazione	83
3.4.2.c	Analisi discriminante lineare	84
3.4.2.d	Prime considerazioni seconda parte	85
3.5	Domanda 3	90
3.5.1	Classificazione mediante il modello di regressione logistico	91
3.5.2	Albero di classificazione	94
3.5.3	Analisi discriminante lineare	96
3.5.4	Considerazioni domanda 3	97

4 Conclusioni finali

4.1	Conclusione domanda 1	101
4.2	Conclusione domanda 2	102
4.3	Conclusione domanda 3	104

CAPITOLO 1

INTRODUZIONE

1.1 SCM Group spa

Le origini del Gruppo SCM risalgono al 1952, quando nell'azienda creata da Franco Aureli e Nicola Gemmani, che fino ad allora aveva prodotto aratri, seminatrici e getti in ghisa, entra il giovane ingegnere Giuseppe Gemmani, figlio del secondo fondatore. A lui viene affidato il compito di progettare una macchina per il legno, destinata all'artigianato, purché fosse innovativa rispetto a quelle esistenti sul mercato che allora era dominato da consolidate aziende tedesche, inglesi e francesi. Nacque così l'Invincibile e, con essa SCM, primo passo per la costituzione dell'attuale Gruppo SCM.

La scelta di produrre in serie consente vantaggi di costo che si dimostrano determinanti per sostenere la sfida con i già affermati produttori tedeschi, inglesi e francesi. La qualità delle macchine, rese disponibili ad un prezzo competitivo, permette di attrarre in breve tempo una rete capillare di concessionari ed agenti, prima in Italia, poi in Europa ed infine in 120 paesi del mondo.

Parallelamente allo sviluppo della rete commerciale, si amplia la gamma di prodotti. Negli anni '70 il Gruppo produce già macchine automatiche ad elevata capacità produttiva destinate anche alla media e grande industria.

Dagli inizi degli anni '80 l'elettronica diventa parte integrante del prodotto, aumentando le prestazioni delle macchine. La costituzione di filiali e joint-ventures con partner locali in vari paesi del mondo, permette di rispondere con rapidità ed efficienza alla domanda di servizi di consulenza ed assistenza.

Le filiali sono aziende commerciali, localizzate in zone strategiche, per la vendita dei prodotti appartenenti al Gruppo SCM. Le joint ventures, invece, sono aziende associate con il Gruppo SCM, distribuite in tutto il mondo per garantire un adeguato supporto commerciale e di servizio alla clientela (assistenza tecnica e servizio ricambi).

Dalla seconda metà degli anni '80 si ha la forte espansione del Gruppo che, con l'acquisizione di aziende già affermate nel settore, completa la propria offerta di macchine automatiche e sistemi integrati destinati alla grande industria del mobile. Tra il 1987 e il 1992 infatti, sei aziende leader nei rispettivi segmenti di mercato entrano a far parte del Gruppo SCM. In questo modo la gamma dei prodotti copre tutte le esigenze di lavorazione dalla piccola alla grande azienda, dalla lavorazione del legno massiccio a quella del pannello, dalla macchina a CNC ai sistemi integrati ad alta flessibilità.

Negli anni '90 la presenza sui mercati mondiali si rafforza, aumenta il numero di filiali e joint venture rivolte al presidio dei mercati strategici in tutto il mondo: attualmente sono presenti in paesi come Germania, Francia, Spagna, Gran Bretagna, Olanda, Romania, Polonia, Russia, Canada, USA, Singapore, Hong Kong, Cina e Giappone. Si consolida la leadership in settori complementari quali quello delle fusioni in ghisa e della componentistica, si potenziano il settore della ricerca e della formazione professionale, vengono stretti accordi di collaborazione con la Georgia Tech, negli USA, con l'Istituto di Tecnologia del Legno della British Columbia University di Vancouver, con il Fontys di Eindhoven in Olanda e con la Facoltà di Ingegneria dell'Università di Bologna.

Recentemente sono entrate a far parte del Gruppo anche le aziende del Gruppo CMS, attive sia nel campo della tecnologia del legno che in quello delle macchine utensili per il marmo, il vetro e la plastica. Oggi il Gruppo SCM occupa 3.000 dipendenti, ha un fatturato consolidato di 450 milioni di Euro, 18 stabilimenti di produzione, 23 filiali estere, una percentuale di esportazione che supera il 70%.

Tutte le unità produttive operano secondo sistemi di qualità certificata e la maggior parte delle aziende ha già ricevuto la certificazione ISO9001.

ORGANIGRAMMA SCM Group



Figura 1.1. Schema contenente tutto ciò che comprende il Gruppo SCM dai vari marchi alle filiali.

1.2 I MARCHI

Il Gruppo SCM opera attraverso una serie di marchi importanti:

SCM, è l'azienda "storica" del Gruppo con una gamma completa di macchine per la lavorazione del massiccio e del pannello. Le macchine che produce sono destinate ad un target medio-basso, piccole e medie imprese che non hanno né un elevato fatturato né l'esigenza di compiere un numero elevato di operazioni al giorno.

MINIMAX, produce una gamma completa per il laboratorio artigianale, adatto in particolare per gli hobbysti. Di conseguenza il target, a cui sono destinati i prodotti di questo marchio, è basso.

MORBIDELLI, macchine e sistemi di foratura.

GABBIANI MACCHINE, sistemi di sezionatura.

DMC, macchine e linee di levigatura.

IDM e **STEFANI**, bordatrici, squadrabordatrici e linee di squadrabordatura.

ROUTECH, centri di lavoro ad elevata produttività (CNC working centres).

MAHROS, sistemi automatici di movimentazione.

Gli ultimi sette marchi elencati hanno un target alto come grandi società, con un elevato fatturato, un elevato numero di dipendenti, le capacità di produrre un numero elevato di pezzi in una dimensione temporale piuttosto ristretta, e quindi la necessità di compiere un numero elevato di operazioni al giorno.

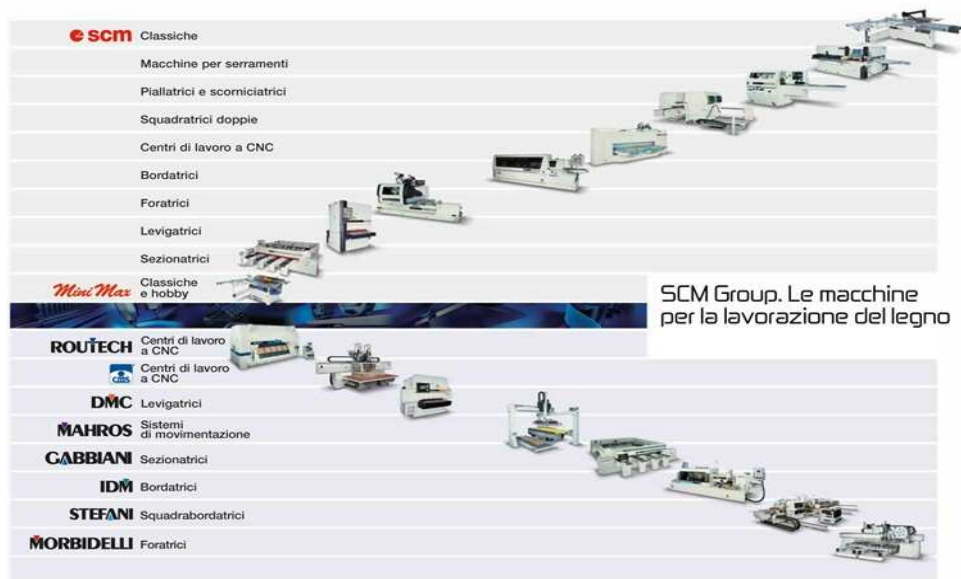


Figura 1.2. Marchi e relativi prodotti del Gruppo SCM.

Il Gruppo, oltre ai marchi elencati precedentemente, comprende anche: due fonderie, per le fusioni di getti in ghisa, un'azienda per le lavorazioni di carpenteria metallica e una delle più importanti aziende italiane del settore elettromeccanico, specializzata nella progettazione e realizzazione di impianti per l'automazione di macchine operatrici.

Il controllo di più aziende diversificate tra loro, ma complementari, ha permesso di adottare una politica di integrazione del processo produttivo, che ha sicuramente incentivato lo sviluppo dell'intero pacchetto aziendale.

Il Gruppo SCM non si dedica solamente all'industria del legno, sia massiccio che pannello, ma comprende anche una sezioni specializzate nella realizzazione di macchine per la lavorazione del marmo e del vetro, nonché laboratori di ricerca impegnati a rilevare e ridurre i livelli di rumorosità e di inquinamento ambientale in molti comparti industriali.

1.3 PROFILO AZIENDALE

STEFANI

STEFANI nasce nel 1950, dedicandosi inizialmente alla produzione di macchine utensili per la lavorazione del legno e del pannello.

Nei primi anni '60 l'azienda compie una scelta produttiva radicale, dedicandosi, prima in Italia e fra le prime al mondo, alla produzione di bordatrici, che si impongono immediatamente all'attenzione mondiale per la loro originalità progettuale e per le qualità tecnico/meccaniche.

L'orientamento alla ricerca e all'innovazione espresso con la produzione delle prime bordatrici monospalla e doppie, ha prodotto, negli anni successivi, anche le prevedibili evoluzioni: squadrabordatrici modulari per la lavorazione del pannello, linee di squadrabordatura con sistema automatizzato di carico e scarico, centri di lavoro per pannelli finiti a partire dal pannello nobilitato di grande formato.

L'ingresso nel Gruppo SCM, avvenuto nel 1992, ha consentito all'azienda di potenziare la propria struttura finanziaria, sfruttare le sinergie produttive, commerciali e tecnologiche, tipiche di un gruppo industriale, e accedere più facilmente ai mercati internazionali.

STEFANI ha sempre investito notevolmente in risorse umane e tecnico-progettuali, creando così un'azienda tecnologicamente avanzata, nella quale vengono progettate e realizzate soluzioni all'avanguardia, offrendo anche un servizio di assistenza post vendita, tramite tecnici specializzati e competenti nel settore. STEFANI vanta referenze di notevole spessore, annoverando tra i propri clienti i maggiori produttori italiani e mondiali di cucine, mobili per la casa, mobili per ufficio e produttori di porte.

1.3.1 I PRODOTTI

La STEFANI e l' IDM si inseriscono principalmente nel settore merceologico delle macchine per la bordatura e la squadrabordatura, di cui producono e/o commercializzano la più completa gamma, in grado di soddisfare tutte le esigenze, dai laboratori artigianali alle medie e grandi imprese. Una macchina bordatrice permette la rifilatura di semilavorati in legno (pannelli), mediante l'applicazione di bordi di varia natura e materiale. Le squadrabordatrici, oltre all'operazione di bordatura, squadrano il pannello.

La bordatura e la squadrabordatura del pannello richiede strutture rigide e resistenti nonché soluzioni tecniche di elevata precisione e affidabilità, in modo da garantire la migliore qualità di finitura in qualunque condizione di impiego. Infatti a questo tipo di macchine viene richiesto di ottenere elevate quantità di produzione, anche differenziate, senza perdere in precisione di lavorazione, sia nella fase di squadratura che in quella di bordatura del pannello.

1.4 PROFILO AZIENDALE

IDM

IDM nasce negli anni '50 come officina per riparazioni meccaniche, ma ben presto decide di orientare la propria attività verso la produzione di macchine per il legno, in concomitanza con lo sviluppo di importanti aziende mobiliere. In particolare, IDM si specializza nella produzione di bordatrici e squadrabordatrici, programmate elettronicamente e in grado di lavorare, in linea, pannelli di vario tipo e dimensioni.

Le soluzioni innovative ed il grande livello tecnologico raggiunto consentono a IDM di servire un'industria mobiliere d'avanguardia ed impostare un tipo di produzione funzionale e mirata.

1.4.1 STEFANI - IDM

Oggi i due marchi sono accorpati in un'unica entità produttiva, nello stabilimento STEFANI del Gruppo SCM di Thiene (VI), dove vengono progettati e realizzati sia i prodotti a marchio STEFANI che IDM. Lo stabilimento è diviso in due unità produttive di circa 12.000 mq e di 4.000 mq e comprende una zona adibita ad ufficio di 2000 mq occupando in tutto 280 persone. L'unità principale è altamente specializzata per il montaggio di macchine di notevole dimensioni: ampi reparti di montaggio dove si monta la fascia alta del prodotto (squadrabordatrici) ed inoltre vengono assemblati i gruppi operatori per tutte le linee di prodotto, il tutto tramite l'utilizzo di attrezzature dedicate per ottimizzare il ciclo produttivo.

I prodotti a marchio IDM comprendono bordatrici monolaterali e squadrabordatrici, mentre il marchio STEFANI comprende bordatrici monolaterali e doppie, squadrabordatrici e linee complete di squadrabordatura. I prodotti a marchio STEFANI forniscono prestazioni tecnologicamente più elevate (velocità, produttività, lavorazioni possibili) rispetto ai prodotti IDM.

STRUTTURA INTERNA STEFANI-IDM

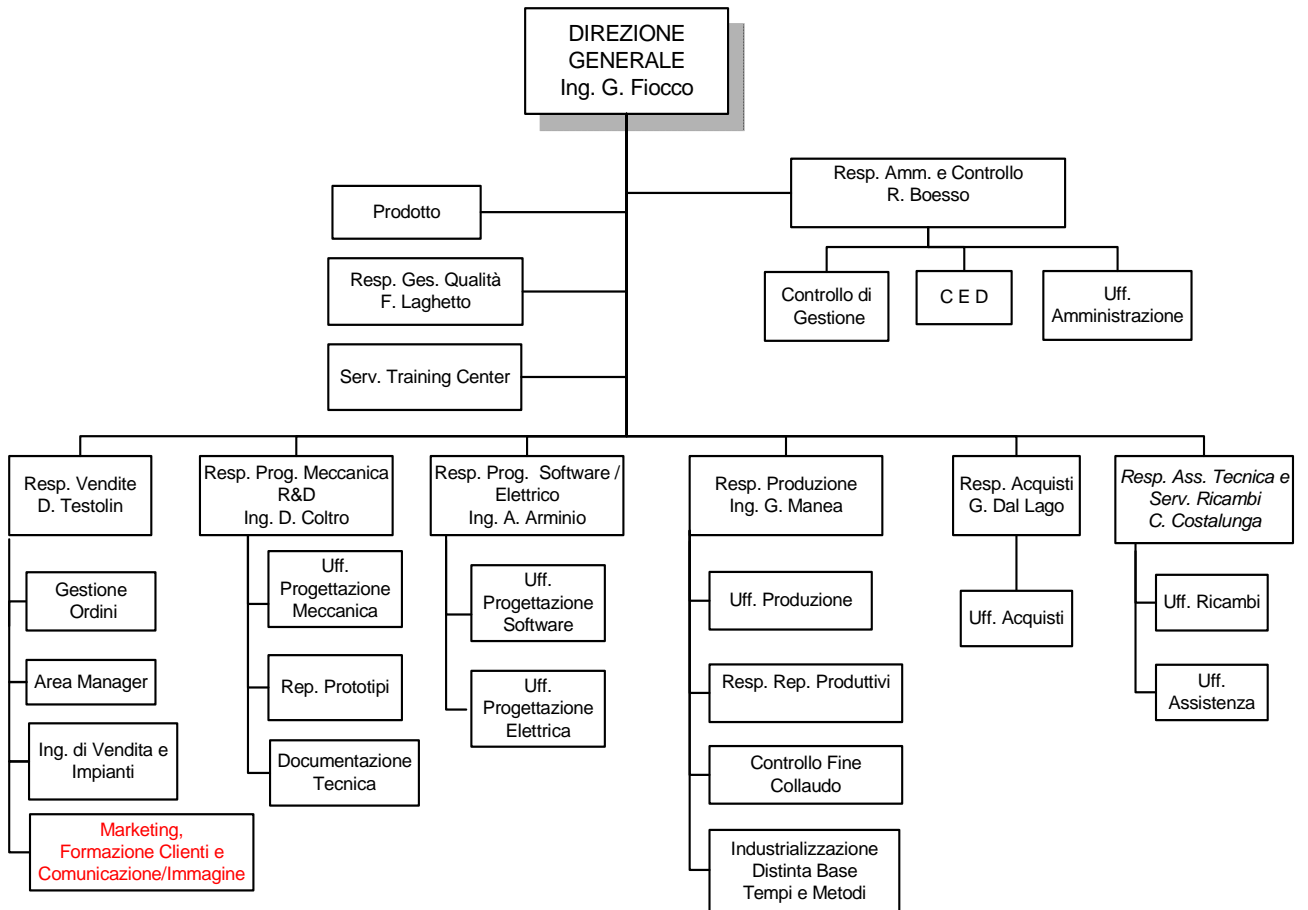


Figura 1.3. Struttura interna del marchio con indicati i responsabili di reparto e in rosso la collocazione dello stage.

1.5 DEFINIZIONE DELLE MANSIONI AZIENDALI

Di seguito verranno descritte le mansioni aziendali all'interno dei marchi STEFANI ed IDM, cioè dell'unità di business bordatura e squadrabordatura del Gruppo SCM, che compongono l'organigramma precedente.

Il Direttore di unità di business, in conformità con la missione ricevuta dal Direttore Generale, organizza la propria attività interna coordinando tra loro gli enti che la costituiscono, facendo riferimento ai responsabili.

A) Prodotto

1. Questo primo ente garantisce un forte presidio dei prodotti di competenza assicurando la congruenza del posizionamento sul mercato.
2. Formula le specifiche prestazionali e qualitative dei prodotti e concorda con le funzioni tecniche e con la responsabilità della direzione i piani di sviluppo.
3. Prepara la configurazione dei listini del prodotto, formando le specifiche prestazionali e qualitative dei prodotti in configurazione aperta e concorda con le funzioni tecniche i piani di sviluppo sotto la responsabilità della direzione.
4. Assicura nel riesame dello sviluppo prodotto il conseguimento degli obiettivi.

B) Gestione Qualità

La gestione della qualità, relativamente al proprio business, si avvale del responsabile della assicurazione qualità che svolge i seguenti compiti.

1. Per quanto riguarda tutti gli enti dipendenti dal direttore di unità di business è responsabile della gestione, dello sviluppo e della progettazione del sistema qualità, che tiene costantemente sotto controllo con verifiche ispettive interne, allo scopo di verificare l'applicazione dei criteri organizzativi stabiliti.

A seguito di tali attività, ove riscontra evidenti scostamenti dai requisiti richiesti, propone adeguate azioni correttive e poi ne verifica attuazione ed efficacia.

2. È il referente di tutte le attività riguardanti la qualità, presidiando la qualità del processo, del prodotto finito e del servizio reso alle strutture distributive e ai clienti finali, attraverso momenti di verifica e confronto con le altre funzioni aziendali.

3. Collabora con i vari reparti al monitoraggio e miglioramento dei processi aziendali.

4. Collabora con il coordinatore di gestione qualità per introdurre nuovi metodi per la qualità partecipando a gruppi di lavoro, in cui si attiva eseguendo i compiti assegnati e controllando che i risultati che ne derivano siano conformi alla politica aziendale.

C) Training Center

Crea documentazione per formare ed informare personale interno e tecnici esterni.

Documentazione di riferimento: Norma UNI EN ISO 9001-00

D) Vendite

1. Assicura univocità di politiche commerciali.

2. Conseguo gli obiettivi di budget definiti.

3. Assicura il rispetto delle condizioni di vendita e dei listini.

4. Gestisce le strutture di vendita dirette, e ne promuove il potenziamento e sviluppa la professionalità degli operatori.

Il reparto 'Vendite' è così composto.

I. Gestione Ordini

- Riceve gli ordini dai clienti, verifica il rispetto delle condizioni generali di vendita.

- Verifica le autorizzazioni dei responsabili commerciali in caso di condizioni commerciali particolari.
- Inserisce gli ordini nel sistema informatico.
- Pianifica le spedizioni con i vettori e la dogana gestendone le pratiche.

II. Area Manager

- Tengono i contatti con la rete distributiva, con l'obiettivo di definire gli ordini, garantendo chiarezza e completezza delle specifiche di contratto.
- Collaborano con l'ingegneria di vendita per l'emissione della richiesta di preventivo per ordini speciali.

III. Ingegneria di Vendita e impianti

- Assicura ai commerciali una costante assistenza per i prodotti.
- Prepara preventivi per ordini speciali e garantisce l'effettuazione dell'esame di fattibilità, attuando momenti di confronto con gli enti competenti. (Progettazione – Produzione – Prodotto).

IV. Marketing, Formazione Clienti e Comunicazione/Immagine

- Tiene sotto costante monitoraggio l'evoluzione delle applicazioni dei prodotti e lo scenario competitivo relativo.
- Imposta i piani di marketing.
- Gestisce il lancio commerciale dei nuovi prodotti relativamente alla promozione, documentazione tecnica commerciale e aggiorna gli uomini di vendita.
- Assicura la formazione della rete di vendita e distributiva, sotto l'aspetto della conoscenza del prodotto e l'utilizzo degli strumenti di vendita.

E) Progettazione Tecnica

1. Sviluppa nuovi prodotti e azioni d'aggiornamento/miglioramento dei prodotti esistenti.
2. Promuove la ricerca di soluzioni tecniche innovative.

3. Garantisce il rispetto delle normative in atto in termini di sicurezza e responsabilità del costruttore.
4. Emette i documenti per la realizzazione ed il controllo del prodotto.
5. È costantemente orientato al miglioramento del prodotto.
6. Cura la preparazione della documentazione tecnica dei prodotti con riferimento a manuali d'uso e cataloghi ricambi.
7. Pianifica le proprie attività nel rispetto dei tempi previsti dal lancio commerciale, così da fornire la documentazione richiesta al momento della consegna.
8. Mantiene aggiornati gli enti post-vendita, ricambi, collaudo, produzione-acquisti in termini di documentazione relativa a modifiche tecniche.

L'area 'Progettazione Tecnica' si suddivide in due settori principali:

I. Progettazione Meccanica e R&D:

Coordina le attività di:

a)Progettazione meccanica

b)Reparto Prototipi

Effettua i test dei modelli sperimentali dei nuovi prodotti.

c)Documentazione Tecnica

Coordina la produzione di documentazione tecnica per uso:formazione, produzione/qualità, marketing, cataloghi ricambi, libretti uso manutenzione.

II. Progettazione Software-Elettrica:

Coordina le attività di:

a)progettazione software

b)progettazione elettrica

F) Produzione:

1. Dà il proprio contributo nella definizione delle specifiche dei nuovi prodotti per gli aspetti di costi,volumi di produzione, standard qualitativi.
2. Garantisce la qualità e l'affidabilità e i tempi di consegna dei prodotti in uscita.

3. Persegue obiettivi di contenimento di costi.
4. Razionalizza le strutture produttive presidiando e potenziando il patrimonio impiantistico amministrato.
5. Garantisce la corretta gestione dei dati dei magazzini materie prime e semilavorati.
6. Garantisce il rispetto delle norme in tema di sicurezza e tutela dell'ambiente.

Questo ente coordina:

I. Industrializzazione

- Effettua studi sui costi preventivi nella fase di sviluppo prodotto.
- Pianifica e standardizza la produzione basandosi sulle caratteristiche del prodotto disegnate.
- Definisce ed emette cicli di lavoro.
- Realizza e mantiene aggiornate le distinte basi.
- Progetta attrezzature particolari che favoriscono la stabilità e la ripetibilità del processo.
- Modifica il processo produttivo in funzione di particolari richieste.
- Collabora alle fasi di definizione delle specifiche di prodotto ed all'esame di fattibilità per ordini speciali.

II. Manutenzione

Si occupa della manutenzione ordinaria e straordinaria degli impianti, delle apparecchiature, dei centri di lavoro, dei mezzi di manutenzione.

III. Controllo Qualità

- Collabora con 'progettazione' e 'industrializzazione' per definire i metodi di controllo.
- Verifica l'accettabilità delle forniture.
- Registra l'andamento qualitativo delle forniture.
- Collabora con i reparti produttivi per l'analisi delle non conformità.

IV. Programmazione Produzione

- Programma l'approvvigionamento dei materiali e la produzione interna, garantendo il raggiungimento degli obiettivi prefissati rispetto a volumi, costi, tempi e qualità.
- Effettua un monitoraggio dello stato d'avanzamento della produzione.
- Collabora all'esame della fattibilità degli ordini speciali.

