

# TNS EX-A-MINE TextMiner

Strategien gegen den  
Information Overload





## Text Mining extrahiert wesentliche Informationen

- In den letzten Jahren hat sich der Information Overload vervielfacht (netzbasierende Kommunikation, nahezu unbegrenzte Speicherkapazitäten, ...)
  - Für Unternehmen wird es zunehmend schwieriger, relevante Informationen zu filtern
  - Offene Antworten bei Befragungen: hoher Codierungsaufwand, keine systematische Quantifizierung
- ➔ **Lösung: automatisierte Inhaltsanalyse durch den EX·A·MINE TextMiner**



### ■ Call Center Kontakte

- Welches sind die Hauptanliegen und -probleme der Kunden, wenn sie mit dem Call Center in Verbindung treten?
- Welche Beschwerden und welche Problemkonstellationen treten häufig gemeinsam auf?

### ■ Offene Antworten bei Befragungen

- Welche Sachverhalte bewerten die Probanden positiv, welche negativ?
- Welche Begriffe werden häufig in Kombination verwandt und welche lassen sich zu übergeordneten Kategorien zusammenfassen?

### ■ Analyse von Interviews und Dokumenten

- Auf welche Themenkategorien lassen sich (auch längere) Texte / Dokumente / Interviews / Emails automatisiert zusammenfassen?

