

Soziometrie und soziale Netzwerkanalyse für Lehrkräfte und pädagogische Fachkräfte

Eine Einführung in die soziometrische Datenanalyse und soziale Netzwerkanalyse
mit dem Statistikprogramm R

Pawel R. Kulawiak

Inhaltsverzeichnis

1	Vorwort	4
2	Soziometrie und soziale Netzwerke mit R	7
3	Was ist R?	10
3.1	Statistische Datenanalyse	10
3.2	Datenimport	12
3.3	Datenvisualisierung	13
3.4	Der Pipe-Operator <code> ></code> bzw. <code>%>%</code>	14
4	Was ist Soziometrie?	16
4.1	Ethik	17
4.2	Datenschutz	17
4.3	Geschichte der Soziometrie	17
5	Literatur	18

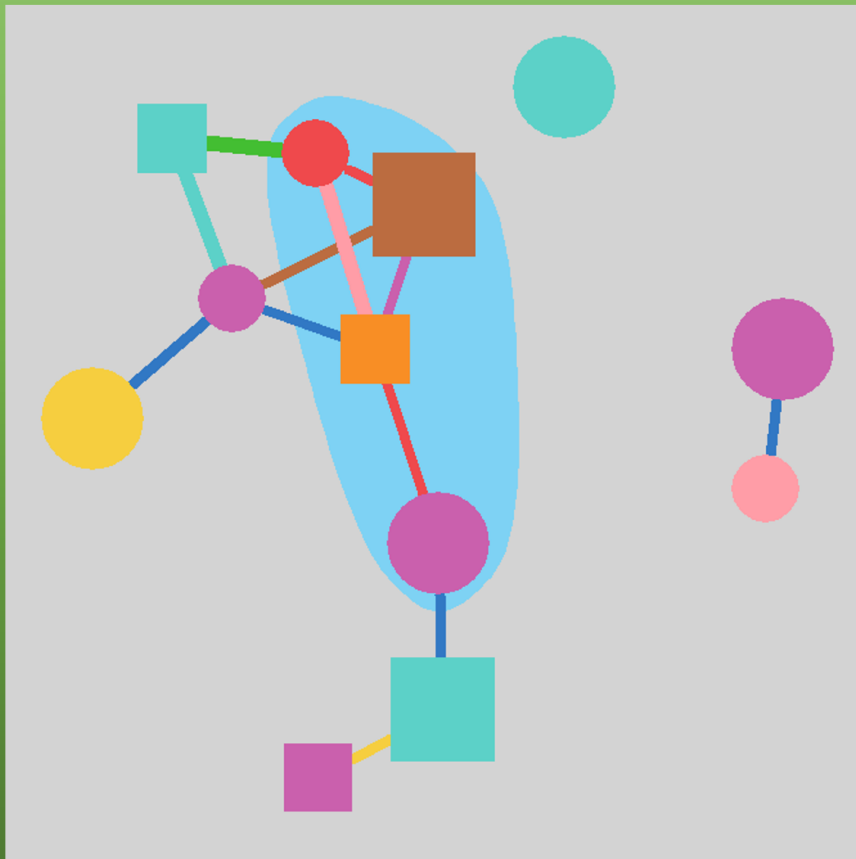
Kapitel 1

Vorwort

Soziometrie und soziale Netzwerkanalyse für Lehrkräfte und pädagogische Fachkräfte

*Eine Einführung in die soziometrische Datenanalyse und
soziale Netzwerkanalyse mit dem Statistikprogramm R*

Pawel R. Kulawiak



Wichtiger Hinweis: Dieses Buch ist in Entstehung begriffen (under construction) und wird sukzessive erweitert sowie verändert.

Letzte Aktualisierung: 2022-10-20 18:34:30

Link zur Online-Version: <https://pawelkulawiak.github.io/SozNetBuchR/>

Link zur PDF-Version: https://pawelkulawiak.github.io/SozNetBuchR/_main.pdf

Pädagogische Einrichtungen (Schulen, Kitas, Jugendfreizeiteinrichtungen, Sportvereine, therapeutische Wohngemeinschaften für Jugendliche, usw.) sind Orte der sozialen Interaktion, d.h. Kinder, Jugendliche, junge Erwachsene und pädagogische Fachkräfte pflegen und gestalten soziale Beziehungen miteinander sowie untereinander. Dieses Buch möchte wissenschaftliche sowie statistische Methoden zur Analyse sozialer Beziehungsstrukturen vorstellen und richtet sich daher explizit an pädagogische Fachkräfte (z.B. Lehrkräfte, Sonderpädagog*innen, Erzieher*innen, Sozialarbeiter*innen, usw.). Das vorliegende Buch orientiert sich dementsprechend an Beispielen aus der pädagogischen Praxis (vor allem aus der Schulpraxis).

