

# whitepaper

## **Data Mining im Direktmarketing mit dem prudsys DISCOVERER**



## Inhalt

<b>Einführung</b>	<b>3</b>
80/20 Regel	3
<b>Data Mining im Direktmailing</b>	<b>3</b>
Ende des Massenmarketing	3
<b>Optimierung von Direktmailingaktionen durch Data Mining</b>	<b>4</b>
Phase 1: Profilerstellung	4
Phase 2: Qualifizierung einer Kundendatenbank	5
Interpretierbare Regel	5
Nichtinterpretierbare Regel	6
Minimales Responspotential	7
Gains-Chart	8
Siehe 80/20-Regel	8
<b>Kosten und Profit einer Direktmailingaktion</b>	<b>9</b>
Ausstattungsichte	9
Responserate	9
Kosten-Umsatz-Relation	10
Deckungsbeitrag	10
Break-Even-Point	10
<b>Data Mining für den Außendienst (Ausblick)</b>	<b>16</b>

## Einführung

In vielen Dingen des Lebens gilt, dass 80% eines Resultats mit 20% des Gesamtaufwandes realisierbar sind. Diese auch als Pareto-Prinzip bekannte Regel ist sicherlich nur eine mehr oder weniger vernünftige Schätzung der tatsächlichen Realität. Dennoch lässt sich das Pareto-Prinzip an sehr vielen Beispielen demonstrieren.

Besonders interessant ist das Pareto-Prinzip im unternehmerischen Bereich. Hier werden z. B. in der Regel 80% des Gesamtumsatzes einer Firma durch 20% der Kunden realisiert. Die verbleibenden 20% des Umsatzes verteilen sich dann auf die übrigen 80% der Kunden. Anders gesagt, oft werden die verbleibenden 20% des Umsatzes mit 80% des Gesamtaufwandes erkaufft. Ein Preis, der nicht in allen Fällen lohnend ist. Eine Kundensegmentierung in Top-Kunden (Kunden aus dem 20% Bereich, die den größten Teil des Umsatzes liefern) und sonstige Kunden ist erforderlich, um die Effizienz einer Marketingaktion zu erhöhen. In der Regel ist sogar eine noch genauere Kundensegmentierung sinnvoll, da damit viel individueller auf die Bedürfnisse der Kunden der einzelnen Segmente eingegangen werden kann und somit die zur Verfügung stehenden Mittel für Marketing und Vertrieb gezielter eingesetzt werden können.

### 80/20 Regel

## Data Mining im Direktmailing

Im Direktmailing ist die 80/20 Regel in verstärktem Maße zu beobachten. Bei einer Direktmailingaktion mit einer Auflagenhöhe von 1.000.000 Stück und einer Responserate von ca. 2%, reagieren nur etwa 20.000 angeschriebene Haushalte auf das Angebot. Um genügend Umsatz zu erreichen, muss die Auflagenhöhe der Marketingaktion also sehr hoch gewählt werden. Um dabei profitabel zu sein, müssen die Kosten für eine solche Marketingaktion aber möglichst gering gehalten werden. Dies führte zu einer Entwicklung, die auch unter dem Namen Massenmarketing bekannt ist. Ein wesentlicher Hauptgrund für die geringen Responseraten bei Direktmarketingaktionen liegt aber im Konzept des Massenmarketings selbst. Nicht selten sind Direktmailingaktionen zum Scheitern verurteilt, da aufgrund der geringen Responsequote die hohen Gesamtkosten einer solchen Aktion nicht wieder eingespielt werden können.

Werden Mittel für Marketing nach dem Gießkannenprinzip gestreut, so kann nicht erwartet werden, dass das offerierte Angebot zu den einzelnen Kunden wirklich passt. Diese Angebote wandern zumeist ungelesen in den Papierkorb, da sie nicht auf die wirklichen Bedürfnisse des einzelnen Kunden eingehen. Der dabei entstandene Schaden ist oftmals sogar höher als die Kosten des Anschreibens, da Kunden nicht selten verärgert werden.

Das Überfluten aller Haushalte; respektive Kunden mit Produktofferten und die daraus resultierenden kostenintensiven Streuverlusten gehören der Vergangenheit an, wenn eine Individualisierung durch Kundensegmentierung z. B. durch die Erstellung von Kunden- oder Responseprofilen vorgenommen wird. Mit Data Mining kann das Massenmarketing in individualisierte Bahnen gelenkt werden, was zu gezielteren Marketingaktionen mit höherer Responsequote führt, da das offerierte Angebot viel besser zum Kunden passt.

Fassen wir zusammen: In der Regel spricht ein bestimmtes Angebot nur einen kleinen Teil der Haushalte; respektive Kunden an (Responseraten von 2% oder weniger sind üblich). Warum sollte also ein spezielles Angebot an alle Haushalte einer Adressdatenbank versendet werden? Die entscheidenden Fragen dabei sind:

- Wer sind die typischen Käufer eines bestimmten Produktes?
- Wie können diese typischen Käufer beschrieben werden?
- Welche Haushalte bzw. Kunden entsprechen noch dieser Beschreibung?

Zur Beantwortung dieser und anderer Fragen, liefert Data Mining die passenden Werkzeuge und Methoden. Die Vorgehensweise selbst ist dabei recht einfach.

### Ende des Massenmarketing

## Optimierung von Direktmailingaktionen durch Data Mining

**Phase 1: Profilerstellung** Data Mining im Direktmarketing erfolgt in der Regel in zwei Phasen (Abbildung 1):  
 In der ersten Phase erstellt das Data-Mining-Tool ein Kundenprofil bezüglich des anzubietenden Produktes oder der Dienstleistung. Dazu sind die Erfahrungen (Response-Reaktionen) aus früheren Direktmailingaktionen zu gleichen oder zu ähnlichen Produkten/Dienstleistungen notwendig (Response-Datenbank).  
 Liegen solche Erfahrungsdaten nicht vor, liefert eine Testaktion mit einer kleinen Auflage (5.000-20.000) repräsentativer Haushalte die gewünschten Response-Daten. Tests dieser Art werden bei neuen Produkten oder bei Direktmarketingaktionen in neuen Märkten grundsätzlich im Vorfeld durchgeführt. Diese Tests liefern wichtige Informationen darüber, wie der Kunde das Angebot annimmt und ermöglichen gleichzeitig die Verwendung des Resultats für die Response-Datenbank, auf Grund derer mittels Data Mining ein Kundenprofil bzgl. der Responsequote erstellt werden kann.

