

*IBM SPSS Modeler  
Social Network Analysis 17  
Benutzerhandbuch*

**IBM**

**Hinweis**

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 25 gelesen werden.

**Produktinformation**

Diese Ausgabe bezieht sich auf Version 17, Release 0, Modifikation 0 von IBM(r) SPSS(r) Modeler und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs  
*IBM SPSS Modeler Social Network Analysis 17, User's Guide*,  
herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2015

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:  
TSC Germany  
Kst. 2877  
Januar 2015

---

# Inhaltsverzeichnis

<b>Vorwort</b> . . . . .	<b>v</b>
Informationen zu IBM Business Analytics . . . . .	v
Technical Support . . . . .	v

## **Kapitel 1. IBM SPSS Modeler Social Network Analysis** . . . . . **1**

Informationen zu IBM SPSS Modeler Social Network Analysis . . . . .	1
Analyse sozialer Netze . . . . .	2
Anzeigen von Netzen . . . . .	2
Beschreiben von Netzen . . . . .	3
Knoten von IBM SPSS Modeler Social Network Analysis . . . . .	5
Knotenregisterkarten. . . . .	5
Anzeigen einer Ausgabevorschau . . . . .	6
Analysieren von Daten . . . . .	6
Anwendungen. . . . .	7
Datenstruktur . . . . .	7
Scripteigenschaften . . . . .	8

## **Kapitel 2. Gruppenanalyse** . . . . . **11**

Gruppenanalyse - Übersicht . . . . .	11
Ermitteln der sozialen Ähnlichkeit. . . . .	11
Partitionierung in Gruppen . . . . .	12

Beschreiben von Gruppen und Gruppenmitgliedern. . . . .	13
Gruppenanalyseknoten . . . . .	13
Angeben von Daten für die Gruppenanalyse . . . . .	14
Festlegen von Erstellungsoptionen für die Gruppenanalyse . . . . .	14
Anzeigen von Statistiken zur Gruppenanalyse. . . . .	15
Ausgabe für die Gruppenanalyse . . . . .	15

## **Kapitel 3. Diffusionsanalyse** . . . . . **19**

Diffusionsanalyse - Übersicht . . . . .	19
Beispiel für Diffusionsanalyse . . . . .	19
Diffusionsanalyseknoten . . . . .	20
Angeben von Daten für die Diffusionsanalyse . . . . .	20
Festlegen von Erstellungsoptionen für die Diffusionsanalyse . . . . .	21
Anzeigen von Statistiken zur Diffusionsanalyse . . . . .	22
Ausgabe für die Diffusionsanalyse. . . . .	22

## **Bemerkungen.** . . . . . **25**

Marken. . . . .	26
-----------------	----

## **Index** . . . . . **29**



---

## Vorwort

IBM® SPSS Modeler Social Network Analysis verarbeiten Informationen zu Beziehungen zwischen Personen in Felder, die die Rolle einer Einzelperson in einem sozialen Netz beschreiben, wodurch soziale Informationen in Vorhersagemodelle aufgenommen werden können. In diesem Handbuch wird die Verwendung der Knoten von IBM SPSS Modeler Social Network Analysis in der IBM SPSS Modeler-Umgebung beschrieben, die Ihnen die Aufnahme der Knoten in Ihre Streams ermöglicht. Wenn die Knotenausgabe mit Feldern kombiniert wird, die für Maßzahlen zu Einzelpersonen stehen, ergibt sich ein vollständigeres Profil der Ergebnisse für die Einzelpersonen.

---

## Informationen zu IBM Business Analytics

Die Software IBM Business Analytics liefert umfassende, einheitliche und korrekte Informationen, mit denen Entscheidungsträger die Unternehmensleistung verbessern können. Ein umfassendes Portfolio aus Anwendungen für Business Intelligence, Vorhersageanalyse, Finanz- und Strategiemangement sowie Analysen bietet Ihnen sofort klare und umsetzbare Einblicke in die aktuelle Leistung und gibt Ihnen die Möglichkeit, zukünftige Ergebnisse vorherzusagen. Durch umfassende Branchenlösungen, bewährte Vorgehensweisen und professionellen Service können Unternehmen jeder Größe die Produktivität maximieren, Entscheidungen automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt IBM SPSS Predictive Analytics-Software Unternehmen dabei, zukünftige Ereignisse vorherzusagen und proaktiv Maßnahmen zu ergreifen, um bessere Geschäftsergebnisse zu erzielen. Kunden aus Wirtschaft, öffentlichem Dienst und staatlichen Lehr- und Forschungseinrichtungen weltweit nutzen IBM SPSS-Technologie als Wettbewerbsvorteil für die Kundengewinnung, Kundenbindung und Erhöhung der Kundenumsätze bei gleichzeitiger Eindämmung der Betrugsmöglichkeiten und Minderung von Risiken. Durch die Einbindung von IBM SPSS-Software in ihre täglichen Abläufe wandeln sich Organisationen zu "Predictive Enterprises", die Entscheidungen auf Geschäftsziele ausrichten und automatisieren und einen messbaren Wettbewerbsvorteil erzielen können. Wenn Sie weitere Informationen wünschen oder Kontakt zu einem Mitarbeiter aufnehmen möchten, besuchen Sie die Seite <http://www.ibm.com/spss>.

---

## Technical Support

Kunden mit Wartungsvertrag können den Technical Support in Anspruch nehmen. Kunden können sich an den Technical Support wenden, wenn sie Hilfe bei der Arbeit mit IBM Produkten oder bei der Installation in einer der unterstützten Hardwareumgebungen benötigen. Den Technical Support erreichen Sie über die Website von IBM unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihre Organisation und Ihre Supportvereinbarung angeben.



---

# Kapitel 1. IBM SPSS Modeler Social Network Analysis

---

## Informationen zu IBM SPSS Modeler Social Network Analysis

Zahlreiche Ansätze zur Modellierung von Verhalten konzentrieren sich auf die jeweilige Einzelperson. Dabei wird eine Reihe von Daten über Einzelpersonen verwendet, um ein Modell zu generieren, das Schlüsselindikatoren des Verhaltens verwendet, um dieses vorherzusagen. Wenn bei einer Einzelperson Werte für die Schlüsselindikatoren vorkommen, die dem Auftreten des Verhaltens zugeordnet sind, kann die Aufmerksamkeit gezielt auf diese Person gerichtet werden, um das betreffende Verhalten zu verhindern.

Betrachten Sie beispielsweise Ansätze zur Modellierung der Abwanderung, also wenn ein Kunde seine Geschäftsbeziehung mit einem Unternehmen beendet. Die Kosten, Kunden zu halten, sind bedeutend niedriger als die Kosten, sie zu ersetzen. Dies macht die Fähigkeit, abwanderungsgefährdete Kunden zu erkennen, so wichtig. Analysten verwenden häufig eine Reihe von Leistungskennzahlen (Key Performance Indicators, KPI), um Kunden zu beschreiben. Dazu gehören auch demografische Informationen und das jüngste Telefonieverhalten der einzelnen Kunden. Bei Vorhersagemodellen, die auf diesen Feldern beruhen, werden Änderungen im Telefonieverhalten der Kunden, die mit dem Telefonieverhalten von Kunden übereinstimmen, die früher abgewandert sind, verwendet, um Personen zu identifizieren, bei denen ein erhöhtes Abwanderungsrisiko besteht. Kunden, bei denen ein hohes Risiko ermittelt wurde, erhalten zusätzlichen Kundendienst oder Serviceoptionen, um zu versuchen, sie zu halten.

Bei diesen Methoden werden soziale Informationen übersehen, die das Verhalten von Kunden erheblich beeinflussen können. Informationen über ein Unternehmen und darüber, was andere Personen tun, breiten sich über die Beziehungen aus und beeinflussen die Personen. Infolgedessen können Personen durch ihre Beziehungen zu anderen Personen die Entscheidungen und Handlungen der anderen beeinflussen. Analysen, die nur Maßzahlen für Einzelpersonen enthalten, vernachlässigen wichtige Faktoren, die Aussagekraft für die Vorhersage besitzen.

IBM SPSS Modeler Social Network Analysis begegnet diesem Problem, indem Beziehungsinformationen verarbeitet und in zusätzliche Felder eingetragen werden, die in Modelle aufgenommen werden können. Diese abgeleiteten Leistungskennzahlen messen soziale Merkmale für Einzelpersonen. Die Verbindung dieser sozialen Eigenschaften mit Maßzahlen für Einzelpersonen bietet einen besseren Überblick über Einzelpersonen und kann folglich die Vorhersagegenauigkeit Ihrer Modelle verbessern.

IBM SPSS Modeler Social Network Analysis besteht aus zwei Hauptkomponenten:

- Zur IBM SPSS Modeler-Umgebung hinzugefügte Knoten von IBM SPSS Modeler Social Network Analysis, die die Einbeziehung von Verfahren zur sozialen Analyse in Streams ermöglichen.
- IBM SPSS Modeler Server Social Network Analysis ergänzt IBM SPSS Modeler Server um die Verarbeitung der Knotenspezifikationen. IBM SPSS Modeler Server Social Network Analysis verarbeitet effizient enorme Mengen von Netzdaten, die Millionen von Einzelpersonen und Beziehungen umfassen können, zu einer relativ kleinen Anzahl von Feldern zur weiteren Analyse.

So ermittelt IBM SPSS Modeler Social Network Analysis beispielsweise die Personen in einem Netz, die am meisten durch die Abwanderung bestimmter Personen beeinflusst werden. Außerdem können Sie Gruppen von Personen in einem Netz ermitteln, die einem erhöhten Abwanderungsrisiko unterliegen. Indem Sie Leistungskennzahlen für diese Effekte in Ihre Modelle aufnehmen, können sie deren Gesamtleistung verbessern.