Dieses kleine Buch (bzw. Büchlein) bietet dabei eine Einführung in die Soziometrie und soziale Netzwerkanalyse für Lehrkräfte und pädagogische Fachkräfte. Ein besonderes Augenmerk liegt hierbei auf der Verarbeitung sowie Analyse soziometrischer Daten und sozialer Netzwerkdaten mit dem Statistikprogramm R (insbesondere mit dem R-Zusatzpaket *igraph*). Dieses Buch ist keine umfassende Einführung in das Statistikprogramm R, sondern fokussiert sich vornehmlich auf die Analyse von sozialen Netzwerkdaten mit R. Grundkenntnisse der statistischen Datenanalyse und Datenverarbeitung mit R sind bei der Lektüre dieses Buches von Vorteil. Dennoch soll dieses Buch ein niedrighschwelliges Angebot sein und bietet deshalb viele R-Hilfestellungen und Verweise auf R-Einstiegsliteratur.

Hinweise und Anregungen zum Buch nehme ich gerne entgegen (kulawiak@uni-potsdam.de).

Ich wünsche Ihnen viel Spaß bei der Lektüre dieses kleinen Büchleins.

Berlin, Oktober 2021

Pawel R. Kulawiak

Kapitel 2

Soziometrie und soziale Netzwerke mit R

In diesem Buch erlernen Sie die Analyse sozialer Netzwerkdaten mit dem Statistikprogramm R (R Core Team 2021). Ich möchte Ihnen daher zu Beginn (ohne weitere Erläuterungen zu R) die Vorzüge der Programmiersprache R bei der Analyse sozialer Netzwerkdaten demonstrieren.

Fallbeispiel

Eine Grundschullehrkraft hat eine kleine Klasse der Jahrgangsstufe 1 übernommen (7 Kinder). Wenn die Schüler*innen der gesamten Schule in der großen Pause auf dem Schulhof zusammen spielen, dann ist es allerdings sehr unübersichtlich und die Lehrkraft kann nicht alle Aktivitäten der Kinder gleichermaßen im Auge behalten. Die Lehrkraft möchte daher einen genaueren Einblick in das “soziale Miteinander” ihrer neuen Klasse erhalten und startet deshalb eine **soziometrische Befragung** der einzelnen Schüler*innen: *Mit wem aus deiner Klasse hast du in der letzten großen Pause gespielt?* Die Lehrkraft notiert die Antworten der Kinder in einer Tabelle.

Tabelle 2.1: Antworten einer soziometrischen Befragung

Befragtes Kind		Antworten
Susi	...hat gespielt mit...	Max, Ali, Til
Eda	...hat gespielt mit...	Lena
Ella	...hat gespielt mit...	-
Lena	...hat gespielt mit...	Eda
Max	...hat gespielt mit...	Ali, Susi
Ali	...hat gespielt mit...	Max, Susi
Til	...hat gespielt mit...	Susi

Die Tabelle mit den Antworten bietet schon eine gewisse Übersichtlichkeit, da zum Beispiel ersichtlich wird, dass Ella bei der Befragung keine Mitschüler*innen benannt hat und auch von keinen Mitschüler*innen benannt wurde (Ella hat mit niemandem gespielt und niemand hat mit Ella gespielt). Die Lehrkraft möchte die soziometrischen Daten weiter verarbeiten und überträgt die Antworten der Kinder deshalb in das Statistikprogramm R. Zuvor wird mit den ersten beiden R-Befehlen (bzw. R-Funktionen) das R-Zusatzpaket *igraph* installiert und anschließend geladen. *igraph* (Csardi und Nepusz 2006) ist ein spezifisches R-Zusatzpaket zur Verarbeitung und Analyse von Netzwerkdaten.

```
install.packages("igraph") # R-Zusatzpaket igraph installieren
library(igraph) # R-Zusatzpaket igraph laden
```