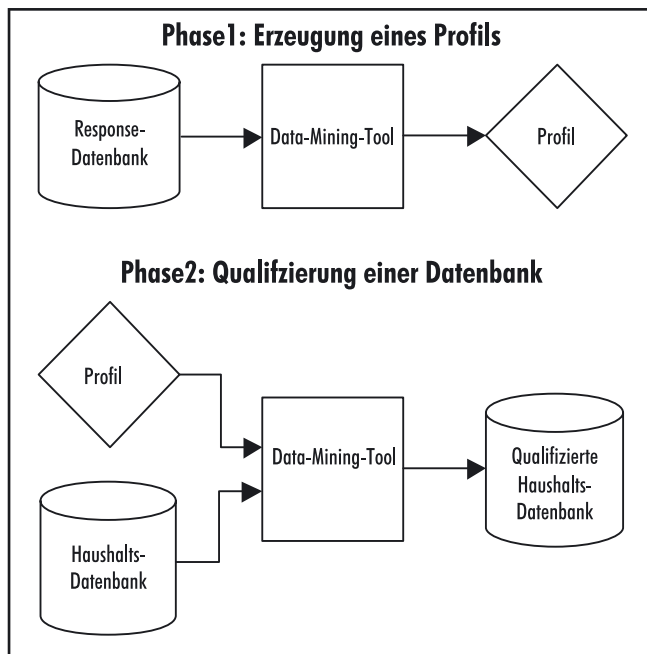


Abbildung 1:  
Data-Mining-Phasen  
beim Direktmarketing

In der zweiten Phase wird das erstellte Profil verwendet, um jedem Haushalt der Kundendatenbank ein Responsepotential zuzuordnen. Dazu werden die Haushalte den Segmenten des mittels Data Mining gewonnenen Profils zugeordnet. Jedes Segment besitzt dabei ein bestimmtes Responsepotential. Die Zuordnung der einzelnen Haushalte zu den Segmenten liefert somit eine Schätzung über das Responseverhalten der einzelnen Haushalte.

Alle Haushalte können anschließend nach diesem Potential absteigend sortiert werden. Es obliegt nun dem Leiter der Direktmarketingaktion zu entscheiden, bis zu welchem Potential er die Haushalte in die Marketingaktion einbeziehen will. Anhand der Kosten einer Direktmarketingaktion und der durchschnittlich erzielten Gewinne einer Bestellung kann das optimale Responsepotential eines Haushalts exakt bestimmt werden, bis zu dem eine Marketingaktion kostendeckend ist. Alle Haushalte, die ein geringeres Responsepotential besitzen, sollten bei der Direktmarketingaktion nicht berücksichtigt werden.

Diese Vorauswahl wird im Allgemeinen als Kundenqualifizierung bezeichnet. Die dabei entstehenden Adresslisten sind bzgl. eines bestimmten Produkts oder einer bestimmten Dienstleistung qualifiziert, da das durchschnittliche Responsepotential dieser Haushalte im Allgemeinen deutlich höher ist als das durchschnittliche Potenzial aller Haushalte.

Das zur Erstellung der Profile notwendige Erfahrungswissen wird dem Data-Mining-Tool als Datenbank (Response-datenbank) zur Verfügung gestellt. Jeder Datensatz entspricht einem Haushalt, der im Rahmen der (Test-)Direktmarketingaktion angeschrieben wurde. Jeder Haushalt muss durch aussagekräftige Merkmale beschrieben werden. Welche Merkmale das sind, hängt ganz vom Produkt oder der angebotenen Dienstleistung ab.

Typische Merkmale sind z. B.:

- personen- und haushaltsbezogene Informationen (z.B. Struktur des Haushaltes, Lifestyle, Alter),
- Wohngebietscharakteristiken (z.B. Postleitzahlen, Wohngegend, soziales Umfeld),
- Interessen der Haushalte (z.B. Auto, Reisen, Tiere, Kultur, Sport),
- Informationen über die Beschäftigungen (z.B. Beruf, Einkommen),
- Konsuminteressen,
- Immobilien,
- Kreditkartentypen.

Darüber hinaus muss in der Response-datenbank zu erkennen sein, ob der betreffende Haushalt auf die Aktion reagiert hat oder nicht (Response). Wie viele Haushalte die Response-datenbank enthalten muss, um ein geeignetes Profil zu erhalten, hängt vom Anwendungsgebiet und der Qualität der Daten ab. Bereits mit 5.000 repräsentativen Datensätzen können sehr gute Profile erstellt werden. Oftmals sind aber auch mehr Datensätze notwendig, da bei 5.000 Haushalten und einer durchschnittlichen Responserate von 2% nur ca. 100 Reagierer vorhanden sind. Diese sind aber nicht in jedem Fall repräsentativ für alle potenziellen Reagierer.

Die erzeugten Profile bestehen aus einer mehr oder weniger großen Anzahl von Segmenten, die durch Regeln beschrieben werden können. Ein Beispiel für ein solches Segment könnte sein:

*Alle Haushalte, die aus mehr als 4 Personen bestehen, deren Jahreseinkommen zwischen 70.000 und 90.000 liegt, die zwei Autos fahren und ein Eigenheim besitzen.*

Dies entspricht der Regel:

*CountPerson > 4 and Income >= 70.000 and Income <= 90.000 and CountCars = 2 and houseowner = true*

Nicht in jedem Fall lassen sich die verschiedenen Segmente so einfach beschreiben.

## Phase 2: Qualifizierung einer Kundendatenbank

### Kundenqualifizierung

### Interpretierbare Regel

**Nichtinterpretierbare Regel**

Folgende Regel wäre aber ebenso denkbar:

$$Income * 0,000393652 - CountPerson * 3,452 + CountCars * 9,2345 > 15,72354$$

Die Form der Regeln und ihre Interpretierbarkeit hängt von den verwendeten Data Mining Verfahren ab.