# Text Mining unterstützt analytisches CRM

## Generierung geschäftskritischen Wissens

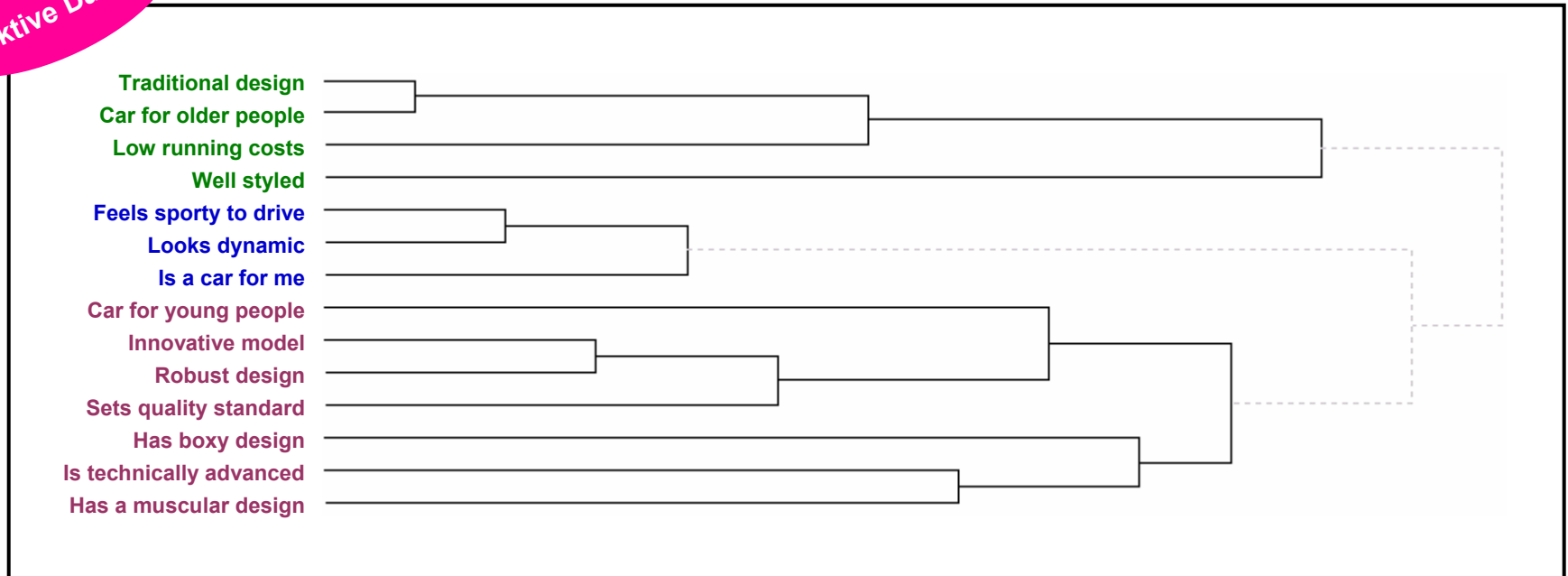


- EX·A·MINE TextMiner quantifiziert qualitative Daten mittels modernster Data Mining-Algorithmen
  - Automatisierte Kategorisierung von Text durch Wörterbücher oder manuelle Codierung
  - Vielfältige explorative Datenanalysen und grafische Aufbereitung der Ergebnisse
  - Analyse von Zusammenhängen mit Variablen wie Geschlecht, Alter etc.
  - Gewonnene Erkenntnisse können mit Hilfe verschiedener multivariater Verfahren weiterverarbeitet werden, (z.B. Clusteranalysen, MDS, Mappings)
- ➔ **Generierung geschäftskritischen Wissens, das auch für das analytische CRM nutzbar wird**

# Ergebnisbeispiel: Hierarchische Clusteranalyse



(fiktive Daten)