Anschließend werden alle Antworten der Kinder in der Form *Susi -> Max* (Susi ...*hat gespielt mit*... Max) an die Funktion `graph_from_literal()` übergeben (eine Funktion des Zusatzpaketes *igraph*). Da Ella mit niemandem gespielt hat, erfolgt für Ella ausschließlich die Übergabe des Namen Ella. Alle Antworten der Kinder werden dann mit dem Zuweisungspfeil `<-` als Objekt **netzwerk** gespeichert.

```
netzwerk <- graph_from_literal(Susi -> Max, Susi -> Ali, Susi -> Til, Eda -> Lena,
                              Ella, Lena -> Eda, Max -> Ali, Max -> Susi,
                              Ali -> Max, Ali -> Susi, Til -> Susi)
```

Mit der Funktion `print_all()` können wir jederzeit die Informationen aus dem Objekt **netzwerk** abrufen (also die Antworten der Kinder). Hierfür übergeben wir das Objekt **netzwerk** an die entsprechende Funktion. Die Antworten der Kinder erscheinen in der Form *Susi->Max* (Susi ...*hat gespielt mit*... Max).

```
print_all(netzwerk)
```

```
## IGRAPH 123f954 DN-- 7 10 --
## + attr: name (v/c)
## + edges from 123f954 (vertex names):
## [1] Susi->Max Susi->Ali Susi->Til Max ->Susi Max ->Ali Ali ->Susi
## [7] Ali ->Max Til ->Susi Eda ->Lena Lena->Eda
```

Ein Bild sagt mehr als tausend Worte. Abschließend möchten wir daher die Antworten der Kinder visualisieren (Netzwerkvisualisierung). Wir übergeben daher das Objekt **netzwerk** an die Funktion `plot()` (*plotten*) und definieren noch ein paar Eigenschaften der Abbildung, z.B. die Größe der Kreissymbole mit dem Argument `vertex.size = 30`.

```
par(mar = c(0, 0, 0, 0)) # Ränder (margins) der Grafik auf null setzten
plot(netzwerk, vertex.size = 30)
```

Die Grafik ist eine Repräsentation der Antworten der Kinder (mit Pfeilen und Kreissymbolen). In der Grafik sehen wir auf den ersten Blick spannende soziale Netzwerkstrukturen: Eine Triade (Max, Ali, Susi), zwei Dyaden (Eda und Lena sowie Susi und Til) und das isolierte Kind ohne Spielpartner*innen (Ella).

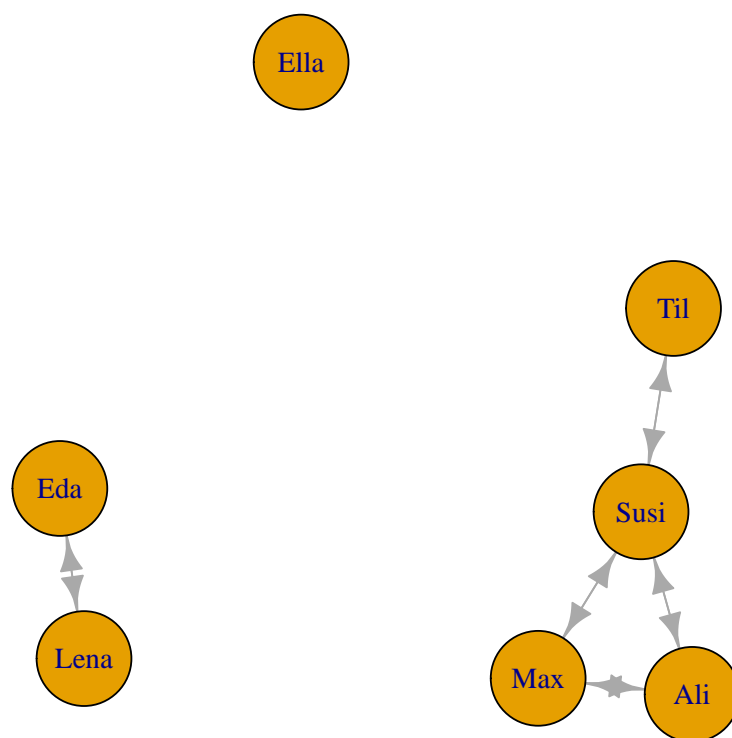


Abbildung 2.1: Soziales Netzwerk: Mit wem aus deiner Klasse hast du in der letzten großen Pause gespielt?

Kapitel 3

Was ist R?