Bei der Erstellung der Profile werden die Segmente nicht einfach zufällig bestimmt. Sie werden so gewählt, dass das Responsepotential der in dem Segment enthaltenen Haushalte entweder deutlich höher oder deutlich niedriger als das durchschnittliche Responsepotential aller Haushalte ist. Folgendes Beispiel, welches im gesamten Dokument benutzt wird, verdeutlicht dies:



**Beispiel:**

Ein Versandhändler wirbt neue Kunden, indem er ihnen einen Werbebrief (Mailing) zuschickt. Dazu besitzt er Adressen von Personen, die noch nicht Kunde bei ihm sind. Um die Kosten des Mailings zu minimieren, sollen mittels Data Mining die anzuschreibenden Personen qualifiziert werden, um somit die Adressen herauszufiltern, die eine hohe Reagierwahrscheinlichkeit haben. Dazu führt er ein Testmailing mit 10.000 zufällig ausgewählten Adressen durch. Er speichert für jede angeschriebene Person, ob diese reagiert hat oder nicht. Ausgehend davon benutzt er den prudsys DISCOVERER, um einen Klassifikator zu generieren, der die Adressen in Bezug auf die Reagierwahrscheinlichkeit bewertet. Dieser wird auf die insgesamt 28.890 weiteren Adressen in der Datenbank angewendet.

Die Daten zu diesem Beispiel, einschließlich der Aufgabenstellung und der DISCOVERER-Projektdatei *dmc2005.pdp*, finden sich im DISCOVERER-Beispielverzeichnis im Unterverzeichnis *dmc2000*.

Um nun den Klassifikator zu generieren, werden die 10.000 Datensätze in je 5.000 Datensätze zum Generieren und Evaluieren des Klassifikators aufgesplittet. Es sollen nun die 5.000 Datensätze der Trainingsmenge betrachtet werden; die 5.000 Datensätze der Evaluierungsmenge dienen dem optimalen Pruning des Entscheidungsbaummodells.

Aus den Personen der Trainingsmenge konnten 279 neue Kunden gewonnen werden. Die Responserate lag damit bei 5,58%. Alle zur Verfügung stehenden Informationen über die 5.000 Haushalte sowie deren Responseverhalten wurden in der Datenbanktabelle *dmc2000\_train* abgespeichert. Diese Haushalte wurden mit dem DISCOVERER in 9 Segmente zerlegt. (In der Modellbildung einfach „Achsenparallele Trennfunktion“ auswählen und danach den Button „Alles entwickeln“ einmal drücken.) Tabelle 1 gibt einen Überblick über die Größe und das Responsepotential der entstandenen 9 Segmente.

Segmentnummer	Anzahl der Personen	Anzahl der Reagierer	Anzahl der Nichtreagierer	Responsepotential
6	89	34	55	38,20%
12	273	63	210	23,08%
11	433	48	385	11,09%
14	431	35	396	8,12%
10	245	17	228	6,94%
16	437	24	413	5,49%
8	248	9	239	3,63%
15	2.764	49	2.715	1,77%
5	80	0	80	0%
Gesamt	5.000	279	4.721	5,58%

Tabelle 1 Segmente der Responseudatenbank

Auf die Beschreibung der Segmente in Form von Regeln wurde verzichtet, da dies hier nicht relevant ist. Diese können

jedoch im DISCOVERER betrachtet werden bzw. sind in Abbildung 4 dargestellt. Um eine bessere Übersicht zu erhalten, wurden die Segmente nach dem Responsepotential sortiert. Jetzt ist es die Aufgabe des Leiters der Direktmailingaktion zu entscheiden, welche Segmente ein genügend hohes Responsepotential besitzen, um in der richtigen Direktmailingaktion berücksichtigt zu werden. Dazu können die Kosten der Mailingaktion sowie die durchschnittlichen Gewinne eines neuen Kunden herangezogen werden.

Ein Anschreiben kostet im Durchschnitt 6 €. Reagiert die angeschriebene Person, werden durchschnittliche Einnahmen von 95 € erzielt (Anschreibekosten mit eingerechnet). Fixkosten für die Mailingaktion werden in diesem Beispiel nicht berücksichtigt.

Das Segment mit dem höchsten Responsepotential ist Segment 6. Es enthält 89 Personen, von denen 34 reagiert haben. Dieses Kundensegment ist sehr lukrativ. Dem Aufwand von 330 € ( $55 \times 6$  €) stehen Einnahmen von 3.230 € ( $34 \times 95$  €) gegenüber, somit beträgt der Gewinn 2.900 €. Es ist klar, daß dieses Segment in die richtige Direktmailingaktion einbezogen werden sollte. Sind die 5.000 Testhaushalte wirklich repräsentativ, so kann ein ähnliches Responseverhalten (38,20%) für das entsprechende Segment aller anderen Personen angenommen werden. Bei einer Adressdatenbank mit 28.890 Personen werden ca. 514 Personen diesem Segment angehören, von denen ca. 196 Personen reagieren werden. Dies entspricht einem Gewinn von ca. 16.756 €.

Das Segment 5 besitzt mit 0 % das geringste Responsepotential. Diesem Segment gehören 80 Personen an, von denen keine reagierte. Hier ist der Aufwand natürlich größer als die „Einnahmen“. Trotzdem wurden aber 480 € ( $80 \times 6$  €) investiert, welche komplett den Verlust bilden. Bezogen auf die eigentliche Direktmarketingaktion entspricht dies einem Verlust von ca. 2.774 €.

Jedes der Segmente kann auf diese Weise analysiert werden. Tabelle 2 enthält eine Zusammenfassung der Gewinne und Verluste der einzelnen Segmente.