---

## Analyse sozialer Netze

Soziale Netze bestehen aus einer Menge von Einzelpersonen und den Beziehungen zwischen diesen Personen. Bei der Analyse sozialer Netze werden diese Beziehungen untersucht, um Einzelpersonen und Gruppen als Bestandteile einer sozialen Struktur zu beschreiben. Einzelpersonen interagieren miteinander, und diese Interaktionsmuster bieten Einblicke in die beteiligten Personen. Beziehungen ermöglichen den Informationsfluss in einem Netz, sodass eine Person eine andere beeinflussen kann. Die Wichtigkeit der Beziehungsinformationen hebt die Analyse sozialer Netze von anderen Ansätzen ab. Anstatt sich jeder Person einzeln zuzuwenden, ist die Untersuchungseinheit eine Dyade, bestehend aus zwei Einzelpersonen und ihren Beziehungen.

Beziehungen in einem Netz können in gerichtete und ungerichtete Beziehungen unterteilt werden. Bei einer *gerichteten Beziehung* wird eine Person als Initiator bzw. Quelle der Beziehung gekennzeichnet und die andere als Empfänger bzw. Ziel. Ein Telefonanruf ist beispielsweise eine gerichtete Beziehung, bei der eine Person eine andere anruft. Im Gegensatz dazu können bei *ungerichteten Beziehungen* die Rollen "Quelle und Ziel" nicht definiert werden. In diesem Fall beteiligen sich beide Parteien gleichermaßen an der Beziehung. Ein Gespräch miteinander ist ein Beispiel für eine ungerichtete Beziehung.

Eine weitere Eigenschaft zur Unterscheidung von Beziehungen besteht darin, ob die Beziehung dichotom oder bewertet ist. Bei *dichotomen Beziehungen* ist die einzige verfügbare Information, ob die Beziehung zwischen zwei Personen besteht oder nicht. Für jede Dyade im Netz ist die Beziehung entweder vorhanden oder abwesend. *Bewertete Beziehungen* dagegen beinhalten eine Gewichtung, die die Stärke der Beziehung anzeigt. Durch die Gewichtungen können die Beziehungen miteinander verglichen werden.

In der Tabelle "Beziehungstypen" werden Beispiele für die Kreuzklassifikation von Beziehungen anhand von Richtung und Skala aufgeführt. Bei den gerichteten Beziehungen ist *Hans* die Quelle der Beziehung und *Marie* das Ziel. Bei den ungerichteten Beziehungen gibt es keinen Hinweis darauf, wer die Beziehung initiiert hat. Bei den bewerteten Beziehungen wird die Länge der Konversation als Beziehungsgewichtung verwendet, während die dichotomen Beziehungen entweder auftraten oder nicht.

Tabelle 1. Beziehungstypen.

Richtung	Skala	Beispiel
Ungerichtet	Dichotom	Hans und Marie sprachen miteinander.
Ungerichtet	Bewertet	Hans und Marie sprachen 20 Minuten lang miteinander.
Gerichtet	Dichotom	Hans rief Marie an.
Gerichtet	Bewertet	Hans rief Marie für ein 20-minütiges Gespräch an.

Weitere Informationen zur Analyse sozialer Netze finden Sie in den ausführlichen Büchern zu diesem Thema<sup>1</sup>.

## Anzeigen von Netzen

Soziale Netze werden üblicherweise mithilfe von *Soziogrammen* dargestellt<sup>2</sup>. Bei dieser visuellen Darstellung entsprechen die einzelnen Personen Punkte bzw. Knoten in einem Raum. Linien bzw. Kanten, die die Punkte verbinden, stellen Beziehungen zwischen den Einzelpersonen dar. Wenn die Beziehungen gerichtet sind, enthalten die Kanten einen Pfeil zur Angabe der Richtung. Wenn die Beziehungen Gewichtungen aufweisen, werden die entsprechenden Werte durch Beschriftungen für die Kanten angegeben. Die folgende Grafik zeigt ein Netz für sieben Einzelpersonen.

---

1. Wasserman, S., and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

2. Moreno, J. L. 1934. *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama*. Washington, D.C.: Nervous and Mental Disease Publishing Co..

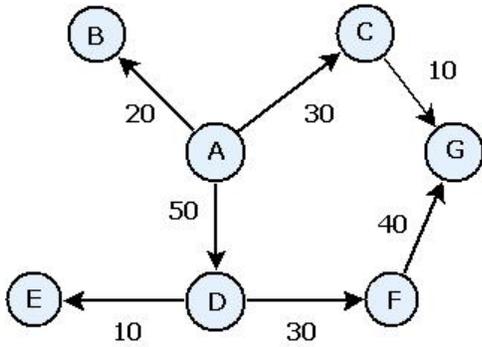


Abbildung 1. Beispiel für soziales Netz

Angenommen, das Netz stellt die von einzelnen Personen getätigten Telefonanrufe dar, wobei die Beziehungsgewichtungen die Länge der Anrufe angeben. In diesem Fall rief Person A drei Personen an, wobei sie den größten Teil der Zeit mit Person D sprach.

Dieses Netz ist deutlich kleiner als die in der Praxis vorkommenden Netze. Die durch einfache Soziogrammen dargestellten Konzepte können jedoch als Verallgemeinerung von Netzen jeder Größe und Komplexität dienen.

## Beschreiben von Netzen

Informationen zu Netzen, Gruppen und Einzelpersonen müssen in beschreibende Merkmale extrahiert werden, die Kreuzvergleiche und die Aufnahme in Vorhersagemodelle gestatten. Netze müssen zu einer endlichen Menge von analysierbaren Leistungskennzahlen zusammengeschmolzen werden. Beispielsweise kann es erforderlich sein, Netze oder Gruppen von Knoten in einem Netz miteinander zu vergleichen. Alternativ kann es sinnvoll sein, Einzelpersonen im Netz miteinander zu vergleichen, um die wichtigsten Personen zu ermitteln.

Zwei häufig zur Beschreibung sozialer Netze verwendete Maßzahlen sind **Dichte** und **Grad**. Beide Statistiken geben die Konnektivität an, die erstere bezieht sich jedoch auf das gesamte Netz bzw. auf Netzuntergruppen, während die letztere die Einzelpersonen im Netz beschreibt.

### Netzdichte

Für jede gegebene Menge an Knoten in einem Netz ist stets eine endliche Anzahl an Beziehungen möglich. Die einzelnen Knoten können als Quelle oder Ziel einer Beziehung mit jedem anderen Knoten dienen. Betrachten wir ein Netz, das aus drei Knoten, A, B und C besteht. In der folgenden Tabelle sind alle möglichen gerichteten Beziehungen zwischen den Knoten aufgeführt.

Tabelle 2. Mögliche gerichtete Beziehungen für drei Knoten.

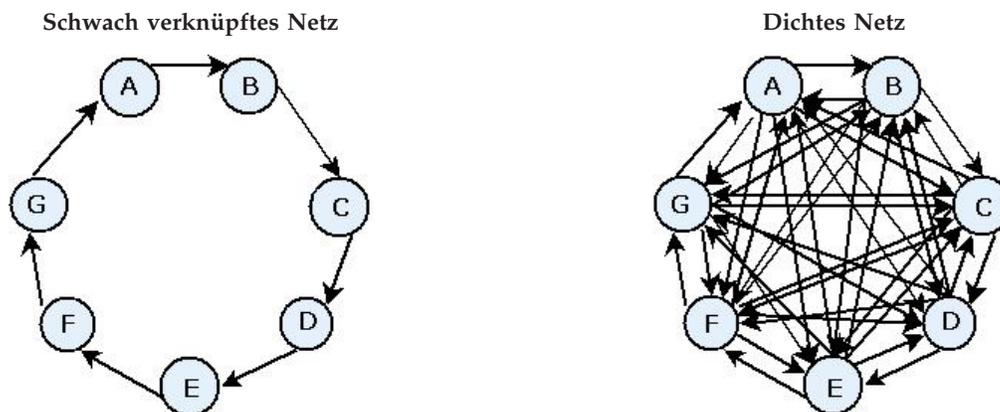
Quelle	Ziel
A	B
A	C
B	A
B	C
C	A
C	B

Jeder Knoten ist die Quelle einer Beziehung mit den beiden anderen Knoten. In der Praxis müssen jedoch nicht alle möglichen Beziehungen tatsächlich vorhanden sein. Einige Knoten weisen möglicherweise überhaupt keine direkte Beziehung mit anderen Knoten auf. Außerdem werden manche gerichteten Beziehungen möglicherweise nicht erwidert.

Die *Dichte*-Statistik steht für den Anteil der möglichen Beziehungen im Netz, die tatsächlich vorhanden sind. Der Wert reicht von 0 bis 1, wobei die Untergrenze einem Netz ohne jegliche Beziehungen entspricht und die Obergrenze einem Netz mit allen möglichen Beziehungen. Je näher der Wert an 1 liegt, desto dichter ist das Netz und desto höher ist die Kohäsion der Knoten im Netz.

Informationen können in dichten Netzen leichter fließen als in schwach verknüpften Netzen. Die Tabelle "Schwach verknüpfte und dichte Netze" zeigt zwei Netze an, die jeweils aus sieben Knoten bestehen. Das schwach verknüpfte Netz enthält nur sieben der möglichen 42 Beziehungen zwischen den Knoten. Dies ergibt eine Dichte von 0,17. Das dichte Netz dagegen enthält alle möglichen Beziehungen und weist eine Dichte von 1,0 auf.

Tabelle 3. Schwach verknüpfte und dichte Netze.



Im schwach verknüpften Netz müssen Informationen fünf andere Knoten passieren, um von Knoten A zu Knoten G zu gelangen. Im Gegensatz dazu können die Informationen im dichten Netz direkt von Knoten A zu Knoten G fließen.

### Knotengrad

Die wichtigsten Personen in einem Netz sind häufig diejenigen, die an der größten Anzahl an Beziehungen beteiligt sind. Diese Einzelpersonen erlangen Informationen aus einer Vielzahl von Quellen und verbreiten diese Informationen an eine große Anzahl anderer Personen weiter. Im Gegensatz dazu sind Personen, die nur an wenigen Beziehungen beteiligt sind, nicht in der Lage, eine große Anzahl anderer Personen im Netz zu beeinflussen.

Der *Grad* eines Knotens, der als Gesamtzahl der Beziehungen definiert ist, an denen der Knoten beteiligt ist, ermöglicht Vergleiche zwischen den Netzteilnehmern. Einzelpersonen mit höheren Gradwerten sind aktiver als Einzelpersonen mit niedrigeren Werten. Beim Gradwert wird die Richtung der Beziehungen ignoriert. Es wird lediglich ein Gesamtmaß der Aktivität für den Knoten angegeben.