R (R Core Team 2021) ist eine herausragende, freie und zukunftsweisende Programmiersprache zur Datenverarbeitung und Datenanalyse. Mit dem vorangegangenen Beispiel haben wir einen ersten Einblick in die Verarbeitung, Analyse und Visualisierung von sozialen Netzwerkdaten mit R erhalten. An dieser Stelle kann aber keine umfassende Einführung in das Statistikprogramm R erfolgen. Hierfür konsultieren Sie bitte das hervorragende Buch “R kompakt - Der schnelle Einstieg in die Datenanalyse” (Wollschläger 2021). Dort erfahren Sie alle Grundlagen zur Installation und Ausführung von R sowie zur Datenverarbeitung, Datenanalyse und Datenvisualisierung mit R. In diesem Kapitel erfolgt dennoch eine kurze sowie exemplarische Einführung in die wichtigsten Grundlagen von R.

3.1 Statistische Datenanalyse

Wenn wir die Sozialbeziehungen in sozialen Gruppen betrachten wollen, so erscheint eine statistische Beschreibung der sozialen Gruppe sinnvoll. Wir sind z.B. interessiert an den persönlichen Merkmalen und Eigenschaften der Schüler*innen in einer Klasse (z.B. Alter, Geschlecht, Verhaltensweisen, usw.). Informationen dieser Art werden in Datensätzen gespeichert.

Tabelle 3.1: Beispieldatensatz

Name	Alter	Geschlecht	Introvertiertes Verhalten (IV)
Susi	6	0	4
Eda	6	0	1
Ella	7	0	4
Lena	6	0	3
Max	5	1	2
Ali	6	1	1
Til	6	1	2

Mit R erstellen wir die 4 Variablen (`name`, `alter`, `geschlecht`, `IV`) als Objekte mit dem Zuweisungspfeil `<-`. Die Werte der jeweiligen Variablen werden dabei mit der Funktion `c()` zusammengefasst (combine values into a vector).

```
name <- c("Susi", "Eda", "Ella", "Lena", "Max", "Ali", "Til")
alter <- c(6, 6, 7, 6, 5, 6, 6)
geschlecht <- c(0, 0, 0, 0, 1, 1, 1)
IV <- c(4, 1, 4, 3, 2, 1, 2)
```

Mit R kann man sich stets alle Objekte anschauen, indem man sie einfach “aufruft”. Wir betrachten die Variable introvertiertes Verhalten (`IV`).

```
IV # Variable aufrufen und anzeigen
```

```
## [1] 4 1 4 3 2 1 2
```

Die 4 Variablen werden nun in einem Datensatz zusammengefasst. Dabei werden die Variablen an die Funktion `data.frame()` übergeben. Wir erstellen mit der Funktion `data.frame()` ein neues Objekt (den Beispieldatensatz). Der Beispieldatensatz wird anschließend aufgerufen und angezeigt.

```
beispieldatensatz <- data.frame(name, alter, geschlecht, IV)
beispieldatensatz # Beispieldatensatz aufrufen und anzeigen
```

```
##   name alter geschlecht IV
## 1 Susi     6           0  4
## 2 Eda     6           0  1
## 3 Ella    7           0  4
## 4 Lena    6           0  3
## 5 Max     5           1  2
## 6 Ali     6           1  1
## 7 Til     6           1  2
```

Nun möchten wir die Daten analysieren. Wir betrachten das introvertierte Verhalten (`IV`) der Kinder. Mit `beispieldatensatz$IV` können wir die entsprechende Variable des Beispieldatensatzes aufrufen. Wir berechnen den Mittelwert mit der Funktion `mean()` und den getrimmten Mittelwert indem wir zusätzlich das Argument `trim = 0.2` festlegen. Die Variable `beispieldatensatz$IV` wird dabei an die Funktion `mean()` übergeben.

```
beispieldatensatz$IV # Variable des Beispieldatensatzes aufrufen und anzeigen
```

```
## [1] 4 1 4 3 2 1 2
```

```
mean(beispieldatensatz$IV) # Mittelwert
```

```
## [1] 2.428571
```

```
mean(beispieldatensatz$IV, trim = 0.2) # Getrimmter Mittelwert
```

```
## [1] 2.4
```

3.2 Datenimport

In den allermeisten Fällen sind die zu analysierenden Daten als Excel- oder CSV-Dateien auf der Festplatte oder in einer Cloud gespeichert. Für die Analysen mit R müssen diese Daten zunächst eingelesen werden (Datenimport).