V. Responsabile reparti Produttivi

Supporta e coordina i reparti produttivi.

VI. Ricevitoria/Magazzino

- Assicura la qualità nelle fasi di ricevimento e spedizioni, movimentazione ed immagazzinamento.
- Esegue la contabilità di magazzino.

VII. Officina-Macchine Utensili

- Esegue la lavorazione dei particolari interni basandosi sulla documentazione tecnica.
- Tiene sotto controllo i parametri di processo, aggiustandoli qualora si discostassero.

VIII. Montaggio

- Assembla i componenti di fornitura esterna e quelli lavorati internamente secondo gli ordini e le distinte e la documentazione fino al realizzo del prodotto.
- Verifica in autocontrollo la corretta esecuzione delle operazioni di sua competenza, basandosi sulle indicazioni riportate nei cicli di montaggio e controllo.
- Organizza al meglio le risorse umane di cui dispone e l'addestramento relativo, coerenza con le linee aziendali di produttività, puntualità di consegna, rispondenza del prodotto finito ai requisiti richiesti.

IX. Verniciatura

- Esegue il trattamento delle superfici.

X. Collaudo Finale

- Assicura e verifica che le macchine STEFANI ed IDM escano secondo le specifiche di cartellino e il check list di controllo.
- Presidia il collaudo finale delle macchine con il cliente, nel caso non fosse presente il cliente, il responsabile del collaudo finale si assume tutte le responsabilità del corretto funzionamento.
- Segnala le non conformità riscontrate agli enti interessati.

G) Acquisti

1. Sceglie i fornitori, avvalendosi dei criteri valutativi previsti.
2. Cura la completezza e precisione degli ordini d'acquisto, verificando che siano trasmesse al fornitore tutte le specifiche tecniche necessarie per la corretta interpretazione e realizzazione.
3. Esegue analisi sull'andamento delle prestazioni, e gestisce le azioni correttive/preventive sui fornitori.

H) Ricambi ed Assistenza

1. Garantisce la qualità e l'immediata disponibilità dei componenti richiesti in seguito ad inconvenienti riscontrati su macchine in esercizio.
2. Gestisce la spedizione dei componenti e le richieste di materiale reso per analisi qualitativa.
3. Assiste il cliente nella fase successiva alla consegna della macchina per quanto riguarda l'installazione, istruzioni d'utilizzo e manutenzione, verifica il rispetto della funzionalità in caso di guasti e gestisce la garanzia.
4. Monitorizza difettosità coordinandosi con 'responsabile progettazione'.
5. Pianifica le attività di installazione e verifica il rispetto dei tempi medi previsti.
6. È responsabile dell'ottenimento dei verbali di collaudo.

I) Amministrazione e Controllo

1. Analizza e verifica le informazioni essenziali sul business e predispone commenti a supporto nella relazione al business manager.
2. Redige un riassunto di gestione (reporting) fornendo supporto ed assistenza ai destinatari.
3. Garantisce la coerenza e veridicità dei dati ed analisi degli scostamenti rispetto al budget e periodi precedenti, ricercandone le cause e proponendo le scelte di correzione.
4. Governa e coordina il processo di pianificazione e budget di unità, in coerenza con i piani strategici di business.
5. Verifica le procedure e i sistemi contabili e gestionali.
6. Garantisce l'applicazione delle procedure interne.
7. Sovrintende la contabilità generale di 'unità' e dei registri IVA.
8. Cura la predisposizione del bilancio civilistico.
9. Sovrintende la contabilità fornitori definendo le regole di registrazione in accordo con le disposizioni di gruppo.
10. È il referente per le attività dell' 'area manager' per quanto riguarda il presidio del margine, le condizioni di pagamento e le regole contrattuali in genere fino al momento di emissione dell'ordine.
11. Sovrintende alla tenuta della cassa ed alla registrazione delle note spese.

Esso coordina:

I. Controllo di gestione

- Controlla i margini di contribuzione ed i costi del sistema aziendale.
- Controllo e predisposizione del report dei costi fissi di Unità e per ciascuna area aziendale.
- Elabora il costo industriale del personale per la predisposizione dei report di periodo.
- Analisi delle varianti industriali: prezzo materiali, prezzo manodopera, resa manodopera, resa materiali (scarti), spesa (costi indiretti e variabili reparti produttivi), assorbimento (sulle ore come precedente).
- Flash mensile completo di magazzino gestionale.
- Magazzini fiscali e fondi obsolescenza.

- Definizione standard annuali.
- Analisi scostamenti rispetto al budget.

II. Ufficio clienti/fornitori

- Gestione delle anagrafiche clienti/fornitori/area zona/area manager.
- Contabilizza le fatture in uscita e gestisce l'incasso dei crediti.
- Contabilizza le fatture in entrata, gestisce i pagamenti e controlla che le fatture siano coerenti agli ordini d'acquisto ed ai carichi di magazzino.

III. Ufficio personale

- Gestisce le pratiche relative a personale (assunzioni, dimissioni, paghe, denunce infortuni, denunce di periodo).
- Controllo ed acquisizione delle presenze giornaliere di 'unità', registrazione e permessi vari/malattie.
- Flash paghe e organici mensili.
- Controllo e inserimento dei dati mensili per elaborazione paghe.
- Controllo e consegna mensile delle cedoline paga.
- Registrazione contabile mensile dei costi del personale, e pagamenti ad enti ed associazioni (associazione industriali,sindacati ecc.)

IV. Servizi amministrativi

- Gestione cassa di 'unità'.
- Verifica e registrazione dei costi per viaggi e trasferte.

V. Contabilità generale

- Gestione banche e tesoreria.
- Verifica e quadratura bilanci civilisti e gestionali.
- Denunce varie di periodo.
- Liquidazione mensile IVA.
- Controllo del piano dei conti.
- Registrazione di apertura e chiusura in sede di bilancio.

VI. CED

- Gestisce i Sistemi Informativi Standard aziendali.
- Gestisce l'hardware in termini di assistenza, riparazione ed amministrazione.

Il progetto di 'Customer Satisfaction', a cui partecipo, si svolge all'interno dell'ufficio marketing in quanto gran parte dei risultati che si andranno ad ottenere riguardano la comunicazione dell'immagine e i concorrenti dell'azienda, formando in questo modo una analisi di mercato. La ricerca potrà essere utile anche ad altri enti facenti parte dell'azienda, in particolare all'ufficio commerciale e al servizio assistenza tecnica, i quali verranno valutati dalle aziende prescelte attraverso il progetto.

CAPITOLO 2

DISEGNO COMPLESSIVO DI RICERCA

Eseguire una ricerca, sia essa di tipo economico-commerciale o più strettamente scientifica, non è una operazione semplice, come si potrebbe pensare in un primo momento, ma piuttosto complessa. Per arrivare al risultato finale, cioè aumentare la propria conoscenza su un fatto, un oggetto, un fenomeno in genere, è necessario seguire alcune fasi che sono concatenate tra loro. In particolare, tenendo conto delle decisioni prese in una prima fase, si sceglie il modo di procedere successivamente.

Le fasi che possono considerarsi standard per la formazione di un disegno di ricerca sono elencate di seguito.

1° FASE definizione degli scopi della ricerca

2° FASE scelta della popolazione di interesse e conseguente formulazione di un campione

3° FASE determinazione dell'insieme di domande da sottoporre ai soggetti, valutando gli oggetti e i concetti che si vogliono esaminare

4° FASE modalità di somministrazione delle domande e della raccolta dati

5° FASE tabulazione, pulitura e codifica dei dati

6° FASE analisi dei dati e loro interpretazione

7° FASE commenti e considerazioni sui risultati ottenuti

Come si può intuire dall'elenco precedente con 'disegno di ricerca' ci si riferisce a tutti gli aspetti inerenti alla pianificazione e conduzione di una ricerca, comprese anche le fasi di analisi dei dati raccolti e presentazione finale dei risultati.

La definizione degli scopi della ricerca costituisce la primaria operazione, in quanto da questa hanno origine e fanno riferimento tutte le altre, prime su tutte la formulazione di un campione e la definizione delle domande, riferite agli oggetti che si vogliono indagare. La raccolta dei dati in una certa modalità viene decisa in base alla considerazione congiunta delle caratteristiche delle domande e del campione.

Nella quinta fase si attua la codifica delle risposte alle domande, permettendo l'attuazione di un'analisi secondo i canoni richiesti dai vari strumenti statistici, sesta fase. In questo modo si possono commentare e trarre indicazioni dai risultati ottenuti, che di per sé è lo scopo per cui una ricerca viene effettuata.

Il processo di disegno di ricerca deve essere considerato iterativo più che lineare. In questo modo, tramite ragionamenti o osservazioni ottenuti in una fase successiva, si possono modificare decisioni prese in una fase precedente.

2.1 LO SCOPO

L'azienda STEFANI del Gruppo SCM spa tramite questa ricerca vuole valutare come è la percezione nel Triveneto (Veneto, Friuli-Venezia Giulia, Trentino-Alto Adige) dei marchi STEFANI ed IDM, mercato di casa ma tecnologicamente molto avanzato. Il tutto avviene effettuando poche domande, rivolte sia a clienti attuali che potenziali, riguardanti se stessa, i concorrenti e l'azienda intervistata. In questo modo si riesce ad ottenere un giudizio sulle macchine proprie ma anche su quelle della concorrenza, e quindi valutare concretamente il lavoro svolto. Inoltre si possono individuare i clienti potenziali più interessanti i quali possono essere contattati dalla

STEFANI per porre loro delle offerte sulle macchine o perlomeno essere invitati ufficialmente dall'azienda ad alcune fiere dedicate alle macchine per il legno oppure direttamente in sede per toccare con mano i prodotti che vengono offerti alla clientela.

Dalle risposte si possono ottenere anche altre informazioni importanti che verranno elencate e discusse in seguito, dopo aver ottenuto delle indicazioni più determinate sui dati a disposizione per effettuare delle analisi.

2.2 POPOLAZIONE, CAMPIONAMENTO E QUALITA' DEI DATI

Un campione di pochi soggetti ma formato in modo appropriato permette di fare inferenze valide circa l'intera popolazione di interesse, cioè si possono generalizzare i dati raccolti e quindi le informazioni raccolte mediante una ricerca.

Le inferenze fatte su un campione rappresentativo casuale possono essere generalizzate alla sola popolazione dalla quale il campione è stato ricavato.

Un metodo di campionamento adeguato è essenziale per la qualità dei dati raccolti soprattutto in quei casi in cui alcune variabili di interesse non si distribuiscono in modo casuale nella popolazione.

Lo scopo del campione è quello di rappresentare il più fedelmente possibile le caratteristiche della popolazione totale, in modo che un gruppo non sia sovrastimato o sottostimato, ma in modo numericamente inferiore.

In questo progetto si è dapprima estratto dal database aziendale quella che può essere considerata la popolazione di riferimento, cioè tutte quelle aziende note all'interno del Gruppo SCM localizzate nel Triveneto. È stata fatta una prima estrazione per provincia ottenendo 8240 clienti, tra attuali e potenziali, che potevano ricoprire il ruolo di soggetti intervistati.

Dalla popolazione si è deciso di formare un campione formato da 1149 clienti. La forma di campionamento utilizzata è definita 'di comodo'. Si ha questa forma quando i soggetti campionati hanno la caratteristica di essere

facilmente contattabili e hanno almeno una caratteristica nota che interessa al committente della ricerca.

La caratteristica nota sfruttata in questo caso è il settore produttivo cioè si è formato un campione con quei soggetti che, oltre a contenere almeno un valore in questa variabile all'interno della scheda cliente presente nel database, presentavano valori che possono essere compatibili con l'utilizzo di macchine bordatrici e squadrabordatrici prodotte da STEFANI e IDM. I settori produttivi selezionati sono i seguenti:

- arredamenti e strutture navali e per camper
- componenti per mobili
- falegnameria generica – mobili su misura
- ingressi – soggiorni – camere da letto
- mobili per il bagno
- mobili per la cucina
- mobili in kit
- mobili per l'ufficio
- porte interne

L'utilizzo di campioni formati in questo modo viene criticato per varie ragioni: si trascurano problemi quali la tendenziosità dei dati raccolti e la validità ecologica delle variabili indagate. Inoltre viene spesso usato più per pigrizia che per impossibilità di attuare un disegno campionario corretto.

È spesso necessario l'utilizzo di un metodo adeguato di campionamento per ottenere dati e quindi considerazioni generalizzabili alla popolazione di interesse.

Per verificare se l'estrazione per settore, più che una vera e propria forma di campionamento, riflette le caratteristiche della popolazione di interesse si è andati a verificare se alcune variabili si distribuiscono allo stesso modo anche nella popolazione formata. Una prima variabile di interesse è sicuramente la provincia, cioè è importante verificare se la proporzione del numero di clienti per provincia presente nella popolazione di interesse è rispettata nella popolazione ricavata tramite campionamento.

2.2.1 VERIFICA PER PROVINCIA

SCALA COLORI PER RAPPRESENTAZIONE

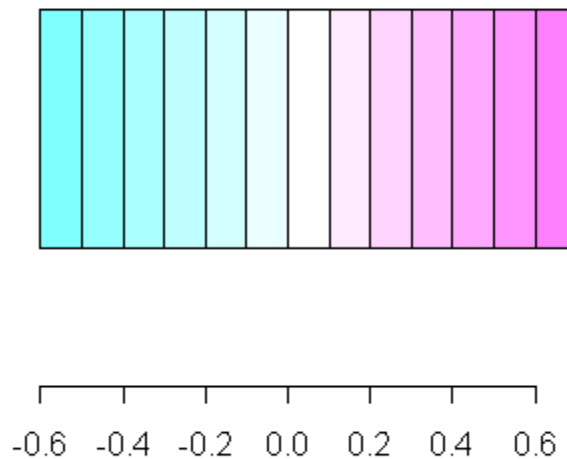


Figura 2.1. Scala colori che verrà utilizzata nelle figure successive.

È stata usata questa rappresentazione in quanto concettualmente serviva l'idea di un colore chiaro (bianco) per indicare che la proporzione della provincia era uguale sia nella popolazione che nel campione formato. I due distinti colori servono per distinguere la sovrarappresentanza (viola) dalla sottorappresentanza (azzurro). L'intensità del colore permette di verificare di quanto una determinata provincia si discosta dalla situazione ottimale cioè logaritmo del rapporto tra la proporzione della provincia nel campione e nella popolazione uguale a 0 indicato dal colore bianco, tramite una scala di 6 possibili variazioni. I valori assegnati ad ogni spicchio della torta indicano il valore del logaritmo del rapporto.

CONFRONTO TRA IL CAMPIONE E LA POPOLAZIONE

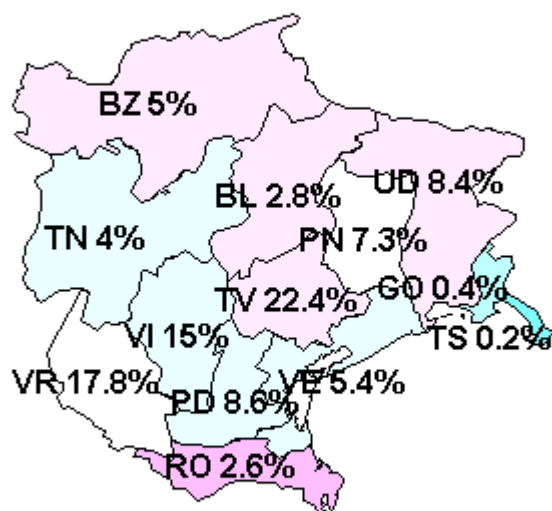


Figura 2.2. Confronto tra campione e popolazione utilizzando la scala colori e percentuale delle province all' interno del campione.

La mappa precedente contiene alcune informazioni rilevanti sulla struttura del campione. Le percentuali riportate indicano la proporzione delle province all'interno del campione. Il colore dato alle province non è casuale ma segue la scala colori di Figura 2.1.

Come si può vedere dalla mappa la provincia di Gorizia ma soprattutto la provincia di Trieste sono sottorappresentate nel campione. Viceversa la provincia di Rovigo risulta essere sovrarappresentata. Inoltre solamente due province, Verona e Pordenone, sono di colore bianco. Questo sta ad indicare che per molte province il campionamento effettuato non rispetta la popolazione anche se non in modo notevole, visto che l'intensità dei colori per la maggior parte delle province è bassa.

Gli scostamenti tra popolazione e campione in base alla provincia si ripercuotono però anche su altre variabili direttamente collegabili come le città e il cap.

I risultati ottenuti non sono certamente ideali, ma considerando il tipo di campionamento utilizzato ci si poteva attendere anche una situazione ben peggiore.

2.2.2 VERIFICA PER TIPO DI CLIENTE

Un'altra ipotesi importante da verificare anche per il tipo di campionamento effettuato è sicuramente l'andamento del tipo di cliente. All'interno del database aziendale esiste la variabile 'tipo di cliente' nella quale lo si classifica in base al rapporto che ha con il Gruppo SCM. Le quattro possibili categorie di clienti sono:

- A : CLIENTE GESTIONALE

Il cliente è diretto, le macchine gli sono state vendute direttamente dalle aziende che costituiscono il Gruppo SCM.

- B : CLIENTE

Il cliente è indiretto, le macchine gli sono state vendute tramite rappresentanti, agenti o società di vendita.

- C : PROSPECT FORTE

Il cliente è potenziale con contatto, gli sono state fatte delle offerte ma non c'è stata una vendita.

- D : PROSPECT DEBOLE

Il cliente è potenziale, ha un'attività che può essere coerente con l'utilizzo delle macchine del Gruppo SCM ma non c'è stato né un contatto né quindi una vendita.

POPOLAZIONE

A	B	C	D
2943	1735	8	3554

Tabella 2.1. Tipo cliente all'interno della popolazione.

CAMPIONE

A	B	C	D
493	359	3	298

Tabella 2.2. Tipo cliente all'interno del campione.

DISTRIBUZIONE NELLA POPOLAZIONE

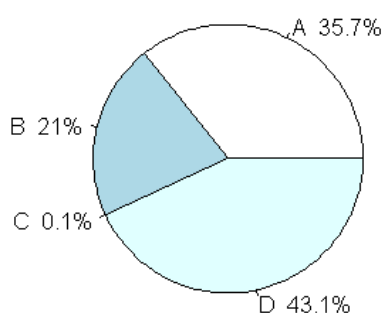


Figura 2.3. Grafico a torta del tipo di cliente nella popolazione.

DISTRIBUZIONE NEL CAMPIONE

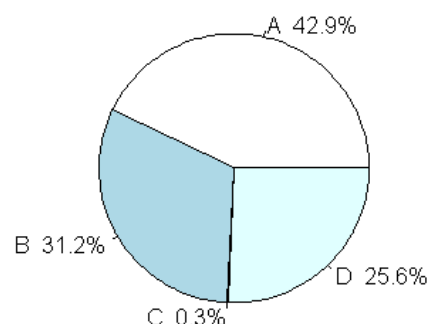


Figura 2.4. Grafico a torta del tipo di cliente nel campione.

Si ha una notevole diversità di valori della variabile 'tipo di cliente' passando dalla popolazione al campione. Questo era prevedibile proprio per il tipo di campionamento effettuato cioè in base al settore produttivo e quindi anche alla sua conoscenza. È intuitivamente molto più facile che nel database appaia il settore di un cliente già noto piuttosto che di una azienda con cui non si hanno mai avuto contatti. Come si può constatare dai grafici precedenti si ha un notevole aumento dei clienti di tipo A e B e la conseguente diminuzione dei clienti D passando dalla popolazione al campione. I clienti C non hanno una particolare rilevanza in nessuno dei due gruppi.

2.3 FORMULAZIONE DEL QUESTIONARIO

Il direttore generale, il responsabile del commerciale del Triveneto, assieme ad altri membri dell'azienda, hanno deciso in azienda di raccogliere le informazioni necessarie per un'analisi del mercato tramite un breve questionario, metodo più semplice per raccogliere le informazioni di interesse. Le domande poste ai soggetti intervistati sono solamente sei, non perché ci siano pochi argomenti da trattare ma per non rendere troppo pesante la compilazione del questionario, data già l'elevata riluttanza delle persone per questo tipo di operazioni. Un altro motivo per cui le domande sono poche, ma comunque sufficienti per compiere un'indagine, verrà spiegato durante il metodo di somministrazione e raccolta del questionario. Si analizza di seguito ogni singola domanda descrivendone in particolare l'oggetto, la struttura e il grado di intrusività.

1. La prima informazione viene ricavata tramite una domanda generale a scelta forzata con risposta dicotomica e verte sulla conoscenza dei prodotti a marchio STEFANI ed IDM.

2. La seconda domanda riguarda l'acquisto di prodotti STEFANI ed IDM e ha una struttura più complessa in quanto si compone di una prima domanda filtro di formato dicotomico in cui se la risposta è affermativa si accede ad altre domande dello stesso argomento mentre se negativa si passa alla macrodomanda successiva. Le sottodomande a cui si accede richiedono un giudizio sull'affidabilità e sul servizio assistenza del prodotto acquistato, che deve essere preventivamente specificato, tramite domande a risposta forzata su una scala di giudizio ordinale a 4 intervalli.

Rispondere a questa intera domanda risulta complicato per il soggetto in quanto per prima cosa deve fare un riferimento con la memoria agli ultimi cinque anni, visto che è espressamente richiesto dalla domanda. Inoltre le ultime due sottodomande sono di soggettive e quindi risulta molto più elaborato per una persona esprimere un giudizio, una opinione piuttosto che rispondere a una domanda di tipo oggettivo.

3. La domanda tre ha la stessa struttura e difficoltà nel rispondere della domanda due ma questa volta riguarda l'acquisto di prodotti da marchi concorrenti alla STEFANI.

4. La struttura di questa domanda è simile alle due precedenti, visto che la prima parte, che verte sulla conoscenza degli agenti STEFANI ed IDM, è costituita da una domanda filtro che permette di accedere, se la risposta è affermativa, a una sottodomanda che riguarda il numero di visite mensili che ogni soggetto riceve dagli agenti dell'azienda.

5. Nella quinta domanda l'azienda intervistata deve rispondere a riguardo del proprio settore produttivo. La domanda è chiusa e a risposta multipla cioè il soggetto può scegliere tra più opzioni elencate, visto che sono possibili più settori produttivi per un'azienda in genere. Tra le alternative di risposta sono presenti quelle che più spesso hanno a che vedere con l'utilizzo di bordatrici e squadrabordatrici ma è stato inserito anche il valore 'altro' con la possibilità di specificare da parte del soggetto dell'opzione che a lui sembra più indicata. Questo è stato fatto perché il numero di opzioni di risposta sarebbe troppo elevato e alcune scelte verrebbero indicate solamente da un numero ristretto di soggetti. Per questo la lista formulata indica solamente le opzioni di scelta principale per l'utilizzo di prodotti STEFANI ed IDM.

6. L'ultima domanda del questionario è una domanda aperta e riguarda il numero di dipendenti impiegati dall'azienda intervistata in modo da poter avere un riferimento sulla sua grandezza. Questa domanda è interessante perché i prodotti STEFANI ed IDM vengono utilizzati solamente da medie e grandi aziende quindi avere un riferimento sulla dimensione del rispondente risulta essere di fondamentale importanza.

2.4 MODALITA' DI SOMMINISTRAZIONE

Come indicato precedentemente la modalità di somministrazione ha influito sulla formulazione del questionario in particolare sul numero di domande. Infatti si è deciso che il questionario doveva essere autocompilato dalla azienda cliente la quale lo riceveva tramite lettera postale. All'esterno delle busta appariva il marchio STEFANI mentre all'interno c'erano due fogli. Un primo foglio di introduzione aveva la funzione di spiegare all'azienda prescelta quale era il ruolo del secondo foglio cioè il questionario vero e proprio.

Il ritorno delle risposte doveva avvenire via fax. L'impaginazione doveva avvenire tutta in un solo foglio in modo tale da rendere le operazioni di risposta più semplici possibili. Questa decisione di carattere pratico, sommata con un'intestazione grafica che in qualche modo riuscisse a colpire chi doveva rispondere e l'inserimento dell'informativa sulla privacy ai sensi del decreto legislativo 196/2003, ha tolto spazio 'fisico' alle domande.