Bei gerichteten Beziehungen können Sie bei der Zählung der Beziehungen berücksichtigen, ob es sich bei einem Knoten um Quelle oder Ziel handelt. Der Wert *Indegree* für einen Knoten ist die Anzahl der Beziehungen, bei denen der betreffende Knoten das Ziel ist. Entsprechend gibt *Outdegree* die Anzahl der Beziehungen an, bei denen der betreffende Knoten als Quelle fungiert. In der folgenden Tabelle werden die Grad-, Indegree- und Outdegree-Werte für die einzelnen Knoten in der Abbildung "Beispiel für soziales Netz" angegeben.

Tabelle 4. Beispielwerte für Grad, Indegree und Outdegree.

Knoten	Grad	Indegree	Outdegree
A	3	0	3
B	1	1	0
C	2	1	1

Tabelle 4. Beispielwerte für Grad, Indegree und Outdegree (Forts.).

Knoten	Grad	Indegree	Outdegree
D	3	1	2
E	1	1	0
F	2	1	1
G	2	2	0

"Indegree" gilt häufig als Maßzahl für das Ansehen bzw. Prestige. Bei höheren Indegree-Werten endet eine größere Anzahl an Beziehungen bei dem betreffenden Knoten. Anders ausgedrückt: Die betreffenden Personen werden von einer großen Anzahl von anderen Personen kontaktiert. Viele andere Knoten initiieren Beziehungen mit dem betreffenden Knoten. Outdegree dagegen gilt als Maßzahl der Zentralität. Bei höheren Werten hat eine höhere Anzahl an Beziehungen ihren Ursprung bei diesem Knoten. Die betreffenden Personen nehmen Kontakt mit einer großen Anzahl an anderen Personen auf.

Bei den Knoten im Beispielnetz geben die Gradwerte an, dass die Knoten A und D die höchste und die Knoten B und E die niedrigste Aktivität aufweisen. Die Indegree-Werte zeigen, dass Knoten G das höchste Prestige aufweist. Gemäß den Outdegree-Werten ist Knoten A am zentralsten.

## Knoten von IBM SPSS Modeler Social Network Analysis

Neben den vielen Standardknoten, die im Lieferumfang von IBM SPSS Modeler enthalten sind, können Sie auch mit Knoten von IBM SPSS Modeler Social Network Analysis arbeiten, um die Ergebnisse von Analysen sozialer Netze in Ihre Streams mit aufzunehmen. In der Tabelle Knoten von IBM SPSS Modeler Social Network Analysis werden diese Knoten beschrieben, die in der Palette "Datenquellen" gespeichert sind.

Tabelle 5. Knoten von IBM SPSS Modeler Social Network Analysis.

Knoten	Symbol	Beschreibung
Gruppenanalyse		Der Gruppenanalyseknoten importiert Daten aus Kommunikationsdatensätzen aus einer Textdatei mit fester Feldlänge, ermittelt Gruppen von Knoten innerhalb des Netzes, die von den Datensätzen definiert sind, und generiert Leistungskennzahlen für die Gruppen und Einzelpersonen im Netz. Weitere Informationen finden Sie im Thema „Gruppenanalyse - Übersicht“ auf Seite 11.
Diffusionsanalyse		Der Diffusionsanalyseknoten importiert Daten aus Kommunikationsdatensätzen aus einer Textdatei mit fester Feldlänge, verbreitet einen durch die Datensätze definierten Effekt im Netz und generiert Leistungskennzahlen, die die Ergebnisse des Effekts auf einzelne Knoten zusammenfassen. Weitere Informationen finden Sie im Thema „Diffusionsanalyse - Übersicht“ auf Seite 19.

## Knotenregisterkarten

Die Knoten von IBM SPSS Modeler Social Network Analysis bieten folgende Registerkarten zur Definition und Vorschau der Analyse:

- **Registerkarte "Daten"**. Dient zur Identifikation der Datei mit den Informationen zum sozialen Netz.
- **Registerkarte "Erstellungsoptionen"**. Dient zur Definition der Einstellungen für die Analysen.
- **Registerkarte "Analyse"**. Dient zur Anzeige einer Übersicht der vorläufigen Ausgaben, die Hinweise dafür bietet, wie die Eingabeeinstellungen auf der Registerkarte "Daten" so geändert werden können, dass sie die optimalen Ergebnisse liefern.

Außerdem bieten die Knoten folgende Registerkarten, die bei allen Quellenknoten in IBM SPSS Modeler zu finden sind:

- **Registerkarte "Filter"**. Dient zum Entfernen oder Umbenennen der von den Knoten erstellten Ausgabefelder. Die Registerkarte bietet dieselben Funktionen wie der Filterknoten.
- **Registerkarte "Typen"**. Dient zum Festlegen der Messniveaus für die von den Knoten erstellten Ausgabefelder. Die Registerkarte bietet dieselben Funktionen wie der Typknoten.
- **Registerkarte "Anmerkungen"**. Dient zum Umbenennen von Knoten, zur Bereitstellung benutzerdefinierter QuickInfos und zum Speichern ausführlicher Anmerkungen.

Weitere Informationen zu diesen allgemeinen Registerkarten finden Sie in der IBM SPSS Modeler-Dokumentation.

## Anzeigen einer Ausgabevorschau

Typischerweise wird die Ausgabe der Knoten als Eingabe für ein Vorhersagemodell verwendet. Wenn Sie die Ausgabe anzeigen möchten, könnten Sie einen Tabellenknoten zum Stream hinzufügen. Die Datenmenge ist jedoch extrem groß, was die Tabellenerstellung zeitaufwendig macht. Außerdem wird ihre Brauchbarkeit durch die Anzahl der Zeilen in der Tabelle begrenzt.

Als Alternative können Sie eine Vorschautabelle generieren, die eine Stichprobe der zu erstellenden Ausgabe anzeigt. Die Vorschau zeigt die generierten Felder für eine begrenzte Anzahl an Zeilen. Die Anzahl an Zeilen ist in den Streameigenschaften definiert. Führen Sie zur Vorschau der Knotenausgabe folgende Schritte aus:

1. Öffnen Sie den Knoten.
2. Geben Sie die Dateneinstellungen auf der Registerkarte "Daten" an.
3. Definieren Sie die Analyseparameter auf der Registerkarte "Erstellungsoptionen".
4. Klicken Sie auf **Vorschau**.

Es wird ein Vorschauenfenster mit den Ergebnissen angezeigt. Außerdem wird durch die Vorschau der Ausgabe die Registerkarte "Analyse" des Knotens mit einem zusammenfassenden Überblick über die Ergebnisse ausgefüllt.

## Analysieren von Daten

Die Analyseeinstellungen, die zu optimalen Ergebnissen führen, werden normalerweise in einem iterativen Verfahren ermittelt. Sie definieren die Einstellungen, führen die Analyse aus und überprüfen die Ergebnisse. Wenn die Ergebnisse nicht die gewünschte Aussagekraft aufweisen, ändern Sie die Einstellungen und führen die Analyse erneut durch.

Führen Sie zur Analyse der Eingabedaten für den Knoten folgende Schritte aus:

1. Öffnen Sie den Knoten.
2. Geben Sie die Dateneinstellungen auf der Registerkarte "Daten" an.
3. Definieren Sie die Analyseparameter auf der Registerkarte "Erstellungsoptionen".
4. Klicken Sie auf **Daten analysieren**.

Wenn in den Erstellungsoptionen angegeben ist, dass Übersichtsstatistiken angezeigt werden sollen, werden die Ergebnisse auf der Registerkarte "Analyse" angezeigt.

Wenn Sie eine Analyse wiederholen möchten, klicken sie auf **Analyse löschen**, um die aktuellen Ergebnisse zu löschen, bevor Sie auf **Daten analysieren** klicken.

---

## Anwendungen

Hier einige Anwendungen, bei denen IBM SPSS Modeler Social Network Analysis besonders vorteilhaft sein kann:

- **Abwanderungsvorhersage.** Gruppenmerkmale können einen Einfluss auf die Abwanderungsquoten haben. Durch Konzentration auf Personen in Gruppen, bei denen ein erhöhtes Abwanderungsrisiko besteht, kann die Abwanderung möglicherweise vermieden werden. Außerdem können Sie Personen ermitteln, die ein hohes Abwanderungsrisiko aufgrund des Informationsflusses aufweisen, der von bereits abgewanderten Personen ausgeht.
- **Nutzen von Gruppenführern.** Gruppenführer üben einen starken Einfluss auf andere Gruppenmitglieder aus. Wenn die Abwanderung eines Gruppenführers vermieden werden kann, verringert sich vermutlich auch die Abwanderungsrate bei den Gruppenmitgliedern. Umgekehrt kann der Versuch, einen Gruppenführer von einem Mitbewerber abzuwerben, die Abwanderungsquote der Gruppenmitglieder erhöhen, die dem betreffenden Mitbewerber zuzurechnen sind und gleichzeitig die Abwanderungsquote der Ihrem Unternehmen zuzurechnenden Gruppenmitglieder verringern.
- **Marketing.** Gruppenführer können zur Einführung neuer Produkte oder Serviceangebote verwendet werden. Der Einfluss des Gruppenführers kann die Wahrscheinlichkeit erhöhen, dass andere Gruppenmitglieder das Angebot erwerben. Mithilfe der Diffusionsanalyse können Sie die Personen ermitteln, die am meisten von den Gruppenführern beeinflusst werden und Ihre Marketingstrategien gezielt auf diese ausrichten.

Im Lieferumfang von IBM SPSS Modeler Social Network Analysis sind zwei Demonstrationsstreams enthalten, die zeigen, wie Sie die Ergebnisse der Analyse sozialer Netzwerke in Ihre Streams einschließen können. Die Datendateien und Beispielstreams werden im Ordner *Demos* im Installationsverzeichnis des Produkts installiert.