Hier finden Sie den bekannten Beispieldatensatz als CSV-Datei (Link zur CSV-Datei¹) und als Excel-Datei (Link zur Excel-Datei²). Speichern Sie die beiden Dateien auf Ihrer Festplatte, z.B. auf dem Desktop.

Mit R gibt es viele Wege um Daten einzulesen. Hier nur ein Beispiel für die CSV-Datei. Sie müssen den Dateipfad (Speicherort und Dateinamen mit Dateiendung) an die Funktion `read.csv2()` übergeben.

```
daten <- read.csv2("C:/Users/pawel/Desktop/daten.csv")
```

Das Einlesen einer Excel-Datei gelingt mit der Funktion `read_excel()`. Zuvor muss allerdings das entsprechende R-Zusatzpaket *readxl* (Excel) (Wickham und Bryan 2019) installiert und geladen werden.

```
install.packages("readxl") # R-Zusatzpaket readxl (Excel) installieren
library(readxl) # R-Zusatzpaket readxl (Excel) laden
daten <- read_excel("C:/Users/pawel/Desktop/daten.xlsx")
```

Wir können auch auf die Angabe des Dateipfades verzichten und stattdessen mit der Funktion `file.choose()` das Dateiverzeichnis nach der entsprechenden CSV- oder Excel-Datei durchsuchen.

```
daten <- read_excel(file.choose())
```

Die Daten können aber auch direkt aus dem Internet eingelesen werden. Hierfür benötigen wir lediglich den Link zur Datei (dies kann auch ein geteilter Link zu einer Datei in einer Cloud sein).

```
daten <- read.csv2("https://figshare.com/ndownloader/files/31108225")
```

In allen Fällen war das einlesen des Beispieldatensatzes erfolgreich und wir können den Datensatz aufrufen.

```
daten
```

¹<https://figshare.com/ndownloader/files/31108225>

²<https://figshare.com/ndownloader/files/31108216>

```
##   name alter geschlecht IV
## 1 Susi     6           0  4
## 2 Eda      6           0  1
## 3 Ella     7           0  4
## 4 Lena     6           0  3
## 5 Max      5           1  2
## 6 Ali      6           1  1
## 7 Til      6           1  2
```

3.3 Datenvisualisierung

R eignet sich vorzüglich zur Visualisierung von Daten. Wir erstellen ein Histogramm der Variable IV.

```
hist(daten$IV)
```

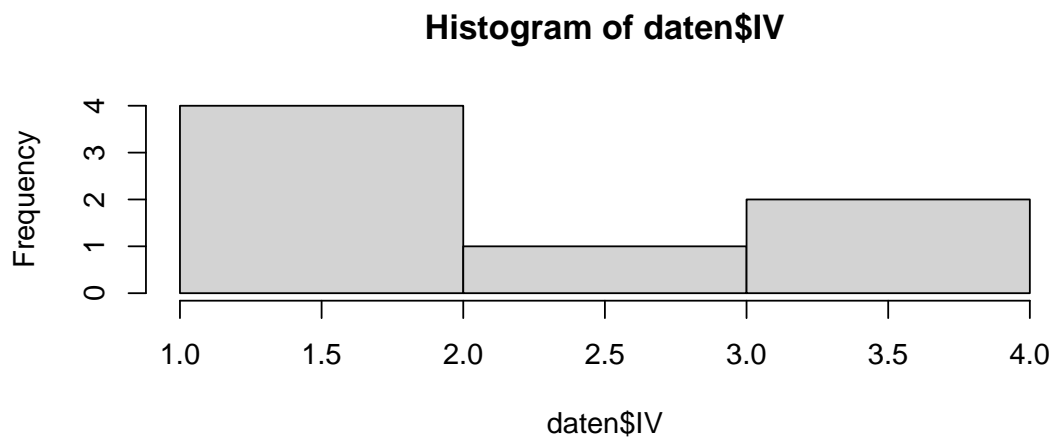


Abbildung 3.1: Histogramm

Mittlerweile sind die Grafikbefehle des Zusatzpaketes *ggplot2* (Wickham 2016) der Standard bei der Erstellung von Grafiken mit R. Das Zusatzpaket *ggplot2* ist in der Paketsammlung *tidyverse* enthalten. *tidyverse* (Wickham u. a. 2019) ist wiederum eine Zusammenstellung vieler extrem nützlicher Zusatzpakete (z.B. Werkzeuge fürs Datenmanagement).

```
install.packages("tidyverse")
library(tidyverse)
```