ANDAMENTO RITORNO QUESTIONARIO

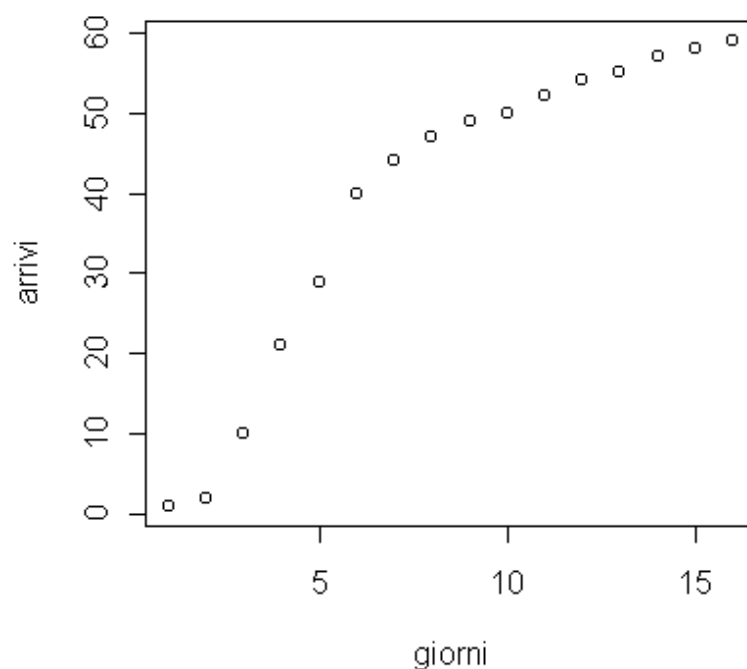


Figura 2.5. Arrivo cumulativo dei questionari via fax.

La figura precedente indica quello che è stato l'andamento, piuttosto scarso, del numero di risposte ricevuto dopo 16 giorni dall'arrivo del primo questionario. Il numero totale di risposte è stato di 59, quindi c'è stato un tasso di risposta del nemmeno 5%, piuttosto basso per tentare di effettuare un' analisi adeguata.

Macchine e sistemi di squadrabordatura
per l'industria del mobile
Combined edgebanding and squaring
machines and systems
for the furniture industry

SCM GROUP Spa
STEFANI
Via del Lavoro, 1/3
PO BOX 168
36016 Thiene (VI) - Italia

Tel. +39/0445/359511
Fax +39/0445/359599
www.scmgroup.com
stefani@scmgroup.com

STEFANI

Spett. «RAG_SOC_____»
c.a. DIREZIONE DI STABILIMENTO
«INDIRIZZO»
«CAP» «CITTA» («PR»)
ITALIA

Thiene, 21.03.05

Egregio Signore,

allo scopo di migliorare l'efficienza dei nostri servizi al Cliente, stiamo conducendo un'indagine tesa valutare la percezione dei marchi STEFANI ed IDM.

Le chiediamo, perciò, di dedicare gentilmente qualche minuto alla compilazione del sondaggio riportato nel foglio allegato e di farcelo pervenire via fax, con il consenso alla raccolta dei dati personali ai sensi del decreto legislativo 196/2003.

RingraziandoLa anticipatamente per la Sua attenzione e collaborazione, La preghiamo di inviare il sondaggio compilato al numero di fax: **0445-359599**.

Cordiali saluti.

Ing. Giannantonio Fiocco
(Direttore Generale)

Sede legale: Via Emilia, 77 - 47900 Rimini (RN) - Capitale sociale € 50.000.000,00 i.v. - Reg. Impr. Rn, C.F. e P.Iva 00126480409 - R.E.A. 72510 RN - M/RN 005359

Figura 2.6. Foglio di introduttivo con indicati i motivi della somministrazione e le modalità di compilazione e di restituzione del questionario.

Macchine e sistemi di squadrabordatura per l'industria del mobile



Spett. «RAG_SOC ____»

c.a. DIREZIONE DI STABILIMENTO

«INDIRIZZO»

«CAP» «CITTA» («PR»)

ITALIA

1. Conosce i prodotti a marchio STEFANI ed IDM? _____ SI NO

2. Ha acquistato prodotti a marchio STEFANI o IDM negli ultimi cinque anni? _____ SI NO

Se si, che modello? _____

a. Come valuta l'affidabilità dei prodotti STEFANI ed IDM? _____
 Elevata Buona Sufficiente Insufficiente

b. Come valuta il servizio di assistenza tecnica STEFANI ed IDM? _____
 Elevato Buono Sufficiente Insufficiente

3. Ha acquistato bordatrici, squadrabordatrici o linee di squadrabordatura non a marchio STEFANI - IDM negli ultimi cinque anni? _____ SI NO

Se si, che marchio? _____ Che modello? _____

a. Come valuta l'affidabilità dei prodotti acquistati? _____
 Elevata Buona Sufficiente Insufficiente

b. Come valuta il servizio di assistenza tecnica del marchio acquistato? _____
 Elevato Buono Sufficiente Insufficiente

4. Conosce gli agenti di zona per i marchi STEFANI ed IDM? _____ SI NO

Se si, quante visite riceve mensilmente dagli agenti STEFANI ed IDM? _____

5. Qual è il settore produttivo della Sua azienda?

Mobili per ufficio Cucine Mobili per bagno

Mobili per la casa Componenti per mobili Altro _____

6. Quanti dipendenti sono impiegati nella Sua azienda? _____

INFORMATIVA AI SENSI DEL D.LGS.196/2003

In ottemperanza agli adempimenti richiesti dal D.Lgs. 196/2003 ed in conformità al codice europeo di condotta e di autodisciplina per l'uso dei dati personali nelle attività di direct marketing, La informiamo che: titolare del trattamento dei dati è SCM GROUP SPA; i dati raccolti presso di Voi verranno conservati esclusivamente per fini commerciali e non saranno soggetti né a comunicazione, né a diffusione; la gestione dei dati è elettronica, le applicazioni usate sono pacchetto Excel e database Oracle. La risposta alle domande contenute nel questionario che accompagna questa informativa è facoltativa. Se riterrate di voler collaborare alla Nostra analisi Vi preghiamo di farci pervenire questo medesimo modulo, con la casellina sottostante barrata. Ai sensi dell'articolo 7 del Decreto Legislativo 196/2003 avete il diritto di conoscenza, cancellazione, rettifica, aggiornamento, integrazione e opposizione al trattamento dei dati stessi. Il responsabile del trattamento dei dati raccolti è il Sig. Antonio Nicolini, domiciliato presso SCM GROUP SPA, via Emilia, n. 77, 47900 Rimini.

Presto il consenso alla raccolta dati (barrare casella)

Timbro e firma del legale rappresentante del destinatario _____

STEFANI- IDM - SCM GROUP spa - Via del Lavoro, 1/3 - 36016 Thiene (VI) - Italy - Tel. +39/0445/359511 - Fax +39/0445/359599

Figura 2.7. Questionario consegnato via lettera alle aziende con tutte le specifiche indicate precedentemente.

Dati gli scarsi risultati ottenuti si è deciso di passare a una seconda fase di raccolta dei dati tramite l'utilizzo del telefono.

PIANO DI COMPORTAMENTO

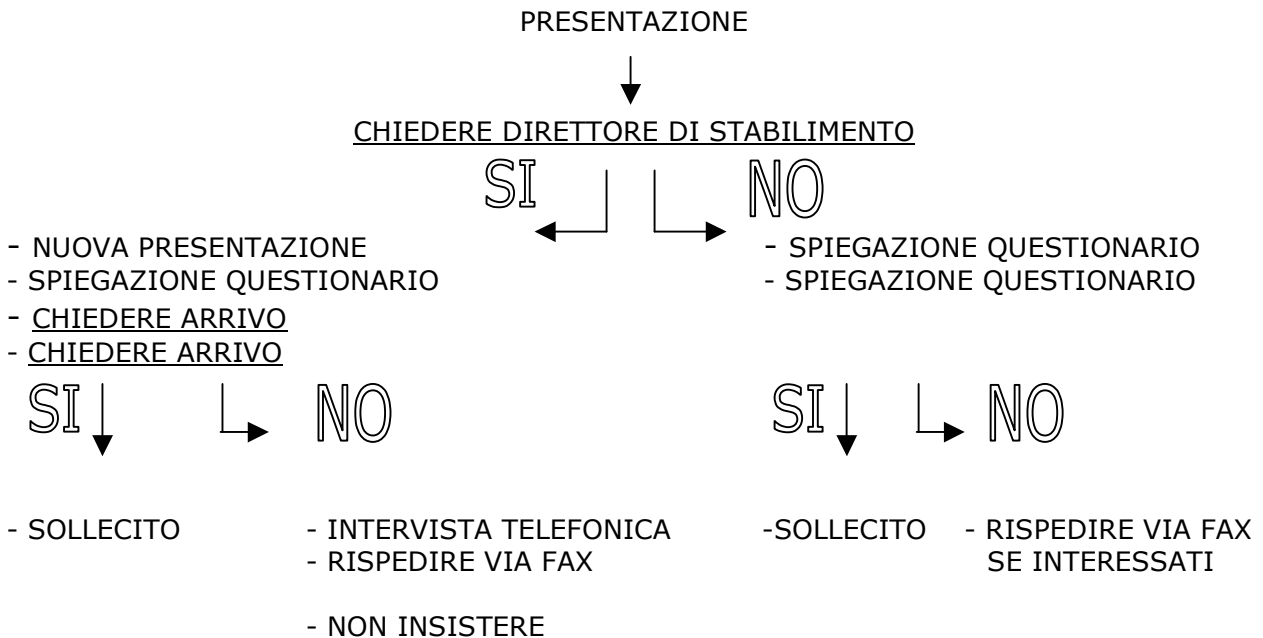


Figura 2.8. Piano di comportamento per le interviste telefoniche.

Dapprima c'è stata una fase di addestramento cercando di valutare quali potessero essere le alternative di risposte dei soggetti qualora si chiedessero informazioni a riguardo del questionario. Dalle riflessioni fatte è seguito il piano di comportamento presentato schematicamente nella Figura 2.8. Presentarsi come 'Ufficio marketing della Stefani SCM Group' e chiedere del direttore di stabilimento. Se il direttore viene passato ripresentarsi e chiedere informazioni in merito al questionario con breve spiegazione della sua funzione. Se il questionario è arrivato in azienda chiedere un semplice sollecito alla risposta mentre se non è arrivato si procede in tre modi diversi a seconda della disponibilità del soggetto: intervista telefonica, specificando che richiede poco tempo, rispedire il questionario alla 'cortese attenzione' e al numero di fax indicato dal soggetto oppure nel caso non fosse interessato non insistere.

Nel caso in cui non è possibile parlare con il direttore di stabilimento ma si è certi che il questionario è arrivato in azienda chiedere al soggetto di

sollecitare la risposta alla persona a cui è stato consegnato, mentre se non è arrivato rispedire il questionario alla 'cortese attenzione' e al numero di fax indicato sempre dal soggetto.

2.4.1 RISPOSTE TOTALI OTTENUTE

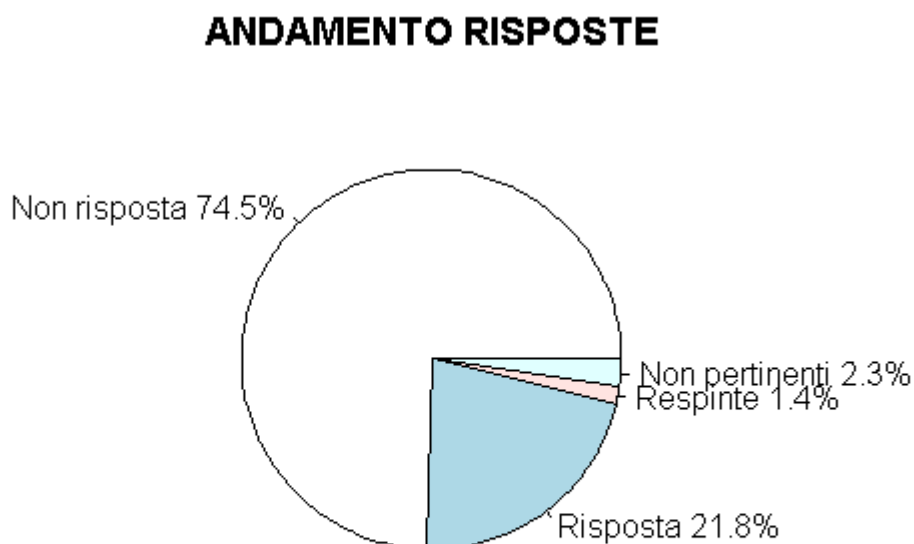


Figura 2.9. Grafico a torta del numero di risposte totali ottenute

Come si può evincere dal grafico precedente il numero di risposte è aumentato durante la seconda fase arrivando a 250. Procedendo con la raccolta dei questionari sono emerse alcune problematiche in merito al campione selezionato. Contattando direttamente alcune aziende si è visto che esse possono essere definire 'non pertinenti' per il genere di questionario somministrato in quanto o l'attività è cessata o in via di chiusura oppure trattano solamente la vendita di mobili, o prodotti simili, e non la produzione. Il numero totale di queste aziende che si è riusciti a contattare è di 27. Con il termine 'respinte' si intendono tutte quelle lettere, contenenti il questionario, che sono ritornate in azienda per i più svariati motivi: trasferimento del cliente, indirizzo inesistente, azienda chiusa, ecc. Il numero di lettere ritornate è stato in tutto di 16.

Il ritorno in azienda di alcuni questionari, l'impossibilità di contattare telefonicamente tutte le aziende, dato che dopo tre tentativi di chiamata molte ancora non rispondevano, il ritardo delle poste, testimoniato da numerose aziende contattate, fanno pensare che un discreto numero di componenti del campione non hanno risposto perché non contattato.

CONFRONTO TRA LE RISPOSTE E IL CAMPIONE

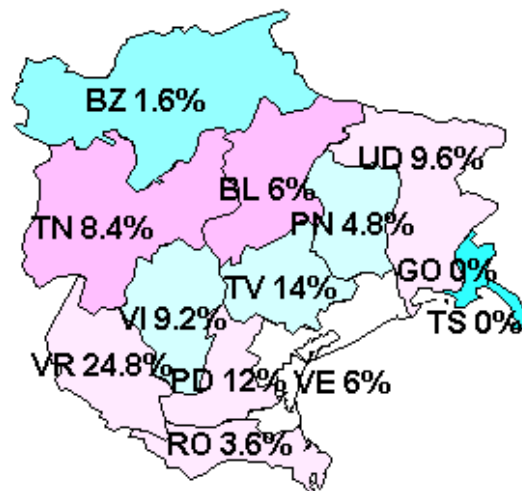


Figura 2.10. Confronto tra le risposte ottenute e il campione, percentuali delle province all'interno dei dati raccolti

CONFRONTO TRA IL CAMPIONE E LA POPOLAZIONE

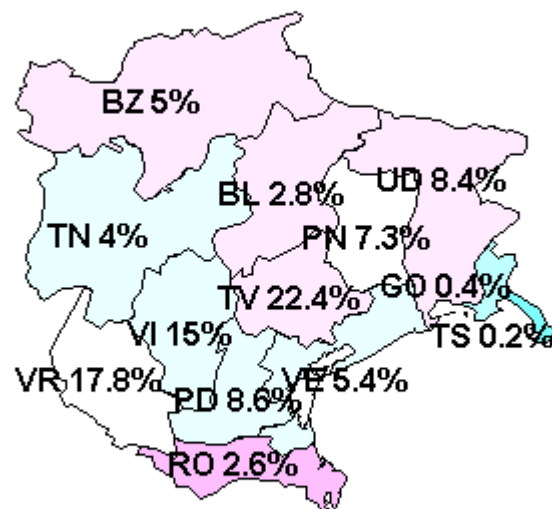


Figura 2.2. Presentata precedentemente a pagina 26.

La prima mappa della pagina precedente rappresenta il confronto tra la proporzione della provincia nelle risposte e nel campione formato. In particolare le percentuali riportate indicano come si distribuisce la numerosità delle province in base alle 250 risposte che si è riusciti a raccogliere. I colori indicano quanto queste percentuali si discostano da quelle del campione, le quali sono indicate nel secondo grafico.

Confrontando le mappe salta subito all'occhio che se la situazione si poteva considerare accettabile, per quanto riguarda la percentuale delle province, nel passaggio dalla popolazione al campione lo stesso non si può dire nel passaggio da campione a numero di risposte ricevute. Infatti, in questo caso, si ha una maggiore varietà di intensità di colori che sta ad indicare che la situazione non è proprio, proporzionalmente parlando, ideale. Solamente la provincia di Venezia rispetta le proporzioni, visto che è di colore bianco, ma bisogna tenere anche presente che nella formulazione del campione è stata sottorappresentata anche se non in maniera forte. Sottorappresentate in entrambe le situazioni sono state la provincia di Vicenza ma in maniera piuttosto forte le province di Gorizia e di Trento. Questo ha portato al drastico risultato di non ottenere nessuna risposta da aziende presenti in queste ultime due province. Discorso opposto riguarda le province di Belluno, Rovigo, Udine che da sovrarappresentate nel campione si trovano ancora più presenti nelle risposte. Alcune province si sono invertite le caratteristiche: nel campione le province di Padova e Trento erano in proporzione con la popolazione poco presenti ma nel numero di risposte addirittura troppo presenti. Discorso inverso invece per le province di Bolzano e di Treviso. Un caso particolare riguarda le province di Pordenone e di Verona le quali non avevano risentito del passaggio dalla popolazione al campione. Durante la raccolta delle risposte, però, si sono ottenute troppe poche risposte dalla prima provincia e troppe dalla seconda alterando le uniche proporzioni perfette che si erano verificate creando il campione.

Tutto questo stravolgimento delle proporzioni per provincia fa capire quanto sia difficile se non addirittura impossibile riuscire a mantenere un equilibrio passando dalla popolazione al campione al numero di risposte.

Un'altra variabile precedentemente utilizzata per esaminare lo scostamento tra la distribuzione della popolazione e quella del campione è stata il 'tipo di cliente'. Si verifica ora come questa variabile si comporta all'interno delle 250 risposte raccolte.

A	B	C	D
149	80	0	21

Tabella 2.3. Distribuzione della variabile tipo cliente all'interno delle risposte raccolte mediante il questionario.

Come si può verificare ai grafici a torta riportati di seguito nel passaggio dalla popolazione al campione e poi alle risposte c'è stato un progressivo aumento dei clienti B ma soprattutto dei clienti di tipo A comportando la diminuzione dei clienti di tipo D che nelle risposte ottenute sono in numero veramente marginale. È un altro sintomo che le risposte ottenute non rispecchiano pienamente le caratteristiche della popolazione.

DISTRIBUZIONE NELLE RISPOSTE

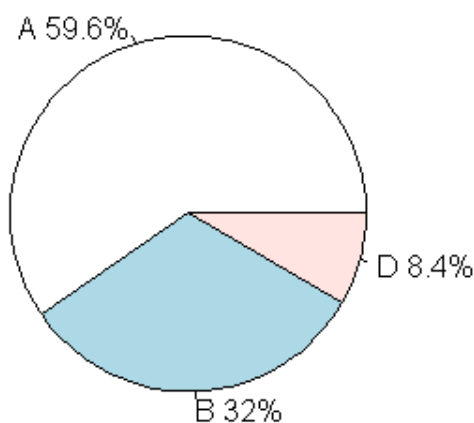
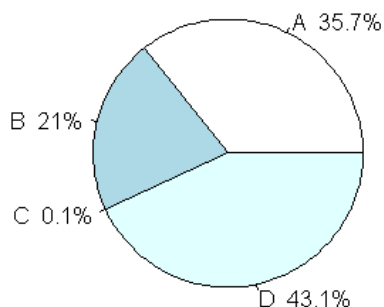


Figura 2.11. Grafico a torta che rappresenta la distribuzione del tipo di cliente nei dati raccolti con il questionario.

DISTRIBUZIONE NELLA POPOLAZIONE



DISTRIBUZIONE NEL CAMPIONE

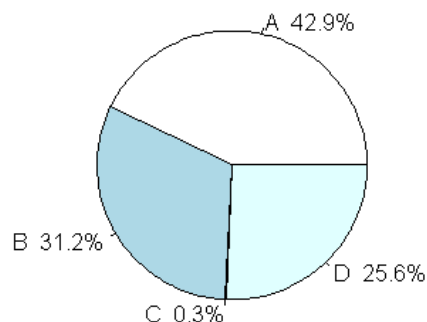


Figura 2.3 e 2.4 presentate precedentemente a pagina 28.

2.5 FATTORI CHE POSSONO COMPROMETTERE LA QUALITA' DEI DATI

Un disegno ottimale di campionamento, fatto che sicuramente non si è verificato in questo progetto, non è una condizione sufficiente affinché la ricerca condotta sia buona, poiché vi sono molteplici fattori che possono minacciare la qualità dei dati introducendo distorsioni nella loro validità e adeguatezza della misura di rappresentazione della popolazione oggetto di indagine.

I seguenti fattori possono abbassare la qualità dei dati:

- Cattiva qualità dello strumento di misura utilizzato.
- Contesto psicofisico inadeguato allo svolgimento del questionario.
- Effetti distorcenti sulla veridicità e attendibilità delle risposte date dai soggetti.
- Insufficiente professionalità dell'intervistatore nella cooperazione con i soggetti.
- Inadeguatezza della codifica dell'analisi dei dati raccolti.
- Alto tasso di non risposta totale cioè un elevato numero di soggetti non risponde per i più svariati motivi al questionario. Questo implica che il campione non è più nè rappresentativo nè casuale anche se in questo

particolare caso il campione è già distorto a causa del metodo di campionamento.

- Tasso di non risposta parziale cioè il numero di casi in cui non vi è risposta a una o più domande del questionario. Il numero di risposte mancanti può differire molto da una domanda all'altra per esempio a seconda del grado di intrusività della domanda

2.5.1 CAUSE DELLA NON RISPOSTA AL QUESTIONARIO

Utilizzando il mezzo telefonico si è riusciti ad avere un contatto diretto con i soggetti che dovevano venire intervistati. Questo non ha permesso solamente la compilazione di numerosi questionari tramite l'intervista telefonica ma anche di entrare nel pensiero delle persone e capire i motivi della non risposta al questionario.

- Molte aziende non hanno ricevuto il questionario.
- Questionario indirizzato alla 'direzione di stabilimento'. Nelle grandi aziende, in particolare, si perde facilmente se non c'è un riferimento a una determinata persona con nome e cognome.
- Riluttanza ai troppi questionari in genere.
- Troppa posta pubblicitaria, quindi il questionario non risalta.
- Non si possiedono prodotti STEFANI o IDM, quindi non si risponde
- Paura che con qualche stratagemma si sia costretti ad acquistare le macchine.

CAPITOLO 3

ANALISI

3.1 TABULAZIONE , PULITURA E CODIFICA DEI DATI

Per poter compiere una analisi, e quindi arrivare a quegli obiettivi per cui una determinata ricerca è stata attuata, non è sufficiente raccogliere i dati in modo grezzo ma essi devono essere messi in un determinato formato. Senza che si perda informazione nel passaggio tra dati raccolti e dati codificati, questa procedura è necessaria per poter utilizzare al meglio quegli strumenti che permettono il raggiungimento degli scopi previsti.

Di seguito verranno riportate tutte le variabili codificate che potranno essere utilizzate nella fase di analisi. In particolare, per ogni variabile, verrà specificato da dove è stata ricavata, il proprio significato, il tipo di codificata adottato e un 'riassunto' dei valori.

"provincia"

Variabile qualitativa che indica la provincia in cui è ubicata l'azienda. Questa variabile è stata ricavata dal database dell'azienda.

BL	BZ	PD	PN	RO	TN	TV	UD	VE	VI	VR
15	4	30	12	9	21	35	24	15	23	62

"conoscenza_prodotti"

Variabile qualitativa ricavata dalla domanda 1 del questionario e indica la conoscenza o meno dei prodotti a marchio STEFANI-IDM.

NO	SI
58	192

"acquisto_prodotti"

Variabile qualitativa ricavata dalla domanda 2 del questionario e indica se è avvenuto l'acquisto di prodotti a marchio STEFANI-IDM negli ultimi cinque

anni da parte delle aziende intervistate.