- *DA demo streams.str* ist ein Beispiel für die Analyse von Daten zum Ermitteln der 300 Einzelpersonen mit der höchsten Abwanderungswahrscheinlichkeit.
- *GA demo streams.str* ist ein Beispiel für die Verwendung von KPIs (Key Performance Indicators) zum Vorhersagen der Abwanderung von Gruppen und Einzelpersonen sowie zum Ermitteln von bestimmten Einzelpersonen als Zielpersonen für eine Marketing-Kampagne.

---

## Datenstruktur

Informationen zu den Personen in dem zu untersuchenden Netz können über eine Vielzahl von Dateien, Datenbanken und Systemen in Ihrem gesamten Unternehmen verteilt sein. Um das Netz mithilfe von IBM SPSS Modeler Social Network Analysis analysieren zu können, müssen Sie die relevanten Datensätze und Felder aus den Datenquellen extrahieren und sie für die Eingabe in die Knoten formatieren.

Die Analyseknotten benötigen Kommunikationsdatensätze, die in einer einzelnen Textdatei mit fester Breite gespeichert sind. Jede Zeile der Datei entspricht einer Beziehung. Dabei sind die Daten in folgenden Spalten organisiert:

- Die Kennung für die Person, die die Beziehung initiiert.
- Die Kennung für die Person, die das Ziel der Beziehung ist.
- Eine optionale Gewichtung für die Beziehung.

Alle Daten müssen numerisch sein, wobei als Kennungen für Einzelpersonen nur ganze Zahlen in Frage kommen. Optional können Feldnamen als erste Datenzeile der Datei mit aufgenommen werden. Diese Datenstruktur wird in der Tabelle "Beispiele für Kommunikationsdatensätze" verdeutlicht.

Beachten Sie, dass die Headerdatensätze ebenfalls entweder Zahlen oder Leerwerte verwenden müssen. Ein Header mit dem Wert *EF BB BF* beispielsweise würde einen Fehler verursachen.

Tabelle 6. Beispiele für Kommunikationsdatensätze.

Quelle	Ziel	Stärke
1000	5642	243
2190	8444	831
0299	9419	559

Die Gewichtungswerte können jeder beliebigen Maßzahl entsprechen, die Sie verwenden möchten, um die Bedeutung der Beziehung in Bezug auf die anderen Beziehungen darzustellen. Bei Anrufrufen sind übliche Gewichtungen die Anrufdauer bzw. die Anruhfrequenz. Beachten Sie, dass dies für die Diffusionsanalyse gilt. Die Gruppenanalyse unterstützt nur Anruhfrequenzen.

Wenn sich die Analyse auf ein Subset des Anrufverlaufs beziehen soll, müssen Sie bei der Erstellung der Eingabedatei das betreffende Subset verwenden. Sie können die Analyse beispielsweise auf die letzten Monate oder auf die aktuellsten Anrufe für eine Person eingrenzen, indem Sie nur diese Daten in die Eingabetextdatei aufnehmen.

## Scripteigenschaften

Die Scripterstellung in IBM SPSS Modeler Social Network Analysis ist ein leistungsstarkes Tool, mit dem Prozesse an der Benutzerschnittstelle automatisiert werden. Scripts können dieselben Arten von Aktionen durchführen, die Sie mit einer Maus oder einer Tastatur durchführen. So können Sie Aufgaben automatisieren, die bei einer manuellen Durchführung sehr viele Wiederholungen verlangen oder sehr viel Zeit beanspruchen. Weitere Informationen zur Verwendung von Scripts finden Sie im Handbuch *Scripting Automation.pdf*, das mit IBM SPSS Modeler bereitgestellt wird.

Eigenschaften für Diffusionsknoten

In der folgenden Tabelle sind die Scripteigenschaften für den Diffusionsknoten aufgeführt.

Eigenschaftsname	Datentyp	Eigenschaftsbeschreibung
input_data_file_name	Zeichenfolge	
calling_field	Feld	
called_field	Feld	
frequency_weight_field	Feld	
read_field_names	boolesch	
diffusion_list_file_name	Zeichenfolge	
spreading_factor	Double	Standardwert: 50. Mindestwert: 1. Maximalwert: 99.
max_number_iterations	ganze Zahl	Standardwert: 100. Mindestwert: 1.
accuracy_threshold	Double	Standardwert: 0.01. Mindestwert: 0.001.
calculate_statistics	boolesch	

Eigenschaften für Gruppenanalyseknoten

In der folgenden Tabelle sind die Scripteigenschaften für den Gruppenanalyseknoten aufgeführt.

<b>Eigenschaftsname</b>	<b>Datentyp</b>	<b>Eigenschaftsbeschreibung</b>
input_data_file_name	<i>Zeichenfolge</i>	
calling_field	<i>Feld</i>	
called_field	<i>Feld</i>	
frequency_weight_field	<i>Feld</i>	
read_field_names	<i>boolesch</i>	
coverage_threshold	<i>Double</i>	Standardwert: 10. Mindestwert: 1. Maximalwert: 99.
min_group_size	<i>ganze Zahl</i>	Standardwert: 2. Mindestwert: 2.
max_group_size	<i>ganze Zahl</i>	Standardwert: 100. Mindestwert: 2.
calculate_statistics	<i>boolesch</i>	



---

## Kapitel 2. Gruppenanalyse

---

### Gruppenanalyse - Übersicht

Bei der Gruppenanalyse werden die Interaktionsmuster von Einzelpersonen in einem Netz verwendet, um Gruppen ähnlicher Einzelpersonen zu ermitteln. Merkmale dieser Gruppen beeinflussen das Verhalten der einzelnen Gruppenmitglieder. Beispielsweise weisen kleine Gruppen mit vielen Beziehungen zwischen den Mitgliedern und starken Gruppenführern ein höheres Abwanderungsrisiko auf, selbst wenn kein Mitglied der Gruppe tatsächlich abgewandert ist.<sup>3</sup> Vorhersagemodelle, die sowohl Maßzahlen für Gruppen als auch Maßzahlen für Einzelpersonen beinhalten, weisen eine höhere Leistungsfähigkeit auf als Modelle, die nur die letzteren enthalten.

Die Gruppenanalyse besteht aus folgenden allgemeinen Schritten:

1. Ermitteln der Beziehungsstärken, die soziale Proximität widerspiegeln. Weitere Informationen finden Sie im Thema „Ermitteln der sozialen Ähnlichkeit“.
2. Partitionieren des Netzes in Gruppen auf der Grundlage der Beziehungsstärke (unter Berücksichtigung von Größenbeschränkungen). Weitere Informationen finden Sie im Thema „Partitionierung in Gruppen“ auf Seite 12.
3. Erstellen von Profilen für Gruppen und Einzelpersonen, einschließlich der Ermittlung von Gruppenführern. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.

### Ermitteln der sozialen Ähnlichkeit

Die Mitglieder einer Gruppe sollten eine größere Ähnlichkeit zueinander aufweisen als zu Personen, die nicht zu der Gruppe gehören. Bei der Netzanalyse hängt die Ähnlichkeit von zwei Knoten von ihren Beziehungen ab. Für jeden beliebigen Knoten in einem Netz gibt es ein Set von Knoten, die die Ziele gerichteter Beziehungen mit dem betreffenden Knoten sind. Bei Telekommunikationsdaten entspricht dieses Set allen Personen, die von einer bestimmten Person kontaktiert wurden. Wenn zwei Personen dasselbe Set von Personen kontaktieren, gelten diese Personen als einander ähnlich. Je stärker sich die Sets der Beziehungsziele für zwei Personen überschneiden, desto ähnlicher sind sie.

Betrachten Sie das Netz in der Abbildung "Beispielnetz mit zehn Knoten".

---

3. Richter, Y., E. Yom-Tov, and N. Slonim. 2010. Predicting customer churn in mobile networks through analysis of social groups. In: *Proceedings of the 2010 SIAM international conference on data mining*. Columbus, OH: SDM 2010.

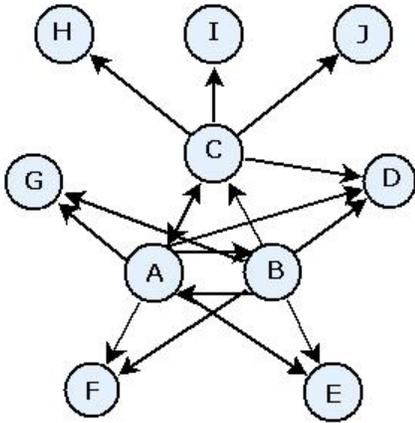


Abbildung 2. Beispielnetz mit zehn Knoten

In der Tabelle "Zielknoten" sehen Sie die Zielknoten für Beziehungen, die ihren Ursprung bei den Knoten A, B und C haben.

Tabelle 7. Zielknoten.

Quellenknoten	Zielknoten
A	B, C, D, E, F, G
B	B, C, D, E, F, G
C	A, D, H, I, J

Die Knoten A und B haben fünf gemeinsame Zielknoten. Die Knoten A und C dagegen nur einen einzigen. Die Knoten A und B weisen also eine größere soziale Ähnlichkeit auf als die Knoten A und C.

Beziehungsgewichtungen, wie beispielsweise Anrufdauer oder -häufigkeit erfassen nicht die Ähnlichkeit der Knoten in einem Netz. Um das Konzept der Ähnlichkeit mit aufzunehmen verwendet die Gruppenanalyse gegenseitige Informationen<sup>4</sup> als Beziehungsgewichtung. Diese Statistik gibt die Wahrscheinlichkeit an, dass zwei Knoten mit demselben Knoten verbunden sind. Wenn die Beziehungen im Netz vordefinierte Gewichtungen aufweisen, wie beispielsweise die Anruhfrequenz bei Telekommunikationsdaten, schließen die gegenseitigen Informationen diese Gewichtungen entsprechend ein.

## Partitionierung in Gruppen

Gruppen sollten aus Personen bestehen, die hohe Beziehungsgewichtungen mit anderen Gruppenmitgliedern aufweisen, wobei die Gewichtungen die Ähnlichkeit der Knoten in der Beziehung messen. Daher beginnt die Gruppenidentifikation durch Weglassen der schwächeren Beziehungen in einem Netz. Der *Abdeckungsschwellenwert* steuert diesen Prozess, indem er den Anteil der stärksten Beziehungen angibt, der beibehalten werden soll. So führt beispielsweise ein Abdeckungsschwellenwert von 0,4 dazu, dass die stärksten 40 % der Beziehungen für die Gruppenidentifizierung verwendet werden, während die restlichen 60 % der Beziehungen weggelassen werden.