Wir erstellen ein Histogramm mit den Grafikbefehlen des Zusatzpaketes *ggplot2*.

```
ggplot(data = daten, mapping = aes(IV)) + geom_histogram()
```

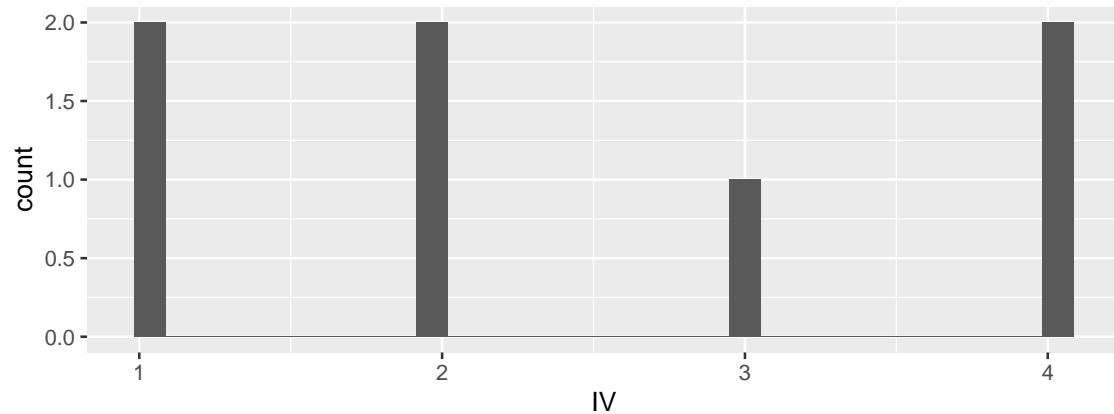


Abbildung 3.2: Histogramm mit ggplot2

3.4 Der Pipe-Operator `|>` bzw. `%>%`

R-Funktionen können ineinander geschachtelt werden. Dies ist z.B. notwendig bei der Visualisierung einer Häufigkeitstabelle. Zunächst erstellen wir die benötigte Häufigkeitstabelle der Variable `alter`.

```
table(daten$alter)
```

```
##
## 5 6 7
## 1 5 1
```

Zur Visualisierung der Häufigkeitstabelle wird `table(daten$alter)` an den Befehl `pie()` (Kreisdiagramm/Tortendiagramm) übergeben.

```
pie(table(daten$alter), main = "Alter in Jahren")
```

Diese Schachtelung von Befehlen (`table()` innerhalb von `pie()`) ist unübersichtlich. Eine objektorientierte Programmierung ist übersichtlicher (auch dies führt zur erfolgreichen Visualisierung).

```
table_alter <- table(daten$alter)
pie(table_alter, main = "Alter in Jahren")
```

Noch übersichtlicher ist die "Weitergabe" der Häufigkeitstabelle an den Grafik-Befehl mittels Pipe-Operator `|>` bzw. `%>%` (auch dies führt zur erfolgreichen Visualisierung). Dabei ist es in den meisten Fällen unerheblich ob `|>` oder `%>%` genutzt wird. `%>%` entstammt aus der Paketsammlung *tidyverse* (Wickham u. a. 2019) und bietet einige Vorteile gegenüber `|>`. Insgesamt ist der Pipe-Operator (egal ob `|>` oder `%>%`) ein sehr nützliches Werkzeug.

```
table(daten$alter) |> pie(main = "Alter in Jahren")
```

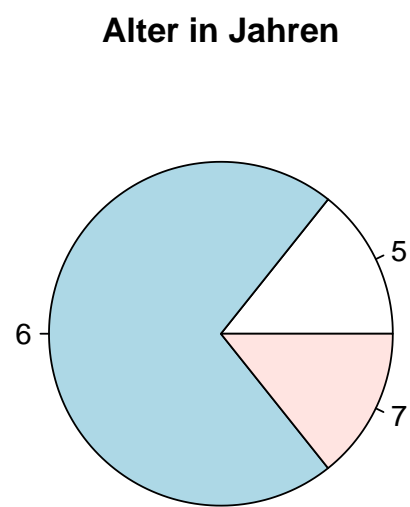


Abbildung 3.3: Kreisdiagramm

Kapitel 4

Was ist Soziometrie?