Das mindestens notwendige Responsepotential (in Prozent) eines profitablen Segments kann wie folgt berechnet werden:

#### Minimales Responsepotential

$$\text{Resp}_{\min} = \frac{100 \times \text{Kosten}}{(\text{Kosten} + \text{Einnahmen})}$$

Die Kosten beziehen sich dabei auf die Kosten eines Anschreibens ohne Antwort. In diesem Beispiel 6 €. Die Einnahmen bezieht sich auf den erzielten Erlös eines Reagierers. In diesem Fall sind dies 95 €. Das minimale Responsepotential beträgt somit 5,94%. Werden alle 28.890 Million Haushalte angeschrieben, kann ein Verlust von 10.522 € erwartet werden. Werden stattdessen nur die profitablen Segmente 6, 12, 11, 14 und 10 ausgewählt, kann ein Gewinn von 63.967,5 € erwartet werden. Grafisch kann dies in einem Gains-Chart verdeutlicht werden (**Abbildung 2**).

Segmentnummer	Responsepotential	Anzahl der Personen Testmail	Gewinn in € Testmail	Anzahl der Personen Mail	Gewinn in € Mail
6	38,20%	89	2.900,00	514	16.756,00
12	23,08%	273	4.725,00	1.578	27.301,00
11	11,09%	433	2.250,00	2.502	13.000,00
14	8,12%	431	949,00	2.490	5.483,50
10	6,94%	245	247,00	1.416	1.427,00
16	5,49%	437	-198,00	2.525	-1.144,00
8	3,63%	248	-579,00	1.433	-3.345,50
15	1,77%	2.764	-11.635,00	15.970	-67.227,00
5	0%	80	-480,00	462	-2.773,50
Gesamt	5,58%	5.000	-1821,00	28.890	-10.522,00

Tabelle 2: Gewinne und Verluste in den einzelnen Segmenten

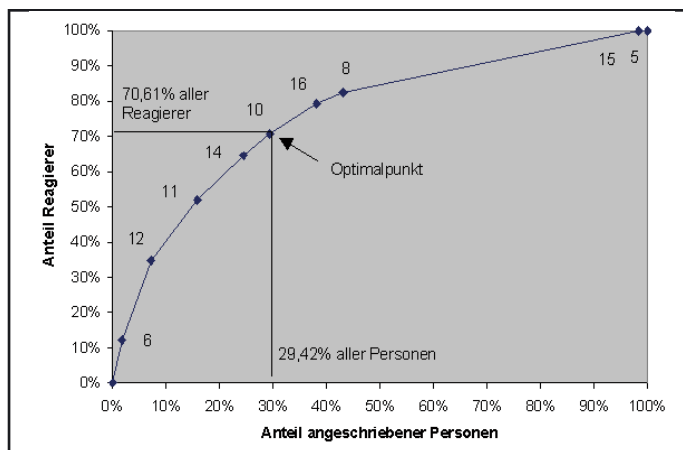


Abbildung 2 Gains-Chart

**Gains-Chart** Gains-Charts gehören zur Gruppe der Business-Charts. Dazu zählen z.B. auch ROC-Chart und Lift-Chart. Im Gains-Chart kann der Aufwand dem Nutzen einer Marketingaktion in einer Kurve gegenübergestellt werden. Die X-Achse stellt dabei den prozentualen Anteil angeschriebener Personen dar (Aufwand). Der Nutzen wird in Form des Anteils potentieller Reagierer (Ertrag) auf der Y-Achse dargestellt.

**Siehe 80/20-Regel** Charakteristisch für die Kurve des Gains-Chart ist der steile Anstieg am Anfang, der sich mit zunehmendem Aufwand allmählich abschwächt. Werden z.B. die 7% der Haushalte mit dem höchsten Responsepotential angeschrieben, erreicht man damit über 34% aller Reagierer. In diesem Bereich ist der Anstieg der Kurve am steilsten. Der Optimalwert für dieses Beispiel wird erreicht, wenn 29,42% aller Haushalte angeschrieben werden (Segmente 6, 12, 11, 14 und 10). Mit diesem Aufwand können bereits 70,61% aller Reagierer erreicht werden. Um die restlichen 29,39% der Reagierer zu erreichen (Segmente 16, 8, 15 und 5), ist ein weiterer Aufwand von 70,58% notwendig (80/20-Regel). Der Anstieg der Kurve ist in diesem Bereich nur noch gering.

Je geringer der Anstieg der Gains-Chart-Kurve ist, um so größer ist der Aufwand weitere Reagierer zu erreichen, da das Responsepotential in diesen Bereichen gering ist. Ab einem bestimmten Punkt lohnt der weitere Aufwand nicht mehr, da der Mehrgewinn die zusätzlich anfallenden Kosten nicht mehr rechtfertigt. Der Erfolg einer Data-Mining-Analyse im Direktmarketing hängt entscheidend davon ab, wie genau dieser Optimalpunkt ermittelt werden kann. Können exakte Kosten für ein Mailing und der durchschnittliche Gewinn eines Reagierers angegeben werden, kann dieser Punkt sehr genau bestimmt werden.

## Kosten und Profit einer Direktmailingaktion

Kosten und Profit spielen bei der Planung einer Direktmailingaktion eine entscheidende Rolle. Sie entscheiden über den Erfolg oder Misserfolg einer Marketingaktion. Im Vorfeld einer Mailingaktion muss anhand der Kosten und prognostizierten Nettoumsätze entschieden werden, ob die Marketingaktion profitabel wird und ein Return on Investment erwartet werden kann. Die Entscheidungen, die der Leiter einer Direktmailingaktion zu fällen hat, betreffen zum einen die Festlegung des Werbebudgets, die Ausstattungsdichte von Werbemitteln sowie deren Verteilung auf die unterschiedlichen Kundensegmente.

In diesem Abschnitt werden die verschiedenen Kosten, die beim Direktmailing unterschieden werden, erläutert. Begriffe wie Ausstattungsdichte, Deckungsbeiträge, Kosten-Umsatz-Relation und Break-Even-Point werden behandelt. Anhand eines ausführlichen Beispiels wird gezeigt, wie eine Direktmailingaktion geplant und deren Erfolg prognostiziert werden kann.