Die verbleibenden Beziehungen können einige sehr kleine oder sehr große Gruppen ergeben, die begrenzten Nutzen für Vorhersagen aufweisen. Um zu verhindern, dass solche Gruppen in die Analyse aufgenommen werden, können Sie Mindest- und Höchstgrößen für die Gruppen angeben. Gruppen, die kleiner sind als der Mindestwert, werden vollständig weggelassen. Gruppen, die größer sind als der Höchstwert, werden dagegen in kleinere Gruppen aufgeteilt, deren Größe innerhalb des zulässigen Bereichs liegt. Die nach dem Weglassen der schwachen Beziehungen und dem Durchsetzen der Größenbeschränkungen verbleibenden Gruppen werden als *Kerngruppen* bezeichnet.

4. Cover, T. M., and J. A. Thomas. 2006. *Elements of Information Theory, 2nd edition*. New York: John Wiley and Sons, Inc.

Das Entfernen von Beziehungen aus dem ursprünglichen Netz kann dazu führen, dass manche Personen zu keiner Kerngruppe gehören. Diese Personen weisen jedoch möglicherweise Beziehungen mit Mitgliedern einer Gruppe auf, die eine Aufnahme in die Gruppe rechtfertigen. Einzelpersonen werden zu Gruppen hinzugefügt, wenn es viele relativ starke Beziehungen mit den Kernmitgliedern der Gruppe gibt, vorausgesetzt die für Gruppen geltenden Größenbeschränkungen werden nicht verletzt. Die endgültigen Gruppen bestehen somit aus einer Menge von Kernmitgliedern sowie Mitgliedern, die aufgrund der Konnektivität mit dem Kern mit aufgenommen wurden.

## Beschreiben von Gruppen und Gruppenmitgliedern

Neben Dichte, Indegree und Outdegree dienen noch weitere Statistikwerte zur Beschreibung von Gruppendynamiken. Insbesondere Berechtigungs- und Verteilungsscores bieten Maßzahlen für den sozialen Status der Personen innerhalb der Gruppen. Die Rolle der einzelnen Personen in einer Gruppe ist von entscheidender Bedeutung für den Versuch, das Verhalten der Gruppe und ihrer Mitglieder vorherzusagen.

Der *Berechtigungsscore* für einen Knoten misst die Tendenz anderer Knoten in der Gruppe, eine Verbindung mit ihm herzustellen. Wenn viele Personen eine bestimmte Person kontaktieren und möglicherweise um Informationen oder um die Meinung dieser Person bitten, kommt der betreffenden Person die Rolle einer Berechtigung zu. Die Berechtigungsscores für die Knoten in der Gruppe, die den stationären Wahrscheinlichkeiten für eine Zufallsbewegung entsprechen, die erneut im Gruppennetz beginnt, reichen von 0 bis 1. Je näher der Berechtigungsscore bei 1 liegt, desto größere Berechtigung hat der Knoten innerhalb der Gruppe. Der Knoten mit dem höchsten Berechtigungsscore innerhalb der Gruppe wird als *Berechtigungsführer* der Gruppe bezeichnet. Die Division des größten Scores innerhalb einer Gruppe durch den kleinsten führt zu einer Maßzahl für die allgemeine Stärke des Berechtigungsführers.

Im Gegensatz dazu misst der *Verteilungsscore* für einen Knoten die Tendenz des Knotens, eine Verbindung mit anderen Knoten in der Gruppe herzustellen. Wenn eine bestimmte Person mit vielen anderen Personen in der Gruppe Kontakt aufnimmt, kann diese Person einen erheblichen Einfluss auf die Meinungen der gesamten Gruppe haben. Die Verteilungsscores für die Knoten in der Gruppe, die den stationären Wahrscheinlichkeiten für eine Zufallsbewegung entsprechen, die in umgekehrter Richtung erneut im Gruppennetz beginnt, reichen von 0 bis 1. Je näher der Verteilungsscore bei 1 liegt, desto mehr stellt der Knoten Verbindungen mit den anderen Gruppenmitgliedern her. Der Knoten mit dem höchsten Verteilungsscore innerhalb der Gruppe wird als *Verteilungsführer* der Gruppe bezeichnet. Die Division des größten Scores innerhalb einer Gruppe durch den kleinsten führt zu einer Maßzahl für die allgemeine Stärke des Verteilungsführers.

---

## Gruppenanalyseknoten

Der Gruppenanalyseknoten, der über die Palette "Datenquellen" verfügbar ist, ermittelt Sets von Einzelpersonen in einem Netz, die eine hohe soziale Ähnlichkeit untereinander aufweisen, und ermittelt den relativen Sozialstatus der Personen in den Gruppen. Gruppen, die relativ klein sind, zahlreiche Verbindungen zwischen den Mitgliedern aufweisen und einen starken Führer haben, sind aufgrund der Gruppendynamiken einem erhöhten Abwanderungsrisiko ausgesetzt. Sie können anhand der Ergebnisse Gruppenscores für die Abwanderung ermitteln, die in Kombination mit den Einzelscores für die Abwanderung die Fähigkeit zur Vorhersage der Abwanderung gegenüber Modellen erhöhen, die ausschließlich auf Einzelscores beruhen.

Angesichts der großen Datenmenge, aus denen das Netz üblicherweise besteht, und der Natur der sozialen Interaktionen, die Gruppeneigenschaften zugrunde liegen, ist die Gruppenanalyse ein zeitaufwendiger Vorgang. Üblicherweise werden die Ergebnisse der Gruppenanalyse in einer Datenbank oder einer Datei gespeichert, die als Eingangsdaten für ein Vorhersagemodell verwendet wird. Die Gruppenergebnisse werden in relativ unregelmäßigen Abständen aktualisiert, beispielsweise monatlich, während das Modell wesentlich häufiger aktualisiert werden kann.

**Anforderungen.** Für den Knoten ist eine Textdatei mit fester Breite erforderlich, die das soziale Netz über drei Felder definiert. Ein Feld gibt die Quelle für jede gerichtete Beziehung an, ein Feld gibt das Ziel für jede gerichtete Beziehung an und das dritte Feld gibt eine optionale Stärke für jede Beziehung an. Alle Beziehungen im Netz müssen gerichtet sein.

## Angeben von Daten für die Gruppenanalyse

Auf der Registerkarte "Daten" des Fensters für den Gruppenanalysequellenknoten können Sie die Eingabedatei angeben, die die Beziehungen zwischen den Knoten im Netz enthält.

**Datei.** Geben Sie den Namen der Datei oder des Ordners mit den Kommunikationsdatensätzen ein. Zur Auswahl eines Namens aus dem Dateisystem können Sie einen Dateinamen eingeben oder auf die Schaltfläche mit den Auslassungspunkten (...) klicken. Der Pfad wird angezeigt, sobald Sie einen Namen ausgewählt haben, und der entsprechende Inhalt wird mit Trennzeichen angezeigt. Wenn Sie einen Ordner angeben, werden die Kommunikationsdatensätze aller in dem Ordner befindlichen Dateien für die Analyse aneinandergereiht. Dabei müssen alle Dateien in dem Ordner die gleiche Struktur haben.

**Feldnamen aus Datei lesen.** Diese standardmäßig ausgewählte Option behandelt die erste Zeile der Datendatei als Namen für die Spalte. Wenn die erste Zeile kein Header ist, inaktivieren Sie diese Option, um jedem Feld automatisch einen generischen Namen zuzuweisen, beispielweise *Feld1* und *Feld2*.

Einstellungen für Netzdefinition

Die Einstellungen für die Netzdefinition legen die Rollen der Felder fest.

**Felder.** Verwenden Sie die Pfeilschaltflächen, um Elemente aus dieser Liste manuell den verschiedenen Rollenfeldern zuzuweisen. Die Symbole zeigen das gültige Messniveau für jedes Rollenfeld an. Klicken Sie auf die Schaltfläche **Alle**, damit alle Felder in der Liste ausgewählt werden, oder klicken Sie auf die Schaltfläche für ein individuelles Messniveau, um alle Felder mit diesem Messniveau auszuwählen.

**Quelle.** Wählen Sie ein Feld als Ursprung der gerichteten Beziehung aus.

**Ziel.** Wählen Sie ein Feld als Ziel der gerichteten Beziehung aus.

**Gewichtung.** Optional können Sie ein Feld auswählen, das angibt, wie oft die Beziehung im Netz vorkommt. Die Analyse gewichtet die Datensätze bei der Ermittlung der Beziehungsstärken entsprechend.

## Festlegen von Erstellungsoptionen für die Gruppenanalyse

Auf der Registerkarte "Erstellungsoptionen" des Fensters für den Gruppenanalysequellenknoten können Sie Optionen für die Ermittlung von Gruppen im Netz definieren.

Einstellungen der Gruppenanalyse

Die Einstellungen der Gruppenanalyse wirken sich auf die Größe und die relative Stärke der gebildeten Gruppen aus.

**Schwellenwert für Abdeckung.** Dient zur Festlegung des Anteils der stärksten Netzbeziehungen, der in der Analyse verwendet werden soll. Ein Abdeckungsschwellenwert von 0,2 beispielsweise gibt an, dass nur die Beziehungen, deren Gewichtungen zu den stärksten 20 % aller Gewichtungen gehören, verwendet werden. Durch diesen Parameter, der von 0 bis 1 reicht, kann die Analyse auf die stärksten Beziehungen konzentriert werden, die innerhalb von Gruppen vorkommen sollten. Höhere Schwellenwerte ergeben Gruppen mit schwächeren Beziehungen zwischen ihren Mitgliedern. Weitere Informationen finden Sie im Thema „Partitionierung in Gruppen“ auf Seite 12.

**Minimale Gruppengröße.** Dient zur Angabe einer Untergrenze für die Gruppengröße. Es werden keine Gruppen ausgegeben, die kleiner sind als dieser Wert.