Das Prinzip der soziometrischen Befragung wurde bereits im Eingangsbeispiel dargestellt. Die Mitglieder einer sozialen Gruppe (z.B. Schüler*innen in einer Klassengemeinschaft oder Lehrkräfte in einem Kollegium) werden anhand einer bestimmten Frage nach ihrem “Verhältnis” zu den anderen Gruppenmitgliedern befragt. Die gestellte Frage wird auch als *soziometrisches Kriterium* bezeichnet.

Tabelle 4.1: Beispiele für soziometrische Kriterien:
(Schüler*innen-Ebene)

Soziometrisches Kriterium	Dimension/ gemessenes Konstrukt
Mit wem aus deiner Klasse hast du in der letzten großen Pause gespielt?	Gemeinsame Aktivitäten (Spiel)
Mit wem aus deiner Klasse spielst du gerne?	Beliebtheit als Spielpartner*in
Mit wem aus deiner Klasse spielst du nicht so gerne?	Unbeliebtheit als Spielpartner*in
Mit wem aus deiner Klasse hast du in der letzten Woche gemeinsam gelernt (außerhalb der Schule)?	Außerschulische Lernaktivitäten
Mit wem aus deiner Klasse triffst du dich auch außerhalb der Schule?	Außerschulische Kontakte (allgemein)
Mit wem aus deiner Klasse triffst du dich auch außerhalb der Schule zum Spielen?	Außerschulische Kontakte (Spielen)
Wen aus deiner Klasse magst du gerne?	Allgemeine Beliebtheit
Wen aus deiner Klasse magst du nicht so gerne?	Allgemeine Unbeliebtheit
Neben wem aus deiner Klasse würdest du gerne sitzen?	Beliebtheit als Sitzpartner*in

Tabelle 4.1: Beispiele für soziometrische Kriterien: (Schüler*innen-Ebene)

Soziometrisches Kriterium	Dimension/ gemessenes Konstrukt
Neben wem aus deiner Klasse würdest du nicht so gerne sitzen?	Unbeliebtheit als Sitzpartner*in
Gibt es Mitschüler*innen, die deine Sachen mit Absicht kaputtmachen?	Abweichendes Verhalten (Sachbeschädigung)
Gibt es Mitschüler*innen, die dich mit Absicht schlagen, schubsen, treten, usw.?	Abweichendes Verhalten (physische Gewalt)

Tabelle 4.2: Beispiele für soziometrische Kriterien: (Lehrkraft-Ebene)

Soziometrisches Kriterium	Dimension/ gemessenes Konstrukt
Mit welchen Kolleg*innen haben Sie im aktuellen Schuljahr Förderplangespräche geführt?	Teamarbeit (Förderplanung)
Mit welchen Kolleg*innen haben Sie in den letzten vier Wochen gemeinsam unterrichtet (Teamteaching)?	Teamarbeit (Teamteaching)
Mit welchen Kolleg*innen haben Sie in den letzten vier Wochen kollegiale Fallberatungen durchgeführt?	Teamarbeit (kollegiale Fallberatung)

Das soziometrische Kriterium sollte mit Bedacht formuliert werden: Welche Dimension einer Sozialbeziehung soll fokussiert werden (z.B. Beliebtheit, Teamarbeit, usw.)? Soll ein konkreter zeitlicher Rahmen gesetzt werden (z.B. gestern, heute, morgen, aktuelles Schuljahr, in der Zukunft, usw.)? Soll ein konkretes Setting abgefragt werden (z.B. in der Pause, im Kunstunterricht, auf dem Ausflug, usw.)?

Soziometrische Befragungen können schriftlich (Fragebogen) oder mündlich (Interview) durchgeführt werden.

4.1 Ethik

4.2 Datenschutz

4.3 Geschichte der Soziometrie

Kapitel 5

Literatur

- Csardi, Gabor, und Tamas Nepusz. 2006. „The igraph software package for complex network research“. *InterJournal Complex Systems*: 1695. <https://igraph.org>.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, u. a. 2019. „Welcome to the tidyverse“. *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, und Jennifer Bryan. 2019. *readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- Wollschläger, Daniel. 2021. *R kompakt: Der schnelle Einstieg in die Datenanalyse*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-63075-4>.