Die Ausstattungsdichte ist einer der entscheidendsten Faktoren für den Erfolg einer Marketingaktion. Die Ausstattungsdichte ergibt sich aus der Relation zwischen der Gesamtauflage eines Werbemittels wie z.B. einem Katalog oder einem Werbemailing und der Gesamtanzahl aller potentiell zur Verfügung stehenden Haushalte einer Kundendatenbank. Die Kosten einer Direktmailingaktion werden unmittelbar durch die Ausstattungsdichte beeinflusst. Hohe Ausstattungsdichten führen zwangsläufig auch zu hohen Kosten.

Die Responderate entspricht dem Anteil an Haushalten, die auf die Mailingaktion mit einer Bestellung, dem Abschluss eines Abonnements oder Ähnlichem reagieren. Bei unqualifizierten Direktmailingaktionen liegen die Responderaten oftmals bei weniger als 2%. Die Responderate hat unmittelbaren Einfluss auf den zu erwartenden Nettoumsatz und damit den Profit der Mailingaktion.

Zwischen der Ausstattungsdichte und der Responderate besteht ein unmittelbarer Zusammenhang. Die Haushalte einer Kundendatenbank sind bezüglich ihrer Responsewahrscheinlichkeit nicht gleich. Beschränkt man eine Mailingaktion ausschließlich auf die Haushalte mit einer hohen Responsewahrscheinlichkeit, so führt dies zu einer hohen Responderate. Die Ausstattungsdichte ist dabei gering. Auf der anderen Seite führt eine geringere Ausstattungsdichte zu einer absolut gesehen geringen Anzahl von Reagierern und damit zu einem geringen Nettoumsatz.

Wird die Ausstattungsdichte deutlich erhöht, führt dies zu einer geringen Responderate, da auch Haushalte mit einer geringen Responsewahrscheinlichkeit angeschrieben werden. Die absolute Anzahl der Reagierer steigt, was zu einer Umsatzerhöhung führt. Mit der Erhöhung der Ausstattungsdichte erhöhen sich aber auch die Kosten dieser Aktion deutlich. Der zusätzliche Umsatz muss mit deutlich erhöhten Kosten bezahlt werden. Beide Fälle haben unter bestimmten Bedingungen ihre Existenzberechtigung.

Ist der durchschnittlich erzielte Gewinn eines Reagierers im Gegensatz zu den Kosten eines Mailings hoch, sollte eine hohe Ausstattungsdichte gewählt werden, da die großen Streuverluste durch die Reagierer abgedeckt werden können. Ist der Gewinn eines Reagierers sehr hoch, kann es sogar sinnvoll sein, auf eine Kundenqualifizierung ganz zu verzichten. Die Ausstattungsdichte beträgt dann 100%.

Ist hingegen der erwartete durchschnittliche Gewinn eines Reagierers im Gegensatz zu den Kosten gering, sollte eine geringe Ausstattungsdichte gewählt werden. Im Extremfall kann es sogar sinnvoll sein, auf die gesamte Mailingaktion zu verzichten, wenn das Kosten/Nutzen-Verhältnis besonders ungünstig ist. Die Ausstattungsdichte würde somit 0% betragen.

### Ausstattungsdichte

### Responderate

**Kosten-Umsatz-Relation**

Ein wichtiges Maß für die Bewertung der Effizienz des Werbemittelleinsatzes ist die Kosten-Umsatz-Relation (KUR), die den prozentualen Werbemittelanteil am Nettoumsatz ausdrückt:

$$KUR = \frac{100\% \times \text{Werbekosten}}{\text{Nettoumsatz}}$$

Je geringer die KUR ist, um so geringer ist der notwendige Werbemittelleinsatz, um einen bestimmten Nettoumsatz zu erzeugen. Eine KUR von beispielsweise 10% bedeutet, dass mit 10 € Werbemittel 100 € Nettoumsatz erzielt werden können.

**Deckungsbeitrag**

In Abhängigkeit von den Nettoerlösen und den Kosten für die Werbemittel ergeben sich die Deckungsbeiträge I und II (DB I bzw. DB II). Sie werden wie folgt definiert:

$$DB I = \text{Nettoerlös} - \text{variable Kosten}$$

$$DB II = \text{Nettoerlös} - \text{variable Kosten} - \text{Werbekosten}$$

**Abbildung 3** stellt den Zusammenhang zwischen der Ausstattungsdichte, dem Nettoumsatz und den Deckungsbeiträgen grafisch dar. Der Nettoumsatz steigt im Anfangsbereich stark an. In diesem Bereich befinden sich die Kunden mit der größten Responsewahrscheinlichkeit. Mit größer werdender Ausstattungsdichte kommen immer mehr Kunden mit einer geringeren Responsewahrscheinlichkeit hinzu. Der Anstieg der Nettoumsatzkurve flacht damit ab. Die Werbungskosten steigen hingegen linear an. Im Gegensatz zum Nettoumsatz beginnt die Kurve der Werbungskosten nicht bei 0, da die Fixkosten einer Werbeaktion hier berücksichtigt werden müssen. Der Deckungsbeitrag I verhält sich proportional zur Nettoumsatzentwicklung. Aus der Differenz des DB I und den Werbungskosten ergibt sich der DB II. Aufgrund der Werbefixkosten ist dieser anfangs negativ.

**Break-Even-Point**

Die erste gestrichelte Linie markiert den Break-Even-Point, die Schwelle der Wirtschaftlichkeit. Ab diesem Punkt deckt der Rücklauf einer Direktwerbeaktion die entstandenen Kosten voll ab.

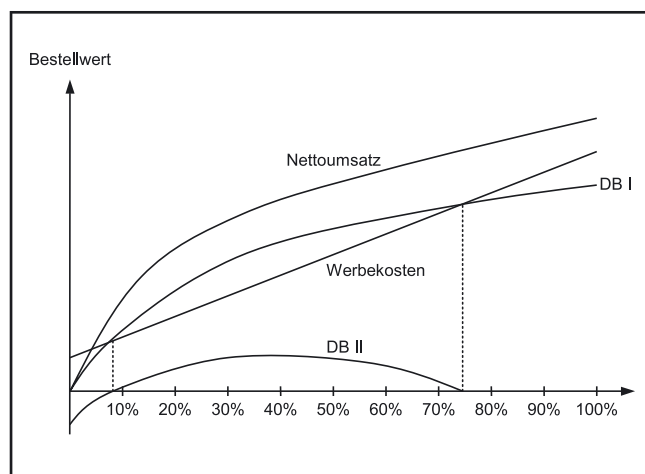


Abbildung 3  
Zusammenhang zwischen  
Ausstattungsdichte,  
Nettoumsatz und  
Deckungsbeiträgen

Die zweite gestrichelte Linie markiert den Punkt, ab dem die Direktmarketingaktion nicht mehr wirtschaftlich ist, da die steigenden Kosten von den fallenden Umsätzen nicht mehr gedeckt werden können.