**Maximale Gruppengröße.** Dient zur Angabe einer Obergrenze für die Gruppengröße. Gruppen, die größer sind als dieser Wert, werden in kleinere Gruppen unterteilt.

**Übersichtsstatistiken berechnen und anzeigen.** Wenn diese Option ausgewählt ist, leitet der Knoten nicht nur die Ausgabe der Leistungskennzahlen ab, sondern berechnet auch Übersichtsstatistiken für die Analyse und zeigt sie an. Die Berechnung dieser Statistiken kann bei großen Netzen die Leistungsfähigkeit des Knotens beeinträchtigen.

## Anzeigen von Statistiken zur Gruppenanalyse

Auf der Registerkarte "Analyse" des Fensters für den Gruppenanalysequellenknoten wird eine überblicksartige Zusammenfassung der im Netz ermittelten Gruppen angezeigt. In der Tabelle "Gruppenanalyse - Übersichtsstatistiken" werden die verfügbaren Übersichtsstatistiken für die Gruppen aufgeführt.

*Tabelle 8. Gruppenanalyse - Übersichtsstatistiken.*

Statistik	Beschreibung
Gesamtzahl der Knoten in Gruppen	Anzahl der in den ermittelten Gruppen enthaltenen Knoten.
Gesamtzahl der Verknüpfungen in Gruppen	Anzahl der in den ermittelten Gruppen enthaltenen Verknüpfungen.
Gesamtzahl der Gruppen	Anzahl der im Netz ermittelten Gruppen.
Mittlere Gruppengröße	Durchschnittliche Anzahl an Knoten in einer Gruppe.
Mittlere Gruppendichte	Durchschnittlicher Anteil direkter Beziehungen zwischen Knoten in einer Gruppe. Weitere Informationen finden Sie im Thema „Netzdicke“ auf Seite 3.
Mittlerer Anteil an Kernmitgliedern	Durchschnittlicher Anteil an Knoten in einer Gruppe, bei denen es sich um Kernknoten für die Gruppe handelt. Weitere Informationen finden Sie im Thema „Partitionierung in Gruppen“ auf Seite 12.
Mittlere Dichte der Kerngruppe	Durchschnittlicher Anteil direkter Beziehungen zwischen Kernknoten in einer Gruppe.
Mittlerer Indegree	Durchschnittliche Anzahl eingehender Verknüpfungen. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.
Mittlerer Outdegree	Durchschnittliche Anzahl ausgehender Verknüpfungen. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.

Wählen Sie eine bestimmte Statistik in der Tabelle "Übersichtsstatistiken" aus, um die Verteilung der Werte, die Standardabweichung und die Schiefe der Statistik anzuzeigen.

Die Überprüfung dieser Statistiken erleichtert die Festlegung der Einstellungen für die Gruppenanalyse. Idealerweise sollten die Ergebnisse relativ kleine Gruppen mit hohen Dichtewerten zeigen. Wenn beispielsweise einige sehr große Gruppen vorhanden sind, sollten Sie vielleicht die maximale Gruppengröße verringern, um sie in kleinere Gruppen aufzuteilen. Umgekehrt gilt: Bei kleinen Werten für die Gruppendichte sollte eventuell der Schwellenwert für die Abdeckung verringert werden, um die Untersuchung auf stärkere Beziehungen einzugrenzen.

## Ausgabe für die Gruppenanalyse

Der Gruppenanalyseknoten generiert verschiedene Felder, die die Gruppen und die einzelnen Personen in den Gruppen beschreiben. Sie können bestehende Modelle und Daten mit diesen Leistungskennzahlen erweitern, um die von Ihren Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Gruppeneinflüsse ergänzen.

In der Tabelle "Leistungskennzahlen für Gruppen" werden die Felder für die Leistungskennzahlen für die in der Analyse ermittelten Gruppen aufgeführt.

Tabelle 9. Leistungskennzahlen für Gruppen.

Feld	Beschreibung
GAG_GroupNumber	Eindeutige Kennung für eine Gruppe.
GAG_Size	Anzahl der Einzelpersonen in einer Gruppe.
GAG_Density	Anteil direkter Beziehungen zwischen Einzelpersonen in einer Gruppe. Weitere Informationen finden Sie im Thema „Netzdichte“ auf Seite 3.
GAG_KernelDensity	Anteil direkter Beziehungen zwischen zum Kern gehörenden Einzelpersonen in einer Gruppe.
GAG_CoreNodesFraction	Anteil an Einzelpersonen in einer Gruppe, die zum Kern der Gruppe gehören. Weitere Informationen finden Sie im Thema „Partitionierung in Gruppen“ auf Seite 12.
GAG_MaxRankType1	Maximaler Berechtigungsscore für ein beliebiges Gruppenmitglied. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAG_MinRankType1	Minimaler Berechtigungsscore für ein beliebiges Gruppenmitglied.
GAG_MaxMinRankRatioType1	Verhältnis zwischen größtem und kleinstem Berechtigungsscore. Dieser Wert gibt die Berechtigungsstärke des Gruppenführers an.
GAG_MaxRankType2	Maximaler Verteilungsscore für ein beliebiges Gruppenmitglied. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAG_MinRankType2	Minimaler Verteilungsscore für ein beliebiges Gruppenmitglied.
GAG_MaxMinRankRatioType2	Verhältnis zwischen größtem und kleinstem Verteilungsscore. Dieser Wert gibt die Verteilungsstärke des Gruppenführers an.

In der Tabelle "Leistungskennzahlen für Einzelpersonen" werden die Felder für die Leistungskennzahlen für die einzelnen Personen im Netz aufgeführt.

Tabelle 10. Leistungskennzahlen für Einzelpersonen.

Feld	Beschreibung
GAI_NodeNumber	Eindeutige Kennung für eine Einzelperson.
GAI_CoreNode	Kennung, die angibt, ob die Einzelperson zum Kern einer Gruppe gehört oder nicht. Weitere Informationen finden Sie im Thema „Partitionierung in Gruppen“ auf Seite 12.
GAI_RankType1	Berechtigungsscore für die Einzelperson. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAI_RankOrderType1	Rangordnung in der Gruppe auf der Grundlage der Berechtigungsscores.
GAI_RankType2	Verteilungsscore für die Einzelperson. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAI_RankOrderType2	Rangordnung in der Gruppe auf der Grundlage der Verteilungsscores.
GAI_InDegree	Anzahl an Beziehungen, bei denen die Einzelperson das Ziel der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.
GAI_OutDegree	Anzahl an Beziehungen, bei denen die Einzelperson die Quelle der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.

Table 10. Leistungskennzahlen für Einzelpersonen (Forts.).

Feld	Beschreibung
GAI_GroupLeaderType1	Gibt an, ob der Knoten ein Berechtigungsführer ist, dessen Führungsscore aus eingehenden Links abgeleitet ist. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAI_GroupLeaderConfidenceType1	Die Konfidenz, dass es sich bei dem Knoten um einen Berechtigungsführer handelt.
GAI_GroupLeaderType2	Gibt an, ob der Knoten ein Verteilungsführer ist, dessen Führungsscore aus ausgehenden Links abgeleitet ist. Weitere Informationen finden Sie im Thema „Beschreiben von Gruppen und Gruppenmitgliedern“ auf Seite 13.
GAI_GroupLeaderConfidenceType2	Die Konfidenz, dass es sich bei dem Knoten um einen Verteilungsführer handelt.



---

## Kapitel 3. Diffusionsanalyse

---

### Diffusionsanalyse - Übersicht

Durch die Diffusionsanalyse werden Einzelpersonen ermittelt, die am meisten durch andere Einzelpersonen in einem sozialen Netz beeinflusst werden. Dabei wird der Effekt als *diffundierte Energie* quantifiziert. Bei diesem Prozess wird ein Ansatz für die Aktivierungsausbreitung verfolgt, bei dem sich ein Effekt iterativ von Netzknoten auf ihre unmittelbaren Nachbarn ausbreitet, wobei er bei der Bewegung von Knoten zu Knoten kleiner wird<sup>5</sup>. Wenn ein Knoten Energie erhält, wird er aktiviert und überträgt einen Teil dieser Energie an alle Nachbarn, die Ziele direkter Beziehungen mit dem betreffenden Knoten sind.

Der *Ausbreitungsfaktor* definiert den Anteil der von einem aktivierten Knoten übertragenen Energie. Die restliche Energie verbleibt im Knoten. Alle Knoten, die diese Energie erhalten, übertragen wiederum denselben Anteil an ihre Nachbarn, was zu einem Abklingvorgang bei der Energieübertragung führt. Bei großen Ausbreitungsfaktoren wird mehr Energie gesendet, wodurch die Energie Knoten erreichen kann, die weiter von den ursprünglich aktivierten Knoten entfernt liegen, bevor der Prozess vollständig abklingt. Kleine Ausbreitungsfaktoren führen zu rasch abklingenden Diffusionsprozessen, bei denen die übertragene Energie relativ nahe an den ursprünglichen Knoten bleibt.

Die Gesamtmenge der von einem aktivierten Knoten diffundierten Energie wird auf alle Knoten verteilt, die Ziele gerichteter Beziehungen mit dem Knoten sind. Welche Menge, die einzelnen Knoten erhalten, hängt von der Stärke der Beziehung mit dem aktivierten Knoten ab. Der an einen bestimmten Knoten gesendete Bruchteil ist gleich der Beziehungsgewichtung dividiert durch die Gesamtsumme der Gewichtungen für alle Beziehungen, in denen der aktivierte Knoten die Quelle ist. Folglich erhalten Nachbarn, die Beziehungen mit höheren relativen Gewichtungen aufweisen, mehr Energie als Nachbarn mit geringeren relativen Gewichtungen.

Der Diffusionsprozess endet, wenn eine der folgenden Bedingungen gegeben ist:

- Die aktivierten Knoten sind nicht die Quelle einer gerichteten Beziehung.
- Die übertragene Energiemenge liegt unter dem *Schwellenwert für Genauigkeit*, einem Grenzwert für die zu übertragende Energie, die erforderlich ist, damit der Prozess fortgesetzt wird.
- Die Anzahl der Iterationen erreicht eine bestimmte Obergrenze.