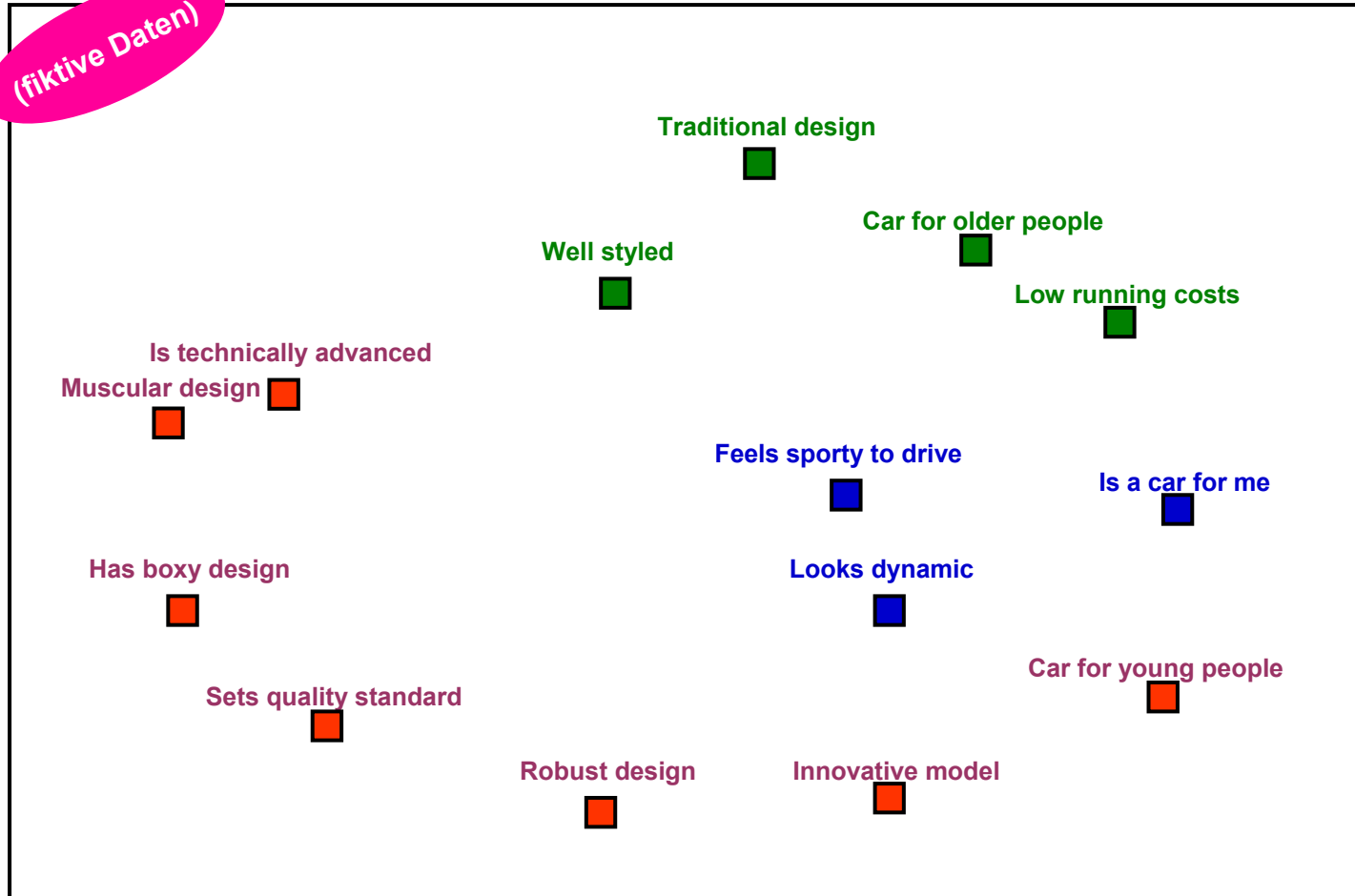


- Das **Dendrogramm** zeigt Wortähnlichkeiten bzw. gemeinsames Auftreten von Begriffen.

# Ergebnisbeispiel: Multidimensionale Skalierung



(fiktive Daten)



- **Multidimensionale Skalierung** wird verwendet, um Ähnlichkeiten von Schlüsselbegriffen zu analysieren.