NO SI

235 15

"affidabilita"

Variabile qualitativa ricavata dalla domanda 2.a del questionario e indica la valutazione dell'affidabilità dei prodotti acquistati a marchio STEFANI-IDM negli ultimi cinque anni.

Questa è la prima variabile, tra quelle disponibili, che presenta la modalità 'NON PERT'. La presenza di questo valore in generale significa che la variabile in questione è derivata da una sottodomanda di una domanda filtro, alla quale l'azienda intervistata ha risposto in modo negativo. Per fare un esempio concreto una azienda che non ha acquistato prodotti STEFANI-IDM non può valutarne l'affidabilità. Infatti il numero di 'NON PERT' è pari al numero di volte che compare la modalità 'NO' nella variabile 'acquisto_prodotti'.

BUONA	SUFFICIENTE	INSUFFICIENTE	NON PERT
7	7	1	235

"servizio_assistenza"

Variabile qualitativa ricavata dalla domanda 2.b del questionario e indica la valutazione del servizio assistenza STEFANI-IDM riferito al prodotto acquistato negli ultimi cinque anni.

BUONO	SUFFICIENTE	INSUFFICIENTE	NON PERT
9	4	2	235

"acquisto_altri"

Variabile qualitativa ricavata dalla domanda 3 del questionario e indica se è avvenuto l'acquisto di bordatrici o squadrabordatrici da aziende concorrenti negli ultimi 5 anni.

NO SI

192 58

"grandi"

Variabile qualitativa ricavata dalla domanda 3.1 del questionario che indica quante macchine sono state acquistate da marchi concorrenti che possono essere definiti grandi (IMA, HOMAG, HOLZER, BIESSE) negli ultimi cinque anni.

0	1	2	3
236	10	3	1

"piccole"

Variabile qualitativa ricavata dalla domanda 3.1 del questionario che indica , quando ha valore 1, se sono state acquistate da marchi che possono essere definiti piccoli (FRAVOL, BI-MATIC, CELASCHI, ...) negli ultimi cinque anni.

0	1
235	15

"SCM"

Variabile qualitativa ricavata dalla domanda 3.1 del questionario che indica, quando ha valore 1, se sono state acquistate macchine dal marchio SCM negli ultimi cinque anni.

0	1
221	29

"affidab_c"

Variabile qualitativa ricavata dalla domanda 3.a del questionario e indica la valutazione dell'affidabilità dei prodotti acquistati da marchi concorrenti negli ultimi cinque anni.

ELEVATA	BUONA	SUFFICIENTE	INSUFFICIENTE	NON PERT	NA
9	37	8	2	192	2

"assistenza_c"

Variabile qualitativa ricavata dalla domanda 3.b del questionario e indica la valutazione del servizio assistenza offerto dai marchi concorrenti.

ELEVATO BUONO	SUFFICIENTE	INSUFFICIENTE	NON PERT	MAI USATO	NA
2	31	10	3	192	10
					2

"conoscenza_agenti"

Variabile qualitativa ricavata dalla domanda 4 del questionario e indica la conoscenza o meno degli agenti di zona STEFANI-IDM.

NO	SI	NA
176	72	2

"visite_mensili"

Variabile qualitativa ricavata dalla domanda 4.1 del questionario e indica il numero di visite mensili ricevute dagli agenti di zona STEFANI-IDM.

minore1	magu1min2	mag2umin3	magu4min5	NON CONOSCO	NA
31	18	6	2	176	17

Il valore 'minore1' indica che l'azienda ha risposto con una frazione mentre per esempio 'magu1min2' indica che le visite mensili sono maggiori di uno ma minori di due. Bisogna tenere presente che le categorie di una variabili devono essere tra loro mutuamente escludenti.

"mobili_ufficio"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo 'mobili per ufficio'.

0	1
227	23

"cucine"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo 'cucine'.

0	1
216	34

"mobili_bagno"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo 'cucine'.

0 1
225 25

"mobili_casa"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo 'mobili per la casa'.

0 1
166 84

"componenti_mobili"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo 'componenti per mobili'.

0 1
218 32

"altro"

Variabile qualitativa ricavata dalla domanda domanda 5 e indica se l'azienda intervistata si occupa del settore produttivo definito come 'altro'.

Le modalità di questa variabile hanno il seguente significato.

'0' non si occupa del settore altro.

'1' si occupa del settore altro senza nessuna specifica.

'F' falegnameria generica.

'MIS' mobili in stile.

'MSM' mobili su misura.

'S' serramenti.

0	1	F	MIS	MSM	S
102	49	11	7	35	46

"n_dipendenti"

Variabile quantitativa ricavata dalla domanda 6 e indica il numero di dipendenti impiegati all'interno dell'azienda intervistata.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA
1.00	3.00	6.00	17.57	15.00	350.00	8.00

"tipo_cliente"

Variabile qualitativa ricavata dal database dell'azienda e classifica, come indicato precedentemente, il cliente in base al suo rapporto con il Gruppo SCM.

A	B	D
149	80	21

"fiera"

Variabile qualitativa ricavata dal database dell'azienda e indica, quando ha valore 1, se c'è stato un contatto in fiera nel maggio 2002 tra l'azienda intervistata e il Gruppo SCM.

0	1
88	162

Dato l'elevato numero di variabili a disposizione, rapportato anche alla bassa numerosità dei dati pari a 250, spesso alcune variabili precedentemente menzionate non verranno utilizzate all'interno dei modelli statistici che si andranno a stimare. Visti i dati a disposizione c'è una forte correlazione tra variabili, in particolare tra quelle ricavate dalle domande 'filtro' del questionario e quelle a cui si accede rispondendo ad esse.

Di seguito sono riportate questi tipi di variabili in particolare quelle che verranno maggiormente utilizzate all'interno dei modelli elencando però anche quelle che derivano da esse e che saranno di conseguenza impiegate di meno o addirittura non verranno mai impiegate.

acquisto_prodotti → affidabilità, servizio_assistenza.

acquisto_altri → grandi, piccole, SCM, affidab_c, assistenza_c.

conoscenza_agenti → visite_mensili.

3.2 MODELLO PER LA PREVISIONE DEL NUMERO DI DIPENDENTI

Osservando attentamente i valori di alcune variabili si può notare che molte di loro contengono dei dati mancanti, codificati tramite il valore 'NA'.

Per risolvere il problema dei dati mancanti ci sono vari metodi: alcuni possono essere drastici, come togliere dai dati tutte quelle osservazioni che contengono valori mancanti con la conseguente perdita di tutte quelle informazioni che queste osservazioni possono dare, oppure più lievi come sostituire la media totale o la media a base a un determinato fattore agli 'NA'. Nelle analisi successive non verranno utilizzate quelle osservazioni che contengono dei valori mancanti all'interno di variabili qualitative. Per quanto riguarda le variabili quantitative è stato applicato un modello lineare con variabile risposta le variabili quantitative con dati mancanti in modo tale da poter predire e quindi sostituire i valori mancanti.

In questo particolare contesto di analisi l'unica variabile di tipo quantitativo presente nei dati è 'n_dipendenti' che indica il numero di dipendenti presenti in ogni azienda intervistata. Come si può vedere dalla precedente descrizione di questa variabile sono presenti in tutto otto valori mancanti. Si è quindi stimato un modello utilizzando tutte le osservazioni tranne quelle contenenti i valori mancanti e in particolare sfruttando quelle variabili che descrivono la struttura dell'azienda, e quindi oggettive, e non soggettive, come opinioni e giudizi su un particolare concetto che possono essere molto diversi anche tra aziende strutturalmente molto simili.

Le variabili esplicative utilizzate sono state le seguenti.

- provincia
- mobili_ufficio
- cucine

- mobili_bagno
- mobili_casa
- componenti_mobili
- altro
- tipo_cliente
- fiera

3.2.1 RICHIAMI SUL MODELLO LINEARE

Per risolvere il problema dei dati mancanti si dovrà adattare un modello con formulazione generale del tipo

$$y=f(x; \beta)+ \varepsilon \quad (3.1)$$

dove y rappresentano tutte le osservazioni della variabile 'n_dipendenti', senza considerare quelle osservazioni con valori mancanti, x indica le variabili esplicative che sono state utilizzate, quindi quelle elencate precedentemente, β indica i parametri di regressione che si andranno a stimare. Bisogna fare alcune considerazione a parte su ε che rappresenta il termine di errore. Ci sono alcune ipotesi che lo caratterizzano, le ipotesi del secondo ordine, secondo cui la sua media è nulla mentre la sua varianza è costante ma ignota pari a σ^2 .

Il criterio dei minimi quadrati porta alla stima di β tramite la minimizzazione della funzione obbiettivo, detta devianza, che indica lo scostamento tra i valori interpolati stimati tramite il modello e i valori realmente osservati.

$$D(\beta)=\sum\{y_i- f(x;\beta)\}^2 \quad (3.2)$$

Dividendo il valore della devianza per il numero di osservazioni meno il numero di parametri ($n-p$) si riesce a ottenere la stima della varianza dell'errore e da essa gli errori standard delle componenti di $\hat{\beta}$ necessari per poter attuare procedure inferenziali. In particolare supponendo per ε una

distribuzione normale, sempre con media pari a zero e varianza costante, si può calcolare il livello di significatività osservato del valore normalizzato della stima. Tutto ciò permette di capire quanto influente sia una variabile all'interno del modello ed eventualmente toglierla o attuare altri procedimenti per il miglioramento dell'adattamento del modello ai dati.

3.2.2 STIMA DEL MODELLO E PREVISIONE DEI DATI MANCANTI

Il primo modello formulato (M1) è quello che in gergo viene definito 'completo' in cui tutte le variabili che erano possibili utilizzare sono state inserite all'interno del modello.

MODELLO COMPLETO

```
M1<- lm (n_dipendenti~
provincia+mobili_ufficio+cucine+mobili_bagno+mobili_casa+
componenti_mobili+altro+tipo_cliente+fiera, data=dati_dipendenti)
summary(M1)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	28.826	9.990	2.885	0.004303	**
provinciaBZ	8.075	16.607	0.486	0.627282	
provinciaPD	-3.204	9.390	-0.341	0.733277	
provinciaPN	39.047	12.433	3.141	0.001920	**
provinciaRO	-1.299	12.806	-0.101	0.919296	
provinciaTN	1.389	10.107	0.137	0.890817	
provinciaTV	33.149	9.511	3.485	0.000594	***
provinciaUD	4.184	9.691	0.432	0.666353	
provinciaVE	17.691	10.872	1.627	0.105163	
provinciaVI	6.636	9.869	0.672	0.502033	
provinciaVR	-3.645	8.570	-0.425	0.671039	
mobili_ufficio1	17.119	7.124	2.403	0.017098	*
cucine1	8.713	6.336	1.375	0.170527	
mobili_bagno1	-10.047	7.266	-1.383	0.168161	
mobili_casa1	-9.616	5.303	-1.813	0.071136	.
componenti_mobili1	-23.240	6.481	-3.586	0.000414	***
altro1	-23.019	6.247	-3.685	0.000289	***
altroF	-29.990	10.185	-2.945	0.003584	**
altroMIS	-21.997	12.180	-1.806	0.072292	.
altroMSM	-16.945	7.411	-2.286	0.023188	*
altroS	-24.031	6.748	-3.561	0.000453	***

```

tipo_clienteB      -10.507      4.661   -2.254  0.025188 *
tipo_clienteD      2.137       8.444    0.253  0.800402
fiera1             3.589       5.190    0.691  0.490001
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 28.78 on 218 degrees of freedom
Multiple R-Squared: 0.3666,    Adjusted R-squared: 0.2997
F-statistic: 5.485 on 23 and 218 DF,  p-value: 2.73e-12

```

Nonostante molte variabili risultino significative il modello ottenuto non è soddisfacente. Il coefficiente di determinazione R^2 , che è un indice di adattamento del modello ai dati, è abbastanza contenuto. Risulta significativo comunque il test F che confronta il modello specificato con il modello a sola intercetta. Ovviamente in questo caso l'intercetta da sola non riesce a spiegare tutto il fenomeno.

MODELLO COMPLETO CON TRASFORMAZIONE DELLA VARIABILE RISPOSTA

Visti i risultati, sicuramente migliorabili, si è passati ad una fase successiva in cui la variabile risposta è data dalla trasformazione della variabile 'n_dipendenti' tramite la funzione logaritmo, sperando che questo porti alla stima di un miglior modello. La funzione logaritmo viene spesso utilizzata in quanto permette di correggere l'eteroschedasticità dei residui, supposti per ipotesi del secondo ordine omoschedastici. In questo particolare contesto di analisi, essendo la variabile risposta solamente a valori positivi, ha permesso di far 'lavorare' il modello lineare in un contesto più adatto alle loro caratteristiche cioè con il numero di dipendenti che ora può variare da $(-\infty, +\infty)$.

```

M2<- lm (log(n_dipendenti)~
provincia+mobili_ufficio+cucine+mobili_bagno+mobili_casa+
componenti_mobili+altro+tipo_cliente+fiera, data=dati_dipendenti)
summary(M2)

```

```

Coefficients:
(Intercept)      Estimate Std. Error t value Pr(>|t|)
                2.620962   0.299445   8.753 5.79e-16 ***

```

provinciaBZ	0.848692	0.497760	1.705	0.089615	.
provinciaPD	0.003577	0.281446	0.013	0.989871	
provinciaPN	1.366317	0.372663	3.666	0.000309	***
provinciaRO	0.334592	0.383834	0.872	0.384325	
provinciaTN	0.077391	0.302941	0.255	0.798605	
provinciaTV	1.013606	0.285068	3.556	0.000462	***
provinciaUD	0.421837	0.290457	1.452	0.147849	
provinciaVE	0.700572	0.325882	2.150	0.032674	*
provinciaVI	0.219264	0.295816	0.741	0.459360	
provinciaVR	-0.294983	0.256860	-1.148	0.252054	
mobili_ufficio1	0.236646	0.213526	1.108	0.268965	
cucinel	0.318107	0.189919	1.675	0.095376	.
mobili_bagno1	-0.273235	0.217794	-1.255	0.210983	
mobili_casal	-0.113069	0.158936	-0.711	0.477588	
componenti_mobili1	-0.546967	0.194246	-2.816	0.005311	**
altrol	-0.646362	0.187242	-3.452	0.000668	***
altroF	-1.005328	0.305262	-3.293	0.001155	**
altroMIS	0.026998	0.365059	0.074	0.941114	
altroMSM	-0.640460	0.222127	-2.883	0.004329	**
altroS	-0.676856	0.202257	-3.347	0.000964	***
tipo_clienteB	-0.958101	0.139714	-6.858	7.11e-11	***
tipo_clienteD	-0.430046	0.253088	-1.699	0.090709	.
fieral	-0.164848	0.155574	-1.060	0.290495	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8627 on 218 degrees of freedom
Multiple R-Squared: 0.5542, Adjusted R-squared: 0.5072
F-statistic: 11.78 on 23 and 218 DF, p-value: < 2.2e-16

Il modello M2 porta a risultati sicuramente migliori rispetto al precedente, dato l'aumentare della significatività di alcune variabili. Alcune variabili, però, non sono influenti quindi tramite una selezione all'indietro si è cercato di stimare quel modello che utilizza solamente quelle significative e quindi occupano un ruolo importante per la previsione del numero di dipendenti.

MODELLO RIDOTTO CON TRASFORMAZIONE DELLA VARIABILE RISPOSTA

```
M3<- lm (log(n_dipendenti)~
provincia+componenti_mobili+altro+tipo_cliente,
data=dati_dipendenti)
summary(M3)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.4258	0.2520	9.626	< 2e-16	***
provinciaBZ	0.8787	0.4961	1.771	0.077892	.
provinciaPD	0.0349	0.2799	0.125	0.900892	
provinciaPN	1.5642	0.3562	4.392	1.74e-05	***
provinciaRO	0.4247	0.3800	1.118	0.264922	
provinciaTN	0.1251	0.3000	0.417	0.677153	
provinciaTV	1.1353	0.2803	4.050	7.07e-05	***
provinciaUD	0.4525	0.2884	1.569	0.118056	
provinciaVE	0.8333	0.3203	2.602	0.009894	**
provinciaVI	0.2926	0.2939	0.996	0.320410	
provinciaVR	-0.2267	0.2541	-0.892	0.373254	
componenti_mobili1	-0.4657	0.1819	-2.561	0.011100	*
altrol	-0.6021	0.1688	-3.566	0.000443	***
altroF	-0.9832	0.2884	-3.410	0.000771	***
altroMIS	0.1059	0.3462	0.306	0.760065	
altroMSM	-0.6292	0.1922	-3.273	0.001232	**
altroS	-0.6775	0.1754	-3.862	0.000148	***
tipo_clienteB	-1.0071	0.1359	-7.412	2.55e-12	***
tipo_clienteD	-0.3586	0.2343	-1.531	0.127216	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8647 on 223 degrees of freedom
Multiple R-Squared: 0.5419, Adjusted R-squared: 0.505
F-statistic: 14.66 on 18 and 223 DF, p-value: < 2.2e-16

Tramite la selezione delle variabili si è arrivati a determinare il cosiddetto 'modello ridotto' che contiene solamente quelle variabili che veramente sono significative. Risulta molto importante constatare che non ci sono sostanziali differenze tra il modello completo e il modello ridotto con variabile risposta trasformata. In particolare la devianza, il coefficiente di determinazione e il test F hanno valori pressoché identici. Guardando i test sulla significatività delle variabili si può notare che i fattori che maggiormente influenzano il numero di dipendenti sono la provincia, il settore produttivo 'componenti per mobili', il settore produttivo 'altro' e il tipo di cliente. Queste indicazioni potrebbero essere molto utili qualora risultasse significativa la variabile 'n_dipendenti' all'interno di altri modelli che si andranno a stimare in seguito.

Previsione del logaritmo degli otto valori mancanti in 'n_dipendenti'.

log(n dipendenti)

1	2	3	4	5	6	7	8
2.55	1.56	1.91	1.49	1.84	1.84	3.98	2.38

Tramite la funzione esponenziale le previsioni fatte vengono trasportate da scala logaritmica alla scala originaria. I valori sottostanti sono stati immessi all'interno dei dati e sostituiti nelle osservazioni in cui erano presenti valori mancanti.

n dipendenti

1	2	3	4	5	6	7	8
12.85	4.80	6.80	4.48	6.29	6.29	54.05	10.85

La variabile 'n_dipendenti' che è stata usata all'interno delle analisi ora ha le seguenti caratteristiche.

`summary(n_dipendenti)`

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	3.00	6.00	17.43	15.00	350.00

3.3 DOMANDA 1

1.a: Quantificare in che percentuale il marchio STEFANI-IDM è conosciuto.

1.b: Individuare quali sono i fattori che influiscono sulla conoscenza del marchio.

Lo scopo di questa prima analisi non è di tipo predittivo ma è quello di verificare quali sono le variabili, tra quelle che si hanno a disposizione, che maggiormente influenzano la conoscenza del marchio STEFANI-IDM. Per raggiungere questo scopo si sono applicati ai dati dei modelli di classificazione i quali, ognuno con le proprie specifiche, sono riusciti ad indicare quali sono i maggiori fattori che determinano la condizione specificata.

La variabile risposta, in questo primo caso, è 'conoscenza_prodotti'. Le variabili che si sono utilizzate non sono state tutte le altre a disposizione ma è stata necessaria una prima selezione intuitiva. L'acquisto dei prodotti e la conoscenza degli agenti di zona STEFANI-IDM sono ovvie conseguenze della conoscenza dei prodotti, per questo le variabili che codificano questi concetti non sono state utilizzate all'interno dell'analisi.

Come operazione preliminare si è verificato, tramite un grafico a torta, la percentuale di conoscenza del marchio, che comunque risulta già essere una informazione importante.

Conoscenza marchio STEFANI-IDM

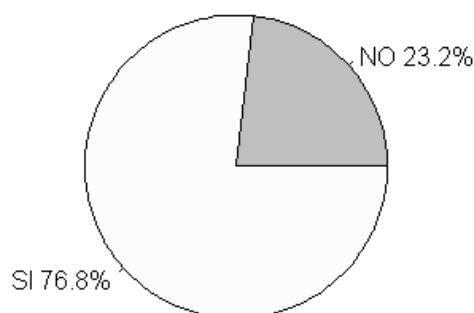


Figura 3.1. Grafico preliminare sulla conoscenza dei prodotti STEFANI-IDM.

La percentuale di aziende intervistate che conosce i prodotti a marchio STEFANI-IDM è piuttosto elevata.

Di seguito sono riportate le analisi con cui si sono andati a verificare quali sono i fattori che principalmente incidono sulla conoscenza o sulla non conoscenza dei prodotti.

3.3.1 CLASSIFICAZIONE MEDIANTE IL MODELLO DI REGRESSIONE LOGISTICO

3.3.1.a RICHIAMI SULLA REGRESSIONE LOGISTICA

La regressione logistica permette di mettere in relazione la probabilità Π che un certo evento si verifichi con un insieme di variabili esplicative. Per tentare di rispondere alla richiesta posta alla domanda 1 la probabilità di interesse utilizzata è stata quella della conoscenza, e la sua complementare non conoscenza, dei prodotti a marchio STEFANI-IDM.

Posto con $\eta(x)$ una combinazione di variabili esplicative lineari nei parametri come in formula generale seguente

$$\eta(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_2 x_2 \quad (3.3)$$

in cui le x_i indicano le variabili a disposizione mentre i β_i i coefficienti di determinazione che si devono stimare tramite il modello.

Definita la funzione logistica

$$L(\eta) = e^\eta / (1 + e^\eta) \quad (3.4)$$

il modello di regressione logistica indica la relazione

$$\Pi(x) = L(\eta(x)) \quad (3.5)$$

attraverso cui si evince che la probabilità dell'evento di interesse, in questo primo caso la conoscenza dei prodotti a marchio STEFANI-IDM, dipende

dall'insieme delle variabili esplicative x utilizzate all'interno del modello tramite il predittore lineare $\eta(x)$.

3.3.1.b DETERMINAZIONE DEL MODELLO LOGISTICO

MODELLO LOGISTICO COMPLETO

Il primo modello di classificazione che si è utilizzato è stato quello logistico completo, con tutte le variabili che erano ragionevolmente possibili utilizzare per questo primo contesto di analisi. Le variabili che sono state utilizzate in questa prima fase sono elencate all'interno del modello sottostante.

```
Log1<- glm (conoscenza_prodotti ~
provincia+acquisto_altri+ mobili_ufficio+cucine+mobili_bagno+
mobili_casa+componenti_mobili+altro+n_dipendenti+tipo_cliente+
fiera, family=binomial, data=dati)
summary(Log1)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.91052	0.95831	0.950	0.34205	
provinciaBZ	0.66515	1.36381	0.488	0.62575	
provinciaPD	0.66840	0.80723	0.828	0.40766	
provinciaPN	0.54594	1.39075	0.393	0.69465	
provinciaRO	-0.49819	0.96071	-0.519	0.60406	
provinciaTN	-0.71765	0.78776	-0.911	0.36230	
provinciaTV	1.09780	1.02341	1.073	0.28341	
provinciaUD	-0.01469	0.77649	-0.019	0.98490	
provinciaVE	-0.71714	0.90463	-0.793	0.42793	
provinciaVI	0.12030	0.81521	0.148	0.88268	
provinciaVR	0.71977	0.69269	1.039	0.29876	
acquisto_altriSI	1.60661	0.56783	2.829	0.00466	**
mobili_ufficio1	-0.44356	0.71770	-0.618	0.53655	
cucinel	-0.40803	0.65169	-0.626	0.53124	
mobili_bagno1	-0.21185	0.67265	-0.315	0.75279	
mobili_casal	-0.19290	0.56028	-0.344	0.73062	
componenti_mobili1	0.73414	0.73145	1.004	0.31553	
altro1	-0.36939	0.59413	-0.622	0.53412	
altroF	-0.28966	0.93565	-0.310	0.75688	
altroMIS	-1.56280	1.01828	-1.535	0.12485	
altroMSM	-0.04613	0.72516	-0.064	0.94928	
altroS	0.44244	0.65825	0.672	0.50150	
n_dipendenti	0.03330	0.01848	1.802	0.07157	.
tipo_clienteB	-0.03767	0.41095	-0.092	0.92696	

```

tipo_clienteD      -1.36208    0.70451  -1.933  0.05319 .
fieral            -0.37094    0.54984  -0.675  0.49991
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(Dispersion parameter for binomial family taken to be 1)

```

Null deviance: 270.84 on 249 degrees of freedom
Residual deviance: 221.87 on 224 degrees of freedom
AIC: 273.87

```

Number of Fisher Scoring iterations: 6

MODELLO LOGISTICO RIDOTTO

Dato che diverse variabili, tramite il test di Wald, non risultano significative si è passati a una selezione all'indietro delle variabili determinando in questo modo il modello logistico finale costituito solamente da quelle variabili che risultano significative. Saranno proprio queste variabili a definire quali fattori influiscono sulla conoscenza dei prodotti a marchio STEFANI-IDM.

```

Log2<- glm (conoscenza_prodotti ~
acquisto_altri+n_dipendenti+ tipo_cliente, family=binomial,
data=dati)
summary(Log2)

```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.81235	0.29628	2.742	0.00611	**
acquisto_altriSI	1.35672	0.50596	2.681	0.00733	**
n_dipendenti	0.03588	0.01575	2.278	0.02275	*
tipo_clienteB	-0.33845	0.36471	-0.928	0.35341	
tipo_clienteD	-1.10766	0.51292	-2.160	0.03081	*

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(Dispersion parameter for binomial family taken to be 1)

```

Null deviance: 270.84 on 249 degrees of freedom
Residual deviance: 240.43 on 245 degrees of freedom
AIC: 250.43

```

Number of Fisher Scoring iterations: 6

Dall' utilizzo del modello di regressione i fattori che maggiormente influiscono la conoscenza dei prodotti a marchio STEFANI-IDM sono l'acquisto di altri prodotti da parte dell'azienda intervistata e il numero di dipendenti impiegati. Nel modello è stata utilizzata la variabile 'n_dipendenti' e non il suo logaritmo in quanto risulta più significativa. Importante è anche il tipo di cliente secondo la classificazione all'interno del Gruppo SCM.