Nach Abschluss der Diffusion weisen die Knoten mit der größten Menge an diffundierter Energie die höchste Empfindlichkeit für den Effekt auf, der den Prozess einleitete. Wenn der Prozess beispielsweise bei Knoten beginnt, die abwandern, besteht bei den Knoten mit der höchsten Energie das höchste Risiko für eine Abwanderung. Sie können diesen Knoten spezielle Aufmerksamkeit zukommen lassen, um ihre Abwanderung zu vermeiden.

### Beispiel für Diffusionsanalyse

Betrachten Sie das Netz in „Anzeigen von Netzen“ auf Seite 2. Das Netz besteht aus sieben Knoten, die direkte Beziehungen unterschiedlicher Stärke zueinander aufweisen.

Die Tabelle "Beispiel für Diffusionsprozess" zeigt einen Diffusionsprozess im Netz mit einem Ausbreitungsfaktor von 0,80. Zunächst enthält Knoten A die gesamte Energie, der willkürlich der Wert 1,00 zugewiesen wurde. In Schritt 1 wird dieser Knoten aktiviert, wodurch 80 % seiner Energie auf seine drei Nachbarn verbreitet werden, die Ziele von Beziehungen sind, während 20 % beim ursprünglichen Knoten verbleiben. Die Beziehungen haben eine Gesamtgewichtung von 100. Die Beziehung mit Knoten D

---

5. Dasgupta, K., R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjee, A. A. Nanavati, and A. Joshi. 2008. Social ties and their relevance to churn in mobile telecom networks. In: *Proceedings of the 11th international conference on extending database technology*. New York, NY: ADM.

macht die Hälfte dieses Gesamtwerts aus, dieser Knoten enthält somit die Hälfte der diffundierten Energie, also 0,40. Die Beziehung mit Knoten B macht 20 % der Gesamtgewichtung aus, was den Wert 0,16 für die diffundierte Energie ergibt. Knoten C erhält die restliche Energie, 0,24, das sind 30 % der aus Knoten A stammenden Energiemenge.

Tabelle 11. Beispiel für Diffusionsprozess.

Schritt	Knoten A	Knoten B	Knoten C	Knoten D	Knoten E	Knoten F	Knoten G
0	1.00	0	0	0	0	0	0
1	0.20	0.16	0.24	0.40	0	0	0
2	0.20	0.16	0.05	0.08	0.08	0.24	0.19
3	0.20	0.16	0.05	0.08	0.08	0.08	0.38

Nachdem die Knoten B, C und D in Schritt 1 Energie erhalten haben, werden sie in Schritt 2 aktiviert und diffundieren Energie an ihre Zielnachbarn (sofern vorhanden). Knoten B verfügt über keine Zielknoten und diffundiert daher keine Energie. Knoten C dagegen verfügt über ein Ziel und diffundiert 80 % seiner Energie an den betreffenden Knoten, Knoten G. Knoten D schließlich verbreitet 80 % seiner Energie an die Knoten E und F. Dabei enthält F das Dreifache von E, da er eine stärkere Beziehung mit D aufweist.

In Schritt 3 werden die Knoten E, F und G aktiviert. Knoten E und G haben keine Zielknoten und diffundieren daher keine Energie. Knoten F dagegen diffundiert 80 % seiner Energie an seinen einzigen Zielknoten, G. Dieser Knoten weist nun einen Gesamtenergiewert von 0,38 auf, der sich aus der von den Knoten C und F diffundierten Energiemengen zusammensetzt.

Nun wird Knoten G aktiviert. Er verfügt jedoch nicht über Zielknoten für die Diffusion, weshalb der Prozess endet. Die von Knoten A herrührende Energie hat sich über das Netz ausgebreitet, wobei Knoten G die größte Energiemenge erhielt. Wenn die in das Netz eingeführte Energie für Abwanderung steht, wäre Knoten G am meisten von einer Abwanderung von Knoten A betroffen.

---

## Diffusionsanalyseknoten

Der Diffusionsanalyseknoten, der über die Palette "Datenquellen" verfügbar ist, verbreitet einen Effekt aus einer angegebenen Menge von Einzelpersonen in einem sozialen Netz. Dabei werden die Netzbeziehungen verwendet, um die Personen zu ermitteln, die am stärksten von dem Effekt betroffen sind. Wenn es sich bei dem Effekt beispielsweise um Abwanderung handelt, gibt der Knoten diejenigen Personen an, bei denen die Wahrscheinlichkeit am höchsten ist, dass sie abwandern, wenn bestimmte andere Personen im Netz abwandern. Sie können bestehende Modelle und Daten mit der Knotenausgabe erweitern, um die von diesen Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Diffusionseinflüsse ergänzen.

**Anforderungen.** Für den Knoten sind zwei Textdateien mit fester Breite erforderlich, die die zu analysierenden Daten enthalten. Die erste Datei definiert das soziale Netz mittels dreier Felder. Ein Feld gibt die Quelle für jede gerichtete Beziehung an, ein Feld gibt das Ziel für jede gerichtete Beziehung an und das dritte Feld gibt eine optionale Stärke für jede Beziehung an. Alle Beziehungen im Netz müssen gerichtet sein. Die zweite Datei enthält eine Liste mit Kennzeichnern, die angeben, von wo der Effekt beginnen soll.

## Angeben von Daten für die Diffusionsanalyse

Auf der Registerkarte "Daten" des Fensters für den Diffusionsanalysequellenknoten können Sie die Eingabedatei angeben, die die Beziehungen zwischen den Knoten im Netz enthält.

**Datei.** Geben Sie den Namen der Datei oder des Ordners mit den Kommunikationsdatensätzen ein. Zur Auswahl eines Namens aus dem Dateisystem können Sie einen Dateinamen eingeben oder auf die Schalt-

fläche mit den Auslassungspunkten (...) klicken. Der Pfad wird angezeigt, sobald Sie einen Namen ausgewählt haben, und der entsprechende Inhalt wird mit Trennzeichen angezeigt. Wenn Sie einen Ordner angeben, werden die Kommunikationsdatensätze aller in dem Ordner befindlichen Dateien für die Analyse aneinandergereiht. Dabei müssen alle Dateien in dem Ordner die gleiche Struktur haben.

**Feldnamen aus Datei lesen.** Diese standardmäßig ausgewählte Option behandelt die erste Zeile der Datendatei als Namen für die Spalte. Wenn die erste Zeile kein Header ist, inaktivieren Sie diese Option, um jedem Feld automatisch einen generischen Namen zuzuweisen, beispielweise *Feld1* und *Feld2*.

Einstellungen für Netzdefinition

Die Einstellungen für die Netzdefinition legen die Rollen der Felder fest.

**Felder.** Verwenden Sie die Pfeilschaltflächen, um Elemente aus dieser Liste manuell den verschiedenen Rollenfeldern zuzuweisen. Die Symbole zeigen das gültige Messniveau für jedes Rollenfeld an. Klicken Sie auf die Schaltfläche **Alle**, damit alle Felder in der Liste ausgewählt werden, oder klicken Sie auf die Schaltfläche für ein individuelles Messniveau, um alle Felder mit diesem Messniveau auszuwählen.

**Quelle.** Wählen Sie ein Feld als Ursprung der gerichteten Beziehung aus.

**Ziel.** Wählen Sie ein Feld als Ziel der gerichteten Beziehung aus.

**Gewichtung.** Wählen Sie ein Feld aus, das für die Beziehungsgewichtung steht. Bei Kommunikationsdatensätzen könnte die Gewichtung einer Anzahl von Anrufmerkmalen entsprechen, die die Beziehungsstärke widerspiegeln, beispielsweise Anrufdauer oder Anruhfrequenz.

## Festlegen von Erstellungsoptionen für die Diffusionsanalyse

Auf der Registerkarte "Erstellungsoptionen" des Fensters für den Diffusionsanalysequellenknoten können Sie Optionen für den Diffusionsprozess definieren.

Liste der Diffusionsstartwerte

Die Liste der Diffusionsstartwerte gibt die Netzknoten an, von denen aus die Energie diffundieren soll. Bei einer Abwanderungsanalyse beispielsweise gibt diese Liste die Netzknoten an, die bereits abgewandert sind. Alternativ könnte diese Liste Knoten enthalten, bei denen das Abwanderungsrisiko besonders hoch ist, um die Auswirkungen auf das Netz für den Fall zu untersuchen, dass diese Knoten tatsächlich abwandern.

**Datei.** Geben Sie den Namen der Textdatei an, die die Knotenkennungen enthält. Dabei entspricht jede Kennung einer Zeile in der Datei. Zur Auswahl einer Datei können Sie einen Dateinamen eingeben oder auf die Schaltfläche mit den Auslassungspunkten (...) klicken.

Einstellungen der Diffusionsanalyse

Die Einstellungen für die Diffusionsanalyse legen fest, wie aggressiv die Diffusion ist und wann der Prozess endet.

**Ausbreitungsfaktor.** Dient zur Angabe des Anteils der Energie, den ein Knoten in jedem Schritt des Diffusionsprozesses überträgt. Mit kleineren Werten wird die Energie auf die Knoten übertragen, die dem Knoten, der die Energie diffundiert, am nächsten liegen. Entsprechend ermöglichen größere Werte, dass sich die Energie auf weiter entfernt liegende Knoten auswirkt.

**Maximalzahl der Iterationen.** Dient zur Angabe einer Obergrenze für die Anzahl an Diffusionsiterationen. Der Diffusionsprozess endet automatisch, wenn diese Obergrenze erreicht wird.

**Schwellenwert für Genauigkeit.** Dient zur Angabe der kleinsten Änderung in der Knotenenergie, die eine Fortsetzung des Diffusionsprozesses rechtfertigt. Wenn die Änderung in der Energie für jeden Knoten kleiner ist als dieser Wert, endet der Diffusionsprozess.

**Übersichtsstatistiken berechnen und anzeigen.** Wenn diese Option ausgewählt ist, leitet der Knoten nicht nur die Ausgabe der Leistungskennzahlen ab, sondern berechnet auch Übersichtsstatistiken für die Analyse und zeigt sie an. Die Berechnung dieser Statistiken kann bei großen Netzen die Leistungsfähigkeit des Knotens beeinträchtigen.