# Ergebnisbeispiel: Kreuztabellen

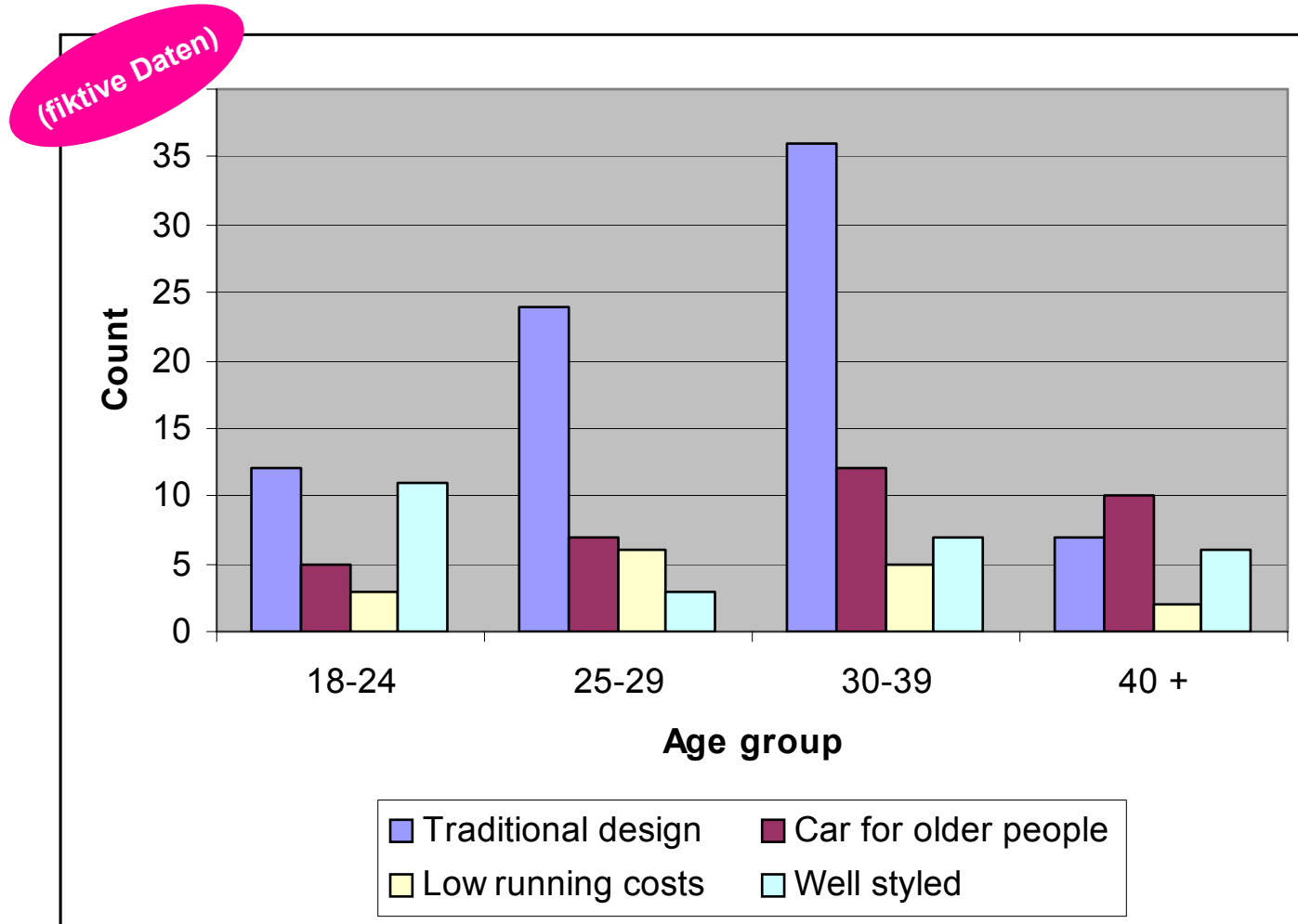


(fiktive Daten)

	Age groups			
	18-24	25-29	30-39	40 +
Traditional design	12	24	36	7
Car for older people	5	7	12	10
Low running costs	3	6	5	2
Well styled	11	3	7	6
Feels sporty to drive		3	1	1
Looks dynamic	1		3	1
Is a car for me	11	14	13	5
Car for young people	5	6	5	4
Innovative model	3	4	7	5
Robust design	12	19	26	2