3.3.2 ALBERO DI CLASSIFICAZIONE

3.3.2.a RICHIAMI SULL'ALBERO DI CLASSIFICAZIONE

La struttura dell'albero, sia di classificazione che di regressione, è basata sull'idea di approssimazione di una funzione attraverso dei gradini cioè con funzioni costanti a tratti in un certo numero di intervalli. Questo può ovviamente avvenire anche quando si ha la variabile esplicativa $y=f(x)$ con x insieme di variabili esplicative maggiori di uno. Questo tipo di modello è rappresentabile attraverso un albero binario costituiti da nodi, rami e foglie. Ogni nodo rappresenta una affermazione di tipo logico riferita a una delle qualsiasi variabili esplicative x_i utilizzate e indica una suddivisione della variabile tramite intervalli a cui è associato un valore che compare nell'albero.

Partendo dalla radice dell'albero e scendendo verso il basso, a seconda delle affermazioni ottenute passando per il nodo, si scorrono tutti i vari rami fino ad arrivare alle foglie che indicano il valore della funzione approssimata.

Ci sono vari passaggi da eseguire per arrivare alla creazione dell'albero definitivo. Per prima cosa c'è la fase di crescita che viene effettuata solamente con un certo campione dei dati a disposizione. Essa consiste in una suddivisione iterativa dell'albero in modo tale da far diminuire la devianza facendo in modo che il numero delle foglie finali sia pari ad n_{c1} cioè la numerosità del campione di crescita. In questo modo le foglie finali conterranno solamente una osservazione.

La fase successiva è detta di 'potatura' dell'albero, che viene effettuata su un campione di dati diverso da quello si stima, dato che un albero costruito in questo modo è paragonabile a una funzione che passa per tutti i punti

osservati, che quindi non è di nessuna utilità. Lo scopo di un modello predittivo non è quello di adattarsi perfettamente ai dati, con cui è stato stimato, ma interpretare un fenomeno di fondo che possa andare bene per quel tipo di dati in generale. In particolare in questa fase verranno tolti tutti quei nodi, e i conseguenti rami e foglie, il cui inutilizzo non porta all'aumento eccessivo della devianza o per gli alberi di classificazione dell'errore di classificazione.

Utilizzando questo metodo di classificazioni non ci sono procedure inferenziali che permettano una selezione delle variabili significative. Comunque c'è una selezione automatica, tramite l'operazione di potatura, delle variabili solamente che non è possibile stilare un ordine di importanza.

3.3.2.b DETERMINAZIONE DELL'ALBERO DI CLASSIFICAZIONE

STIMA DELL'ALBERO COMPLETO – FASE DI CRESCITA DELL'ALBERO

Questo albero è stato stimato con le stesse variabili utilizzato nel modello completo di regressione logistica e ovviamente la variabile risposta è sempre la conoscenza dei prodotti a marchio STEFANI-IDM

```
set.seed(91110)
T1<- tree (conoscenza_prodotti ~
provincia+acquisto_altri+mobili_ufficio+cucine+mobili_bagno+mobil
i_casa+componenti_mobili+ altro+n_dipendenti+tipo_cliente+fiera,
data=dati)
plot(T1)
text(T1)
```

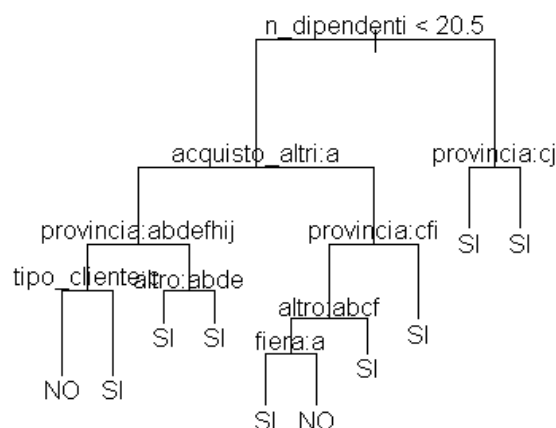


Figura 3.2. Albero di classificazione completo.

ALBERO DEFINITIVO – FASE DI POTATURA

La fase di potatura dell'albero non è avvenuta in modo usuale ma è stata sfruttata la 'cross validation'. Questo procedimento consiste nel dividere i dati in k campioni e usarne k-1 per la stima e 1 per la verifica. In questo ambito di analisi tutti gli alberi che sono stati stimati hanno k pari a dieci. È stato necessario l'utilizzo di questo procedimento data la bassa numerosità campionaria e quindi la necessità di stabilizzare i risultati ottenuti dato che, con il metodo usuale, si ottenevano risultati diversi ogni volta che un albero veniva adattato ai dati.

```
T2<-cv.tree(T1, ,prune.tree)
plot(T2)
```

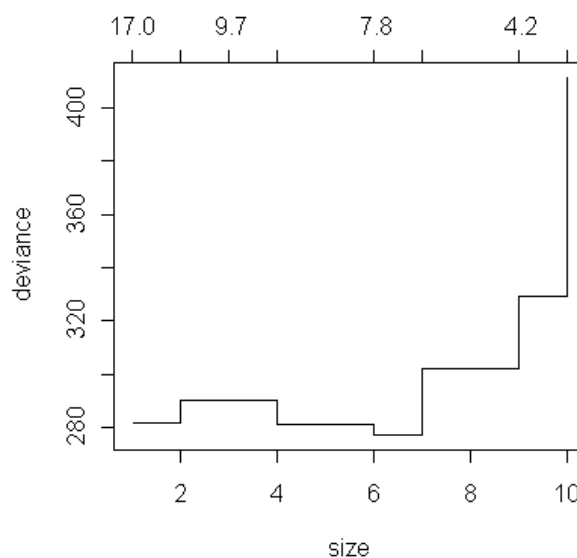


Figura 3.3. Devianza per la grandezza dell'albero.

La figura 3.3 è riferita alla fase di potatura e riporta il valore della devianza a seconda della grandezza dell'albero. Quindi è stata scelta la grandezza dell'albero che porta al minor valore di devianza ed è stato riportato l'albero così ricavato nella figura 3.4.

```
T3<-cv.tree(T1,best=T2$size[T2$dev==min(T2$dev)])
plot(T3)
text(T3)
```

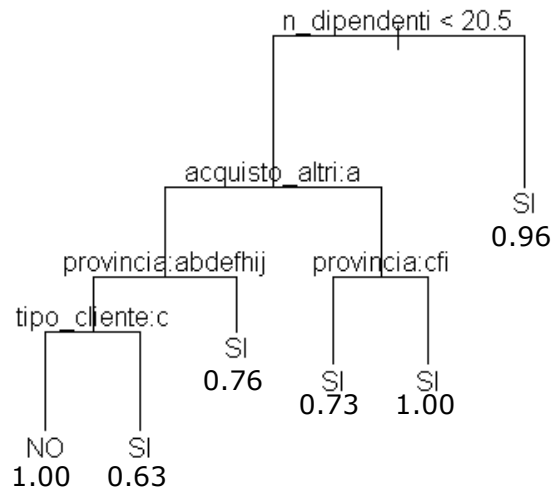



Figura 3.4. Albero potato e definitivo e indicazione della probabilità associata alle foglie

L'albero di classificazione indica, senza dare un ordine preciso di importanza, che la conoscenza dei prodotti a marchio STEFANI-IDM dipende dal numero di dipendenti dell'azienda intervistata, dall'aver acquistato prodotti da marchi concorrenti, dalla provincia in cui è ubicata l'azienda e dal tipo di cliente. I numeri presenti nell'albero potato indicano le stime della probabilità associata alla classe che, pur non cambiando, spesso viene associata in modo diverso da foglia a foglia.

3.3.3 ANALISI DISCRIMINANTE LINEARE

3.3.3.a RICHIAMI ALL'ANALISI DISCRIMINANTE LINEARE

L'analisi discriminante è l'unica tecnica di pura classificazione che è stata utilizzata in questo particolare ambito di analisi. Essa viene condotta per definire una modalità di assegnazione dei casi a differenti gruppi, in funzione di k variabili fra loro correlate, che per semplicità indicheremo con X . Solitamente l'interesse è rivolto a definire un modello che consenta di assegnare un nuovo caso ad un gruppo. In questo caso, tramite opportuni procedimenti, è stata utilizzata per determinare quelle variabili che

maggiormente influenzano un particolare fenomeno, costituito per ogni analisi effettuata solamente da due classi.

Facendo un discorso in termini generali si ha che la popolazione complessiva è composta da k subpopolazioni, aventi funzioni di densità di probabilità rispettivamente $p_1(x), \dots, p_k(x)$, per la distribuzione di X , e con peso Π_1, \dots, Π_k rispetto al totale della popolazione. La densità per la popolazione complessiva è data da

$$p(x) = \sum \Pi_k p_k(x) \quad (3.6)$$

A priori la probabilità che un soggetto non ancora classificato appartenga a una delle k subpopolazioni è data da Π_k . Se per quel soggetto è noto il valore assunto dalle variabili contenute in X , allora tramite il teorema di Bayes si può determinare la probabilità che quel soggetto appartenga al k -mo gruppo. Di conseguenza si può attuare un confronto tra classi attraverso le rispettive funzioni discriminanti date da

$$d_k(x_0) = \log \Pi_k + \log p_k(x_0) \quad (3.7)$$

in cui x_0 rappresenta i valori delle variabili esplicative che precedentemente sono state indicate con X . Quel valore di k che massimizza d_k rappresenta il gruppo a cui assegnare il nuovo soggetto.

Il primo elemento della formula (3.7) è semplicemente stimabile con $\Pi_k = n_k/n$ in cui n_k indica la numerosità della k -ma classe mentre n la numerosità totale delle osservazioni. Per quanto riguarda la densità $p_k(x)$, una delle ipotesi che si può fare, è che sia normale multipla con parametri dipendenti da k . Supposte tutte le matrici di varianza uguali alla stessa σ^2 la funzione discriminante risultante è una funzione lineare in x da cui il nome analisi discriminante lineare.

Si può arrivare alla determinazione della funzione lineare senza ricorrere all'ipotesi di normalità multipla ma utilizzando solamente le ipotesi del secondo ordine. Questo ha permesso di utilizzare questa tecnica di classificazione in questo particolare contesto, in cui la maggior parte delle

variabili esplicative che sono state utilizzate sono di tipo qualitativo e quindi ben lontane dall'assunto della normalità multipla.

3.3.b DETERMINAZIONE DEL MODELLO TRAMITE ANALISI DISCRIMINANTE LINEARE

MODELLO COMPLETO

Anche in questo caso è stato utilizzato la cross validation sempre a causa della bassa numerosità dei dati, pari a 250 osservazioni. In questo caso questo metodo viene utilizzato per stimare, tramite n-1 osservazioni, il modello e tramite l'oggetto 'class' del comando lda viene riportata la classe stimata dell'osservazione rimanente.

Di seguito viene riportato il modello completo stimato con la relativa 'tabella sommario' che riporta come vengono classificate tutte le osservazioni possedute in particolare la percentuale di errore totale e le percentuali di falsi negativi e di falsi positivi.

```
ldalcv<- lda (conoscenza_prodotti ~  
provincia+acquisto_altri+mobili_ufficio+cucine+mobili_bagno+  
mobili_casa+componenti_mobili+altro+n_dipendenti+  
tipo_cliente+fiera, CV=TRUE, data=dati)
```

```
tabella.sommario(ldalcv$class,dati$conoscenza_prodotti)
```

osserv

```
pred NO  SI  
NO    9  18  
SI   49 174
```

```
errore totale: 0.268
```

```
falsi positivi/negativi: 0.219 0.666
```

Il passaggio successivo alla determinazione di questo modello sarà quello di togliere tutte quelle variabili che non risultano essere legate con la variabile risposta. In questo modo nel modello finale resteranno solo quelle variabili che realmente influenzano il fenomeno oggetto di studio che in questo caso sarà la conoscenza dei prodotti a marchio STEFANI-IDM. Dato che per

questo particolare procedimento non c'è nessuna metodo inferenziale per la selezione delle variabili, il percorso seguito per la determinazione delle variabili più influenti è stato il seguente. Per prime venivano tolte quelle che non facevano modificare l'errore totale e le percentuali di falsi negativi e falsi positivi. Il passo successivo era quello di togliere una unica variabile che faceva diminuire maggiormente una delle tre percentuali di errore senza andare ad intaccare in modo considerevole il valore delle altre. Si ristimava il modello con le variabili rimaste e il ciclo di operazioni ricominciava fino a che non si arrivava a trovare il modello ridotto in cui il numero di variabili era decisamente minore mentre le percentuali di errore dovevano diminuire o al massimo aumentare in modo minimo rispetto al modello completo stimato in precedenza.

MODELLO RIDOTTO

```
lda2cv1<-lda(conoscenza_prodotti ~  
acquisto_altri+n_dipendenti+ tipo_cliente, CV=TRUE, data=dati)
```

```
tabella.sommario(lda2cv1$class,dati$conoscenza_prodotti)
```

```
  osserv
```

```
pred NO  SI
```

```
  NO  10   9
```

```
  SI  48 183
```

```
errore totale:  0.228
```

```
falsi positivi/negativi:  0.207 0.473
```

Anche attraverso l'utilizzo dell'analisi discriminante lineare si arriva alla conclusione che le variabili che maggiormente influenzano la conoscenza dei prodotti sono l'acquistato prodotti da altri marchi, il numero di dipendenti impiegati e il tipo di cliente.

3.3.4 CONSIDERAZIONI DOMANDA 1

Tutti e tre i metodi utilizzati indicano che le variabili che maggiormente influenzano la conoscenza dei prodotti a marchio STEFANI-IDM sono l'acquisto di bordatrici e squadrabordatri da altri marchi, il numero di dipendenti e il tipo di cliente. Di seguito è stato approfondito, tramite anche analisi grafiche, il legame tra queste variabili e la conoscenza dei prodotti tentando di estrarre informazioni di carattere pratico che possono essere utili per l'azienda.

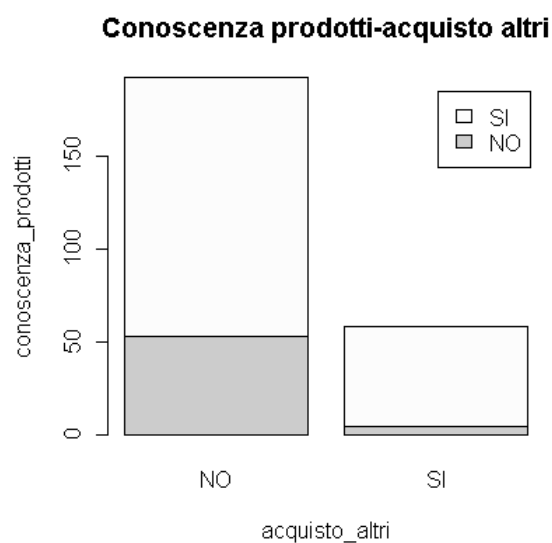


Figura 3.5. Grafico per la relazione tra la conoscenza dei prodotti a marchio STEFANI-IDM e l'acquisto di altri prodotti.

La figura precedente riporta l'acquisto di prodotti non a marchio STEFANI-IDM stratificato per la conoscenza dei prodotti. Si nota che, in proporzione, le aziende che negli ultimi cinque anni hanno acquistato prodotti da aziende concorrenti conoscono maggiormente l'azienda STEFANI-IDM rispetto a chi non ha comprato questo genere di prodotto. Una prima considerazione importante da fare è che, dato che l'acquisto dei prodotti è riferito nel questionario agli ultimi cinque anni, non è automaticamente detto che al momento dell'acquisto del prodotto dalla concorrenza l'azienda conoscesse già STEFANI-IDM ma magari la conoscenza è avvenuta in un periodo successivo.

Questa relazione tra le variabili 'acquisto_altri' e 'acquisto_prodotti' è probabilmente spiegabile dal fatto che le aziende che necessitano della lavorazione della bordatura, prima di fare una spesa piuttosto consistente, guardano sul mercato cosa i vari marchi offrono e in questo modo perlomeno vengono a conoscenza delle varie aziende. Chi non utilizza questo tipo di macchine probabilmente non conosce neanche le cosiddette possibili 'marche' proprio perché non è interessato.

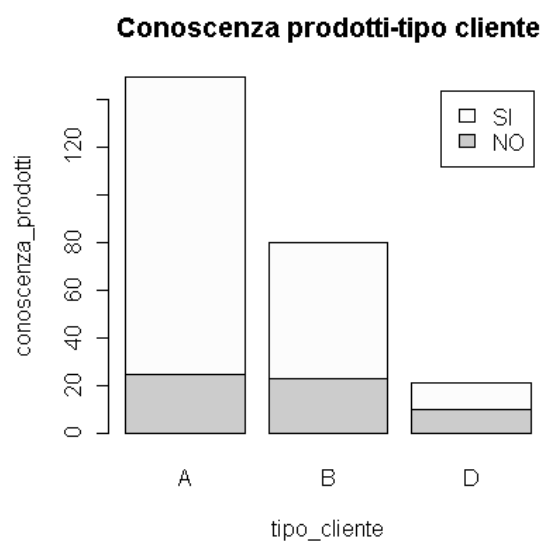


Figura 3.6. Grafico per la relazione tra la conoscenza dei prodotti a marchio STEFANI-IDM e il tipo di cliente.

Il grafico precedente riporta il tipo di cliente stratificato per la conoscenza dei prodotti a marchio STEFANI-IDM. Il risultato è piuttosto scontato in quanto i clienti di tipo 'D' del gruppo SCM, cioè clienti potenziali senza nessun contatto, sono quelli che in proporzione conoscono il marchio STEFANI-IDM di meno. Quindi l'aver acquistato un prodotto del Gruppo SCM, o direttamente in sede di un marchio o tramite un agente, fa aumentare la conoscenza dei prodotti STEFANI-IDM.

Conoscenza prodotti-numero dipendenti

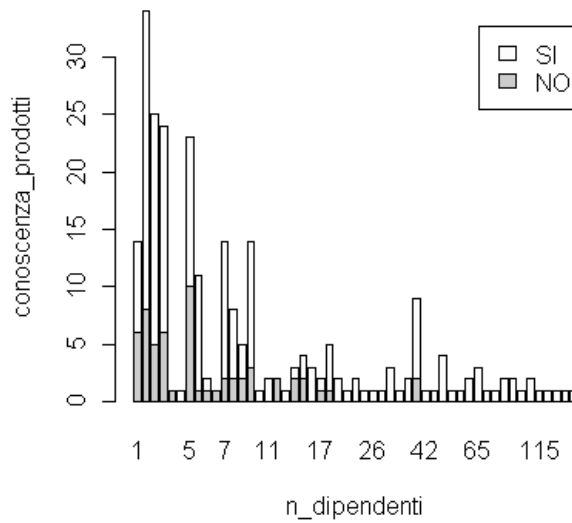


Figura 3.7. Grafico per la relazione tra numero dipendenti e conoscenza prodotti STEFANI-IDM

Il grafico precedente riporta il numero di dipendenti stratificato per la conoscenza dei prodotti a marchio STEFANI-IDM. Si nota molto bene il passaggio della 'non conoscenza' dei prodotti quando il numero di dipendenti è basso al progressivo aumento della conoscenza dei prodotti quando il numero di dipendenti si eleva. Questo è spiegabile dal fatto che all'aumentare del numero di dipendenti aumenta la quantità di produzione e quindi di lavorazioni che comporta la necessità di acquistare prodotti tecnologicamente avanzati. L'acquisto, come descritto precedentemente, costringe l'azienda a guardarsi attorno e quindi comporta almeno la conoscenza delle aziende che offrono una certa tecnologia. Inoltre teoricamente maggior produzione dovrebbe coincidere con maggiori vendite e maggiori guadagni creando quella disponibilità economica per poter acquistare macchine di un certo livello.

3.4 DOMANDA 2

Cause che comportano l'acquisto di macchine dalla concorrenza e non STEFANI-IDM.

Per rispondere alla domanda precedente è stato necessario dividere il problema attraverso due distinti procedimenti di analisi. Tramite il primo si è tentato di ricavare i fattori che influiscono sull'acquisto di prodotti da marchi concorrenti da parte delle aziende. Il secondo procedimento è consistito nell'individuare le variabili significative per l'acquisto di prodotti a marchio STEFANI-IDM. Sono state confrontate le caratteristiche dei due diversi tipi di acquisto traendone delle considerazioni interessanti.

3.4.1 PRIMA PARTE – ACQUISTO PRODOTTI NON STEFANI-IDM

Acquisto altri prodotti

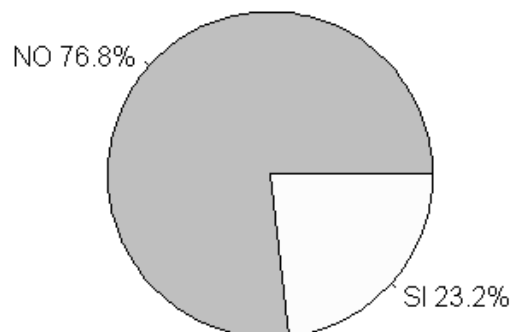


Figura 3.8. Percentuale di acquisto di prodotti non a marchio STEFANI-IDM.

La percentuale di aziende intervistate che ha acquistato prodotti dalla concorrenza non è molto elevata. Questa prima analisi non è significativa senza un confronto diretto con i risultati della STEFANI-IDM quindi ulteriori considerazioni verranno fatte in seguito.

Tipo prodotti

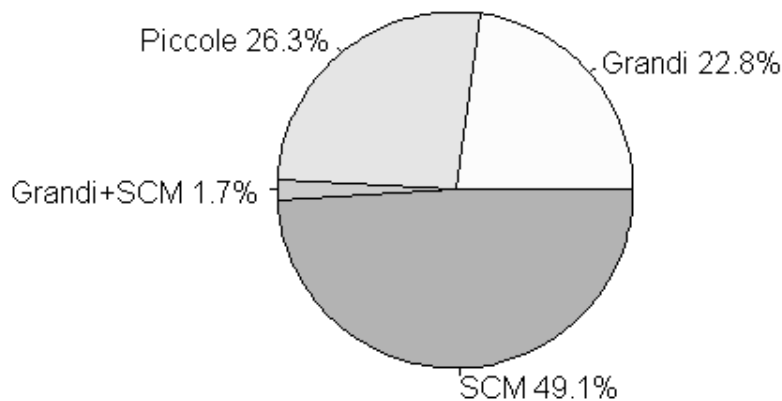


Figura 3.9. Tipi di macchine acquistate.