## Anzeigen von Statistiken zur Diffusionsanalyse

Auf der Registerkarte "Analyse" des Fensters für den Diffusionsanalysequellenknoten wird eine überblicksartige Zusammenfassung der Diffusionsergebnisse angezeigt. In der Tabelle "Diffusionsanalyse - Übersichtsstatistiken" werden die verfügbaren Übersichtsstatistiken aufgeführt.

*Tabelle 12. Diffusionsanalyse - Übersichtsstatistiken.*

Statistik	Beschreibung
Gesamtzahl der Knoten im Netz	Anzahl der Knoten im Netz.
Gesamtzahl der Verknüpfungen im Netz	Anzahl der Verknüpfungen im Netz
Gesamtzahl der Diffusionsstartwerte im Netz	Anzahl der als Startwerte für den Diffusionsprozess verwendeten Knoten.
Mittlerer Einfluss	Durchschnittsmenge an diffundierter Energie, die Einzelpersonen zugeordnet ist.
Mittlerer Indegree	Durchschnittliche Anzahl an Beziehungen, bei denen eine Einzelperson das Ziel der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.
Mittlerer Outdegree	Durchschnittliche Anzahl an Beziehungen, bei denen eine Einzelperson die Quelle der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.

Wählen Sie eine bestimmte Statistik in der Tabelle "Übersichtsstatistiken" aus, um die Verteilung der Werte, die Standardabweichung und die Schiefe der Statistik anzuzeigen.

## Ausgabe für die Diffusionsanalyse

Der Diffusionsanalyseknoten generiert verschiedene Felder, die die einzelnen Personen im Netz beschreiben. Sie können bestehende Modelle und Daten mit diesen Leistungskennzahlen erweitern, um die von Ihren Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Diffusionseinflüsse ergänzen.

In der Tabelle "Diffusionsanalyse - Leistungskennzahlen" werden die Felder für die Leistungskennzahlen für die einzelnen Personen in der Analyse aufgeführt.

*Tabelle 13. Diffusionsanalyse - Leistungskennzahlen.*

Feld	Beschreibung
DA_NodeNumber	Eindeutige Kennung für eine Einzelperson.
DA_DiffusedEnergy	Menge an diffundierter Energie, die der Einzelperson zugeordnet ist. Bei der Abwanderungsanalyse weisen höhere Werte auf eine größere Abwanderungsneigung hin als niedrigere Werte.
DA_InDegree	Anzahl an Beziehungen, bei denen die Einzelperson das Ziel der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.

*Tabelle 13. Diffusionsanalyse - Leistungskennzahlen (Forts.).*

<b>Feld</b>	<b>Beschreibung</b>
DA_OutDegree	Anzahl an Beziehungen, bei denen die Einzelperson die Quelle der Beziehung ist. Weitere Informationen finden Sie im Thema „Knotengrad“ auf Seite 4.



---

## Bemerkungen

Diese Informationen wurden für weltweit angebotene Produkte und Dienstleistungen erarbeitet.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing  
IBM Europe, Middle East & Africa  
Tour Descartes  
2, avenue Gambetta  
92066 Paris La Defense  
France

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Software Group  
ATTN: Licensing  
200 W. Madison St.  
Chicago, IL; 60606  
USA

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer kontrollierten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Gewährleistung, dass diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können davon abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

---

## Marken

IBM, das IBM Logo und [ibm.com](http://ibm.com) sind Marken oder eingetragene Marken der IBM Corp in den USA und/oder anderen Ländern. Weitere Produkt- und Servicenamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite "Copyright and trademark information" unter [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder der Tochtergesellschaften des Unternehmens in den USA und anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA, anderen Ländern oder beidem.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA und/oder anderen Ländern.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Java und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.

Weitere Produkt- und Servicenamen können Marken von IBM oder anderen Unternehmen sein.

---

## Glossar

---

### A

*Ausbreitungsfaktor.* Der Anteil an Energie, den ein Knoten in einem Schritt des Diffusionsprozesses an seine Nachbarn überträgt.

### B

*Berechtigungsführer.* Die Einzelperson in einer Gruppe, die den höchsten Berechtigungsscore aufweist.

*Berechtigungs-score.* Maß, das angibt, in welchem Umfang die anderen Mitglieder einer Gruppe eine Verbindung zu einer Einzelperson herstellen. Berechtigungscores verknüpfen die Bedeutung einer Einzelperson mit der Anzahl der Beziehungen, die bei dieser Person enden.

*Bewertete Beziehung.* Eine Verbindung zwischen zwei Knoten, denen eine Gewichtung zugeordnet ist, das die Stärke der Beziehung angibt.

### D

*Dichotome Beziehung.* Eine Verbindung zwischen zwei Knoten, die nur einen von zwei Werten annehmen kann, die für das Vorhandensein bzw. Fehlen der Verbindung stehen.

*Dichte.* Ein Maß für die Netzkohäsion, das als Anzahl der beobachteten Beziehungen dividiert durch die Anzahl der möglichen Beziehungen definiert ist.

*Diffundierte Energie.* Effekt, der von bestimmten Knoten in einem Netz ausgeht und sich über das gesamte Netz ausbreitet, wobei seine Größe mit zunehmender Entfernung von den Ursprungsknoten abnimmt.

### G

*Gerichtete Beziehung.* Eine Verbindung zwischen zwei Knoten, die von einem Knoten (der Quelle) ausgeht und beim anderen Knoten (dem Ziel) endet.

*Grad.* Maß für die Knotenaktivität in einem Netz, definiert als Gesamtzahl der Beziehungen, an denen der Knoten beteiligt ist.

### I

*Indegree.* Maß für das Prestige für einen Knoten in einem Netz, das aus gerichteten Beziehungen besteht, definiert als Gesamtzahl der Beziehungen, in denen der Knoten das Ziel ist.

### K

*Kerngruppe.* Eine Gruppe von Knoten, die verbleiben, nachdem schwächere Beziehungen aus einem Netz weggelassen und Größenbeschränkungen auf die Gruppe angewendet wurden.

### O

*Outdegree.* Maß der Zentralität für einen Knoten in einem Netz, das aus gerichteten Beziehungen besteht, definiert als Gesamtzahl der Beziehungen, in denen der Knoten die Quelle ist.

---

## S

*Schwellenwert für Abdeckung.* Der Anteil der stärksten Beziehungen, der für die Gruppenidentifikation verwendet wird.

*Schwellenwert für Genauigkeit.* Ein Stoppkriterium für einen Diffusionsprozess, das der kleinsten Änderung in der Knotenenergie entspricht, die die Fortsetzung der Diffusion rechtfertigt.

*Soziogramm.* Eine visuelle Darstellung eines sozialen Netzes, in der mit Linien verbundene Punkte für Personen stehen, die eine Beziehung zueinander haben.

---

## U

*Ungerichtete Beziehung.* Eine Zuordnung zwischen zwei Knoten.

---

## V

*Verteilungsführer.* Die Einzelperson in einer Gruppe, die den höchsten Verteilungsscore aufweist.

*Verteilungsscore.* Maß, das angibt, in welchem Umfang eine Einzelperson eine Verbindung zu anderen Mitglieder einer Gruppe herstellt. Verteilungsscores verknüpfen die Bedeutung einer Einzelperson mit der Anzahl der Beziehungen, die von dieser Person ausgehen.

---

# Index

## A

- Analyse (Registerkarte) 5
  - Diffusionsanalyse 22
  - Gruppenanalyse 15
- Analysieren von Daten 6
- Anmerkungen (Registerkarte) 5
- Anzeigen einer Ausgabevorschau 6
- Ausbreitungsfaktor 19, 21
- Ausführen von Streams
  - mit IBM SPSS Modeler Server Social Network Analysis 5

## B

- Berechtigungsführer 13, 15
- Berechtigungscores 13, 15
- Bewertete Beziehungen 2
- Beziehung
  - bewertet 2
  - Richtung 2

## D

- Daten (Registerkarte) 5
  - Diffusionsanalyse 20
  - Gruppenanalyse 14
- Dichotome Beziehungen 2
- Dichte 3, 15
- Diffundierte Energie 22
- Diffusionsanalyse 5, 19
  - Anforderungen 20
  - Ausbreitungsfaktor 19, 21
  - Optionen 20
  - Schwellenwert für Genauigkeit 19, 21
  - Startwerte 21

## E

- Eigenschaften
  - Scripts 8
- Erstellungsoptionen (Registerkarte) 5
  - Diffusionsanalyse 21
  - Gruppenanalyse 14

## F

- Filter (Registerkarte) 5
- Filtern der Ausgaben 5

## G

- GAG\_GroupNumber 15
- Gerichtete Beziehungen 2
- Gewichtungsfelder 7, 14, 20
- Grad 4
- Gruppenanalyse 5
  - Anforderungen 13

- Gruppenanalyse (*Forts.*)
  - Gruppengrößen 12, 14
  - Schwellenwert für Abdeckung 12, 14
- Gruppengrößen 15
  - Grenzwerte 12, 14

## I

- Indegree 4, 15, 22

## K

- Kerndichte 15
- Kerngruppen 12
- Kernknoten 15
- Kommunikationsdatensätze 7

## L

- Leistungskennzahlen 15, 22
- Liste der Diffusionsstartwerte 21
- Löschen von Ergebnissen 6

## M

- Messniveaus 5
- Mittlere Gruppendichte 15
- Mittlere Gruppengröße 15
- Mittlere Kerngruppendichte 15
- Mittlerer Anteil an Kernmitgliedern 15
- Mittlerer Einfluss 22
- Mittlerer Indegree 15, 22
- Mittlerer Outdegree 15, 22

## O

- Outdegree 4, 15, 22

## Q

- Quellenfelder 7

## S

- Schwellenwert für Abdeckung 12, 14
- Schwellenwert für Genauigkeit 19, 21
- Scripts
  - Eigenschaften 8
- Startwerte
  - für Diffusionsanalyse 21

## T

- Typen (Registerkarte) 5

## U

- Ungerichtete Beziehungen 2

## V

- Verteilungsführer 13, 15
- Verteilungsscores 13, 15

## Z

- Zielfelder 7