Sets quality standard	2	3	7	1
Has boxy design	1		8	
Is technically advanced	3	2	4	2
Has a muscular design	2	7	7	6

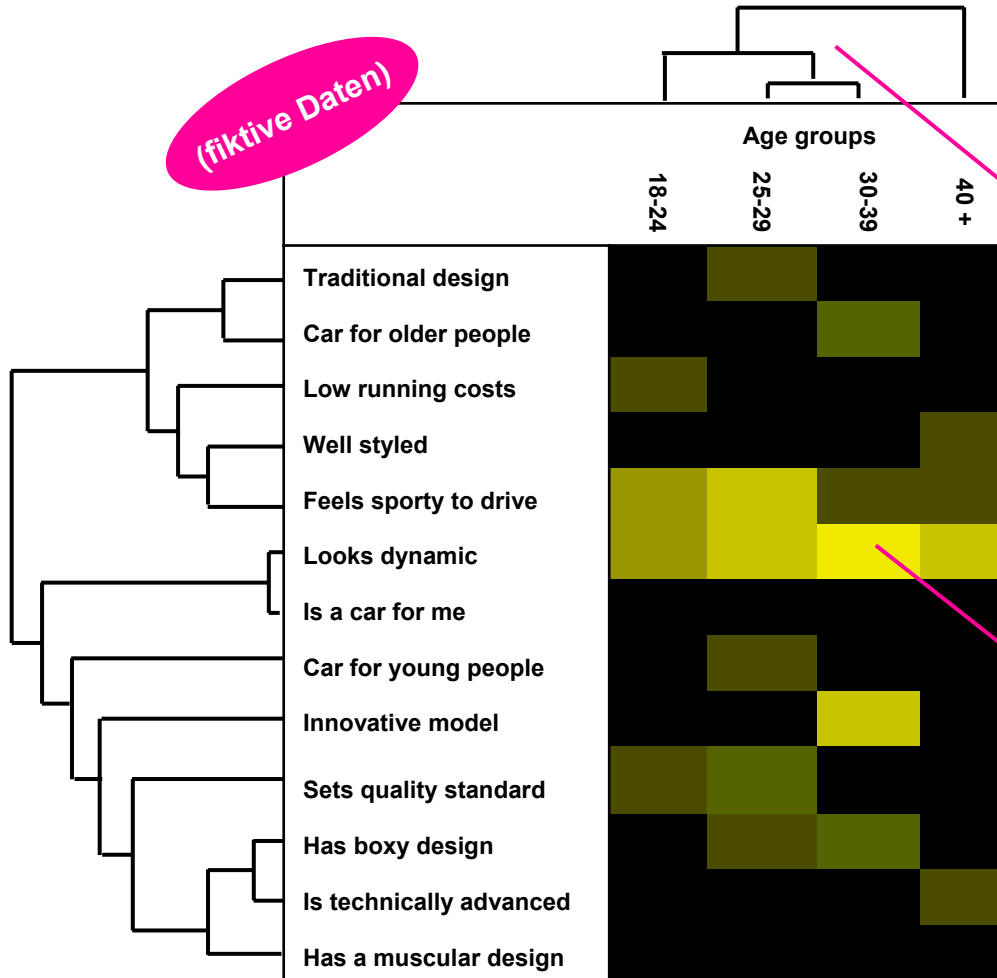
- Mithilfe von **Kreuztabellen** können Beziehungen zwischen Wörtern oder Begriffskategorien und Ausprägungen einer unabhängigen Variable untersucht werden (in diesem Fall Altersgruppen, man könnte aber auch Automarken, Geschlecht etc. verwenden).