Gli acquisti avvenuti riguardano per la maggior parte macchine piccole, visto che anche il marchio SCM produce macchine piccole. Bassa è la probabilità di acquistare macchine di dimensioni diverse.

3.4.1.a MODELLO DI CLASSIFICAZIONE LOGISTICO

La stima del modello completo è stato il primo passo da cui si è partiti per determinare quelle variabili che influenzano l'acquisto di prodotti non a marchio STEFANI-IDM. Non ci sono state controindicazioni particolari su quali variabili utilizzare come esplicative, tranne ovviamente quelle fortemente correlate con la variabile risposta che sono 'grandi', 'piccole', 'SCM', 'affidab_c' e 'assistenza_c'.

MODELLO LOGISTICO COMPLETO

```
Logdom2 <- glm(acquisto_altri ~  
provincia+conoscenza_prodotti+acquisto_prodotti+
```

```

conoscenza_agenti+mobili_ufficio+cucine+mobili_bagno+mobili_casa+
componenti_mobili+altro+n_dipendenti+tipo_cliente+fiera,
family=binomial, data=dati)
summary(Logdom2)

```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-5.017e+00	1.417e+00	-3.540	0.00040	***
provinciaBZ	-1.344e+01	1.084e+03	-0.012	0.99010	
provinciaPD	2.581e+00	1.214e+00	2.125	0.03355	*
provinciaPN	1.108e+00	1.460e+00	0.759	0.44787	
provinciaRO	1.772e+00	1.432e+00	1.238	0.21574	
provinciaTN	1.985e+00	1.319e+00	1.505	0.13244	
provinciaTV	1.190e+00	1.297e+00	0.917	0.35912	
provinciaUD	1.256e+00	1.238e+00	1.015	0.31015	
provinciaVE	2.164e+00	1.322e+00	1.637	0.10169	
provinciaVI	2.069e+00	1.257e+00	1.647	0.09961	.
provinciaVR	8.596e-01	1.169e+00	0.735	0.46214	
conoscenza_prodottiSI	1.555e+00	5.969e-01	2.606	0.00916	**
acquisto_prodottiSI	-1.016e+00	7.856e-01	-1.293	0.19597	
conoscenza_agentiSI	4.947e-01	4.334e-01	1.142	0.25366	
mobili_ufficio1	2.738e-01	6.384e-01	0.429	0.66795	
cucinel	8.188e-01	5.384e-01	1.521	0.12832	
mobili_bagno1	-2.575e-01	6.986e-01	-0.369	0.71244	
mobili_casal	7.661e-01	4.705e-01	1.628	0.10347	
componenti_mobili1	2.960e-01	6.009e-01	0.493	0.62229	
altrol	3.346e-01	6.033e-01	0.555	0.57914	
altroF	1.500e+00	8.552e-01	1.754	0.07948	.
altroMIS	1.656e+00	1.140e+00	1.453	0.14617	
altroMSM	2.106e+00	6.601e-01	3.190	0.00142	**
altroS	-7.005e-01	7.330e-01	-0.956	0.33926	
n_dipendenti	6.472e-03	5.423e-03	1.193	0.23270	
tipo_clienteB	9.514e-01	4.761e-01	1.998	0.04566	*
tipo_clienteD	-1.505e+00	1.146e+00	-1.313	0.18914	
fieral	-6.364e-01	4.786e-01	-1.330	0.18359	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 267.38 on 247 degrees of freedom
Residual deviance: 209.79 on 220 degrees of freedom
AIC: 265.79

Number of Fisher Scoring iterations: 15

MODELLO RIDOTTO

Il primo passo compiuto nella determinazione del modello ridotto è stato quello di togliere la variabile 'conoscenza_prodotti' anche se risultava, come si può vedere anche dal summary precedente, significativa. Nel rispondere alla domanda 1 si è già constatato il forte legame tra questa variabile e 'acquisto_altri', in questo caso variabile risposta. Tenendo comunque presente questa considerazione, nella stima dei modelli successivi non è stata utilizzata come variabile risposta ma verrà comunque tenuta presente nelle conclusioni finali.

Le altre variabili, quelle non significative per il modello logistico, sono state tolte attraverso una selezione all'indietro, ottenendo i seguenti risultati.

```
Logdom3 <- glm(acquisto_altri ~  
provincia+conoscenza_agenti+altro, family=binomial, data=dati)  
summary(Logdom3)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3.1775	1.1216	-2.833	0.00461	**
provinciaBZ	-13.7885	1124.9371	-0.012	0.99022	
provinciaPD	2.3493	1.1558	2.033	0.04209	*
provinciaPN	1.5354	1.3132	1.169	0.24233	
provinciaRO	1.8392	1.3738	1.339	0.18066	
provinciaTN	1.8974	1.2544	1.513	0.13038	
provinciaTV	1.4525	1.1978	1.213	0.22525	
provinciaUD	1.3153	1.1968	1.099	0.27174	
provinciaVE	2.4442	1.2427	1.967	0.04920	*
provinciaVI	1.9893	1.1993	1.659	0.09717	.
provinciaVR	1.0836	1.1233	0.965	0.33469	
conoscenza_agentiSI	0.7232	0.3908	1.851	0.06421	.
altrol	-0.2105	0.5149	-0.409	0.68270	
altroF	0.7733	0.7369	1.049	0.29403	
altroMIS	0.6347	0.9313	0.682	0.49553	
altroMSM	1.4717	0.5109	2.881	0.00397	**
altroS	-1.0850	0.6356	-1.707	0.08781	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 267.38 on 247 degrees of freedom  
Residual deviance: 233.24 on 231 degrees of freedom  
AIC: 267.24
```

La provincia in cui è situata l'azienda intervistata, il settore produttivo 'altro' e la condizione di conoscere gli agenti di zona STEFANI-IDM sono i fattori indicati da questo metodo di classificazione come influenti per l'acquisto di prodotti non a marchio STEFANI-IDM.

3.4.1.b ALBERO DI CLASSIFICAZIONE

STIMA DELL'ALBERO COMPLETO – FASE DI CRESCITA DELL'ALBERO

Come già indicato precedentemente, anche in questo modello la variabile che indica se l'azienda che ha risposto al questionario conosce i prodotti a marchio STAFANI-IDM è stata omessa tra le variabili esplicative perché già presunta significativa. In questo modo si è tentato di non 'oscurare' altre variabili che possono fornire comunque indicazioni importanti.

```
set.seed(5123)
library(tree)
t1 <- tree(acquisto_altri ~
provincia+ acquisto_prodotti+conoscenza_agenti+ mobili_ufficio+
cucine+mobili_bagno+mobili_casa+componenti_mobili+altro+
n_dipendenti+tipo_cliente+fiera, data=dati)
plot(t1)
text(t1)
```

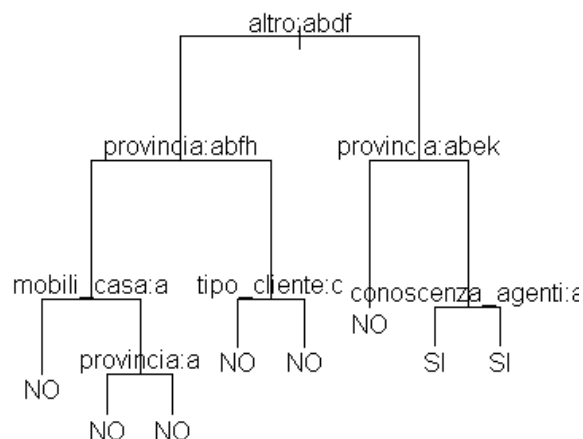


Figura 3.10. Albero di classificazione completo.

ALBERO DEFINITIVO – FASE DI POTATURA

Come nella domanda precedente, anche in questo caso l'utilizzo della cross validation ci ha permesso di perfezionare la stima dell'albero, che è molto sensibile al basso numero di osservazioni, e quindi arrivare a conclusioni più stabili.

```
t2<-cv.tree(t1, , prune.tree)
plot(t2)
```

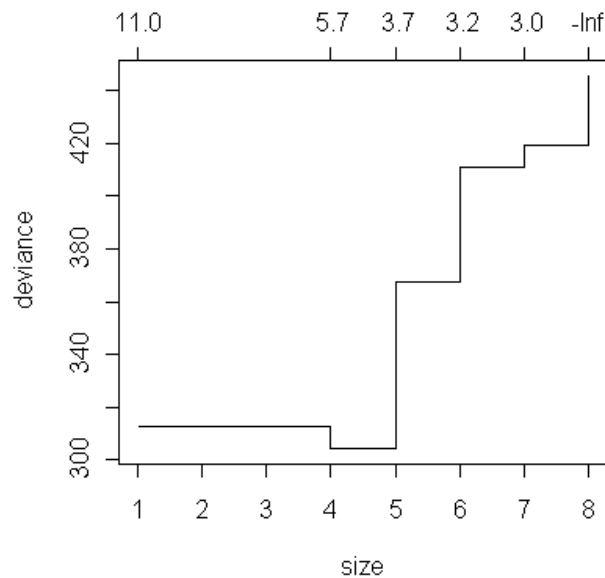


Figura 3.11. Devianza per la grandezza dell'albero

```
t3<-cv.tree(t1,best=4)
t3<-cv.tree(t1,best=t2$size[t2$dev==min(t2$dev)])
plot(t3)
text(t3)
```

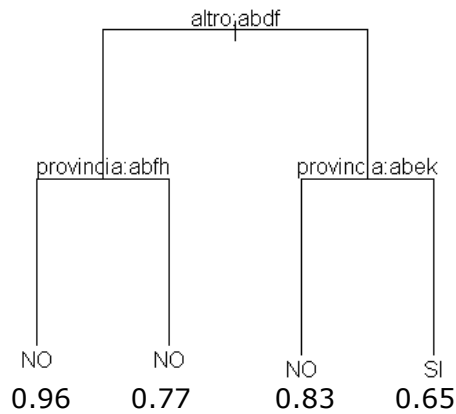


Figura 3.12. Albero potato e definitivo e indicazione della probabilità associata alle foglie

L'albero di classificazione appena stimato ci ha portato alla conclusione che le variabili altro e provincia sono quelle che influiscono sull'acquisto di prodotti dalla concorrenza da parte delle aziende.

3.4.1.c ANALISI DISCRIMINANTE LINEARE

MODELLO COMPLETO

Le variabili presenti nel modello completo sono le stesse utilizzate nella stima dei due modelli precedenti. La differenza sta nel numero di osservazioni utilizzate. Nella variabile 'conoscenza_agenti' sono presenti due valori mancanti. Nella stima del modello queste due osservazioni non sono state prese in considerazione proprio per una opzione interna del comando utilizzato. Ciò non ha permesso di determinare le percentuali di errore dato che le classi stimate e la variabile risposta non avevano la stessa dimensione e quindi non confrontabili. Si sono dovute omettere le due osservazioni per poter calcolare le percentuali di errore, che sono il punto di riferimento per la determinazione delle variabili significative.

```

Ad1<-lda(acquisto_altri ~
provincia+conoscenza_agenti+acquisto_prodotti+mobili_ufficio+cucine+mobili_bagno+mobili_casa+componenti_mobili+altro+n_dipendenti+
tipo_cliente+fiera, CV=TRUE, data=dati)
  
```

```
tabella.sommario(ad1$class,dati$acquisto_altri)
```

```
  osserv
```

```
pred NO  SI
```

```
NO 176  50
```

```
SI  15   7
```

```
errore totale:  0.262
```

```
falsi positivi/negativi:  0.681 0.221
```

MODELLO RIDOTTO

```
ad1GLM <- lda(acquisto_altri ~ provincia+conoscenza_agenti+  
altro, CV=TRUE, data=dati248)
```

```
tabella.sommario(ad1GLM$class,dati248$acquisto_altri)
```

```
  osserv
```

```
pred NO  SI
```

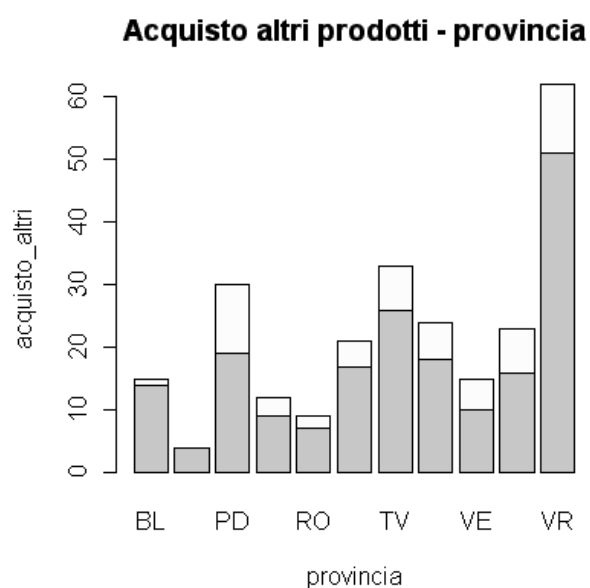
```
NO 180  47
```

```
SI  11  10
```

```
errore totale:  0.233
```

```
falsi positivi/negativi:  0.523 0.207
```

3.4.1.d PRIME CONSIDERAZIONI PRIMA PARTE



BL BZ PD PN RO TN TV UD VE VI VR

Figura 3.13. Acquisto prodotti dalla concorrenza per la provincia.

Nel grafico sopra riportato è presentata la relazione tra la provincia e l'acquisto di prodotti non a marchio STEFANI-IDM. Il colore grigio indica il non acquisto mentre il colore bianco l'acquisto di prodotti. Nelle province di Padova, Venezia e Vicenza sono situate quelle aziende che si servono maggiormente di prodotti offerti dalla concorrenza. Le altre province sono tutte proporzionalmente molto simili con un numero non molto elevato di compere , ad eccezione della provincia di Bolzano in cui non sono avvenute.

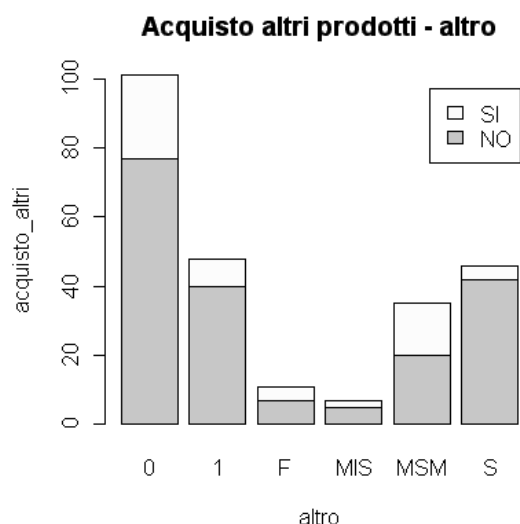


Figura 3.14. Acquisto prodotti dalla concorrenza per il settore produttivo 'altro'.

La significatività della variabile 'altro' sta ad indicare che molte aziende operanti nel settore dei mobili su misura acquistano le macchine dalla concorrenza, che per il tipo di dati raccolti risulta produrre per la maggior parte macchine cosiddette 'piccole'. Lo stesso vale per le falegnamerie generiche, indicata con 'F' nella figura precedente, ma la numerosità delle osservazioni è troppo bassa per dichiarare questa considerazione affidabile. La categoria dei serramentisti, indicata con 'S', è quella che ha comperato il minor numero di macchine bordatrici, probabilmente perché il settore produttivo non necessita del tipo di lavorazione che questo tipo macchina offre oppure non richiede un ricambio generazionale del prodotto dato che bisogna sempre tenere ben presente che nel questionario l'acquisto è riferito agli ultimi cinque anni.

Acquisto altri prodotti - conoscenza agenti

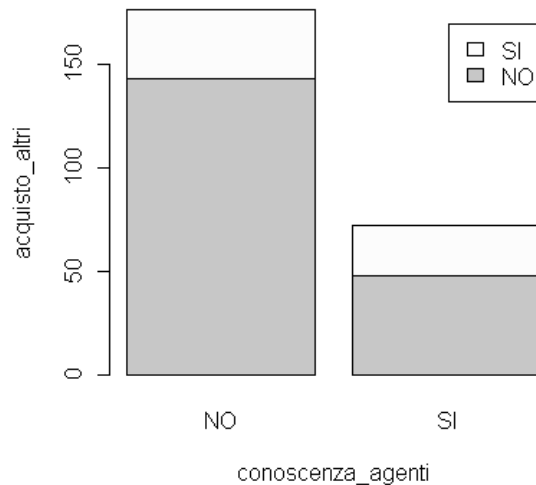


Figura 3.15. Acquisto prodotti per la conoscenza degli agenti di zona STEFANI-IDM.

La Figura 3.15 precedente porta a un risultato piuttosto inaspettato: il fatto che le aziende intervistate conoscano gli agenti di zona STEFANI-IDM fa in qualche modo aumentare la possibilità che esse acquistino prodotti dalla concorrenza. Si potrebbero fare alcune ipotesi in merito. Le aziende vengono contattate sia dagli agenti STEFANI-IDM che da quelli di marchi concorrenti che propongono loro le macchine per la lavorazione della bordatura di cui necessitano. Il risultato può essere che queste aziende, che sono per lo più di piccole dimensioni visti i dati raccolti, preferiscano l'acquisto di un certo tipo di macchine (per dimensione, prezzo e tipo di lavorazione) non offerte dall'azienda STEFANI-IDM. Nonostante l'offerta non sia stata accettata l'azienda ha comunque avuto un contatto con gli agenti.

3.4.2 SECONDA PARTE - ACQUISTO PRODOTTI STEFANI-IDM

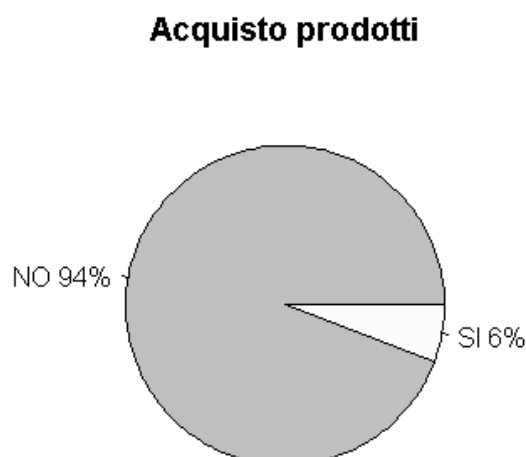


Figura 3.16. Percentuale di prodotti acquistati STEFANI-IDM.

La percentuale di aziende intervistate che ha acquistato prodotti a marchio STEFANI-IDM è piuttosto bassa soprattutto se confrontata con la percentuale di acquisti di prodotti da aziende concorrenti, riportata nella Figura 3.8. Per il tipo di mercato a cui punta il marchio STEFANI-IDM sarebbe più adeguato confrontare le vendite con le aziende così definite 'grandi' proprio per la similitudine della tipologia di prodotti offerti. Prendendo come riferimento le Figure 3.8. – 3.9. si può determinare che una percentuale leggermente inferiore al 6% delle aziende intervistate ha acquistato macchine da aziende grandi. Ovviamente con il termine generico grandi si intendono più aziende quindi, per i non molti dati a disposizione e per il sottomercato preso in considerazione, la STEFANI-IDM ha effettuato maggiori vendite rispetto alle dirette concorrenti.

3.4.2.a DETERMINAZIONE DEL MODELLO LOGISTICO

MODELLO LOGISTICO COMPLETO

Come accaduto in precedenza anche per la stima di questo modello è stato necessario togliere alcune variabili. La prima variabile tolta è stata

'conoscenza_prodotti' dato che una aziende che acquista un prodotto anche lo conosce. Come nella stima degli altri modelli alcune variabili fortemente correlate con quelle presenti nel modello completo sono state omesse. In seguito, per motivi di carattere pratico, si è dovuta togliere anche la variabile 'provincia' altrimenti il comandi impostati fornivano dei 'warning' e non si poteva procedere con l'analisi.

```
Logdom22<- glm (acquisto_prodotti ~
acquisto_altri+conoscenza_agenti+mobili_ufficio+cucine+mobili_bagno+mobili_casa+componenti_mobili+altro+n_dipendenti+tipo_cliente+fiera, family=binomial, data=dati)
summary(Logdom22)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3.315e+00	1.015e+00	-3.267	0.00109	**
acquisto_altriSI	-6.514e-01	7.819e-01	-0.833	0.40475	
conoscenza_agentiSI	2.352e+00	9.010e-01	2.610	0.00905	**
mobili_ufficio1	-6.981e-02	8.297e-01	-0.084	0.93295	
cucine1	8.095e-01	7.049e-01	1.148	0.25082	
mobili_bagno1	-1.124e-01	9.400e-01	-0.120	0.90481	
mobili_casa1	2.037e-01	6.953e-01	0.293	0.76954	
componenti_mobili1	-4.279e-02	8.429e-01	-0.051	0.95951	
altro1	-9.470e-01	1.189e+00	-0.797	0.42566	
altroF	-1.739e+01	4.951e+03	-0.004	0.99720	
altroMIS	-1.702e+01	6.498e+03	-0.003	0.99791	
altroMSM	-1.682e+01	2.887e+03	-0.006	0.99535	
altroS	-1.771e+01	2.283e+03	-0.008	0.99381	
n_dipendenti	4.134e-03	8.280e-03	0.499	0.61756	
tipo_clienteB	1.331e+00	1.206e+00	1.104	0.26953	
tipo_clienteD	-1.358e-01	1.410e+00	-0.096	0.92328	
fiera1	-1.883e+00	9.779e-01	-1.925	0.05422	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 113.235 on 247 degrees of freedom
Residual deviance: 69.273 on 231 degrees of freedom
AIC: 103.27

Number of Fisher Scoring iterations: 19

MODELLO RIDOTTO

Durante l'operazione di selezione all'indietro, per determinare il modello ridotto con solamente le variabili significative, si è tentato di sostituire il numero di dipendenti con il suo logaritmo e la variabile trasformata è risultata significativa al contrario della non trasformata.

```
Logdom22<-glm(acquisto_prodotti ~  
conoscenza_agenti+log(n_dipendenti)+fiera, family=binomial,  
data=dati)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-4.9697	1.0744	-4.626	3.73e-06	***
conoscenza_agentiSI	2.2273	0.8190	2.719	0.00654	**
log(n_dipendenti)	0.5732	0.2762	2.075	0.03797	*
fiera1	-1.4057	0.7180	-1.958	0.05025	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 113.235 on 247 degrees of freedom
Residual deviance: 76.132 on 244 degrees of freedom
AIC: 84.132

Number of Fisher Scoring iterations: 7

Dal modello logistico stimato risultano significative per l'acquisto di prodotti a marchio STEFANI-IDM la conoscenza degli agenti, la variabile 'n_dipendenti' trasformata mediante la funzione logaritmo e la variabile fiera.

3.4.2.b ALBERO DI CLASSIFICAZIONE

Per questo particolare tipo di analisi l'albero di classificazione, che si è provato a stimare, non ha portato a conclusioni che potessero essere utili agli obiettivi preposti. Per questo motivo in questo caso questo metodo di classificazione non viene preso in considerazione.

3.4.2.c ANALISI DISCRIMINANTE LINEARE

MODELLO COMPLETO

Per la stima di questo modello non vengono utilizzate le variabili 'conoscenza_prodotti' e 'provincia' per i motivi già specificati in precedenza.

```
ad1<-lda(acquisto_prodotti ~
acquisto_altri+conoscenza_agenti+mobili_ufficio+cucine+mobili_bag
no+mobili_casa+componenti_mobili+altro+ n_dipendenti+
tipo_cliente+fiera, CV=TRUE, data=dati248)
tabella.sommario(ad1$class,dati248$acquisto_prodotti)
  osserv
pred NO  SI
  NO 224  13
  SI   9   2
errore totale:  0.088
falsi positivi/negativi:  0.818 0.054
```

MODELLO RIDOTTO

```
ad3<-lda(acquisto_prodotti ~
conoscenza_agenti+log(n_dipendenti)+fiera, CV=TRUE, data=dati248)
tabella.sommario(ad3$class,dati248$acquisto_prodotti)
  osserv
pred NO  SI
  NO 225  11
  SI   8   4
errore totale:  0.076
falsi positivi/negativi:  0.667 0.047
```

Anche la stima di questo modello indica che i fattori maggiormente influenti per l'acquisto dei prodotti a marchio STEFANI-IDM sono la conoscenza degli agenti, il contatto in fiera tra l'azienda e uno dei marchi del Gruppo SCM e il logaritmo del numero di dipendenti.