Beispiel: Es soll nun das eingangs betrachtete Beispiel des Versandhändlers weiter vertieft werden. Hierbei solle die Mailingaktion dem Vertrieb eines Buches dienen. Dazu stand eine Adressdatenbank zur Verfügung, die durch zusätzliche Adressen eines Adressverlages angereichert wurde. Insgesamt waren damit 38.890 Adressen für diese Aktion verfügbar. Im Vorfeld wurden im Rahmen eines Tests bereits 10.000 Personen angeschrieben, wovon wieder die Trainingsmenge von 5.000 Personen betrachtet werden soll. 279 Bestellungen erfolgten auf diesen Test. Die Responserate lag somit bei 5,58%. Der Verkaufspreis des Buches wurde auf 141,71 € festgelegt. Zur Präsentation des Buches wurde von einer Werbeagentur ein Werbekonzept entworfen, welches folgende Bestandteile enthält:



- einen doppelseitigen A4-Brief,
- einen zusätzlichen lift-letter (Kurzbrief) im Format A6,
- eine vierfarbige Broschüre mit sechs Fotografien,
- einen zweifarbigen Briefumschlag mit Fenster und
- eine einfarbige Rückantwortkarte mit Briefumschlag.

Die für dieses Werbekonzept entstandenen Fixkosten stellt **Tabelle 3** dar.

Aufgaben		Kosten
Kreativität:	Texte entwerfen und redigieren	8.250 €
	Design	6.000 €
	Typographie	450 €
	Scannen von Bildern	350 €
Bilder und Illustrationen	Fotografie	950 €
	Modell	0 €
	Bildbearbeitung	450 €
	Illustration	0 €
Sonstige Kosten	Administration	2.000 €
Gesamtkosten		18.450 €

Tabelle 3  
Werbefixkosten  
für eine Direktmailingaktion

Die variablen Kosten sind in **Tabelle 4** dargestellt. Bei den Gesamtkosten wurde eine Auflagenhöhe von 38.890 Stück zugrundegelegt.

Die variablen Kosten, die bei einer Bestellung anfallen, sind in **Tabelle 5** dargestellt. Die hier dargestellten Portokosten beziehen sich auf die Rückportokosten der Antwortbriefe. Darüber hinaus wurde eine Servicenummer für die Entgegennahme von Bestellungen eingerichtet.

Die aufgelisteten variablen Kosten und die Kosten einer Bestellung können aufgrund von Rabattstaffelungen je nach Auflagenhöhe variieren. Dies wurde in diesem vereinfachten Beispiel nicht berücksichtigt.

Aufgaben		Gesamtkosten	Kosten für 1000 Mails
Druck	Farbseparierung	4.500 €	115,71 €
	Brief	1.100 €	28,28 €
	Kurzbrief	800 €	20,57 €
	Bestellformular/Antwortkarte	800 €	20,57 €
	Broschüre	90.000 €	2.314,22 €
	Katalog	0 €	0 €
	Briefumschlag	3.000 €	77,14 €
	Briefumschlag für Rückantwort	1.000 €	25,71 €
	Sonstiges	51.140 €	1.314,99 €
Adressen	Adressmiete	29.000 €	745,69 €
	Bereinigen, Korrektur, Vervollständigung und Doublettenabgleich	3.000 €	77,14 €
Lettershop	Einfügung von Namen	1.500 €	38,58 €
	Adressierung	2.000 €	51,42 €
	Portooptimierung/Sortierung und Versand	1.500 €	38,58 €
Portokosten	Postausgang	16.000 €	411,42 €
Telefonmarketing	Telefonkosten für Verkaufsgespräche	28.000 €	719,98 €
Gesamtkosten		233.340 €	6.000,00 €

Tabelle 4  
Variable Werbekosten einer Direktmailingaktion bei einer Auflagenhöhe von 38.890 Stück

Aufgaben		Kosten pro Bestellung
Portokosten	Posteingang	0,65 €
Telefonmarketing	Telefonkosten für 0800-Service Nummer	0,76 €
Produktkosten	Herstellungskosten bzw. Einkaufspreis	36,80 €
	Versand, Verpackung und Versicherung	2,50 €
Gesamtkosten		40,71 €

Tabelle 5  
Kosten für eine Bestellung inklusive der Herstellungskosten

Die drei Kostentabellen können als Direktmailing-Checklisten verwendet werden. Einige Positionen sind für dieses Beispiel nicht notwendig, an diesen Stellen wurden keine Kosten eingetragen.

Die 5.000 Personen wurden mit dem DISCOVERER und einem Entscheidungsbaumverfahren in 9 Segmente zerlegt.

**Abbildung 4** zeigt den generierten Entscheidungsbaum mit den 9 Segmenten. In **Tabelle 6**, welche der **Tabelle 1** entspricht, sind nochmals alle 9 Segmente, nach dem Responsepotential sortiert, aufgelistet.