# Ergebnisbeispiel: Balkendiagramme



- Ausgehend von Kreuztabellen können **Balkendiagramme** verwendet werden, um die Beziehungen zwischen Begriffen und kategorialen Variablen zu analysieren.

# Ergebnisbeispiel: Hitzebilder (Heatmaps)



Die **Dendrogramme** links und oben zeigen die Ähnlichkeiten zwischen Begriffsnennungen und Altersgruppen.

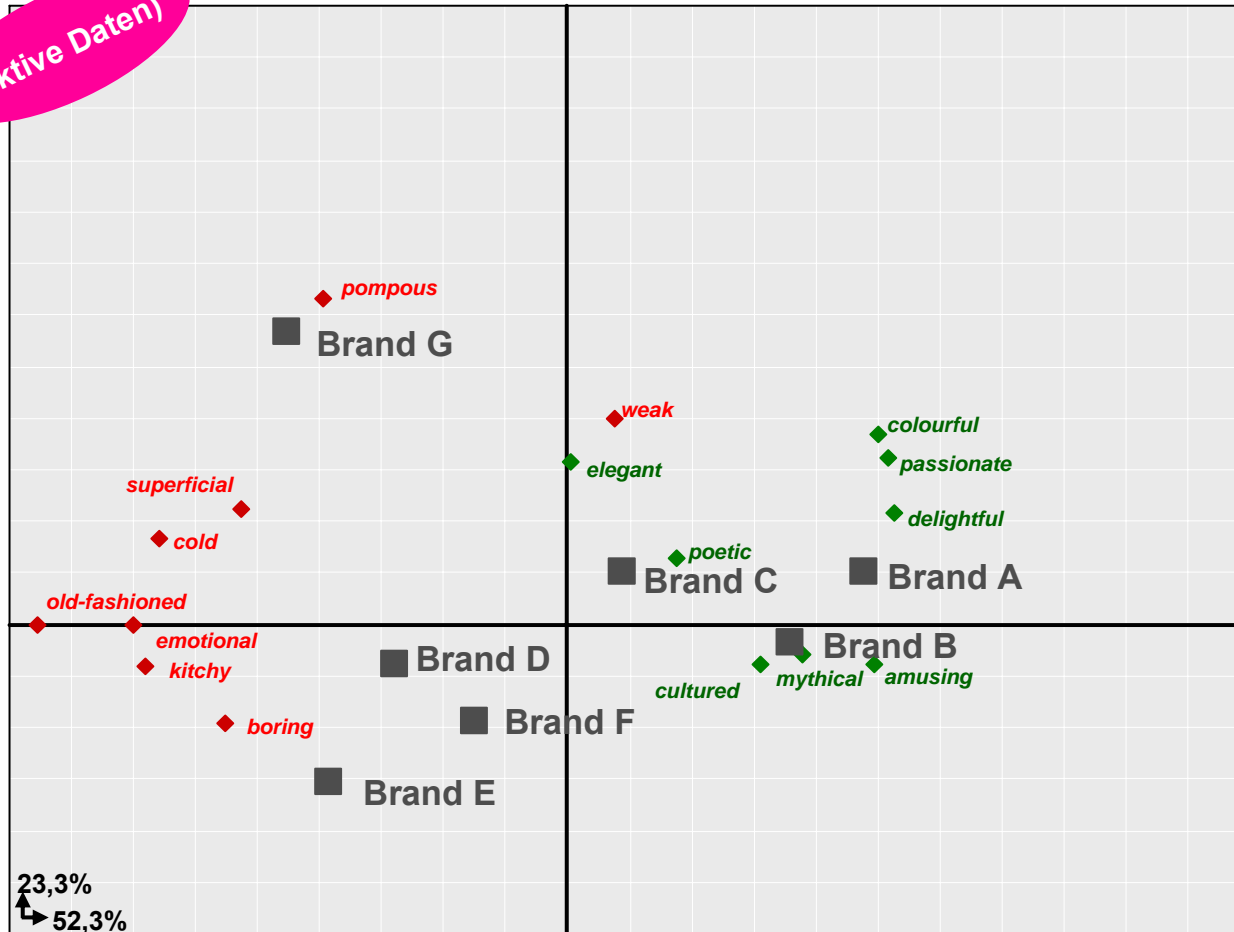
**Lesebeispiel:** ein helles Gelb bedeutet, dass der Begriff „looks dynamic“ hauptsächlich mit der Altersgruppe zwischen 30 und 39 assoziiert wird.