3.4.2.d PRIME CONSIDERAZIONI SECONDA PARTE

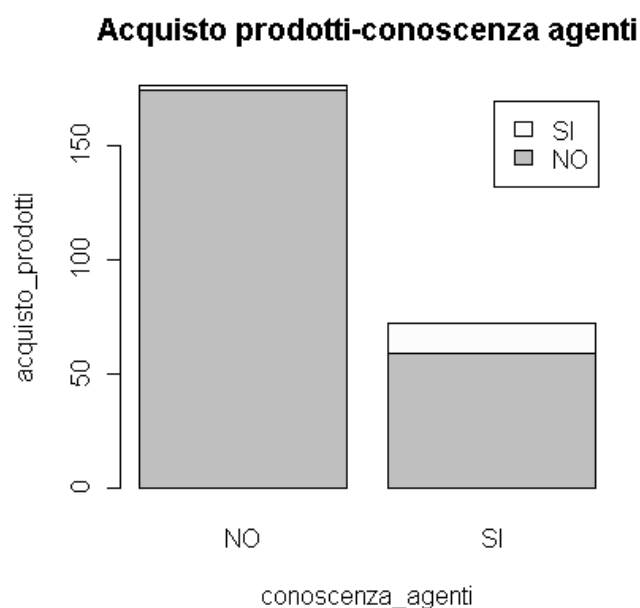


Figura 3.17. Grafico che rappresenta le relazione tra l’acquisto dei prodotti e la conoscenza degli agenti di zona STEFANI-IDM.

Il grafico sottolinea come le aziende che hanno acquistato prodotti a marchio STEFANI-IDM conoscano gli agenti di zona tanto da indicare che un acquisto, o comunque il rapporto tra azienda cliente e azienda fornitore, ha quasi sempre come intermediario un agente che presumibilmente non ha solamente lo scopo della vendita. L’agente in questione sembrerebbe più una persona che va direttamente dalle aziende che per esempio ‘aspettarle’, come per esempio nelle fiere del 2002 visto che dalla Figura 3.18 si ricava che i contatti in fiera con uno dei marchi del Gruppo SCM fanno diminuire la probabilità di acquisto dei prodotti. Risultato piuttosto sorprendente dato il fatto che lo scopo delle fiere, oltre quello di esporre e illustrare i propri prodotti, è quello della vendita. Le aziende intervistate dimostrano in larga maggioranza di aver preso parte alla fiera ma probabilmente c’è stato un maggior contatto verso gli altri marchi del gruppo rispetto alla STEFANI-IDM.

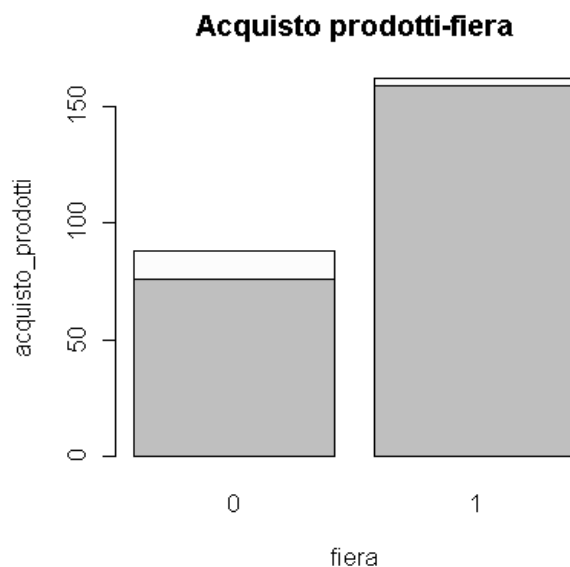


Figura 3.18. Relazione tra l'acquisto dei prodotti STEFANI-IDM e il contatto in fiera con il gruppo SCM.

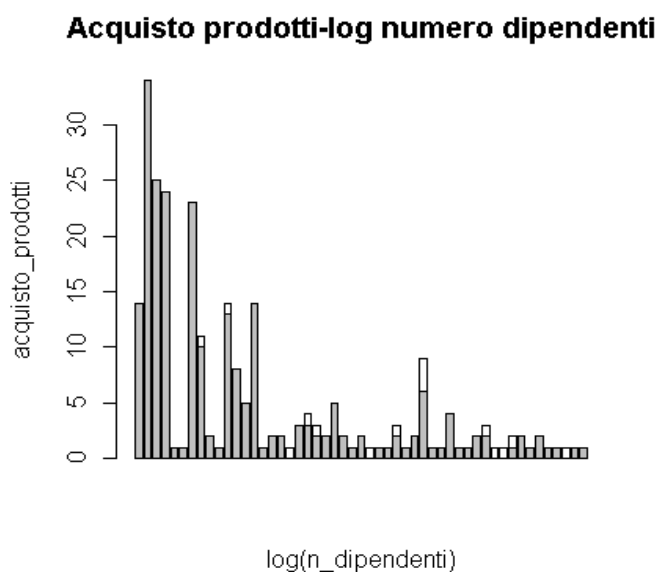


Figura 3.18. Relazione tra l'acquisto dei prodotti STEFANI-IDM e il logaritmo del numero di dipendenti.

Dal grafico precedente si può intuire come all'aumentare del logaritmo del numero di dipendenti aumenta il numero di acquisti di prodotti a marchio STEFANI-IDM. Si può solamente intuire questa osservazione guardando il grafico in quanto ci sono troppe poche osservazioni di aziende che non hanno acquistato prodotti rispetto a quelle che ne hanno acquistato. La

trasformazione della variabile tramite la funzione logaritmo funge da 'lisciatore' mano a mano che il numero di dipendenti aumenta.

Prima di iniziare le analisi nel paragrafo 3.2.2 è stato ricavato un modello per determinare il numero di dipendenti in quelle osservazioni in cui erano presenti valori mancanti. Da questo modello si è ricavato che il logaritmo del numero di dipendenti è molto legato la provincia in cui un'azienda è situata, il settore produttivo componenti per mobili e 'altro' e il tipo di cliente secondo la classificazione interna del Gruppo SCM. Probabilmente queste variabili sono state 'offuscate' dall'utilizzo della trasformata di 'n_dipendenti'.

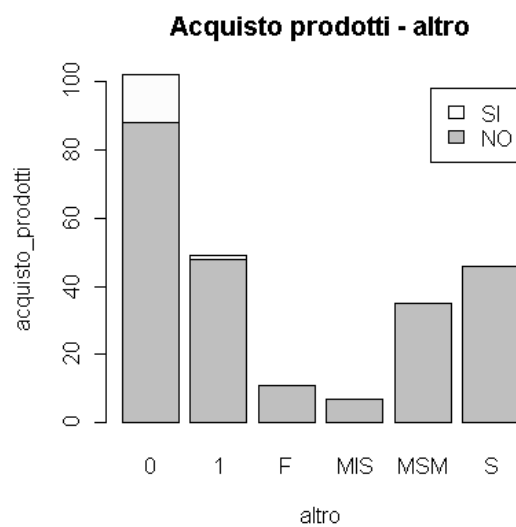
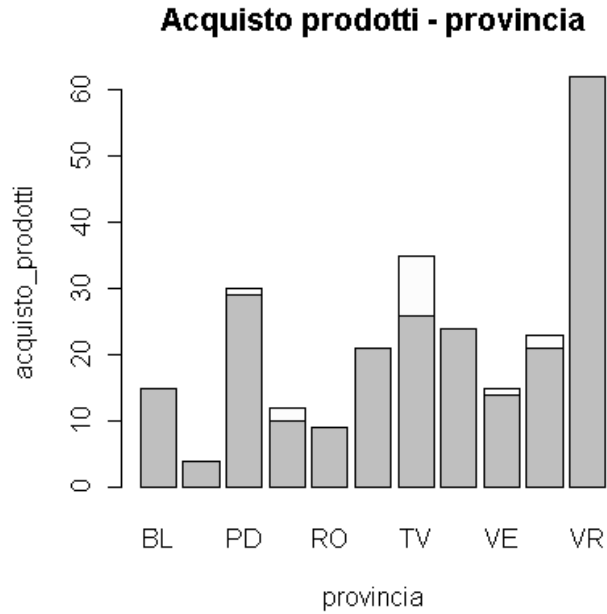


Figura 3.19. Relazione tra l'acquisto dei prodotti STEFANI-IDM e il settore produttivo 'altro'.

Il settore produttivo 'altro' sembrerebbe comunque influenzare la conoscenza dei prodotti dato che, se pur poche, le aziende che hanno acquistato prodotti a marchio STEFANI-IDM non hanno attività di tipo 'altro', che per il tipo di attività svolte, spesso impiega un basso numero di dipendenti.



BL BZ PD PN RO TN TV UD VE VI VR

Figura 3.20. Relazione tra l'acquisto dei prodotti STEFANI-IDM e la provincia.

Sempre dal modello stimato si ricava che le aziende della provincia di Treviso, Pordenone e Venezia sono quelle con un maggior numero di dipendenti impiegati. Come si può notare dalla figura precedente sono anche quelle con maggior acquisto di prodotti STEFANI-IDM.

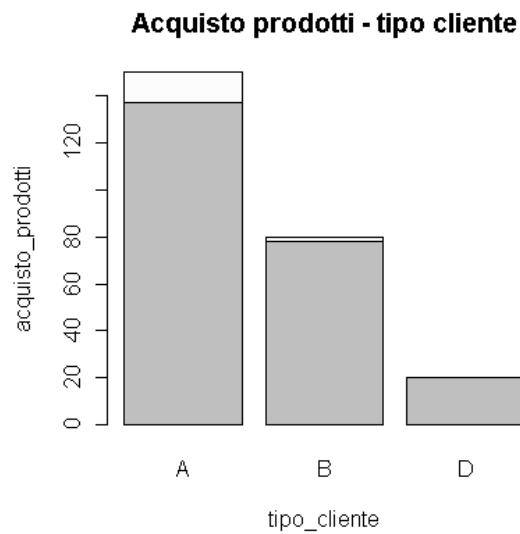


Figura 3.21. Relazione tra l'acquisto dei prodotti STEFANI-IDM e il tipo di cliente.

Dal modello ricavato si ricava che i clienti di tipo B hanno un numero di dipendenti tendenzialmente minore rispetto ai clienti di tipo A. Per questo motivo l'acquisto di prodotti è avvenuto solamente da parte dei clienti di tipo A mentre quelli di tipo D non hanno fatto acquisti proprio per definizione in quanto sono clienti solamente potenziali.

3.5 DOMANDA 3

Caratteristiche delle aziende che vengono contattate dagli agenti di zona STEFANI-IDM.

Come approccio al quarto quesito è stata fatta una analisi preliminare di tipo descrittivo della variabile 'conoscenza_agenti', per avere una prima indicazione sulle caratteristiche di questa variabile che, già di per se, risulta molto utile.

All'interno del questionario, collegata alla domanda filtro sugli agenti, c'era la richiesta di indicare quante visite ricevono mensilmente le aziende intervistate da parte degli agenti STEFANI-IDM. La variabile ricavata da questa domanda comporta comunque una informazione interessante ma non può essere utilizzata all'interno dell'analisi per la sua elevata correlazione con la variabile risposta 'conoscenza_agenti'.

Conoscenza agenti STEFANI-IDM

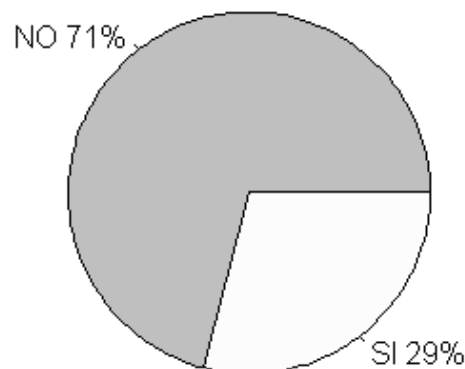


Figura 3.22. Grafico a torta sulla conoscenza degli agenti STEFANI-IDM.

La figura appena riportata indica che un numero piuttosto considerevole delle aziende intervistate non ha mai fatto la conoscenza degli agenti di zona o

comunque le visite fatte non sono state così impresse ai membri dell'azienda da potergli far rispondere in modo affermativo alla domanda.

Visite mensili agenti STEFANI-IDM

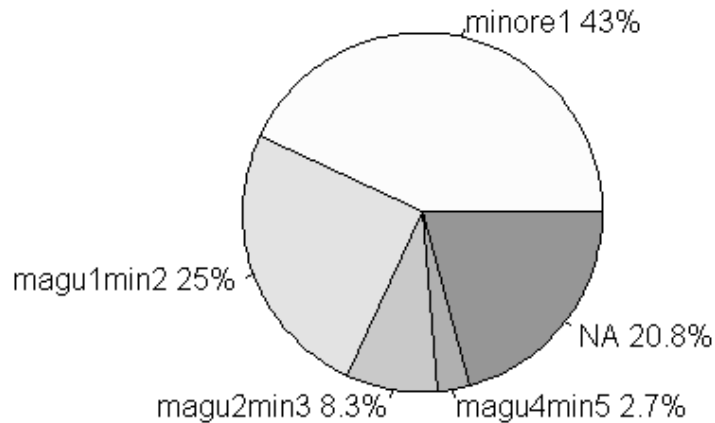


Figura 3.23. Grafico a torta del numero di visite mensili effettuate dagli agenti STEFANI-IDM alle aziende intervistate.

La Figura 3.23 visualizza quante visite mensili ricevano le aziende che hanno dichiarato di conoscere gli agenti di zona STEFANI-IDM. Una parte considerevole riceve meno di due visite al mese mentre un ruolo importante lo giocano anche le non risposte. Questo probabilmente indica la difficoltà delle aziende nel rispondere alla domanda e quindi di ricordare, e successivamente quantificare, le visite ricevute.

3.5.1 CLASSIFICAZIONE MEDIANTE IL MODELLO DI REGRESSIONE LOGISTICO

MODELLO LOGISTICO COMPLETO

Nel modello seguente non è stata inserita la variabile 'conoscenza_prodotti' perché è scontato che chi conosce gli agenti conosce i prodotti. Inoltre non sono state inserite alcune variabili fortemente correlate con quelle presenti nel modello sottostante.

```

glm1 <- glm (conoscenza_agenti~
provincia+acquisto_prodotti+acquisto_altri+mobili_ufficio+
cucine+mobili_bagno+mobili_casa+componenti_mobili+altro+
n_dipendenti+tipo_cliente+fiera, family=binomial, data=dati)
summary(glm1)

```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.603e+00	1.057e+00	-2.463	0.01376	*
provinciaBZ	-1.326e+01	1.065e+03	-0.012	0.99007	
provinciaPD	2.150e-01	9.971e-01	0.216	0.82928	
provinciaPN	3.536e+00	1.340e+00	2.639	0.00832	**
provinciaRO	5.226e-01	1.260e+00	0.415	0.67846	
provinciaTN	7.421e-01	1.164e+00	0.637	0.52382	
provinciaTV	9.583e-01	9.886e-01	0.969	0.33240	
provinciaUD	1.960e+00	9.921e-01	1.975	0.04822	*
provinciaVE	2.000e+00	1.125e+00	1.778	0.07539	.
provinciaVI	1.337e+00	9.747e-01	1.372	0.17004	
provinciaVR	-6.502e-02	9.487e-01	-0.069	0.94536	
acquisto_prodottiSI	3.068e+00	1.100e+00	2.789	0.00528	**
acquisto_altriSI	6.347e-01	4.515e-01	1.406	0.15977	
mobili_ufficiol	-2.709e-03	6.611e-01	-0.004	0.99673	
cucinel	-7.167e-01	6.126e-01	-1.170	0.24197	
mobili_bagnol	9.787e-01	6.880e-01	1.423	0.15485	
mobili_casal	1.470e+00	5.103e-01	2.881	0.00396	**
componenti_mobilil	1.950e+00	6.404e-01	3.045	0.00232	**
altrol	-7.047e-01	6.752e-01	-1.044	0.29667	
altroF	-3.855e-02	1.017e+00	-0.038	0.96975	
altroMIS	6.867e-01	1.289e+00	0.533	0.59414	
altroMSM	6.501e-01	7.566e-01	0.859	0.39024	
altroS	-7.188e-02	6.915e-01	-0.104	0.91721	
n_dipendenti	2.972e-03	5.766e-03	0.516	0.60620	
tipo_clienteB	-1.003e+00	4.883e-01	-2.054	0.04002	*
tipo_clienteD	-2.823e+00	1.257e+00	-2.247	0.02464	*
fieral	-1.191e-01	4.812e-01	-0.248	0.80445	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 298.81 on 247 degrees of freedom
Residual deviance: 194.07 on 221 degrees of freedom
AIC: 248.07

Number of Fisher Scoring iterations: 15

MODELLO LOGISTICO RIDOTTO FINALE

Prima di iniziare la procedura di selezione all'indietro è stata tolta anche la variabile 'acquisto_prodotti' dal modello completo in quanto già in una analisi precedente è stato verificato che risulta significativo il suo rapporto con la conoscenza degli agenti ma anche perché era intuitivo che fosse così proprio per il ruolo dell'agente.

Tramite la selezione all'indietro si è arrivati ad ottenere il seguente modello.

```
glm1<-glm(conoscenza_agenti~  
provincia+mobili_casa+componenti_mobili+altro + tipo_cliente,  
family=binomial, data=dati)  
summary(glm1)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.0589	0.9114	-2.259	0.02388	*
provinciaBZ	-13.5868	1058.2828	-0.013	0.98976	
provinciaPD	0.4142	0.9313	0.445	0.65651	
provinciaPN	3.1790	1.1675	2.723	0.00647	**
provinciaRO	0.5698	1.2376	0.460	0.64525	
provinciaTN	0.6129	1.1311	0.542	0.58794	
provinciaTV	1.7203	0.9063	1.898	0.05767	.
provinciaUD	1.8621	0.9505	1.959	0.05011	.
provinciaVE	2.2004	1.0568	2.082	0.03732	*
provinciaVI	1.5918	0.9310	1.710	0.08731	.
provinciaVR	-0.1901	0.9135	-0.208	0.83517	
mobili_casa1	1.2470	0.4284	2.911	0.00360	**
componenti_mobili1	1.4276	0.5428	2.630	0.00854	**
altro1	-1.1041	0.6076	-1.817	0.06919	.
altroF	-0.6454	0.9393	-0.687	0.49199	
altroMIS	-0.1607	1.2651	-0.127	0.89894	
altroMSM	0.4351	0.6358	0.684	0.49375	
altroS	-0.6893	0.6069	-1.136	0.25608	
tipo_clienteB	-0.8707	0.4445	-1.959	0.05013	.
tipo_clienteD	-2.3194	1.1128	-2.084	0.03713	*

```
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 298.81  on 247  degrees of freedom  
Residual deviance: 209.04  on 228  degrees of freedom  
AIC: 249.04
```

Number of Fisher Scoring iterations: 15

Dai risultati dei test del modello di regressione logistico risultano significative, per la conoscenza degli agenti di zona STEFANI-IDM, la provincia in cui l'azienda è ubicata, il tipo di cliente secondo la classificazione all'interno del gruppo SCM, i settori produttivi mobili per la casa, componenti per mobili e altro.

3.5.2 ALBERO DI CLASSIFICAZIONE

Prendendo spunto dalle considerazioni fatte in precedenza con il modello di regressione logistico, nel modello iniziale per la stima dell'albero di classificazione non sono state inserite le variabili 'conoscenza_prodotti' e 'acquisto_prodotti'.

STIMA DELL'ALBERO COMPLETO – FASE DI CRESCITA

```
t1 <- tree (conoscenza_agenti~
provincia+acquisto_altri+mobili_ufficio+cucine+mobili_bagno+
mobili_casa+componenti_mobili+altro+n_dipendenti+ tipo_cliente+
fiera,data=dati2)
plot(t1)
text(t1)
```

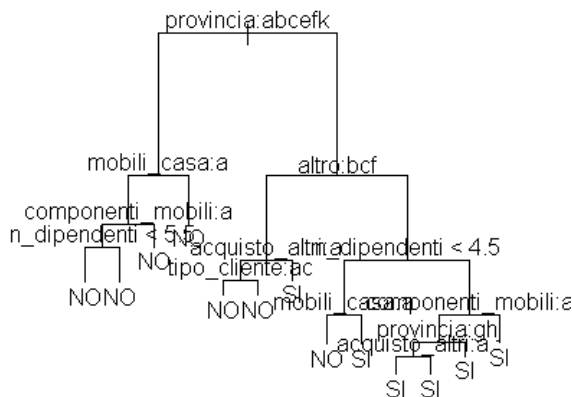


Figura 3.24. Albero di classificazione completo.

ALBERO DEFINITIVO – FASE DI POTATURA

Anche in questo caso come nella analisi precedenti la fase di potatura avviene tramite il comando 'cv.tree' che sfrutta il processo di convalida incrociata.

```
t2 <- cv.tree(t1, , prune.tree)
plot(t2)
```

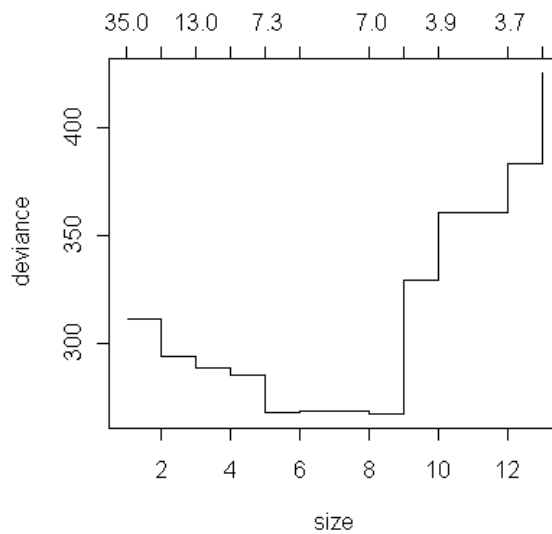


Figura 3.25. Devianza per la grandezza dell'albero.