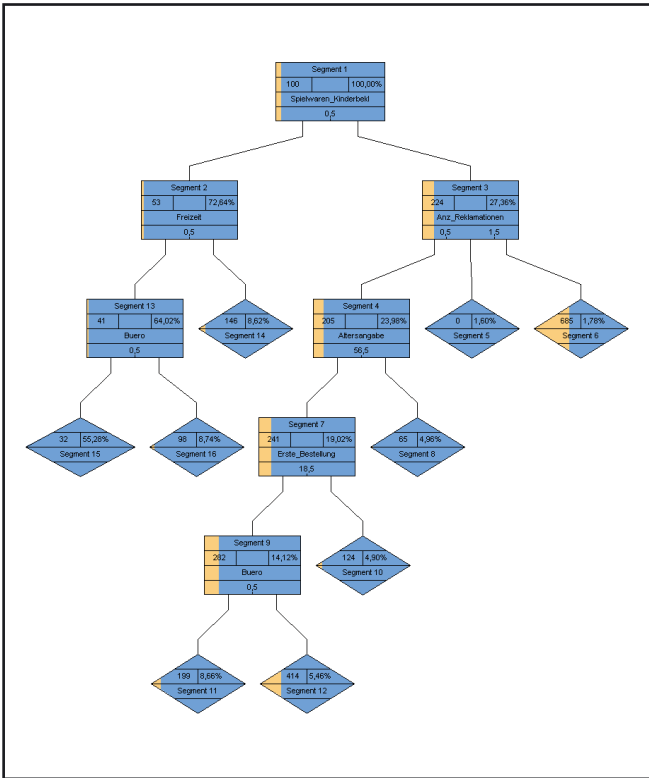


Abbildung 4  
Entscheidungsbaum mit  
9 Segmenten

Segmentnummer	Anzahl der Personen	Anzahl der Reagierer	Anzahl der Nichtreagierer	Responserate
6	89	34	55	38,20%
12	273	63	210	23,08%
11	433	48	385	11,09%
14	431	35	396	8,12%
10	245	17	228	6,94%
16	437	24	413	5,49%
8	248	9	239	3,63%
15	2.764	49	2.715	1,77%
5	80	0	80	0%
Gesamt	5.000	279	4.721	5,58%

Tabelle 6  
Segmente  
der Responserdatenbank

Ausgehend von den Ergebnissen der 5.000 Haushalte der Testaktion kann auf das Kaufverhalten der 28.890 Personen der Adressdatenbank geschlossen werden. Dem Segment 6 wurden 89 Testpersonen zugeordnet. Davon bestellten 34 Haushalte das Buch. Dies entspricht einer Responserate von 38,20%. Hochgerechnet auf 28.890 Personen werden ca. 514 Personen dem Segment 15 angehören. Bei einer Responserate von 38,20% ist mit ca. 196 Bestellungen zu rechnen. Werden in der Direktmailingaktion ausschließlich Personen aus diesem Segment berücksichtigt, ergibt sich eine Ausstattungsdichte von 1,78%.

Die Werbekosten, die sich aus den Werbefixkosten von 18.450 € und den variablen Werbekosten für 514 Mailings (3.084 €) zusammensetzen, belaufen sich auf 21.534 €. Der Nettoumsatz ergibt sich aus dem Erlös von 196 Buchverkäufen (27.775,16 €). Die variablen Herstellungskosten für das Buch belaufen sich auf 7.979,16 €. Damit ergibt sich ein Realverlust von 1.738 €.

Erweitert man nun die Auswahl um das Segment 12, erhöht sich die Ausstattungsdichte auf 7,24%. Es fallen 31.002 € Werbekosten und 22.797,60 € Herstellungskosten an. Der Deckungsbeitrag erhöht sich auf 25.558 €. Mit diesen beiden Segmenten wird der Break-Even-Point bereits erreicht. **Tabelle 7** stellt die Kosten und den Erlös für alle Segmente dar.

Segmentnummer	Ausstattungs-dichte	Werbekosten	Nettoumsatz	DB I	DB II
6	1,78%	21.534,00 €	27.775,16 €	19.796,00 €	-1.738,00 €
12	7,24%	31.002,00 €	79.357,60 €	56.560,00 €	25.558,00 €
11	15,90%	46.014,00 €	118.611,27 €	84.537,00 €	38.523,00 €
14	24,52%	60.954,00 €	147.236,69 €	104.939,00 €	43.985,00 €
10	29,42%	69.450,00 €	161.124,27 €	114.837,00 €	45.387,00 €
16	38,16%	84.600,00 €	180.821,96 €	128.876,00 €	44.276,00 €
8	43,12%	93.198,00 €	188.190,88 €	134.128,00 €	40.930,00 €
15	98,40%	189.018,00 €	228.294,81 €	162.711,00 €	-26.307,00 €
5	100,00%	191.790,00 €	228.294,81 €	162.711,00 €	-29.079,00 €

Tabelle 7  
Ausstattungs-  
dichte, Kosten  
und Profit der  
Direktmailing-  
aktion

Der maximale Deckungsbeitrag von 45.387,00 € kann erreicht werden, wenn die Haushalte der Segmente 6, 12, 11, 14 und 10 in die Direktmailingaktion einbezogen werden. Für die KUR ergibt sich ein Wert von 43,10%. **Abbildung 5** zeigt den Einfluss der Ausstattungsdichte auf die Nettoumsätze, Werbekosten und Deckungsbeiträge.

Die Bestimmung der optimalen Ausstattungsdichte ist bei komplexeren Marketingaktionen oftmals schwierig, da die Anzahl der verschiedenen Segmente nicht selten mehr als 50 oder 100 Segmente beträgt. Der DISCOVERER enthält ein Modul, das Sie bei der Analyse Ihrer Direktmailingaktionen unterstützt. **Abbildung 6** zeigt dieses Modul für die 5000 Personen der Trainingsmenge.