- **Hitzebilder** bieten eine weitere Möglichkeit, Beziehungen in großen Tabellen grafisch zu untersuchen. In einem Hitzebild werden **relative Häufigkeiten** durch verschiedene Farbtöne oder Farbhelligkeiten dargestellt.

# Ergebnisbeispiel: Korrespondenzanalyse



(fiktive Daten)

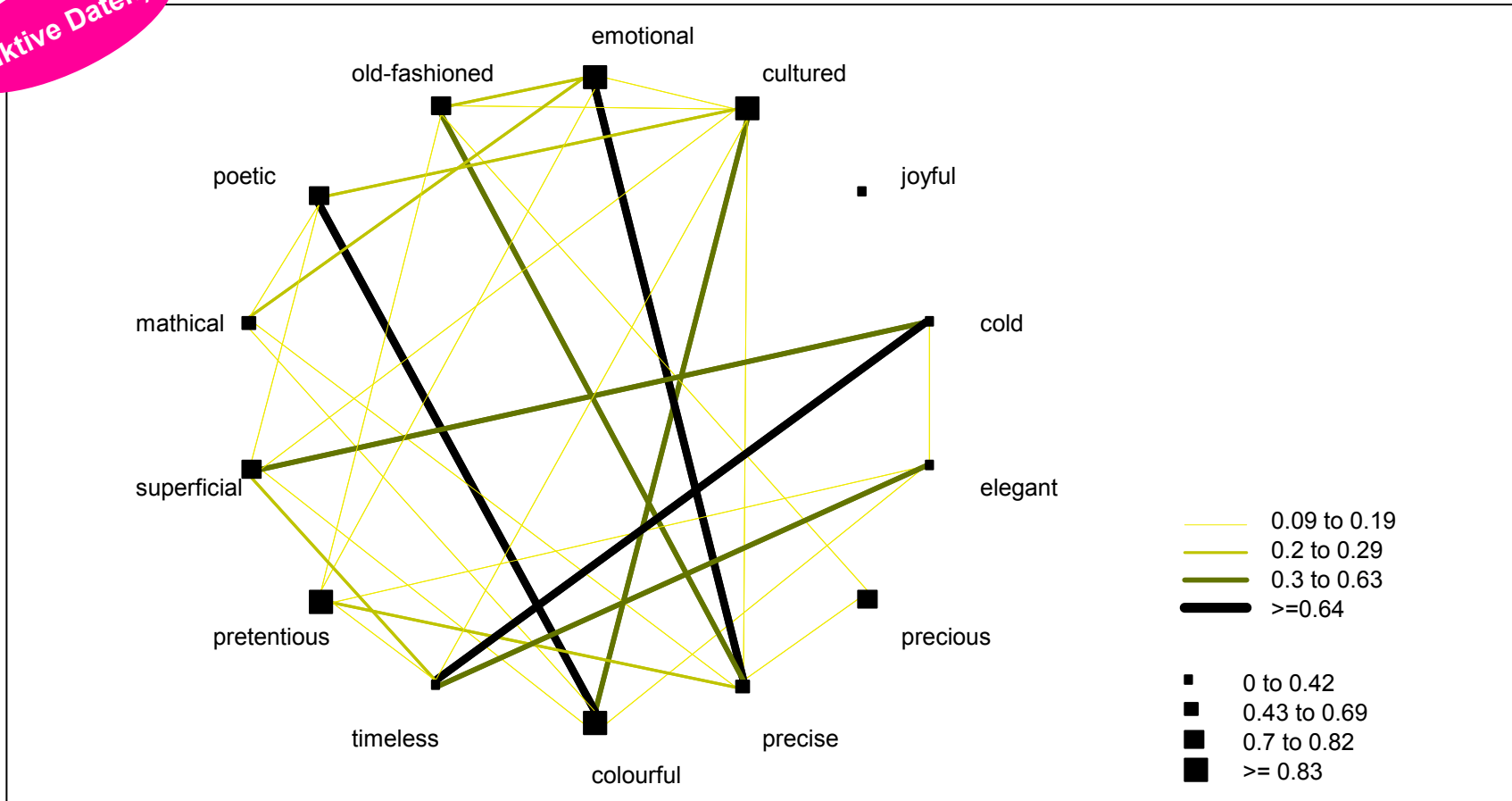


- Die **Korrespondenzanalyse** kann dazu verwendet werden, Beziehungen zwischen Objekten (Marken, Produkte) und Begriffen grafisch darzustellen.

# Ergebnisbeispiel: Netzknoten



(fiktive Daten)



- **Netzknoten** zeigen, welche Begriffe häufig zusammen genannt werden. Dicke Linien deuten auf starke, dünne Linien auf schwache Zusammenhänge hin.

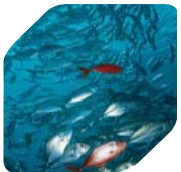
# TNS EX•A•MINE im Überblick



- **Competence Centre** für Data Mining und Data Fusion von TNS Infratest



- Unterstützung aller TNS Marktgesellschaften in den Bereichen **Data Mining, Data Fusion** und **Holistic Segmentation** für analytisches CRM und Database Marketing



- **Hoch spezialisierte (Senior) Consultants** und (Senior) Data Analysts aus den Fachgebieten Wirtschaftswissenschaften, Statistik und EDV



- Langjährige **branchenübergreifende und internationale Projekterfahrung** mit komplexen Analysedesigns und Studienkonzepten
- Toolbox mit **state-of-the-art Verfahren** aus den Bereichen der klassischen multivariaten Statistik und des Data Mining
- Enge Kontakte zu Forschungseinrichtungen und Softwareindustrie; fortlaufende Anpassung der Methoden und Tools an den **aktuellen Stand der Forschung**

# Die EX·A·MINE Algorithmen-Toolbox

## ■ **Multivariate Statistik**

- Logistische, kategoriale, lineare Regression, EM-Algorithmus
- Multivariate Adaptive Regression Splines (MARS)
- Ridge Regression, Robust Regression
- Clusteranalyse, Latent Class Analyse

## ■ **Entscheidungsbäume/-regeln, maschinelles Lernen**

- C&RT, C5.0, QUEST, CHAID, Assoziationsregeln
- MART – Multiple Additive Regression Trees, Random Forest
- Nearest Neighbours / Instance based learning EX·A·MINE Profiler

## ■ **Künstliche Neuronale Netze**

- Cascade Correlation Learning Architecture, MLP, SOM

## ■ **Hybride Methoden**

- Automatisierte OLAP Navigation und Suche
- Genetische Algorithmen zur Variablenauswahl
- Neuro Fuzzy Algorithmen, interaktive Datenvisualisierung

# Ihre Ansprechpartner

Holistic  
Customer  
Understanding  
[EX•A•MINE  
Services]

Dr. Stefan Tuschl  
089 – 5600 1107  
stefan.tuschl@tns-infratest.com



Dr. Robert Hartl  
089 – 5600 1320  
robert.hartl@tns-infratest.com



Susanne Menzler  
089 – 5600 1768  
susanne.menzler@tns-infratest.com



Dr. Melanie Merzenich  
089 – 5600 1172  
melanie.merzenich@tns-infratest.com



Marcus Ott  
089 – 5600 1165  
marcus.ott@tns-infratest.com



Dr. Andreas Unterreitmeier  
089 – 5600 1423  
andreas.unterreitmeier@tns-infratest.com