```
t3 <- prune.tree(t1, best=t2$size[t2$dev==min(t2$dev)],
newdata=dati2)
plot(t3)
```

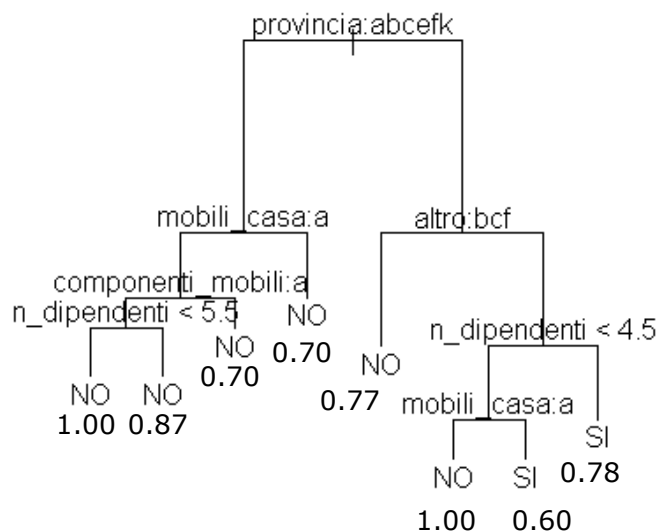


Figura 3.26. Albero potato e definitivo e indicazione della probabilità associata alle foglie

Dalla potatura dell'albero si è ricavato una selezione automatica delle variabili significative per quanto concerne la conoscenza degli agenti STEFANI-IDM. Le variabili che risultano significative quindi sono la provincia, il numero di dipendenti e i settori produttivi 'altro' e mobili per la casa. Il settore produttivo componenti per mobili non risulta significativo in quanto il nodo a cui è legata questa variabile conduce, tramite i rami, a due classi uguali, in questo caso alla classe 'NO' cioè la non conoscenza degli agenti.

3.5.3 ANALISI DISCRIMINANTE LINEARE

Anche in questo caso come nell'albero di classificazione non vengono introdotte nel modello completo le variabili riguardanti l'acquisto di prodotti e la conoscenza degli agenti STEFANI-IDM.

MODELLO COMPLETO

```
lda4cv<-lda(conoscenza_agenti~  
provincia +acquisto_altri+mobili_ufficio+ cucine  
+mobili_bagno+mobili_casa+componenti_mobili+altro+n_dipendenti+  
tipo_cliente+fiera, CV=TRUE, data=dati)
```

```
tabella.sommario(lda4cv$class,dati$conoscenza_agenti)
```

```
      osserv  
pred NO  SI  
  NO 156  34  
  SI  20  38  
errore totale:  0.217  
falsi positivi/negativi:  0.344  0.178
```

Il modello completo ha portato a dei discreti risultati per quanto riguarda le percentuali di errore ma il numero di variabili esplicative che sono state utilizzate è troppo elevato per cui, con il metodo spiegato in precedenza, sono state selezionate le variabili maggiormente significative.

MODELLO DEFINITIVO

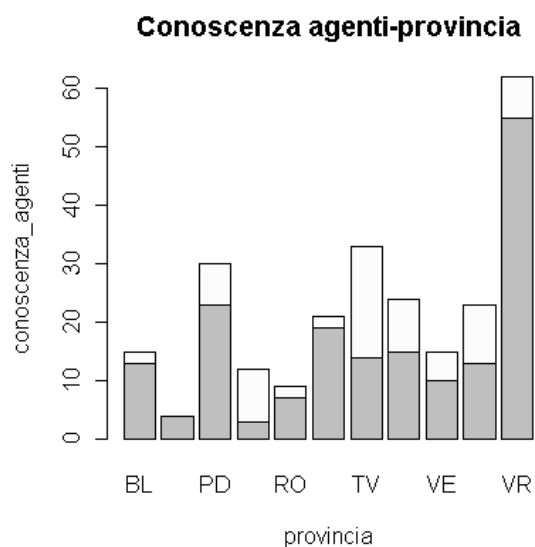
```
lda4cvDEF1<-lda(conoscenza_agenti~
provincia+altro+tipo_cliente, CV=TRUE, data=dati)

tabella.sommario(lda4cvDEF$class,dati$conoscenza_agenti)
  osserv
pred NO  SI
  NO 159  35
  SI  17  37
errore totale:  0.209
falsi positivi/negativi:  0.314 0.180
```

E' avvenuto un sensibile miglioramento nel passaggio dal modello completo al modello ridotto, come testimoniato dalla leggera diminuzione della percentuali di errore. La cosa più importante, però, è stata la notevole diminuzione delle variabili esplicative visto che si è riusciti a passare da un modello a undici variabili a un modello a quattro. Queste, che quindi sono risultate significative, sono la provincia, il tipo di cliente, il settore produttivo 'altro' e quello mobili per la casa .

3.5.4 CONSIDERAZIONI DOMANDA 3

Raccogliendo i risultati, che tutte e tre i tipi di metodi utilizzati forniscono, si è ottenuto che le variabili 'provincia' e il settore produttivo 'altro' sono significative per tutti i modelli stimati. Le variabili 'mobili_casa' e 'tipo_cliente' sono significative solamente per due metodi ma sono state considerate ugualmente influenti e quindi analizzate più in dettaglio.



BL BZ PD PN RO TN TV UD VE VI VR

Figura 3.27. Conoscenza degli agenti nelle province.

Nell'istogramma precedente sono riportate le province stratificate per la conoscenza degli agenti. Il colore grigio indica la non conoscenza degli agenti STEFANI-IDM mentre il bianco la conoscenza. Si nota immediatamente, come già visto in precedenza, che gli agenti non sono molto conosciuti ma la proporzione non è per nulla uguale in tutte le province. Nella provincia di Bolzano nessuno conosce gli agenti, anche se bisogna tenere presente la bassa numerosità che contraddistingue questa provincia. Anche nella provincia di Verona gli agenti sono poco conosciuti. Le province di Belluno, Padova, Rovigo e Trento non sono statisticamente differenti tra di loro, ma anche in questo caso gli agenti sono poco conosciuti. C'è una discreta conoscenza nelle province di Udine, Venezia e Vicenza che in proporzione sono piuttosto simili. Molto sopra alla media sono le province di Treviso ma soprattutto di Pordenone in cui la maggior parte delle aziende conoscono gli agenti STEFANI-IDM.

Questa prima analisi non rende giustizia all'analisi preliminare in cui gli agenti non risultavano molto conosciuti. Infatti le province di Verona, Padova e Trento sono tra quelle con minore numerosità ma minore conoscenza. Probabilmente se le province di Treviso e Pordenone avessero avuto una numerosità più elevata la conoscenza complessiva degli agenti cambierebbe

di molto. Questa è una considerazione che rivela ancora di più quanto sia importante effettuare un buon campionamento per ottenere dei risultati il più coerenti possibili con la realtà.

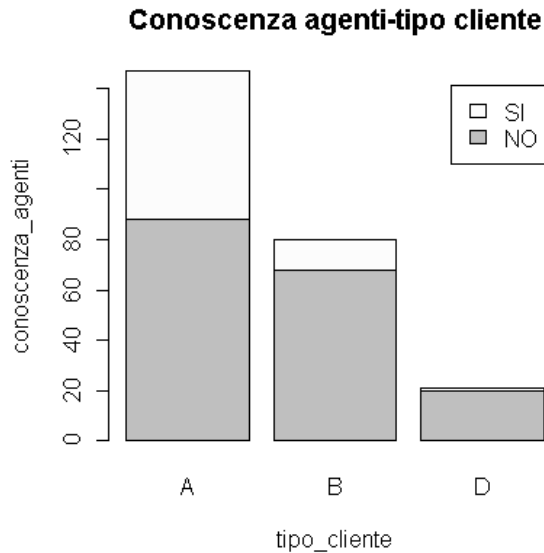


Figura 3.28. Conoscenza degli agenti per il tipo di cliente intervistato.

Come era logico aspettarsi i clienti potenziali quindi quelli di tipo D, proprio perché potenziali, non conoscono gli agenti STEFANI-IDM. Un aspetto insolito si ricava però confrontando i clienti di tipo A con quelli di tipo B. I clienti diretti, sempre riferiti al Gruppo SCM, conoscono maggiormente gli agenti rispetto ai clienti indiretti cioè quelli che acquistano tramite rappresentanti o società di vendita.

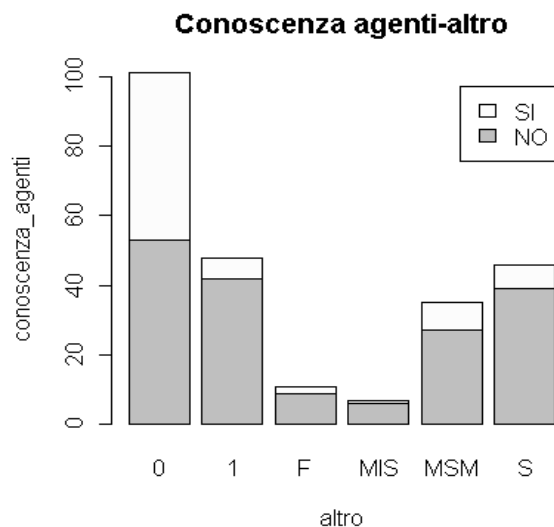


Figura 3.29. Conoscenza degli agenti in base al settore produttivo 'altro'

Prima di trarre delle conclusioni bisogna ricordare che una azienda può avere più settori produttivi tra cui 'altro'. Per come è stata studiata questa alternativa di risposta all'interno del questionario, essa sta ad indicare quei settori che solitamente non necessitano di macchine tecnologicamente elevate per l'operazione di bordatura. Tutti gli altri settori presenti come alternative di risposta nella domanda 5 del questionario di pag. 35 sono quelli che solitamente utilizzano macchine bordatrici.

Dal grafico del settore produttivo 'altro', stratificato per la conoscenza degli agenti, si ottiene che le aziende, che non hanno risposto indicando come settore produttivo 'altro', sono quelle che maggiormente conoscono gli agenti di zona STEFANI-IDM. Una azienda deve avere almeno un settore produttivo di conseguenza i settori forniti come alternative di risposta nel questionario sono quelli che permettono il maggior contatto agente-azienda. Di questi settori quello che risulta incidere maggiormente sulla conoscenza degli agenti è mobili per la casa, come si può verificare dalla Figura 3.30. sottostante. Una azienda su due che svolge anche questa attività conosce gli agenti STEFANI-IDM.

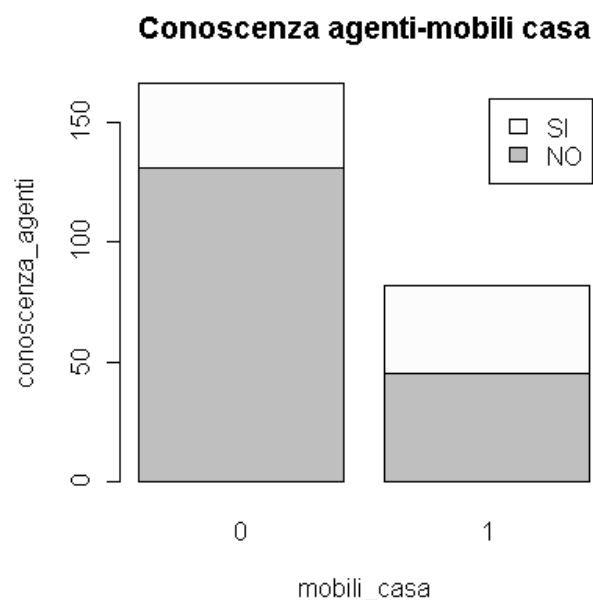


Figura 3.30. Relazione tra il settore mobili per la casa e la conoscenza degli agenti STEFANI-IDM.

CAPITOLO 4

CONCLUSIONI FINALI

4.1 CONCLUSIONI DOMANDA 1

Ricordiamo che la prima domanda era riferita ai fattori che influiscono sulla conoscenza dei prodotti a marchio STEFANI-IDM. Tramite la stima dei tre modelli stimati è risultato che le variabili significative sono l'acquisto di altri prodotti, il tipo di cliente e il numero di dipendenti. I grafici relativi a queste tre variabili sono riportate nella sezione 3.3.4 di pagina 66. Tutti e tre dimostrano che l'azienda è piuttosto conosciuta.

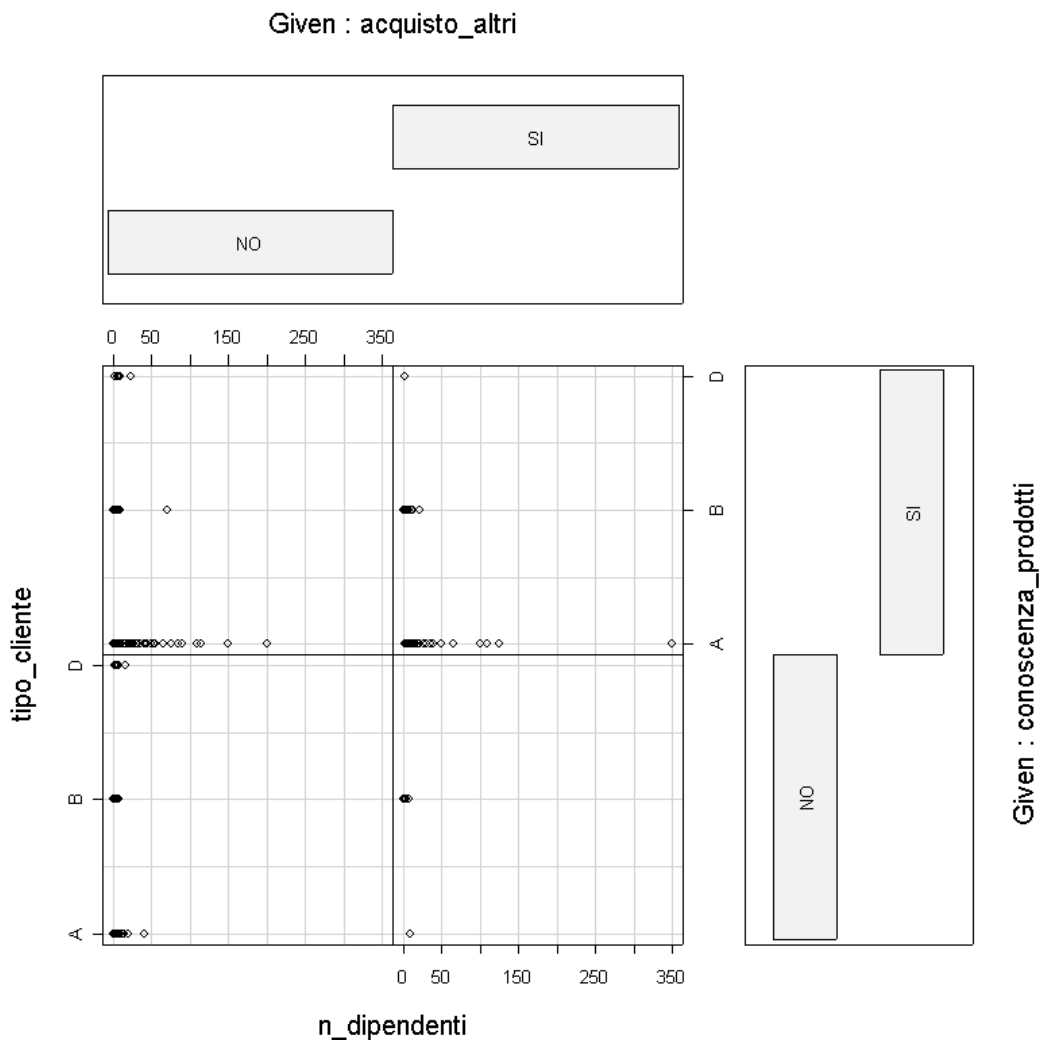


Figura 4.1. Grafico che individua la relazione tra le tre variabili significative più la variabile risposta in questo caso 'conoscenza_prodotto'.

Dal grafico precedente si nota che sono numerose le aziende che conoscono i prodotti. Quelle che non li conoscono sono di dimensioni piccole, in particolare il numero di dipendenti aumenta passando dalla categoria D alla categoria A passando per la categoria B e di conseguenza aumenta anche la conoscenza passando per questi tre tipi di clienti. La grande maggioranza di aziende che ha acquistato macchine non a marchio STEFANI-IDM conosce il marchio probabilmente perché tiene in considerazione comunque l'azienda prima di effettuare l'acquisto definitivo.

4.2 CONCLUSIONI DOMANDA 2

La domanda 2 richiedeva di determinare le cause che comportano acquisti di prodotti dalla concorrenza e non STEFANI-IDM. Dai risultati ottenuti dalle due sottoanalisi si ottiene che alcune variabili sembrano influenzare entrambi gli acquisti. Da notare che certe variabili sono definite lo stesso significative anche se in realtà per quanto riguarda l'acquisto di prodotti STEFANI-IDM sono 'derivate' dal logaritmo del numero di dipendenti. Si va ora a confrontare graficamente le indicazioni che vengono date da queste variabili 'comuni' che sono la provincia, il settore produttivo 'altro' e la conoscenza degli agenti di zona. Le figure riportate di seguito sono state già presentate in precedenza.

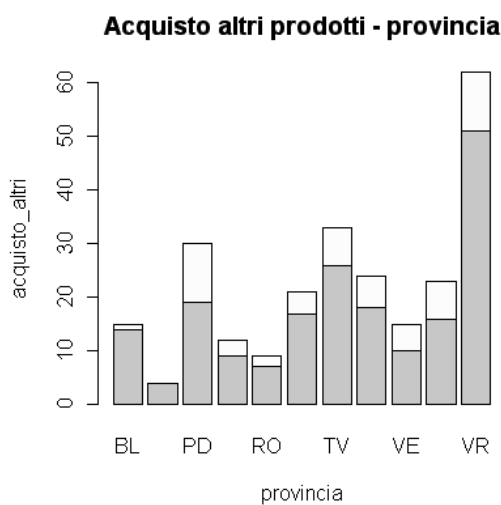


Figura 3.13.

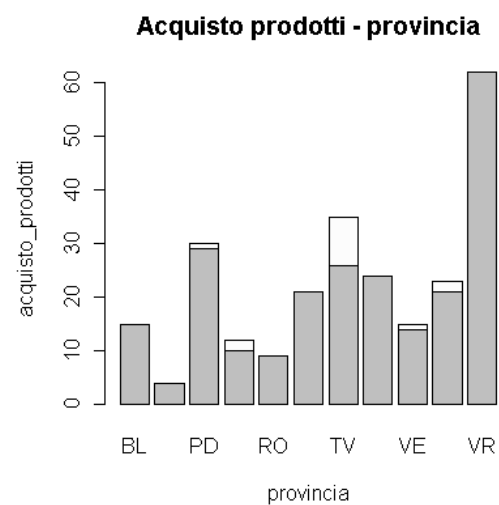


Figura 3.20.

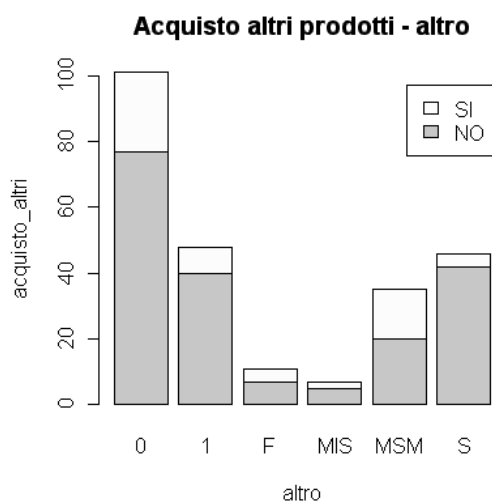


Figura 3.14

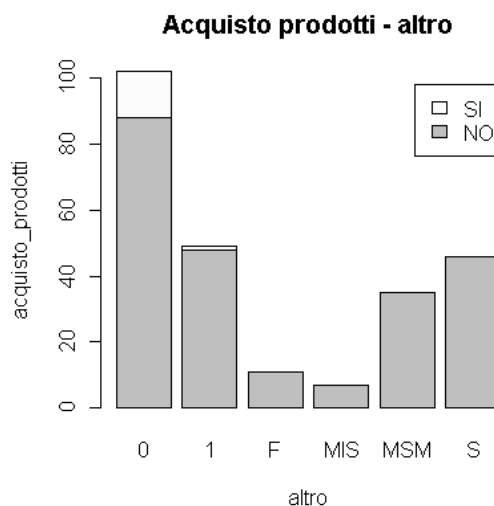


Figura 3.19

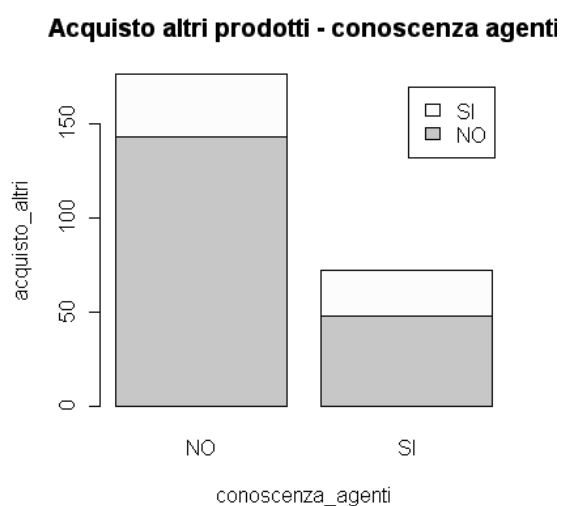


Figura 3.15

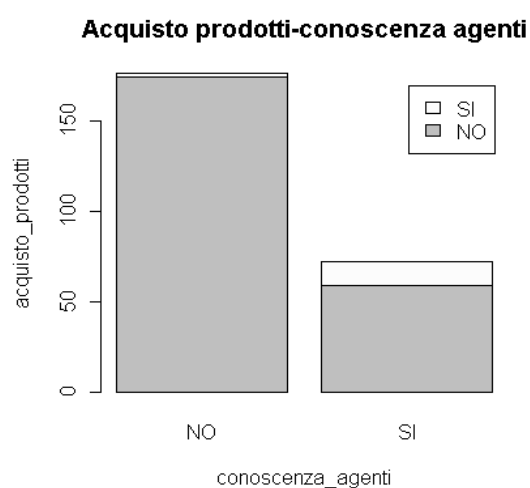


Figura 3.17

Dalle analisi fatte sembrerebbe quasi che la STEFANI-IDM tenti di accaparrarsi una parte del mercato ma non ci riesca e vada a coprire solo una parte di esso. Dal confronto delle figure tra 3.13 e 3.20 si vede nettamente come i prodotti STEFANI-IDM siano acquistati nella maggior parte nelle province di Treviso e di Pordenone mentre le aziende concorrenti hanno una distribuzione logistica più vasta che si estende in tutto il Triveneto ad eccezione di Bolzano. Stessa considerazione si riesce a ricavare dal settore produttivo 'altro' in cui praticamente la STEFANI-IDM non opera

mentre i marchi concorrenti sono ben presenti anche e non solo nel settore 'altro'.

In precedenza è stato fatto notare il tentativo di espandersi nel mercato di STEFANI-IDM per il fatto che il marchio è molto conosciuto ma soprattutto le aziende che acquistano prodotti dalla concorrenza lo conoscono, e quindi lo valutano, e una buona fetta conosce anche gli agenti. Questo potrebbe significare che determinate aziende, in particolare quelle medio piccole, prendono visione in qualche modo dei prodotti STEFANI-IDM ma poi ripiegano su altri marchi. La spiegazione di questo fatto non può essere data dalle caratteristiche del prodotto offerto che non è in qualche modo compatibile con quello richiesto dalle piccole-medie aziende.

4.3 CONCLUSIONI DOMANDA 3

Per ribadire i concetti precedentemente espressi i risultati della domanda 3 vengono in aiuto. Come testimoniato dalla Figura 3.17 l'acquisto del prodotto implica la conoscenza degli agenti di zona STEFANI-IDM. Dalle considerazioni ricavate dalla domanda 3 si ha che le 'piste' maggiormente battute dagli agenti sono quelle della provincia di Treviso e Pordenone, dei clienti di tipo A che tendenzialmente sono quelli con maggior numero di dipendenti e il settore produttivo diverso da altro. Sembrerebbe esserci un tentativo di tentare di aggredire anche altra parte del mercato, perché anche in altre province e nel settore 'altro' gli agenti sono comunque conosciuti ma, in riferimento agli ultimi cinque anni, la conoscenza non è supportata dalla vendita.

CONSIDERAZIONI PERSONALI

L'esperienza di stage svolta presso l'azienda STEFANI-IDM e le successive analisi compiute per la realizzazione di questa 'relazione finale' hanno costituito una esperienza personale molto utile, in particolare riferimento al percorso di studi affrontato, perché mi ha reso partecipe di una analisi statistica dalla raccolta dei dati alle conclusioni finali.

Lo stage mi ha permesso di toccare con mano alcuni strumenti incontrati in questi tre anni di studio e verificare come effettivamente possono venire utilizzati per scopi aziendali. Mi riferisco all'utilizzo del database, alla formulazione di un campione, al rispetto delle proporzioni tra popolazione-campione-dati raccolti, impresa che ritengo ardua per qualsiasi tipo di analisi che necessiti la stretta collaborazione di soggetti. La cosa che ritengo più interessante è stata quella di aver ricevuto delle richieste di analisi di tipo generale da parte dell'azienda a cui ho tentato di rispondere mediante gli strumenti di analisi appresi, provando il significato pratico dello studio. Probabilmente solamente vivendo l'analisi 'da dentro' e non come mezzo per risolvere un esercizio si riesce a cogliere appieno le potenzialità e l'utilizzo degli strumenti che si hanno a disposizione.

Una piccola nota di rammarico sta nel fatto di non essere riuscito a raccogliere un numero maggiore di questionari e quindi di osservazioni per rendere migliore il punto di partenza dell'analisi. I dati così raccolti non sono numerosi e sono anche sproporzionati rispetto alle caratteristiche della popolazione, non aiutando la stabilità dei risultati. Mi auguro comunque che il lavoro svolto e le considerazioni fatte possano servire a chi di dovere, come è stata utile a me, per motivi teorico-pratici, questa esperienza.