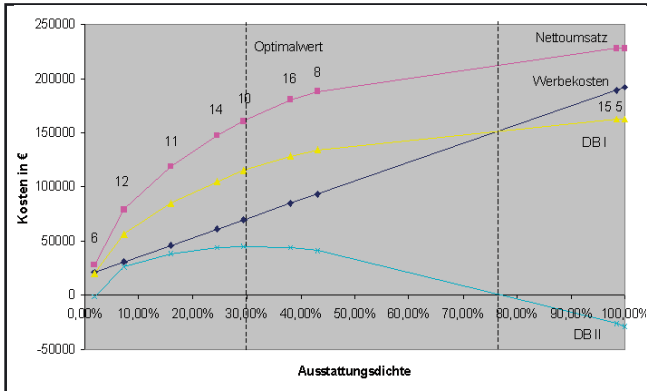


Abbildung 5  
Ausstattungsichte und  
realisierte Gewinne

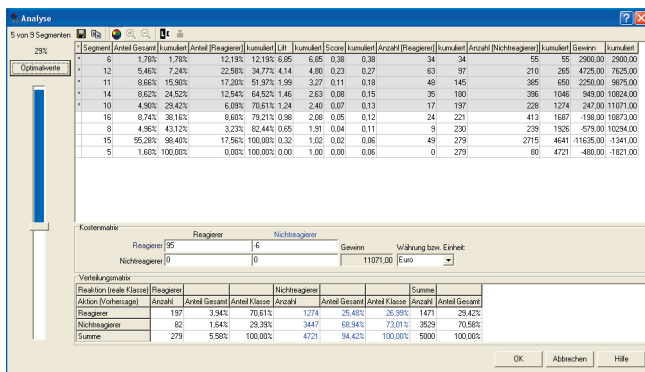


Abbildung 6  
Kostenanalyse mit dem  
prudsys DISCOVERER

Im oberen Teil der Abbildung wird eine Tabelle dargestellt, die für verschiedene Ausstattungsdichten die Anzahl der Reagierer und Nichtreagierer, den Anteil aller potentiellen Reagierer, den Lift sowie weitere Kennzahlen anzeigt. Der Lift stellt ein weiteres Maß für die Bewertung der Segmente dar und gehört zur Gruppe der Business-Charts.

Über den linken Schieberegler kann die Ausstattungsdichte per Hand verändert werden. In der Verteilungsmatrix werden die Veränderungen der Ausstattungsdichte, der Reagierer und Nichtreagierer sofort angezeigt. Für diese Direktmarketingaktion wurde die optimale Ausstattungsdichte von 29% benutzt. Dies entspricht 1.471 Mailings. Die Anzahl der Reagierer und Nichtreagierer ist in der dritten Zeile der Verteilungsmatrix dargestellt. Von den 1.471 angeschriebenen Personen haben 197 Personen das Buch bestellt, während 1.274 Personen nicht reagiert haben. Die vierte Zeile stellt die Verteilung der restlichen nichtangeschriebenen Personen dar. Darin enthalten sind 82 Reagierer sowie 3.447 Nichtreagierer.

Im oberen Teil können für die verschiedenen Konstellationen Kosten definiert werden. Grundsätzlich können vier verschiedene Fälle betrachtet werden:

1. Eine Person, die angeschrieben wurde, bestellt das Buch (Reagierer).
2. Eine Person, die angeschrieben wurde, bestellt kein Buch (Nichtreagierer).
3. Eine Person, die nicht angeschrieben wurde, hätte auf die Mailingaktion reagiert und das Buch bestellt.
4. Eine Person, die nicht angeschrieben wurde, hätte auch nicht auf die Mailingaktion reagiert.

Im ersten Fall kann ein Nettoumsatz von 141,71 € verzeichnet werden. Abzüglich der Werbekosten für das Mailing (6 €) und den Herstellungskosten des Buches (40,71 €) kann eine Nettoeinnahme von 95 € erwartet werden.

Im zweiten Fall fallen nur die Kosten des Mailings an, da keine Bestellung erfolgt. Die Kosten betragen somit 6 €.

Die Fälle drei und vier verursachen keinerlei Kosten oder Gewinne. Als Kosten können 0 € verwendet werden.

Insgesamt kann somit ein Gewinn von 11.071,00 € erzielt werden.

Die Fixkosten der Werbeaktion werden in diesem Modul nicht berücksichtigt. Der Erlös der Mailingaktion vermindert sich somit noch um diese Fixkosten. Zur optimalen Bestimmung der Ausstattungsdichte kann der Schalter *Optimalwerte* verwendet werden. Anhand der definierten Kosten wird die optimale Ausstattungsdichte berechnet. In vorliegenden Fall liegt sie bei 29%.

## Data Mining für den Außendienst (Ausblick)

Die Idee des Einsatzes von Data Mining im Direktmailing lässt sich auch auf andere Bereiche des Direktmarketings wie z. B. dem Außendienst anwenden.

Auch im Außendienst ist es nicht sinnvoll ein bestimmtes Produkt allen Kunden gleichermaßen zu offerieren. Eine Kundensegmentierung und damit eine Qualifizierung der Kunden ist hier besonders wichtig, da die Kosten eines Kundenbesuchs natürlich deutlich höher sind als die anfallenden Kosten eines Mailings bei einer Postwurfsendung. Oftmals belaufen sich die Kosten eines Kundenbesuchs mit einer individuellen Beratung auf mehr als 250,00 €.

Für den Außendienst kann Data Mining Fragen folgender Art beantworten:

- Welche Kunden interessieren sich für ein bestimmtes Produkt?
- Welchen Preis sind die Kunden bereit für ein bestimmtes Produkt zu zahlen?
- Wo liegen Cross- und Up-Selling Potenziale?

Die Vorgehensweise bei der Durchführung einer Data-Mining-Analyse ist hier ähnlich zur Vorgehensweise beim Direktmailing:

1. Bei Bedarf Durchführung einer Testaktion z. B. einer Umfrage bei einer repräsentativen Kundengruppe. Die Rücklaufquote beläuft sich dabei erfahrungsgemäß auf ungefähr 5 bis 10 Prozent.
2. Erstellung eines Profils mit einem Data-Mining-Tool durch Segmentierung.
3. Kostenanalyse der verschiedenen Segmente anhand ihrer Reagierwahrscheinlichkeit.
4. Qualifizierung (Zuordnung zu den Segmenten) der Kunden z. B. zur Bestimmung, welche Kundengruppen in der Zukunft durch den Außendienst und welche nur telefonisch durch den Innendienst betreut werden.



prudsys AG  
Bergstraße 61, D-09113 Chemnitz  
Tel.: 0371 27093-0  
Fax: 0371 27093-90  
WWW: [www.prudsys.de](http://www.prudsys.de)  
Hotline: [support@prudsys.de](mailto:support@prudsys.de)

© 2006 prudsys AG.  
Alle Rechte vorbehalten.